Master’s Thesis
submitted in partial fulfilment of the requirements for the
Erasmus+: Erasmus Mundus Joint Masters Degree programme in
Astronomy and Astrophysics

CARMENES radial velocity with data driven methods

Douglas Rodrigues Alves

Institute for Astrophysics

Master of Science Thesis of
the Institute for Astrophysics at the
Georg-August-Universität Göttingen

September 10, 2019
Georg-August-Universität Göttingen
Institute for Astrophysics

Friedrich-Hund-Platz 1
37077 Göttingen
Germany

☎ +49 (0)551 39-5042
✉ +49 (0)551 39-5043
✉ sekr@astro.physik.uni-goettingen.de
},'https://www.uni-goettingen.de/en/203293.html

First Supervisor: Prof. Dr. Ansgar Reiners
Second Supervisor: Dr. Mathias Zechmeister
I hereby declare that I have written this thesis independently without any help from others and without the use of documents or aids other than those stated. I have mentioned all used sources and cited them correctly according to established academic citation rules.

Göttingen, September 10, 2019
Abstract

The radial velocity (RV) technique is utilized in several exoplanet detection surveys, being responsible for the discovery of about one thousand planets. Technology breakthroughs allowed spectrographs to measure signals in the 1 m s\(^{-1}\) level, thus expanding the possibility to explore the low-mass planetary regime. Yet, data analysis faces significant challenges to extract RVs at its highest precision due to stellar activity or atmospheric contamination. Therefore, in this work, we focus on the removal of telluric spectra from the CARMENES survey in order to compute the highest precision RVs from data. For this task we utilize a recently developed python-based code called wobble that models stellar- and telluric spectrum while simultaneously extracting the RVs which are thereafter compared to serval – the official CARMENES high-precision RV pipeline. We select four targets with different properties for inquiring wobble’s performance in different stellar regimes. Barnard’s star which is relatively quiet, Teegarden’s star and J08413+594 presenting visible H\(\alpha\) activity and Iot Cyg, a standard telluric star that contains fewer stellar lines. We were able to acknowledge that the code performance on archival CARMENES spectra display minor issues concerning continuum normalization and sub-optimal modeling in Échelle orders containing heavy telluric contamination. Moreover, the results imply that significant stellar activity may induce poor selection of wobble’s regularization amplitudes. Yet, the code was able to calculate RV time-series in agreement with serval’s for all the stars in the sample, and particularly for J08413+594, it reaches slightly better RVs with \(r_{\text{msRV}} = 46.03\) m s\(^{-1}\) compared to serval \(r_{\text{msRV}} = 46.44\) m s\(^{-1}\).

Key-words: <Radial Velocity>, <Exoplanets>, <Telluric modeling>, <wobble>, <Barnard’s star>, <J08413+594>, <HIP 95853>, <Teegarden’s star>, <serval>
Acknowledgement

I would like to thank my family, in special my mother for the unconditional support given to me when I decided to be an Astrophysicist. Thank you for the understanding when I said I was going to study in a master program far away from home, thank you for everything. To my sister who is always in good mood, spreading happiness wherever she passes in. Thank you so much for coming over and bringing mother and niece along so that we could spend Christmas together. It was just the best days ever, I wish you stayed more! To the kids I love so much: my niece Luisa Rodrigues and my two nephews Théo Martins and Heitor Rodrigues. Your energies, joy, and laughter are contagious, thank you for filling my life with happiness and for the warm welcoming whenever I come visit you.

To my co-supervisor Dr. Mathias Zechmeister who I am greatly indebted for all the times I knocked at his door and he patiently explained me all the aspects of this research. Thank you for this and the countless advice in programming (I have learned a lot with you!). I would like to thank my supervisor Prof. Ansgar Reiners who welcomed me in the CARMENES group; without you this work would have not been possible. Thank you for the instructions on how we should proceed with the research and also the exoplanet class in my 3rd semester, it was crucial to my knowledge of this subject and, most importantly, it was fun (specially the quizzes).

I would like to thank my Astromundus friends, specially the ones that spent almost the 2-year program with me. Your friendship, the trips together and the countless gathering to study for exams softened the burdens of this master program. To my non-astromundus friends, specially to the one taking her PhD in Chile who I am deeply indebted to for many valuable advice concerning the academia, thank you so much Ana.

Last but not least, I would like to thank the ASTROMUNDUS program, the professors who were responsible to broaden the horizon of my comprehension of Astronomy and the secretary Andrea Rumpold who has been very helpful throughout our entire stay in Europe. It’s been an immense honor to participate in this joint master program. Studying in different universities across Europe, allowed me to have great experiences both personal and professional. I am certain that I will start a PhD more prepared thanks to what Astromundus provided me with.
Agradecimentos

Eu gostaria de agradecer a minha família, em especial minha mãe pelo apoio incondicional dado a mim quando eu decidi ser um Astrofísico. Obrigado pela compreensão quando eu disse que estudaria em um programa de mestrado longe de casa, obrigado por tudo. A minha irmã que está sempre de bom humor, espalhando felicidade onde quer que ela passe. Muito obrigado por ter vindo e trazido mamãe e sobrinha para passarmos o Natal juntos. Foram os melhores dias, gostaria que vocês tivessem ficado mais tempo! Às crianças que eu amo muito: minha sobrinha Luísa Rodrigues e meus sobrinhos Théo Martins and Heitor Rodrigues. Suas energias, alegria e sorrisos são contagiantes, obrigado por preencher minha vida com felicidade e pela acolhedora recepção sempre que eu visito vocês.

Ao meu co-orientador Dr. Mathias Zechmeister por todas as vezes em que eu bati a sua porta e você pacientemente me explicava todos os aspectos desta pesquisa. Obrigado por isso e pelos incontáveis conselhos sobre programação (eu aprendi bastante com você!). Gostaria de agradecer meu orientador Prof. Ansgar Reiners quem me acolheu no grupo CARMENES; sem você esse trabalho não teria sido possível. Obrigado pelo direcionamento da pesquisa e pelas aulas de exoplanetas no meu 3º semestre, o curso foi crucial para o meu conhecimento sobre o assunto e, mais importante, foi bastante divertido (especialmente os quizzes).

Gostaria de agradecer aos meus colegas do astromundus, especialmente aos que passaram quase que os 2 anos de programa comigo. Suas amizades, as viagens juntos e as incontáveis reuniões de estudo para as provas amorteceram o peso desse programa de mestrado. Às meus amigos em geral, especialmente a que está num doutorado no Chile, a quem eu sou muito grato pelos bons conselhos com relação ao meio acadêmico, obrigado Ana.

Por último mas não menos importante, gostaria de agradecer ao programa ASTROMUNDUS, aos professores que foram responáveis por expandir os horizontes da minha compreensão da Astronomia e a secretária Andrea Rumpold que foi bastante solícita durante toda a nossa estadia na Europa. Foi uma grande honra participar deste programa de mestrado conjunto. Estudar em diferentes universidades na Europa, me possibilitou a ter grandes experiências tanto pessoais quanto profissionais. Eu estou certo de que iniciarei um doutorado mais preparado graças ao que o Astromundus me proporcionou.
## Contents

1 Introduction  
1.1 The Search for Planets Beyond the Solar System .......................... 1  
1.2 The Radial Velocity Technique ............................................ 3  
1.3 CARMENES SURVEY ...................................................... 7  

2 The Data Driven Method  
2.1 The wobble code .............................................................. 9  
2.2 Lp Regularization ............................................................ 11  

3 Methodology  
3.1 Target and Data Selection .................................................. 15  
3.2 Parameter Selection .......................................................... 15  

4 Results  
4.1 Barnard star ................................................................. 19  
4.2 J08413+594 .................................................................. 30  
4.3 Iot Cyg star ................................................................. 41  
4.4 Teegarden’s star ............................................................... 51  

5 Conclusion ................................................................. 57  

Bibliography ................................................................. 62  

A Appendix  
A.1 Barnard’s star ............................................................ 65  
A.2 J08413+594 ............................................................... 75  
A.3 Teegarden’s star ........................................................... 82  

xiii
Chapter 1

Introduction

1.1 The Search for Planets Beyond the Solar System

One of the most fundamental questions of science is whether complex life has also evolved in other planets. Therefore, an appropriate path to probe for such a mystery is through the search of habitable worlds, i.e., the ones that could harbor life as we know it. From just a few tenths of years ago, imaging-, transit-, and radial velocity amongst other techniques were introduced and improved over the years, making the detection of low-mass planets possible. The Radial Velocity (RV) method \cite{Struve1952} was the first of its kind to detect an exoplanet (51 Peg b) around a star like the Sun \cite{Mayor1995} in October 1995, and therein, as the number of catalogued planets increased more attention to the extra-solar planet field was granted. Thus, many astronomical surveys were put forward both from Ground (HARPS, ESPRESSO, CARMENES) and space (TESS, CoRoT, KEPLER), allowing remarkable discoveries regarding stellar system configuration (e.g., type-s and type-p planet configuration, very compact planetary systems, hot-jupiters, etc.) as well as unveiling the internal and atmospheric properties of planets to some extent. To date, about two decades after the first confirmed planet, the number of known planets is about four thousand\footnote{see https://exoplanetarchive.ipac.caltech.edu/} (see histogram in Figure 1.1). As the instruments yield more precise measurements, the more the low-mass planet regime is explored; consequently the number of Earth-mass planets in the habitable zone (HZ) are likely to increase in the following decades. Figure 1.2 displays a mass-period diagram where the number of planets detected by means of different techniques is represented by different colours. A simple scan at this diagram implies that more massive planets exist in nature in detriment of rocky ones. Nevertheless, the true population of low-mass planets is misrepresented due to detection bias from both the instrument and technique utilized. The two major techniques for planetary detection, transit and RV, favor the larger and more massive planets because they induce higher signal amplitudes which usually lie above the uncertainty limit, hence making their detection less complicated. Several enhancements are made to spectrographs
Chapter 1. Introduction

Figure 1.1: Accumulated number of planets detected throughout the years by multiple techniques represented by different colors.

such as temperature stabilization, iodine cell techniques, hollow cathode lamps, etc. in order to improve the planetary detectability, once that any source of error reduces the RV precision or even induce RV signals leading to false-positive discoveries. Therefore, the goal to detect planets similar to the Earth, and more importantly in the HZ of stars, is tightly woven to apparatus precision. Since late 1990’s, persistent efforts were carried out to detect planets orbiting around FGK-type stars once they are, to some extent, similar to the Sun. This reasoning has been changing though, Earth-like planets also exist around M-dwarfs, and they are relatively easier to detect by RV technique compared to the FGK stars. For example, the Earth induces a semi-amplitude RV signal of $\sim 10 \text{ cm s}^{-1}$ to the Sun at semi-major axis 1 AU, and approximately 3 times that value at 0.1 AU. A red-dwarf with $M_\star \sim 0.1 M_\odot$ hosting the same planet at its HZ (typically at $\sim 0.1$ AU) would have a semi-amplitude signal of $\sim 3 \times 30 \text{ cm s}^{-1}$ because Doppler shifts scales with $1/\sqrt{M_\star}$. The current instrumental precision revolves around the $1 \text{ m s}^{-1}$, thus it would be possible to detect a planet in the HZ of an M-dwarf whereas the detection of low-mass planets in the temperate zone of a G-type star is beyond the current limit. Based on that, astronomical surveys such as CARMENES run observation campaigns for the search of exoplanets around M-dwarfs. Therefore, as these surveys contribute to the extra-solar planetary field, we move towards answering relevant scientific and philosophical questions related to the possibility of complex life in other planets.
1.2. THE RADIAL VELOCITY TECHNIQUE

Figure 1.2: Mass-Period diagram displaying the total number of planets discovered by distinct techniques.

1.2 The Radial Velocity Technique

When a star is orbited by a planet, the barycenter of the system is displaced from the stellar geometrical center, making the star perform a small orbit around the former. This motion, as observed from the Earth, consists of two components: the transverse and the radial. The measurement of the latter, also called the line-of-sight stellar motion, is only possible due to the Doppler effect. Therefore, the RV technique consists of the estimation of how much the unobserved planet induces shifts in the radial component of the stellar light. Through this technique several planetary parameters could be calculated such as the planet period $P$, semi-major axis $a$, eccentricity $e$, and if the orbit inclination $i$ is known, its mass $M_p$ is estimated, otherwise only a minimum mass $M_p \sin i$ is measured. Figure 1.3 shows an example of a planet orbiting a star described uniquely by its six orbital parameters (OP) $a$, $e$, $i$, $\Omega$, $\varpi$ and $M$ where the 1st three were already defined, and the remaining parameters are respectively the longitude of ascending node, longitude of periapsis and the mean anomaly. The relationship between the plane of the orbit $(x,y,z)$ and the reference plane $(X,Y,Z)$ is given by the Eq. 1.1, 1.2 and 1.3

$$X = r \{ \cos \Omega \cos(\omega + f) - \sin \Omega \sin(\omega + f) \cos i \}$$  \hspace{1cm} (1.1)
Figure 1.3: Test particle orbiting about the center of mass. Coordinates $(x,y,z)$ represent the object position whereas $(X,Y,Z)$ the reference coordinates with $Z$ parallel to the line-of-sight. The angles $\omega$, $i$ and $\Omega$ represented in the image are some of the orbital elements describing the orbit. Credits: [Seager, 2010]
1.2. THE RADIAL VELOCITY TECHNIQUE

\[ Y = r \sin \Omega \cos (\omega + f) + \cos \Omega \sin (\omega + f) \cos \iota \]  
(1.2)

\[ Z = r \sin (\omega + f) \sin \iota \]  
(1.3)

with \( f = \theta - \varpi \) being the equation relating the true anomaly \( f \) to the true longitude \( \theta \) and \( \varpi = \Omega + \omega \), where \( \omega \) is the argument of perihelion. The radial velocity is simply its total velocity projected in the line-of-sight (Z-direction) described by Eq. 1.4,

\[ v_r = \frac{d}{dt} r \cdot Z = V_Z + \frac{m_p}{m_* + m_p} \frac{d}{dt} Z \]  
(1.4)

with \( Z \) being the barycentric velocity. In order to arrive at the final RV equation, we need to compute the term \( \frac{d}{dt} Z \). Taking the time derivative of equation 1.3 we find,

\[ \frac{d}{dt} Z = \frac{2 \pi a \sin \iota}{P \sqrt{1 - e^2}} \cos (\omega + f) + e \cos \omega \]  
(1.5)

and combining Eq. 1.4 and 1.5 we have the RV equation

\[ v_r = V_z + K (\cos (\omega + f) + e \cos \omega) \]  
(1.6)

where the 1st term on the right-side \( V_z \) is the barycentric proper motion and the semi-amplitude \( K \) could be written as \( K = \frac{M_p}{M_* + M_p} \frac{2 \pi a \sin \iota}{P \sqrt{1 - e^2}} \). Under the assumption that the planetary mass is negligible compared to the stellar, and by using the 3rd Kepler’s law to remove the dependence on semi-major axis \( a \), the semi-amplitude \( K \) could be rewritten as \( K = 0.6395 \text{ m s}^{-1} \frac{(\text{1 day})^\frac{\frac{1}{2}}{P} \frac{M_p \sin \iota}{M_*}}{\sqrt{1 - e^2}} \).

The radial velocity planetary detection consists on measuring \( v_r \) variation in time caused by the Doppler shifts of spectral lines (Eq. 1.7) due to the presence of a planet (e.g., Fig. 1.4), the semi-amplitude \( K \) contains the minimum mass, \( M_p \sin \iota \), \( P \) and \( e \), hence the best Keplerian curve that fits the data yields the parameters describing the planetary orbit. It is important to point out that the best fit to the data is not enough to confirm a planet, there are many instrumental and stellar effects that need to be ruled out in order to claim that a signal has planetary origin. For instance, effects such as stellar activity (e.g., spots, flares, plages, so forth) or spurious instrumental signals could mimic planet signals if not carefully subtracted from the data\(^2\).

\[ \lambda = \lambda_0 - \frac{\frac{1}{2} k \cdot v}{\sqrt{1 - \frac{v^2}{c^2}}} \]  
(1.7)

Figure 1.4: Radial velocity phase chart displaying the presence of a planet around the star 51 Peg – the first exoplanet discovered to orbit a G-type star. Data (black closed-circles) are well fitted by the curve. Credits: [Mayor and Queloz, 1995]
1.3 CARMENES SURVEY

The CARMENES (Calar Alto high-Resolution search for M dwarfs with Exo-earths with Near-infrared and optical Échelle Spectrographs) consortium is part of a joint program by German and Spanish institutions situated at the Calar Alto Astronomy Observatory in the province of Andalucia south of Spain. The apparatus comprises a 3.5 m telescope, Échelle spectrographs, a Fabry-Pérot interferometer for nightly tracking, fibers, lamps for calibration, red-ward sensitive detectors, among other instruments (refer to Table 1.1 for more details). For data reduction, CARMENES utilizes the SpEctrum Radial Velocity AnaLyser (serval) \cite{zechmeister2018} in order to extract precise RVs with accuracy to the 1 m s\(^{-1}\) level with long term stability. The survey operates in both visual (VIS) and near-infrared channel (nIR) covering the wavelength range: 520–960 nm in VIS mode with spectral resolution \(R \sim 94\,600\), and 960-1710nm in nIR with \(R \sim 80\,400\). It started its observation campaign in January 2016 aiming to track about 300 M0-M9 type stars covering the mass range \(M_\star < 0.60\,M_\odot\) (Fig. 1.5). Therefore, a broad range of stellar spectral types from M4V and later to M0V are under daily observation. Therefore, CARMENES main goal is to detect low-mass planets orbiting M-dwarf stars with the RV technique. Besides its own planetary search campaign, CARMENES may also assists other missions such as TESS and the upcoming PLATO for follow-ups, hence either confirming planets candidates or disregarding them as false positives. Therefore, missions such as CARMENES aid the global efforts toward the detection of low-mass planets, allowing the number of Earth planets per stars \(\eta_\oplus\) to be better constrained. For a detailed description of the project refer to the official webpage\footnote{https://carmenes.caha.es/index.html}.

<table>
<thead>
<tr>
<th>Table 1.1: CARMENES specification adapted from the official webpage (footnote)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VIS channel</td>
</tr>
<tr>
<td>Wavelength coverage, (\Delta \lambda^*)</td>
</tr>
<tr>
<td>Detector</td>
</tr>
<tr>
<td>Wavelength calibration</td>
</tr>
<tr>
<td>Working temperature, (T)(_{\text{work}})</td>
</tr>
<tr>
<td>Spectral resolution, (R)</td>
</tr>
<tr>
<td>Mean sampling</td>
</tr>
<tr>
<td>Mean inter-fibre spacing</td>
</tr>
<tr>
<td>Cross disperser</td>
</tr>
<tr>
<td>Reflective optics coating</td>
</tr>
<tr>
<td>No. of orders</td>
</tr>
<tr>
<td>Échelle grating</td>
</tr>
<tr>
<td>Target fibre field of view</td>
</tr>
<tr>
<td>A&amp;G system field of view</td>
</tr>
<tr>
<td>A&amp;G system band</td>
</tr>
</tbody>
</table>
Figure 1.5: Histogram displaying the CARMENES SURVEY sample (in white) and its sub-sample containing stars with Hα in emission (in red). The histograms are by spectral type (top-left), stellar mass (top-right) and J magnitude (bottom-center). Adapted from [Reiners et al., 2018]
Chapter 2

The Data Driven Method

2.1 The wobble code

wobble\cite{Bedell2019} is a python package designed to compute precise differential RVs as it simultaneously infers from the data what component belongs to the star and the Earth’s atmosphere. This distinction is only possible because the Earth orbital motion varies over the year, causing a relative motion of the stellar spectrum with respect to the telluric, which is fixed in the observatory reference frame. wobble uses a data driven approach, meaning that external information (e.g., molecules column density, layer temperature and pressure, amongst other parameters) is not necessary for the process of components modeling. Since wobble well-functioning is dependent on large data sets, it utilizes a library called tensorflow\cite{Abadi2016} which has a large number of highly optimized routines to perform linear algebra (LA) calculations, thus making the wobble more efficient.

The searches for low-mass exoplanets demand extremely-high precision RVs, hence, apart from the fine stability of all astronomical equipment components, RV extraction must be performed in the best manner so that no loss of precision occurs. One of the major issues regarding precision degradation is the influence of tellurics and micro-tellurics presented in the data, where its impact revolves around the $1 \text{ m s}^{-1}$ level\cite{Cunha2014}, thus affecting the planetary detection in the Earth-mass regime. There are several means to compute RVs, and for each of them, loss of precision is almost inevitable. For instance, calculation of RVs by prior masking of telluric regions lead to loss of stellar features in the covered area, thus the RV precision is compromised. In addition, micro-tellurics are hardly ever well masked due to their unpredictability, this adds to the uncertainties in the RV data. Another technique utilizes fast rotating standard telluric stars which could, in principle, reproduce the atmospheric transmission spectrum if observed near in time and position to the main target. Nevertheless, the former contains stellar features that could

\footnote{The description of this piece of software is consistent with its first version. Please refer to \cite{Bedell2019} and https://github.com/megbedell/wobble for a thorough discussion of the code.}
be introduced to the latter when the removal of atmospheric features is performed, thus degrading the RV precision; besides the amount of extra time in observing standard telluric stars. Therefore, a reasonable method to mitigate the flaws of poor removal of atmospheric contamination is through the modeling of both components simultaneously, therefore we utilize \textit{wobble} for this task. Assuming that the data were perfectly reduced, \textit{wobble} makes a quality evaluation prior to running. First, it masks the data points that falls below a certain user-defined signal-to-noise (SN) and minimum flux by assessing each epoch by order. All these parameters may be changed by users depending on their data to be utilized, e.g., for CARMENES we set the minimum SN and minimum flux to be masked approximately zero otherwise in default mode, most of the data at this quality evaluation step is rejected. Subsequently, data is normalized by a polynomial of order \( n \) chosen by user, where all the orders will be normalized by the same polynomial. The following step is the modeling procedure which occurs in natural logarithm basis since stellar and telluric components are multiplied and in log-scale the multiplication operation becomes a summation which is a simpler mathematical operation to perform. The stellar- (Eq. 2.3) and telluric templates (Eq. 2.4) consist of data interpolation through a set of wavelength grid of control points arbitrarily chosen after a Doppler shifts is applied so that the template is in the observatory reference frame. First, \textit{wobble} initializes the averaged stellar template in the manner just described, from the residuals the averaged telluric template is defined. Since each epoch will present trends in the residuals due to the atmospheric lines variability, with the averaged templates fixed, the basis vectors modeled by the term \( W^t z_n \) in eq. 2.4 is initialized and adjusted so that the minimum possible residuals are achieved. Therefore, the basis vectors could be thought of the number of components added to the averaged telluric template so that fine tuning of atmospheric variability in each epoch is reached. The user has control of the exact amount of basis vectors by tweaking the parameters \( K_t, K^* \). Moreover, this procedure to encounter the best templates occur by multiple evaluation of a log-likelihood cost function (eq. 2.1) defined by the \( N_{\text{iter}} \) parameter. A reasonable value for this parameter is set so that the likelihood converges.

\[
\ln L = -\frac{1}{2} \sum_n (y_n - y'_n)^T C^{-1}_n (y_n - y'_n) + \lambda_1 \| \mu^* \|_1^1 + \lambda_2 \| \mu^* \|_2^2 + \lambda_3 \| \mu^t \|_1^1 + \lambda_4 \| \mu^t \|_2^2 + \lambda_5 \| W^t \|_2^2 + \lambda_6 \| W^t \|_2^2 + 1.0 \| z^t \|_1^1
\]  

(2.1)

The variables \( y_n, y'_n, C^{-1}_n, \mu^*, \mu^t, z^t, W^t, \lambda_1, \lambda_2, \ldots, \lambda_6 \) and \( \lambda_6 \) are respectively the data, model, inverse of covariance matrix, average stellar template, average telluric template, basis weights, telluric variability template and the regularization amplitudes associated to the respective parameter\(^2\) (see Eq. 2.1). The model \( y'_n \) in Eq. 2.2

\[
y'_n = y^*_n + y^t_n + \text{noise}
\]  

(2.2)

\(^2\)These \( \lambda \)'s are determined internally from the regularization scripts that must be run prior to \textit{wobble}; however, this is not enforced, i.e., the code is able to run without regularization where all \( \lambda \)s are set to zero. Nonetheless, regularization is an important \textit{wobble} feature that we explore throughout this work.
2.2. LP REGULARIZATION

is comprised of the stellar component (Eq. 2.3),

\[ y^*_n = P(v^{\text{obs}}_n)\mu^*, \]  

(2.3)

where \( P(v^{\text{obs}}_n) \) is a Doppler shift operator and \( v^{\text{obs}}_n \) the observed velocity, and the telluric component (Eq. 2.4),

\[ y^{\text{tel}}_n = A_n (\mu^{\text{tel}} + W^t z_n) \]  

(2.4)

which is scaled by the airmass \( A_n \). Notice that the user has some control of the number of basis vectors that are needed to better capture the telluric variability, e.g., if the parameters \( K_\star \) and \( K_t \) are set to zero, \texttt{wobble} assumes that no variability occurs, hence the model contains only averaged components (see chapter Results Fig. 4.17 for effects in the modeling when these parameters are varied for Iot Cyg and the impact in the RV time-series in Fig. 4.19). Upon conclusion of optimization, the following step consists of combining all the orders into a RV time-series which is computed with equation 2.5

\[ v_{n,r} = v_n + v_r + G(0, \sigma^2_{n,r} + \delta^2) \]  

(2.5)

where \( v_{n,r} , v_n , v_r , G(0, \sigma^2_{n,r} + \delta^2) \) are the RV to be modeled by order and epoch, time-dependent stellar RV, order-dependent RV offset and a assumed Gaussian noise term containing an estimated uncertainty \( \sigma \) added quadratically to a jitter term \( \delta \). Lastly, a file is compiled with all the results and the run history.

2.2 Lp Regularization

Figure 2.1 (left side) illustrates an example the impact regularization causes to a simple fitting exercise. The yellow curve uses a \( \chi^2 \) to minimize the sum of the squared errors whereas the red curve utilizes the same cost function plus an additional term associated to the L2 regularization. Notice that without regularization (yellow curve) the mean square error has a similar magnitude compared to when the regularization (red curve) is used (Fig. 2.1 right side), however, the latter is able to retrieve a smoother template; therefore, regularization could be thought of a technique that desensitizes the fitting curve. Template smoothing is important in science because data contain a significant amount of noise (or other unwanted component); hence a smoother curve prevents the model to fit noise. By tweaking the regularization amplitudes \( \lambda_\text{s} \), one is able to adjust how smooth, or, how much penalty should be associated to a particular parameter. There are several techniques used in optimization problems whose goal is to enhance the final models, here we focus on the \( L_p \) regularization because it is implemented in \texttt{wobble}.

The ‘p’ subscript in \( L_p \) defines the type of regularization to be utilized. If \( p=1 \) we have a lasso- or L1 regression and if \( p=2 \) ridge- or L2 regression. Equation 2.6 represents standard form of a
minimization problem with regularization,

\[
\min \sum_{i=1}^{n} V(f(w_i), y_i) + \lambda R(f)
\]  

(2.6)

where the first term could be thought of a \(\chi^2\) equivalent to the summation term in eq. 2.1, and the second one the regularization term assuming the form \(\lambda \sum \|w\|_1\) for L1 and \(\lambda \sum \|w\|_2\) for L2 with ‘par’ meaning parameter and \(w_i\) a variable (e.g., \(\mu^*, \mu^{tel}, W^{tel}\) in eq. 2.1). The \(\lambda\)s are the amplitudes or weights for a given vector \(w\), hence adjustments of \(\lambda\)s control the importance given to a coefficient in the fitting of data. Equation 2.7 represents the analytical solution for the \(L_2\) regularization problem which illustrates well the significance of the amplitudes.

\[
w_i = (\hat{\mathbf{W}}^T \hat{\mathbf{W}} + \lambda n I)^{-1} (\hat{\mathbf{W}}^T \mathbf{Y})
\]  

(2.7)

If a large value for \(\lambda\) is set, the template’s dependence on \(w_i\) is lost, whereas if the amplitude is zero, Eq. 2.7 represents a standard \(\chi^2\) minimization problem without regularization. Therefore, a reasonable value for the amplitude must be encountered in order to achieve a smooth template without too much penalizing the terms so that relevant features are neglected. Assuming we have data that rely on only two parameters, e.g., \(w_1\) and \(w_2\), Fig. 2.2 illustrates function contours of a generic optimization problem where on the left we have the lasso represented as a 45° flipped square, and on the right the ridge regularization illustrated by the circle with ellipses from the \(\chi^2\)
2.2. LP REGULARIZATION

These shapes are the consequence of the numerical value associated to \( p \). For \( p=1 \), it yields the contours. For \( p=2 \), the L2-norm, \( \| w \|_2 = (\| w_1 \|^2 + \| w_2 \|^2)^{1/2} \). A standard \( \chi^2 \) minimization problem is represented only by the ellipsoidal contours where the optimal parameters lie in the central point, i.e., the minimum of the \( \chi^2 \). However, the additional \( L_p \) terms or the combination of them, will shift the minimum of the new cost function so that it lies in the interception between the ellipse and the contours from L1 and (or) L2. The equivalent of looking for the interception of this new cost function is that, as we increase the sum of the squared error by moving away from the \( \chi^2 \) minimum, we miss the optimal fit to the data points. The level of underfitting cannot be too low otherwise all features are lost, therefore the first intercepting contour level from the \( L_p \) term yields the optimal values for the parameters seen in Fig. 2.1. In other words, moving away from the ellipses central point yields a quadratic loss in the fitting, however, we are moving towards the contours of \( L_2 \) (right panel or \( L_1 \) left panel), hence a quadratic gain for \( L_2 \) is achieved. The optimal solution from the regularization point of view is reached when the interception occurs, i.e., a balance is reached between the \( \chi^2 \) and the \( L_2 \) terms. Each regularization has its own properties, for instance L1 induces sparse solutions making the model convex and less complex whereas L2 is differentiable hence could be solved by simpler algorithms (e.g. gradient descent).

wobble aims to achieve the best possible templates from the interpolation of data, therefore, it is important that the code learns which feature to avoid as it computes the templates. For instance, either noise or the telluric component, if present in the stellar template, automatically degrades the

---

3For a complete description of \( L_p \) regularization and its properties refer to the literature (e.g., [Tibshirani, 1996] and [Hoerl and Kennard, 1970]).
RV precision. Therefore, wobble regularization routine computes the $\lambda$s amplitudes through a cross-validation scheme where 85% of data is utilized as training set where only the regularization amplitudes associated to the averaged templates $\mu^*$ and $\mu^{tel}$ are probed from a default grid. The 15% of data remaining is used as a validation set so that the $\lambda$s associated to the variability parameter $W^{tel}$ is adjusted.
Chapter 3

Methodology

3.1 Target and Data Selection

Target selection was based on stellar properties as well as amount of available data since the code performance relies on big data sets (>100 epochs). We select Barnard’s star because it is relatively quiet, thus less spectral variability occurs. This property may favor wobble well-functioning. Moreover, the developers of wobble tested the code on Barnard’s star with HARPS data, hence we aim to compare the code performance between HARPS and CARMENES where applicable. Teegarten’s star was selected because it presents evident Hα activity, therefore we may analyze wobble behavior in this regime. We have also opted for a bright standard telluric star named Iot Cyg. It has fewer stellar features compared to the other targets, therefore, although it may represent a challenge to the code, we could test its limitations. We have also included the star J08413+594 which is known to host two exoplanets where the RVs semi-amplitude is quite evident ($K \sim 47 \text{ m s}^{-1}$). The idea is to retrieve the orbital parameters and compare it with the literature. The data selection consists of reducing the number of epochs if a particular target has more than 300 spectra. This was necessary once wobble uses up several Gigabytes of RAM and swap memory. We noticed that a run with 30 orders could only support about 300 epochs, otherwise the lack of memory would make the code stop. Therefore, when necessary, I remove the epochs containing the lowest SNR.

3.2 Parameter Selection

A major step prior to executing the code is the selection of suitable parameters so that the best possible modeling occurs; for instance the continuum normalization polynomial degree, number of basis vectors $K$, and $K_t$ and number of iteration $N_{\text{iter}}$. wobble utilizes $\sigma$-clipping in order to localize the continuum level and then it fits a polynomial function to it. Both the $\sigma$ range and the function degree must be set by the user. Data normalization
Figure 3.1: Continuum normalization illustration of CARMENES Barnard’s star sample. A polynomial fit is performed (blue line) to encounter the continuum. The same epoch at orders 10 (left) and 25 (right) is displayed.

is enforced in wobble, hence in the case that orders display nearly-normalized condition or needed only an offset adjustment, a polynomial of 1st degree is preferred. The way continuum normalization was built to the code does not allow the user to choose different polynomial functions to distinct orders, therefore I had to carefully select the orders in which data shape were somewhat similar. Due to the amount of time wobble takes for normalization, a few orders with high- and low SNR were selected (e.g., order 10, 11, 25, 28, 38, 40) and visually checked whether any epoch cannot be properly normalized or if the polynomial degree utilized performed well. In case the selected polynomial degree performs poorly, a different one is chosen. If most of the epochs in multiples orders are well fit, we delete the few poorly normalized epochs and use the picked σ clipping range and polynomial for the complete run. Figure 3.1 illustrates the continuum normalization procedure of wobble on Barnard’s star. Notice that poor and good normalization arise from the fact that orders are different in shape, specially the ones containing large portion of atmospheric features. The usage of the same polynomial function (in this case order 4) degrades the continuum normalization (in Results we discuss continuum normalization issues for each target when necessary). Fortunately, most of the targets had orders which were uniform in shape hence only an offset adjustment was necessary.

The number of parameter contained in wobble is quite large, hence the parameter selection related to the modeling was based on the few ones that seemed more relevant to the performance of wobble, that is, the number of iteration $N_{\text{iter}}$, the basis vector $K_*$ and $K_t$ and the regularization amplitudes $\lambda$. From these, only the latter should be performed by wobble itself due to the complexity and vast parameter space searches as described in section 2.1. The $N_{\text{iter}}$ parameter is selected by analyzing the convergence of the likelihood (e.g., see Fig. 3.3), however, if no convergence is reached we utilize higher magnitudes such as $N_{\text{iter}} = 400$ or 500 depending on the amount of data since memory shortage is possible to occur. For the parameters $K_*$ and $K_t$, we

\footnote{If the overall shape of epochs in each order are significantly different, a high order polynomial must be chosen, thus it is likely that the polynomial fits stellar or telluric features.}
3.2. PARAMETER SELECTION

Figure 3.2: Continuation of Fig. 3.1 for orders 38 (top-left), 40 (top-right) and 52 (bottom-left). A poor continuum normalization at order 25 for a different epoch is illustrated on the bottom-right.

search the optimal number of basis vectors for a grid where \( K_\star \) is either 0 or 1, not higher once we assume small variability to the stellar spectrum. For \( K_i \) we take values from 0 to 3 since the telluric transmission spectrum indeed varies depending on multiple factors such as airmass, atmospheric composition, so forth. The selected grids are not strict, hence based on each modeling results we could experiment other basis vectors values.
Figure 3.3: Evaluation of the likelihood (y-axis) for each iteration (x-axis) from wobble runs on Iot Cyg star with parameters $K_*=0$, $K_t=3$ and $N_{\text{iter}}=200$. Notice that for a 200 loops order 25 shows convergence, however order 38 needs a higher $N_{\text{iter}}$ to stabilize.
Chapter 4

Results

4.1 Barnard star

Barnard star is a quiet M-dwarf with $V_{\text{mag}} = 9.5$ and $M = 0.14 \, M_\odot$. It is located about 6 ly away and has the highest proper motion to date. CARMENES data set is comprised of 347 spectra, hence a few epochs were removed due to memory constrains (see methodology). The continuum normalization step showed that a polynomial of 4\textsuperscript{th} degree was suitable for most of the orders where only 20 epochs were poorly normalized due to either lower than average flux (similarly to Fig. 3.1 bottom-right) or function misfit.

Prior to running \texttt{wobble} for the entire data set, we test it in smaller portions of data so to probe the parameters: $N_{\text{iter}}$ for likelihood convergence and a reasonable basis vector grid in which to run the complete sample. We concluded that the former should have $N_{\text{iter}} = 200$ whereas the later $K_\star = 0$ and $K_t = 0, 1, 3$. We included $K_t = 3$ in order to probe the code behavior for when high number of basis vectors are used despite the risk of overfitting as explained in the developers paper \cite{Bedell et al., 2019}. Subsequently, we utilize the code to calculate the optimal regularization amplitudes and run \texttt{wobble} multiple times with different number of basis vectors with and without\footnote{No regularization means that the $\lambda$’s in Eq. 2.1 are set to zero, including the $\lambda = 1.0$ multiplied to the $z_n$ term; thus a $\chi^2$ is the actual function in the minimization problem.} the regularization. Figure 4.1 illustrates the RV time-series calculated from \texttt{wobble} (in black) in comparison with \texttt{serval} (in red) for distinct parameters. It shows that the lowest RV rms takes place for the run with regularization for the parameters $K_t = K_\star = 0$ (upper right panel). This is in agreement with the developers study on Barnards where HARPS data \cite{Bedell et al., 2019} was utilized; i.e, they have concluded that these basis vectors retrieves better modeling. Despite that, \texttt{serval} RV time-series ($rms_{\text{RV}} = 2.69 \, \text{ms}^{-1}$) seems better than \texttt{wobble}'s ($rms_{\text{RV}} = 2.89 \, \text{ms}^{-1}$) for the best run in Fig. 4.1 panel (b). In order to improve \texttt{wobble} RV calculation we should mitigate any issue in the particular orders that had poor modeling, therefore, we investigate all the modeled orders for the run that reached the best RV rms and compare to the same run without
regularization. This is necessary so that we could confirm that the regularization amplitudes were indeed optimal. Figures 4.2 (order 32) and 4.3 (order 38) are examples of good and poor modeled orders. In fact other orders presented poor RVs similarly to order 38 where from the total (30 orders) 10 were poorly modeled (e.g., orders 29, 30, 37, 43, etc.). Therefore, from the analysis of all orders, it seems that there is a correlation between `wobble's` performance and the amount of telluric contamination.

\[\text{see appendix for more orders}\]
4.1. BARNARD STAR

Figure 4.1: Barnards RV time-series extracted with wobble for the same $N_{\text{iter}} = 200$, $K_s = 0$ and different $K_t$, on the left and right panels the RVs were calculated without and with regularization of templates. For all plots we compare serval (red) and wobble data (black).
CHAPTER 4. RESULTS

Figure 4.2: Wobble run without and with regularization. On top (a) and (b), Barnard spectra at order 32 where data is shown in black, stellar model in red and telluric template in blue with $N_{\text{iter}} = 200$, $K_0 = 0$ and $K_T = 0$, at the bottom (c) and (d), the RV at the same order. On the left and right panels the RVs were calculated without and with regularization of templates. For all the plots we compare the rms of SERVAL (red) and wobble data (black).
4.1. BARNARD STAR

Figure 4.3: Same as Fig. 4.2 for order 38.
CHAPTER 4. RESULTS

From these two orders we notice clearly that regularization prevented most of the tellurics to be modeled in order 32, however, since order 32 is not much contaminated, there is little difference between the runs with/without regularization in terms of in the rms (Fig. 4.2 and 4.3 panels (c)). For order 38 though, we see clearly that the modeling fails in both cases. The regularization amplitudes applied to this order seemed to have made some corrections to the overall shape of the telluric component, however, from the residuals the run without regularization had a better outcome (panel (a)) which is confirmed by the $\text{rms}_\text{RV}$ (bottom panel). In general wobble performs in such a way that the poorly modeled orders are assigned a lower weight, hence the final RV time-series in Fig. 4.1 panel (b) were calculated based more on the good orders. Therefore, regularization is preferred despite the sub-optimal behavior for 10 orders compared to no regularization whatsoever.

As mentioned throughout this work, the code seems to face issues with continuum normalization as well as calculation of optimal regularization amplitudes. Since both the Barnard’s star HARPS data and the script utilized by the developers is available on their github page\footnote{https://github.com/megbedell/wobble}, I ran wobble with the same settings as in their work so that we could compare if similar issues appear in data from distinct surveys. Figure 4.4 display a few orders from wobble runs on HARPS data. Notice that I intentionally selected good and poor epochs in order to show that the same issues we faced with CARMENES data were present concerning modeling (Fig. 4.4 panel (a)) and continuum regularization (Fig. 4.4 panel (b)). Besides that, the telluric component presents a flat shape for most of the orders, this characteristic is induced by the heavy penalties applied to by the regularization amplitudes. These penalties also force the template to miss atmospheric features that are not evident, i.e., the ones which are usually less than 10–20% deep. We acknowledge that the results from our runs on HARPS data retrieve better modeling despite the mentioned similar issues. However, it is likely that wobble performed better with HARPS data because its wavelength range is bluewards compared to CARMENES, thus, in the lack of significant atmospheric contamination, less poor modeling takes place. Figures 4.6 and 4.7 illustrate CARMENES orders 21 and 25 which overlap with DRS-HARPS orders 63 and 67. Both runs had the same basis vectors and number of iteration, but different $\lambda$’s for the regularization amplitudes. Notice that the epoch from HARPS captures less atmospheric features compared to CARMENES without regularization (Fig. 4.6 panel (a) and (b)), and if we compare DRS-HARPS with the CARMENES run with regularization (Fig. 4.6 panel (a) and (c)), both cases capture almost no atmospheric shallow features and the stellar component misses most of the data on the left edge as seen in panel (c). These poor modeling effects are, as mentioned above, due to heavy regularization penalties, thus it is possible that this routine is calculating non-optimal $\lambda$’s for orders containing more telluric lines independently of surveys. Notice that the residuals in Fig. 4.5 panel (a) and 4.6 panel (a) imply that the wrong regularization amplitude may have been the reason for poor modeling of these orders, specially if we compare it to the run with not regularization (Fig. 4.6 panel (b)) where less atmospheric features are present in the residuals. Therefore, from running wobble with HARPS data it was possible to realize that the issues we faced with the code is probably not exclusively due to CARMENES.
Figure 4.4: Results from *wobble* run using developers script, regularization files and HARPS data.
CHAPTER 4. RESULTS

(a) DRS-HARPS Order 35

(b) DRS-HARPS Order 40

(c) DRS-HARPS Order 42

Figure 4.5: Continuation of Fig. 4.4
4.1. **BARNARD STAR**

data, but something else related to data or the manner **wobble** calculates these regularization amplitudes.

As an attempt to improve our results, I have carried out a few tests based on the following hypothesis:

1. Independently of regularization, perhaps the poor modeling may be influenced by the lower SNR edges of Échelle orders. Therefore, I trimmed 10% from both edges for the entire sample. This procedure does not lead to information loss because these corners are contained in the nearby orders with a higher SNR. The runs with and without regularization showed no improvement in neither RV time-series rms nor the models.

2. Deletion of poor orders may allow **wobble** to compute better final RV time-series. I have run the code selecting only the orders where RV rms $\leq 10$ ms$^{-1}$. I assumed that perhaps the orders with high RV spread is influencing significantly the final RV time-series. However, the outcome from the RV time-series had a higher rms compared to when no orders were deleted. This shows that indeed the orders with larger RV spread contribute much less to the final RV time-series, hence decreasing the number of orders would in principle make the final RVs rms larger. Moreover, a few orders deleted had 2 or 3 times the established RV rms ($RV_{rms} = 10$ ms$^{-1}$) for cut selection, and it is not desirable to exclude them since we wish to evaluate the performance of **wobble** exactly in orders that present large atmospheric features.

We may conclude that, **wobble** calculates RV time-series that are in accordance with **serval**, however, it still produces poor modeling with or without regularization, which for Barnard’s star, the former yielded the lowest RV rms but yet its performance was sub-optimal.
Figure 4.6: Comparison of overlapping orders between HARPS and CARMENES surveys. Same parameters were used in both runs, except for the \( \lambda \)s since each data set will probe the optimal amplitude accordingly.
4.1. BARNARD STAR

Figure 4.7: Continuation of Fig. 4.6
4.2 J08413+594

J08413+594 (or GJ 3512) is an M-dwarf of brightness $J \approx 9$ located at $d \approx 10$ pc that hosts two planets\[4\]. Following the same methodology, we use orders in the range 20–50, with a total sample of 159 spectra. The continuum normalization process was carried out with a polynomial of 1st degree where 36 epochs were deleted. The grid utilized for probing the basis vectors has $K_\star = 0$ for stellar component and $K_t = 0$ or 1 for telluric template. This range of values were chosen based on prior tests where higher magnitudes yielded poor modeling, specially for the stellar component. However, if any evidence suggests that other basis vectors must be tried, we proceed in this direction. For the number of iteration $N_{\text{iter}}$, we notice that a few orders do not show convergence of the likelihood function. This has been observed mostly for orders that contain several telluric bands, yet other similar orders may or may not converge; the exact reason for this behavior is still unclear to us. Therefore, $N_{\text{iter}} = 400$ is preferred since most of the orders already converged at this point. Having these initial parameters selected, we analyze the results, and for the data that retrieves the final RV time-series with lowest rms, I estimate a few planetary parameters such as mass, eccentricity and orbital period by fitting the best Keplerian curve\[5\].

We start by running wobble multiple times with and without regularization so to probe which basis vectors yield good modeling of data and consequently low RV spread. Figure 4.8 illustrates the RV time-series calculated with wobble (in black) compared to serval (in red) for distinct basis vectors. The lowest RV rms achieved had basis vectors $K_\star = 0$ and $K_t = 1$ without the usage of regularization (Fig. 4.8 panel (c)). Notice in panel (d) that the $\lambda$ amplitudes calculated were probably sub-optimal. Therefore we investigate the modeling of orders for the runs displayed in Fig. 4.8 panel (c) and (d) in order to comprehend why the regularization may have failed. Figure 4.9 show order 25 where the top and middle panels represent the modeling without and with regularization whereas at the bottom their respective RVs are compared to serval. J08413+594 shows a visible H$\alpha$ variability for several epochs, and because the stellar component is only composed of the average template ($K_\star = 0$), it fails to properly model the H$\alpha$ line. Concerning the telluric component, it fits no real atmospheric lines and, in panel (a) due to the big hump from the stellar template, it assumes an absorption-like shape so that the sum of both components\[6\] yields the lowest possible residuals. In an attempt to better model the H$\alpha$ feature, I ran the code with $K_\star = 1$ and $K_t = 1$, yet the stellar template still poorly modeled the H$\alpha$ feature inducing a $rms_{\text{RV}} = 52.66$ ms$^{-1}$ at this order whereas with only the average template ($K_\star = 0$) we have $rms_{\text{RV}} = 51.03$ ms$^{-1}$ (Fig. 4.9 panel (c)). The run with regularization where the RV spread is even larger (Fig. 4.9 panel (d)) could be explained by the high penalties imposed specially to the stellar template by the $\lambda$ amplitudes, hence making it miss most of the data. The same behavior is observed in

\[4\] CARMENES paper in preparation
\[5\] The python script utilized for this exercise can be found at https://adamdempsey90.github.io/python/radial_velocities/radial_velocities.html
\[6\] wobble works in log-space
4.2. J08413+594

Figure 4.8: wobble RV time-series (black circles) for J08413+594 calculated with $N_{\text{iter}} = 400$, $K_i = 0$ and distinct $K_t$ basis vectors. RVs had barycentric motion, instrumental drifts and nightly zero point corrections. serval is also shown (in red) for comparison.
Figure 4.9: Comparison between runs with and without regularization and their by-order RV time-series. serval results are shown in red for comparison. The parameters utilized are: $K_s = 0$, $K_t=1$ and $N_{\text{iter}}=400$. 

(a) No regularization

(b) Regularization

(c) RV no regularization

(d) RV with regularization
many orders\footnote{see more orders in appendix} therefore stellar activity and the poor selection of regularization amplitudes might be connected, at least in orders that spectra variability is evident. As it became evident that the regularization routine behaviour for both Barnard’s star and J08413+594 is sub-optimal, we attempt to search for reasonable $\lambda$’s manually. A few hypotheses were tested: 

**Hypothesis 1** – I search the amplitudes that yield good modeling based on a few orders selected from the range 20–50. For instance, I picked orders 22, 26, 32, 38, 42 and 46, these orders will represent the nearby ones, e.g., the $\lambda$’s picked based on orders 22 and 26 will be assigned to the range 20–29, the $\lambda$’s based on orders 32 and 38 will be assigned to the range 30–40, and the $\lambda$’s from orders 42 and 46 will be assigned to the range 41–50. From these 6 orders mentioned above, I tweaked only the averaged templates amplitudes $\lambda_n$, with $n=1,2,3,4$ (see Eq. 2.1). I change their magnitudes by trial and error for each order (see Figures 4.10–4.13). I concluded that, based on orders 22 and 26, the amplitudes should be assigned $\lambda_n=10$ for all $n$. From orders 32 and 38, $\lambda_n=10$ for $n \in 1,2$ and $\lambda_n=10^4$ for $n \in 3,4$ and from orders 42 and 46, $\lambda_n=10^2$ for $n \in 1,2$ and $\lambda_n=10^4$ for $n \in 3,4$. The selected $\lambda$’s were based on visual goodness of fit. Notice that the higher the magnitude of the amplitude, the more penalty is applied, thus pushing the template to the continuum level. From this test we were able to select $\lambda$’s that indeed enhanced the modeling process compared to when wobble runs with regularization routine. This confirms our assumption that the regularization routine calculated non-optimal amplitudes. Yet wobble data modeling with no regularization still yields the best outcomes whatsoever.

**Hypothesis 2** – While running the code with regularization, some orders are indeed well modeled while others are not. Therefore modifying the amplitudes only for the non-optimal orders may yield better modeling. Upon tweaking only these particular amplitudes from poor orders, the results were comparable to the ones from Hypothesis 1 i.e., we get better modeling and final RV time-series compared to the runs with regularization, however, no regularization still performs better. I must point out that both the assumptions above are far from ideal since the optimal amplitudes should not be selected manually but through wobble regularization routine which performs a machine learning process so to calculate each $\lambda$ for a given order. Moreover, I tested only the amplitudes related to the averaged components, which are four variables out of the countless parameters taken into account by wobble in the regularization step.
Figure 4.10: Different values for regularization amplitudes showing the effect it causes in the modeling of data for order 26
Figure 4.11: Continuation of Figure 4.10
Figure 4.12: Continuation of Figure 4.10
Figure 4.13: Different values for regularization amplitudes showing the effect it causes in the modeling of data for order 38
Figure 4.14: Continuation of Figure 4.13
Figure 4.15: Continuation of Figure 4.13
CHAPTER 4. RESULTS

Figure 4.16: At the top, the best Keplerian fit to the RV time-series data. The RVs came from the wobble run with parameters $N_{\text{iter}} = 400$, $K_{\star} = 0$ and $K_t = 1$ without regularization. On the bottom, Lomb-Scargle Periodogram showing a strong peak at $P = 204.0\,\text{d}$.

The orbital parameters were retrieved by a Keplerian curve fit for one planet where a Lomb-Scargle periodogram revealed a strong peak at $P \sim 201.51\,\text{d}$ (Fig. 4.16). No stellar jitter was assumed once that wobble takes it into account in the RV time-series calculation. Notice that the residuals shows some possible evidence of a 2nd planet. However, performing another Keplerian fit on the residuals fails to retrieve realistic orbital parameters (e.g., $e = 1.0$). Therefore, by this simple procedure, I was able to retrieve one planet with $M_p \sin i = 0.67 \pm 0.44\,M_J$ and eccentricity $e=0.457$. 
4.3 Iot Cyg star

According to the astronomical database SIMBAD, Iot Cyg (HIP 95853) is a fast rotating star of spectral type A5Vn with brightness $V = 3.8$ mag located at $\sim 39$ pc away. The effect of high rotation [Royer et al., 2007] ($v \sin i = 240$ kms$^{-1}$) and its high temperature [Zorec and Royer, 2012] ($T_{\text{eff}} = 8260$ K) yield broader and fewer spectral lines compared to the cooler M-dwarfs, thus making Iot Cyg an interesting candidate on which to probe wobble capability in such a regime.

Iot Cyg data set contains a total of 266 spectrum hence no data cut was necessary (see methodology). I selected orders in the range 25–50 once they presented overall similar shapes and contain higher SNR compared to the remaining orders, thus causing less issues in the continuum normalization procedure. Indeed, only 8 epochs were deleted from the sample after a polynomial of 1st degree was utilized. For the number of iteration $N_{\text{iter}}$, a reasonable estimation based on test runs turned out to be $N_{\text{iter}} = 200$. In order to search the number of basis vectors $K_\star$ and $K_t$, a grid of values with $K_\star \in 0,1$ and $K_t \in 0,1,2,3,4$ was set up. The reason for the shorter span of $K_\star$ is based on the assumption that stellar spectrum is less variable than the telluric’s, thus an increase in the number of basis vectors may lead to potential data overfitting.

The studies on previous targets imply that wobble’s performance may be affected by distinct amounts of atmospheric features, therefore I selected orders 25 and 38 so that we could analyse the data modeling for different basis vectors. Figures 4.17 and 4.18 display order 25 for distinct basis vectors $K_\star$ and $K_t$ and Fig. 4.19 represents the RV time-series at the same order. In principle, the code is able to localize many atmospheric features in the data independently of basis vectors selection, however a few unrealistic features are present in the telluric components such as the artifacts near 6620 Å and several bumps. Concerning the streaky line near 6620 Å, CCD pixel defects are translated to not-a-number (nan) values which are not well handle by the code, thus we set these ‘nan’s to the zero level flux.8 The stellar component captured a few atmospheric features mostly in the H$\alpha$ wings (Fig. 4.17 and 4.18 panel (b)) and presented some misalignment w.r.t to data (e.g., see Fig 4.17 panel (a) and (b)). Moreover, we observe several telluric features in the residuals of many runs for order 25. These issues are reflected directly in the quality of RVs as seen in Fig. 4.19. Therefore, a reasonable basis vectors for this target based on order 25 would be $K_\star = 0$ and $K_t = 3$ because it presents the best modeling and lowest RV rms. However, at order 38 (Figure 4.20 and 4.21), we notice that all the runs had telluric template with many emission-like features whereas the stellar component fits great amounts of atmospheric lines independently of basis vectors. As acknowledged for other targets, the presence of large amounts of telluric features in CARMENES redward orders is probably the reason for the difficulty in the modeling process. The regularization routine is supposed to aid in these situations by enforcing adequate penalties to both templates, hence dumping emission-like features and adjusting the variability weights so that better fit data. Therefore, we ran wobble with regularization and parameters $N_{\text{iter}} = 200$, $K_\star =}$

---

8 We ran wobble placing ‘nan’s at different levels such as median, 10% below and above the median. Yet, setting them to zero yielded lower RV rms.
(a) $K_\star = 0$ and $K_t = 0$.

(b) $K_\star = 0$ and $K_t = 1$.

(c) $K_\star = 0$ and $K_t = 2$.

Figure 4.17: wobble modeling without regularization for iot Cyg at order 25 for fixed $N_{\text{iter}} = 200$ and different basis-vector components. The artifacts in the telluric component at 6620 Å is due to bad data points.
4.3. IOT CYG STAR

(a) $K_\star = 0$ and $K_t = 3$

(b) $K_\star = 0$ and $K_t = 4$

(c) $K_\star = 1$ and $K_t = 3$

Figure 4.18: Continuation from Fig. 4.17
(a) $K_\star = 0$ and $K_t = 0$

(b) $K_\star = 0$ and $K_t = 1$

(c) $K_\star = 0$ and $K_t = 2$

(d) $K_\star = 0$ and $K_t = 3$

(e) $K_\star = 0$ and $K_t = 4$

(f) $K_\star = 1$ and $K_t = 3$

Figure 4.19: Iot Cyg RV time-series from order 25 extracted with wobble for fixed $N_{\text{iter}} = 200$ and different basis vectors. The RVs are only berv corrected since no further corrections terms were available.
4.3. IOT CYG STAR

\[ K_\star = 0 \text{ and } K_t = 0. \]
\[ K_\star = 0 \text{ and } K_t = 1. \]
\[ K_\star = 0 \text{ and } K_t = 2. \]

Figure 4.20: wobble modeling without regularization for iot Cyg at order 25 for fixed \( N_{\text{iter}} = 200 \) and different basis vector components. The artifacts in the telluric component at 6620 Å is due to bad data points originated from CCD pixel defects.
(a) $K_\star = 0$ and $K_t = 3$

(b) $K_\star = 0$ and $K_t = 4$

(c) $K_\star = 1$ and $K_t = 3$

Figure 4.21: Continuation from Fig. 4.17
0 and \( K_1 = 3 \) and compared to the runs without it. The results in panel (a) of Figs. 4.17 and 4.23 show that, with regularization, both telluric and stellar components are smoother specially in the low SN edges, and the artifact near 6620 Å is also well handled. For order 38 (Figs. 4.21 panel (a) and 4.23 panel (b)) we observe, in the run with regularization, that the amplitudes were not enough to neither mitigate the emission-like features in the telluric component or the misfitting in the stellar component, thus yielding poor residuals compared to the run without regularization. The consequence is a large RV rms (Fig. 4.23 panel (d)) which confirms that the models computed worse radial velocities. Therefore, we acknowledge that a few improvements were achieved with regularization, yet not enough to consider that it performs well.

Although I have only displayed two orders for the analysis above, the entire set was scrutinized, and in general this reasoning is applicable to the other orders and epochs. This could be confirmed by the RV time-series from all the runs presented above (Fig. 4.24 and 4.25) where the lowest RV \( \text{wrms} \) is achieved for when we utilize the basis vectors \( K_\star = 0 \) and \( K_1 = 3 \) without regularization. As mentioned in the target selection section, Iot Cyg has fewer stellar features which could represent a potential challenge to \textit{wobble}. This is noticeable through the large error bars in most of the RV measurements independently of order, epoch, basis vectors or regularization. Therefore, we may conclude that although \textit{wobble} regularization shows a slightly better performance for orders such as 25, for the remaining orders it has a sub-optimal behavior.
CHAPTER 4. RESULTS

(a) \(K_0 = 0\) and \(K_1 = 0\)

(b) \(K_0 = 0\) and \(K_1 = 1\)

(c) \(K_0 = 0\) and \(K_1 = 2\)

(d) \(K_0 = 0\) and \(K_1 = 3\)

(e) \(K_0 = 0\) and \(K_1 = 4\)

(f) \(K_0 = 1\) and \(K_1 = 3\)

Figure 4.22: Iot Cyg RV time-series from order 25 extracted with \texttt{wobble} for fixed \(N_{\text{iter}} = 200\) and different basis vectors. The RVs are only berv and drift corrected since no further corrections were available.
4.3. IOT CYG STAR

Figure 4.23: \textit{wobble} executed with regularization routine.
(a) $K_\star = 0$ and $K_1 = 0$

(b) $K_\star = 0$ and $K_1 = 1$

(c) $K_\star = 0$ and $K_1 = 2$

(d) $K_\star = 0$ and $K_1 = 3$

(e) $K_\star = 0$ and $K_1 = 4$

(f) $K_\star = 1$ and $K_1 = 3$

Figure 4.24: Iot Cyg RV time-series extracted with wobble without regularization for fixed $N_{\text{iter}} = 200$ and different basis vectors. RVs are only berv corrected since no further corrections terms were available.
4.4 TEEGARDEN’S STAR

Figure 4.25: RV time-series with regularization and parameters $K_\star = 0$ and $K_t = 3$.

4.4 Teegarden’s star

Teegarden star (or J02530+168) is an M-dwarf with $M_\star = 0.08 M_\odot$ and brightness $J \approx 8$. It is located at $\sim 12$ ly away and host two planets [Zechmeister et al., 2019]. Similarly as to the other targets, we run wobble for the basis vector grid $K_\star = 0$ and $K_t \in \{0, 1, 3\}$. The $N_{\text{iter}}$ parameter was set to 400 based on the likelihood convergence. On the continuum normalization step, I utilized a polynomial function of 4th degree where 26 epochs were excluded from the total 243. The final RV time-series for the runs with $K_\star = 0$ and distinct $K_t$ (Fig. 4.26) implies that the regularization shown on the right column fails in comparison to the runs without it (left column). Besides that, increasing $K_t$ from 0 to 1 yielded RV data with lower rms whereas the run with $K_t = 3$ calculates RVs for which the rms escalates, hence this is evidence that we must probe $K_t = 2$. Indeed, the results (Fig. 4.27) extracted better RVs ($RV_{\text{rms}} = 3.58 \, \text{ms}^{-1}$). We do not compare this results to the outcomes from the run with regularization because the calculation of regularization amplitudes take several hours, and previous Teegarden’s star runs (Fig. 4.26 right column) implies that the RV data is likely to fail as well. Therefore, we compare the outcomes for the runs with parameters $K_\star = 0$ and $K_t \in \{1, 2\}$ without regularization in order to probe the modeling differences that made wobble with $K_t = 2$ perform better. Figures 4.28 and 4.29 display orders 25 and 43 for the same arbitrarily chosen epochs. Notice that both orders have poor modeling with non-realistic features in stellar and telluric components. Order 25 (Fig. 4.28) shows wiggly telluric

---

9I choose $K_t = 3$ to test slightly higher magnitudes, if the RV rms decreases the more we increase the basis vectors, I would run wobble with more numbers of basis vectors.
component whereas the stellar template misses most of the data points at the bottom (panel (a) and (b)). The latter effect could be due to the Hα variability observed in several epochs. Since these runs had $K_*=0$, the stellar template is less flexible, thus an attempt to model constantly varying spectra may force the model to shift upwards. We tested wobble with parameters $K_* = 1$ and $K_t = 2$ in an attempt to better model the Hα feature by giving some variability degree to the stellar component, yet no improvement was achieved where the RV at order 25 had $rms_{RV} = 33.3 \text{ms}^{-1}$ and total $rms_{RV} = 3.79 \text{ms}^{-1}$. At Order 43 (Fig. 4.28), it is clear from the residuals in panel (a) that the telluric template was not able to capture the atmospheric features, thus several spikes are shown whereas in panel (b) the residuals seem smoother thus retrieving better RVs (panel (d)). this order 43 (and a few others) showed a lower RV spread compared to serval (in red), however, more orders had larger spread, hence the final RV time-series rms observed in Fig. still favors serval.

Therefore, for Teegarden’s star the parameters that allowed the templates to reproduce the data more faithfully were: $K_* = 0$, $K_t = 2$ and $N_{iter} = 400$ with no aid of regularization. Moreover, we acknowledge from Teegarden’s star and J08413+594 results that perhaps activity may represent a challenge to the calculation of wobble regularization amplitudes. Although we did not focus much on the regularization, I show in appendix the run with parameters $K_* = 0$ and $K_t = 1$ where the wobble amplitudes seem to have penalized the templates, yet not as significant as in J08413+594.

---

10 Assuming that this reasoning is valid, epochs with Hα in absorption would in principle shift the stellar template downwards, unfortunately no Hα in such a condition was possible to encountered visually in our sample.

11 see more orders in appendix
Figure 4.26: Teegarden’s star RV time-series calculated for the parameters $N_{iter} = 400$, $K_* = 0$ and different $K_t$. For all the plots we compare the RV rms of SERVAL (red) with wobble’s (black).
Figure 4.27: wobble with lowest RV spread. This run was executed without regularization and parameters $N_{\text{iter}} = 400$, $K_* = 0$ and $K_t = 2$. 
4.4. TEEGARDEN’S STAR

Figure 4.28: Order 25 modeling with parameters $N_{\text{iter}} = 400$, fixed $K_* = 0$ and $K_t = 1$ (top panel) and $K_t = 2$ (middle panel) without regularization. Their respective RV time-series (black) are shown at the bottom compared to \texttt{serval} (red) RV data.
CHAPTER 4. RESULTS

(a) $K_t = 1$

(b) $K_t = 2$

(c) RV from panel (a)

(d) RV from panel (b)

Figure 4.29: Same as Fig. 4.28 for order 43.
Chapter 5

Conclusion

Wobble developers tested the code with HARPS data whose spectral range lies in the 383–690 nm. The goal of this master thesis was to test the code’s capability in a different wavelength regime located at 520–960 nm where abundant atmospheric features are present, thus modeling the telluric transmission spectrum turned out to be more difficult. For this study we selected three M-dwarfs with slightly different properties and an A-type standard telluric star. We probed the influence that each wobble’s parameter had in the data modeling. We noticed that only three of them ($N_{\text{iter}}$, $K_\star$ and $K_t$) played a crucial role, i.e., the results implied that different star needed distinct set of parameters in order to be well modeled. Another important implementation of wobble was extensively tested: the regularization routine. In principle it should aid the templates to circumvent unwanted features, yet it fails usually for orders where large variability (e.g., Hα activity) takes place or significant amount of atmospheric features are present. The likelihood function at orders that meet these criteria often do not converge, and if so, it takes longer than the other orders.

Barnard’s star results with parameters $N_{\text{iter}} = 200$, $K_\star = 0$ and $K_t = 0$ had the lowest RV spread ($\text{rms}_{\text{RV}} = 2.89 \text{ m s}^{-1}$) for when regularization was utilized. This implies that the data were better modeled in comparison with the runs without regularization (where $\text{rms}_{\text{RV}} = 3.24 \text{ m s}^{-1}$). Therefore the analysis performed on Barnard’s star implies that regularization is important although the rms differences are not large. It is important to point out that wobble performance with regularization for other targets were judged sub-optimal, therefore it may be that the quietness property of Barnard had some leverage on the modeling concerning stellar activity. Yet, if we compare wobble best RV calculation ($\text{rms}_{\text{RV}} = 2.89 \text{ m s}^{-1}$) with serval’s ($\text{rms}_{\text{RV}} = 2.69 \text{ m s}^{-1}$), it implies that the latter still performs better.

J08413+594 was the only star in which wobble presented lower RV spread ($\text{rms}_{\text{RV}} = 46.03 \text{ m s}^{-1}$) compared to serval ($\text{rms}_{\text{RV}} = 46.44 \text{ m s}^{-1}$). This run had parameters $N_{\text{iter}} = 200$, $K_\star = 0$ and $K_t = 0$ and no regularization. The comparison between runs with and without regularization showed that many orders were poorly modeled with regularization. This happened due to the strong penalties imposed by the amplitudes which could be explained from the code’s assumption that
stellar variability must be low compared to telluric changes. For instance, the order containing Hα presented significant flux change which was likely interpreted as something other than a stellar feature. Therefore large amplitudes were applied, yielding poor modeling and large RV data spread (e.g., order 25 with $rms_{RV} = 51.03 \text{ m s}^{-1}$ and without it $rms_{RV} = 72.44 \text{ m s}^{-1}$ regularization). Although this may explain the large penalties in order 25, a few other orders had similar effects, yet no apparent activity was observed.

Iot Cyg results imply that stars with fewer stellar lines represent a remarkable challenge to wobble. Since most of the orders had only stellar continuum, we expected that the telluric component would be severely penalized whereas the stellar component would fit the continuum entirely. For orders with heavy atmospheric contamination, the contrary should take place, i.e., the telluric suffers less constrains compared to the stellar template. Yet, the regularization was unable to select optimal amplitudes, allowing the fit of atmospheric features by the stellar component. Therefore, the RV time-series for Iot Cyg seems quite inaccurate independently of basis vectors. Despite that, we selected $K_\star = 0$ and $K_t = 3$ as a suitable parameters compared to the other trials. I should point out that the order 25 in particular contained a nice looking Hα line that wobble was able to model. This orders presented RV time-series ($rms_{RV} = 26.96 \text{ m s}^{-1}$) without regularization and with it ($rms_{RV} = 26.28 \text{ m s}^{-1}$). Although the other orders had poor modeling in general, runs without regularization computed RVs with lower spread, hence the final RVs with no regularization were better ($rms_{RV} = 138.67 \text{ m s}^{-1}$) in comparison to without it ($rms_{RV} = 56.44 \text{ m s}^{-1}$). Due to the sub-optimal behavior of the code while modeling the orders, the RVs error bars were significantly large, hence we are uncertain about the RVs accuracy for Iot Cyg. Perhpas, more studies on this type of star must be carried out in order to confirm that wobble may perform poorly for stars with fewer stellar lines which is the case of A-type stars.

Teegarden's star had better RV data ($rms_{RV} = 3.58 \text{ m s}^{-1}$) for $N_{iter} = 400$, $K_\star = 0$ and $K_t = 2$ without regularization. This target had a visible Hα activity level that may have influenced wobble regularization similarly as to J08413+594, i.e., the templates were highly penalized for the runs with regularization causing a large spread in the RV time-series. Whether the reason for the sub-optimal amplitudes selected by wobble is connected to stellar activity is still uncertain, however, Barnard’s star being the only quiet M-dwarf in our sample did not suffer with severe penalization for runs with regularization, hence yielding outcomes consonant with the usage of this routine.

Therefore, from this work with wobble we acknowledge issues related to continuum normalization and data modeling. Perhaps the future version of wobble could use a different type of normalization, otherwise a possible way to circumvent this problem is through an external data continuum normalization. The peculiarities with respect to the poor data modeling revolves around the orders with heavy atmospheric contamination where wobble fails most of the time. Moreover, as mentioned above, activity may play a role in the selection of amplitudes in the regularization routine. Despite the problems we came across, wobble is able to calculate radial velocities consistent with serval. The three M-dwarfs RV time-series had RV rms differences within the 2 m s$^{-1}$ independently of basis vectors utilized, although serval had been superior
most of the time. Therefore, we conclude that **wobble** is a promising piece of software that could perform well with CARMENES survey, yet some optimization must be implemented in the code.
Bibliography


61


Appendix A

Appendix

A.1 Barnard’s star

More results from the runs with parameters $N_{iter} = 200$, $K_s = 0$ and $K_t = 0$ are shown below with and without regularization. Figures A.1 and A.3 display the runs without and with regularization whereas Figs. A.5–A.9 display the models for several orders with the same mentioned parameters. Notice that different orders have distinct regularization amplitudes which may influence the modeling at a given order. The orders containing significant amount of tellurics presented more RV spread as one could infer from all the images. Concerning the continuum normalization, notice that at order 37 (3rd and 4th panel of Fig. A.8) wobble fails due to the telluric band located at the right edge of this order.
Figure A.1: Barnard’s RV time-series results with regularization for several order. We compare wobble (black circles) with serval (red squares) where berv, drift, nzp corrections were applied.
A.1. BARNARD'S STAR

Figure A.2: Continuation of Fig. A.1

\[ \text{RV} - \text{RV}_{\text{rms}} \text{ (m} \text{s}^{-1}) \]

- BARNARD CARM-VIS ORDER 31
  - Wobble rv rms = 3.97
  - SERVAL rv rms = 4.02

- BARNARD CARM-VIS ORDER 35
  - Wobble rv rms = 21.01
  - SERVAL rv rms = 3.16

- BARNARD CARM-VIS ORDER 37
  - Wobble rv rms = 35.49
  - SERVAL rv rms = 6.87

- BARNARD CARM-VIS ORDER 39
  - Wobble rv rms = 3.77
  - SERVAL rv rms = 3.26

- BARNARD CARM-VIS ORDER 41
  - Wobble rv rms = 37.32
  - SERVAL rv rms = 2.31

- BARNARD CARM-VIS ORDER 43
  - Wobble rv rms = 159.36
  - SERVAL rv rms = 25.17

- BARNARD CARM-VIS ORDER 44
  - Wobble rv rms = 135.83
  - SERVAL rv rms = 23.96

- BARNARD CARM-VIS ORDER 45
  - Wobble rv rms = 18.64
  - SERVAL rv rms = 9.06
Figure A.3: wobble run similar to Fig. A.2 without regularization
Figure A.4: Continuation of Fig. A.3
Figure A.5: From top-down 1st and 2nd panels show the wobble runs for orders 17 without and with regularization respectively. Order 21 is shown in 3rd and 4th panel without and with regularization displayed in the same format. These models retrieved the RVs by order such as in Fig. [A.1]
Figure A.6: Same as Fig. A.5 for order 21 and 25.
Figure A.7: Same as Fig. A.5 for order 29 and 30.
Figure A.8: Same as Fig. A.5 for order 33 and 37.
Figure A.9: Same as Fig. A.5 for order 41 and 45.
A.2 J08413+594

As mentioned in the results for J08413+594, wobble regularization calculates amplitudes that penalize severely the components, particularly the stellar template. Notice that in orders such as in Figs. A.11d, A.13d, and A.15d, the stellar component is not able to fit the data, thus the telluric template fit most of the data in an attempt to retrieve minimum residuals. The consequence is observed as a large spread in the RV rms in panel (d).
Figure A.10: Comparison between runs with and without regularization and their by-order RV time-series for order 22. *serval* results are shown in red for comparison. The parameters utilized are: $K_\ast = 0$, $K_t=1$ and $N_{\text{iter}}=400$. 

(a) No regularization

(b) regularization

(c) RV no Regularization

(d) RV with regularization
Figure A.11: Same as in Figure A.10 for order 24
Figure A.12: Same as in Figure A.10 for order 32.
Figure A.13: Same as in Figure A.10 for order 32
Figure A.14: Same as in Figure A.10 for order 38.
A.2. J08413+594

Figure A.15: Same as in Figure A.10 for order 47
A.3 Teegarden’s star

Figures A.16 and A.17 show more orders from the runs with fixed $K_\star = 0$ and distinct $K_\varepsilon$ without regularization. Figures A.18–A.20 display the results with regularization and parameters $K_\star = 0$ and $K_\varepsilon = 1$. The comparison between Teegarden’s star and J08413+594 which present H$\alpha$ activity to Barnard’s star, a quiet M-dward imply that activity may influence how wobble calculates the regularization amplitudes. Notice that Barnard’s star templates are less penalized compared to J08413+594 and Teegarden’s star.
Figure A.16: Order 30 modeling with parameters $N_{\text{iter}} = 400$, fixed $K_* = 0$ and $K_t = 1$ (top panel) and $K_t = 2$ (middle panel) without regularization. Their respective RV time-series (black) are shown at the bottom compared to $\text{}$.

\[ (a) \quad K_t = 1 \]

\[ (b) \quad K_t = 2 \]

\[ (c) \quad \text{RV from panel (a)} \]

\[ (d) \quad \text{RV from panel (b)} \]
Figure A.17: Same as in Fig. A.17 for order 38
A.3. TEEGARDEN’S STAR

Figure A.18: wobble run with parameters $N_{\text{iter}} = 400$, $K_\star = 0$ and $K_1 = 1$ with regularization for orders 21, 24, 25 and 27.
Figure A.19: Same as Fig. A.18 for orders 29, 32, 36 and 38.
Figure A.20: Same as Fig. A.18 for orders 39, 41, 42 and 45.