Universidade de São Paulo
Instituto de Astronomia, Geofísica e Ciências Atmosféricas
Departamento de Astronomia

Thiago Ferreira dos Santos

Unveiling Planetary Systems in Solar Twins with High-Precision Radial Velocity

Revelando Sistemas Planetários em Gêmeas Solares com Velocidade Radial de Alta Precisão

São Paulo

2023

Thiago Ferreira dos Santos

Unveiling Planetary Systems in Solar Twins with High-Precision Radial Velocity

Revelando Sistemas Planetários em Gêmeas Solares com Velocidade Radial de Alta Precisão

Thesis presented to the Department of Astronomy of the Instituto de Astronomia, Geofísica e Ciências Atmosféricas at the Universidade de São Paulo as a partial requirement to obtain the title of Master of Science.

Area: Astronomy
Advisor: Jorge Luis Meléndez Moreno, PhD

Versão Corrigida.
O original encontra-se disponível na Unidade.

São Paulo
2023

Para Evany, Luiz Carlos, Maria Luiza, e Maria das Graças.

Acknowledgements

Aos meus pais, Evany e Luiz Carlos, minha irmã, Maria Luiza, e minha tia, Maria das Graças, principais impulsionadores dos meus objetivos e aspirações.

Aos professores Jorge Meléndez, Roberto Saito, e Dante Minniti, cujos ensinamentos e dedicação abriram novos horizontes em minha busca pelo conhecimento.

À Anne Rathsam, *ein Licht zwischen den Schatten*, cuja amizade e ensinamentos imprimiram uma marca indelével em minha alma, moldando-me como ser humano e pesquisador.

Aos amigos do grupo SAMPA, Jhon Yana Galarza, Yuri Netto, Diego Lorenzo-Oliveira, Geisa Ponte, Anne Rathsam, Gabriela Silva, Giulia Martos, Débora Alves, Letícia Laurindo, Liana Li, Shejeelamal Jameela, Henrique Reggiani, e Leonardo dos Santos.

Aos *chaspas* Adolfo Scheidt, Alexandher Kill, Alisson Pereira, Anthony Bet, Amanda Macário, Diogo Souza, Francisco Neubauer, Franck Peffan, Gabriel de Carlos Felicio dos Santos Barbosa, Gabriel Duarte, Giovanna Tramontin, Guilherme Tirelli, Gustavo Café, Hamilton Haskel, Helcio Felippe, Ian Diaz, Janayna Mendes, João Paulo Boger, João Pering, José Ricardo, Juarez Dal'Acqua, Kevin Voigt, Leonardo Souto, Luiz Portilho, Muryel Guolo, Pedro Sanson, Pedro Ibarra, Pedro Jung, Pedro Krick, Rômulo Cenci, e Valéria Mattiello.

À CAPES pela concessão da bolsa de Mestrado.[1]

---

[1] This dissertation was written in LaTeXusing the class `IAGTHESIS` for thesis and dissertations from IAG.

Si un día me faltas no seré nada
Y al mismo tiempo lo seré todo
Porque en tus ojos están mis alas
Y está la orilla donde me ahogo

Carlos Varela

Resumo

A busca por exoplanetas testemunhou um aumento drástico nos últimos anos com o surgimento de novos instrumentos e técnicas precisas de análise de dados. Observações recentes mostram que planetas gigantes frios tendem a se formar preferencialmente ao redor de estrelas com metalicidade solar, enquanto planetas potencialmente habitáveis de baixa massa podem ser comuns ao redor de estrelas que hospedam planetas gigantes frios semelhantes a Júpiter. Neste estudo, apresentamos a detecção de 16 candidatos a sistemas exoplanetários utilizando dados de velocidade radial de alta precisão do espectrógrafo ESO/HARPS, focando em estrelas gêmeas solares como parte do programa Solar Twin Planet Search. De modo geral, examinamos correlações entre as velocidades radiais e indicadores de atividade magnética estelar, além de modulações temporais, para identificar potenciais sinais exoplanetários nos dados. Em seguida, empregamos modelos Keplerianos e regressão por processos gaussianos quasi-periódicos em alguns casos para determinar as características físicas e orbitais dos candidatos encontrados. Os resultados apresentados aqui fornecem informações valiosas sobre a formação e evolução de sistemas planetários ao redor de estrelas gêmeas solares, destacando fenômenos de engolimento planetário e a presença de companheiros externos não observados que podem explicar a evolução dinâmica de exoplanetas interiores, principalmente do tipo gigante de gelo. Identificamos vários análogos de Júpiter e de Netuno, estabelecendo uma base com um catálogo de alvos a serem observados pela próxima geração de telescópios terrestres e espaciais em busca de exoplanetas análogos à Terra.

Abstract

The search for exoplanets has witnessed significant growth in recent years with the dawn of new instruments and precise data analysis techniques. Recent observations suggest that cold giant planets are likely to form around stars possessing metallicity levels comparable to those of the Sun, while low–mass, potentially habitable planets may be common around stars that host cold Jupiter–like planets. This study presents the detection of 16 exoplanetary systems from the analysis of high–precision radial velocity time series data from the ESO/HARPS spectrograph, concentrating on solar twin stars as part of the Solar Twin Planet Search programme. We examined correlations between magnetic stellar activity indicators and radial velocity measurements, alongside timing variations, to detect any potential exoplanetary candidates' modulations. Then, we employed Keplerian models and quasi–periodic Gaussian process regression in some cases, to probe the physical and orbital features of the candidates found. Our results offer significant insights into the formation and evolution of planetary systems around solar twin stars, uncovering planetary engulfment phenomena and the existence of outer companions that may contribute to the dynamic evolution of mainly inner ice giant exoplanets. Our findings include several Neptune–like and Jupiter analogues, laying the ground with a catalogue of targets to be observed by the next generation of ground– and space–based telescopes seeking Earth–analogue exoplanets.

List of Figures

List of Tables

Contents

Chapter 1

Introduction

1.1 Needles in a Haystack: Towards Jupiter and Earth Analogues

Upon the unyielding quest to fathom what lies beyond the boundaries of the Solar System, we embark on an arduous quest to uncover places that harbour the ethereal essence of life, in which within the halls of contemporary planetary science and astrobiology, the sacred grail manifests as the elusive quest for habitable worlds. Our yearning intensifies to discover planets where life can establish in forms both familiar and beyond the proper scope of human imagination. It is within this profound pilgrimage that the discovery and characterisation of Earth and Jupiter analogues, especially orbiting Sun-like stars, emerge.

Jupiter is a giant planet located in the outer region ($a \sim 5.2$ AU) of the Solar System, yet, its impact extends far beyond its orbit. Due to its size and gravitational pull, it played a crucial role in the formation of the inner rocky planets, including the Earth. The importance of finding such objects, especially those orbiting Sun-like stars, is evident in light of our own Solar System's architecture, which suggests that Jupiter played an important role by (*i*) limiting the migration to inner orbits of outer massive bodies (e.g., Izidoro et al. 2015; Batygin and Laughlin 2015; see Fig. 1.1), therefore preserving the inner rocky planet region, and (*ii*) enhancing planetary habitability by sowing short-period orbits with planetesimals composed of water and other smaller bodies rich in volatiles formed beyond the ice line (the boundary where water and other volatile compounds freeze into solid ice due to low temperatures; e.g., see Figs. 9 and 10 in Raymond and Izidoro 2017).

Since its formation, about 4.6 billion years ago, the Earth has undergone several phases, called aeons, which can broadly be categorised as the Hadean (4.6 to 4.0 billion years ago), Archean

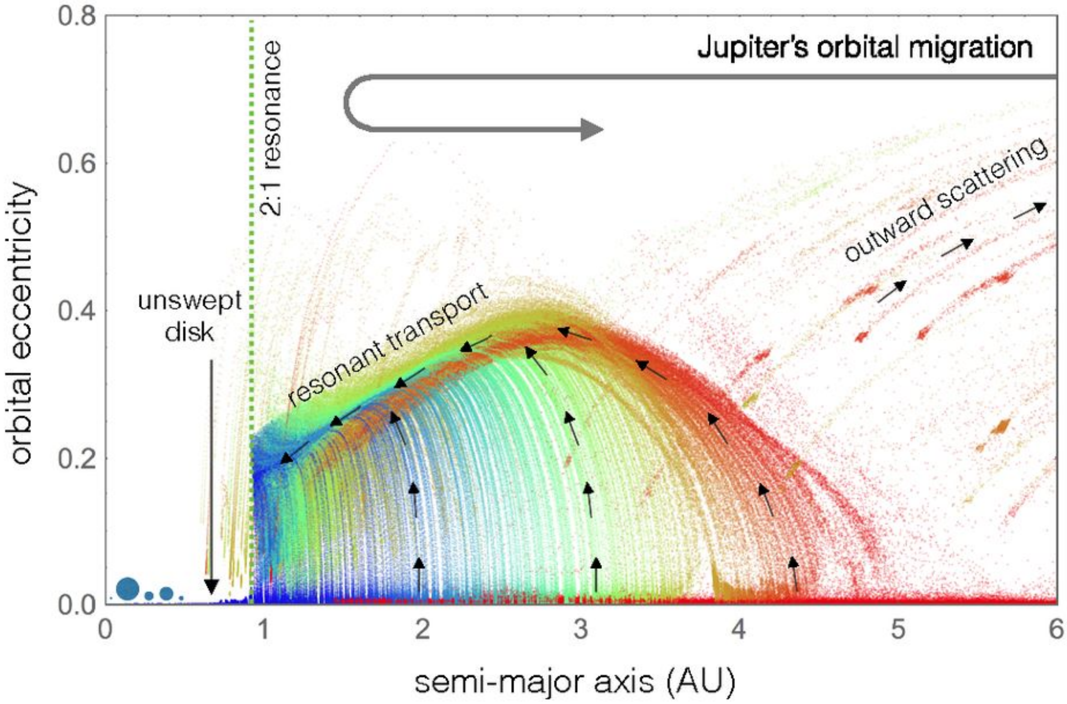

Figure 1.1: During Jupiter's inward orbital migration to approximately 1.5 AU, certain planetesimals undergo orbital instability, resulting in a noticeable increase in their orbital eccentricity. Consequently, a portion of these simulated planetesimals is expelled from the Solar System, while others enter a 2:1 mean motion resonance (MMR). This migration of Jupiter has a significant impact on the inner region of the Solar System by effectively clearing it, thus creating favourable conditions for the development of rocky planets. Adapted from Batygin and Laughlin (2015).

(4.0 to 2.5 billion years ago), Proterozoic (2.5 bl. to 540 million years ago), and Phanerozoic (540 million years ago to present day). Each of these is characterised by unique geological, biological, and environmental changes that have shaped the Earth into the planet we know today (Faure 2007; Benton 2008, and references therein). During the Hadean aeon, Jupiter likely deflected some asteroids and comets away from the inner region of the Solar System, reducing the frequency of impacts on the early Earth. Jupiter may have also diverted Mars-sized objects that collided with Earth, leading to the formation of the Moon. As Earth's crust solidified and continents were formed, Jupiter's strong gravity continued to deflect celestial bodies during the Archean aeon, contributing to the orbital stability of the inner planets. Furthermore, Jupiter's gravitational pull potentially delivered water and organic molecules to Earth, crucial for the development of life on Earth given the toxic composition of its atmosphere during that phase. In the Proterozoic aeon, Jupiter's gravitational influence stabilised the Solar System, preventing catastrophic collisions and allowing for the emergence of complex life forms. Lastly, in the Phanerozoic aeon, Jupiter's gravity continued to shape planetary orbits, protecting Earth from potential collisions and enabling the evolution and flourishing of life.

Earth analogues are of particular interest in planetary science since they may possess the right conditions in which life exists[1]. At the same time, life definition is a complex subject and it is outside the scope of this writing but, in general, life, as we know, requires the presence of liquid water, a stable atmosphere, and a source of energy, all of which can be facilitated by the planet's distance from its host star (where for Sun-like stars ranges from 0.75 to 1.76 AU; Kopparapu et al. 2014).

We define an Earth analogue as a planet that shares similar characteristics to our own Earth, which includes (a) a rocky composition similar to Earth, mainly by Fe/Ni-metal plus $\mathrm{MgSiO_3}-$rock (Zeng et al., 2016, 2019); (b) radius between 0.5 and 1.6 $\mathrm{R}_\oplus$, since exoplanets with sizes below or above these limits are *(i)* unable to retain an atmosphere caused by intense X-ray and ultraviolet radiation from their host stars (Owen and Wu, 2017) and *(ii)* have too low density to be comprised of iron and silicates alone, i.e., are not of rocky composition (Rogers, 2015); and (c) masses similar to 1 $\mathrm{M}_\oplus$. Even planets with lower masses (e.g., Mars-like 0.1 $\mathrm{M}_\oplus$) could have a rocky composition, but they are also more likely to have thinner atmospheres, which could affect their ability to support life. On the other hand, planets with higher masses, such as super-Earths, may have denser atmospheres and could be more similar to mini-Neptunes rather than Earth-like planets (e.g., Zeng et al. 2016). Likewise, in terms of effective temperature, the definition may vary, according to the own definition of life; (d) at the lower end of the temperature spectrum, some extremophiles (e.g., *Rhodotorula glutinis*) are capable of surviving at temperatures as low as $-20°$C, or even lower (Clarke et al., 2013). At the upper end of the temperature spectrum, archæa such as *Pyrolobus fumarii*, *Pyrococcus furiosus* and *Methanopyrus kandleri* can live in extremely hot environments, with temperatures exceeding $100°$C, such as in a hydrothermal vent or volcanic hot springs (Marion et al., 2003).

Jupiter analogues ($P \in [5, 18]$ years, $m \sin i \in [0.3, 3]$ M_{Jup}, $e < 0.3$) form preferentially around stars with solar-like metallicities and, strikingly, it has been suggested that low-mass plan-

---

[1] Although it is thrilling to discover rocky exoplanets that resemble the Earth, the complexity of the planet formation process requires a multidisciplinary approach to determine their habitability. Orbiting in the habitable zone around a star does not necessarily means that a planet is habitable. And it is also important to highlight that Earth's aeons are unique to our planet, and cannot be by any means directly applied to exoplanets. While we can observe exoplanets and gather information about their composition, atmosphere, and other characteristics, we cannot directly infer their past. Nevertheless, the search for Jupiter analogues in exoplanetary systems is still crucial. These gas giants are believed to play a fundamental role in shaping the architecture of their planetary systems, and Sun-like stars that host such a kind of planets present the best opportunities for further studies on Solar System analogues.

ets could be common around stars that host a cold Jupiter, whose orbital periods are greater than 100 days (Buchhave et al. 2018, and references therein). Moreover, Howard et al. (2012) demonstrated that small Earth-sized planets are more common than giant planets, and because there were, at the time of that publication, more planets in distant orbits than short-period orbits, it was estimated that about 40% of this group have lower-mass planetary companions. These findings open precedents to the search for Earth-like planets around solar-type stars by primarily looking for Jupiter-like planets (e.g. Petigura et al. 2013; He and Weiss 2023; Ferreira et al. 2023, and references therein).

1.2 The Solar Twin Planet Search (STPS) Observing Programme

The aphorism in exoplanet research, *know thy star, know thy planet*, reminds us of the complex interplay between a star and its orbiting planets. An in-depth comprehension of the host star's properties, from gravitational forces to electromagnetic radiation, is essential to understand the nuances of its orbiting planets. For instance, the stellar luminosity and its effective temperature in a simpler case are closely related to the potential habitability of its orbiting planets, by determining a range called the habitable zone, where liquid water can exist on the surface of orbiting (exo)planets (Kopparapu et al., 2014). Furthermore, stellar brightness and levels of magnetic activity sometimes put a threshold into the detectability of exoplanets and, in some cases, such properties are used to refute previously discovered exoplanets (e.g., Queloz et al. 2001). Ultimately, the star-planet connection also works in reverse, with the properties of the planets providing clues about the stars they orbit. Signatures of stellar spots, for instance, may appear as temporary bumps in brightness in a transiting light curve (Silva-Valio et al., 2010; Netto and Valio, 2020; Valio and Araújo, 2022), and the planet's composition can provide insights into the chemical makeup of the protoplanetary disc from where it formed, which in turn provides clues about the early conditions of its stellar system (Schiller et al., 2018). How and where a planet may have formed may be also imprinted in such information; *in situ* (Batygin et al., 2016) or far away followed by inward migration (e.g., Reggiani et al. 2022).

In light of the star-planet connection, solar twins offer the best conditions to study exoplanets due to several factors. Firstly, as their name suggests, they are similar to our Sun in terms of effective temperature (inside the intervals ±100 K from the solar value; $T_{\text{eff},\odot} = 5777\ K$), surface gravity ($\pm 0.1$ dex; $\log g_\odot = 4.44$ dex), and metallicity ([Fe/H]; ± 0.1 dex from $[\mathrm{Fe/H}]_\odot = 0.00$), which in turn translates to similar masses and radii as of the Sun (see Fig. 1.2). It is important to mention that these limits may vary from author to author (e.g., Porto de Mello and da Silva 1997; Ramírez et al. 2014; Yana Galarza et al. 2021), and the ones used in this work are the most common definition. Some stars, as seen in the aforementioned figure, do not appear to belong to these established limits, precisely because of the non-uniformity of the definition concerning analogues and solar twins. Secondly, solar twin stars are relatively bright (although distant solar twins have been searched recently; Giribaldi et al. 2019; Lehmann et al. 2023), which makes them *highly observable* and easy to follow up. Lastly, solar twins usually exhibit greater stability and less magnetic activity compared to other types of stars, making them easier to study the orbital and physical properties of their orbiting planets with accuracy. This reduces the impact of noise from the stellar activity and allows for a clearer signal from the exoplanet to be observed and modelled (e.g., refer to Section 4.3).

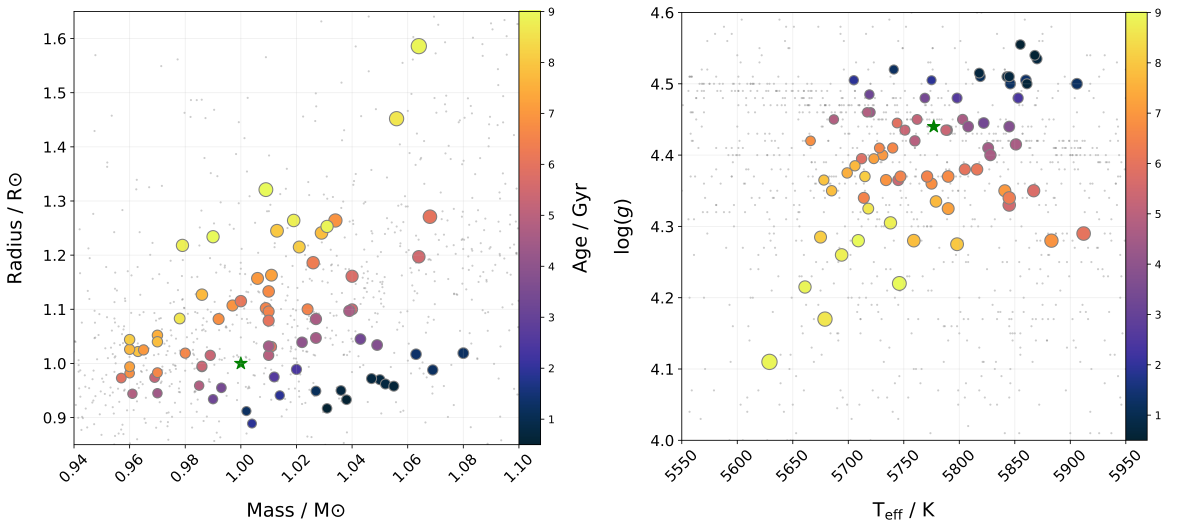

Figure 1.2: Masses and radius on the left panel of 79 solar twins in the STPS programme with measured parameters. The sizes of each data point represent the corresponding star's luminosity (spanning $0.78 - 2.26\ L_\odot$), while the colour indicates its age (ranging $0.5 - 9$ Gyr). The grey points in the background show known stars hosting exoplanets from the SWEET-Cat 2.0 catalogue (Sousa et al., 2021), as of April 8, 2023. The position of the Sun is highlighted with a green star. On the right panel, it displays the effective temperature and surface gravity for the same set of stars.

The stars for the Solar Twin Planet Search (STPS) programme were selected using a variety of criteria, including colour-based selection methods (e.g., Meléndez and Ramírez 2007; Ramírez et al. 2009; Meléndez et al. 2009) and previously identified solar twins found in publicly available spectral databases (e.g., Valenti and Fischer 2005; Meléndez et al. 2014; Bensby et al. 2014). It comprises more than 100 nights of observations with the ESO/High Accuracy Radial Velocity

Planet Searcher (HARPS) spectrograph, located at La Silla Observatory, Chile, and the main objective of this project is to establish a connection, wherever possible, between the properties of a star, such as its chemical composition, and the characteristics of any orbiting exoplanets, if they are discovered via radial velocity (RV) analysis. The initial set of stars analysed for this project consisted of 89 stars, whose fundamental properties were previously determined (Spina et al., 2018; Bedell et al., 2018; Botelho et al., 2019), and it was gradually reduced based on the number of RVs and chemical abundances previously estimated.

Each star analysed was observed generally with exposures lasting 900 seconds. This minimum exposure time is necessary to average over multiple cycles of the asteroseismic p-mode oscillations, which typically have a cycle period of approximately five minutes and RV amplitude of approximately 1 m s^{-1} for stars similar to the Sun (e.g., Kjeldsen and Bedding 1995; see also Section 3.3). Since the observed targets are relatively bright, it was possible to achieve usually a signal-to-noise ratio (SNR) > 100 px^{-1} at 600 nm for each spectrum, ensuring a photon noise contribution under $\sim$1 m s^{-1} to the overall RV uncertainty.

The data for this project is publicly available in the ESO Advanced Data Products Archive (ADP[2]), and once retrieved, we perform a series of initial procedures aimed at obtaining a reliable estimate of the true nature of a periodic RV signal before the process of Keplerian modelling starts. This includes:

1. Determine a prominent modulation period of a dataset through Fourier/Least-Squares methods (refer to Chapter 4 for a theoretical outline);
2. Verify non-statistically significant correlations between the RV measurements and primary stellar magnetic activity tracers, such as the Cross-Correlation Function Bisector Span (CCF BIS) and CCF Full Width at Half-Maximum (CCF FWHM), which are estimated within the ESO/HARPS standard pipeline reduction, and also the $S-$index (Ca II H&K lines; $\lambda = 396.85$ and 393.37 nm), Hα ($\lambda = 656.46$ nm), the Na doublet ($\lambda = 589.59$ nm and $\lambda = 588.99$) and He I line ($\lambda = 587.56$ nm), which are estimated with the ACTIN Python package (Gomes da Silva et al. 2018; refer to Chapter 3 for an in-depth discussion about each magnetic activity indicator);

---

[2] `http://archive.eso.org/wdb/wdb/adp/`

3. Verify periodic modulations that may arise due to stellar magnetic activity with a period equal to the planet's orbital period by performing a Fourier/Least-Squares analysis on the magnetic activity tracers.

   If the planet's prominent period is not detected in the magnetic activity tracers, it can provide strong evidence for the planet's existence;

4. And lastly, the proposed planet's orbital period is tested via the Nyquist frequency sampling test to assure that it does not come from instrumental modulations (see Subsection 4.1.1).

Chapter 2

Exoplanets

In this chapter, we provide a brief overview of key insights and progress made in the field of exoplanetary systems research, ranging from theoretical concepts to factual detection and direct observations. The core principles governing Keplerian orbits are reviewed in Section 2.2.

2.1 Insights into Exoplanet Diversity and Formation

The dawn of new technologies and the rising of modern data analysis techniques led to the transition from just theoretical speculations to factual observations of exoplanets in the early 1900s. Observations of heavy elements like calcium in the van Maanen 2 star spectra (van Maanen, 1917) – a nowadays known white dwarf – may have been the first evidence of detecting exoplanets. It turns out that such heavy elements could not have been found in the stellar atmosphere. Due to the intense gravitational field of white dwarf stars, those elements would have been immediately wrecked, and the theory of polluted white dwarfs offers one plausible explanation for such phenomena; rocky material from exoplanets could be pulled and driven toward the star, whose intense gravitational field pulverises them into small dust particles. As it contains heavy elements from the ancient exoplanet, it would lead them to fall on the star's photosphere remaining for a relatively short period to be therefore spectroscopically analysed (Harrison et al., 2018; Putirka and Xu, 2021; Xu and Bonsor, 2021; Trierweiler et al., 2022).

Similar polluted spectroscopic behaviour was also recently observed in Sun-like stars, where it was found out by Spina et al. (2021) that Sun-like stars experience planetary engulfment events with a probability ranging 25%–33%, and that the fall of planets can change the surface chemical pattern of the central stars (see also Nagar et al. 2020 upon differences in the composition of binary systems in the presence of exoplanets). The elemental composition of stars can provide important insights into the formation and diversity of planetary systems: Stars hosting highly dynamic planetary systems, for instance, may have undergone more violent interactions during their formation, leading to a higher concentration of heavy elements in their atmosphere (Miller and Fortney, 2011; Benítez-Llambay et al., 2015; Meléndez et al., 2017). On the other hand, stars hosting planetary systems that resemble our Solar System may present a lower abundance of heavy elements, and perhaps, may be less likely to form planets with similar compositions (Guillot et al., 2006; Zhao et al., 2018).

The elemental composition of stars can provide important insights into the formation and diversity of planetary systems: Stars hosting highly dynamic planetary systems may have undergone more violent interactions during their formation, leading to a higher concentration of heavy elements in their atmosphere (Meléndez et al., 2009). On the other hand, stars hosting planetary systems that resemble our Solar System may present a lower abundance of heavy elements and perhaps may be less likely to form planets with similar compositions.

Exoplanets are in fact difficult to observe directly, unlike the many planets in our own Solar System. About 99% of the currently known exoplanets have been discovered without even analysing photons from the planet itself – a star can be dimmed (as in the transit method) or increased in apparent flux for a short period (microlensing), and also revolve around its equilibrium point out of our line-of-sight revealing changes in its *colour* (radial velocity) –[1].

The first extrasolar planet, γ Cephei Ab, was observed and reported in the late 1980s using RV data from a binary star system (Campbell et al., 1988). However, confirmation of its existence came years later due to the great uncertainties in the data collected at that time (Hatzes et al., 2003). Another planet, HD 114762 b, was also detected using RV data, but its large mass and high orbital eccentricity led to scepticism about its classification as a brown dwarf (Latham et al., 1989). The first discovered exoplanet was found orbiting a pulsar (PSR B1257+12; Wolszczan and Frail 1992), and numerous subsequent discoveries and observations of exoplanets provided insights into their nature and the diversity of exoplanetary systems, but one significant finding

---

[1] The NASA Exoplanet Archive (`https://exoplanetarchive.ipac.caltech.edu/`) has recorded a total of 5338 exoplanets as of May 1, 2023, using various detection methods. The most commonly used methods are Transit, which accounts for 3986 known exoplanets, and Radial Velocity, which detected 1036 exoplanets. Other detection methods include Microlensing (187), Direct Imaging (66), Pulsar Timing Variations (9), Eclipse Timing Variations (17), Orbital Brightness Modulation (9), Transit Timing Variations (25), Astrometry (2), and Disk Kinematics (1).

during this time was the detection of hot Jupiters – massive gas giants that orbit very close to their host stars –, such as 51 Peg b (Mayor and Queloz, 1995). These planets were unexpected to exist based on our understanding of the formation and evolution of the Solar System, and their discovery paved the way for the emergence of theories of planetary migration. Observations of hot Jupiters also provided important information about the atmospheres and compositions of exoplanets, such as the detection of water vapour, carbon dioxide, methane, and various metal and alkali atoms (Sudarsky et al. 2003; Fortney et al. 2021; Quanz et al. 2021, and references therein). This variation suggests that hot Jupiters may have formed through different mechanisms, such as *in-situ* formation (Batygin et al., 2016; Poon et al., 2021), or migration from farther out of their planetary systems after or before the protoplanetary disk dissipates (Dawson and Johnson, 2018; Ito and Ohtsuka, 2019; Reggiani et al., 2022; Wu et al., 2023).

The process of both planet and star formation is complex and depends on various factors such as initial conditions, composition, density, and gravitational interactions. Stars form from the gravitational collapse of molecular clouds made up of primarily gas H and He. These clouds collapse in the centre first due to their high density, fragmenting into small regions of around 0.1 pc in size, containing masses of $10 - 50$ $\mathrm{M}_{\odot}$ (Cameron, 1978). As matter moves towards the nearly formed star, pressure and temperature increase, causing it to emit infrared radiation. The continuous transfer of matter from the molecular cloud to the protostar increases its mass, leading to thermonuclear reactions in the core, which balance the external gravitational pressure and keep it in hydrostatic equilibrium (Nakamoto and Nakagawa, 1994). The conservation of angular momentum during the collapse results in the formation of a protoplanetary disc that can evolve into a planetary system via gravitational instabilities (Kratter and Lodato, 2016) or the accretion model (Safronov, 1972; Lissauer, 1993). The latter model involves the coagulation of small solid grains of dust and ice into planetary embryos and later protoplanets (Morbidelli and Raymond, 2016; Raymond and Morbidelli, 2022). When the thermal velocity of gas drops below a planet's escape velocity, it accumulates around the planet until its gravity becomes strong enough to hold it, forming a gaseous atmosphere. Observations, such as the increased rate of giant planet formation at higher stellar metallicities (Fischer and Valenti, 2005; Ghezzi et al., 2010) and the composition similarities among the terrestrial planets in the Solar System (Guillot, 2005), support this model. Additionally, the high-mass nuclei in Jupiter and Saturn could also be a result of this process (Militzer et al., 2008).

Within the conventional picture, rocky and gas giants begin to form from extremely abundant micrometre-sized dust particles in protoplanetary discs (Haisch et al., 2001), which eventually grow by agglomeration until reaching the metre-sized barrier (Weidenschilling, 1977; Birnstiel et al., 2012). When these particles are sufficiently concentrated relative to the gas around them, gravitational instabilities trigger drifting self-gravity limiting particle clusters, forming kilometre-sized objects known as planetesimals (Youdin and Goodman, 2005; Johansen et al., 2009; Yang et al., 2018). Subsequently, these planetesimals grow by accreting other planetesimals (e.g., Lambrechts and Johansen 2014), or even smaller particles of a millimetre or centimetre sizes known as pebbles (Kokubo and Ida, 2000). The more planetesimals are formed, their orbits became dynamically excited, inducing random velocities between nearby objects, thus reducing the efficiency of matter accretion from the protoplanetary disc (e.g., Leinhardt et al. 2009). When planetesimals reach a mass of approximately lunar mass, they are called planetary embryos (Johansen and Lambrechts, 2017). Furthermore, when these objects reach the pebbles' isolation mass (Bitsch et al., 2018), not only does the growth of the planetary embryo cease, but the flow of pebbles toward the central star is also reduced or prevented, therefore depriving objects in inner orbits to grow. Over time, planetesimal accretion results in larger and larger objects. In the outer region, if a rocky body reaches a mass of around 10 $M_\oplus$, it can accumulate a gas envelope, forming a giant planet on a timescale of less than 10 Myr before the protoplanetary disc dissipates. In the inner region, planetesimal accretion results in the formation of rocky planets on a longer timescale, on the order of 50 to 100 Myr.

It might be reasonable to assume that every star has at least one planet in orbit, but the reality is that many stars have multiple planets in orbit around them (e.g. Beaugé et al. 2012; Akeson et al. 2013; Sousa et al. 2021). Exoplanets are ubiquitous, but a tiny fraction of them resemble the planets found in our own Solar System (see Figure 2.1). To put this into perspective, if we were to observe the Sun as an exoplanet-host star, we would only be able to detect Jupiter using current instrumentation and temporal observation baseline (Morbidelli and Raymond, 2016; Raymond et al., 2020), which places the Solar System as atypical at around a 1% level (Rowan et al., 2016). Even the existence of Jupiter could be met with scepticism because its modulations would be of a nearly-perfect sinusoidal shape ($e = 0.048$), which is consistent with cycles of stellar activity (see Section 3.3).

One of the most striking features in exoplanet demographics is the clear division between small, rocky planets and large, gaseous planets. This divide is known as the *radius gap*, occurring at around $1.7 - 1.9\ R_{\oplus}$, or $5 - 10\ M_{\oplus}$, depending on the composition of the planet (Owen and Wu, 2013; Lopez and Fortney, 2014; Fulton et al., 2017; Martinez et al., 2019; Venturini et al., 2020; Loaiza-Tacuri et al., 2023).

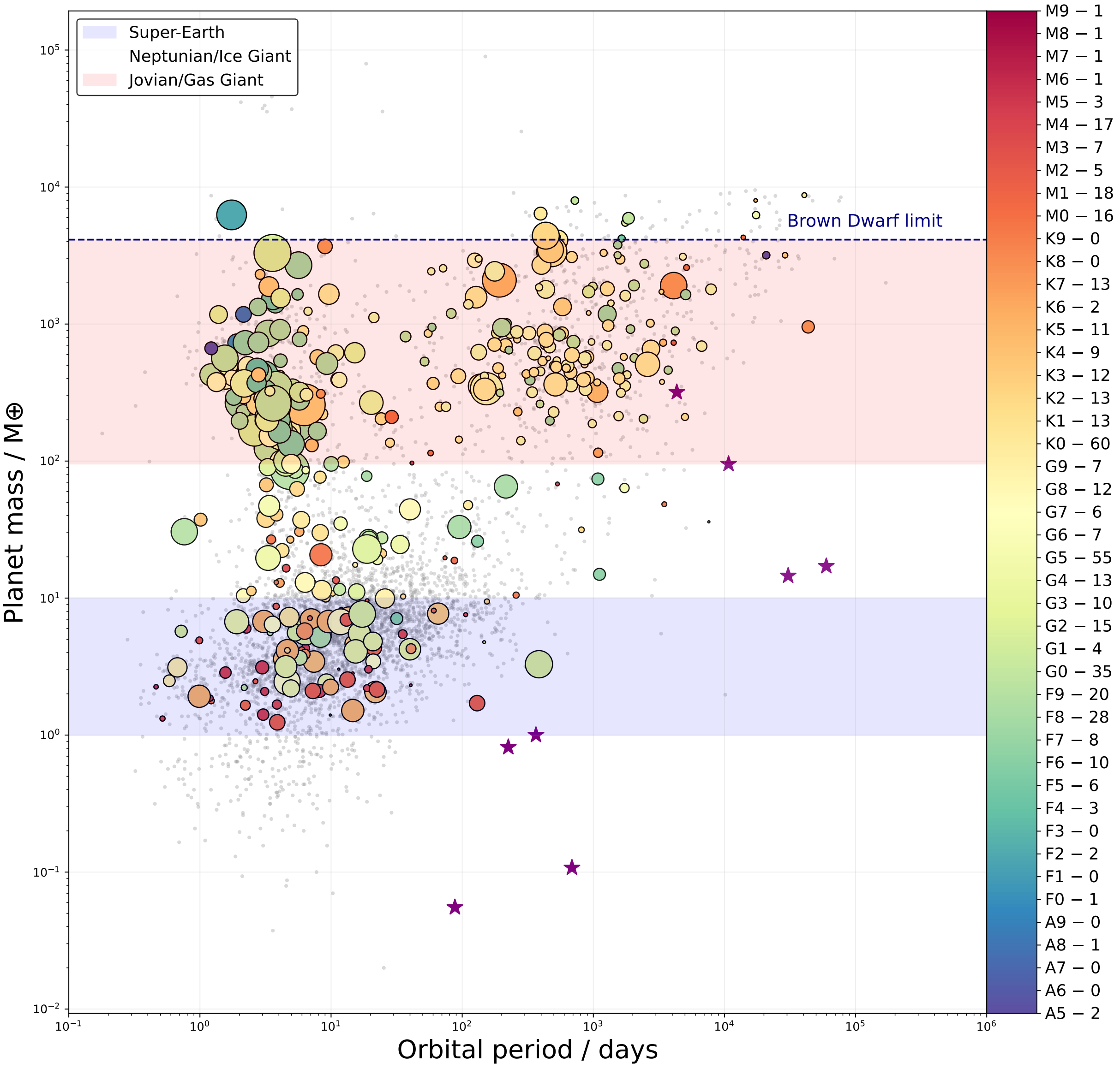

Figure 2.1: Masses (y-axis in Earth units) and orbital period (x-axis in days) of known exoplanets (as of June 20, 2023; data from the NASA Exoplanet Archive). Shaded blue, white, and red regions indicate distinct planetary types by mass, respectively, Super-Earths, Ice Giants and Gas Giants. The spectral types of host stars, from A–type to M–type, are colour-coded if known, and planets with unknown host stars' spectral types are shown on a grey-scale background. The Solar System planets are shown as purple-coloured stars for reference. Circle size represents the system's distance from us (ranging from 1.30 pc to 8 kpc), with larger circles indicating closer distances. The blue horizontal line represents the threshold for objects more likely to be brown dwarfs, whose masses are greater than about 13 Jupiter masses.

It was recently proposed that ice giant planets (M $\in [10, 95]$ M⊕) may be more prevalent than their gas giant counterparts. For instance, planets with radii ranging $1.5 - 4$ $R\oplus$ were found to be more common from the analysis of *Kepler* data (Fressin et al., 2013). These studies are consistent with theoretical models of planet formation, which suggest that the ice giants may form more easily in the outer regions of protoplanetary disks where there is a greater abundance of ice (e.g., Dodson-Robinson et al. 2009). Furthermore, it was also recently found that main-sequence stars in the Milky Way host planets with larger sizes and masses than those of Earth and Neptune, usually called Super-Earths ($M \in [1, 10]$ $M\oplus$) and Mini-Neptunes ($M \in [10, 20]$ $M\oplus$), with orbital periods shorter than 100 days (Howard et al., 2012; Petigura et al., 2013). These kinds of planets appear to be common, with occurrence rates estimated from 30% to 50% of all stars orbited by at least one such planet. It has been also proposed, for instance, that super-Earths and Mini-Neptunes may represent the building blocks of larger, more massive planets like Jupiter, and that their presence may be a key ingredient in the planet formation process (e.g., Hori and Ikoma 2011; Morbidelli et al. 2012). On the other hand, other studies have suggested that these planets may have formed through a different mechanism, such as *in-situ* accretion of solids or gravitational instability in the disk (e.g., Boss 2006; Chiang and Laughlin 2013).

The majority of low-mass planets tend to be found around M-dwarf stars, especially due to their clear signal and constraints of gas giant planets to form around such a type of star, that albeit rare, are not impossible (e.g., Johnson et al. 2012; Hartman et al. 2015; Bayliss et al. 2018; Gan et al. 2022, and references therein). Giant planets ($M \in [0.30, 13]$ $M_{\rm Jup}$) tend to be found around FG-type stars, and mostly orbiting stars in the vicinity of the Solar System. Demographically, giant planets exist around approximately 10% of Sun-like stars (Cumming et al., 2008; Mayor et al., 2011), and their occurrence rate appears to be strongly dependent on the metallicity of the host star, with a higher occurrence rate observed for stars with higher levels of metals (Santos et al., 2003; Fischer and Valenti, 2005b,a; Ghezzi et al., 2010, 2018; Dressing and Charbonneau, 2015); plus, the presence of giant planets can have a significant impact on the architecture and orbital stability of planetary systems, configuring a key factor in determining the habitability of rocky planets in a planetary system (c.f. Section 1.1; Raymond et al. 2016; Pu et al. 2017; Barnes et al. 2017).

2.2 On the Lagrangian Formulation of Keplerian Orbits

Late in the 17th century, Newton (1687) demonstrated that a planet's orbit around a central mass comprises an elliptical path as one of the consequences of the gravitational law attraction dependency on the inverse-square of the distance r between them[2]. Restricted to conservative forces of central nature, with the system's potential energy expressed by $\phi(r)$, an angle coordinate θ representing a rotation about a fixed axis must be cyclic (Landau and Lifshitz, 1975; Goldstein et al., 2002). For this configuration, no net torque remains on the system under any perturbation, therefore, the system's angular momentum, $\mathbf{L} = \mathbf{r} \ \times \ p$, is conserved. Additionally to the radial and polar coordinates, the motion of a planet with mass m revolving a central star/object of mass M should be also described by the zenith angle φ, however, by choosing the polar axis $\mathbf{z} = z\hat{\mathbf{k}}$ in the same direction of $\mathbf{L}$, φ takes a value of $\pi/2$, so it can be ignored in the discussion to follow. The system's Lagrangian is expressed according to

$$\mathcal{L} = \frac{1}{2}m\left(r^2 - r\dot{\theta}^2\right) - \phi(r), \tag{2.1}$$

from which we obtain the equation of motion

$$m\ddot{r} - mr\dot{\theta}^2 = f(r)$$

straight from Euler-Lagrange's equation of r. Associated with the azimuth angle θ, a determination of its canonical momentum yields one of our first integrals of interest,

---

[2] *Propositio* XI in *Philosophiæ naturalis principia mathematica*; *Liber Primus. De motu corporum* (On the motion of bodies); *Revolvatur corpus in ellipsi* (If a body revolves in an ellipse); *requiritur lex vis centripetæ tendentis ad umbilicum ellipseos* (it is required to find the law of the centripetal force tending to the focus of the ellipse). Interestingly, early in a manuscript addressed to E. Halley in 1689, named *De motu corporum in gyrum* (On the motion of bodies in an orbit), Newton refers to elliptical orbits *as Kepler supposed*, albeit nowhere in *Liber Primus*, does he mention Kepler's findings. Originally, *Gyrant ergo planetæ majores in ellipsibus habentibus umbilicum in centro solis* (Therefore, the major planets revolve in ellipses, having the focus in the centre of the Sun), *et radijs ad Solem ductis describunt areas temporibus proportionales* (and the rays directed to the Sun describes areas proportional to the times), *omninò ut supposuit Keplerus* (just as Kepler supposed).

$$mr^2\dot{\theta} = \ell,$$

with ℓ representing the constant magnitude of the angular momentum $\|\mathbf{L}\|$, which in turn is associated to $f(r)$ to obtain a second-order differential equation solely dependent of r as

$$m\ddot{r} - \left(\frac{\ell^2}{mr^3}\right) = f(r).$$

From the former,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(-\frac{1}{2}m\dot{r}^2 + \frac{\ell^2}{2mr^2} + \phi(r)\right) = 0,$$

and since we are dealing with conservative forces, it follows, from the general energy conservation theorem, that one of the constants of motion can be expressed as

$$\varepsilon = \frac{1}{2}m\dot{r}^2 + \frac{\ell^2}{2mr^2} + \phi(r). \tag{2.2}$$

Working with equation (2.2) for $\dot{r}$, we obtain the differential

$$\mathrm{d}t = \mathrm{d}r\sqrt{\frac{2}{m}\left(\varepsilon - \phi(r) - \ell^2/2mr^2\right)},$$

which we shall later express as time-independent, rather than azimuth angle-dependent, making use of the canonical momentum associated with the polar angle. Given a particular force law between two massive bodies, the orbit equation is described in terms of the constant system's energy ε and the angular momentum magnitude ℓ via the differential

$$\mathrm{d}\theta = \ell \,\mathrm{d}r \,/mr^2 \sqrt{\frac{2}{m}\,(\varepsilon - \phi(r) - l^2/2mr^2)},$$

or conveniently in its integral form under substitution $r^{-1} \ \to \ \xi$, plus considering a generic power-law for the system's potential energy $\phi(r) = \kappa r^{n+1}$, in such a way for the gravitational framework $n = -2$ and $\kappa = -GmM$.

$$\theta = \theta_0 + \int_{\xi_0}^{\xi} \frac{\mathrm{d}\xi}{\sqrt{\frac{2m\varepsilon}{\ell^2} - \frac{2m\kappa\xi}{\ell^2} - \xi^2}}. \tag{2.3}$$

Due to its cyclic property, the later integral is taken as indefinite, and denoting θ_0 by ν as it commonly appears in the literature yields a standard form with fairly known solution[3]

$$\theta = \nu - \arccos\left(\frac{\frac{\ell^2\xi}{m\kappa} - 1}{\sqrt{1 + \frac{2\varepsilon\ell^2}{m\kappa^2}}}\right). \tag{2.4}$$

Back to the radial variable r, the orbit equation is found to be

$$\frac{1}{r} = \frac{m\kappa}{\ell^2}\left[1 + \sqrt{1 + \frac{2\varepsilon\ell^2}{m\kappa^2}}\cos(\theta - \nu)\right]. \tag{2.5}$$

In comparison to the general form of the cone equation in polar spherical coordinates, read out as $r^{-1} \equiv \alpha\,[1 + e\cos(\theta - \nu)]$, with $\alpha \in \mathbb{R}$, we found out that ν represents the planet's orbit periapse longitude, or in other words, the angle location of the closest approach between two orbiting bodies concerning a reference direction. It is defined as the true anomaly the quantity

---

[3] *c.f.* equation 2.261 in Gradshteyn et al. (2007)

$$\int \frac{\mathrm{d}x}{\sqrt{\alpha + \beta x + \gamma x^2}} = \frac{1}{\sqrt{-\gamma}} \arccos\left(\frac{-\beta + 2\gamma x}{\sqrt{\beta^2 - 4\gamma\alpha}}\right).$$

$f \equiv \theta - \nu$, a time-dependent parameter, and the orbit's eccentricity becomes equal by definition to $e \triangleq \sqrt{1 + (2\varepsilon\ell^2/m\kappa^2)}$.

Bounded motion can manifest as *negative energies* in the context of the one-dimensional potential for inverse-square laws of force, being defined for elliptical orbits the major axis as a straight line passing through the ellipse's foci connecting the *apoapsis* (the point of greatest separation between two bodies) and the *periapse* of the orbit, with one half of it called the orbital semi-major axis a.

On these apsidal distances, by definition, the radial velocity is null concerning the mean, and it follows from (2.2) that the two distances form a set of roots of

$$r^2 + (\kappa/\varepsilon)\, r - \left(\ell^2/2m\varepsilon\right) = 0.$$

Therefore, shall we write the semi-major axis as $a = -\kappa/2\varepsilon$, and subsequently, the orbit's eccentricity as a function of the semi-major axis takes the form $e = \sqrt{1 - (\ell^2/m\kappa a)}$, from which we obtain the distance of the planet's orbital path around a star in terms of its orientation following (see Fig. 2.2)

$$r = \frac{a\,(1 - e^2)}{1 + e\cos f}. \tag{2.6}$$

Consider the plane $\mathrm{R} = (X, Y, Z)^\intercal$, with Z pointing towards the line of sight of the external observer. It is tilted by i (called orbital inclination) concerning the plane $r = (x, y, z)^\intercal$ – where by convention, $i = 0$ denotes a face-on system, while it is called edge-in systems those with $i = \pi/2$, or values approaching it, with the former, albeit not directly measured by radial velocity observations, one of the requirements for Doppler spectral shifts to be observed –. To consider the position r from the orbital plane to the observer's, it is required a few rotations: (a) about the $\mathrm{z}-$axis so that the $\mathrm{x}-$axis coincides with the line-of-nodes at the ascending node – the point at which the planet passes from below the plane-of-the-sky to above; on the other hand, it is called the descending node; physically, *redshift* becomes *blueshift*, and vice-versa –, generating a $\mathrm{x}'-$axis; (b) a rotation about the $\mathrm{x}'-$axis, so that the planes are now coincident, and (c) a rotation

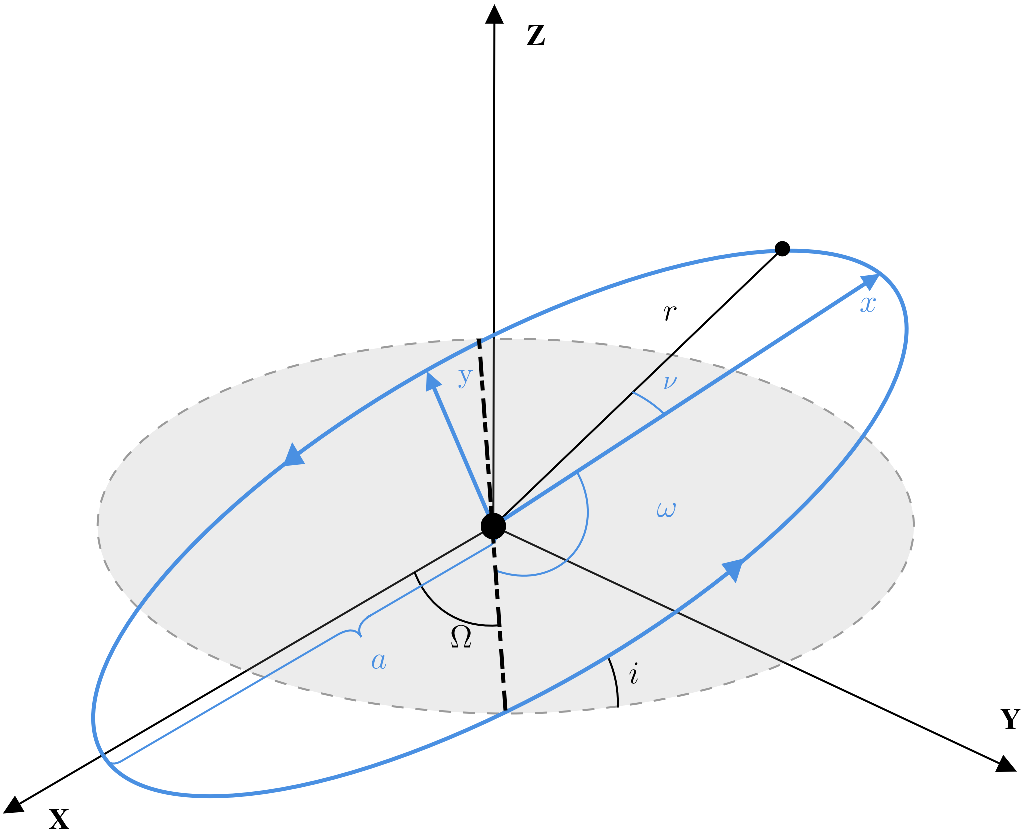

Figure 2.2: Orbital elements of a single planet around a star in the centre of the XYZ-frame. These elements include the true anomaly (ν), which represents the angle between the current position of the massive body and its position at the periastron, as well as the semi-major axis (a), which is the distance between the centre of mass and the farthest point on the body's orbit. Moreover, it includes the orbital inclination (i), which described the angle between the orbital plane (solid blue line) and the equatorial plane (dashed grey line), and the right ascension of the ascending node (Ω), which is the angle between the vernal equinox and the point where the orbit crosses the equatorial plane from south to north – the "bottom" of the XY-plane to it's "top"). Additionally, the argument of periapsis (ω) represents the angle between the line of nodes (a line connecting the ascending to the descending node; black dashed line in the graph) and the point of closest approach (periastron). In the absence of external perturbations, such as those caused by massive outer migrating companions or stellar flybys, these orbital parameters might remain constant (e.g., Dawson and Murray-Clay 2013; Rodet et al. 2021, and references therein). The $Z-$axis is oriented towards the line of sight of the external observer, and the distance between the massive body and the centre of mass is denoted by r (refer to the orbit equation 2.6).

about the $\mathrm{z}''-$axis generated in the previous step.

From the general operation for matrix rotation $\mathrm{R} = \left(\Gamma_{\mathrm{z}}^{\,\Omega}\,\Gamma_{\mathrm{x}}^{\,i}\,\Gamma_{\mathrm{z}}^{\,\omega}\right)\cdot r$, with the rotation matrices $\Gamma_{\mathrm{(axis)}}^{\,\mathrm{(angle)}}$ constructed following kinematics of rigid body motion theory (*c.f.* Goldstein et al. 2002), the position of a planet is found to be

$$Z = r \sin(f + \omega) \sin i. \tag{2.7}$$

Note that X and Y orientations are superfluous as we are interested in the radial velocity along the observer's line of sight, which is straightforwardly determined by evaluating the time-dependent differential $\partial_t Z$.

Radial velocity measurements may assume both positive and negative values, depending on the star's direction of motion. Historically, positive values have been traditionally associated with the redshift of a star. As an object moves away from us, its emitted wavelengths shift towards longer values, resulting in a perceived reddening. On the other hand, negative values are attributed to a blueshift in the star's spectra. A blueshift occurs when an object moves towards us, causing the wavelengths to shift towards shorter values (Murray and Correia, 2010).

$$V_r := \partial_t Z = \sin i \left[\dot{r} \sin(f + \omega) + r \dot{f} \cos(f + \omega)\right]. \tag{2.8}$$

From the orbit equation 2.6 we obtain $\dot{r}$, and making use of Kepler's second law[4],

---

[4] Originally published in *Astronomia nova* (Kepler, 1609), the distance law, or area law, is presented both in *Capvt* XXII (pp. 165–167). *Virtuem quæ Planetam movet in circulum* (The force that moves a planet circularly), *attenuari cum disceffu a fonte* (weakens with distance from the source) and in *Capvt* LIX (pp. 285) *Demonstratio, quod orbita MARTIS, librati in diametro epicycli, fiat perfecta ellipsis* (Demonstration that the orbit of Mars, balanced on the diameter of the epicycle, is a perfect ellipse), however its modern form is solely presented in *Epitome Astronomiæ Copernicanæ* (Kepler, 1618) *Liber Qvintvs* (pp. 668); *Dictum quidem est in superioribus* (It has indeed said in the foregoing), *divisâ orbitâ in particulas minutissimias æquales* (if the orbit is divided into smallest equal parts), *accelerete iis moras planetæper eas, in proportione interuallorum inter eas & Solem* (the times of the planet in them increase in the ratio of the distance between them and the Sun). Formally, it states that planets move at different velocities depending on their orbital positions, and algebraically, $\dot{S} = r\dot{v}/2$, with S representing the area of the ellipse formed by the planet's orbit.

$$V_r = \left(\frac{2\pi a \sin i}{P\sqrt{1-e^2}}\right)\left[\cos(f+\omega) + e\cos(f)\right]. \tag{2.9}$$

The multiplying term in front of the brackets in equation (2.8) is called the radial velocity semi-amplitude, or Doppler amplitude, denoted by K for historical reasons. Throughout the analysis of time-series radial velocity data, we normally do not infer the exoplanet's semi-major axis, however, the square of the planet's orbital period is proportional to the cube length of the semi-major axis of its orbit[5].

It becomes opportune to express K in terms of an exoplanet's orbital period P, which is possible to be estimated from Fourier analysis (*c.f.* Section 4.1), and also by recognising that the semi-major axis of the star around the system's CM is the module of the vector pointing from this one to the star itself. In practice, considering a stationary CM, and by representing the star's distance and the planet's distance to it by a_1 and a_2, whose masses are M and m, respectively, the equality $Ma_1 = ma_2$ holds. Consequently, the Doppler amplitude is found to be

$$K = \frac{2\pi \sin i}{P\sqrt{1-e^2}}\left[\frac{m}{M}\left(\frac{G(m+M)P^2}{4\pi^2}\right)^{\frac{1}{3}}\right]. \tag{2.11}$$

For comparison, $K \sim 12.4$ m s^{-1} for a Jupiter-like planet around a Sun-like star at Jupiter's orbit, while for a planet similar to Earth in mass and orbital paths, $K \sim 0.10$ m s^{-1}. The success of the radial velocity method, which led to the first detection of an exoplanet orbiting a main-sequence star back in the mid-1990 decade is not due to observing variations in the star's position,

---

[5] Published in *Harmonices Mvndi Libri V* (Kepler, 1619) (pp. 189–190); *Sed res est certissima extactissimaque quod proportio qua est inter binorum quorumcunque Planetarum tempora periodica, sit præcise sesquialtera proportionis mediarum distantiarum, id est Orbium ipsorum* (But it is absolutely certain and exact that the proportion between the periodic times of any two planets is precisely the sesquialternate proportion [3:2] of their mean distances, that is, of the actual spheres). Algebraically, the ratio of the planet's semi-major axis cubed and the planet's orbital period squared is constant, which in turn is proportional to the sum of the masses of the planet the star

$$\frac{a^3}{P^2} = \frac{G}{4\pi^2}(M+m). \tag{2.10}$$

but rather to noting changes in its velocity by looking at oscillations in the wavelength of the star's spectral lines caused by its orbit around the CM.

Our starting point in this section was to determine an expression for r in terms of θ by eliminating a time-dependency (*viz.* equation 2.6); and rotation properties yields to the radial velocity equation along the external observer's line-of-sight (*viz.* equation 2.8). We now aim at determining an expression for V_r in terms of r, which is the essence of Kepler's problem.

Back to our time-differential $\mathrm{d}t$ at the beginning of this section, we might re-write as

$$t = \sqrt{m/2} \int_{r_0}^{r} \mathrm{d}r \,/ \sqrt{\varepsilon - (\kappa/r) - (\ell^2/2mr^2)},$$

that can be integrated by making use of a new auxiliary variable E, called eccentric anomaly – which, along with the mean and eccentric anomaly, defines the position of a body that is moving along an elliptic Keplerian orbit – through a substitution $r = a\,(1 - e\cos E)$. In comparison to equation (2.5), we note that E completes a full revolution as θ goes from 0 to 2π, i.e., the most extreme point of a planetary orbit occurs at $E = 0$, while the point of closest approach occurs at $E = \pi$. In terms of the orbit's semi-major axis and the orbital eccentricity, t can be re-written as

$$t = \sqrt{(ma^3)/\kappa} \int_{0}^{E} (1 - e\cos E)\,\mathrm{d}E\,.$$

We introduce the frequency of revolution $\omega = (2\pi/P) = \sqrt{\kappa/(ma^3)}$, with the right-hand-side obtained via Kepler's third law. Evaluating the time-dependent differential of r in equation 2.6, with the aid of Kepler's second law, we obtain

$$\dot{E} = \frac{\omega}{1 - e\cos E}, \tag{2.12}$$

whose general solution assumes the form down below by taking t_0 as a constant of integration, and recognising the initial condition $E(t = t_0) = 0$.

$$\omega(t - t_0) = E - e \sin E. \tag{2.13}$$

The expression (2.13) is known as Kepler's equation, with the left-hand-side customarily denoted by $\mathcal{M}$, called the mean anomaly, also ranging from 0 to 2π, with t_0 called the time of periastron passage, or time of inferior conjunction. Noteworthy, $\mathcal{M} = f = 0$ with $t = t_0$, or $t = t_0 + P$ (at the periastron passage), and $\mathcal{M} = f = \pi$ with $t = t_0$, or $t = t_0 + \pi/2$ (at the apastron passage).

On a quest to the planet's orbital position over time, equation (2.13) can be first inverted to obtain an expression for E. Note, by comparing equation (2.6) to $r = a\,(1 - e \cos E)$, we get

$$1 + e \cos f = (1 - e^2)/(1 + e \cos E),$$

which can be re-written after a few algebraic manipulations to obtain an expression involving the orbital true anomaly;

$$\tan\left(\frac{f}{2}\right) = \sqrt{\left(\frac{1+e}{1-e}\right)} \tan\left(\frac{E}{2}\right). \tag{2.14}$$

The description presented above took into account a single two-body problem, while for the case of multi-planetary systems, the Keplerian orbits must be summed together, and following Fulton et al. (2018), the total RV motion of the star due to all companions is found to be

$$V_r = \sum_{k=1}^{N_{\text{pl.}}} V_{\text{k,r}} + \gamma + \dot{\gamma}(t - t_0) + \ddot{\gamma}(t - t_0)^2 + \mathcal{O}(\partial_n \gamma; \Delta t^n).,\tag{2.15}$$

where it is considered perturbation terms in the analysis with orbital periods much longer than the observational baseline (*c.f.* Section 4.1), represented by $\dot{\gamma}$ – the radial velocity slope –, and $\ddot{\gamma}$ – the radial velocity curvature –, where also t_0 representing an arbitrary abscissa epoch. $\mathcal{O}(\partial_n\gamma; \Delta t^n)$ in this equation represents additional higher-order perturbation contributions to the radial velocity. We incorporated a Keplerian model using the Python package framework RADVEL[6] (Fulton et al., 2018), which describes the motion of an orbiting object around a central star assuming that the object moves in an elliptical path, with its velocity and position determined by six orbital parameters (see Fig. 2.2 and Chapter 5): a, e, i, Ω, ω, and t_0. RADVEL fits the Keplerian model using a basis of five orbital elements: the orbital period P, time of inferior conjunction T_c, orbital eccentricity e, Doppler-induced semi-amplitude K, and periastron argument ω. Synthesising radial velocities involves solving equations 2.8, 2.13, and 2.14. Our commonly used basis in the Keplerian fits presented in Chapter 5 is *P, T_c, $\sqrt{e}\cos\omega$, $\sqrt{e}\sin\omega$, and* K, because this imposes flat priors on all of the orbital elements, avoids biasing $K > 0$, and helps to speed MCMC convergence.

2.2.1 The Relativistic Doppler Effect

Minor changes in the *colour* of stars as long as they are in motion through the external observer's line of sight may be perceptible (Doppler, 1842). Shall we define initially a wave vector in a four-dimensional space, making use of Einstein's standard notation, according to $\kappa^\mu \equiv (\omega/c, \kappa)$, where the frequency of a plane wave is represented by ω, and κ is a three-dimensional wave-vector. Under Lorentz group transformations, the phase of the wave as the invariant quantity $\kappa_\mu x^\mu := \phi - \omega t - \langle \kappa, \mathbf{r} \rangle$, with $\langle \cdot, \cdot \rangle$ denoting the inner product between κ and the coordinates four-vector of the source in Σ, with usual metric. Let v be the velocity of a source in its inertial frame Σ_0, being at rest relative to Σ (Landau and Lifshitz, 1975). Under the general formula of four-vectors transformations,

$$\kappa^{(0)0} = \frac{\kappa^0 - \beta\kappa^1}{\sqrt{1-\beta^2}} \triangleq \gamma\left(\kappa^0 - \beta\kappa^1\right), \tag{2.16}$$

introducing in the right-hand-side the terms $\gamma = \left(1-\beta^2\right)^{-1/2}$ – the Lorentz factor –, and $\beta = \frac{v}{c}$, the relative velocity of a source escalated by the speed of light in vacuum.

---

[6] `https://radvel.readthedocs.io/`

It followed from the special theory of relativity that changes in frequency ω of a wave emitted by a source moving concerning the observer, as compared to the "true" frequency ω_0 of the same source moving in the reference frame Σ_0 – in which is at rest – may be perceptible, and vice-versa (Einstein, 1905). The frequency and wavelength of a wave are not invariant but instead depend on the motion of the observer and the source relative to one another. The frequency of a wave emitted by a moving source appears to be shifted to higher or lower frequencies as described by the relativistic Doppler effect.

From κ^μ and $\kappa^{(0)0}$, with the aid of the defined terms κ^0 and κ^1, by also denoting η as the angle between the direction of emission of the wave and the direction of motion of the source that arises from the aforementioned inner product, we obtain an expression for the "new" detected frequency of a wave in the source frame – the relativistic Doppler effect – according to

$$\omega_0 = \gamma\omega\left[1 - \beta\cos\eta\right], \tag{2.17}$$

or similarly, the frequency of a wave in the receiver/observer frame,

$$\omega = \frac{\omega_0}{\gamma\left[1 + \beta\cos\eta\right]}. \tag{2.18}$$

To obtain the maximum effect of the Doppler effect, a specific alignment is required between the velocity vector and the line-of-sight of the external observer. The angle between the direction of the wave and the direction of motion of the source, represented by the cosine of the angle η, is responsible for this alignment. In the case of detecting exoplanetary systems, if the system is viewed face-on, it is impossible to determine if gravitational perturbations are due to an orbiting planet, as the star's RV relative to Earth will not change. However, it is still possible to observe RV variations of a star, which may be due to other phenomena such as stellar spots, faculæ, plages, or flares, as discussed in Hatzes (1999), Gray (1999), and Reiners (2009), and Section 3.3.

Furthermore, it is important to note that the greater the velocity of the source (such as a star) relative to a rest one (the external observer), the larger the difference between the shifted frequency ω and the nominal frequency ω_0. This is due to the $\gamma-$ and $\beta-$dependencies and the assumption that $\omega = c/\lambda$. As a result, if a planet is large enough to cause significant gravitational perturbations in the star's orbit around the system's centre of mass, its (radial) velocity will be significantly altered. This could result in a reddening or bluishness observed at a certain wavelength λ of the spectrum.

Chapter 3

The State-of-Art in High Precision Spectroscopy

Échelle spectroscopy involves using a high-resolution spectrometer to split the starlight into its wavelength components, which can then be analysed to determine the chemical composition, fundamental parameters of the star, and also reveal orbiting planets by analysing the Doppler shifts of spectral lines. The analysis of spectral data is, however, not always straightforward. Stellar oscillations and magnetic activity can create modulations in time-series data that can be mistaken for planetary signals, and comprehending the underlying physics behind these effects is crucial for accurately detecting and characterising exoplanets as discussed in the following.

3.1 The HARPS Échelle Spectrograph

A spectrograph is an instrument that breaks light into its component wavelengths. The instrument consists of a *diffraction grating*, which is responsible for this splitting, and a *detector*, which records the resulting spectrum. The High Accuracy Radial Velocity Planet Searcher (HARPS; Mayor et al. 2003) covers the optical range ($\lambda \in [378.1, 691.2]$ nm; with a gap between 530 nm and 560 nm) with high resolving power ($R \sim 115\,000$; the ability to separate two closely-spaced spectral lines; $R = \lambda/\Delta\lambda$). The instrument is mounted on the European Southern Observatory's (ESO) 3.6-meter telescope at La Silla Observatory in Chile, and it is installed in a pressure– and temperature–controlled enclosure, which allows a high instrumental precision on the order of 1 m s^{-1}. It uses a single-mode optical fibre to collect the light from a target star and feed it to the spectrograph, and for each spectrum obtained, the instrument simultaneously observes a thorium-argon ($^{90}\mathtt{Th}$-$^{36}\mathtt{Ar}$) hollow-cathode lamp as a reference source for calibration. Alternatively, a Fabry-Perot etalon can be used as a standard calibration source for the simultaneous reference method. The data are processed in near real-time with the HARPS ESO pipeline through cross-correlation (c.f. Section 3.2) with a G2V spectral template for each star observed (Pepe et al., 2002) within the STPS observing programme.

Light coming from an astronomical source reaches a slit and diverges. In classical spectrographs, a collimator that matches the same focal ratio of the telescope is used to parallelize these diverging beams to a wavelength-dispersing instrument. In counterpoint to these classical spectrographs, échelle gratings have high dispersion orders coupled to a cross-disperser, which directs the incident light into perpendicular directions splitting the spectral orders (Hatzes, 2019). Similarly to other diffraction gratings, an échelle grating is composed of several slits that are closely spaced at a width comparable to the wavelength of the diffracted light, however, the blaze – a design feature of a diffraction grating that maximises the amount of light that is diffracted in a particular spectral order or wavelength range – is optimised for multiple overlapping high orders (Harwit, 2003; Palmer, 2006).

RVs of a star are measured by determining the Doppler shift of its spectral lines (refer to Subsection 2.2.1 for a theoretical outline). If a star moves towards or away from an observer, the wavelengths of its spectral lines will be shifted towards the blue or red end of the spectrum, respectively. This spectral shift is proportional to the star's RV, given by

$$\frac{\Delta\lambda}{\lambda} = \frac{v}{c}. \tag{3.1}$$

where $\Delta\lambda$ is the shift in the wavelength of the spectral line, λ is the rest wavelength of the line, v is the RV of the star, and c is the speed of light in vacuum.

3.2 Cross-Correlation Function (CCF) and Wavelength Calibration

The Cross-Correlation Function (CCF) is a standard procedure for signal processing, which is also known as the sliding inner product as it measures how similar two series are by determining the displacement of one to the other. The CCF is similar to a convolution of two functions, and it is often used to estimate RVs from a spectrum by comparing it to a known reference spectrum (Valenti and Fischer 2005; Lovis and Fischer 2010; alternatively, see also the line-by-line algorithm presented by Artigau et al. 2022).

To define the CCF, we consider a stellar spectrum denoted by $\mathcal{S}(x)$ and a template spectrum, called a mask or noise-free template, denoted by $\mathcal{T}(x)$. The CCF is given by the integral or sum of the product of the complex conjugate of $\mathcal{S}(x)$ and $\mathcal{T}(x+\delta x)$, where δx is the lag of the CCF,

$$(\mathcal{S} * \mathcal{T})\,(x) = \int_{-\infty}^{+\infty} \mathcal{S}^{\dagger}(x) \cdot \mathcal{T}(x+\delta x)\,\mathrm{d}x = \sum_{x=1}^{N} \mathcal{S}(x) \cdot \mathcal{T}(x+\delta x)\,\mathrm{d}x\,, \tag{3.2}$$

where the left-hand-side is obtained since a stellar spectrum is physically a discretely sampled real function ($\mathcal{S}, \mathcal{T} \in \mathbb{C}^{\mathrm{N}}$), the number of pixels in a CCD is finite, and δx is called the lag of the CCF. Note that the CCF differs from the convolution by the sign of the argument in the second function (e.g., Bracewell 2000, $(f * g) := \int_{-\infty}^{+\infty} f(\tau) g(t-\tau)\,\mathrm{d}\tau\,.$). When $\mathcal{S}$ and $\mathcal{T}$ are identical, the CCF reaches its maximum value, which corresponds to the shift value that maximises the similarity between the two spectra. The peak of the CCF provides an estimate of the velocity shift between the target spectrum and the template spectrum. This shift, denoted by Δv, can be related to the lag δx in the CCF by

$$\Delta v = -\frac{\delta x \cdot c}{\lambda}, \tag{3.3}$$

where c is the speed of light, and λ is the central wavelength of the spectral feature being used for the cross-correlation. The negative sign in the equation indicates that the velocity shift is in the opposite direction to the lag in the CCF.

Accurate RV measurements require the use of a template spectrum that closely resembles the properties of the target star. For example, when analysing solar twin stars, templates for early or late-type stars can lead to greater uncertainties in their CCF due to variations in emission and absorption lines, rotation-induced broadening, and other dominant line-broadening mechanisms (Gray, 2008). The strength of the CCF function is enhanced as the template spectrum matches more closely with that of the target star's properties (e.g., Sousa et al. 2008; Ramírez et al. 2009). Hence, for the utmost RV precision, it is vital to select a template spectrum similar to the targeted star.

3.3 *Don't get fooled by the star!*

Spectral lines are created when light is either emitted or absorbed by atoms and/or molecules (Rybicki and Lightman, 1979; Hubeny and Mihalas, 2014). If the atom absorbs energy, its electrons move to higher energy levels (or orbitals). When these electrons go back to their original energy levels, they release the extra energy as light (or photons), that has a particular frequency ν, or wavelength ($\lambda = c/\nu$), that depends on the particular energy difference (ΔE) between the two energy levels involved in the transition ($\lambda = c/\nu = hc/\Delta E$, where c is the speed of light, and h is Planck's constant). Each atom has its particular electronic structure and corresponding energy levels, therefore yielding distinct spectral lines.

Star spectra are created when photons, undergo radiative transport until they exit the star. This transport involves processes such as absorption, re-emission, and scattering, which can change the frequency and direction of the photons during radiative transport. The resulting spectrum of the star can provide valuable information about its chemical composition, temperature, and density. From these photons, the observed flux is obtained, which measures the net energy flow over time in a particular frequency range. Even with a high-resolution instrument, precise RV observations of a star can exhibit data scattering much higher than expected solely from internal sources of noise, which is often referred to as RV jitter, or simply jitter (Butler et al., 2003; Jeffers et al., 2014; Cloutier et al., 2017). These additional sources of noise arise from intrinsic stellar variabilities (flows and inhomogeneities), such as dark spots, plages, granulation and pulsations on the stellar surface, but may also arise from unknown systematic errors creeping into the observations (see Fig. 3.1).

Time-dependent sources of noise can be quantified as the difference between the true stellar flux $\phi(t)$ and the observed flux $\psi(t)$. The noise component $\zeta(t)$ can be decomposed into physical noise $\zeta_{\text{phys}}(t)$, arising from intrinsic stellar variability, and systematic noise $\zeta_{\text{sys}}(t)$, arising, for instance, from instrumental effects, imperfect calibration, or other observational biases. The observed flux becomes $\psi(t) = \phi(t) + \zeta_{\text{phys}}(t) + \zeta_{\text{sys}}(t)$. Considerations about these undesirable but non-removable-at-all sources of noise/errors are important for two main reasons; *(i)* it provides better estimates of orbital parameters uncertainties, which can be computed using standard statistical tools such as Markov Chain Monte Carlo (MCMC) methods (see Section 4.2.1), and *(ii)* it provides estimates of intrinsic noise present in the data, allowing one to distinguish RV planetary variations from stellar variability itself, which is crucial for understanding the nature

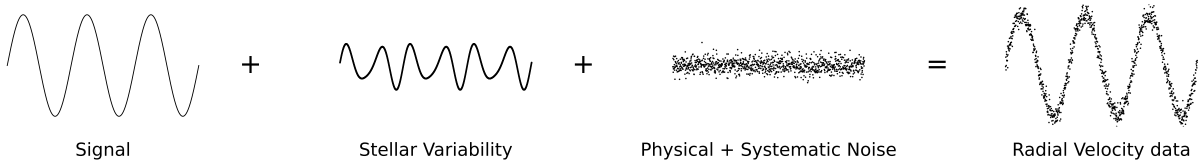

Figure 3.1: Different components – but not restricted to just two – that contribute to the observed RV signal of a star. The first panel shows the true RV signal, followed by the variability in the RV due to spots and plages on the surface of the star (= $A_{\rm spot}\sin(2\pi\nu_{\rm spot}t+\phi_{\rm spot})+A_{\rm plage}\sin(2\pi\nu_{\rm plage}t+\phi_{\rm plage})$, where $A_{\rm spot/plage}$ represents the amplitude of the spot/plage signal (in units of velocity), $\nu_{\rm spot/plage}$ is the frequency of the spot/plage signal (in cycles per unit of time), $\phi_{\rm spot/plage}$ is the phase of the spot/plage signal; in radians), the physical and systematic noise that affects the RV measurements and the observed RV data, which is the sum of all the components.

and characteristics of exoplanets. In practice, modelling these noise sources can be done using various techniques, including Gaussian process regression (see Section 4.3).

RV detection of exoplanets is hindered by astrophysical processes on the surface of stars that produce stochastic signals, the jitter, which can masquerade or even mimic planetary signals (Boisse et al., 2011; Haywood et al., 2014; Vanderburg et al., 2016; Aigrain et al., 2016; Oshagh et al., 2017). At the end of the last millennium, for F, G, and K stars, Saar and Donahue (1997) demonstrated empirically that RV jitter should increase as a function of the star's effective temperature and differential rotation velocity ($v\sin i$) by modelling effects of spots and convection. Following the metrics presented by Wright (2005), for 52 stars of the 89 in our sample, we provide empirical estimates of jitter from available photometric stellar parameters: V and B magnitudes from the Hipparcos and Tycho Catalogues (Perryman et al., 1997), and the $S-$index chromospheric activity from Gomes da Silva et al. (2021).

To quantify the excess of variations in RV, we utilized a jitter metric denoted by σ'. This metric was defined as the square root of the difference between the variance of the RV observations, $\sigma^2_{\rm rv}$, and the variance of the measurement errors, $\sigma^2_{\rm meas}$. The corrected RVs were obtained by removing linear trends caused by long-period planetary or stellar companions. Specifically, we utilized the model $v_r = \alpha + \beta t$ for all time series, where $v_{r,i}$ denotes the RV of the star at time t_i. These were calculated following

$$\sigma^2_{\rm rv} = \sum_{i=0}^{N}\frac{(v_{r,i}-\alpha-\beta t_i)^2}{N-2}. \tag{3.4}$$

Here, N is the total number of observations, and α and β are the intercept and slope of the linear fit. To measure the level of jitter between active and inactive stars of a given spectral type, we utilised an activity metric called ΔF_{CaII}, which differs from the usual R'_{HK} metric[1]. F_{CaII} is given by

$$F_{\mathrm{CaII}} = SC_{\mathrm{cf}}T_{\mathrm{eff}}^4 \times 10^{-14}, \tag{3.5}$$

where S is the flux at the Ca II H and K lines, and C_{cf} is a correction factor that depends on the colour index $(B-V)$ for main-sequence stars with $(B-V) \in [0.3, 1.6]$. The logarithm of C_{cf} is given by the regression:

$$\log C_{\mathrm{cf}}(B-V) = 0.25(B-V)^3 - 1.33(B-V)^2 + 0.43(B-V) + 0.24, \tag{3.6}$$

and Table 3.1, adapted from Wright (2005), provides the lower boundary of the surface flux F_{CaII} as a function of the colour index $(B-V)$. The equation for $\log C_{\mathrm{cf}}$ may present degeneracy due to the use of $(B-V)$ that is closely related to a star's effective temperature and its metallicity (e.g., Rocha-Pinto and Maciel 1998; Lorenzo-Oliveira et al. 2016), so alternatively, the colour index can be replaced by T_{eff} following Lorenzo-Oliveira et al. (2018):

$$\log C_{\mathrm{cf}}(T_{\mathrm{eff}}) = (-1.70 \times 10^{-7})T_{\mathrm{eff}}^2 + (2.25 \times 10^{-3})T_{\mathrm{eff}} - 7.31 \tag{3.7}$$

To calculate the activity metric ΔF_{CaII}, we subtract $F_{\mathrm{CaII,min}}$, the minimum value of F_{CaII} for a given spectral type, from F_{CaII}.

Figure 3.2 illustrates the relationship between activity and jitter metric for the sample of solar twin stars with measured parameters stated beforehand. We observe a clear $(r = 0.49)$ and

---

[1] The R'_{HK} metric measures the amount of emission in the Ca II H and K line cores, normalised by nearby continuum regions and corrects for a photospheric component to the Ca II H and K lines. In contrast, ΔF_{CaII} is a flux density, rather than a fraction, and is constructed using an additional factor of effective temperature (T_{eff}^4)

$(B-V)$	$F_{\mathrm{CaII,min}}$	$(B-V)$	$F_{\mathrm{CaII,min}}$	$(B-V)$	$F_{\mathrm{CaII,min}}$
0.4	5.24	0.8	0.645	1.2	0.079
0.45	3.80	0.85	0.489	1.25	0.063
0.5	2.88	0.9	0.363	1.3	0.052
0.55	2.29	0.95	0.281	1.35	0.044
0.6	1.82	1.0	0.218	1.4	0.038
0.65	1.41	1.05	0.174	1.45	0.042
0.7	1.09	1.1	0.135	1.5	0.029
0.75	0.831	1.15	0.105	1.6	0.025

Table 3.1 - Lower boundary of the Surface Flux $\mathrm{F}_{\mathrm{CaII}}$ as a function of the colour index $(B-V)$. Extracted from Wright (2005).

significant ($p = 0.0004$) trend between the two quantities, i.e., most active stars consistently exhibited higher noise levels. Interestingly, two stars below the unitary $\Delta\mathrm{F}_{\mathrm{CaII}}$ threshold also displayed jitter levels exceeding σ'_{RV} of 15$-$20 m s^{-1}. Following the interpretation of Wright (2005), these stars, HIP5301 ($\sigma'_{\mathrm{RV}} = 19.3$ m s^{-1}) and HIP15527 ($\sigma'_{\mathrm{RV}} = 34.2$ m s^{-1}), should be hosting planets. Note that Fig. 3.2 should be not used as a rule-of-thumb for vetting stars that do not deserve follow-up. It in fact provides a useful estimate of the noise levels one can expect when analysing spectroscopically such stars. Two possible alternative interpretations of such high noise levels in some stars are the lack of a consistently large number of RV measurements – *you don't make statistics with just two points* – or an indication that such systems are young.

Although we do not focus on modelling convection in solar-type stars, our interest is estimating the amplitude in the RV and the time scales involved in causing the oscillations. Classical examples of stellar oscillators are Cepheids and RR Lyræ, that present acoustic oscillation when pressure acts as a restoring force (Osaki, 1977; Dziembowski, 1977; Van Hoolst et al., 1998; Netzel and Smolec, 2022). These oscillations are known as $p-$mode oscillations and have a well-known timing modulation shape. For solar-type stars, $p-$mode oscillations show a power spectrum of evenly spaced modes in frequency, presenting specific non-radial modes that are modelled in the field of asteroseismology (e.g. Aerts et al. 2010; Basu and Chaplin 2018).

Following Kjeldsen and Bedding (1995), the expected RV amplitude v_{osc} and the maximum modulation frequency ν_{max} for a star with mass M, radius R, effective temperature T_{eff}, and luminosity L, can be estimated from:

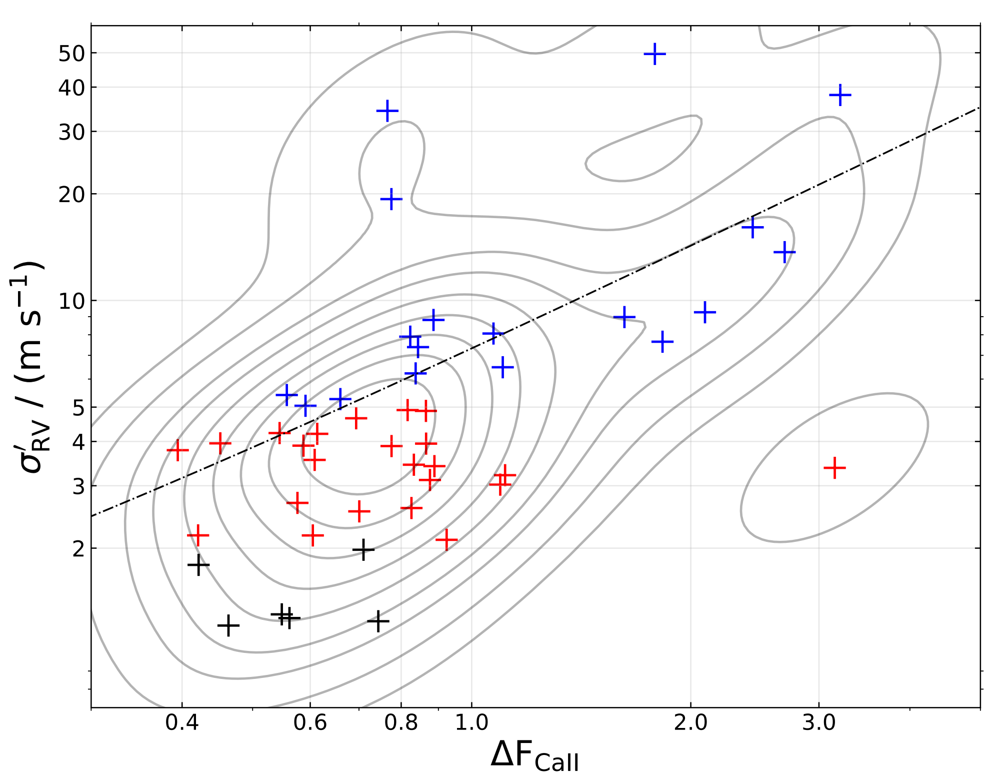

Figure 3.2: Correlation between $\Delta\mathrm{F}_{\mathrm{CaII}}$ (Ca II Flux Variations) and σ'_{RV} (RV jitter) for the sample of solar twin stars analysed in this thesis. The dashed line indicates the linear best-fit for the two variables, whilst black, red and blue crosses indicate stars with $\sigma'_{\mathrm{RV}} < 2$, $\in \ [2,5]$, and > 5 m s^{-1} , respectively.

$$v_{\mathrm{osc}} = (0.234 \pm 1.4 \text{ m s}^{-1}) \times \left(\frac{L}{L\odot}\right)\left(\frac{M}{M\odot}\right)^{-1}; \tag{3.8}$$

$$\tag{3.9}$$

$$\nu_{\mathrm{max}} = (0.183 \text{ min}^{-1}) \times \left(\frac{M}{M\odot}\right)\left(\frac{R}{R\odot}\right)^{-2}\left(\frac{T_{\mathrm{eff}}}{5777 \text{ K}}\right)^{-1/2}. \tag{3.10}$$

For the sample of solar twins analysed in this project, we found out that the average RV amplitude due to intrinsic stellar oscillations was $\langle v_{\mathrm{osc}} \rangle = 0.27 \pm 0.06$ m s^{-1} – similar to the RV amplitude due to oscillations in 18 Scorpii (Bazot et al., 2011) –, and an average maximum modulation frequency of $\langle \nu_{\mathrm{max}} \rangle = 0.17 \pm 0.03$ min^{-1}.

These results are consistent with previous studies of solar-type stars (e.g., Kjeldsen and Bedding 1995; Bedding et al. 2011), which demonstrate that $p-$mode oscillations in such stellar type have typical amplitudes of a few meters per second and maximum frequency in the range of $0.1-1$ min^{-1}. Observationally, Luhn et al. (2020) demonstrated that RV jitter, or the variations in a star's RV over time, is closely linked to stellar evolution. According to their findings, stars evolve through different stages of RV jitter. In younger stars, RV jitter is mainly driven by magnetic activity, while in older stars, it is convectively driven and dominated by granulation and oscillations. This implies that the causes of RV jitter can vary significantly depending on a star's age and evolutionary stage.

We note that values of RV contributions due to stellar noise for solar-type stars are significantly small, down below $5\times$ the typical uncertainty of HARPS' measurements, albeit similar in magnitude order $2\times$ larger than those RV variations provoked by terrestrial planets as discussed in the previous chapter, which typically has amplitudes of tens of cm s^{-1} or more. In this sense, stellar jitter may affect the detection mainly of Earth-sized planets but may have a minor effect on the detection of Jovian planets. We also note that these oscillations are much larger in amplitude than the typical noise level of stellar spectra, which is typically a few cm s^{-1} for Sun-like stars.

The detection of Earth-like planets around Sun-like stars is proving to be an arduous task, requiring robust data analysis techniques and instruments with extreme precision and stability. While HARPS is capable of detecting some low-mass planets, the majority of Earth-like planets remain undetectable with current technology. Nevertheless to say, the next generation of instruments – with some of them already commissioned – such as the Near Infra Red Planet Searcher (ESO/NIRPS; Cabral et al. 2022), the Echelle Spectrograph for Rocky Exoplanet- and Stable Spectroscopic Observations (ESO/ESPRESSO; Pepe et al. 2010), The Carnegie Planet Finder Spectrograph (PFS; Crane et al. 2010), the Keck Planet Finder (KPF; Gibson et al. 2020), and the EXtreme-PREcision Spectrograph (EXPRES; Petersburg et al. 2020) hold great premises for enabling the detection of Earth-twin planets in the future.

3.3.1 Stellar Magnetic Activity Tracers

One can implement a period search in an RV time series and detect a periodic signal. The next step is to determine the significance of the signal and its nature (instrumental, stellar variability, or planetary). The dominant source of RV jitter in late-type main-sequence stars is generally stellar magnetic activity, which includes periodic phenomena (e.g., dark spots and plages), or stochastic ones (e.g., flares). The underlying physics behind these phenomena is outside the scope of this project, however, we are interested in their influences on RV measurements. In the best case, stellar activity may act by simply adding noise to the data, but in the worst scenario, it can mimic planetary signals. The study of activity can be conducted using various tracers, each offering different insights: the Bisector Inverse Slope of the CCF (CCF BIS), the Full Width at Half Maximum of the CCF (CCF FWHM), the $S-$index, Na doublet, He I, and the Balmer lines.

The CCF BIS measures the asymmetry of the spectral line profile, calculated by bisecting (cutting across) it and measuring the slope of the bisector. For HARPS' spectra, it is done within its dedicated pipeline. The relationship between the CCF BIS and RV measurements is direct and proportional: A positive (negative) CCF BIS indicates that the line profile is skewed towards the red (blue) side, which means that the object is moving away (towards) the observer's line of sight. Due to its sensibility to variations in the photospheric velocity during magnetic activity, the CCF BIS is often used to detect the presence of stellar pulsations or rotationally-induced activity, in some cases, ruling out the presence of orbiting exoplanets around a star (e.g. Henry 2000; Queloz et al. 2001).

If $f(\lambda)$ is the spectral intensity as a function of wavelength λ, and $g(\lambda)$ is the reference/template spectrum, recall that the CCF of these two spectra is given by (c.f. eq. 3.2):

$$\mathrm{CCF}(v) = \int_{-\infty}^{\infty} f(\lambda) g(\lambda + v)\,\mathrm{d}\lambda \tag{3.11}$$

where v is the relative velocity between the two spectra. The bisector is defined as the line that connects the midpoints of the tops of the CCF peaks on both sides. It is divided into a fixed number of segments, and the velocity at the midpoint of each segment is calculated. The difference in velocity between the first and last segments gives an estimate of the CCF BIS:

$$\mathrm{CCF\ BIS} = \frac{1}{n} \sum_{i=1}^{n} \frac{v_2(i) - v_1(i)}{\Delta v} \tag{3.12}$$

where v_1 and v_2 are the velocities at two points on the bisector line, Δv is the velocity range over which the bisector line is calculated, and n is the number of points on the bisector line (see Fig. 3.3 for a visual representation).

Similarly to the CCF BIS, the CCF FWHM provides a measure of the spectral line broad-

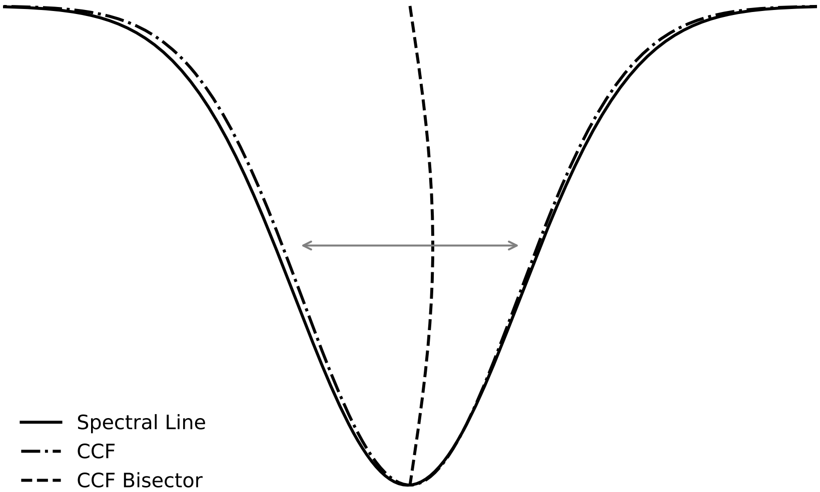

Figure 3.3: Schematic representation of a spectral line profile, the cross-correlation function (CCF; see Section 3.2, and the CCF Bisector Span (CCF BIS). The line profile is represented by the solid black line, which was simulated using the equation $\mathcal{A}\left(-\frac{(x-x_0)^2}{2\sigma^2}\right)\left[1+\beta\left(\frac{x-x_0}{\sigma}\right)\right]$, where $\mathcal{A}$ denotes the amplitude of the line, x_0 the central RV of the line, σ the width (standard deviation) of the line profile, and β is an asymmetric term. The CCF profile, shown as the dashed-dotted black line, is calculated using a Gaussian template following $\exp\left(-\frac{1}{2}\left(\frac{x}{\delta}\right)^2\right)$, where δ is the template profile width. The CCF BIS is depicted by the black dashed line, and the FWHM is represented as the left-right grey arrow. The determination of the RV entails fitting the spectral line with a reference spectrum (a mask; see eq. 3.2) and estimating the displacement in the peak position. It is important to highlight that Keplerian motion, induced by an orbiting companion, should only result in lateral shifts to the spectral lines (blueshift and redshift; Doppler 1842), while activity-based motion will also alter the shape of the spectral line.

ening due to various physical processes, such as turbulence, rotation, and magnetic activity (see Section 3.3). It is also estimated within the HARPS' dedicated pipeline, and a simple expression for the CCF FWHM is given by $2\left(\sqrt{2\log 2}\right)\sigma$, where σ is the standard deviation of the line profile, which quantifies the spread of the intensity values within the line profile. Additionally, for magnetically active stars, the FWHM can be significantly larger than that of inactive stars due to the presence of chromospheric emission lines, and it is also used to estimate the rotation velocity of a star, as rotation is closely related to the broadening of a spectral line. The rotation velocity can be estimated from the FWHM using the expression $v_{\rm rot} = ({\rm FWHM}/\lambda)\,(c/2)$, where $v_{\rm rot}$ is the rotation velocity. Following dos Santos et al. (2016), it is important to observe that broadening effects of macro-turbulence must be also considered, and subsequently, $v\sin i = \sqrt{(0.73\pm0.02)\left({\rm FWHM}^2 - v_{\rm macro}^2 - (5.97\pm0.01)^2\right)}$.

One of the most important stellar activity tracers is the $S-$index, widely used as a measurement of the chromospheric emission of a star. It is known that the Ca II H&K lines (396.85 nm and 393.37 nm, respectively) are highly sensitive to the presence of magnetic fields in the chromosphere (Duncan et al., 1991), and the $S-$index is defined as the ratio of the integrated flux in the chromospheric emission cores to that in the nearby continuum, expressed as:

$$S = \frac{\int_{\lambda_1}^{\lambda_2} F(\lambda) d\lambda}{\int_{\lambda_3}^{\lambda_4} F(\lambda) d\lambda}, \tag{3.13}$$

where λ_1 and λ_2 are the wavelength limits for the chromospheric emission cores and λ_3 and λ_4 are the limits for the nearby continuum. $F(\lambda)$ is the flux density at wavelength λ. Note, the presence of magnetic fields in a star's chromosphere can lead to heating of the plasma and emission in the Ca II H&K lines, resulting in higher $S-$index values for more active stars. Moreover, long-term variations in the $S-$index can provide information about the evolution of magnetic activity in a star. In particular, the $S-$index can be used to study the *activity cycle*, characterised by variations of magnetic activity over several years to decades (Baliunas et al., 1995).

As part of the Balmer series of hydrogen spectral lines, the Hα is another important stellar activity indicator (Hall, 2008; Ibañez Bustos et al., 2023). It is produced when electrons make a transition from $n = 3$ to the $n = 2$ energy level, releasing photons of 656.28 nm wavelength. The Hα line can be observed in absorption when the chromosphere is magnetically heated, due to flares or prominences, and its strength is typically expressed as the equivalent width of the line, denoted by $\mathrm{EW}(H\alpha)$. This quantity is calculated as the ratio of the difference in flux between the Hα line and the nearby continuum region, divided by the flux in the continuum region: $\mathrm{EW}(H\alpha) = (F_{\mathrm{cont}} - F_{H\alpha})/F_{\mathrm{cont}}$, where F$_{\mathrm{cont}}$ is the flux in the nearby continuum region, and F$_{H\alpha}$ is the flux in the Hα line.

Stellar flares are also often studied using the He I (587.56 nm; with EW$= 0.1-0.3\,\mathring{A}$ for solar-type stars) and Na doublet lines (589.0 nm and 589.6 nm). The presence of strong magnetic fields, generated by the dynamo effect in the convective envelopes of stars, can enhance the strength of the He I line. Thus, the strength of this line can be used as a measure of a star's magnetic activity, with stronger lines indicating higher levels of activity. The Na line, on the other hand, is highly sensitive to the presence of cool, dense gas in the chromospheres of stars. As for He I, stronger Na lines indicate higher levels of stellar activity (Zirin, 1988; Oza et al., 2019).

Chapter 4

The Framework in Time-Series Data Analysis

In this chapter, we will delve into the mathematical and statistical framework that is behind the analysis and interpretation of time-series data, with a particular emphasis on RVs. Our primary focus will be on the Generalised Lomb-Scargle method, a technique that is used to detect periodic signals within unevenly sampled data. We will also introduce Bayesian inference, a technique that involves using prior knowledge and data to update our beliefs and make predictions about the probabilities of the data's behaviour. Finally, we will explore Gaussian Process regression, a method that establishes the relationship between variables and predicts outputs based on inputs, while providing a measure of uncertainty for each prediction.

4.1 The Generalised Lomb-Scargle (GLS) Method – An Overview

Consider a time-series $\Lambda = \{t_k, y_k, \sigma_k\}_{k=1}^{N}$, where t_k is defined as $t_k = t_0 + h\Delta t$ and h is a positive real number. To estimate the periodicity of this time series, we can use the power spectrum and discrete Fourier transform (Bracewell, 1965), which involves calculating:

$$P_S(\omega) = \frac{1}{2\pi N} \left| \sum_{k=1}^{N} y_k \, \mathrm{cis}(\omega t_k) \right|^2 \tag{4.1}$$

where $\mathrm{cis}(x) = \cos(x) + i\sin(x),\ i \in \mathbb{C}$, ω is the angular frequency, and N is the length of the time-series. This simplest model assumes a uniform distribution of events over time, however, when the distribution is not uniform, the results obtained from the Equation 4.1 resembles a χ^2 distribution, which is commonly used in classical hypothesis testing to differentiate between

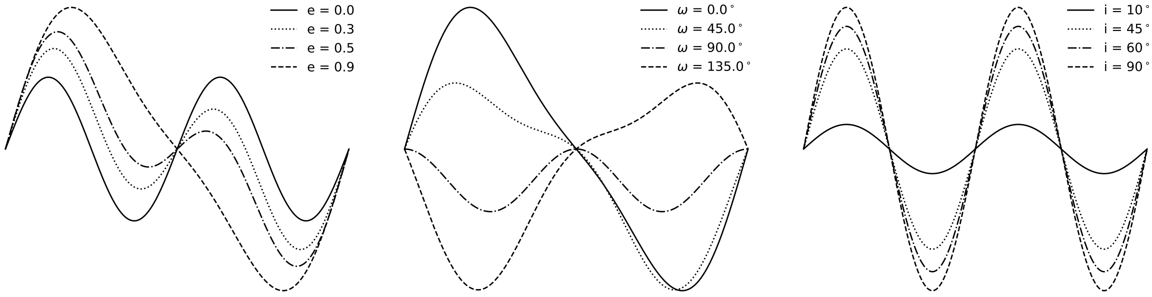

Figure 4.1: The impact of changing orbital eccentricity e, periastron argument ω, and orbital inclination i on RV signals for a star with a 10-d orbiting planet. Eccentricity changes cause variations in the distance between the planet and its host star. For high eccentricities, the RV amplitude becomes larger and induces a shorter period of the signal. On the other hand, lower values of e result in a smaller RV amplitude and a longer period for the signal. As seen in the first panel, the shape of the RV curve also changes, differing from a perfect sine as the orbital eccentricity increases. Changes in ω lead to shifts in the timing of the maximum and minimum RVs. If the planet's closest approach to the star coincides with the line of sight between the star and the observer, the RV signal will have its maximum amplitude, acquiring a *perfect* sinusoidal shape. Conversely, if the closest approach occurs when the planet is farthest from the observer, the RV amplitude will be at its minimum. Finally, in the last panel, the orbital inclination also affects the signal, with a lower orbital inclination resulting in a smaller amplitude.

periodic and non-periodic signals in non-uniform data. This method is known as the Lomb-Scargle periodogram (LS; Lomb 1976; Scargle 1982).

One key disadvantage of the LS method is that it often constrains the assumption that the signals are sinusoids with a constant frequency (VanderPlas, 2018). In reality, many signals exhibit non-sinusoidal behaviour and varying frequencies. For example, the RV signal induced by an orbiting planet can be affected by the orbital eccentricity, periastron argument and orbital inclination (see Fig. 4.1), which can cause the signal to deviate from a perfect sinusoid shape with a constant frequency. To address these limitations, the Generalised Lomb-Scargle (GLS) method was developed (Bretthorst, 2001; Zechmeister and Kürster, 2009).

In the GLS method, the signal is represented as a linear combination of basis functions, which can be chosen to be orthogonal, non-orthogonal, or even non-linear. The amplitudes and phases of these basis functions are determined by minimising the difference between the model and the data using a least-squares method. The simplest approach to the analysis of RV observations consists in a decomposition of its measurements over time following the sine-like model $y(x) = \alpha\cos(\omega t) + \beta\sin(\omega t) + \gamma$ at some frequency ω. By fitting a sinusoid function to the data, we mean obtaining the best optimal values for the parameters set $\eta := \{\alpha, \beta, \gamma\}$, employing a minimisation of the model function, that is, $\partial_\eta \chi^2 = 0$, where,

$$\chi^2 = \sum_{i=1}^{N} \frac{[y_i - y(t_i)]^2}{\sigma_i^2} := W \sum_{i=1}^{N} r_i [y_i - y(t_i)]^2, \tag{4.2}$$

with N representing the number of observations, y_i the observed RV at times t_i, $y(t_i)$ is the model prediction at time t_i, σ_i is the measurement uncertainty, and $r_i = (W\sigma_i^2)^{-1}$ is a weight that normalises the contribution of each observation. This minimisation yields a system of linear equations that can be written in matrix form (Zechmeister and Kürster, 2009),

$$\begin{aligned}
\partial_\alpha \chi^2 &= 0 = 2W \sum_{i=1}^{N} r_i [y_i - y(t_i)] \cos(\omega t_i), && (4.3)\\
\partial_\beta \chi^2 &= 0 = 2W \sum_{i=1}^{N} r_i [y_i - y(t_i)] \sin(\omega t_i), && (4.4)\\
\partial_\gamma \chi^2 &= 0 = 2W \sum_{i=1}^{N} r_i [y_i - y(t_i)] && (4.5)
\end{aligned}$$

$$\begin{pmatrix} \hat{YC} \\ \hat{YS} \\ Y \end{pmatrix} = \begin{pmatrix} \hat{CC} & \hat{CS} & C \\ \hat{CS} & \hat{SS} & S \\ C & S & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}. \tag{4.6}$$

Note that the parameter γ can be written in terms of the parameters α and β according to $\gamma = Y - \alpha C - \beta S$, which reduces our problem to a bi-dimensional case. Furthermore, denoting for two entities A and B, $\widehat{AB} - A \cdot B = AB$, we can write compactly as

$$\begin{pmatrix} YC \\ YS \end{pmatrix} = \begin{pmatrix} CC & CS \\ CS & SS \end{pmatrix} \cdot \begin{pmatrix} \alpha \\ \beta \end{pmatrix}. \tag{4.7}$$

This matrix can be inverted to obtain the values of α and β, which are used to calculate the goodness of fit by evaluating the χ^2 quantity. Therefore, the solution for the set of parameters η:

$$\alpha = \frac{YC \cdot SS - YS \cdot CS}{CC \cdot SS - CS^2}, \text{ and } \beta = \frac{YS \cdot CC - YS \cdot CS}{CC \cdot SS - CS^2}. \tag{4.8}$$

Getting back to the first set of linear equations we obtained (4.3–4.5), we must first observe that $\sum r_i[y_i - y(t_i)] = 0$ by summing all minimisation terms. By multiplying both the left and right sides of this expression by $y(t_i)$, and at the end substituting the previous values for the parameters α and β found, we obtain:

$$\frac{\chi^2}{W} = YY - \frac{YC^2 \cdot SS}{CC \cdot SS - CS^2} - \frac{YS^2 \cdot CC}{CC \cdot SS - CS^2} + 2\frac{(YC \cdot CS \cdot YS)}{CC \cdot SS - CS^2}. \tag{4.9}$$

For a given time-series observations Λ, the Lomb-Scargle periodogram is defined as a measure of the power spectral density of the signal, and it can be expressed in terms of the matrix elements,

$$\Theta_{\mathrm{LS}}(\omega) = \frac{SS \cdot YC^2 + CC \cdot YS^2 - 2\,(CS \cdot YC \cdot YS)}{YY \cdot (CC \cdot SS - CS^2)}, \tag{4.10}$$

4.1.1 The Kotel’nikov-Whittaker-Nyquist-Shannon (KWNS) theorem

The Kotel’nikov-Whittaker-Nyquist-Shannon theorem (KWNS) is a fundamental concept in signal processing that relates the sampling rate of a signal to its frequency content. It states that *to reconstruct a continuous-time signal from its discrete-time samples, the sampling rate must be at least twice the maximum frequency component of the signal* (e.g., Davenport et al. 2012).

Upon examination of a planet in orbit to a star, the observed radial velocity signal usually exhibits a sinusoidal-like behaviour in the time domain, however, owing to the limitations of ground-based observations conducted during nighttime, only a limited segment of the orbital motion is observable, resulting in the manifestation of a Heaviside step function in the time domain ($H(x) := \int_{-\infty}^{+\infty} \delta(s)\,\mathrm{d}s$). Consequently, the convolution of these two functions leads to

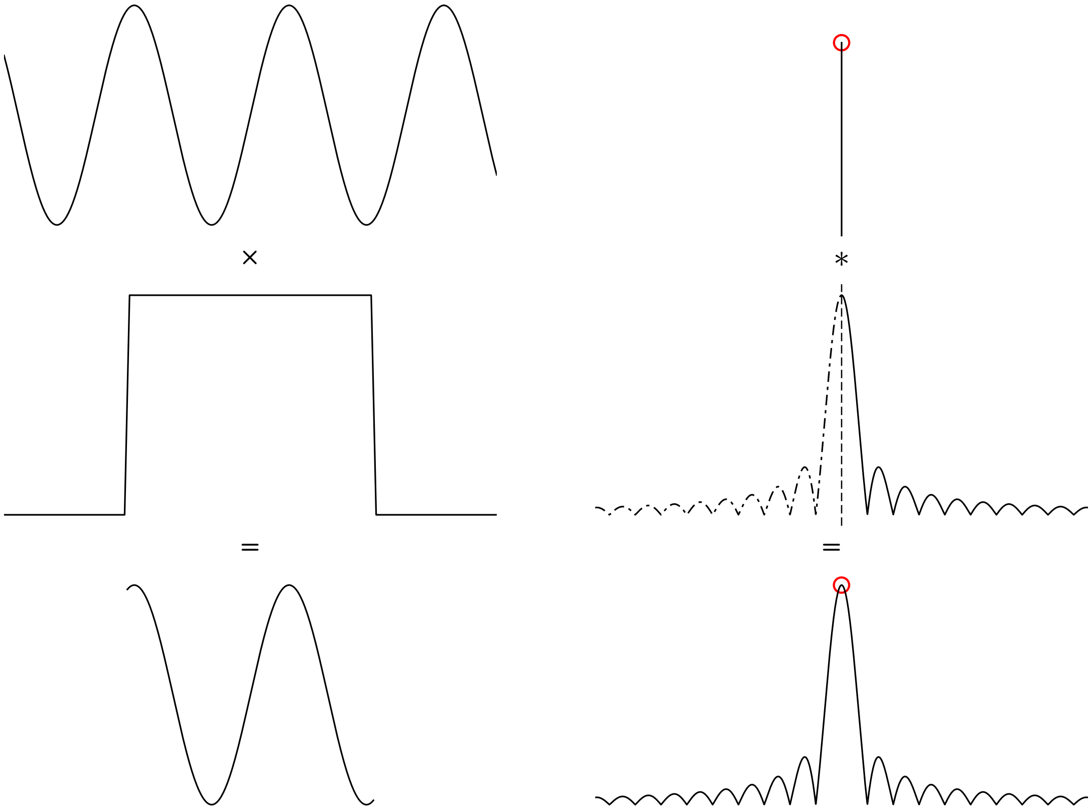

Figure 4.2: In time domains (left panels), radial velocity observations are seen as continuous functions in time (first panel), and owing to the limitations of ground-based observations conducted during nighttime (second panel), i.e., just a portion of the observable window, the observed final data (third panel) will consist of a truncated sine function. In the frequency domain (right panels), however, the Fourier transform of a sine function is equivalent to a Dirac$-\delta$ function (first panel), while the Fourier transform of the window/Heaviside function (second panel) is equivalent to a $sinc$ function peaked at 0. In this panel, the negative values of $sinc$ are depicted as dashed lines. This same function appears, in the Fourier domain (third panel), as a $sinc$ function peaked at the period of the modulation (red dot). Adapted from Figure 7.9 in Hatzes (2019).

the extraction of a partial representation of the radial velocity signal owing to a sinusoidal-like shape in the time domain. Shifting the focus to the frequency domain, whereby a Fourier transform is employed, it is observed that the radial velocity signal transforms into a Dirac function ($\delta(x-\alpha) = (2\pi)^{-1} \int_{-\infty}^{+\infty} e^{it(x-\alpha)}\,\mathrm{d}x$), characterised by a single spike at the prominent modulation period of the signal, while conversely, the Heaviside function transforms into a $sinc$ function ($\sin(x)/x$). Following the convolution in the time domain, the resultant $sinc$ function in the frequency domain is centred around the prominent period modulation present in the data, with smaller peaks evenly spaced representing the data modulation aliases (usually P/n or $n \times P$, with $n = 2, 3, 4, \dots$). This modulation, whose underlying nature could be attributed to stellar, systematic, or planetary factors, necessitates further analysis to discern its origin with certainty (see Fig. 4.2).

Yet in the context of RV analysis, the maximum frequency component from a periodogram corresponds to the period of an orbiting planet, and the KWNS theorem is important to consider as it sets limits on the necessary sampling rate for accurately determining the properties of a planet's orbit. Insufficient sampling rates can result in the planet's periodic signal being mistaken for a lower frequency signal, known as aliasing. This problem is especially pronounced for planets with short orbital periods, as their frequency components can overlap with noise signals from the instruments or the environment (1 synodic day, 1 synodic month, or even 1 sidereal year). To avoid aliasing, the sampling rate must be at least twice the highest frequency component present in the signal, or twice the frequency corresponding to the planet's orbital period for RV data (Whittaker, 1915; Shannon, 1949).

Mathematically, the KWNS theorem is expressed as $\nu_s \geq 2\nu_{\max}$, where ν_{s} is the sampling frequency (Oppenheim et al., 1999). If the sampling rate is lower than the Nyquist rate ($2\nu_{\max}$), then the signal will be under-sampled, and the high-frequency components will be aliased into lower frequencies, causing distortions and loss of information on a periodogram. The sampling rate must exceed the Nyquist rate for the samples to accurately represent the function. These inequalities describe the so-called Nyquist-Raabe criterion.

The Nyquist rate is expressed as $\nu_{\text{Nyquist}} = 1/(2\Delta t)$, where Δt is the time interval between consecutive measurements. For RV data, the Nyquist rate is related to the maximum frequency component of the signal, which is determined by the planet's orbital period. Therefore, the minimum sampling rate required to precisely determine the planet's orbital parameters decreases as the orbital period increases. The Nyquist rate sets limits on the necessary sampling rate for accurately determining the properties of a planet's orbit, and the KWNS theorem is crucial in analysing RV data. From the periodogram, we can determine if the two periods are aliases of each other. Given P_1 and P_2, it is firstly done by estimating the Nyquist frequency, given by $1/(2(P_2 - P_1))$. Moreover, compute the alias period for P_1 according to $P_1 - (1/(1/P_1 - 2 \times \text{Nyquist frequency}))$. Finally, determine whether the difference between the alias period and P_2 is less than the Nyquist frequency, which indicates that P_1 and P_2 are aliases.

4.1.2 Confidence Estimates

Confidence estimation techniques play a critical role in quantifying the associated uncertainty of statistical results. In the context of RV time series, these methods provide a range of values, which is likely to include the true value of a parameter of interest (e.g., the modulation period of the data) with a certain level of confidence. One of these methods is known as *bootstrapping* (Efron, 1982), which involves generating multiple datasets by re-sampling the original with replacement, each of which is of the same size as the given original one. Bootstrapping is especially useful when the underlying data distribution is unknown, or when it is difficult to model (Ivezić et al., 2014), so it allows for the estimation of the sampling distribution without requiring any assumption about the population distribution.

The most straightforward sequential process for implementing bootstrapping comprises:

1. Generate a sample from the population of interest (the original dataset);

2. Re-sample the data with replacement to create multiple datasets of the same size as the original one;

3. Calculate a statistic of interest, which in our case involves comparing varying with non-varying models with the Bayesian Information Criterion (BIC; see eq. 4.18).

   In this case, the BIC for each new sample generated in the previous step is computed by first fitting a non-varying model to the original dataset, and then fitting a varying model to each of the new samples. In this proposed method, it becomes possible to differentiate a periodic signal within the data by its sinusoidal waveform, irrespective of whether it arises from instrumental or intrinsic stellar noise, or is of planetary origin.

4. Repeat the previous step multiple times (e.g., 300 or more).

   The quantity of iterations necessary is contingent on the size of the dataset and the accuracy desired in the estimation. For small datasets, this is a reasonable number. The computational cost involved in this phase is significant, therefore, an increase in the number of repetitions results in a lengthier execution time, sometimes leading to over-fitting;

5. Making use of the re-sampled dataset, estimate the sampling distribution of the statistic of interest, which can be used to calculate the standard deviation of this distribution;

6. And finally, construct the 95% and 99% confidence intervals for the population mean weight, which is given by

$$\mathrm{CI} = \bar{X} \pm t_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right), \tag{4.11}$$

where $\bar{X}$ is the sample mean, s is the standard deviation of the distribution of mean weights obtained from bootstrapping, n is the sample size, and $t_{\alpha/2}$ is the critical value from the t-distribution with $n-1$ degrees of freedom and significance level $\alpha/2$ (Rees, 1989).

4.1.3 The Bayesian Generalised Lomb-Scargle (BGLS) method

The Bayesian Generalised Lomb-Scargle (BGLS) method is a statistical extension of the GLS approach that offers a probabilistic framework to estimate model parameters and their uncertainties, as well as comparing models with varying frequencies (Mortier et al., 2015; Olspert et al., 2018). In the BGLS framework, Bayes' theorem is employed, which posits that the posterior probability of a model, given a set of observed data, is proportional to the likelihood of the data given the model, multiplied by the prior probability of the model (i.e., $\pi(\boldsymbol{\theta}|y_i,t_i) \propto \pi(y_i|\boldsymbol{\theta},t_i)\pi(\theta)$; see also Section 4.2).

The likelihood function in this case is given by:

$$\pi(y_i|\boldsymbol{\theta},t_i) = \prod_{i=1}^{N} \mathcal{N}\left(y_i|\sum_{k=1}^{K} a_k \cos\left(\omega_k t_i + \phi_k\right), \sigma_i^2\right) \tag{4.12}$$

where $\boldsymbol{\theta}$ represents the model parameters (a_k, ω_k, and ϕ_k), $\mathcal{N}$ is the normal distribution, and σ_i is the uncertainty in the i-th observation. Similarly to the GLS method, the power spectral density (PSD) of the signal can be estimated from the posterior distribution of the model parameters, i.e.,

$$P(\omega) = \frac{1}{2\pi}\int_{\boldsymbol{\theta}} \pi(\boldsymbol{\theta}|y_i,t_i) \sum_{k=1}^{K} a_k^2\, \delta(\omega-\omega_k)\, \mathrm{d}\boldsymbol{\theta}\,, \tag{4.13}$$

where δ is the Dirac delta function.

Like any statistical method, the BGLS approach has its drawbacks, with certain assumptions and limitations. First, the method assumes that the underlying signal can be represented as a sum of sinusoidal functions with fixed frequencies and phases, which is not always an appropriate assumption for certain types of data that do not exhibit clear periodicity, or have complex, non-linear dynamics (see for example Fig. 4.1). Moreover, the accuracy of the BGLS method relies on the choice of prior knowledge or assumptions about the distribution of model parameters, such as the amplitudes, frequencies, and phases of the sinusoidal functions. The selection of appropriate priors can be challenging sometimes and may require careful experimentation. For instance, the choice of prior distributions may depend on the physical properties of the system being studied, such as the mass and orbital period of the exoplanet, or the rotation rate and magnetic activity of the star. Therefore, it is important to carefully select priors that reflect the available knowledge and data and to test the sensitivity of the results to different prior assumptions.

4.2 On the Doctrine of Chances

The accuracy of information we gather from observations heavily relies on our knowledge of statistics and its influence on our methodology. From a Bayesian perspective, we treat all unknown quantities as stochastic variables whose values depend on the outcome of random phenomena or an experiment that is yet to be conducted. We use probability statements to describe unobserved parameters, which are considered model parameters θ, with the set of possible values represented by Ω. When dealing with observed data, such as RVs of a star over time, denoted by the set $\mathcal{D}$, if a particular measurement $y \in \mathcal{D}$ results from observations at some point $x \in \mathcal{D}$, then the distribution of y is determined by an observational model that relies on the underlying function value ϕ, which is defined as $f(x)$. This dependence is represented by a joint likelihood function denoted by $\pi(y|\phi)$. Here, $\pi(\alpha|\beta)$ refers to the probability of α given β (Wasserman, 2010).

$$\pi(y|x, \phi) = \prod_{i=1}^{n} \pi(y_i|x_i, \phi_i), \tag{4.14}$$

As a widely used practice, we compute the logarithm of the likelihood function instead of the original function. This approach offers several advantages, including (a) simplifying the computation of the maximum likelihood estimator (MLE) by exploiting the fact that the logarithm is a monotonically increasing function. Therefore, the maximum of the logarithm is equivalent to the maximum of the original function, (b) computing the logarithm helps avoid numerical problems that could arise when using small probabilities, which may result in floating-point underflow, and (c) it can help prevent over-fitting by reducing the impact of extreme outliers that could cause the likelihood to become very small.

Different from frequentist methods – the somehow natural approach to data analysis –, the concept of probability in Bayesian inference is associated with the level of certainty about statements. This approach allows us to question *how do I think my data is?* before commencing the analysis. Despite the difference in problem-solving methods, the outputs of both approaches are comparable. Bayesian inference produces probability density functions (PDFs), which describe the probability of a stochastic variable taking a particular value. A random variable X is defined as continuous if there exists a function f_X that is positive semi-definite, normalised, and, for every real number $\alpha \leq \beta$, the probability that X lies between α and β is given by the integral of f_X over the interval $[\alpha, \beta]$. When X is finite and countable, it is considered a discrete random variable, and the probability that X takes a specific value is determined by the probability mass function (PMF). The concept of a stochastic variable is crucial in Bayesian inference as it serves as a connection between the spaces and events present in the background data. The reader may refer to Section 4.2.1 for more details on sample spaces and events.

The goal is to identify a model that can explain the information in the data accurately while avoiding noise or false information. Using the product rule (also known as the chain rule in probability theory), the probability of two hypotheses A and B given some evidence of C can be expressed as the product of the probability of A given C and the probability of B given both A and C, i.e., $\pi(\mathrm{AB}|\mathrm{C}) = \pi(\mathrm{A}|\mathrm{C})\pi(\mathrm{B}|\mathrm{AC})$. This followed from Aristotelian logic, a system of deductive reasoning often used in Bayesian inference to derive conditional probabilities[1]. We can rearrange the equation to obtain the probability of B and A given C as $\pi(\mathrm{BA}|\mathrm{C}) = \pi(\mathrm{B}|\mathrm{C})\pi(\mathrm{A}|\mathrm{BC})$.

---

[1] Aristotelian logic, or syllogistic logic, is a system of deductive reasoning. In Bayesian inference, it is often used to derive conditional probabilities, i.e., specifically, the product rule (also known as the chain rule) of probability states that $\pi(A \cap B \mid C) = \pi(A \mid B \cap C) \cdot \pi(B \mid C)$. This equation can be rearranged to obtain $\pi(B \cap A \mid C) = \pi(B \mid A \cap C) \cdot \pi(A \mid C)$, which is another way of expressing the joint probability of A and B given C. This simple rule allows us to update our beliefs about the probability of events as we receive new information.

Moreover, by combining the joint probability that A and B are true given that C is true with the joint probability that B and A are true given that C is true, we obtain:

$$\pi(\mathrm{A}|\mathrm{BC}) = \frac{\pi(\mathrm{A}|\mathrm{C})\pi(\mathrm{B}|\mathrm{AC})}{\pi(\mathrm{B}|\mathrm{C})}. \tag{4.15}$$

Equation 4.15 is known as Bayes' theorem[2], used to measure the goodness of a model in fitting data. The denominator term calculates the marginal likelihood, but in most cases, it cannot be computed analytically. Exceptions occur when the parameters being marginalised are conjugate priors to the data distribution or when $\pi(\theta|x)$ takes the same distribution form as $\pi(\theta)$. Analytical solutions are possible in such cases. However, numerical integration techniques are necessary for all other cases (Parviainen, 2018).

In its modern form, let $\mathrm{A}_1, \ldots, \mathrm{A}_k$ be events such that $\pi(\mathrm{A}_j) > 0$, $i = 1, 2, \ldots, k$. Assume that A_j forms a partition of the sample space C. Let B be any event. Therefore, the conditional probability of A, given B, is,

$$\underbrace{\pi(\mathrm{A}_j|\mathrm{B})}_{\text{posterior}} = \frac{\overbrace{\pi(\mathrm{B}|\mathrm{A}_j)}^{\text{likelihood}} \times \overbrace{\pi(\mathrm{A}_j)}^{\text{prior}}}{\underbrace{\sum_{k=1}^{j} \pi(\mathrm{A}_j) \times \pi(\mathrm{B}|\mathrm{A}_j)}_{\text{marginal likelihood}}}, \quad \text{with } \pi(\mathrm{B}|\mathrm{A}) = \pi(\mathrm{A} \cup \mathrm{B})/\pi(\mathrm{A}) \text{ if } \pi(\mathrm{A}) \neq 0. \tag{4.16}$$

Bayesian inference faces a critical hurdle when it comes to calculating the posterior probability distribution, which informs our updated beliefs about an unknown quantity of interest. While

---

[2] Thomas Bayes' (1701–1761) theorem was posthumously announced in 1763 in *An Essay towards solving a Problem in the Doctrine of Chances*, edited and published by Richard Price FRS (1723–1791). The theorem's symbolic form does not appear in the essay, rather Mr Bayes stated conditional probabilities in the third, fourth and fifth propositions (pp. 370–418), namely, *The probability that two subsequent events will both happen is a ratio compounded of the probability of the 1st, and the probability of the 2nd on supposition the 1st happens*; *If there be two subsequent events be determined every day, and each day the probability of the 2nd is b/N and the probability of both P/N, and I am to receive N if both of the events happen the 1st day on which the 2nd does; I say, according to these conditions, the probability of my obtaining N is P/b*; and *If there be two subsequent events, the probability of the 2nd b/N and the probability of both together P/N, and it being 1st discovered that the 2nd event has also happened, the probability I am right is P/b* (Bayes and Price, 1763).

conjugate priors offer an analytical solution by producing posterior distributions with the same functional form, this is not always possible. As a result, the majority of cases require numerical computation, often through Markov Chain Monte Carlo (MCMC) sampling (c.f. Section 4.2.1). A major advantage of this approach is the ability to incorporate prior information into the analysis, making it valuable in cases where there is a limited amount of data. The prior distribution can act as a regulator, preventing over-fitting. Additionally, Bayesian inference enables model comparisons, where we can compare the posterior probabilities of different models to determine the best statistical explanation for the data.

Model comparisons can also be achieved using the Akaike information criterion (AIC) or Bayesian information criterion (BIC), developed respectively by Akaike (1974) and Schwarz (1978). Both AIC and BIC are used to compare models that fit the same dataset but differ in complexity or the number of parameters they contain.

The AIC is defined as:

$$\mathrm{AIC} = -2\log(\mathcal{L}) + 2k, \tag{4.17}$$

where $\mathcal{L}$ is the likelihood of the data given the model, and k is the number of parameters in the model. We highlight AIC penalises models for the number of parameters they contain so that models with more parameters will have a higher AIC value. The likelihood term in this equation is used to assess the goodness of fit of the model so that models with a better fit will have a lower AIC value. Similarly, the BIC is defined as:

$$\mathrm{BIC} = -2\log(\mathcal{L}) + k\log(n), \tag{4.18}$$

where n is the sample size. Model comparison with these two criteria is relatively simple: the model with the lowest AIC or BIC value is considered the best, as it provides the best balance between model complexity and goodness of fit. It is important to observe, however, that when comparing two models, a difference of 2 in either value is typically considered substantial evidence in favour of one model over the other. A difference of 6 or more is considered very strong evidence in favour of the model with the lower BIC or AIC value. Differences of less than 2 are often considered as weak evidence for model selection, so either hypothetical models A or B can represent the statistical behaviour of a data (Burnham and Anderson, 2004).

Model comparison encompasses a critical aspect wherein AIC and BIC exhibit distinct underlying assumptions and objectives. AIC prioritises simpler models over complex ones and assumes that the authentic model lies among the ones being compared, with the goal being to minimise the expected information loss. Conversely, BIC relies on Bayesian model selection, i.e., it takes into account the prior probability of each model's veracity and presumes that the true model may be not among the ones being compared. BIC assumes that the true model may not be one of the candidate models being compared, and the objective is to identify the model that maximises the PDF based on the available data.

One potential criticism of Bayesian inference is that it relies heavily on the choice of the prior distribution. The prior distribution can have a significant impact on the posterior distribution, in particular, if the data are limited (see Section 4.2.2). As a result, the results can be subjective, and different analyses may arrive at different conclusions based on their choice of prior. Furthermore, the estimates involving Bayesian inference can be computationally expensive and time-consuming, especially when the models are complex and the dataset is large. Bayesian inference also requires the specification of the likelihood function, which is often not straightforward to obtain, and the Markov Chain Monte Carlo (MCMC) algorithms used to obtain the posterior distribution can be difficult to implement and tune. On the other hand, frequentist methods, such as maximum likelihood estimation, do not require the specification of prior distributions, and their results are generally considered to be more objective. Still, frequentist methods may not perform well when the sample size is small or when the data are noisy, and they may not provide estimates of the uncertainty associated with the estimates. Bayesian inference is a powerful statistical framework that allows us to update our beliefs about model parameters in the light of new data. While it has some limitations, such as its dependence on the prior distribution and computational complexity, it provides a flexible and robust approach to statistical inference.

4.2.1 Markov Chain Monte Carlo (MCMC)

Markov Chain simulations, also known as Markov Chain Monte Carlo (MCMC), refer to a group of sampling methods, such as the Affine Invariant sampler (Goodman and Weare, 2010), Metropolis-Hastings (Hastings, 1970), Gibbs (Geman and Geman, 1984) and the Hamiltonian No U-Turn samplers (NUTS; Hoffman et al. 2014). These techniques are used to generate samples from a probability distribution of a stochastic variable X_n, which only depends on its previous state X_{n-1}. The state space is assumed to be discrete, and the sampling is done using a transition distribution, represented by $T_n(\lambda_n|\lambda_{n-1})$. MCMC is particularly useful in situations where it is not practical, or computationally expensive, to sample a small subset of Ω that represents the entire population, and when the prior distribution of model parameters, $\pi(\theta|\mu)$, is complex (Gelman et al., 2013).

The Affine Invariant (AI) sampler, in particular, is an MCMC algorithm that efficiently samples high-dimensional posterior distributions. Traditional MCMC methods can struggle to sample from distributions with high correlation between parameters, and in contrast, the AI sampler overcomes this issue by using a set of *stretch moves* to sample the posterior distribution in an invariant way to affine transformations. The algorithm operates by using an ensemble of N walkers, where each walker represents a point in the parameter space. During each iteration, a pair of walkers is randomly selected, and a new point is proposed using a stretch move, which involves stretching or compressing the current point along a randomly chosen axis, with the stretching factor chosen from a distribution that depends on the scale of the distribution at the current point. Details about the AI sampler are given by Goodman and Weare (2010) and Foreman-Mackey et al. (2013).

The acceptance probability for the proposed move is given by the Metropolis-Hastings ratio:

$$\alpha = \min\left\{1, \frac{\pi(\theta')}{\pi(\theta)}\frac{q(\theta|\theta')}{q(\theta'|\theta)}\right\}, \tag{4.19}$$

where $\pi(\theta)$ is the posterior distribution, θ is the current point, θ' is the proposed point, and $q(\theta'|\theta)$ and $q(\theta|\theta')$ are the proposal distributions for the new and old points, respectively.

In MCMC, a sequence of random samples, also known as a Markov chain, is generated iteratively. At each step of the simulation, the approximate distribution is updated to ensure convergence to the desired final distribution. One of the important aspects of MCMC is to determine when the chain has converged to the target distribution. This is essential to ensure that the estimates obtained from the chain are accurate and reliable. The Gelman-Rubin (GR) statistic, denoted by $\hat{R}$, is one of the widely used criteria to assess convergence. It compares the estimate for the marginal posterior variance V for parameter θ to the mean within-chain variance W. For a set of well-converged chains, these two quantities should be equal, thus GR statistics values approaching 1 are optimal, i.e. the chains have converged and the estimates are reliable. With the estimated scale reduction for M chains with N steps,

$$\sqrt{\hat{R}} = \sqrt{\frac{N-1}{N} + \frac{M+1}{NM}\frac{N}{M-1}\sum_{m=1}^{M}(\hat{\theta}_m - \hat{\theta})^2 \cdot M\sum_{m=1}^{M}\hat{\sigma}^2_m}, \tag{4.20}$$

where $\hat{\theta}$ is the parameter estimated using all samples, $\hat{\theta}_m$ the mean from a simple chain, and $\hat{\sigma}^2_m$ is the parameter variance estimated from a single chain.

An MCMC algorithm starts a chain from some point in the parameter space Ω, moving to another point if it rejects or accepts the movement based on the posterior density ratio between both locations in view to ensure an asymptotic convergence of the samples in a chain to the posterior distribution. A process is said to be Markov, or to have the Markov property, if, for all $n \geq 1$, the probability distribution X_{n+1} is determined by the state X_n of the process at time n, and does not depend on the past value of X_k, for $k = 0, 1, \ldots, n-1$ (Gelman et al., 2013).

To ensure that a Markov Chain Monte Carlo (MCMC) simulation has accurately converged to the desired stationary distribution, it is important to assess the quality of the generated samples. Several convergence diagnostics have been proposed to evaluate different aspects of convergence, and their combined use can provide complementary information about the quality of the simulation (Foreman-Mackey et al., 2013). Among the convergence diagnostics commonly used are the following three: (a) The minimum auto-correlation time factor for chains to be considered well-mixed and the MCMC run to be halted (*minAfactor*; set at 40 within RadVel). It is a diagnostic that measures the effective sample size of the MCMC simulation and provides an estimate of the minimum number of independent samples that should be retained from the chain to accurately represent the posterior distribution. This diagnostic is calculated as the maximum value of k such that the auto-correlation at lag k is greater than or equal to $1/2$. (b) The maximum relative change in the auto-correlation time to be considered well-mixed and the MCMC run to be halted (*maxArchange*; set at 0.03 in RadVel). As the simulation approaches convergence, it is expected that the largest changes in parameter values will become smaller. This diagnostic is calculated as the maximum absolute change in the parameter values between successive iterations. (c) The minimum number of independent samples (*minTz*). It measures the minimum number of iterations required for the simulation to converge to the stationary distribution. It is based on the assumption that, as the simulation approaches convergence, the parameter values will start to fluctuate around their true values. This diagnostic is calculated as the minimum number of iterations required for the standard deviation of the parameter values to fall below a predefined threshold.

4.2.2 The importance and impact of priors choice

The prior distribution serves as a starting point for Bayesian analysis and influences the posterior distribution (e.g., Ghaderinezhad and Ley 2018, 2019), which represents the updated beliefs about the parameters after observing the data. The prior distribution can assume either informative or uninformative form (Gelman et al., 2013). When there is a prior knowledge about the parameters of interest, an informative prior can be beneficial, resulting in a posterior distribution that is more concentrated around small sample sizes, leading to a more precise estimation of the parameter. On the other hand, an uninformative prior is useful when there is little knowledge about the parameter of interest. It allows the data to drive the inference – *let the data speak by itself* –, and the posterior distribution will be mainly determined by the likelihood function (c.f. eq. 4.33).

As always, the choice of an uninformative prior can also have its drawbacks. In some cases, an uninformative prior can be diffuse, and assign a significant probability to extreme values of the parameter, leading to posterior distributions that are overly broad and imprecise. This problem can be mitigated usually by employing a regularised prior (for instance, the Gaussian Process prior; see Section 4.3), which balances the need for prior information with the need for the data to drive the inference. In such cases, the distribution is chosen to have a penalty term that discourages extreme or unrealistic values of the parameters, for instance, in exoplanet RV modelling, it relies on imposing the eccentricity of the model to be below 1, or close to 0 for shorter-period orbits due to orbital circularisation expected for such a kind of planets. The result is a posterior distribution that balances the fit to the data with the constraint imposed by the prior distribution.

The choice of prior distribution also has a significant impact on the model selection process (e.g., Keller 1990; Bognanni and Herbst 2018). One situation of this kind is the comparison of Keplerian and non-Keplerian models. The first model assumes that the RV measurements are due to the orbital motion of the exoplanet, and the model parameters include the mass, period, eccentricity, and argument of the periastron of it. On the other hand, the non-Keplerian model assumes that the RV measurements are due to some other physical process, such as stellar activity, and the model parameters include the amplitude and timescale of the non-Keplerian signal. When comparing these two models, an uninformative prior distribution for the model indicator variable would assign equal probability to both models. However, if we have prior knowledge that the Keplerian model is more likely to be true, we can use an informative prior to reflect this belief. For context, we may know that the host star is relatively inactive and has a low level of stellar jitter, which makes the Keplerian model more plausible. In this case, we can use a prior distribution that assigns a higher probability to the Keplerian model. This will result in a posterior distribution that is more likely to favour the Keplerian model over the non-Keplerian model.

Another example of the impact of prior choice on the RV fitting of exoplanets is the choice of the prior distribution for the eccentricity parameter. The eccentricity is a key parameter in the Keplerian model, as it determines the shape of the exoplanet's orbit, and may even correlate with its formation history (e.g., Roberts et al. 2015; Rodet et al. 2017). However, the eccentricity parameter is often poorly constrained by the RV data, which can lead to biases in the estimation of the exoplanet's properties. One possible approach to mitigate these biases is to use a prior distribution for the eccentricity parameter that is informed by other observations, or theoretical considerations. For example, we may choose to use a uniform prior distribution that assumes that any value of eccentricity is equally likely, while others may use a prior that favours lower values of eccentricity, such as a Beta distribution,based on theoretical models of planet formation. Not as a rule-of-thumb, short-orbit exoplanets tend to have circularised orbits due to the tidal forces from their host star. This is known as tidal circularisation (Ferraz-Mello, 2013; Teitler and Königl, 2014; Ferraz-Mello et al., 2015), and it occurs because the gravitational forces between the planet and its host star create tidal bulges on both objects. These bulges cause a transfer of angular momentum between the planet and its orbit, which can lead to the planet's orbit becoming more circular over time (e.g., Ford and Rasio 2008; Barker 2016; Grunblatt et al. 2018).

In the case of null hypothesis[3] testing, the choice of prior is also important. An informative prior for the null hypothesis can be used to represent prior knowledge that the null hypothesis is true, and can help to avoid false positive results. However, an informative prior can also bias the results towards the null hypothesis and lead to false negative results. On the other hand, an uninformative prior for the null hypothesis can be useful when there is little prior knowledge about the hypothesis. This approach can avoid bias towards either hypothesis and let the data drive the inference. However, an uninformative prior can also lead to weak evidence against the null hypothesis, making it challenging to reject the null hypothesis even when it is false. Therefore, the choice of the prior distribution for the null hypothesis depends on prior knowledge and the research question. In general, if there is strong prior evidence that the null hypothesis is true, it may be appropriate to use an informative prior. However, if there is little prior knowledge, an uninformative prior may be more appropriate (Kruschke, 2013; Zaslavsky, 2013).

There are several strategies for choosing a prior distribution in Bayesian inference. One approach is to use an uninformative prior, such as a uniform distribution (also known as a rectangular distribution), whose probability density function with support over the interval [a, b] is given by:

$$\pi(\theta) = \begin{cases} (b-a)^{-1} & \text{for } a \le x \le b, \\ 0 & \text{for } x < a \text{ or } x < b. \end{cases} \tag{4.21}$$

where θ is the parameter of interest or a Jeffrey's prior, which is invariant under re-parameterisation and is proportional to the square root of the determinant of the Fisher information matrix,

$$\pi_{\mathcal{J}}(\theta) \propto \sqrt{\det \mathcal{I}(\theta)}, \quad I_{ij}(\theta) = -\mathbb{E}\left[\partial^2_{\theta_i,\theta_j} \log \pi(X|\theta)\right], \tag{4.22}$$

---

[3] The null hypothesis embodies the standard stance or presumption, that there exists no substantial relationship or disparity between two or more variables, or populations, under scrutiny, thereby implying that any observed relationship is attributable to sheer chance or the vagaries of sampling.

where $\pi(X|\theta)$ is the probability density function of the observed data, and $\mathbb{E}[\cdot]$ denotes the expected value operator. The Jeffreys' prior is often used in cases where there is no prior knowledge about the parameter, or when the prior distribution should not influence the inference in a significant way (Jeffreys, 1946; Kass and Wasserman, 1996).

4.3 Gaussian Process Regression

RV measurements are obtained by observing the Doppler shift of spectral lines in the star's spectrum, which are caused by the gravitational influence of an orbiting planet (c.f. Chapter 2). However, the RV signal can also be affected by the intrinsic variability of the star itself, which is known as stellar activity (c.f. Section 3.3). These modulations arise from various physical processes, such as magnetic activity, stellar rotation, and convection, which introduce noise and systematic errors in the data, and hamper the detection/orbital modelling of exoplanets (e.g., Ferraz-Mello et al. 2011). To overcome this challenge, Gaussian Process (GP) regression has emerged as a powerful technique for modelling stellar activity, and separating it from the exoplanet signal in RV data (e.g., Haywood et al. 2014; Rajpaul et al. 2015; Camacho et al. 2023, and references therein). In this section, we will explore the use of GP regression to model stellar activity and its application to RV measurements for exoplanet detection.

GP regression is a probabilistic framework that allows us to model the data from a mean function and a covariance function (or kernel hereafter; Gelman et al. 2013). The kernel function specifies the correlation between data points and captures the smoothness and complexity of the underlying process. To apply GP regression to stellar activity modelling, we need to choose an appropriate kernel function that captures the correlation structure of the data. Commonly used kernels for modelling stellar activity include the Matern, quasi-periodic, and harmonic kernel functions.

Let us consider that any finite set of points in a function follows a multivariate Gaussian distribution. In other words, given a set of input points $\mathbf{x} = x_1, x_2, \ldots, x_n$, the function values $f(\mathbf{x}) = f(x_1), f(x_2), \ldots, f(x_n)$ follow a multivariate Gaussian distribution with mean vector $\boldsymbol{\mu} = \mu(x_1), \mu(x_2), \ldots, \mu(x_n)$ and covariance matrix $\mathbf{K}(\mathbf{x}, \mathbf{x}')$:

$$\begin{bmatrix} f(x_1) & f(x_2) & \dots & f(x_n) \end{bmatrix}^\intercal \sim \mathcal{N}\left(\mu, \mathbf{K}(x, x')\right). \tag{4.23}$$

The covariance matrix $\mathbf{K}(\mathbf{x}, \mathbf{x}')$ determines the degree of similarity between any two points in the function. It can be expressed in terms of a kernel function $k(\mathbf{x}, \mathbf{x}')$ that measures the similarity between any two points in the input space:

$$\mathbf{K}(x, x') = \begin{bmatrix} k(x_1, x_1') & k(x_1, x_2') & \dots & k(x_1, x_n') \\ k(x_2, x_1') & k(x_2, x_2') & \dots & k(x_2, x_n') \\ \vdots & \vdots & \ddots & \vdots \\ k(x_n, x_1') & k(x_n, x_2') & \dots & k(x_n, x_n') \end{bmatrix} \tag{4.24}$$

As previously mentioned, the choice of kernel function depends on the nature of the function being modelled and the prior knowledge about its properties. Once the covariance matrix is defined, we can use it to make predictions about the function at any given input point $\mathbf{x}_*$, given the observed function values $\mathbf{f}$ and their corresponding input points $\mathbf{x}$. The predictive distribution of the GP is given by:

$$\pi(f_\star | x_\star, x, f) = \mathcal{N}\left(\mu_\star, \Sigma_\star\right), \tag{4.25}$$

where $\boldsymbol{\mu}*$ is the mean of the predictive distribution and $\boldsymbol{\Sigma}*$ is its covariance matrix. These are given by:

$$\mu_\star = k_\star^\intercal \left(\mathbf{K} + \sigma_n^2 \mathbb{I}\right)^{-1} f, \tag{4.26}$$

and

$$\mathbf{\Sigma}_\star = k(\mathbf{x}, \mathbf{x}') - \mathbf{k}_\star^\intercal(\mathbf{K} + \sigma_n^2 \mathbb{I}_n)^{-1}\mathbf{k}_\star, \tag{4.27}$$

where $k(\cdot,\cdot)$ is the kernel function, $\mathbf{K}$ is the covariance matrix of the training data, $\mathbf{k}_\star$ is the vector of covariances between the training data and the test points, σ_n^2 is the noise variance, and $\mathbb{I}_n$ is the identity matrix of order n. The first term $k(\mathbf{x}_\star, \mathbf{x})$ is the covariance between the test points themselves, which measures the uncertainty in the predictions at different locations in the input space. The second term in this expression, $\mathbf{k}_\star^\intercal(\mathbf{K} + \sigma_n^2\mathbb{I})^{-1}\mathbf{k}_\star$, is the cross-covariance between the test points and the training data, which captures the similarity between the new data and the observed data. This term is used to adjust the uncertainty of the predictions based on the similarity between the training and test data (see Rasmussen and Williams 2006; Bishop and Nasrabadi 2006).

Stellar activity problem-solving involves modelling these phenomena as a function of time, using a kernel function that captures the periodicity and smoothness of the activity signal (Angus et al., 2018; Netto et al., 2021). For observed RVs v_i, at time t_i, $v_i = \gamma + f_i + \varepsilon_i$, where γ is the systemic velocity, f_i is the stellar activity signal, and ϵ_i is the measurement noise. We can then write the GP model for the stellar activity signal by:

$$f_i = \sum_{j=1}^{m} a_j \text{exp}\left(-\frac{(t_i - t_j)^2}{2\lambda^2}\right), \tag{4.28}$$

where a_j and t_j are the amplitude and time of the $j-$th Gaussian component, and λ determines the smoothness of the signal. The covariance between the activity signals at two different times, t_i and t_j, is given by the quasi-periodic kernel,

$$k(t_i, t_j) = \eta_1^2 \text{exp}\left[-\frac{|t_i - t_j|}{\eta_2^2} - \frac{\sin^2\left(\frac{\pi|t_i - t_j|)}{\eta_3}\right)}{2\eta_4^2}\right], \tag{4.29}$$

a function of the hyper-parameters Doppler-induced magnetic activity of the signal (η_1), the time-scale for growth and decay of active regions (η_2), the recurrence time-scale for active regions (η_3; which is often comparable to the star's rotation period; Rajpaul et al. 2015), and a smoothing parameter (η_4; Haywood et al. 2014; Aigrain and Foreman-Mackey 2022). This kernel model functions that have a smooth underlying trend with periodic modulations on top of it[4].

Consider a set of n independent – when there is no correlation between the sample items – and identically distributed – the distribution does not fluctuate, and the terms in the sample are taken from the same probability distribution – stochastic measurements (IID) $\mathcal{D} := \{X_1, X_2, \dots X_n\}$ with a PDF $\pi(x|\theta)$. The likelihood function is defined as (Wasserman, 2010),

$$\mathcal{L}_n(\theta) = \prod_{i=1}^{n} \pi(x_i|\theta), \tag{4.33}$$

representing the probability that exists events X_i that fall in any particular range, or discrete set, of values specified for that variable, that is, the joint density of the data. Noteworthy, $\mathcal{L}_n(\theta)$:

---

[4] Note that the Matern kernel with $\nu = 5/2$ and the quasi-periodic kernel are correlated through a spectral representation (Abramowitz and Stegun, 1972; Rasmussen and Williams, 2006), i.e., it can be expressed as a linear combination of Fourier components with coefficients given by the square root of the Gamma function evaluated at $\nu + 1/2 = 3$, which is $\sqrt{3/\pi}$. The quasi-periodic kernel, on the other side, can be defined as the product of the Matern kernel with $\nu = 1/2$ and a periodic kernel. The last, in particular, modulates the smoothness of the Matern kernel and models the periodic variability on top of the underlying trend. For $\nu = 5/2$, the Matern kernel can be expressed by:

$$k_{\mathrm{M}}(t_i, t_j) := \sigma_f^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\sqrt{2\nu}\frac{d}{\rho}\right)^{\nu} K_\nu \left(\sqrt{2\nu}\frac{d}{\rho}\right) = \sigma_f^2 \left(1 + \frac{\sqrt{5}d}{\eta} + \frac{5d^2}{3\eta^2}\right) \exp\left(-\frac{\sqrt{5}d}{\eta}\right), \tag{4.30}$$

where K_ν is the hyperbolic Bessel function of the second kind, ρ and ν are positive parameters of the covariance, σ_f is the amplitude of the kernel, $d = \|t_i - t_j\|_2$ is the Euclidean distance between the two input points, and η is the length scale parameter. The quasi-periodic kernel can be given by:

$$k(t_i, t_j) = \eta_1^2 k_{\mathrm{M}}(t_i, t_j) k_{\mathrm{P}}(t_i, t_j), \tag{4.31}$$

where η_1 controls the amplitude of the kernel, and $k_{\mathrm{P}}(t_i, t_j)$ is the periodic kernel, expressed by:

$$k_{\mathrm{P}}(t_i, t_j) = \exp\left(-\frac{2}{\eta_2^2} \sin^2\left(\frac{\pi|t_i - t_j|}{\eta_3}\right)\right), \tag{4.32}$$

with η_2 controlling the length scale of the periodicity, and η_3 controlling the rate of decay of it. We observe that the quasi-periodic kernel combines a smooth, stationary kernel (Matern with $\nu = 1/2$) with a periodic kernel to model both the underlying trend and the periodic variability in a given dataset.

$\theta \in \Omega \to \mathbb{R}^+$.

We define the maximum likelihood estimator (MLE; see also Section 4.1), henceforth $\hat{\theta}$, as the value of the parameter set that maximises $\mathcal{L}_n(\theta)$, that is, $\hat{\theta} = \mathrm{argmax}[\mathcal{L}_n(\theta)]$. Since both the probability density function (PDF) and the probability mass function (PMF) are differentiable functions of θ, we seek the solution to the equation $\partial_\theta \mathcal{L}_n(\theta) = 0$. To address computational limitations, it is often preferred to work with the natural logarithm of the likelihood function, known as the log-likelihood, denoted as $l_n(\theta) \doteq \log \mathcal{L}_n(\theta)$ (e.g., Foreman-Mackey et al. 2017). Using the log-likelihood is advantageous because it allows us to work with values that are more manageable numerically. For example, in the case of N data-set measurements, each with a probability of γ, the likelihood function's value would be on the order of $\mathcal{O}(\gamma^N)$. Such values are often much smaller than what can be accurately represented using standard floating-point notation on computers. Furthermore, since the likelihood function is a monotonically increasing function, it is one-to-one with the log-likelihood function. Therefore, maximising either function leads to the same parameter estimates (Wasserman, 2010; Craig et al., 2019).

To infer the hyper-parameters of the GP model for stellar activity, we can use the MLE approach. The MLE estimates the hyper-parameters that maximise the likelihood of the observed data given the GP model. This log-likelihood function is given by:

$$\mathcal{L}(y|X,\theta) = -\frac{1}{2}\mathbf{y}^\intercal(\mathbf{K}+\sigma_n^2\mathbb{I})^{-1}\mathbf{y} - \frac{1}{2}\log|\mathbf{K}+\sigma_n^2\mathbb{I}| - \frac{n}{2}\log 2\pi. \tag{4.34}$$

where $\mathbf{y}$ is the vector of observed target values, $\mathbf{X}$ is the matrix of input data, $\boldsymbol{\theta}$ is the vector of hyper-parameters, $\mathbf{K}$ is the covariance matrix of the training data, σ_n^2 is the noise variance, and n is the number of data points. The first term in the equation is the data fit term, which measures the agreement between the model predictions and the observed data. The second term is the complexity penalty, which discourages over-fitting by penalising complex models. The third term is a constant term that does not depend on the model parameters. Maximising the log-likelihood function concerning the hyper-parameters gives the maximum likelihood estimates of the parameters.

To compute the MLE, one can use numerical optimization techniques such as gradient descent or quasi-Newton methods. The optimal hyper-parameters can then be used to predict the stellar activity signal and subtract it from the observed RVs. One of the main advantages of using the Gaussian Process for exoplanet detection is its ability to model correlated noise, which can improve the sensitivity and reliability of exoplanet detection. Correlated noise arises from various sources, including instrumental effects, stellar activity, and other astrophysical sources. By modelling the correlated noise using Gaussian Process, one can better separate the exoplanet signal from the noise and improve the precision of exoplanet parameters (see for example Rajpaul et al. 2015; Jones et al. 2017; Yu et al. 2018; Ferreira and Meléndez 2022; Ferreira et al. 2023, and references therein).

Chapter 5

Exoplanetary System Candidates: Signals and Insights

In this chapter, we present the results for the search of exoplanet candidates around solar twin stars as part of the Solar Twin Planet Search programme, totalling 16 new candidates. We analysed ESO/HARPS RVs collected through $7 - 12$ years – both obtained by the STPS team and other observing runs –, which are publicly available in the ESO ADP Archive and examined the chemical and physical processes underlying their host stars, which in some cases indicated planetary engulfment events[1], insights into their formation and dynamical evolution pathways given the high orbital eccentricity of some candidates, and allowed to disentangling stellar activity from exoplanetary nature signals by employing Gaussian Process regression to some cases where a proposed planet's orbital period might coincide with stellar modulations. Our findings are summarised in Table 5.1, followed by radial velocity curves of the discovered exoplanets host stars in Figures 5.1 and 5.2, and a detailed description of the analyses of each exoplanetary system candidate are provided in Sections 5.1 and 5.2.

---

[1] Omission of analysis on planetary engulfment episodes in this chapter implies that they have not been detected.

ID	P (days)	a (AU)	$m \sin i$ ($\mathrm{M_{Jup}}$)	e	$\mathrm{T_0}$ (days)	K ($\mathrm{m\,s^{-1}}$)	$\dot{\gamma}$ ($\mathrm{m\,s^{-1}\,d^{-1}}$)	$\mathrm{M_{eng}}$ ($\mathrm{M_\oplus}$)	$\mathrm{M_{add}}$ ($\mathrm{M_{Jup}}$)
HIP1954 b	$62.307^{+0.099}_{-0.094}$	$0.304^{+0.00032}_{-0.00030}$	$0.045^{+0.015}_{-0.013}$	0.0	$2456087^{+0.113}_{-0.117}$	$2.4^{+0.83}_{-0.76}$	$-4 \times 10^{-5}{}^{+0.0001}_{-0.0008}$	3.5	–
HIP5301 b	2473±11	3.533±0.011	$1.864^{+0.044}_{-0.046}$	0.039 ± 0.023	2455452^{+21}_{-19}	$28.78^{+0.67}_{-0.71}$	–	0.7	–
HIP30158 b	1050^{+29}_{-27}	$1.997^{+0.037}_{-0.035}$	$0.269^{+0.018}_{-0.017}$	0.064±0.058	2457083^{+27}_{-21}	$5.55^{+0.37}_{-0.38}$	–	–	–
HIP30502 b	$62.4^{+5.75}_{-5.13}$	$0.304^{+0.0164}_{-0.0190}$	$0.093^{+0.021}_{-0.021}$	0.0	$2456087^{+113.52}_{-106.3}$	$4.9^{+1.14}_{-1.09}$	–	6.0	–
HIP43297 b	$816.1^{+8.3}_{-9.0}$	$1.718^{+0.012}_{-0.013}$	2.0±0.21	0.67±0.07	$2457826.7^{+7.5}_{-8.7}$	$58.1^{+8.2}_{-8.7}$	–	17.8	–
HIP44935 b	$507^{+8.3}_{-7.5}$	$1.248^{+0.014}_{-0.012}$	$0.235^{+0.050}_{-0.034}$	0.4±0.18	2457264^{+41}_{-35}	$6.5^{+2.2}_{-1.1}$	$-0.0013^{+0.0013}_{-0.0014}$	2.7	–
HIP49756 b	$884^{+9.8}_{-9.9}$	$1.809^{+0.013}_{-0.014}$	$0.248^{+0.026}_{-0.023}$	0.34±0.15	2456193^{+34}_{-36}	$5.48^{+0.68}_{-0.58}$	–	–	–
HIP64673 b	$45.05^{+0.737}_{-0.824}$	$0.2534^{+0.00272}_{-0.00278}$	$0.09^{+0.02}_{-0.02}$	0.0	$2456087^{+4.981}_{-5.827}$	$2.4^{+1.16}_{-1.17}$	$-3\times10^{-6}\ {}^{+0.00095}_{-0.00129}$	–	–
HIP72043 b	1023^{+40}_{-36}	$2.004^{+0.053}_{-0.048}$	$0.861^{+0.074}_{-0.096}$	0.31±0.17	2457973^{+83}_{-73}	$18.0^{+1.5}_{-1.6}$	–	–	–
HIP87769 b	1419^{+100}_{-120}	$2.5^{+0.12}_{-0.14}$	$0.337^{+0.11}_{-0.08}$	0.46±0.17	2456088±100	$6.9^{+2.0}_{-1.8}$	$-0.2456^{+0.0059}_{-0.0053}$	1.7	23.41
HIP102152 b	$22.93^{+0.0103}_{-0.01023}$	$0.156^{+4.70\times10^{-5}}_{-4.65\times10^{-5}}$	$0.0256^{+0.0028}_{-0.0029}$	0.0	$2456489^{+0.0101}_{-0.0102}$	$1.87^{+0.204}_{-0.218}$	–	–	–
HIP104045 b	$2315^{+378.18}_{-243.49}$	$3.459^{+0.367}_{-0.247}$	$0.498^{+0.075}_{-0.073}$	0.0	$2455245^{+291.22}_{-375.81}$	$7.5^{+1.07}_{-1.04}$	$-0.00001^{+0.0005}_{-0.0012}$	3.5	–
HIP104045 c	$316^{+52.29}_{-98.71}$	$0.917^{+0.098}_{-0.202}$	$0.135^{+0.040}_{-0.024}$	0.0	2456313^{+1187}_{-543}	$4.0^{+1.10}_{-1.03}$	$-0.00001^{+0.0005}_{-0.0012}$	3.5	–
HIP108158 b	7141^{+47}_{-48}	7.309±0.033	$0.63^{+0.24}_{-0.27}$	0.63±0.21	2452917^{+150}_{-190}	$9.3^{+2.2}_{-2.5}$	$-0.00650^{+0.0012}_{-0.0019}$	–	1.78
HIP114615 b	$24.657^{+0.027}_{-0.037}$	0.16726±0.00031	$0.093^{+0.015}_{-0.013}$	0.3±0.17	$2455452^{+1.8}_{-1.0}$	$6.9^{+1.1}_{-1.0}$	–	–	–
HIP117367 b	$56.2^{+2.05}_{-0.47}$	$0.291^{+0.0070}_{-0.00163}$	$0.052^{+0.006}_{-0.006}$	0.0	$2456855^{+11.96}_{-7.39}$	$2.72^{+0.36}_{-0.35}$	$-1\times10^{-7}\ {}^{+4.07\times10^{-5}}_{-4.96\times10^{-5}}$	–	–

Table 5.1 - Orbital and physical properties of the 16 exoplanetary system candidates discovered in this project. Columns indicate the orbital period, orbital semi-major axis, planet's minimum mass, orbital eccentricity, time of inferior conjunction, Doppler-induced semi-amplitude, radial velocity slope, planetary mass engulfed (see text), and mass of an additional unobserved-with-current-data companion (see also Fig. 5.3).

Table 5.2 presents the photometric parameters of the 15 solar twins found to host planets analysed in this project, which includes magnitude measurements from the *Gaia* DR3 catalogue (van Leeuwen et al., 2022; Gaia Collaboration et al., 2022) – G band ($330 - 1050$ nm), BP ($330 - 680$ nm) and RP ($630 - 1050$ nm) bands – and from the Hipparcos catalogue (Perryman et al., 1997) – B ($390 - 490$ nm) and V ($500 - 590$ nm) –. Additionally, the table includes near-infrared J (1.25 μm), H (1.65 μm), and Ks (2.17 μm) bands from 2MASS (Skrutskie et al., 2006) and mid-infrared in four mid-infrared bands centred at 3.4, 4.6, 12, and 22 μm WISE (Cutri et al., 2021) magnitudes. Moreover, the masses, effective temperatures, surface gravity, age and metallicity of the stars presented in Table 5.3 were retrieved from Spina et al. (2018), whilst the stellar radius is from the TESS Input Catalogue (TIC v8.2; Paegert et al. 2022). The kinematic properties of the resulting planet hosting solar twin stars are presented in Table 5.4 – locations, parallax, proper motion, and the Renormalised Unit Weight Error (RUWE[2]) – were retrieved from the *Gaia* DR3 catalogue. Lastly, Table 5.5 lists [X/Fe] abundances of 17 elements obtained from Bedell et al. (2018) with an accuracy of 2$\%$ for the resulting planet hosts discovered in this work. These chemical abundances were utilized to unveil planet engulfment events in some cases, following the method outlined in Galarza et al. (2016) and Galarza et al. (2021).

HIP1954: A planet with a similar mass to Neptune has been found in circular orbit through joint Keplerian modelling and Gaussian Process regression. In addition, the chemical abundances revealed that it may have engulfed rocky material equivalent to 3.5 M$\oplus$.

HIP5301: A Jupiter-like planet has been previously reported in 2010, and we present a re-analysis of this system with an increase of $\sim$ 200$\%$ in the number of analysed RVs, and a bigger time-span. The presence of a Jupiter analogue was confirmed, and similar to above, the host star may have undergone little planet engulfment process (0.7 $M_\oplus$ of rocky material).

HIP30158: A planet with a mass similar to Saturn in circular orbit has been discovered.

HIP30502: It is orbited by a planet with a mass twice that of Neptune in a circular orbit. The host star's chemical abundances revealed that it may have engulfed 6.0 M$\oplus$ of rocky material.

---

[2] The RUWE is based on the residuals of the astrometric solutions and compares the expected errors to the actual errors. It assists to identify and flag any data points that may be affected by errors, such as binary or variable stars and improve the overall accuracy and precision of the *Gaia* mission data.

HIP43297: Through Keplerian modelling, the existence of a gas giant planet with a mass twice that of Jupiter was discovered. This planet has an extremely eccentric orbit ($e = 0.67$). Analysis of the chemical abundances of the host star demonstrated that it may have engulfed about 18 $\mathrm{M}_\oplus$ of rocky material. The high mass of the rocky engulfed material, associated with the eccentric orbit of the planet indicates a significant dynamism in the timing evolution of this planetary system. The same planet may be even responsible for the orbital excitation of an inner ex-companion that surpassed the Roche limit of the central star.

HIP44935: It has a planet with a mass similar to Saturn in an eccentric orbit. Through the analysis of chemical abundances, the host star may have engulfed 2.7 $\mathrm{M}_\oplus$ of rocky material.

HIP49756: Hosts a planet with a mass similar to Saturn in eccentric orbit as well.

HIP64673: The central star is orbited by a planet twice as massive as Neptune in a circular orbit.

HIP72043: The Keplerian analysis revealed a Jupiter-like planet in a relatively eccentric orbit.

HIP87769: The star is orbited by a Saturn-mass planet in a relatively eccentric orbit. The chemical abundances of the host star, suggest that it may have engulfed $\sim$1.7 M$\oplus$ of rocky material. Furthermore, the presence of a brown dwarf with a mass of 23.4 $\mathrm{M}_{\mathrm{Jup}}$ in a more distant orbit was estimated from the RV curve inclination ($\dot{\gamma} = -0.2456$). Only treated as a hypothesis that warrants further investigation, this brown dwarf may be responsible for the orbital excitation of the Saturn-like planet in inner orbits, which in turn may be responsible for the migration of the inner ex-companion that surpassed the Roche limit of the central star.

HIP102152: We implemented a joint Keplerian and Gaussian Process regression to model the RVs of the host star, which is the oldest of the analysed sample to present a warm mini-Neptune ($m \sin i = 8 \pm 1$ M$\oplus$) candidate under circular orbit.

HIP104045: Keplerian plus Gaussian Process analysis revealed two candidates: a Jupiter analogue and a cold super-Neptune. The results were published in Ferreira et al. (2023).

HIP117367: The Keplerian model alone was statistically insufficient to explain the RV modulations, so the analysis was performed with Gaussian Process regression to account for intrinsic stellar activity. There is no statistically significant correlation between the RVs and both activity proxies CCF BIS and CCF FWHM nor timing modulations at the proposed planet's orbital period

($P = 58.4d$). If confirmed with desired and encouraged follow-up observations, the system hosts a warm Neptune-like under circular orbit ($P = 55.95$ days, $m \sin i = 17.42$ M⊕).

HIP114615: A planet with a mass about two times that of Neptune in a relatively eccentric orbit.

HIP108158: The RV modulations correspond to a gas giant planet in a relatively eccentric orbit. An external planet with mass 1.78 $\mathrm{M_{Jup}}$ has been inferred through the inclination of the RVs ($\dot{\gamma} = -0.0065$).

Star	G	$(BP-RP)$	B	$B-V$	J	H	K	$W1$	$W2$	$W3$	$W4$
HIP1954	7.12	0.83	7.87	0.60	6.11±0.02	5.78±0.03	5.72±0.02	5.63±0.06	5.59±0.03	5.71±0.01	5.70±0.04
HIP5301	8.29	0.82	9.05	0.60	7.27±0.02	6.99±0.05	6.88±0.02	6.83±0.04	6.88±0.02	6.88±0.02	6.91±0.08
HIP30158	8.35	0.84	9.05	0.68	7.29±0.02	7.02±0.03	6.91±0.02	6.90±0.04	6.92±0.02	6.93±0.02	6.99±0.07
HIP30502	8.50	0.82	9.66	1.00	7.47±0.03	7.14±0.03	7.07±0.02	7.04±0.03	7.07±0.02	7.08±0.02	7.11±0.07
HIP43297	7.30	0.85	8.10	0.65	6.24±0.02	5.95±0.05	5.84±0.02	5.89±0.05	5.74±0.02	5.81±0.01	5.81±0.04
HIP44935	8.57	0.81	9.37	0.65	7.55±0.02	7.26±0.03	7.17±0.02	7.14±0.03	7.18±0.02	7.18±0.02	7.28±0.12
HIP49756	7.39	0.81	8.10	0.55	6.38±0.02	6.08±0.03	6.00±0.02	5.94±0.05	5.93±0.02	6.04±0.01	5.96±0.04
HIP64673	8.19	0.78	8.95	0.62	7.18±0.02	6.97±0.04	6.83±0.02	6.82±0.03	6.86±0.02	6.87±0.02	6.97±0.07
HIP72043	7.37	0.80	8.16	0.64	6.33±0.02	6.07±0.02	6.00±0.02	5.92±0.06	5.89±0.02	5.98±0.01	5.91±0.03
HIP87769	8.30	0.82	9.11	0.68	7.17±0.03	6.83±0.04	6.76±0.03	6.66±0.04	6.74±0.02	6.73±0.01	6.59±0.06
HIP102152	9.02	0.83	9.79	0.61	7.96±0.02	7.68±0.03	7.59±0.02	7.55±0.02	7.59±0.02	7.61±0.02	7.92±0.20
HIP104045	8.19	0.80	8.95	0.61	7.20±0.02	6.91±0.03	6.83±0.02	6.76±0.04	6.82±0.02	6.83±0.02	6.85±0.09
HIP108158	7.27	0.85	8.14	0.71	6.23±0.02	5.93±0.04	5.79±0.01	5.78±0.05	5.72±0.03	5.81±0.01	5.77±0.04
HIP114615	9.45	0.80	10.21	0.61	8.46±0.03	8.19±0.03	8.09±0.02	8.03±0.02	8.07±0.02	8.08±0.02	8.70±0.35
HIP117367	7.55	0.79	8.34	0.62	6.56±0.04	6.29±0.06	6.20±0.02	6.19±0.05	6.18±0.02	6.22±0.01	6.22±0.05

Table 5.2 - Photometry for the planet-hosting solar twins discovered in this project, which includes *Gaia* G magnitude and $BP-RP$ colour, Hipparcos B magnitude and $B-V$ colour, as well as near-infrared 2MASS J, H, and K magnitudes, plus AllWISE mid-infrared bands.

Star	$M/M\odot$	$R/R\odot$	T_{eff}/K	$\log(g)$	$L/L\odot$	Age/Gyr	[Fe/H]
HIP1954	0.970±0.002	0.945±0.052	5720±2	4.460±0.008	0.861±0.022	4.400±0.500	−0.090±0.003
HIP5301	0.960±0.001	0.994±0.040	5723±3	4.395±0.011	0.959±0.023	7.200±0.400	−0.074±0.003
HIP30158	0.963±0.005	1.022±0.053	5678±4	4.365±0.011	0.994±0.025	7.900±0.300	−0.004±0.003
HIP30502	0.965±0.005	1.025±0.055	5731±4	4.400±0.013	1.013±0.026	7.000±0.400	−0.057±0.004
HIP43297	1.014±0.005	0.941±0.040	5705±4	4.505±0.009	0.840±0.020	1.900±0.500	0.082±0.003
HIP44935	1.009±0.003	1.102±0.062	5771±4	4.370±0.012	1.223±0.034	6.500±0.300	0.038±0.004
HIP49756	1.010±0.002	1.032±0.056	5789±3	4.435±0.009	1.075±0.030	4.500±0.400	0.023±0.003
HIP64673	1.068±0.006	1.271±0.050	5912±5	4.290±0.014	1.783±0.055	6.000±0.400	−0.017±0.004
HIP72043	1.026±0.005	1.186±0.049	5845±4	4.340±0.011	1.488±0.040	6.300±0.300	−0.026±0.003
HIP87769	1.039±0.003	1.097±0.058	5828±3	4.400±0.010	1.231±0.033	4.600±0.600	0.072±0.004
HIP102152	0.978±0.004	1.083±0.062	5718±4	4.325±0.011	1.133±0.030	8.600±0.300	−0.016±0.003
HIP104045	1.027±0.005	1.047±0.042	5826±3	4.410±0.010	1.142±0.032	4.500±0.500	0.051±0.003
HIP108158	1.021±0.002	1.215±0.051	5675±4	4.285±0.011	1.384±0.034	8.100±0.300	0.055±0.003
HIP114615	1.027±0.005	0.949±0.040	5819±5	4.510±0.009	0.928±0.026	1.100±0.700	−0.063±0.004
HIP117367	1.040±0.001	1.161±0.048	5867±3	4.350±0.010	1.442±0.041	5.700±0.300	0.024±0.003

Table 5.3 - Stellar fundamental properties (mass, radius, effective temperature, surface gravity, luminosity, age and metallicity) for the planet-hosting solar twins discovered in this project (see text).

Star	α/deg	δ/deg	π/mas	μ_{α^*}/mas yr^{-1}	μ_δ/mas yr^{-1}	RUWE
HIP1954	6.17888	−53.98412	35.262±0.017	210.591±0.013	−28.155±0.015	0.993
HIP5301	16.95363	−8.23361	19.454±0.023	192.581±0.025	20.077±0.024	0.957
HIP30158	95.19632	−44.78881	18.679±0.016	187.973±0.018	81.731±0.020	0.934
HIP30502	96.18235	−43.40530	17.210±0.013	3.608±0.017	83.413±0.017	0.971
HIP43297	132.30154	3.48502	32.978±0.024	−148.238±0.025	60.209±0.020	1.004
HIP44935	137.29402	10.82097	15.253±0.018	181.376±0.020	−164.130±0.019	0.956
HIP49756	152.34672	2.37105	27.818±0.039	−202.983±0.042	−6.533±0.042	1.032
HIP54582	167.55453	−7.39017	30.804±0.021	−238.340±0.028	−173.313±0.018	1.029
HIP64673	198.83367	−55.34219	14.936±0.024	−230.807±0.018	−135.691±0.018	1.169
HIP72043	221.04844	18.46138	23.788±0.026	−59.034±0.027	−165.142±0.027	1.167
HIP87769	268.93903	11.80375	17.379±0.039	15.499±0.033	−48.216±0.033	1.655
HIP102152	310.47852	−27.21602	12.771±0.016	175.628±0.016	−15.593±0.013	1.039
HIP104045	316.18333	−4.82976	18.591±0.030	−139.585±0.029	−194.291±0.022	1.060
HIP108158	328.68528	−73.44043	26.087±0.017	−196.563±0.018	−437.609±0.022	1.019
HIP114615	348.29094	−30.44986	11.540±0.013	8.620±0.011	−13.337±0.012	0.914
HIP117367	356.96993	4.17534	22.242±0.074	352.541±0.078	−31.590±0.047	1.185

Table 5.4 - Kinematic properties of a group of solar twins hosting exoplanets identified in this project. The data is sourced from the *Gaia* DR3, which includes location, parallaxes (in milliarcseconds), proper motion components (in milliarcseconds per year), and the system's RUWE.

HIP	[C/H]	[O/H]	[Na/H]	[Mg/H]	[Al/H]	[Si/H]	[S/H]	[Ca/H]	[Sc/H]
1954	-0.130±0.013	-0.065±0.008	-0.102±0.010	-0.076±0.009	-0.098±0.007	-0.086±0.003	-0.086±0.003	-0.064±0.003	-0.096±0.008
5301	-0.084±0.013	-0.045±0.014	-0.077±0.013	-0.053±0.005	-0.063±0.013	-0.077±0.004	-0.034±0.011	-0.061±0.004	-0.075±0.021
30158	0.021±0.007	0.055±0.011	0.012±0.008	0.030±0.008	0.050±0.002	0.009±0.002	0.035±0.021	0.009±0.005	-0.015±0.008
30502	-0.117±0.007	-0.053±0.012	-0.087±0.014	-0.035±0.009	-0.018±0.009	-0.052±0.003	-0.117±0.018	-0.034±0.006	-0.059±0.013
43297	-0.056±0.012	0.027±0.011	0.017±0.010	0.062±0.017	0.057±0.004	0.051±0.003	0.036±0.022	0.094±0.006	0.058±0.012
44935	0.037±0.014	0.043±0.009	0.069±0.008	0.067±0.003	0.086±0.003	0.056±0.003	0.045±0.010	0.045±0.005	0.049±0.013
49756	0.035±0.018	0.023±0.006	0.034±0.011	0.034±0.007	0.031±0.004	0.029±0.003	0.024±0.007	0.031±0.004	0.032±0.005
64673	0.055±0.014	0.034±0.009	0.024±0.008	0.013±0.005	0.016±0.004	0.007±0.004	0.001±0.013	-0.014±0.006	-0.010±0.013
72043	0.015±0.006	0.024±0.014	-0.012±0.003	0.033±0.013	0.033±0.002	0.007±0.004	-0.012±0.013	-0.007±0.004	0.031±0.005
87769	0.064±0.014	0.066±0.015	0.095±0.010	0.083±0.008	0.081±0.004	0.072±0.003	0.096±0.010	0.092±0.004	0.085±0.016
102152	-0.001±0.008	0.035±0.010	-0.044±0.014	0.015±0.010	0.010±0.004	-0.013±0.003	-0.028±0.022	-0.010±0.005	-0.032±0.011
104045	0.009±0.008	0.035±0.008	0.026±0.009	0.040±0.004	0.038±0.004	0.045±0.002	0.044±0.028	0.059±0.005	0.026±0.012
108158	0.156±0.005	0.198±0.011	0.088±0.008	0.193±0.011	0.210±0.012	0.118±0.002	0.151±0.022	0.102±0.005	0.138±0.010
114615	-0.092±0.012	-0.025±0.009	-0.102±0.015	-0.064±0.012	-0.084±0.012	-0.068±0.003	-0.044±0.028	-0.044±0.006	-0.068±0.013
117367	0.049±0.004	0.063±0.012	0.066±0.011	0.056±0.007	0.048±0.013	0.046±0.004	0.038±0.018	0.025±0.004	0.038±0.006

Table 5.5 - Chemical Abundances retrieved from Bedell et al. (2018) for the planet-hosting solar twins discovered in this project.

HIP	[Ti/H]	[V/H]	[Cr/H]	[Mn/H]	[Co/H]	[Ni/H]	[Cu/H]	[Zn/H]
1954	-0.062±0.003	-0.086±0.003	-0.082±0.005	-0.107±0.003	-0.108±0.003	-0.118±0.003	-0.109±0.010	-0.122±0.006
5301	-0.067±0.005	-0.085±0.005	-0.074±0.006	-0.092±0.005	-0.089±0.005	-0.084±0.003	-0.080±0.004	-0.093±0.008
30158	0.016±0.005	0.007±0.006	-0.003±0.005	-0.001±0.005	0.008±0.004	0.003±0.003	0.037±0.014	0.026±0.013
30502	-0.028±0.005	-0.059±0.005	-0.061±0.004	-0.085±0.005	-0.066±0.006	-0.074±0.003	-0.064±0.002	-0.093±0.005
43297	0.093±0.006	0.091±0.007	0.098±0.006	0.065±0.006	0.043±0.005	0.048±0.004	0.017±0.011	0.034±0.030
44935	0.058±0.005	0.054±0.006	0.041±0.005	0.044±0.004	0.056±0.005	0.052±0.004	0.076±0.006	0.057±0.021
49756	0.034±0.004	0.036±0.004	0.027±0.004	0.029±0.005	0.036±0.004	0.027±0.003	0.027±0.005	0.007±0.015
64673	-0.014±0.006	-0.017±0.007	-0.030±0.005	-0.034±0.006	0.009±0.006	-0.014±0.004	0.015±0.015	-0.013±0.007
72043	0.004±0.005	-0.003±0.005	-0.034±0.005	-0.055±0.004	0.004±0.004	-0.025±0.003	0.005±0.006	-0.012±0.006
87769	0.078±0.005	0.088±0.007	0.084±0.005	0.080±0.005	0.070±0.005	0.080±0.004	0.067±0.004	0.083±0.006
102152	-0.007±0.005	-0.025±0.006	-0.021±0.004	-0.043±0.005	-0.034±0.008	-0.033±0.004	-0.014±0.004	-0.012±0.011
104045	0.048±0.005	0.042±0.005	0.048±0.004	0.045±0.004	0.037±0.004	0.048±0.003	0.031±0.006	0.022±0.010
108158	0.160±0.006	0.133±0.008	0.070±0.005	0.033±0.004	0.128±0.004	0.082±0.005	0.164±0.026	0.191±0.017
114615	-0.048±0.007	-0.066±0.007	-0.061±0.007	-0.099±0.005	-0.086±0.007	-0.105±0.005	-0.120±0.033	-0.090±0.014
117367	0.026±0.005	0.027±0.005	0.018±0.004	0.027±0.003	0.050±0.005	0.037±0.003	0.065±0.010	0.044±0.004

Table 5.5 - Continued from Table 5.5.

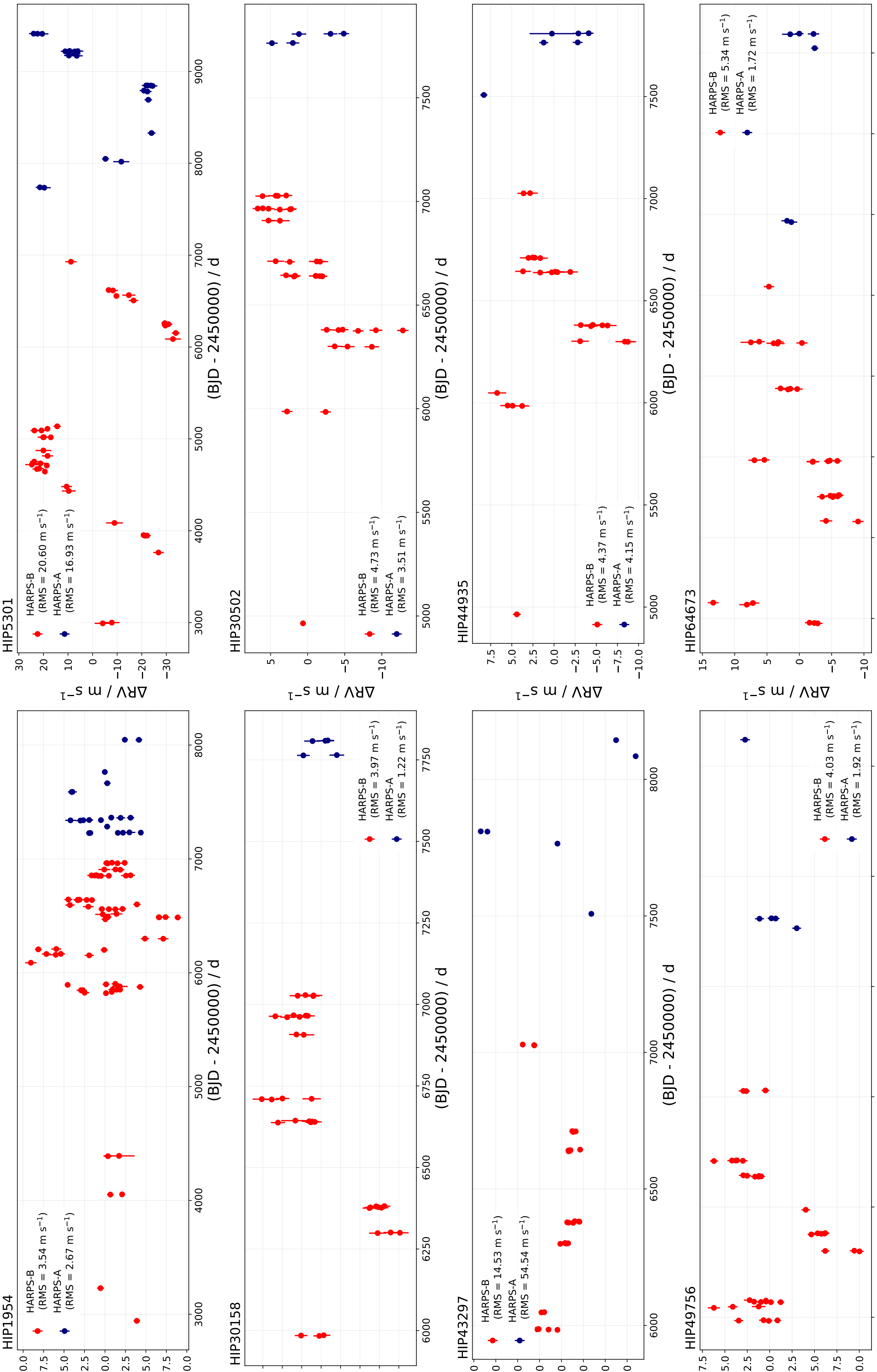

Figure 5.1: RV curves for HIP1954, HIP5301, HIP30158, HIP30502, HIP43297, HIP44935, HIP49756, and HIP64673. Red and blue dots represent the data taken before and after the ESO/HARPS' upgrade in mid-2015, respectively (if existent), and the RMS is also indicated in the legend.

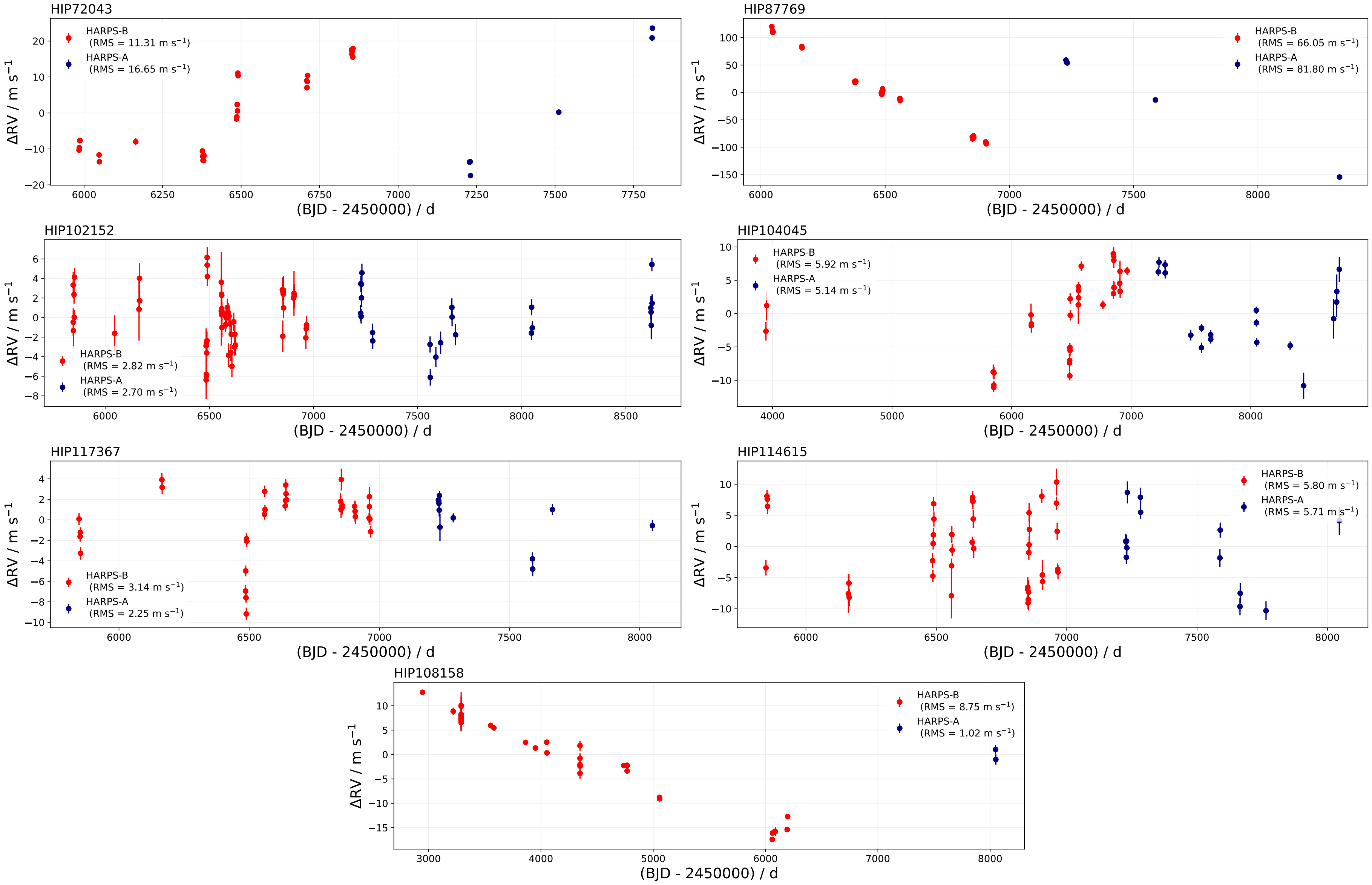

Figure 5.2: RV curves for HIP72043, HIP87769, HIP102152, HIP104045, HIP117367, HIP114615, and HIP108158. Red and blue dots represent the data taken before and after the ESO/HARPS' upgrade in mid-2015, respectively (if existent), and the RMS is also indicated in the legend.

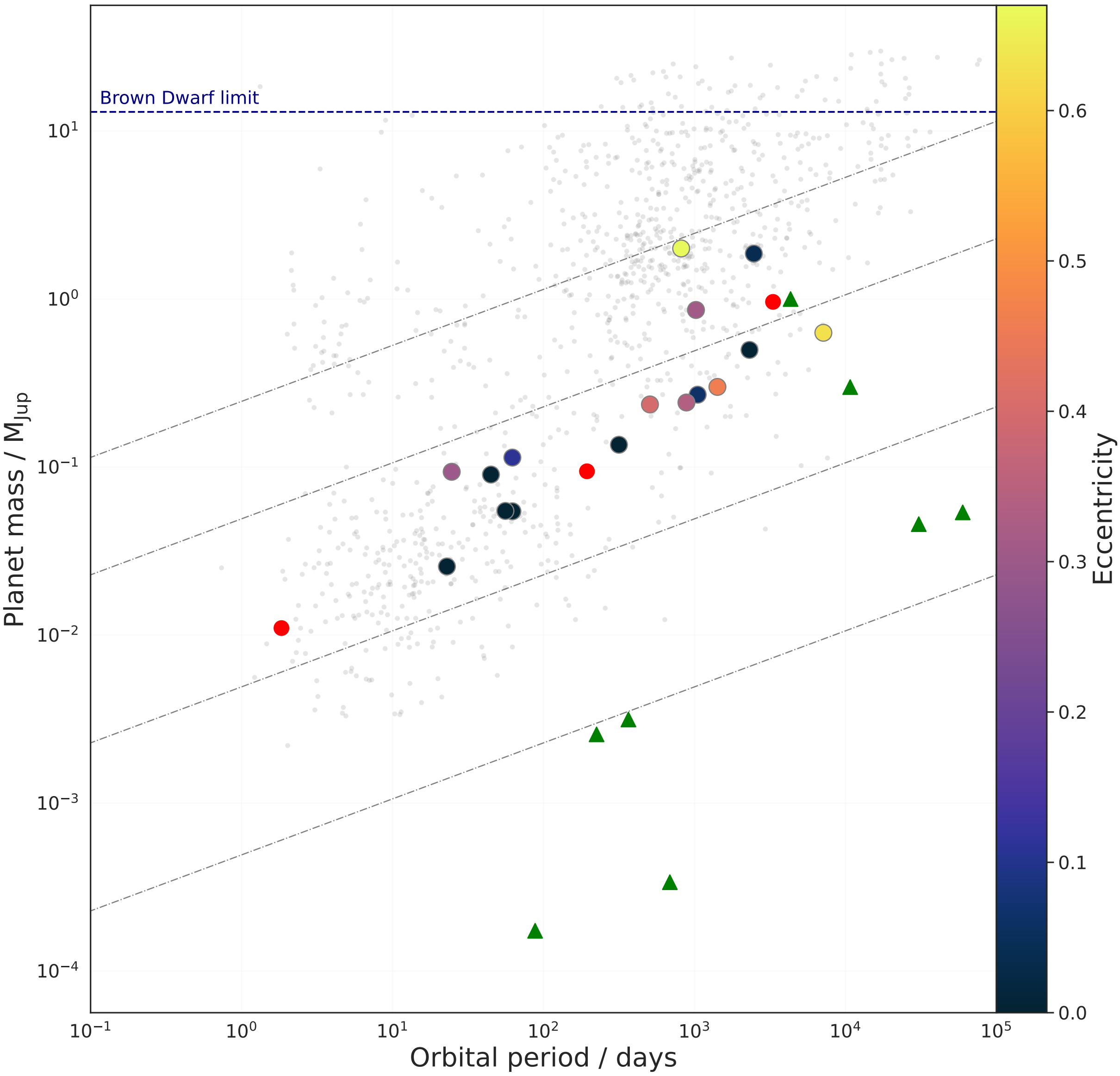

Figure 5.3: Mass-orbital period relationship for the exoplanet candidates found in this writing. Colours indicate the orbital eccentricity, while green triangles represent planets in the Solar System for reference, and red-filled circles represent previously known STPS planets (HIP11915 *b*, HIP68468 *b* and *c*). The dashed lines indicate Doppler amplitude limits $K = [0.1, 1, 10, 50]$ m s^{-1}). Exoplanets known to date (June 20, 2023; NASA Exoplanet Archive [`https://exoplanetarchive.ipac.caltech.edu/`]) discovered only by RV method are represented in grey.

5.1 From Warm to Cold Ice Giants and a Mini-Neptune

5.1.1 HIP1954 b: A warm Neptunian planet under circular orbit

HIP1954 was observed by the ESO/HARPS spectrograph during 88 epochs between October 2003 and October 2017 (see Section 3.1 for a detailed description of the instrument). Data is publicly available at the ESO Advanced Data Products Archive (ADP[3]). From its Renormalised Unit Weight Error (RUWE; Fitton et al. 2022) of 0.993, we observe that HIP1954 is unlike to be part of a binary system. Data collected before the HARPS upgrade in mid-2015 indicates a root-mean-square (RMS) value of 3.54 m s^{-1}, while for data collected after the upgrade, RMS = 2.67 m s^{-1}. The kinematic (locations, parallax, proper motion, and RUWE), physical (mass, radius, effective temperature, surface gravity, luminosity, age, and metallicity), photometric (*Gaia*, 2MASS and WISE magnitudes) and chemical properties (abundances of 17 elements following Bedell et al. 2018) are presented in Tables 5.4, 5.3, 5.2, and 5.5, respectively, for reference (as well as for the next exoplanet candidates presented in this chapter).

With a Generalised Lomb-Scargle algorithm (GLS; see Section 4.1) in a 0.1-d to 20 000-d grid (hereafter standard GLS grid), we found a prominent signal close to 62.3 days. The change in the Bayesian Information Criterion (BIC; Ivezić et al. 2014) with a non-varying noise model corresponds to $\|\Delta \mathrm{BIC}\| = 1625.39$, which demonstrates the high-statistical significance of this signal under sine-like assumptions. False-Alarm Probability (FAP; Ivezić et al. 2014) tests were performed via bootstrapping re-sampling, where we found a probability to obtain a higher peak than 0.364 (associated with the 62.3-d period) with GLS-powers of 0.248 and 0.214 in the RV periodogram have 1% and 5% probability, respectively. With the Nyquist frequency sampling test, (see Subsection 4.1.1), we found out that the proposed planet's orbital period is not an alias resulting from instrumental modulations, i.e., it is not an alias from one sidereal day (1.0-d), one synodic month (29.5-d), nor one sidereal year (365.25-d).

The non-rigid nature of a star causes differential rotation, so, if a spot or plage, for instance, is located at a higher or lower latitude, it will take generally longer than the equator to complete a full revolution. As a result, RV measurements taken at different times may reflect such stellar modulations that hamper or even mimic planetary signals. An inspection of the activity proxies

---

[3] `http://archive.eso.org/wdb/wdb/adp/`

CCF BIS, CCF FWHM, $S-$index and Hα (see Section 3.3) at the proposed planet's orbital period revealed, for all, no time or phase modulations, and we did not observe significant correlations of these with the RVs (see Fig. 5.4). To quantify the RV-activity proxies correlations significance, we applied the Pearson $r-$correlation test. While for CCF BIS and CCF FWHM we note weak correlations (where $r < 0.30$), moderate correlations ($r \in [0.3, 0.7]$) are seen between RVs for $S-$index and Hα. Therefore, we modelled stellar magnetic activity effects with Gaussian Process (GP; see Section 4.3), and considering that HIP1954 b companion follow Keplerian orbits (see Section 2.2), we proceeded with the orbital fit making use of `RadVel`[4], which implements a maximum a posteriori probability density via Markov chain simulations (see Section 4.2).

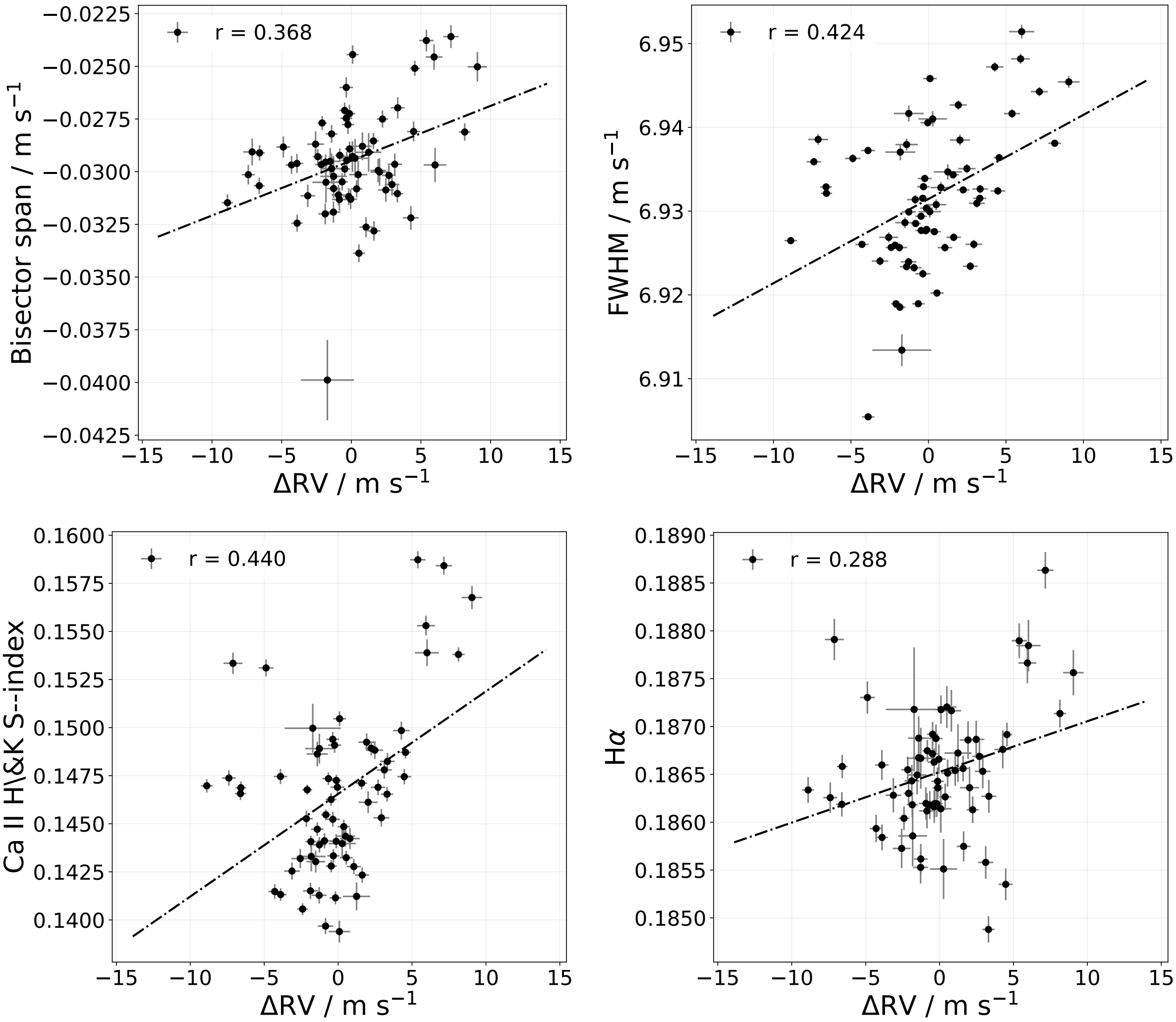

Figure 5.4: RV measurements and their relationship with stellar activity indicators for HIP1954.

---

[4] https://radvel.readthedocs.io/en/latest/

We chose a standard Keplerian basis consisting of a set of five orbital elements (hereafter standard Keplerian basis): orbital period P, time of inferior conjunction T_c, Doppler-induced semi-amplitude K, plus orbital eccentricity and periastron argument coupled in the orthogonal basis $\sqrt{e}\cos\omega$ and $\sqrt{e}\sin\omega$, which ensures that the model can be compared with a circular one, as the last parameter cannot be defined for zero-eccentricity orbits (e.g., Lucy and Sweeney 1971). Accounting for instrumental shifts, we set jitter noise corrections in the fit for each instrumental epoch to freely vary ($\sigma_{\mathrm{HARPS-A/B}}$). These are added in quadrature to the RV uncertainties and are usually associated with stellar modulations. Moreover, considering the presence of an outer companion, from which only a small portion of its orbit could be observed, we set as free parameters the barycentre motion $\dot{\gamma}$ and a curvature $\ddot{\gamma}$ parameter term accounting for a non-linear trend of the RV model over its entire baseline of approximately 14 years. Furthermore, to evaluate the impact of stellar activity on the variations in RV, we employed a quasi-periodic covariance kernel. This kernel function depends on four hyper-parameters: the Doppler-induced magnetic activity of the signal (η_1), the time-scale for the growth and decay of active regions (η_2), the recurrence time-scale for active regions (η_3, which is typically similar to the star's rotation period[5]), and a smoothing parameter (η_4). Table 5.6 summarises the HIP1954 b model's prior parameters together with orbital solutions.

The MCMC sampling was implemented with 50 walkers over 1000 steps and 4 ensembles following `RadVel`'s genetic routine (see also Subsection 4.2.1). We required burn-out chains with Gelman-Rubin statistics $\hat{R} > 1.01$ and maximum $\hat{R}$ at the same value. The model was minimised with `scipy.optimize` function employing the Nelder-Mead (N-M; Virtanen et al. 2020) algorithm over 200 iterations.

We compared statistically an eccentric and a circular model, both accounting for non-zero quantities $\dot{\gamma}$ and $\ddot{\gamma}$, finding out that a circular model with RV slope is favoured by a Bayesian/Akaike Information Criterion (BIC/AIC; see Section 4.2) variation $||\Delta\mathrm{BIC/AIC}|| = 9.28/6.08$ (nominal BIC circular = 625.61, AIC circular = 598.33, with $\ln Z$ = -290.47), over an eccentric model. Therefore, following Liddle (2007), we conclude that a circular model with RV slope is bonafide in representing the data modulations of planetary nature for HIP1954 b, consisting in a warm

---

[5] In the absence of a measured star's rotation period, it was estimated from its velocity along our line-of-sight ($v\sin i$), and its radius, providing $P_{\mathrm{rot}} \sim 2\pi R/v\sin i$.

Neptune under circular orbit (see Fig. 5.5).

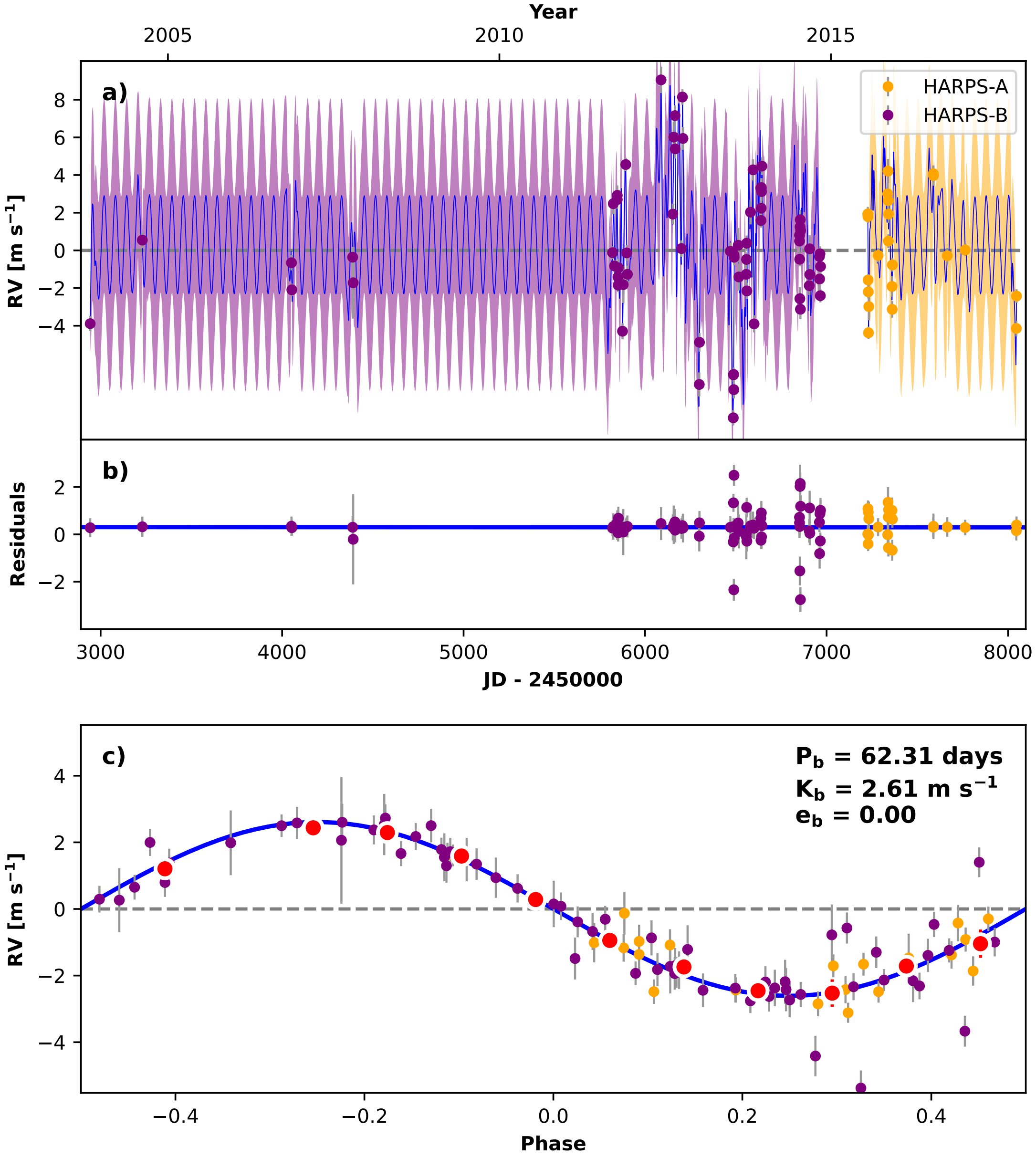

Figure 5.5: Best-fit 1-planet Keplerian plus stellar activity model for HIP1954. The uncertainties include the RV jitter ($\sigma_{\mathrm{HARPS-A/B}}$). In each panel, the purple and yellow dots/curves are the HARPS data before and after the upgrade in mid-2015. (a) The GP model is plotted as a thin blue line, while the purple and yellow curves are Keplerian models. (b) Residuals to the best fit. (c) RV phase diagrams to the ephemeris of planet b. RVs are binned in 0.08 phase units as red-filled circles. The phase-folded model for planet b is shown in blue, along with N-M Keplerian orbital parameters (period, eccentricity and planet-induced Doppler semi-amplitude).

Similarly to HIP104045 (host of a Jupiter analogue and a cold-Super Neptune; Ferreira et al. 2023), HIP1954 reveals a chemical anomaly that may be correlated with moderate planet engulfment episodes[6]. Fig. 5.6 demonstrates a relative refractory elements enhancement for the host star (black dots/curve), with linear-best-fit slope value of $\alpha = 1.231 \times 10^{-5}$ dex. Under validation of planet's accretion scenario, we employed the TERRA[7] code (Galarza et al., 2016). From the host star mass and metallicity, we first estimated its convective zone mass as 0.024 $M_\odot$. The red dots/curve in Fig. 5.6 presents the [X/C] ratios as a function of the condensation temperature T_C ($\alpha_{\rm model} = 2.407 \times 10^{-5}$ dex) and the effect due to the accretion of 3.5 $M_\oplus$ of rocky material with Earth-like composition into the convective zone of HIP1954. The process of planet engulfment is one of the mechanisms that can shape the formation and evolution of planetary systems, providing initial insights into the star-planet connection among solar twin stars. While planet-planet interactions and gravitational scattering can cause a planet to collide and merge (e.g., Raymond et al. 2008; Chatterjee et al. 2008; Li and Lai 2020), migration of planets towards the star is also a common scenario that could lead to planet engulfment (Quillen and Holman, 2000). Eventually, during its inward migration, a planet may reach a distance where it is no longer able to maintain its orbit and enters the Roche limit ($r_{\rm Roche} = a\left[2(M_{\rm star}/M_{\rm planet})\right]^{1/3}$), which is the critical distance from the host star where the planet's gravitational forces are overcome by the tidal forces of the star. At this point, the planet can experience significant tidal heating and deformation, which can lead to the release of gravitational potential energy and the loss of its outer layers[8]. Subsequently, this scenario could result in the accretion of rocky disintegrating material into the star's convective zone, as observed in the case of HIP1954 (whose $r_{\rm Roche} \sim 0.0022$ AU).

---

[6]

[7] `https://github.com/ramstojh/terra`

[8] Tidal heating and deformation can occur when a planet orbits close to its host star. The energy from the star's gravitational pull may cause the planet to deform and generates heat within it. This amount of tidal heating depends on both the planet and host star mass and radius, plus the distance between the star and its orbiting planet,

$$\Delta E = k\left(\frac{\left(GM_p^2 \cdot R_p \cdot \omega\right)^{5/2} \cdot Q}{(3 \cdot k_2 \cdot M_p \cdot R_p^2)}\right), \tag{5.1}$$

where ΔE is the tidal heating energy per unit time, k is a constant that depends on the planet's interior structure, G is the gravitational constant, M_p is the planet's mass, R_p is the planet's radius, ω is the planet's angular frequency, Q is the planet's tidal quality factor, and k_2 is the planet's Love number (e.g., Jackson et al. 2010; Ginzburg and Sari 2017, and references therein).

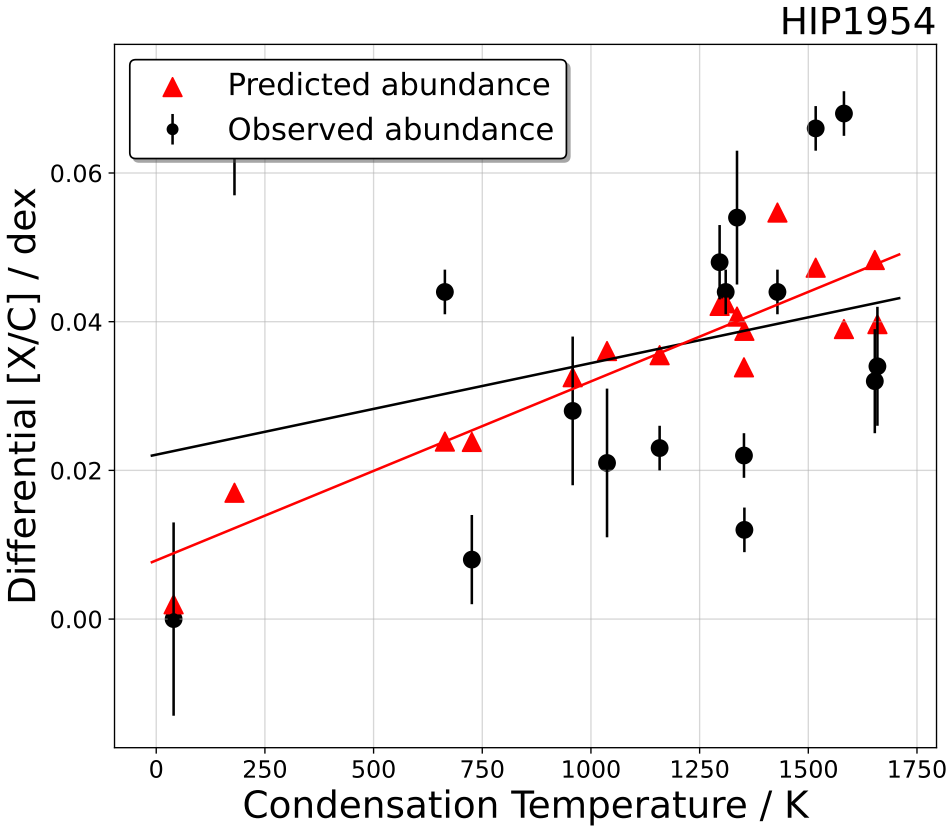

Figure 5.6: Comparison of the differential abundance of HIP1954 (black dots plus linear fit) as a function of dust condensation temperature in the protoplanetary disc, and the predicted abundances (red dots plus linear fit) estimated from a planetary accretion of 3.5 $M_{\oplus}$.

Parameter	**Prior**	**Credible Interval**	**Description**	**Units**
P	$\mathcal{N}(62.3, 0.1)$	$62.307^{+0.099}_{-0.094}$	Orbital period	days
T_c	$\mathcal{N}(2456088, 0.1)$	$2456087^{+0.113}_{-0.117}$	Time of inf. conj.	days
K	$\mathcal{N}(5, 1)$	$2.4^{+0.83}_{-0.76}$	RV semi-amplitude	m s^{-1}
$\sqrt{e}\cos\omega$	–	$\equiv 0$	1st comp. of the orthogonal basis	–
$\sqrt{e}\sin\omega$	–	$\equiv 0$	2nd comp. of the orthogonal basis	–
$\sigma_{\mathrm{HARPS-A}}$	$]-\infty, +\infty[$	$0.0002^{+0.0094}_{-0.0107}$	Instrumental noise after upgrade	m s^{-1}
$\sigma_{\mathrm{HARPS-B}}$	$]-\infty, +\infty[$	$0.0002^{+0.0097}_{-0.0095}$	Instrumental noise before upgrade	m s^{-1}
$\dot{\gamma}$	$]-\infty, +\infty[$	$-4\times10^{-5}\,{}^{+0.0001}_{-0.0008}$	RV Slope	m s^{-1} d^{-1}
$\ddot{\gamma}$	$]-\infty, +\infty[$	$-2\times10^{-11}\,{}^{+3.38\times10^{-10}}_{-5.74\times10^{-11}}$	RV Curvature	m s^{-1} d^{-2}
η_1	$\mathcal{J}(0.01, 100)$	$4.45^{+0.77}_{-0.56}$	Stellar activity's semi-amplitude	m s^{-1}
η_2	$\mathcal{N}(30, 1/\sqrt{2})$	$29.9^{+0.80}_{-0.79}$	Time-scale for growth and decay	days
η_3	$\mathcal{N}(24.1, 0.2)$	$24.04^{+0.24}_{-0.23}$	Recurrence time-scale for active regions	days
η_4	$\mathcal{N}(1.07, 1.4)$	$0.461^{+0.01}_{-0.01}$	Smoothing parameter	–
$m\sin i$	–	$14.6^{+5.06}_{-4.63}$	Planet b minimum mass	$\mathrm{M}_{\oplus}$
a	–	$0.304^{+0.00032}_{-0.00030}$	Planet b orbital semi-major axis	AU

Table 5.6 - Best-fit 1-planet Keplerian + GP parameters and derived quantities ($m\sin i$ and a) for HIP1954 system.

5.1.2 HIP30502 b: An warm Neptunian planet under circular orbit

HIP30502 was observed using HARPS during 40 epochs from May 2009 and February 2017. Its RUWE value of 0.971 indicates that the host star is not a binary and that its astrometry is reliable. The RMS of data before and after the HARPS upgrade are 4.727 and 3.513 m s^{-1}.

We utilised a GLS algorithm over the standard grid and found a significant signal close to 62 days with high statistical significance compared to a non-varying noise model ($\|\Delta\mathrm{BIC}\| = 841$). A FAP bootstrapping re-sampling test resulted in a low probability of obtaining a higher peak than 0.692 (associated with the prominent period obtained) with GLS powers of 0.463 and 0.365 associated with 1% and 5% significance levels, respectively. Additionally, we tested the resulting period with the Nyquist frequency sampling test, which ruled out the possibility of the proposed planet's orbital period being an alias from instrumental modulations. For HIP30502, we examined six different stellar activity indicators, namely CCF BIS, CCF FWHM, $S-$index, Hα, Na doublet, and He I (see Fig. 5.7). While moderate correlations with RVs are visible for the majority of the indicators, we observe that the $S-$index and the RVs present the weakest of the correlations, and a moderate-to-strong correlation is observed for the RVs and He I. This can indicate the presence of stellar activity such as starspots or plages, which can mimic the signal of a planet and lead to false-positive detections. Therefore, we modelled stellar magnetic activity effects with Gaussian Process as a proxy incorporated into the Keplerian model.

The Keplerian plus GP model employed in this case followed similarly to the case of HIP1954 (c.f., Section 5.1.1), where we implemented the standard Keplerian basis and in addition, a quasi-periodic kernel to evaluate the impact of stellar activity on the RVs (see Fig. 5.8). The MCMC sampling was implemented with 50 walkers over 1000 steps and 4 ensembles following `RadVel`'s genetic routine (see also Subsection 4.2.1). We required burn-out chains with Gelman-Rubin statistics $\hat{R} > 1.01$ and maximum $\hat{R}$ at the same value. The model was minimised with `scipy.optimize` function employing the Nelder-Mead (N-M) algorithm over 200 iterations. Table 5.13 summarises the HIP30502 b model's prior parameters together with orbital solutions.

In close resemblance to the case of HIP1954, HIP30502's chemical abundance pattern reveals anomalies that also seem to be strongly correlated with planet engulfment episodes. The star's convective zone mass was estimated as 0.026 M$\odot$ with the TERRA code. Fig. 5.9 demonstrates the [X/C] ratios as a function of the condensation temperature T_C for refractory/volatile elements

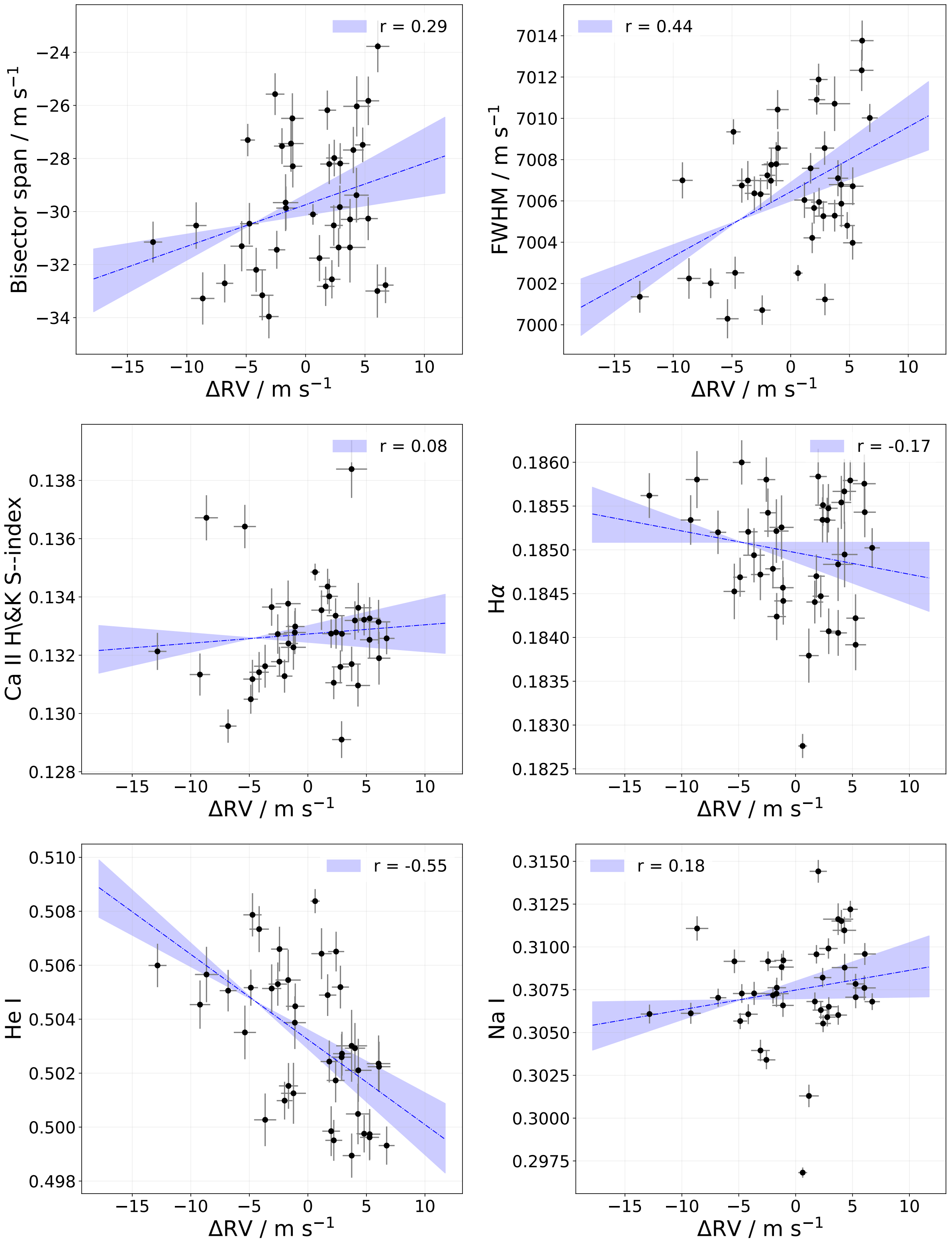

Figure 5.7: RV measurements and their relationship with stellar activity indicators for HIP30502.

(Black dots/curve; slope $\alpha = 4.02 \times 10^{-5}$ dex), and the effect due to the accretion of 6.0 $M_\oplus$ of rocky material ($\alpha_{\rm model} = 3.54 \times 10^{-5}$ dex; a mixture of 5.60 $M_\oplus$ of chondrite-like material plus

Parameter	**Prior**	**Credible Interval**	**N-M minimisation**	**Units**
P	$\mathcal{N}(62.0, 5.0)$	$62.4^{+5.75}_{-5.13}$	63.8626	days
T_c	$\mathcal{N}(2456966, 100)$	$2456087^{+113.52}_{-106.3}$	2456980	days
K	$\mathcal{N}(5, 1)$	$4.9^{+1.14}_{-1.09}$	4.93229	m s^{-1}
$\sqrt{e}\cos\omega$	–	$-0.003^{+0.497}_{-0.565}$	-8.30×10^{-6}	–
$\sqrt{e}\sin\omega$	–	$0.001^{+0.623}_{-0.5971}$	4.18×10^{-5}	–
$\sigma_{\text{HARPS-A}}$	$]-\infty, +\infty[$	$0.0002^{+0.0106}_{-0.0105}$	4×10^{-5}	m s^{-1}
$\sigma_{\text{HARPS-B}}$	$]-\infty, +\infty[$	$0.0002^{+0.0111}_{-0.0106}$	3×10^{-5}	m s^{-1}
$\dot{\gamma}$	$]-\infty, +\infty[$	$-0.001^{+0.0026}_{-0.0041}$	-3.39×10^{-7}	m s^{-1} d^{-1}
$\ddot{\gamma}$	$]-\infty, +\infty[$	$4\times10^{-10}\,{}^{+1.08\times10^{-9}}_{-1.70\times10^{-9}}$	7.77×10^{-14}	m s^{-1} d^{-2}
η_1	$\mathcal{J}(0.01, 100)$	$8.8^{+1.8}_{-2.5}$	5.15	m s^{-1}
η_2	$\mathcal{N}(30, 1/\sqrt{2})$	$29.9^{+0.78}_{-0.78}$	29.97	days
η_3	$\mathcal{N}(26.19, 2.0)$	$22^{+1.92}_{-2.12}$	26.22	days
η_4	$\mathcal{N}(1.07, 1.4)$	$0.462^{+0.010}_{-0.011}$	0.454	–
$m\sin i$	–	$29.6^{+6.75}_{-6.81}$	–	$\text{M}_\oplus$
a	–	$0.304^{+0.0164}_{-0.0190}$	–	AU

Table 5.7 - Best-fit 1-planet Keplerian + GP parameters and derived quantities ($m\sin i$ and a) for HIP30502 system.

0.40 $\text{M}_\oplus$ of Earth-like composition material) into the convective zone for HIP30502.

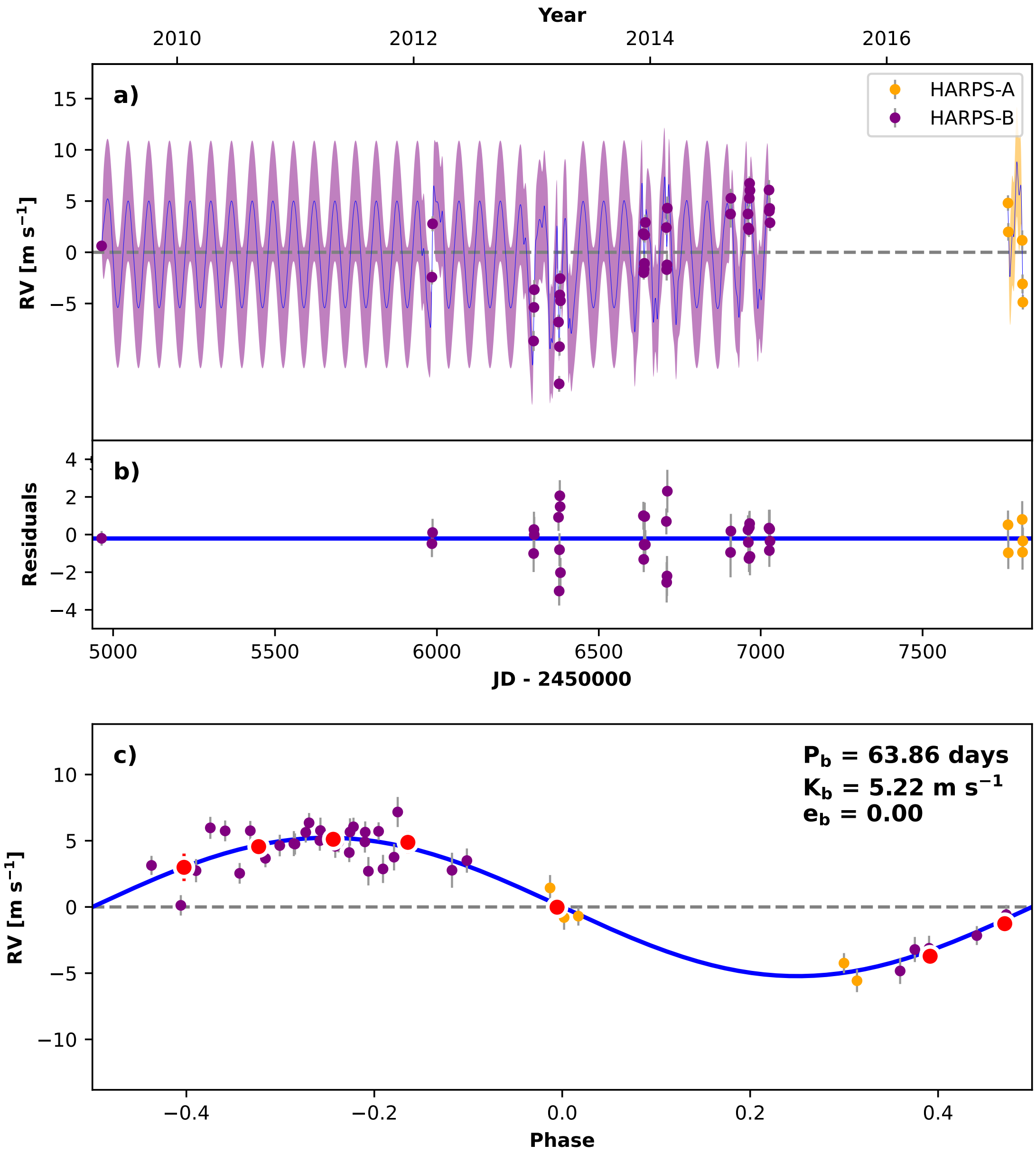

Figure 5.8: Best-fit 1-planet Keplerian plus stellar activity model for HIP30502. The uncertainties include the RV jitter ($\sigma_{\mathrm{HARPS-A/B}}$). In each panel, the purple and yellow dots/curves are the HARPS data before and after the upgrade in mid-2015. (a) The GP model is plotted as a thin blue line, while the purple and yellow curves are Keplerian models. (b) Residuals to the best fit. (c) RV phase diagrams to the ephemeris of planet b. RVs are binned in 0.08 phase units as red dots. The phase-folded model for planet b is shown in blue, along with N-M Keplerian orbital parameters (period, eccentricity and planet-induced Doppler semi-amplitude).

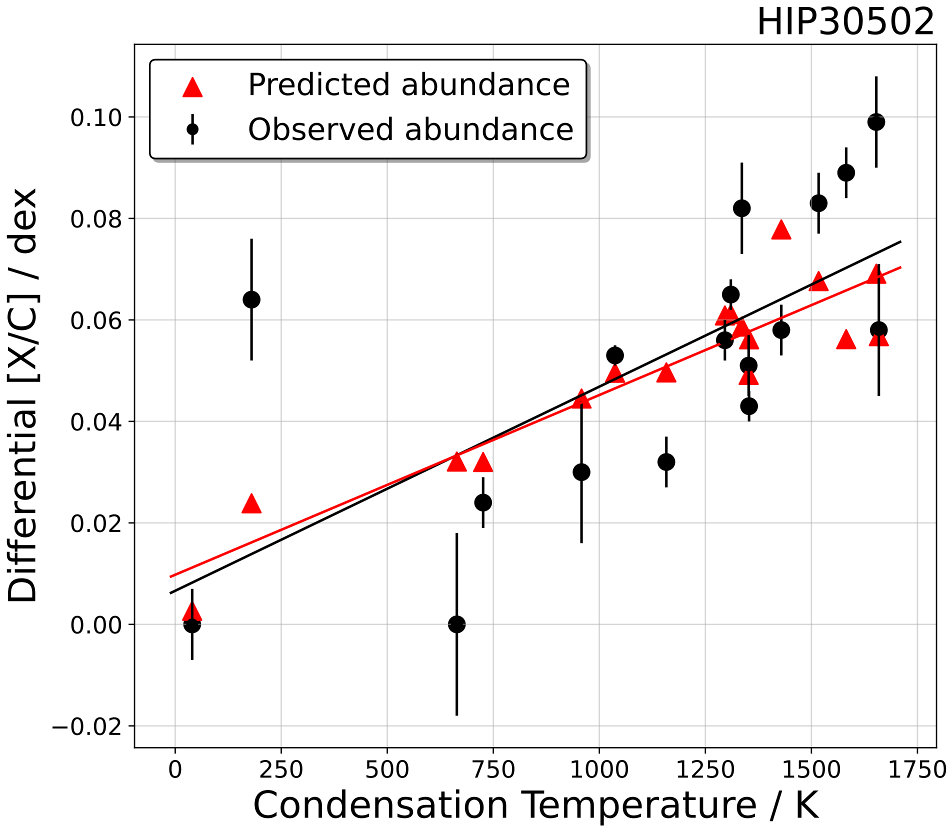

Figure 5.9: Comparison of the differential abundance of HIP30502 (black dots plus linear fit) as a function of dust condensation temperature in the protoplanetary disc, and the predicted abundances (red dots plus linear fit) estimated from a planetary accretion of 6.0 $M_\oplus$.

5.1.3 *HIP64673 b: A Neptunian planet under circular orbit*

HIP64673 has 39 HARPS epochs from February 2012 to February 2017. The system's RUWE of 1.169 may indicate that the astrometry for this system is of lower quality compared to other stars observed by *Gaia*, which could be due to various factors such as measurement uncertainties, systematic errors, or instrumental effects. We highlight that it is not necessarily indicative of the presence of a sub-stellar mass companion to the host star. Pre- and pos-upgrade data have RMS of 5.343 and 1.715 m s^{-1}, respectively.

We detected a signal with a period of $\sim$44.9 days using a standard GLS grid, which had a high statistical significance under sine-like assumptions ($\Delta\mathrm{BIC} = 692$ in comparison to a non-varying noise model). A FAP bootstrapping re-sampling test resulted in a low probability of obtaining a higher peak than 0.53 (associated with the 44.9-d period) with GLS powers of 0.467 and 0.490 with 1% and 5% significance levels, respectively. In addition, we tested the resulting period with the Nyquist test, which ruled out the possibility of the proposed planet's orbital period being an alias from instrumental modulations (e.g., one sidereal day, one synodic month, one sidereal year), or as a source of astrophysical phenomena (e.g., a Sun's activity cycle of nearly eleven years). For HIP64673, we examined two stellar activity indicators: CCF BIS and CCF FWHM (see Fig. 5.10). Although there is a weak correlation between the RVs and CCF FWHM ($r = 0.19$), this correlation is not statistically significant with a $p-$value of 0.25. We did not observe any timing or phase modulations in either activity indicator at the proposed planet's orbital period. Therefore, even if there is some contribution from stellar activity to the RVs, it could be not strong enough to mask a planet's signal.

We used the standard Keplerian to model the RVs of HIP64673, which consisted of five orbital elements. Moreover, accounting for instrumental shifts, we allowed jitter noise corrections to freely vary for each instrumental epoch. We also included the barycentre motion $\dot{\gamma}$ and a curvature term $\ddot{\gamma}$ to account for non-linear trends in the RV model given the presence of an outer companion from which only a small portion of its orbit could be observed. To directly evaluate the impact of stellar activity on the variations in RV, we employed a quasi-periodic covariance kernel in close similarity to the case of HIP1954 given that the prominent orbital period found could be of an intrinsic stellar nature, as it is of the order of a dozen days (see Section 5.1.1). Table 5.8 summarises the HIP64673 model's prior parameters together with orbital solutions.

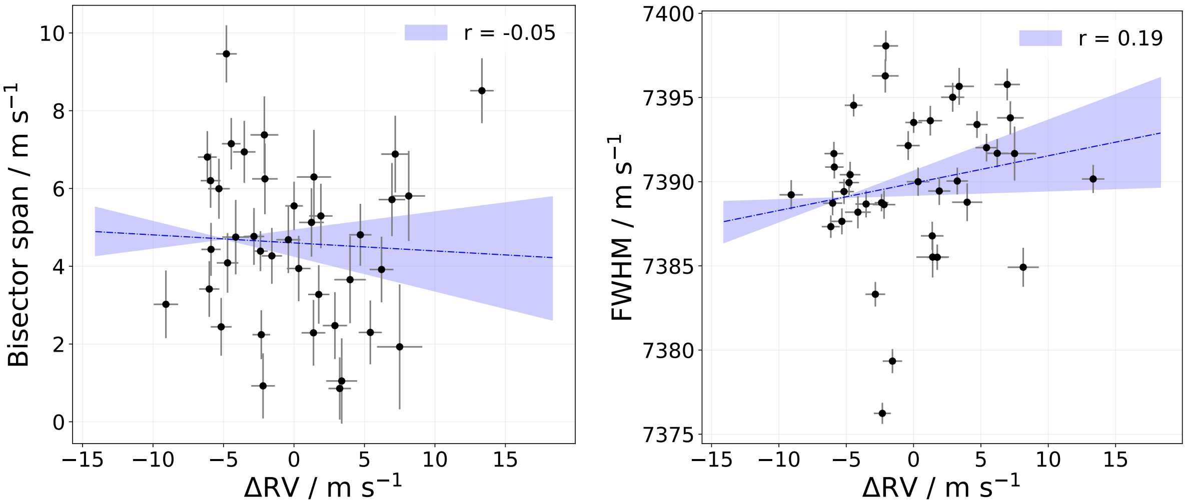

Figure 5.10: RV measurements and their relationship with stellar activity indicators for HIP64673.

Parameter	Prior	Credible Interval	N-M minimisation	Units
P	$\mathcal{N}(44.9, 10)$	$45.05^{+0.737}_{-0.824}$	45.1208	days
T_c	$\mathcal{N}(2456049, 50)$	$2456087^{+4.981}_{-5.827}$	2456060	days
K	$\mathcal{N}(5, 1)$	$2.4^{+1.16}_{-1.17}$	5.086	m s^{-1}
$\sqrt{e}\cos\omega$	–	$\equiv$0	–	–
$\sqrt{e}\sin\omega$	–	$\equiv$0	–	–
$\sigma_{\mathrm{HARPS-A}}$	$]-\infty, +\infty[$	$-5 \times 10^{-5}\ ^{+0.0073}_{-0.0081}$	-3.63$\times$10^{-5}	m s^{-1}
$\sigma_{\mathrm{HARPS-B}}$	$]-\infty, +\infty[$	$0.0004^{+0.0101}_{-0.01030}$	0.00046	m s^{-1}
$\dot{\gamma}$	$]-\infty, +\infty[$	$-3\times10^{-6}\ ^{+0.00095}_{-0.00129}$	-2.04$\times$10^{-6}	m s^{-1} d^{-1}
$\ddot{\gamma}$	$]-\infty, +\infty[$	$-1\times10^{-12}\ ^{+5.28\times10^{-10}}_{-3.90\times10^{-10}}$	1×10^{-12}	m s^{-1} d^{-2}
η_1	$\mathcal{J}(0.01, 100)$	$5.9^{+2}_{-1.36}$	6.45134	m s^{-1}
η_2	$\mathcal{N}(30, 1/\sqrt{2})$	$30^{+0.855}_{-0.889}$	30.6875	days
η_3	$\mathcal{U}(15, 50)$	$15.59^{+1.35}_{-0.45}$	15.95	days
η_4	$\mathcal{N}(1.07, 1.4)$	$0.462^{+0.012}_{-0.012}$	0.4606	–
$m\sin i$	–	$28.80^{+6.99}_{-6.97}$	–	M$_\oplus$
a	–	$0.2534^{+0.00272}_{-0.00278}$	–	AU

Table 5.8 - Best-fit 1-planet Keplerian + GP parameters and derived quantities ($m\sin i$ and a) for HIP64673 system.

We used an MCMC sampling approach with 50 walkers for 1000 steps and 8 ensembles following `RadVel`'s genetic routine. We required burn-out chains with Gelman-Rubin statistics $\hat{R} > 1.01$ and maximum $\hat{R}$ at the same value. The model was minimised with `scipy.optimize` function employing the Nelder-Mead (N-M; Virtanen et al. 2020) algorithm over 200 iterations. We compared statistically an eccentric and a circular model, both accounting for non-zero quantities $\dot{\gamma}$ and $\ddot{\gamma}$, finding out that a circular model with RV slope is favoured by a BIC/AIC variation of $\|\Delta\mathrm{BIC/AIC}\| = 37.68/40.66$ (nominal BIC circular = 289.40, AIC circular = 282.33, with $\ln Z$ = -130.52), over an eccentric model. Therefore, following Liddle (2007), we conclude that a circular model with RV slope is bonafide in representing the data modulations of planetary nature for HIP64673 b, consisting in a warm Neptune-like under circular orbit (see Fig. 5.11).

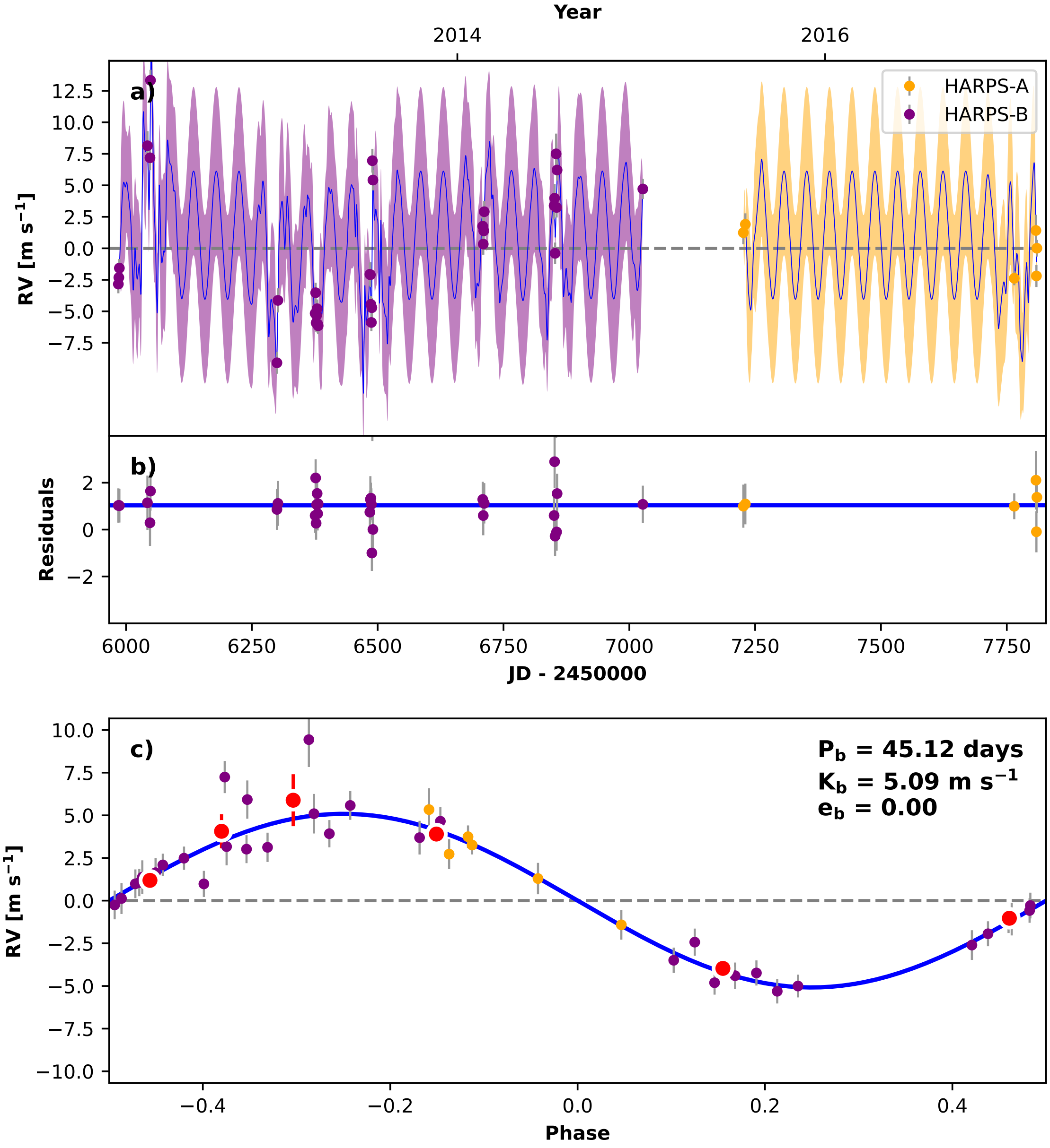

Figure 5.11: Best-fit 1-planet Keplerian plus stellar activity model for HIP64673. The uncertainties include the RV jitter ($\sigma_{\mathrm{HARPS-A/B}}$). In each panel, the purple and yellow dots/curves are the HARPS data before and after the upgrade in mid-2015. (a) The GP model is plotted as a thin blue line, while the purple and yellow curves are Keplerian models. (b) Residuals to the best fit. (c) RV phase diagrams to the ephemeris of planet b. RVs are binned in 0.08 phase units as red dots. The phase-folded model for planet b is shown in blue, along with N-M Keplerian orbital parameters (period, eccentricity and planet-induced Doppler semi-amplitude).

5.1.4 HIP102152 b: A mini-Neptune under circular orbit

Between October 2011 and May 2019, HARPS observed HIP102152 for a total of 79 epochs, comprising a temporal baseline of 7.6 years. The system's RUWE value of 1.039 suggests that it does not belong to a spectroscopic binary. The data's RMS before HARPS upgrade in mid-2015 was 2.820 m s^{-1}, whilst post-upgrade is 2.697 m s^{-1}.

Using GLS over the standard grid, we detected a significant signal with a period of $\sim$22.94 days (with a $\Delta\mathrm{BIC} = 149.50$ under sine-like assumptions). A FAP bootstrapping re-sampling test indicated a low probability of obtaining a higher peak than 0.275 (associated with the prominent period) with GLS powers of 0.258 and 0.247 with 1% and 5% significance levels, respectively. We also tested the proposed planet's orbital period with the Nyquist test, which ruled out the possibility of it being an alias from instrumental modulations or a source of astrophysical phenomena (e.g., one sidereal day, one synodic month, one sidereal year, and a Sun's activity cycle timescale). We observed a weak anti-correlation between the RVs and CCF BIS ($r = -0.27$), and a weak correlation for the other activity indicators (CCF FWHM, $S-$index, and Balmer lines up to Hϵ; see Fig. 5.12). Additionally, we did not observe any timing or phase modulations in either activity indicator at the proposed planet's orbital period. Therefore, we can conclude that it is unlikely that stellar activity could mask a planet's signal, even if it contributed to the RV measurements.

To model the RVs we used a standard Keplerian consisting of five orbital elements and allowed jitter noise corrections to vary for each instrumental epoch. We also influenced the barycentre motion $\dot{\gamma}$ and curvature term $\ddot{\gamma}$ to account for non-linear trends in the RV model due to the presence of an outer companion that could only be partially observed. Furthermore, we used a quasi-periodic covariance kernel similar to that used in some cases presented in this chapter due to its ability to capture the inherent irregular and sinusoidal nature of stellar activity signals, providing a more accurate representation of the underlying radiative processes and improving the precision and reliability of exoplanet detection and characterisation. The best-fit model's prior parameters and orbital solutions are summarised in Table 5.9.

Using an MCMC sampling approach with 50 walkers for 1000 steps and 4 ensembles following `RadVel`'s genetic routine, we employed the Nelder-Mead algorithm to minimise the model over 200 iterations. We compared a total of four models: circular vs. eccentric, with and without

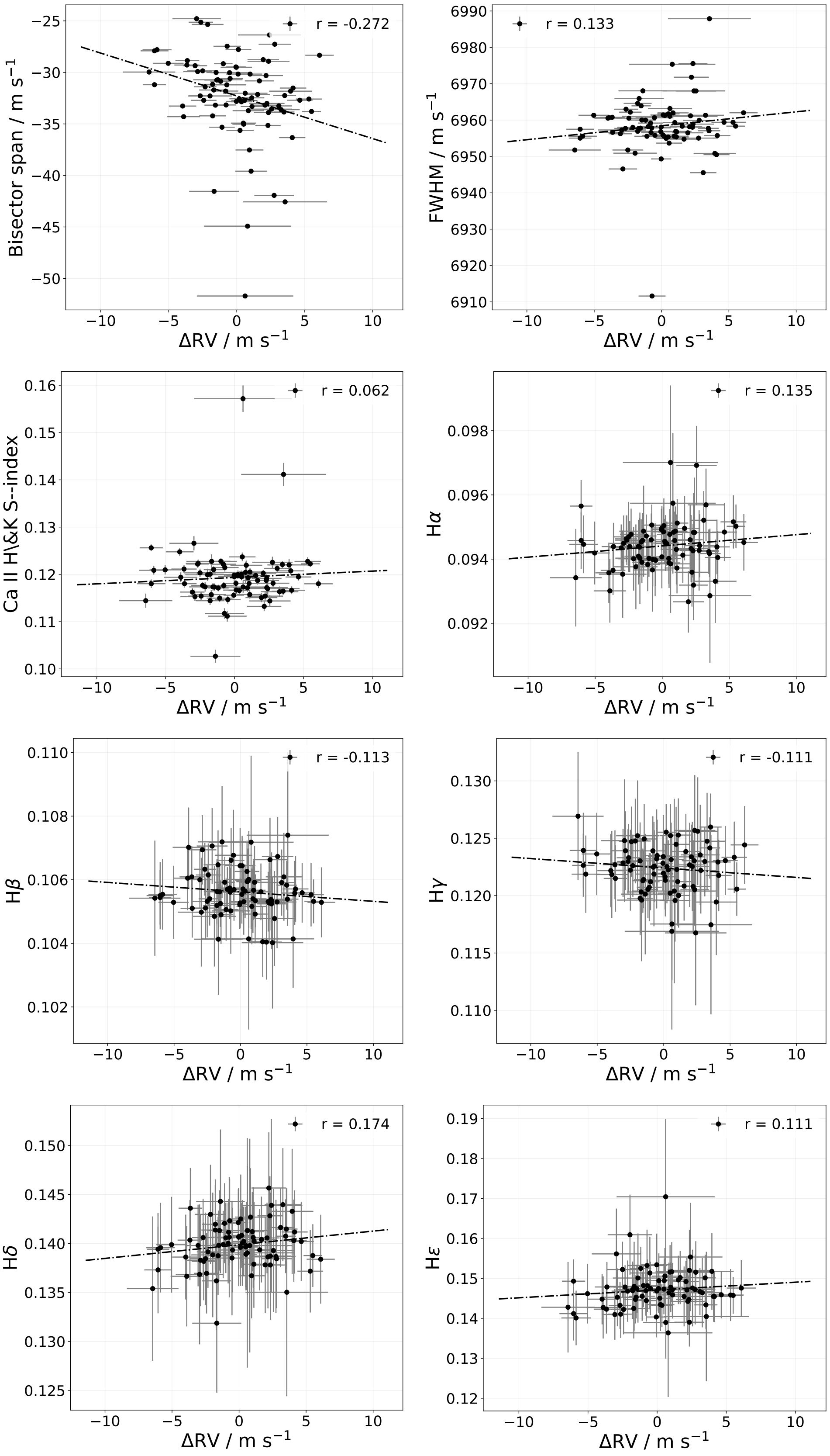

Figure 5.12: RV measurements and their relationship with stellar activity indicators for HIP102152.

Parameter	**Prior**	**Credible Interval**	**N-M minimisation**	**Units**
P	$\mathcal{N}(22.92, 0.01)$	$22.93^{+0.0103}_{-0.01023}$	22.9272	days
T_c	$\mathcal{N}(2456490, 50)$	$2456489^{+0.0101}_{-0.0102}$	2456490	days
K	$\mathcal{N}(2, 1)$	$1.87^{+0.204}_{-0.218}$	1.78999	m s^{-1}
$\sqrt{e}\cos\omega$	–	$\equiv$0	–	–
$\sqrt{e}\sin\omega$	–	$\equiv$0	–	–
$\sigma_{\rm HARPS-A}$	$]-\infty, +\infty[$	$-0.003^{+0.118}_{-0.114}$	-0.0028	m s^{-1}
$\sigma_{\rm HARPS-B}$	$]-\infty, +\infty[$	$0.002^{+0.115}_{-0.120}$	0.0037	m s^{-1}
$\dot{\gamma}$	$]-\infty, +\infty[$	$\equiv 0$	$\equiv 0$	m s^{-1} d^{-1}
$\ddot{\gamma}$	$]-\infty, +\infty[$	$\equiv 0$	$\equiv 0$	m s^{-1} d^{-2}
η_1	$\mathcal{J}(0.01, 100)$	$3.68^{+0.57}_{-0.45}$	3.4847	m s^{-1}
η_2	$\mathcal{N}(30, 1/\sqrt{2})$	$29.62^{+0.706}_{-0.715}$	29.6198	days
η_3	$\mathcal{N}(35.7, 1.4)$	$35.8^{+1.38}_{-1.39}$	35.9079	days
η_4	$\mathcal{N}(1.07, 1.4)$	$0.46^{+0.0102}_{-0.0103}$	0.462929	–
$m\sin i$	–	$8.145^{+0.892}_{-0.951}$	–	M$_\oplus$
a	–	$0.156^{+4.70\times10^{-5}}_{-4.65\times10^{-5}}$	–	AU

Table 5.9 - Best-fit 1-planet Keplerian + GP parameters and derived quantities ($m\sin i$ and a) for HIP102152 system.

RV slope ($\dot{\gamma}$). Our results demonstrate that a circular model without RV slope better represented the data modulations of planetary nature for HIP102152 b, a mini-Neptune planet under a circular orbit (nominal $\|\mathrm{BIC/AIC}\| = 464/445.28$ and $\ln Z = -222.73$ (see Fig. 5.13).

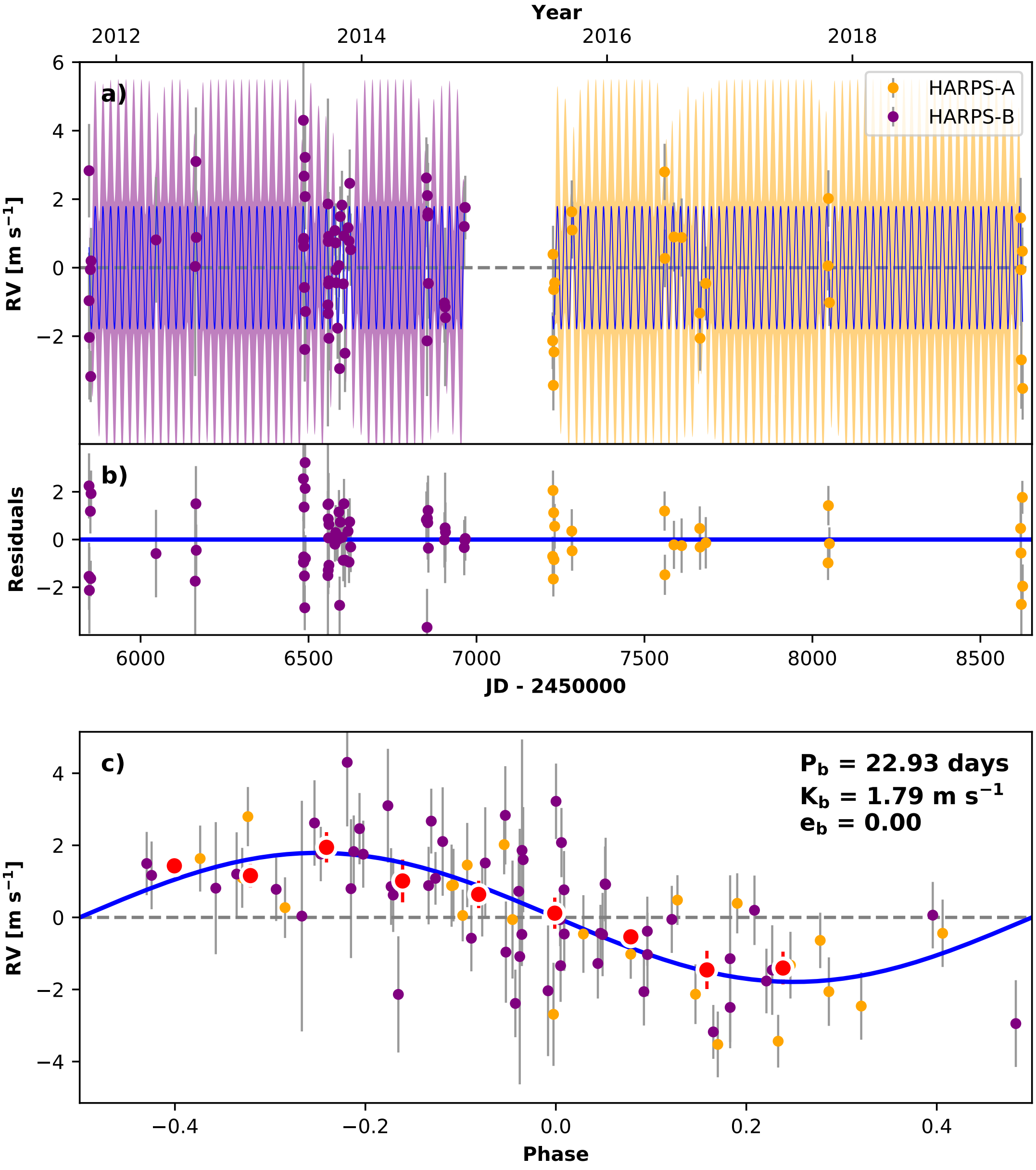

Figure 5.13: Best-fit 1-planet Keplerian plus stellar activity model for HIP102152. The uncertainties include the RV jitter ($\sigma_{\mathrm{HARPS-A/B}}$). In each panel, the purple and yellow dots/curves are the HARPS data before and after the upgrade in mid-2015. (a) The GP model is plotted as a thin blue line, while the purple and yellow curves are Keplerian models. (b) Residuals to the best fit. (c) RV phase diagrams to the ephemeris of planet b. RVs are binned in 0.08 phase units as red dots. The phase-folded model for planet b is shown in blue, along with N-M Keplerian orbital parameters (period, eccentricity and planet-induced Doppler semi-amplitude).

5.1.5 HIP114615 b: An warm Neptunian planet under eccentric orbit

From October 2011 to October 2017, HARPS observed HIP114615 for 54 epochs. The RUWE is 0.914, suggesting that a binary is unlikely and that the astrometry is reliable. The data before and after mid-2015 shows RMS values of 5.797 and 5.709 m s^{-1}. A GLS analysis identifies a signal close to 24.7 days. The signal was found to have a high statistical significance under sine-like assumptions, as indicated by a BIC change of $||\Delta\mathrm{BIC}|| = 493$ compared to a non-varying noise model. FAP tests with bootstrapping re-sampling also confirmed the significance of the signal, which revealed that the likelihood of obtaining a higher peak than 0.491 (related to the 24.7-d period) with GLS powers of 0.366 and 0.310 in the RV periodogram was 1% and 5%, respectively. Furthermore, we ruled out an alias resulting from instrumental modulations, using the Nyquist test with one sidereal day, one synodic month, one sidereal year, and a Sun's time activity cycle of eleven years. We examined six activity proxies: CCF BIS, CCF FWHM, $S-$index, Hα, the Na doublet ($\lambda = 589.0$ and $\lambda = 589.6$ nm) and He I lines ($\lambda = 586.9$ nm). We found a moderate correlation between the $S-$index and the RV measurements ($r = 0.53$), CCF FWHM ($r = 0.36$), and He I ($r = -0.36$), however, there were no timing or phase modulations detected in any of these activity proxies at the proposed planet's orbital period. This suggests that the host star's activity weakly interferes with the planet's signal. Therefore, we used a standard Keplerian with RadVel.

We implemented a standard Keplerian, leaving free jitter terms for both HARPS epochs $\sigma_{\mathrm{HARPS-A/B}}$ – that are added in quadrature to the RV uncertainties –, as well as the orbital eccentricity coupled with the periastron argument in the orthogonal basis. Table 5.10 summarises the model priors and orbital solutions. The MCMC sampling was also implemented following a genetic routine, with 50 walkers over 10000 steps and 8 ensembles. Chains with $\hat{R}$ greater than 1.03 were discarded, the minimum auto-correlation time factor to deem chains as well-mixed and halt the MCMC (minAfactor) was set as default (40), the maximum relative change in auto-correlation time to deem chains well-mixed (maxArchange) was also set as default (0.03), and the minimum number of independent samples (minTz) to consider well-mixed was set as 1000 (default). In practice, the final values for these parameters were minAfactor = 42.396, maxArchange = 0.018, and minTz = 5858.940, with the model's maximum $\hat{R} = 1.005$, indicating a good convergence of the model to the RV modulations. We compared an eccentric model with a circular model using the BIC criterion and found no significant difference between them ($\Delta\mathrm{BIC} = 0.98$). However,

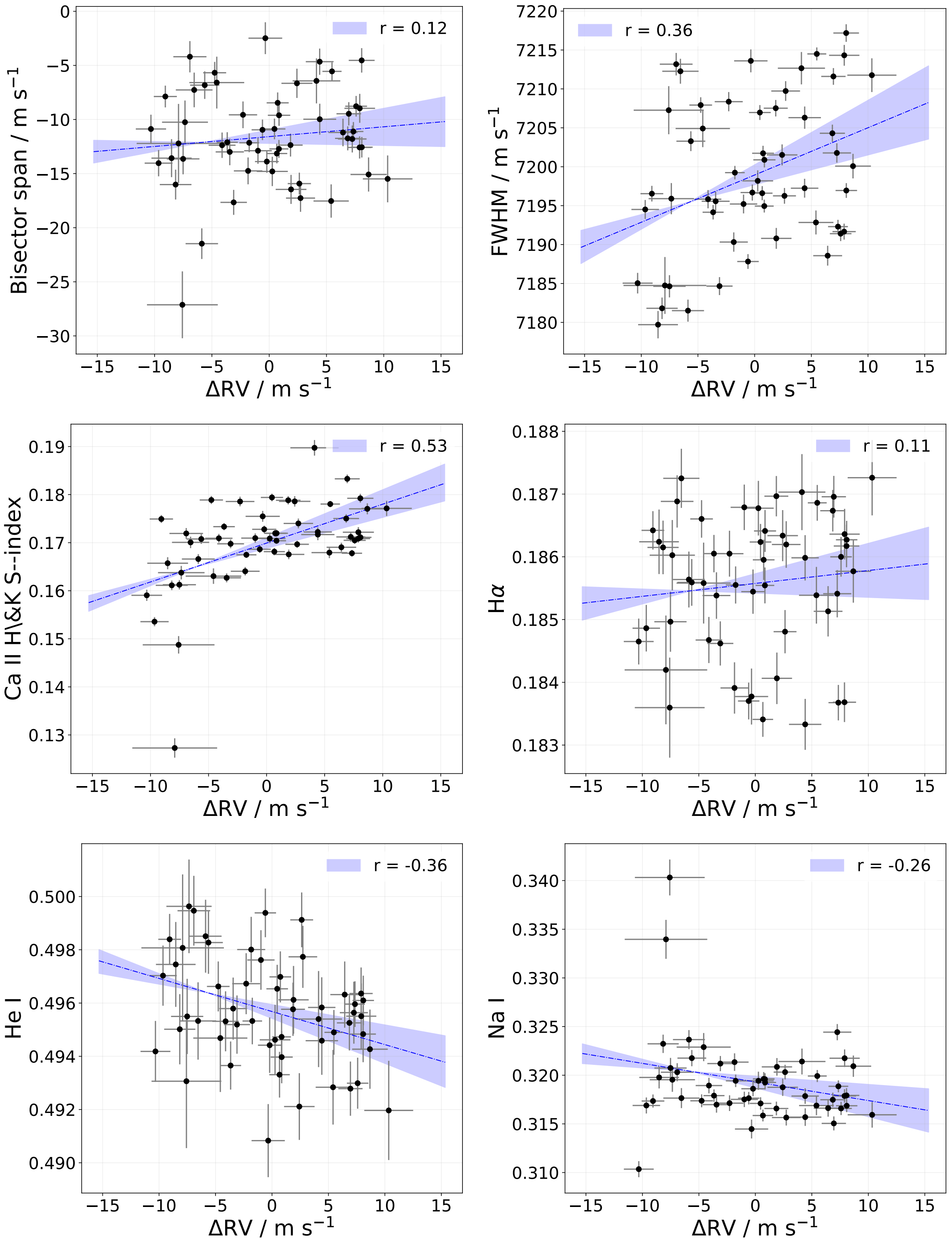

Figure 5.14: RV measurements and their relationship with stellar activity indicators for HIP114615.

when we used the AIC criterion, we observed a slightly higher preference for the eccentric model, with $\Delta\text{AIC} = 3.30$. Even though those models are statistically indistinguishable, we opted for the eccentric model since it presented a better convergence over the three criteria stated before (minAfactor, maxArchange, and minTz). We then conclude that an eccentric model is bonafide in representing the data modulations of HIP114615 b, consisting of a warm ice giant planet under eccentric orbit (see Fig. 5.15).

Parameter	**Prior**	**Credible Interval**	**Units**
P	$\mathcal{N}(24.70, 1.0)$	$24.657^{+0.027}_{-0.037}$	days
T_c	$\mathcal{N}(2456961, 100)$	$2455452^{+1.8}_{-1.0}$	days
K	$\mathcal{U}(0.1, 20)$	$6.9^{+1.1}_{-1.0}$	m s^{-1}
$\sqrt{e}\cos\omega$	$-$	$0.42^{+0.18}_{-0.37}$	$-$
$\sqrt{e}\sin\omega$	$-$	$0.20^{+0.22}_{-0.27}$	$-$
$\sigma_{\mathrm{HARPS-A}}$	$]-\infty, +\infty[$	$4.3^{+1.5}_{-1.0}$	m s^{-1}
$\sigma_{\mathrm{HARPS-B}}$	$]-\infty, +\infty[$	$-3.51^{+0.47}_{-0.60}$	m s^{-1}
$m\sin i$	$-$	$29.8^{+4.8}_{-4.3}$	M$_\oplus$
a	$-$	0.16726±0.00031	AU

Table 5.10 - Best-fit 1-planet Keplerian parameters and derived quantities ($m\sin i$ and a), for HIP114615b.

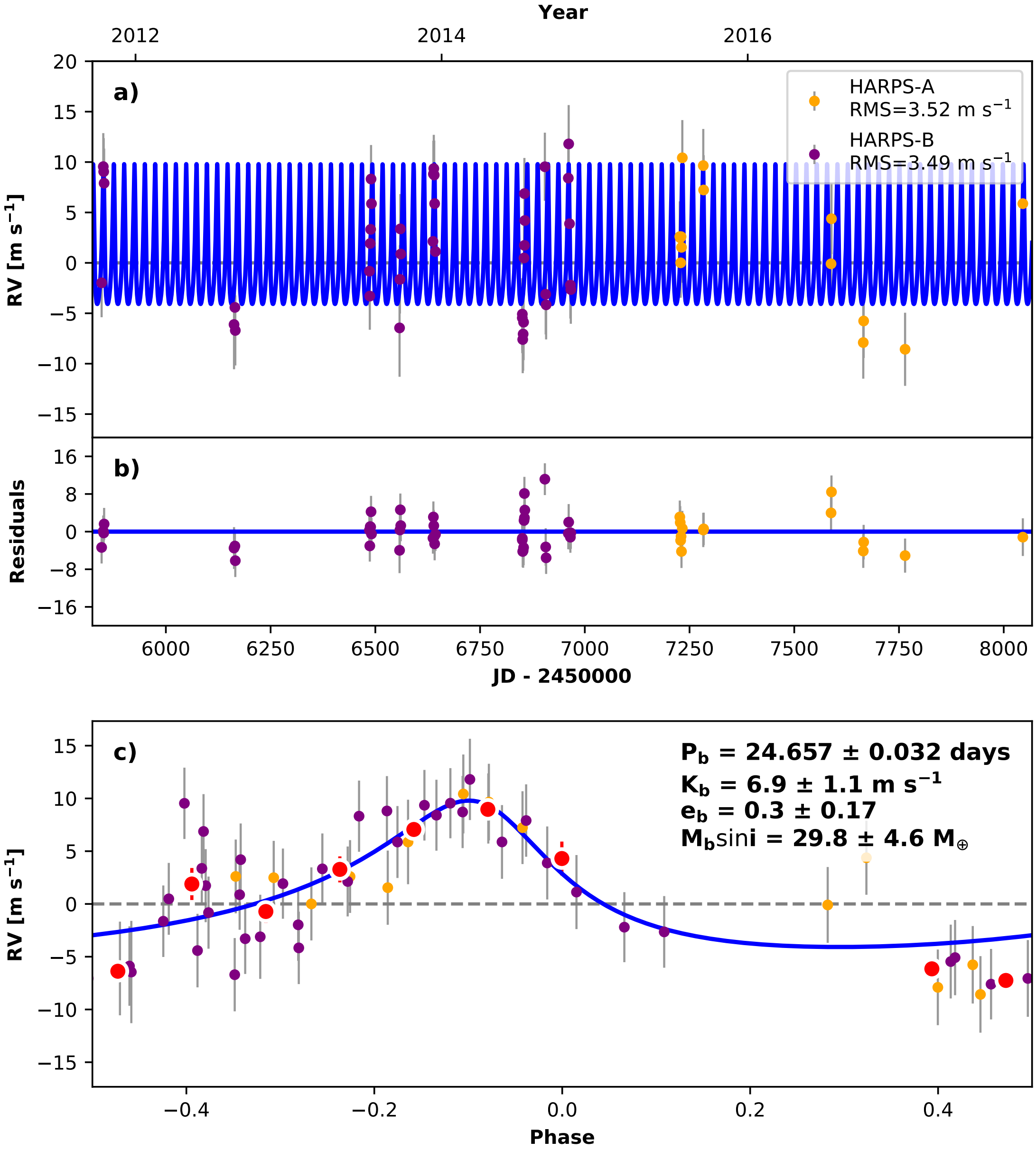

Figure 5.15: Best-fit 1-planet Keplerian orbital model for HIP114615. In all panels, the purple and yellow dots/curves are the HARPS observations before and after its upgrade in mid-2015, respectively. Label in panel (a) indicates the data RMS. (b) Residuals to the best fit 1-planet model. (c) RV phase diagram to the ephemeris of planet *b*. The RVs are binned in 0.08 phase units as red dots. The phase-folded model for planet *b* is shown in blue, along with the best-fit Keplerian orbital parameters (period, eccentricity and planet-induced Doppler amplitude), plus the planet's minimum mass.

5.1.6 HIP117367 b: A warm Neptunian planet under circular orbit

From October 2011 to October 2017, HARPS observed HIP117367 for 45 epochs. The system's RUWE is somewhat high, at 1.185, which suggests that the astrometric data for this system may be of lower quality compared to other stars observed by *Gaia*. There is no report of known sub-stellar companions of this system in the literature. The RMS values before and after the HARPS' upgrade are 3.143 and 2.249 m s^{-1}, respectively.

Using GLS we detected a signal with a period of ∼58.4 days. This signal had a high statistical significance under sine-like assumptions, i.e., when compared to a non-varying noise model, $\Delta\mathrm{BIC} = 655$. We conducted a FAP bootstrapping re-sampling test and found a low probability of obtaining a higher peak than 0.590 with GLS powers of 0.499 and 0.342 with 1% and 5% significance levels, respectively. Additionally, we tested the proposed planet's orbital period with the Nyquist test and ruled out the possibility of the period being an alias from instrumental modulations (e.g., one sidereal day, one synodic month, one sidereal year), or as a source of astrophysical phenomena (e.g., a Sun's activity cycle). We examined two indicators of activity for HIP117367, CCF BIS and CCF FWHM (see Fig. 5.16). We did not find any significant correlation between the RVs and either of these indicators, with correlation coefficients of 0.11 and 0.15, respectively. We also did not see any timing or phase modulations at the proposed planet's orbital period for either indicator.

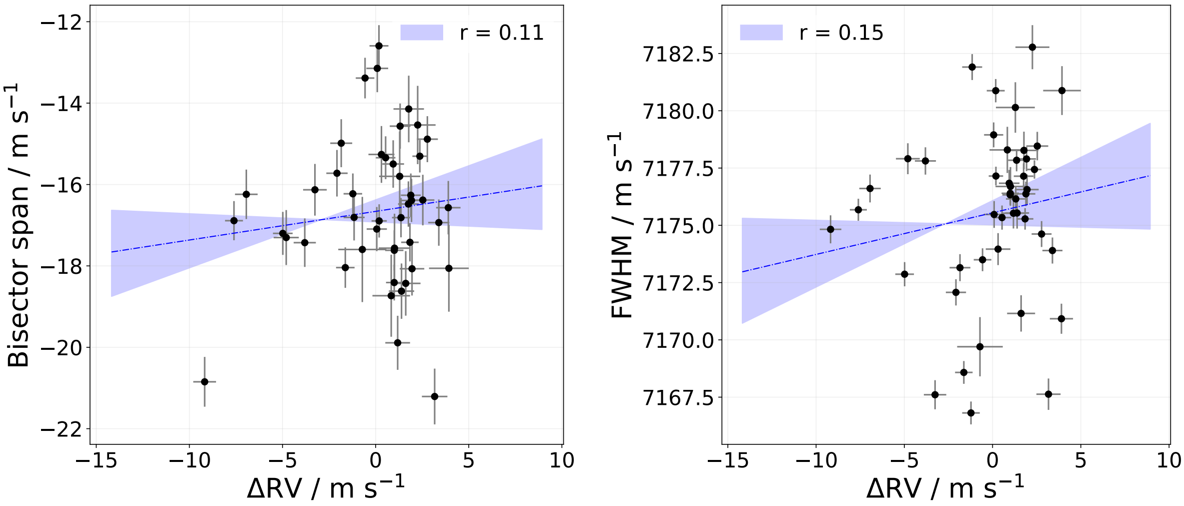

Figure 5.16: RV measurements and their relationship with stellar activity indicators for HIP117367.

We used the standard Keplerian to model the RVs, which consisted of five orbital elements. Accounting for instrumental shifts, we allowed jitter noise corrections to freely vary for each instrumental epoch. We also included the barycentre motion $\dot{\gamma}$ and a curvature term $\ddot{\gamma}$ to account for non-linear trends in the RV model given the presence of an outer companion from which only a small portion of its orbit could be observed. To evaluate the impact of stellar activity on the variations in RV, we employed a quasi-periodic covariance kernel in close similarity to the case of HIP1954 (see Section 5.1.1). Table 5.11 summarises the HIP117367 model's prior parameters together with orbital solutions.

Parameter	**Prior**	**Credible Interval**	**N-M minimisation**	**Units**
P	$\mathcal{N}(58.4, 1)$	$56.2^{+2.05}_{-0.47}$	55.9695	days
T_c	$\mathcal{N}(2456854, 10)$	$2456855^{+11.96}_{-7.39}$	2456850	days
K	$\mathcal{N}(2.7, 0.2)$	$2.72^{+0.36}_{-0.35}$	2.74	m s^{-1}
$\sqrt{e}\cos\omega$	$-$	$\equiv$0	$-$	$-$
$\sqrt{e}\sin\omega$	$-$	$\equiv$0	$-$	$-$
$\sigma_{\mathrm{HARPS-A}}$	$]-\infty, +\infty[$	$-5 \times 10^{-5}\ {}^{+0.0073}_{-0.0081}$	3.25$\times10^{-5}$	m s^{-1}
$\sigma_{\mathrm{HARPS-B}}$	$]-\infty, +\infty[$	$0.0004^{+0.0101}_{-0.01030}$	7.98$\times10^{-5}$	m s^{-1}
$\dot{\gamma}$	$]-\infty, +\infty[$	$-1\times10^{-7}\ {}^{+4.07\times10^{-5}}_{-4.96\times10^{-5}}$	-6.06$\times10^{-8}$	m s^{-1} d^{-1}
$\ddot{\gamma}$	$]-\infty, +\infty[$	$-3\times10^{-15}\ {}^{+2.01\times10^{-11}}_{-1.65\times10^{-11}}$	-6.13$\times10^{15}$	m s^{-1} d^{-2}
η_1	$\mathcal{J}(0.01, 100)$	$3.0^{+0.95}_{-0.97}$	2.878	m s^{-1}
η_2	$\mathcal{N}(30, 1/\sqrt{2})$	$30.05^{+0.952}_{-0.977}$	29.8963	days
η_3	$\mathcal{N}(50.20, 5.0)$	$45.2^{+7.16}_{-7.29}$	49.40	days
η_4	$\mathcal{N}(1.07, 1.4)$	$0.464^{+0.013}_{-0.0132}$	0.465451	$-$
$m\sin i$	$-$	$16.80^{+2.20}_{-2.18}$	$-$	M$_\oplus$
a	$-$	$0.291^{+0.0070}_{-0.00163}$	$-$	AU

Table 5.11 - Best-fit 1-planet Keplerian + GP parameters and derived quantities ($m\sin i$ and a) for HIP117367 system.

We used an MCMC sampling with 50 walkers for 1000 steps and 8 ensembles following `RadVel`'s genetic routine. We required burn-out chains with Gelman-Rubin statistics $\hat{R} > 1.01$ and maximum $\hat{R}$ at the same value. The model was minimised with `scipy.optimize` function employing the Nelder-Mead (N-M) algorithm over 200 iterations. We compared statistically an eccentric and a circular model, both accounting for non-zero quantities $\dot{\gamma}$ and $\ddot{\gamma}$, finding out that a circular model with RV slope is favoured by a BIC/AIC variation of $||\Delta\mathrm{BIC/AIC}|| = 8.87/10.07$ (nominal BIC circular = 291.84, AIC circular = 280.09, with $\ln Z$ = -126.10), over an eccentric model. Therefore, we conclude that a circular model with RV slope is bonafide in representing the data modulations of planetary nature for HIP117367 b, consisting of a warm Neptune-like planet under circular orbit (see Fig. 5.17). .

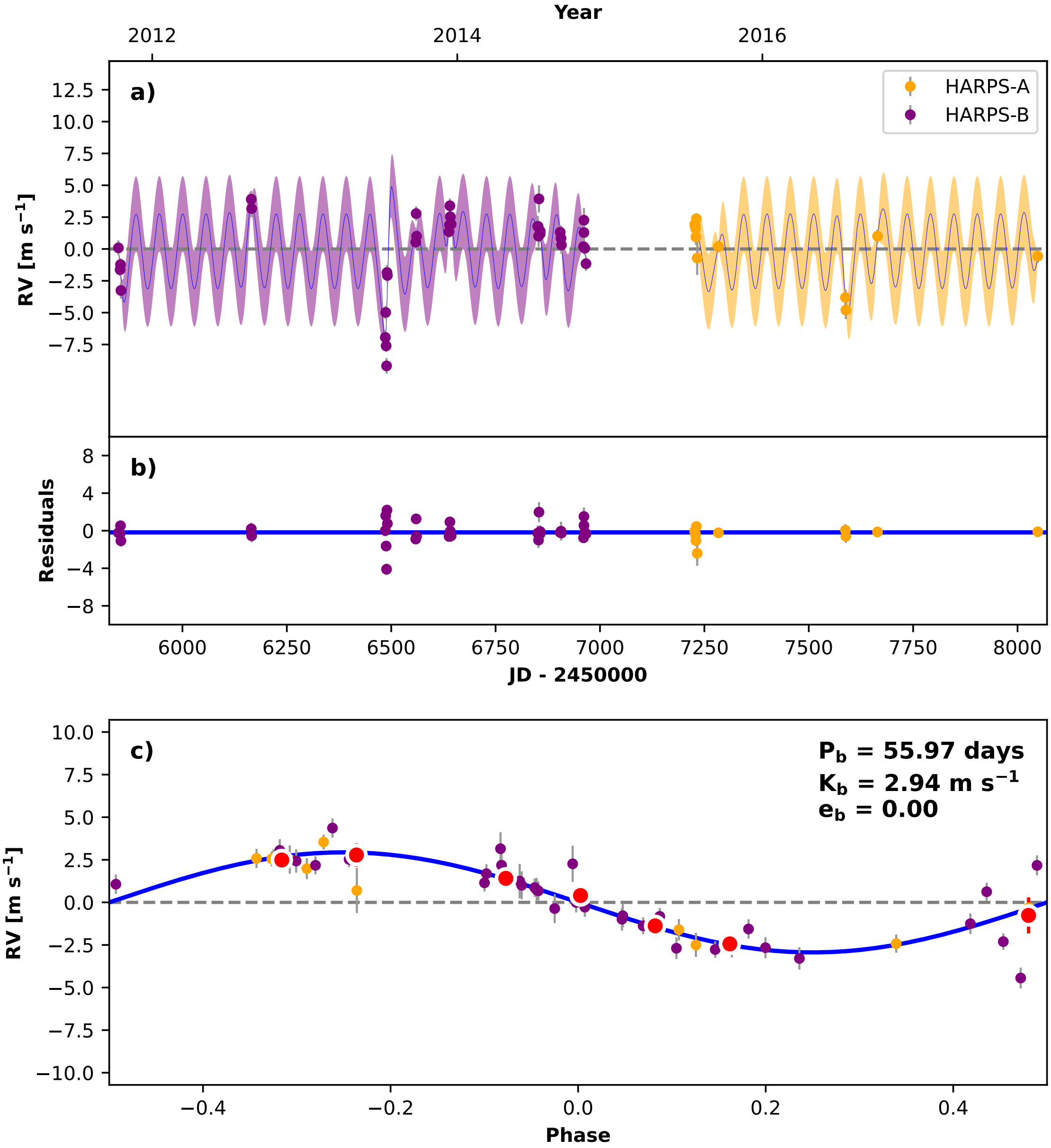

Figure 5.17: Best-fit 1-planet Keplerian plus stellar activity model for HIP117367. The uncertainties include the RV jitter ($\sigma_{\mathrm{HARPS-A/B}}$). In each panel, the purple and yellow dots/curves are the HARPS data before and after the upgrade in mid-2015. (a) The GP model is plotted as a thin blue line, while the purple and yellow curves are Keplerian models. (b) Residuals to the best fit. (c) RV phase diagrams to the ephemeris of planet b. RVs are binned in 0.08 phase units as red dots. The phase-folded model for planet b is shown in blue, along with N-M Keplerian orbital parameters (period, eccentricity and planet-induced Doppler semi-amplitude).

5.2 Gas Giant Exoplanets

5.2.1 HIP5301 b: A Jupiter analogue

HARPS gathered 63 epochs of HIP5301 between December 2003 and July 2021. The RUWE of HIP5301 is 0.957, indicating that the host star is unlikely a binary. The data before mid-2015 showed an RMS value of 20.59 m s-1, while after the upgrade is 16.93 m s^{-1}, indicating the presence of an orbiting planet at first sight. Previous studies by Naef et al. (2010) suggested the presence of a massive companion ($P = 2496 \pm 176$ d, $m \sin i = 1.56^{+1.11}_{-1.10}$ $\mathrm{M_{Jup}}$, and $e = 0.10^{+0.11}_{-0.04}$) from 22 RV measurements until ~2010. We present a re-analysis of this system improving the orbital solutions with an increase of >200% of RV measurements and a bigger timing baseline.

We employed a GLS algorithm over the standard grid, identifying a discernible signal in the vicinity of 2439 days. The signal was found to have high statistical significance under sine-like assumptions, as indicated by a BIC change of $||\Delta\mathrm{BIC}|| = 8768$ compared to a non-varying noise model. FAP tests with bootstrapping re-sampling also confirmed the significance of the signal which revealed that the likelihood of obtaining a higher peak than 0.961 (related to the 2414-d period) with GLS powers of 0.417 and 0.383 in the RV periodogram was 1% and 5%, respectively. Moreover, we ruled out the possibility of the proposed planet's orbital period being an alias resulting from instrumental modulations testing it with the Nyquist frequency sampling test with one sidereal day, one synodic month, one sidereal year, and a Sun's time activity cycle of roughly eleven years. We examined six activity proxies for HIP5301: CCF BIS, CCF FWHM, $S-$index, Hα, the Na doublet ($\lambda = 589.0$ and $\lambda = 589.6$ nm) and He I lines ($\lambda = 586.9$ nm). At the proposed planet's orbital period, we found no evidence of time or phase modulations. Furthermore, we did not observe any significant correlations between these activity tracers and HIP5301 RV measurements (see Fig. 5.18), indicating that the host star is relatively "quiet". Pearson's $r-$correlation index only shows a moderate anti-correlation between RVs and the CCF FWHM ($r = -0.49$), while for all the other indicators, the RVs present weak correlations. We performed the orbital fit with `RadVel` for HIP5301 b solely using a Keplerian model given these findings.

We implemented the standard Keplerian basis, leaving as free parameters jitter terms for both HARPS epochs $\sigma_{\mathrm{HARPS-A/B}}$, that is added in quadrature to the RV uncertainties. Table 5.12 summarises the model priors and orbital solutions. The MCMC sampling, in this case,

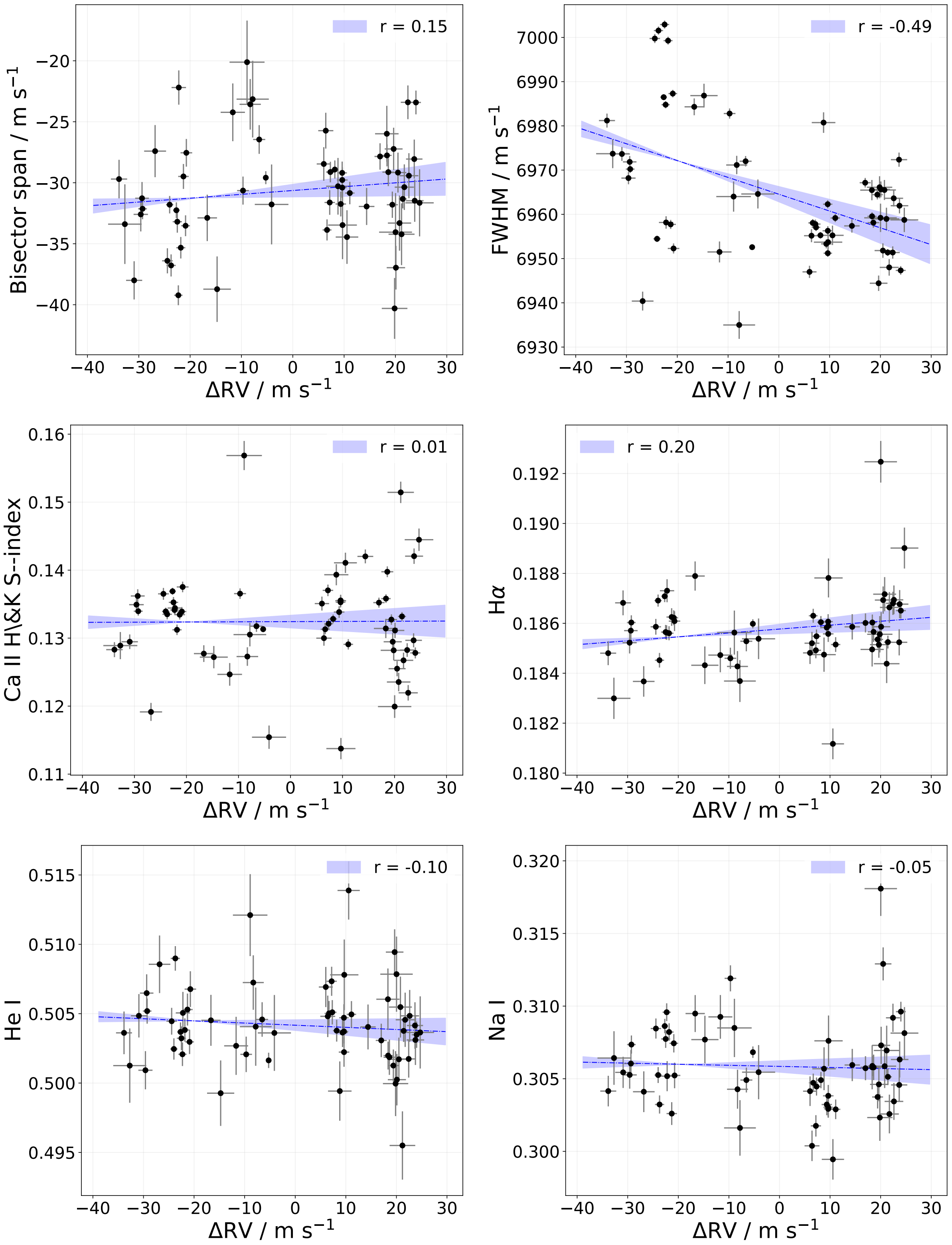

Figure 5.18: RV measurements and their relationship with stellar activity indicators for HIP5301.

was also implemented following `RadVel`’s genetic routine, with 50 walkers, over 10000 steps and 8 ensembles. The chains with $\hat{R}$ greater than 1.03 were discarded, the minimum auto-correlation time factor to deem chains as well-mixed and halt the MCMC (minAfactor) was set as default (40), the maximum relative change in auto-correlation time to deem chains and well-mixed (maxArchange) was also set as default (0.03), and the minimum number of independent samples (minTz) to consider well-mixed was set as 1000 (default). In practice, the final values for these parameters were minAfactor = 34.895, maxArchange = 0.007, and minTz = 6474.595, with maximum $\hat{R} = 1.004$, indicating a good convergence of the model to the modulations of the observational data. Model comparison indicates a slight favour for the eccentric model by $\Delta\mathrm{BIC}/\mathrm{AIC} = 0.44/2.47$ over a circular model (nominal BIC eccentric = 345.37, AIC eccentric = 329.48, with $\ln\mathcal{L} = -197.27$). Even though those models are statistically indistinguishable, we opted for the eccentric model since it presented a better convergence over the three criteria stated before (minAfactor, maxArchange, and minTz). We then conclude that an eccentric model is bonafide in representing the data modulations of HIP5301 b, consisting of a Jupiter analogue (minimum mass ranging between 0.3 and 3 $\mathrm{M_{Jup}}$, orbital semi-major axis between 3 and 7 AU (orbital periods between ~5 to ~18 years), and orbital eccentricity $e \leq 0.3$; see Fig. 5.19).

Parameter	**Prior**	**Credible Interval**	**Units**
P	$\mathcal{N}(2439.02, 100)$	2473±11	days
T_c	$\mathcal{N}(2454719, 100)$	2455452^{+21}_{-19}	days
K	$\mathcal{U}(0.01, 30)$	$28.78^{+0.67}_{-0.71}$	m $\mathrm{s^{-1}}$
$\sqrt{e}\cos\omega$	$\mathcal{U}(-1, 1)$	$-0.12^{+0.11}_{-0.08}$	–
$\sqrt{e}\sin\omega$	$\mathcal{U}(-1, 1)$	$0.12^{+0.08}_{-0.12}$	–
$\sigma_{\mathrm{HARPS-A}}$	$]-\infty, +\infty[$	$3.59^{+0.85}_{-0.72}$	m $\mathrm{s^{-1}}$
$\sigma_{\mathrm{HARPS-B}}$	$]-\infty, +\infty[$	$-1.11^{+2.89}_{-0.80}$	m $\mathrm{s^{-1}}$
$m\sin i$	–	$1.864^{+0.044}_{-0.046}$	$\mathrm{M_{Jup}}$
a	–	3.533±0.011	AU

Table 5.12 - Best-fit 1-planet Keplerian parameters and derived quantities ($m\sin i$ and a), for HIP5301b.

Unlike HIP1954, HIP5301 presents a moderate depletion of refractory/volatile elements when compared to the Sun (with T_C−plane slope of $\alpha = -5.055 \times 10^{-6}$ dex). From the host star's mass and metallicity, we estimated its convective zone as 0.026 $\mathrm{M_{\odot}}$. The red dots in Fig. 5.20 demonstrate the [X/C] ratios as a function of condensation temperature ($\alpha_{\mathrm{model}} = 5.854 \times 10^{-6}$ dex), accounting for the impact of 0.7 $\mathrm{M_{\oplus}}$ rocky material into the convective zone for HIP5301 ($\chi^2 = 5.63$), but considering the large scatter, we could consider that the star has about the same composition as the Sun.

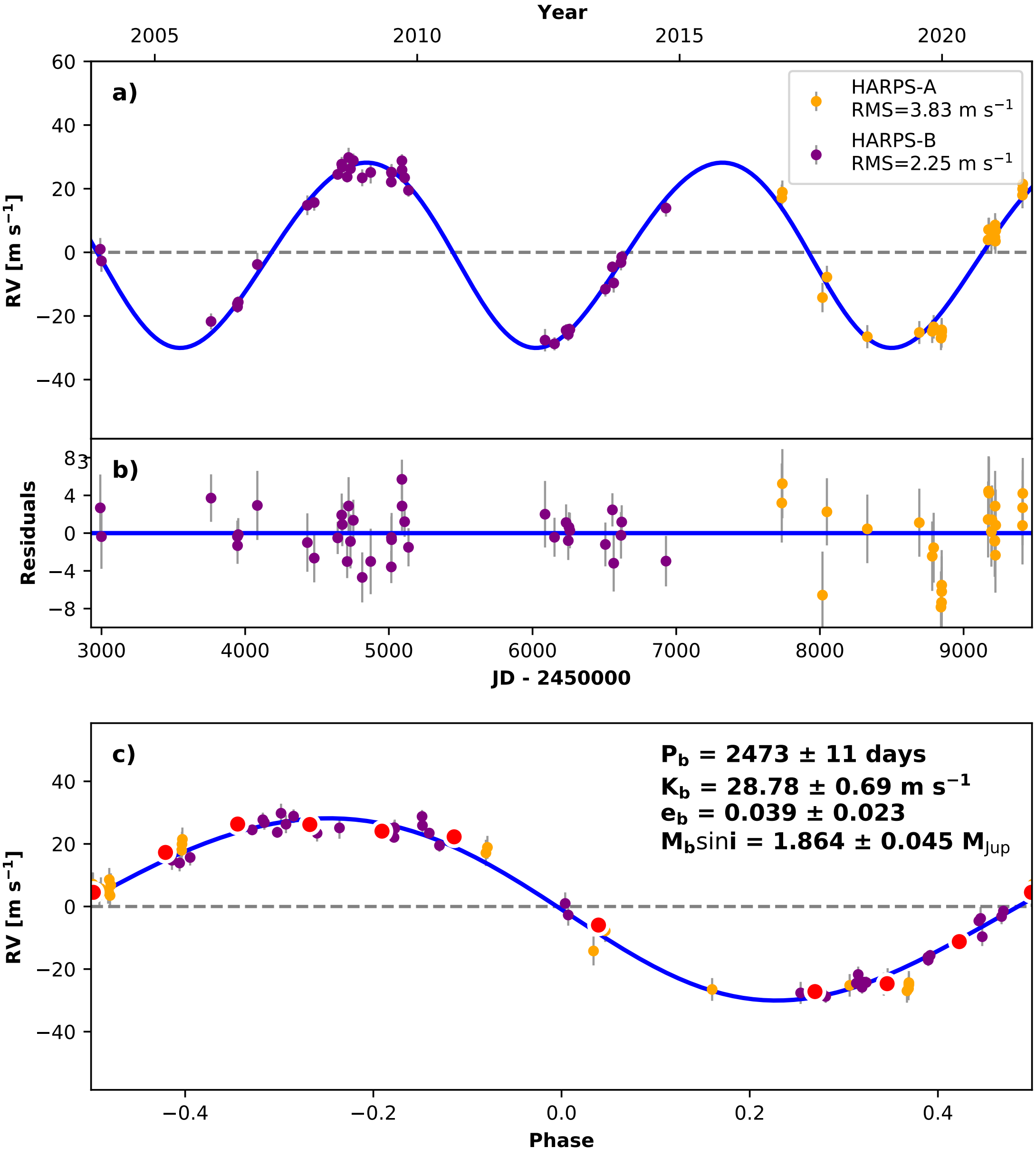

Figure 5.19: Best-fit 1-planet Keplerian orbital model for HIP5301. In all panels, the purple and yellow dots/curves are the HARPS observations before and after its upgrade in mid-2015, respectively. Label in panel (a) indicates the data RMS. (b) Residuals to the best fit 1-planet model. (c) RV phase diagram to the ephemeris of planet *b*. The RVs are binned in 0.08 phase units as red dots. The phase-folded model for planet *b* is shown in blue, along with the best-fit Keplerian orbital parameters (period, eccentricity and planet-induced Doppler amplitude), plus the planet's minimum mass.

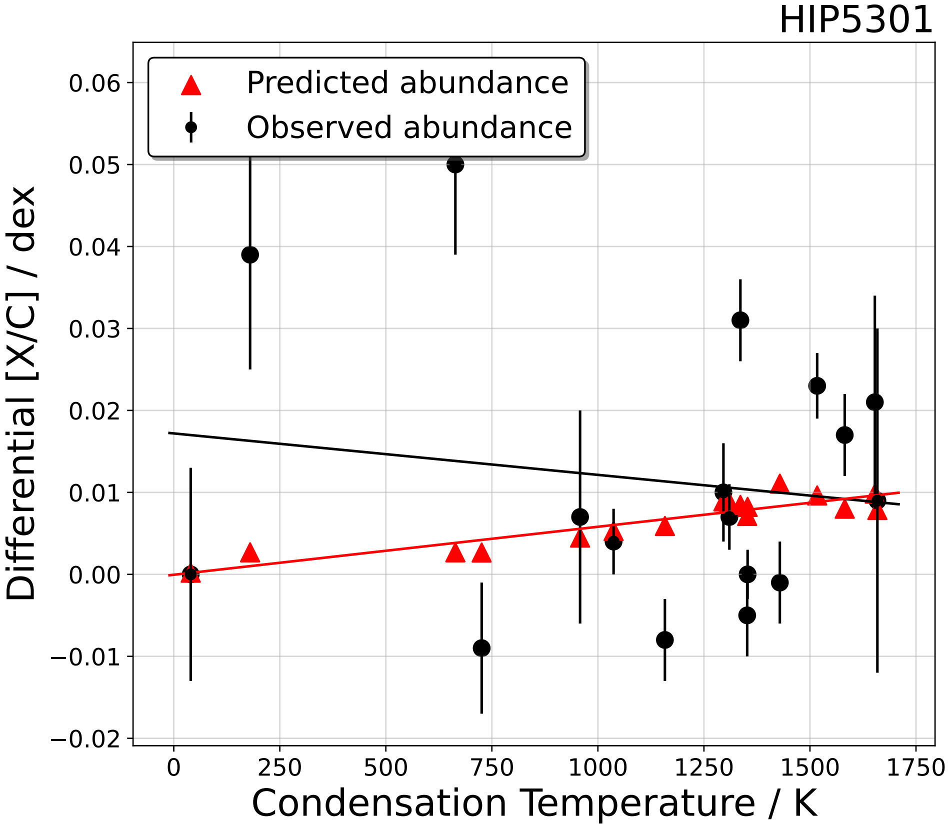

Figure 5.20: Comparison of the differential abundance of HIP5301 (black dots plus linear fit) as a function of dust condensation temperature in the protoplanetary disc, and the predicted abundances (red dots plus linear fit) estimated from a planetary accretion of 0.7 $M_{\oplus}$. Note that within the uncertainties, HIP5301 has about solar composition.

5.2.2 HIP30158 b: A cold gas giant planet under circular orbit

ESO/HARPS spectrograph observed HIP30158 during 39 epochs between February 2012 and February 2017. The host star's RUWE is 0.934, indicating that it is unlikely to be part of a spectroscopic binary system and that its astrometric measurements are accurate. The RMS of data before and after the HARPS upgrade are 3.970 and 1.216 m s^{-1}, respectively. Utilising a GLS algorithm over the standard grid, we detected a significant signal close to 1137 days. It demonstrated high statistical significance in comparison to a non-varying noise model ($\|\Delta\mathrm{BIC}\| = 850$), and a FAP bootstrapping re-sampling test revealed a low probability of obtaining a higher peak than 0.849 (associated with the 1137-d period) with GLS powers or 0.458 and 0.402 in the RV periodogram corresponding to significance levels of 1% and 5%, respectively. From the Nyquist test, we ruled out the possibility of the proposed planet's orbital period being an alias from instrumental modulations. For HIP30158, we examined a set of six different stellar activity indicators: CCF BIS, CCF FWHM, $S-$index, Hα, Na doublet and He I (see Fig. 5.21), from which we observed moderate correlations between the RVs and CCF FWHM, Hα, and He I, while weak correlations are seen for the other activity indicators. Despite that, at the proposed planet's orbital period, none of the six activity indicators demonstrated timing or phase modulations, which indicates that, albeit stellar activity may contribute to some scatter in the RVs, it does not hamper the planet's signal significantly.

We employed the Keplerian orbital basis, with jitter terms for both HARPS epochs $\sigma_{\mathrm{HARPS-A/B}}$ set as free parameters. Table 5.13 summarises the model priors and orbital solutions. We highlight that the jitter terms' best-fit values before and after HARPS' upgrade in mid-2015 are relatively small compared to the planet's induced Doppler semi-amplitude. Similarly to HIP5301, the MCMC sampling was implemented with default `RadVel`'s configuration (50 walkers over 10000 steps and 8 ensembles). The final values of the convergence criteria for the best-fit model found were minAfactor = 29.137, maxArchange = 0.006, maximum $\hat{R}$ = 1.005, and minTz = 5394.844, demonstrating a good convergence of the model to the modulations of the observed data. Model comparison indicates a strong disfavour of a highly eccentric model ($\Delta\mathrm{BIC}/\mathrm{AIC} = 6.71/5.98$) over a circular model (nominal BIC circular = 152.38, AIC circular = 144.35, $\ln\mathcal{L} = -82.66$). We then conclude that HIP30158 b consists of a cold Saturn-mass planet under a nearly circular orbit (see Fig. 5.22).

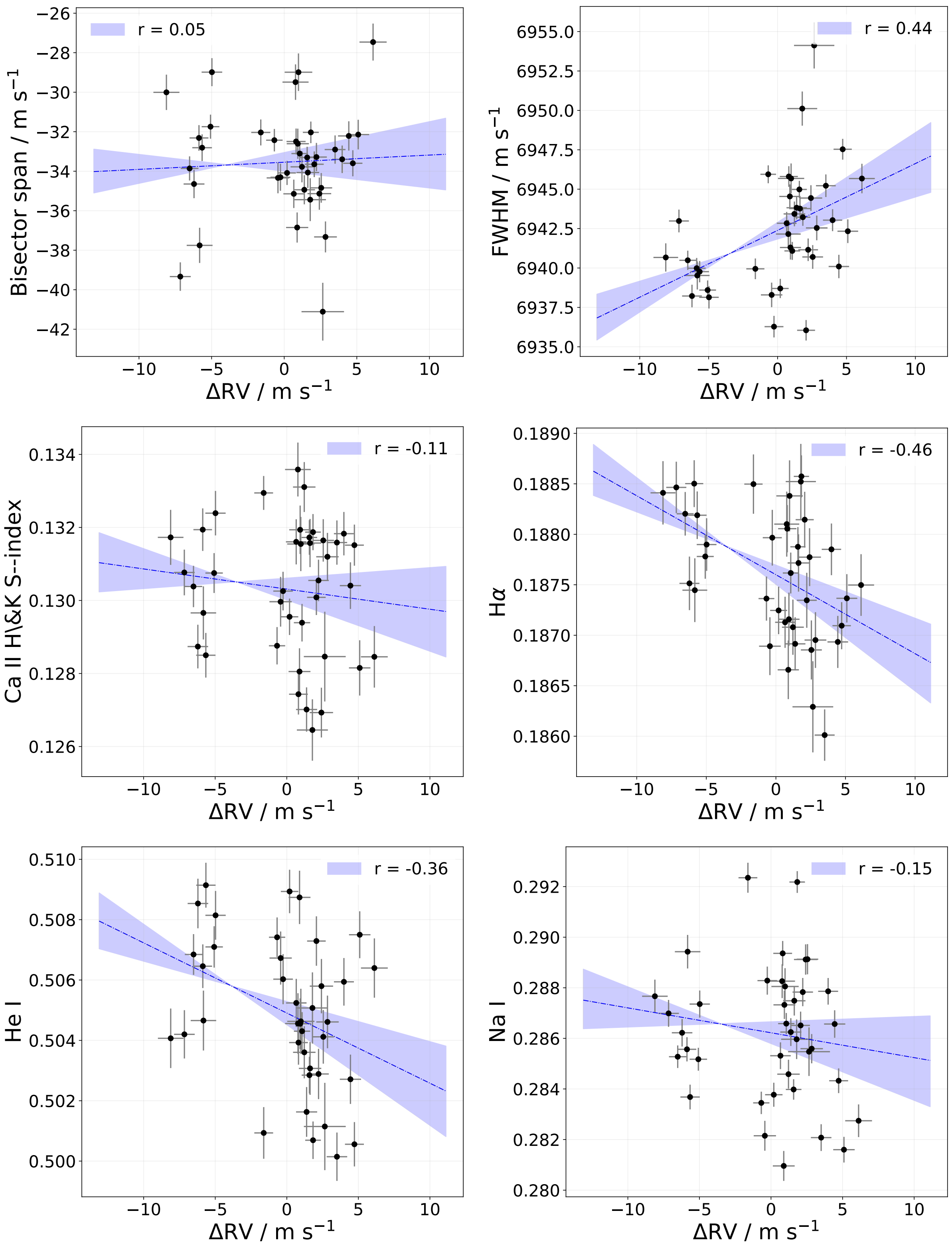

Figure 5.21: RV measurements and their relationship with stellar activity indicators for HIP30158.

Parameter	**Prior**	**Credible Interval**	**Units**
P	$\mathcal{N}(1137, 100)$	1050^{+29}_{-27}	days
T_c	$\mathcal{N}(2456709, 100)$	2457083^{+27}_{-21}	days
K	$\mathcal{N}(5.0, 1)$	$5.55^{+0.37}_{-0.38}$	$\mathrm{m\ s^{-1}}$
$\sqrt{e}\cos\omega$	$\mathcal{U}(-1, 1)$	$-0.06^{+0.20}_{-0.20}$	–
$\sqrt{e}\sin\omega$	$\mathcal{U}(-1, 1)$	$-0.06^{+0.20}_{-0.19}$	–
$\sigma_{\mathrm{HARPS-B}}$	$]-\infty, +\infty[$	$1.04^{+0.25}_{-0.24}$	$\mathrm{m\ s^{-1}}$
$\sigma_{\mathrm{HARPS-A}}$	$]-\infty, +\infty[$	0.0 ± 2.2	$\mathrm{m\ s^{-1}}$
$m \sin i$	–	$85.7^{+5.8}_{-5.7}$	$\mathrm{M}_{\oplus}$
a	–	$1.997^{+0.037}_{-0.035}$	AU

Table 5.13 - Best-fit 1-planet Keplerian parameters and derived quantities ($m \sin i$ and a), for HIP30158b.

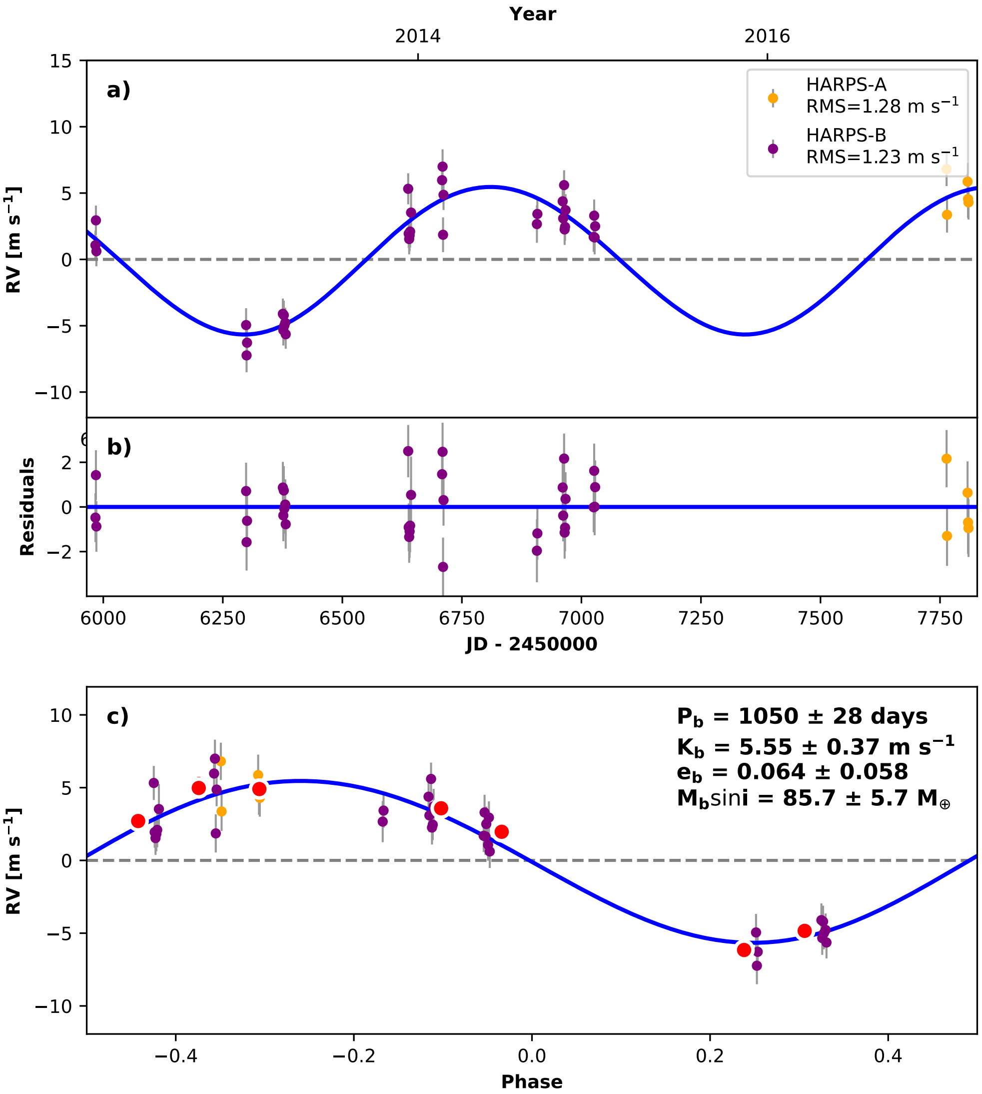

Figure 5.22: Best-fit 1-planet Keplerian orbital model for HIP30158. In all panels, the purple and yellow dots/curves are the HARPS observations before and after its upgrade in mid-2015, respectively. Label in panel (a) indicates the data RMS. (b) Residuals to the best fit 1-planet model. (c) RV phase diagram to the ephemeris of planet b. The RVs are binned in 0.08 phase units as red dots. The phase-folded model for planet b is shown in blue, along with the best-fit Keplerian orbital parameters (period, eccentricity and planet-induced Doppler amplitude), plus the planet's minimum mass.

5.2.3 HIP43297 b: A highly eccentric cold Jupiter-like planet

The ESO/HARPS spectrograph recorded data on HIP43297 for 35 epochs from February 2012 to January 2018. The host star's RUWE value of 1.004 suggests it is unlikely to belong to a spectroscopic binary system. The data gathered before mid-2015 had an RMS value of 14.533 m s^{-1}, while the post-upgrade data displayed an RMS value of 54.543 m s^{-1}, indicating that the object is experiencing significant fluctuations in its RV over time.

A strong signal with a period of approximately 817 days was detected using the GLS algorithm on a standard grid, and it was highly statistically significant under sine-like assumptions, as indicated by a BIC change of $||\Delta\mathrm{BIC}|| = 76233$ in comparison to a non-varying noise model. The FAP tests confirmed the significance of the signal, with a likelihood of obtaining a higher peak than 0.961 in the RV periodograms with GLS-powers of 0.473 and 0.420 with 1% and 5% significance levels, respectively. We ruled out instrumental modulations as the source of the proposed planet's orbital period by employing the Nyquist frequency sampling test. For HIP43297, we tested six activity proxies (CCF BIS, CCF FWHM, $S-$index, Hα, Na doublet and He I). Due to the high RMS of the RV measurements, it is observed that the correlations have a relative spread, which somehow makes it difficult to find a linear relationship between two parameters (see Fig. 5.23). We observed weak-to-moderate correlations for all activity tracers, ranging $|r| = 0.24$ to $|r| = 0.43$. Such correlations for CCF BIS ($p = 0.13$), CCF FWHM ($p = 0.159$), Hα ($p = 0.154$) and $S-$index ($p = 0.07$) present no statistical relevance, albeit the null-hypothesis for Na ($p = 0.009$) and He ($p = 0.033$) are to be considerable. Given the high apparent eccentricity and high Doppler-induced semi-amplitude, we employed a Keplerian model alone not coupled with a Gaussian Process regression (e.g., Vogt et al. 2015). Following Frink et al. (2002), the unique shape observed in the radial velocity curve of a highly eccentric orbit is challenging to replicate through alternative phenomena such as stellar oscillations or rotational modulation. Consequently, the distinctive pattern exhibited by a highly eccentric RV curve as the case of HIP43297 b can be considered as compelling evidence supporting the presence of a planetary companion.

We used the standard Keplerian basis and introduced jitter terms for both HARPS epochs as free parameters. Note that, in this case, the eccentricity coupled with the periastron argument in the orthogonal basis was also set as free parameters, i.e., no priors were assigned, however, we designated the orbital eccentricity to be below 0.99 (as the case for bounded orbits). The

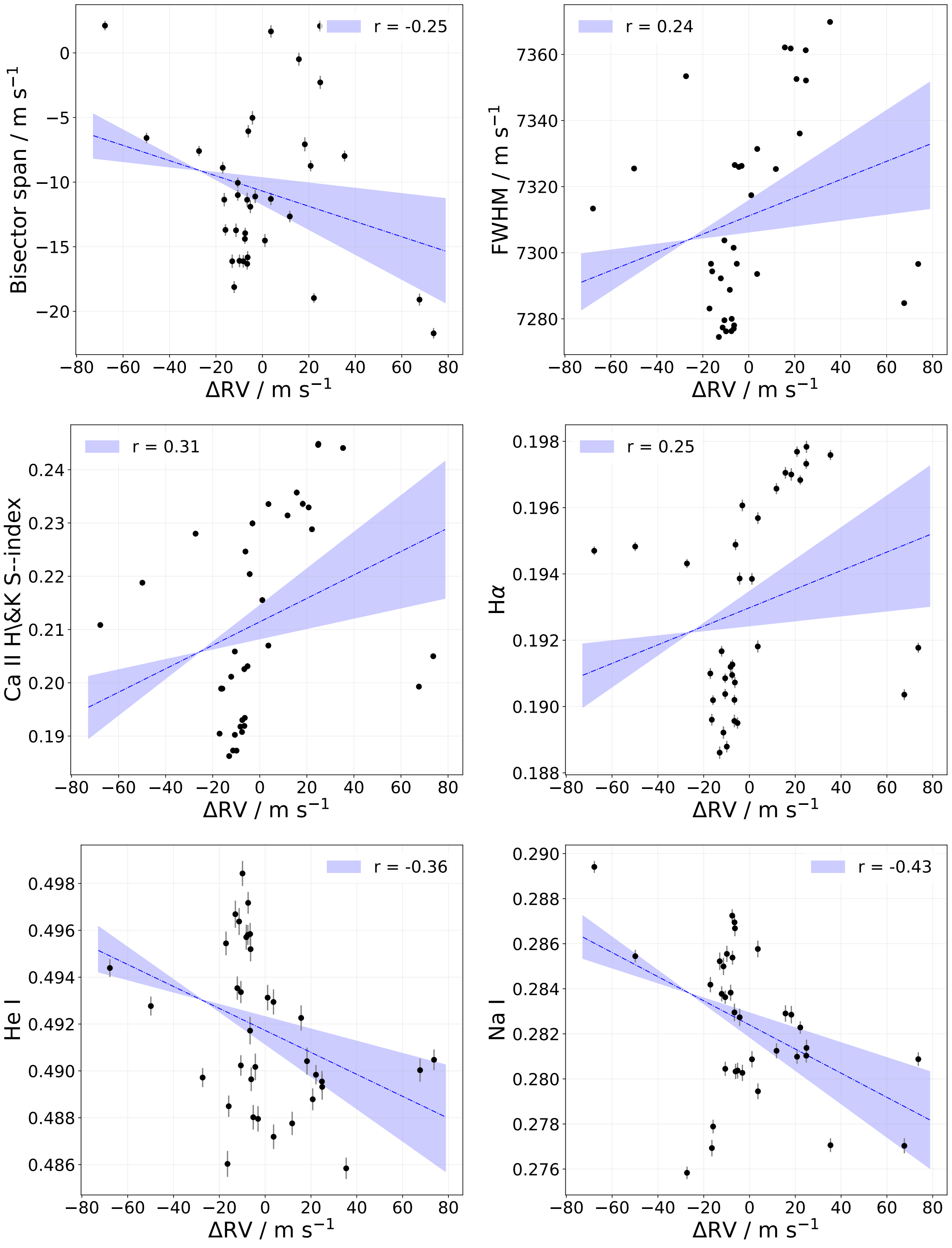

Figure 5.23: RV measurements and their relationship with stellar activity indicators for HIP43297.

MCMC sampling was performed similarly to the case of HIP5301 (50 walkers over 10000 steps and 8 ensembles), where we obtained minAfactor = 41.896, maxArchange = 0.009, and minTz

= 12886.557. We obtained a good convergence of the model with a maximum $\hat{R} = 1.002$. A comparison between a circular and an eccentric model indicated that the last is favoured by a $\Delta\mathrm{BIC}/\mathrm{AIC} = 42.16/42.21$ (nominal BIC eccentric = 268.38, AIC eccentric = 261.59, and $\ln\mathcal{L} = -131.66$). Therefore, we obtained that the RV data modulations of HIP43297 b consist of a highly eccentric cold Jupiter-like planet (see Fig. 5.24 and Table 5.14).

Parameter	**Prior**	**Credible Interval**	**Units**
P	$\mathcal{N}(817.5, 10)$	$816.1^{+8.3}_{-9.0}$	days
T_c	$\mathcal{N}(2457809, 10)$	$2457826.7^{+7.5}_{-8.7}$	days
K	$\mathcal{N}(50, 10)$	$58.1^{+8.2}_{-8.7}$	m s^{-1}
$\sqrt{e}\cos\omega$	–	$0.79^{+0.04}_{-0.05}$	–
$\sqrt{e}\sin\omega$	–	$0.17^{+0.07}_{-0.07}$	–
$\sigma_{\mathrm{HARPS-A}}$	$]-\infty, +\infty[$	16^{+34}_{-58}	m s^{-1}
$\sigma_{\mathrm{HARPS-B}}$	$]-\infty, +\infty[$	$-6.86^{+0.92}_{-1.2}$	m s^{-1}
$m\sin i$	–	2.0±0.21	$\mathrm{M_{Jup}}$
a	–	$1.718^{+0.012}_{-0.013}$	AU

Table 5.14 - Best-fit 1-planet Keplerian parameters and derived quantities ($m\sin i$ and a), for HIP43297b.

From HIP43297's chemical abundance pattern, we observe that the host star may have experienced planet engulfment episodes in the past. From its mass and metallicity, the mass of its convective zone was estimated to be 0.027 $\mathrm{M}_{\odot}$. The [X/C] ratios as a function of the condensation temperature T_C for refractory/volatile elements (black dots/curve; slope $\alpha = 6.062 \times 10^{-5}$ dex) are shown in Fig. 5.25. The figure also displays the impact of accreting 17.8 M$\oplus$ of rocky material ($\alpha_{\mathrm{model}} = 6.57 \times 10^{-5}$ dex; a mixture of 17.60 M$\oplus$ of chondrite-like-material plus 0.2 M$\oplus$ of Earth-like composition material) into the convective zone of HIP43297.

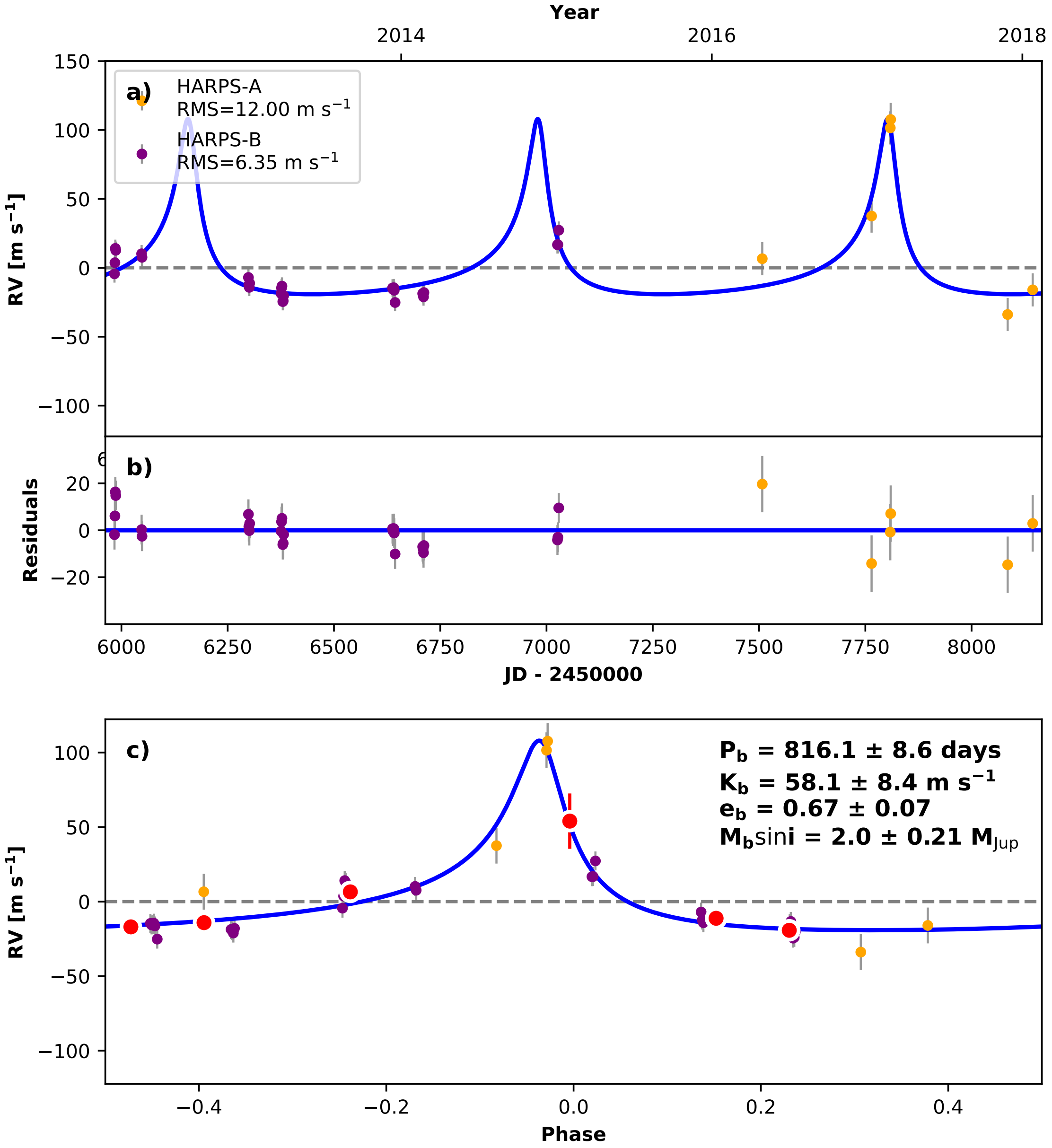

Figure 5.24: Best-fit 1-planet Keplerian orbital model for HIP43297. In all panels, the purple and yellow dots/curves are the HARPS observations before and after its upgrade in mid-2015, respectively. Label in panel (a) indicates the data RMS. (b) Residuals to the best fit 1-planet model. (c) RV phase diagram to the ephemeris of planet *b*. The RVs are binned in 0.08 phase units as red dots. The phase-folded model for planet *b* is shown in blue, along with the best-fit Keplerian orbital parameters (period, eccentricity and planet-induced Doppler amplitude), plus the planet's minimum mass.

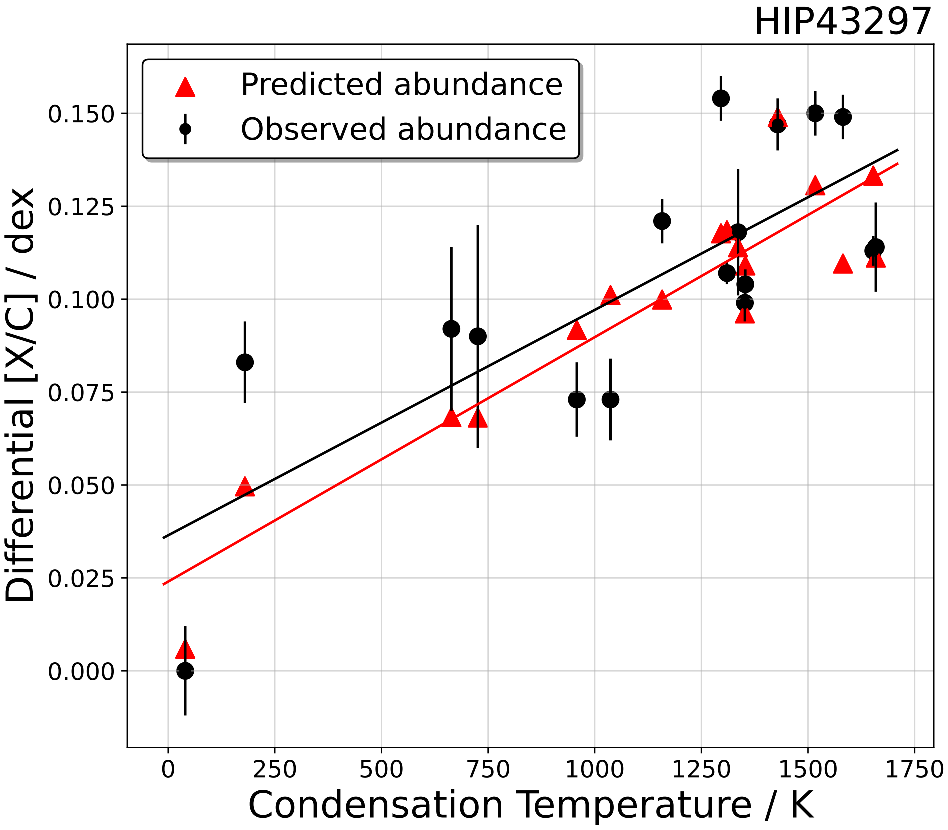

Figure 5.25: Comparison of the differential abundance of HIP43297 (black dots plus linear fit) as a function of dust condensation temperature in the protoplanetary disc, and the predicted abundances (red dots plus linear fit) estimated from a planetary accretion of 17.8 $M_{\oplus}$.

5.2.4 HIP44935 b: A cold gas giant planet under eccentric orbit

Between May 2009 and February 2017, HIP44935 was observed during 31 epochs using HARPS. Based on the RUWE value of 0.956, it is improbable that the host star is part of a binary, and its Gaia astrometry is reliable. The data before and after the upgrade have RMS values of 4.368 and 4.154 m s^{-1}, respectively.

Via the GLS algorithm on a standard grid, we detected a signal with a period of ∼506 days. The signal exhibited high statistical significance under sine-like assumptions, as evidenced by a BIC change of $||\Delta\mathrm{BIC}|| = 649$ relative to a non-varying noise model. We also conducted FAP tests with bootstrapping re-sampling and obtained that the likelihood of gathering a higher peak than 0.637 (which is associated with the 506-d period) in the RV periodogram with GLS powers of 0.600 and 0.551 has significance levels of 1% and 5%, respectively. Albeit noisy, we ruled out instrumental variations or astrophysical phenomena as the source of the proposed planet's orbital period via the Nyquist test. We evaluated six activity proxies for HIP44935 (c.f. Fig. 5.26): CCF BIS, CCF FWHM, $S-$index, Hα, Na doublet and He I. For all activity proxies, we observed weak or moderate-to-weak correlations, with moderate correlations between RVs and CCF FWHM ($r = 0.34$), and He I ($r = -0.32$). Despite that, at the proposed planet's orbital period, we found no evidence of time or phase modulations. Therefore, we fitted the orbit using a Keplerian with `RadVel`.

We chose the standard Keplerian basis, setting jitter noise terms for both HARPS epochs as free parameters. Accounting for an unobserved-with-current-data companion, we also set as a free parameter the systems barycentre motion $\dot{\gamma}$. Table 5.15 summarises the model priors and orbital solutions. The MCMC sampling was carried out using `RadVel`'s genetic routine, with 50 walkers, over 10000 steps and 8 ensembles. We ensured a good convergence of the model (minAfactor = 40.546, maxArchange = 0.009, and minTz = 13256.056), with a maximum $\hat{R}$ of 1.002. Models comparison demonstrated that an eccentric model (nominal BIC = 160.47, AIC = 156.14, and $\ln\mathcal{L} = -81.20$) is favoured by a $\Delta\mathrm{BIC}/\mathrm{AIC} = 3.37/4.37$ concerning the same eccentric model accounting for varying quantity $\dot{\gamma}$, $\Delta\mathrm{BIC}/\mathrm{AIC} = 6.03/5.20$ over a circular model with fixed $\dot{\gamma}$, and $\Delta\mathrm{BIC}/\mathrm{AIC} = 8.11/7.52$ over a circular model with varying $\dot{\gamma}$. Therefore, we state that an eccentric model without RV slope provides a suitable representation of the data, which corresponds to a cold Saturn-like planet under a relatively high eccentricity orbit (see Fig. 5.28).

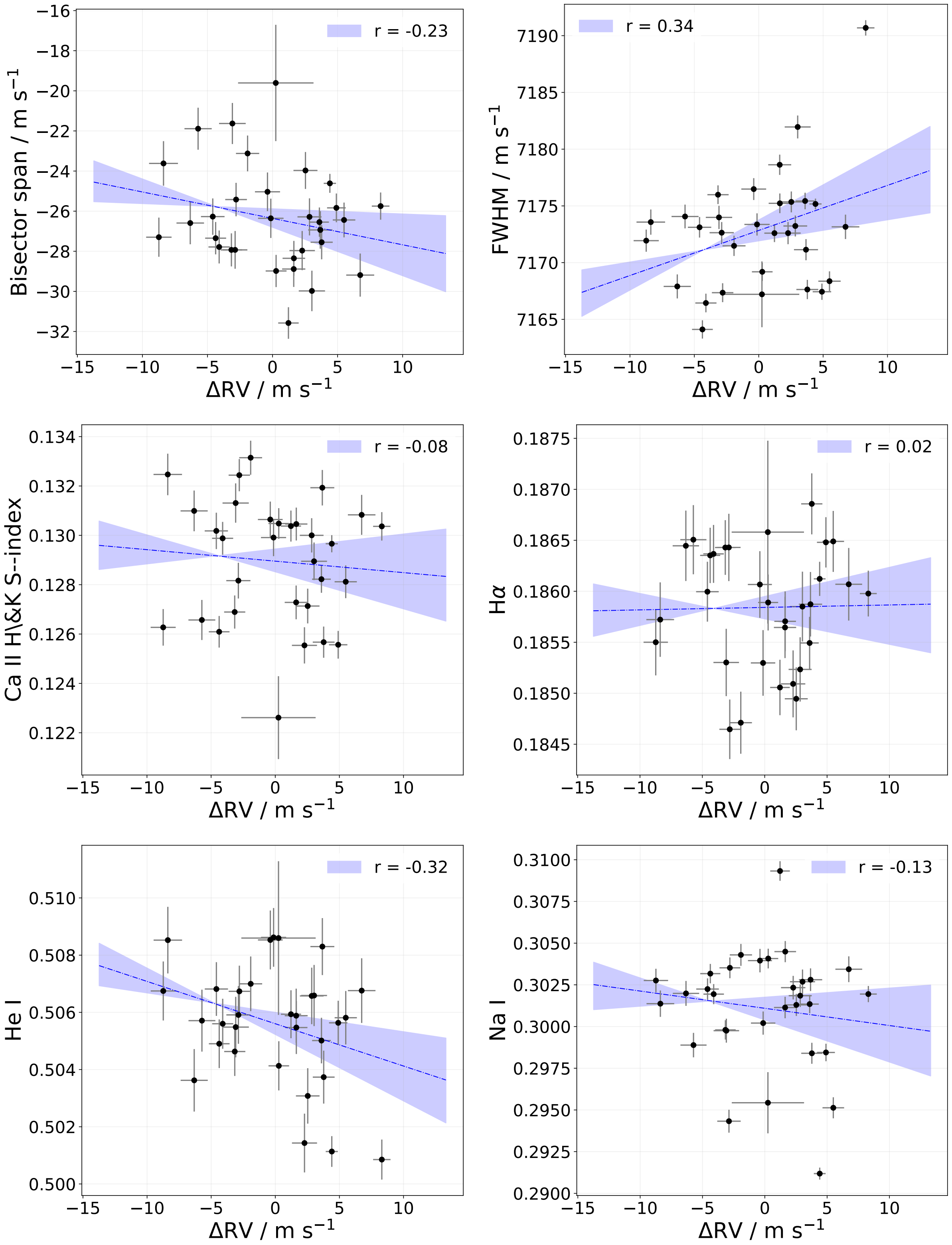

Figure 5.26: RV measurements and their relationship with stellar activity indicators for HIP44935.

Based on the chemical abundance pattern of HIP44935, we can deduce that the host star may also have undergone planet engulfment events in the past. By firstly estimating the convective

Parameter	**Prior**	**Credible Interval**	**Units**
P	$\mathcal{N}(506.2, 10)$	$507^{+8.3}_{-7.5}$	days
T_c	$\mathcal{N}(2457507, 100)$	2457264^{+41}_{-35}	days
K	$\mathcal{N}(10, 5)$	$6.5^{+2.2}_{-1.1}$	m s^{-1}
$\sqrt{e}\sin\omega$	$\mathcal{N}(0, 0.3)$	$-0.28^{+0.21}_{-0.19}$	–
$\sqrt{e}\cos\omega$	$\mathcal{N}(0, 0.3)$	$-0.52^{+0.21}_{-0.15}$	–
$\dot{\gamma}$	$]-\infty, +\infty[$	$-0.0013^{+0.0013}_{-0.0014}$	m s^{-1} d^{-1}
$\ddot{\gamma}$	–	$\equiv 0$	m s^{-1} d^{-2}
$\sigma_{\rm HARPS-A}$	$]-\infty, +\infty[$	$1.3^{+2.2}_{-4.6}$	m s^{-1}
$\sigma_{\rm HARPS-B}$	$]-\infty, +\infty[$	$2.14^{+0.51}_{-0.40}$	m s^{-1}
$m\sin i$	–	75^{+16}_{-11}	M$_\oplus$
a	–	$1.248^{+0.014}_{-0.012}$	AU

Table 5.15 - Best-fit 1-planet Keplerian parameters and derived quantities ($m\sin i$ and a), for HIP44935b.

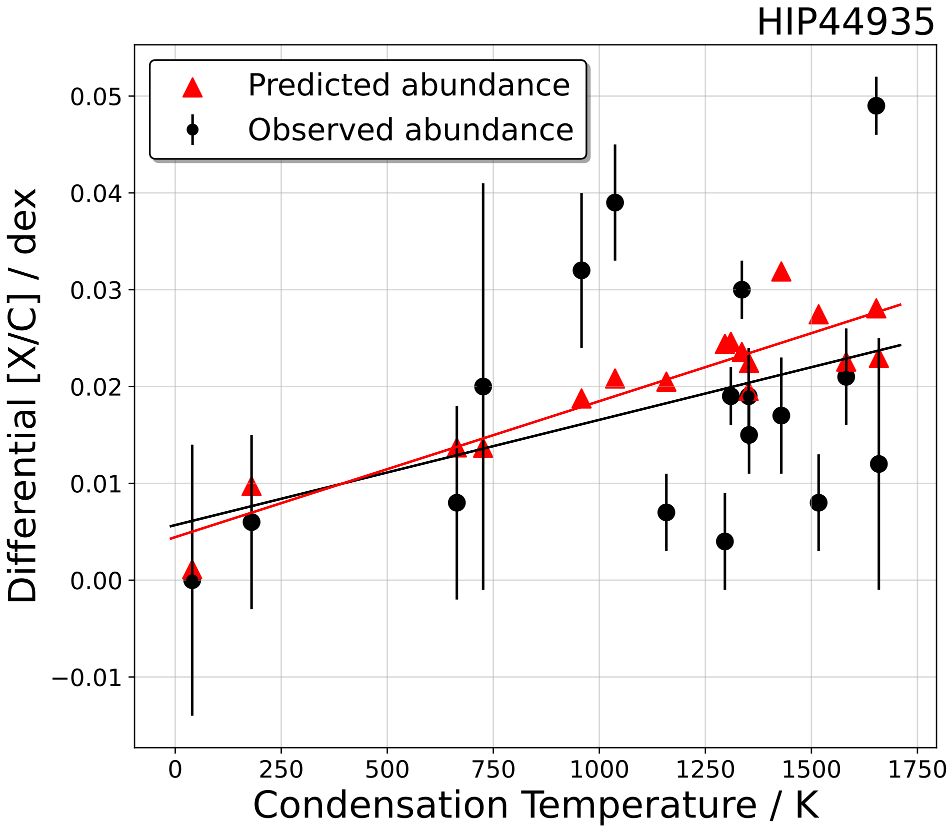

Figure 5.27: Comparison of the differential abundance of HIP44935 (black dots plus linear fit) as a function of dust condensation temperature in the protoplanetary disc, and the predicted abundances (red dots plus linear fit) estimated from a planetary accretion of 2.7 M$_\oplus$.

zone to be 0.024 M$_\odot$, we observe the [X/C] ratios for refractory/volatile elements as a function of condensation temperature T_C in Fig. 5.27 (slope $\alpha = 1.087 \times 10^{-5}$ dex). The figure also illustrates the impact of adding 2.7 M$_\oplus$ of Earth-like rocky composition material ($\alpha_{\rm model} = 1.404 \times 10^{-5}$ dex) into HIP44935's convective zone.

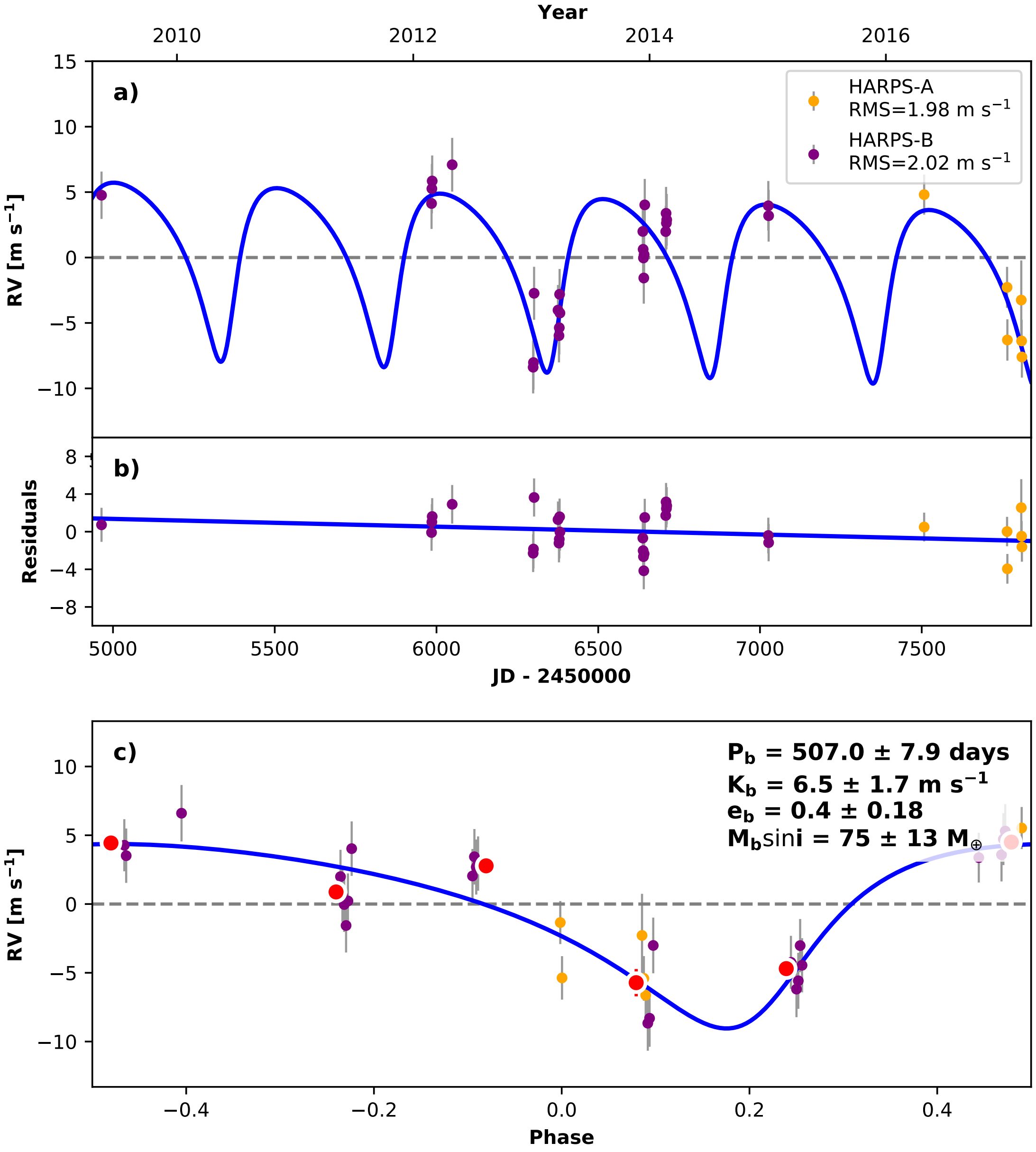

Figure 5.28: Best-fit 1-planet Keplerian orbital model for HIP44935. In all panels, the purple and yellow dots/curves are the HARPS observations before and after its upgrade in mid-2015, respectively. Label in panel (a) indicates the data RMS. (b) Residuals to the best fit 1-planet model. (c) RV phase diagram to the ephemeris of planet *b*. The RVs are binned in 0.08 phase units as red dots. The phase-folded model for planet *b* is shown in blue, along with the best-fit Keplerian orbital parameters (period, eccentricity and planet-induced Doppler amplitude), plus the planet's minimum mass.

5.2.5 HIP49756 b: A gas giant planet under eccentric orbit

HARPS observed HIP49756 for 43 epochs between February 2012 and May 2019. The host star is unlikely to be part of a binary, based on its RUWE value of 1.032, and its *Gaia* astrometry is reliable. The data before and after mid-2015 have RMS values of 4.026 and 1.917 m s^{-1}, respectively.

Using the GLS algorithm on a standard grid, we detected a signal with a period of ∼878 days. The signal had a high statistical significance under sine-like assumptions, as demonstrated by a BIC change of $\|\Delta\mathrm{BIC}\| = 1795$ compared to a non-varying model. We conducted FAP tests with bootstrapping re-sampling and determined that the likelihood of obtaining a higher peak than 0.663 (which is associated with the 878-d period) in the RV periodogram with GLS powers of 0.400 and 0.357 have significance levels of 1% and 5%, respectively. Despite the relatively noisy GLS periodogram, we ruled out instrumental modulations as the origin of the suggested planet's orbital period via a Nyquist frequency sampling test. We assessed six activity proxies for HIP49756 (see Fig. 5.29): CCF BIS, CCF FWHM, $S-$index, Hα, the Na doublet, and He I. We observed weak correlations between RVs and all tracers but the Na I doublet, whose Pearson's $r-$coefficient is 0.37. Nonetheless, we found no evidence of timing or phase modulations at the proposed planet's orbital period. Therefore, with the lack of statistical significance of RV-activity tracers correlations, and no significant variation at the planet's orbital period, we proceeded with a Keplerian fit making use of `RadVel`.

Similarly to HIP44935 (Section 5.2.4), we used the standard Keplerian, with jitter noise terms for both HARPS epochs as a free parameter. We did not include a barycentre motion $\dot{\gamma}$ in this case. Moreover, the orbital eccentricity and periastron argument in the coupled orthogonal basis were also set as free parameters, i.e., no priors were assigned. Table 5.16 summarises the model priors and orbital solutions obtained. We employed `RadVel`'s genetic routine for MCMC sampling with 50 walkers, 10000 steps and 8 ensembles, ensuring a good model convergence (minAfactor = 42.379, maxArchange = 0.024, and minTz = 4816.731), with a maximum $\hat{R}$ of 1.006. Our model comparison indicates that, concerning a circular model, the eccentric model is favoured by a difference $\Delta\mathrm{BIC}/\mathrm{AIC} = 2.35/3.63$ (nominal BIC eccentric = 205.92, AIC = 195.52, with $\ln\mathcal{L} = -99.72$). Therefore, we conclude that an eccentric model provides a suitable representation of the modulations in HIP49756's data, corresponding to a cold Saturn-like planet with a moderate orbital eccentricity (see Fig. 5.30).

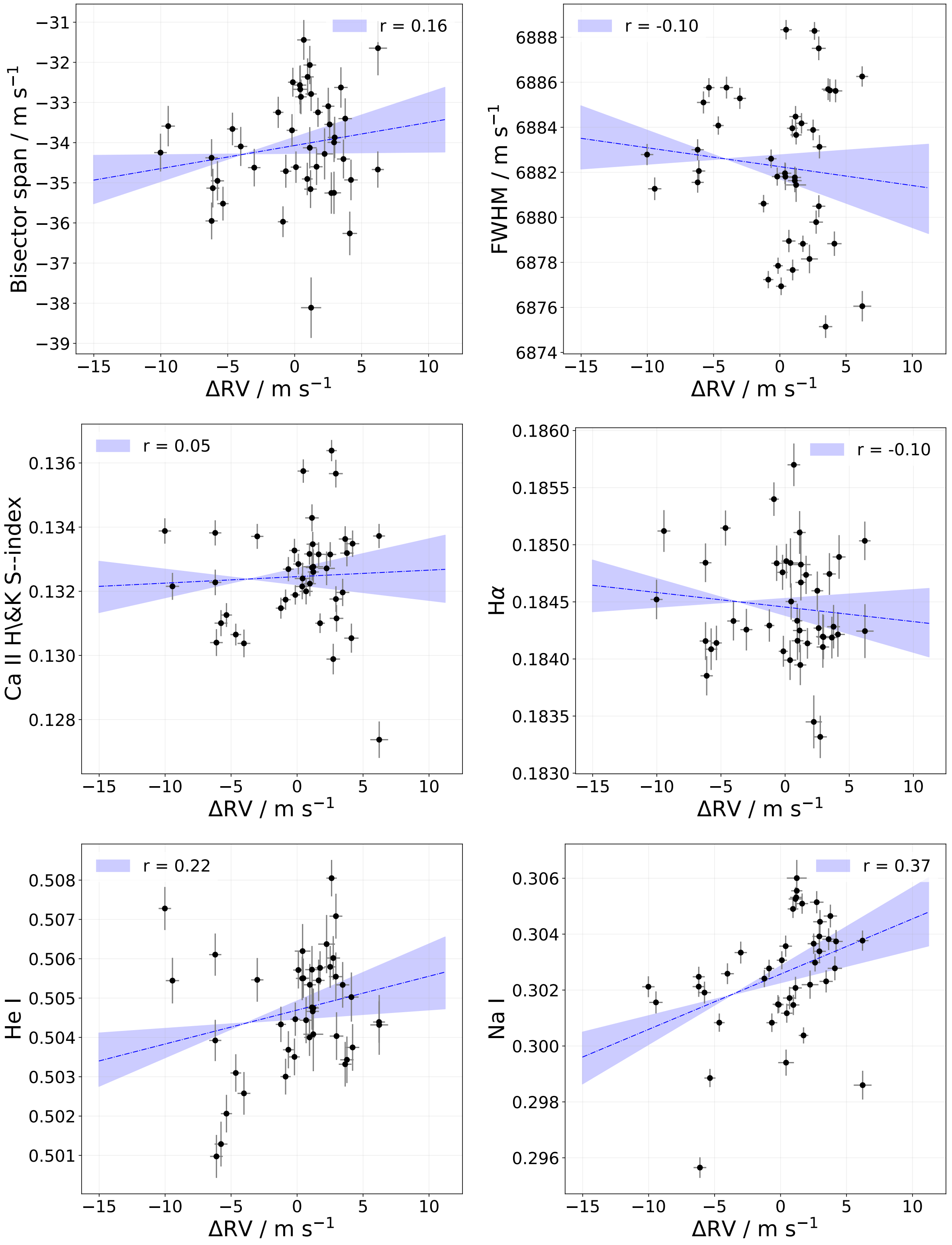

Figure 5.29: RV measurements and their relationship with stellar activity indicators for HIP49756.

Parameter	Prior	Credible Interval	Units
P	$\mathcal{N}(878.3, 10)$	$884^{+9.8}_{-9.9}$	days
T_c	$\mathcal{N}(2456042, 100)$	2456193^{+34}_{-36}	days
K	$\mathcal{U}(0.01, 30)$	$5.48^{+0.68}_{-0.58}$	m s^{-1}
$\sqrt{e}\cos\omega$	$-$	$-0.55^{+0.19}_{-0.11}$	$-$
$\sqrt{e}\sin\omega$	$-$	$0.11^{+0.15}_{-0.18}$	$-$
$\sigma_{\mathrm{HARPS-A}}$	$]-\infty, +\infty[$	$1.93^{+0.30}_{-0.24}$	m s^{-1}
$\sigma_{\mathrm{HARPS-B}}$	$]-\infty, +\infty[$	$1.7^{+2.0}_{-4.9}$	m s^{-1}
$m\sin i$	$-$	$77.5^{+8.3}_{-7.6}$	$M_\oplus$
a	$-$	$1.809^{+0.013}_{-0.014}$	AU

Table 5.16 - Best-fit 1-planet Keplerian parameters and derived quantities ($m\sin i$ and a), for HIP49756b.

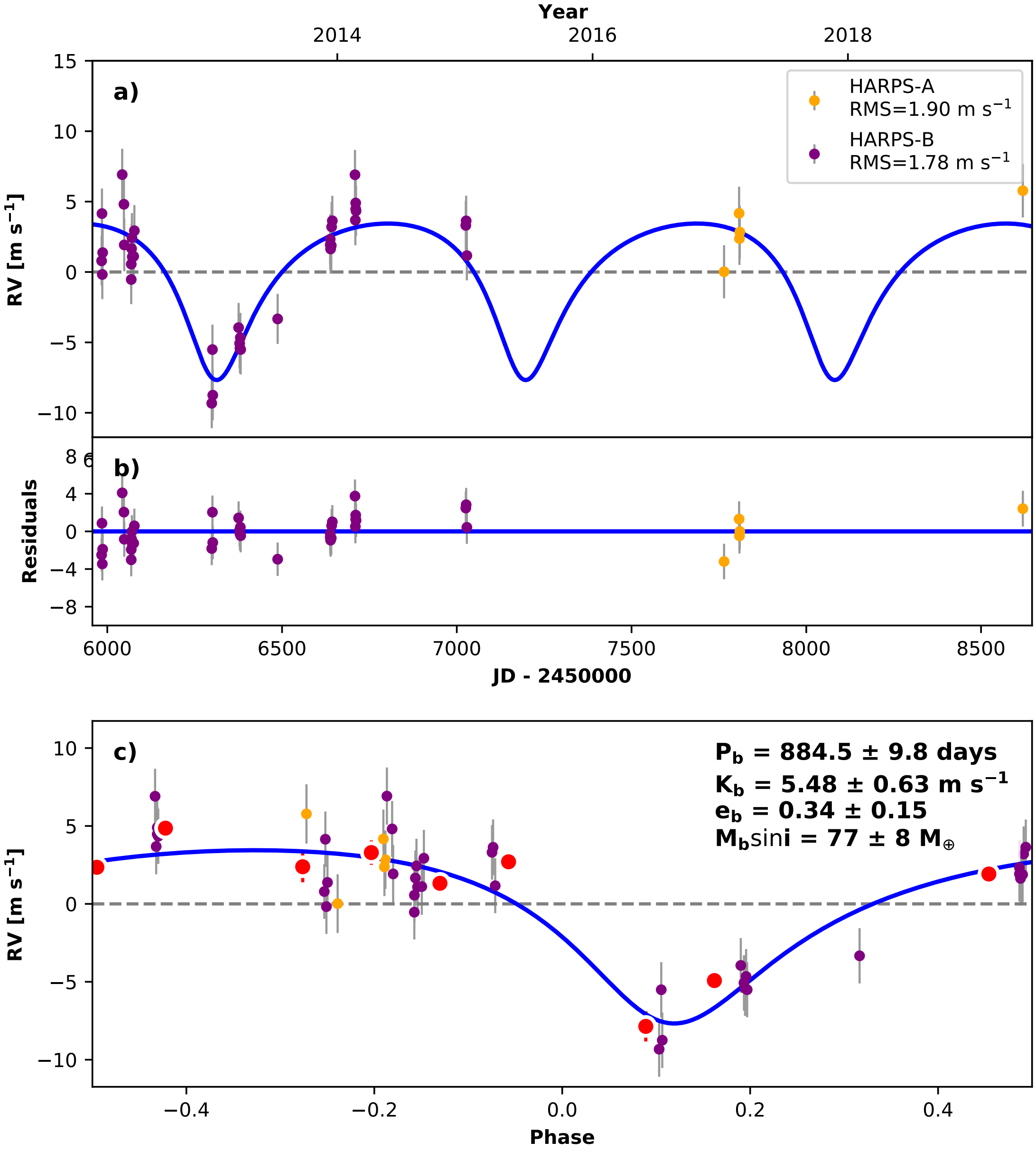

Figure 5.30: Best-fit 1-planet Keplerian orbital model for HIP49756. In all panels, the purple and yellow dots/curves are the HARPS observations before and after its upgrade in mid-2015, respectively. Label in panel (a) indicates the data RMS. (b) Residuals to the best fit 1-planet model. (c) RV phase diagram to the ephemeris of planet b. The RVs are binned in 0.08 phase units as red dots. The phase-folded model for planet b is shown in blue, along with the best-fit Keplerian orbital parameters (period, eccentricity and planet-induced Doppler amplitude), plus the planet's minimum mass.

5.2.6 HIP72043 b: A cold Jupiter-like planet under eccentric orbit

From February 2012 to February 2017, the ESO/HARPS spectrograph monitored HIP72043 during 34 epochs. The RUWE value of HIP72043 is high, at 1.167, which likely suggests (a) that the astrometric data for this particular star may be of lower quality compared to other stars observed by *Gaia*, and/or (b) the presence of a sub-stellar companion. A previous study by Kervella et al. (2019) set the presence of a $\sim$0.07 $M_\odot$ (76.74 M_{Jup}, normalised at 1 AU) companion orbiting HIP72043 by analysing anomalies in the proper motion of the star with the *Gaia* data. Albeit the system's proper motion is relatively high ($\mu_\alpha^* = -59.034 \pm 0.027$ mas yr^{-1}, $\mu_\delta = -165.142 \pm 0.027$ mas yr^{-1}), it is consistent with the other Solar Twin stars reported in this writing.

We employed a GLS algorithm over the standard grid, identifying a statistically significant ($\|\Delta\mathrm{BIC}\| = 11\,880$ compared to a non-varying noise model) signal close to 1093 days. FAP tests confirmed the significance of the signal, which revealed that the likelihood of obtaining a higher peak than 0.854 (associated with the 1093-d period) is small with GLS powers of 0.484 and 0.427 with 1% and 5% of significance. Moreover, we ruled out the possibility that the proposed planet's orbital period is an alias from instrumental modulations testing it with the Nyquist frequency sampling test with one sidereal day, one synodic month, one sidereal year, and a Sun's time activity cycle of roughly eleven years. For HIP72043 we examined a set of six activity tracers (see Fig. 5.31): CCF BIS, CCF FWHM, the $S-$index, Hα, the Na doublet and the He I line. We observe moderate correlations between the RV measurements and CCF BIS ($r = -0.40$), CCF FWHM ($r = -0.30$), $S-$index ($r = -0.49$), and Na ($r = 0.36$). However, at the proposed planet's orbital period, we found no evidence of timing or phase modulations, which suggests that, even in the presence of stellar magnetic activity, we expect a clear sign of planetary modulation.

The standard Keplerian basis was used to model HIP72043 variations. Noise jitter terms for both HARPS epochs were set as free parameters, as of the orbital eccentricity and periastron argument coupled in the orthogonal basis ($\sqrt{e}\sin\omega$, $\sqrt{e}\cos\omega$). Table 5.17 summarises the model priors and the orbital solutions retrieved (see Fig. 5.17). The MCMC sampling was conducted using a genetic routine with 50 walkers, over 10000 steps and 8 ensembles. The model achieved convergence with minAfactor = 42.100, maxArchange = 0.022, minTz = 12503.322, and maximum $\hat{R}$ of 1.002. The model had a nominal BIC of 217.37, AIC of 210.99 and $\ln\mathcal{L}$ of -114.59.

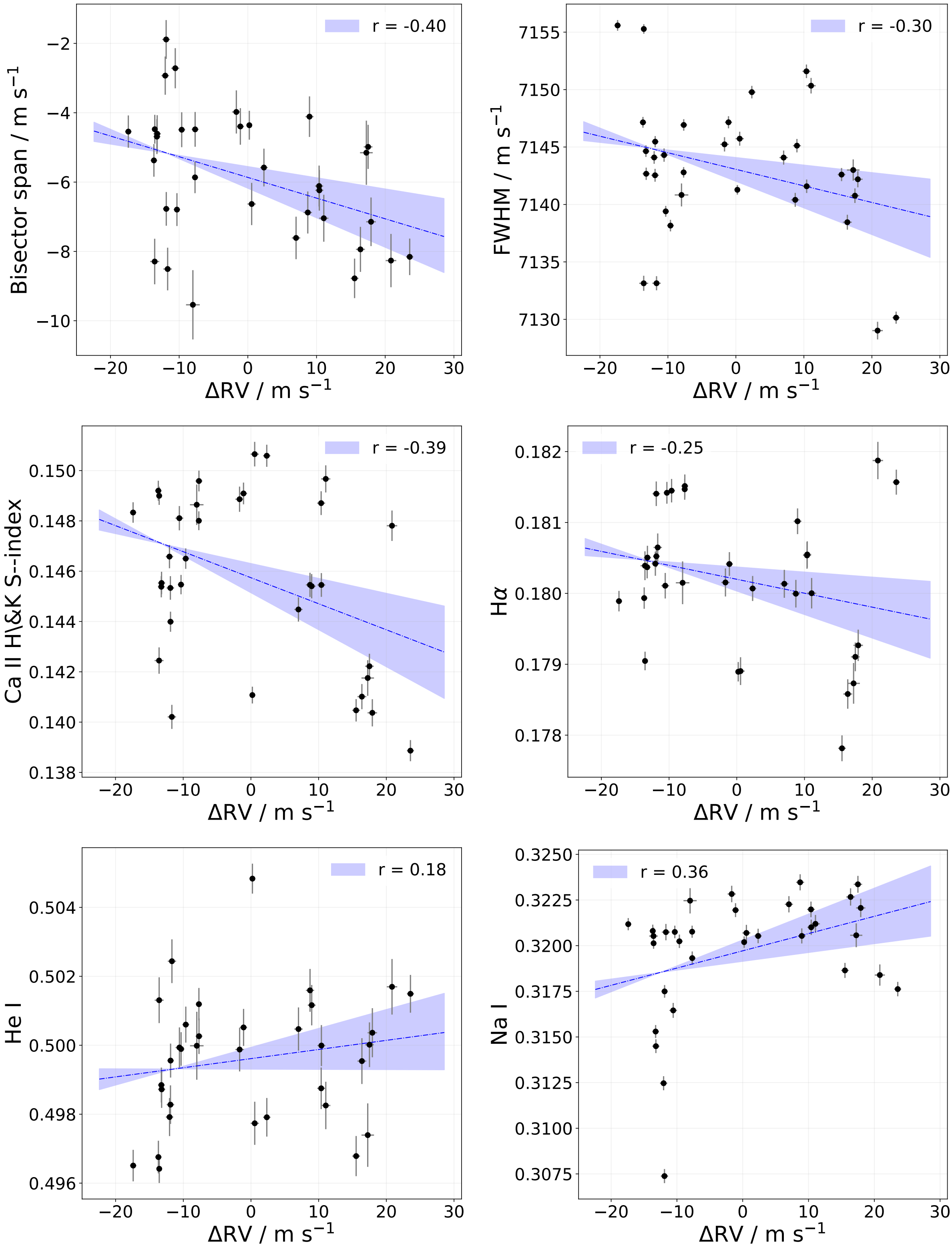

Figure 5.31: RV measurements and their relationship with stellar activity indicators for HIP72043.

Parameter	**Prior**	**Credible Interval**	**Units**
P	$\mathcal{N}(1093.3, 100)$	1023^{+40}_{-36}	days
T_c	$\mathcal{N}(2457809.0, 100)$	2457973^{+83}_{-73}	days
K	$\mathcal{U}(0.01, 40)$	$18.0^{+1.5}_{-1.6}$	m s^{-1}
$\sqrt{e}\cos\omega$	$-$	$0.24^{+0.18}_{-0.19}$	$-$
$\sqrt{e}\sin\omega$	$-$	$-0.34^{+0.10}_{-0.17}$	$-$
$\sigma_{\mathrm{HARPS-A}}$	$]-\infty, +\infty[$	$5.32^{+0.95}_{-0.73}$	m s^{-1}
$\sigma_{\mathrm{HARPS-B}}$	$]-\infty, +\infty[$	$1.5^{+3.0}_{-5.7}$	m s^{-1}
$m\sin i$	$-$	$0.861^{+0.074}_{-0.096}$	M$_{\mathrm{Jup}}$
a	$-$	$2.004^{+0.053}_{-0.048}$	AU

Table 5.17 - Best-fit 1-planet Keplerian parameters and derived quantities ($m\sin i$ and a), for HIP72043b.

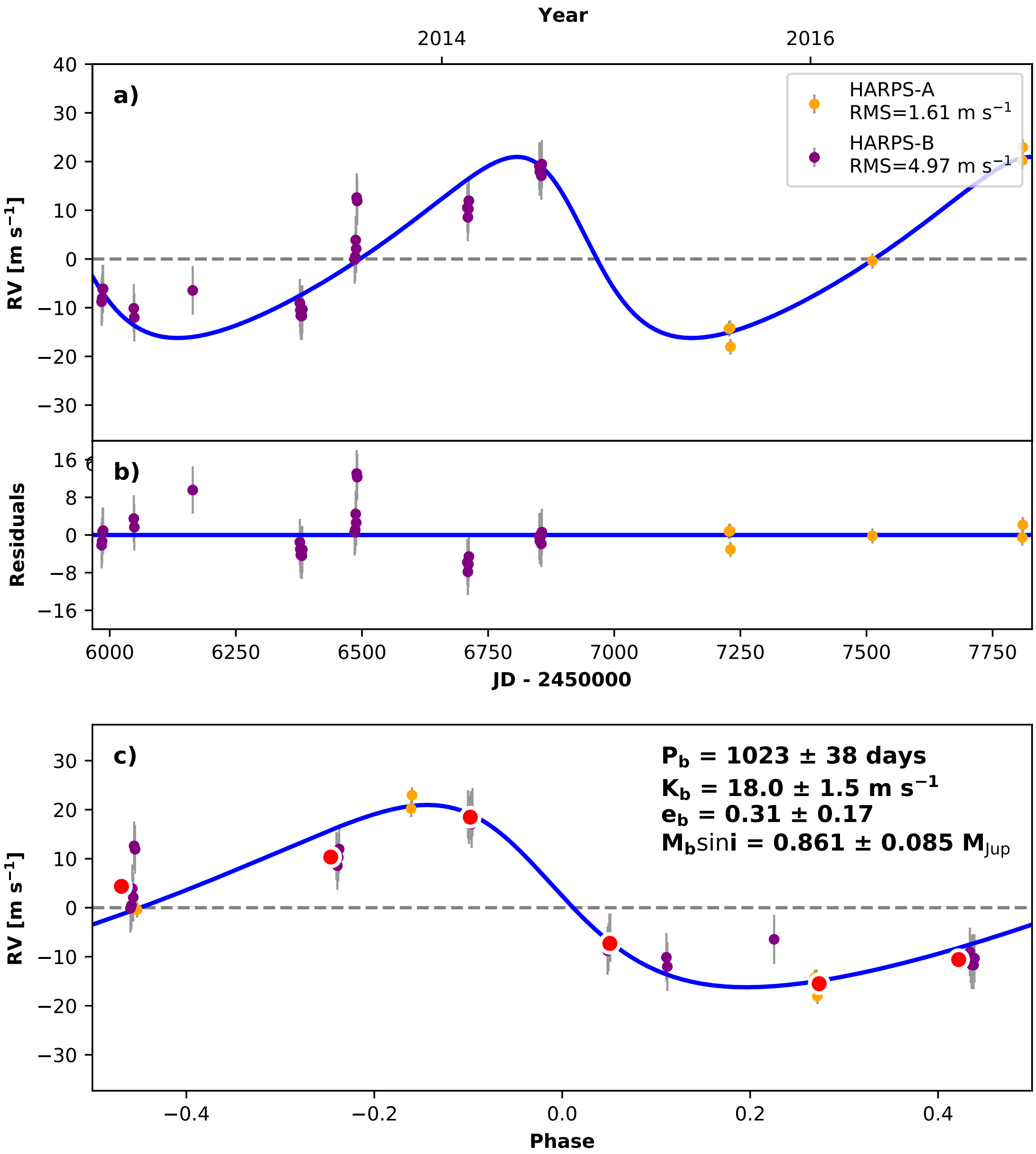

Figure 5.32: Best-fit 1-planet Keplerian orbital model for HIP72043. In all panels, the purple and yellow dots/curves are the HARPS observations before and after its upgrade in mid-2015, respectively. Label in panel (a) indicates the data RMS. (b) Residuals to the best fit 1-planet model. (c) RV phase diagram to the ephemeris of planet b. The RVs are binned in 0.08 phase units as red dots. The phase-folded model for planet b is shown in blue, along with the best-fit Keplerian orbital parameters (period, eccentricity and planet-induced Doppler amplitude), plus the planet's minimum mass.

5.2.7 HIP87769 b: A cold gas giant planet under eccentric orbit

Between April 2012 and July 2018, the ESO/HARPS spectrograph observed HIP87769 across 34 epochs. The system's RUWE value of 1.655 indicates that the likelihood of the system being part of a binary system is high, which is also seen from the strong RV slope in the raw (prior modelling) RV curve. The RMS value of the data collected before mid-2015 was 66.051 m s^{-1}, while data collected after the HARPS' upgrade had an RMS value of 81.797 m s^{-1}, which indicates that the object is experiencing significant fluctuations in its RV over time, possibly due to the presence of an orbiting companion. Indeed, there is a known sub-stellar companion in this system with orbital period $P > 81$ years and $m \sin i > 0.30$ M$_\odot$ (314 M$_{\rm Jup}$), which may account for the strong RV slope observed in the data (dos Santos et al., 2017).

A GLS algorithm was employed on the standard grid, resulting in an extremely high significant ($\|\Delta\text{BIC}\| = 146\ 952$ compared to a non-varying model) signal close to 1413 days. A FAP bootstrapping re-sampling test resulted in a low probability of obtaining a higher peak than 0.815 (associated with the prominent period obtained) with GLS powers of 0.499 and 0.431 with 1% and 5% significance levels, respectively. To ensure that the proposed planet's orbital period was not an alias from instrumental modulations, we tested it with the Nyquist frequency sampling test. For HIP87769, we examined two activity indicators proxies (CCF BIS and CCF FWHM; see Fig. 5.33), identifying a moderate ($r = -0.36$) correlation between the RVs and CCF BIS, but weak correlation with CCF FWHM. None of these two activity indicators, however, demonstrated strong timing or phase modulations at the proposed planet's orbital period.

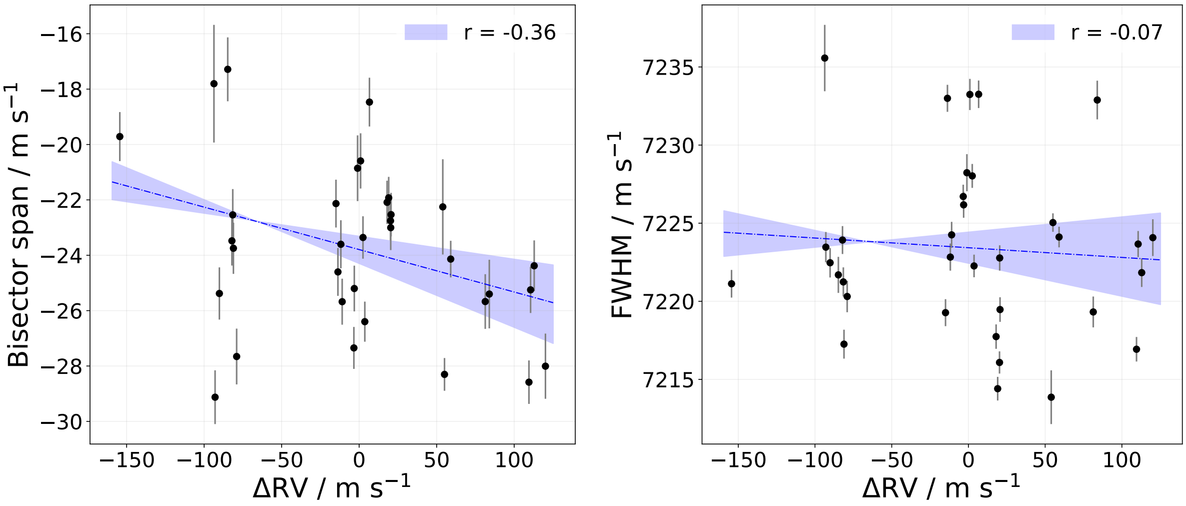

Figure 5.33: RV measurements and their relationship with stellar activity indicators for HIP87769.

We also employed the standard Keplerian basis and treated jitter noise terms for both HARPS epochs as free parameters. Additionally, we considered the barycentre motion of the system $\dot{\gamma}$, as well as its curvature $\ddot{\gamma}$ as a free parameter to account for any unobserved companion. The priors and orbital solutions for the best-fit model are summarised in Table. 5.18. We performed the MCMC sampling using `RadVel`'s genetic routine with 50 walkers over 10000 steps and 8 ensembles, and we ensured a good convergence of the model (minAfactor = 38.402, maxArchange = 0.003, minTz = 8898.103, and maximum $\hat{R} = 1.003$). By treating the RV curvature $\ddot{\gamma}$, as well as its slope $\dot{\gamma}$ would result in a nearly indistinguishable model ($\Delta\mathrm{BIC/AIC} = 0.06/0.13$) concerning the best-fit model (nominal BIC = 213.18, AIC = 206.94, and $\ln\mathcal{L} = -178.37$). It also is somewhat unfavourable a model in which the orbital eccentricity is held constant ($\Delta\mathrm{BIC/AIC} = 2.26/2.28$). Hence, we conclude that an eccentric model with RV slope is a suitable representation of the data modulations of HIP87769, corresponding to a cold Saturn-like planet with a relatively high orbital eccentricity (see Fig. 5.34).

Parameter	**Prior**	**Credible Interval**	**Units**
P	$\mathcal{N}(1413.7, 100)$	1419^{+100}_{-120}	days
T_c	$\mathcal{N}(2456043, 100)$	2456088±100	days
K	$\mathcal{N}(6, 10)$	$6.9^{+2.0}_{-1.8}$	$\mathrm{m\ s^{-1}}$
$\sqrt{e}\sin\omega$	$\mathcal{N}(0, 0.3)$	$-0.46^{+0.27}_{-0.20}$	–
$\sqrt{e}\cos\omega$	$\mathcal{N}(0, 0.3)$	$-0.35^{+0.43}_{-0.22}$	–
$\dot{\gamma}$	$]-\infty, +\infty[$	$-0.2456^{+0.0059}_{-0.0053}$	$\mathrm{m\ s^{-1}\ d^{-1}}$
$\ddot{\gamma}$	$]-\infty, +\infty[$	$2\times10^{-5}\ {}^{+4.1\times10^{-6}}_{-4\times10^{-6}}$	$\mathrm{m\ s^{-1}\ d^{-2}}$
$\sigma_{\mathrm{HARPS-A}}$	$]-\infty, +\infty[$	$1.7^{+9.4}_{-13}$	$\mathrm{m\ s^{-1}}$
$\sigma_{\mathrm{HARPS-B}}$	$]-\infty, +\infty[$	$3.13^{+0.79}_{-0.54}$	$\mathrm{m\ s^{-1}}$
$m\sin i$	–	$0.337^{+0.11}_{-0.08}$	$\mathrm{M_{Jup}}$
a	–	$2.5^{+0.12}_{-0.14}$	AU

Table 5.18 - Best-fit 1-planet Keplerian parameters and derived quantities ($m\sin i$ and a), for HIP87769*b*.

HIP87769 also exhibits a chemical peculiarity that may be related to moderate planet engulfment events. Fig. 5.35 illustrates an enhancement in refractory elements for the host star (represented by black dots/curve), with a linear-best-fit slope of $\alpha = 7.44 \times 10^{-6}$ dex. To validate the planet accretion hypothesis, we used the TERRA code, in which based on the host star's mass and metallicity, we firstly estimate its convective zone mass as 0.022 $\mathrm{M_{\odot}}$. The red dots/curve in Fig. 5.35 represents the [X/C] ratios as a function of the condensation temperature T_C ($\alpha_{\mathrm{model}} = 9.02 \times 10^{-6}$ dex), taking into account the effect of 1.7 of Earth-like rocky

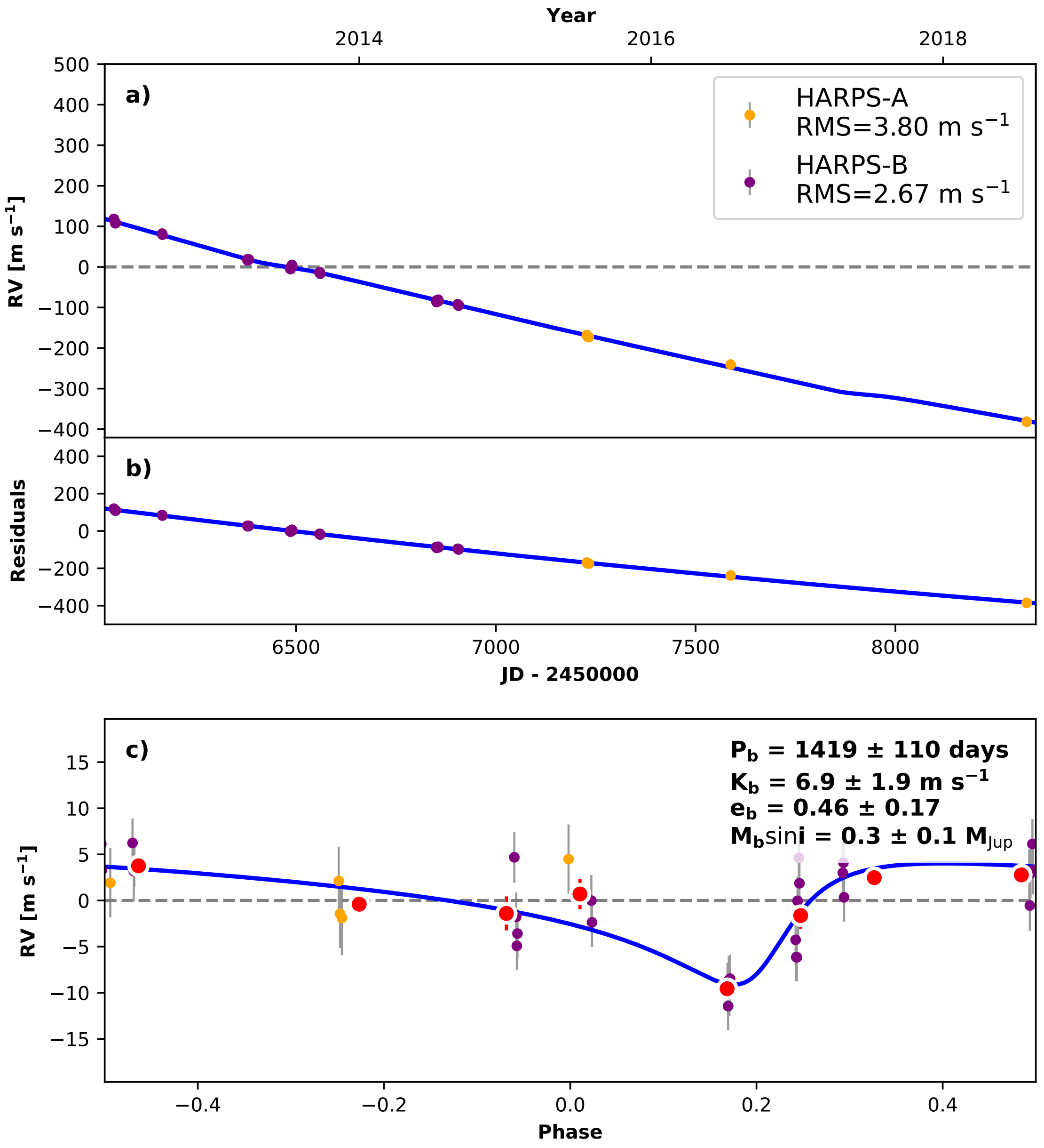

Figure 5.34: Best-fit 1-planet Keplerian orbital model for HIP87769. In all panels, the purple and yellow dots/curves are the HARPS observations before and after its upgrade in mid-2015, respectively. Label in panel (a) indicates the data RMS. (b) Residuals to the best fit 1-planet model. (c) RV phase diagram to the ephemeris of planet *b*. The RVs are binned in 0.08 phase units as red dots. The phase-folded model for planet *b* is shown in blue, along with the best-fit Keplerian orbital parameters (period, eccentricity and planet-induced Doppler amplitude), plus the planet's minimum mass.

composition material into the host star's convective zone[9].

---

[9] It is important to highlight the divergence between the model and observations for HIP87769, particularly for elements with condensation temperatures ranging from $T_c = 600 - 1100$ K during planetary engulfment episodes. This divergence can be attributed to several factors, such as that (*i*) elements in this temperature range may exhibit weaker absorption lines, making their detectability and measurability a challenge and inaccurate, (*ii*) line blending,

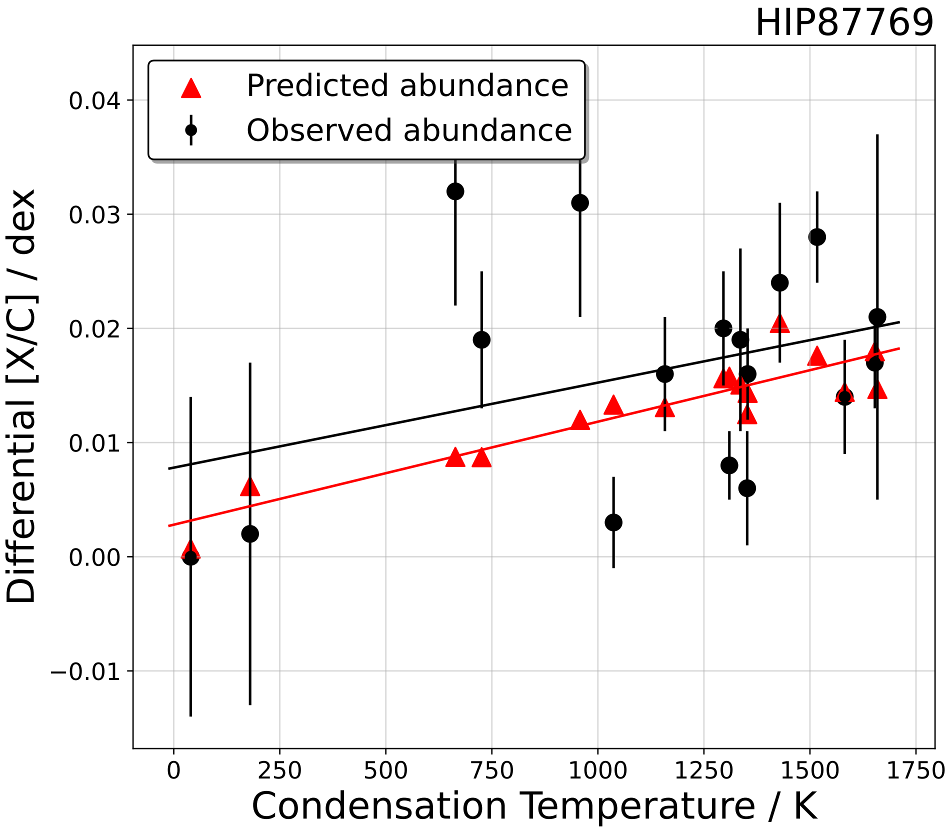

Figure 5.35: Comparison of the differential abundance of HIP87769 (black dots plus linear fit) as a function of dust condensation temperature in the protoplanetary disc, and the predicted abundances (red dots plus linear fit) estimated from a planetary accretion of 1.7 $M_\oplus$.

In cases where a linear trend is observed in the data, the most conservative approach is to provide a minimum mass for the putative companion (either a massive planet or a brown dwarf) that gives rise to this feature when the orbital eccentricity is close to 0.5 (so it is considered a dynamically excited planet), and the periastron argument is close to $\pi/2$. Following Wright and Howard (2009) and Feng et al. (2015), with $\tau \equiv 1.25 \times W$ (where W is the time-span of the observations), and $K \sim \tau\dot{\gamma}$, this minimum mass can be estimated through

$$m_{\min} \sim (0.0164\ \mathrm{M_{Jup}}) \left(\frac{\tau}{yr}\right)^{4/3} \left|\frac{\dot{\gamma}}{\mathrm{m\ s^{-1}\ yr^{-1}}}\right| \left(\frac{M_\star}{M_\odot}\right)^{2/3} \tag{5.2}$$

---

and (*iii*) temperature gradients that are not accounted for the model, both within the engulfed object of about 1.7 $M_\oplus$, and its surrounding medium, that can result in variations in line strengths and shapes, introducing slight inaccuracies in the determination of elemental abundances.

For the retrieved value of $\dot{\gamma}$, and from the RV baseline of approximately 6.25 years, we estimated with eq. 5.2 the presence of a companion with a minimum mass of 23.41 $\mathrm{M_{Jup}}$ (∼0.022 $\mathrm{M_{\odot}}$; thus a brown dwarf) orbiting HIP87769, from which only a small portion of its orbit is observable. It is important to highlight that this minimum mass just represents a lower limit estimated, i.e., it assumes an orbital inclination of 90 degrees. The true mass of the companion, if existent, could be higher if the orbit is more elliptical, or if its inclination is lower. Furthermore, other factors such as stellar activity and instrumental errors can affect the accuracy of the RV measurements, and consequently the calculated mass. Follow-up of these kinds of targets is necessary and encouraged, not only to confirm the measured planet's nature, but also to confirm the existence of the companion (either with more RV measurements, transits if the companion is at short period orbits and $i \sim 90^{\circ}$, or even direct imaging if the orbiting companion is at longer distances from its host star).

The detection of an outer companion in a planetary system, whether it presents high orbital eccentricity or not, has been known to have consequential implications on the dynamics of the inner planets. Specifically, the gravitational influence of an outer companion can lead to complex interactions among the planets, which can result in a manifold of changes in their orbital parameters. Several models and theories have emerged in recent years to explain the high eccentricity of the exoplanets found. One such effect of these interactions is the von Zeipel-Lidov-Kozai (ZLK) mechanism, a well-known phenomenon where a planet's orbit can undergo significant changes in eccentricity due to perturbations of an outer companion from exchanges of angular momentum between the pairs (e.g., Lidov 1962; Kozai 1962). In this case,

$$\dot{\mathbf{J}} = \mathbf{\Omega}(\mathbf{J}), \text{ and } \dot{\mathbf{e}} = \frac{1}{n}\left(\mathbf{h} \times \mathbf{\Omega}(\mathbf{J})\right), \tag{5.3}$$

where $\mathbf{J}$ is a planet's angular momentum vector, $\mathbf{e}$ is its eccentricity, n is the mean motion, and $\mathbf{h}$ is the planet's angular momentum per unit mass. The function $\mathbf{\Omega}(\mathbf{J})$ represents the perturbing torque due to the outer companion, and it is given by:

$$\mathbf{\Omega}(\mathbf{J}) = -\frac{3}{4}\frac{GM_c}{a^3}\frac{(\mathbf{J}\cdot\mathbf{e}_c)\mathbf{e}_c - \mathbf{J}}{(1-e_c^2)^{3/2}}, \tag{5.4}$$

where G is the gravitational constant, M_c is the mass of the outer companion, a is the semi-major axis of the planet's orbit, and $\mathbf{e}_c$ is the eccentricity of the outer companion (Shevchenko, 2017). From the equation for $\mathbf{\Omega}$, we observe that the strength of the perturbation on the inner planet is directly proportional to the outer companion's mass. In addition, the perturbing torque is inversely proportional to the orbital semi-major axis of the inner planet, and as such, the closer the planet moves towards the outer companion, the stronger the orbital perturbation becomes. Lastly, the alignment of the planet's orbital momentum with the eccentricity vector of the outer companion increases the magnitude of the perturbing torque. We highlight that the ZLK mechanism can be confirmed in a planetary system through observations of changes in the eccentricity and orbital inclination of the planet's orbit (e.g., Pont et al. 2009; Naoz et al. 2011, and references therein). In particular, if a planet's eccentricity or inclination is observed to undergo significant changes over time, and these changes can be attributed to the gravitational influence of an outer companion, this may be strong evidence of the ZLK mechanism in action. These studies require complex mathematical modelling and advanced analysis which is outside the scope of this writing. It is important, however, to recognise that the complex interplay between the planets in a system and the outer companion is not always straightforward, and can involve various additional factors such as gravitational resonances and secular effects (e.g., Batygin et al. 2011; Bryan et al. 2016; Millholland and Laughlin 2019, and references therein). As such, a thorough investigation into the dynamical evolution of the system is necessary to corroborate this hypothesis and provide further insight into the underlying mechanisms governing the observed planetary system behaviour.

5.2.8 HIP104045 b: A Jupiter analogue under circular orbit

HIP 104045 was observed between July 2006 and September 2019 with the *High Accuracy Radial Velocity Planet Searcher* (HARPS; Mayor et al. 2003) échelle spectrograph connected to the 3.6–m telescope of the European Southern Observatory in La Silla, Chile. HARPS covers the optical range ($\lambda \in [378.1, 691.2]$–nm) with high resolving power ($R = 115\ 000$) and it is installed in a pressure– and temperature–controlled enclosure, which allows a high instrumental precision. For each spectrum obtained, the instrument simultaneously observes a Thorium-Argon hollow-cathode lamp as a reference source, and the data are processed in near real-time with the HARPS ESO pipeline through cross-correlation with a G2 spectral type mask.

A total of 39 observations from October 2011 to October 2016 were obtained for the STPS programme using an exposure time of 15 minutes, while 12 additional data points were obtained by our group or other programmes from the ESO Advanced Data Products Archive[10], comprising 2 observations in July/August 2006, plus 10 data points from October 2017 to September 2019.

We implemented a Generalised Lomb-Scargle algorithm (GLS; Ferraz-Mello 1981; Zechmeister and Kürster 2009; VanderPlas 2018) in a $0.1-$d to $20\ 000-$d grid, finding out initially two prominent signals close to 2145–d and 306–d. In addition to a constant Doppler offset, we implemented Bayesian GLS regressions accounting for intrinsic harmonic, offset and with/without linear trends (BGLST/BGLSnT) in the data following the method outlined in Olspert et al. (2018). Both Bayesian periodograms presented in Fig. 5.36 support, in principle, the detection of the two periodic signals in the data. However, by implementing a Nyquist frequency sampling test, we found out that the longer period represents an alias of the shorter one, whilst the latter in turn is a close alias of one sidereal year as verified using the blind-search methodology outlined in Rosenthal et al. (2021). From the same Nyquist test and blind-search framework, we found that the longer period is neither an alias resulting from instrumental (e.g., alias from one sidereal day, 1.0–d, nor one sidereal year, 365.25–d), nor from stellar modulations. We highlight periodic signals close to 29–d in the GLS periodogram. The rotation of a star is not uniform across its surface due to its non-rigid nature, so other latitudes rotate differently than the equator. If a spot or plage, for instance, is located at a higher latitude, it will take longer to complete a full rotation than a spot or plage at a lower latitude (e.g., Robertson et al. 2014; Lorenzo-Oliveira et al. 2019; do

---

[10] `https://archive.eso.org/wdb/wdb/adp/`

Nascimento et al. 2020). Consequently, presumably, radial velocity measurements taken at different times will reflect the presence of the spots or plages at different rotational phases, as seen in the several peaks near this period in the periodogram. Presently, the hypothesis positing that these near one-month periods originate from stellar rotation is adopted as a working assumption and handled within the Gaussian Process (GP) model, and further comprehensive investigation will be addressed in future works (Carvalho Silva et al., *in prep.*).

False-alarm probability (FAP) tests were performed via both bootstrap re-sampling (Ivezić et al., 2014) and extreme value statistics method (Baluev, 2008), where we found out a probability to obtain a higher peak with GLS–powers of 0.447 and 0.386 in the radial velocity periodograms with 1% and 5% of probability, respectively, for bootstrapping, while GLS–powers of 0.350 and 0.300 are associated to FAPs of 1% and 5% using Baluev's approach. A single–period FAP test was performed at the proposed planet's orbital period, revealing a FAP of 5.209×10^{-6} for the prominent ~ 2145 days period, ruling out their spurious nature. Additionally, we also computed the change in the Bayesian Information Criterion (BIC) with a non-varying noise model following Ivezić et al. (2014) (*c.f.* equation 10.54) finding out that the GLS–power of 2145 period corresponds to $\|\Delta\mathrm{BIC}\| = 1233.6$, which demonstrates the high-statistical significance of the signal under sine-like assumptions. An inspection of the activity proxies at the proposed planet's orbital period also revealed, for all, no time or phase modulations (see Fig. 5.36). Albeit a strong peak is visible in the Hβ periodogram close to the proposed planet's orbital period, it is down below the 15% FAP level threshold significance. Such time correlations would be expected if the period were the result of magnetic stellar activity cycles (e.g. atmospheric expansion or pulsations, which are well-known correlated with Hβ; Sobolev 1960; Doazan and Peton 1970), and although activity cycles trend analogous to the Sun's 11–year cycle (in addition to its lower and upper order harmonics) being also expected to cause a suppression of convective blueshift, consequently changing the observed radial velocity of Sun-like stars in similar time-scales (e.g. Saar and Donahue 1997; Hall et al. 2007), we did not also observe significant correlations of these with HIP 104045 radial velocity measurements. In present-day analysis, the most common stellar activity tracer is the S$-$index, which represents the flux ratio of the core of the Ca II H&K lines to the continuum nearby (Lorenzo-Oliveira et al., 2018; Gomes da Silva et al., 2018; Dumusque et al., 2021), and may even correlate with Hα emission (Cincunegui et al., 2007).

To quantify the RV-activity proxies correlations significance, we applied the Pearson r correlation and the null-hypothesis significance testing $p-$value. Fig. 5.37 demonstrates that solely Hδ correlate moderately and significantly with the radial velocity residuals (for moderately, $|r| \geq 0.3$, and significantly, $p < 0.05$), while all other tracers demonstrated weak correlations with the radial velocity residuals ($|r| < 0.3$). Magnetic stellar activity effects were taken under consideration to determine the best-fit orbital solution of planet HIP 104045 b.

Taking into account magnetic stellar activity effects with Gaussian Process (GP) regression, and that HIP104045 companions follow Keplerian orbits, we proceeded with the orbital fit making use of RadVel, which implements a maximum a posteriori probability density via Markov chain simulations (Fulton et al., 2018). Table 5.19 summarises the model prior's parameters used in this analysis, which are described in the following.

We choose a fitting basis as a set of five Keplerian parameters: orbital period P, time of inferior conjunction T_c, orbital eccentricity, and periastron argument coupled with the orthogonal basis $\sqrt{e}\cos\omega$ and $\sqrt{e}\sin\omega$, and the Doppler-induced radial velocity semi-amplitude K. This form for the eccentricity and periastron argument as Laplace parameters ensures that the model can be compared with a circular one, as the last parameter is undefined for zero-eccentricity orbits (Lucy and Sweeney, 1971; Hatzes, 2019). We imposed Gaussian priors on the orbital period and Uniform prior on the time of inferior conjunction based on the results of the GLS and BGLST/BGLSnT periodograms, with mean the period associated with the maximum power in the periodogram, and the mid-point of the data, respectively, while the Doppler amplitude was set to follow a Uniform distribution based on the observed data modulations. The Laplace parameters were set as free, with an initial assumption of a circular orbit, and constrained to be smaller than 0.99 as default within the RadVel routine.

Accounting for instrumental shifts, we set the mean RV measurements as a free parameter, as well as instrumental noise jitter corrections in the fit for each instrument epoch, before and after the ESO/HARPS' upgrade in mid-2015. The latter is added in quadrature to the RV uncertainties and is associated with stellar modulations (c.f. equation 5.5). Furthermore, considering the presence of an outer companion (that cannot be precisely constrained with the current data), from which only a small portion of its orbit was observed, we set to freely vary in the model a barycentre motion parameter $\dot{\gamma}$, plus a curvature $\ddot{\gamma}$ term accounting for a non-linear trend of the RV model over its entire baseline of approximately 13 years.

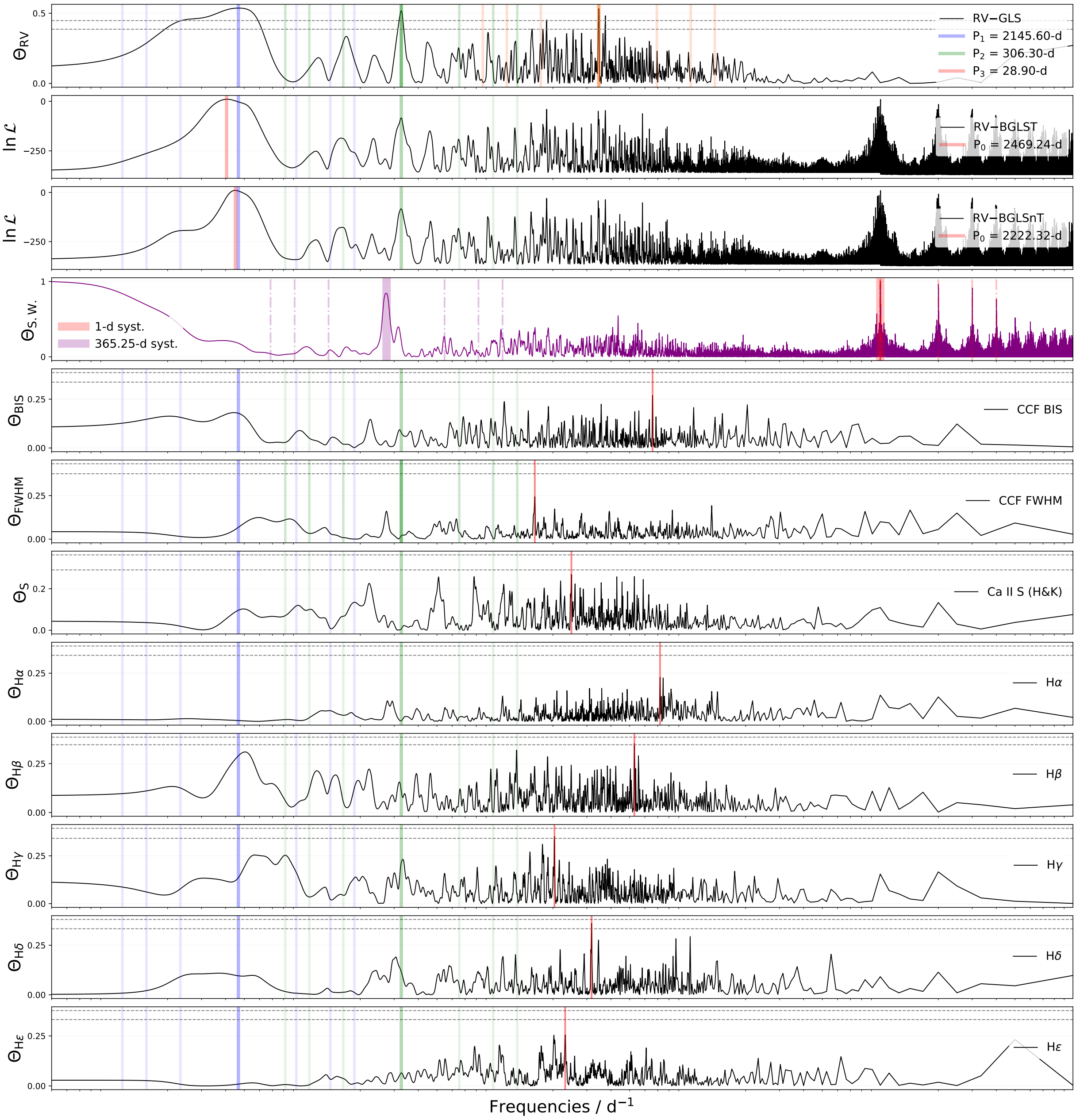

Figure 5.36: Lomb-Scargle periodograms of the radial velocity data and different activity tracers of HIP 104045 (CCF BIS, CCF FHWM, Ca II H&K S–index, and the Balmer lines up to Hϵ). The second and third panels are the BGLST and BGLSnT periodograms (see text). The blue and green vertical line marks the two prominent modulation periods found, along with their lower– ($P_{\rm orb}$/2, $P_{\rm orb}$/3, $P_{\rm orb}$/4), and upper-order ($P_{\rm orb}\times 2$, $P_{\rm orb}\times 3$, $P_{\rm orb}\times 4$) harmonics represented as faint blue and green vertical lines, while the red line represents the periodogram peak for each indicator. The orange vertical line in the GLS periodogram marks a third prominent modulation period in the data (see text), along with its lower- and upper-order harmonics. The dashed horizontal line shows the power level for which the FAPs are less than 1% and 5% (higher and lower, respectively). The fourth panel, in purple, shows the spectral function for HIP 104045 observations, highlighting the one sidereal day (1.0–d, along with lower– and upper–order harmonics), and one sidereal year (365.25–d, along with its lower– and upper-order harmonics) aliases.

The total RV motion of a star due to the total number of companions/planets ($\mathrm{N}_{\rm pl.}$) is given by the sum of the Keplerian orbits (see Section 2.2), and by assuming Gaussianity on the uncertainties distribution, the likelihood is found to be (*viz.* equation (6) in Fulton et al. 2018):

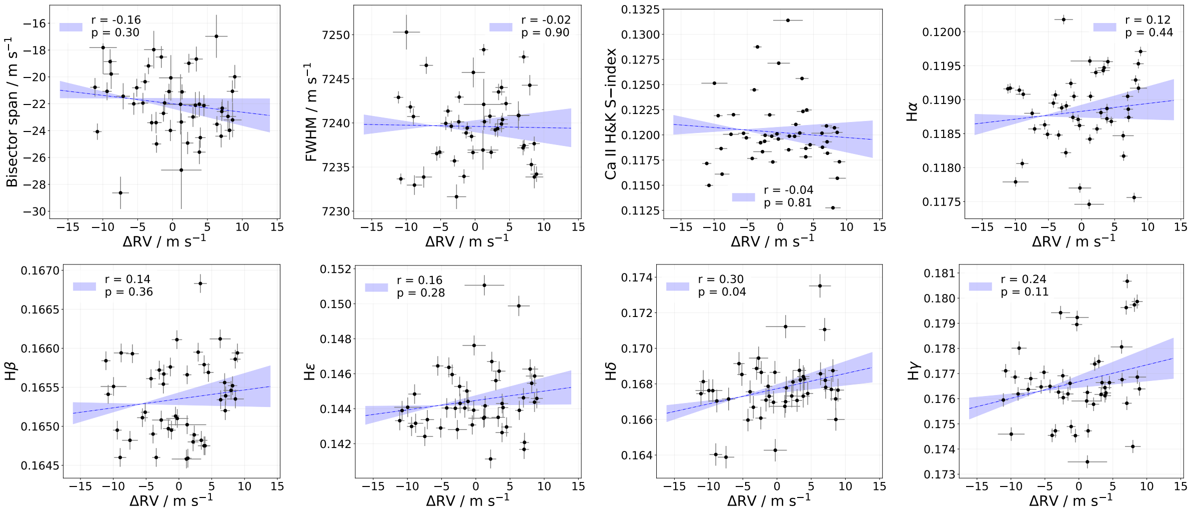

Figure 5.37: RV measurements and the relationship with the several stellar activity indicators for HIP104045. The blue line represents the best-fit linear correlation with 16% and 84% quantiles estimated from MCMC sampling (Salvatier et al., 2015), while in the legend it is indicated the Pearson−r correlation index and the associated p−value.

$$\ln \mathcal{L}_i(\Theta) = -\frac{1}{2}\sum_j \frac{\left(V_{r,j} - y_j\right)^2}{e_j^2 + \sigma_i^2} - \ln\sqrt{2\pi\left(e_j^2 + \sigma_i^2\right)}, \tag{5.5}$$

a function of the radial velocity model $V_{r,j}$ (c.f. Eq. 2.15), each RV measurement y_j, associated uncertainties e_j, and Gaussian white noise σ_i for each HARPS epoch (after and before the upgrade).

Stellar activity relies on the presence of dark spots, plages, granulation, faculæ, and even stellar flares, which are generated through magnetic turbulence in a star's outer convective zone, or in the interface between the radiative and convective envelopes for late-type stars (Hansen et al., 2004; Kippenhahn et al., 2013). Considerations about these undesirable but non-removable sources of noise are important because (a) it provides better estimates of orbital parameters uncertainties, and (b) provide estimates of intrinsic noise present in the data, allowing one to disentangle planetary-like modulations from intrinsic stellar variability. Under a Gaussian Process model, the likelihood function depends on hyper-parameters $\theta \in \Theta$ in the form

$$\ln \mathcal{L}(\Theta) = -\frac{n}{2}\ln(2\pi) - \frac{1}{2}\left[\mathcal{R}\,\Sigma^{-1}\,\mathcal{R}^{\intercal}\right] - \frac{1}{2}\left[\log(\det\Sigma)\right], \tag{5.6}$$

where $\mathcal{R} := y - m$, the residuals vector after the mean function of the data is subtracted, n is the number of data points, and Σ is the covariance matrix (also called *kernel*; see Section 4.2). We used a quasi-periodic covariance kernel to assess the effects of stellar activity on the RV modulations (see eq. 4.29), and following the methodologies outlined in Angus et al. (2018) and Netto et al. (2021), we choose a Jeffrey's prior onto the stellar activity's Doppler-induced semi-amplitude, and Gaussian priors onto the remaining GP hyper-parameters.

The MCMC sampling was implemented following RadVel's genetic routine with 50 walkers over 3000 steps and 8 ensembles. We required burn-out chains with Gelman-Rubin statistics $\hat{R} > 1.01$ and maximum $\hat{R}$ at the same value. The model was minimised with scipy.optimize function employing the Nelder-Mead (N-M) algorithm over 200 iterations (Virtanen et al., 2020). Table 5.20 summarises the N-M minimisation results and orbital solutions. We statistically compared an eccentric and circular model, both accounting for non-zero quantities $\dot{\gamma}$ and $\ddot{\gamma}$, finding out that a circular model with radial velocity slope is favoured by a Bayesian/Akaike Information Criterion (BICc/AICc) variation $\mathrm{BICc/AICc} = -3.17/-3.35$ (nominal $\mathrm{BICc}_{\mathrm{circular}} = 384.81$ and $\mathrm{AICc}_{\mathrm{circular}} = 370.06$, with $\ln Z = -193.49$) over an eccentric model. Therefore, following Liddle (2007), we conclude that a circular model with radial velocity slope but without significant curvature is bonafide in representing the data modulations of planetary nature for HIP 104045 b.

From the retrieved value of $\dot{\gamma}$ (see eq. 5.2), and from the radial velocity baseline of approximately 13.1 years, we estimated the presence of a companion with a minimum mass of $56.67\ M_{\oplus}$ ($\sim 0.1783\ M_{\mathrm{Jup}}$) orbiting HIP 104045, from which only a small portion of its orbit is observable. It is important to highlight that this minimum mass just represents a lower limit estimated, i.e., it assumes an orbital inclination of 90 degrees. The true mass of the companion, if existent, could be higher if the orbit is more elliptical, or if its inclination is lower. Furthermore, other factors such as stellar activity and instrumental errors can affect the accuracy of the RV measurements, and consequently the calculated mass.

HIP104045 presents a quite similarity to the Sun in terms of its chemical abundance pattern as reported in Bedell et al. (2018), which has a lower amount of refractory/volatile elements when compared to solar twins (Meléndez et al., 2009). Interestingly, in comparison to the other planet's hosting stars observed with the STPS programme – HIP11915 (host of a Jupiter twin exoplanet) and HIP68468 (host of a hot Neptune and a hot super-Earth; Meléndez et al. 2017) –

HIP104045 reveals a chemical anomaly that may be correlated with moderate planet engulfment episodes (e.g. Spina et al. 2015; Galarza et al. 2021). The left panel in Fig. 5.38 demonstrates a relative refractory elements enhancement for HIP68468 (with the linear best-fit slope value of $\alpha = 1.52(\pm 0.29) \times 10^{-5}$ dex) and HIP104045 ($\alpha = 1.06(\pm 0.32) \times 10^{-5}$ dex), while a relative depletion for HIP11915 ($\alpha = -1.01(\pm 0.17) \times 10^{-5}$ dex) was estimated. The correlation for HIP101045 is slightly more evident by correcting the abundances with Galactic Chemical Evolution (GCE) effects, although this correction presents minimum effect given the similar ages between HIP104045 and the Sun. HIP104045 sits about midway between the refractory-poor solar pattern and the refractory-rich composition of most solar twins (Meléndez et al., 2009). This moderate enhancement of refractories in HIP104045 means that it may have engulfed some rocky planet material.

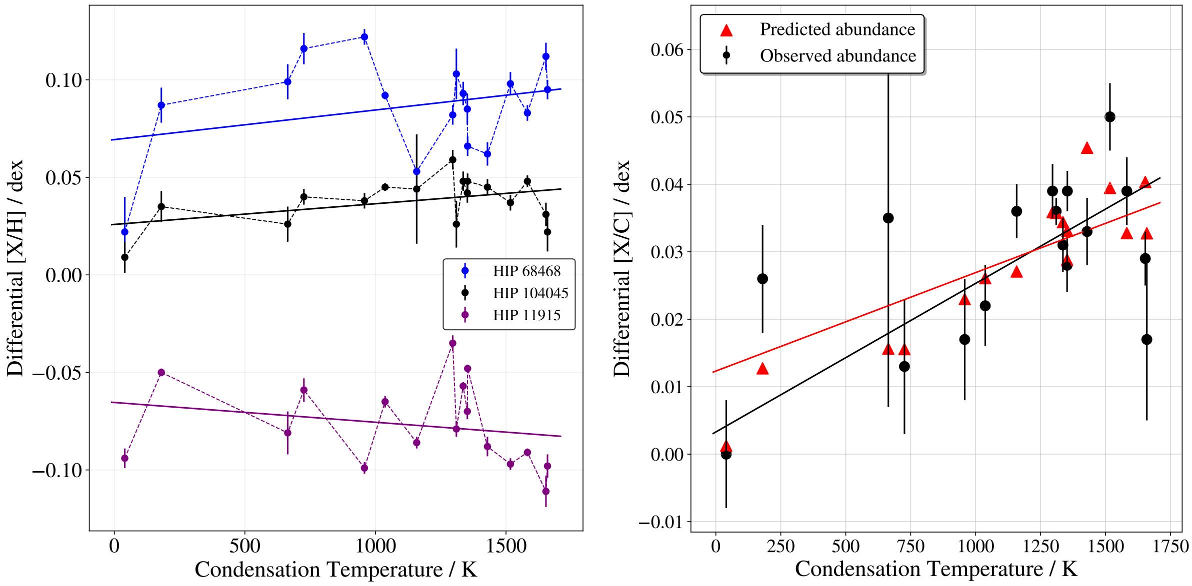

Figure 5.38: Left: Comparison of the differential chemical abundances of HIP68468 (a solar twin that hosts two planets and presents chemical signatures of planet engulfment episodes), HIP104045, and HIP11915 (a solar twin that hosts a Jupiter twin) as a function of the element's condensation temperatures in the protoplanetary disc (Lodders, 2003). The solid lines represent the linear best fit for each star. *Right*: Comparison of the differential abundance of HIP 104045 (black dots plus linear fit) as a function of dust condensation temperature in the protoplanetary disc, and the predicted abundances (red dots plus linear fit) estimated from a planetary accretion of 3.5 M$\oplus$.

Under validation of the planet's accretion scenario, we employed the TERRA code[11] (Galarza et al., 2016). From the star's mass and [Fe/H] abundance, we estimated the convective mass of HIP 104045 as 0.022 $M_\odot$, and the right panel in Fig. 5.38 presents the [X/C] ratios as a function

---

[11] `https://github.com/ramstojh/terra`

of the condensation temperature and the effect due to the accretion of 3.5 $M_\oplus$ of rocky material (a mixture of 2.8 $M_\oplus$ of chondrite-like material plus 0.7 $M_\oplus$ of Earth-like composition material) into the convective zone for HIP 104045.

The phased diagram for HIP 104045 b is present in Fig. 5.39. We adopt as a working definition of a Jupiter analogue around a Sun-like star a definition similar to that presented by Rowan et al. (2016) and Wittenmyer et al. (2016), which is based on a gas giant planet with a dynamical role similar to Jupiter and orbiting a Sun-like star: minimum mass ranging between 0.3 and 3 $M_{\rm Jup}$, orbital semi-major axis between 3 and 7 AU (orbital periods between ∼5 to ∼18 years), and orbital eccentricity $e \leq 0.3$. Based on the refractory composition of HIP 104045, which is situated between that of the refractory-poor Sun and most of the refractory-rich solar twins (Meléndez et al., 2009), we initially postulated that its planetary system might resemble our own Solar System – gas and ice giants populating the outer region ($a > 3$ AU), and the inner populated by rocky planets –. This implied that HIP 104045 could be somewhat lacking in refractory elements compared to most solar twins, yet enriched in refractories relative to the Sun. Considering the aforementioned planet characterisation criteria, HIP 104045 b is a Jupiter analogue, and therefore, HIP 104045 may keep a planetary architecture roughly similar to the Solar System, and perhaps also host inner rocky planets (e.g. Izidoro et al. 2015). Following the methodologies outlined by Kopparapu et al. (2014) and Kane et al. (2016), we estimate the conservative habitable zone boundaries for HIP 104045 as 0.998−1.758 AU, while the optimistic habitable zone boundaries span 0.790−1.855 AU. Moreover, we observed a significant barycentre motion of the system, as $\dot{\gamma} = -0.0007 \pm 0.0015$ m s^{-1} d^{-1} (under reference epoch BJD = 2456855.77), which may be the result of an accelerated motion of the host star/the system barycentre towards Earth, or dramatically, due to dynamical interaction with an outer massive companion ($m \sin i \sim 56.67\ M_\oplus$), albeit further analysis and additional follow-up is required to confirm this claim (e.g. Bouma et al. 2020; Turner et al. 2022).

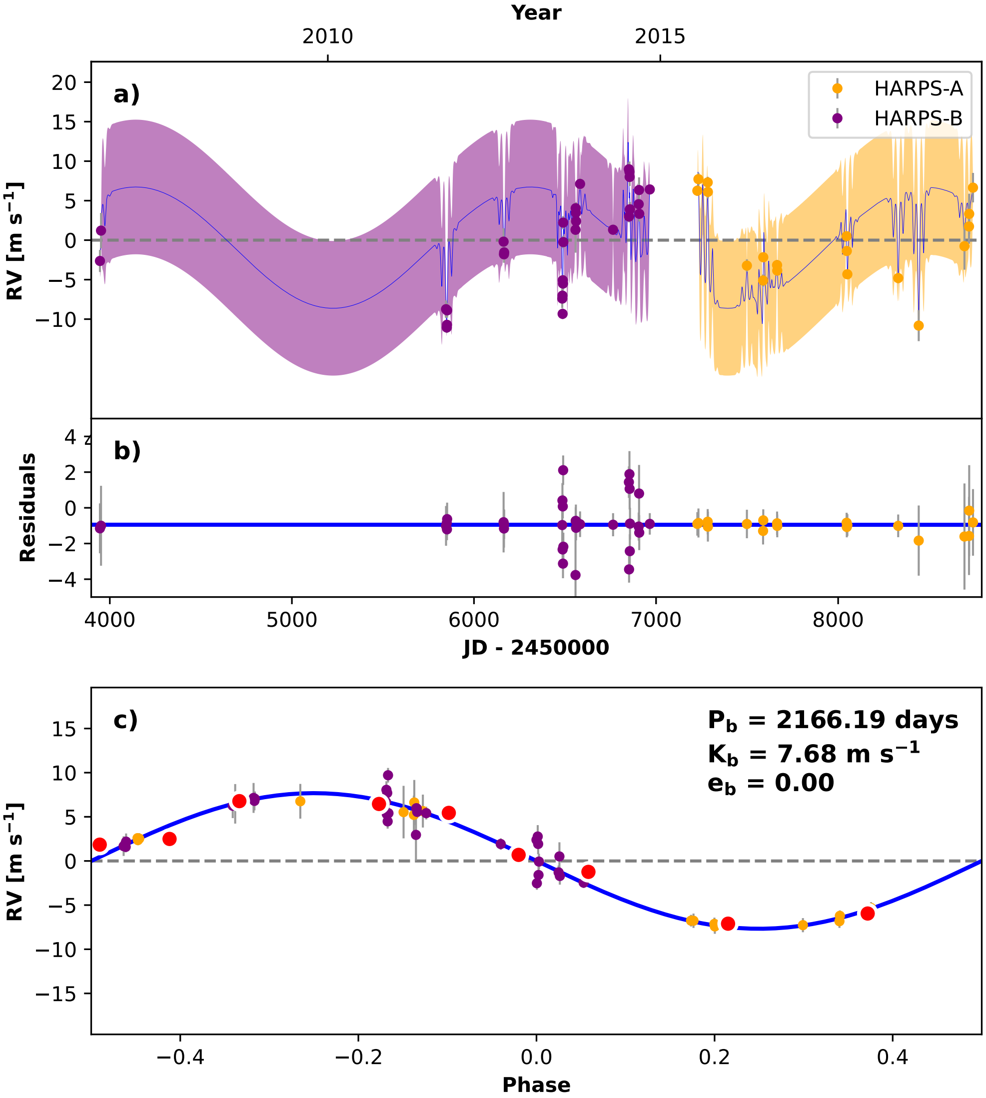

Figure 5.39: Best-fit 1-planet model for HIP 104045. The uncertainties include the RV jitter terms ($\sigma_{\mathrm{HARPS-A/HARPS-B}}$) listed in Table 5.20 added in quadrature with the measured uncertainties for all RVs. In each panel, the purple and yellow dots/curves are the HARPS data before and after the upgrade in mid-2015, respectively. (a) The GP model is plotted as a thin blue line, while the purple and yellow curves are Keplerian models. (b) Residuals to the fit. (c) Radial velocity phase diagrams to the ephemeris of planet b. RVs are binned in 0.08 phase units in red, the phase-folded model for planet b is shown in blue, and N-M Keplerian orbital parameters (planet period, eccentricity held constant and planet-induced Doppler-amplitude) are presented on the upper right corner.

Parameter	Units	Priors	Description	Range
$P_{\mathrm{b,orb}}$	days	$\mathcal{N}$(2145.6, 1000)	Orbital period	–
$T_{\mathrm{b,c}}$	JD	$\mathcal{U}$(2454851, 2458851)	Time of inferior conjunction for planet b	–
K_{b}	m s^{-1}	$\mathcal{N}$(8.0, 1.0)	Planet b Doppler-induced semi-amplitude	–
$[\sqrt{e}\cos\omega]_b$	–	–	Planet b 1st. Laplace parameter	$]-1,1[$
$[\sqrt{e}\sin\omega]_b$	–	–	Planet b 2nd. Laplace parameter	$]-1,1[$
μ_{A}	m s^{-1}	–	Mean radial velocity after HARPS' upgrade in mid-2015.	$]-\infty,\infty[$
μ_{B}	m s^{-1}	–	Mean radial velocity before HARPS' upgrade in mid-2015.	$]-\infty,\infty[$
σ_{A}	m s^{-1}	–	Instrumental noise (added in quadrature) after HARPS' upgrade	$]-\infty,\infty[$
σ_{B}	m s^{-1}	–	Instrumental noise (added in quadrature) before HARPS' upgrade	$]-\infty,\infty[$
$\dot{\gamma}$	m s^{-1} d^{-1}	–	Radial velocity slope	$]-\infty,\infty[$
$\ddot{\gamma}$	m s^{-1} d^{-2}	–	Radial velocity curvature	$]-\infty,\infty[$
η_1	m s^{-1}	$\mathcal{J}$(0.01, 100)	Stellar activity's Doppler-induced semi-amplitude	–
η_2	days	$\mathcal{N}$(30, 5.0)	Time-scale for growth and decay of active regions	–
η_3	days	$\mathcal{N}$(24.6, 5.0)	Recurrence time-scale for active regions	–
η_4	–	$\mathcal{N}$(0.46, 1.4)	Smoothing parameter	–

Table 5.19 - HIP 104045 parameters space probed by the MCMC sampling. $\mathcal{U}(\dots)$ stands for an uniform distribution, $\mathcal{N}(\dots)$ for a Gaussian distribution, and $\mathcal{J}(\dots)$ stands for a Jeffrey's distribution.

Parameter	**Credible Interval**	**N-M minimisation**	**Units**
$P_{\rm b,orb}$	$2484.0^{+402.13}_{-355.68}$	2166.12	days
$T_{\rm b,c}$	$2457524^{+112.60}_{-130.06}$	2456850	days
$K_{\rm b}$	$7.00^{+1.002}_{-0.995}$	6.99565	m s^{-1}
$e_b^\dagger$	$\equiv 0$	–	–
$\dot{\gamma}$	$-0.0007^{+0.0011}_{-0.0020}$	8.837×10^{-8}	m s^{-1} d^{-1}
$\ddot{\gamma}$	$-3.0\times10^{-10}\,{}^{+8.5\times10^{-10}}_{-4.8\times10^{-10}}$	$-$-2.04×10^{-13}	m s^{-1} d^{-2}
$\sigma_{\rm HARPS-A}$	$-0.0002^{+0.0104}_{-0.0106}$	-3.645×10^{-5}	m s^{-1}
$\sigma_{\rm HARPS-B}$	$-1\times10^{-5}\,{}^{+0.0097}_{-0.091}$	-3.644×10^{-5}	m s^{-1}
η_1	$4.4^{+1.37}_{-0.79}$	8.15631	m s^{-1}
η_2	$30.2^{+5.67}_{-5.59}$	31.7801	days
η_3	$21.1^{+3.01}_{-1.18}$	25.8463	days
η_4	$0.464^{+0.01}_{-0.011}$	0.4653	–
$m_b \sin i$	$0.5006^{+0.077}_{-0.075}$	–	$M_{\rm Jup}$
a_b	$3.625^{+0.381}_{-0.354}$	–	AU

Table 5.20 - Keplerian orbital, instrumental shifts/noise and intrinsic stellar activity parameters for HIP 104045 b, along with the derived parameters: planet's minimum mass and orbital semi-major axis. *Note*: † Given $\mathcal{S} := \sqrt{e}\sin\omega$ and $\mathcal{C} := \sqrt{e}\cos\omega$, $e \triangleq \mathcal{S}^2 + \mathcal{C}^2$.

5.2.9 HIP108158 b: A highly eccentric Jupiter and a hidden companion

The ESO/HARPS spectrograph conducted 37 observations of HIP108158 from November 2003 to October 2017. The RUWE of the host star is 1.019, which suggests that the system is unlikely to belong to a sub-stellar binary companion, albeit the strong visual slope seen in the RV data indicates the presence of outer massive companions. The RMS values for the data collected before and after the HARPS upgrade in mid-2015 are 8.755 m s^{-1} and 1.020 m s^{-1}, respectively.

We used the GLS algorithm on the standard grid and obtained a significant signal with a period of approximately 7341 days. Its significance was high under sine-like assumptions, as shown by a BIC change of $||\Delta\mathrm{BIC}|| = 7493$ compared to a non-varying model. FAP tests using bootstrapping re-sampling indicated that the likelihood of obtaining a higher peak than 0.964 (which is associated with the 7343-d period) in the RV periodogram was 1% and 5% for GLS powers of 0.664 and 0.608, respectively. Even though the GLS periodogram appeared noisy due to the non-temporal regularity of data points collected, we ruled out instrumental modulations as the source of the suggested planet's orbital period through a Nyquist frequency sampling test over one sidereal day, one synodic month, one sidereal year, and a Sun's typical activity cycle period. We investigated two activity proxies for HIP108158: CCF BIS and CCF FWHM (see Fig. 5.40). A weak correlation was found between RVs and the first indicator ($r = -0.07$), and albeit a strong correlation is seen for the RVs and CCF FWHM ($r = -0.71$); no timing or phase modulation at the proposed planet's orbital period were observed, which favours the interpretation that, albeit stellar activity may be present and contribute to the RV observations, it may not mimic or masquerade a planetary signal.

We utilised the standard Keplerian basis to model the variations in HIP108158, with orbital eccentricity coupled with the periastron argument in the orthogonal basis, RV slope $\dot{\gamma}$, as well as jitter noise terms for both HARPS epochs $\sigma_{\mathrm{HARPS-A/B}}$ left as free parameters. Table 5.21 presents the model priors and orbital solutions obtained. The MCMC was set with 50 walkers, 10000 steps and 8 ensembles within `RadVel`'s genetic routine. We discarded chains with $\hat{R}$ greater than 1.03 and ensured well-mixed chains by setting default values for the minimum auto-correlation time factor (minAfactor = 40), a maximum relative change in auto-correlation time (maxArchange = 0.03), and a minimum number of independent samples (minTz = 1000). The final values were minAfactor = 20.233, maxArchange = 0.015, and minTz = 702.069, with a maximum $\hat{R}$ of 1.042,

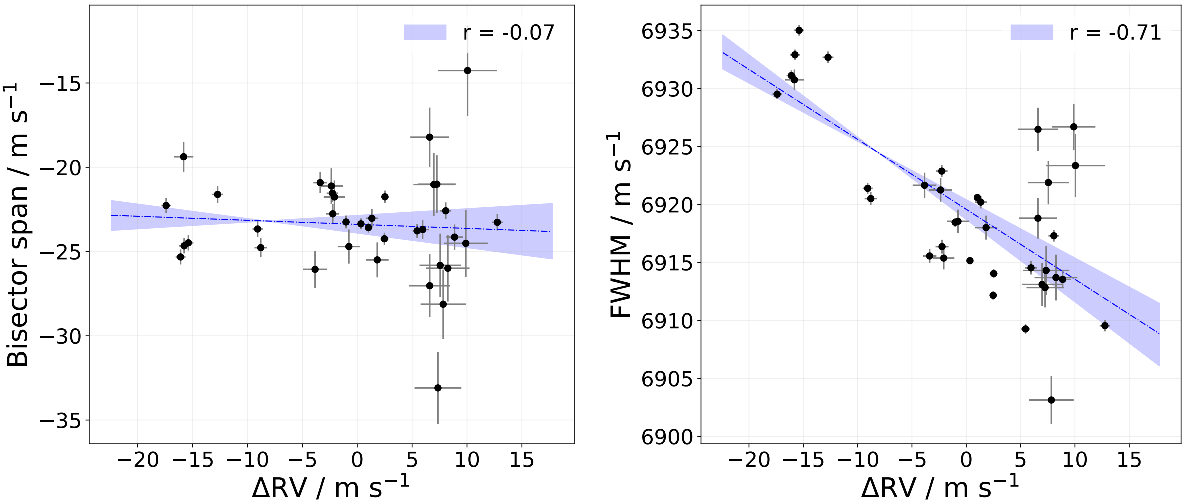

Figure 5.40: RV measurements and their relationship with stellar activity indicators for HIP108158.

indicating a good convergence of the model to the observational data.

Model comparisons were performed by varying or setting as a free parameter the orbital eccentricity and RV slope. A model whose RV slope was not set as a free parameter was ruled out by a $\Delta\mathrm{BIC/AIC} = 12.33/10.62$ in comparison to a $\dot{\gamma}-$free model. In contrast, a model whose eccentricity was not set as a free parameter was ruled out by a $\Delta\mathrm{BIC/AIC} = 12.92/11.30$ in comparison to a $e-$free model. Our analysis concludes that an eccentric model with RV slope, both set as free parameters accurately represents the modulations in HIP108158, indicating the presence of a Jupiter-like companion under a highly eccentric orbit, and the possibility of an outer massive companion that may account for the strongly observed RV slope present in the data (see Fig. 5.41)

Parameter	**Prior**	**Credible Interval**	**Units**
P	$\mathcal{N}(7142, 50)$	7141^{+47}_{-48}	days
T_c	$\mathcal{N}(2452945, 200)$	2452917^{+150}_{-190}	days
K	$\mathcal{N}(10, 2)$	$9.3^{+2.2}_{-2.5}$	m s^{-1}
$\sqrt{e}\cos\omega$	$-$	$0.24^{+0.18}_{-0.19}$	$-$
$\sqrt{e}\sin\omega$	$-$	$-0.34^{+0.10}_{-0.17}$	$-$
$\dot{\gamma}$	$]-\infty, +\infty[$	$-0.0065^{+0.0012}_{-0.0019}$	m s^{-1} d^{-1}
$\ddot{\gamma}$	$]-\infty, +\infty[$	$\equiv 0$	m s^{-1} d^{-2}
$\sigma_{\mathrm{HARPS-A}}$	$]-\infty, +\infty[$	$1.38^{+0.33}_{-0.58}$	m s^{-1}
$\sigma_{\mathrm{HARPS-B}}$	$]-\infty, +\infty[$	$1\pm220^{\dagger}$	m s^{-1}
$m\sin i$	$-$	$0.63^{+0.24}_{-0.27}$	M$_{\mathrm{Jup}}$
a	$-$	7.309 ± 0.033	AU

Table 5.21 - Best-fit 1-planet Keplerian parameters and derived quantities ($m\sin i$ and a), for HIP108158b. *Note*: †With only two measurements, the uncertainties are unreliable.

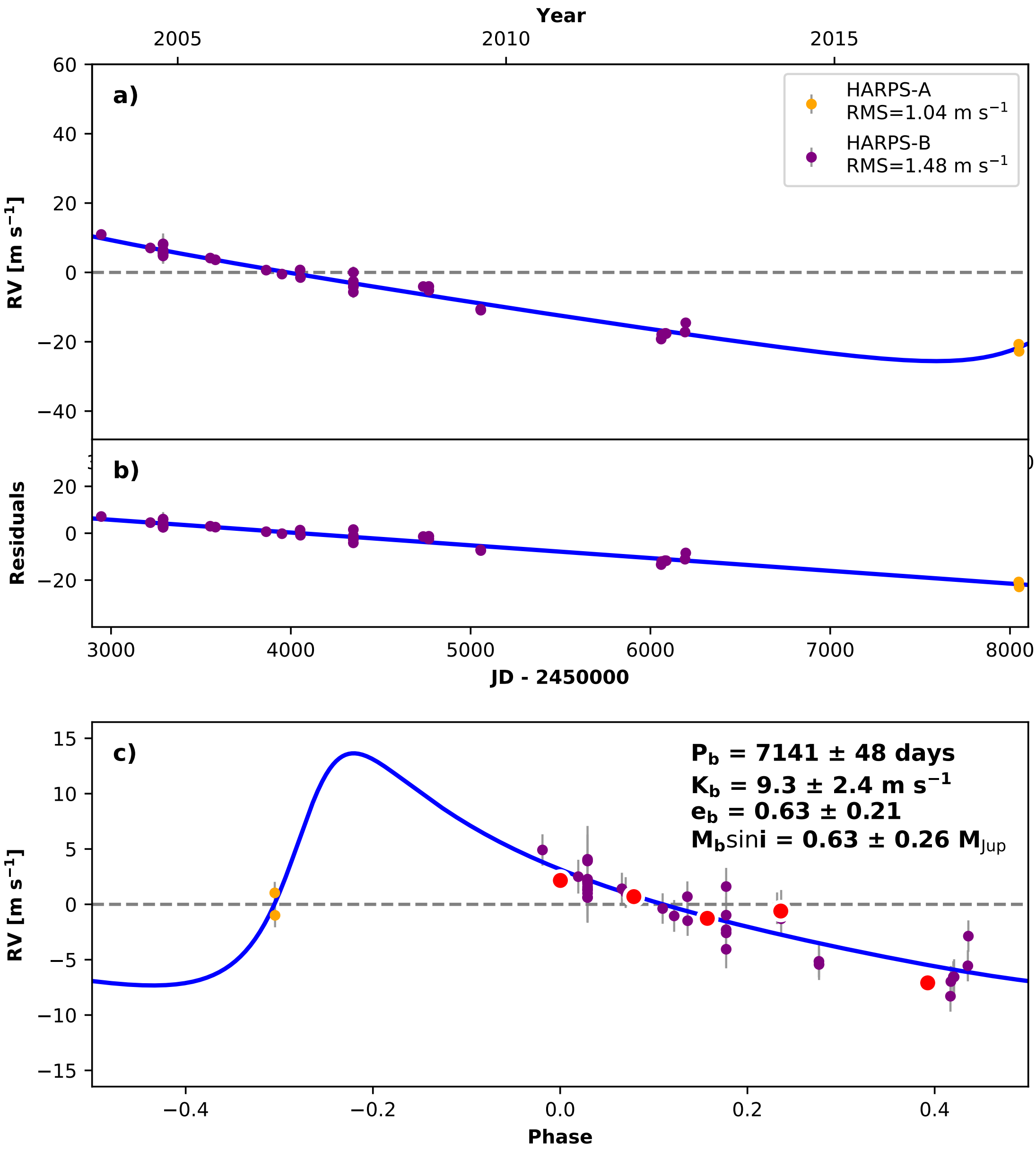

Figure 5.41: Best-fit 1-planet Keplerian orbital model for HIP108158. In all panels, the purple and yellow dots/curves are the HARPS observations before and after its upgrade in mid-2015, respectively. Label in panel (a) indicates the data RMS. (b) Residuals to the best fit 1-planet model. (c) RV phase diagram to the ephemeris of planet b. The RVs are binned in 0.08 phase units as red dots. The phase-folded model for planet b is shown in blue, along with the best-fit Keplerian orbital parameters (period, eccentricity and planet-induced Doppler amplitude), plus the planet's minimum mass.

The high eccentricity ($e = 0.63$) for HIP108158 b is a remarkable feature that, similarly to the case of HIP54582 and HIP87769, warrants further investigation. Previous research has shown that gravitational perturbations from an outer, massive companion can excite an inner planet's orbit and lead to a high eccentricity value (Terquem and Papaloizou, 2002; Luhman et al., 2007). Therefore, one possible explanation for the high eccentricity of planet b is the influence of an outer companion. The data shows a strong RV slope with a value of -0.0065, which further supports the hypothesis of the existence of this outer companion. However, due to the current baseline of almost 14 years, the orbit of the outer companion cannot be fully observed.

Similarly for the cases of HIP54582 and HIP87769, the presence of this outer companion has a significant implication for the formation and the dynamical evolution of HIP108158's planetary system. It is possible that the two planets formed at different distances from the star and migrated to their current positions due to gravitational interactions with the protoplanetary disc (Lin et al., 1996; Barnes et al., 2009), or they may even have formed at similar distances and experienced orbital dynamical instabilities, such as via ZLK mechanism, which led to the ejection of the outer planet and the excitation of the inner planet's eccentricity. Further investigation and modelling are necessary to understand the complex interplay between the two planets and the outer companion in the HIP108158 planetary system and to uncover the underlying mechanisms governing the observed planetary system behaviour.

When a linear trend is observed in the data, a conservative approach is to estimate a minimum mass for a potential companion (such as a massive planet or a brown dwarf) that could produce such a trend for $e \sim 0.5$ and $\omega \sim 90°$. From eq. 5.2, we found a minimum mass of about 1.78 $\mathrm{M_{Jup}}$ (~ 0.0017 $\mathrm{M_{\odot}}$) for the hypothetical planet c. Note that this minimum mass just represents a lower limit, i.e., it assumes an orbital inclination of 90 degrees. The true mass of the companion could be higher if the orbit is more elliptical or the inclination angle is lower. In addition, other factors, such as stellar activity and instrumental errors, can affect the accuracy of the RV measurements, and thus the calculated mass. To confirm or rule out the presence of an outer companion and investigate its role in shaping the HIP108158 planetary system, further observations and follow-up studies will be necessary. One possible avenue for investigation is to obtain more precise RV measurements with longer observational baselines, which would provide a more detailed picture of the system's dynamics and help to constrain the properties of the outer companion. In addition, theoretical studies could be conducted to simulate the dynamical evolution of the HIP108158 system and explore various scenarios for the formation and evolution of its planets. Together, these observational and theoretical efforts could shed light on the complex processes that have shaped the HIP108158 planetary system and provide insights into the broader processes of planet formation and evolution.

Chapter 6

Non-Detection of Exoplanets around Solar Twin Stars

In this chapter, we present a summary of the non-detection of exoplanets around solar twin stars, also part of the Solar Twin Planet Search programme. Similar to the stars presented in the previous chapter, the following were observed by the ESO/HARPS spectrograph through the last decade and, in some cases, there are simply an insufficient number of RVs for analysis, while in other cases a hypothetical planet's orbital period was found at the same frequency into one, or more, activity indicator or instrumental feature, which ruled out the interpretation of the signal as of planetary nature. The RV curves are shown in Figures 6.1 to 6.11.

6.1 Young Main-Sequence Stars, Binary Systems, and Variable Stars

HIP3203. The system is classified as a Young Stellar Object (YSO; Torres et al. 2006; Spina et al. 2018), and a prominent modulation period of 14.7 days was found via Lomb-Scargle periodogram analysis. We observed a strong anti-correlation between the RV measurements and the CCF BIS ($r = -0.82$), which likely indicates that the host star is active, and the RV modulations are sourced from stellar activity. With current data, it was not possible to infer the existence of an orbiting planet within the system.

HIP4909. The system is classified as a young main-sequence star (0.6 Gyr; Spina et al. 2018). There is a prominent modulation period of 11.7 days observed via implementation of the Lomb-Scargle periodogram method, and we observe a moderate anti-correlation between the RVs and CCF BIS ($r = -0.63$), which indicates that the modulations are likely to be from stellar magnetic activity. With current data, it is not also possible to infer the existence of an orbiting planet within the system.

HIP9349. The system is also classified as a young main-sequence star based on its age of 0.6 Gyr estimated by Spina et al. (2018). We observed a moderate correlation between the RVs and CCF FWHM ($r = 0.61$), which indicates that the modulations observed in the star are likely to be sourced from stellar activity.

HIP10303. This star is part of a wide binary system named 66 Cet (Abt, 1988). A prominent modulation period of 13.4 days was found in the RV time series with the Lomb-Scargle method, albeit the RV modulations at this period do not indicate any sinusoidal shape typically associated with the presence of orbiting planets, or stellar magnetic activity.

HIP22263. The system is classified as a BY Dra variable named 58 Eri (Samus' et al., 2017). There is solely a strong RV-CCF FWHM correlation ($r = 0.77$), which may indicate that the RV variations are associated with stellar activity, befitting with that stellar type. The prominent period found with the Lomb-Scargle method implementation is 34.40 days.

HIP29525. It is classified as an RS CVn Variable star, and a young main-sequence star (0.8 Gyr; Spina et al. 2018). RV observations present an interesting modulation at 22.1 days, however, there is a strong anti-correlation between these RVs and BIS (r = -0.8), which cast doubt about these modulations being of planetary nature. Despite this cut-off, we could indeed fit a Keplerian model to the data, finding out a best-fit model at $P = 21.91 \pm 0.12$ days with orbital eccentricity $e = 0.363 \pm 0.091$. Follow-up of this system is encouraged.

HIP30037. It is classified as a spectroscopic binary system, and there was already a known sub-stellar companion reported as part of the Solar Twin Planet Search observing programme (dos Santos et al., 2017). Complementarily, we observe a strong anti-correlation between RVs and CCF FWHM ($r = -0.78$).

HIP36515. This object is classified as a young main-sequence star (0.5 Gyr; Spina et al. 2018, and the prominent period found with the Lomb-Scargle is relatively short (6.1 d) to claim any nature for the RV modulations, rather it is from intrinsic stellar activity, instrumental noise, or planetary-like. Under this modulation period, CCF BIS present slightly sine-like variations.

HIP42333. The star is classified as a BY Draconis Variable star (V401 Hya; Samus' et al. 2017), and also as a young main-sequence star (1 Gyr; Spina et al. 2018). Both RV and FWHM measurements present sinusoidal variations at the prominent Lomb-Scargle period found at 19.30

days, and the stellar activity indicator proxies CCF BIS and CCF FWHM correlate moderately and significantly with the RV measurements, respectively, $r = -0.53$ ($p = 0.0005$) and $r = 0.69$ ($p = 1.75 \times 10^{-6}$). Interestingly, the RV Lomb-Scargle period corresponds to an inferior second-order harmonic of CCF BIS, and a superior first-order harmonic of CCF FWHM, which likely indicates that stellar magnetic activity plays a role in the RV modulations.

HIP62039. The host star is part of a wide binary system (dos Santos et al., 2017), with a companion of $m \sin i{>}19$ $\mathrm{M_{Jup}}$. We observe a moderate correlation between RVs and activity indicator proxies: CCF BIS ($r = 0.34$), CCF FWHM ($r = 0.61$), and $S-$index ($r = 0.31$). The Keplerian best-fit model resulted in a relatively eccentric RV modulation of planetary nature, however, it is important to emphasise the high RV RMS before HARPS' upgrade in mid-2015 (3.27 m s^{-1}), and none of the stellar activity indicator proxies modulates sinusoidally at the proposed planet's orbital period. The existence of an inner planet in this system is dubious in our analysis.

HIP101905. This is a young main-sequence star (1.2 Gyr; Spina et al. 2018). We observe that there is a significant and strong correlation between FWHM and RV measurements ($r = 0.69$, $p = 9.63 \times 10^{-9}$), which indicates stellar activity (following the system's age). There is no correlation at all between RVs and CCF BIS ($r = 0$). Moreover, we observed a certain scatter in the RVs observations (again, following the system's age). Therefore, it is not possible to infer the presence of an orbiting planet in such a system with current data.

HIP102040. This is also a young main-sequence star (2.4 Gyr; Spina et al. 2018). Albeit a prominent period of 43.10 days was detected with a Lomb-Scargle implementation, the RV curve does not present any characteristic modulation that could be attributed to the presence of an orbiting planet and/or stellar magnetic activity.

HIP105184. This is a young main-sequence star (0.8 Gyr; Spina et al. 2018). We note that the resulting Lomb-Scargle modulation period found is an inferior fourth-order harmonic of the prominent period found in the CCF BIS periodogram, which likely indicates that the RV modulations are the result of stellar activity, per the stellar age.

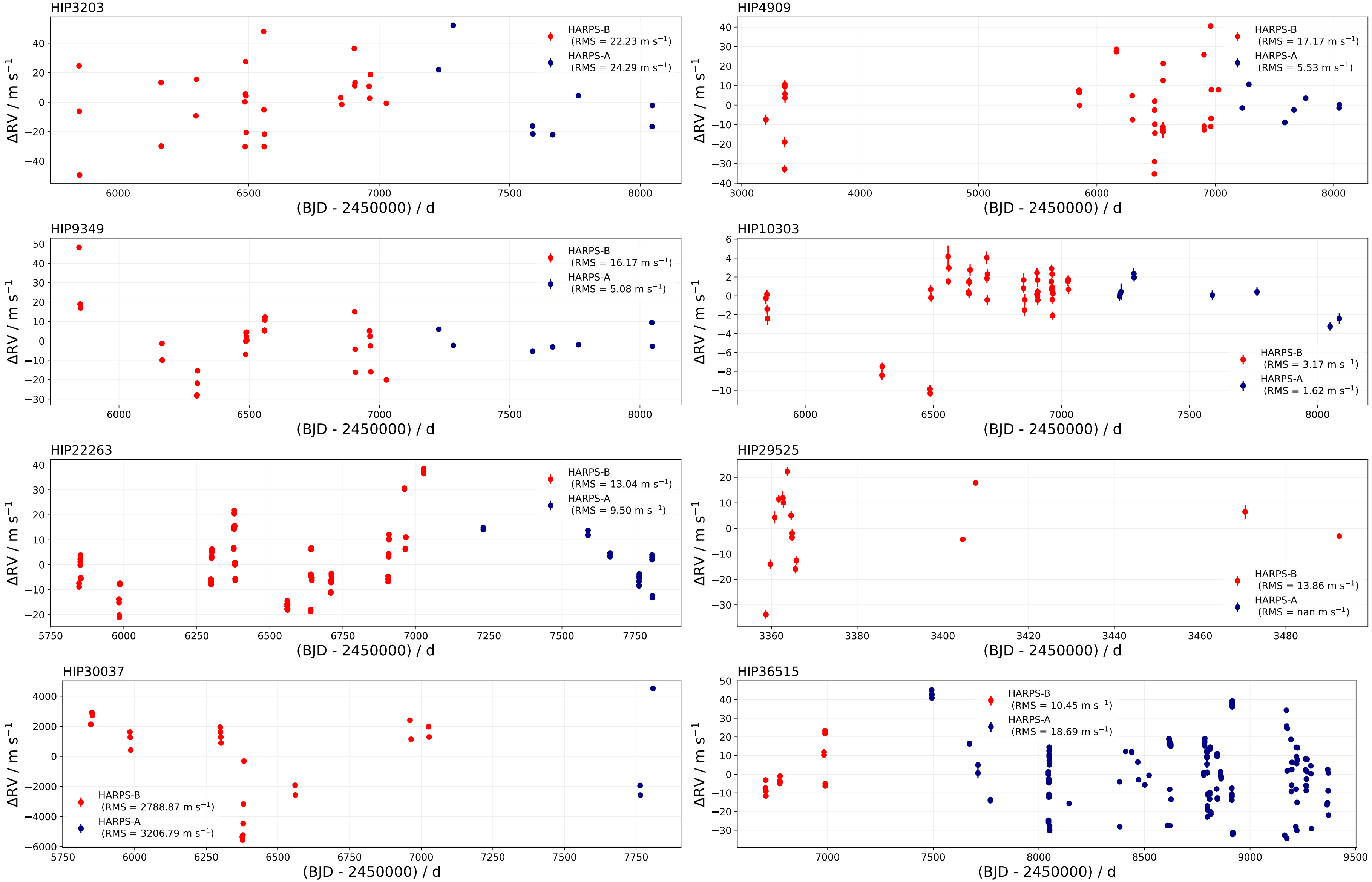

Figure 6.1: RV curves for HIP3203, HIP4909, HIP9349, HIP10303, HIP22263, HIP29525, HIP30037, and HIP36515. Red and blue dots represent the data taken before and after the ESO/HARPS' upgrade in mid-2015, respectively (if existent), and the RMS is also indicated in the legend.

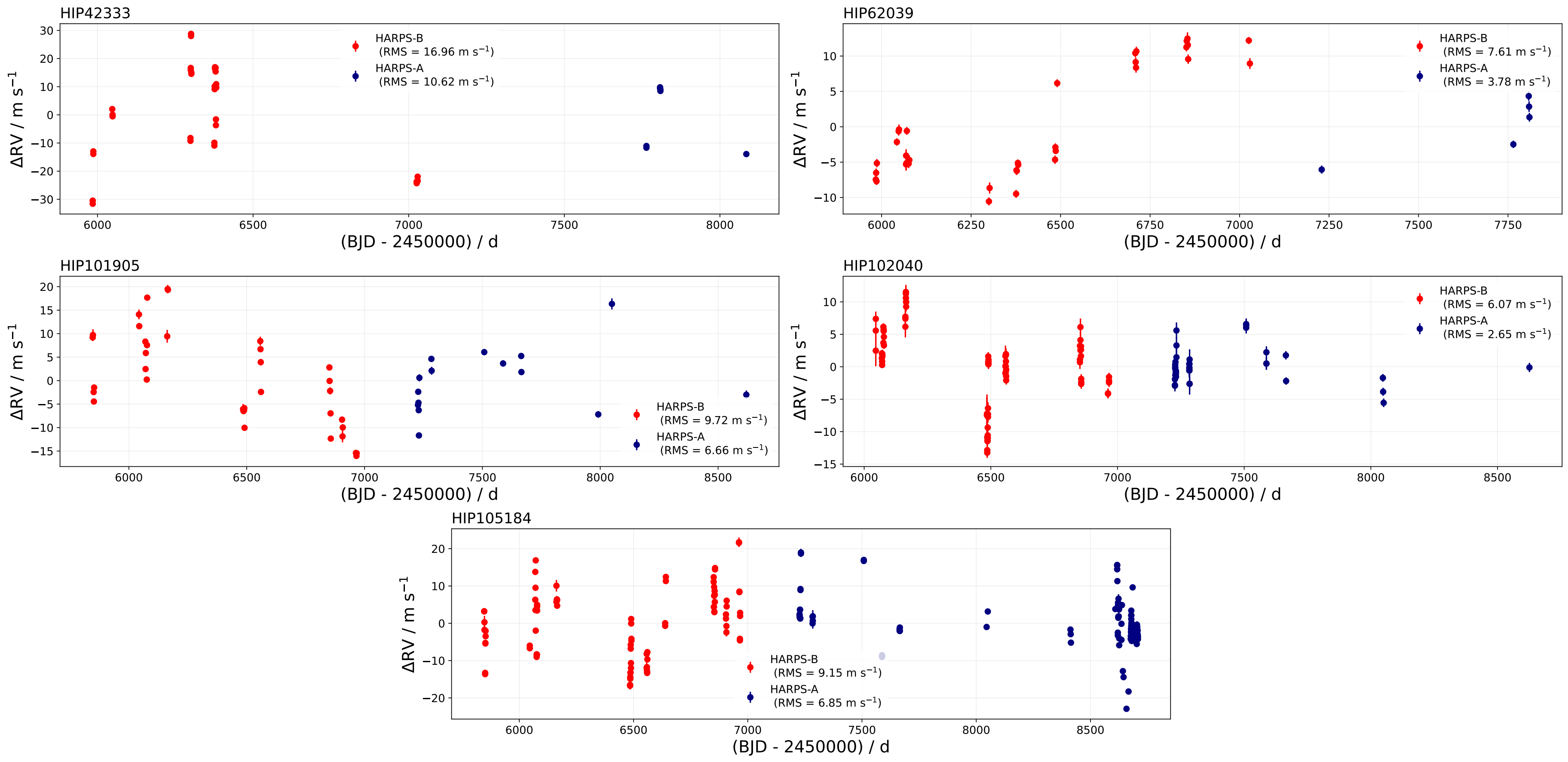

Figure 6.2: RV curves for HIP42333, HIP62039, HIP101905, HIP102040, and HIP105184. Red and blue dots represent the data taken before and after the ESO/HARPS' upgrade in mid-2015, respectively if existent, and the RMS is also indicated in the legend.

6.2 Radial Velocity Trends due to Massive Companions

HIP6407. There are only seven RV observations of the system in the ESO/HARPS ADP. With that few number of points, it is impossible to provide a concise and robust analysis of the system, however, its follow-up is strongly encouraged since it presents a strong visual slope, which may indicate that the system is part of a yet-unknown/unreported spectroscopic binary, or that is orbited by a massive planet-like object.

HIP19911. It is notable a strong RV slope, however, with few number of eight observations, we cannot conclude a robust analysis. A follow-up of the system is strongly required and encouraged, since this RV slope may indicate that the system is part of a spectroscopic binary (dos Santos et al., 2017).

HIP54102. We observe a strong trend in the RV data that would be consistent with a long-period massive companion given the Doppler-induced semi-amplitude of $K \sim 100$ m s^{-1}, however, no clear periodic signal was detected with the Lomb-Scargle method.

HIP54582. With the implementation of the Lomb-Scargle periodogram method, we obtained a prominent modulation period of 3977 days consistent with a Jupiter analogue ($m \sin i = 1.21$ M$_{\rm Jup}$) under nearly eccentric orbit $e = 0.17 \pm 0.13$. In addition, we observed a strong linear trend in the data that, in principle, is compatible with the presence of an orbiting outer massive ($m \sim 3.51$ M$_{\rm Jup}$) under eccentric orbit $e = 0.5$ from the value of the resulting RV slope ($\dot{\gamma} = 0.0325$ m s^{-1} d^{-1}). We point out that the absence of an inflexion point makes it difficult to pinpoint the existence of both orbiting objects with certainty, and therefore, follow-up of this system, either to confirm or refute their existence, is encouraged.

HIP65708. We observe a strong linear trend in the data, which could be the result of a planet, or a sub-stellar companion, in distant orbit ($P = 406.90$ days). It should be noted that there is a strong and statistically significant anti-correlation between the RV measurements and CCF FWHM ($r = -0.93$; $p = 8.79 \times 10^{-5}$), whilst no correlation at all is observed for CCF BIS ($r = 0.02$; $p = 0.94$). This correlation cast doubt on the interpretation of the data modulations as of planetary nature.

HIP73241. There is a relatively strong linear trend in the data that could be the result of an unknown companion, however, no Keplerian model was capable to explain with current data RV modulations, largely due to the lack of an inflexion point in the data (a point where the RVs reach the minima or maxima). Follow-up of this system is encouraged.

HIP79578. There is a strong linear trend in the RV data, as well as high RMS for measurements taken before and after the ESO/HARPS in mid-2015. Without an inflexion point, it is not possible to infer the nature of the modulations, however, seen from the high Doppler-induced amplitude (~ 400 m s^{-1}) and the curve shape, we estimate the existence of a massive object under eccentric orbit around HIP79578. Follow-up of this system is encouraged.

HIP83276. We observe a strong linear trend in the data, which may account for the massive companion since the system is classified as a spectroscopic binary (Pourbaix et al., 2004), however, there are just seven RV observations recorded at the ESO/HARPS ADP. It should be only treated as a hypothesis that it likely seems that the final RV points are *flat*, which could represent observations at one of the inflexion points, and if follow-up were performed, the RV amplitude could be less than these observed. Follow-up of the system is strongly recommended.

HIP97358. We observe a strong linear trend in the data, however, there are only six RV measurements recorded at the ESO/HARPS ADP.

HIP109110. The RV data presents a relatively strong linear trend, however, just five RV observations are recorded in the ESO/HARPS ADP. This object is also classified as a BY Dra Variable (Samus' et al., 2017).

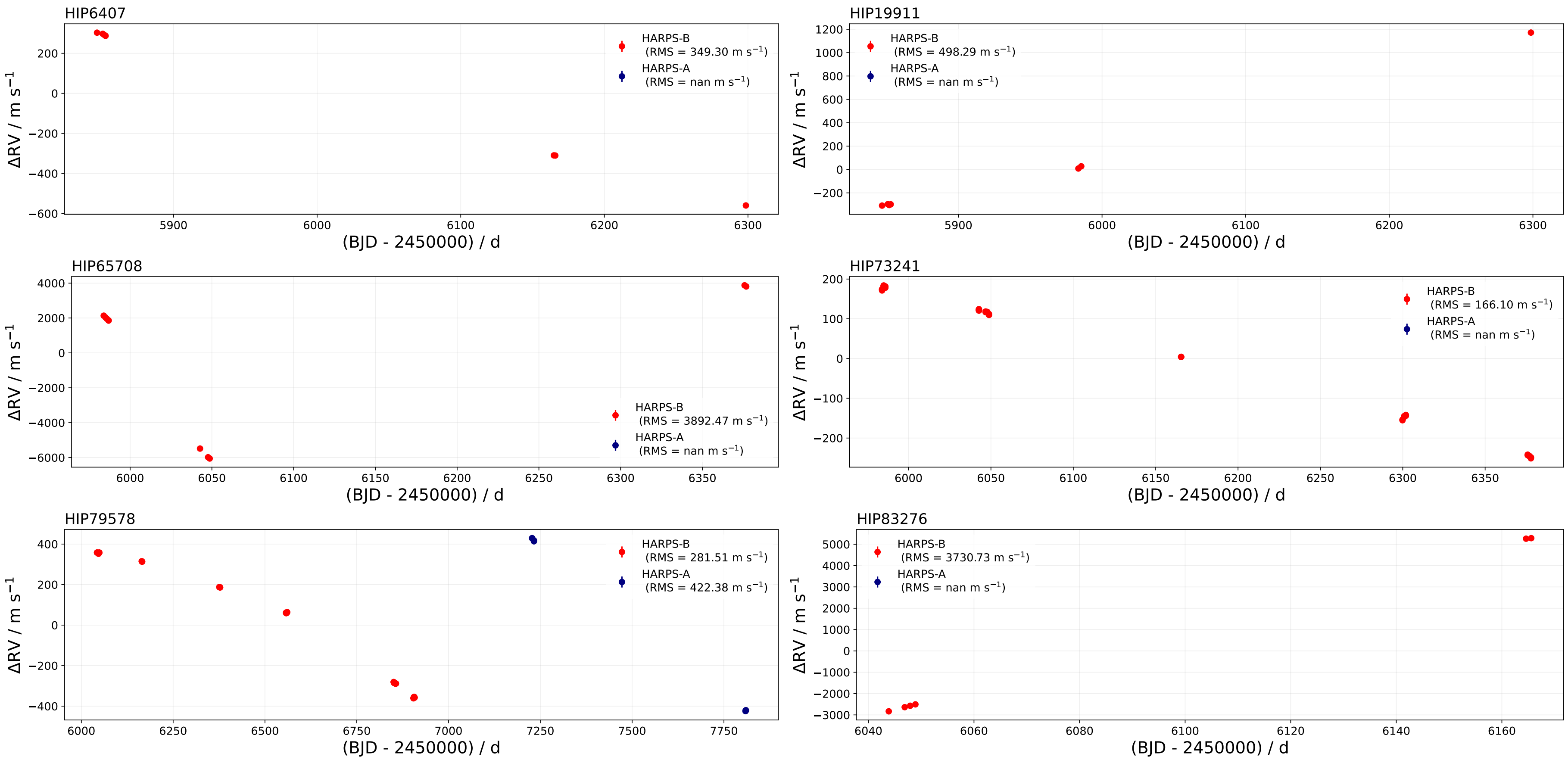

Figure 6.3: RV curves for HIP6407, HIP19911, HIP65708, HIP73241, HIP79578, and HIP83276. Red and blue dots represent the data taken before and after the ESO/HARPS' upgrade in mid-2015, respectively (if existent), and the RMS is also indicated in the legend.

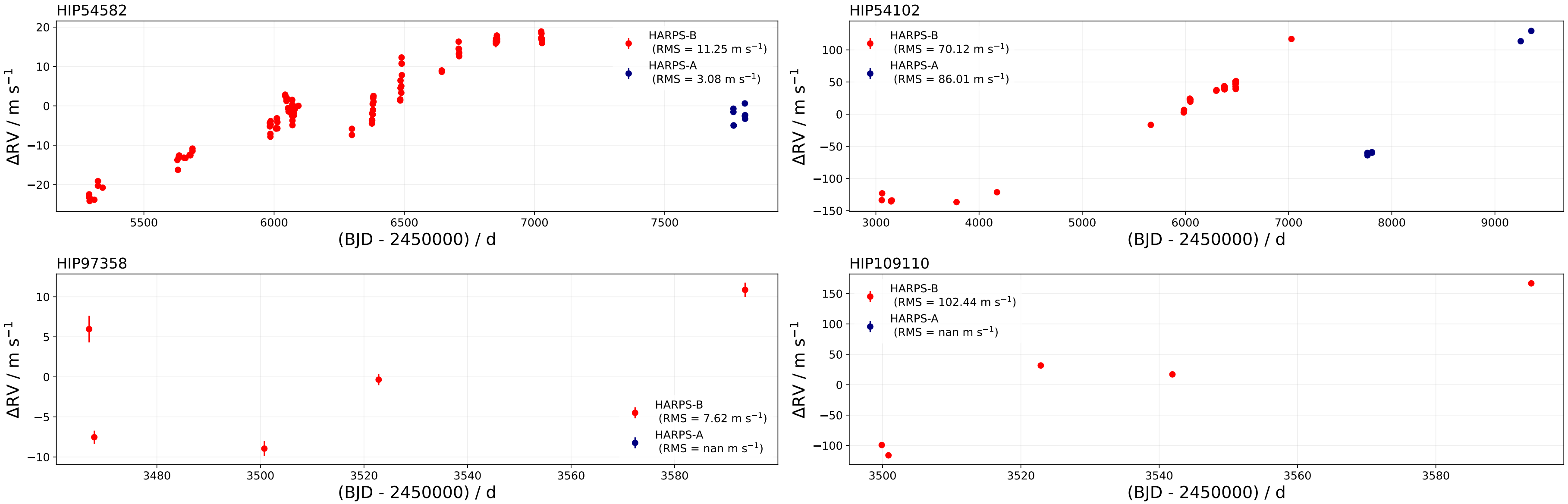

Figure 6.4: RV curves for HIP54102, HIP54582, HIP97358, and HIP109110. Red and blue dots represent the data taken before and after the ESO/HARPS' upgrade in mid-2015, respectively (if existent), and the RMS is also indicated in the legend.

6.3 Uninformative Radial Velocities Modulations

HIP8507. There are no clear modulations in the data that indicate whether they are from stellar magnetic activity or any orbiting exoplanet. A prominent period of 20.80 days was found with the implementation of the Lomb-Scargle algorithm, and none of the two activity indicators of the host star, CCF BIS and CCF FWHM, modulates coherently at the RV period found.

HIP15527. Both data before and after the ESO/HARPS upgrade in mid-2015 present high RMS: 34.49 m s^{-1} and 35.94 m s^{-1}, respectively. An attempt to model a Keplerian to the data with RadVel indicated the presence of a 1.0 $M_{\rm Jup}$ companion at $P = 599.39$ days, and orbital eccentricity of $e = 0.98$, yet, the MCMC simulations were not satisfactory, as the chains did not pass convergence tests (with $\hat{R} = 1.743$; see Subsection 4.2.1). Especially to obtain RV measurements in the inflexion points, follow-up of this object is strongly encouraged.

HIP18844. A Lomb-Scargle method implementation revealed a prominent modulation period at $P = 3613.5$ days, and given the strong RV slope observed in the data, we plausibly attribute the existence of an outer massive companion orbiting HIP18844. The ESO/HARPS ADP has recorded 17 RV observations of the host star over around 2000 days. We note that there is a gap of 1500 days between the last two sets of observations, and the first 500 days have three sets of observations that are relatively spread out. The most recent observations show a smaller RV variation (ΔRV) than those in the third set of observations. Nevertheless, due to the gap in the data, it is impossible to determine if the third set of observations was taken near or at an RV inflexion point. Therefore, we cannot determine the nature of the modulations with the current data.

HIP28066. Whilst indeed we could fit both a Keplerian and Gaussian Process regression, the existence of an orbiting planet to HIP28066 is dubious. We observed a weak RV-CCF FWHM correlation ($r = 0.26$), and no RV-CCF BIS correlation ($r = 0.01$). The hypothetical planet's orbital period found is 39.9 days, and similar sine-like modulations are observed in the CCF FWHM time series at this period, which may indicate that 39.9 days is a stellar activity modulation. There is no sufficient statistical evidence to claim that the RV modulations are planetary.

HIP29432. We detected a modulation period of 79.3 days in the RV time series with the Lomb-Scargle algorithm, however, there is no clear sign that the modulations are from the presence of an orbiting planet, or that it could be attributed to intrinsic stellar activity. Complementarily, we observe moderate correlations between the RV measurements and CCF BIS ($r = 0.50$) and CCF FWHM ($r = 0.40$), both with high statistical significance under $p-$value hypothesis testing.

HIP30476. With current data, we cannot infer the nature of the RV modulations detected at the prominent period of $P = 61.50$ days since there is no clear sign of sinusoidal modulations that are typically attributed to orbiting exoplanets or intrinsic stellar magnetic activity.

HIP33094. The star is relatively old (8.9 Gyr; Spina et al. 2018), and besides the moderate correlation between the RVs and CCF BIS ($r = 0.38$), at the prominent modulation period found with a Lomb-Scargle implementation ($P = 171.70$ days), both RVs and CCF FWHM modulate equally. With current data, it is not possible to infer the presence of a planet, nor claim that the RV modulations are intrinsically stellar.

HIP34511. Despite there is no correlation between the RVs and CCF BIS/FWHM, the joint Keplerian and Gaussian Process regression best-fit model was statistically insignificant in representing the data modulations.

HIP36512. The ESO/HARPS RV observations are *flat*, i.e., there is no clear sign of sinusoidal variations at the prominent period of $P = 39.4$ days discovered with a Lomb-Scargle algorithm implementation.

HIP38072. There is no clear sign of periodicity detected with the Lomb-Scargle method.

HIP40133. With current data, we cannot infer the nature of the RV variations at the prominent modulation period of $P = 38.6$ days detected with the Lomb-Scargle method. Additionally, we did not observe strong correlations between the RVs and the activity indicators CCF BIS ($r = 0.10$) and CCF FWHM ($r = 0.21$).

HIP44713. The ESO/HARPS RV observations are *flat*, and there is too much scatter. At the prominent Lomb-Scargle period found ($P = 87.90$ days), CCF BIS modulates sinusoidally. Furthermore, we observe moderate correlations between the RVs and both CCF BIS ($r = 0.57$) and CCF FWHM ($r = 0.34$), therefore, there is no sufficient and statistically significant evidence to claim that the RV modulations could be planetary.

HIP54287. The ESO/HARPS RV observations are flat. Furthermore, the system is part of a spectroscopic binary system (Kervella et al., 2019). We generally avoid studying spectroscopic binary systems when looking for exoplanets, especially making use of RV data, because in a spectroscopic binary system, two stars orbit around a common centre of mass, and their motion causes periodic Doppler shifts in the observed spectrum. These shifts can be mistaken for the signal of an exoplanet, even hampering it, leading to false detections or inaccurate measurements of the exoplanet's parameters.

HIP55409. There are just three RV observations recorded at the ESO/HARPS ADP.

HIP67620. This is a spectroscopic binary system (dos Santos et al., 2017) probably with an M-dwarf companion (Pacheco et al., 2019). An inflexion point in the data is nonexistent, making it impossible to fit a Keplerian model to the data.

HIP69645. There is too much scatter in the RV observations, and no clear sinusoidal variations at the prominent period of 26.7 days were detected with a Lomb-Scargle periodogram implementation. We observe a moderate RV-CCF BIS correlation ($r = 0.45$), and therefore, the RV modulation nature of this object is uncertain; it could be the result of intrinsic stellar activity, and the modulation period found could be associated with the star's rotation period given its age estimated at 5.3 Gyr (Spina et al., 2018).

HIP77052. The ESO/HARPS RV observations are divided into several sets, with too much scatter among the observations in each set. The data RMS before HARPS' upgrade is 5.48 m s^{-1}, and for the data collected after the upgrade, $\mathrm{RMS} = 3.65$ m s^{-1}. We observe some strong peaks in the CCF FWHM and CCF BIS activity indicators periodograms, but these peaks are relatively close to the proposed planet's orbital period found with Lomb-Scargle for the RVs. The best-fit Keplerian model does not seem to be so *best-fit* due to possible active regions of the star that may account for the observed RVs modulations.

HIP77883. The RVs are flat and the nature of the modulations at the prominent period of 50.9 days detected with the Lomb-Scargle is uncertain given the lack of coherent sine-like variations.

HIP79715. The RVs are flat and the nature of the modulations at the prominent period of 17.20 days detected with the Lomb-Scargle is uncertain given the lack of coherent sine-like variations.

HIP81746. Albeit 15 RV observations are recorded at the ESO/HARPS ADP for this system, the nature of the modulations is uncertain and inconsistent.

HIP85042. The ESO/HARPS RV observations are divided into several sets, with too much scatter among the observations in each set similar to the case of HIP77052. With current data, we cannot infer the nature of the RV modulations, being planetary, instrumental, or intrinsic stellar activity.

HIP89650. The RVs are flat and the nature of the modulations at the prominent period of 50.8 days detected with the Lomb-Scargle is uncertain given the lack of coherent sine-like variations.

HIP95962. The RVs are flat and the nature of the modulations at the prominent period of 39.40 days detected with the Lomb-Scargle is uncertain given the lack of coherent sine-like variations.

HIP103983. The ESO/HARPS ADP records only eight RV observations for this system, and its spectrum is contaminated by bright companions (see dos Santos et al. 2017).

HIP109821. The RVs are flat and the nature of the modulations at the prominent period of 4522.30 days detected with the Lomb-Scargle is uncertain given the lack of coherent sine-like variations.

HIP115577. The RVs are flat and the nature of the modulations at the prominent period of 1537 days detected with the Lomb-Scargle is uncertain given the lack of coherent sine-like variations.

HIP118115. There is no statistically significant correlation between the RV measurements and the two intrinsic stellar activity indicator proxies (CCF BIS and CCF FWHM), nor timing variations at the prominent Lomb-Scargle period ($P = 115.40d$). There is too much scatter in the data, so the presence of a planet in such a system is dubious.

HIP19855. The ESO/HARPS ADP record only 5 RV observations of this system.

HIP29568. The HARPS data are flat, and the nature of the modulations at the prominent period of 10.20 days detected with the Lomb-Scargle is uncertain given the lack of coherent sine-like variations.

HIP44890. HARPS recorded only 8 RV observations of this system.

HIP91287. The RVs are flat and the nature of the modulations at the prominent period of 2922.30 days detected with the Lomb-Scargle is uncertain given the lack of coherent sine-like variations.

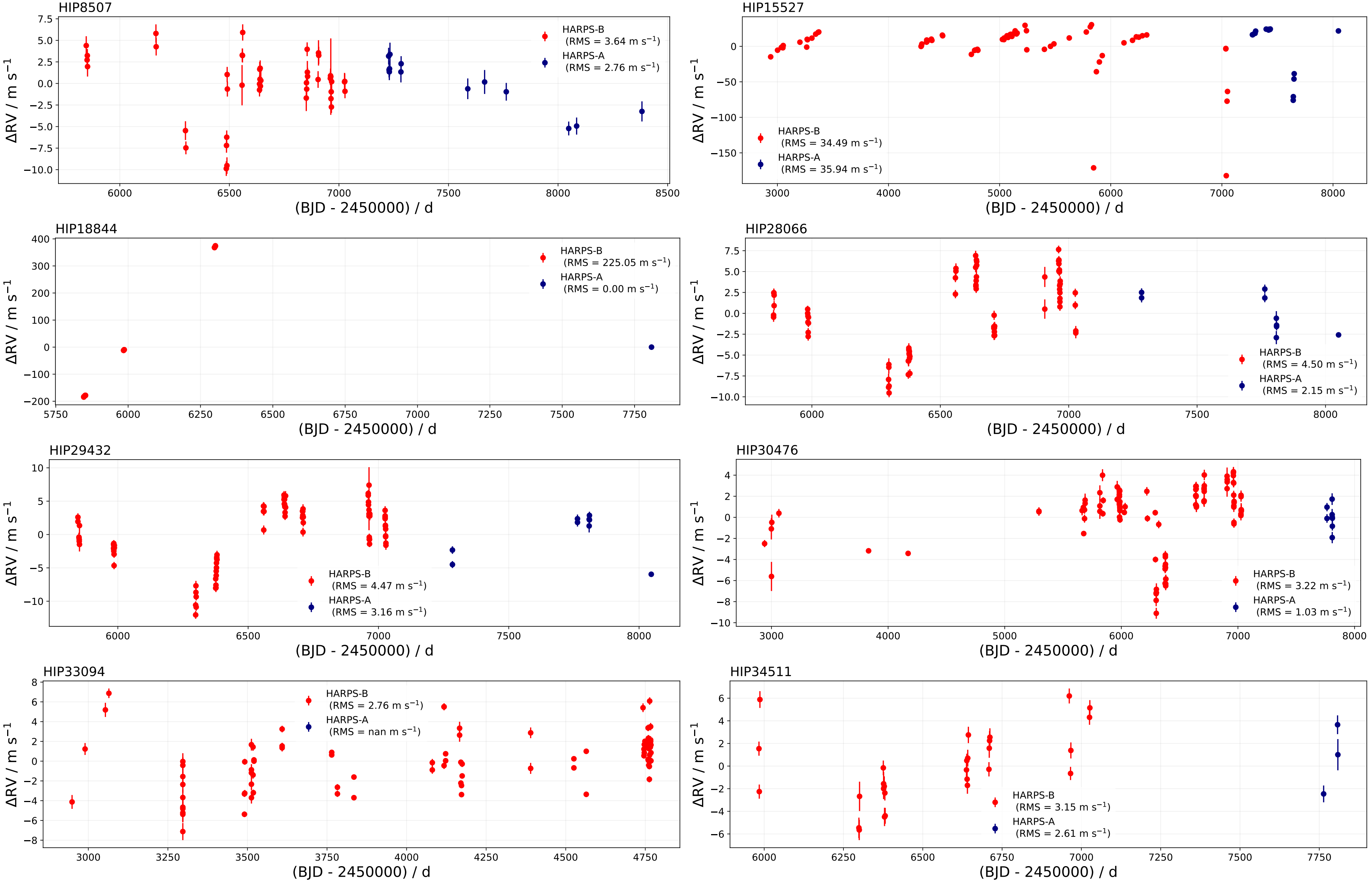

Figure 6.5: RV curves for HIP8507, HIP15527, HIP18844, HIP28066, HIP29432, HIP30476, HIP33094, and HIP34511. Red and blue dots represent the data taken before and after the ESO/HARPS' upgrade in mid-2015, respectively (if existent), and the RMS is also indicated in the legend.

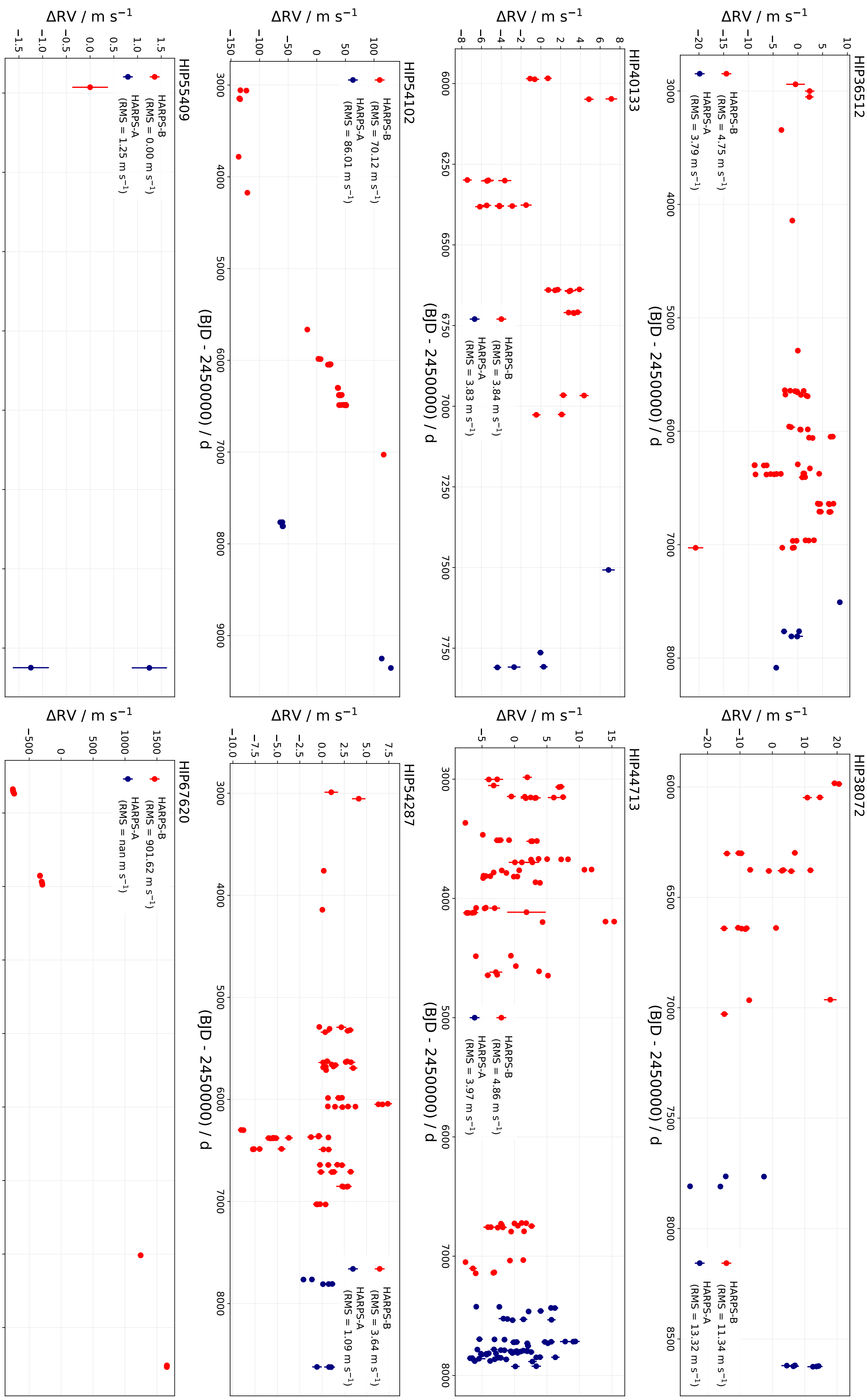

Figure 6.6: RV curves for HIP36512, HIP38072, HIP40133, HIP44713, HIP54102, HIP54287, HIP55409, and HIP67620. Red and blue dots represent the data taken before and after the ESO/HARPS' upgrade in mid-2015, respectively (if existent), and the RMS is also indicated in the legend.

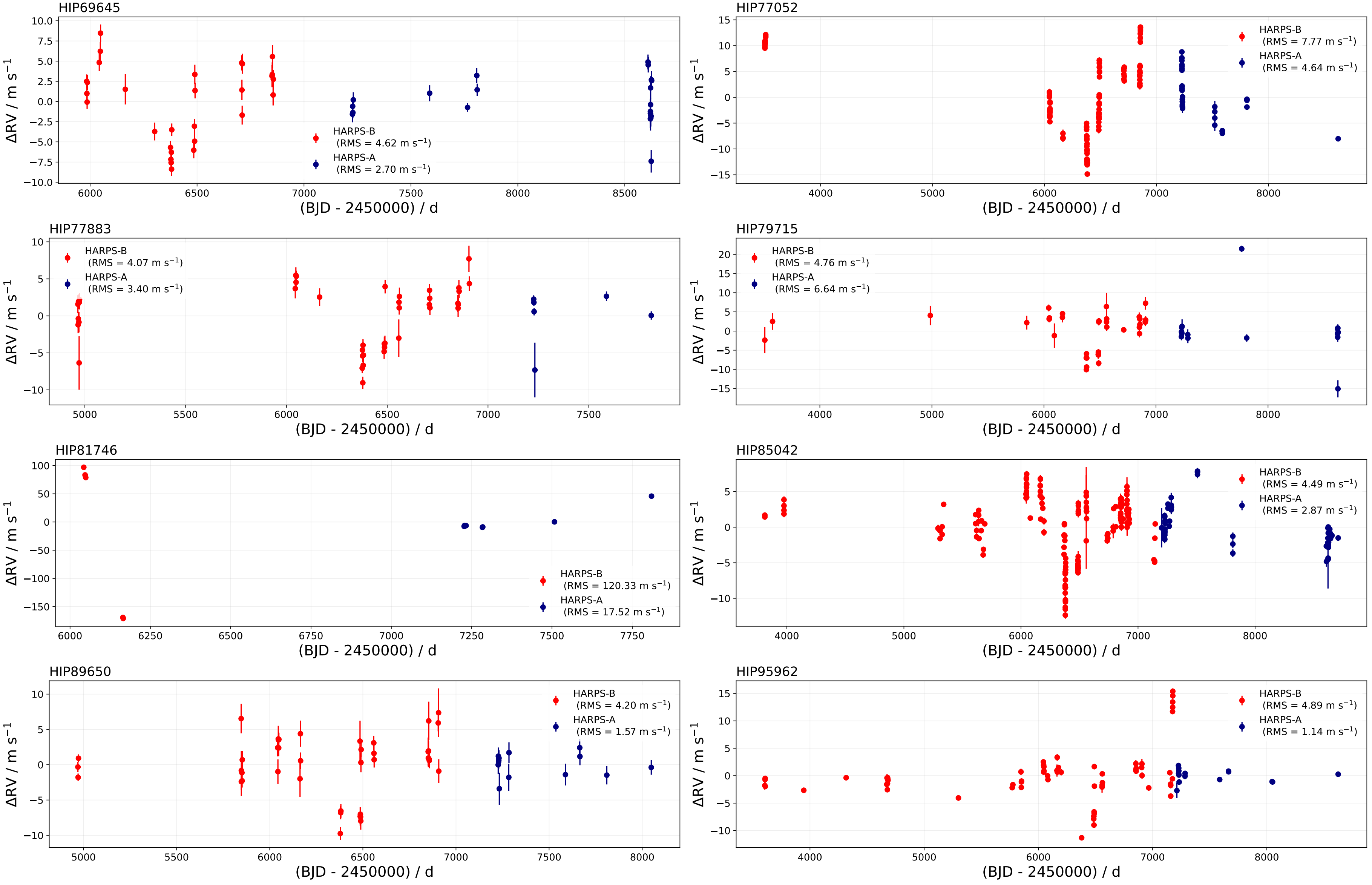

Figure 6.7: RV curves for HIP69645, HIP77052, HIP77883, HIP79715, HIP81746, HIP85042, HIP89650, and HIP95962. Red and blue dots represent the data taken before and after the ESO/HARPS' upgrade in mid-2015, respectively (if existent), and the RMS is also indicated in the legend.

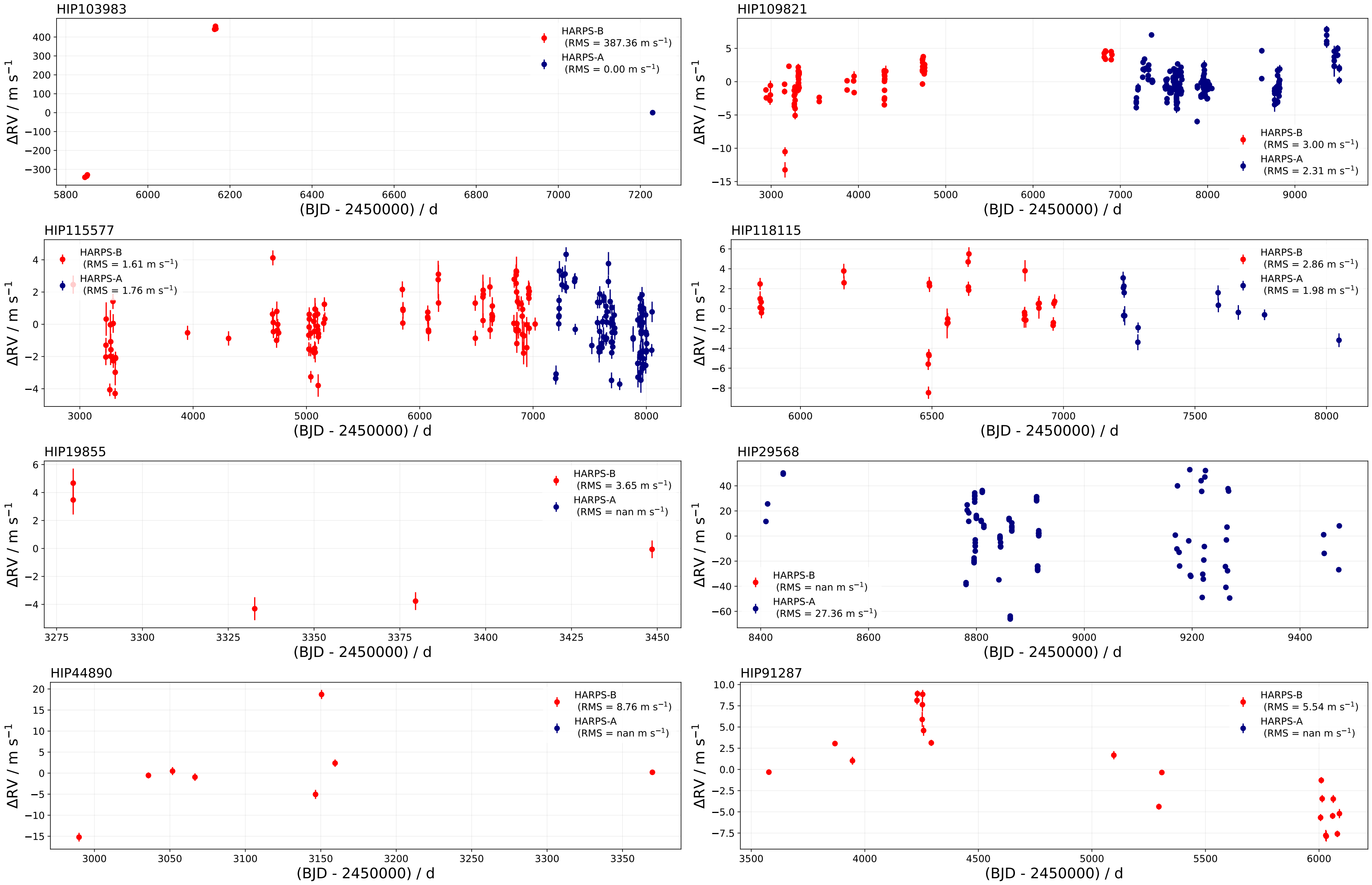

Figure 6.8: RV curves for HIP103983, HIP109821, HIP115577, HIP118115, HIP19855, HIP29568, HIP44890, and HIP91287. Red and blue dots represent the data taken before and after the ESO/HARPS' upgrade in mid-2015, respectively (if existent), and the RMS is also indicated in the legend.

6.4 Correlations Between Radial Velocities and Stellar Magnetic Activity

HIP3311. Given the moderate correlation between the RVs and CCF BIS ($r = 0.43$) and CCF FWHM ($r = 0.39$), strong and statistically significant peaks in these stellar activity indicator proxies periodograms, plus no sign of sinusoidal variations at the prominent Lomb-Scargle period found ($P = 1059d$), the RV modulations seem to be of unknown intrinsic stellar nature.

HIP6455. This case is very similar to the one presented for HIP3311, i.e., given the moderate correlation between the RVs and CCF BIS ($r = 0.50$) and CCF FWHM ($r = 0.49$), strong and statistically significant peaks in these stellar activity indicator proxies periodograms, plus no sign of sinusoidal variations at the prominent Lomb-Scargle period found ($P = 2024.7d$), the RV modulations seem to be of unknown intrinsic stellar nature.

HIP7585. We observed moderate correlations between the RV measurements and the HARPS standard's activity indicators CCF BIS ($r = 0.34$) and CCF FWHM ($r = 0.53$). Albeit a sine-like variation is indeed observed in the RV time series, the same modulations were found in the CCF FWHM, which strongly suggests that the RV modulations are sourced from stellar activity.

HIP10175. With the implementation of the Lomb-Scargle algorithm, we found a prominent period of 1160.40 days in the RV time series. Despite this, we also observed a strong correlation between RV and both CCF FWHM ($r = 0.75$) and moderate correlation for RV and CCF ($r = 0.46$), which cast doubt about the interpretation of the signal modulation to be planetary. Periodic variations at the hypothetical planet's orbital period were also observed in both activity indicators, which suggests that the spectral modulations are the result of intrinsic stellar variations.

HIP14501. The RV curve presents a high slope, which is consistent with the presence of a planet, or a sub-stellar companion, at a distant orbit. Indeed, Maire et al. (2020) reports the presence of an eccentric ($e = 0.59$) cold T-type brown dwarf with $P = 398$ years from direct imaging observations. Whilst we could indeed fit the RV variations with a Keplerian model for an inner planet, these modulations are likely to be sourced from stellar magnetic modulations. The prominent period found with the Lomb-Scargle method was 7724 days (~ 21.16 years), and the orbital eccentricity was 0.059. Besides, there is a strong anti-correlation between the RVs and CCF FWHM ($r = -0.77$), as well as timing corresponding modulations at the prominent period, and a moderate RV-CCF BIS correlation ($r = 0.5$). As discussed in Chapter 2, if we observe the solar system as an exoplanetary one, and detect RV modulations induced by Jupiter, it would cast doubt since its period is consistent with a Sun's activity cycle of roughly eleven years, and its orbital eccentricity is close to zero. Therefore, follow-up observations of HIP14501 are strongly encouraged.

HIP14614. We observed a strong correlation between RV and CCF FWHM ($r = 0.90$), which likely indicates that, even with a detection of a prominent modulation period in the data, we cast doubt into its exoplanetary nature. Albeit a weak RV-CCF BIS correlation is seen ($r = 0.1$), it is not possible to infer the true nature of the RV modulations as the result of stellar activity, neither from a planet.

HIP22504. There are moderate anti-/correlations between the RV measurements and the stellar activity indicator proxies CCF BIS ($r = 0.53$), and CCF FWHM ($r = -0.55$), which may indicate the presence of active regions in the star's surface, and these are contributing too much for the RV modulations.

HIP44997. We observe a moderate RV-CCF BIS ($r = 0.35$; $p = 0.02$) and weak RV-CCF FWHM ($r = 0.10$; $p = 0.54$) correlations. We could indeed model a Keplerian to the RV data, however, this best-fit model is not statistically sufficient to explain the RV modulations, therefore, we cannot claim the presence of a planet with current data in such a system. It deserves to note that the RV data collected before the HARPS upgrade in mid-2015 presented an $\mathrm{RMS} = 2.53$ m s^{-1}, while for data collected after that upgrade, $\mathrm{RMS} = 0.64$ m s^{-1}.

HIP50534. There are no clear sine-like variations of the RV measurements at the prominent Lomb-Scargle resulting period. Interestingly, there is a strong correlation between RVs and CCF FWHM ($r = 0.74$), which may indicate that the host star is active and that the RV modulations are of intrinsic stellar nature.

HIP52369. There is a strong RV-CCF FWHM correlation ($r = 0.71$), and a weak RV-CCF BIS ($r = 0.23$). Two strong peaks in the CCF BIS and CCF FWHM periodograms are noticeable close to the prominent RV Lomb-Scargle period ($P = 2315.5d$), which may indicate that the host star is active and that stellar activity may be playing a role in the RV modulations. The true nature of this system is unknown with current data.

HIP64713. We observe a weak RV-CCF FWHM correlation ($r = 0.19$), and both modulate sinusoidally at the hypothetical planet's orbital period of 1084.7 days. The Keplerian best-fit model resulted in a high eccentricity modulation, which may indicate that the RV variations are true of planetary nature. If confirmed with further observations, it would represent an eccentric cold Jupiter analogue ($P = 1094$ days, $m \sin i = 0.327$ $\mathrm{M_{Jup}}$, $e = 0.44$). Follow-up of this target is desired since the Keplerian model is relatively dubious in our analysis.

HIP73815. There is a moderate scatter in the RV modulations that cast doubt on the interpretation of the signal's nature. We observed no temporal variations in the several activity indicator proxies (CCF BIS, CCF FWHM, $S-$index, Na doublet, He I, and Hα) with the proposed planet's orbital period ($P = 91.2$ days) found with the Lomb-Scargle periodogram technique. We note, however, that in the Na doublet periodogram, the prominent period corresponds to a superior first-order harmonic of the RVs period, and in the CCF BIS periodogram, it corresponds to an inferior second-order harmonic, which likely indicates the presence of active regions in the star's surface, and these are hampering any planetary signal, if existent. No significant correlation was observed between RV and these activity indicators, and there is a relatively little scatter in the RV observations, where we attribute the high error in the orbital eccentricity modelled as a consequence of the high RMS before HARPS' upgrade in mid-2015 (2.75 m s^{-1}). If confirmed with follow-up observations, HIP73815 modulations could be the result of an eccentric warm sub-Saturn ($e = 0.13$, $P = 91.38$ days, and $m \sin i = 55.6$ M$\oplus$), albeit it seems that this is a case of stellar activity.

HIP74389. We observe a moderate correlation between the RVs and both CCF BIS and CCF FWHM, respectively, $r = 0.61$, $p = 1.11 \times 10^{-5}$ and $r = 0.69$, $p = 1.65 \times 10^{-7}$. Moreover, we detected a prominent period of 271.70 days in the RV measurements with a Lomb-Scargle implementation, however, we also observe that at this same period, CCF FWHM modulates coherently, which likely indicates that the RV observations of HIP74389 are not by Keplerian nature, but rather a result of stellar magnetic activity.

HIP74432. The prominent RV Lomb-Scargle period found was 20.10 days. CCF FWHM activity indicator modulates sinusoidally at this period, and there are no clear sine-like variations of the RVs with this same Lomb-Scargle period. There is a weak ($r = -0.12$) RV-CCF BIS anti-correlation and a moderate ($r = 0.30$) RV-CCF FWHM correlation. We model jointly a possible planet modulation and stellar activity with Gaussian Process regression, however, the results are statistically insignificant, so it was not possible to infer the real nature of the RV modulations.

HIP96160. Initially, we treated this system as one hosting an orbiting exoplanet candidate, however, further analysis castled doubt into that interpretation. At the proposed planet's orbital period of $P = 22.7$ days, there are no significant sine-like modulations of the CCF BIS and CCF FWHM activity indicator proxies, nor periodic variations of such indicators, which in principle excludes the RV modulations at this period as of intrinsic stellar nature. We observed, however, with a Nyquist test that the prominent orbital period is an alias of $\sim$29.5 days, which represents one synodic month, indicating that the true orbital period may be an instrumental alias, and not of planetary nature. We point out that the 22.7 days modulation period can be also related to the star's rotation period.

HIP97507. A strong correlation between the RVs and CCF FWHM is seen ($r = 0.74$), but no correlation at all for CCF BIS ($r = -0.04$). The proposed planet's orbital period is a close inferior third-order harmonic of a prominent CCF FWHM period, which indicates the presence of active regions on the star's surface. If confirmed with desired follow-up observations, the RVs modulations are the result of a warm super-Neptune under relative eccentric orbit ($P = 69.17$ days, $e = 0.24$, $m \sin i = 43$ M$\oplus$).

HIP108468. There is no clear sine-like variation at the prominent Lomb-Scargle period ($P = 36.40$ days) among RVs and CCF FWHM or CCF BIS. Weak correlations are visible between RVs and CCF BIS ($r = 0.17$) and CCF FWHM ($r = 0.28$). Interestingly, the prominent Lomb-Scargle period found in the RV time-series data equals the one found in the CCF BIS time-series data, casting doubt if the modulations are planetary. Given the system's age of 7.0 Gyr (Spina et al., 2018), it is probable that the prominent modulation period found is the star's rotation period (see e.g., Spada et al. (2017); Lorenzo-Oliveira et al. (2019, 2020)).

HIP116906. The Keplerian best-fit model does not seem statistically significant, albeit a good eye fit is notable. There is a strong slope in the RVs and no correlation at all with the BIS/FWHM stellar activity indicator proxies. Follow-up observations of this object are desired, and if confirmed, the system hosts a cold eccentric Jupiter ($P = 438$ days, $e = 0.12 \pm 0.12$ – not well constrained –, and $m \sin i = 0.913$ M$_{\text{Jup}}$).

HIP116937. Strong correlations between the RV measurements and CCF BIS ($r = 0.82$) and CCF FWHM ($r = 0.91$) are visible. Timing variations of these activity proxies are also visible at the prominent Lomb-Scargle period found in the RV data ($P = 1375.40$ days). These factors cast doubt if the RV modulations of the system are planetary.

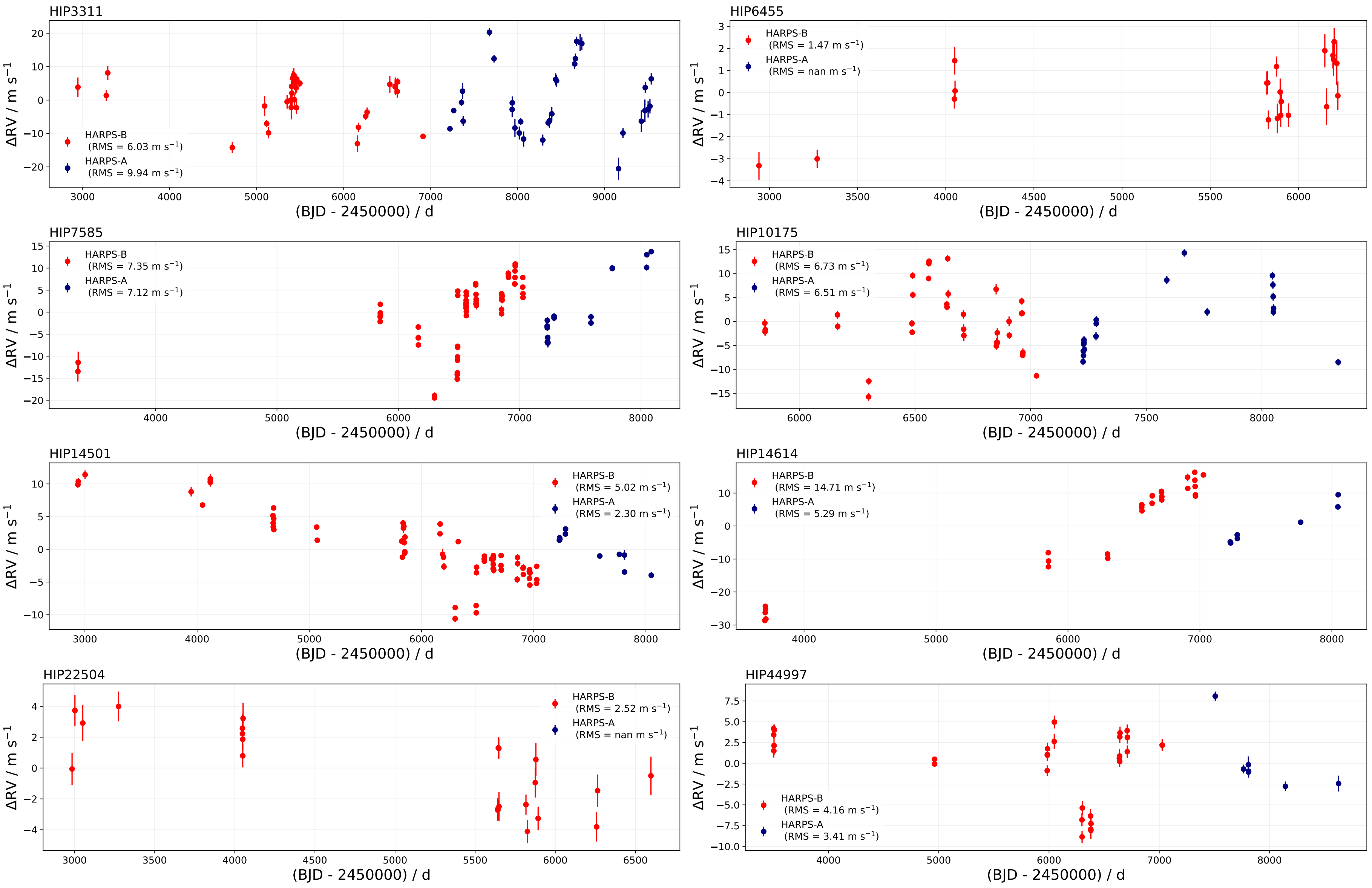

Figure 6.9: RV curves for HIP3311, HIP6455, HIP7585, HIP10175, HIP14501, HIP14614, HIP22504, and HIP44997. Red and blue dots represent the data taken before and after the ESO/HARPS' upgrade in mid-2015, respectively (if existent), and the RMS is also indicated in the legend.

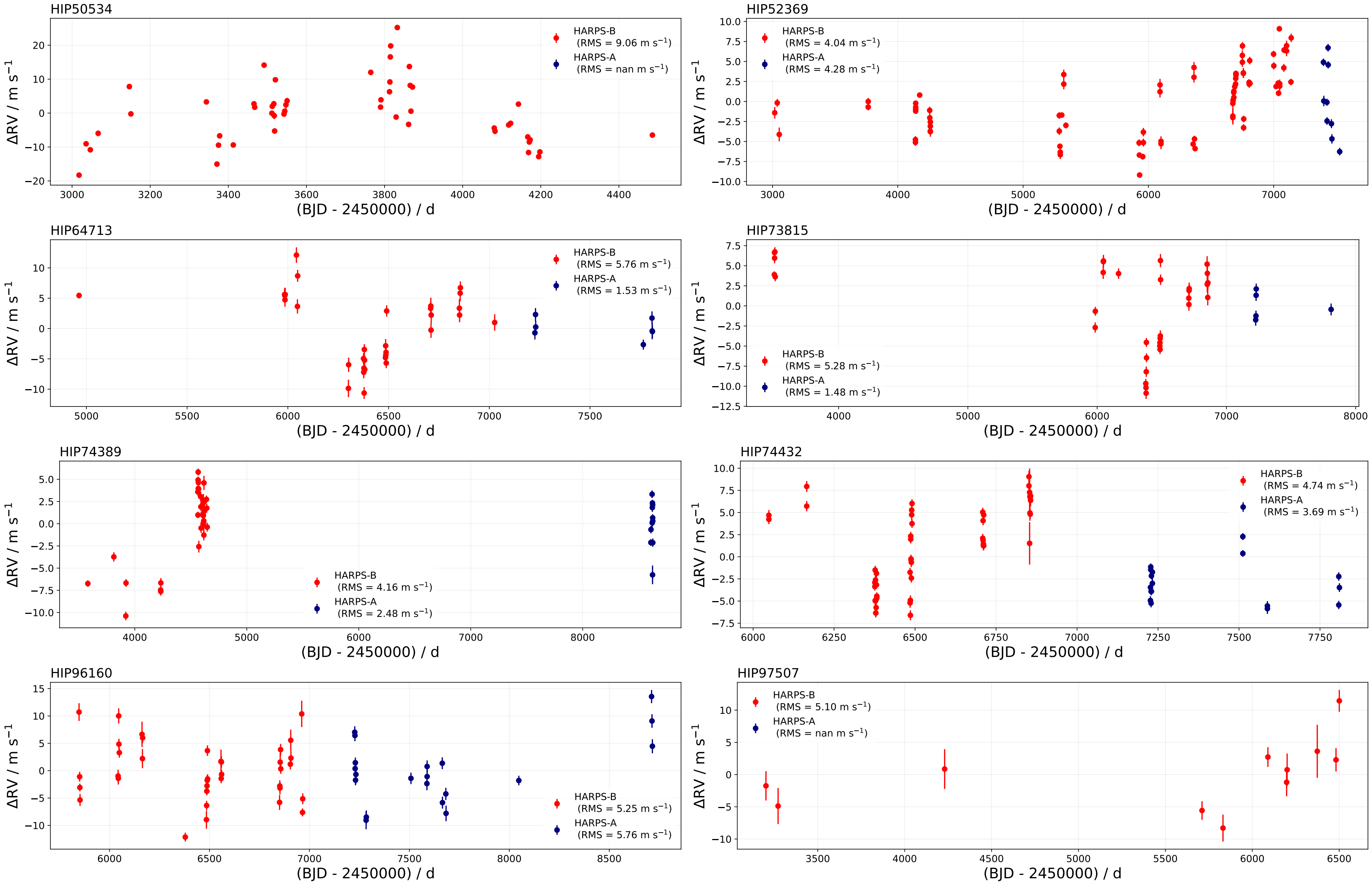

Figure 6.10: RV curves for HIP50534, HIP52369, HIP64713, HIP73815, HIP74389, HIP74432, HIP96160, and HIP97507. Red and blue dots represent the data taken before and after the ESO/HARPS' upgrade in mid-2015, respectively if existent, and the RMS is also indicated in the legend.

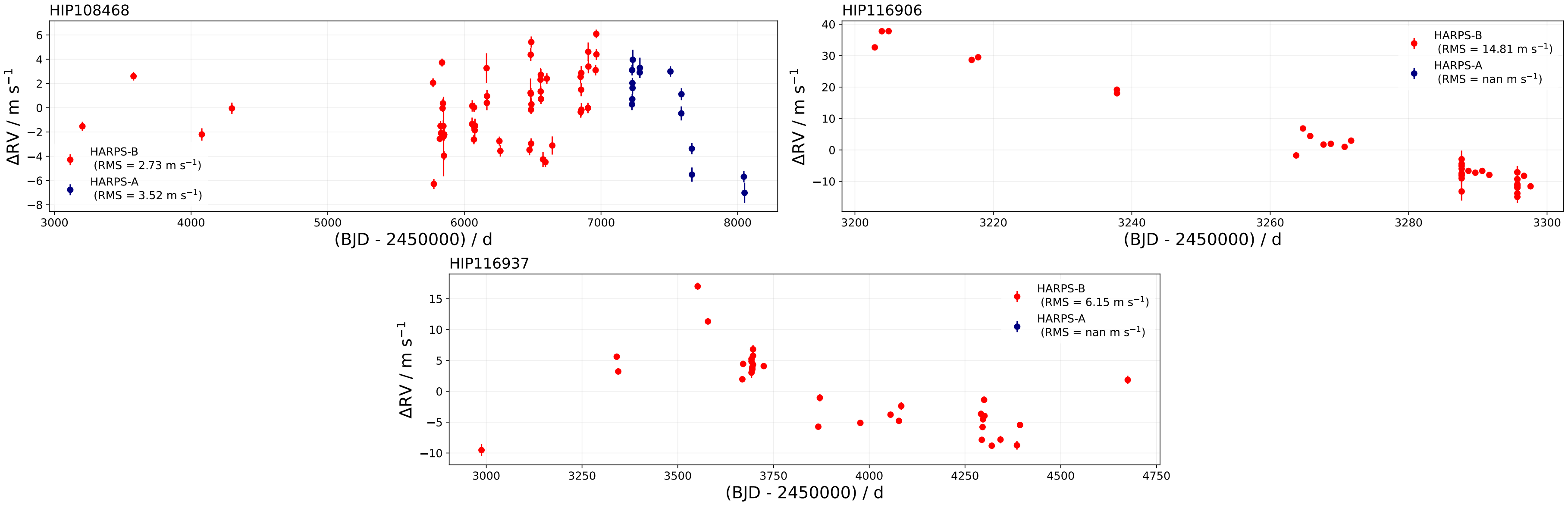

Figure 6.11: RV curves for HIP108468, HIP116906, and HIP116937. Red and blue dots represent the data taken before and after the ESO/HARPS' upgrade in mid-2015, respectively if existent, and the RMS is also indicated in the legend.

Chapter 7

Concluding Remarks and Future Perspectives

In this work, we detected 16 exoplanet candidates using radial velocity data obtained with the ESO/HARPS spectrograph as part of the Solar Twin Planet Search (STPS) observing programme, plus additional campaigns (see Table 5.1 and Fig. 5.3; Ramírez et al. 2014; dos Santos et al. 2017; Bedell et al. 2018; Spina et al. 2018; Botelho et al. 2019; Lorenzo-Oliveira et al. 2019), highlighting subtle features such as evidence of planetary engulfment episodes in some cases (e.g., HIP1954, HIP5301, HIP30502, HIP43297, HIP44935, and HIP87769), as well as the presence of outer companions that might be the responsible to trigger relatively high orbital eccentricities of inner planets (e.g., HIP87769, and HIP108158). Concomitantly, we sought to establish a connection, wherever possible, between the chemical and physical properties of a star and the characteristics of any orbiting exoplanet detected via RV analysis.

Given exoplanet and sub-stellar companion detection, the potential avenues for future research and follow-up are in HIP6407, HIP14501, HIP19911, HIP29525, HIP44997, HIP73241, HIP73815, HIP83276, HIP64713, HIP62039, HIP116906, HIP117367, and HIP97507, as outlined throughout Chapter 6.

The initial step in this thesis involved the implementation of the Lomb-Scargle periodogram method for identifying periodic signals in the RV data, which have a complex nature. Timing and statistical correlations between these measurements and a series of stellar magnetic activity indicators (CCF BIS, CCF FWHM, $S-$index, Hα, Na D and He I) allowed us to determine the likelihood of an RV signal being of planetary origin. This was a crucial step because stellar activity can produce periodic signals in the data that might be extremely difficult to distinguish from planetary signals. Since instrumental modulations and the presence of stellar magnetic activity can mimic, or masquerade, RV signals of exoplanets – which could lead to false detections –, we implemented in some cases a joint Keplerian and quasi-periodic Gaussian Process regression to disentangle these two signals, and additional statistical analyses were performed to ensure that the detected RV signals were truly likely to be planetary.

In light of the planetary engulfment events aforementioned, they represent a significant phenomenon that occurs when a planet comes perilously close to its parent star and breaches what is called the Roche limit. It marks the boundary beyond which a planet's gravitational forces are insufficient to withstand the tidal forces exerted by the host star. This gravitational disparity results in the planet being torn apart by the star's tidal forces, potentially leading to the creation of a debris disc around the star, which is likely to be rich in refractory elements (e.g., Ti, Al, Cr, Fe, Ni, and V) and poor in volatiles (e.g., C, N, and O). This debris is composed of fragments of the planet that failed to resist the tidal forces, which eventually accrete onto the star's photosphere, making the end of the planet's existence. Planetary engulfment phenomena are of particular interest since they provide valuable insights into the complex mechanisms that govern the orbital evolution of exoplanets, and our study provides additional evidence that a non-negligible number of planetary systems around Sun-like stars have experienced an extremely dynamical past, culminating with the fall of planetary-like material into the host star.

In addition to detecting planetary engulfment events, our study also revealed the presence of outer companions in some exoplanetary systems. Treated as a hypothesis that deserves further investigation, these outer objects may be the ones responsible for the high eccentricity of inner planets, which can have important implications for the habitability and long-term stability of exoplanetary systems. Potential theoretical explanations for these phenomena include: (a) Planet-planet scattering; which is a phenomenon that occurs when two, or more planets interact with each other gravitationally, exchanging angular momentum, and resulting in the ejection of one or more planets from the system, or their subsequent settling into relatively highly eccentric orbits. (b) The von Zeipel-Lidov-Kozai mechanism; is another important phenomenon that can affect the dynamics of exoplanetary systems. In this mechanism, in particular, there are also exchanges of angular momentum that occur when a planet is perturbed by an outer companion, leading to changes in the planet's orbital inclination and eccentricity. These changes can have significant implications for the habitability of the planet, as they can affect the planet's climate and atmospheric composition. And (c) Resonance trapping; which is yet another phenomenon that can affect the formation and evolution of exoplanetary systems. This process occurs when the gravitational interactions between planets cause them to settle into stable orbits that are in a resonant relationship with each other. Such resonances can have important implications for the formation and migration of planets within the system, as they can affect the stability of their orbits and even the distribution of material within the disc. Even with our studies providing evidence of outer companions and their influence on inner objects in some exoplanetary systems, further research is required and encouraged, to confirm or refute these findings, as well as to explore these underlying mechanisms. In particular, investigations should focus on theoretical explanations, which will deepen our understanding of the formation and evolution of exoplanetary systems and provide insights into the potential habitability and long-term stability of these systems. Continued efforts in this field will also pave the way for future exoplanet discoveries and our ongoing search for habitable worlds beyond our own Solar System.

The methods and techniques presented in this writing can serve as a starting point for future research on the dynamics and evolution of exoplanetary systems, particularly those around stars that are similar to our Sun. These findings can also help to establish a connection between stars and their orbiting planets. For instance, the use of higher precision RV measurements, a longer temporal baseline, and advanced data analysis techniques can lead to the detection of planets with longer orbital periods.

As of the present writing, we are observing 45 solar twin stars using the ESO/HARPS spectrograph as part of the *How rare are analogues of the Solar System?* project (PI: Lorenzo Spina, with main co-Is Thiago Ferreira and Jorge Meléndez). The sample represents bonafide solar twin stars with stellar parameters such as ages, masses, and metallicities being very similar to those of the Sun. We will expand the current baseline from an average of 7 to 12 years, which is ideal for discovering Jupiter analogues, and could contribute to the identification of more Earth-like exoplanets, which are of particular interest in the quest for habitable worlds (see Section 1.1). Furthermore, we will also exploit the full potential of HARPS spectra by studying the chemical patterns of these stars, which may be linked to the presence of planets (MSc thesis of Giulia Martos), leverage the inherent nature of lithium as a proxy for stellar evolution, as its lower abundance in stars can potentially serve as a key indicator of the processes involved in planet formation (PhD thesis of Anne Rathsam), and examine stellar rotation by assessing their influence on spectral broadening, modelling the impact of stellar activity on radial velocities through magnetic indices, which hold significance in characterising exoplanets (PhD thesis of Gabriela Silva).

Bibliography

Abramowitz M., Stegun I. A., Handbook of Mathematical Functions, 1972

Abt H. A., Maximum Separations among Cataloged Binaries, ApJ, 1988, vol. 331, p. 922

Aerts C., Christensen-Dalsgaard J., Kurtz D. W., Asteroseismology, 2010

Aigrain S., Angus R., Barstow J., Rajpaul V., Gillen E., Parviainen H., Pope B., Roberts S., McQuillan A., Gibson N., Mazeh T., Pont F., Zucker S., The Effects Of Stellar Activity On Detecting And Characterising Planets. In 19th Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun (CS19) , Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun, 2016, p. 12

Aigrain S., Foreman-Mackey D., Gaussian Process regression for astronomical time-series, arXiv e-prints, 2022, p. arXiv:2209.08940

Akaike H., A New Look at the Statistical Model Identification, IEEE Transactions on Automatic Control, 1974, vol. 19, p. 716

Akeson R. L., Chen X., Ciardi D., Crane M., Good J., Harbut M., Jackson E., Kane S. R., Laity A. C., Leifer S., Lynn M., McElroy D. L., Papin M., Plavchan P., Ramírez S. V., Rey R., von Braun K., Wittman M., Abajian M., Ali B., Beichman C., Beekley A., Berriman G. B., Berukoff S., Bryden G., Chan B., Groom S., Lau C., Payne A. N., Regelson M., Saucedo M., Schmitz M., Stauffer J., Wyatt P., Zhang A., The NASA Exoplanet Archive: Data and Tools for Exoplanet Research, PASP, 2013, vol. 125, p. 989

Angus R., Morton T., Aigrain S., Foreman-Mackey D., Rajpaul V., Inferring probabilistic stellar rotation periods using Gaussian processes, MNRAS, 2018, vol. 474, p. 2094

Artigau É., Cadieux C., Cook N. J., Doyon R., Vandal T., Donati J.-F., Moutou C., Delfosse X., Fouqué P., Martioli E., Bouchy F., Parsons J., Carmona A., Dumusque X., Astudillo-Defru N., Bonfils X., Mignon L., Line-by-line Velocity Measurements: an Outlier-resistant Method for Precision Velocimetry, AJ, 2022, vol. 164, p. 84

Baliunas S. L., Donahue R. A., Soon W. H., Horne J. H., Frazer J., Woodard-Eklund L., Bradford M., Rao L. M., Wilson O. C., Zhang Q., Bennett W., Briggs J., Carroll S. M., Duncan D. K., Figueroa D., Lanning H. H., Misch

T., Mueller J., Noyes R. W., Poppe D., Porter A. C., Robinson C. R., Russell J., Shelton J. C., Soyumer T., Vaughan A. H., Whitney J. H., Chromospheric Variations in Main-Sequence Stars. II., ApJ, 1995, vol. 438, p. 269

Baluev R. V., Assessing the statistical significance of periodogram peaks, MNRAS, 2008, vol. 385, p. 1279

Barker A. J., Non-linear tides in a homogeneous rotating planet or star: global simulations of the elliptical instability, MNRAS, 2016, vol. 459, p. 939

Barnes R., Deitrick R., Luger R., Driscoll P. E., Quinn T., Fleming D. P., Guyer B., Habitability of planets on eccentric orbits: limits of the mean flux approximation, Monthly Notices of the Royal Astronomical Society, 2017, vol. 469, p. 4017

Barnes R., Jackson B., Raymond S. N., West A. A., Greenberg R., The HD 40307 Planetary System: Super-Earths or Mini-Neptunes?, ApJ, 2009, vol. 695, p. 1006

Basu S., Chaplin W. J., Asteroseismic Data Analysis. Foundations and Techniques, 2018

Batygin K., Bodenheimer P. H., Laughlin G. P., In Situ Formation and Dynamical Evolution of Hot Jupiter Systems, ApJ, 2016, vol. 829, p. 114

Batygin K., Laughlin G., Jupiter's decisive role in the inner Solar System's early evolution, Proceedings of the National Academy of Science, 2015, vol. 112, p. 4214

Batygin K., Morbidelli A., Tsiganis K., Formation and evolution of planetary systems in presence of highly inclined stellar perturbers, A&A, 2011, vol. 533, p. A7

Bayes M., Price M., An Essay towards Solving a Problem in the Doctrine of Chances. By the Late Rev. Mr. Bayes, F. R. S. Communicated by Mr. Price, in a Letter to John Canton, A. M. F. R. S., Philosophical Transactions of the Royal Society of London Series I, 1763, vol. 53, p. 370

Bayliss D., Gillen E., Eigmüller P., McCormac J., Alexander R. D., Armstrong D. J., Booth R. S., Bouchy F., Burleigh M. R., et al. NGTS-1b: a hot Jupiter transiting an M-dwarf, MNRAS, 2018, vol. 475, p. 4467

Bazot M., Ireland M. J., Huber D., Bedding T. R., Broomhall A. M., Campante T. L., Carfantan H., Chaplin W. J., Elsworth Y., Meléndez J., Petit P., Théado S., Van Grootel V., Arentoft T., Asplund M., Castro M., Christensen-Dalsgaard J., Do Nascimento J. D., Dintrans B., Dumusque X., Kjeldsen H., McAlister H. A., Metcalfe T. S., Monteiro M. J. P. F. G., Santos N. C., Sousa S., Sturmann J., Sturmann L., ten Brummelaar T. A., Turner N., Vauclair S., The radius and mass of the close solar twin 18 Scorpii derived from asteroseismology and interferometry, A&A, 2011, vol. 526, p. L4

Beaugé C., Ferraz-Mello S., Michtchenko T. A., Multi-planet extrasolar systems — detection and dynamics, Research in Astronomy and Astrophysics, 2012, vol. 12, p. 1044

Bedding T. R., Mosser B., Huber D., Montalbán J., Beck P., Christensen-Dalsgaard J., Elsworth Y. P., García R. A., Miglio A., Stello D., White T. R., De Ridder J., Hekker S., Aerts C., Barban C., Belkacem K., Broomhall A.-M., Brown T. M., Buzasi D. L., Carrier F., Chaplin W. J., di Mauro M. P., Dupret M.-A., Frandsen S., Gilliland R. L., Goupil M.-J., Jenkins J. M., Kallinger T., Kawaler S., Kjeldsen H., Mathur S., Noels A., Silva Aguirre V., Ventura P., Gravity modes as a way to distinguish between hydrogen- and helium-burning red giant stars, Nature, 2011, vol. 471, p. 608

Bedell M., Bean J. L., Meléndez J., Spina L., Ramírez I., Asplund M., Alves-Brito A., dos Santos L., Dreizler S., Yong D., Monroe T., Casagrande L., The Chemical Homogeneity of Sun-like Stars in the Solar Neighborhood, ApJ, 2018, vol. 865, p. 68

Benítez-Llambay P., Masset F., Koenigsberger G., Szulágyi J., Planet heating prevents inward migration of planetary cores, Nature, 2015, vol. 520, p. 63

Bensby T., Feltzing S., Oey M. S., Exploring the Milky Way stellar disk. A detailed elemental abundance study of 714 F and G dwarf stars in the solar neighbourhood, A&A, 2014, vol. 562, p. A71

Benton M., The History of Life: A Very Short Introduction. Very Short Introductions, OUP Oxford, 2008

Birnstiel T., Klahr H., Ercolano B., A simple model for the evolution of the dust population in protoplanetary disks, A&A, 2012, vol. 539, p. A148

Bishop C. M., Nasrabadi N. M., Pattern recognition and machine learning. vol. 4, Springer, 2006

Bitsch B., Morbidelli A., Johansen A., Lega E., Lambrechts M., Crida A., Pebble-isolation mass: Scaling law and implications for the formation of super-Earths and gas giants, A&A, 2018, vol. 612, p. A30

Bognanni M., Herbst E., A sequential Monte Carlo approach to inference in multiple[U+2010]equation Markov[U+2010]switching models, Journal of Applied Econometrics, 2018, vol. 33, p. 126

Boisse I., Bouchy F., Hébrard G., Bonfils X., Santos N., Vauclair S., Disentangling between stellar activity and planetary signals, A&A, 2011, vol. 528, p. A4

Boss A. P., Formation of gas and ice giant planets, The Astrophysical Journal, 2006, vol. 641, p. 1148

Botelho R. B., Milone A. d. C., Meléndez J., Bedell M., Spina L., Asplund M., dos Santos L., Bean J. L., Ramírez I., Yong D., Dreizler S., Alves-Brito A., Yana Galarza J., Thorium in solar twins: implications for habitability in rocky planets, MNRAS, 2019, vol. 482, p. 1690

Bouma L. G., Winn J. N., Howard A. W., Howell S. B., Isaacson H., Knutson H., Matson R. A., WASP-4 Is Accelerating toward the Earth, ApJ, 2020, vol. 893, p. L29

Bracewell R., The Fourier Transform and its applications, 1965

Bracewell R. N., The Fourier transform and its applications, 2000

Bretthorst G. L., Generalizing the Lomb-Scargle periodogram. In Bayesian Inference and Maximum Entropy Methods in Science and Engineering , vol. 568 of American Institute of Physics Conference Series, 2001, p. 241

Bryan M. L., Knutson H. A., Howard A. W., Ngo H., Batygin K., Crepp J. R., Fulton B. J., Hinkley S., Isaacson H., Johnson J. A., Marcy G. W., Wright J. T., Statistics of Long Period Gas Giant Planets in Known Planetary Systems, ApJ, 2016, vol. 821, p. 89

Buchhave L. A., Bitsch B., Johansen A., Latham D. W., Bizzarro M., Bieryla A., Kipping D. M., Jupiter Analogs Orbit Stars with an Average Metallicity Close to That of the Sun, ApJ, 2018, vol. 856, p. 37

Burnham K. P., Anderson D. R., Multimodel inference: understanding AIC and BIC in model selection, Sociological methods & research, 2004, vol. 33, p. 261

Butler R. P., Marcy G. W., Vogt S. S., Fischer D. A., Henry G. W., Laughlin G., Wright J. T., Seven New Keck Planets Orbiting G and K Dwarfs, ApJ, 2003, vol. 582, p. 455

Cabral A., Wehbé B., Blind N., Coelho J., Abreu M., Segovia A., Sordet M., Wildi F., Bouchy F., Santos N. C., Doyon R., Lo Curto G., Zins G., Kolb J., NIRPS - the Near Infra-Red Planet Searcher: design, integration, and tests of the atmospheric dispersion corrector. In Ground-based and Airborne Instrumentation for Astronomy IX , vol. 12184 of Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 2022, p. 121844M

Camacho J. D., Faria J. P., Viana P. T. P., Modelling stellar activity with Gaussian process regression networks, MNRAS, 2023, vol. 519, p. 5439

Cameron A. G. W., Physics of the Primitive Solar Accretion Disk, Moon and Planets, 1978, vol. 18, p. 5

Campbell B., Walker G. A. H., Yang S., A Search for Substellar Companions to Solar-type Stars, ApJ, 1988, vol. 331, p. 902

Chatterjee S., Ford E. B., Matsumura S., Rasio F. A., Dynamical Outcomes of Planet-Planet Scattering, ApJ, 2008, vol. 686, p. 580

Chiang E., Laughlin G., The formation of super-earths and mini-neptunes with giant impacts, Monthly Notices of the Royal Astronomical Society, 2013, vol. 431, p. 3444

Cincunegui C., Díaz R. F., Mauas P. J. D., Hα and the Ca II H and K lines as activity proxies for late-type stars, A&A, 2007, vol. 469, p. 309

Clarke A., Morris G. J., Fonseca F., Murray B. J., Acton E., Price H. C., A Low Temperature Limit for Life on Earth, PLoS ONE, 2013, vol. 8, p. e66207

Cloutier R., Doyon R., Menou K., Delfosse X., Dumusque X., Artigau É., On the Radial Velocity Detection of Additional Planets in Transiting, Slowly Rotating M-dwarf Systems: The Case of GJ 1132, AJ, 2017, vol. 153, p. 9

Craig A., Hogg R., McKean J., Introduction to mathematical statistics eighth edition edn. Pearson, 2019

Crane J. D., Shectman S. A., Butler R. P., Thompson I. B., Birk C., Jones P., Burley G. S., The Carnegie Planet Finder Spectrograph: integration and commissioning. In Ground-based and Airborne Instrumentation for Astronomy III , vol. 7735 of Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 2010, p. 773553

Cumming A., Butler R. P., Marcy G. W., Vogt S. S., Wright J. T., Fischer D. A., The Keck Planet Search: Detectability and the Minimum Mass and Orbital Period Distribution of Extrasolar Planets, PASP, 2008, vol. 120, p. 531

Cutri R. M., Wright E. L., Conrow T., Fowler J. W., Eisenhardt P. R. M., Grillmair C., Kirkpatrick J. D., Masci F., McCallon H. L., Wheelock S. L., Fajardo-Acosta S., Yan L., Benford D., Harbut M., Jarrett T., Lake S., Leisawitz D., Ressler M. E., Stanford S. A., Tsai C. W., Liu F., Helou G., Mainzer A., Gettngs D., Gonzalez A., Hoffman D., Marsh K. A., Padgett D., Skrutskie M. F., Beck R., Papin M., Wittman M., VizieR Online Data Catalog: AllWISE Data Release (Cutri+ 2013), VizieR Online Data Catalog, 2021, p. II/328

Davenport M. A., Duarte M. F., Eldar Y. C., Kutyniok G., Introduction to compressed sensing. In Compressed Sensing , 2012

Dawson R. I., Johnson J. A., Origins of Hot Jupiters, ARA&A, 2018, vol. 56, p. 175

Dawson R. I., Murray-Clay R. A., Giant Planets Orbiting Metal-rich Stars Show Signatures of Planet-Planet Interactions, ApJ, 2013, vol. 767, p. L24

do Nascimento J. D. J., de Almeida L., Velloso E. N., Anthony F., Barnes S. A., Saar S. H., Meibom S., da Costa J. S., Castro M., Galarza J. Y., Lorenzo-Oliveira D., Beck P. G., Meléndez J., Rotation of Solar Analogs Crossmatching Kepler and Gaia DR2, ApJ, 2020, vol. 898, p. 173

Doazan V., Peton A., A study of the Hbeta profiles and pulsation of the envelope of the Be star HD 218393., A&A, 1970, vol. 9, p. 245

Dodson-Robinson S. E., Veras D., Ford E. B., Beichman C. A., Showman A. P., Formation of ice giant planets, The Astrophysical Journal, 2009, vol. 707, p. 79

Doppler C., Über das farbige Licht der Doppelsterne und einiger anderer Gestirne des Himmels., 1842

dos Santos L. A., Meléndez J., Bedell M., Bean J. L., Spina L., Alves-Brito A., Dreizler S., Ramírez I., Asplund M., Spectroscopic binaries in the Solar Twin Planet Search program: from substellar-mass to M dwarf companions, MNRAS, 2017, vol. 472, p. 3425

dos Santos L. A., Meléndez J., do Nascimento J.-D., Bedell M., Ramírez I., Bean J. L., Asplund M., Spina L., Dreizler S., Alves-Brito A., Casagrande L., The Solar Twin Planet Search. IV. The Sun as a typical rotator and evidence for a new rotational braking law for Sun-like stars, A&A, 2016, vol. 592, p. A156

Dressing C. D., Charbonneau D., The Occurrence of Potentially Habitable Planets Orbiting M Dwarfs Estimated from the Full Kepler Dataset and an Empirical Measurement of the Detection Sensitivity, The Astrophysical Journal, 2015, vol. 807, p. 45

Dumusque X., Cretignier M., Sosnowska D., Buchschacher N., Lovis C., Phillips D. F., Pepe F., Alesina F., Buchhave L. A., Burnier J., Cecconi M., Cegla H. M., et al. Three years of HARPS-N high-resolution spectroscopy and precise radial velocity data for the Sun, A&A, 2021, vol. 648, p. A103

Duncan D. K., Vaughan A. H., Wilson O. C., Preston G. W., Frazer J. e. a., CA II H and K Measurements Made at Mount Wilson Observatory, 1966–1983, ApJS, 1991, vol. 76, p. 383

Dziembowski W., Oscillations of giants and supergiants., Acta Astron., 1977, vol. 27, p. 95

Efron B., The Jackknife, the Bootstrap and other resampling plans, 1982

Einstein A., Zur Elektrodynamik bewegter Körper, Annalen der Physik, 1905, vol. 322, p. 891

Faure Gunter M. T. M., Introduction to Planetary Science: The Geological Perspective, 2007

Feng Y. K., Wright J. T., Nelson B., Wang S. X., Ford E. B., Marcy G. W., Isaacson H., Howard A. W., The California Planet Survey IV: A Planet Orbiting the Giant Star HD 145934 and Updates to Seven Systems with Long-period Planets, ApJ, 2015, vol. 800, p. 22

Ferraz-Mello S., Estimation of Periods from Unequally Spaced Observations, AJ, 1981, vol. 86, p. 619

Ferraz-Mello S., Tidal synchronization of close-in satellites and exoplanets. A rheophysical approach, Celestial Mechanics and Dynamical Astronomy, 2013, vol. 116, p. 109

Ferraz-Mello S., Tadeu Dos Santos M., Beaugé C., Michtchenko T. A., Rodríguez A., On the mass determination of super-Earths orbiting active stars: the CoRoT-7 system, A&A, 2011, vol. 531, p. A161

Ferraz-Mello S., Tadeu dos Santos M., Folonier H., Csizmadia S., do Nascimento J. D. J., Pätzold M., Interplay of Tidal Evolution and Stellar Wind Braking in the Rotation of Stars Hosting Massive Close-In Planets, ApJ, 2015, vol. 807, p. 78

Ferreira T., Meléndez J., HIP 102152b: A low-mass planet candidate around an old solar twin. In Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun , Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun, 2022, p. 11

Ferreira T., Meléndez J., Lorenzo-Oliveira D., Bean J. L., Spina L., Bedell M., A Jupiter analogue and a cold Super-Neptune orbiting the solar-twin star HIP 104045, arXiv e-prints, 2023, p. arXiv:2303.01358

Fischer D. A., Valenti J., The Planet-Metallicity Correlation, ApJ, 2005, vol. 622, p. 1102

Fischer D. A., Valenti J. A., The metallicity signature of transiting extrasolar planets, The Astrophysical Journal Letters, 2005a, vol. 622, p. 1102

Fischer D. A., Valenti J. A., The Planet-Metallicity Correlation, The Astrophysical Journal, 2005b, vol. 622, p. 1102

Fitton S., Tofflemire B. M., Kraus A. L., Disk Material Inflates Gaia RUWE Values in Single Stars, Research Notes of the American Astronomical Society, 2022, vol. 6, p. 18

Ford E. B., Rasio F. A., Origins of Eccentric Extrasolar Planets: Testing the Planet-Planet Scattering Model, ApJ, 2008, vol. 686, p. 621

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

Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., emcee: The MCMC Hammer, PASP, 2013, vol. 125, p. 306

Fortney J. J., Dawson R. I., Komacek T. D., Hot Jupiters: Origins, Structure, Atmospheres, Journal of Geophysical Research (Planets), 2021, vol. 126, p. e06629

Fressin F., Torres G., Charbonneau D., Bryson S. T., Christiansen J., Dressing C. D., Jenkins J. M., Walkowicz L. M., Batalha N. M., O'rourke J. G., et al., The false positive rate of Kepler and the occurrence of planets, The Astrophysical Journal, 2013, vol. 766, p. 81

Frink S., Mitchell D. S., Quirrenbach A., Fischer D. A., Marcy G. W., Butler R. P., Discovery of a Substellar Companion to the K2 III Giant ι Draconis, ApJ, 2002, vol. 576, p. 478

Fulton B. J., Petigura E. A., Blunt S., Sinukoff E., RadVel: The Radial Velocity Modeling Toolkit, PASP, 2018, vol. 130, p. 044504

Fulton B. J., Petigura E. A., Howard A. W., Isaacson H., Marcy G. W., Cargile P. A., Hebb L., Weiss L. M., Johnson J. A., Morton T. D., Sinukoff E., Crossfield I. J. M., Hirsch L. A., The California-Kepler Survey. III. A Gap in the Radius Distribution of Small Planets, AJ, 2017, vol. 154, p. 109

Gaia Collaboration Vallenari A., Brown A. G. A., Prusti T., de Bruijne J. H. J., Arenou F., Babusiaux C., Biermann M. e. a., Gaia Data Release 3: Summary of the content and survey properties, arXiv e-prints, 2022, p. arXiv:2208.00211

Galarza J. Y., López-Valdivia R., Meléndez J., Lorenzo-Oliveira D., Evidence of Rocky Planet Engulfment in the Wide Binary System HIP 71726/HIP 71737, ApJ, 2021, vol. 922, p. 129

Galarza J. Y., Meléndez J., Cohen J. G., Serendipitous discovery of the faint solar twin Inti 1, A&A, 2016, vol. 589, p. A65

Gan T., Lin Z., Wang S. X., Mao S., Fouqué P., Fan J., Bedell M., Stassun K. G., Giacalone S., Fukui A., Murgas F. e. a., TOI-530b: a giant planet transiting an M-dwarf detected by TESS, MNRAS, 2022, vol. 511, p. 83

Gelman A., Carlin J., Stern H., Dunson D., Vehtari A., Rubin D., Bayesian Data Analysis, Third Edition. Chapman & Hall/CRC Texts in Statistical Science, Taylor & Francis, 2013

Geman S., Geman D., Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images, IEEE Transactions on Pattern Analysis and Machine Intelligence, 1984, vol. 6, p. 721

Ghaderinezhad F., Ley C., A general measure of the impact of priors in Bayesian statistics via Stein's Method, arXiv e-prints, 2018, p. arXiv:1803.00098

Ghaderinezhad F., Ley C., , 2019 in Tang N., ed., , Bayesian Inference on Complicated Data. IntechOpen Rijeka Chapt. 1

Ghezzi L., Cunha K., Smith V. V., de Araújo F. X., Schuler S. C., de la Reza R., Stellar Parameters and Metallicities of Stars Hosting Jovian and Neptunian Mass Planets: A Possible Dependence of Planetary Mass on Metallicity, ApJ, 2010, vol. 720, p. 1290

Ghezzi L., Montet B. T., Johnson J. A., Retired A Stars Revisited: An Updated Giant Planet Occurrence Rate as a Function of Stellar Metallicity and Mass, ApJ, 2018, vol. 860, p. 109

Gibson S. R., Howard A. W., Rider K., Roy A., Edelstein J., Kassis M., Grillo J., Halverson S., Sirk M. M., Smith C., Allen S. e. a., Keck Planet Finder: design updates. In Ground-based and Airborne Instrumentation for Astronomy VIII , vol. 11447 of Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 2020, p. 1144742

Ginzburg S., Sari R., Tidal heating of young super-Earth atmospheres, MNRAS, 2017, vol. 464, p. 3937

Giribaldi R. E., Porto de Mello G. F., Lorenzo-Oliveira D., Amôres E. B., Ubaldo-Melo M. L., Faint solar analogues at the limit of no reddening. Precise atmospheric parameters from moderate low-resolution spectra, A&A, 2019, vol. 629, p. A33

Goldstein H., Poole C., Safko J., Classical mechanics, 2002

Gomes da Silva J., Figueira P., Santos N., Faria J., ACTIN: A tool to calculate stellar activity indices, The Journal of Open Source Software, 2018, vol. 3, p. 667

Gomes da Silva J., Santos N. C., Adibekyan V., Sousa S. G., Campante T. L., Figueira P., Bossini D., Delgado-Mena E., Monteiro M. J. P. F. G., de Laverny P., Recio-Blanco A., Lovis C., Stellar chromospheric activity of 1674 FGK stars from the AMBRE-HARPS sample. I. A catalogue of homogeneous chromospheric activity, A&A, 2021, vol. 646, p. A77

Goodman J., Weare J., Ensemble samplers with affine invariance, Communications in Applied Mathematics and Computational Science, 2010, vol. 5, p. 65

Gradshteyn I. S., Ryzhik I. M., Jeffrey A., Zwillinger D., Table of Integrals, Series, and Products, 2007

Gray D. F., Stellar Rotation and Precise Radial Velocities. In IAU Colloq. 170: Precise Stellar Radial Velocities , vol. 185 of Astronomical Society of the Pacific Conference Series, 1999, p. 243

Gray D. F., The observation and analysis of stellar photospheres 3 edn. Cambridge University Press, 2008

Grunblatt S. K., Huber D., Gaidos E., Lopez E. D., Barclay T., Chontos A., Sinukoff E., Van Eylen V., Howard A. W., Isaacson H. T., Do Close-in Giant Planets Orbiting Evolved Stars Prefer Eccentric Orbits?, ApJ, 2018, vol. 861, p. L5

Guillot T., THE INTERIORS OF GIANT PLANETS: Models and Outstanding Questions, Annual Review of Earth and Planetary Sciences, 2005, vol. 33, p. 493

Guillot T., Santos N. C., Pont F., Iro N., Melo C., Ribas I., A correlation between the heavy element content of transiting extrasolar planets and the metallicity of their parent stars, A&A, 2006, vol. 453, p. L21

Haisch Karl E. J., Lada E. A., Lada C. J., Disk Frequencies and Lifetimes in Young Clusters, ApJ, 2001, vol. 553, p. L153

Hall J. C., Stellar Chromospheric Activity, Living Reviews in Solar Physics, 2008, vol. 5, p. 2

Hall J. C., Lockwood G. W., Skiff B. A., The Activity and Variability of the Sun and Sun-like Stars. I. Synoptic Ca II H and K Observations, AJ, 2007, vol. 133, p. 862

Hansen C. J., Kawaler S. D., Trimble V., Stellar interiors : physical principles, structure, and evolution, 2004

Harrison J. H. D., Bonsor A., Madhusudhan N., Polluted white dwarfs: constraints on the origin and geology of exoplanetary material, MNRAS, 2018, vol. 479, p. 3814

Hartman J. D., Bayliss D., Brahm R., Bakos G. Á., Mancini L., Jordán A., Penev K., Rabus M., Zhou G., Butler R. P., Espinoza N., de Val-Borro M., Bhatti W., Csubry Z., Ciceri S., Henning T., Schmidt B., Arriagada P., Shectman S., Crane J., Thompson I., Suc V., Csák B., Tan T. G., Noyes R. W., Lázár J., Papp I., Sári P., HATS-6b: A Warm Saturn Transiting an Early M Dwarf Star, and a Set of Empirical Relations for Characterizing K and M Dwarf Planet Hosts, AJ, 2015, vol. 149, p. 166

Harwit M., Astrophysical Concepts 4th edn. Springer Science & Business Media, 2003

Hastings W. K., Monte Carlo Sampling Methods using Markov Chains and their Applications, Biometrika, 1970, vol. 57, p. 97

Hatzes A. P., Radial-Velocity Variations from Starspots. In IAU Colloq. 170: Precise Stellar Radial Velocities , vol. 185 of Astronomical Society of the Pacific Conference Series, 1999, p. 259

Hatzes A. P., The Doppler Method for the Detection of Exoplanets, 2019

Hatzes A. P., Cochran W. D., Endl M., McArthur B., Paulson D. B., Walker G. A. H., Campbell B., Yang S., A Planetary Companion to γ Cephei A, ApJ, 2003, vol. 599, p. 1383

Haywood R. D., Collier Cameron A., Queloz D., Barros S. C. C., Deleuil M., Fares R., Gillon M., Lanza A. F., Lovis C., Moutou C., Pepe F., Pollacco D., Santerne A., Ségransan D., Unruh Y. C., Planets and stellar activity: hide and seek in the CoRoT-7 system, MNRAS, 2014, vol. 443, p. 2517

He M. Y., Weiss L. M., Inner Planetary System Gap Complexity is a Predictor of Outer Giant Planets, arXiv e-prints, 2023, p. arXiv:2306.08846

Henry G. W., Search for Transits of a Short-Period, Sub-Saturn Extrasolar Planet Orbiting HD 46375, ApJ, 2000, vol. 536, p. L47

Hoffman M. D., Gelman A., et al., The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo., J. Mach. Learn. Res., 2014, vol. 15, p. 1593

Hori Y., Ikoma M., Formation of low-mass planets by type I migration and the effects of turbulence, Monthly Notices of the Royal Astronomical Society, 2011, vol. 416, p. 1419

Howard A. W., Marcy G. W., Bryson S. T., Jenkins J. M., Rowe J. F., Batalha N. M., Borucki W. J., Christiansen J. L., Cochran W. D., Latham D. W., et al., Planet occurrence within 0.25 au of solar-type stars from Kepler, The Astrophysical Journal Supplement Series, 2012, vol. 201, p. 15

Howard A. W., Marcy G. W., Bryson S. T., Jenkins J. M., Rowe J. F., Batalha N. M. e. a., Planet Occurrence within 0.25 AU of Solar-type Stars from Kepler, ApJS, 2012, vol. 201, p. 15

Hubeny I., Mihalas D., Theory of Stellar Atmospheres, 2014

Ibañez Bustos R. V., Buccino A. P., Flores M., Martinez C. F., Mauas P. J. D., Correlation between activity indicators: Hα and Ca II lines in M-dwarf stars, A&A, 2023, vol. 672, p. A37

Ito T., Ohtsuka K., The Lidov-Kozai Oscillation and Hugo von Zeipel, Monographs on Environment, Earth and Planets, 2019, vol. 7, p. 1

Ivezić Ž., Connolly A. J., VanderPlas J. T., Gray A., Statistics, Data Mining, and Machine Learning in Astronomy: A Practical Python Guide for the Analysis of Survey Data, 2014

Izidoro A., Raymond S. N., Morbidelli A., Hersant F., Pierens A., Gas Giant Planets as Dynamical Barriers to Inward-Migrating Super-Earths, ApJ, 2015, vol. 800, p. L22

Jackson B., Miller N., Barnes R., Raymond S. N., Fortney J. J., Greenberg R., The roles of tidal evolution and evaporative mass loss in the origin of CoRoT-7 b, MNRAS, 2010, vol. 407, p. 910

Jeffers S. V., Barnes J. R., Jones H. R. A., Reiners A., Pinfield D. J., Marsden S. C., Is it possible to detect planets around young active G and K dwarfs?, MNRAS, 2014, vol. 438, p. 2717

Jeffreys H., An Invariant Form for the Prior Probability in Estimation Problems, Proceedings of the Royal Society of London Series A, 1946, vol. 186, p. 453

Johansen A., Lambrechts M., Forming Planets via Pebble Accretion, Annual Review of Earth and Planetary Sciences, 2017, vol. 45, p. 359

Johansen A., Youdin A., Klahr H., Zonal Flows and Long-lived Axisymmetric Pressure Bumps in Magnetorotational Turbulence, ApJ, 2009, vol. 697, p. 1269

Johnson J. A., Gazak J. Z., Apps K., Muirhead P. S., Crepp J. R., Crossfield I. J. M., Boyajian T., von Braun K., Rojas-Ayala B., Howard A. W., Covey K. R., Schlawin E., Hamren K., Morton T. D., Marcy G. W., Lloyd J. P., Characterizing the Cool KOIs. II. The M Dwarf KOI-254 and Its Hot Jupiter, AJ, 2012, vol. 143, p. 111

Jones D. E., Stenning D. C., Ford E. B., Wolpert R. L., Loredo T. J., Gilbertson C., Dumusque X., Improving Exoplanet Detection Power: Multivariate Gaussian Process Models for Stellar Activity, arXiv e-prints, 2017, p. arXiv:1711.01318

Kane S. R., Hill M. L., Kasting J. F., Kopparapu R. K., Quintana E. V., Barclay T., Batalha N. M., Borucki W. J., Ciardi D. R., Haghighipour N., Hinkel N. R., Kaltenegger L., Selsis F., Torres G., A Catalog of Kepler Habitable Zone Exoplanet Candidates, ApJ, 2016, vol. 830, p. 1

Kass R. E., Wasserman L., The Selection of Prior Distributions by Formal Rules, Journal of the American Statistical Association, 1996, vol. 91, p. 1343

Keller J. P., Selection of input and output variables as a model reduction problem, 1990 American Control Conference, 1990

Kepler J., Astronomia nova., 1609

Kepler J., Epitome astronomiae copernicanae usitata forma quaestionum & responsionum conscripta, inque VII. libros digesta, 1618

Kepler J., Harmonices Mvndi Libri V.: Qvorvm Primus Geometricvs, De Figurarum Regularium, quae Proportiones Harmonicas constituunt, ortu demonstrationibus. Secundus Architectonicvs, seu ex Geometria Figvrata, De Figurarum Regularium Congruentia in plano vel solido: Tertius propriè Harmonicvs, De Proportionum Harmonicarum ortu ex Figuris, 1619

Kervella P., Arenou F., Mignard F., Thévenin F., Stellar and substellar companions of nearby stars from Gaia DR2. Binarity from proper motion anomaly, A&A, 2019, vol. 623, p. A72

Kippenhahn R., Weigert A., Weiss A., Stellar Structure and Evolution, 2013

Kjeldsen H., Bedding T. R., Amplitudes of stellar oscillations: the implications for asteroseismology., A&A, 1995, vol. 293, p. 87

Kokubo E., Ida S., Formation of Protoplanets from Planetesimals in the Solar Nebula, Icarus, 2000, vol. 143, p. 15

Kopparapu R. K., Ramirez R. M., SchottelKotte J., Kasting J. F., Domagal-Goldman S., Eymet V., Habitable Zones around Main-sequence Stars: Dependence on Planetary Mass, ApJ, 2014, vol. 787, p. L29

Kozai Y., Secular perturbations of asteroids with high inclination and eccentricity, AJ, 1962, vol. 67, p. 591

Kratter K., Lodato G., Gravitational Instabilities in Circumstellar Disks, ARA&A, 2016, vol. 54, p. 271

Kruschke J. K., Bayesian estimation supersedes the t test, Journal of Experimental Psychology: General, 2013, vol. 142, p. 573

Lambrechts M., Johansen A., Forming the cores of giant planets from the radial pebble flux in protoplanetary discs, A&A, 2014, vol. 572, p. A107

Landau L. D., Lifshitz E. M., The classical theory of fields, 1975

Latham D. W., Mazeh T., Stefanik R. P., Mayor M., Burki G., The unseen companion of HD114762: a probable brown dwarf, Nature, 1989, vol. 339, p. 38

Lehmann C., Murphy M. T., Liu F., Flynn C., Smith D., Berke D. A., Survey for Distant Solar Twins (SDST) - III. Identification of new solar twin and solar analogue stars, MNRAS, 2023, vol. 521, p. 148

Leinhardt Z. M., Richardson D. C., Lufkin G., Haseltine J., Planetesimals to protoplanets - II. Effect of debris on terrestrial planet formation, MNRAS, 2009, vol. 396, p. 718

Li J., Lai D., Planetary Spin and Obliquity from Mergers, ApJ, 2020, vol. 898, p. L20

Liddle A. R., Information criteria for astrophysical model selection, MNRAS, 2007, vol. 377, p. L74

Lidov M. L., The evolution of orbits of artificial satellites of planets under the action of gravitational perturbations of external bodies, Planet. Space Sci., 1962, vol. 9, p. 719

Lin D. N. C., Bodenheimer P., Richardson D. C., Orbital migration of the planetary companion of 51 Pegasi to its present location, Nature, 1996, vol. 380, p. 606

Lissauer J. J., Planet formation., ARA&A, 1993, vol. 31, p. 129

Loaiza-Tacuri V., Cunha K., Smith V. V., Martinez C. F., Ghezzi L., Schuler S. C., Teske J., Howell S. B., A Spectroscopic Analysis of a Sample of K2 Planet-Host Stars: Stellar Parameters, Metallicities and Planetary Radii, arXiv e-prints, 2023, p. arXiv:2301.03961

Lodders K., Solar System Abundances and Condensation Temperatures of the Elements, ApJ, 2003, vol. 591, p. 1220

Lomb N. R., Least-squares frequency analysis of unequally spaced data, Ap&SS, 1976, vol. 39, p. 447

Lopez E. D., Fortney J. J., Understanding the Mass-Radius Relation for Sub-neptunes: Radius as a Proxy for Composition, ApJ, 2014, vol. 792, p. 1

Lorenzo-Oliveira D., Freitas F. C., Meléndez J., Bedell M., Ramírez I., Bean J. L., Asplund M., Spina L., Dreizler S., Alves-Brito A., Casagrande L., The Solar Twin Planet Search. The age-chromospheric activity relation, A&A, 2018, vol. 619, p. A73

Lorenzo-Oliveira D., Meléndez J., Ponte G., Galarza J. Y., The ancient main-sequence solar proxy HIP 102152 unveils the activity and rotational fate of our Sun, MNRAS, 2020, vol. 495, p. L61

Lorenzo-Oliveira D., Meléndez J., Yana Galarza J., Ponte G., dos Santos L. A., Spina L., Bedell M., Ramírez I., Bean J. L., Asplund M., Constraining the evolution of stellar rotation using solar twins, MNRAS, 2019, vol. 485, p. L68

Lorenzo-Oliveira D., Porto de Mello G. F., Schiavon R. P., The age-mass-metallicity-activity relation for solar-type stars: comparisons with asteroseismology and the NGC 188 open cluster, A&A, 2016, vol. 594, p. L3

Lovis C., Fischer D., Radial Velocity Techniques for Exoplanets, Annual Review of Astronomy and Astrophysics, 2010, vol. 48, p. 581

Lucy L. B., Sweeney M. A., Spectroscopic binaries with circular orbits., AJ, 1971, vol. 76, p. 544

Luhman K. L., Patten B. M., Marengo M., Schuster M. T., Hora J. L., Ellis R. G., Stauffer J. R., Sonnett S. M., Winston E., Gutermuth R. A., Megeath S. T., Backman D. E., Henry T. J., Werner M. W., Fazio G. G., Discovery of Two T Dwarf Companions with the Spitzer Space Telescope, ApJ, 2007, vol. 654, p. 570

Luhn J. K., Wright J. T., Howard A. W., Isaacson H., Astrophysical Insights into Radial Velocity Jitter from an Analysis of 600 Planet-search Stars, AJ, 2020, vol. 159, p. 235

Maire A. L., Molaverdikhani K., Desidera S., Trifonov T., Mollière P., D'Orazi V., Frankel N., Baudino J. L., Messina S., Müller A., et al. Orbital and spectral characterization of the benchmark T-type brown dwarf HD 19467B, A&A, 2020, vol. 639, p. A47

Marion G. M., Fritsen C. H., Eicken H., Payne M. C., The Search for Life on Europa: Limiting Environmental Factors, Potential Habitats, and Earth Analogues, Astrobiology, 2003, vol. 3, p. 785

Martinez C. F., Cunha K., Ghezzi L., Smith V. V., A Spectroscopic Analysis of the California-Kepler Survey Sample. I. Stellar Parameters, Planetary Radii, and a Slope in the Radius Gap, ApJ, 2019, vol. 875, p. 29

Mayor M., Marmier M., Lovis C., Udry S., Ségransan D., Pepe F., Benz W., Bertaux J. L., Bouchy F., Dumusque X., Lo Curto G., Mordasini C., Queloz D., Santos N. C., The HARPS search for southern extra-solar planets XXXIV. Occurrence, mass distribution and orbital properties of super-Earths and Neptune-mass planets, arXiv e-prints, 2011, p. arXiv:1109.2497

Mayor M., Pepe F., Queloz D., Bouchy F., Rupprecht G., Lo Curto G., Avila G., Benz W., Bertaux J. L., Bonfils X., Dall T., Dekker H., Delabre B., Eckert W., Fleury M., Gilliotte A., Gojak D., Guzman J. C., Kohler D., Lizon J. L., Longinotti A., Lovis C., Megevand D., Pasquini L., Reyes J., Sivan J. P., Sosnowska D., Soto R., Udry S., van Kesteren A., Weber L., Weilenmann U., Setting New Standards with HARPS, The Messenger, 2003, vol. 114, p. 20

Mayor M., Queloz D., A Jupiter-mass companion to a solar-type star, Nature, 1995, vol. 378, p. 355

Meléndez J., Asplund M., Gustafsson B., Yong D., The Peculiar Solar Composition and Its Possible Relation to Planet Formation, ApJ, 2009, vol. 704, p. L66

Meléndez J., Bedell M., Bean J. L., Ramírez I., Asplund M., Dreizler S., Yan H.-L., Shi J.-R., Lind K., Ferraz-Mello S., Galarza J. Y., dos Santos L., Spina L., Maia M. T., Alves-Brito A., Monroe T., Casagrande L., The Solar Twin Planet Search. V. Close-in, low-mass planet candidates and evidence of planet accretion in the solar twin HIP 68468, A&A, 2017, vol. 597, p. A34

Meléndez J., Ramírez I., HIP 56948: A Solar Twin with a Low Lithium Abundance, ApJ, 2007, vol. 669, p. L89

Meléndez J., Ramírez I., Karakas A. I., Yong D., Monroe T. R., Bedell M., Bergemann M., Asplund M., Tucci Maia M., Bean J., do Nascimento José-Dias J., Bazot M., Alves-Brito A., Freitas F. C., Castro M., 18 Sco: A Solar Twin Rich in Refractory and Neutron-capture Elements. Implications for Chemical Tagging, ApJ, 2014, vol. 791, p. 14

Militzer B., Hubbard W. B., Vorberger J., Tamblyn I., Bonev S. A., A Massive Core in Jupiter Predicted from First-Principles Simulations, ApJ, 2008, vol. 688, p. L45

Miller N., Fortney J. J., The Heavy-element Masses of Extrasolar Giant Planets, Revealed, ApJ, 2011, vol. 736, p. L29

Millholland S., Laughlin G., Obliquity-driven sculpting of exoplanetary systems, Nature Astronomy, 2019, vol. 3, p. 424

Morbidelli A., Raymond S. N., Challenges in planet formation, Journal of Geophysical Research (Planets), 2016, vol. 121, p. 1962

Morbidelli A., Tsiganis K., Batygin K., Crida A., Gomes R., Building Terrestrial Planets, Annual Review of Earth and Planetary Sciences, 2012, vol. 40, p. 251

Mortier A., Faria J. P., Correia C. M., Santerne A., Santos N. C., BGLS: A Bayesian formalism for the generalised Lomb-Scargle periodogram, A&A, 2015, vol. 573, p. A101

Murray C. D., Correia A. C. M., , 2010 in Seager S., ed., , Exoplanets. pp 15–23

Naef D., Mayor M., Lo Curto G., Bouchy F., Lovis C., Moutou C., Benz W., Pepe F., Queloz D., Santos N. C., Ségransan D., Udry S., Bonfils X., Delfosse X., Forveille T., Hébrard G., Mordasini C., Perrier C., Boisse I., Sosnowska D., The HARPS search for southern extrasolar planets. XXIII. 8 planetary companions to low-activity solar-type stars, A&A, 2010, vol. 523, p. A15

Nagar T., Spina L., Karakas A. I., The Chemical Signatures of Planetary Engulfment Events in Binary Systems, ApJ, 2020, vol. 888, p. L9

Nakamoto T., Nakagawa Y., Formation, Early Evolution, and Gravitational Stability of Protoplanetary Disks, ApJ, 1994, vol. 421, p. 640

Naoz S., Farr W. M., Lithwick Y., Rasio F. A., Teyssandier J., Hot Jupiters from secular planet-planet interactions, Nature, 2011, vol. 473, p. 187

Netto Y., Lorenzo-Oliveira D., Meléndez J., Yana Galarza J., Haywood R. D., Spina L., dos Santos L. A., Radial-velocity Precision of ESPRESSO Through the Analysis of the Solar Twin HIP 11915, AJ, 2021, vol. 162, p. 160

Netto Y., Valio A., Stellar magnetic activity and the butterfly diagram of Kepler-63, A&A, 2020, vol. 635, p. A78

Netzel H., Smolec R., Asteroseismology of RR Lyrae stars with non-radial modes, MNRAS, 2022, vol. 515, p. 3439

Newton I., Philosophiae Naturalis Principia Mathematica. Auctore Js. Newton, 1687

Olspert N., Pelt J., Käpylä M. J., Lehtinen J., Estimating activity cycles with probabilistic methods. I. Bayesian generalised Lomb-Scargle periodogram with trend, A&A, 2018, vol. 615, p. A111

Oppenheim A. V., Schafer R. W., Buck J. R., Discrete-Time Signal Processing. Prentice-Hall Upper Saddle River, NJ, 1999

Osaki Y., Stability of Cepheid-Type Stars against Nonradial Oscillations, PASJ, 1977, vol. 29, p. 235

Oshagh M., Santos N. C., Figueira P., Barros S. C. C., Donati J. F., Adibekyan V., Faria J. P., Watson C. A., Cegla H. M., Dumusque X., Hébrard E., Demangeon O., Dreizler S., Boisse I., Deleuil M., Bonfils X., Pepe F., Udry S., Understanding stellar activity-induced radial velocity jitter using simultaneous K2 photometry and HARPS RV measurements, A&A, 2017, vol. 606, p. A107

Owen J. E., Wu Y., Kepler Planets: A Tale of Evaporation, ApJ, 2013, vol. 775, p. 105

Owen J. E., Wu Y., The Evaporation Valley in the Kepler Planets, ApJ, 2017, vol. 847, p. 29

Oza A. V., Johnson R. E., Lellouch E., Schmidt C., Schneider N., Huang C., Gamborino D., Gebek A., Wyttenbach A., Demory B.-O., Mordasini C., Saxena P., Dubois D., Moullet A., Thomas N., Sodium and Potassium Signatures of Volcanic Satellites Orbiting Close-in Gas Giant Exoplanets, ApJ, 2019, vol. 885, p. 168

Pacheco T. A., Hector Garcia L., Meléndez J., On the Nature of the HD120690 Companion: A Solar Twin-M Dwarf Binary, Research Notes of the American Astronomical Society, 2019, vol. 3, p. 149

Paegert M., Stassun K. G., Collins K. A., Pepper J., Torres G., Jenkins J., Twicken J. D., Latham D. W., VizieR Online Data Catalog: TESS Input Catalog version 8.2 (TIC v8.2) (Paegert+, 2021), VizieR Online Data Catalog, 2022, p. IV/39

Palmer C. W., Diffraction Grating Handbook 6th edn. Newport Corporation, 2006

Parviainen H., , 2018 Bayesian Methods for Exoplanet Science. p. 149

Pepe F., Mayor M., Galland F., et al. The HARPS search for southern extra-solar planets, A&A, 2002, vol. 388, p. 632

Pepe F. A., Cristiani S., Rebolo Lopez R., Santos N. C., Amorim A., Avila G., Benz W., Bonifacio P., Cabral A., Carvas P., Cirami R. e. a., ESPRESSO: the Echelle spectrograph for rocky exoplanets and stable spectroscopic observations. In Ground-based and Airborne Instrumentation for Astronomy III , vol. 7735 of Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 2010, p. 77350F

Perryman M. A. C., Lindegren L., Kovalevsky J., Hoeg E., Bastian U., Bernacca P. L., Crézé M., Donati F., Grenon M., Grewing M., van Leeuwen F., van der Marel H., Mignard F., Murray C. A., Le Poole R. S., Schrijver H., Turon C., Arenou F., Froeschlé M., Petersen C. S., The HIPPARCOS Catalogue, A&A, 1997, vol. 323, p. L49

Petersburg R. R., Ong J. M. J., Zhao L. L., Blackman R. T., Brewer J. M., Buchhave L. A., Cabot S. H. C., Davis A. B., Jurgenson C. A., Leet C., McCracken T. M., Sawyer D., Sharov M., Tronsgaard R., Szymkowiak A. E., Fischer D. A., An Extreme-precision Radial-velocity Pipeline: First Radial Velocities from EXPRES, AJ, 2020, vol. 159, p. 187

Petigura E. A., Howard A. W., Marcy G. W., Prevalence of Earth-size planets orbiting Sun-like stars, Proceedings of the National Academy of Science, 2013, vol. 110, p. 19273

Petigura E. A., Howard A. W., Marcy G. W., Prevalence of Earth-size planets orbiting Sun-like stars, Proceedings of the National Academy of Sciences, 2013, vol. 110, p. 19273

Pont F., Hébrard G., Irwin J. M., Bouchy F., Moutou C., Ehrenreich D., Guillot T., Aigrain S., Bonfils X., Berta Z., Boisse I., Burke C., Charbonneau D., Delfosse X., Desort M., Eggenberger A., Forveille T., Lagrange A. M., Lovis C., Nutzman P., Pepe F., Perrier C., Queloz D., Santos N. C., Ségransan D., Udry S., Vidal-Madjar A., Spin-orbit misalignment in the HD 80606 planetary system, A&A, 2009, vol. 502, p. 695

Poon S. T. S., Nelson R. P., Coleman G. A. L., In situ formation of hot Jupiters with companion super-Earths, MNRAS, 2021, vol. 505, p. 2500

Porto de Mello G. F., da Silva L., HR 6060: The Closest Ever Solar Twin?, ApJ, 1997, vol. 482, p. L89

Pourbaix D., Tokovinin A. A., Batten A. H., Fekel F. C., Hartkopf W. I., Levato H., Morrell N. I., Torres G., Udry S., S_B^9: The ninth catalogue of spectroscopic binary orbits, A&A, 2004, vol. 424, p. 727

Pu B., Wu Y., Yuan Y., Lai D., Formation and evolution of planetary systems: the impact of high-angular momentum material, The Astrophysical Journal Letters, 2017, vol. 845, p. L13

Putirka K. D., Xu S., Polluted white dwarfs reveal exotic mantle rock types on exoplanets in our solar neighborhood, Nature Communications, 2021, vol. 12, p. 6168

Quanz S. P., Absil O., Benz W., Bonfils X., Berger J.-P., Defrère D., van Dishoeck E., Ehrenreich D., Fortney J., Glauser A., Grenfell J. L., Janson M., Kraus S., Krause O., Labadie L., Lacour S., Line M., Linz H., Loicq J., Miguel Y., Pallé E., Queloz D., Rauer H., Ribas I., Rugheimer S., Selsis F., Snellen I., Sozzetti A., Stapelfeldt K. R., Udry S., Wyatt M., Atmospheric characterization of terrestrial exoplanets in the mid-infrared: biosignatures, habitability, and diversity, Experimental Astronomy, 2021

Queloz D., Henry G. W., Sivan J. P., Baliunas S. L., Beuzit J. L., Donahue R. A., Mayor M., Naef D., Perrier C., Udry S., No planet for HD 166435, A&A, 2001, vol. 379, p. 279

Quillen A. C., Holman M., Production of Star-grazing and Star-impacting Planetesimals via Orbital Migration of Extrasolar Planets, AJ, 2000, vol. 119, p. 397

Rajpaul V., Aigrain S., Osborne M. A., Reece S., Roberts S., A Gaussian process framework for modelling stellar activity signals in radial velocity data, MNRAS, 2015, vol. 452, p. 2269

Ramírez I., Meléndez J., Asplund M., Accurate abundance patterns of solar twins and analogs. Does the anomalous solar chemical composition come from planet formation?, A&A, 2009, vol. 508, p. L17

Ramírez I., Meléndez J., Bean J., Asplund M., Bedell M., Monroe T., Casagrande L., Schirbel L., Dreizler S., Teske J., Tucci Maia M., Alves-Brito A., Baumann P., The Solar Twin Planet Search. I. Fundamental parameters of the stellar sample, A&A, 2014, vol. 572, p. A48

Rasmussen C. E., Williams C. K. I., Gaussian Processes for Machine Learning, 2006

Raymond S. N., Barnes R., Armitage P. J., Gorelick N., Mean Motion Resonances from Planet-Planet Scattering, ApJ, 2008, vol. 687, p. L107

Raymond S. N., Barnes R., Veras D., Armitage P. J., Gorelick N., The dynamical architecture and habitable zones of the quintet of exoplanets orbiting HD 219134, The Astrophysical Journal, 2016, vol. 831, p. 196

Raymond S. N., Izidoro A., Origin of water in the inner Solar System: Planetesimals scattered inward during Jupiter and Saturn's rapid gas accretion, Icarus, 2017, vol. 297, p. 134

Raymond S. N., Izidoro A., Morbidelli A., , 2020 in Meadows V. S., Arney G. N., Schmidt B. E., Des Marais D. J., eds, , Planetary Astrobiology. p. 287

Raymond S. N., Morbidelli A., Planet Formation: Key Mechanisms and Global Models. In Demographics of Exoplanetary Systems, Lecture Notes of the 3rd Advanced School on Exoplanetary Science , vol. 466 of Astrophysics and Space Science Library, 2022, p. 3

Rees D., Essential Statistics. Springer US, 1989

Reggiani H., Schlaufman K. C., Healy B. F., Lothringer J. D., Sing D. K., Evidence that the Hot Jupiter WASP-77 A b Formed Beyond Its Parent Protoplanetary Disk's H2O Ice Line, AJ, 2022, vol. 163, p. 159

Reiners A., Activity-induced radial velocity jitter in a flaring M dwarf, A&A, 2009, vol. 498, p. 853

Roberts Lewis C. J., Mason B. D., Neyman C. R., Wu Y., Riddle R. L., Shelton J. C., Angione J., Baranec C., Bouchez A., Bui K., Burruss R., Burse M., Chordia P., Croner E., Das H., Dekany R. G., Guiwits S., Hale D., Henning J., Kulkarni S., Law N., McKenna D., Milburn J., Palmer D., Punnadi S., Ramaprakash A. N., Roberts J. E., Tendulkar S. P., Trinh T., Troy M., Truong T., Zolkower J., Know the Star, Know the Planet. IV. A Stellar Companion to the Host Star of the Eccentric Exoplanet HD 8673b, AJ, 2015, vol. 149, p. 144

Robertson P., Mahadevan S., Endl M., Roy A., Stellar activity masquerading as planets in the habitable zone of the M dwarf Gliese 581, Science, 2014, vol. 345, p. 440

Rocha-Pinto H. J., Maciel W. J., Metallicity effects on the chromospheric activity-age relation for late-type dwarfs, MNRAS, 1998, vol. 298, p. 332

Rodet L., Beust H., Bonnefoy M., Lagrange A. M., Galli P. A. B., Ducourant C., Teixeira R., Origin of the wide-orbit circumbinary giant planet HD 106906. A dynamical scenario and its impact on the disk, A&A, 2017, vol. 602, p. A12

Rodet L., Su Y., Lai D., On the Correlation between Hot Jupiters and Stellar Clustering: High-eccentricity Migration Induced by Stellar Flybys, ApJ, 2021, vol. 913, p. 104

Rogers L. A., Most 1.6 Earth-radius Planets are Not Rocky, ApJ, 2015, vol. 801, p. 41

Rosenthal L. J., Fulton B. J., Hirsch L. A., Isaacson H. T., Howard A. W., Dedrick C. M., Sherstyuk I. A., Blunt S. C., Petigura E. A., Knutson H. A., Behmard A., Chontos A., Crepp J. R., Crossfield I. J. M., Dalba P. A., Fischer D. A., Henry G. W., Kane S. R., Kosiarek M., Marcy G. W., Rubenzahl R. A., Weiss L. M., Wright J. T., The California Legacy Survey. I. A Catalog of 178 Planets from Precision Radial Velocity Monitoring of 719 Nearby Stars over Three Decades, ApJS, 2021, vol. 255, p. 8

Rowan D., Meschiari S., Laughlin G., Vogt S. S., Butler R. P., Burt J., Wang S., Holden B., Hanson R., Arriagada P., Keiser S., Teske J., Diaz M., The Lick-Carnegie Exoplanet Survey: HD 32963—A New Jupiter Analog Orbiting a Sun-like Star, ApJ, 2016, vol. 817, p. 104

Rybicki G. B., Lightman A. P., Radiative processes in astrophysics, 1979

Saar S. H., Donahue R. A., Activity-Related Radial Velocity Variation in Cool Stars, ApJ, 1997, vol. 485, p. 319

Safronov V. S., Evolution of the protoplanetary cloud and formation of the earth and planets., 1972

Salvatier J., Wiecki T., Fonnesbeck C., Probabilistic Programming in Python using PyMC, arXiv e-prints, 2015, p. arXiv:1507.08050

Samus' N. N., Kazarovets E. V., Durlevich O. V., Kireeva N. N., Pastukhova E. N., General catalogue of variable stars: Version GCVS 5.1, Astronomy Reports, 2017, vol. 61, p. 80

Santos N. C., Israelian G., Mayor M., On the correlation between metallicity and the presence of giant planets around stars, Astronomy & Astrophysics, 2003, vol. 398, p. 363

Scargle J. D., Studies in astronomical time series analysis. II - Statistical aspects of spectral analysis of unevenly spaced data, ApJ, 1982, vol. 263, p. 835

Schiller M., Bizzarro M., Fernandes V. A., Isotopic evolution of the protoplanetary disk and the building blocks of Earth and the Moon, Nature, 2018, vol. 555, p. 507

Schwarz G., Estimating the Dimension of a Model, Annals of Statistics, 1978, vol. 6, p. 461

Shannon C. E., Communication in the Presence of Noise, IEEE Proceedings, 1949, vol. 37, p. 10

Shevchenko I. I., The Lidov-Kozai Effect - Applications in Exoplanet Research and Dynamical Astronomy. vol. 441, 2017

Silva-Valio A., Lanza A. F., Alonso R., Barge P., Properties of starspots on CoRoT-2, A&A, 2010, vol. 510, p. A25

Skrutskie M. F., Cutri R. M., Stiening R., Weinberg M. D., Schneider S., Carpenter J. M., Beichman C., Capps R., Chester T., Elias J., Huchra J., Liebert J., Lonsdale C., Monet D. G., Price S., Seitzer P., Jarrett T., Kirkpatrick J. D., Gizis J. E., Howard E., Evans T., Fowler J., Fullmer L., Hurt R., Light R., Kopan E. L., Marsh K. A., McCallon H. L., Tam R., Van Dyk S., Wheelock S., The Two Micron All Sky Survey (2MASS), AJ, 2006, vol. 131, p. 1163

Sobolev V. V., Moving Envelopes of Stars, 1960

Sousa S. G., Adibekyan V., Delgado-Mena E., Santos N. C., Rojas-Ayala B., Soares B. M. T. B., Legoinha H., Ulmer-Moll S., Camacho J. D., Barros S. C. C., Demangeon O. D. S., Hoyer S., Israelian G., Mortier A., Tsantaki M., Monteiro M. A., SWEET-Cat 2.0: The Cat just got SWEETer. Higher quality spectra and precise parallaxes from Gaia eDR3, A&A, 2021, vol. 656, p. A53

Sousa S. G., Santos N. C., Mayor M., Udry S., Casagrande L., Israelian G., Pepe F., Queloz D., Monteiro M. J. P. F. G., Spectroscopic parameters for 451 stars in the HARPS GTO planet search program. Stellar [Fe/H] and the frequency of exo-Neptunes, A&A, 2008, vol. 487, p. 373

Spada F., Demarque P., Kim Y. C., Boyajian T. S., Brewer J. M., The Yale-Potsdam Stellar Isochrones, ApJ, 2017, vol. 838, p. 161

Spina L., Meléndez J., Karakas A. I., dos Santos L., Bedell M., Asplund M., Ramírez I., Yong D., Alves-Brito A., Bean J. L., Dreizler S., The temporal evolution of neutron-capture elements in the Galactic discs, MNRAS, 2018, vol. 474, p. 2580

Spina L., Palla F., Randich S., Sacco G., Jeffries R., Magrini L., Franciosini E., Meyer M. R., Tautvaišienė G., Gilmore G., Alfaro E. J., Allende Prieto C., Bensby T., Bragaglia A., Flaccomio E., Koposov S. E., Lanzafame A. C., Costado M. T., Hourihane A., Lardo C., Lewis J., Monaco L., Morbidelli L., Sousa S. G., Worley C. C., Zaggia S., The Gaia-ESO Survey: chemical signatures of rocky accretion in a young solar-type star, A&A, 2015, vol. 582, p. L6

Spina L., Sharma P., Meléndez J., Bedell M., Casey A. R., Carlos M., Franciosini E., Vallenari A., Chemical evidence for planetary ingestion in a quarter of Sun-like stars, Nature Astronomy, 2021, vol. 5, p. 1163

Sudarsky D., Burrows A., Hubeny I., Theoretical Spectra and Atmospheres of Extrasolar Giant Planets, ApJ, 2003, vol. 588, p. 1121

Teitler S., Königl A., Why is there a Dearth of Close-in Planets around Fast-rotating Stars?, ApJ, 2014, vol. 786, p. 139

Terquem C., Papaloizou J. C. B., Dynamical relaxation and the orbits of low-mass extrasolar planets, MNRAS, 2002, vol. 332, p. L39

Torres C. A. O., Quast G. R., da Silva L., de La Reza R., Melo C. H. F., Sterzik M., Search for associations containing young stars (SACY). I. Sample and searching method, A&A, 2006, vol. 460, p. 695

Trierweiler I. L., Doyle A. E., Melis C., Walsh K. J., Young E. D., Exomoons as Sources of White Dwarf Pollution, ApJ, 2022, vol. 936, p. 30

Turner J. D., Flagg L., Ridden-Harper A., Jayawardhana R., Characterizing the WASP-4 System with TESS and Radial Velocity Data: Constraints on the Cause of the Hot Jupiter's Changing Orbit and Evidence of an Outer Planet, AJ, 2022, vol. 163, p. 281

Valenti J. A., Fischer D. A., Spectroscopic Properties of Cool Stars (SPOCS). I. 1040 F, G, and K Dwarfs from Keck, Lick, and AAT Planet Search Programs, ApJS, 2005, vol. 159, p. 141

Valenti J. A., Fischer D. A., Spectroscopic Properties of Cool Stars (SPOCS). I. 1040 F, G, and K Dwarfs from Keck, Lick, and AAT Planet Search Programs, The Astrophysical Journal Supplement Series, 2005, vol. 159, p. 141

Valio A., Araújo A., Stellar Obliquity from Spot Transit Mapping of Kepler-210, ApJ, 2022, vol. 940, p. 132

Van Hoolst T., Dziembowski W. A., Kawaler S. D., Unstable non-radial modes in radial pulsators: theory and an example, MNRAS, 1998, vol. 297, p. 536

van Leeuwen F., de Bruijne J., Babusiaux C., Busso G., Castañeda J., Ducourant C., Fabricius e. a., , 2022 Gaia DR3 documentation

van Maanen A., Two Faint Stars with Large Proper Motion, PASP, 1917, vol. 29, p. 258

Vanderburg A., Plavchan P., Johnson J. A., Ciardi D. R., Swift J., Kane S. R., Radial velocity planet detection biases at the stellar rotational period, MNRAS, 2016, vol. 459, p. 3565

VanderPlas J. T., Understanding the Lomb-Scargle Periodogram, ApJS, 2018, vol. 236, p. 16

Venturini J., Guilera O. M., Haldemann J., Ronco M. P., Mordasini C., The nature of the radius valley. Hints from formation and evolution models, A&A, 2020, vol. 643, p. L1

Virtanen P., Gommers R., Oliphant T. E., Haberland M., Reddy T., Cournapeau D., Burovski E., Peterson P., Weckesser W., Bright J., van der Walt S. J., Brett M., Wilson J., Millman K. J., Mayorov N., Nelson A. R. J., Jones E., Kern R., Larson E., Carey C. J., Polat İ., Feng Y., Moore E. W., VanderPlas J., Laxalde D., Perktold J., Cimrman R., Henriksen I., Quintero E. A., Harris C. R., Archibald A. M., Ribeiro A. H., Pedregosa F., van Mulbregt P., SciPy 1. 0 Contributors SciPy 1.0: fundamental algorithms for scientific computing in Python, Nature Methods, 2020, vol. 17, p. 261

Vogt S. S., Burt J., Meschiari S., Butler R. P., Henry G. W., Wang S., Holden B., Gapp C., Hanson R., Arriagada P., Keiser S., Teske J., Laughlin G., Six Planets Orbiting HD 219134, ApJ, 2015, vol. 814, p. 12

Wasserman L., All of Statistics: A Concise Course in Statistical Inference. Springer, 2010

Weidenschilling S. J., Aerodynamics of solid bodies in the solar nebula., MNRAS, 1977, vol. 180, p. 57

Whittaker E. T., , 1915 XVIII.—On the Functions which are represented by the Expansions of the Interpolation-Theory

Wittenmyer R. A., Butler R. P., Tinney C. G., Horner J., Carter B. D., Wright D. J., Jones H. R. A., Bailey J., O'Toole S. J., The Anglo-Australian Planet Search XXIV: The Frequency of Jupiter Analogs, ApJ, 2016, vol. 819, p. 28

Wolszczan A., Frail D. A., A planetary system around the millisecond pulsar PSR1257 + 12, Nature, 1992, vol. 355, p. 145

Wright J. T., Radial Velocity Jitter in Stars from the California and Carnegie Planet Search at Keck Observatory, PASP, 2005, vol. 117, p. 657

Wright J. T., Howard A. W., Efficient Fitting of Multiplanet Keplerian Models to Radial Velocity and Astrometry Data, ApJS, 2009, vol. 182, p. 205

Wu D.-H., Rice M., Wang S., Evidence for Hidden Nearby Companions to Hot Jupiters, AJ, 2023, vol. 165, p. 171

Xu S., Bonsor A., Exogeology from Polluted White Dwarfs, Elements, 2021, vol. 17, p. 241

Yana Galarza J., López-Valdivia R., Lorenzo-Oliveira D., Reggiani H., Meléndez J., Gamarra-Sánchez D., Flores M., Portal-Rivera J., Miquelarena P., Ponte G., Schlaufman K. C., Vargas Auccalla T., Searching for new solar twins: The Inti survey for the Northern Sky, MNRAS, 2021, vol. 504, p. 1873

Yang C.-C., Mac Low M.-M., Johansen A., Diffusion and Concentration of Solids in the Dead Zone of a Protoplanetary Disk, ApJ, 2018, vol. 868, p. 27

Youdin A. N., Goodman J., Streaming Instabilities in Protoplanetary Disks, ApJ, 2005, vol. 620, p. 459

Yu J., Huber D., Bedding T. R., Stello D., Predicting radial-velocity jitter induced by stellar oscillations based on Kepler data, MNRAS, 2018, vol. 480, p. L48

Zaslavsky B., Bayesian Hypothesis Testing in Two[U+2010]Arm Trials with Dichotomous Outcomes, Biometrics, 2013

Zechmeister M., Kürster M., The generalised Lomb-Scargle periodogram. A new formalism for the floating-mean and Keplerian periodograms, A&A, 2009, vol. 496, p. 577

Zeng L., Jacobsen S. B., Sasselov D. D., Petaev M. I., Vanderburg A., Lopez-Morales M., Perez-Mercader J., Mattsson T. R., Li G., Heising M. Z., Bonomo A. S., Damasso M., Berger T. A., Cao H., Levi A., Wordsworth R. D., Growth model interpretation of planet size distribution, Proceedings of the National Academy of Science, 2019, vol. 116, p. 9723

Zeng L., Sasselov D. D., Jacobsen S. B., Mass-Radius Relation for Rocky Planets Based on PREM, ApJ, 2016, vol. 819, p. 127

Zhao L., Fischer D. A., Brewer J., Giguere M., Rojas-Ayala B., Planet Detectability in the Alpha Centauri System, AJ, 2018, vol. 155, p. 24

Zirin H., Astrophysics of the sun, 1988