Università degli Studi di Padova

Dipartimento di Fisica e Astronomia "Galileo Galilei"

Master's Degree in Astrophysics and Cosmology

Master's thesis

Homogeneous radial velocity analysis of ultra-short period planets

Supervisor
Prof. Luca Malavolta
Università degli Studi di Padova

Candidate
Júlia Mestre
Student ID: 2080657

Academic year
2023/2024

"Real worlds, surreal life."
— Kevin Parker

Abstract

Ultra-short period (USP) planets, which orbit their host stars in less than a day, represent one of the most intriguing areas of development in exoplanetary science. These planets can offer unique insights into planetary formation, migration, and survival under extreme conditions, such as intense stellar radiation and strong tidal forces. With a growing catalogue of USP planets from missions like Kepler, CoRoT, and TESS, as well as spectroscopic data from high-resolution spectrographs such as HARPS, HARPS-N, and ESPRESSO, we are now capable of characterizing these planets with the indirectly obtained radii, masses, and bulk densities. However, usually only some system configurations are tested based on the Generalized Lomb-Scargle Periodogram (GLS) and the effect of stellar activity is not properly considered, which can lead to inaccurate results. This thesis aims to homogeneously re-analyse a sample of USP planets using radial velocity (RV) data by testing different models, finding more likely configurations and obtaining or improving the previously obtained planetary parameters. To this end, stellar parameters are uniformly determined through isochrone fitting, while RV and stellar activity data are gathered from literature and available catalogues. To mitigate the impact of stellar noise in the RV data, Gaussian processes (GP) are used within a Bayesian framework implemented in PyORBIT, a Python package that supports both Markov Chain Monte Carlo (MCMC) and Nested Sampling algorithms for robust parameter estimation and model selection. Multiple models are tested, including different methods for stellar activity mitigation using GPs and varying number of planets for each planetary system. We also verify the equivalence of two mathematical implementations of the same kernel for GPs. The statistical relevance of each model is calculated with the Bayesian evidence and information criteria. The final set of planetary parameters are obtained with MCMC algorithms. The derived properties of the USP sample are presented, together with the orbital architecture of their planetary systems. The sample is situated in a mass-radius plot to contextualize it within planetary composition theoretical models.

Resumo

Planetas de período ultra-curto (USP), que orbitam as suas estrelas em menos de um dia, representam uma das áreas mais intrigantes atualmente a ser desenvolvida em ciências exoplanetárias. Estes planetas proporcionam perspetivas únicas sobre a formação de planetas, processos de migração e de sobrevivência sob condições extremas, como exposição a radiação solar intensa e forças de maré acentuadas. Com o crescente catálogo de planetas USP descobertos por missões como Kepler, CoRoT e TESS além dos dados espectroscópicos provenientes de espetrógrafos de alta-resolução como HARPS, HARPS-N e ESPRESSO, é possível caracterizar estes planetas com a massa, raio e densidade determinados indiretamente. Contudo, geralmente apenas algumas configurações dos sistemas planetários são testadas com base em periodogramas generalizados Lomb-Scargle, e o efeito da atividade estrelar não é propriamente considerado, o que pode levar a resultados imprecisos. Esta tese tem como objetivo reanalisar uma amostra de planetas USP utilizando dados de velocidade radial (RV). Diferentes modelos são explorados para identificar configurações mais prováveis além de refinar os parâmetros planetários previamente obtidos. Para tal, os parâmetros estrelares são determinados uniformemente através do uso de curvas isócronas, enquanto os dados de velocidade radial e de atividade estrelar são recolhidos da literatura ou de catálogos disponíveis. De modo a mitigar o impacto do ruído estrelar nos dados de velocidade radial, processos gaussianos (GP) são modelados num esquema estatístico baseado no teorema de Bayes implementado no PyORBIT. Este é um pacote de Python que suporta tanto Markov Chain Monte Carlo (MCMC) como Nested Sampling, permitindo a obtenção robusta de parâmetros e a seleção do modelo mais adequado. Diversos modelos são testados, incluindo diferentes maneiras de mitigar a atividade estrelar através do uso de GPs bem como modelos com diferentes números de planetas para cada sistema planetário. Também verificamos a equivalência de duas implementações matemáticas do mesmo kernel para GPs. Os conjuntos finais de parâmetros planetários são obtidos com algoritmos MCMC. As propriedades desta amostra de planetas são apresentadas, juntamente com a arquitetura orbital dos sistemas planetários. A amostra é situada num gráfico de massa-raio para contextualizá-la com modelos teóricos de composição interior planetária.

Contents

List of figures

List of tables

List of abbreviations

BIS ..... Bisector Inverse Span
BF ..... Bayes Factor
CCD ..... Charge-Coupled Device
CCF ..... Cross Correlation Function
CoRoT ..... Convection Rotation and planetary Transits
ES ..... Exponential Sine
ESP ..... Exponential Sine Periodic
ESPRESSO ..... Échelle SPectrograph for Rocky Exoplanets and Stable Spectroscopic Observations
FWHM ..... Full Width at Half Maximum
GLS ..... Generalized Lomb-Scargle
GP ..... Gaussian Process
HARPS ..... High Accuracy Radial Velocity Planet Searcher
HARPS-N ..... High Accuracy Radial Velocity Planet Searcher North
HIRES ..... High Resolution Échelle Spectrometer
MCMC ..... Markov Chain Monte Carlo
NS ..... Nested Sampling
QP ..... Quasi-periodic
RV ..... Radial Velocity
SE ..... Squared Exponential
SEP ..... Squared Exponential Periodic
SNR ..... Signal to Noise Ratio
TESS ..... Transiting Exoplanet Survey Satellite
USP ..... Ultra-Short Period

1 Introduction

Exoplanetary science has developed rapidly over the past three decades. While technically the first detection of an exoplanet occurred in 1992 around a pulsar [1], the field is generally considered to have begun in 1995 with the discovery of an exoplanet orbiting a Sun-like star, 51 Pegasi b [2]. Early discoveries were dominated by "Hot Jupiters"— massive gas giants in close orbits around their stars. However, advancements in instrumentation and data analysis have enabled the detection of Earth-like planets within the habitable zones of their stars. Over the years, the focus has expanded from merely detecting exoplanets to characterizing them and their atmospheres. Multiple detection methods have been developed, including radial velocity (RV) measurements, transit photometry, astrometry, direct imaging, transit timing variations, and gravitational microlensing [3]. Each technique, with its inherent observational biases, contributes uniquely to our understanding of the diversity and formation processes of planetary systems. In parallel, specialized space- and ground-based observatories have been dedicated to exoplanetary studies, resulting in the detection of over 5,700 exoplanets to date*, a number that continues to grow rapidly. Among these methods, the transit technique has been the most prolific, followed by the RV technique. The RV method is particularly valuable because it relies on gravitational effects to determine a planet's mass, one of the most crucial parameters for characterizing an exoplanet. When combined with photometric transit observations that provide the planet's radius, the bulk density can be calculated, offering key insights into the planet's internal structure and surface gravity.

As our observational capabilities and techniques have advanced, we have not only discovered Earth-like planets and giant gas worlds but also identified a variety of unusual and extreme exoplanetary types. Among these, ultra-short period (USP) planets are an interesting group. These are planets that orbit extremely close to their host stars, completing a full orbit in less than a day. Such tight orbits expose these planets to intense stellar radiation and tidal forces, creating environments that challenge our understanding of planetary formation, migration, and survival.

*according to the NASA Exoplanet Archive.

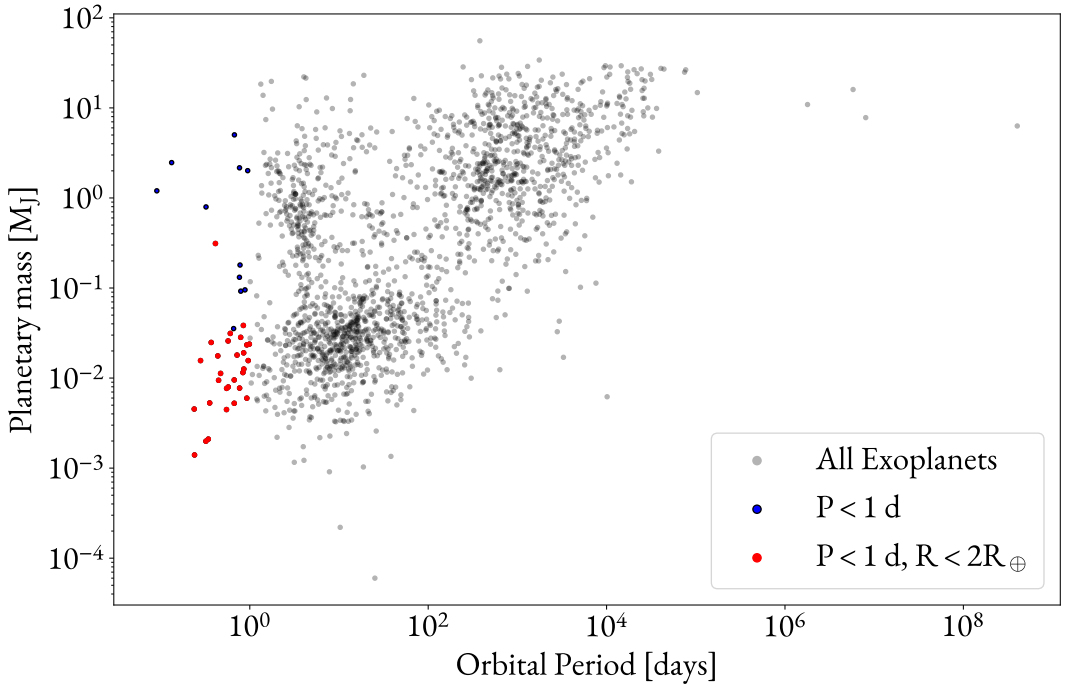


Figure 1.1: Logarithmic plot of orbital period vs. planetary mass of all confirmed exoplanets. Red points represent ultra-short period planets. Data used to produce these plots sourced from NASA Exoplanet Archive.

1.1 Ultra-Short Period Planets

Ultra-short period planets, or USPPs, are exoplanets that orbit their stars with a very short orbital period (< 1 day) and have a small radius ($R < 2\ R_\oplus$). They occur in around 0.5% of Sun-like stars [4].

In 2006, the Sagittarius Window Eclipsing Extrasolar Planet Search (SWEEPS) transit survey was conducted using the Hubble Space Telescope. It detected five giant planet candidates with an orbital period shorter than one day that were dubbed ultra-short period planets or USPPs [5]. However, what is considered the first confirmed transiting USP planet is CoRoT-7b, detected in 2009 with the European mission CoRoT. With an orbital radius of 1.68 $R_\oplus$ and orbital period of 0.85 days, it was the first exoplanet shown to have a rock or metal-dominated composition [6]. The planet 55 Cnc e, initially observed in 2004 with the McDonald Observatory [7], was confirmed as a USP planet after revision to its orbital parameters [8][9]. The launch of NASA's Kepler Space Telescope in 2009 marked a turning point. Designed to find Earth-sized planets in or near the habitable zone, Kepler also identified a wide variety of exoplanets throughout the 4 years of observations on about 200,000 stars, including USPPs [10]. Discovered in 2011, Kepler-10 b was one of the first rocky planets found by Kepler, with an orbital period of just 20 hours [11]. Two years later, Kepler-78 b was found to orbit its star in just 8.5 hours [12]. While the searches of the Kepler telescope did not target USP planets, Sanchis-Ojeda et al. 2014 [4] managed to isolate the periodic USP planetary transits via Fourier transforms to the light curves of Kepler target stars, increasing

the number of USP planet candidates. Kepler's extended mission K2 [13] continued to amass USP candidates. The most recent mark was in 2018 with the launch of TESS, a mission focused on the brightest stars near Earth and that, unlike Kepler/K2, covers almost the entire sky [14]. One big advantage of this wide-field brightest stars approach is that it makes it possible to have RV follow-up studies on the detected USP planets and obtain mass measurements.

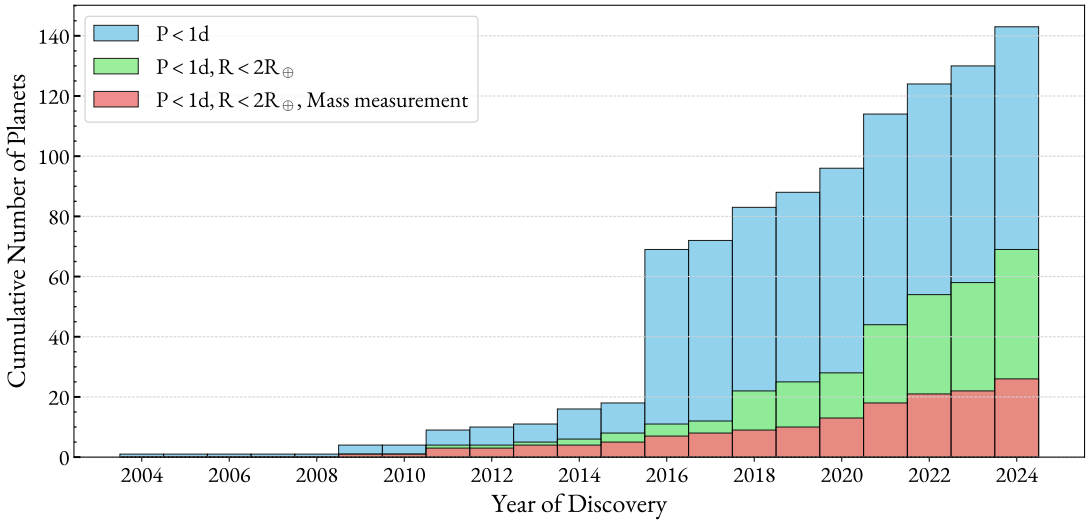


Figure 1.2: Cumulative number of ultra-short period planets detected. USPPs that have radial velocity measurements and thus a mass measurement are in red. Data used to produce this plot sourced from the .

Many open questions remain about this population of planets, one of the most significant being their formation. The orbits of these planets are so small that it is very unlikely that an Earth-mass core formed there, since it is inside the dust sublimation radius [15][16]. Some theories on how these planets came to be at such short distances to their host stars include low-eccentricity orbital migration [17] and origin *in situ* with tidal dissipation [18]. Their close proximity to their stars also causes strong tidal forces that can lead to tidal locking, orbital decay and circularization of the planetary orbits. Furthermore, the intense stellar radiation and gravitational forces USPPs endure result in extreme surface and atmospheric conditions. Initially, it was thought that they lost completely the H-He atmospheres and were bare rocky cores. It has been shown that an incident flux of $F/F_{\oplus} > 650$ in sub-Neptunes can lead to the total evaporation of atmospheres [19]. It could thus be that USP planets are the bare cores of sub-Neptunes, which would then tell us about what lies beyond the atmospheres of the latter. The same study revealed a paucity of hot super-Earths with radius greater than 2 $R_{\oplus}$. This value could represent the maximum radius of the rocky interior of a sub-Neptune, in agreement with theoretical models [20]. On the other hand, there have been discoveries of planets with unusually low densities — TOI-561 b [21] — as well as unusually high ones — K2-229 b [22]. The first case could be explained by a water/low-density envelope while the second has been explained by an iron-core [23]. It is possible that some USPPs retain residual atmospheres, and while interior models are useful for studying the true composition of the planets, there are degeneracies that can only be resolved through atmospheric observations. In this regard, transmission spectroscopy with JWST will

be highly valuable, as well as studying phase curve variations [24]. In any case, precise mass measurements are essential, firstly to get a more accurate bulk density and secondly to infer the atmospheric properties and internal composition [25].

The properties that USP planets present are advantageous to their study. Firstly, the geometric probability for transits increases the shorter the orbital radius (or period) is [26]. So even if USP planets are rare, they are much easier to detect than an Earth-sized planet with a much longer period [27]. Secondly, since they are so close to their host star they also produce bigger radial velocity amplitudes — the RV technique favours shorter periods with the signal amplitude scaling with $P^{-1/3}$. In this way, it is our best bet in order to detect Earth-like planets in other systems with our present-day precision limits. There are currently more than 60 confirmed USP planets according to the NASA Exoplanet Archive, almost all discovered with the transit method. Out of these, almost 30 have both mass and radii measurements (Figure 1.2). Accurate measurements of both are essential to get the planetary density and composition. High precision spectrographs like HARPS that can go to the sub-ms^{-1} domain are needed for detection and mass measurements. Still, the presence of other RV variations, in particular due to stellar activity, can make it challenging to distinguish signals from the star and the planet. A Bayesian framework can help remedy this.

To properly study the ultra-short period planet population, it is crucial to have a homogeneous analysis method. The biggest problem is that planetary parameters are obtained indirectly and depend both on stellar parameters and on the method of dealing with time-correlated noise, specially stellar noise, in the radial velocity data. The stellar parameters are model-dependent, as they are obtained by matching the star's properties to theoretical predictions, and the method of handling time-correlated noise varies considerably between research groups. Discrepancies in the results can arise when different statistical methods are applied to the datasets, particularly when calculating Bayes' factors. Such inconsistencies can lead to unreliable or conflicting interpretations of the USP planet population. In this study, we focus on a sample of 16 USP planets selected for their well-characterized datasets. While this sample size may not be statistically robust on its own, it serves as an important starting point for exploring this population.

1.2 Thesis structure

This work is divided in 5 chapters. In the current chapter, I provided a brief overview of ultra-short period planets, explaining the history of discovery, their unique characteristics and advantages to their study. In **Chapter 2**, the main concepts and techniques used in this research are described. Firstly, the radial velocity (RV) method and the process of extracting the planet-induced signals is explained. This section also discusses stellar activity and its indicators, which help distinguish stellar noise from planetary signals. Additionally, the fundamentals of Bayesian inference are outlined, including concepts like Gaussian Processes (GP), Markov Chain Monte Carlo (MCMC), nested sampling algorithms and model selection techniques. In **Chapter 3**, the methodology section, I describe the instruments used in this study, namely HARPS, HARPS-N, and HIRES, detailing their capabilities for obtaining high-precision RV measurements. I also outline the sample selection process and how the stellar and planetary parameters were determined. This includes an explanation of the workflow and the use of PyORBIT, which can use both MCMC analysis and nested sampling for parameter estimation. In **Chapter 4**, I confirm the validity of using the ESP kernel as a good approximation of the QP kernel and select one model for each system

after assessing the Bayesian evidence, information criteria, and the posterior distributions. I discuss the derived planetary parameters for the USP sample and situate it in a mass-radius plot alongside theoretical planetary model composition curves. In **Chapter 5**, I present the main conclusions of this work.

2

Theoretical background

In this chapter, I provide the theoretical foundation necessary for understanding the key methodologies used throughout this thesis. I begin by introducing the radial velocity method (Section 2.1), covering its basic principles, the process of measuring stellar spectra, and the techniques used to extract precise radial velocities. I then address stellar activity (Section 2.2), describing the types of noise it introduces, along with the indices and indicators that can be used to assess and correct for these effects. Finally, in Section 2.3, I explain the fundamentals of Bayes' theorem, Gaussian Processes, parameter estimation, and model selection, all of which form the basis for the statistical analysis applied in subsequent chapters.

2.1 Radial Velocity Method

2.1.1 Basic Principle

The RV method for detecting exoplanets is conceptually simple. The presence of a planet around a star will cause a periodic reflex motion in the star as they both move around their common centre of mass. The star's spectrum will shift towards shorter wavelengths as it moves towards the observer (blueshift) and larger wavelengths as it moves away from the observer (redshift). By measuring these Doppler shifts, we obtain the radial component of the velocity of the star, from which we can get information on the planet and the orbit.

A stellar spectrum is filled with thousands of absorption lines, produced by ions, atoms, and molecules in the upper layers of the stellar atmosphere absorbing photons. The rest wavelengths of these lines are well-known from atomic physics. By denoting λ_0 as the wavelength of the star's light if the star was at rest and λ as the wavelength of the star's light that undergoes a Doppler shift, we observe a shift in the lines of the stellar spectrum equal to $\Delta\lambda = \lambda - \lambda_0$.

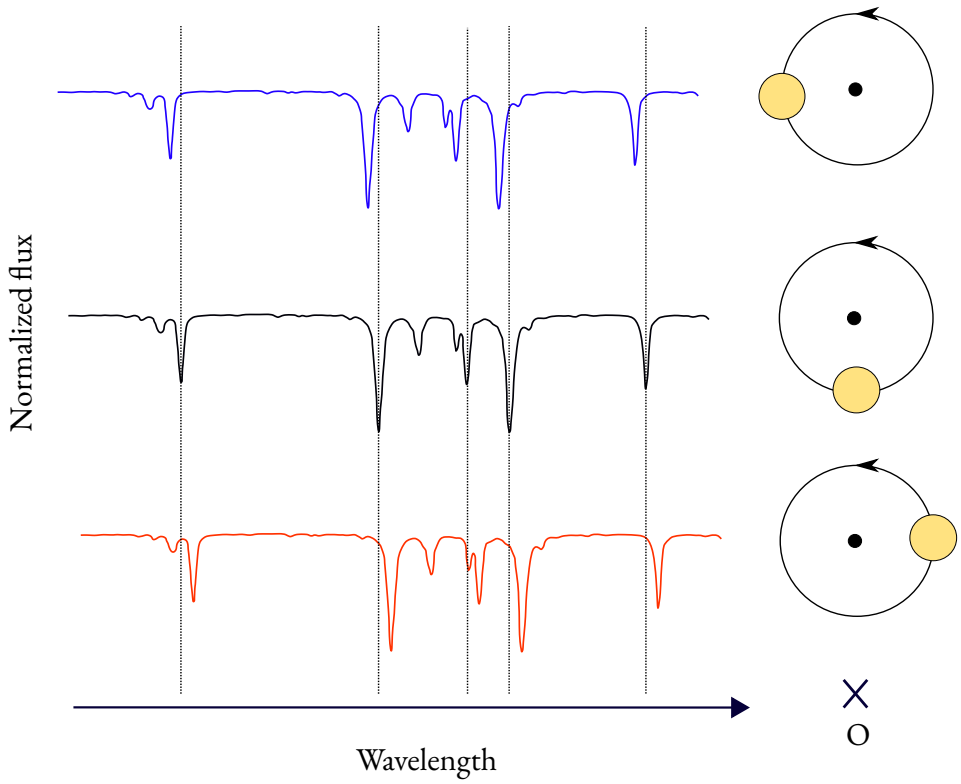


Figure 2.1: Illustration of the Doppler effect on a portion of a stellar spectrum. The spectrum moves towards bluer or redder wavelengths relative to the position of reference spectral lines (grey dashed) as the star orbits around the star-planet centre of mass. The observer is located at the point O.

The wavelength shifts relate to the RV, in a non-relativistic approximation, as

$$\frac{\Delta\lambda}{\lambda_0} = \frac{RV}{c} \tag{2.1}$$

By capturing the star's spectrum at multiple instances, we can track the relative radial velocity of the star over time, denoted as $RV(t)$.

To derive precise radial velocity values from the spectrum, several corrections must be applied, such as wavelength calibration and barycentric correction, which will be discussed later. For now, focusing purely on the gravitational influence, the RV induced by a single planet at any given time is described by the following Keplerian signal:

$$RV_p(t) = K_p \left[\cos(\nu_p(t) + \omega_p) + e_p \cos(\omega_p)\right] . \tag{2.2}$$

Here, K_p is the RV semi-amplitude, ν_p the true anomaly, e_p the orbital eccentricity, and ω_p the argument of peri-

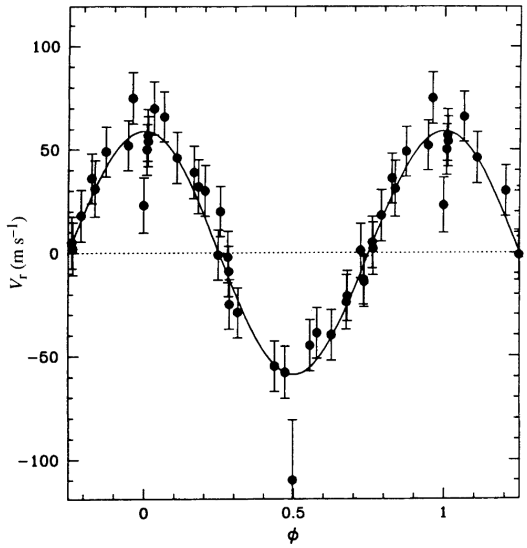


Figure 2.2: The original radial velocity curve of the star 51 Peg that revealed the first exoplanet around another solar-type star. The radial velocity is in function of the orbital phase. Credits: Mayor & Queloz 1995.

astron. The true anomaly is given by

$$\tan\left(\frac{\nu_p(t)}{2}\right) = \sqrt{\frac{1+e_p}{1-e_p}}\tan\left(\frac{E_p(t)}{2}\right). \tag{2.3}$$

To obtain the eccentric anomaly $E_p(t)$, Kepler's equation needs to be numerically solved:

$$E_p(t) - e_p \sin E_p(t) = \frac{2\pi(t - T_{P,p})}{P_p} \equiv M_p(t). \tag{2.4}$$

This equation defines the mean anomaly $M_p(t)$ and it depends on the planet's orbital period P_p and on $T_{P,p}$. This last one, $T_{P,p}$, defines the time of passage at periastron. This leaves us with five observables that can be fit using RV data: e_p, P_p, $T_{P,p}$, ω_p, and K_p.

To generalize the RV measurements to a multi-planet system, the most correct plan of action would be to consider planet-planet interactions which would cause deviations from the Keplerian model. But in practice, for most systems, it is a reasonable approximation to neglect their interactions and simply add the individual Keplerian contributions. We should also add a constant offset and a linear term to account for the stellar barycentric motion with respect to the centre of the galaxy. For a star with n planets the RV measurements will be

$$RV_{Kepler}(t) = c_0 + c_1 t + \sum_{p=1}^{n} RV_p(t). \tag{2.5}$$

The RV variation semi-amplitudes over the course of an orbital period are related to the stellar and planetary

masses, $M_\star$ and M_p,

$$K_p = \left(\frac{2\pi G}{P_p}\right)^{1/3} \frac{M_p \sin i_p}{(M_\star + M_p)^{2/3}} \frac{1}{(1-e_p^2)^{1/2}}, \tag{2.6}$$

where G is the gravitational constant and i_p is the orbital inclination. This angle is measured between the normal of the orbital plane and the observer's line of sight. When i_p is near 0°, the orbit is nearly face-on, making the radial velocity component difficult to detect. Conversely, when i_p approaches 90°, the system is nearly edge-on, maximizing the observable Doppler effect. By using the approximation that $M_p \ll M_\star$, we can get the planetary mass if we know the stellar mass $M_\star$. The stellar mass can be obtained from comparing spectroscopic, photometric and astrometric data with grids of stellar evolutionary models. Note that we do not measure the true planetary mass M_p, only the minimum mass, $M_p \sin i_p$. This is because the angle of projection of the orbital plane is unknown and an edge-on system ($\sin i_p = 1$) is assumed. However, if the planet is transiting we can measure i_p and obtain the true planetary mass.

From Equation 2.6 we can immediately identify the observational biases inherent in RV measurements. Assuming the planetary mass is negligible compared to the mass of the parent star, we have an RV semi-amplitude proportional to $M_p/(M_\star^{2/3} P_p^{1/3})$. This indicates that it is easier to detect high-mass planets that are close to their star (or equivalently, have shorter orbital periods, as per Kepler's third law), particularly around lower-mass stars. To detect an Earth twin at 1 AU we would need to go to $0.1\,\mathrm{ms}^{-1}$ precision while for an Earth twin closer to its star, at 0.1 AU, around $1\,\mathrm{ms}^{-1}$ [28] is sufficient. Current high-resolution spectrographs can reach the latter precision, and this is the reason why many Earth-like planets detected with the RV method have very short periods.

A sample of RV measurements will contain both periodic variations from planetary signals as well as instrumental and astrophysical noise, which are harder to model. The astrophysical noise comes mostly from stellar activity and it is the current limitation for finding small planets.

2.1.2 Measuring the spectrum

To go from light to measurable radial velocities, multiple instrumental steps are required. For a sub ms^{-1} precision in the RVs we need to measure wavelength shifts of around 10^{-5} Å (for the mid-visible $\lambda = 500$ nm). This, in turn, corresponds to miniscule shifts in the width of a spectral line or a fraction of a pixel on a typical CCD detector. The signal-to-noise ratio (SNR) of a single spectral line is too low to reach this precision. For this reason, thousands of spectral lines are used. From an instrumental point of view, a broad spectral range, a high SNR, and high resolution ($R = \lambda/\Delta\lambda$) are required to get very precise RVs.

Firstly, light is collected in the telescope and is directed into the spectrograph either by a slit or by optical fibres. Traditionally, the slit, a narrow opening at the focal plane of the telescope, was used, which allowed a specific portion of the star's light to enter and then be dispersed by a diffraction grating to produce a spectrum. However, the slit presents two disadvantages: there can be variable illumination due to changes in atmospheric conditions or guiding errors, and the position of the slit relative to the spectrograph optics can change. The first affects the stability of the spectral lines while the second leads to a shift in the wavelength calibration, both reducing the precision and accuracy of RV measurements. Consequently, optical fibres are now used in modern high-precision instruments. The fibre is positioned at the focal plane of the telescope, where it collects the starlight

and transmits it to the spectrograph, usually located in a stable environment away from the telescope itself. These fibres scramble the incoming light before injecting it to the spectrograph. Thus, variations in illumination caused by seeing conditions or guiding errors are averaged out. Additionally, since the spectrograph can be located in a highly controlled environment (vibration, temperature, and pressure are all considered), mechanical flexure is minimized. This leads to more stable and accurate RV measurements [29].

When it reaches the spectrograph, the incoming light will be dispersed and the spectrum will be recorded in a detector, usually a Charge-Coupled Device (CCD). For diffraction, échelle gratings are used, a specialized type of diffraction grating. With a second dispersing element, these gratings are able to create a two-dimensional spectrum where each row corresponds to a different diffraction order. What we want is the flux of the star as a function of wavelength so, from the dispersed light in a 2D format, the different orders are extracted, calibrated, and merged into a continuous one-dimensional spectrum.

An important calibration step worth mentioning is the calculation of the wavelength solution. The photons fall as pixels in the detector and we must translate pixel positions into wavelengths and, as such, reference lines are needed. The two main calibration sources are iodine cells and ThAr hollow cathode lamps. The iodine cell technique involves placing a cell filled with iodine gas which imprints a dense set of absorption lines onto the starlight before it enters the spectrograph. The wavelength range is between 5000 Å to 6200 Å. Since these lines are also subjected to the the same instrumental shifts as the starlight, the wavelength scale can be calibrated during the observation itself. A downside is that it later requires sophisticated modelling to separate the stellar spectrum from the iodine spectrum, especially since changes in iodine pressure lead to complex line shapes [29]. Moreover, a range of 1200 Å is quite limited.

Unlike the iodine cell, which imprints lines directly on the star's spectrum, the ThAr technique involves taking separate calibration exposures. The thorium and argon gases in the lamp, when electrically excited, emit light with a well-known set of spectral lines across a wide range of wavelengths. The thorium lines are particularly dense and stable. In this type of calibration, the spectrograph has two channels, a science and reference one. Before observing a star, both channels are illuminated with light from the ThAr lamp while, during observations, only the reference channel is. The two channels should be affected by instrumental drifts, namely variations in the air refraction index and mechanical flexures, in the same way. So the reference channel keeps track of any variation, which is useful for calibration. The calibration is over a broader wavelength range compared to the iodine cell technique, but ThAr lamps have a limited lifetime. Therefore, attention must be paid to see if the lines are moving or broadening over time. Other types of lamps can be used, either one that emits white light such as a tungsten lamp or one with reference emission spectral lines similarly to the mentioned ThAr lamps [30]. Two recent developments have been the Fabry-Perot interferometer and the laser frequency comb. They can be used in conjunction with traditional calibration methods, such as ThAr lamps or iodine cells, to enhance the overall precision and accuracy of spectrograph calibration, like it has been done with HARPS [29].

A barycentric correction is applied as well. The observed wavelengths are in the Earth's reference frame and both the Earth's rotation and movement around the Sun will introduce Doppler shifts. These need to be considered to isolate the true RV of the star. After this correction, the spectrum is in the inertial frame co-moving with the Solar System's centre of mass [31]. The correction is made with precise ephemeris data that describe the Earth's position and motion relative to the Solar System's barycentre. It also accounts for the time it takes light to travel from the star to the Earth, ensuring that the measurements are based on the star's actual position and velocity.

There are additional instrumental effects. Telluric absorption lines coming from the Earth's atmosphere contaminate the spectrum and should be de-trended [28][29]. Others effects like colour variation of the spectrum with airmass and sunlight scattered by the moon are harder to remove but are being studied [30].

2.1.3 Extracting Precise Radial Velocities

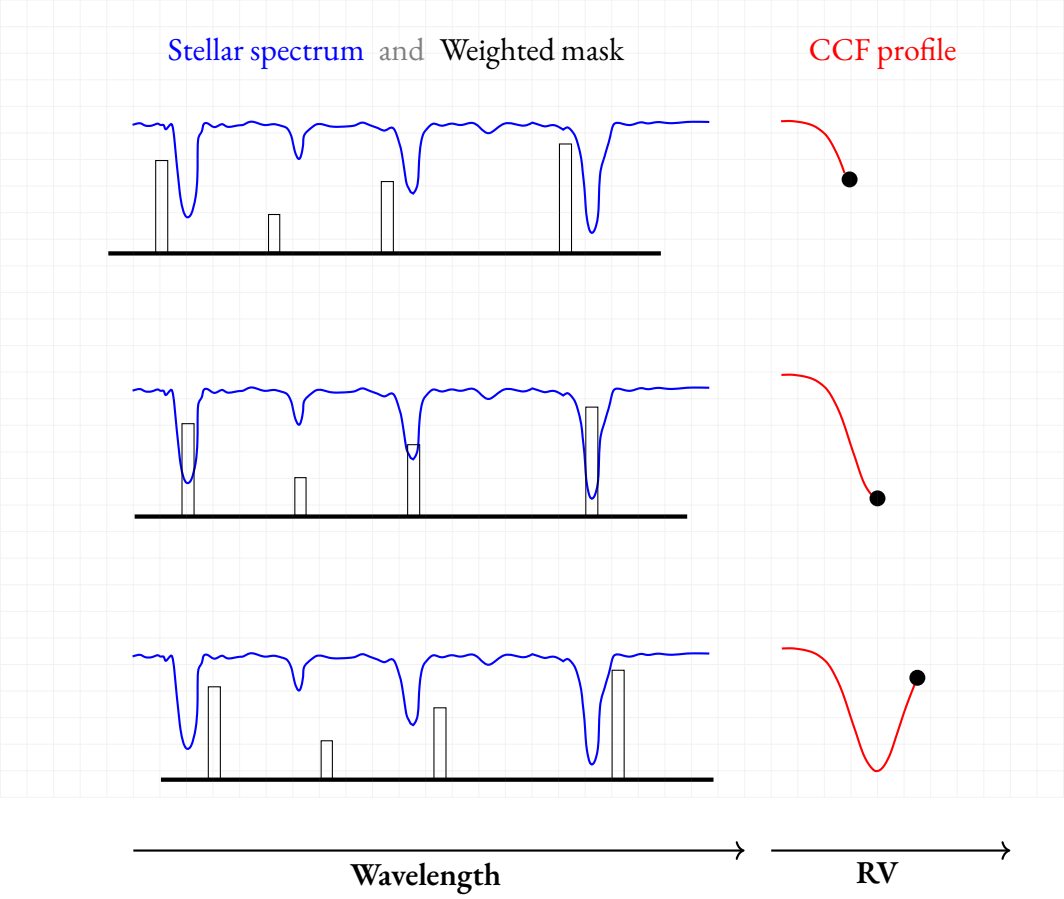


Figure 2.3: Schematic of the CCF method with a weighted mask. On the left, the cross-correlation of the observed stellar spectrum (blue) with a weighted mask (black), with the weights representing the line depths. On the right, the resulting CCF profile. The minimum of the CCF profile, fitted by a Gaussian, is taken as the radial velocity of the star. Real masks are constructed for each spectral type and contain thousands of lines.

There are different ways of extracting the velocity of each spectrum. One way would be to measure the RV of each line and do a weighted average to get the overall stellar RV, but that is computationally intensive. Currently, the most widely used technique consists in cross-correlating the measured spectrum with a mask [32][33]. This numerical mask will be based on where the spectral lines should appear, according to the type of star being observed. The measured spectrum is multiplied by the numerical mask that is Doppler shifted. This process is

repeated for a range of velocity shifts, resulting in a cross-correlation function (CCF) that shows how well the observed spectrum matches the mask at each velocity. This is illustrated in Figure 2.3. The minimum of the CCF, fitted with a Gaussian function, is taken as the radial velocity of the star. In this way, this method can be viewed as computing an average line shape because it effectively combines information from many individual spectral lines into a single, averaged profile:

$$CCF(v) = \sum_i \int Sp(\lambda) M_i(\lambda_v) d\lambda, \tag{2.7}$$

where $\lambda_v = \lambda\sqrt{\frac{1-v/c}{1+v/c}}$, $Sp(\lambda)$ is the spectrum and M is the mask shifted by the velocity v for each absorption line i.

The mask isolates the regions of the spectrum that contribute most to the RV signal, typically focusing on strong, well-defined lines. Lines that are likely contaminated are excluded. A simple binary mask consists of a series of 0 and 1 regions, with the 1 value zones at the wavelengths of the selected spectral lines. It can also be weighted, with larger weight being given to lines with a larger relative depth, as they in principle should contain more information on the radial velocity [33].

Many masks have been constructed for different spectral types such as G2, G7, K5, and M2. However, the CCF method becomes worse for later type stars whose lines are blended. Other ways of extracting the RVs include using a template spectrum and Taylor expanding the observed spectrum as a function of velocity [34] or building a forward model of the spectrum [35][36].

2.2 Stellar activity

Stellar activity can cause deformations on the spectral lines that can be mistaken as Doppler shifts caused by an orbiting planet. It is quite important therefore to be able to understand and mitigate these activity-induced variations to differentiate them from planetary signals. It is especially so in the case of USP which mostly need sub ms^{-1} precision in order to be detected.

2.2.1 Types of Stellar Noise

Different sources of stellar noise or astrophysical noise will cause variations in different timescales [28]. On the shorter term, in the timescale of minutes, we have p-mode oscillations. These are asteroseismic oscillations in main-sequence solar-mass stars and the superposition of these modes can induce RV variations of 1 ms^{-1} or more. Fortunately, by choosing an exposure time that is more than 1 to 2 typical oscillation periods, this source of noise is mostly averaged out. On the range of minutes to about 1 day, there are different phenomena of granulation (granulation, mesogranulation, and supergranulation). They are caused by convective motions of increasingly larger structures in the outer layers of stars with a convective envelope. The convective motion in each cell in the envelope is uncorrelated with the others so the overall signal averages out to about the ms^{-1} scale.

Then, on longer time scales, we have the effect of the magnetic activity. Magnetic fields produce cool spots and bright plages that modify the spectral line profiles as they move across the stellar disk, as illustrated in Figure 2.4. This movement is due to the rotation of the star so these modulations can be quasi-periodic at the rotation period

or its harmonics. This will typically be of the order of days but could also be hours or months depending on the age and spectral type of the star. For example, younger stars typically have shorter rotation periods and larger RV variations making it harder to detect planets around them [37]. Furthermore, stars exhibit differential rotation and the starspots are not permanent features, they appear and disappear over time, which further complicates the observed modulations.

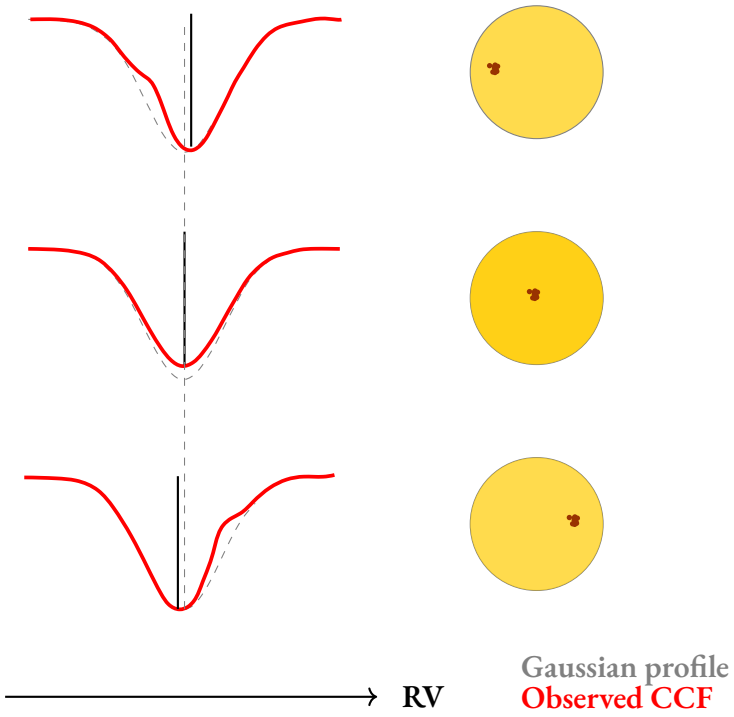


Figure 2.4: Illustration of the effect of a stellar spot on the observed CCF as the star rotates. The blockage of flux by starspots introduces asymmetries in the CCF, leading to variations in the measured radial velocity.

Finally, the longest time scale variations are due to magnetic activity cycles. Similarly to the Sun, which has a activity cycle of around 11 years, stars also experience activity cycles whose period can range from years to decades [38]. In this sense, there will be times of maximum and minimum activity which in turn will make the RV jitter vary.

Some short term activity can be mitigated by the observational strategy but the longer term effects are harder to model [39]. Overall, since the stellar activity signals are incoherent and vary quasi-periodically, there is currently no general analytical model we can use [29]. To identify stellar activity signals in the RV data and disentangle them from exoplanetary reflex motion, we can use activity proxies that are not sensitive to the presence of planets. These can be either generated by chromospheric emissions, such as the S-index, or derived from the shape of the cross-correlation function, which will be analysed in the following subsections.

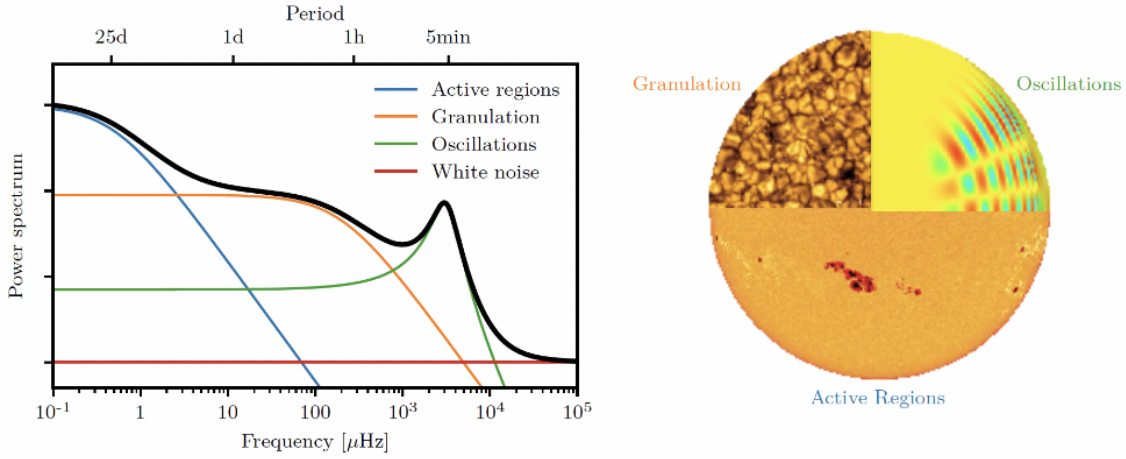


Figure 2.5: Illustration of the timescales associated to stellar activity. The timescales are represented by the power spectrum of the different activity-induced signals. Credits: Faria (2018).

2.2.2 Activity indicators (S-index and $\log R'_{\mathrm{HK}}$)

Certain spectral lines, like the Ca II H & K and the Na I D lines, are recognized as reliable indicators of magnetic stellar activity. This is because in active regions the temperature difference between stellar plages and their surroundings alters the rate of atomic transitions.

The Mount Wilson S-index consists in the Ca II core flux, normalized by the flux in the bandpasses to the red and blue sides of the Ca II lines. Specifically, the K and H lines have a triangular shape with a Full Width at Half Maximum (FWHM) of 1.09 Å and are located at 3933.664 Å and 3968.470 Å, respectively. The V and R continuum passbands have a width of 20 Å and are centered at 3901.070 Å and 4001.070 Å, respectively.

The S-index can be defined as

$$S_{\mathrm{MWO}} = \alpha \frac{N_{\mathrm{H}} + N_{\mathrm{K}}}{N_R + N_V}, \tag{2.8}$$

where N_x represents the counts of the band x, and α is a proportionality constant that equates measurements made by the HKP-2 and HKP-1 spectrophotometer. This value is usually set at 2.40 [40], or 2.3 [41]. Additionally, for spectrographs like HARPS, the value needs to be multiplied by 8 to account for the fact that in Mount Wilson the line core is exposed 8 times more than the continuum bandpasses.

The scale defined by this index became the standard way of measuring chromospheric activity. The S-index contains both photospheric and chromospheric components. Ideally, we would measure solely the chromospheric flux, the one directly related to the magnetic field [42]. Therefore, in order to be able to usefully compare the activity level between stars, we should subtract the photospheric component and normalize the chromospheric flux to the total (bolometric) luminosity of the star. A more useful comparative quantity, the R'_{HK} [43], is defined as

$$R'_{\mathrm{HK}} = R_{\mathrm{HK}} - R_{\mathrm{phot}}. \tag{2.9}$$

The surface flux ratio $R_{\rm HK}$ is defined as:

$$R_{\rm HK} = \frac{N_{\rm H} + N_{\rm K}}{N_{\rm tot}} = \frac{\mathcal{F}_{\rm HK}}{\sigma T_{\rm eff}^4}, \tag{2.10}$$

where $\mathcal{F}_{\rm HK} = F_H + F_K$ is the total flux per cm^2 at the stellar surface in the H and K bands, σ is the Stefan-Boltzmann constant, and T_{eff} is the effective temperature.

The photospheric flux ratio, in turn, is given by

$$R_{\rm phot} = \frac{\mathcal{F}_{\rm phot}}{\sigma T_{\rm eff}^4}, \tag{2.11}$$

where $\mathcal{F}_{\rm phot}$ is the flux in the wings of the H and K lines, where we can neglect the chromospheric contribution.

The surface flux ratio $R_{\rm HK}$ is obtained with a conversion from the S-index while the photospheric flux ratio $R_{\rm phot}$ is given by an empirical relation. Specifically, Noyes et al. 1984 calculated $R_{\rm HK}$ using the equation

$$R_{\rm HK} = 1.34 \times 10^{-4} C_{\rm cf}\, S_{\rm MWO}, \tag{2.12}$$

where $C_{\rm cf} \equiv C_{\rm cf}\big(B - V\big)$ is a colour-dependent conversion factor that corrects flux variations in the continuum passbands and normalizes to the bolometric luminosity. Multiple relations have been proposed in different colour ranges.

The original Middelkoop et al. 1982 [44] relation is valid in the colour range $0.45 < B - V < 1.2$ and is given by

$$\log C_{\rm cf} = 1.13(B-V)^3 - 3.91(B-V)^2 + 2.84(B-V) - 0.47. \tag{2.13}$$

Rutten et al. 1984 [45] extended this relation to $0.3 < B - V < 1.6$ and obtained

$$\log C_{\rm cf} = 0.25(B-V)^3 - 1.33(B-V)^2 + 0.43(B-V) + 0.24. \tag{2.14}$$

For the $R_{\rm phot}$, Noyes et al. 1984 [43] used the relation from Hartmann et al. 1984 [46],

$$\log R_{\rm phot} = -4.898 + 1.918(B-V)^2 - 2.893(B-V)^3, \tag{2.15}$$

valid for $0.44 < (B - V) < 0.82$.

2.2.3 Asymmetry indices (BIS, FWHM, Contrast)

The stellar cross-correlation function between the spectrum and a chosen spectral type mask changes in time. Particularly, we can analyse the bisector inverse span (BIS), the full width at half maximum (FWHM) and the contrast of the CCF to obtain information on the stellar activity. The bisector of a spectral line is constructed by tracing the midpoints of horizontal segments across the line profile, from the line's core to its wings. In an ideal line profile, the bisector has a constant curved shape independent of time but, in the presence of stellar activity, the curvature of this profile changes with time. The BIS value is the difference between the wavelength

corresponding to the midpoint at the continuum (top) and the wavelength corresponding to the midpoint at the maximum absorption (bottom), converted to a velocity. The FWHM represents the width of the line at half the maximum depth and it increases due to the presence of active regions on the stellar surface. The contrast is the depth of the spectral lines relative to the surrounding continuum and changes are related to the overall brightness of the star, which can also fluctuate with activity, for example with the presence of starspots.

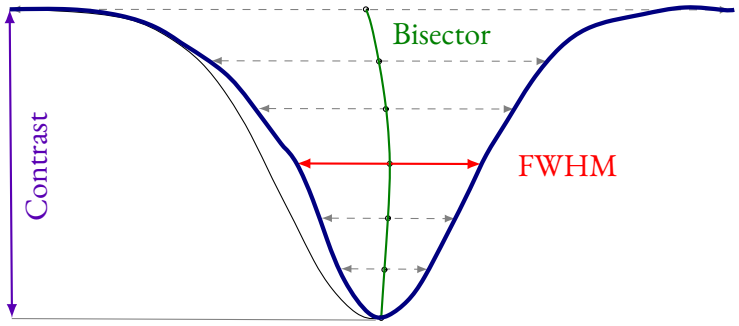


Figure 2.6: Illustration of CCF shape parameters. Bisector (green), full width at half maximum (red) and contrast (purple). The bisector divides the spectral line into two equal parts of equivalent width. The CCF profile gets distorted due to stellar activity and this distortion changes with time depending on the position of the stellar region on the surface of the star (see Figure 2.4). We can use the BIS, FWHM and contrast to measure these variations.

2.3 Bayesian inference

We have seen that radial velocity data can be noisy and influenced by various factors such as stellar activity, instrumental noise, and other planets in the system. Our goal is to refine the planetary parameters of detected USP planets. For that, we need to determine the number of planets that can be confidently detected in each system, and their orbital elements.

One way to extract the planetary signal is to use the (Generalized) Lomb-Scargle periodogram [47][48][49], which shows the dominating frequencies in the dataset. It can be applied iteratively to identify the orbital periods of multiple planets. A drawback is that aliasing may occur, leading to mistaken physical signals [30].

Bayesian inference provides a framework for incorporating prior knowledge and uncertainties, allowing for more reliable interpretation of RV data. The following section covers Bayesian inference methods and how they can be used both to estimate the parameters of a model and to decide between different models.

2.3.1 Bayes' theorem

Starting with some observed data $\mathbf{D}$ and a given model $\mathcal{M}$ conditioned on this data, our goal is to estimate the posterior probability distribution function of a set of parameters $\boldsymbol{\Theta}$. Bayes' theorem can be written as

$$p(\boldsymbol{\Theta}|\mathbf{D},\mathcal{M}) = \frac{p(\mathbf{D}|\boldsymbol{\Theta},\mathcal{M})p(\boldsymbol{\Theta}|\mathcal{M})}{p(\mathbf{D}|\mathcal{M})}. \tag{2.16}$$

Here, $p(\boldsymbol{\Theta}|\mathbf{D},\mathcal{M})$ is the posterior probability distribution of the parameters, $p(\mathbf{D}|\boldsymbol{\Theta},\mathcal{M})$ the likelihood, $p(\boldsymbol{\Theta}|\mathcal{M})$ the prior and $p(\mathbf{D}|\mathcal{M})$ the evidence. We can drop the $\mathcal{M}$ when discussing parameters of only one model. These quantities can also be written as:

$$\begin{aligned} p(\boldsymbol{\Theta}|\mathbf{D},\mathcal{M}) &\equiv \mathcal{P}(\boldsymbol{\Theta}), \\ p(\mathbf{D}|\boldsymbol{\Theta},\mathcal{M}) &\equiv \mathcal{L}(\boldsymbol{\Theta}), \\ p(\boldsymbol{\Theta}|\mathcal{M}) &\equiv \pi(\boldsymbol{\Theta}), \\ p(\mathbf{D}|\mathcal{M}) &\equiv \mathcal{Z}, \end{aligned} \tag{2.17}$$

For parameter estimation of a given model the important quantity is the posterior probability distribution. Since the evidence $\mathcal{Z}$ is independent of the parameters $\boldsymbol{\Theta}$, it can be viewed as a normalization factor and ignored. But as it will be seen in subsequent sections, the evidence is essential for model selection and it will need to be calculated then.

The likelihood function, describing how likely the data is for a certain set of parameters $\boldsymbol{\Theta}$ in a given model $\mathcal{M}$, can be assumed to be Gaussian:

$$\mathcal{L}(\Theta) = \frac{1}{\sqrt{2\,\pi\,|\mathbf{K}|}} \exp\left(-\tfrac{1}{2}(\mathbf{y}-\mathbf{m})^{\mathrm{T}}\mathbf{K}^{-1}(\mathbf{y}-\mathbf{m})\right). \tag{2.18}$$

Here, $\mathbf{K}$ is a covariance matrix. We could assume it is a diagonal matrix, which translates to white uncorrelated noise. However, to account for the contamination from instrumental errors and astrophysical noise, it is better to model these contaminating signals as correlated Gaussian noise using a non-diagonal matrix. In this way, the matrix $\mathbf{K}$ will be specified with a kernel, which gives the correlation between the value of the stellar RV at time t and $t + \Delta t$. The data $\mathbf{y}$ represents the RV time series and the mean function $\mathbf{m}$ is given by the physical model (the sum of Keplerians and an affine function seen in Equation 2.5). The data $\mathbf{y}$ can also be a concatenation of the RV time series with ancillary time series.

The logarithmic likelihood is more convenient to work with and can be written as

$$\ln\mathcal{L}(\Theta) = -\frac{N}{2}\ln 2\pi - \tfrac{1}{2}\ln|\mathbf{K}| - \frac{1}{2}\mathbf{r}^{\mathrm{T}}\mathbf{K}^{-1}\mathbf{r}, \tag{2.19}$$

with $\mathbf{r} = \mathbf{y} - \mathbf{m}$ the vector of residuals and N the number of observations of types considered (N_{RV} or $N_{\mathrm{RV}} + N_{\mathrm{auxiliary}}$)

2.3.2 Gaussian Processes

We can use Gaussian Processes (GP) to represent the RV signal and ancillary data. GPs have been increasingly used in the past few years to mitigate and correct the effects of stellar-induced RV variability [50][51]. Particularly, to properly characterize planets, quasi-periodic GP models have been used [52][53].

A Gaussian Process is a collection of random variables, any finite number of which have a joint Gaussian distribution. Mathematically, we can say that the joint probability distribution over any finite sample $\mathbf{y} = \{y_i\}_{i=1,\ldots,N}$ from the GP is a multi-variate Gaussian,

$$p(\mathbf{y}) = \mathcal{N}(\mathbf{m}, \mathbf{K}), \tag{2.20}$$

where $\mathbf{m}$ is the mean vector and $\mathbf{K}$ the covariance matrix.

The elements of the mean vector and covariance matrix are given by the mean function m and the covariance function k, respectively,

$$m_i = m(\mathbf{x}_i, \theta), \tag{2.21}$$

$$K_{ij} = k(\mathbf{x}_i, \mathbf{x}_j, \varphi), \tag{2.22}$$

where $\mathbf{x}_i$ is the set of independent variables corresponding to the i^{th} sample. The parameters θ and φ of the mean function and the kernel are the hyper-parameters of the GP and can be given a physical meaning.

In this case, as we have an RV time series, there is just one independent variable, time t. The observations are given by $y_i = m(t_i, \theta) + \varepsilon_i$ where ε_i is the noise on the i^{th} observation. The noise vector is what is drawn from the multivariate Gaussian distribution with covariance matrix $\mathbf{K}$. We can rewrite the last equations as

$$m_i = m(t_i, \theta), \tag{2.23}$$

$$K_{ij} = k(t_i, t_j, \varphi) + \delta_{ij}\,(\sigma_i^2 + S^2), \tag{2.24}$$

where δ_{ij} is the Kronecker delta, σ_i^2 represents the reported RV uncertainties (such as the known photon noise or instrumental errors) and S^2 represents an additional unknown noise term called jitter. The mean function represents the expected deterministic planetary-induced radial velocity at any given time (Equation 2.5). The covariance function, which gives the covariance element between any two sample times t_i and t_j, encodes assumptions about the smoothness, periodicity, and variability of the RV data. The choice of this kernel function is very important for the GP model and several kernels have been tested for RV time series data. The kernel choice will be discussed in Subsection 2.3.4.

GPs present several advantages: firstly, they are non-parametric models, meaning they do not assume a specific form for the underlying function. This makes them highly flexible as they can adapt to the specific characteristics of the data, providing better fits than pre-defined models. Different kernels can capture various aspects of the data, such as smooth trends, periodic signals or abrupt changes. Additionally, RV observations are often taken at irregular intervals and can be sparse due to observational constraints. Most times, GPs can interpolate between observations and, in the case of limited data, use the information about the covariance structure of the data to make reliable predictions.

GPs are theoretically robust: the Central Limit Theorem states that the sum of a large number of independent random variables, regardless of their original distributions, tends toward a Gaussian distribution so Gaussian models are often a reasonable approximation.

There are other ways of constructing stochastic process models that are also non-parametric. The reason why GPs are extremely useful for this type of RV analysis, however, is that they make inference easier by allowing for exact marginalization over the function space analytically.

2.3.3 Gaussian Process Models

Trained GP

In a standard Gaussian process model, each dataset, such as the RV measurements and ancillary time series, is described with its own independent covariance matrix, while sharing some hyper-parameters across these datasets. This allows for the independent treatment of noise and variations in each dataset while maintaining commonalities through shared hyper-parameters.

The same kernel can be used to model the periodic signals due to stellar activity across the RV and ancillary datasets. For example, with the quasi-periodic kernel, $P_{\rm rot}$, $P_{\rm dec}$ and $O_{\rm amp}$ are common GP parameters, but the amplitude h is different for each dataset.

Multidimensional GP

In a multidimensional GP framework, there is a joint covariance matrix structure that accounts for the correlations between different datasets. This means that the noise and variations in one dataset can influence the predictions in another dataset. This approach makes a stronger assumption about the relationship between the RVs and the ancillary time series data.

In the original GP framework [54], this relationship is a linear combination of an underlying Gaussian process $G(t)$ and its first derivative $\dot{G}(t)$, allowing for the simultaneous modelling of the RV data and activity indicators:

$$\begin{aligned} \Delta RV &= V_{\rm c}\, G(t) + V_{\rm r}\, \dot{G}(t) && (2.25)\\ \log R'_{\rm HK} &= L_{\rm c}\, G(t) && (2.26)\\ \mathrm{BIS} &= B_{\rm c}\, G(t) + B_{\rm r}\, \dot{G}(t) && (2.27) \end{aligned}$$

These equations were obtained by building upon the $FF^{'}$ framework developed by Aigran et al. 2012 [50]. They presented simple formulas using $F(t)$, the fraction of the visible stellar hemisphere covered in spots, and $F'(t)$, its time derivative, to predict RV variations even when the stellar rotation period is not known. The RV variations caused by spots are modelled considering both the spots' flux and the convective blueshift effects as

$$\Delta RV(t) = \Delta RV_{\rm rot}(t) + \Delta RV_{\rm c}(t). \tag{2.28}$$

The first term comes from the fact that the presence of spots attenuates the flux of a portion of the rotating stellar disk, perturbing the disk-averaged RV. In the most simple case of a simple spot, the RV changes according to the projected area of the spot and the RV of the stellar surface at the location of the spot as

$$\Delta RV_{\rm rot}(t) = -F(t)\; V_{\rm eq} \cos\delta\, \sin\varphi(t) \sin i \tag{2.29}$$

with

$$V_{\rm eq} = \frac{2\pi R_\star}{P_{\rm rot}} \tag{2.30}$$

Here, ∂ is the latitude of the spot relative to the star's rotational equator, $\varphi(t)$ is the phase of the spot relative to the line of sight and i the stellar inclination. $P_{\rm rot}$ can be taken as the equatorial rotational velocity of the star or modelled as a function of the spot latitude.

The second term in Equation 2.28 has to do with the the fact that spots are associated with magnetized areas. In these areas where the magnetic field is stronger convection is partially suppressed, leading to a reduction in the convective blueshift. It can be modelled as

$$\Delta RV_{\rm c}(t) = F(t)\,\Delta V_{\rm c}\,\kappa\,\cos\beta(t), \tag{2.31}$$

where κ is the ratio of the area of the magnetized area to the spot surface, ΔV_c is the difference between the convective blueshift in the unspotted photosphere and the magnetized area, and $\beta(t)$ is the angle between the spot normal and the line of sight.

Although so far only a single spot and associated magnetized area were considered, as a first-order approximation, this formalism can be used to describe the combined effects of multiple active regions in the visible hemisphere. Equation 2.28 is rewritten in Rajpaul et al. 2015 [54] as

$$\Delta RV = V_{\rm r}\,F(t)\,F'(t) + V_{\rm c}\,F^2(t), \tag{2.32}$$

with V_r and V_c as free parameters. Since photometry alone is insensitive to some spot configurations, two activity indicators are added to help constrain the activity induced RV variations. The $\log R'_{\rm HK}$index is related to active regions so it should behave as the second term, of the convective blueshift effect. The bisector span is related to both the convective blueshift effect and the velocity of the active region's surface. This leads to the equations

$$\log R'_{\rm HK} = L_{\rm c}\,F^2(t) \tag{2.33}$$

$$\text{BIS} = B_{\rm r}\,F(t)F'(t) + B_{\rm c}\,F^2(t) \tag{2.34}$$

which are then transformed into Equations 2.25, 2.26 and 2.27 by using the GP variable $G(t) \equiv F^2(t)$.

2.3.4 Kernel choice

Selecting an appropriate kernel is important to accurately capture the underlying patterns in stellar activity. Valid kernel functions can be constructed by adding or multiplying simpler kernel functions [55]. The following section discusses a commonly used kernel and one that approximates it.

Quasi periodic kernel

The quasi-periodic (QP), also called squared-exponential periodic (SEP), is built by multiplying a periodic term, specifically an exponential of a sine-squared function, with a decaying envelope — a squared exponential function. This function is controlled by 4 hyper-parameters and is able to produce signals with different ampli-

tudes, periods, evolution time, and degree of harmonic complexity. It is given by

$$\gamma_{qp}^{(G,G)}(t,t') = \eta_1^2 \exp\left\{ -\frac{\sin^2\left[\pi(t-t')/P\right]}{2\lambda_{\rm p}^2} - \frac{(t-t')^2}{2\lambda_{\rm e}^2} \right\}, \tag{2.35}$$

where P and $\lambda_{\rm p}$ are the period and length scale of the periodic component of the variations, $\lambda_{\rm e}$ is the evolutionary time-scale and η_1 is the amplitude of the covariance function.

This kernel function has some physical motivations, since it implies that the underlying signal should be quasi-periodic. There are also some physical interpretations of the hyper-parameters, namely correlations between P and the stellar rotation period and between the $\lambda_{\rm p}$ length scale and spot evolution time. The $\lambda_{\rm e}$ parameter correlates moderately with the rotation period and spot evolution time [52].

Performing these direct interpretations is not always theoretically sound, primarily because there are degeneracies between the hyper-parameters [54]. However, the QP remains a popular kernel since active stellar regions produce RV variations that can be easily modelled with it.

The ESP approximation

A critical consideration when using GPs is their computational cost. The likelihood computation requires solving a linear system with the full covariance matrix of the measurements. For a dataset of n RV measurements, the covariance matrix of the noise has a size of $n \times n$ and the operations needed for the likelihood calculation, such as inversion and calculating the determinant, scale as $\mathcal{O}(n^3)$. Storing these matrices is also costly, with the memory footprint scaling as $\mathcal{O}(n^2)$. When using RV data and 2 additional spectroscopic time series, such as the BIS and $\log R'_{\rm HK}$, it becomes even worse: the matrix has a size of $3n \times 3n$ and the cost scales as $\mathcal{O}((3n)^3)$. On top of that, sampling methods used to estimate posterior distributions for individual hyper-parameters of interest, like MCMC or nested sampling, require a large number (typically $> 10^4$) of likelihood evaluations. Since multiple models with different configurations for each planetary system will be tested, a less computationally intensive method becomes essential.

One effective way to reduce computational cost is to represent the covariance matrix in a semi-separable form, which allows the use of specialized algorithms designed for sparse matrices, such as banded or block-diagonal matrices. The S+LEAF [56][57] model does just that by introducing a matrix which is the sum of one favoured matrix (S) with a general class of sparse matrix LEAF.

A general symmetric semi-separable matrix S is defined as

$$S = \mathrm{diag}(A) + \mathrm{tril}\left(UV^T\right) + \mathrm{triu}\left(VU^T\right), \tag{2.36}$$

and it can represent a correlated noise model. The rank of the S matrix is $r = 2n_c$, where n_c is the number of components in the correlated noise model:

$$k(\Delta t) = \sum_{s<n_c} \left(a_s \cos(\nu_s \Delta t) + b_s \sin(\nu_s \Delta t)\right) e^{-\lambda_s \Delta t}. \tag{2.37}$$

The LEAF matrices, on the other hand, are a class of matrices that are close to being diagonal (banded, block-diagonal, staircase, etc.). They must be symmetric so

$$F_{ij} = F_{ji} = 0 \quad \text{for } j < i - b_i, \tag{2.38}$$

with b_i the number of non-zero entries left to the diagonal at line i. These matrices are related to the calibration errors. There are periodic calibrations, typically done once per night, and the measurements taken during that time have the same calibration error. This leads to block-diagonal calibration noise matrix where one block represents one night.

The S+LEAF matrix is thus

$$C = \text{diag}(A) + \text{tril}\left(UV^T\right) + \text{triu}\left(VU^T\right) + F. \tag{2.39}$$

Here, A represents the diagonal part of C while U and V are $n \times r$ matrices that represent the symmetric semi-separable part of C, with n the number of points and $r = 2n_c$ the number of components of the semi-separable part. F is the symmetric LEAF part of C with the diagonal of F filled with zeros. b_i is the number of non-zero entries left to the diagonal at line i of F and $\overline{b}$ is the average band width of F. The memory footprint of this model scales as $\mathcal{O}\left(\left(r + \overline{b}\right) n\right)$ and the computational cost as $\mathcal{O}\left(\left(r^2 + r\overline{b} + \overline{b^2}\right) n\right)$.

S+LEAF can also be used for the multidimensional GP framework. S+LEAF2, released in 2022, generalized the S+LEAF method to GPs trained simultaneously on multiple time series. However, to make use of S+LEAF or S+LEAF2, a new kernel needs to be introduced. The quasi-periodic kernel or SEP talked in Subsection 2.3.4, wanted for the way it can model the quasi-periodic signals in RV data, is not semi-separable. A semi-separable representation is needed in order to model the covariance matrix with S+LEAF2 and have faster computations times. Fortunately, there are ways of constructing a semi-separable kernel that reproduces the main characteristics of the SEP kernel.

The SEP kernel can be approximated by

$$k_{\text{SEP}}(\Delta t) \approx k_{\text{SE}}(\Delta t) k_{\text{P}}(\Delta t) \tag{2.40}$$

where k_{SE} is the squared-exponential kernel and k_P is the periodic part, given respectively by:

$$k_{\text{SE}}(\Delta t) = \sigma^2 \exp\left(-\frac{\Delta t^2}{2\rho^2}\right) \tag{2.41}$$

$$k_{\text{P}}(\Delta t) = \frac{1 + f\cos\left(\nu\Delta t\right) + \frac{f^2}{4}\cos\left(2\nu\Delta t\right)}{1 + f + \frac{f^2}{4}} \tag{2.42}$$

Though twice mean square differentiability is not obligatory, it is wanted to generate a smoother model, as we are dealing with G and its derivative $\dot{G}$ in our GP multidimensional framework. We define the exponential-sine (ES) kernel as

$$k_{\text{ES}}(\Delta t) = \sigma^2 e^{-\lambda\Delta t}\left(1 + \frac{1 - 2\mu^{-2}}{3}\left(\cos(\mu\lambda\Delta t) - 1\right)\frac{1 - 2\mu^{-2}}{3} + \mu^{-1}\sin(\mu\lambda\Delta t)\right). \tag{2.43}$$

We can then build upon this last one with a periodic term, constructing the exponential-sine periodic (ESP) kernel. It essentially follows the same structure as the SEP kernel but with an ES term, which is semi-separable, instead of

an SE term, which is not:

$$k_{\mathrm{ESP}}(\Delta t) = k_{\mathrm{ES}}(\Delta t) k_{\mathrm{P}}(\Delta t). \tag{2.44}$$

The ESP kernel is semi-separable since it is the product of two semi-separable terms. It is positive definite with rank 15 and it has been shown to reproduce the SEP kernel very closely. Despite the high rank, it still greatly reduces the computation time compared to the original full covariance matrix way with the SEP kernel [57].

2.3.5 Parameter estimation methods

After defining the model and selecting an appropriate kernel function, we can start to estimate the parameters that best describe the data. The following steps outline the general process of parameter estimation.

First, we select the kernel function along with initial guesses for hyper-parameters, often with incorporated priors. The choice of prior represents the original beliefs about the parameters before looking at the data at hand. The motivations are usually what is physically possible, and what is known from previous data/studies. For instance, uniform priors are used when we want uninformative priors but still need to impose physical boundaries on the parameters. We compute the GP covariance matrix for the data and evaluate the log-likelihood of the model. We then compute the posterior distribution. For this, sampling methods such as MCMC or nested sampling are employed where the samples are used to estimate posterior distributions for individual hyper-parameters of interest.

2.3.6 Markov Chain Monte Carlo

Computing the exact posterior distribution of the elements involves integrating over a high-dimensional space and is analytically intractable. With Markov Chain Monte Carlo (MCMC) algorithms, we can approximate this distribution by generating samples from it without the normalizing constant, the evidence, which is an integral over the parameter space.

The key idea of MCMC is to construct a Markov chain whose equilibrium distribution is the desired target distribution. The Markov property is that the future state depends only on the current state and not on the sequence of events that preceded it. Initial guesses for the parameters are defined, and a jump in parameters is attempted. The move is accepted or rejected based on an acceptance rule. This process continues with subsequent jumps, stopping after a very large number of jumps (e.g., 10^6), ensuring convergence and the removal of burn-in points. Once the chain samples are produced, a histogram of the samples projected into the parameter subspace can be taken to obtain the properties of the distribution of that parameter.

The Metropolis-Hastings method is a simple MCMC method whose acceptance criterion is given by the ratio of the target posterior densities at the new and current sample points multiplied by a a transition distribution [58]. The transition distribution can be defined to allow for asymmetric jumps, which is good for skewed or multimodal data. It is more general than the simple Metropolis algorithm where the transition distribution is factored out and that therefore is only good for data that is not skewed.

Recent advancements in MCMC algorithms have focused on improving convergence speed, which is crucial for reducing computational costs. One such method that has gained popularity is the affine-invariant ensemble

sampling algorithm `emcee` [59][60]. This approach uses an ensemble of "walkers" to explore the parameter space simultaneously. For each walker, a new position is proposed by "stretching" a vector between its current position and a randomly selected walker from the rest of the ensemble. This vector is in the N-dimensional space, N being the number of parameters. The new position is then accepted or rejected based on a probability proportional to the ratio of the posterior densities at the new and current positions. This process is repeated sequentially for every walker in the ensemble. To further speed up the sampling, the ensemble can be divided into two subsets, S_0 and S_1. First, the positions of all walkers in subset S_0 are updated using the positions of walkers in subset S_1. Then, the positions of walkers in S_1 are updated based on the newly updated positions of S_0.

There are multiple reasons why this type of sampler is particularly well-suited for estimating the planetary parameters from radial velocity data. Firstly, there are many correlated parameters in the data (orbital eccentricity and argument of periastron, for example). This would require a careful tuning of the proposal distribution in traditional MCMC methods, but for affine-invariant samplers is not as much of a problem — due to the affine-invariance it can handle distributions with complex geometries. Secondly, the posterior may exhibit a multimodal distribution, for instance a system containing multiple planets could have different possible configurations of planets. In addition, the number of parameters can be very high, especially for models with multiple planets. Affine-invariant ensemble samplers are good at escaping from local minima and exploring multiple modes even at those high dimensions simultaneously.

2.3.7 Model selection

Bayesian inference methods are used to estimate the parameters of a certain model. But they can also provide a powerful framework for model selection, such as models with different covariance matrix structure (due to different kernel functions or frameworks) or with a different number of planets in a planetary system. Unlike the traditional frequentist approach, which focuses on rejecting the null hypothesis (such as the hypothesis that no planets exist based on radial velocity data), Bayesian methods enable the computation of quantitative evidence in favour of different models.

The Bayesian evidence or marginal likelihood is given by

$$\mathcal{Z} = p(\mathbf{D}|\mathcal{M}) = \int \mathcal{L}(\mathbf{\Theta})\pi(\mathbf{\Theta})\mathrm{d}^D\mathbf{\Theta}, \tag{2.45}$$

where D is the number of dimensions of the parameter space. The ratio of the Bayesian evidences of two models is the Bayes factor. To compare two models, the model posterior odds ratio can be calculated,

$$\frac{p(\mathcal{M}_1|\mathbf{D})}{p(\mathcal{M}_2|\mathbf{D})} = \frac{\mathcal{Z}_1}{\mathcal{Z}_2}\frac{p(\mathcal{M}_1)}{p(\mathcal{M}_2)}, \tag{2.46}$$

which is given by the Bayes factor ($B_{12} = \mathcal{Z}_1/\mathcal{Z}_2$) multiplied by the model prior odds ratio. The latter is set to 1 if there is no prior reason for preferring one model over the other. In that case, comparing the posterior odds ratio (POR) or the Bayes factor is equivalent. If the Bayes' factor is bigger than unity then $\mathcal{M}_1$ is preferred over $\mathcal{M}_2$. The larger the value of the Bayes factor the stronger the evidence of model $\mathcal{M}_1$ compared to $\mathcal{M}_2$. If models with varying planet counts are being evaluated, the $n+1$, $n+2$, $n+3$, and so on are compared to the n-planet

model. When a certain threshold is reached, the model is selected. This threshold is given by empirical scales, such as the one suggested by Kass and Raftery (1995) [61], shown in Table 2.1, which is a modification on the one by Jeffreys (1961) [62]. The scale presented uses the natural logarithm of the Bayes factor, which is useful for numerical accuracy reasons since evidences can span multiple orders of magnitude.

$2 \ln B_{12}$	Bayes factor (B_{12})	Evidence Interpretation
0 to 2	1 to 3	Weak
2 to 6	3 to 20	Positive
6 to 10	20 to 150	Strong
> 10	> 150	Very strong

Table 2.1: Interpretation of Bayes factors (Kass and Raftery, 1995).

2.3.8 Information criteria

To evaluate the best model efficiently, besides the Bayes Factor, it is often useful to employ computationally inexpensive methods that balance model fidelity against overfitting. Information criteria, such as the Akaike Information Criterion (AIC) [63] and the Bayesian Information Criterion (BIC) [64], are commonly used for this purpose.

The AIC comes from information theory and it identifies the model that minimizes the expected Kullback-Leibler (KL) divergence [65] between the true data-generating process and the candidate model. It is defined as

$$\mathrm{AIC} = -2 \ln \mathcal{L}(\hat{\theta}) + 2k, \tag{2.47}$$

where $\mathcal{L}(\hat{\theta})$ is the maximized value of the likelihood function under the model and k represents the number of free parameters. Smaller AIC values suggest higher model probability. By minimizing it, you can obtain a balance between data alignment and model complexity. Importantly, this balance does not require the true model to exist within the candidate set, making the AIC useful for analyses where the true model is unknown or misspecified, as in our case.

In practical applications, the AIC is often supplemented by its corrected version, the AICc, which accounts for small sample sizes, given by

$$\mathrm{AICc} = \mathrm{AIC} + \frac{2k(k+1)}{n-k-1}, \tag{2.48}$$

where n is the sample size. The AICc is recommended over the AIC when $n < 40k$ [66]. One potential limitation of the AIC is its fixed penalty term, $2k$, which can favour overly complex models in large-sample scenarios.

An alternative criterion is the BIC, whose penalty term scales with the sample size. It is defined as:

$$\mathrm{BIC} = -2 \ln \mathcal{L}(\hat{\theta}) + k \ln(n). \tag{2.49}$$

As with the AIC, smaller BIC values indicate a more favourable model. The first term encapsulates model fit, favouring models with higher likelihoods, while the second term constitutes a penalty proportional to the model's

complexity. Notably, the $\ln(n)$ term imposes a stricter penalty on complex models as the sample size increases. This aligns with Occam's razor, favouring simpler models in large-sample contexts. In practice, this means the BIC is likely to favour models with fewer planets compared to the AIC, particularly for datasets with a large number of RV observations. The BIC is derived with a lot of assumptions, namely a large sample size. In that case, it is asymptotically equivalent to the logarithm of the Bayes factor.

Both of these information criteria are based off the best-fit location; as a consequence, they might be misleading in cases where the likelihood has a multi-modal shape, as can happen in RV analyses with multiple planets or strong stellar activity [67].

2.3.9 Nested sampling

The Bayesian evidence $\ln(\mathcal{Z})$ is generally considered the most robust and comprehensive method for model selection since it integrates over the entire parameter space and inherently penalizes overfitting by spreading probability mass across parameter values for overly complex models.

The computation of the integral of Equation 2.45 cannot be performed analytically in our cases and it is extremely challenging numerically due to the high dimensionality. One technique is thermodynamic integration, which uses a modified form of MCMC sampling. Nonetheless, an order of 10^6 MCMC samples per chain are needed to determine the evidence, making it about an order of magnitude more costly than parameter estimation [68][69].

Nested sampling algorithms are a numerical approach recently developed to compute the Bayesian evidence [70]. Since it produces samples from the posterior PDFs of model parameters for the integral estimation, it can also be used for parameter estimation.

NS algorithms refactor the multidimensional integral of the evidence over position $\boldsymbol{\Theta}$ into a one dimensional integral taken over the prior volume X of the enclosed parameter space. Starting from equation

$$\mathcal{Z} = \int \mathcal{L}(\boldsymbol{\Theta})\pi(\boldsymbol{\Theta})\mathrm{d}^D\boldsymbol{\Theta} \tag{2.50}$$

we define a prior volume $X(\lambda)$ where the likelihood is above a certain threshold λ:

$$X(\lambda) = \int_{\mathcal{L}(\boldsymbol{\Theta})>\lambda} \pi(\boldsymbol{\Theta})\mathrm{d}^D\boldsymbol{\Theta}. \tag{2.51}$$

The priors are normalized and they need to be integrable. We start with the full volume $X = 1$ when $\lambda = 0$ and as we increase the value λ, the volume shrinks, eventually reaching 0 in the case that the maximum-likelihood value $\mathcal{L}_{\max}$ is a singular point.

The evidence integral then becomes

$$\mathcal{Z} = \int_0^\infty X\mathrm{d}\lambda, \tag{2.52}$$

or, rewritten in terms of $\mathcal{L}(X)$, which is the iso-likelihood contour defining the edge of the prior volume X,

$$\mathcal{Z} = \int_0^1 \mathcal{L}(X)\mathrm{d}X. \tag{2.53}$$

$\mathcal{L}(X)$ is the inverse of $X(\lambda)$, and it decreases monotonically as a function of X. We can determine $\mathcal{L}_i$ at a discrete number N of points X_i and approximate the evidence numerically, for example with a Riemann sum,

$$\mathcal{Z} = \sum_{i=1}^{N} \mathcal{L}(\Theta_i) \times w_i, \tag{2.54}$$

where

$$w_i = \frac{1}{2}(X_{i-1} - X_{i+1}) \tag{2.55}$$

with $0 < X_N < \ldots < X_1 < X_0 = 1$.

Once the evidence $\mathcal{Z}$ is estimated, the posterior distributions can also be estimated using the generated points,

$$p_j = \frac{\mathcal{L}_j w_j}{\mathcal{Z}}. \tag{2.56}$$

The likelihood at a given position $\mathcal{L}(\Theta)$ is simple to obtain, while for the associated prior volume $X(\Theta)$ only noisy estimates can be obtained. In particular, we need to generate samples Θ_{i+1} from the prior distribution $\pi(\Theta)$ with increasing likelihoods $\mathcal{L}(\Theta_{i+1}) > \mathcal{L}(\Theta_i)$. For this, N "live points" are drawn, from which the point with the smallest likelihood is removed and replaced by a new point with a likelihood $\mathcal{L} > \mathcal{L}_0$. This will make the prior volume shrink at each iteration into smaller and smaller nested shells of increasing likelihood. When a certain precision in the evidence is reached the algorithm stops.

The most challenging part of the process is guaranteeing the condition of increasing likelihood ($\mathcal{L}(\Theta_{i+1}) > \mathcal{L}(\Theta_i)$) at every step. One approach is ellipsoidal nested sampling proposed by Mukherjee et al. 2006 [71]. Ellipsoidal nested sampling consists in approximating the iso-likelihood contour by a D-dimensional ellipsoid determined from the covariance matrix of the current set of active points. Then, new points are selected from the prior within this ellipsoidal bound until one meets the criterion of increasing likelihood. The MultiNest python package [69] adds some improvements to this approach. Firstly, the parameter space is transformed into a D-dimensional unit hypercube from which samples are drawn uniformly, which is easier to do than to sample from an arbitrary prior. Plus, at each step of sampling, active points are partitioned into clusters which are then enclosed in ellipsoids. This is much more efficient for highly degenerate and multimodal problems. However, this algorithm maintains a constant number of live points at all times and as such the rate of posterior integration is constant.

Another algorithm developed recently is that of `dynesty` [72], which instead uses a dynamical number of live points. Like MultiNest, `dynesty` generates samples from a prior normalized to the volume, this normalization being done to increase the acceptance rate since the volume continuously decreases at each step. It also samples from a D-dimensional unit cube, with D the number of parameters $\boldsymbol{\Theta}$. Therefore, the original problem of sampling from the posterior directly is likewise transformed into one of repeatedly sampling uniformly within the transformed constrained prior:

$$\pi'_\lambda(\Phi) \equiv \begin{cases} 1/X(\lambda) & \mathcal{L}(\Theta = \mathcal{T}(\Phi)) \geq \lambda \\ 0 & \text{otherwise.} \end{cases} \tag{2.57}$$

The main difference is that the number of live points vary during the runtime. Specifically, it allows the number of live points, which depend on the prior volume, $K(X)$, to follow an importance function $K(X) \propto \mathcal{I}(X)$, given by

$$\mathcal{I}(X) = f^{\mathcal{P}}\mathcal{I}^{\mathcal{P}}(X) + (1 - f^{\mathcal{P}})\mathcal{I}^{\mathcal{Z}}(X). \tag{2.58}$$

The posterior importance function $\mathcal{I}^{\mathcal{P}}(X)$ is defined to be the probability density function of the importance weight which is used to estimate the evidence integral:

$$\mathcal{I}^{\mathcal{P}}(X) \equiv \hat{p}_i \equiv [\mathcal{L}(\Theta_{i-1}) + \mathcal{L}(\Theta_i)] \times \left[\hat{X}_{i-1} - \hat{X}_i\right]. \tag{2.59}$$

The evidence importance function is

$$\mathcal{I}^{\mathcal{Z}}(X) \equiv \frac{1 - \mathcal{Z}(X)/\mathcal{Z}}{\int_0^1 (1 - \mathcal{Z}(X)/\mathcal{Z})\mathrm{d}X}. \tag{2.60}$$

Giving a bigger relative amount of importance on estimating the posterior $f^{\mathcal{P}}$ translates into allocating more live points in regions where the posterior mass $\mathcal{L}(X)\mathrm{d}X$ is higher. On the other hand, a bigger importance given to $\mathcal{I}^{\mathcal{Z}}(X)$ translates into allocating more live points when we have not integrated over much of the posterior and are thus less confident in the estimated value of $\mathcal{Z}$. We do not have the actual values of X nor of $\mathcal{I}(X)$, only noisy estimators which come from generating samples from the posterior. The `dynesty` algorithm starts with performing a static baseline run to obtain the location of the posterior mass as well as an evaluation of the importance function. More points are allocated where $\mathcal{I}_i$ is larger. These new points are merged into the previous sample set and the steps are repeated until the stopping criterion is reached. The stopping criterion now instead of being just about the evidence estimate is a criterion based on both the uncertainty in the current posterior and on the evidence estimates. It is given by

$$\mathcal{S} = s^{\mathcal{P}}\mathcal{S}^{\mathcal{P}} + (1 - s^{\mathcal{P}})\,\mathcal{S}^{\mathcal{Z}} < \varepsilon, \tag{2.61}$$

where ε is the chosen tolerance, $\mathcal{S}^{\mathcal{P}}$ is the posterior stopping criterion, $\mathcal{S}^{\mathcal{Z}}$ is the evidence stopping criterion, and $s^{\mathcal{P}}$ is the relative amount of weight given to each one of these.

`dynesty` implements five bounding methods: unit cube, single or multiple ellipsoids, overlapping balls or overlapping cubes. By default, `dynesty` uses multiple ellipsoids to construct the bounding distribution, which, again, is more effective for multimodal distributions. Yet, they are more conservative than the MultiNest approach when separating live points into clusters. To generate samples the options are uniform sampling, random walks, multivariate slice sampling, and Hamiltonian slice sampling. The chosen method is given by the number of dimensions of the problem since each one works better in different regimes.

3 Methods

In this chapter, I discuss the methodology used for the determination of the USP planets' parameters. I begin by describing the instrumentation that was used to collect data on the targets, specifically the HARPS, HARPS-N, and HIRES spectrographs. Following this, I detail the criteria for sample selection and how the stellar parameters of this sample were re-analysed in order to ensure consistency in our dataset. The updated stellar parameters of the sample are presented. I then discuss the sources and methods to gather the radial velocity and stellar activity data. Finally, I explain the planetary parameters' derivation, including the workflow in `PyORBIT` and the sampling methods used.

3.1 Instrumentation

3.1.1 HARPS

The High Accuracy Radial velocity Planet Searcher or HARPS is a high-precision échelle spectrograph installed at the 3.6-meter telescope at the La Silla Observatory in Chile [73]. It was designed to achieve a radial velocity precision of 1 ms^{-1}, enabling the detection of exoplanets with masses similar to that of Earth. It has a spectral resolution of R $\sim$ 115000 and a wavelength range of 3780 - 6910 Å. The spectrograph is housed in a temperature-controlled (17°$\pm$0.01°) vacuum vessel to minimize environmental variations [29].

3.1.2 HARPS-N

HARPS-N is essentially a northern hemisphere counterpart to HARPS, located at the 3.58-meter Telescopio Nazionale Galileo (TNG) in La Palma, Canary Islands [74]. It shares many design features with HARPS but is optimized for observing from the northern hemisphere. It has a spectral resolution of R $\sim$ 115000 and wavelength

coverage between 3830 and 6900 Å. It is also housed in a temperature-controlled vacuum vessel for extreme stability. Installed in 2012 and with levels of precision as good or slightly better than HARPS [75], it has since become a critical tool for exoplanet research in the northern hemisphere.

Both HARPS and HARPS-N utilize thorium-argon (ThAr) lamps for wavelength calibration. The ThAr lamps produce a rich spectrum of lines that are used to precisely determine the instrument's wavelength scale, ensuring that the radial velocity measurements remain accurate over long periods, as described in Subsection 2.1.2. A Fabry-Pérot is used for simultaneous calibration for stars brighter than $V = 11$.

3.1.3 HIRES

The HIgh Resolution Échelle Spectrometer (HIRES, [76]) is located at the 10-meter Keck I Telescope. Its wavelength range is between 3000 and 11000 Å and it can reach a maximum resolution of about 85000. HIRES was majorly upgraded in 2004 with a new CCD and other optical improvements [77] so now it can reach a maximum precision in radial velocities of about 1 ms^{-1} [78]. Differently than HARPS and HARPS-N, it uses an iodine reference cell for wavelength calibration. It is not in a vacuum enclosure but the temperature is between 0° and 2°, making it a stable instrument [29].

3.1.4 CARMENES

CARMENES (Calar Alto high-Resolution search for M-dwarfs with Exoearths with Near-infrared and optical Échelle Spectrographs) was built for the 3.5m telescope at the Calar Alto Observatory. It consists of two separated spectrographs, one of visible-light covering the wavelength ranges from 5200 to 1050 Å and another for the near-infrared from 9500 to 1710 Å. Each channel has a spectral resolution of 94600 and 80400, respectively [79]. The spectrographs are housed in vacuum tanks and simultaneous calibration is performed with an emission-line lamp or with a Fabry-Pérot etalon. In this manner, it is possible to reach 1 ms^{-1} radial velocity precision. The CARMENES first light with the two NIR and VIS channels working simultaneously occurred in November 2015 [80].

3.2 Sample selection

The first step in performing the homogeneous analysis is to select a sample of USP planets, according to their definition of $P < 1$ day and $R < 2\ R_{\oplus}$. Bright targets were needed since those are the ones that have had spectroscopic radial velocity measurements so I selected stars with a magnitude in the Johnson system of $V < 14$. Moreover, I looked for planets that were observed with either HARPS, HARPS-N or ESPRESSO since these spectrographs provide activity indices that are needed for the GP methods of using three spectroscopic series simultaneously. All of these constraints were applied to the NASA Exoplanet Archive*.

*`https://exoplanetarchive.ipac.caltech.edu/cgi-bin/TblView/nph-tblView?app=ExoTbls&config=PS`

Some of these targets were M-dwarfs which had very high error bars in the activity indices. As was discussed in Rainer et al. 2020 [81], the online HARPS DRS pipeline uses a M2 mask which results in strangely shaped CCF's bisectors without much physical meaning. To analyse M-dwarfs some works use template-matching algorithms like TERRA or serval [34][36], which can provide better RV measurements but they do not provide activity indicators such as the bisector span or the FWHM. For that reason, I decided to exclude M-dwarfs. The threshold between the K and the M stars was done with help of the table from Pecaut and Mamajek 2013 [82].

The planet TOI-1416 b with a period of slightly above 1 day (1.067 days) technically did not belong in the list. The 1 day cutoff, however, is somewhat arbitrary as up to 10 days there is no sharp distinction in the planetary occurrence rate or mean metallicity of the host stars [83]. Thus, I kept the planet as it satisfied the other criteria.

The final sample is made of 16 systems with USPPs with mass measurements.

3.3 Stellar Parameters Re-analysis

After selecting the sample, I needed to collect the stellar parameters since they are extremely important to deduce the planetary parameters. For example, the planetary mass directly depends on the stellar mass as $M_p \propto M_\star^{2/3}$. A homogeneous analysis of the stellar parameters will then also help in the homogeneity of the final planetary parameters.

To determine the stellar mass and radius I used the `isochrones` python package (version 1.2.2, [84]) with a posterior sampling performed by MultiNest [68][69]. As inputs, I included the effective temperature, surface gravity and metallicity ($T_{\rm eff}$, $\log(g)$, [Fe/H]). These values come from literature, mostly from HARPS, HARPS-N, ESPRESSO, but also from other spectrographs (CARMENES, HIRES, APF, etc.) given that reliable stellar parameters such as $T_{\rm eff}$ can often be obtained even from spectrographs with moderate resolution. Where literature provided only two of the three parameters, priority was given in decreasing order to $T_{\rm eff}$, [Fe/H], and $\log(g)$, allowing for inclusion as long as $T_{\rm eff}$ was available. To avoid biasing the analysis toward methods with artificially low internal errors, I standardized uncertainties to 60 K for $T_{\rm eff}$, 0.1 for $\log(g)$, and 0.03 for [Fe/H] when the reported errors were smaller than these values, as was done in Sousa et al. 2011 [85]. Otherwise, I retained the original values. In addition to these three photospheric parameters, I included the parallax of the target from the newest Gaia DR3 catalogues [86] [87] and multiple magnitudes in several bands, since `isochrones` also performs a Spectral Energy Distribution (SED) fit. These were the AAVSO Photometric All-Sky Survey B and V magnitudes [88], the Two Micron All Sky Survey J, H and K magnitudes [89] and, where available, the Wide-field Infrared Survey Explorer W1, W2, W3 and W4 magnitudes [90]. To perform the isochrone fitting, a library of theoretical isochrones is needed. Two stellar models were used: MESA Isochrones and Stellar Tracks (MIST, [91][92][93]) and the Dartmouth Stellar Evolution Database (DART, [94]). The identifying information and input parameters used for the analysis are shown in the Appendix, Table A.1.

I varied the spectroscopic parameters ($T_{\rm eff}$, $\log(g)$, [Fe/H]), with each set derived from a different source in the literature, and tested two stellar models (MIST and DART). Each of these ensembles was run with 1000 live points. The stellar age was also constrained to be between 0.1 and 13.7 Gyr. In some systems, the stellar age has been determined through independent methods (e.g., gyrochronology), so I applied a Gaussian prior to the age to refine the fit. Both MIST and DART models produced consistent results, so I utilized both sets. I then combined all the posteriors and extracted the final values and errors by taking the median and the 16th and 84th percentiles.

Combining posteriors from different sources—each using slightly varied methodologies—results in a final uncertainty slightly larger than those from individual fits, yet likely improves overall accuracy. In this way, the values and errors of radius, mass and density for each star, were obtained (Table 3.1). While stellar age estimates were also calculated, they generally show a larger spread due to the challenges in constraining age precisely. Figure 3.1 shows the combined posterior distributions of the mass and radius and for the age and density for star Kepler-78. Similar graphs were obtained for the rest of the targets.

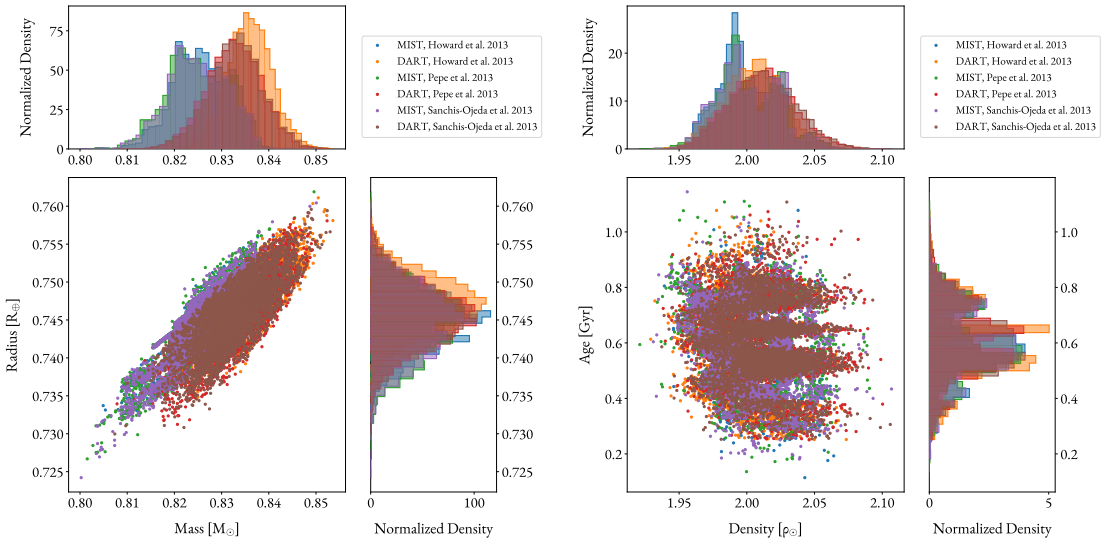


Figure 3.1: Posterior distributions of mass and radius (left) and of the density and age (right) for Kepler-78 from isochrone fitting. Each colour corresponds to a certain stellar evolution model (DART or MIST) with input parameters from a certain literature source. The combined value for the mass is $0.829 \pm 0.007\, M_{\odot}$ and for the radius $0.745 \pm 0.004\, R_{\odot}$. The combined value for the density is $2.00 \pm 0.02\, \rho_{\odot}$ and for the age 0.60 ± 0.13 Gyr.

In Figure 3.2, I show the mass and radius of the sample of stars with values from literature and the values derived from combining the posteriors from the isochrone fits. We can see that the uncertainties for both the stellar masses and radii decreased considerably.

In Figure 3.3, the spectral type and metallicity of the sample of stars is shown.

Table 3.1: Derived stellar parameters of the sample from isochrone fitting.

System	M_s [$M_\odot$]	R_s [$R_\odot$]	$T_{\rm eff}$ [K]	[Fe/H] [dex]	Age [Gyr]	Density [$\rho_\odot$]	Spectral type	Spectroscopic Reference
TOI-500	0.764 ± 0.009	0.717 ± 0.005	4614 ± 37	0.18 ± 0.06	$5.0^{+0.2}_{-0.2}$	2.071 ± 0.025	K	[95]
Kepler-78	0.830 ± 0.007	0.745 ± 0.004	5110 ± 38	-0.10 ± 0.05	$10.6^{+1.5}_{-1.3}$	2.004 ± 0.023	K	[96], [97], [12]
TOI-561	0.778 ± 0.022	0.839 ± 0.005	5474 ± 47	$-0.33^{+0.08}_{-0.05}$	$12.1^{+1.2}_{-2.2}$	1.317 ± 0.045	G	[98], [99], [21]
TOI-1807	0.776 ± 0.007	0.700 ± 0.004	4782 ± 28	-0.05 ± 0.03	$0.29^{+0.07}_{-0.05}$	2.265 ± 0.022	K	[37]
K2-229	0.852 ± 0.019	0.776 ± 0.005	5241 ± 37	-0.09 ± 0.03	$1.82^{+2.04}_{-1.05}$	1.832 ± 0.057	K	[22], [100], [101]
TOI-431	0.828 ± 0.020	0.752 ± 0.005	4859 ± 41	0.15 ± 0.04	$1.65^{+3.98}_{-1.17}$	1.949 ± 0.070	K	[102]
TOI-1416	0.817 ± 0.004	0.784 ± 0.004	4986 ± 27	0.07 ± 0.04	$6.90^{+0.10}_{-0.10}$	1.697 ± 0.021	K	[103]
Kepler-10	0.930 ± 0.023	1.087 ± 0.009	5687 ± 72	$-0.08^{+0.10}_{-0.07}$	$9.48^{+0.65}_{-1.02}$	0.726 ± 0.026	G	[104],[105], [11]
CoRoT-7	0.918 ± 0.012	0.832 ± 0.005	5687 ± 72	$0.08^{+0.08}_{-0.05}$	$1.78^{+0.74}_{-0.70}$	1.596 ± 0.027	G	[106], [107], [105], [108], [6]
HD 3167	0.877 ± 0.034	0.866 ± 0.007	5375 ± 66	$0.01^{+0.06}_{-0.09}$	$7.03^{-2.73}_{+2.64}$	1.349 ± 0.074	G	[109], [110], [100], [101], [111], [112], [113]
K2-141	0.727 ± 0.022	0.706 ± 0.006	4534 ± 71	$0.11^{+0.16}_{-0.10}$	$10.18^{2.60}_{-5.51}$	2.056 ± 0.068	K	[114], [115]
HD 80653	1.170 ± 0.031	1.208 ± 0.008	5996 ± 49	$0.26^{+0.05}_{-0.06}$	$2.74^{0.83}_{-0.86}$	0.663 ± 0.025	F	[116]
K2-106	0.939 ± 0.039	0.971 ± 0.009	5584 ± 73	$0.05^{+0.07}_{-0.05}$	$7.22^{+2.62}_{-2.65}$	1.025 ± 0.064	G	[117],[110], [100], [101], [118], [119], [120]
HD 213885	1.023 ± 0.027	1.144 ± 0.008	5953 ± 49	-0.04 ± 0.03	$5.79^{+0.99}_{-0.96}$	0.683 ± 0.027	F	[121]
HD 20329	0.951 ± 0.019	1.111 ± 0.006	5702 ± 42	$0.01^{+0.03}_{-0.08}$	$9.04^{+0.80}_{-0.91}$	0.696 ± 0.020	G	[122]
K2-131	0.812 ± 0.030	0.755 ± 0.006	5133 ± 55	-0.13 ± 0.06	$3.33^{+4.44}_{-2.41}$	1.895 ± 0.087	K	[117], [123], [100], [101]

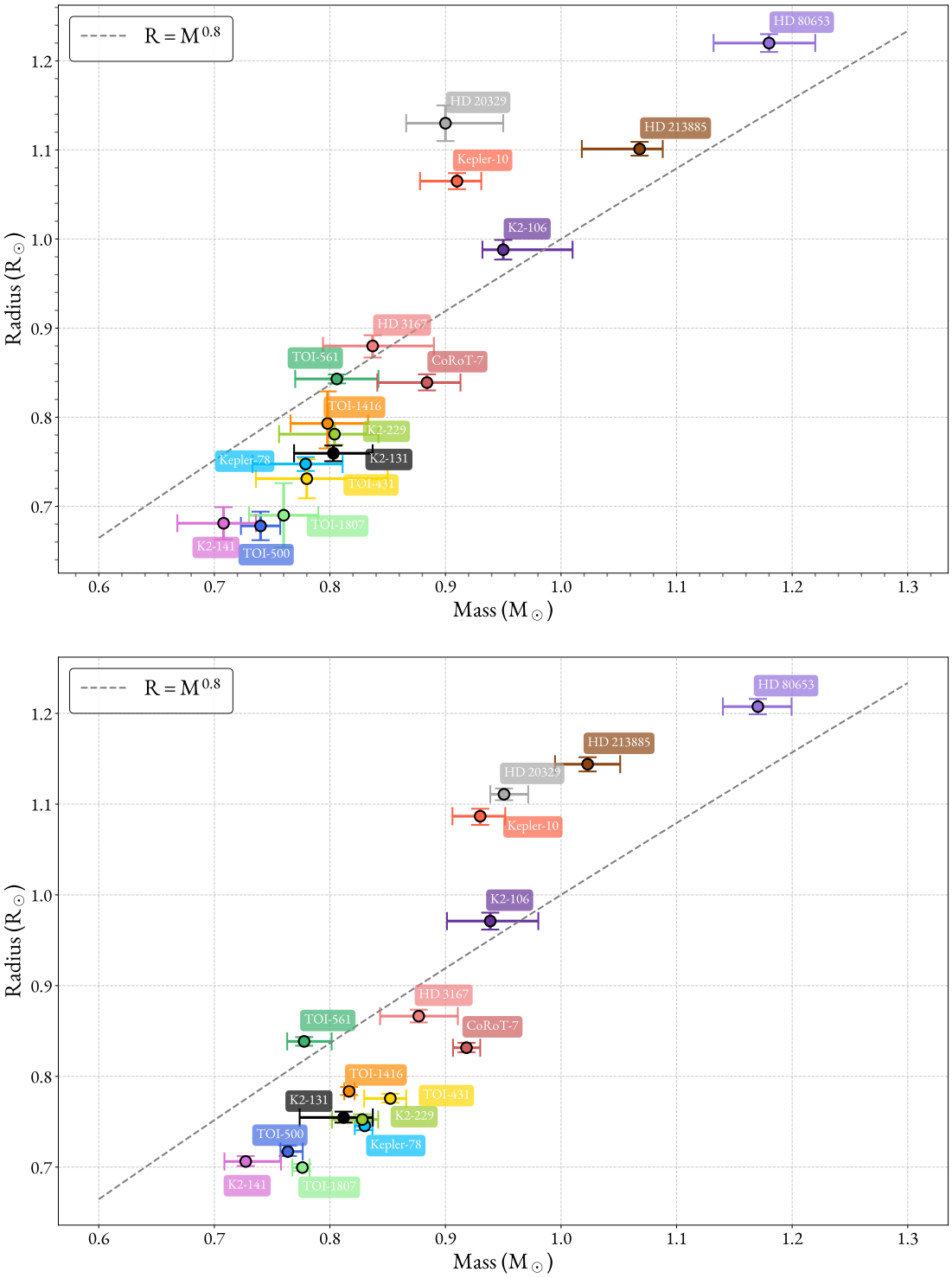


Figure 3.2: Mass-radius plot of the sample of stars with literature values (top) and the derived values from this analysis (bottom). A simple empirical line representing the mass-radius relation for low-mass main-sequence is included for reference. References for the literature plot: Serrano et al. 2022 [95], Bonomo et al. 2023 [117], Lacedelli et al. 2022 [98], Nardiello et al. 2022 [37], Dai et al. 2019 [15], Osborn et al. 2021 [102], Deeg et al. 2023 [103], Espinoza et al. 2020 [121], Murgas et al. 2022 [122]

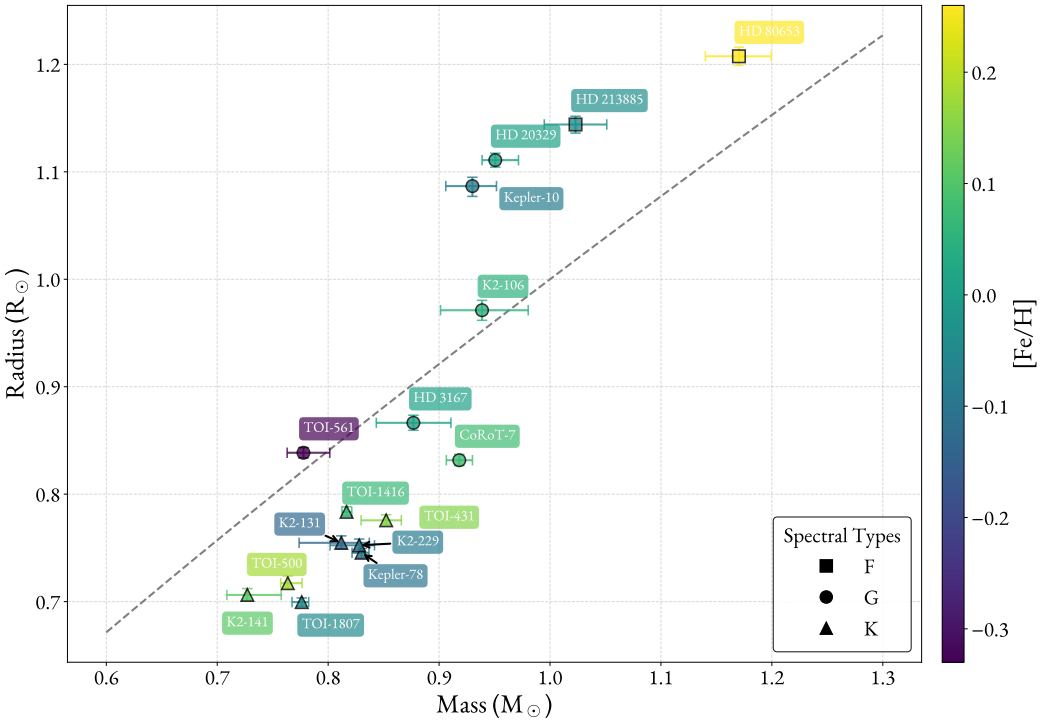


Figure 3.3: Mass-radius plot of the sample of stars with the derived values from this analysis colour-coded according to the star's metallicity. The different shapes represent the stellar type. A simple empirical line representing the mass-radius relation for low-mass main-sequence is included for reference.

3.4 Collecting RV and Stellar Activity Data

For most systems, I took the the radial velocity data and the activity indicators (BIS, FWHM and S-index or $\log R'_{\rm HK}$) directly from reference papers in case they had used the HARPS DRS or equivalent softwares.

The raw spectra obtained with HARPS are automatically processed using the Data Reduction Software pipeline developed by the HARPS Consortium, and this has been the case since operations began in 2003. It consists of bias and flat field corrections, wavelength calibration and calculation of the cross correlation function from which the radial velocities and stellar activity indicators are extracted. In 2015, there was a major upgrade to the fibre link between the telescope and the spectrograph which led to a slight shift in the RV zero point. Accordingly, there was a recalibration done to the data reduction; consequently, data obtained before 29 May 2015 is processed with version 3.5 of the pipeline while data obtained after that is processed with version 3.8 of the pipeline. For this reason, datasets from before and after that date should have different offsets.

The first public data release of 20 years of HARPS Radial Velocity catalogue was made on December 12, 2023. The data is accessible through the ESO Archive Science Portal (version 2.3.3). Within this portal, users can filter the dataset based on several parameters, such as target name, instrument used and observation data. After selecting

and downloading the filtered files, it is possible to extract the wanted variables, for example, the Barycentric Julian Date (HIERARCH DRS BJD), Radial Velocity (HIERARCH DRS CCF RVC), Cross-Correlation Function Mask (HIERARCH DRS CCF MASK), and many others.

Regarding HARPS, for some systems there were observations done after the papers' release, and subsequently new data in the ESO Archive Portal. For those, I used the `ACTIN2` python tool combined with `pyrhk`[124][125]. `ACTIN2` can easily extract data from the different spectrographs' FITS files and calculate spectroscopic activity indices for different lines. In particular, I used the Ca II H and K lines to calculate the S-index and converted it to $\log R'_{\mathrm{HK}}$ with the `pyrhk` python package. The Rutten relation (Equation 2.14) was used since the colour range where it is valid applied to our sample composed of FGK stars.

For certain HARPS-N targets, I utilized the Data & Analysis Center for Exoplanets (DACE, [126]), a web platform developed by PlanetS and hosted at the University of Geneva. DACE provides access to both public and private observational data, including radial velocities. The HARPS-N data from Guaranteed Time Observations (GTO) were processed using the latest Data Reduction Software (DRS), offering enhanced precision and fewer systematic issues compared to previously published data. In some cases there were also new observations after a paper's release.

Radial velocity data from HIRES was sourced exclusively from literature. These datasets only include RV and associated uncertainties, with no activity indices. Despite this limitation, their high precision aids in the overall analysis as within `PyORBIT` it is possible to have different sized datasets.

HARPS and HARPS-N are very similar to each other, hence data of a target from these instruments were stored in the same file but assigned distinct jitter and offset parameters. I treated HIRES data, when available, as an independent dataset. These sets of instruments have different spectral ranges, so even though the underlying Gaussian process should be the same, the coefficients could be different.

There are multiple variables associated with each dataset in `PyORBIT`, such as activity coefficients, jitter and offset; as such, it was only worth adding a dataset from a computational and informational point of view if there were 10 or more points.

When the error of the BIS or FWHM was not explicitly given, it was calculated by simply doubling the error of the RV measurement. I discarded points with an error 5σ away from the median error. Those points were flagged in all 3 datasets since if, for example, there was a very low SNR at one moment, all datasets should be affected.

Following the process outlined above, I obtained the datasets, consisting of RV, BIS, $\log R'_{\mathrm{HK}}$ and associated errors, for all targets. Table 3.2 shows the spectrograph(s) used to observe each system, the number of points from each spectrograph and the data source.

Table 3.2: RV and auxiliary data size and source for each target of the sample used.

System	Instruments	Number of points	Data Reference
TOI-500	HARPS	197	[95]
Kepler-78	HARPS-N, HIRES	117 + 83	[117], [97], [96]
TOI-561	HARPS-N, HIRES	168 + 62	DACE, [98], [21], [99]
TOI-1807	HARPS-N	161	[37]
K2-229	HARPS, HARPS-N	12+120	ACTIN, DACE, [22], [100]
TOI-431	HARPS	162	[102]
TOI-1416	HARPS-N, HIRES, CARMENES	96+12+34	[103]
Kepler-10	HARPS-N, HIRES	292+40	[117], [104], [11]
CoRoT-7	HARPS	180	ACTIN, [53], [127]
HD 3167	HARPS, HARPS-N, HIRES	53+215+55	ACTIN, DACE, [117], [109], [111], [112]
K2-141	HARPS, HARPS-N	28+49	ACTIN,[117], [115], [114]
HD 80653	HARPS-N	208	[117], [116]
K2-106	HARPS, HARPS-N, HIRES	20+43+35	[118], [117], [119]
HD 213885	HARPS	47	ACTIN, [121]
HD 20329	HARPS-N	120	[122]
K2-131	HARPS-N	83	[117], [123]

3.5 Deriving the Planetary parameters

Planetary parameters were derived using a Bayesian approach implemented in `PyORBIT`, requiring as inputs the stellar mass — determined through isochrone fitting, as detailed in Section 3.3 — along with the RV and stellar activity data, described in Section 3.4.

3.5.1 Workflow

To refine the planetary parameters, I followed a structured approach, delineated in Figure 3.4. First, I performed an MCMC analysis using the planets listed in the discovery paper to obtain initial estimates of the planetary parameters, as well as the offset and jitter parameters for each spectroscopic time series. Next, I conducted nested sampling analyses with boundaries constrained by the initial MCMC results to prevent the sampler from exploring low-likelihood regions of parameter space. I started with the number of confirmed transiting planets and incrementally increased the number of planets until the model no longer showed a significant improvement over the previous one. MCMC analyses were also performed for the same number of models as in the NS. The choice of the best model for each planetary system relied on evaluating the Bayesian evidence, information criteria as well as the posterior distribution. The final planetary parameters were derived from this selected model.

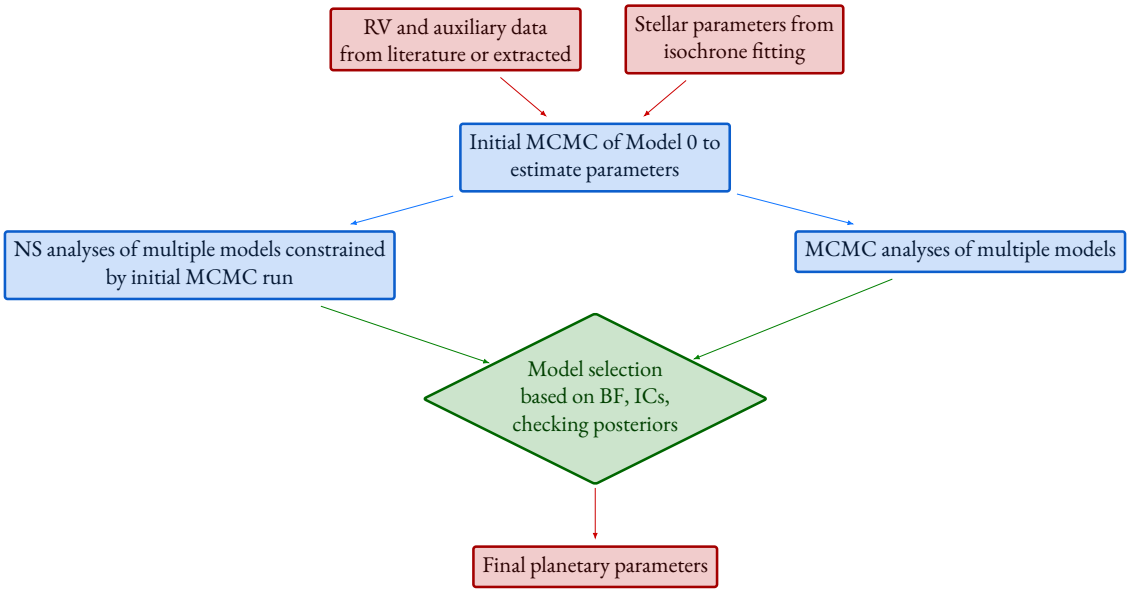


Figure 3.4: The general workflow used for determining the planetary parameters using MCMC and nested sampling. Model 0 represents a model with only the transiting planets.

3.5.2 PyORBIT

`PyORBIT`[†] is a python package used to model planetary transits and RVs [128][115]. It can simultaneously model stellar activity on the light curves or RVs by using Gaussian Processes. For this analysis I used the latest version, 10.8.

Multiple RV datasets and activity indicator datasets can be added, each with its own unique offset and jitter flag. Different types of prior distributions are available, offering flexibility in setting up the model. The model specifies the number of planets, the parametrization of the orbital parameters, the type of orbit, and the parameter ranges for each planet. Stellar parameters, especially the stellar mass, are also required to accurately derive the planetary parameters. For handling stellar activity, there are many kernel options and it is possible to use a Trained or a Multidimensional GP. In terms of sampling, the choice between Markov Chain Monte Carlo (MCMC) and nested sampling algorithms is available and both provide posterior probability density distributions for the parameters.

`PyORBIT` can also use photometry data which allows for precise constraints on the planet's radius and, if multiple transits are observed, on the orbital period as well. Both photometric and radial velocity models share common parameters, such as the orbital period and the central time of transit. Whether modelled separately or jointly, the derived planetary radius remains consistent, and any variations in the mass estimate typically stem from the radial velocity data. Modelling both photometric and radial velocity data together requires a sophisticated model, careful selection of transit data points, and a more time-consuming analysis. For this reason, only radial velocity data was used in the analysis. Nevertheless, information coming from photometry from past papers was included by adding Gaussian priors on some parameters such as the central time of transit T_0 and period P.

3.5.3 NS and MCMC configurations

The Bayesian evidence was computed using the `dynesty` nested sampling algorithm [72], described in Subsection 2.3.9. For all analyses 2000 live points were used. Each dataset in the model had constrained jitter and offset parameters that came from the preliminary MCMC run to make it converge faster.

When using MCMC, first a global optimisation of the parameters is performed with the differential evolution code `PyDE` and the resulting parameters are used as a starting point for the MCMC analysis. The MCMC algorithm chosen for this analysis is the python package `emcee` [59], a type of affine-invariant ensemble sampler [60], described in Subsection 2.3.6.

The first MCMC run is performed with a number of points which is not enough for convergence, instead its purpose is to get estimates of the jitter and offset parameters needed to constrain the parameter space for the nested sampling analysis.

The subsequent runs are done with a very high number of chains, 10^5, to properly sample the posterior distribution of the parameters. There are different ways to diagnose the convergence of the MCMC. A way to check if the analysis is to be "trusted" is to check the autocorrelation of the chains. This is done automatically in the `emcee` package with the integrated autocorrelation time τ_f. Due to the Markovian property of any MCMC method, the

[†]Available at `https://github.com/LucaMalavolta/PyORBIT`.

iterations in the chain will always be somewhat correlated. This means it is necessary to compute how many iterations are effectively independent.

To initiate the sampler, the number of burn-in points ($n_{\rm burn}$) must be specified. These points reflect the initial phase of the sampling process, during which the sampler has not yet converged on the target distribution. Consequently, the burn-in points are removed to prevent them from biasing the derived parameters. Generally, using about 25% of the chain for burn-in is sufficient. The parameter $n_{\rm thin}$ must also be set, the number by which the total number of points is divided by to minimize correlation between points. The autocorrelation function (ACF) is then computed on the thinned chains, with each parameter having a unique integrated autocorrelation time. The authors of the `emcee` paper found that with parallel chains, each chain should ideally run for a duration approximately 50 times the integrated autocorrelation time, denoted as τ_f. In `PyORBIT`, a warning is triggered if the chains are considered too short, along with a suggested minimum run length for adequate sampling.

We can check the trace plots which show the progress of the chain throughout time. Ideally, they are wiggly with a lot of "up and down movement", which is a sign that the chains are mixing well and are not stuck in any one part of the parameter space. Another valuable diagnostic is the Gelman-Rubin statistic, which compares the variance within a single chain to the variance between multiple chains to see if the chains have "forgotten" their initial conditions. Typically, values less than 1.1 indicate that convergence is likely achieved. All these diagnostics, the trace plots, GR plots and tables with ACF values, are provided by `PyORBIT`.

3.5.4 Priors for planetary parameters

The first model tested for each system is one where only the transiting planets are considered. Table B.1 in the Appendix shows the priors for this model for each target. Besides this one, models with increasing number of planets were tested. These extra planets had the following priors: uniform prior in the period $\mathcal{U}(2, 500)$ and in the RV semiamplitude $\mathcal{U}(0.01, 100)$ and a half-normal prior in the eccentricity $\mathcal{HN}(0.00, 0.098)$. The latter prior choice comes from Van Eylen et al. 2019 [129].

For each system, 3 spectroscopic time series were included (RV, BIS, $\log R'_{\rm HK}$), which were considered independent — the "Trained" — or assumed to share an underlying GP process — the "Multidimensional", described in Subsection 2.3.3.

For the GP coefficients, the same non-informative uniform priors were used for all systems, shown in Table 3.3. In the Trained case, each spectroscopic dataset from each instrument has its own coefficient ($H_{amp}^{RV,HARPS}$, $H_{amp}^{BIS,HARPS}$, $H_{amp}^{\log R'_{\rm HK},HARPS}$, $H_{amp}^{RV,HIRES}$, etc.), but the same boundary was used. In the Multidimensional approach, it is already explicitly divided by spectroscopic dataset but, again, each instrument will have its own coefficient (ex: V_r^{HARPS}, V_r^{HIRES}). Additionally, with the Multidimensional model there is a coupling between the two coefficients (X_r and X_c). To solve this degeneracy, one coefficient was forced to be positive, in this case V_r.

Table 3.3: Priors for the GP coefficients in the Trained and Multidimensional case.

Trained GP	
$H_{amp}\,(\mathrm{m\,s^{-1}})$	$\mathcal{U}(0,100)$
Multidimensional GP	
$V_c\,(\mathrm{m\,s^{-1}})$	$\mathcal{U}(-100,100)$
$V_r\,(\mathrm{m\,s^{-1}})$	$\mathcal{U}(0,100)$
$B_c\,(\mathrm{m\,s^{-1}})$	$\mathcal{U}(-100,100)$
$B_r\,(\mathrm{m\,s^{-1}})$	$\mathcal{U}(-100,100)$
$L_c\,(\mathrm{m\,s^{-1}})$	$\mathcal{U}(-1,1)$

4 Results

In this chapter, I present the results obtained. The effectiveness of the ESP kernel approximation compared to the QP is tested. The Bayesian evidences, information criteria and posteriors for the various models are assessed to choose the most likely configuration of each planetary system. The final planetary parameters are presented. A homogeneous mass-radius plot of the sample of USP planets is shown.

4.1 SPLEAF's ESP kernel validity

Firstly, I aimed to verify whether the ESP kernel implemented in S+LEAF produces results consistent with the QP kernel implemented in `tinyGP`. This comparison was motivated by the significantly reduced computational cost of the ESP kernel, achieved through its semi-separable matrix representation. The test runs were performed for the Kepler-78 system.

The posterior distributions of all parameters using `tinyGP`'s QP and S+LEAF's ESP kernels were compared for all models tested. These were the 1-planet, 2-planet and 3-planet models for both the Trained and Multidimensional GP models and for both the `dynesty` and `emcee` samplers. Corner plots were created using `pygtc`*[130]. I show here the results for one in particular: the Multidimensional GP 1-planet model with sampling done by `dynesty`. Figure 4.1 presents the posteriors of the planetary and stellar activity parameters. The two contour levels correspond to the 68% and 95% containment ranges for the two-dimensional projections. The blue contours correspond to the ESP kernel of S+LEAF, while the orange contours represent the QP kernel of `tinyGP`. As can be seen, S+LEAF's ESP kernel and `tinyGP`'s QP give very good agreement between each other. Corner plots with the same 1-planet model and Multidimensional GP but with sampling done for `emcee` instead of `dynesty` also show also very good agreement. Similarly, the 1-planet model using a Trained GP showed very good agreement

*Available at `https://github.com/SebastianBocquet/pygtc`.

for both `emcee` and `dynesty` sampling.

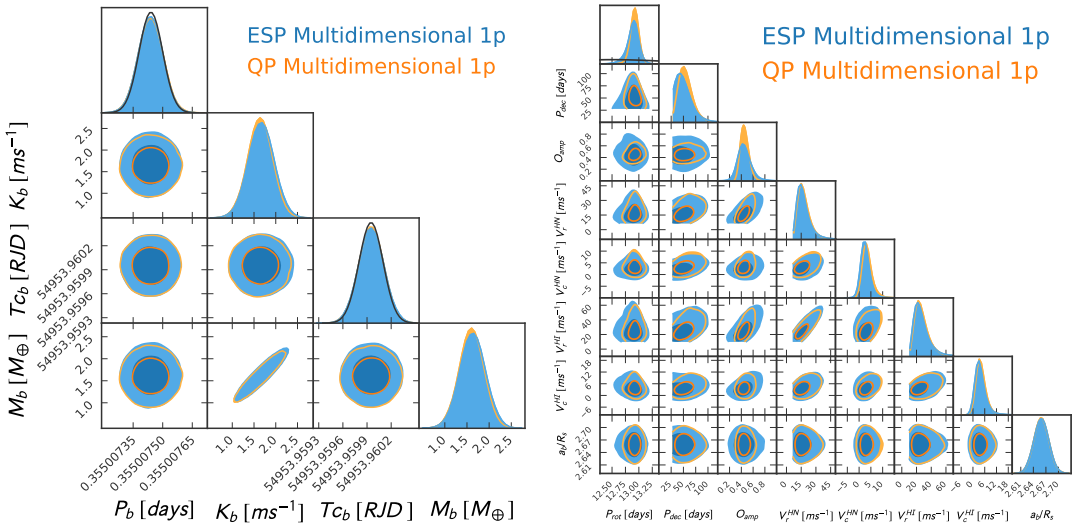


Figure 4.1: Kepler-78's main planetary (left) and stellar activity parameters (right) posterior distributions from the 1-planet Multidimensional GP with dynesty. The two distributions come from the ESP and QP kernel implemented in S+LEAF and `tinyGP`, respectively.

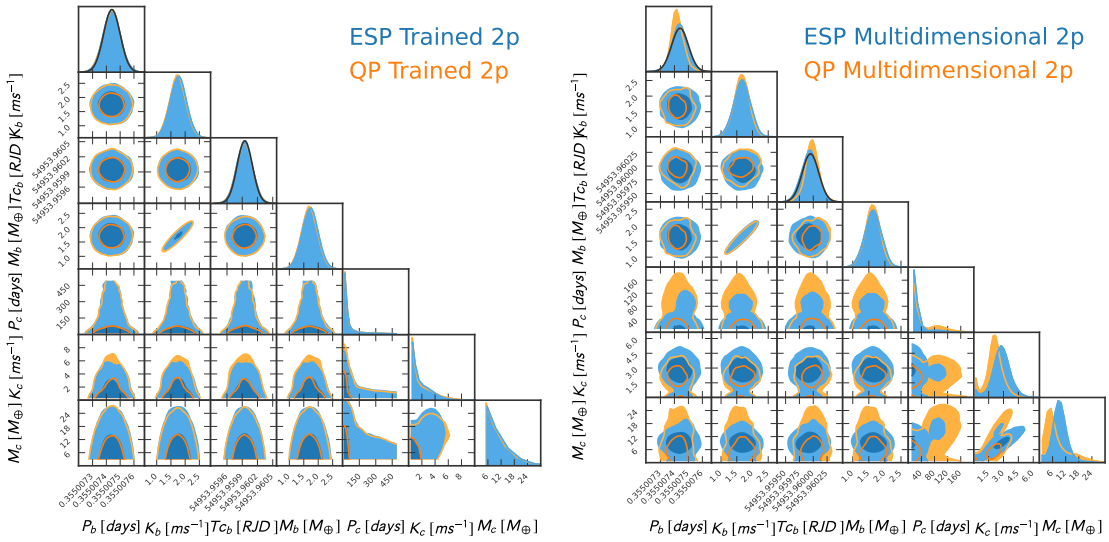


Figure 4.2: Kepler-78's main planetary parameters posterior distributions for the 2-planet model. Left: With Trained GP and `emcee`. Right: With Multidimensional GP and dynesty. The two distributions come from the ESP and QP kernel implemented in S+LEAF and `tinyGP`, respectively.

As follows, in the 1-planet model, which is the current one according to literature, S+LEAF and `tinyGP` are equivalent across samplers and across GP models.

For the 2-planet model, the ESP and QP are again perfectly superimposed both for `emcee` and `dynesty` but only for the Trained case. In the Multidimensional case, they start to diverge slightly on the 2nd planet and activity parameters. In Figure 4.2, the planetary parameters of two 2-planet models are shown. In the 3-planet model, the ESP and QP results show more significant discrepancies in the activity parameters and the parameters of the second and third planets. This divergence is expected, as there is likely only one planet, Kepler-78 b, in this system, making it challenging to constrain the "extra" parameters introduced.

Overall, we are able to get the same posteriors distributions independently of the sampler and of the GP model for the 1-planet model and very similar distributions in the 2-planet model. For this reason, SPLEAF's ESP kernel is considered a very good approximation to `tinyGP`'s QP kernel with the added benefit of much lower computational cost. As such, it was implemented for all the remaining targets.

4.2 Model selection criteria comparison

In addition to analysing the parameter posteriors, I examined whether the same model was preferred based on Bayesian evidence and information criteria. This comparison was conducted for the Kepler-78 system, utilizing both S+LEAF's ESP kernel and `tinyGP`'s QP kernel.

Both the Trained and Framework GP were tested across the 1-planet, 2-planet and 3-planet models. The BIC, AIC and AICc were computed for the samples generated by both the `emcee` and `dynesty` samplers. The bayesian evidence $\ln(\mathcal{Z})$ can only be obtained with the dynesty nested sampler. The differences to the 1-planet model were calculated for each one of the model selection criteria. For the information criteria, $-(\mathrm{IC}_{n+1} - \mathrm{IC}_n)$ is plotted, and for the Bayesian evidence $2(\ln(\mathcal{Z})_{n+1} - \ln(\mathcal{Z})_n)$. These adjustments ensure that all criteria are aligned in direction, since higher model probabilities correspond to lower BIC/AIC/AICc values and higher $\ln(\mathcal{Z})$. In this manner, values above 0 indicate a model more likely than the 1-planet model. The results for the Trained and Multidimensional GP are shown in Figure 4.3.

For this particular system, Kepler-78, the uncertainties in the information criteria are too large to make a decision. It is interesting to note that from the $\ln(z)$ for the Framework GP we could be led to select the 2-planet or even the 3-planet model. But as can be seen from the posterior distributions, the period is very badly constrained and the semi-amplitude for the 2nd and 3rd planet is extremely close to 0, which suggests that there are likely no additional planets beyond Kepler-78 b. The Trained case seems to give more accurate results, penalizing all models more complex than the 1-planet one.

It can also be observed that `emcee` and `dynesty`, represented by solid colours and hatched patterns respectively, generally lead to the same model selection outcomes. The main exceptions are the AIC/AICc values in the Framework case.

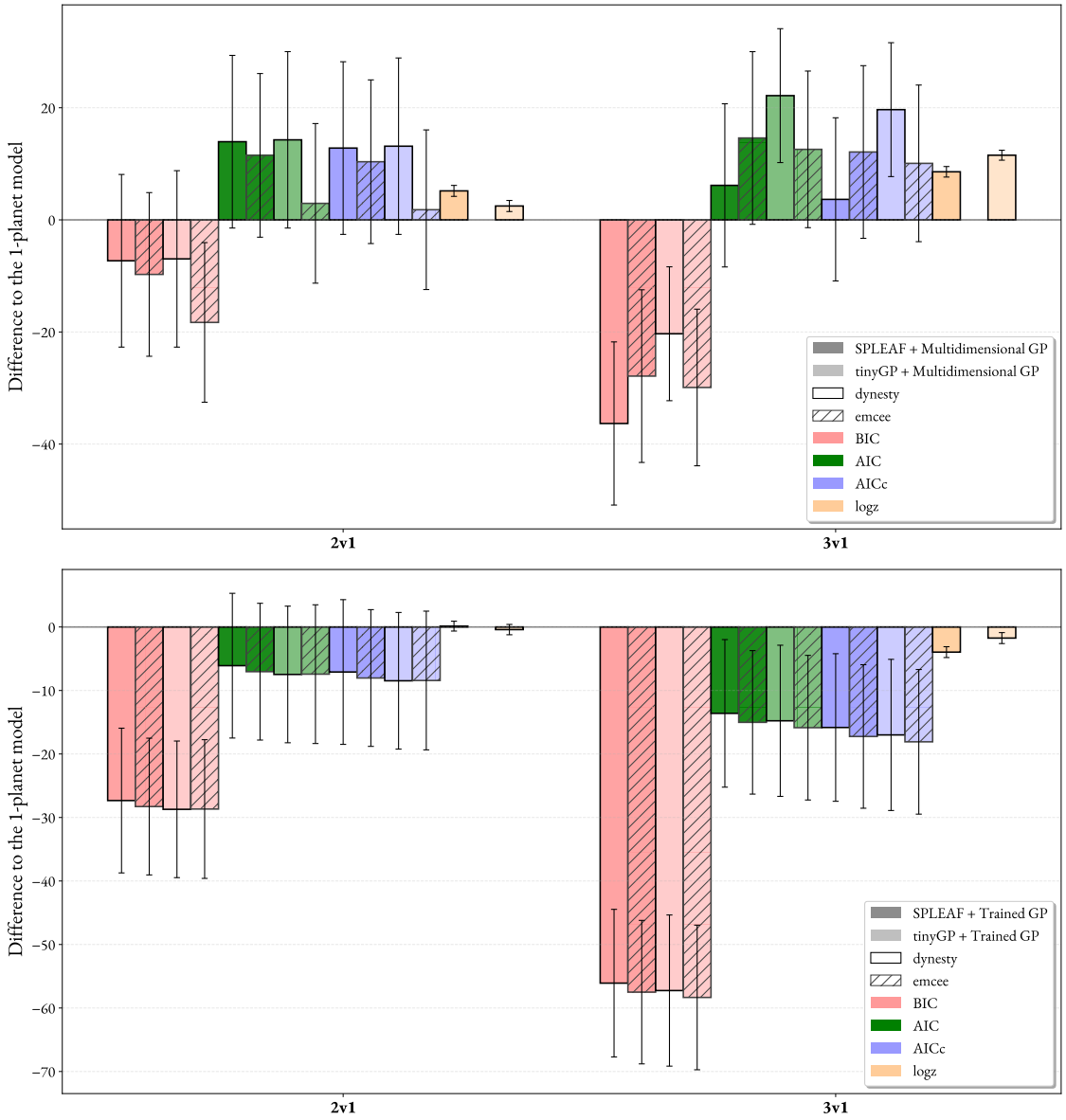


Figure 4.3: Difference in evidence to the 1-planet model for Kepler-78's system. Top: Using the Multidimensional GP. Bottom: Using the Trained GP. Darker and lighter shades of the same colour correspond to `S+LEAF` and `tinyGP`, which implement the ESP and the QP kernel, respectively. Solid colours correspond to nested sampling performed by `dynesty` while hatched bars correspond to MCMC performed by `emcee`. The information criteria metrics BIC, AIC, AICc were calculated for both samplers while $\ln(z)$ is only given by nested sampling.

4.3 Model selection for the targets

To select the best model, in principle, one could calculate the Bayesian evidence for all models (the 1-planet, 2-planet, 3-planet, for example) and select the one that gives a difference substantial enough to be considered "Strong" evidence. In Figure 4.4 this is represented, where the green, red and blue lines represent the Positive, Strong and Very Strong thresholds from Kass and Raftery's Table 2.1. In this simple example for K2-131 the 1-planet model is favoured.

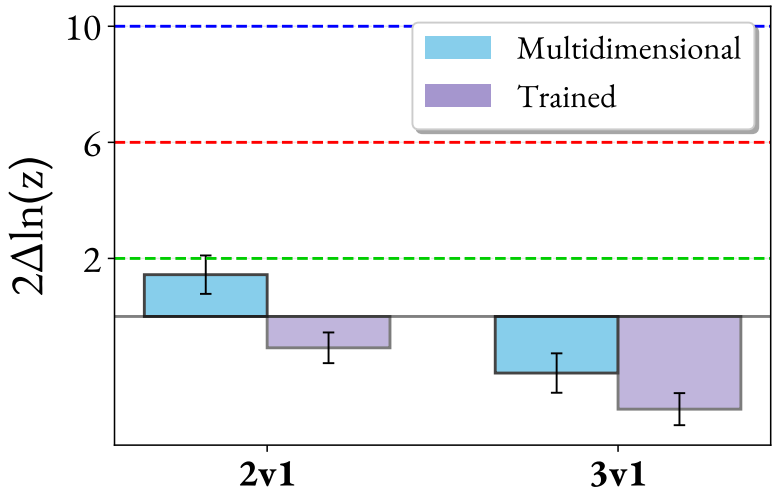


Figure 4.4: Difference in Bayesian evidence for K2-131's system using the Trained and Framework GP. A positive value suggest higher model probability relative to the 1-planet model. The green, orange and red lines represent the empirical Positive, Strong and Very Strong model selection thresholds from Kass and Raftery's Table 2.1.

Alternatively, a more computationally efficient approach involves selecting the model that minimizes the BIC or AIC, followed by a final MCMC or nested sampling run with a sufficient number of chains to better constrain the parameters. This single, large final run is often preferred due to the computational time and cost involved. However, as was discussed in the previous Section 4.2, there seemed to be some problems with this theoretically straightforward approach of model selection. Firstly, the information criteria (ICs) did not always align with each other or with the Bayesian evidence. This is illustrated in Figure 4.5 for Kepler-10's system. Secondly, some favoured models either by the ICs, the Bayesian evidence, or both, in reality had poor constrains on the added parameters. Given that the ESP kernel in S+LEAF was shown to provide a good approximation for the QP kernel and significantly reduced computational costs, it became feasible to perform both MCMC and nested sampling runs for all target configurations. This allowed for a more thorough approach, where instead of relying solely on model selection criteria, the focus was shifted to examining the parameter distributions across different models for each target.

I observed that in some targets, a more complex model was favoured by one or more model selection criterion, yet the periods of the added signal did not match between the Trained and Framework models, which suggests that

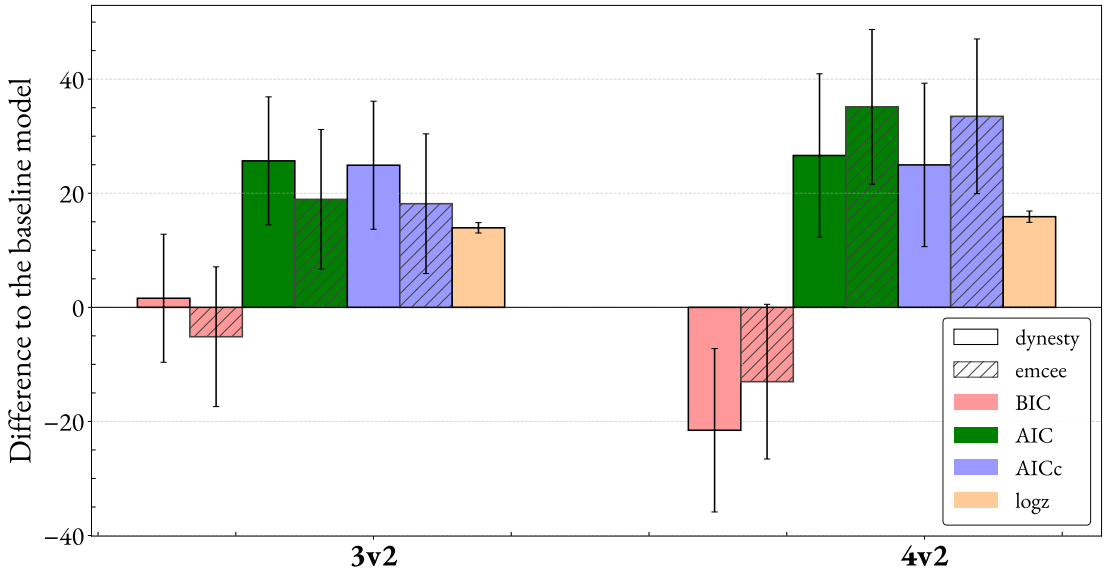


Figure 4.5: Difference in Bayesian evidence and Information Criteria metrics to the baseline 2-planet model for Kepler-10's system. Using SPLEAF's ESP kernel and the Multidimensional GP. A positive value suggest higher model probability relative to the 2-planet model.

the added signal is not likely to be planetary in nature. I would expect that the parameters coming from Trained and Framework to be consistent, that is, they could have different levels of detection, but the period should match since the periodicity of the underlying signal should be the same.

For some other targets, where a more complex model was favoured by one or more model selection criterion, the period of the added signal was found to be a harmonic of the stellar rotation period, such as $1/2$ or $2/3$. This could mean that stellar activity is still being incorrectly modelled as a planetary signal, despite the GP attempting to account for the stellar activity. This issue is particularly pronounced in young stars, where stellar activity signals often have a sinusoidal shape that closely resembles planetary signals. These "extra" signals in some cases had a sigma level of detection in the semi-amplitude of around 5 so they could be mistaken as planets were they not checked against the stellar rotation period. For those occurrences, a global fit incorporating both photometry and radial velocity data, along with more refined data treatment, would be beneficial.

4.4 Derived planetary parameters

Finally, after checking the consistency of the derived planetary parameters across models with varying number of planets, for both `emcee` and `dynesty` sampling and stellar activity modelling Trained/Framework, a set of parameters for each system was obtained. The set of parameters coming from s+leaf's ESP kernel using the Multidimensional GP Framework and `emcee` sampling are displayed in Table 4.1.

Table 4.1: Derived planetary parameters from MCMC run of chosen model (Part 1)

Parameter	TOI-500	Kepler-78	TOI-561	TOI-1807	K2-229	TOI-431
Planet b						
P (days)	0.5481584 ± 0.0000080	$0.355007438 \pm 0.000000059$	$0.44656881 \pm 0.00000080$	0.549380 ± 0.000012	0.584244 ± 0.000013	0.4901990 ± 0.0000042
K (ms^{-1})	1.34 ± 0.19	1.67 ± 0.29	2.07 ± 0.18	2.39 ± 0.40	2.09 ± 0.42	4.18 ± 0.30
e (Fixed)	0	0	0	0	0	0
Ω (Fixed)	90	90	90	90	90	90
T_c ($BJD_{TDB}-2400000$)	58468.39103 ± 0.00057	54953.95995 ± 0.00015	59317.74976 ± 0.00049	58899.34491 ± 0.00080	57582.8862 ± 0.0011	58627.5344 ± 0.0066
a/R_s	3.592 ± 0.029	2.659 ± 0.017	2.700 ± 0.030	3.705 ± 0.023	3.587 ± 0.035	3.191 ± 0.031
i (Fixed)	82.09 ± 2.53	75.2 ± 2.6	87.2 ± 2.1	82.2 ± 2.0	81.6 ± 10.8	84.3 ± 1.3
$M(M_\oplus)$	1.45 ± 0.21	1.69 ± 0.28	2.10 ± 0.19	2.67 ± 0.45	2.52 ± 0.51	4.65 ± 0.34
L	123.5 ± 1.3	-	206.41 ± 0.43	167.7 ± 1.5	269.5 ± 3.0	57.6 ± 4.9
Planet c						
P (days)	6.6364 ± 0.0051	-	10.778831 ± 0.000034	-	8.32799 ± 0.00040	4.8502 ± 0.0089
K (ms^{-1})	1.86 ± 0.19	-	1.86 ± 0.22	-	2.47 ± 0.75	1.38 ± 0.24
e	$0.055^{+0.059}_{-0.039}$	-	$0.042^{+0.048}_{-0.030}$	-	0.27 ± 0.19	0.059 ± 0.041
Ω	-63^{+82}_{-100}	-	13^{+82}_{-120}	-	112^{+44}_{-101}	-3^{107}_{-143}
T_c ($BJD_{TDB}-2400000$)	58601.79 ± 0.20	-	59238.46290 ± 0.00080	-	57586.3351 ± 0.0017	58630.59 ± 0.20
a/R_s	18.94 ± 0.15	-	22.55 ± 0.25	-	21.09 ± 0.20	14.71 ± 0.14
$M(M_\oplus)$	4.55 ± 0.48	-	5.44 ± 0.66	-	6.6 ± 1.9	3.27 ± 0.57
L	350.9 ± 9.2	-	$244.6^{+2.9}_{-4.6}$	-	259 ± 23	183 ± 14
Planet d						
P (days)	26.213 ± 0.021	-	25.71240 ± 0.00020	-	-	12.4591 ± 0.0026
K (ms^{-1})	8.94 ± 0.21	-	3.05 ± 0.27	-	-	3.37 ± 0.28
e	0.038 ± 0.022	-	$0.122^{+0.061}_{-0.056}$	-	-	$0.041^{+0.053}_{-0.029}$
Ω	45 ± 37	-	-117^{+13}_{-27}	-	-	-81^{+110}_{-97}
T_c ($BJD_{TDB}-2400000$)	58613.21 ± 0.24	-	59318.9659 ± 0.0040	-	-	58627.547 ± 0.033
a/R_s	47.33 ± 0.38	-	40.26 ± 0.44	-	-	27.59 ± 0.27
$M(M_\oplus)$	34.68 ± 0.87	-	11.8 ± 1.1	-	-	10.93 ± 0.91
L	266.0 ± 2.4	-	231.0 ± 3.1	-	-	73.9 ± 0.91
Planet e						
P (days)	102.21 ± 0.80	-	76.98 ± 0.20	-	-	-
K (ms^{-1})	2.73 ± 0.28	-	2.57 ± 0.23	-	-	-
e	0.317 ± 0.058	-	$0.060^{+0.063}_{-0.042}$	-	-	-
Ω	-31 ± 15	-	153^{+85}_{-98}	-	-	-
T_c ($BJD_{TDB}-2400000$)	$58626.9^{+3-5}_{-2.9}$	-	59231.8 ± 1.7	-	-	-
a/R_s	117.2 ± 1.1	-	83.65 ± 0.93	-	-	-
$M(M_\oplus)$	15.4 ± 1.4	-	14.4 ± 1.3	-	-	-
L	$314.7^{+6.5}_{-7.0}$	-	304.1 ± 5.6	-	-	-
Stellar activity						
$P_{\rm rot}$ (days)	$41.7^{+2.8}_{-1.1}$	$12.898^{+0.087}_{-0.110}$	$40.8^{+3.4}_{-12}$	8.757 ± 0.064	18.92 ± 0.28	31.01 ± 0.43
$P_{\rm dec}$ (days)	$121_{-26}+76$	47^{+19}_{-13}	101^{+28}_{-19}	$18.41^{+1.4}_{-0.64}$	$40.7^{+5.6}_{-2.2}$	$70.5^{+10}_{-6.1}$
O_{amp}	$0.887^{+0.082}_{-0.14}$	$0.442^{-0.120}_{-0.092}$	$0.896^{-0.078}_{-0.15}$	0.445 ± 0.053	0.74 ± 0.11	$0.61^{+0.17}_{-0.11}$

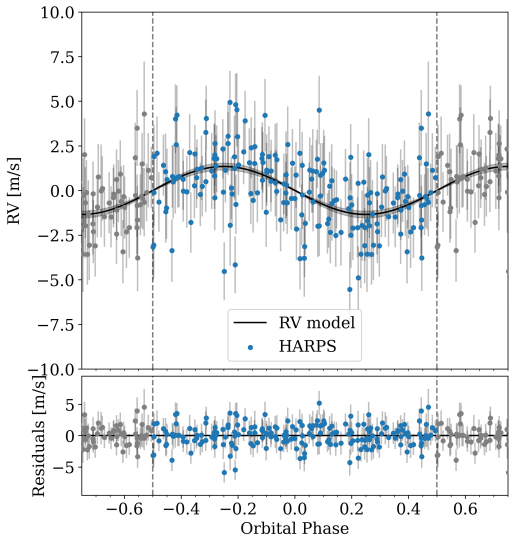
Table 4.1: Derived planetary parameters from MCMC run of chosen model (Part 2)

Parameter	TOI-1416	Kepler-10	CoRoT-7	HD 3167	K2-141
Planet b					
P (days)	1.0697605 ± 0.0000036	$0.83749070 \pm 0.00000020$	0.853574 ± 0.000012	0.9596296 ± 0.0000091	$0.280324956 \pm 0.000000068$
K (ms^{-1})	2.39 ± 0.29	2.39 ± 0.21	3.99 ± 0.47	3.35 ± 0.18	5.31 ± 0.91
e (Fixed)	0	0	0	0	0
Ω (Fixed)	90	90	90	90	90
T_c ($\mathrm{BJD}_{TDB}-2400000$)	59689.4061 ± 0.0027	55034.08687 ± 0.00018	54398.0767 ± 0.0015	57394.37447 ± 0.00044	57744.07151 ± 0.00010
a/R_s	5.248 ± 0.030	3.356 ± 0.038	4.423 ± 0.033	4.521 ± 0.069	2.297 ± 0.030
i (Fixed)	85.7 ± 1.7	84.8 ± 3.9	80.98 ± 0.51	83.4 ± 7.7	86.3 ± 3.6
$M(M_\oplus)$	3.36 ± 0.41	3.37 ± 0.30	5.66 ± 0.67	4.80 ± 0.30	4.43 ± 0.76
L	288.4 ± 1.2	244.13 ± 0.31	154.7 ± 3.6	127.2 ± 2.1	79.48 ± 0.15
Planet c					
P (days)	-	45.294302 ± 0.000048	3.69697 ± 0.00028	29.8454 ± 0.0012	7.748994 ± 0.000015
K (ms^{-1})	-	1.95 ± 0.24	5.15 ± 0.35	2.72 ± 0.19	1.7 ± 1.1
e	-	0.00	$0.048^{+0.050}_{-0.034}$	0.00	$0.063^{+0.070}_{-0.045}$
Ω	-	90	8^{+69}_{-117}	90	124^{+112}_{-139}
T_c ($\mathrm{BJD}_{TDB}-2400000$)	-	55062.26649 ± 0.00081	55003.433 ± 0.076	57394.9788 ± 0.0012	58371.07417 ± 0.00066
a/R_s	-	48.00 ± 0.54	11.752 ± 0.088	44.72 ± 0.69	21.00 ± 0.27
$M(M_\oplus)$	-	10.3 ± 1.3	11.74 ± 0.81	12.10 ± 0.90	$4.2^{+3.0}_{-2.5}$
L	-	39.164 ± 0.025	113.1 ± 4.8	187.85 ± 0.29	$131.1^{+7.3}_{-5.6}$
Planet d					
P (days)	-	150.92 ± 0.50	$8.5342^{+0.026}_{-0.0045}$	8.5342 ± 0.026	-
K (ms^{-1})	-	1.52 ± 0.30	4.00 ± 0.72	1.30 ± 0.17	-
e	-	0.22 ± 0.11	$0.079^{+0.072}_{-0.054}$	$0.058^{+0.063}_{-0.041}$	-
Ω	-	148^{27}_{-29}	-157 ± 60	-168^{+118}_{-126}	-
T_c ($\mathrm{BJD}_{TDB}-2400000$)	-	$58072.2^{5.1}_{-6.5}$	$55008.28^{+0.56}_{-8.2}$	$58003.44^{+0.83}_{-0.37}$	-
a/R_s	-	107.1 ± 1.2	21.21 ± 0.16	19.42 ± 0.30	-
$M(M_\oplus)$	-	11.7 ± 2.2	12.2 ± 2.2	3.80 ± 0.49	-
L	-	297 ± 11	96.9 ± 9.2	308^{+13}_{-39}	-
Planet e					
P (days)	-	-	-	79.53 ± 0.22	-
K (ms^{-1})	-	-	-	1.27 ± 0.21	-
e	-	-	-	$0.070^{+0.072}_{-0.049}$	-
Ω	-	-	-	2^{+76}_{-88}	-
T_c ($\mathrm{BJD}_{TDB}-2400000$)	-	-	-	$58045.5^{+2.3}_{-2.4}$	-
a/R_s	-	-	-	86.0 ± 1.3	-
$M(M_\oplus)$	-	-	-	7.8 ± 1.3	-
L	-	-	-	239.5 ± 8.3	-
Stellar activity					
P_{rot} (days)	22.88 ± 0.32	$27.66^{+0.59}_{-1.2}$	22.03 ± 0.24	$24.50^{+0.61}_{-0.52}$	$11.71^{+0.11}_{-0.19}$
P_{dec} (days)	$47.4^{+2.6}_{-1.3}$	$59.9^{+12}_{-3.5}$	$45.3^{+1.9}_{-1.0}$	$53.8^{+7.3}_{-3.6}$	506^{+360}_{-372}
O_{amp}	0.73 ± 0.12	$0.72^{+0.20}_{-0.16}$	$0.924^{+0.051}_{-0.068}$	$0.919^{+0.058}_{-0.092}$	0.63 ± 0.22

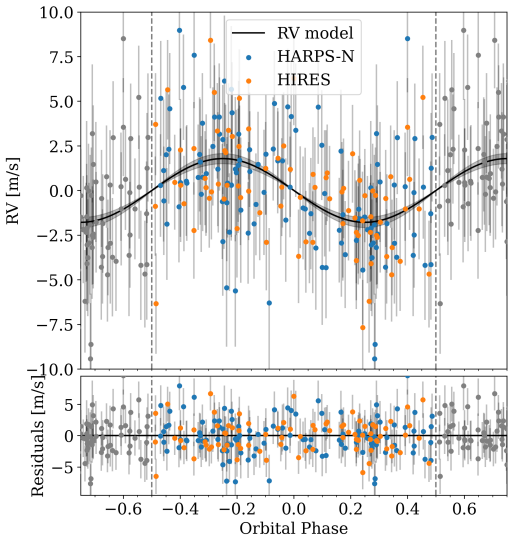
Table 4.1: Derived planetary parameters from MCMC run of chosen model (Part 3)

Parameter	HD 80653	K2-106	HD 213885	HD 20329	K2-131
Planet b					
P (days)	0.7195764 ± 0.0000086	0.5712905 ± 0.0000095	1.007913 ± 0.00010	0.9260625 ± 0.0000025	0.3693017 ± 0.0000064
K (ms^{-1})	3.43 ± 0.30	6.53 ± 0.65	5.03 ± 0.52	6.82 ± 0.80	8.2 ± 1.3
e (Fixed)	0	0	0	0	0
Ω (Fixed)	90	90	90	90	90
T_c ($\mathrm{BJD}_{TDB}-2400000$)	58134.42443 ± 0.00069	57394.011400 ± 0.00010	58379.96470 ± 0.00010	60233.3678 ± 0.0018	57582.9359 ± 0.0011
a/R_s	2.947 ± 0.066	2.923 ± 0.049	3.725 ± 0.042	3.542 ± 0.032	2.670 ± 0.039
i (Fixed)	86.3 ± 2.4	86.4 ± 4.1	80.09 ± 0.62	76.01 ± 0.46	85.0 ± 10.0
$M(M_\oplus)$	5.35 ± 0.47	8.18 ± 0.84	8.13 ± 0.84	10.4 ± 1.2	8.2 ± 1.3
L	52.5 ± 5.2	$351.4^{+5.6}_{-346}$	46.08 ± 0.70	330.62 ± 0.99	$7.2^{+4.0}_{-3.6}$
Planet c					
P (days)	860 ± 12	13.33957 ± 0.00095	-	-	-
K (ms^{-1})	189 ± 12	2.90 ± 0.85	-	-	-
e	0.8382 ± 0.0096	0.135 ± 0.074	-	-	-
Ω	42.7 ± 1.6	-151^{+35}_{-47}	-	-	-
T_c ($\mathrm{BJD}_{\mathrm{TDB}}-2400000$)	59829 ± 12	57405.7315 ± 0.0044	-	-	-
a/R_s	332.906 ± 4.867	23.88 ± 0.40	-	-	-
$M(M_\oplus)$	1712^{+71}_{-63}	10.2 ± 3.0	-	-	-
L	58.3 ± 1.6	299.1 ± 8.1	-	-	-
Stellar activity					
P_{rot} (days)	$29.27^{+0.55}_{-1.0}$	$21.60^{+0.22}_{-0.17}$	$9.495^{+0.093}_{-0.14}$	$38.10^{+1.10}_{-0.76}$	9.152 ± 0.054
P_{dec} (days)	$64.2^{+9.6}_{-4.4}$	63 ± 11	62^{+76}_{-25}	574^{+293}_{-318}	$40.1^{+13}_{-10.0}$
O_{amp}	0.78 ± 0.16	$0.016^{-0.085}_{-0.014}$	$0.933^{+0.050}_{-0.097}$	$0.44^{+0.28}_{-0.42}$	$0.474^{+0.084}_{-0.075}$

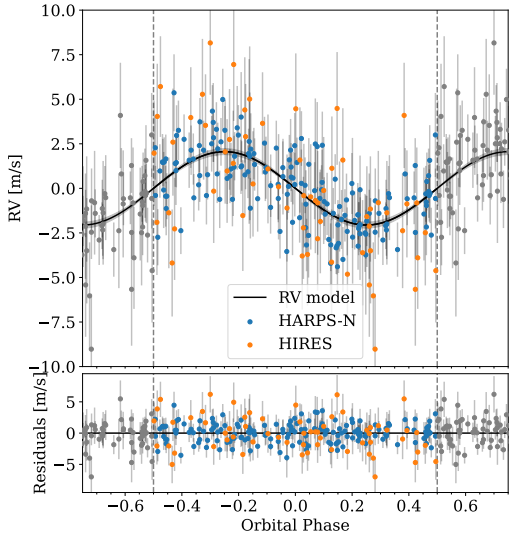
A Keplerian model is created for each planet by computing the expected radial velocities based on the derived orbital parameters for each planet. The model is then plotted with the observed data to see how well the Keplerian orbit matches the RV measurements. These are shown in Figure 4.6, where they were folded with respect to the transit time. Black lines represent the modelled RV curves and the shaded region shows the range of RV values across the model samples, representing the uncertainty in the model. The points represent the observed RV measurements with different colours representing the different instruments. Residuals show the difference between the modelled and observed RV measurements.



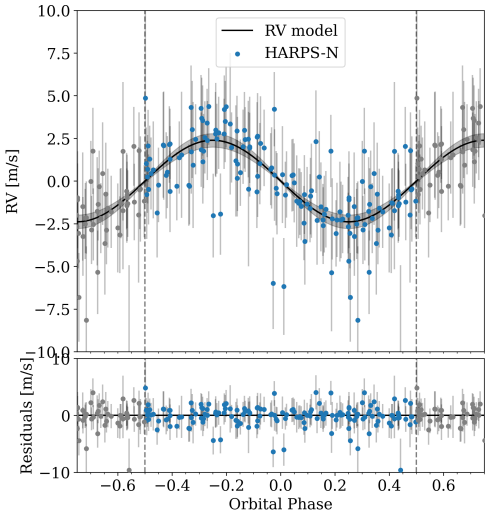
(a) TOI-500 b. Model: Multidimensional GP, 4-planet, emcee sampling.



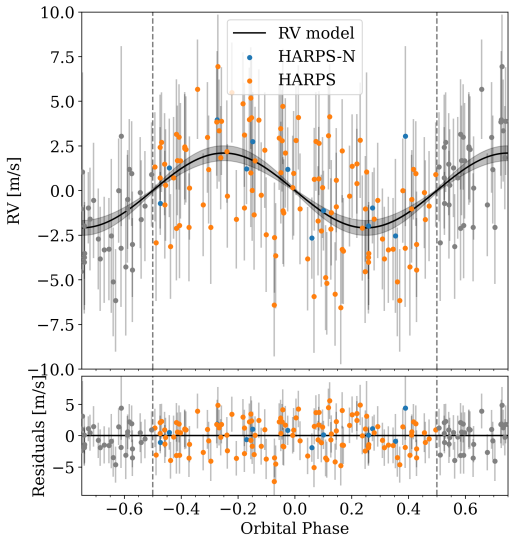
(b) Kepler-78 b. Model: Multidimensional GP, 1-planet, emcee sampling.



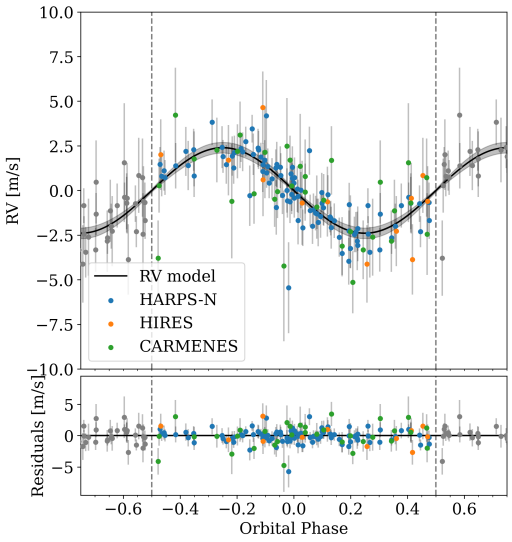
(c) TOI-561 b. Model: Multidimensional GP, 4-planet, emcee sampling.



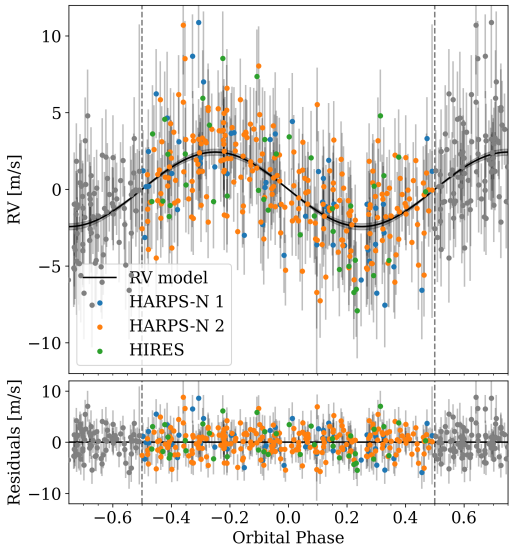
(d) TOI-1807 b. Model: Multidimensional GP, 1-planet, emcee sampling.



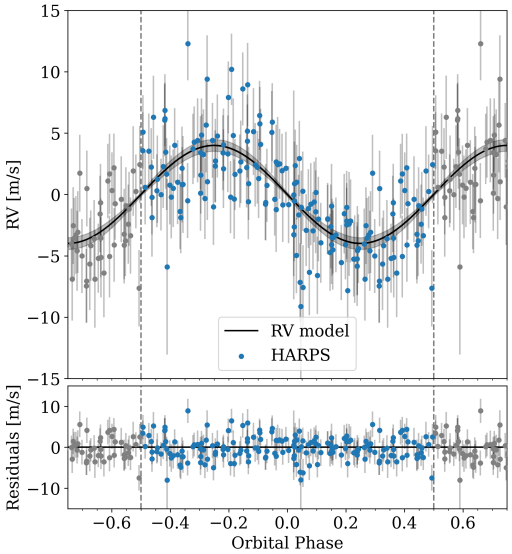
(e) K2-229 b. Model: Multidimensional GP, 2-planet, emcee sampling.



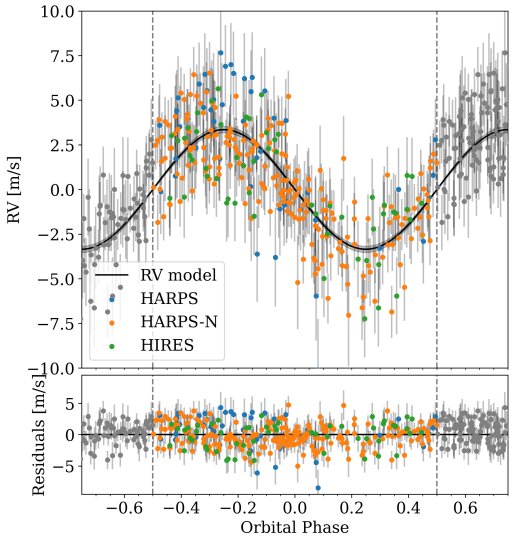
(f) TOI-1416 b. Model: Multidimensional GP, 1-planet, emcee sampling.



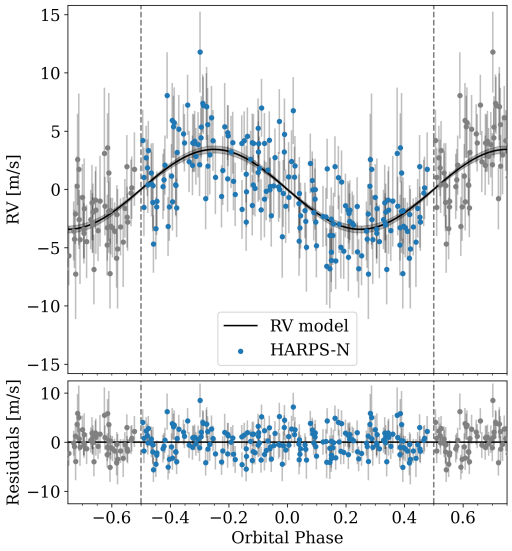
(g) Kepler-10 b. Model: Multidimensional GP, 3-planet, emcee sampling.



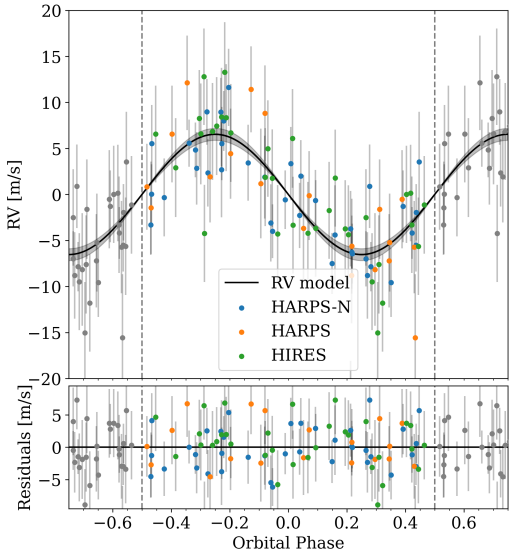
(h) CoRoT-7 b. Model: Multidimensional GP, 3-planet, emcee sampling.



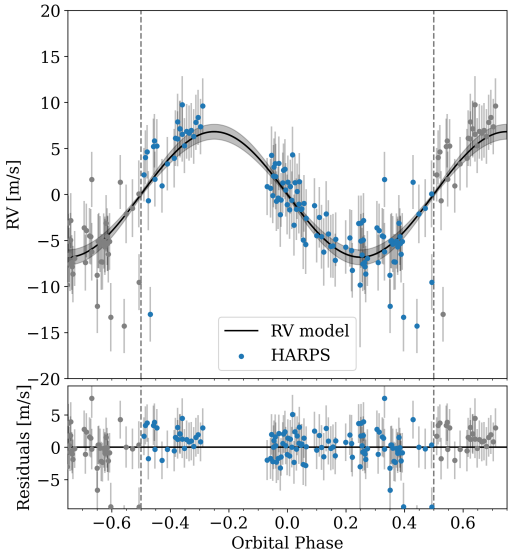
(i) HD 3167 b. Model: Multidimensional GP, 4-planet, emcee sampling.



(j) HD 80653 b. Model: Multidimensional GP, 2-planet, emcee sampling.



(k) K2-106 b. Model: Multidimensional GP, 2-planet, emcee sampling.



(l) HD 20329 b. Model: Multidimensional GP, 1-planet, emcee sampling.

Figure 4.6: RV plots of USP planets using the derived parameters.

In Figure 4.7, the RV and BIS data of K2-229 are plotted with different instruments in different colours. The grey area represents the GP model.

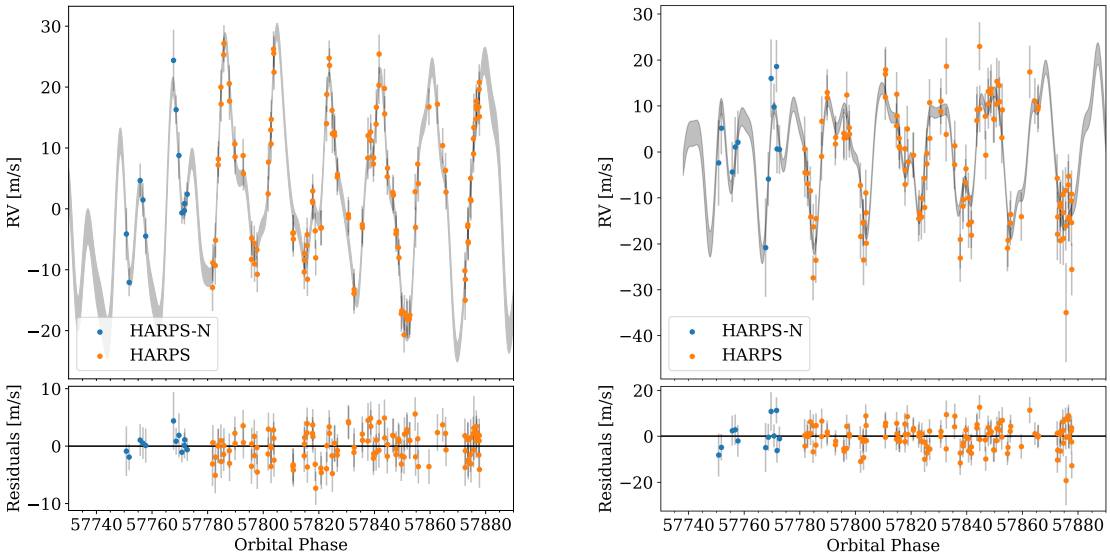


Figure 4.7: GP model of the combined effect of planetary motion and stellar activity coming from RV (left) and BIS (right) measurements of K2-229.

4.5 Comparison of configurations with literature

For most systems, I obtained a configuration that closely matches those reported in the literature. However, for a few systems, some derived parameters did not match or it was not possible to detect a confirmed planet at all. In this section, I take note of some observed mismatches.

For TOI-500, the derived period for the 4th planet was 102.21 ± 0.80 compared to 61.30 ± 0.28 from Serrano et al. 2022 [95].

For TOI-1807, though a second planetary signal is detected with a 8 and 6 sigma level on the semi-amplitude for the Multidimensional and Trained GP, respectively, the period of this signal has a substantial difference between the two. With `emcee`, the Multidimensional gives a period of 8.7527 ± 0.0053 and the Trained gives 5.5068 ± 0.0056. This difference was also reported by Nardiello et al. 2022 [37]. For this reason, I reported only the 1-planet model.

For HD 3167, the 4th planet was detected using `emcee` with at least 6σ for both Multidimensional and Trained but it was not detected with `dynesty`. With `emcee` the period obtained with Multidimensional was 79.53 ± 0.22 while with Trained was 115.79 ± 0.50. The values for the period of the fourth planet in previous papers were 96.630 ± 0.292 and 102.09 ± 0.51 [117] [109].

For K2-141, there are two confirmed transiting planets which were put in the baseline model. In spite of that, in this analysis, the second planet is not correctly recovered, with less than 2σ detection for all `emcee`/`dynesty`

and Trained/Multidimensional combinations. The previous analysis of this system by Malavolta et al. 2018 [115] was done also with both Trained and Multidimensional GPs and different samplers, but on a smaller number of data points, and the mass of the second planet was not recovered. A new analysis using `tinyGP` on the same dataset recovered the mass of the planet with a 4σ confidence, only in the Multidimensional GP case. This is the only case when S+LEAF and `tinyGP` gave different results, and it could represent a benchmark case to investigate the differences between different stellar activity modelling in future works.

For HD 213885, only one planet is detected instead of the two planets reported by Espinoza et al. 2020 [121], even after adding HARPS and FEROS datasets.

Fortunately, across all configurations tested, the USP planets have well-constrained parameters, largely due to the strict priors coming from photometry, as all of these planets are transiting. Consequently, even in models where the extra signals' detection is not sure or there are some mismatches with literature, the USPP's parameters remain consistent across configurations, with similar derived parameters.

4.6 Homogeneous mass-radius analysis of USP planets

The key properties of the USP planets, including their mass, radius and density, are summarized in Table 4.2.

To gain insight into the density and composition of planets, they can be placed on a mass-radius diagram alongside theoretical mass-curves for different planet compositions. I generated mass-radius plots for the sample of 16 USP planets using the package `mr-plotter`[†][23], shown in Figure 4.8. The plot on top uses values drawn from literature with the reference paper in parenthesis. The bottom one uses the values coming from this analysis. The planetary masses come from the RV analysis and are reported in Table 4.1. The planetary radii were calculated with R_p/R_s from literature and the stellar radii from the isochrone fitting reported in Table 3.1.

The derived stellar radii had lower uncertainties than the ones used in previous papers, which translated into lower uncertainties for the majority of the planetary radii. Some exceptions to this were when Rp/Rs had larger reported uncertainties, for example TOI-431. Regarding the derived masses, some have lower uncertainties compared to previous studies (TOI-1807, TOI-500, TOI-561, K2-229 and TOI-1416) while the remaining have similar or slightly larger uncertainties. The latter cases occur primarily when the orbital architecture of the systems could not be fully recovered. The overall scatter in the M-R plot did not decrease, but considering that all of the values were obtained using the same methodology, a direct comparison between planets is more reliable. The fact that there is still scatter might mean that indeed these planets have different internal compositions. This, in turn, could point to different formation mechanisms or evolutionary histories. Most of these planets seem to have a composition ranging between 0 and 33% Fe. Some planets have very low densities, namely TOI-561 b and TOI-1416 b. On the other hand, TOI-431 b and K2-229 b have very high densities, and appear close to a 100 % Fe composition. At a first glance, these differences do not seem to be related to the metallicity of the star, e.g., TOI-561 is a metal-poor star ([Fe/H]=$0.33^{+0.08}_{-0.05}$) while TOI-1416 is slightly supersolar ([Fe/H]=0.07 ± 0.04); a detailed analysis of the star-planet connection is beyond the scope of this thesis. By comparing the values from the literature with those obtained here, we can conclude that even when using the same datasets, altering the treatment of stellar activity

[†]Available at `https://github.com/castro-gzlz/mr-plotter`.

can significantly impact the results. Therefore, population studies are likely to be reliable only if conducted in a homogeneous manner.

USP Planet	$M_p(M_\oplus)$	$R_p(R_\oplus)$	Density (g/cm^3)	Reference for Rp/Rs
TOI-500 b	1.45 ± 0.21	1.234 ± 0.057	$4.25^{+1.00}_{-0.97}$	[95]
Kepler-78 b	1.63 ± 0.28	1.197 ± 0.031	5.35 ± 0.94	[117]
TOI-561 b	2.10 ± 0.19	1.419 ± 0.037	4.1 ± 0.49	[98]
TOI-1807 b	2.67 ± 0.44	1.390 ± 0.046	5.48 ± 1.06	[37]
K2-229 b	2.52 ± 0.51	1.188 ± 0.042	$8.29^{+1.88}_{-1.92}$	[15]
TOI-431 b	4.65 ± 0.35	1.313 ± 0.083	$11.33^{+2.31}_{-2.28}$	[131]
TOI-1416 b	3.36 ± 0.42	1.602 ± 0.047	$4.51^{+0.69}_{-0.68}$	[103]
Kepler-10 b	3.37 ± 0.30	1.505 ± 0.013	$5.45^{+0.50}_{-0.51}$	[15]
CoRoT-7 b	5.66 ± 0.68	1.562 ± 0.096	$8.19^{+1.74}_{-1.85}$	[15]
HD 3167 b	4.80 ± 0.30	$1.618^{+0.082}_{-0.057}$	$6.25^{+1.03}_{-0.76}$	[109]
K2-141 b	4.43 ± 0.76	$1.535^{+0.042}_{-0.034}$	$6.75^{+1.26}_{-1.27}$	[15]
HD 80653 b	5.35 ± 0.47	1.595 ± 0.054	$7.77^{+0.93}_{-0.92}$	[116]
K2-106 b	8.18 ± 0.84	$1.664^{+0.075}_{-0.037}$	$9.79^{1.67}_{-1.20}$	[110]
HD 213885	8.13 ± 0.086	1.814 ± 0.053	$7.51^{+1.05}_{-1.00}$	[121]
HD 20329	10.40 ± 1.20	1.685 ± 0.061	11.99 ± 1.90	[122]
K2-131	8.20 ± 1.30	$1.674^{+0.064}_{-0.054}$	$9.64^{+1.89}_{-1.79}$	[110]

Table 4.2: Derived main properties of the USP planets from this work.

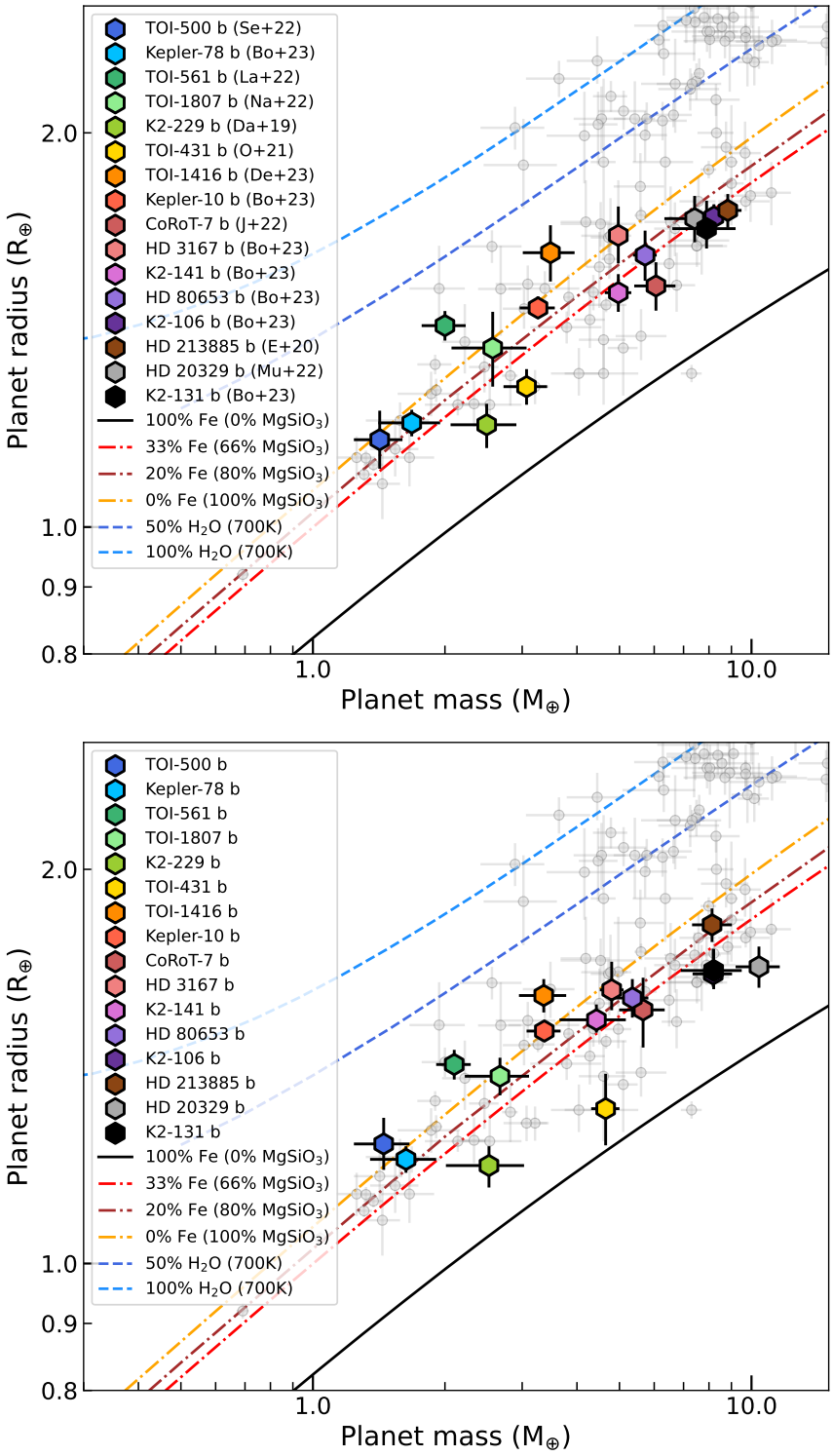


Figure 4.8: Mass-radius plot of my sample of USP planets. Top: Literature values. Bottom: Derived planetary parameters from this analysis. The theoretical mass-radius curves for the planet compositions come from Zeng et al. (2019). Background points in grey come from NASA Exoplanet Archive. References for literature plot in order of appearance: Serrano et al. 2022 [95], Bonomo et al. 2023 [117], Lacedelli et al. 2022 [98], Nardiello et al 2022 [37], Dai et al. 2019 [15], Osborn et al. 2021 [102], Deeg et al. 2023 [103], John et al. 2022 [132], Espinoza et al. 2020 [121], Murgas et al. 2022 [122]. Additional references for Rp/Rs values for bottom plot: Kokori et al. 2023 [131], Bourrier et al. 2022 [109], Frustagli et al. 2020 [116], Adams et al. 2021 [110].

5 Conclusions

This homogeneous analysis of 16 ultra-short period planets has demonstrated both the power and limitations of current methods for characterizing these extreme worlds.

A uniform approach to stellar parameter determination was taken through isochrone fitting, which combined input parameters from multiple sources, two stellar evolution models (MIST and DART) and standardized uncertainty treatments. This provided a homogeneous set of stellar parameters, which is a more consistent foundation for planetary parameter derivation.

The incorporation of activity indicators (BIS and $\log R'_{\rm HK}$), through a Trained or Multidimensional GP, was essential for reliable parameter estimation. The standardized treatment of activity across all systems enabled more reliable comparisons within this sample.

A faster approximation to the standard quasi-periodic kernel but which still retains its main characteristics was tested. The ESP kernel implemented in S+LEAF provided a good approximation to the QP kernel, showing nearly identical posterior distributions for confirmed planets while offering significant computational advantages.

Model selection criteria exhibited concerning inconsistencies across different methods. While information criteria (AIC, BIC, AICc) and Bayesian evidence should theoretically provide clear model selection guidance, it was found that different criteria often disagree on the optimal model and the reported uncertainties in nested sampling evidence calculations appear to be underestimated. Models favoured by traditional selection criteria sometimes yield poorly constrained parameters, suggesting that these criteria alone may be insufficient for model selection. This highlights the importance of combining quantitative model selection with qualitative evaluation of posterior distributions to avoid misinterpreting spurious signals as genuine planets.

When evaluating the posteriors, it was important to look not only at the strength of detection in the semi-amplitude but also to check if the period of the signal was well constrained. Some systems showed discrepancies with literature values, particularly in the detection and characterization of additional planets. These apparent additional signals sometimes did not match when using different stellar activity modelling (Trained/Multidimensional) or they did match but were a harmonic of the stellar rotation period. For these systems, a more in-depth activity

analysis is required.

In most instances, the choice of sampler was indifferent, with sampling done either by `emcee` or `dynesty` giving perfectly consistent results. However, in some cases `emcee` demonstrated greater stability in parameter estimation, making it more reliable than `dynesty` for the final parameter determination.

The USP planet parameters remain consistently well-constrained across different modelling approaches, in part due to the strong constraints coming from transit photometry. The mass-radius analysis of this sample of USP planets reveals that most planets fall within the 0-33% Fe composition range. Several outliers exist, including very low-density planets (TOI-561 b, TOI-1416 b) and high-density planets (TOI-431 b, K2-229 b). There were notable changes in the final derived parameters compared to literature values, highlighting that the treatment of stellar activity can significantly impact the results. When conducting population studies, it is crucial to apply a consistent methodology across all targets. While Gaussian Process (GP) models are computationally intensive — especially for large datasets and when combining multiple spectroscopic time series — they represent the current state-of-the-art approach. Establishing a standardized methodology would be highly beneficial for ensuring reliability and comparability in such studies.

Future work would benefit from the joint analysis of photometric and RV data to further constrain system architectures and from the development of more robust model selection criteria for GP models. A major challenge remains the proper treatment of stellar activity, which continues to introduce significant uncertainties in RV measurements. Targeted studies of stellar activity and its time-correlated noise, potentially building upon the multidimensional GP framework or using more activity indices, could help disentangle these effects from genuine planetary signals. Finally, this homogeneous analysis can be extended to a larger sample of USP planets as more high-precision RV data becomes available. Upcoming missions and spectrographs will offer even greater precision, which will enable the detection of lower-mass USPPs, extend precise RV measurements to fainter stars as well as increase the number of observations for previously detected USP planets. Together, these advancements could help provide more definitive insights into the nature of USP planets and their place in the broader context of planetary formation and evolution.

Bibliography

[1] A. Wolszczan and D. A. Frail, "A planetary system around the millisecond pulsar PSR1257 + 12," *Nature*, vol. 355, no. 6356, pp. 145–147, Jan. 1992, ISSN: 1476-4687. DOI: 10.1038/355145a0.

[2] M. Mayor and D. Queloz, "A Jupiter-mass companion to a solar-type star," *Nature*, vol. 378, no. 6555, pp. 355–359, Nov. 1995, ISSN: 1476-4687. DOI: 10.1038/378355a0.

[3] D. A. Fischer, A. W. Howard, G. P. Laughlin, *et al.*, "Exoplanet Detection Techniques," in *Protostars and Planets VI*, University of Arizona Press, 2014, ISBN: 978-0-8165-3124-0. DOI: 10.2458/azu_uapress_9780816531240-ch031.

[4] R. Sanchis-Ojeda, S. Rappaport, J. N. Winn, M. C. Kotson, A. Levine, and I. E. Mellah, "A STUDY OF THE SHORTEST-PERIOD PLANETS FOUND WITH *KEPLER*," *The Astrophysical Journal*, vol. 787, no. 1, p. 47, May 2, 2014, ISSN: 0004-637X, 1538-4357. DOI: 10.1088/0004-637X/787/1/47.

[5] K. C. Sahu, S. Casertano, H. E. Bond, *et al.*, "Transiting extrasolar planetary candidates in the Galactic bulge," *Nature*, vol. 443, no. 7111, pp. 534–540, Oct. 2006, ISSN: 0028-0836, 1476-4687. DOI: 10.1038/nature05158.

[6] A. Leger, D. Rouan, J. Schneider, *et al.*, "Transiting exoplanets from the CoRoT space mission VIII. CoRoT-7b: The first Super-Earth with measured radius," *Astronomy & Astrophysics*, vol. 506, no. 1, pp. 287–302, Oct. 2009, ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361/200911933. arXiv: 0908.0241 [astro-ph].

[7] B. E. McArthur, M. Endl, W. D. Cochran, *et al.*, "Detection of a NEPTUNE-mass planet in the ρ^{1} Cancri system using the Hobby-Eberly Telescope," *The Astrophysical Journal*, vol. 614, no. 1, pp. L81–L84, Oct. 10, 2004, ISSN: 0004-637X, 1538-4357. DOI: 10.1086/425561. arXiv: astro-ph/0408585.

[8] R. I. Dawson and D. C. Fabrycky, "RADIAL VELOCITY PLANETS DE-ALIASED: A NEW, SHORT PERIOD FOR SUPER-EARTH 55 Cnc e," *The Astrophysical Journal*, vol. 722, no. 1, p. 937, Sep. 2010, ISSN: 0004-637X. DOI: 10.1088/0004-637X/722/1/937.

[9] J. N. Winn, J. M. Matthews, R. I. Dawson, *et al.*, "A Super-Earth Transiting a Naked-Eye Star," *The Astrophysical Journal*, vol. 737, no. 1, p. L18, Aug. 10, 2011, ISSN: 2041-8205, 2041-8213. DOI: 10.1088/2041-8205/737/1/L18. arXiv: 1104.5230 [astro-ph].

[10] W. J. Borucki, D. Koch, G. Basri, *et al.*, "Kepler Planet-Detection Mission: Introduction and First Results," *Science*, vol. 327, p. 977, Feb. 1, 2010, ISSN: 0036-8075. DOI: 10.1126/science.1185402.

[11] N. M. Batalha, W. J. Borucki, S. T. Bryson, *et al.*, "*KEPLER*'S FIRST ROCKY PLANET: KEPLER-10b," *The Astrophysical Journal*, vol. 729, no. 1, p. 27, Mar. 1, 2011, ISSN: 0004-637X, 1538-4357. DOI: 10.1088/0004-637X/729/1/27.

[12] R. Sanchis-Ojeda, S. Rappaport, J. N. Winn, *et al.*, "TRANSITS AND OCCULTATIONS OF AN EARTH-SIZED PLANET IN AN 8.5 hr ORBIT," *The Astrophysical Journal*, vol. 774, no. 1, p. 54, Aug. 16, 2013, ISSN: 0004-637X, 1538-4357. DOI: 10.1088/0004-637X/774/1/54.

[13] S. B. Howell, C. Sobeck, M. Haas, *et al.*, "The K2 Mission: Characterization and Early Results," *Publications of the Astronomical Society of the Pacific*, vol. 126, no. 938, p. 398, Apr. 1, 2014, ISSN: 1538-3873. DOI: 10.1086/676406.

[14] G. R. Ricker, J. N. Winn, R. Vanderspek, *et al.*, "The Transiting Exoplanet Survey Satellite," *Journal of Astronomical Telescopes, Instruments, and Systems*, vol. 1, no. 1, p. 014 003, Oct. 24, 2014, ISSN: 2329-4124. DOI: 10.1117/1.JATIS.1.1.014003. arXiv: 1406.0151 [astro-ph].

[15] F. Dai, K. Masuda, J. N. Winn, and L. Zeng, "Homogeneous Analysis of Hot Earths: Masses, Sizes, and Compositions," *The Astrophysical Journal*, vol. 883, no. 1, p. 79, Sep. 20, 2019, ISSN: 0004-637X, 1538-4357. DOI: 10.3847/1538-4357/ab3a3b. arXiv: 1908.06299 [astro-ph].

[16] A. Isella, L. Testi, and A. Natta, "Large dust grains in the inner region of circumstellar disks," *Astronomy & Astrophysics*, vol. 451, no. 3, pp. 951–959, 3 Jun. 1, 2006, ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361:20054647.

[17] B. Pu and D. Lai, "Low-Eccentricity Migration of Ultra-Short Period Planets in Multi-Planet Systems," *Monthly Notices of the Royal Astronomical Society*, vol. 488, no. 3, pp. 3568–3587, Sep. 21, 2019, ISSN: 0035-8711, 1365-2966. DOI: 10.1093/mnras/stz1817. arXiv: 1901.08258 [astro-ph].

[18] E. J. Lee and E. Chiang, "Magnetospheric Truncation, Tidal Inspiral, and the Creation of Short and Ultra-Short Period Planets," *The Astrophysical Journal*, vol. 842, no. 1, p. 40, Jun. 10, 2017, ISSN: 0004-637X, 1538-4357. DOI: 10.3847/1538-4357/aa6fb3. arXiv: 1702.08461 [astro-ph].

[19] M. S. Lundkvist, H. Kjeldsen, S. Albrecht, *et al.*, "Hot super-Earths stripped by their host stars," *Nature Communications*, vol. 7, no. 1, p. 11 201, Apr. 11, 2016, ISSN: 2041-1723. DOI: 10.1038/ncomms11201.

[20] R. S. Ojeda, "Investigations of close-in exoplanets : Starspot transits, and ultra-short period planets," Thesis, Massachusetts Institute of Technology, 2014.

[21] G. Lacedelli, L. Malavolta, L. Borsato, *et al.*, "An unusually low density ultra-short period super-Earth and three mini-Neptunes around the old star TOI-561," *Monthly Notices of the Royal Astronomical Society*, vol. 501, no. 3, pp. 4148–4166, Jan. 15, 2021, ISSN: 0035-8711, 1365-2966. DOI: 10.1093/mnras/staa3728. arXiv: 2009.02332 [astro-ph].

[22] A. Santerne, B. Brugger, D. J. Armstrong, *et al.*, "An Earth-sized exoplanet with a Mercury-like composition," *Nature Astronomy*, vol. 2, no. 5, pp. 393–400, Mar. 26, 2018, ISSN: 2397-3366. DOI: 10.1038/s41550-018-0420-5. arXiv: 1805.08405 [astro-ph].

[23] A. Castro-González, O. D. S. Demangeon, J. Lillo-Box, *et al.*, "An unusually low-density super-Earth transiting the bright early-type M-dwarf GJ 1018 (TOI-244)," *Astronomy and Astrophysics*, vol. 675, A52, Jul. 1, 2023, ISSN: 0004-6361. DOI: 10.1051/0004-6361/202346550.

[24] R. A. Rubenzahl, F. Dai, A. W. Howard, *et al.*, "The TESS-Keck Survey. XII. A Dense 1.8 R□ Ultra-short-period Planet Possibly Clinging to a High-mean-molecular-weight Atmosphere after the First Gigayear," *The Astronomical Journal*, vol. 167, no. 4, p. 153, Mar. 2024, ISSN: 1538-3881. DOI: 10.3847/1538-3881/ad28bb.

[25] N. E. Batalha, T. Lewis, J. J. Fortney, *et al.*, "The Precision of Mass Measurements Required for Robust Atmospheric Characterization of Transiting Exoplanets," *The Astrophysical Journal*, vol. 885, p. L25, Nov. 1, 2019, ISSN: 0004-637X. DOI: 10.3847/2041-8213/ab4909.

[26] B. S. Gaudi, S. Seager, and G. Mallen-Ornelas, "On the Period Distribution of Close-in Extrasolar Giant Planets," *The Astrophysical Journal*, vol. 623, no. 1, p. 472, Apr. 10, 2005, ISSN: 0004-637X. DOI: 10.1086/428478.

[27] J. N. Winn, R. Sanchis-Ojeda, and S. Rappaport, "Kepler-78 and the Ultra-Short-Period Planets," *New Astronomy Reviews*, vol. 83, pp. 37–48, Nov. 2018, ISSN: 13876473. DOI: 10.1016/j.newar.2019.03.006. arXiv: 1803.03303 [astro-ph].

[28] C. Lovis and D. Fischer, *Radial Velocity Techniques for Exoplanets*. Dec. 1, 2010, pp. 27–53.

[29] D. A. Fischer, G. Anglada-Escude, P. Arriagada, *et al.*, "State of the Field: Extreme Precision Radial Velocities," *Publications of the Astronomical Society of the Pacific*, vol. 128, p. 066 001, Jun. 1, 2016, ISSN: 0004-6280. DOI: 10.1088/1538-3873/128/964/066001.

[30] N. C. Hara and E. B. Ford, "Statistical methods for exoplanet detection with radial velocities," *Annual Review of Statistics and Its Application*, vol. 10, no. 1, pp. 623–649, Mar. 10, 2023, ISSN: 2326-8298, 2326-831X. DOI: 10.1146/annurev-statistics-033021-012225. arXiv: 2308.00701 [astro-ph, stat].

[31] J. T. Wright and J. D. Eastman, "Barycentric Corrections at 1 cm/s for precise Doppler velocities," *Publications of the Astronomical Society of the Pacific*, vol. 126, no. 943, pp. 838–852, Sep. 2014, ISSN: 00046280, 15383873. DOI: 10.1086/678541. arXiv: 1409.4774 [astro-ph].

[32] A. Baranne, M. Mayor, and J. L. Poncet, "Coravel— A new tool for radial velocity measurements," *Vistas in Astronomy*, vol. 23, pp. 279–316, Jan. 1, 1979, ISSN: 0083-6656. DOI: 10.1016/0083-6656(79)90016-3.

[33] F. Pepe, M. Mayor, F. Galland, *et al.*, "The CORALIE survey for southern extra-solar planets VII. Two short-period Saturnian companions to HD 108147 and HD 168746," *Astronomy and Astrophysics*, vol. 388, pp. 632–638, Jun. 1, 2002, ISSN: 0004-6361. DOI: 10.1051/0004-6361:20020433.

[34] G. Anglada-Escudé and R. P. Butler, "THE HARPS-TERRA PROJECT. I. DESCRIPTION OF THE ALGORITHMS, PERFORMANCE, AND NEW MEASUREMENTS ON A FEW REMARKABLE STARS OBSERVED BY HARPS," *The Astrophysical Journal Supplement Series*, vol. 200, no. 2, p. 15, May 2012, ISSN: 0067-0049. DOI: 10.1088/0067-0049/200/2/15.

[35] A. M. Silva, J. P. Faria, N. C. Santos, *et al.*, "A novel framework for semi-Bayesian radial velocities through template matching," *Astronomy & Astrophysics*, vol. 663, A143, Jul. 2022, ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361/202142262.

[36] M. Zechmeister, M. Zechmeister, M. Zechmeister, *et al.*, "Spectrum radial velocity analyser (SERVAL). High-precision radial velocities and two alternative spectral indicators," *Astronomy and Astrophysics*, Oct. 27, 2017. DOI: 10.1051/0004-6361/201731483.

[37] D. Nardiello, L. Malavolta, S. Desidera, *et al.*, "The GAPS Programme at TNG - XXXVII. A precise density measurement of the young ultra-short period planet TOI-1807 b," *Astronomy & Astrophysics*, vol. 664, A163, Aug. 1, 2022, ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361/202243743.

[38] K. Oláh, Z. Kolláth, T. Granzer, *et al.*, "Multiple and changing cycles of active stars: II. Results," *Astronomy & Astrophysics*, vol. 501, no. 2, pp. 703–713, Jul. 2009, ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361/200811304.

[39] A. Vanderburg, P. Plavchan, J. A. Johnson, D. R. Ciardi, J. Swift, and S. R. Kane, "Radial velocity planet detection biases at the stellar rotational period," *Monthly Notices of the Royal Astronomical Society*, vol. 459, no. 4, pp. 3565–3573, Jul. 11, 2016, ISSN: 0035-8711. DOI: 10.1093/mnras/stw863.

[40] D. K. Duncan, A. H. Vaughan, O. C. Wilson, *et al.*, "Ca n H AND K MEASUREMENTS MADE AT MOUNT WILSON OBSERVATORY, 1966-1983," . .,

[41] A. S. Mascareño, R. Rebolo, J. I. G. Hernández, and M. Esposito, "Rotation periods of late-type dwarf stars from time-series high-resolution spectroscopy of chromospheric indicators," *Monthly Notices of the Royal Astronomical Society*, vol. 452, no. 3, pp. 2745–2756, Sep. 21, 2015, ISSN: 0035-8711, 1365-2966. DOI: 10.1093/mnras/stv1441. arXiv: 1506.08039 [astro-ph].

[42] C. Lovis, X. Dumusque, N. C. Santos, *et al.* "The HARPS search for southern extra-solar planets. XXXI. Magnetic activity cycles in solar-type stars: Statistics and impact on precise radial velocities," arXiv e-prints. (Jul. 1, 2011), [Online]. Available: https://ui.adsabs.harvard.edu/abs/2011arXiv1107.5325L (visited on 05/13/2024), pre-published.

[43] R. W. Noyes, L. W. Hartmann, S. L. Baliunas, D. K. Duncan, and A. H. Vaughan, "Rotation, convection, and magnetic activity in lower main-sequence stars," *The Astrophysical Journal*, vol. 279, p. 763, Apr. 1984, ISSN: 0004-637X, 1538-4357. DOI: 10.1086/161945.

[44] F. Middelkoop, "Magnetic structure in cool stars. IV - Rotation and CA II H and K emission of main-sequence stars," *Astronomy and Astrophysics*, vol. 107, pp. 31–35, Mar. 1, 1982, ISSN: 0004-6361.

[45] R. G. M. Rutten, "Magnetic structure in cool stars. VII. Absolute surface flux in CA IIH and K line cores.," *Astronomy and Astrophysics*, vol. 130, pp. 353–360, Jan. 1, 1984, ISSN: 0004-6361.

[46] L. Hartmann, D. R. Soderblom, R. W. Noyes, N. Burnham, and A. H. Vaughan, "An analysis of the Vaughan-Preston survey of chromospheric emission.," *The Astrophysical Journal*, vol. 276, pp. 254–265, Jan. 1, 1984, ISSN: 0004-637X. DOI: 10.1086/161609.

[47] N. R. Lomb, "Least-Squares Frequency Analysis of Unequally Spaced Data," *Astrophysics and Space Science*, vol. 39, pp. 447–462, Feb. 1, 1976, ISSN: 0004-640X. DOI: 10.1007/BF00648343.

[48] J. D. Scargle, "Studies in astronomical time series analysis. II. Statistical aspects of spectral analysis of unevenly spaced data.," *The Astrophysical Journal*, vol. 263, pp. 835–853, Dec. 1, 1982, ISSN: 0004-637X. DOI: 10.1086/160554.

[49] M. Zechmeister and M. Kürster, "The generalised Lomb-Scargle periodogram. A new formalism for the floating-mean and Keplerian periodograms," *Astronomy and Astrophysics*, vol. 496, pp. 577–584, Mar. 1, 2009, ISSN: 0004-6361. DOI: 10.1051/0004-6361:200811296.

[50] S. Aigrain, F. Pont, and S. Zucker, "A simple method to estimate radial velocity variations due to stellar activity using photometry," *Monthly Notices of the Royal Astronomical Society*, vol. 419, no. 4, pp. 3147–3158, Feb. 1, 2012, ISSN: 0035-8711, 1365-2966. DOI: 10.1111/j.1365-2966.2011.19960.x.

[51] M. Perger, G. Anglada-Escudé, I. Ribas, A. Rosich, E. Herrero, and J. C. Morales, "Auto-correlation functions of astrophysical processes, and their relation to Gaussian processes. Application to radial velocities of different starspot configurations," *Astronomy and Astrophysics*, vol. 645, A58, Jan. 1, 2021, ISSN: 0004-6361. DOI: 10.1051/0004-6361/202039594.

[52] B. A. Nicholson and S. Aigrain, "Quasi-periodic Gaussian processes for stellar activity: From physical to kernel parameters," *Monthly Notices of the Royal Astronomical Society*, vol. 515, pp. 5251–5266, Oct. 1, 2022, ISSN: 0035-8711. DOI: 10.1093/mnras/stac2097.

[53] R. D. Haywood, A. Collier Cameron, D. Queloz, *et al.*, "Planets and stellar activity: Hide and seek in the CoRoT-7 system," *Monthly Notices of the Royal Astronomical Society*, vol. 443, no. 3, pp. 2517–2531, Jul. 2014, ISSN: 1365-2966. DOI: 10.1093/mnras/stu1320.

[54] V. Rajpaul, S. Aigrain, M. A. Osborne, S. Reece, and S. Roberts, "A Gaussian process framework for modelling stellar activity signals in radial velocity data," *Monthly Notices of the Royal Astronomical Society*, vol. 452, pp. 2269–2291, Sep. 1, 2015, ISSN: 0035-8711. DOI: 10.1093/mnras/stv1428.

[55] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. The MIT Press, Nov. 23, 2005, ISBN: 978-0-262-25683-4. DOI: 10.7551/mitpress/3206.001.0001.

[56] J.-B. Delisle, N. Hara, and D. Ségransan, "Efficient modeling of correlated noise II. A flexible noise model with fast and scalable methods," *Astronomy & Astrophysics*, vol. 638, A95, Jun. 2020, ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361/201936906. arXiv: 2004.10678 [astro-ph, stat].

[57] J.-B. Delisle, N. Unger, N. C. Hara, and D. Ségransan, "Efficient modeling of correlated noise. III. Scalable methods for jointly modeling several observables' time series with Gaussian processes," *Astronomy & Astrophysics*, vol. 659, A182, Mar. 2022, ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361/202141949. arXiv: 2201.02440 [astro-ph].

[58] D. MacKay, *Information Theory, Inference and Learning Algorithms*. Cambridge University Press, 2003, ISBN: 978-0-521-64298-9.

[59] D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman, "Emcee: The MCMC Hammer," *Publications of the Astronomical Society of the Pacific*, vol. 125, no. 925, pp. 306–312, Mar. 2013, ISSN: 00046280, 15383873. DOI: 10.1086/670067. arXiv: 1202.3665 [astro-ph, physics:physics, stat].

[60] J. Goodman and J. Weare, "Ensemble samplers with affine invariance," *Communications in Applied Mathematics and Computational Science*, vol. 5, no. 1, pp. 65–80, 2010, ISSN: 1559-3940. DOI: 10.2140/camcos.2010.5.65.

[61] R. Kass and A. E. Raftery, "Bayes Factor," 1995.

[62] H. Jeffreys, *Theory of Probability* (International Series of Monographs on Physics). Clarendon Press, 1961.

[63] H. Akaike, "A New Look at the Statistical Model Identification," *IEEE Transactions on Automatic Control*, vol. 19, pp. 716–723, Jan. 1, 1974, ISSN: 0018-9286.

[64] G. Schwarz, "Estimating the Dimension of a Model," *Annals of Statistics*, vol. 6, pp. 461–464, Jul. 1, 1978.

[65] S. Kullback and R. A. Leibler, "On Information and Sufficiency," *The Annals of Mathematical Statistics*, vol. 22, no. 1, pp. 79–86, Mar. 1951, ISSN: 0003-4851, 2168-8990. DOI: 10.1214/aoms/1177729694.

[66] "Information and Likelihood Theory: A Basis for Model Selection and Inference," in *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, K. P. Burnham and D. R. Anderson, Eds., New York, NY: Springer, 2002, pp. 49–97, ISBN: 978-0-387-22456-5. DOI: 10.1007/978-0-387-22456-5_2.

[67] B. E. Nelson, E. B. Ford, J. Buchner, *et al.*, "Quantifying the Bayesian Evidence for a Planet in Radial Velocity Data," *The Astronomical Journal*, vol. 159, p. 73, Feb. 1, 2020, ISSN: 0004-6256. DOI: 10.3847/1538-3881/ab5190.

[68] F. Feroz and M. P. Hobson, "Multimodal nested sampling: An efficient and robust alternative to MCMC methods for astronomical data analysis," *Monthly Notices of the Royal Astronomical Society*, vol. 384, no. 2, pp. 449–463, Jan. 18, 2008, ISSN: 00358711, 13652966. DOI: 10.1111/j.1365-2966.2007.12353.x. arXiv: 0704.3704 [astro-ph].

[69] F. Feroz, M. P. Hobson, and M. Bridges, "MULTINEST: An efficient and robust Bayesian inference tool for cosmology and particle physics," *Monthly Notices of the Royal Astronomical Society*, vol. 398, pp. 1601–1614, Oct. 1, 2009, ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2009.14548.x.

[70] J. Skilling, "Nested sampling for general Bayesian computation," *Bayesian Analysis*, vol. 1, no. 4, pp. 833–859, 2006. DOI: 10.1214/06-BA127.

[71] P. Mukherjee, D. Parkinson, and A. R. Liddle, "A Nested Sampling Algorithm for Cosmological Model Selection," *The Astrophysical Journal*, vol. 638, no. 2, p. L51, Jan. 31, 2006, ISSN: 0004-637X. DOI: 10.1086/501068.

[72] J. S. Speagle, "DYNESTY: A dynamic nested sampling package for estimating Bayesian posteriors and evidences," *Monthly Notices of the Royal Astronomical Society*, vol. 493, pp. 3132–3158, Apr. 1, 2020, ISSN: 0035-8711. DOI: 10.1093/mnras/staa278.

[73] F. Pepe, M. Mayor, B. Delabre, *et al.*, "HARPS: A new high-resolution spectrograph for the search of extrasolar planets," vol. 4008, pp. 582–592, Aug. 1, 2000. DOI: 10.1117/12.395516.

[74] R. Cosentino, C. Lovis, F. Pepe, *et al.*, "Harps-N: The new planet hunter at TNG," in *Ground-Based and Airborne Instrumentation for Astronomy IV*, vol. 8446, SPIE, Sep. 24, 2012, pp. 657–676. DOI: 10.1117/12.925738.

[75] X. Dumusque, "Mitigating stellar signals in the quest for other Earths," Université de Genève, Dec. 10, 2012. DOI: 10.13097/archive-ouverte/unige:26502.

[76] S. S. Vogt, S. L. Allen, B. C. Bigelow, *et al.*, "HIRES: The high-resolution echelle spectrometer on the Keck 10-m Telescope," vol. 2198, p. 362, Jun. 1, 1994. DOI: 10.1117/12.176725.

[77] R. P. Butler, S. S. Vogt, G. Laughlin, *et al.*, "The LCES HIRES/Keck Precision Radial Velocity Exoplanet Survey," *The Astronomical Journal*, vol. 153, p. 208, May 1, 2017, ISSN: 0004-6256. DOI: 10.3847/1538-3881/aa66ca.

[78] L. J. Rosenthal, B. J. Fulton, L. A. Hirsch, *et al.*, "The California Legacy Survey I. A Catalog of 178 Planets from Precision Radial Velocity Monitoring of 719 Nearby Stars over Three Decades," *The Astrophysical Journal Supplement Series*, vol. 255, no. 1, p. 8, Jul. 1, 2021, ISSN: 0067-0049, 1538-4365. DOI: 10.3847/1538-4365/abe23c. arXiv: 2105.11583 [astro-ph].

[79] A. Quirrenbach, P. J. Amado, J. A. Caballero, *et al.*, "CARMENES instrument overview," vol. 9147, 91471F, Jul. 1, 2014. DOI: 10.1117/12.2056453.

[80] A. Quirrenbach, P. J. Amado, I. Ribas, *et al.*, "CARMENES: High-resolution spectra and precise radial velocities in the red and infrared," vol. 10702, 107020W, Jul. 1, 2018. DOI: 10.1117/12.2313689.

[81] M. Rainer, F. Borsa, and L. Affer, "Stellar masks and bisector's shape for M-type stars observed in the GAPS programme with HARPS-N at TNG," 2020, ISSN: 0922-6435. DOI: 10.1007/s10686-020-09654-z.

[82] M. J. Pecaut and E. E. Mamajek, "Intrinsic Colors, Temperatures, and Bolometric Corrections of Pre-main-sequence Stars," *The Astrophysical Journal Supplement Series*, vol. 208, p. 9, Sep. 1, 2013, ISSN: 0067-0049. DOI: 10.1088/0067-0049/208/1/9.

[83] J. N. Winn, R. Sanchis-Ojeda, L. Rogers, *et al.*, "Absence of a Metallicity Effect for Ultra-short-period Planets *," *The Astronomical Journal*, vol. 154, no. 2, p. 60, Aug. 1, 2017, ISSN: 0004-6256, 1538-3881. DOI: 10.3847/1538-3881/aa7b7c.

[84] T. D. Morton, "Isochrones: Stellar model grid package," *Astrophysics Source Code Library*, ascl:1503.010, Mar. 1, 2015.

[85] S. G. Sousa, N. C. Santos, G. Israelian, *et al.*, "Spectroscopic characterization of a sample of metal-poor solar-type stars from the HARPS planet search program. Precise spectroscopic parameters and mass estimation," *Astronomy & Astrophysics*, vol. 526, A99, Feb. 2011, ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361/201015646. arXiv: 1012.1528 [astro-ph].

[86] Gaia Collaboration, T. Prusti, J. H. J. de Bruijne, *et al.*, "The Gaia mission," *Astronomy and Astrophysics*, vol. 595, A1, Nov. 1, 2016, ISSN: 0004-6361. DOI: 10.1051/0004-6361/201629272.

[87] Gaia Collaboration, A. Vallenari, A. G. A. Brown, *et al.*, "Gaia Data Release 3. Summary of the content and survey properties," *Astronomy and Astrophysics*, vol. 674, A1, Jun. 1, 2023, ISSN: 0004-6361. DOI: 10.1051/0004-6361/202243940.

[88] A. A. Henden, D. L. Welch, D. Terrell, and S. E. Levine, "The AAVSO Photometric All-Sky Survey (APASS)," vol. 214, p. 407.02, May 1, 2009.

[89] M. F. Skrutskie, R. M. Cutri, R. Stiening, *et al.*, "The Two Micron All Sky Survey (2MASS)," *The Astronomical Journal*, vol. 131, pp. 1163–1183, Feb. 1, 2006, ISSN: 0004-6256. DOI: 10.1086/498708.

[90] E. L. Wright, P. R. M. Eisenhardt, A. K. Mainzer, *et al.*, "The Wide-field Infrared Survey Explorer (WISE): Mission Description and Initial On-orbit Performance," *The Astronomical Journal*, vol. 140, pp. 1868–1881, Dec. 1, 2010, ISSN: 0004-6256. DOI: 10.1088/0004-6256/140/6/1868.

[91] J. Choi, A. Dotter, C. Conroy, M. Cantiello, B. Paxton, and B. D. Johnson, "Mesa Isochrones and Stellar Tracks (MIST). I. Solar-scaled Models," *The Astrophysical Journal*, vol. 823, p. 102, Jun. 1, 2016, ISSN: 0004-637X. DOI: 10.3847/0004-637X/823/2/102.

[92] A. Dotter, "MESA Isochrones and Stellar Tracks (MIST) 0: Methods for the Construction of Stellar Isochrones," *The Astrophysical Journal Supplement Series*, vol. 222, p. 8, Jan. 1, 2016, ISSN: 0067-0049. DOI: 10.3847/0067-0049/222/1/8.

[93] B. Paxton, L. Bildsten, A. Dotter, F. Herwig, P. Lesaffre, and F. Timmes, "Modules for Experiments in Stellar Astrophysics (MESA)," *The Astrophysical Journal Supplement Series*, vol. 192, p. 3, Jan. 1, 2011, ISSN: 0067-0049. DOI: 10.1088/0067-0049/192/1/3.

[94] A. Dotter, B. Chaboyer, D. Jevremović, V. Kostov, E. Baron, and J. W. Ferguson, "The Dartmouth Stellar Evolution Database," *The Astrophysical Journal Supplement Series*, vol. 178, pp. 89–101, Sep. 1, 2008, ISSN: 0067-0049. DOI: 10.1086/589654.

[95] L. M. Serrano, D. Gandolfi, A. J. Mustill, *et al.*, "A low-eccentricity migration pathway for a 13-h-period Earth analogue in a four-planet system," *Nature Astronomy*, vol. 6, no. 6, pp. 736–750, Apr. 28, 2022, ISSN: 2397-3366. DOI: 10.1038/s41550-022-01641-y. arXiv: 2204.13573 [astro-ph].

[96] A. W. Howard, R. Sanchis-Ojeda, G. W. Marcy, *et al.*, "A rocky composition for an Earth-sized exoplanet," *Nature*, vol. 503, no. 7476, pp. 381–384, Nov. 21, 2013, ISSN: 0028-0836, 1476-4687. DOI: 10.1038/nature12767.

[97] F. Pepe, A. C. Cameron, D. W. Latham, *et al.*, "An Earth-sized planet with an Earth-like density," *Nature*, vol. 503, no. 7476, pp. 377–380, Nov. 21, 2013, ISSN: 1476-4687. DOI: 10.1038/nature12768. pmid: 24172902.

[98] G. Lacedelli, T. G. Wilson, L. Malavolta, *et al.*, "Investigating the architecture and internal structure of the TOI-561 system planets with CHEOPS, HARPS-N, and TESS," *Monthly Notices of the Royal Astronomical Society*, vol. 511, pp. 4551–4571, Apr. 1, 2022, ISSN: 0035-8711. DOI: 10.1093/mnras/stac199.

[99] L. M. Weiss, F. Dai, D. Huber, *et al.*, "The TESS-Keck Survey. II. An Ultra-short-period Rocky Planet and Its Siblings Transiting the Galactic Thick-disk Star TOI-561," *The Astronomical Journal*, vol. 161, no. 2, p. 56, Feb. 1, 2021, ISSN: 0004-6256, 1538-3881. DOI: 10.3847/1538-3881/abd409.

[100] J. H. Livingston, M. Endl, F. Dai, *et al.*, "44 Validated Planets from K2 Campaign 10," *The Astronomical Journal*, vol. 156, no. 2, p. 78, Aug. 1, 2018, ISSN: 0004-6256, 1538-3881. DOI: 10.3847/1538-3881/aaccde. arXiv: 1806.11504 [astro-ph].

[101] A. W. Mayo, A. Vanderburg, D. W. Latham, *et al.*, "275 Candidates and 149 Validated Planets Orbiting Bright Stars in K2 Campaigns 0-10," *The Astronomical Journal*, vol. 155, no. 3, p. 136, Mar. 1, 2018, ISSN: 0004-6256, 1538-3881. DOI: 10.3847/1538-3881/aaadff. arXiv: 1802.05277 [astro-ph].

[102] A. Osborn, D. J. Armstrong, B. Cale, *et al.*, "TOI-431/HIP 26013: A super-Earth and a sub-Neptune transiting a bright, early K dwarf, with a third RV planet," *Monthly Notices of the Royal Astronomical Society*, vol. 507, no. 2, pp. 2782–2803, Sep. 2, 2021, ISSN: 0035-8711, 1365-2966. DOI: 10.1093/mnras/stab2313. arXiv: 2108.02310 [astro-ph].

[103] H. J. Deeg, I. Y. Georgieva, G. Nowak, *et al.*, "TOI-1416: A system with a super-Earth planet with a 1.07 d period," *Astronomy & Astrophysics*, vol. 677, A12, Sep. 2023, ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361/202346370.

[104] X. Dumusque, A. S. Bonomo, R. D. Haywood, *et al.*, "THE KEPLER-10 PLANETARY SYSTEM REVISITED BY HARPS-N: A HOT ROCKY WORLD AND A SOLID NEPTUNE-MASS PLANET*," *The Astrophysical Journal*, vol. 789, no. 2, p. 154, Jun. 2014, ISSN: 0004-637X. DOI: 10.1088/0004-637X/789/2/154.

[105] G. Torres, D. A. Fischer, A. Sozzetti, *et al.*, "IMPROVED SPECTROSCOPIC PARAMETERS FOR TRANSITING PLANET HOSTS," *The Astrophysical Journal*, vol. 757, no. 2, p. 161, Sep. 2012, ISSN: 0004-637X. DOI: 10.1088/0004-637X/757/2/161.

[106] S. C. C. Barros, J. M. Almenara, M. Deleuil, *et al.*, "Revisiting the transits of CoRoT-7b at a lower activity level," *Astronomy and Astrophysics*, vol. 569, A74, Sep. 1, 2014, ISSN: 0004-6361. DOI: 10.1051/0004-6361/201423939.

[107] A. Mortier, N. C. Santos, S. G. Sousa, *et al.*, "New and updated stellar parameters for 90 transit hosts: The effect of the surface gravity□□□□□□," *Astronomy & Astrophysics*, vol. 558, A106, Oct. 2013, ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361/201322240.

[108] J. Southworth, "Homogeneous studies of transiting extrasolar planets. IV. Thirty systems with space-based light curves," *Monthly Notices of the Royal Astronomical Society*, vol. 417, no. 3, pp. 2166–2196, Nov. 1, 2011, ISSN: 00358711. DOI: 10.1111/j.1365-2966.2011.19399.x. arXiv: 1107.1235 [astro-ph].

[109] V. Bourrier, A. Deline, A. Krenn, *et al.*, "A CHEOPS-enhanced view of the HD 3167 system," *Astronomy & Astrophysics*, vol. 668, A31, Dec. 2022, ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361/202243778.

[110] E. R. Adams, B. Jackson, S. Johnson, *et al.*, "Ultra-short-period Planets in K2. III. Neighbors are Common with 13 New Multiplanet Systems and 10 Newly Validated Planets in Campaigns 0–8 and 10," *The Planetary Science Journal*, vol. 2, no. 4, p. 152, Aug. 6, 2021, ISSN: 2632-3338. DOI: 10.3847/PSJ/ac0ea0.

[111] J. L. Christiansen, A. Vanderburg, J. Burt, *et al.*, "Three's Company: An additional non-transiting super-Earth in the bright HD 3167 system, and masses for all three planets," *The Astronomical Journal*, vol. 154, no. 3, p. 122, Sep. 1, 2017, ISSN: 0004-6256, 1538-3881. DOI: 10.3847/1538-3881/aa832d. arXiv: 1706.01892 [astro-ph].

[112] D. Gandolfi, D. Gandolfi, Davide Gandolfi, *et al.*, "The Transiting Multi-planet System HD 3167: A 5.7 M □ Super-Earth and an 8.3 M □ Mini-Neptune," *arXiv: Earth and Planetary Astrophysics*, vol. 154, no. 3, p. 123, Aug. 31, 2017. DOI: 10.3847/1538-3881/aa832a.

[113] A. Vanderburg, A. Bieryla, D. A. Duev, *et al.*, "Two Small Planets Transiting HD 3167," *The Astrophysical Journal Letters*, vol. 829, no. 1, p. L9, Sep. 20, 2016, ISSN: 2041-8205, 2041-8213. DOI: 10.3847/2041-8205/829/1/L9. arXiv: 1607.05248 [astro-ph].

[114] O. Barragán, D. Gandolfi, F. Dai, *et al.*, "K2-141 b - A 5-M□ super-Earth transiting a K7 V star every 6.7 h," *Astronomy & Astrophysics*, vol. 612, A95, Apr. 1, 2018, ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361/201732217.

[115] L. Malavolta, A. W. Mayo, T. Louden, *et al.*, "An Ultra-short Period Rocky Super-Earth with a Secondary Eclipse and a Neptune-like Companion around K2-141," *The Astronomical Journal*, vol. 155, p. 107, Mar. 1, 2018, ISSN: 0004-6256. DOI: 10.3847/1538-3881/aaa5b5.

[116] G. Frustagli, E. Poretti, T. Milbourne, *et al.*, "An ultra-short period rocky super-Earth orbiting the G2-star HD 80653," *Astronomy & Astrophysics*, vol. 633, A133, Jan. 2020, ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361/201936689. arXiv: 2001.02217 [astro-ph].

[117] A. S. Bonomo, X. Dumusque, A. Massa, *et al.*, "Cold Jupiters and improved masses in 38 Kepler and K2 small planet systems from 3661 HARPS-N radial velocities. No excess of cold Jupiters in small planet systems," *Astronomy & Astrophysics*, vol. 677, A33, Sep. 2023, ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361/202346211. arXiv: 2304.05773 [astro-ph].

[118] E. W. Guenther, O. Barragan, F. Dai, *et al.*, "K2-106, a system containing a metal-rich planet and a planet of lower density," *Astronomy & Astrophysics*, vol. 608, A93, Dec. 2017, ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361/201730885. arXiv: 1705.04163 [astro-ph].

[119] E. Sinukoff, A. W. Howard, E. A. Petigura, *et al.*, "K2-66b and K2-106b: Two extremely hot sub-Neptune-size planets with high densities," *The Astronomical Journal*, vol. 153, no. 6, p. 271, Jun. 1, 2017, ISSN: 0004-6256, 1538-3881. DOI: 10.3847/1538-3881/aa725f. arXiv: 1705.03491 [astro-ph].

[120] E. R. Adams, B. Jackson, M. Endl, *et al.*, "Ultra Short Period Planets in K2 with companions: A double transiting system for EPIC 220674823," *The Astronomical Journal*, vol. 153, no. 2, p. 82, Feb. 1, 2017, ISSN: 0004-6256, 1538-3881. DOI: 10.3847/1538-3881/153/2/82. arXiv: 1611.00397 [astro-ph].

[121] N. Espinoza, R. Brahm, T. Henning, *et al.*, "HD 213885b: A transiting 1-day-period super-Earth with an Earth-like composition around a bright ($V=7.9$) star unveiled by TESS," *Monthly Notices of the Royal Astronomical Society*, vol. 491, no. 2, pp. 2982–2999, Jan. 11, 2020, ISSN: 0035-8711, 1365-2966. DOI: 10.1093/mnras/stz3150. arXiv: 1903.07694 [astro-ph].

[122] F. Murgas, G. Nowak, T. Masseron, *et al.*, "HD 20329b: An ultra-short-period planet around a solar-type star found by TESS," *Astronomy & Astrophysics*, vol. 668, A158, Dec. 2022, ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361/202244459. arXiv: 2211.02547 [astro-ph].

[123] F. Dai, J. N. Winn, D. Gandolfi, *et al.*, "The Discovery and Mass Measurement of a New Ultra-short-period Planet: K2-131b," *The Astronomical Journal*, vol. 154, no. 6, p. 226, Nov. 2017, ISSN: 1538-3881. DOI: 10.3847/1538-3881/aa9065.

[124] J. Gomes da Silva, P. Figueira, N. Santos, and J. Faria, "ACTIN: A tool to calculate stellar activity indices," *The Journal of Open Source Software*, vol. 3, p. 667, Nov. 1, 2018. DOI: 10.21105/joss.00667.

[125] J. G. da Silva, N. C. Santos, V. Adibekyan, *et al.*, "Stellar chromospheric activity of 1674 FGK stars from the AMBRE-HARPS sample - I. A catalogue of homogeneous chromospheric activity," *Astronomy & Astrophysics*, vol. 646, A77, Feb. 1, 2021, ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361/202039765.

[126] N. Buchschacher, D. Ségransan, S. Udry, and R. Díaz, "Data and Analysis Center for Exoplanets," vol. 495, p. 7, Sep. 1, 2015.

[127] D. Queloz, F. Bouchy, C. Moutou, *et al.*, "The CoRoT-7 planetary system: Two orbiting super-Earths," *Astronomy & Astrophysics*, vol. 506, no. 1, pp. 303–319, Oct. 2009, ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361/200913096.

[128] L. Malavolta, "PyORBIT: Exoplanet orbital parameters and stellar activity," *Astrophysics Source Code Library*, ascl:1612.008, Dec. 1, 2016.

[129] V. Van Eylen, S. Albrecht, X. Huang, *et al.*, "The Orbital Eccentricity of Small Planet Systems," *The Astronomical Journal*, vol. 157, p. 61, Feb. 1, 2019, ISSN: 0004-6256. DOI: 10.3847/1538-3881/aaf22f.

[130] S. Bocquet and F. W. Carter, "Pygtc: Beautiful parameter covariance plots (aka. giant triangle confusograms)," *The Journal of Open Source Software*, vol. 1, no. 6, Oct. 2016. DOI: 10.21105/joss.00046.

[131] A. Kokori, A. Tsiaras, B. Edwards, *et al.*, "ExoClock Project. III. 450 New Exoplanet Ephemerides from Ground and Space Observations," *The Astrophysical Journal Supplement Series*, vol. 265, p. 4, Mar. 1, 2023, ISSN: 0067-0049. DOI: 10.3847/1538-4365/ac9da4.

[132] A. A. John, A. Collier Cameron, and T. G. Wilson, "The impact of two non-transiting planets and stellar activity on mass determinations for the super-Earth CoRoT-7b," *Monthly Notices of the Royal Astronomical Society*, vol. 515, pp. 3975–3995, Sep. 1, 2022, ISSN: 0035-8711. DOI: 10.1093/mnras/stac1814.

[133] J. H. Livingston, I. J. M. Crossfield, E. A. Petigura, *et al.*, "Sixty Validated Planets from K2 Campaigns 5-8," *The Astronomical Journal*, vol. 156, p. 277, Dec. 1, 2018, ISSN: 0004-6256. DOI: 10.3847/1538-3881/aae778.

A

Stellar parameter priors for isochrone fitting

Table A.1: Stellar parameters of the sample for isochrone fitting (Part 1)

Parameter	TOI-500	Kepler-78	TOI-561	TOI-1807	K2-229	TOI-431
Identifying Information						
TIC ID	377064495	270701667	377064495	180695581	98720809	-
2MASS ID	J09524454+0612589	J19345800+4426539	J09524454+0612589	J13250800+3855210	J12272958-0643188	J05330459-2643286
GAIA ID	3850421005290172416	2078373642776670080	3850421005290172416	1476485996883837184	3583630934082556416	2908664557091200768
R.A. (J2000, h:m:s)	$09^h52^m44.55^s$	$19^h34^m58.01^s$	$09^h52^m44.55^s$	$13^h25^m08.00^s$	$12^h27^m29.58^s$	053304.60
DEC (J2000, d:m:s)	$-47^o35'13.87''$	$44^o26'53.96''$	$06^o12'58.92''$	$+38^o55'20.95''$	$-06^o43'18.77''$	$-26^o43'28.26''$
Input parameters						
Parallax (mas)	21.0928 ± 0.0104	8.0085 ± 0.0104	11.8342 ± 0.0208	23.4804 ± 0.0142	9.7621 ± 0.0215	30.6517 ± 0.0141
B (mag)	11.668 ± 0.050	12.865 ± 0.337	10.965 ± 0.082	11.082 ± 0.057	12.049 ± 0.186	10.104 ± 0.030
V (mag)	10.54 ± 0.03	11.717 ± 0.034	10.252 ± 0.006	10.00 ± 0.03	10.985 ± 0.017	9.12 ± 0.03
J (mag)	8.403 ± 0.024	10.184 ± 0.025	8.879 ± 0.020	8.103 ± 0.023	9.518 ± 0.022	7.305 ± 0.026
H (mag)	7.848 ± 0.038	9.675 ± 0.018	8.504 ± 0.055	7.605 ± 0.018	9.126 ± 0.022	6.846 ± 0.029
K (mag)	7.715 ± 0.026	9.586 ± 0.015	8.394 ± 0.019	$-$	9.050 ± 0.023	6.723 ± 0.021
W_1 (mag)	7.63 ± 0.03	9.529 ± 0.022	8.337 ± 0.023	7.395 ± 0.032	8.969 ± 0.024	6.651 ± 0.073
W_2 (mag)	7.736 ± 0.020	9.584 ± 0.019	8.396 ± 0.020	7.508 ± 0.020	9.036 ± 0.021	6.709 ± 0.023
W_3 (mag)	7.658 ± 0.018	9.580 ± 0.034	8.375 ± 0.023	7.445 ± 0.017	8.989 ± 0.032	6.670 ± 0.017
W_4 (mag)	7.617 ± 0.124	9.208 ± 0.482	7.971 ± 0.260	7.368 ± 0.115	8.616 ± 0.456	6.539 ± 0.058
$T_{eff}(K)$	4440 ± 100, [95]	5121 ± 44, [96] 5058 ± 50, [97] 5089 ± 50, [12]	5372 ± 70, [98] 5326 ± 64,[99] 5455^{+65}_{-47}, [21]	4730 ± 75, [37] 5315^{+35}_{-31}, [100] 5231 ± 50, [101]	5185 ± 32, [22]	4850 ± 75, [102]
$\log g$ (cgs)	4.618 ± 0.017, [95]	4.61 ± 0.06, [96] 4.55 ± 0.10, [97] 4.60 ± 0.10, [12]	4.50 ± 0.12, [98] 4.55 ± 0.10, [96] 4.47 ± 0.01, [98]	4.55 ± 0.10, [37] $4.59^{+0.00}_{-0.01}$, [100] 4.59 ± 0.01, [101]	$4.56^{+0.03}_{-0.05}$, [22]	4.60 ± 0.06, [102]
[Fe/H] (dex)	0.12 ± 0.08, [95]	-0.08 ± 0.04, [96] -0.18 ± 0.08, [97] -0.14 ± 0.08, [12]	-0.40 ± 0.05, [98] -0.41 ± 0.05, [96] $-0.33^{+0.10}_{-0.05}$, [98]	-0.04 ± 0.03, [37] -0.09 ± 0.02, [100] -0.03 ± 0.08, [101]	-0.06 ± 0.02, [22]	0.20 ± 0.05, [102]

Table A.1: Stellar parameters of the sample for isochrone fitting (Part 2)

Parameter	TOI-1416	Kepler-10	CoRoT-7	HD 3167	K2-141
Identifying Information					
TIC ID	-	-	-	-	
2MASS ID	J14274177+4157124	J19024305+5014286	J06434947-0103468	J00345752+0422531	J23233996-0111215
GAIA ID	1491634483976350720	2132155017099178624	3107267177757848576	2554032474712538880	2643952940813536768
R.A. (J2000, h:m:s)	$14^h27^m41.76^s$	$19^h02^m43.06^s$	$06^h43^m49.47^s$	$00^h34^m57.52^s$	$23^h23m^39.97^s$
DEC (J2000, d:m:s)	$+41^o57'12.32''$	$+50^o14'28.70''$	$-01^o03'46.83''$	$+04^o22'53.28''$	$-01^o11'21.44''$
Input parameters					
Parallax (mas)	18.1671 ± 0.0126	5.3698 ± 0.0103	6.2676 ± 0.0139	21.1363 ± 0.0187	16.1303 ± 0.0176
B (mag)	10.932 ± 0.062	12.258 ± 0.067	12.784 ± 0.307	9.767 ± 0.031	12.139 ± 0.329
V (mag)	9.91 ± 0.03	11.043 ± 0.023	11.728 ± 0.027	8.97 ± 0.03	11.454 ± 0.023
J (mag)	8.266 ± 0.024	9.889 ± 0.022	10.301 ± 0.024	7.548 ± 0.032	9.086 ± 0.021
H (mag)	7.815 ± 0.017	9.563 ± 0.025	9.880 ± 0.022	7.203 ± 0.038	8.524 ± 0.053
K (mag)	7.708 ± 0.024	9.496 ± 0.021	9.806 ± 0.021	7.066 ± 0.020	8.401 ± 0.023
W_1 (mag)	7.645 ± 0.028	9.440 ± 0.023	9.760 ± 0.022	6.942 ± 0.053	8.311 ± 0.024
W_2 (mag)	7.739 ± 0.020	9.491 ± 0.020	9.824 ± 0.019	7.09 ± 0.02	8.391 ± 0.020
W_3 (mag)	7.686 ± 0.018	9.443 ± 0.029	9.782 ± 0.064	7.054 ± 0.018	8.311 ± 0.023
W_4 (mag)	7.601 ± 0.097	9.403 ± 0.482	7.926	7.065 ± 0.144	7.930 ± 0.223
$T_{\rm eff}$ (K)	4884 ± 70, [103]	5708 ± 28, [104] 5643 ± 75, [105] 5627 ± 44, [11]	5259 ± 58, [106] 5288 ± 27, [107] 5313 ± 73, [105] 5250 ± 60, [108] 5275 ± 75, [6]	5261 ± 60, [117] 5300 ± 73, [109] 5438 ± 63, [110] 5528 ± 162, [133] 5366.0 ± 50.0, [101] 5261 ± 60, [111] 5286 ± 40, [112] 5370 ± 50, [113]	4373 ± 57, [114] 4599 ± 79, [115] 4570 ± 100, [117]
$\log g$ (cgs)	4.52 ± 0.05, [103]	4.344 ± 0.004, [104] 4.34 ± 0.01, [105] 4.341 ± 0.012, [11]	4.573 ± 0.011, [106] 4.40 ± 0.07, [107] 4.54 ± 0.04, [105] 4.42 ± 0.14, [108]	4.47 ± 0.12, [109] 4.6 ± 0.1, [110] 4.53 ± 0.02, [133] 4.54 ± 0.1, [101] 4.47 ± 0.05, [111] 4.51 ± 0.03, [112] 4.52 ± 0.03, [113]	4.584 ± 0.051, [114] $4.62^{+0.02}_{-0.03}$, [115]
[Fe/H] (dex)	0.080 ± 0.050, [103]	-0.15 ± 0.04, [104] 0.15 ± 0.07, [105] -0.15 ± 0.04, [11]	0.138 ± 0.061, [106] 0.02 ± 0.02, [107] 0.03 ± 0.07, [105]	0.04 ± 0.05, [117] 0.037 ± 0.048, [109] -0.10 ± 0.10, [110] -0.020 ± 0.170, [133] -0.01 ± 0.08, [101] 0.04 ± 0.05, [111] 0.03 ± 0.03, [112] 0.0 ± 0.1, [113]	0.030 ± 0.100, [114] $-0.060^{+0.080}_{-0.100}$, [115] 0.0 ± 0.1, [117]

Table A.1: Stellar parameters of the sample for isochrone fitting (Part 3)

Parameter	HD 80653	K2-106	HD 213885	HD 20329	k2-131
Identifying Information					
TIC ID	-	-	-	-	
2MASS ID	J09212142+1422046	J00521914+1047409	J22355630-5951522	J03164262+1539260	J12110036-0945547
GAIA ID	606477252238780160	2582617711154563968	6407428994690988928	30398648945512960	3580920878437938688
R.A. (J2000, h:m:s)	$09^h21^m21.42^s$	$00h^52^m19.14^s$	$22^h35^m56.32^s$	$03^h16^m42.63^s$	$12^h11^m00.38^s$
DEC (J2000, d:m:s)	$+14^o22'04.55''$	$10^o47'40.91''$	$-59^o51'52.08''$	$+15^o39'26.01''$	$-09^o45'54.78''$
Input parameters					
Parallax (mas)	9.26 ± 0.0223	4.0854 ± 0.018	20.9317 ± 0.0218	15.6585 ± 0.0232	6.4786 ± 0.0212
B (mag)	10.119 ± 0.047	12.576 ± 0.443	8.495 ± 0.001	9.43 ± 0.08	13.054 ± 0.036
V (mag)	9.452 ± 0.003	12.101 ± 0.031	7.95 ± 0.03	8.760 ± 0.012	12.129 ± 0.092
J (mag)	8.315 ± 0.023	10.770 ± 0.023	6.806 ± 0.023	7.492 ± 0.021	10.570 ± 0.026
H (mag)	8.079 ± 0.029	10.454 ± 0.026	6.501 ± 0.034	7.208 ± 0.049	10.122 ± 0.022
K (mag)	8.018 ± 0.021	10.344 ± 0.021	6.419 ± 0.024	7.116 ± 0.024	10.051 ± 0.026
W_1 (mag)	7.959 ± 0.024	10.299 ± 0.023	6.370 ± 0.087	7.054 ± 0.041	9.950 ± 0.024
W_2 (mag)	8.004 ± 0.020	10.355 ± 0.021	6.281 ± 0.024	7.105 ± 0.020	9.9970 ± 0.0210
W_3 (mag)	8.011 ± 0.021	10.380 ± 0.091	6.370 ± 0.015	7.120 ± 0.017	9.907 ± 0.057
W_4 (mag)	7.869 ± 0.204	8.423	6.268 ± 0.067	7.111 ± 0.108	8.667
$T_{\rm eff}$ (K)	5959 ± 61, [117]	5532 ± 62, [117] 5613 ± 39, [110] 5617 ± 86, [133] 5590 ± 50, [101] 5470 ± 30, [118] 5496 ± 46, [119] 5590 ± 51, [120]	5978 ± 50, [121]	5596 ± 50, [122]	5120 ± 71, [117] 5200 ± 100, [15] 5245^{+46}_{-52}, [100] 5113 ± 50, [101]
$\log g$ (cgs)	4.34 ± 0.02, [116]	4.60 ± 0.07, [110] 4.45 ± 0.03, [133] 4.54 ± 0.1, [101] 4.53 ± 0.08, [118] 4.42 ± 0.05, [119] 4.56 ± 0.09, [120]	4.3827 ± 0.0096, [121]	4.40 ± 0.07, [122]	4.62 ± 0.10, [15] 4.62 ± 0.10, [100] 4.5 ± 0.1, [101]
[Fe/H] (dex)	0.255 ± 0.065, [117]	0.11 ± 0.05, [117] 0.01 ± 0.01, [110] 0.13 ± 0.06, [133] 0.08 ± 0.08, [101] -0.025 ± 0.050, [118] 0.06 ± 0.03, [119] 0.025 ± 0.020, [120]	-0.04 ± 0.03, [121]	-0.070 ± 0.060, [122]	-0.040 ± 0.070, [117] -0.02 ± 0.08, [15] -0.02 ± 0.08, [100] -0.15 ± 0.08, [101]

B

Planetary and stellar priors for MCMC and NS analyses

Table B.1: Planetary and stellar parameter priors for the analysis of spectroscopic series with the GP framework (Part 1). These represent "Model 0" where no additional RV signal other than the signals from the transiting planets are considered.

Parameter	TOI-500	Kepler-78	TOI-561	TOI-1807	K2-229	TOI-431
Planet b						
P (days)	$\mathcal{N}(0.5481573, 0.0000089)$	$\mathcal{N}(0.35500744, 0.00000006)$	$\mathcal{N}(0.4465688, 0.0000008)$	$\mathcal{N}(0.549374, 0.000013)$	$\mathcal{N}(0.584240, 0.000014)$	$\mathcal{N}(0.4900657, 0.001)$
T_c ($\mathrm{BJD}_{TDB}-2400000$)	$\mathcal{N}(58468.39103, 0.00057)$	$\mathcal{N}(54953.95995, 0.00015)$	$\mathcal{N}(59317.7498, 0.0005)$	$\mathcal{N}(58899.3449, 0.0008)$	$\mathcal{N}(57582.88613, 0.001057)$	$\mathcal{N}(58627.533, 0.1)$
i (Fixed)	82.09 ± 2.53	75.2 ± 2.6	87.2 ± 2.1	82.2 ± 2.0	81.6 ± 10.8	84.3 ± 1.3
e	-	-	-	-	-	-
K (ms^{-1})	$\mathcal{U}(0.001, 20.0)$	$\mathcal{U}(0.001, 20.0)$	$\mathcal{U}(0.001, 20.0)$	$\mathcal{U}(0.001, 20.0)$	$\mathcal{U}(0.01, 20.0)$	$\mathcal{U}(0.01, 20.0)$
Planet c						
P (days)	-	-	$\mathcal{N}(10.778831, 0.000034)$	-	$\mathcal{N}(8.328010, 0.000405)$	$\mathcal{N}(12.46, 0.01)$
T_c ($\mathrm{BJD}_{TDB}-2400000$)	-	-	$\mathcal{N}(59238.4629, 0.0008)$	-	$\mathcal{N}(57586.33508, 0.00167)$	$\mathcal{N}(58627.545, 0.1)$
i (Fixed)	-	-	89.69 ± 0.31	-	88.33 ± 3.07	-
e	-	-	-	-	-	-
K (ms^{-1})	-	-	$\mathcal{U}(0.01, 20.0)$	-	$\mathcal{U}(0.01, 100.0)$	$\mathcal{U}(0.01, 20.0)$
Planet d						
P (days)	-	-	$\mathcal{N}(25.7124, 0.0002)$	-	-	
T_c ($\mathrm{BJD}_{TDB}-2400000$)	-	-	$\mathcal{N}(59318.966, 0.004)$	-	-	
i (Fixed)	-	-	89.40 ± 0.21	-	-	
e	-	-	$\mathcal{HN}(0.00, 0.098)$	-		
K (ms^{-1})	-	-	$\mathcal{U}(0.01, 20.0)$	-	-	
Planet e						
P (days)	-	-	$\mathcal{U}(50.00, 100.0)$	-	-	
T_c ($\mathrm{BJD}_{TDB}-2400000$)	-	-	$\mathcal{N}(58538.180, 0.005)$	-	-	
i (Fixed)	-	-	-	-	-	
e	-	-	$\mathcal{HN}(0.00, 0.098)$	-	-	
K (ms^{-1})	-	-	$\mathcal{U}(0.01, 20.0)$	-	-	
Stellar parameters						
$M_*(M_\odot)$	$\mathcal{N}(0.764, 0.009)$	$\mathcal{N}(0.829, 0.007)$	$\mathcal{N}(0.782, 0.022)$	$\mathcal{N}(0.776, 0.007)$	$\mathcal{N}(0.848790, 0.019005)$	$\mathcal{N}(0.822403, 0.020389)$
Stellar activity						
P_{rot} (days)	$\mathcal{N}(45, 5)$	$\mathcal{N}(12.5, 1.0)$	$\mathcal{U}(10, 50)$	$\mathcal{U}(8, 10)$	$\mathcal{U}(5, 10)$	$\mathcal{N}(30.5, 0.7)$
P_{dec} (days)	$\mathcal{U}(20, 1000)$	$\mathcal{U}(20, 1000)$	$\mathcal{U}(20, 1000)$	$\mathcal{U}(10, 1000)$	$\mathcal{U}(10, 1000)$	$\mathcal{U}(10, 1000)$
O_{amp}	$\mathcal{U}(0.001, 1.0)$	$\mathcal{U}(0.001, 1.0)$	$\mathcal{U}(0.001, 1.0)$	$\mathcal{U}(0.01, 1.0)$	$\mathcal{U}(0.01, 1.0)$	$\mathcal{U}(0.01, 1.0)$

Table B.1: Planetary and stellar parameter priors for the analysis of spectroscopic series with the GP framework (Part 2). These represent "Model 0" where no additional RV signal other than the signals from the transiting planets are considered.

Parameter	TOI-1416	Kepler-10	CoRoT-7	HD 3167	K2-141
Planet b					
P (days)	$\mathcal{N}(1.0697611, 0.0000037)$	$\mathcal{N}(0.8374907, 0.0000002)$	$\mathcal{N}(0.853585, 0.000024)$	$\mathcal{N}(0.959628, 0.000012)$	$\mathcal{N}(0.280324956, 0.000000067)$
T_c (BJD$_{TDB}-2400000$)	$\mathcal{N}(59689.4057990, 0.0027278)$	$\mathcal{N}(55034.08687, 0.00018)$	$\mathcal{N}(54398.0767, 0.0015)$	$\mathcal{N}(57394.37450, 0.00044)$	$\mathcal{N}(57744.071508, 0.000103)$
i (Fixed)	85.7 ± 1.7	84.8 ± 3.9	80.98 ± 0.51	83.4 ± 7.7	86.3 ± 3.6
e	-	-	-	-	-
K (ms^{-1})	$\mathcal{U}(0.001, 20.0)$	$\mathcal{U}(0.001, 20.0)$	$\mathcal{U}(0.001, 20.0)$	$\mathcal{U}(0.001, 20.0)$	$\mathcal{U}(0.01, 20.0)$
Planet c					
P (days)	-	$\mathcal{N}(45.294301, 0.000048)$	-	$\mathcal{N}(29.8454, 0.0012)$	$\mathcal{N}(7.7489943, 0.0000149)$
T_c (BJD$_{TDB}-2400000$)	-	$\mathcal{N}(55062.26648, 0.00081)$	-	$\mathcal{N}(57394.9788, 0.0012)$	$\mathcal{N}(58371.07415, 0.000652)$
i (Fixed)	-	-	-	-	-
e	-	-	-	-	-
K (ms^{-1})	-	$\mathcal{U}(0.01, 100.0)$	-	$\mathcal{U}(0.01, 100.0)$	$\mathcal{U}(0.01, 100.0)$
Stellar parameters					
$M_\star(M_\odot)$	$\mathcal{N}(0.816, 0.004)$	$\mathcal{N}(0.929, 0.023)$	$\mathcal{N}(0.918, 0.011)$	$\mathcal{N}(0.776, 0.007)$	$\mathcal{N}(0.731098, 0.021942)$
Stellar activity					
$P_{\rm rot}$ (days)	$\mathcal{N}(17.6, 2.0)$	$\mathcal{U}(15, 30)$	$\mathcal{U}(15, 30)$	$\mathcal{N}(23.52, 2.87)$	$\mathcal{U}(10, 20)$
$P_{\rm dec}$ (days)	$\mathcal{U}(20, 1000)$	$\mathcal{U}(20, 1000)$	$\mathcal{U}(20, 1000)$	$\mathcal{U}(10, 1000)$	$\mathcal{U}(10, 1000)$
O_{amp}	$\mathcal{U}(0.001, 1.0)$	$\mathcal{U}(0.001, 1.0)$	$\mathcal{U}(0.001, 1.0)$	$\mathcal{U}(0.01, 1.0)$	$\mathcal{U}(0.01, 1.0)$

Table B.1: Planetary and stellar parameter priors for the analysis of spectroscopic series with the GP framework (Part 3). These represent "Model 0" where no additional RV signal other than the signals from the transiting planets are considered.

Parameter	HD 80653	K2-106	HD 213885	HD 20329	K2-131
Planet b					
P (days)	$\mathcal{N}(0.719573, 0.000021)$	$\mathcal{N}(0.571292, 0.000012)$	$\mathcal{N}(1.0079, 0.0001)$	$\mathcal{U}(0.5, 1.5)$	$\mathcal{N}(0.3693038, 0.0000091)$
T_c (BJD$_{TDB}-2400000$)	$\mathcal{N}(58134.4244, 0.0007)$	$\mathcal{N}(57394.01140, 0.00010)$	$\mathcal{N}(58379.9647, 0.0001)$	$\mathcal{U}(59471.7445, 59472.5445)$	$\mathcal{N}(57582.9360, 0.0011)$
i (Fixed)	86.3 ± 2.4	86.4 ± 4.1	80.09 ± 0.62	76.01 ± 0.46	85.0 ± 10.0
e	-	-	-	-	-
K (ms^{-1})	$\mathcal{U}(0.001, 20.0)$	$\mathcal{U}(0.001, 20.0)$	$\mathcal{U}(0.001, 20.0)$	$\mathcal{U}(0.001, 20.0)$	$\mathcal{U}(0.01, 20.0)$
Planet c					
P (days)	-	$\mathcal{N}(13.33970, 0.00096)$	-	-	$\mathcal{N}(7.7489943, 0.0000149)$
T_c (BJD$_{TDB}-2400000$)	-	$\mathcal{N}(57405.7316, 0.0044)$	-	-	-
i (Fixed)	-	-	-	-	-
e	-	-	-	-	-
K (ms^{-1})	-	$\mathcal{U}(0.01, 100.0)$	-	-	-
Stellar parameters					
$M_\star (M_\odot)$	$\mathcal{N}(1.171, 0.031)$	$\mathcal{N}(0.941, 0.039)$	$\mathcal{N}(1.023, 0.027)$	$\mathcal{N}(0.954, 0.019)$	$\mathcal{N}(0.807, 0.030)$
Stellar activity					
$P_{\rm rot}$ (days)	$\mathcal{U}(10, 30)$	$\mathcal{U}(15, 30)$	$\mathcal{U}(1, 30)$	$\mathcal{U}(10, 40)$	$\mathcal{N}(9.37, 1.859.37, 1.85)$
$P_{\rm dec}$ (days)	$\mathcal{U}(20, 1000)$	$\mathcal{U}(20, 1000)$	$\mathcal{U}(20, 1000)$	$\mathcal{U}(10, 1000)$	$\mathcal{U}(10, 1000)$
O_{amp}	$\mathcal{U}(0.001, 1.0)$	$\mathcal{U}(0.001, 1.0)$	$\mathcal{U}(0.001, 1.0)$	$\mathcal{U}(0.01, 1.0)$	$\mathcal{U}(0.01, 1.0)$

Acknowledgements

I would like to sincerely thank my supervisor, Professor Luca Malavolta, for his unwavering kindness, patience, and insightful guidance throughout our meetings. I'm grateful for my friends from the Azores and from Lisbon who, despite the distance, I still feel close to my heart. I owe a lot to my family for their unconditional love and for always rooting for me. Lastly, I want to thank the amazing people I met in Padova and with whom I spent countless hours throughout the making of this thesis. Their generosity, humour, and support have made these past two years an unforgettable, laughter-filled chapter of my life.