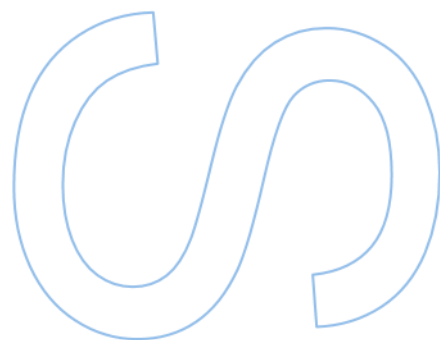
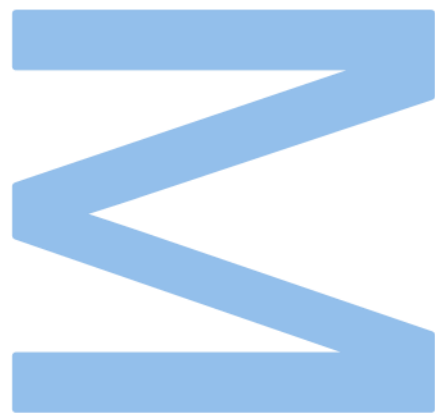


Seeing behind the planets: probing the surface of distant stars using planetary transits

Eduardo Campos Gonçalves

Master in Astronomy and Astrophysics
Department of Physics and Astronomy
Faculty of Sciences of the University of Porto
2024



Seeing behind the planets: probing the surface of distant stars using planetary transits

Eduardo Campos Gonçalves

Dissertation carried out as part of the Master in Astronomy and Astrophysics

Department of Physics and Astronomy
2024

Supervisor

Nuno M. C. Santos, Professor, DFA, FCUP and Researcher,
Institute of Astrophysics and Space Sciences, IA/CAUP

Co-supervisor

Sérgio A. G. Sousa, Researcher, Institute of Astrophysics and Space Sciences, IA/CAUP



UNIVERSIDADE DO PORTO

MASTERS THESIS

**Seeing behind the planets: probing the
surface of distant stars using planetary
transits**

Author:

Eduardo C. GONÇALVES

Supervisor:

Nuno C. SANTOS

Co-supervisor:

Sérgio G. SOUSA

*A thesis submitted in fulfilment of the requirements
for the degree of MSc. Astronomy and Astrophysics*

at the

Faculdade de Ciências da Universidade do Porto
Departamento de Física e Astronomia

November 20, 2024

Acknowledgements

I would like to thank my supervisors Professor Nuno Santos and Researcher Sérgio Sousa for giving the opportunity to work in this project and for offering this fellowship. Their guidance during this year was essential and deeply appreciated and it was a pleasure to work with them.

I would also like to thank those that helped me along my dissertation namely Eduardo Cristo for always being available to guide me for the entire year. I also thank Jorge Martins for the installation of Molecfite and Tomás Silva to help with the telluric correction.

I also thank the course director Professor Catarina Lobo for the support given during these two years.

I thank my parents Sandra Gonçalves and José Gonçalves for the support they have been giving me for my entire life and for being always on my side. Without them I could not achieve anything.

I thank the rest of my family for also supporting me, especially my cousin Sandrina for serving as a role model for me.

I also thank the friends I made during my Master's for improving the time I had these past years.

This project was funded by the European Union (ERC, FIERCE, 101052347). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them. This work was supported by FCT - Fundação para a Ciência e a Tecnologia through national funds and by FEDER through COMPETE2020 - Programa Operacional Competitividade e Internacionalização by these grants: UIDB/04434/2020; UIDP/04434/2020.

UNIVERSIDADE DO PORTO

Abstract

Faculdade de Ciências da Universidade do Porto

Departamento de Física e Astronomia

MSc. Astronomy and Astrophysics

Seeing behind the planets: probing the surface of distant stars using planetary transits

by [Eduardo C. GONÇALVES](#)

For a long time, spectroscopy has been used to analyze and characterize the nature of stars, from their composition to the amount of stellar activity. Spectroscopy has also been used for the detection of exoplanets through the radial velocity method as was the famous discovery of 51 Pegasi b back in 1995. Since then thousands of exoplanets have been discovered using a variety of methods.

As more and more exoplanets are known, interest has shifted to characterizing these bodies and to obtain data from small terrestrial exoplanets. This poses a challenge as the signal they induce in the light we observe is very small and on similar order to stellar activity. In fact, stellar rotation and line of sight phenomena like convective blueshift, are known to affect the shape and strength of spectral lines across a star's surface. The former can be observed through the Rossiter–McLaughlin effect and the latter is shown for example by [Cegla et al. \(2018\)](#).

Our current technology does not allow for the resolution of details in stellar surfaces, however in the last few years a "Doppler shadow" method that relies on planetary transits has been employed to probe the spectrum of a stellar surface partly obscured by a transiting planet. Such method was employed by another work of [Cegla et al. \(2016\)](#) for HD 189733, also by [Dravins et al. \(2017b\)](#) and [Dravins et al. \(2018\)](#) for both HD 209458 and HD 189733 respectively.

Until now these observations have been conducted with the use of Cross Correlation functions (CCFs) due to this method requiring very high resolution spectra. In this dissertation not only do we attempt at recreating the results from [Dravins et al. \(2018\)](#), but with new high resolution data from the ESPRESSO spectrograph, we extend this analysis to individual lines, namely the sodium doublet lines and the H α line.

UNIVERSIDADE DO PORTO

Resumo

Faculdade de Ciências da Universidade do Porto

Departamento de Física e Astronomia

Mestrado em Astronomia e Astrofísica

Vendo atrás dos planetas: sondagem da superfície de estrelas distantes através do uso de trânsitos planetários

por [Eduardo C. GONÇALVES](#)

Por muito tempo, espectroscopia tem sido usada para analisar e caracterizar a natureza das estrelas, desde a sua composição à quantidade de atividade estelar. Espectroscopia tem também sido usada para a deteção de exoplanetas através do método das velocidades radiais, como foi o caso da descoberta famosa de Pegasi b em 1995. Desde então milhares de exoplanetas têm sido descobertos usando uma variedade de métodos.

À medida que mais exoplanetas são conhecidos, o interesse tem mudado para a caracterização destes corpos e a obtenção de dados de pequenos exoplanetas terrestres. Isto apresenta um desafio pois o sinal induzido por eles na luz que observamos é muito pequeno e de ordem semelhante à atividade estelar. De facto, rotação estelar e fenómenos na linha de visão como o desvio para o azul convectivo, são conhecidos por afetar a forma e intensidade das riscas espectrais ao longo da superfície da estrela. O primeiro pode ser observado através do efeito Rossiter-McLaughlin e o último é demonstrado por exemplo em [Cegla et al. \(2018\)](#).

A tecnologia atual não possibilita a resolução de detalhes na superfície estelar, contudo nos últimos anos um método “Doppler shadow” que usa trânsitos planetários tem sido empregado para sondar o espectro de uma superfície estelar parcialmente obscurecida por um planeta em trânsito. Tal método foi empregado por outro trabalho de [Cegla et al. \(2016\)](#) para HD 189733, também por [Dravins et al. \(2017b\)](#) e [Dravins et al. \(2018\)](#) para ambos HD 209458 e HD 189733 respetivamente.

Até agora estas observações têm sido conduzidas através do uso de funções de relação cruzada (CCFs) devido a este método precisar de espectro de resolução muito alta. Nesta

dissertação não só tentamos recriar os dados de [Dravins et al. \(2018\)](#), mas com um novo conjunto de dados de alta resolução do espectrógrafo ESPRESSO, nós estendemos esta análise a riscas individuais, nomeadamente as riscas do dubleto de sódio e a risca $H\alpha$.

Contents

Acknowledgements	iii
Abstract	v
Resumo	vii
Contents	ix
List of Figures	xi
1 Introduction	1
1.1 Exoplanet Detection Methods	2
1.1.1 Transit Photometry Method	2
1.1.2 Radial Velocity Method	5
1.2 The Rossiter–McLaughlin effect	7
1.3 The Doppler Shadow method	8
1.4 Center to limb variations	11
2 Data and Methodology	13
2.1 The Target: HD 189733	13
2.2 The Dataset	15
2.2.1 The ESPRESSO Spectrograph	15
2.2.2 The Observations	15
2.3 Spectral data reduction	17
2.3.1 Telluric line correction with Molecfit	17
2.3.2 Normalization of the spectra	19
2.3.3 Cross Correlation of the spectra	24
2.4 Extraction of local stellar spectra	25
2.4.1 CCF analysis	26
2.4.2 Individual spectral line analysis	39
3 CCFs and individual line analysis results	47
3.1 White light CCF	48
3.2 Iron CCFs	53
3.3 Sodium Doublet lines	59
3.4 H α line	65

4 Discussion of the data - Center to limb variations	71
4.1 The white light CCF	72
4.2 Iron CCFs	74
4.3 Sodium Doublet lines	79
4.4 H α line	80
5 Conclusion	83
A Uncertainties	87
A.1 CCF uncertainty propagation	87
A.2 Individual lines uncertainty propagation	91
B Derivation of μ	93
C Statistical tests	97
D Telluric correction comparison	101
E Iron lines list	103
Bibliography	105

List of Figures

1.1	Schematic of a transit light curve based on the figures of Seager & Mallen-Ornelas (2003).	3
1.2	The Radial Velocity curve of 51 pegasi b an exoplanet with low eccentricity so the curve resembles a sinusoidal taken from Mayor & Queloz (1995) . . .	6
1.3	3 different configurations of alignment of the planet's path across the stellar disk taken from https://wasp-planets.net/tag/rossiter-mclaughlin-effect/ . The angle between the planet's orbital plane and the stellar rotation axis in addition to the impact parameter can affect the shape and strength of the resulting curve.	8
1.4	Representation of the difference between the stellar surface outside of transit and the star during transit. This was inspired by Figure 8 in Dravins et al. (2017a). This representation is also a real relative scale diagram of the multiple positions along the path of transit and the difference in radius between the star HD 189733 A and its planet	9
2.1	This plot taken from Dravins et al. (2017b) contains the transit depth vs visual magnitude of the exoplanets from the Exoplanet Orbit Database (2017) (Han et al. 2014). The diameter of the dots is proportional to the transit duration.	14
2.2	Airmass of the two sets of observations in function of the orbital phases and with the interval of transit. In blue are the observations on the 11/08/2021 and red for the 31/08/2021 with the shaded area representing the extent of the transit, and the darker region excludes the ingress and egress of the transit.	16
2.3	Signal to noise ratios for two orders of the ESPRESSO spectrograph those being the 56th (dots) and 69th (crosses) echelle orders of the spectrum which correspond to the regions around the sodium doublet and H α spectral lines. In blue is the data for the observations on the 11/08/2021 and red for the day 31/08/2021.	17
2.4	Top: locally Normalized spectra of the observations for the day 11/08/2021 between 594 nm and 594.5 nm showcasing telluric lines from Earth's atmospheres which vary over the course of the observations, the orbital phases of the observations are color coded according to the colorbar on the right. Bottom: same description as top, but the spectra were corrected with Molecfit removing the telluric lines.	19

2.5	Top: globally normalized spectra of the observations for the day 11/08/2021 color coded according to the orbital phases of the observations according to the colorbar on the right. Bottom: same description as above, but the spectra were corrected using Molecfit removing the telluric lines, over-fitting of the saturated telluric lines is also visible. The multiple spectra are overlaid on top of each other so the changes are subtle, but noticeable.	20
2.6	Spectra with zoomed in sections, in the middle we see the area whose mean was used for global normalization while on the left and right zoomed in sections we see a dependence on the orbital phases showing that this type of normalization is insufficient.	21
2.7	Local normalization of the sodium doublet lines. At the top we see an example of the sodium doublet after global normalization and before local normalization in yellow with the black region being the continuum from the selected wavelength interval for line fitting. The central plot we see all spectra for the day 11/08/2021 (the same happens for 31/08/2021) colored based on their orbital phases according to the colorbar on the right. Overlaid on top are dashed lines of the linear regressions using the fitting parameters from the fit of the selected wavelength intervals, with the same color as their respective spectra. The bottom plot shows the spectra after local normalization using each respective linear fit.	22
2.8	Local normalization of the H α line. At the top we see an example of the sodium doublet after global normalization and before local normalization in yellow with the black region being the continuum from the selected wavelength interval for line fitting. The central plot we see all spectra for the day 11/08/2021 (the same happens for 31/08/2021) colored based on their orbital phases according to the colorbar on the right. Overlaid on top are dashed lines of the lines using the fitting parameters from the fit of the selected wavelength intervals, with the same color as their respective spectra. The bottom plot shows the spectra after local normalization using each respective linear fit.	23
2.9	Flowchart containing the basic steps in our methodology to get from the starting data to the subtracted CCFs or spectral lines. The top represents the process for the CCFs while the bottom is the process for the individual line analysis. The rounded boxes represent the data at its multiple steps, arrow boxes represent the main processes, the diamond represent the final process which requires multiple sets of data. The black color shows the data between the multiple stages.	26
2.10	(Top day 11/08/2021, bottom day 31/08/2021) Left side: central radial velocities of the CCFs over their orbital phases in red is the values from the in-transit observations while in blue is the out-of-transit observations, with the blue line being a linear fit of them. The shaded area indicates the phases where the transit occurs with the darker region excluding the ingress and egress, error bars are shown colored in black. Right side: same as the left except all observations had the barycentric motion subtracted (blue line). The linear fit as values $m = -1.290 \pm 0.004 km/s$ and $m = -2.1780 \pm 0.0001 km/s$	29

2.11	Barycentric corrected RV values of the weak iron CCFs for the two nights of observation 11/08/2021 left side and 31/08/2021 right side, with the same legend as in Figure 2.10. The linear fit as values $m = -1.29 \pm 0.02 \text{ km/s}$ and $m = -2.0853 \pm 0.0005 \text{ km/s}$ for day 11 and values $m = -1.23 \pm 0.02 \text{ km/s}$ and $m = -2.0952 \pm 0.0006 \text{ km/s}$ for day 31.	30
2.12	Barycentric corrected RV values of the strong iron CCFs for the two nights of observation 11/08/2021 left side and 31/08/2021 right side, with the same legend as in Figure 2.10. The linear fit as values $m = -1.291 \pm 0.008 \text{ km/s}$ and $m = -2.2314 \pm 0.0002 \text{ km/s}$ for day 11 and values $m = -1.22 \pm 0.01 \text{ km/s}$ and $m = -2.2399 \pm 0.0003 \text{ km/s}$ for day 31.	30
2.13	Simulated transit light curves for the two nights of observation, with relative flux with 1 meaning the total stellar flux of the star without transit, the shade area signals where the transit occurs while the darker region excludes the ingress and egress.	31
2.14	The white light local CCFs of the in-transit observations obtained after subtraction, the color of the points represents the orbital phase of the exoplanet at the time of observation, uncertainties are represented by their respective error bars. The top graph is for the observation on the 11/08/2021 while the bottom for the day 31/08/2021.	33
2.15	Tomography of the subtracted white light CCFs, the y axis represents the orbital phase while the x axis the radial velocities of the CCFs, the color represents the relative flux in the scale of the colorbar, which in this case 1 would represent the total flux from the star. The horizontal lines show the start and end of the transit with the dashed lines showing the limits of ingress/egress. In addition the crosses indicate the central radial velocities of the CCF (the same shown in Figure 3.3) with error bars indicating their uncertainties. The top graph is for the observation on the 11/08/2021 while the bottom for the day 31/08/2021.	34
2.16	The weak Fe I lines CCFs of the in-transit observations obtained after subtraction, with the same representation as in Figure 2.14.	35
2.17	Tomography of the subtracted weak Fe I CCFs displayed in the same manner as Figure 2.15.	36
2.18	The strong Fe I lines CCFs of the in-transit observations obtained after subtraction, with the same representation as in Figure 2.14.	37
2.19	Tomography of the subtracted strong Fe I CCFs displayed in the same manner as Figure 2.15.	38
2.20	Barycentric corrected values of the sodium doublet for the two nights of observation 11/08/2021 left side and 31/08/2021 right side, with the same legend as in Figure 2.10.	40
2.21	Barycentric corrected values of the H α line for the two nights of observation 11/08/2021 left side and 31/08/2021 right side, with the same legend as in Figure 2.10. Of note that the error bars have been scaled down by a factor of 100 so we can see the points in more detail, although any deviation is statistically insignificant. This is due to the H α line being very wide and as such the fitting has greater difficulty in finding the center value.	41

2.22	Left side: Master out-of-transit spectrum for the sodium doublet for the two days of observation including day 31 divided by its mean and multiplied by the mean of day 11 for showing the almost perfect alignment. Right side: spectrum of day 31 master out-of-transit sodium doublet divided by day 11 showing minimal noisy difference.	41
2.23	Left side: Master out-of-transit spectrum for the $H\alpha$ line for the two days of observation including day 31 divided by its mean and multiplied by the mean of day 11 which did not show significant difference. Right side: spectrum of day 31 master out-of-transit $H\alpha$ line divided by day 11 showing a difference of the line profile.	42
2.24	The reconstructed sodium doublet lines of the in-transit observations after subtraction, with the same representation as in Figure 2.14, although error bars were excluded for ease of visibility.	43
2.25	Tomography of the reconstructed sodium doublet lines displayed in the same manner as Figure 2.15.	44
2.26	The reconstructed $H\alpha$ line of the in-transit observations after subtraction, with the same representation as in Figure 2.14, although error bars were excluded for ease of visibility.	45
2.27	Tomography of the reconstructed $h\alpha$ line displayed in the same manner as Figure 2.15.	46
3.1	Values of μ in function of the respective orbital phase in circles. The crosses represent the orbital phases where the center of the planet is outside of the stellar disk and μ could not be calculated.	48
3.2	The R^2 values for each of the fits of the local white light CCFs in function of their respective orbital phase.	49
3.3	The sets of data extracted from the subtracted white light CCFs through the use of function fitting. The sets of data are from left to right the central radial velocity of the CCF, the line-width of the CCF profile, and the line-center intensity of the profile. The top row as them represented in function of the orbital phase while the bottom row is in function of μ . The blue data is for the day 11/08/2021 while the red is for the day 31/08/20221 with error bars for their uncertainties. Horizontal solid lines represent for comparison the values from the master out-of-transit which was retrieved in the same way as our subtracted CCFs. The dashed lines represent the linear regressions performed to our data points. The error bars are marked in black and finally the shaded area represents the transit interval with the lighter regions showing the ingress and egress of the transit.	50
3.4	The residuals of each of each of the linear fits done in Figure 3.3. From left to right are the residuals of the center radial velocity, the line width and the line center intensity. The data in blue represents observations for the day 11/08/2021 while the data in red represents data from the day 31/08/2021. The shaded area represents the transit interval with the lighter regions showing the ingress and egress of the transit and error bars are shown in black.	51

3.5	Plots containing distributions of the spearman coefficient and p-values calculated from 10 000 samplings of the data used for each of the 3 linear fits. The top row contains the distribution of spearman coefficients and bottom contains the p-values. From left to right the plots pertain to the linear fits of the center radial velocity, the line width and the line center intensity. . . .	52
3.6	The R^2 values for each of the fits of the local white light CCFs in function of their respective orbital phase. Left: weak iron CCFs. Right: strong iron CCFs.	54
3.7	The sets of data extracted from the subtracted weak Fe I lines CCFs through the use of function fitting, displayed in the same manner as Figure 3.3. . . .	55
3.8	The residuals of each of each of the linear fits done in 3.7, displayed in the same manner as Figure 3.4.	55
3.9	The sets of data extracted from the subtracted strong Fe I lines CCFs through the use of function fitting, displayed in the same manner as Figure 3.3. . . .	56
3.10	The residuals of each of each of the linear fits done in 3.9, displayed in the same manner as Figure 3.4.	56
3.11	Plots containing distributions of the spearman coefficient and p-values calculated from 10000 samplings of the data used for each of the 3 linear fits, displayed in the same manner as Figure 3.5.	58
3.12	Plots containing distributions of the spearman coefficient and p-values calculated from 10000 samplings of the data used for each of the 3 linear fits, displayed in the same manner as Figure 3.5.	59
3.13	The R^2 values for each of the fits of the local sodium doublet lines in function of their respective orbital phase.	60
3.14	The sets of data extracted from the reconstructed local sodium D1 line through the use of function fitting, displayed in the same manner as Figure 3.3.	61
3.15	The residuals of each of each of the linear fits done in 3.14, displayed in the same manner as Figure 3.4.	61
3.16	The sets of data extracted from the reconstructed local sodium D2 line through the use of function fitting, displayed in the same manner as Figure 3.3.	62
3.17	The residuals of each of each of the linear fits done in 3.16, displayed in the same manner as Figure 3.4.	62
3.18	Plots containing distributions of the spearman coefficient and p-values calculated from 10000 samplings of the data used for each of the 3 linear fits, displayed in the same manner as Figure 3.5.	64
3.19	Plots containing distributions of the spearman coefficient and p-values calculated from 10000 samplings of the data used for each of the 3 linear fits, displayed in the same manner as Figure 3.5.	65
3.20	The R^2 values for each of the fits of the local $H\alpha$ lines in function of their respective orbital phase.	66
3.21	The sets of data extracted from the reconstructed $H\alpha$ line through the use of function fitting, displayed in the same manner as Figure 3.3.	67
3.22	The residuals of each of each of the linear fits done in 3.21, displayed in the same manner as Figure 3.4.	67

3.23	Plots containing distributions of the spearman coefficient and p-values calculated from 10000 samplings of the data used for each of the 3 linear fits, displayed in the same manner as Figure 3.5.	69
4.1	Figure 3.3 for ease of view of the results of the white light CCFs.	72
4.2	Direct comparison of the weak iron and strong iron CCFs. Left is Figure 10. of Dravins et al. (2018) while right is our results. The plots from up to bottom contain the line-widths, line-depths, central radial velocities and μ for reference.	75
4.3	Figure 11. of Dravins et al. (2018). From their description: Center-to-limb changes of the parameters for differently strong lines computed from the various simulations representing stars of spectral types F, G, and K. Properties of strong lines are shown by solid curves, of intermediate lines by long-dashed, and of weak lines by short-dashed curves. Line-width measures, line-center intensities, and convective line shifts are from symmetric Gaussian fits to the full synthetic profiles and therefore do not resolve any line asymmetries. Convective line shifts are in absolute units but gravitational redshifts are neglected.	77
4.4	Figure 2 of Cegla et al. (2018). Total RVs from the snapshots of their simulated profiles (lines) and averages over a time-series (points). Original profile results are in the red points, connected by a solid black line, and the reconstructed granulation-only profiles are in green. Also plotted are the oscillation RVs averaged over the entire series (blue points). Plus, a black dotted line at 0 m/s was added to guide the eye.	78
4.5	Plots containing the residuals of the linear fits to the central radial velocities of the strong iron CCFs 3.10 in function of μ . The colors follow the same legend as the referenced Figure.	78
4.6	Figure 2 of Dethier & Bourrier (2023) containing simulated Na I D2 lines at different μ s. The upper plot (case a) represents these lines on a uniform static disk. Case b introduces stellar rotation. Case c the lines are subject to center to limb variations. Case d combines center to limb variations and Doppler shifts due to stellar formation. The bottom plot contains the profiles of the integrated disk for all cases.	79
A.1	Top: interpolation of a strong iron CCF without taking into account uncertainties using formula A.5. Bottom: same as the top except we consider covariances between points using formula A.6. The blue points are the CCF prior to interpolation, while orange is after interpolation. The black error bars are the uncertainties (square root of the diagonal of the covariance matrix) of the CCFs before interpolation with the red error bars being after interpolation. The uncertainties have been scaled up by a factor of 100 for ease of viewing.	90
B.1	Side view perpendicular to our line of sight showing the planet's projection on the stellar surface.	93
B.2	Right side shows a different perspective showing the planet's orbit around a star and the angle α it traces in space.	94
B.3	Front view of the star and planet during a transit.	94

D.1 Tomography of the reconstructed sodium doublet lines without telluric correction displayed in the same manner as Figure 2.15.	102
---	-----

Chapter 1

Introduction

Ever since 1995 ([Mayor & Queloz 1995](#)), thousands of exoplanets have been detected and confirmed through the use of multiple techniques ([Santos et al. 2020](#)). The most common detection method being transit photometry. After detection the existence of these planets can be confirmed through techniques like the radial velocity method.

However, these detection methods are biased towards larger planets that are nearer their host star (subsection [1.1.1](#)). If we are to observe small terrestrial planets around solar type stars with similar orbits to Earth, high signal to noise and high resolution data will be required. Nonetheless even after overcoming that technology barrier one has to deal with stellar surface phenomena that can induce signals that interfere with our observations and cause false positive exoplanet detections ([Santos et al. 2020](#)). These are the major barriers that have to be overcome in order to further develop the study of exoplanets and the search for other earths. To better understand these difficulties it is important to study the spectrum of the stellar surface and how it varies from center to limb, which is what we intend to tackle in this dissertation.

In this work we employ a methodology based on the Doppler shadow method ([Dravins et al. 2017a](#)) to recover the spectra of segments of the stellar surface obscured by its planet.

We improve on the results of how line profiles vary from the stellar disk center to limb obtained by [Dravins et al. \(2018\)](#) for a set of weak Fe I lines and strong Fe I lines. In addition to studying how the typical spectral line of a K2 class star varies with a white light cross correlation function (subsection [2.3.3](#)). In this work we employ new high resolution and high accuracy data from ESPRESSO (subsection [2.2.1](#)) for HD 189733 A.

We also extend this analysis to individual lines like the sodium doublet lines and $H\alpha$. We then analyze how these change along the stellar surface and test the feasibility of the

methodology for future studies.

In this chapter we start by introducing relevant concepts to our work as well as explaining the basics of the Doppler Shadow method. On the second chapter we describe our target of observation and the data used in our work. Moreover, we explain in more detail the methodology employed in implementing the Doppler Shadow method to extract our results. For the third chapter we present the results and offer basic statistical analyses. Finally, in the fourth chapter we discuss the results and our findings.

1.1 Exoplanet Detection Methods

1.1.1 Transit Photometry Method

Transit Photometry is currently the most common detection method, as of writing today it accounts for 4166 confirmed exoplanets with the first transit detection being HD 209458 b in 1999 by [Charbonneau et al. \(2000\)](#), although this planet had already been discovered by the radial velocity method.

This method was first proposed by Frank Rosenblatt in 1971 ([Rosenblatt 1971](#)) and is based on a simple concept. When from our perspective a planet passes in front of its host star covering its surface it is called a transit and we observe its effect as a dip in the flux detected from the host star. The dip in flux is related with the angular sizes of the stellar disk and planetary disk and is proportional to their ratio, and given the vast distances from any star system to Earth this relation is essentially dependent on the ratio of the respective physical radii of the star exoplanet pair. We can then express the maximum dip in flux by the following relation:

$$\frac{\Delta F}{F} = \left(\frac{R_{planet}}{R_{star}} \right)^2 \quad (1.1)$$

Here, ΔF is the decrease in flux, F the total stellar flux outside of transit, and R_{star} and R_{planet} are respectively the stellar radius and planetary radius. This formula gives the maximum relative dip in flux (the same would apply if we used luminosity instead) which is equal to the ratio of the disk areas (in reality this is approximate due to limb darkening). A larger planet will give origin to a stronger and easier to detect signal, for example a Jupiter sized planet with 0.1 times the solar radius on a sun like star would have a transit depth of 1%, while an Earth like planet with 0.01 solar radii would only leave a depth in the order of 0.01% or 100 ppm. Smaller orbital periods will also result in

greater probability of transit and more frequent transits (Santos et al. 2020), as such this method is biased towards giant planets with small periods, the so called hot Jupiters.

Given that exoplanetary transits are one of the main topics of this dissertation it will be relevant to introduce some concepts. Starting by the flux time series or as we will refer from now on as the light curve. A typical light curve will have a similar structure as the one in Figure 1.1

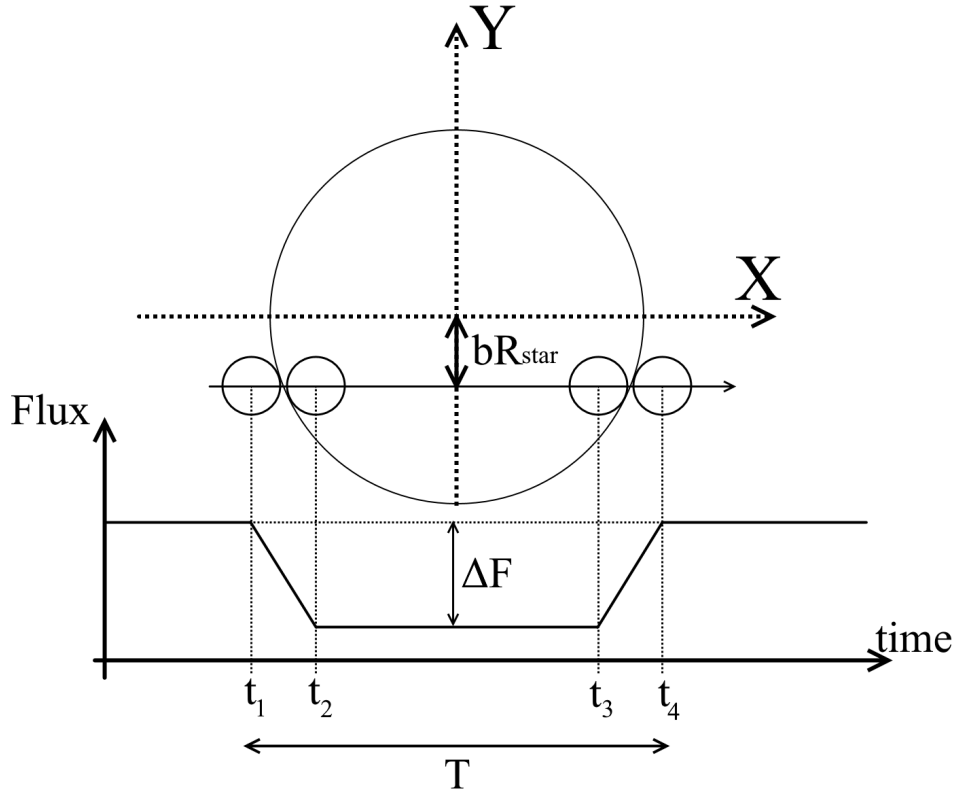


FIGURE 1.1: Schematic of a transit light curve based on the figures of Seager & Mallen-Ornelas (2003).

From Figure 1.1 we see that equation 1.1 only applies to the times between t_2 and t_3 as only then does the planet's disk completely obscure the star, not counting for limb darkening. The intervals between t_1, t_2 and t_3, t_4 are respectively the ingress and egress of the transit when the planet only partially obscures the star.

b is the impact parameter and it quantifies the distance of the planet's path to the center of the star from our line of sight. The impact parameter is related with the orbital inclination (i) and the semi-major axis (a) (for low eccentricities) through:

$$b = \frac{a}{R_{star}} \cos(i) \quad (1.2)$$

The ingress and egress parts of the transit contain information of the exoplanet's orbit as their duration depends on the impact parameter. As a consequence, one can derive the orbital inclination and the ratio between the semi-major axis and stellar radius from the ingress and egress. Something to note is that the shape of the light curve during the ingress or egress is dependent on factors such as the limb darkening which have a complex behavior.

The transit duration (T), between t_1 and t_4 can be given by the following equation for zero eccentricity:

$$T = \frac{P}{\pi} \arcsin \left(\frac{\sqrt{(R_{star} + R_{planet})^2 - (bR_{star})^2}}{a} \right) \quad (1.3)$$

where P is the planet's orbital period. The time of ingress plus egress (T_{ie}) is also given by:

$$T_{ie} = T - \frac{P}{\pi} \arcsin \left(\frac{\sqrt{(R_{star} - R_{planet})^2 - (bR_{star})^2}}{a} \right) \quad (1.4)$$

A complete discussion of the derivation of these equations is available in [Seager & Mallen-Ornelas \(2003\)](#).

One useful quantity that can be defined is orbital phase which is a dimensionless quantity that gives the phase at which the planet is in its orbit and is given by:

$$\phi_t = \frac{t - t_{epoch}}{P_{orbit}} \quad (1.5)$$

where ϕ_t is the orbital phase, t is the time of a given observation and t_{epoch} gives the time at which the planet is at the exact middle of its transit if measured in the same night of the transit of interest. Otherwise an integer number of orbits has to be removed from t_{epoch} and the also for the measured times t .

Another useful quantity is μ which is the cosine of the angle between our line of sight and the normal of the stellar surface at that location. μ is a way to quantify the distance from a position on the stellar disk to its center from our perspective. This is a geometric problem that is similar to the one used to calculate the transit duration and is given by:

$$\mu_t = \frac{\sqrt{R_{star}^2 - (bR_{star})^2 - a^2 \sin^2(2\pi\phi_t)}}{R_{star}} \quad (1.6)$$

A derivation of this equation is available in [appendix B](#).

1.1.2 Radial Velocity Method

The second most common detection method that is also used as a confirmation method is radial velocities. This technique has been used for over 1000 exoplanets so far and also accounts for the first ever detection of an exoplanet of a main sequence star, that being 51 pegasi b in 1995 by Michel Mayor and Didier Queloz ([Mayor & Queloz 1995](#)).

This method relies on observing the Doppler shift in the star's spectrum due to the gravitational interaction between it and its planetary companions. As is commonly known the change in wavelength due to the motion of the source of radiation (the non-relativistic case is enough for our purposes) is given by:

$$\frac{\Delta\lambda}{\lambda} = \frac{v_r}{c} \quad (1.7)$$

where λ is the wavelength of the source if it were at rest, $\Delta\lambda$ is the difference between at rest and observed wavelengths, c is the speed of light and v_r is the speed of the source along our line of sight, i.e., radial velocity (RV).

Since this method measures the star's velocity in the radial direction one can then derive many physical attributes of the star-planet system. We will focus on the simple case of a two body problem which is enough for our purposes (larger multiplanetary systems require more complex statistical analyses).

For a two body problem one can show that the star's movement around the system's barycenter is an ellipse with the same parameters as the planet's orbit around the star. The only difference being the semi-major axis of the star around the barycenter (a_{star}), but this is related to the planet's semi-major axis around the barycenter (a_{planet}) and the one from the planet's orbit in the star's referential (a) through the formula $a = a_{star} + a_{planet}$.

In this situation the radial velocity of the star around the system's barycenter follows a Keplerian curve, examples in [Figure 1.2](#). A Keplerian curve is given by [equation 1.8](#) and a derivation is available in [Murray & Correia \(2011\)](#).

$$v_r = \gamma + K \left(\cos(\omega + f) + e \cos \omega \right) \quad (1.8)$$

where γ is the radial velocity of the system's barycenter, ω is the argument of periastron which is the angle measured from the intersection of the orbit with our reference plane with the periastron in the direction of motion, f is the true anomaly which is the angle between the periastron and the body measured from the focus it orbits around, e is the eccentricity which describes how extreme the ellipse of the orbit is and K is the semi-amplitude of the Keplerian curve and it is given by:

$$K = \left(\frac{2\pi G}{P} \right)^{\frac{1}{3}} \frac{M_{planet} \sin i}{\left(M_{star} + M_{planet} \right)^{\frac{2}{3}}} \frac{1}{\sqrt{1 - e^2}} \quad (1.9)$$

Here, P is the orbital period, G is the gravitational constant, i the orbital inclination, and M_{star} and M_{planet} are respectively the star's and planet's masses. From equation 1.9 it is clear that we need to know the orbital inclination if we want to know the true mass of the exoplanet otherwise we only get a minimal value for its mass. The value for the inclination can be obtained by supplementing the radial velocities with either transit photometry or through the Rossiter-McLaughlin effect as we will see next, although the last one gives the orbital inclination in relation to the stellar inclination.

This method based on spectroscopy allows us to retrieve the mass of the bodies of the system as well as their orbital parameters, being a valuable asset in the study of exoplanets.

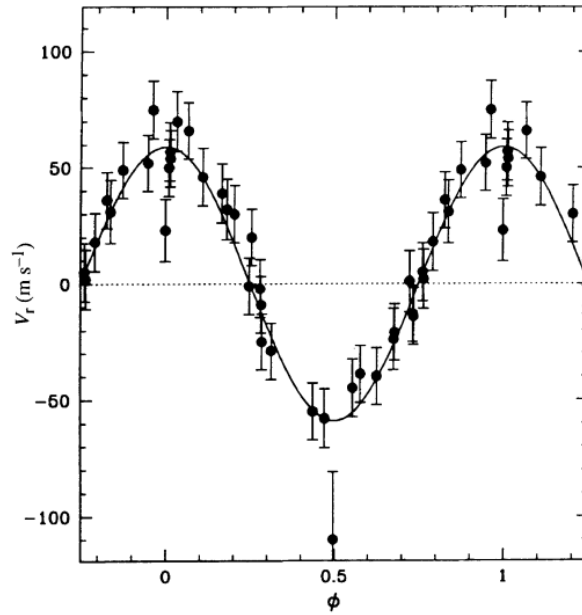


FIGURE 1.2: The Radial Velocity curve of 51 pegasi b an exoplanet with low eccentricity so the curve resembles a sinusoidal taken from [Mayor & Queloz \(1995\)](#)

1.2 The Rossiter–McLaughlin effect

What if we try to apply the radial velocity method during the transit of an exoplanet? This reveals the Rossiter–McLaughlin effect (RM) which is an anomaly in the radial velocity curve of the transit.

Stars like many objects in the universe rotate around their rotation axis and from our perspective we observe the star having a side that is moving away from us and thus has greater redshift, while the opposite side moves in our direction and has greater blueshift and this varies depending on the orientation of the star’s rotation axis with our line of sight and is not the case if they are parallel. The overall Doppler shift then changes in a gradient from one side to the other according to the star’s rotational velocity.

Normally these differences would result in a net zero Doppler shift (ignoring other phenomena in the stellar surface such as convective blueshift) since we observe the integrated spectrum of the entire stellar disk. Adding a transiting planet that obscures successive parts of the stellar disk results in the removal of the respective contribution of each part to the integrated stellar spectrum.

So when the planet crosses over regions of greater blueshift (assuming that the planet’s path is perpendicular to the stellar rotation axis) we observe a redshift on the star’s radial velocity. This cancels out when the planet reaches the middle point of the transit, only to then start obscuring redshifted regions and thus we observe a blueshift until the transit ends. This is shown in Figure 1.3.

Additionally, the Rossiter-McLaughlin effect depends on how the planet’s orbit is oriented, most importantly on the angle between the normal to the orbital plane of the planet and the stellar rotation axis (λ) also known as the spin-orbit angle. If λ is large the planet might only cross one redshifted or blueshifted side as in the third case of Figure 1.3 altering the shape of the RM curve. The impact parameter also alters the amplitude of the RM signal as the planet covers smaller radial velocities in absolute value closer to the poles.

This effect can be used to derive the star’s 3d orientation and quantify the differential rotation of a star as was done for example in [Cegla et al. \(2016\)](#) and in [Cristo, E. et al. \(2024\)](#). This anomaly can also be affected by the star’s limb darkening as shown in [Cristo et al. \(2022\)](#).

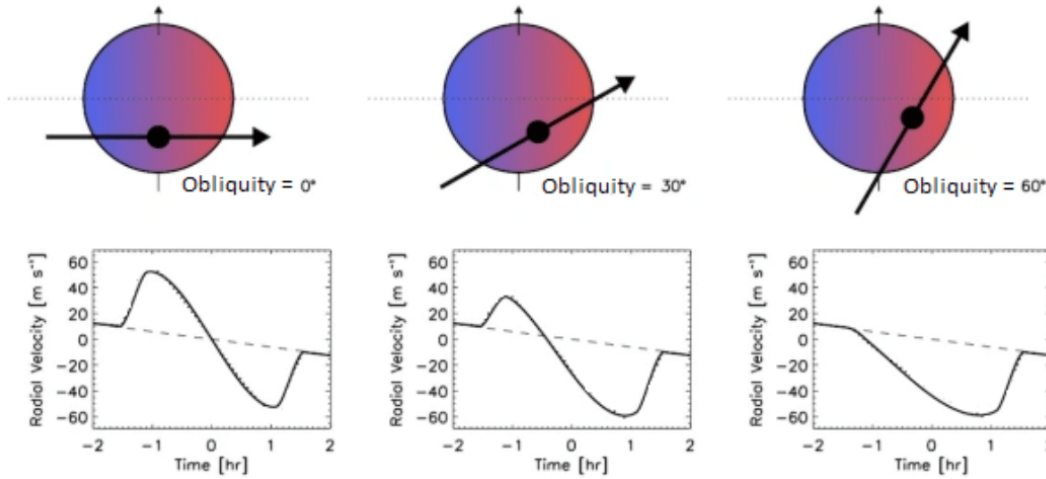


FIGURE 1.3: 3 different configurations of alignment of the planet's path across the stellar disk taken from <https://wasp-planets.net/tag/rossiter-mclaughlin-effect/>. The angle between the planet's orbital plane and the stellar rotation axis in addition to the impact parameter can affect the shape and strength of the resulting curve.

1.3 The Doppler Shadow method

The Doppler Shadow method was first introduced in [Cameron et al. \(2010\)](#). This technique aims to retrieve spectral data from specific sections of the stellar surface which would allow for the study of the physics of stellar surfaces. In addition to studying how stellar rotation and center to limb variations on the stellar disk impact the signals in the stellar spectrum. This would allow for a better understanding and separation of the spectral information of an exoplanet from its host star, since we know for example that center to limb variations can distort the transmission spectra of the sodium doublet lines ([Yan et al. 2017](#)).

This method hinges on using a star with a transiting exoplanet to probe its surface. In contrast to other hypothetical methodologies mentioned in [Dravins et al. \(2017a\)](#) like eclipsing binary stars or some advanced integral field spectrometer, the use of transiting exoplanets is advantageous. As it was concluded that this technique would produce the cleanest and most direct methodology to retrieve high-spatial-resolution spectra of the stellar surface.

Moreover, a given star may have different amounts of convective features on their surface with more massive stars having fewer features compared to smaller ones. Solar type stars have around $\sim 10^6$ granules on their surface. This means that a Jupiter sized planet that covers 1% of a stellar surface would in turn cover about $\sim 10^4$ granules. According to [Dravins et al. \(2017a\)](#) simulations show that this is a statistically significant number of

granules for the spectral profiles from these planet covered regions to become stable, despite the random nature of the granules themselves, even without the need for averaging in time. This results in the convenience that exoplanets are small enough to allow for the resolving of specific areas of the stellar surface, but large enough to not be affected by potential random fluctuations from granules. This of course would start to deteriorate for very small stars or for very large planets.

On a star with a transiting planet as the transit progresses the planet will obscure successive parts of the stellar surface. Resulting in differences we detect on the spectrum of the star in comparison to the spectrum of the star outside of transit due to missing contributions of the obscured regions on the integrated disk spectra. By subtracting the in-transit data from the out-of-transit data we can then retrieve the data of the covered regions. If we have multiple observations along an entire transit, we then get data for successive regions of the stellar surface along the transit chord. Figure 1.4 shows a diagram of the idea behind the subtraction for the stellar system of HD 189733 A which we will study in this work.

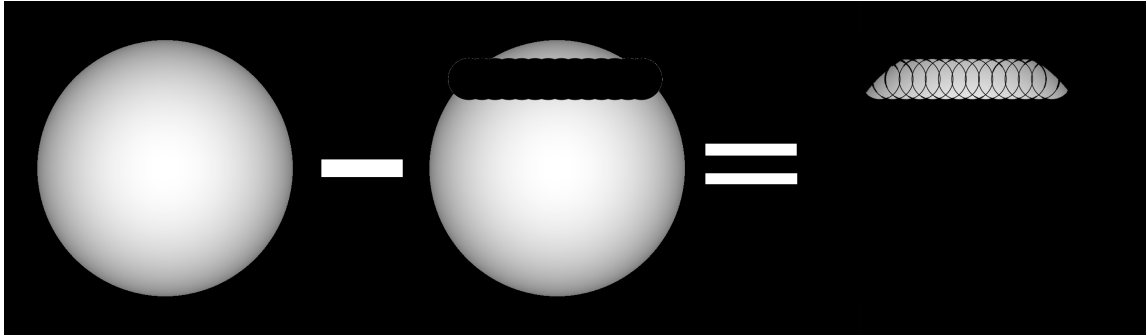


FIGURE 1.4: Representation of the difference between the stellar surface outside of transit and the star during transit. This was inspired by Figure 8 in [Dravins et al. \(2017a\)](#). This representation is also a real relative scale diagram of the multiple positions along the path of transit and the difference in radius between the star HD 189733 A and its planet

Despite its simplicity there are a few difficulties. The first is the performance of a subtraction. If we were to start with a signal to noise ratio of 1000 with a transit depth of 1% our subtracted data would have $1000 - 1000(1 - 0.01) = 10$ S/N, which is a much smaller value. This highlights the need for very high S/N in our initial set of data. One way to counteract this is by averaging over multiple observations at the same position of the transit or by averaging multiple lines by using cross correlations functions (CCFs). The last being done in previous works and which we will replicate, but nevertheless we also extend the analysis to specific spectral lines.

From a purely theoretical perspective the observations required are the absolute irradiance of resolved spectral lines both inside and outside of transit plus the knowledge of the effective area of the planetary disk for the respective moment in-transit (Dravins et al. 2017a). This is due to the obscured flux by the planet being proportional to the effective area of the planet (A_{eff}), which is the projected geometric area of the planet times the local stellar continuum brightness at a given wavelength. The spectral flux (S) normalized in intensity covered by the planet according to Dravins et al. (2017a) could then be retrieved as:

$$S_{covered} = (S_{outside_transit} - S_{inside_transit}) / A_{eff} \quad (1.10)$$

However, the simultaneous measure of high resolution spectra with correct flux variations from ground based telescopes is not practically possible due to variations in atmospheric transmission during the night of observations. So the high resolution spectral measurements need to be complemented with lower resolution space-based photometric data of the transit. The photometric transit has the advantage of already taking into account the effect of the planet's effective area, as it already contains the relative change in flux over the course of the transit. So we can start by normalizing the stellar spectra and later multiply by the correct flux, based on the photometric data before performing the subtraction as given by:

$$F_{covered} = f_{outside_transit} \times F_{photometric_outside_transit} - f_{inside_transit} \times F_{photometric_inside_transit} \quad (1.11)$$

Here, $F_{covered}$ is the resulting subtracted flux or as we will call it the local flux, f denotes the normalized spectral flux and $F_{photometric}$ denotes the photometric flux.

Another detail to point out is that this method assumes we are working on the star's referential. Otherwise when performing the subtraction, the different spectra will be shifted by a radial velocity that is dependent on the observer and the star's motion through space, which will induce us to error.

The Doppler Shadow method allows us to retrieve data from specific sections of the star's surface along a transit chord and will be the core in getting to our results. A more detailed step by step methodology and data manipulation will be given in the next chapter.

1.4 Center to limb variations

Center to limb variations encompass different phenomena that are capable of modifying the observed spectra along the surface of a star. This can include the rotation of the stellar surface, convective blueshift, limb darkening, surface oscillations, among others. Before going into our data analysis and into the results, it is important to understand the phenomena that occur on the surface of stars.

One of the most well-known is the limb darkening. This is an optical phenomenon first observed on the Sun, but also applies to all stars, where the center of the observed disk appears brighter than near the limb. This is well understood and can be derived from the radiative transfer equation whose solution is commonly expressed in the form of a polynomial.

Although limb darkening is a well-known photometric effect, from a spectroscopic perspective it gets more complicated and ties with our work as we observe its resulting variations in flux. This is due to the fact that the behavior of limb darkening changes depending on the observed wavelength.

These phenomena were studied in [Cegla et al. \(2018\)](#) and [Dravins et al. \(2017a\)](#). While presenting the Rossiter-McLaughlin effect we mentioned how different sections of the stellar surface have a different radial velocity due to the star's rotation. This is just the main component. Other components include oscillations around the convective envelope from mainly p-modes at the time scales of a few minutes, which produce amplitudes of up to the order 100 m/s. Luckily these oscillations can be significantly removed by averaging in time with a longer exposure time ([Cegla et al. 2018](#)).

The other main effect is that of convective blueshift. Stars with convection in their surface present a corrugated surface with granules where hot and brighter plasma rises and are thus blueshifted towards us, while in the intergranular lanes where the cooler and less bright plasma descends being redshifted, the difference in flux causes an overall blueshift in the stellar disk.

Taking into account the angle between the normal of the stellar surface to our line of sight analyzed in [Cegla et al. \(2018\)](#) for a star with an effective temperature of 5800 K, near the stellar center the main contribution is the granules with plasma moving towards us creating the convective blueshift effect, while the other plasma flows move near perpendicular to our line of sight. Moving away from stellar center to the stellar limb we start to observe the contributions of the other plasma flows while the hotter rising plasma of

the granules flows diagonal to us. In addition, due to the corrugated surface the granules act like “mountains” and start to obscure each other while further uncovering the lanes between them (Cegla et al. 2018).

In Cegla et al. (2018) a study of the convective blueshift was made using magneto-hydrodynamic simulations of a Fe I 6302 Å line at a temperature of 5800 K. The results were then separated into 4 main components that of the granules, the non-magnetic and magnetic intergranular lanes, and magnetic bright points. The net radial velocity of the line leaving the center of the disk is a blueshift that increases as more convective flows are revealed until around 50 degrees. From there the blueshift decreases and eventually turns into a redshift at the limb where the granules start to point away from the observer which provide the biggest contribution to RVs.

Another general consequence of the corrugation and the line of sight effects is the widening of line profiles and decrease in their depth. This is due to the different plasma flows becoming visible and their different RVs.

Chapter 2

Data and Methodology

In this chapter we discuss the ESPRESSO data of the transit observations of HD 189733 Ab as well as the methodology that we employed. The final goal for the data reduction is to reconstruct the CCFs and spectral lines on the stellar surface along the planet's transit chord. This way we can analyze the variations of the CCFs or spectral line profiles of the corresponding area segments of the star's surface. Since the CCF and spectral line reconstruction are similar we will discuss them in succession, additionally we will showcase the spectral data we analyzed from the ESPRESSO spectrograph and how we corrected for telluric absorption and prepared for the methodology discussed.

2.1 The Target: HD 189733

But first, we justify the choice of HD 189733 as a target.

HD 189733 or V452 Vulpeculae is a binary star system composed of a class K2V ([Paredes et al. 2021](#)) primary star (HD 189733 A) and a smaller companion of class M (HD 189733 AB), so it is important to point out that the actual target is HD 189733 A. In addition, this binary system has a period of 3200 years with an orbit almost perpendicular to our line of sight ([Bakos et al. 2006](#)). This combined with the great distance of the two stars means it is unlikely for our observations to be contaminated by its smaller companion.

This star is 64.5 light years away from the Earth with a radius of $0.766^{+0.007}_{-0.013}$ solar radii ([Triaud et al. 2009](#)), and has an exoplanet known as HD 189733 Ab which was originally detected in 2005 using the radial velocity method by [Bouchy et al. \(2005\)](#). The detection in 2005 was also the third time the Rossiter-McLaughlin effect was ever measured and the first time a transit was inferred before a photometric observation.

HD 189733 Ab is a famous example of a Hot Jupiter, with a radius of $0.1602^{+0.0039}_{-0.0035}$ stellar radii (Cristo, E. et al. 2024) (around 1.22 Jupiter radii) and an orbital period of $2.21857312^{+0.00000036}_{-0.00000076}$ days (Triaud et al. 2009). The planet also presents a low eccentricity (Table 2.2) allowing us to simplify its orbit as being a circle. This planet has been used in multiple studies ranging from measuring its visible color (Evans et al. 2013) to creating a surface heat map and determining its average wind speed (Knutson et al. 2007).

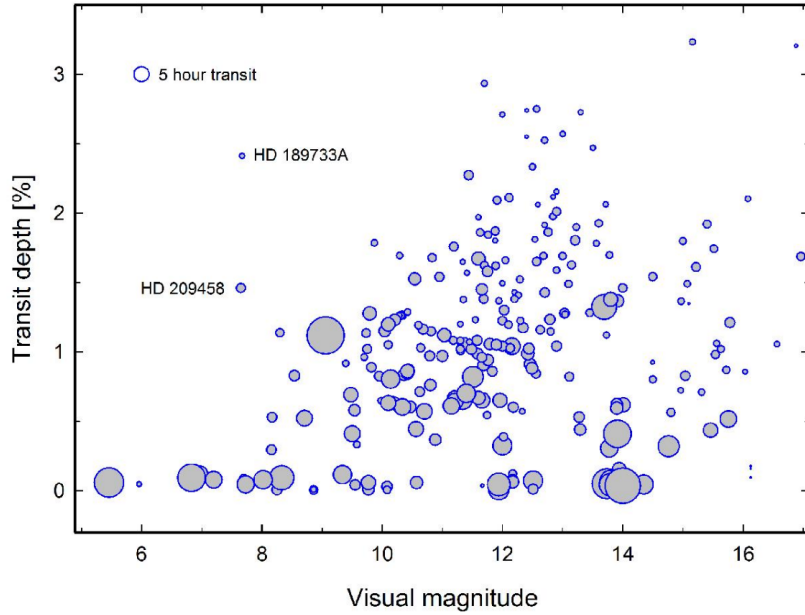


FIGURE 2.1: This plot taken from Dravins et al. (2017b) contains the transit depth vs visual magnitude of the exoplanets from the Exoplanet Orbit Database (2017) (Han et al. 2014). The diameter of the dots is proportional to the transit duration.

One notable feature of this particular star is the possession of one of the largest planetary transit depths of any known exoplanet, while having a high visual magnitude as seen in Figure 2.1. This makes it the best candidate to retrieve the spectra of the stellar surface obscured by the planet as was done in Cegla et al. (2016) and Dravins et al. (2018). This was the reason for choosing HD 189733 as our target as it can provide the strongest signal to noise ratio in our data and allow us to directly compare our results with the works of others and the improvements provided by ESPRESSO spectrograph, although this comes at the cost of lower resolution as we probe larger areas of the stellar disk.

2.2 The Dataset

2.2.1 The ESPRESSO Spectrograph

ESPRESSO is a very high-resolution Spectrograph installed on ESO's Very Large Telescope (VLT) providing ultra-high resolution spectral measurements for the aim of helping exoplanetary research ([Pepe et al. 2021](#)).

ESPRESSO is a highly-stabilized fiber-fed echelle spectrograph that can receive light from 1 to 4 telescopes of the VLT. The instrument is installed at the incoherent combined Coudé focus (ICCF) of the VLT. It is fed by two fibers one for the scientific target and other for simultaneous reference or simultaneous observation covering wavelengths from 378.2 nm up to 788.7 nm.

ESPRESSO also supports 3 main instrumental configurations those being High Resolution 1-UT (HR), Ultra High-Resolution 1-UT (UHR) and Medium Resolution 4-UT (MR). The HR configuration presents a resolving power of 140 000 and maximum RV precision of 10 cm/s using one unit telescope, UHR provides the highest resolving power at 190 000, but a maximum RV precision of 5 m/s and MR collects the most light using all 4 telescopes at the cost of lower resolving power of 70 000 and maximum radial velocity precision of 5 m/s.

2.2.2 The Observations

ESPRESSO was configured to the HR21 observation mode where it has a velocity precision of up to 10 cm/s and a resolving power of 140 000. The HR21 observation mode was chosen as it provides the highest radial velocity precision being ideal for the Doppler shadow method, the "21" refers to how the mode handles sampling of the observed photons.

The data used was collected over two separate nights of observation using the UT1 telescope at the VLT. The first was on the 11/08/2021 between 00:45 and 04:45 UTC where 41 observations were made, and the second on the 31/08/2021 between 00:27 and 04:40 UTC where 43 observations were made with an exposure time of 300 seconds. For each time interval a transit of HD 189733 Ab was observed. These same observations were first used in [Cristo, E. et al. \(2024\)](#) to study the broadband transmission spectrum of HD 189733 Ab.

The spectral data was stored in the form of .fits files containing data organized in "wavelength", "wavelength.air", "flux", "error" and "quality", which characterized the spectra. For our work we used the wavelength in the air, the flux and the error as our uncertainties for the spectra. The .fits files also included important data like time of observation, duration of observation, signal to noise ratio for all echelle orders among others in the header.

An overview of the observations relative to the planet's orbital phases can be seen in Figure 2.2, where we gather that for each of the two days there are 18 observations made during the transit.

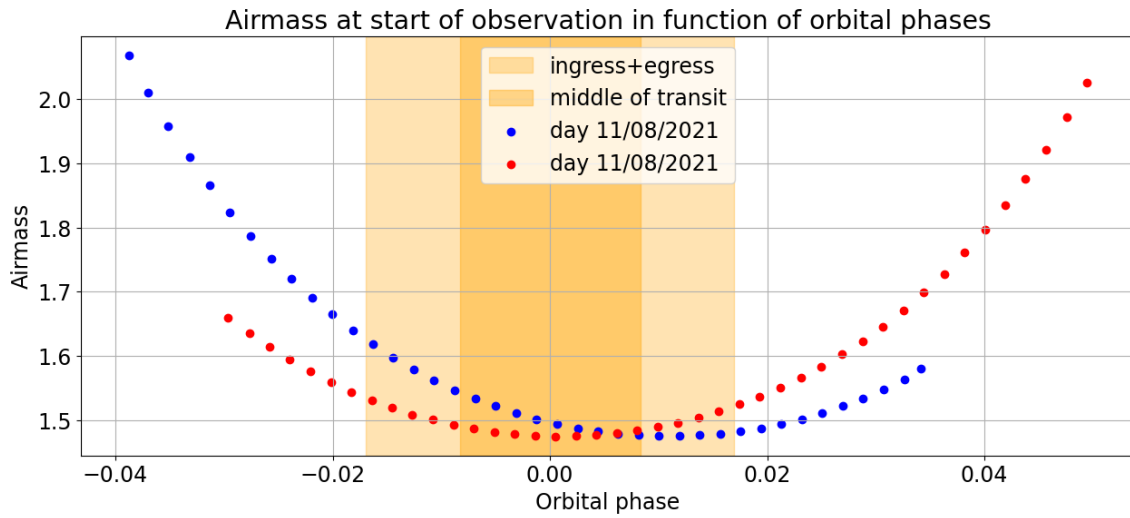


FIGURE 2.2: Airmass of the two sets of observations in function of the orbital phases and with the interval of transit. In blue are the observations on the 11/08/2021 and red for the 31/08/2021 with the shaded area representing the extent of the transit, and the darker region excludes the ingress and egress of the transit.

It is also important to note that on the first night, the signal-to-noise ratio (S/N) dropped significantly in the final observation, leading us to discard this data point. The signal to noise ratio of all the observations for two echelle orders can be seen in Figure 2.3.

The calibration was performed with the data reduction pipeline (DRS) 2.3.1. This deals with problems like calibrating the exposures to their respective wavelengths and performing actions like the heliocentric correction. The ESPRESSO pipeline was used to calculate the Cross Correlation functions of the spectra using a K2 mask also used in [Cristo, E. et al. \(2024\)](#).

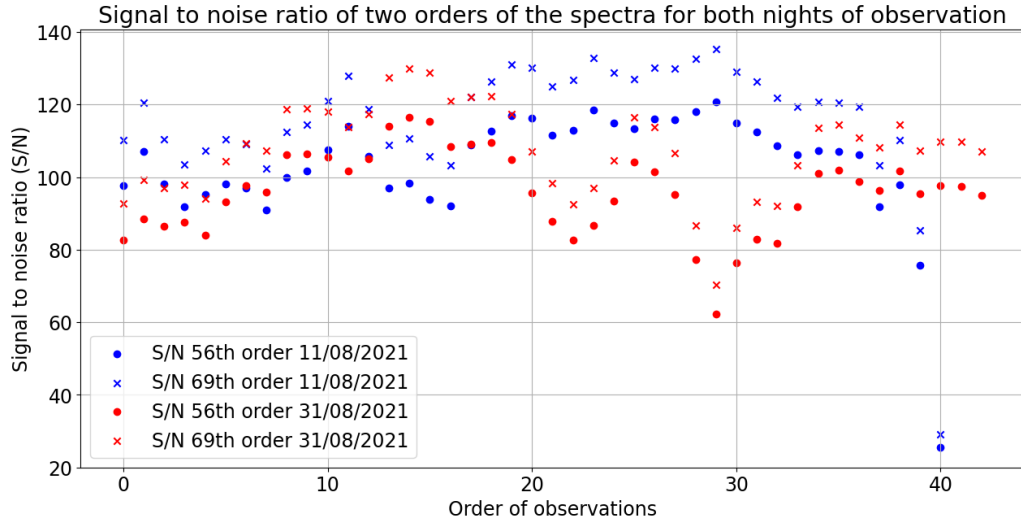


FIGURE 2.3: Signal to noise ratios for two orders of the ESPRESSO spectrograph those being the 56th (dots) and 69th (crosses) echelle orders of the spectrum which correspond to the regions around the sodium doublet and $H\alpha$ spectral lines. In blue is the data for the observations on the 11/08/2021 and red for the day 31/08/2021.

2.3 Spectral data reduction

Before applying the methodology that we will discuss next, we have to prepare the data into a more useful form. The objective of this work is to analyze both the CCFs and the individual spectral lines of the star behind the planet. This involves not only constructing the CCFs, but also normalizing the spectral lines themselves. Additionally, the atmosphere of the Earth introduces absorption and emission signatures, known as telluric contamination, into the observations, which need to be removed.

2.3.1 Telluric line correction with Molecfit

One of the big obstacles for the reconstruction of the local stellar spectra are the absorption lines from Earth's atmosphere primarily oxygen and water vapor. These telluric lines make it difficult to fit a function to the stellar line profiles, increase errors and clutter some of our plots.

To remove the majority of these lines from the regions of interest we decided to use the ESO program Molecfit ([Smette et al. 2015](#); [Kausch et al. 2015](#)) designed for this purpose. Molecfit is a tool to remove absorption features from a spectrum by fitting synthetic Earth transmission spectra with real data. This program reads automatically the keys in the .fits files of the spectra from specific instruments such as ESPRESSO allowing for simpler processing of the files generated by such instruments.

Parameter Key	Value	Description
LIST_MOLEC	H2O, O2	List of molecules to be included in the model.
WAVE_INCLUDE	0.6890, 0.6900, 0.7160, 0.7340, 0.7590, 0.7770	Wavelength ranges to be included.
WLG_TO_MICRON	0.0001	Multiplicative factor applied to the wavelength to express it in micrometres.
WAVELENGTH_FRAME	VAC_RV	Wavelength reference frame.
FTOL	1e-8	Relative chi-square convergence criterion.
XTOL	1e-8	Relative parameter convergence criterion.
FIT_CONTINUUM	1	Flag to enable/disable the polynomial fit of the continuum.
CONTINUUM_N	4	Degree of the polynomial continuum fit to use per range.
FIT_RES_GAUSS	True	Fit resolution by Gaussian.
RES_GAUSS	3.5	Initial value for FWHM of the Gaussian in pixels, at the centre of the spectrum.

TABLE 2.1: Table with the settings used for the telluric correction taken from [Azevedo Silva et al. \(2022\)](#).

When adjusting the settings for the telluric correction we used the ones used in [Azevedo Silva et al. \(2022\)](#), which can be seen in table 2.1. These values were used originally to clean data of WASP-76 b and WASP-121 b, which was also taken from ESPRESSO and ended up performing just as well for our data. When executing the program, we used the same settings configuration for all observation files. Figures 2.4 and 2.5 show a comparison of the spectra before and after telluric correction.

In Figure 2.4 we see the difference of the same section of the spectra for the night of observations on the 11/08/2021 (the same happens for the data of 31/08/2021) before and after correction. For the uncorrected data we see lines whose depth varies over the course of the transit as they are color marked according to their orbital phase. This is a typical behavior of Earth’s atmospheric lines where the amount of absorption is dependent, among other aspects, on the airmass along our line of sight which varies over the course of the night of observations, which lasted for about 4 hours (Figure 2.2). Similar behavior is also documented in Figure C.1 of [Azevedo Silva et al. \(2022\)](#). In contrast the corrected spectra have the telluric lines essentially removed resulting in spectra whose lines are independent of changes in airmass.

Another important detail is that of over-fitting, which is seen in Figure 2.5 where some telluric lines are saturated due to strong absorption primarily the oxygen A band and B band around 760 nm and 690 nm respectively. These regions are of course ignored when Molecfit does fitting of the selected regions containing non saturated O_2 and H_2O lines.

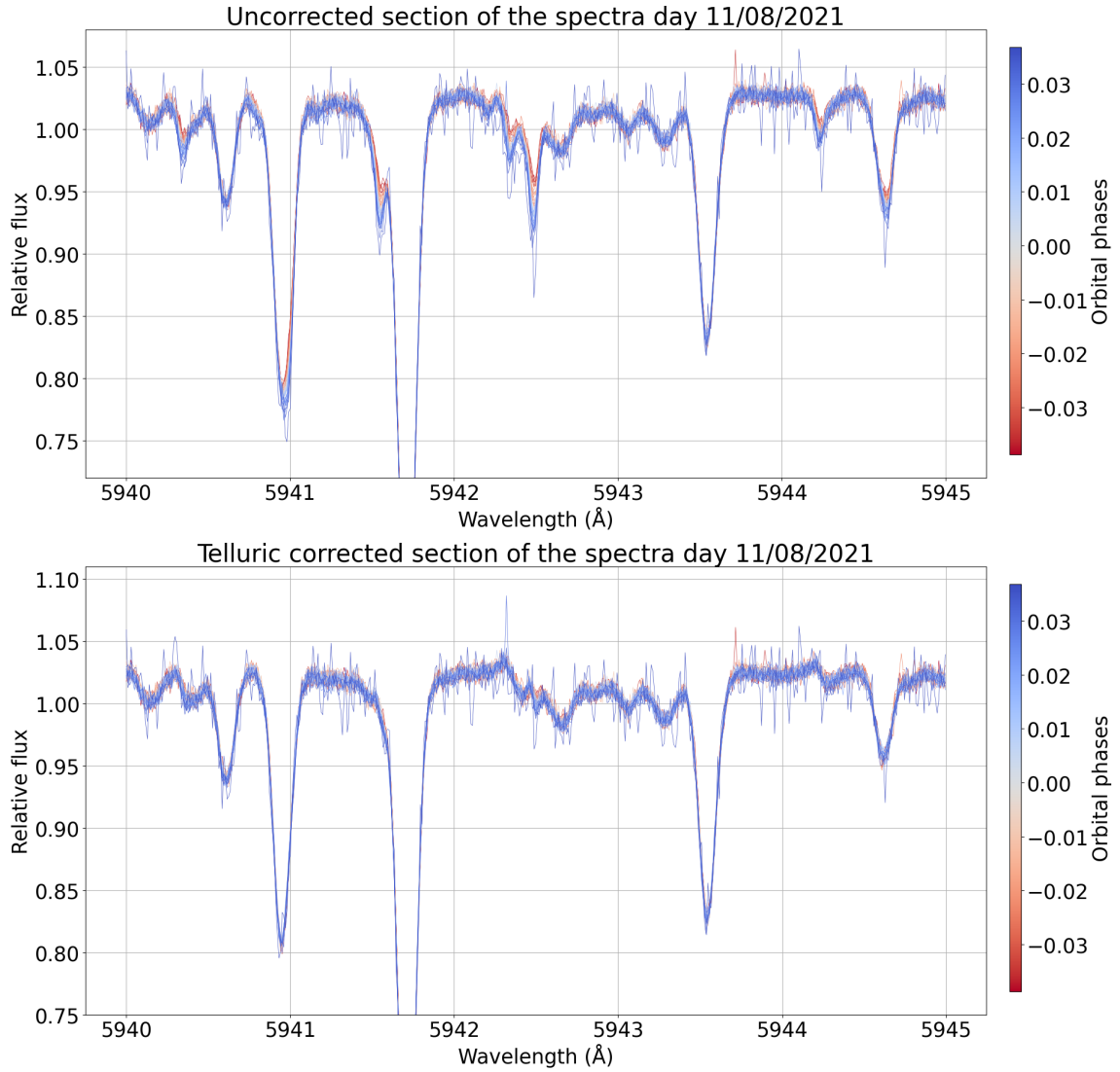


FIGURE 2.4: Top: locally Normalized spectra of the observations for the day 11/08/2021 between 594 nm and 594.5 nm showcasing telluric lines from Earth's atmospheres which vary over the course of the observations, the orbital phases of the observations are color coded according to the colorbar on the right. Bottom: same description as top, but the spectra were corrected with Molecfit removing the telluric lines.

This however results in an over correction rendering the saturated part of the spectrum appear as extreme peaks, which will be masked out for the rest of our analysis.

2.3.2 Normalization of the spectra

Normalization of the spectra is an important part for the reconstruction of individual line profiles as it allows for easy comparison of the chosen lines.

During a planetary transit the planet blocks parts of the stellar surface of different flux and over the course of the observation night the airmass changes. Because of this there

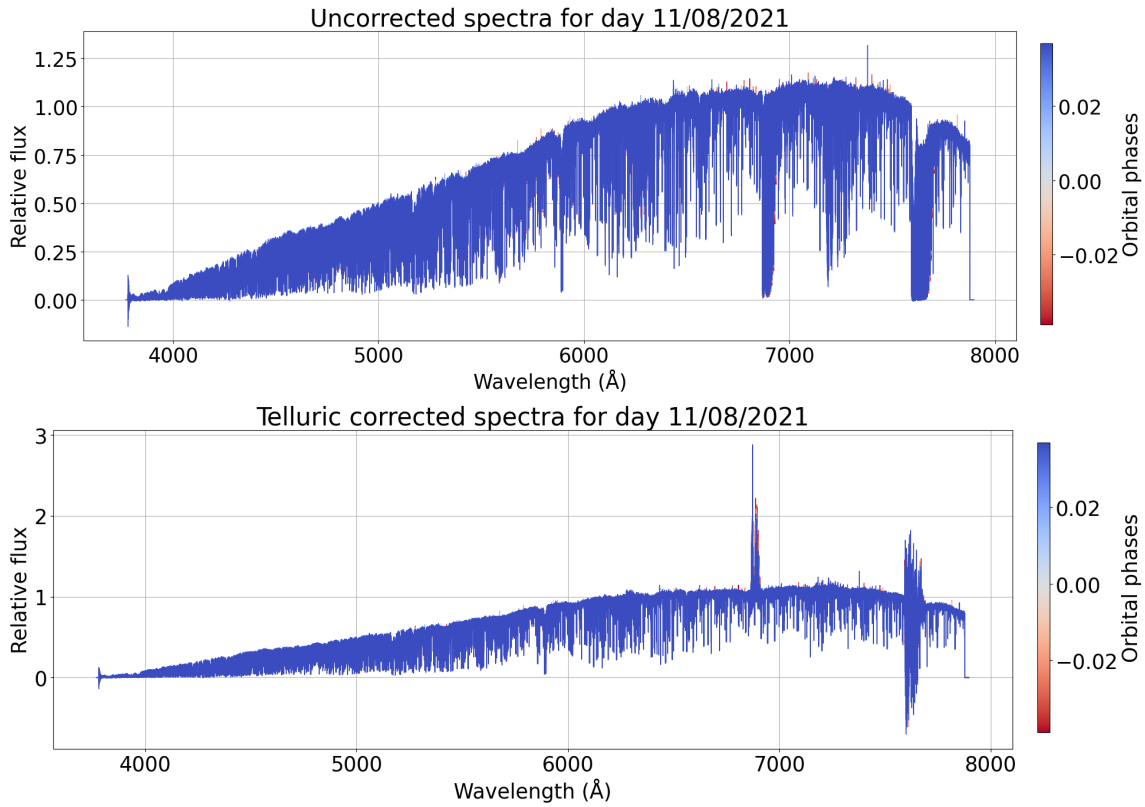


FIGURE 2.5: Top: globally normalized spectra of the observations for the day 11/08/2021 color coded according to the orbital phases of the observations according to the colorbar on the right. Bottom: same description as above, but the spectra were corrected using Molecfit removing the telluric lines, over-fitting of the saturated telluric lines is also visible. The multiple spectra are overlaid on top of each other so the changes are subtle, but noticeable.

is a great variation of the detected flux. This means that the observed spectra possesses very different continuum values over the multiple observations.

Since we only care about the relative change of the flux in the spectra, we do not need to know the absolute value of the continuum, instead we will first normalize every spectrum. For the case of the CCFs we will first perform the cross correlation and then normalize in a different way (section 2.4.1). In a later step we will restore the relative changes between observations using a simulated photometric transit. This is done because the observed flux differences are not representative of changes in the stellar disk but affected by other factors such as change in airmass during observations.

At first a global normalization was attempted by dividing each spectrum by the average value of a relatively flat region with few spectral lines spanning between 640 nm and 680 nm. However, this proved insufficient given that the spectra presented a change in flux continuum over the course of the two observation nights. This is shown in Figure 2.6,

where we see that the region chosen for global normalization does not vary over the multiple observations as expected, but in contrast for smaller wavelengths we see an increase in flux along the orbital phases, with the opposite behavior existing for larger wavelengths revealing a variation in slope of the flux continuum. Because of this, for individual line analysis we will supplement global normalization with local normalization.

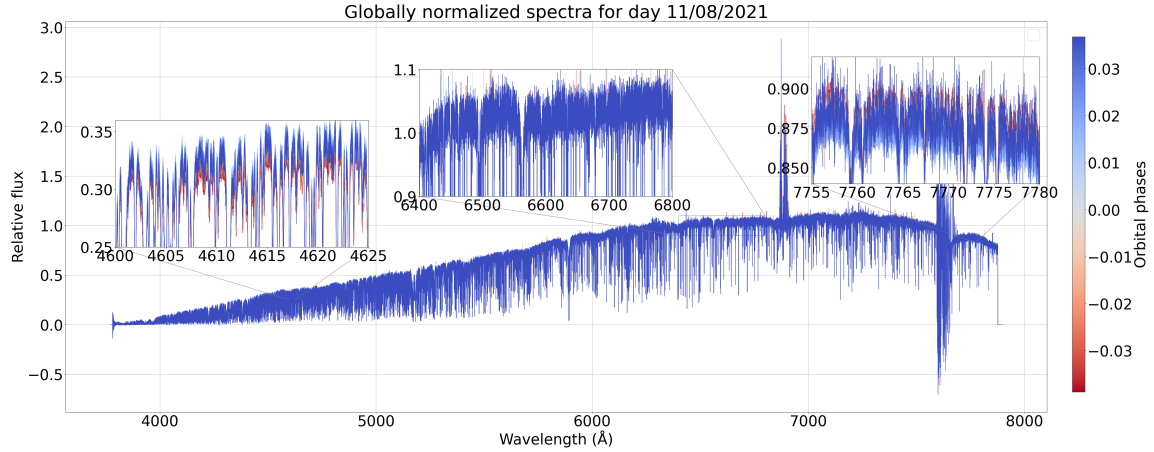


FIGURE 2.6: Spectra with zoomed in sections, in the middle we see the area whose mean was used for global normalization while on the left and right zoomed in sections we see a dependence on the orbital phases showing that this type of normalization is insufficient.

For local normalization two small intervals on both sides of the chosen spectral line were chosen corresponding to the continuum in the zone of the spectral line. A linear fit of these regions was then done and used to normalize the area of the specified spectral line.

In this work we analyzed the $H\alpha$ line and the sodium doublet lines. In the first case the specific intervals for the continuum near the $H\alpha$ line did not have a significant impact on the normalization, being relatively simple. However, for the sodium doublet we spotted an anomaly on the continuum on the bluer side of the lines where the continuum seems to change over the course of the transit. This anomaly forced us to choose a specific continuum region right next to the line for the fit. We are unsure of the cause of such behavior, but given that it is present around the interval of 586 nm to 588 nm and varies with time we suspect to be caused by Chappuis ozone absorption from the Earth's atmosphere ([Brion et al. 1998](#)).

The local normalization of the sodium doublet lines was done simultaneously as their wings partially overlap (Figure 2.7). The selected regions were between [588.20, 588.30] nm and [590.82, 590.92] nm and they were chosen carefully due to the anomaly in the continuum on the smaller wavelengths.

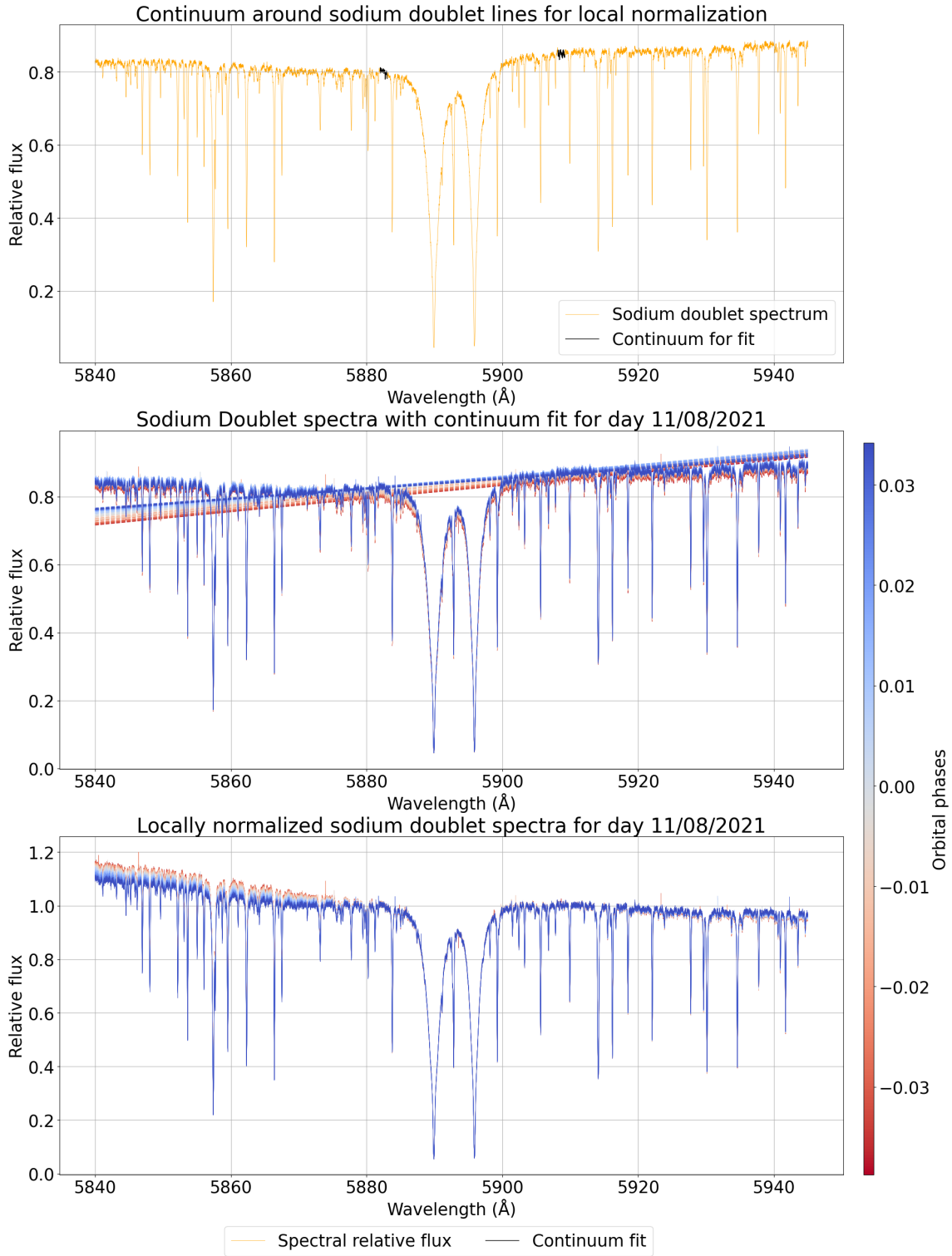


FIGURE 2.7: Local normalization of the sodium doublet lines. At the top we see an example of the sodium doublet after global normalization and before local normalization in yellow with the black region being the continuum from the selected wavelength interval for line fitting. The central plot we see all spectra for the day 11/08/2021 (the same happens for 31/08/2021) colored based on their orbital phases according to the colorbar on the right. Overlaid on top are dashed lines of the linear regressions using the fitting parameters from the fit of the selected wavelength intervals, with the same color as their respective spectra. The bottom plot shows the spectra after local normalization using each respective linear fit.

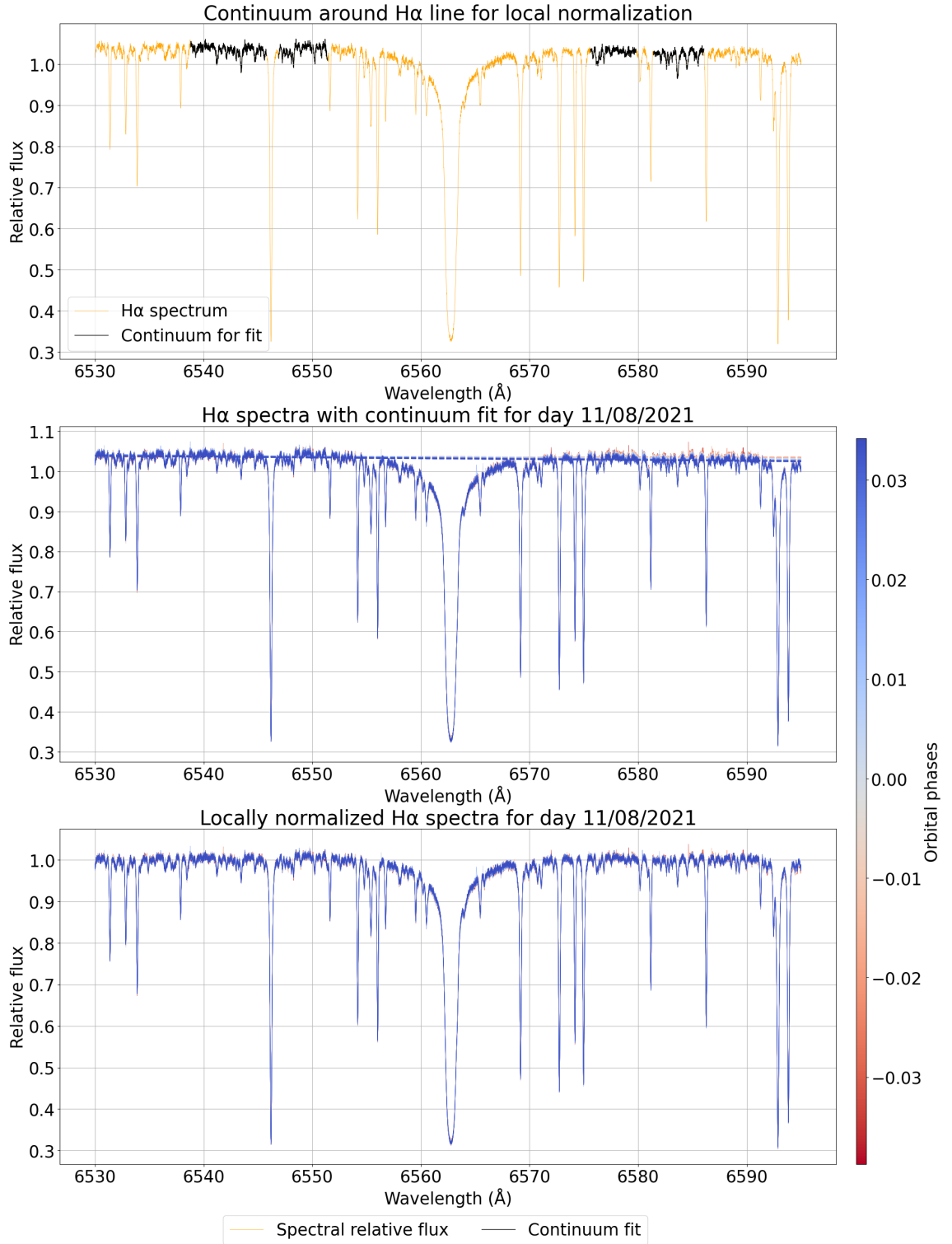


FIGURE 2.8: Local normalization of the H α line. At the top we see an example of the sodium doublet after global normalization and before local normalization in yellow with the black region being the continuum from the selected wavelength interval for line fitting. The central plot we see all spectra for the day 11/08/2021 (the same happens for 31/08/2021) colored based on their orbital phases according to the colorbar on the right. Overlaid on top are dashed lines of the lines using the fitting parameters from the fit of the selected wavelength intervals, with the same color as their respective spectra. The bottom plot shows the spectra after local normalization using each respective linear fit.

For the $H\alpha$ line its local normalization is shown in Figure 2.8 and the wavelength interval ranges did not seem to affect the result significantly since we did not have the same anomaly in the sodium. The selected continuum intervals were in this case [653.880, 654.580] nm, [654.69, 655.140] nm, [657.56, 657.980] nm and [658.140, 658.605] nm the extra gaps avoid catching other spectral lines.

2.3.3 Cross Correlation of the spectra

Cross correlation functions provide the ability to increase the signal to noise ratio of a spectrum at the cost of losing the ability to distinguish between individual lines as they essentially “average” the behavior of multiple chosen lines. CCFs allow us to measure the overall behavior of the combined lines, in addition to being able to retrieve the radial velocity of the source since a spectrum is typically Doppler shifted in relation to the mask.

CCFs are constructed by cross correlating a spectrum with a chosen mask resulting in a flux vs radial velocity function. A common type of mask is a stellar mask that retrieves the overall behavior of the entire spectrum of a star and is made with the stellar type in mind. These masks are typically a theoretical model of the main features of a specific stellar spectrum if it were at rest. Another common type of mask is a binary mask whose flux values consists entirely of either 1 or 0 with 0 being the intervals of wavelength we are not interested in, while 1 is for areas of interest. These masks are normally used when we want to study sets of specific spectral lines. In this case the mask takes the form of a square function where each of the squares are centered to the wavelength of a specific spectral line at rest and their width is given by a chosen radial velocity which is then converted into wavelength.

Our target HD 189733 A is classified as a type K2V star so a K2 mask was used to create “white light” CCFs that allows to extract the star’s RVs whose main components is the systemic velocity, Keplerian velocity the RM effect and possible stellar activity. These CCFs were made by the ESPRESSO pipeline with the K2 mask being provided by ESPRESSO.

In the context of this work we are also interested in the study of more specific groups of lines. So we decided to use the same selected lines by [Dravins et al. \(2018\)](#), with tables E.1 and E.2 having a set for weak Fe I lines and strong Fe I lines respectively. These lines were chosen as they are numerous and form at very high temperatures preventing contamination from the exoplanet’s atmosphere, so we can consider the exoplanet as being opaque ([Dravins et al. 2018](#)). Recreating the iron CCFs allows for a direct comparison of

our results regarding how the profiles of the iron CCFs vary from star center to limb. As opposed to the white light CCFs made with a K2 mask, that were provided to us.

The masks of the weak and strong iron lines CCFs were created using a simple binary mask in the form of a square function. The center of the square having the central wavelengths of the iron lines (Tables E.1 and E.2) with the width of squares given by a radial velocity of 500 m/s due to the resolution of the ESPRESSO spectrograph.

To create the CCFs we used the iCCF python package (Queloz et al. 2001; Nardetto et al. 2006; Léger et al. 2011; Figueira et al. 2013; Santerne et al. 2015; Lanza et al. 2018). The basic formula used in the source code of iCCF for calculating numerically the cross correlation with our discrete data for the spectra in the form of arrays of numbers is of the type:

$$CCF(rv(i)) = \sum_j S(\lambda(j)) \times M\left(\lambda(j) \times \left(1 + \frac{rv(i)}{c}\right)\right) \quad (2.1)$$

Here, $CCF(rv(i))$ is the flux of a CCF point indexed by i with a given radial velocity, $S(\lambda(j))$ is the flux of the spectrum and its wavelength at a point with index j . $M(\lambda(j))$ is the flux and wavelength of the mask at a point with index j , with $rv(i)$ being the specific radial velocity at that point of the CCF, used to shift the mask by the appropriate amount to do the cross correlation.

At the end we got two sets of white light CCFs, which were originally provided by the ESPRESSO pipeline and also used in Cristo, E. et al. (2024). Plus, two sets of weak Fe I CCFs, and two sets of strong Fe I CCFs for the two nights of observations.

2.4 Extraction of local stellar spectra

With the first steps of the local normalization of the spectra for individual line analysis and cross correlation for CCF analysis done. We now present the main methodology to extract the subtracted lines/CCFs or the local lines/CCFs as they embody the local information of a portion of the stellar surface.

Figure 2.9 illustrates the methodology in order to get to the local lines/CCFs from which we can obtain our results. The Doppler shadow method was discussed in section 1.3, but in this section we will explain the multiple steps required to manipulate the data. Starting by the CCFs analysis and followed by individual line analysis.

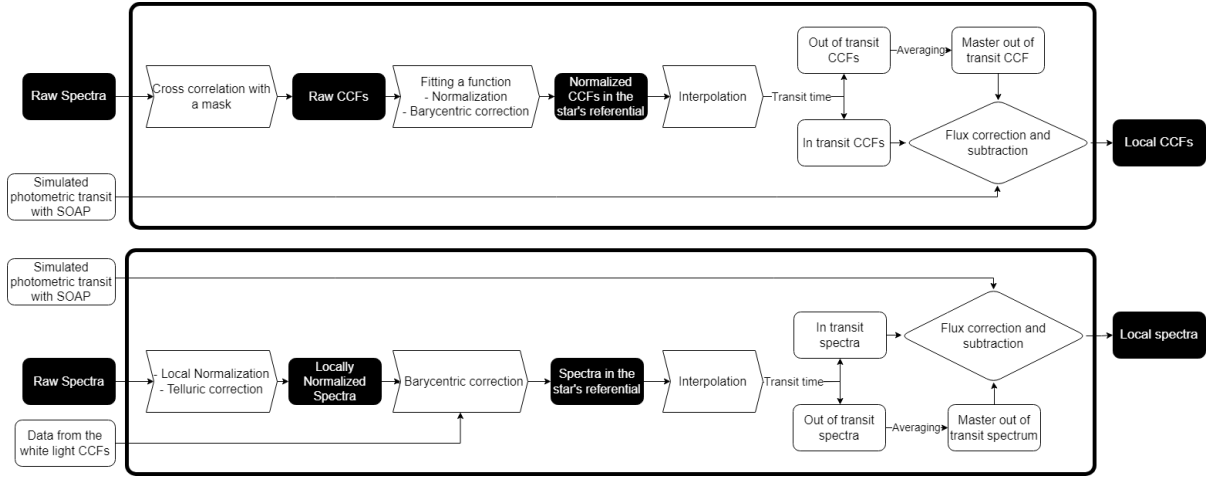


FIGURE 2.9: Flowchart containing the basic steps in our methodology to get from the starting data to the subtracted CCFs or spectral lines. The top represents the process for the CCFs while the bottom is the process for the individual line analysis. The rounded boxes represent the data at its multiple steps, arrow boxes represent the main processes, the diamond represent the final process which requires multiple sets of data. The black color shows the data between the multiple stages.

2.4.1 CCF analysis

One of our objectives is to retrieve the CCFs of areas of the stellar surface covered by the transiting planet. This requires the processing of multiple CCFs from over the course of a transit. The following methodology is based on the Doppler Shadow method.

One thing to mention is that in the case of the iron CCFs, since we already performed the telluric correction for the analysis of individual lines, we decided to use the corrected spectra to perform the cross correlation. Although the same is not the case for the white light CCF, since it was provided by the ESPRESSO pipeline.

We start by taking our CCFs and fit them with a profile function, with the goal of deriving its central and continuum values in order to measure the star's radial velocity and to normalize the CCFs to their continuum. This is due to the total flux varying significantly during the course of the multiple observations (this was discussed further for the raw spectra in subsection 2.3.2).

The function we choose for the fitting is the same used in Dravins et al. (2017a), Dravins et al. (2017b) and Dravins et al. (2018). The function is a modified Gaussian function of the form:

$$f(x) = y_0 + a \exp \left(-\frac{1}{2} \left(\frac{|x - x_0|}{b} \right)^c \right) \quad (2.2)$$

where y_0 is the value of the continuum, x_0 is the central radial velocity of the CCF, a is a scaling parameter, b is the equivalent to the variance of a regular Gaussian function which changes the width and c is a variable exponent. The function does not have a simple formula to compute the FWHM and c adjusts how "steep" the curvature of the function is at its peak. This function is used to provide a direct comparison with the work of [Dravins et al. \(2018\)](#).

For the fit we used from the python package `scipy` the function `scipy.curve_fit` ([Virtanen et al. 2020](#); [Vugrin et al. 2007](#)) as it allows for the input of uncertainties and outputs a covariance matrix. The `scipy` function uses the non-linear least squares method to fit a function. The covariance matrix allows to take the square root of the diagonals used to get the standard deviation of the fitting parameters that we used as uncertainties.

With the raw CCFs continuum values known we normalize the CCFs by dividing by y_0 and propagate their uncertainties.

We then separate the observations into in-transit and out-of-transit observations using the formula 1.3 as a criterion to separate the time of observations. More specifically we convert formula 1.3 to the form:

$$\phi_{dur} = \frac{1}{\pi} \arcsin \left(\frac{\sqrt{\left(1 + \frac{R_{planet}}{R_{star}}\right)^2 - \left(\frac{a}{R_{star}} \cos(i)\right)^2}}{\frac{a}{R_{star}}} \right) \quad (2.3)$$

where ϕ_{dur} was the duration of the transit (T) converted into orbital phases. Having the multiple lengths in units of the stellar radius allows for the use of the values in table 2.2 which were taken from table 4 and E.1 of [Cristo, E. et al. \(2024\)](#).

Additionally, we convert our times of observation (t) into their respective dimensionless orbital phases using equation 1.5. The observations for phases larger in absolute value to $\phi_{dur}/2$ will be outside the transit and the rest will be inside the transit.

Converting times of observation into orbital phases requires knowing the instant where the planet is at the center of its transit (t_{epoch}), which is derived from T_0 present in table 2.2. Given that ESPRESSO records time of observation in the header keyword "HIERARCH ESO QC BJD" using the BJD (Barycentric Julian Date) format we have to add a conversion factor of $0.5 + 2.4 \times 10^6$ to account for the MBJD (Modified Barycentric Julian Date) format of T_0 and finally we add the mid-transit phase shift (converted into days by multiplying

by the orbital period) to account for small changes in the planet’s orbit since 2009. The values $\Delta\phi_{p,0}$ and $\Delta\phi_{p,1}$ refer respectively to the days of the observations on the 11/08/2021 and 31/08/2021.

Parameter	Value	Source
$P_{rot}(days)$	$11.454^{+0.092}_{-0.088}$	Cristo, E. et al. (2024)
$i_{star}(^\circ)$	$71.87^{+5.55}_{-6.91}$	Cristo, E. et al. (2024)
$i_{planet}(^\circ)$	$85.465^{+0.020}_{-0.021}$	Cristo, E. et al. (2024)
$\lambda(^\circ)$	$-1.00^{+0.22}_{-0.23}$	Cristo, E. et al. (2024)
a/R_{star}	$8.7686^{+0.0082}_{-0.0082}$	Cristo, E. et al. (2024)
R_{planet}/R_{star}	$0.1602^{+0.0039}_{-0.0035}$	Cristo, E. et al. (2024)
$\Delta\phi_{p,0}$	$-0.002424^{+0.00036}_{-0.00035}$	Cristo, E. et al. (2024)
$\Delta\phi_{p,1}$	$-0.002300^{+0.00035}_{-0.00044}$	Cristo, E. et al. (2024)
$R_{star}(R_\odot)$	$0.766^{+0.007}_{-0.013}$	Triaud et al. (2009)
$T_{eff}(K)$	4969 ± 43	Sousa et al. (2018)
T_0 (MBJD)	$53988.30339^{+0.000072}_{-0.000039}$	Triaud et al. (2009)
P_{orbit} (days)	$2.21857312^{+0.00000036}_{-0.00000076}$	Triaud et al. (2009)
$[Fe/H]$ (dex)	-0.07 ± 0.02	Sousa et al. (2018)
$\log(g)$	4.60 ± 0.01	Sousa et al. (2018)
$\lambda_{sodiumD1}$ (nm)	588.995095 ± 0.000003	Kramida et al. (2023)
$\lambda_{sodiumD2}$ (nm)	589.592424 ± 0.000003	Kramida et al. (2023)
$\lambda_{H\alpha}$ (nm)	656.279 ± 0.003	Kramida et al. (2023)
e	$0.027^{+0.021}_{-0.018}$	Rosenthal et al. (2021)

TABLE 2.2: Table with the quantities used in this work compiled from multiple works.

For the barycentric correction we plot the central radial velocities (x_0 in the fitting function 2.2) over the course of their transit phases resulting in a plot like the one in Figure 2.10. The plot for the white light CCF is especially useful since it provides the smallest uncertainties and higher S/N of the star’s center of mass motion in space, which is visible for the out-of-transit observations in Figure 2.10. This allows for us to correct the data into the star’s referential (the heliocentric correction was already done by the ESPRESSO pipeline).

While the out-of-transit motion of the star follows a Keplerian curve (Eq. 1.8) at this small scale the motion is best approximated as being uniform. We then used a linear fit of the type $y = mx + b$ to subtract the star’s motion including any additional constant speed the star may have. This places our CCFs in the referential of HD 189733 A. This process of RV correction is shown in Figure 2.10 for the central radial velocities of the profiles, where out-of-transit the velocity is essentially to zero. With Figures 2.11 and 2.12 showing the aftermath for the other types of masks.

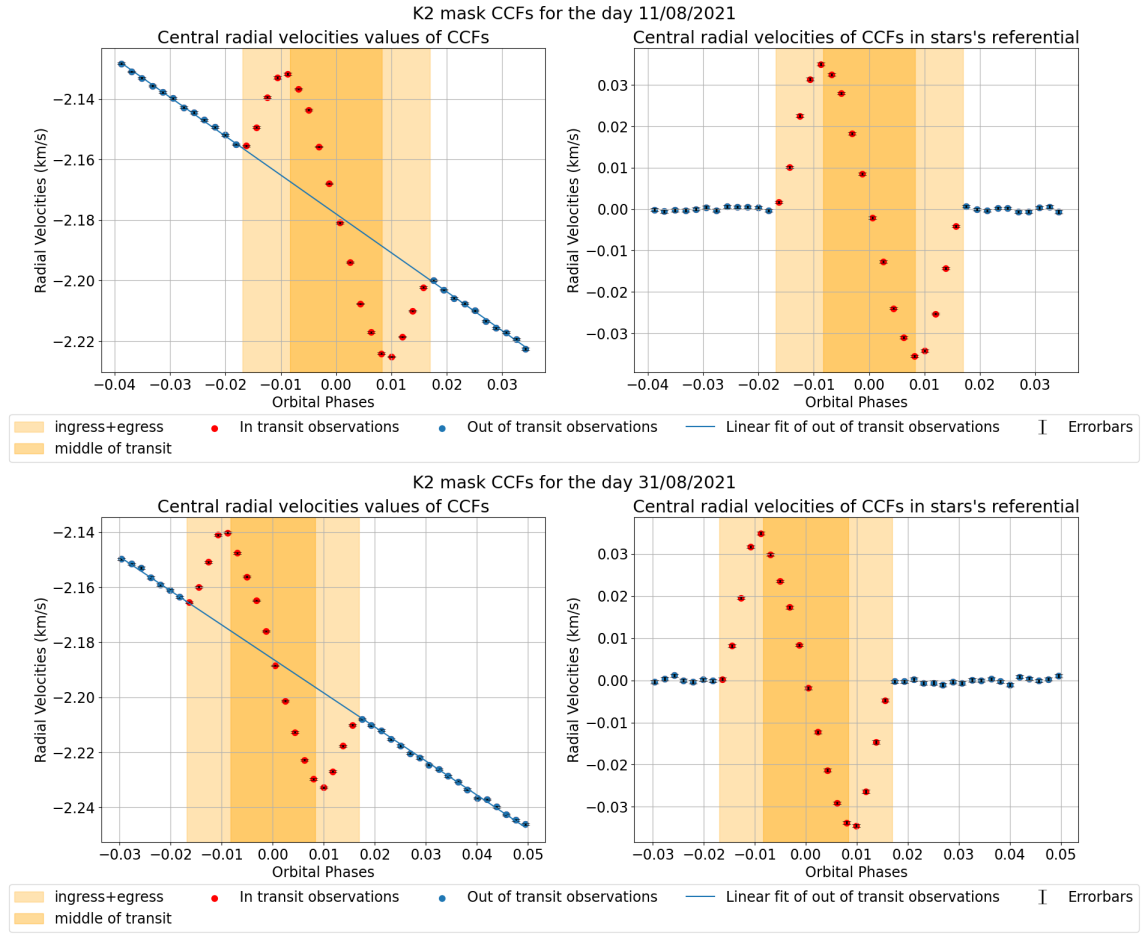


FIGURE 2.10: (Top day 11/08/2021, bottom day 31/08/2021) Left side: central radial velocities of the CCFs over their orbital phases in red is the values from the in-transit observations while in blue is the out-of-transit observations, with the blue line being a linear fit of them. The shaded area indicates the phases where the transit occurs with the darker region excluding the ingress and egress, error bars are shown colored in black. Right side: same as the left except all observations had the barycentric motion subtracted (blue line).

The linear fit as values $m = -1.290 \pm 0.004 \text{ km/s}$ and $m = -2.1780 \pm 0.0001 \text{ km/s}$.

We are left with CCFs on the referential of the star allowing for the better study of the stellar surface. The Rossiter-McLaughlin is also clearly seen as one would expect and its cause will help us to probe the stellar surface. For all of our CCFs we used a similar linear fit for each type of mask and for each day of observation. While for the individual line reconstruction we use the white light values for the barycentric correction since it was the least noisy of the 3 masks used.

However, the barycentric correction results in different abscissas for all the CCFs, so in order to be able to do operations with the CCFs we need to interpolate them. We decided to use linear interpolation as it is the simplest and enough for our purposes. When it came to propagate the uncertainties of the interpolation, it turned out to be a more complex

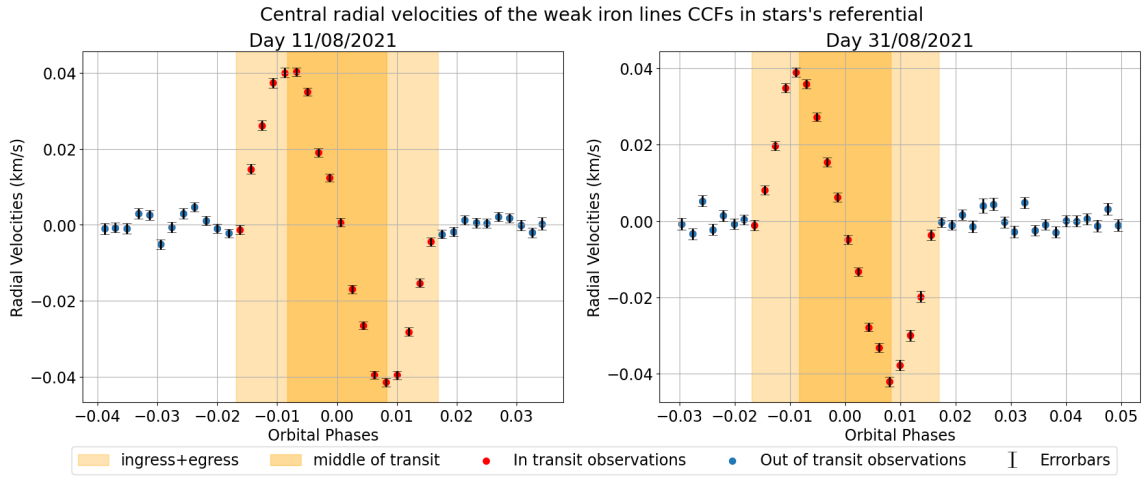


FIGURE 2.11: Barycentric corrected RV values of the weak iron CCFs for the two nights of observation 11/08/2021 left side and 31/08/2021 right side, with the same legend as in Figure 2.10. The linear fit as values $m = -1.29 \pm 0.02 \text{ km/s}$ and $m = -2.0853 \pm 0.0005 \text{ km/s}$ for day 11 and values $m = -1.23 \pm 0.02 \text{ km/s}$ and $m = -2.0952 \pm 0.0006 \text{ km/s}$ for day 31.

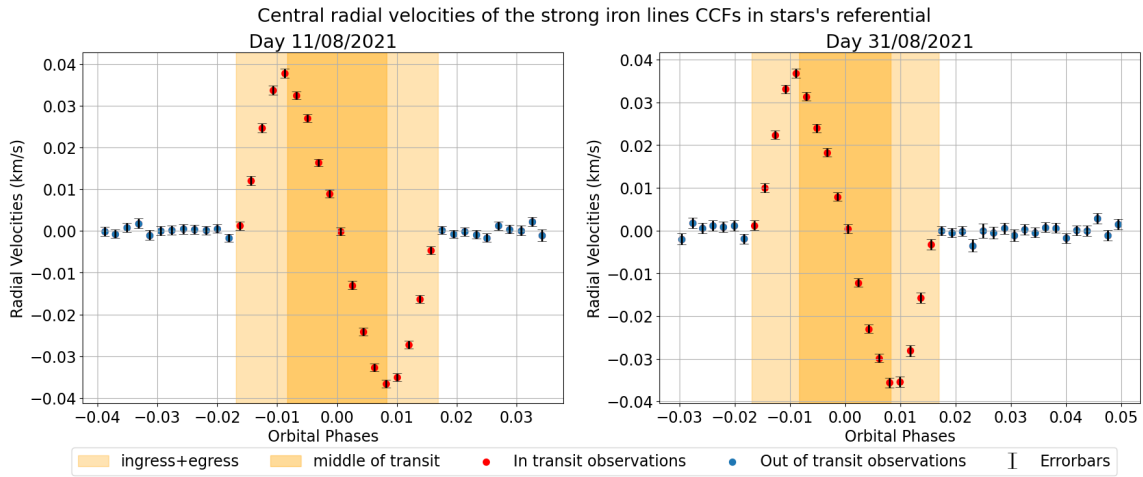


FIGURE 2.12: Barycentric corrected RV values of the strong iron CCFs for the two nights of observation 11/08/2021 left side and 31/08/2021 right side, with the same legend as in Figure 2.10. The linear fit as values $m = -1.291 \pm 0.008 \text{ km/s}$ and $m = -2.2314 \pm 0.0002 \text{ km/s}$ for day 11 and values $m = -1.22 \pm 0.01 \text{ km/s}$ and $m = -2.2399 \pm 0.0003 \text{ km/s}$ for day 31.

subject than at first realized and we leave the details of it in appendix A.

The next step is to compile all interpolated out-of-transit CCFs and average them out producing a Master out-of-transit CCF that represents the star as a whole and not affected by the transiting planet.

We compared the Master out-of-transit of the day 11/08/2021 with the one of day 31/08/2021 by dividing the latter by the former to see if there were any changes as there is a gap of 20 days between observations. This was done for the weak iron and strong iron

CCFs but it did not show any difference larger than 0.0025% for the weak iron lines CCF and 0.0045% for the strong iron lines CCF.

Now we have to correct the in-transit observations to their appropriate relative flux in comparison with the out-of-transit observations. This could be done by taking a real photometric transit of the HD 189733 Ab, but one would have to deal with units, noise and interpolation of the data. Instead we used the SOAP software (Boisse et al. 2012; Oshagh et al. 2013; Dumusque et al. 2014; Akinsanmi et al. 2019; Serrano et al. 2020; Zhao & Dumusque 2023), that is capable of simulating a stellar surface with star spots and planetary transits to simulate a transit light curve like the ones in Figure 2.13, that we used for flux correction. We then multiply each in-transit CCF by the corresponding flux as in equation 1.11. The master out-of-transit CCF would be multiplied by 1, which would be redundant.

The SOAP light curve simulation requires values for the planet's orbital period (P_{orbit}), eccentricity (e) which we took as 0, orbital inclination (i_{planet}), spin-orbit angle (λ), semi-major axis (a), and radius R_{planet} and for the star their rotational period (P_{rot}), stellar inclination (i_{star}), radius (R_{star}), and effective temperature (T_{eff}), which can all be found in Table 2.2. In addition the SOAP simulation also requires the limb darkening coefficients for its quadratic form function, for this we used the LDTK python package (Parviainen & Aigrain 2015; Husser et al. 2013) which calculates the coefficients from the star's effective temperature (T_{eff}), metallicity ($[Fe/H]$), and surface gravity ($\log(g)$) using a Markov chain, while we selected the range of wavelengths of the ESPRESSO spectrograph. Also for the case of individual lines we adjusted the wavelength intervals to the wavelengths of the specified lines, which is one advantage of using simulations over photometric data as it is harder to account for variations in limb darkening with wavelength.

Finally, to get our results we take the multiple in-transit CCFs and subtract them from the master out-of-transit CCF using:

$$F_{local} = f_{Master,out} - f_{in} \times F_{SOAP} \quad (2.4)$$

Where F_{local} is the flux of our local CCFs, $f_{Master,out}$ is the flux of the master out-of-transit, f_{in} are the normalized in transit CCFs and F_{SOAP} is the simulated photometric flux given by SOAP.

This results in our local CCFs from the respective areas of the stellar surface covered by the planet for each observation. An estimate of the CCFs S/N is in the order of 1000

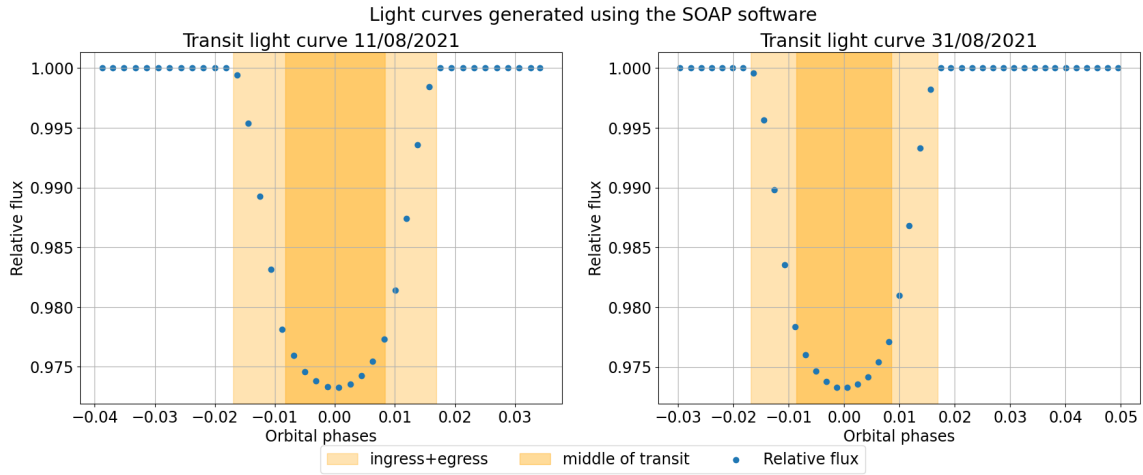


FIGURE 2.13: Simulated transit light curves for the two nights of observation, with relative flux with 1 meaning the total stellar flux of the star without transit, the shade area signals where the transit occurs while the darker region excludes the ingress and egress.

since our transit has a depth of almost 3%. This means that the residual CCFs will have a S/N of roughly $(1000 - 1000 \times (1 - 0.3) = 30)$, which is a much lower signal to noise ratio resulting in noisier data, which highlights the importance of very precise and high S/N observations for this method to work.

Figure 1.4 represents that by subtracting the out-of-transit stellar surface by the in-transit observations for all instances shows the path of transit, and results in the data for the successive areas of the star's surface being retrieved.

With the local CCFs we can then evaluate how the profiles change over the transit chord of the star's surface, one of the ways we did this is by fitting them with a function to gain information about several profile features. Again we used a modified Gaussian (Eq. 2.2) to get the multiple values for the fitting parameters, from which we retrieve mainly the central radial velocity, the line-center intensity in comparison to the continuum and the line-width as defined in Dravins et al. (2018). We also did this to the master out-of-transit CCF to get a reference value for the integrated spectrum of the whole stellar surface to compare with our local CCFs that represent smaller areas.

The local white light CCFs are displayed in Figure 2.14, additionally the local weak iron lines CCFs and strong iron lines CCFs are shown in Figures 2.16 and 2.18 respectively.

Figure 2.15 contains tomographies of the subtracted CCFs, here it becomes clearer compared with Figure 2.14, how the subtracted CCFs profiles vary during the transit as already discussed, within the horizontal black lines, outside of which only show residual

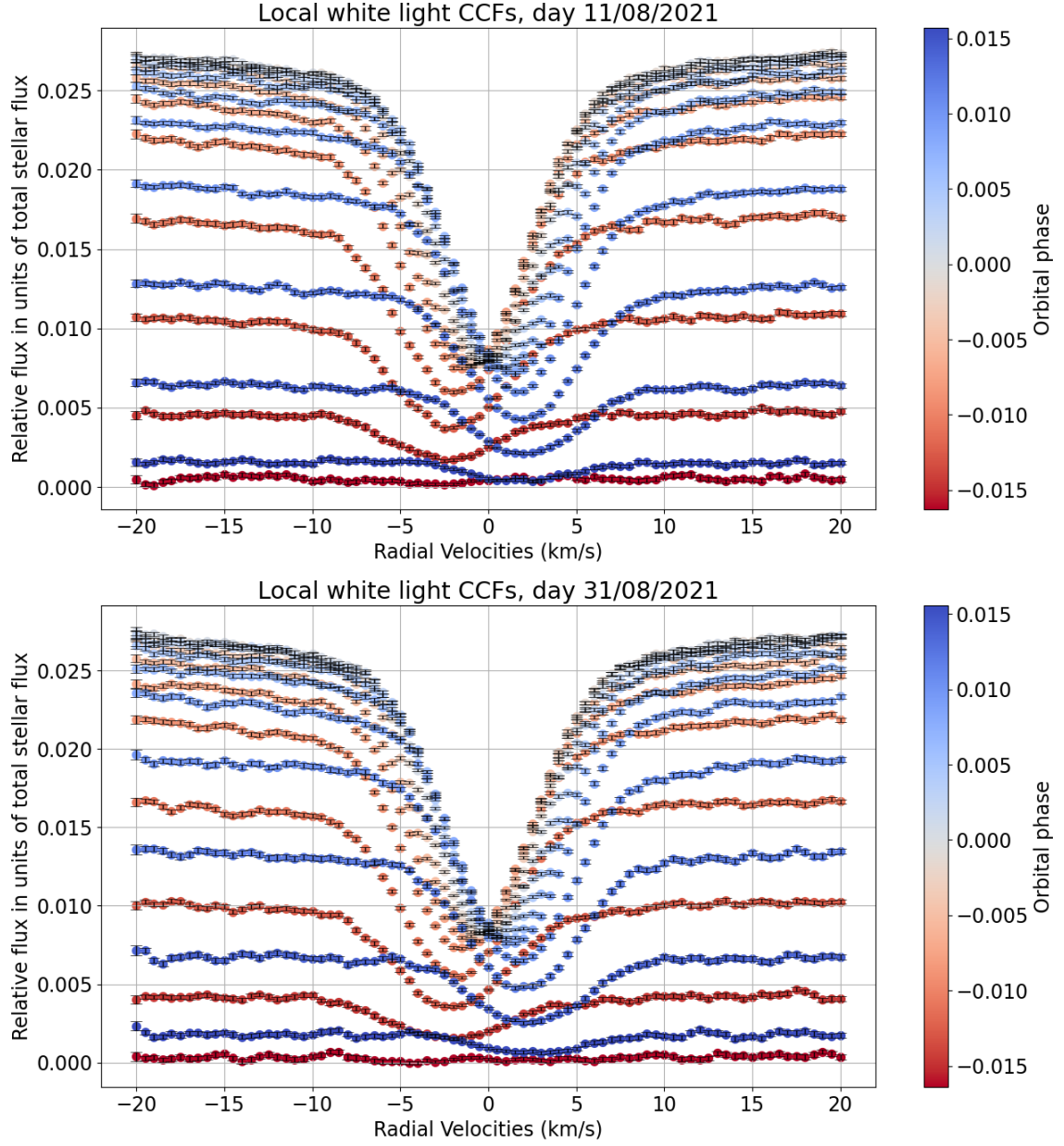


FIGURE 2.14: The white light local CCFs of the in-transit observations obtained after subtraction, the color of the points represents the orbital phase of the exoplanet at the time of observation, uncertainties are represented by their respective error bars. The top graph is for the observation on the 11/08/2021 while the bottom for the day 31/08/2021.

noise as these regions do not contain any information of the stellar surface for obvious reasons. The central radial velocities data found in 3.3 are also overlaid with their respective uncertainties.

Figures 2.17 and 2.19 show tomographies of the weak iron lines CCFS and strong iron lines CCFs respectively over the course of the transit.

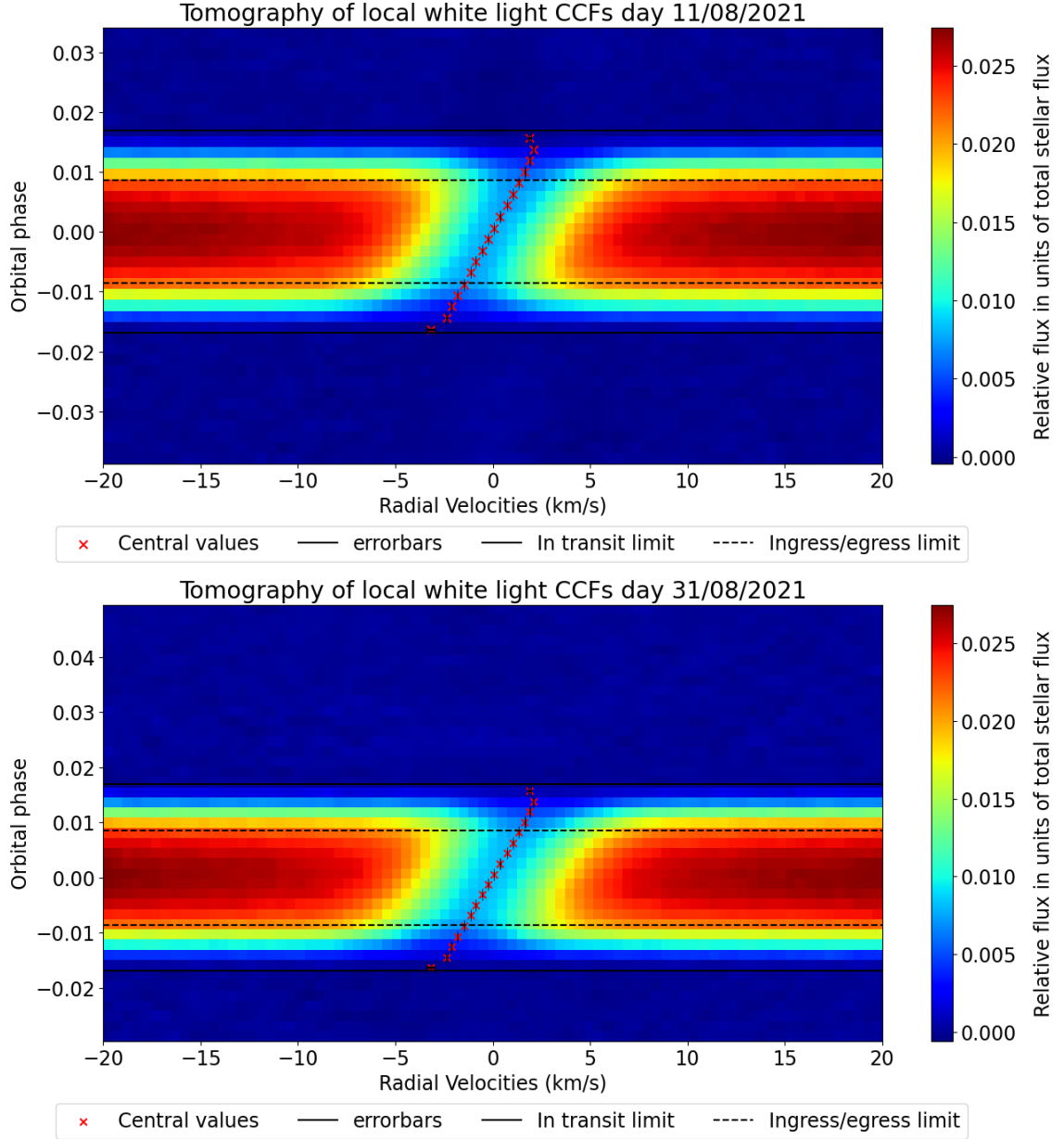


FIGURE 2.15: Tomography of the subtracted white light CCFs, the y axis represents the orbital phase while the x axis the radial velocities of the CCFs, the color represents the relative flux in the scale of the colorbar, which in this case 1 would represent the total flux from the star. The horizontal lines show the start and end of the transit with the dashed lines showing the limits of ingress/egress. In addition the crosses indicate the central radial velocities of the CCF (the same shown in Figure 3.3) with error bars indicating their uncertainties. The top graph is for the observation on the 11/08/2021 while the bottom for the day 31/08/2021.

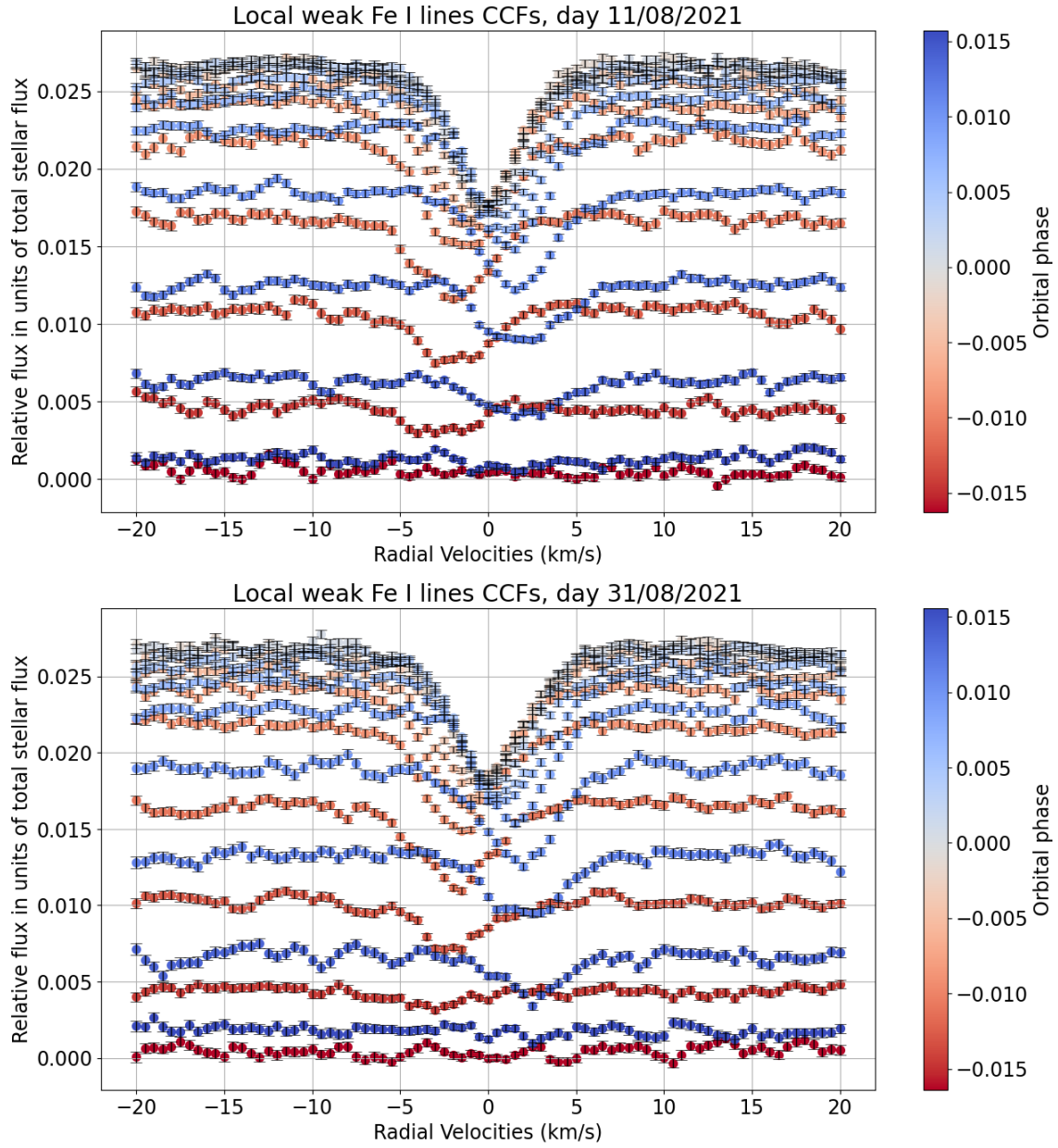


FIGURE 2.16: The weak Fe I lines CCFs of the in-transit observations obtained after subtraction, with the same representation as in Figure 2.14.

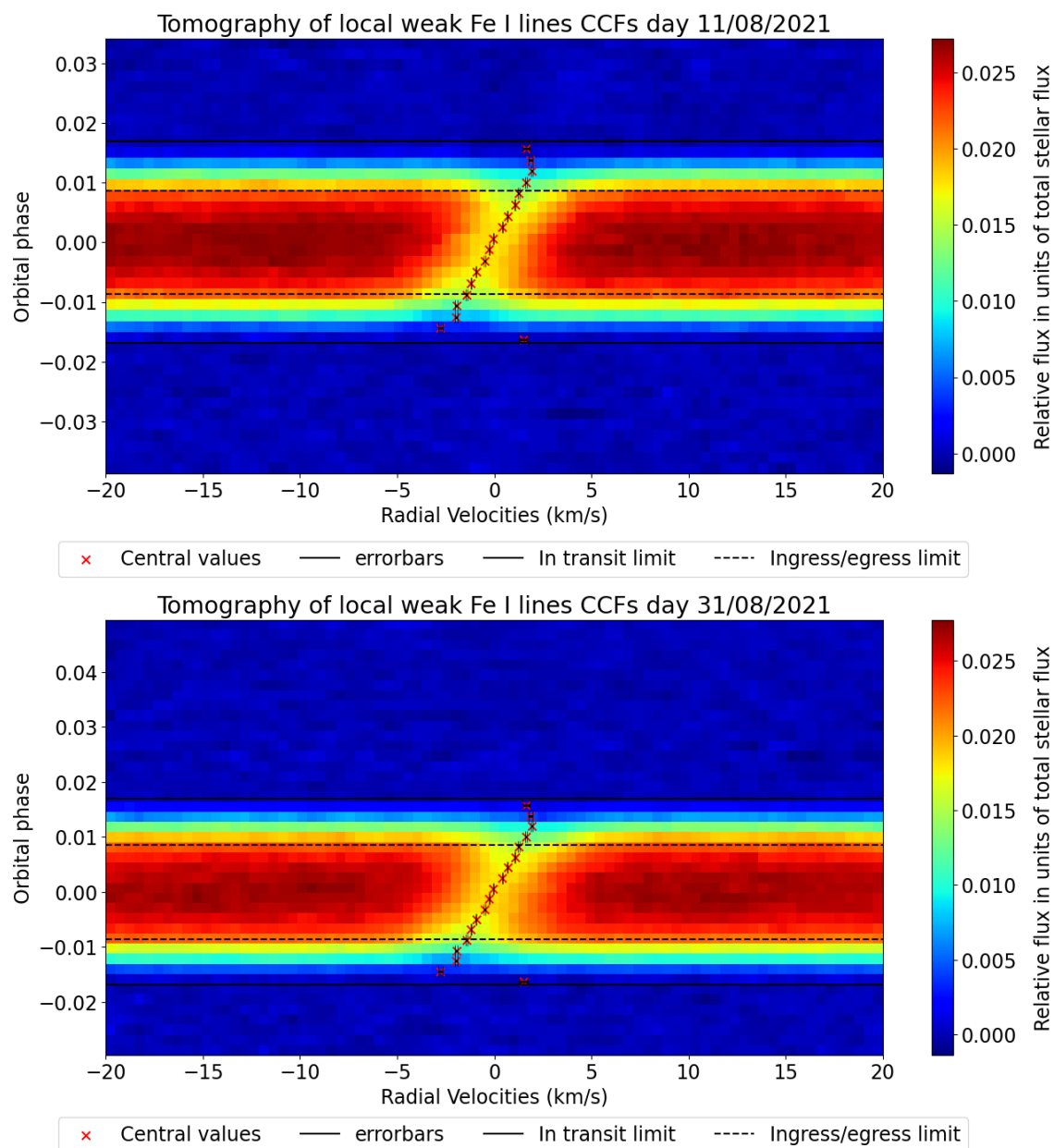


FIGURE 2.17: Tomography of the subtracted weak Fe I CCFs displayed in the same manner as Figure 2.15.

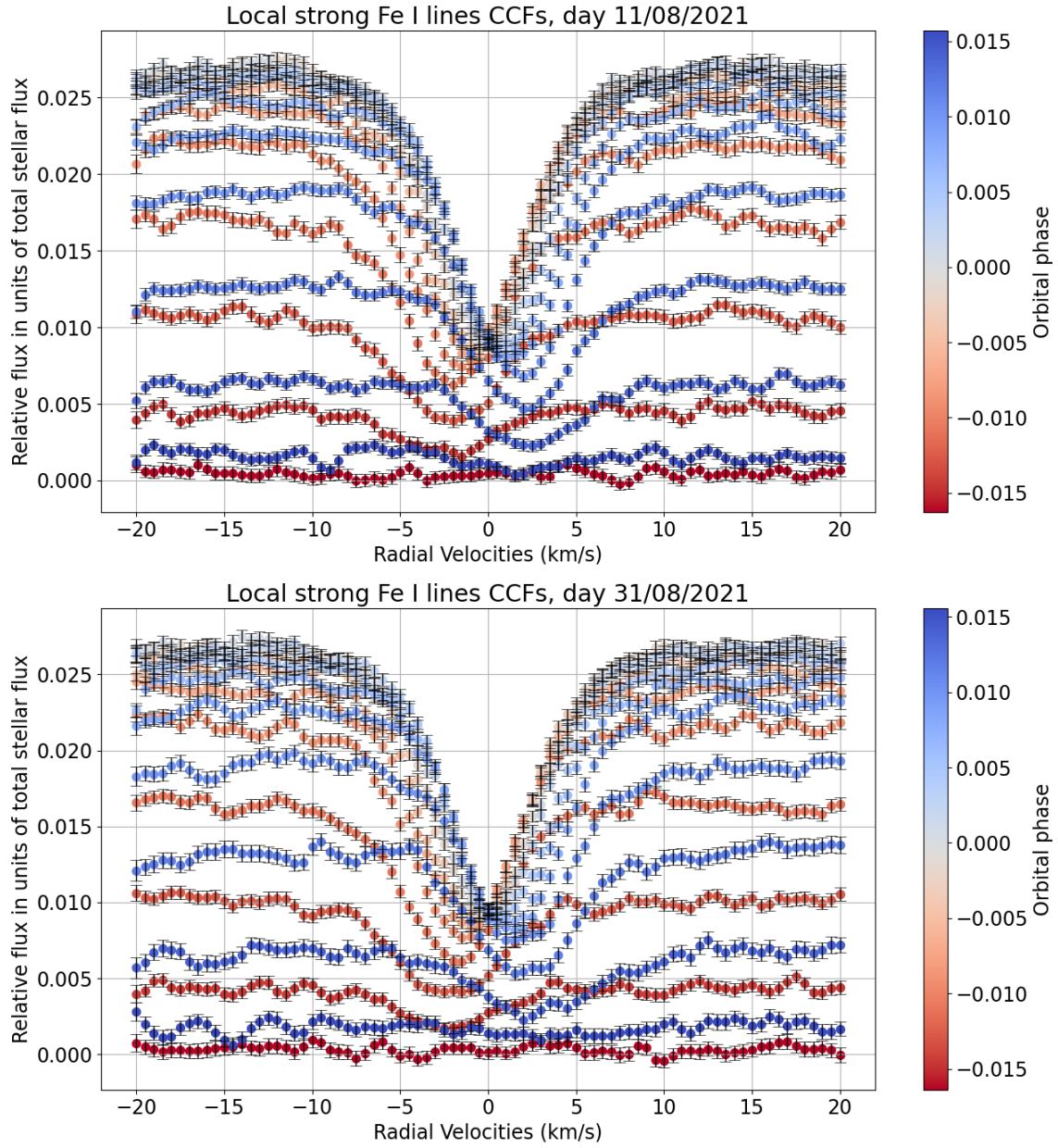


FIGURE 2.18: The strong Fe I lines CCFs of the in-transit observations obtained after subtraction, with the same representation as in Figure 2.14.

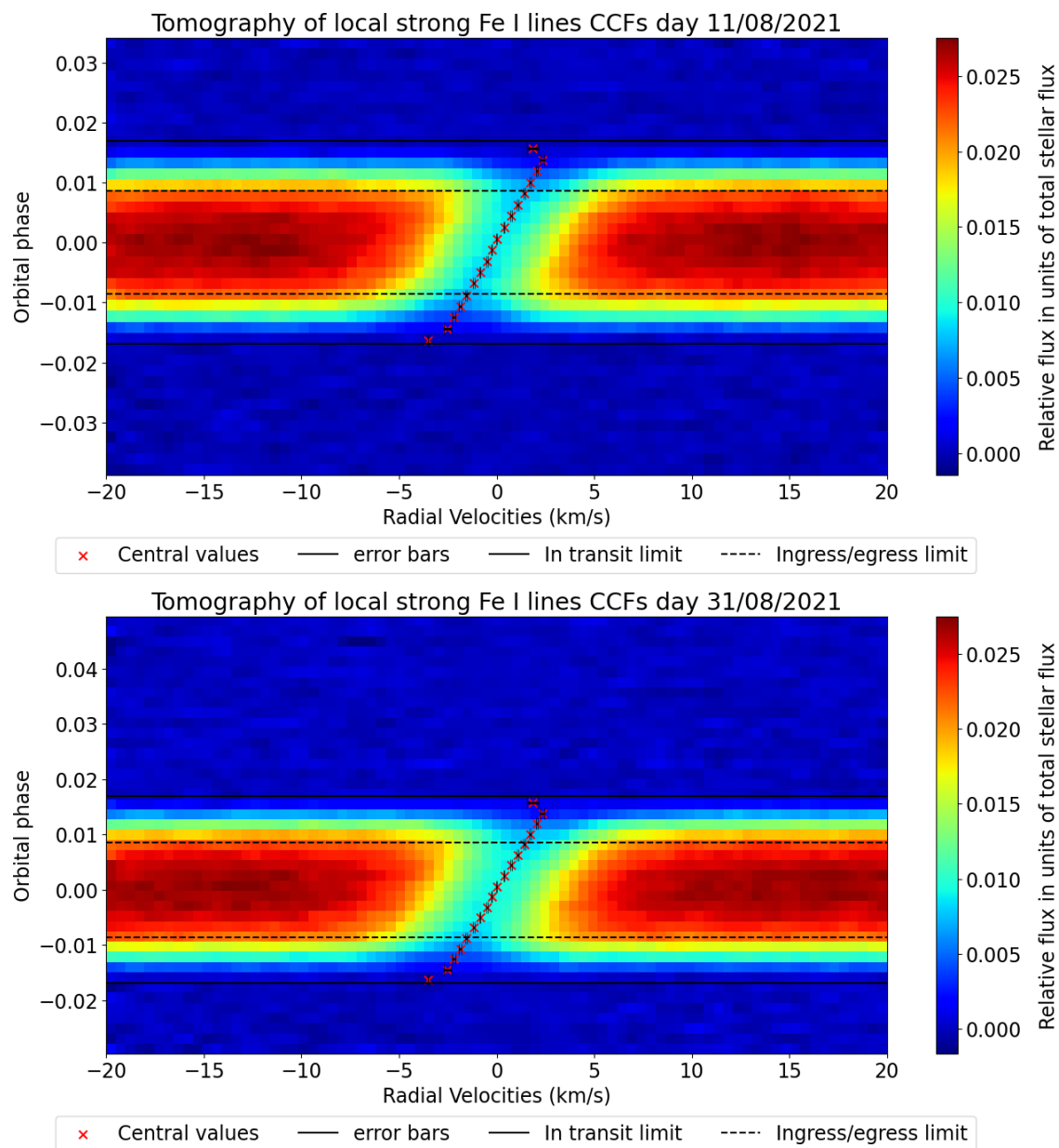


FIGURE 2.19: Tomography of the subtracted strong Fe I CCFs displayed in the same manner as Figure 2.15.

2.4.2 Individual spectral line analysis

The process of reconstruction of chosen individual lines is very similar to the one applied to obtain the subtracted or local CCFs, as such we will describe the main differences between the methodologies applied.

We will start with an already locally normalized part of a spectra that corresponds to the selected wavelength interval of our lines of choice. In a similar way to the CCFs we first start by fitting them with a function, reminding that we selected the $H\alpha$ line and the two sodium doublet lines.

The next step is the RV correction to the stellar rest frame for which we used the RV calculated from the slope values (m) of the linear fit of the star's center of mass motion from the white light CCFs (Figure 2.10) as they provide the most precise data. This correction is done first as we already worked with the CCFs. For the spectra correcting the wavelength to the correct velocity is done using the Doppler factor $\lambda' = \lambda \sqrt{\frac{1+v/c}{1-v/c}}$ where v is the velocity used for correction, i.e., the opposite number of the star's radial velocity.

We do not use the values for the constant radial velocities (b) of the linear fit of the CCF RVs as each spectral line presents a different constant speed in comparison to each other and to the CCFs. For that reason, we will also fit our chosen lines with a profile function, and perform a similar linear fit to the out-of-transit radial velocities of the lines to get their constant speed component, which we will subtract to each respective line. The same fit to a profile function of the lines will be used to get the continuum value which we will use to renormalize the chosen lines even if the difference in flux is small.

For the sodium doublet the modified Gaussian function (Eq. 2.2) that was also used for the CCFs proved to offer a very close fit, Since the sodium doublet lines have very narrow and pointy cores which the variable exponent c can replicate. This would not be possible with the usual fitting functions for spectral lines. In addition, due to the big overlap that the wings of each of the two sodium lines have, it was decided to work with them simultaneously for this reason our fitting function consists of two modified Gaussians that share the same continuum.

The $H\alpha$ line has a very wide and broad core for which we also opted for a modified Gaussian profile, although due to the line being very broad the fitting has trouble in finding its central value. Other profiles were attempted such as the Voigt profile and a regular Gaussian function which did not show much difference in their result, but the modified Gaussian was chosen to be consistent in our methodology.

With the fitting parameters known we can then renormalize them to guarantee there is no difference in flux. For the constant velocity (b) we fitted each of these constant RV individually and subtract them from their respective line (for the sodium doublet lines we average their two values so their separation does not change). We then obtain the corrected line spectra in the stellar rest frame whose radial velocities are seen in Figures 2.20 and 2.21.

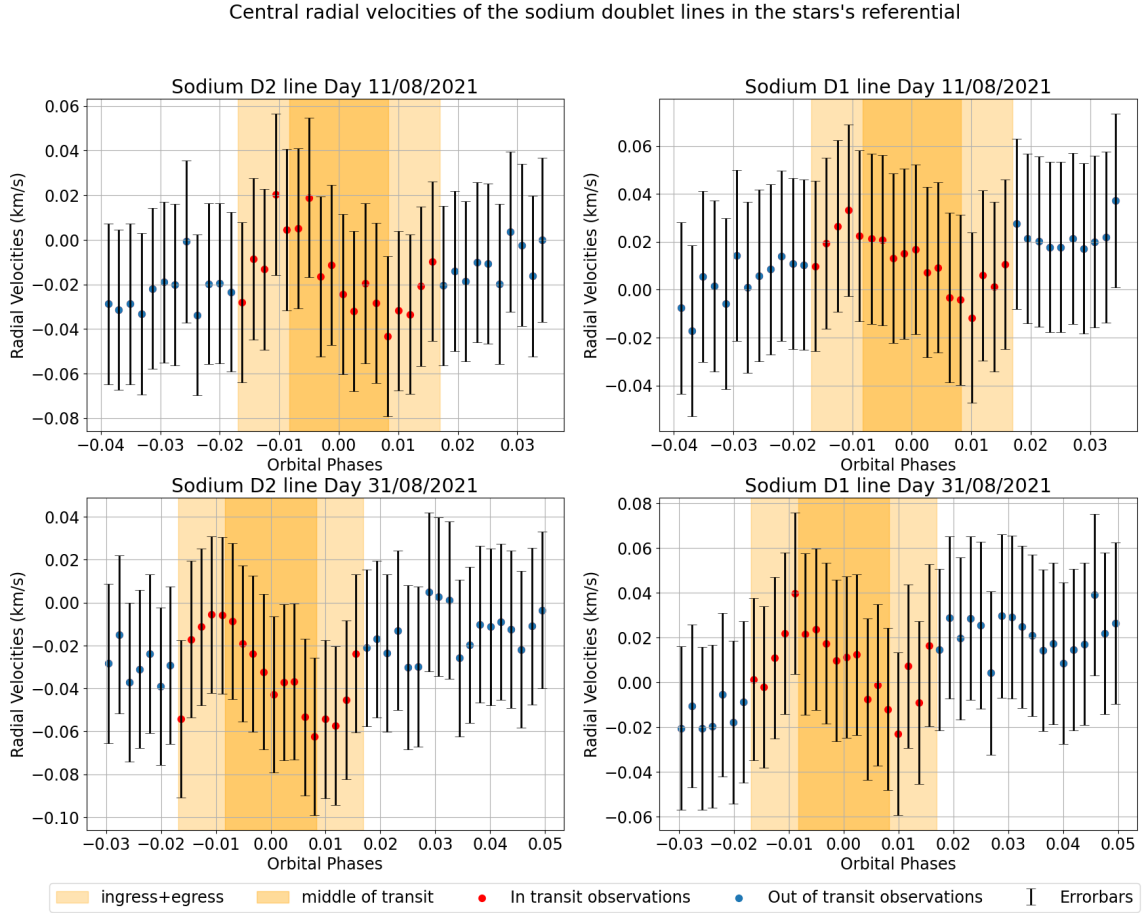


FIGURE 2.20: Barycentric corrected values of the sodium doublet for the two nights of observation 11/08/2021 left side and 31/08/2021 right side, with the same legend as in Figure 2.10.

With the lines in the star's referential we proceeded to interpolate them so operations with the spectra can be made. This was done in the same manner as for the CCFs which is discussed further in appendix A.

We then take each set of out-of-transit lines spectra and average them to create a master out-of-transit for the $H\alpha$ line and sodium doublet for the two nights of observation. We decided to compare both by dividing the master out-of-transit spectra of one night by

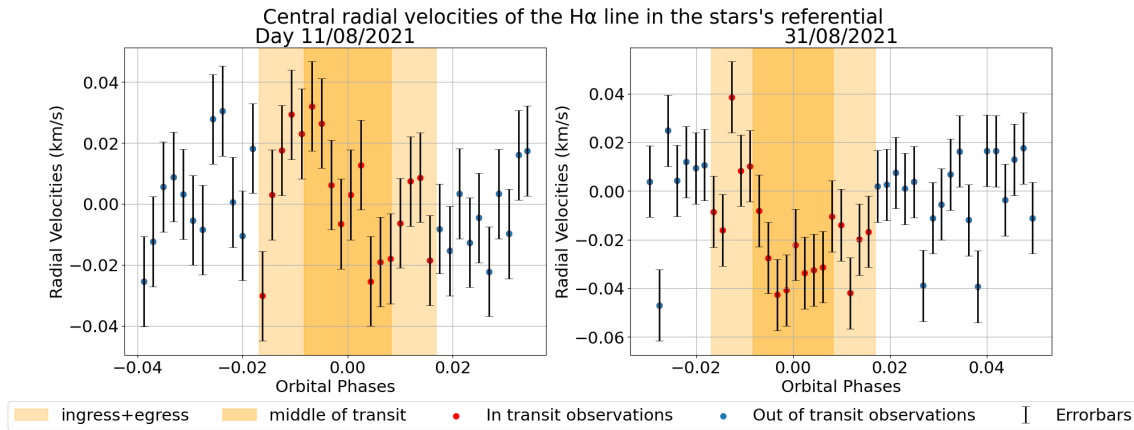


FIGURE 2.21: Barycentric corrected values of the $H\alpha$ line for the two nights of observation 11/08/2021 left side and 31/08/2021 right side, with the same legend as in Figure 2.10. Of note that the error bars have been scaled down by a factor of 100 so we can see the points in more detail, although any deviation is statistically insignificant. This is due to the $H\alpha$ line being very wide and as such the fitting has greater difficulty in finding the center value.

the other which can be seen in Figures 2.22 and 2.23. For the sodium doublet the only difference is off a scaling constant indicating a slight difference in flux with no meaningful variation between the two, while for the $H\alpha$ the continuum is essentially equal but the shape of the profile itself has a significant difference. This could be due to differences in stellar activity as $H\alpha$ is a well-known stellar activity indicator and the two observations are separated by 20 days with HD 189733 A being known as an active star.

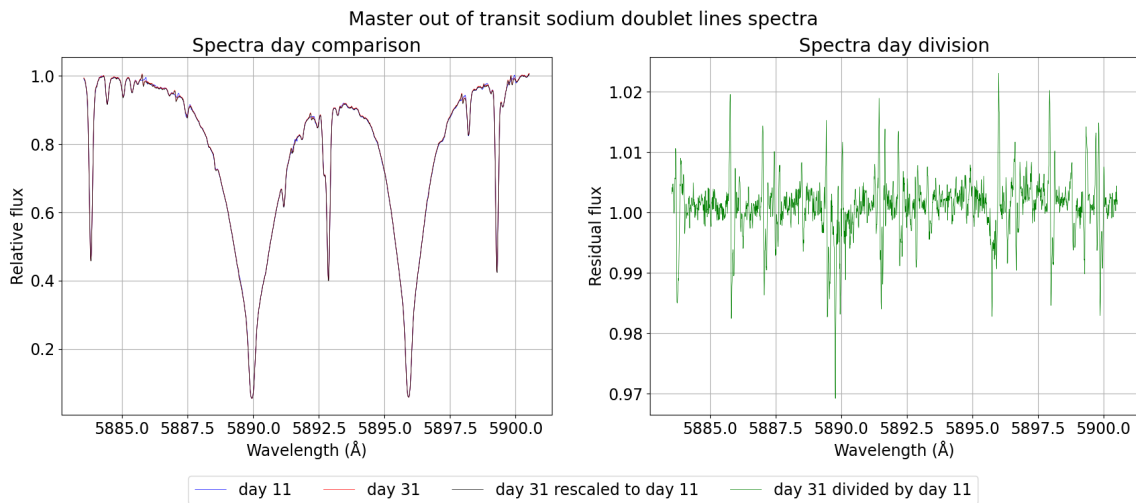


FIGURE 2.22: Left side: Master out-of-transit spectrum for the sodium doublet for the two days of observation including day 31 divided by its mean and multiplied by the mean of day 11 for showing the almost perfect alignment. Right side: spectrum of day 31 master out-of-transit sodium doublet divided by day 11 showing minimal noisy difference.

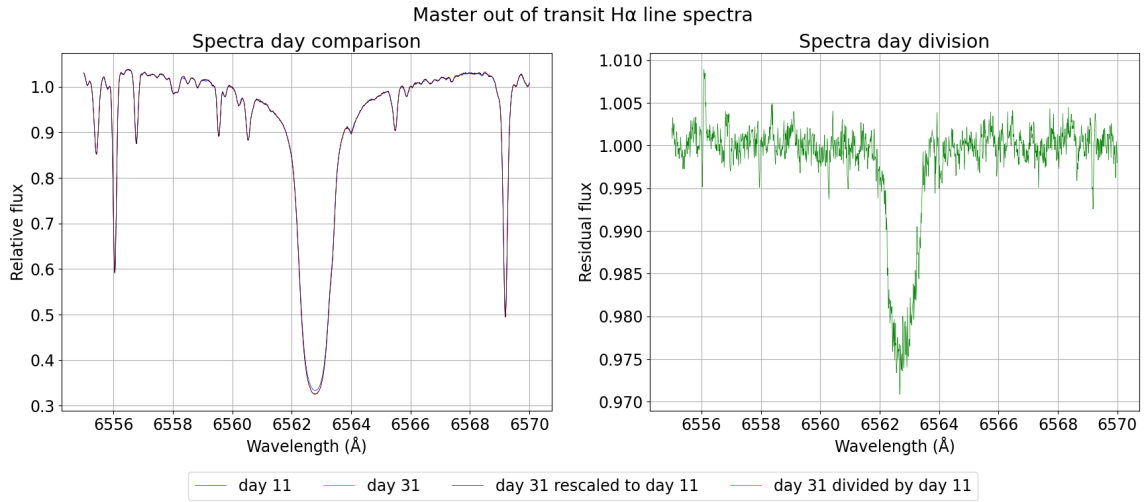


FIGURE 2.23: Left side: Master out-of-transit spectrum for the $H\alpha$ line for the two days of observation including day 31 divided by its mean and multiplied by the mean of day 11 which did not show significant difference. Right side: spectrum of day 31 master out-of-transit $H\alpha$ line divided by day 11 showing a difference of the line profile.

We now correct the in-transit line spectra to their correct flux again using the SOAP software described for the CCFs. The main difference is that since we are covering a smaller interval of the spectrum we had to adjust the quadratic limb darkening coefficients respectively, we adjusted the values for the functions from the `ldtk` python package, choosing an interval of wavelengths for the two sets of lines and adding more samples and a burn in to better increase the accuracy of the Markov chain. For the sodium doublet we chose values between 586nm and 592nm , while for the $H\alpha$ line we chose between 650nm and 660nm . The resulting light curves were similar to the ones in Figure 2.13.

We finally proceeded to the subtraction which is done exactly the same as for the CCFs (Eq. 2.4) resulting in the reconstructed line profiles for the multiple areas of the stellar surface. We only note that because a CCF contains information from multiple lines their uncertainties are lower and S/N is higher when compared to a single line from a spectrum. So our reconstructed line profiles have much higher uncertainties and a lower S/N.

In a similar manner to the residual CCFs we then fit the reconstructed line profiles to better understand their features as well as extracting the same variables from the master out-of-transit for comparison. The results can then be seen in the next Chapter.

The local sodium doublet lines spectra are displayed in Figure 2.24, in addition to the local $H\alpha$ line spectra which are visible in figure 2.26.

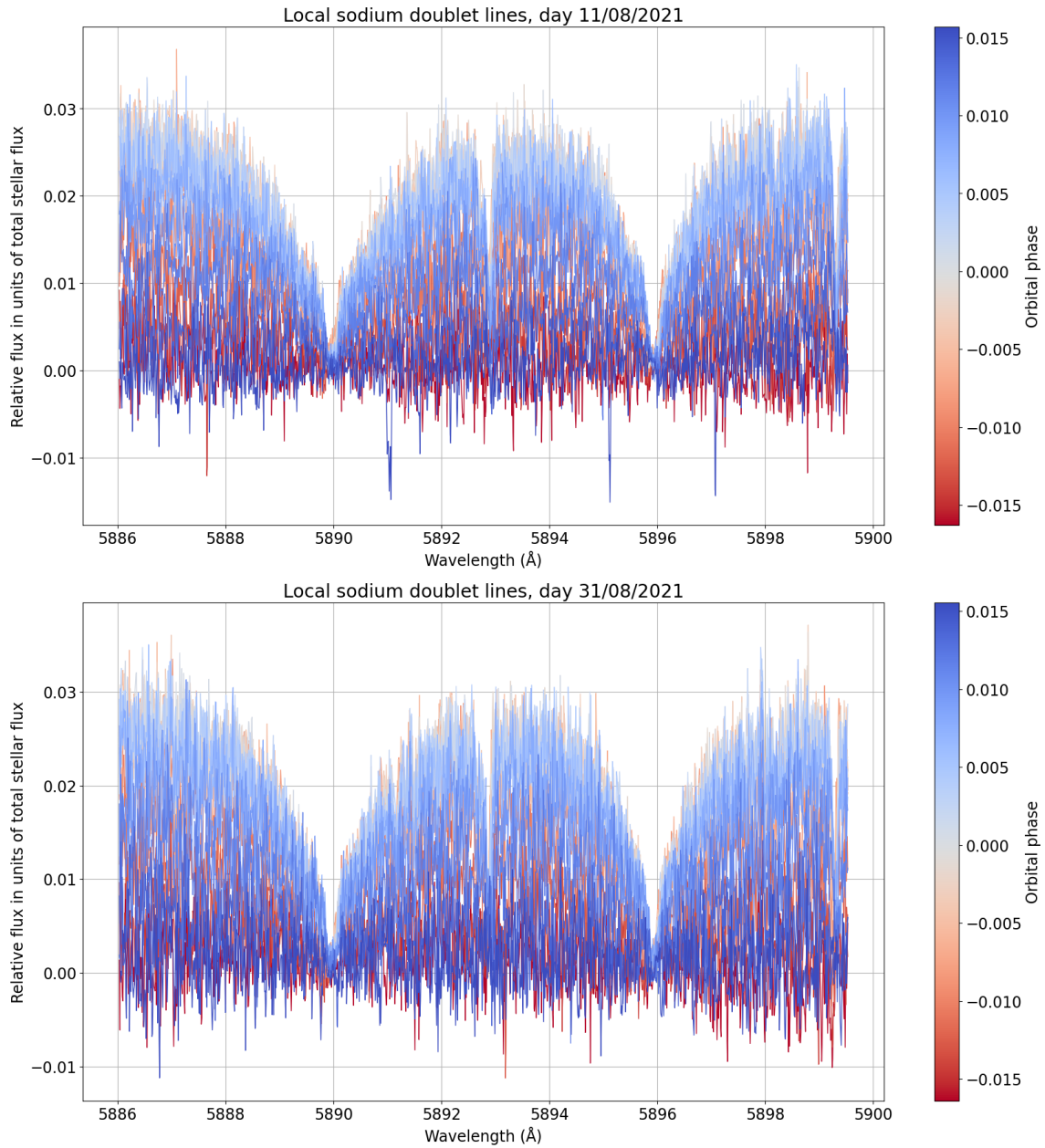


FIGURE 2.24: The reconstructed sodium doublet lines of the in-transit observations after subtraction, with the same representation as in Figure 2.14, although error bars were excluded for ease of visibility.

The tomographies of the reconstructed local sodium doublet lines for all observations are shown in Figure 2.25. These tomographies give an overview of all the line profiles change over the course of the transit located within the horizontal black lines, the more visible aspect being their radial velocity.

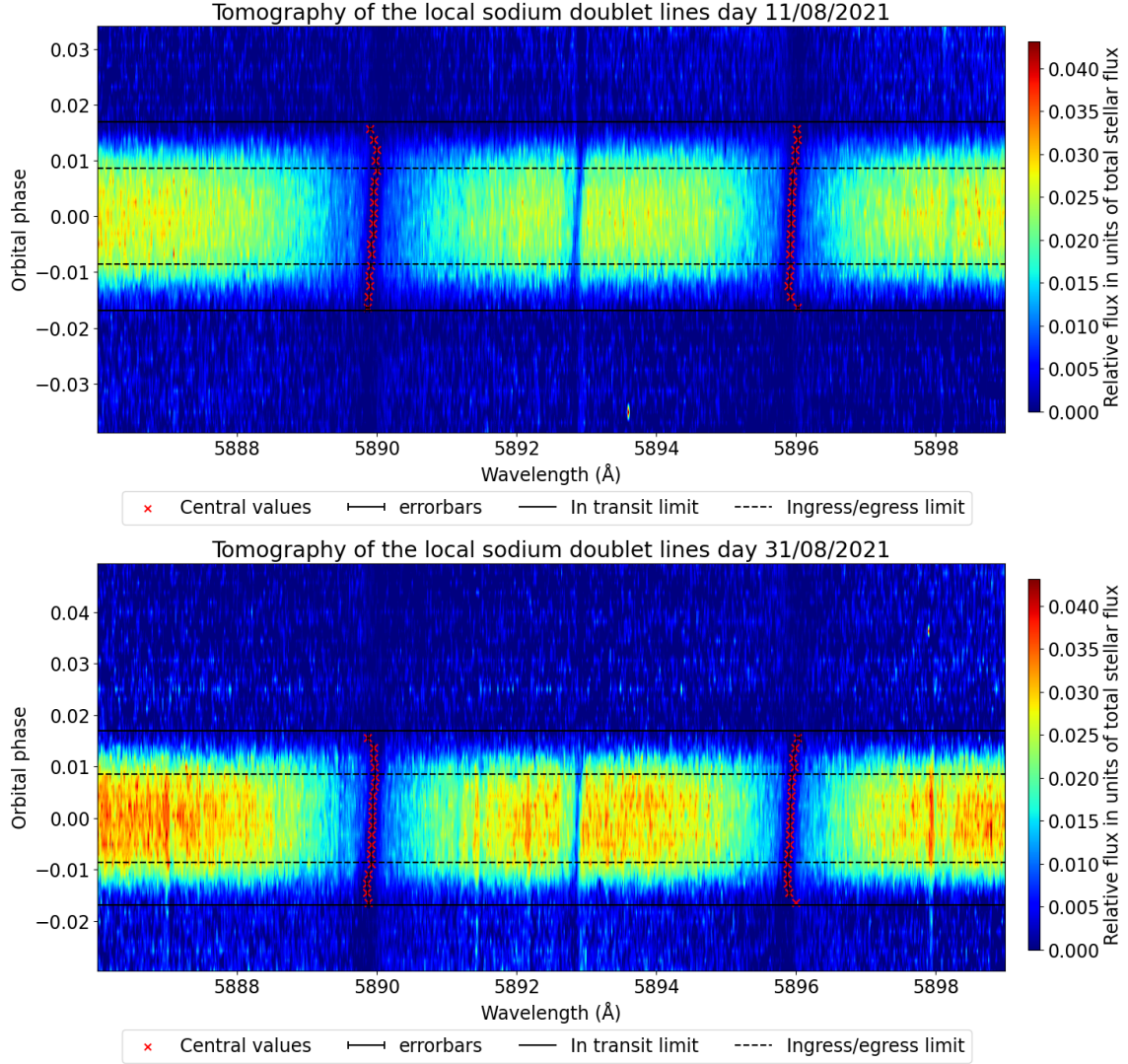


FIGURE 2.25: Tomography of the reconstructed sodium doublet lines displayed in the same manner as Figure 2.15.

We also present the tomographies of the local $H\alpha$ lines for both days in Figure 2.27 from which becomes clear how broad the $H\alpha$ lines are compared to the other profiles which gives an harder time in successively retrieve its describing parameters from fitting.

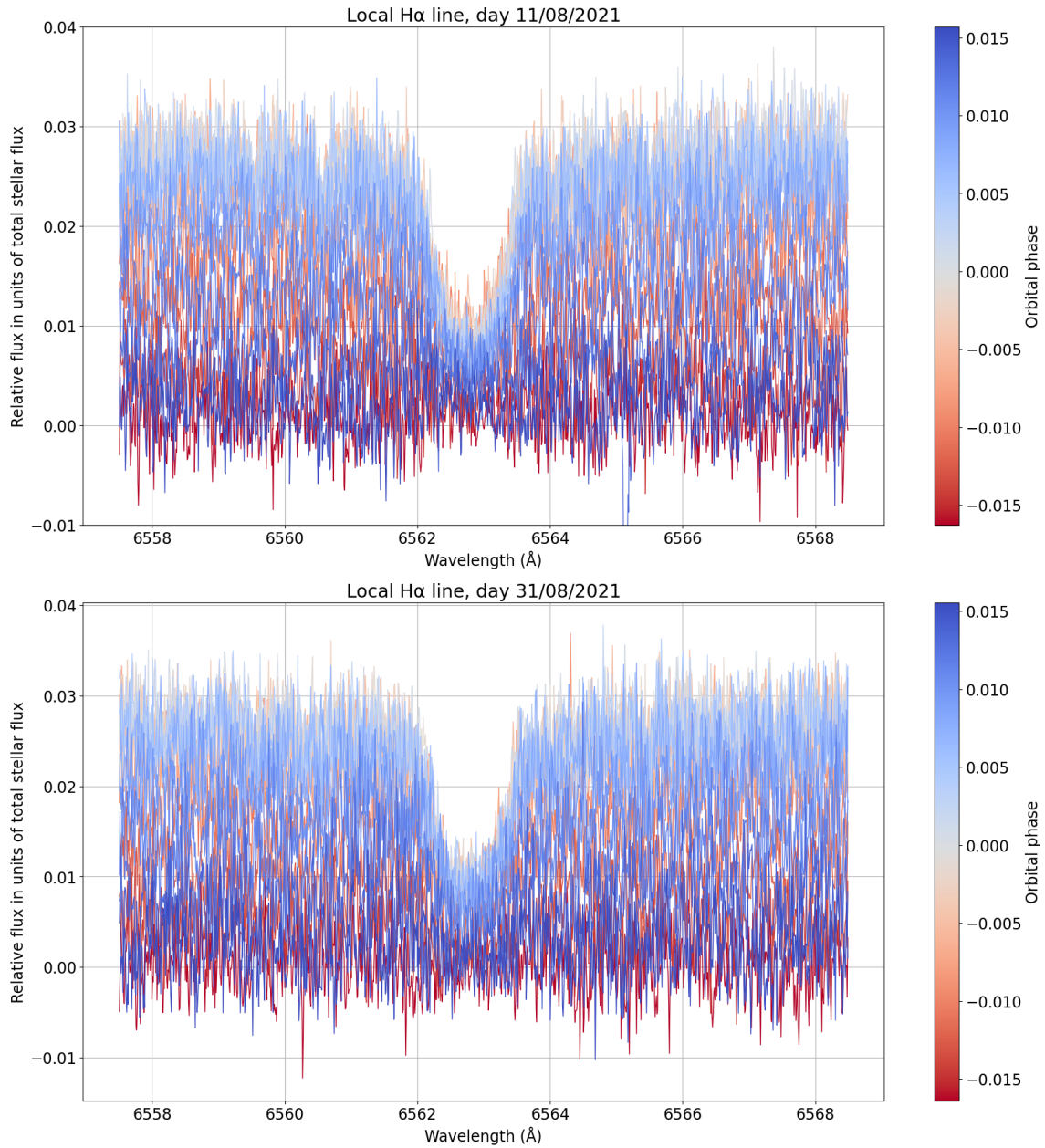


FIGURE 2.26: The reconstructed H α line of the in-transit observations after subtraction, with the same representation as in Figure 2.14, although error bars were excluded for ease of visibility.

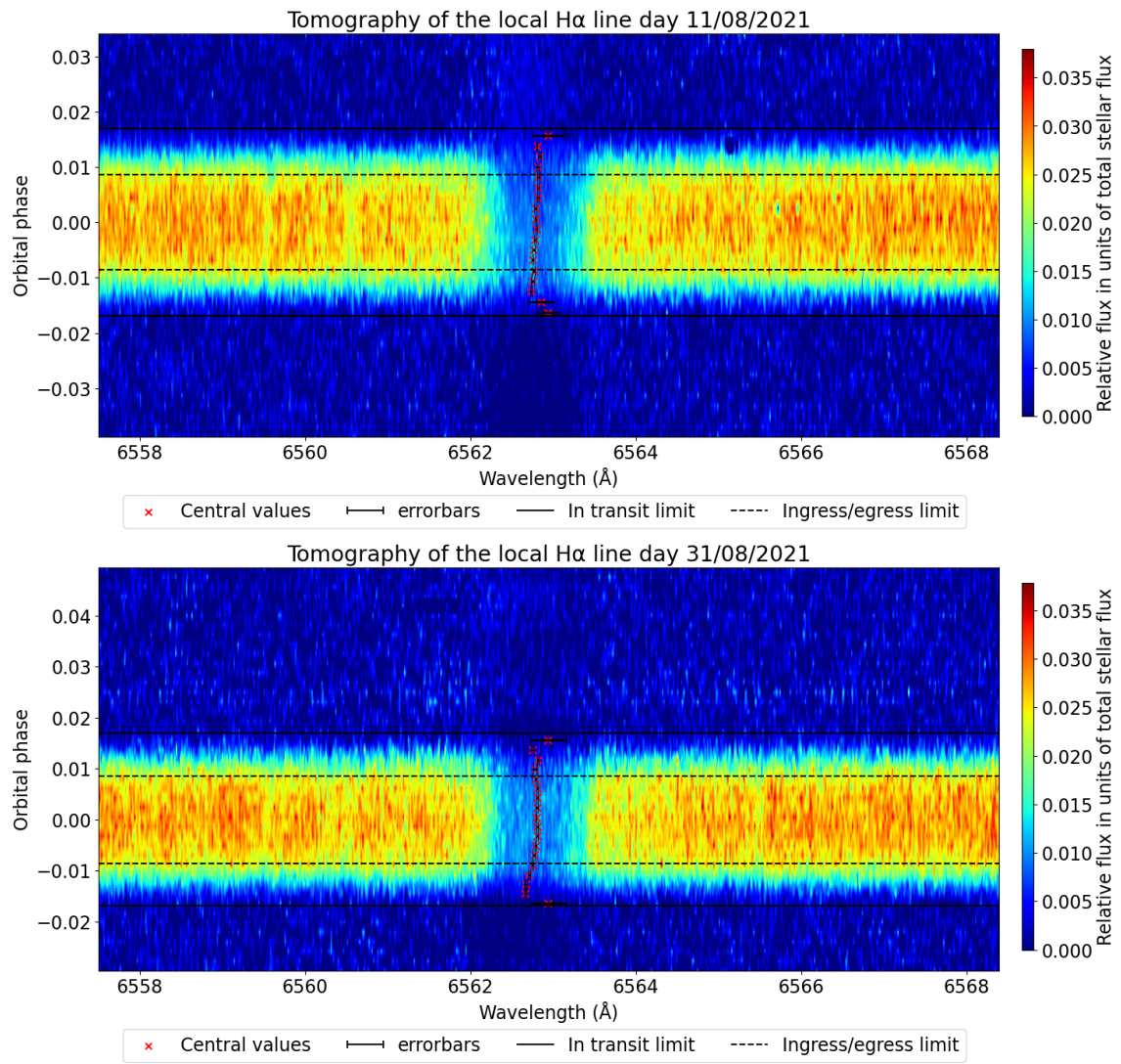


FIGURE 2.27: Tomography of the reconstructed $\text{H}\alpha$ line displayed in the same manner as Figure 2.15.

Chapter 3

CCFs and individual line analysis results

In this chapter we present the results of our data analysis performed in the previous chapter. First we show the data for the white light CCFs, followed by the CCFs of weak and strong Fe I iron lines, and finally the results for the two sodium doublet lines and H α line.

Before going into the results it is important to consider that some of our CCFs/lines were discarded. This was due to the fact that the local data at the end or start of a transit is extremely noisy resulting in very poor fits in contrast with the rest of the data. As such we calculated the R^2 (a description is given in appendix C) of each fit in order to establish a criterion of what data points are to be discarded. We decided to discard data with an R^2 of less than 0.5.

Furthermore, converting the orbital phases into μ results in some points being removed. This is due to some observations being done when the exoplanet's center, whose location is used for the calculation of μ is still outside the stellar disk despite already partly blocking it, which removes some data points during the start and end of the transit. In total the conversion from orbital phases into μ results in 2 data points at the start of transit and 2 at the end to be removed, shown in Figure 3.1.

For each set of data, the profiles were fitted with a modified Gaussian to retrieve their central radial velocity, line-width, and line-center intensity. These variables are respectively taken as x_0 , b and $(1 - \frac{a}{y_0}) * 100\%$ from the parameters of (Eq. 2.2). These variables were chosen so as to provide a direct comparison with the results presented in [Dravins et al. \(2018\)](#).

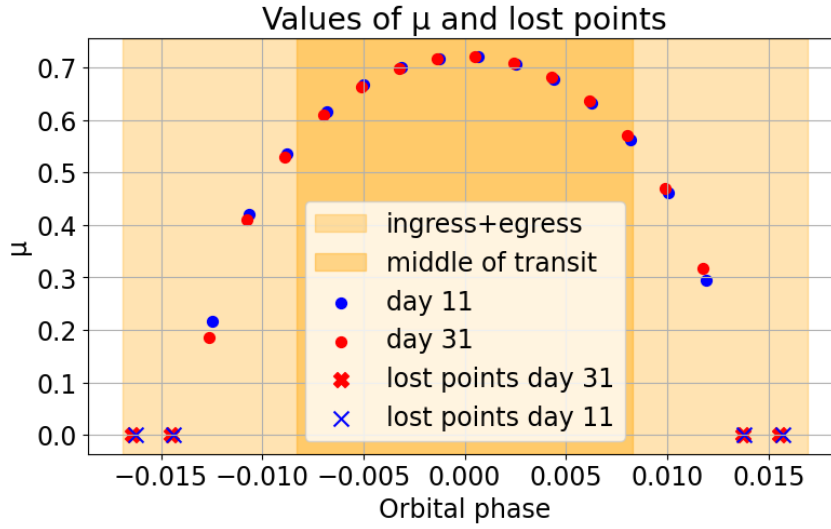


FIGURE 3.1: Values of μ in function of the respective orbital phase in circles. The crosses represent the orbital phases where the center of the planet is outside of the stellar disk and μ could not be calculated.

3.1 White light CCF

As a reminder the white light CCFs are the result of the cross correlation of the spectra and a K2 stellar class mask that was chosen to best represent the star as a whole.

When applying the methodology discussed in section 2.4.1 we first got the local CCFs, which can be seen in Figure 2.14. These are the CCF profiles that represent their respective areas of the stellar surface covered by the exoplanet, in which two clear characteristics of the profiles are evident. The first one is the limb darkening of the stellar surface where the profiles near the center have the greatest continuum flux values while those near the limb (at the start and end of the transit) have much smaller flux values. The smaller continuum flux also partly reflects the smaller information we have of those regions, as they represent moments where the exoplanet does not completely cover the stellar surface, so the flux we get during ingress and egress after subtraction is also partly impacted by this. The second easy to notice effect is the central radial velocity of the profiles which showcase the motion of the stellar surface due to the star's own rotation and are the missing component that causes the Rossiter-McLaughlin effect.

After performing the linear regressions we calculated their R^2 seen in figure 3.2, where we see that we discarded the first fits for both days as they were too noisy and as such their data will not be considered. These would have been removed regardless for the cases where we look at dependence in μ due to the conversion from orbital phases.

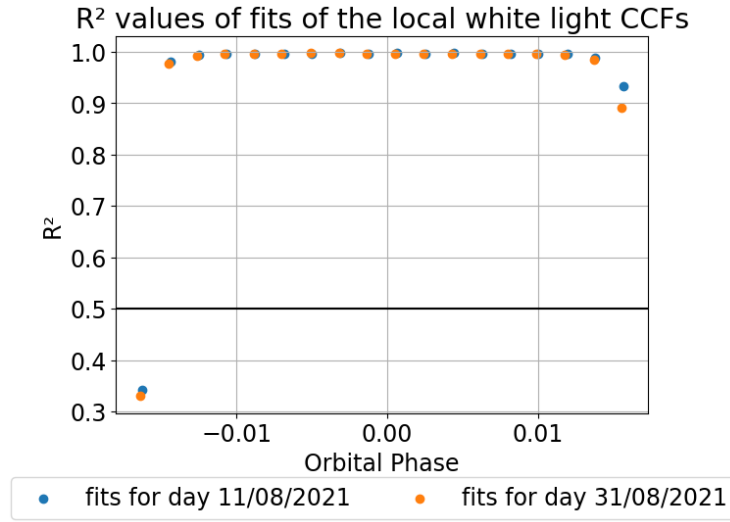


FIGURE 3.2: The R^2 values for each of the fits of the local white light CCFs in function of their respective orbital phase.

Figure 3.3 contains the results with the parameters obtained from fitting with a modified Gaussian function, the Figure contains 6 plots for the three studied parameters with the top ones being in function of orbital phases, while the ones at the bottom are in function of μ . We observe clear trends in our 3 sets of data. For each of the 3 variables we decided to perform a linear fit as a first approximation in order to evaluate if the behavior in our points is statistically significant. Several statistical tests were made to evaluate the quality of the fit. Moreover Figure 3.4 contains the residuals for each of the fits to our data points.

The first variable is central radial velocity in function of orbital phase which, as we already noticed before, increases along the path of the transit due to the rotation of the stellar surface. We verify an almost perfect linear behavior with the orbital phases, however in the residuals small wiggles in the residuals are visible. The linear fit ($y = mx + b$) has values $m = 167 \pm 1$ km/s, $b = -0.027 \pm 0.005$ km/s for 11/08/2021 and $m = 165 \pm 2$ km/s, $b = -0.003 \pm 0.008$ km/s for 31/08/2021.

The line-width in function of μ follows a more uncertain trend with it seemingly decreasing, furthermore all subtracted CCFs are verified to be narrower than the out-of-transit case which represents the integrated stellar surface, as is expected due to rotational broadening of the profiles. A difference in behavior is seen between the two days of observation with a line-width decrease from limb to center for both days. The linear fit has values $m = -0.3 \pm 0.1$ km/s, $b = 3.21 \pm 0.06$ km/s for 11/08/2021 and $m = -0.1 \pm 0.1$ km/s, $b = 3.20 \pm 0.07$ km/s for 31/08/2021.

Finally, the relative depth of the profiles in function of μ shows a clear increase from the limb to nearer the stellar center when seen as a function of μ with all subtracted profiles being deeper than the out-of-transit case. A difference between the two days of observation is also visible with the lines of the first day being mostly deeper than the ones at the 31st of August. The linear fit has values $m = -14 \pm 2$, $b = 39 \pm 1$ for 11/08/2021 and $m = -12 \pm 1$, $b = 38.8 \pm 0.9$ for 31/08/2021.

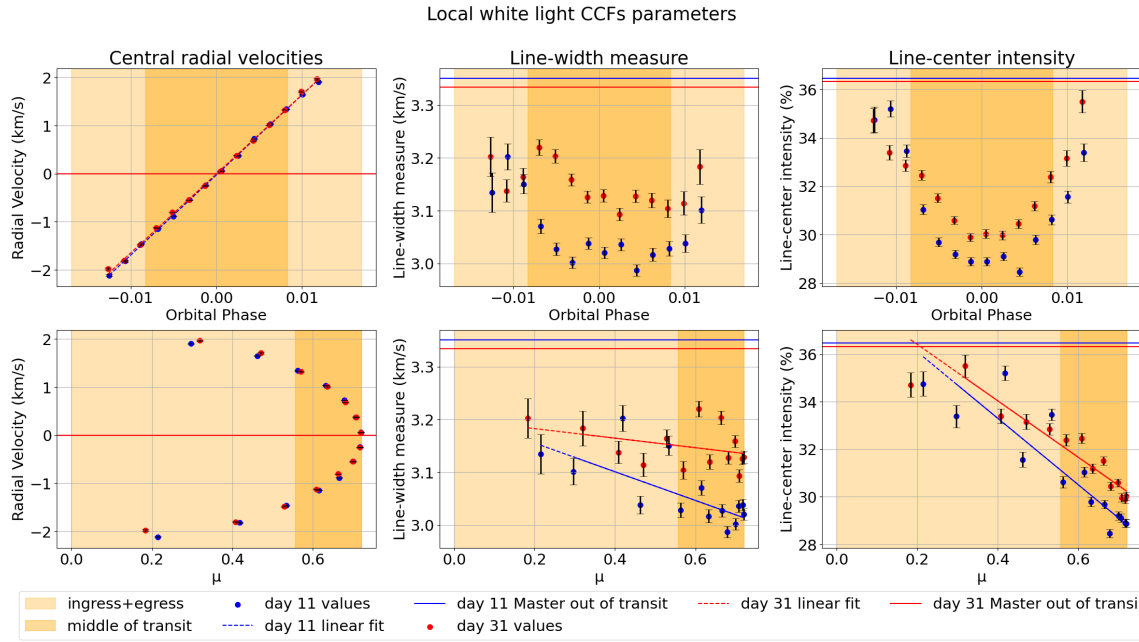


FIGURE 3.3: The sets of data extracted from the subtracted white light CCFs through the use of function fitting. The sets of data are from left to right the central radial velocity of the CCF, the line-width of the CCF profile, and the line-center intensity of the profile. The top row as them represented in function of the orbital phase while the bottom row is in function of μ . The blue data is for the day 11/08/2021 while the red is for the day 31/08/2021 with error bars for their uncertainties. Horizontal solid lines represent for comparison the values from the master out-of-transit which was retrieved in the same way as our subtracted CCFs. The dashed lines represent the linear regressions performed to our data points. The error bars are marked in black and finally the shaded area represents the transit interval with the lighter regions showing the ingress and egress of the transit.

In figure 3.5 we have distributions of spearman correlation coefficients and the base 10 logarithm of the p-values calculated with the `scipy.stats.spearmanr` function. These were estimated by sampling from the respective data points using a Gaussian whose mean is the value of the data point and the standard deviation the uncertainty of the data point and calculating the spearman coefficient and p-value of each set of 10000 samples of the data creating our distributions.

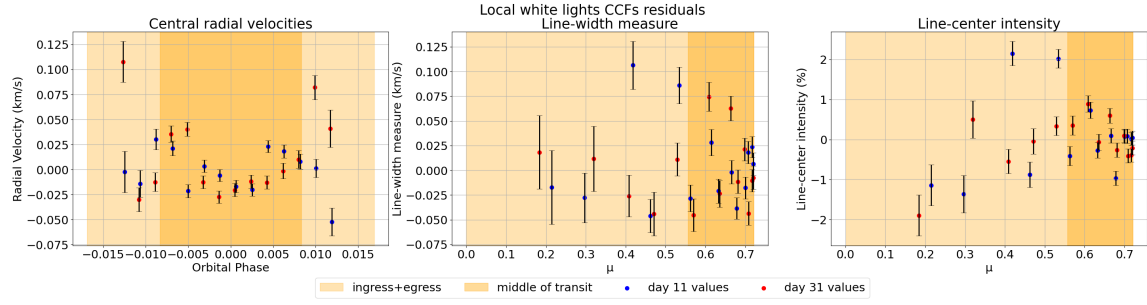


FIGURE 3.4: The residuals of each of each of the linear fits done in Figure 3.3. From left to right are the residuals of the center radial velocity, the line width and the line center intensity. The data in blue represents observations for the day 11/08/2021 while the data in red represents data from the day 31/08/2021. The shaded area represents the transit interval with the lighter regions showing the ingress and egress of the transit and error bars are shown in black.

Coefficients	Central radial velocity		line-width measure		line-center intensity	
	day 11	day 31	day 11	day 31	day 11	day 33
R^2	1.00	0.999	0.469	0.135	0.790	0.854
Pearson	1.00	1.00	-0.690	-0.370	-0.893	-0.939
χ^2 (linear fit)	88	184	88	88	186	67
χ^2 (constant mean)	233276	183277	195	94	1217	723
Spearman	1.00	0.999	-0.692	-0.327	-0.921	-0.969
$\log(P\text{-value})$	-63	-63	-2.217	-0.5967	-5.526	-7.941
Slope (m)	167 ± 1	165 ± 2	-0.3 ± 0.1	-0.1 ± 0.1	-14 ± 2	-12 ± 1
Intercept (b)	-0.027 ± 0.005	-0.003 ± 0.008	3.21 ± 0.06	3.20 ± 0.07	39 ± 1	38.8 ± 0.9

TABLE 3.1: Table containing for each of the 3 sets of data of the local white light CCFs and their respective fits for each day of observation the following quantities: the R^2 of the linear regression, the Pearson correlation coefficient, the χ^2 of the linear regression, the χ^2 of a horizontal line whose height is the mean of the specific data set, the Spearman correlation coefficient and the base 10 logarithm of the p-value associated to it. The linear fit parameters are also shown as m and b.

We also calculated the p-value and the Spearman correlation coefficient of the data itself without sampling in addition to the R^2 of the linear fit, the Pearson coefficient, and the χ^2 of the linear regression as well as the χ^2 of an horizontal constant line whose height is the mean of the data for model comparison, more details about these quantities are in appendix C. These statistical coefficients are summarized in table 3.1.

There are many potential choices in establishing a criterion for evaluating the statistical significance of an observed correlation, a common choice that we opted is a maximum possible p-value of 0.05 ($\log(p) \approx -1.30$) (Fisher 1954). For most cases the single p-value taken directly from the data was enough to establish statistical significance as the distribution which takes into account uncertainties does not impact such conclusions, but in some cases the distribution can be wide enough to cover high p-values due to high uncertainties and for those the correlations were also deemed insignificant despite the single

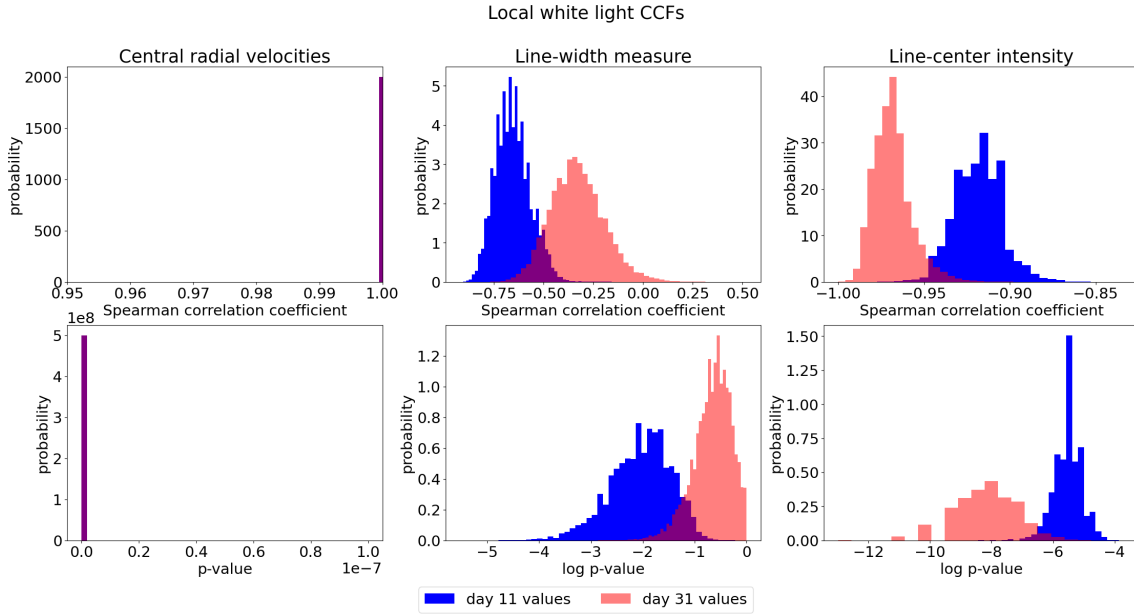


FIGURE 3.5: Plots containing distributions of the spearman coefficient and p-values calculated from 10 000 samplings of the data used for each of the 3 linear fits. The top row contains the distribution of spearman coefficients and bottom contains the p-values. From left to right the plots pertain to the linear fits of the center radial velocity, the line width and the line center intensity.

p-value taken directly from the data being behind the threshold of 0.05. The other statistical tests serve to complement our analysis.

From these results it is clear that the central radial velocities have the strongest linear behavior as to be expected. For both days the Pearson and Spearman coefficients showcase the great linearity and correlation of the data, the χ^2 shows a clear preference for the linear model over the constant mean and the p-value is extremely small. From Figure 3.5 where we take uncertainties into account in sampling we see the very distinct behavior of highly correlated data with the distribution of Spearman coefficients concentrated on one and the p-values concentrated on zero. The R^2 quantifies that the linear fit is highly accurate.

For the line-width there exists a difference between the two days of observation with the data for the first day showing a statistically significant downward trend with μ . This is shown by the Pearson and Spearman coefficients and a small p-value, comparing the χ^2 values also shows a clear preference for the linear fit model over the constant mean. In contrast the data for the 31st does not show a clear trend in the data with the χ^2 only showing a small preference for the linear model, with small Pearson and Spearman coefficients and a large p-value of over 0.05 (over -1.3 in base 10 logarithm). The distributions

on Figure 3.5 show the same story, with the data of the first day having a clear correlation with the distribution of the Spearman coefficients being centered on a non-zero value with the p-values being centered around zero. In opposition the second day gives a much wider Spearman distribution closer to zero with the p-value distribution also being very wide. In addition, both distributions have low R^2 values which shows that the linear fit does not approximate the data well.

And at last the line-center intensities for both days show a strong correlation with a clear downward trend as highlighted by the Spearman and Pearson coefficients. The χ^2 show a clear preference for the linear model and the R^2 show that the fit is decently good. The distributions in Figure 3.5 show for both days a clear correlation, with the Spearman distributions centered around specific non-zero values while the p-values are clearly concentrated around zero.

3.2 Iron CCFs

We now move to the iron CCFs, which consists of two types of CCFs one for the weak Fe I lines and other for the strong Fe I lines, the local CCFs are visible in Figure 2.16 for the former and in figure 2.18 for the latter. Again, just like for the white light CCFs we see the effects of limb darkening and ingress/egress on the continuum flux of the profiles, on top of the variation in radial velocity due primarily to the star's rotation. These features are clear for both types of CCFs, although in comparison to the white light the noise and uncertainties are higher as we are working with a more selective group of lines.

With the local CCFs of both sets of iron lines known, we then performed fitting with a modified Gaussian function. The R^2 values of the fit are present in Figure 3.6 and for both types of masks and both observed transits we discarded the first and last observations from our data sets. Reminding again that when converting from orbital phases to μ the first two and last two observations are lost since the center of the planet is outside the stellar disk.

Figure 3.7 contains the values obtained from the fits of the central radial velocities, line-widths and line-center intensities for the weak iron CCFs in addition to the three linear regressions done as a first approximation in the same manner as for the white light case. The residuals of the linear regressions are visible in Figure 3.8.

Looking at the central radial velocities vs orbital phases we see the same linear behavior as for the white light with small wiggles on the residuals and the points near

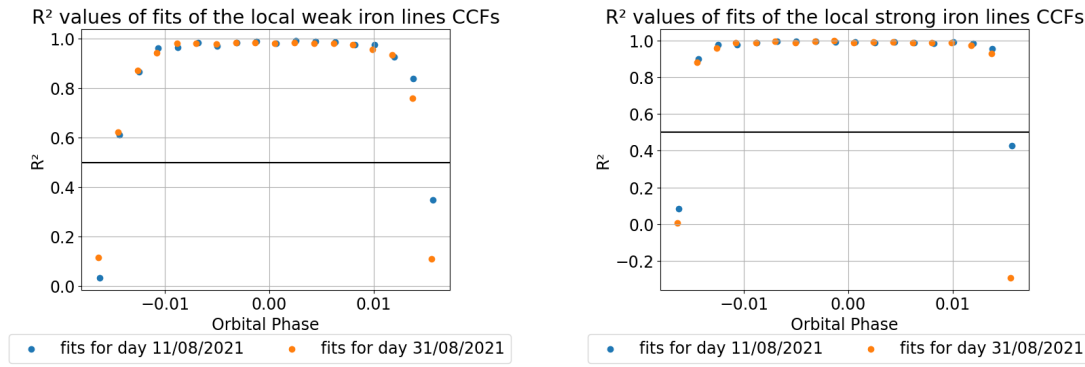


FIGURE 3.6: The R^2 values for each of the fits of the local white light CCFs in function of their respective orbital phase. Left: weak iron CCFs. Right: strong iron CCFs.

the limb being very noisy. The linear fit ($y = mx + b$) has values $m = 167 \pm 3$ km/s, $b = -0.05 \pm 0.02$ km/s for 11/08/2021 and $m = 169 \pm 2$ km/s, $b = 0.02 \pm 0.01$ km/s for 31/08/2021.

The line-widths are more uncertain compared with the white light with differences between the two days of observation, but overall all local profiles are narrower compared to the master out-of-transit value which represents the integrated spectra of the entire stellar surface. The linear fit has values $m = -0.7 \pm 0.3$ km/s, $b = 2.5 \pm 0.2$ km/s for 11/08/2021 and $m = 0.3 \pm 0.3$ km/s, $b = 1.9 \pm 0.2$ km/s for 31/08/2021.

Unlike the white light CCFs the line-center intensities of the weak iron CCFs are noisier and more unclear with differences between the two days, with the first day having a clearer downward trend with μ unlike the second day. The linear fit has values $m = -11 \pm 3$, $b = 74 \pm 2$ for 11/08/2021 and $m = -1 \pm 3$, $b = 68 \pm 2$ for 31/08/2021.

Moving into the strong iron CCFs whose results are in Figure 3.9. This time unlike the weak iron lines the results are less noisy and much better defined as one would expect since the set of strong iron lines are deeper and more well defined. The residuals of the linear fits are represented in Figure 3.10.

The central radial velocities just like in previous cases feature the same linear behavior due to the star's rotation. However, the wiggles visible in the residuals along the transit take the form of two distinct waves symmetric around the center of the transit highlighting additional phenomena on these set of lines. The linear fit as values $m = 172 \pm 2$ km/s, $b = -0.02 \pm 0.01$ km/s for 11/08/2021 and $m = 171 \pm 2$ km/s, $b = 0.00 \pm 0.01$ km/s for 31/08/2021.

Due to the greater accuracy of the data of the stronger lines, the line-width values seem to have a more noticeable decreasing behavior consistent with both days although with

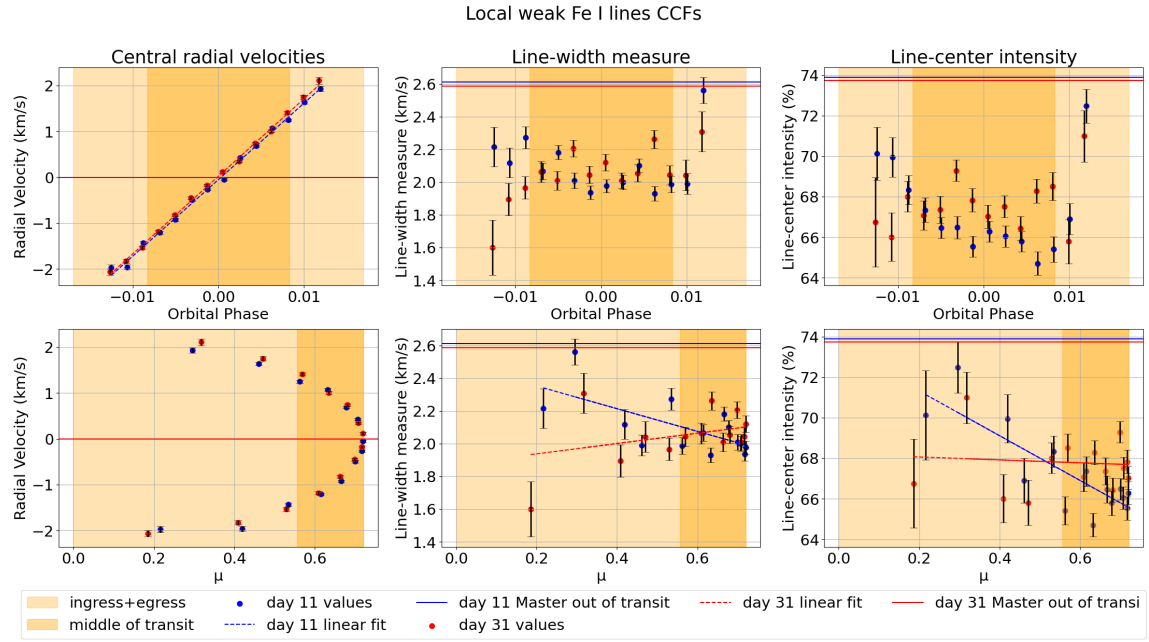


FIGURE 3.7: The sets of data extracted from the subtracted weak Fe I lines CCFs through the use of function fitting, displayed in the same manner as Figure 3.3.

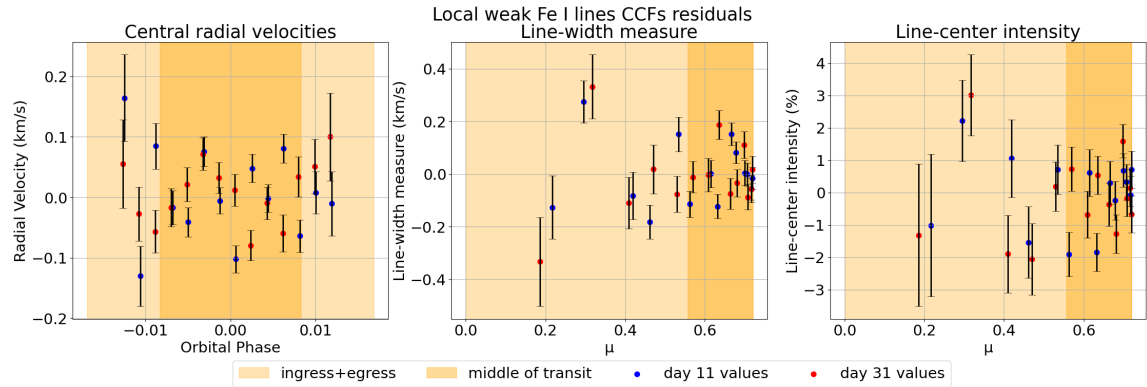


FIGURE 3.8: The residuals of each of each of the linear fits done in 3.7, displayed in the same manner as Figure 3.4.

11th being narrower compared to the data of 31st. The linear fit as values $m = -0.2 \pm 0.2$ km/s, $b = 2.8 \pm 0.1$ km/s for 11/08/2021 and $m = -0.3 \pm 0.2$ km/s, $b = 3.0 \pm 0.1$ km/s for 31/08/2021.

Finally, the line-center intensities are well defined compared to the weak iron CCFs as they show a clear decreasing trend with lines of the first day being slightly deeper compared to the second day, in addition to both being deeper to the master out-of-transit value. The linear fit as values $m = -11 \pm 2$, $b = 41 \pm 2$ for 11/08/2021 and $m = -12 \pm 2$, $b = 43 \pm 1$ for 31/08/2021.

We now look at the statistical tests made for the data of the weak iron CCFs and their

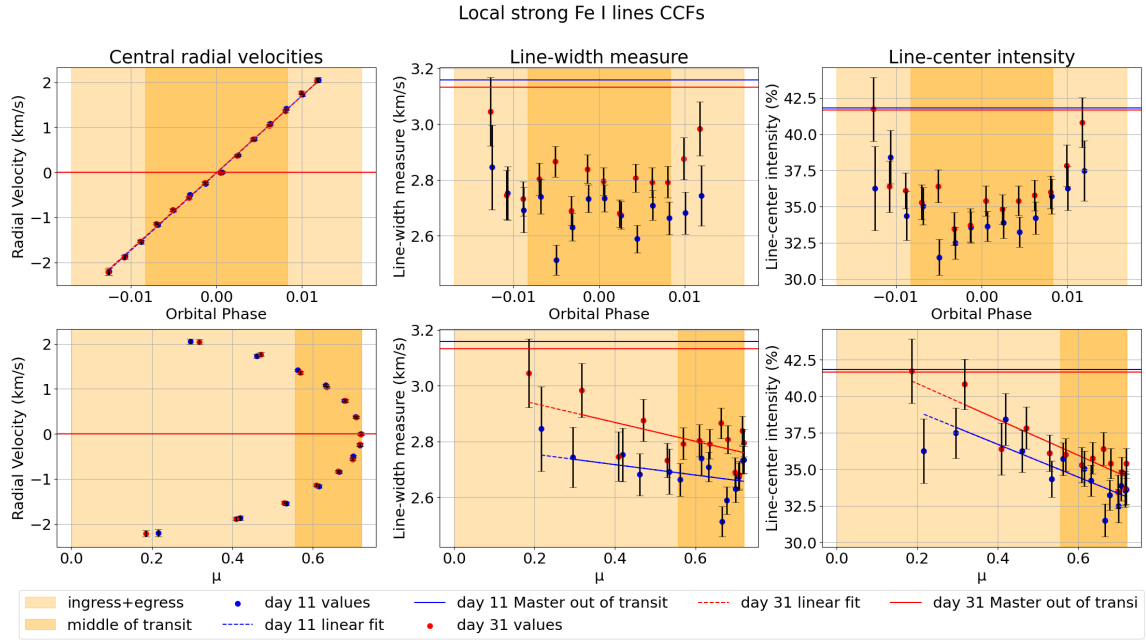


FIGURE 3.9: The sets of data extracted from the subtracted strong Fe I lines CCFs through the use of function fitting, displayed in the same manner as Figure 3.3.

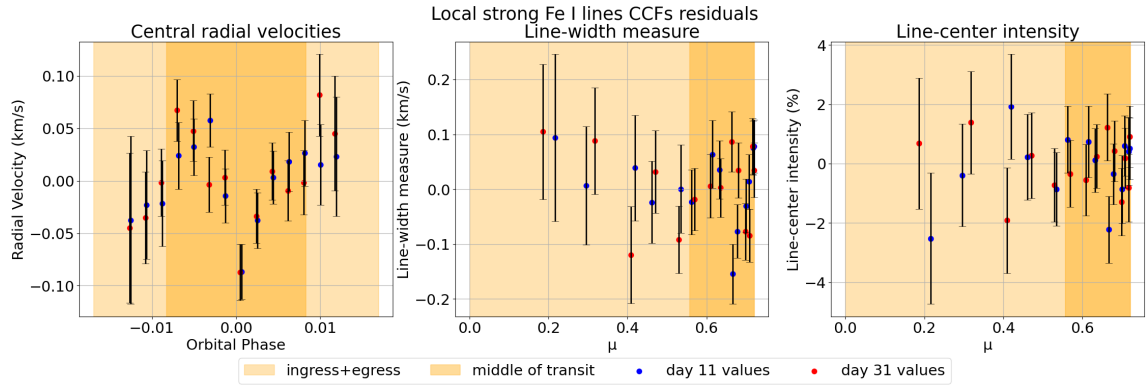


FIGURE 3.10: The residuals of each of each of the linear fits done in 3.9, displayed in the same manner as Figure 3.4.

respective linear regressions. Starting with the weak iron CCFs which has multiple statistical quantities compiled in Table 3.2, plus the Spearman coefficients and p-values distributions, made from 10 000 samples of the sets of data, shown in Figure 3.11.

The central radial velocities are the closest to linear behavior with excellent linear fits whose R^2 is almost one for both days, and consequently the Pearson and Spearman coefficients are essentially one. The χ^2 shows a clear preference to the linear model and so does the p-value which is basically zero. Looking at the distributions taken from sampling taking uncertainties into account we see the same story with Spearman coefficients concentrated around one and p-values concentrated on zero.

Line-widths show weaker results, for both days the fits are not great with R^2 smaller than 0.5. The Pearson and Spearman coefficients are relatively small in absolute value. The χ^2 show only a slight support for the linear model over the constant and the p-values are relatively high. Being more specific the results for the 31st of August are the most inconclusive being too noisy with a log p-value larger than -1.30 both from the table and seen in the distributions, where taking into account uncertainties it shows that no conclusions may be taken for this day. For the 11th of August the tests seem to have a clearer result with a p-value inside the specified limit.

The results for the line-center intensity see also a difference between days of observation. The results for the first day have a decent R^2 of 0.704, a large Pearson and Spearman coefficients, the χ^2 gives a preference for the linear model and a small p-value both from the data itself as well as the sampled distribution, which allows to consider a non-constant model. In contrast the set of data for the second day is inconclusive with a very high log p-value of -0.15 .

Coefficients	Central radial velocity		line-width measure		line-center intensity	
	day 11	day 31	day 11	day 31	day 11	day 33
R^2	0.996	0.998	0.427	0.192	0.704	-0.018
Pearson	0.998	0.999	-0.654	0.472	-0.839	-0.026
χ^2 (linear fit)	72	31	58	37	38	31
χ^2 (constant mean)	16898	13220	99	45	116	31
Spearman	1.00	1.00	-0.609	0.345	-0.684	0.112
log(P-value)	j-63	j-63	-1.68	-0.644	-2.15	-0.153
Slope (m)	167 ± 3	169 ± 2	-0.7 ± 0.3	0.3 ± 0.3	-11 ± 3	-1 ± 3
Interrcept (b)	-0.05 ± 0.02	0.02 ± 0.01	2.5 ± 0.2	1.9 ± 0.2	74 ± 2	68 ± 2

TABLE 3.2: Table containing for each of the 3 sets of data of the local weak iron lines CCFs and their respective fits for each day of observation the following quantities: the R^2 of the linear regression, the Pearson correlation coefficient, the χ^2 of the linear regression, the χ^2 of a horizontal line whose height is the mean of the specific data set, the Spearman correlation coefficient and the base 10 logarithm of the p-value associated to it. The linear fit parameters are also shown as m and b.

Going to the strong iron CCFs these have the statistical tests of their linear regressions in Table 3.3, with the sampled distributions of Spearman coefficients and p-values in Figure 3.12.

Again we see similar results for the central radial velocities for both days with a high quality linear regression with high R^2 values, the χ^2 show a preference for the linear model, the Pearson and Spearman are essentially one and the p-value basically zero. With the description of the sampled distributions being the same as the other CCFs.

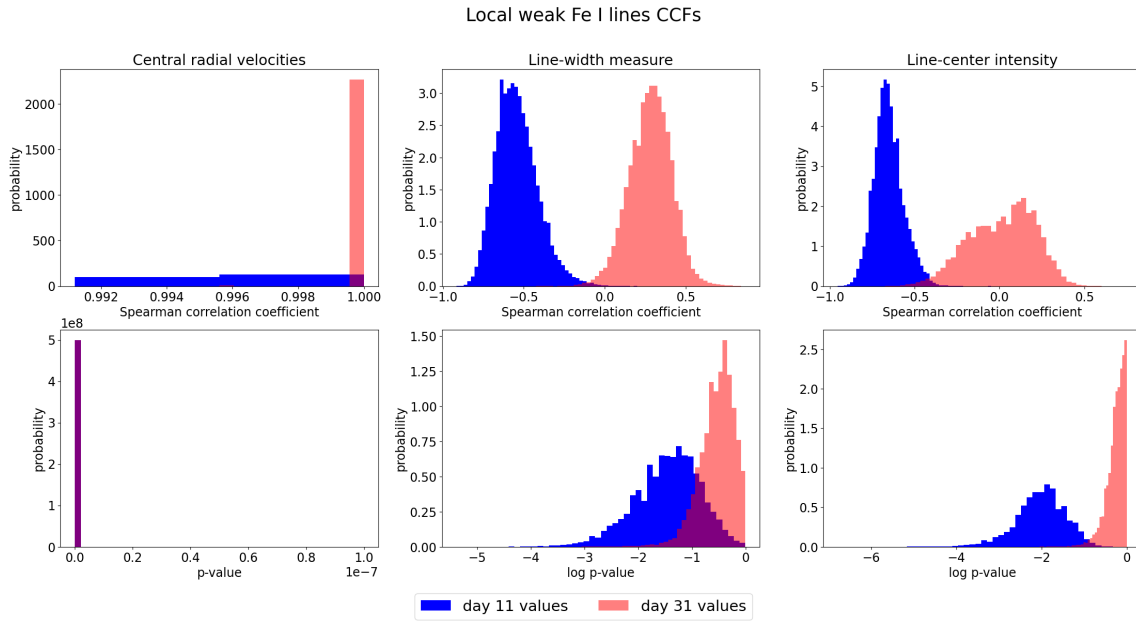


FIGURE 3.11: Plots containing distributions of the spearman coefficient and p-values calculated from 10000 samplings of the data used for each of the 3 linear fits, displayed in the same manner as Figure 3.5.

Although visually looking at the graphs would suggest better results compared to the weak iron CCFs for the line-widths the statistical quantities showcase similar results. The R^2 are small with the second day having improved compared to the weak iron lines, the χ^2 shows a slightly preference of the linear model over the constant mean, the Pearson coefficient is larger than 0.5 but the Spearman is smaller. The p-value is also large enough to reconsider the validity of the result, in the same manner the distributions of the p-values get extremely high when factoring uncertainties which lead to not conclude much beyond the constant model.

At last for the CCFs we look at the line-center intensities whose results have improved compared to the weak iron case especially for the 31st of August. They present high R^2 , Pearson and Spearman coefficients with the χ^2 showing a bigger preference for the linear model. Finally, the distributions of the Spearman coefficient are centered around non zero values with the p-values distribution being concentrated around zero for both days.

Coefficients	Central radial velocity		line-width measure		line-center intensity	
	day 11	day 31	day 11	day 31	day 11	day 33
R^2	0.999	0.998	0.293	0.471	0.611	0.834
Pearson	1.000	0.999	-0.587	-0.706	-0.798	-0.915
χ^2 (linear fit)	22	27	18	17	8	7
χ^2 (constant mean)	14348	15621	21	26	24	31
Spearman	1.00	1.00	-0.473	-0.393	-0.833	-0.820
log(P-value)	-63	-63	-1.06	-0.785	-3.66	-3.48
Slope (m)	172 ± 2	171 ± 2	-0.2 ± 0.2	-0.3 ± 0.2	-11 ± 2	-12 ± 2
Intercept (b)	-0.02 ± 0.01	0.00 ± 0.01	2.8 ± 0.1	3.0 ± 0.1	41 ± 2	43 ± 1

TABLE 3.3: Table containing for each of the 3 sets of data of the local strong iron lines CCFs and their respective fits for each day of observation the following quantities: the R^2 of the linear regression, the Pearson correlation coefficient, the χ^2 of the linear regression, the χ^2 of a horizontal line whose height is the mean of the specific data set, the Spearman correlation coefficient and the base 10 logarithm of the p-value associated to it. The linear fit parameters are also shown as m and b.

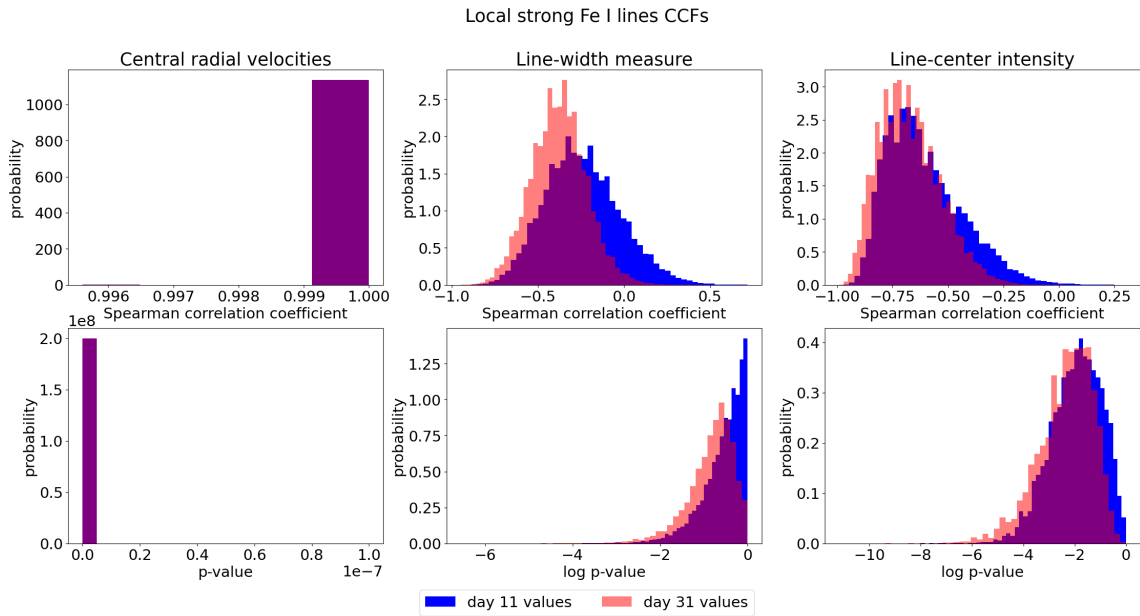


FIGURE 3.12: Plots containing distributions of the spearman coefficient and p-values calculated from 10000 samplings of the data used for each of the 3 linear fits, displayed in the same manner as Figure 3.5.

3.3 Sodium Doublet lines

With the results of the CCFs delivered, we now present the first results for the reconstruction of individual lines using the Doppler shadow method. Starting with the sodium doublet lines chosen due to their strength which allows measurable differences. Obviously as we are working with single lines the results are bound to be noisier and less accurate compared to the CCFs, nevertheless it highlights the potential in possible future applications. The profiles of the reconstructed sodium lines are given in Figure 2.24.

Just like the CCFs we used a modified Gaussian function in Eq. 2.2 from which we used its parameters to extract the central radial velocity, line-width and line-center intensity. The R^2 values of the profile fitting are present in Figure 3.13, from which for day 11/08/2021 we remove the first three and last two observations, while for the other day we remove the first and last three observations.

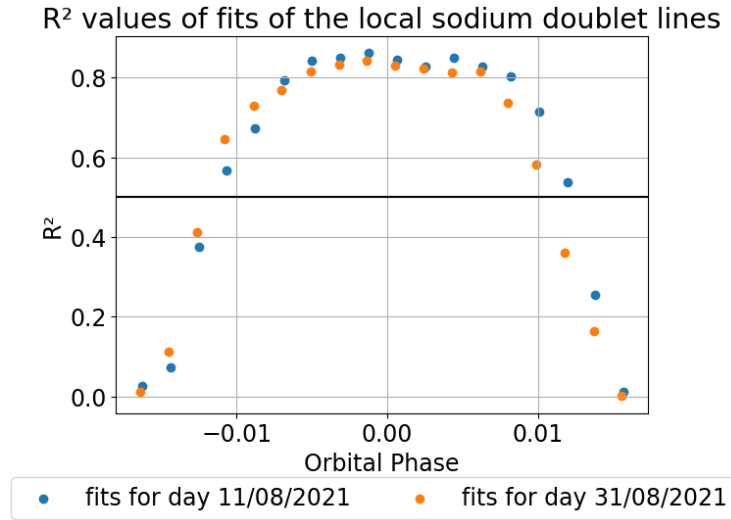


FIGURE 3.13: The R^2 values for each of the fits of the local sodium doublet lines in function of their respective orbital phase.

Starting with the Sodium D1 line centered at 589.6 nm in a vacuum or 589.592424 nm in the air at rest (Table 2.2), the extracted quantities are present in Figure 3.14 together with the respective linear fits used as a first approximation, whose residuals are located in Figure 3.15.

The central radial velocities despite roughly following a linear behavior seem to have some deviation from it, with the residuals showing a clear curvature, additionally a slight difference between the two observed transits suggests that additional effects on the radial velocity related to stellar activity may be present. The slope and intercept values of the linear fit for both days are $m = 133 \pm 15$ km/s, $b = 0.35 \pm 0.09$ km/s for the 11th of August 2021 and $m = 179 \pm 24$ km/s, $b = -0.4 \pm 0.1$ km/s for the 31st of August 2021.

The line-width measure in contrast to the CCFs show an increase with μ instead of a decrease for both days of observation, meaning that the lines are broadening as they near the center of the stellar disk. The parameters of the linear fit are $m = 0.20 \pm 0.08$ Å, $b = 0.17 \pm 0.05$ Å for the 11th of August 2021 and $m = 0.4 \pm 0.1$ Å, $b = 0.06 \pm 0.07$ Å for the 31st of August 2021.

The line-center intensity results are very uncertain and inconclusive with both days being slightly different, with the linear fits being well inside the error bars. The parameters of the linear fit are $m = 2 \pm 10$, $b = -1 \pm 6$ for the 11th of August 2021 and $m = 18 \pm 11$, $b = -12 \pm 7$ for the 31st of August 2021.

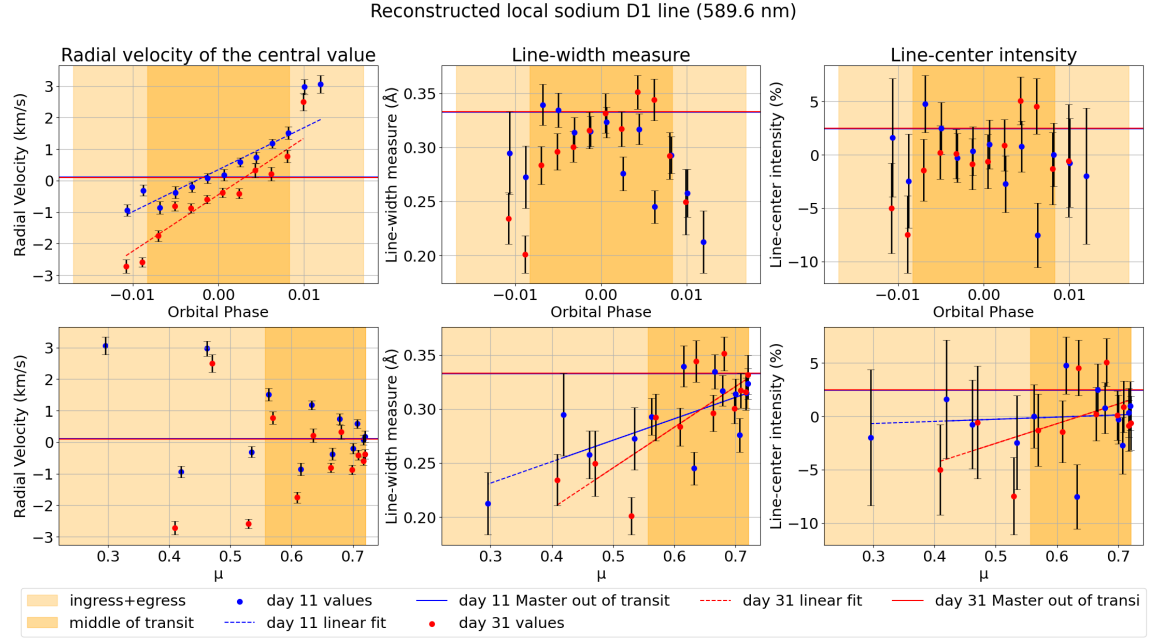


FIGURE 3.14: The sets of data extracted from the reconstructed local sodium D1 line through the use of function fitting, displayed in the same manner as Figure 3.3.

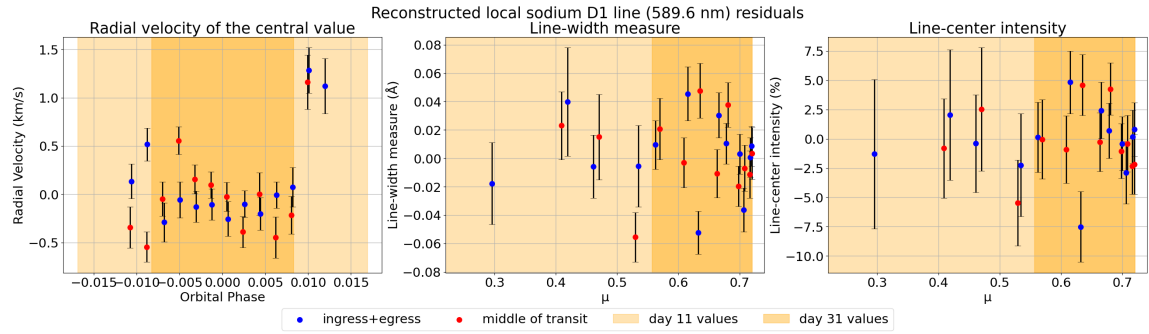


FIGURE 3.15: The residuals of each of each of the linear fits done in 3.14, displayed in the same manner as Figure 3.4.

For the sodium D2 line we got very similar results, these can be seen in Figure 3.16 with their linear fits whose residuals are in Figure 3.17. This line has a rest center wavelength of 589.0 nm in a vacuum or 588.995095 in air (Table 2.2).

The central radial velocity like for the D1 line mostly follows a linear trend, although with a noticeable curve in its residuals with a slight difference between the two days. The

parameters of the linear fit are $m = 188 \pm 26$ km/s, $b = -0.2 \pm 0.1$ km/s for the 11th of August 2021 and $m = 168 \pm 21$ km/s, $b = -0.7 \pm 0.1$ km/s for the 31st of August 2021.

The line-width measure just like the other line presents a growing trend with μ , with a different slope for the two observed transits. The parameters of the linear fit are $m = 0.20 \pm 0.09$ Å, $b = 0.33 \pm 0.06$ Å for the 11th of August 2021 and $m = 0.5 \pm 0.1$ Å, $b = 0.09 \pm 0.08$ Å for the 31st of August 2021.

The line-center intensity again is the most uncertain and inconclusive with large error bars that cover the linear fits. This is reflected in the uncertainties of the slope values from the fits. The parameters of the linear fit are $m = -7 \pm 11$, $b = 7 \pm 7$ for the 11th of August 2021 and $m = 12 \pm 6$, $b = -4 \pm 4$ for the 31st of August 2021.

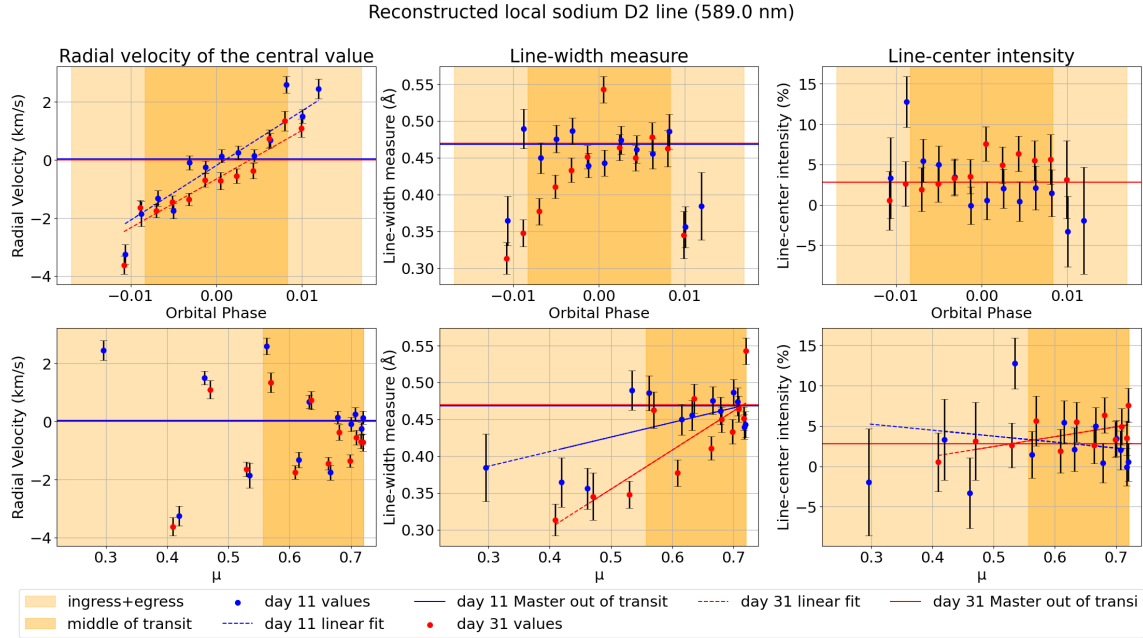


FIGURE 3.16: The sets of data extracted from the reconstructed local sodium D2 line through the use of function fitting, displayed in the same manner as Figure 3.3.

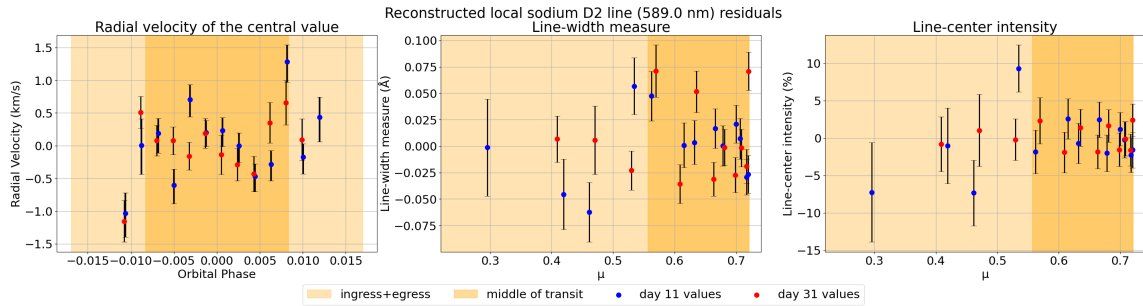


FIGURE 3.17: The residuals of each of each of the linear fits done in 3.16, displayed in the same manner as Figure 3.4.

Now we look at the statistical tests done for the sets of data used for the linear regressions. We will look at both sodium lines. The sodium D1 line with the many statistical quantities calculated compiled on Table 3.4. With the histograms containing the distributions of Spearman correlation coefficients and p-values calculated using 10 000 Gaussian samples of the data sets with the points as the mean and uncertainties as standard deviations in Figure 3.18. For the sodium D2 line the quantities are presented in Table 3.5 and histograms with the Spearman correlation coefficients and p-values are present in Figure 3.19.

Looking at the statistical quantities for the central radial velocities we first notice that the fittings have high R^2 which shows the quality of the fit, for the data itself the Pearson and Spearman coefficients have values of over 0.9 showcasing the high linearity and correlation of the data, the χ^2 gives a clear preference for the linear model and a p-value of almost zero. Looking at the histograms which take into account uncertainties, we notice that the Spearman coefficients are concentrated around non-zero values, while the base 10 logarithm p-values are concentrated on values much lower than -1.3 , which guarantees for sure the linear growing behavior of the data. Although slight differences in slope are present, but these can be impacted by extreme data points near the limb.

The line-widths offer decent linear regressions with the data of both lines of the first day being a bit more uncertain than the data of the second day. The R^2 of the second day is not very large at just over 0.6 while the first day is only around 0.45. The Pearson coefficients are relative large, while the Spearman coefficient of the first day is smaller than for the second day, due to the smaller slope of the first one. The χ^2 shows a preference for the linear model for both transits and the p-values are relatively small with the first day going over the limit of what may be considered statistically relevant. Looking at the distributions of Spearman values we see that both are centered around non-zero values with the distribution of the 11th of August being wider, for the log p-values we see that the second day is more concentrated near smaller values bellow -1.3 , while the first day covers some p-values that go over our chosen limit, as such only the data of the 31st of August should be considered as having a growing trend with μ .

For the line-center intensities the results are very imprecise so not much is to be said other than no noticeable difference over the constant mean model is observed. Both observations have low R^2 for their linear regressions, skipping over the Pearson and Spearman

coefficients taken directly from the data the χ^2 does not show a preference for the linear model. Looking at the distributions of the p-values obtained from sampling we notice that they are very wide with the first day being almost uniform in addition to the p-values being extremely high, so no behavior can be detected.

Coefficients	Central radial velocity		line-width measure		line-center intensity	
	day 11	day 31	day 11	day 31	day 11	day 33
R^2	0.830	0.890	0.453	0.616	-0.001	0.318
Pearson	0.938	0.950	0.677	0.791	0.093	0.567
χ^2 (linear fit)	62	60	30	27	12	11
χ^2 (constant mean)	435	526	48	65	13	16
Spearman	0.984	0.979	0.533	0.762	0.115	0.469
log(P-value)	-8.79	-7.51	-1.22	-2.40	-0.150	-0.905
Slope (m)	133 ± 15	179 ± 24	0.20 ± 0.08	0.4 ± 0.1	2 ± 10	18 ± 11
Intercept (b)	0.35 ± 0.09	-0.4 ± 0.1	0.17 ± 0.05	0.06 ± 0.07	-1 ± 6	-12 ± 7

TABLE 3.4: Table containing for each of the 3 sets of data of the local sodium D1 lines and their respective fits for each day of observation the following quantities: the R^2 of the linear regression, the Pearson correlation coefficient, the χ^2 of the linear regression, the χ^2 of a horizontal line whose height is the mean of the specific data set, the Spearman correlation coefficient and the base 10 logarithm of the p-value associated to it. The linear fit parameters are also shown as m and b.

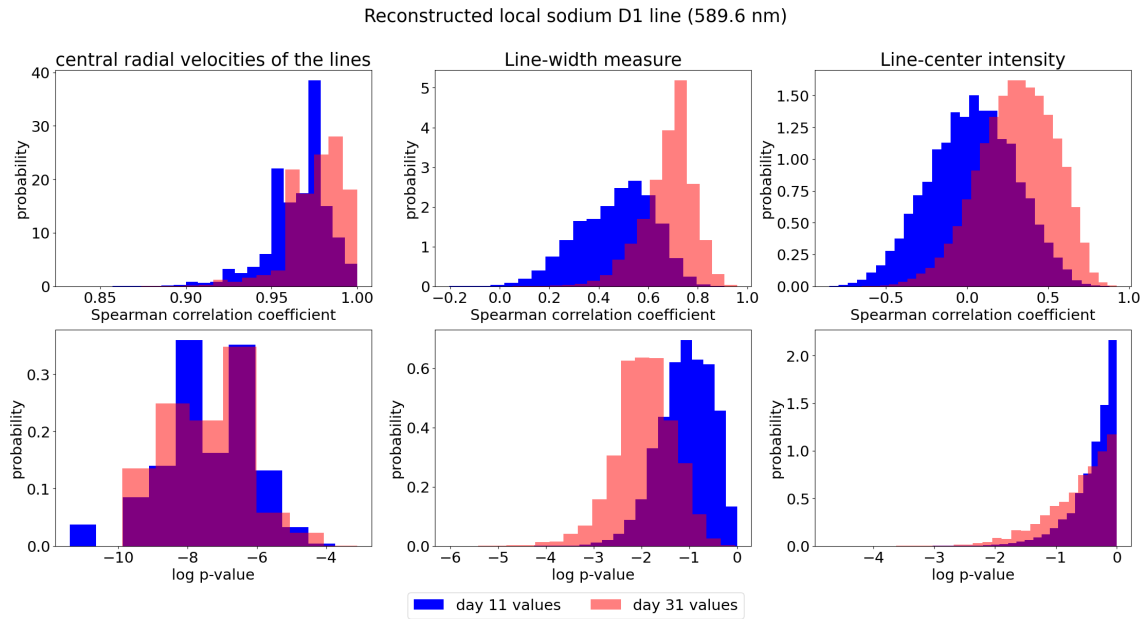


FIGURE 3.18: Plots containing distributions of the spearman coefficient and p-values calculated from 10000 samplings of the data used for each of the 3 linear fits, displayed in the same manner as Figure 3.5.

Coefficients	Central radial velocity		line-width measure		line-center intensity	
	day 11	day 31	day 11	day 31	day 11	day 33
R^2	0.879	0.881	0.464	0.655	-0.160	0.351
Pearson	0.947	0.947	0.693	0.814	0.142	0.594
χ^2 (linear fit)	54	28	23	44	18	5
χ^2 (constant mean)	374	232	36	120	18	7
Spearman	0.967	0.979	0.269	0.741	-0.011	0.615
log(P-value)	-7.15	-7.51	-0.427	-2.24	-0.013	-1.48
Slope (m)	188 ± 26	168 ± 21	0.20 ± 0.09	0.5 ± 0.1	-7 ± 11	12 ± 6
Intercept (b)	-0.2 ± 0.1	-0.7 ± 0.1	0.33 ± 0.06	0.09 ± 0.08	7 ± 7	-4 ± 4

TABLE 3.5: Table containing for each of the 3 sets of data of the local sodium D2 lines and their respective fits for each day of observation the following quantities: the R^2 of the linear regression, the Pearson correlation coefficient, the χ^2 of the linear regression, the χ^2 of a horizontal line whose height is the mean of the specific data set, the Spearman correlation coefficient and the base 10 logarithm of the p-value associated to it. The linear fit parameters are also shown as m and b.

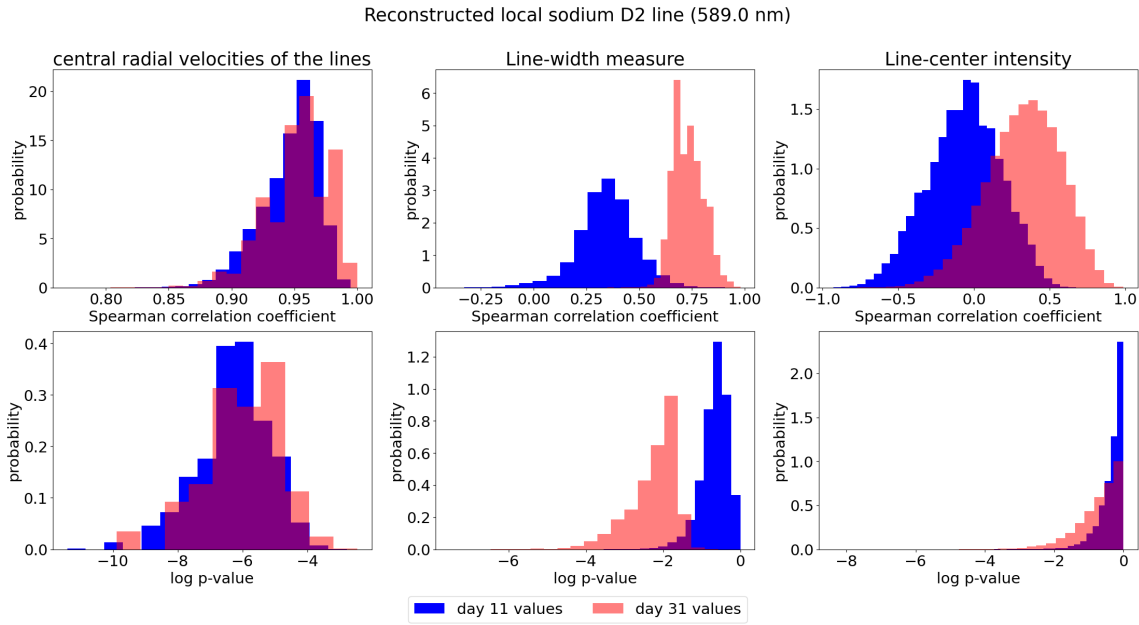


FIGURE 3.19: Plots containing distributions of the spearman coefficient and p-values calculated from 10000 samplings of the data used for each of the 3 linear fits, displayed in the same manner as Figure 3.5.

3.4 H α line

We finally present the results of the H α line which was our other chosen spectral feature for studying. This line is a well-known stellar activity indicator (Cincunegui, C. et al. 2007) and is very broad. For this line we attempted multiple profile functions like the Voigt profile and a regular Gaussian which did not show much of a difference between each other. However, applying the modified Gaussian which was used for all other data sets gave slightly different results and was ultimately chosen to maintain a consistent

methodology. The profiles of the reconstructed $H\alpha$ lines are given in Figure 2.26.

We performed profile fitting with a modified Gaussian and the R^2 of the fittings can be seen in Figure 3.20. Just like the sodium lines and unlike the CCFs the imprecision of the fits is greater, especially near the limb due to us only working with a single line. From the R^2 we discarded for the data of both dates the first four and last three fits.

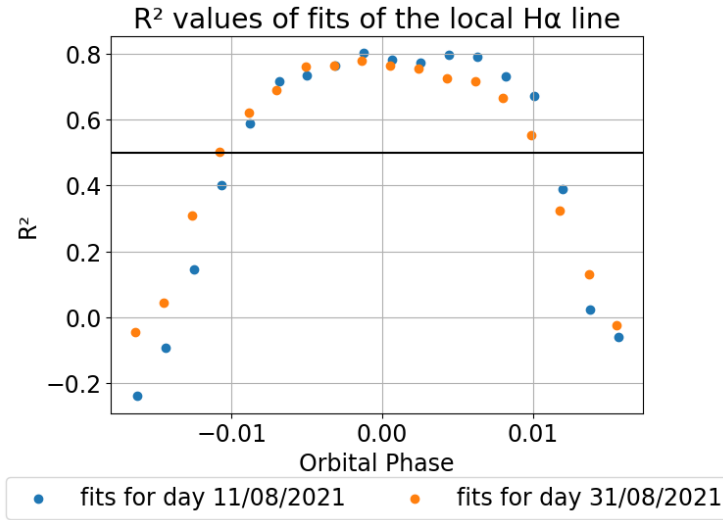


FIGURE 3.20: The R^2 values for each of the fits of the local $H\alpha$ lines in function of their respective orbital phase.

In this section we only look at a single spectral feature, however this line seems to be strongly affected by stellar activity and this is reflected in the results of the extracted parameters from the profile fitting. Again the quantities we look at are the central radial velocities, the line widths and the line center intensities which are provided in Figure 3.21 with their linear regressions as a first approximation and whose residuals are in Figure 3.22.

Starting with the central radial velocities, these unlike all the other previous results seem to deviate the furthest from the linear approximation with the points of the 31st of August deviating the furthest. This is also reflected in the linear fits to the data points and their uncertainties as well as their residuals with the second day residuals presenting a curvature. The linear regressions have values $m = 157 \pm 26$ km/s, $b = 0.0 \pm 0.1$ km/s for the 11th of August 2021 and $m = 76 \pm 40$ km/s, $b = 0.0 \pm 0.2$ km/s for the 31st of August 2021.

The line-widths behavior is more unclear with the linear fits themselves overall being much more uncertain. The linear regressions have values $m = 0.0 \pm 0.1$ Å, $b = 0.52 \pm 0.06$

$\text{\AA}$ for the 11th of August 2021 and $m = 0.4 \pm 0.1 \text{ \AA}$, $b = 0.2 \pm 0.1 \text{ \AA}$ for the 31st of August 2021.

For the line-center intensities a more erratic behavior compared to the sodium lines is visible and the linear regressions do not seem to provide a proper fit. The linear regressions have values $m = -9 \pm 18$, $b = 37 \pm 12$ for the 11th of August 2021 and $m = 22 \pm 11$, $b = 20 \pm 7$ for the 31st of August 2021.

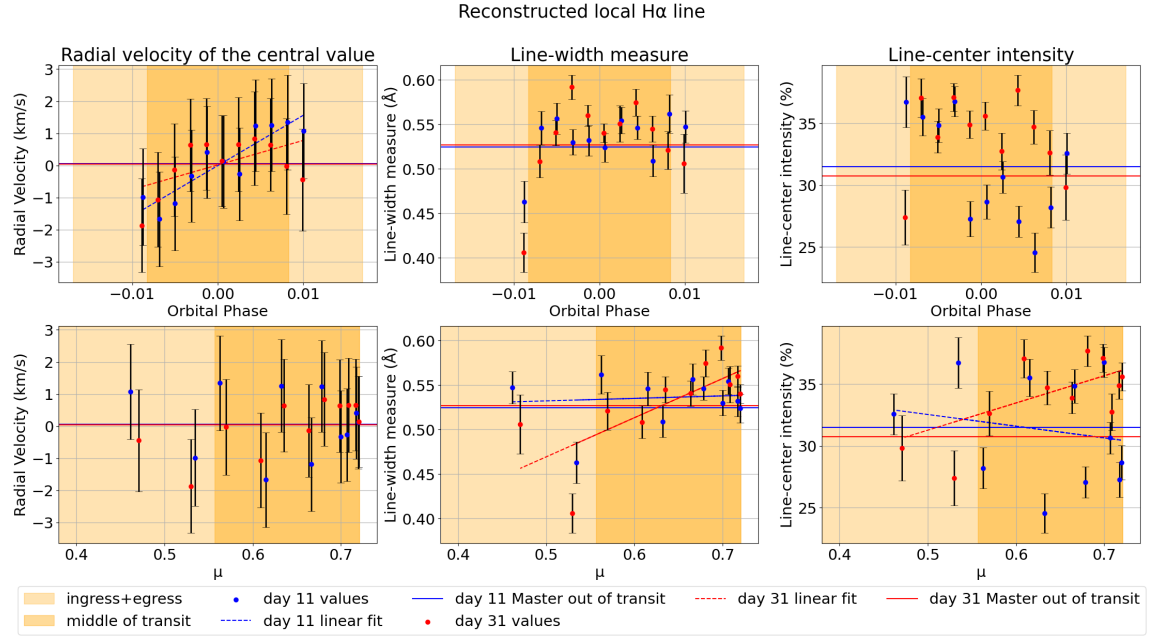


FIGURE 3.21: The sets of data extracted from the reconstructed H α line through the use of function fitting, displayed in the same manner as Figure 3.3.

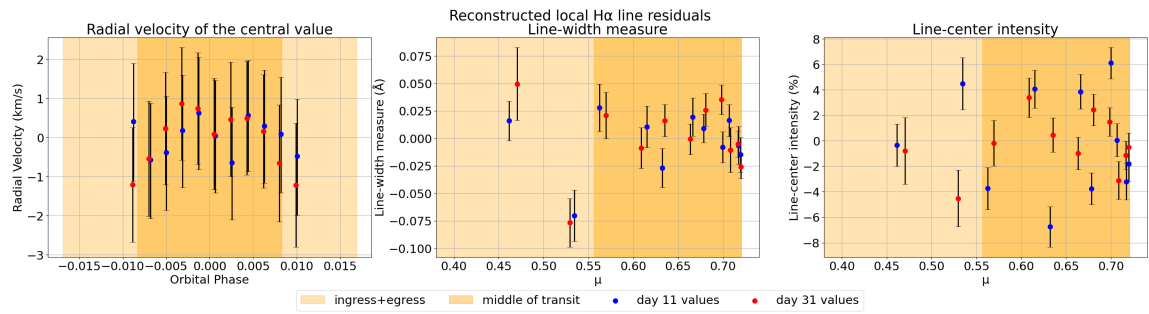


FIGURE 3.22: The residuals of each of each of the linear fits done in 3.21, displayed in the same manner as Figure 3.4.

We now look at the statistical tests done to the data points of the parameters that describe the profiles of the local H α lines, these are presented in Table 3.6 together with the distributions of Spearman correlation coefficients and p-values shown in Figure 3.23.

The central radial velocities are much noisier compared to the previously studied spectral features, additionally the two transits present different results with the second being more uncertain. The first day presents results similar to other lines with a good fit with its high R^2 , a high Pearson and Spearman coefficients showcasing great linearity and a well-defined growing trend, the χ^2 show a clear preference for the linear model and a low p-value that supports the non-constant behavior. However, this falls apart when considering the distributions which take into account uncertainties and end up covering higher p-values resulting in us discarding this correlation. The second day performs worse, which is clear when comparing the distributions, with the log p-value distribution of the first day being concentrated on larger values compared to the second day.

For the line-widths the results are even more uncertain with the second day being of better quality. The first presents a very poor R^2 and its high p-value are enough to conclude that no difference from a constant model is observed (same for the distribution). For the second day despite having better R^2 , larger Spearman and Pearson coefficients and low p-value the χ^2 does not show a preference for the linear model. Even looking at the distribution of p-values its width edges on the limits of what may be considered satisfactory, so the results are inconclusive.

At last line-center intensities are very noisy and inconclusive for both observed transits with the second transit being slightly better, nevertheless when looking at the histograms of p-values of the samples which take into account the uncertainties of the data points, the distributions are very wide and thus cover a significant portion of high p-values and as such no meaningful take away other than a constant model can be taken.

Coefficients	Central radial velocity		line-width measure		line-center intensity	
	day 11	day 31	day 11	day 31	day 11	day 33
R^2	0.813	0.241	0.013	0.510	0.058	0.473
Pearson	0.901	0.498	0.169	0.716	-0.252	0.708
χ^2 (linear fit)	1	2	18	33	84	21
χ^2 (constant mean)	5	3	19	71	87	36
Spearman	0.882	0.391	-0.127	0.709	-0.245	0.582
log(P-value)	-3.48	-0.630	-0.149	-1.84	-0.331	-1.22
Slope (m)	157 ± 26	76 ± 40	0.0 ± 0.1	0.4 ± 0.1	-9 ± 18	22 ± 11
Intercept (b)	0.0 ± 0.1	0.0 ± 0.2	0.52 ± 0.06	0.2 ± 0.1	37 ± 12	20 ± 7

TABLE 3.6: Table containing for each of the 3 sets of data of the local H α lines and their respective fits for each day of observation the following quantities: the R^2 of the linear regression, the Pearson correlation coefficient, the χ^2 of the linear regression, the χ^2 of a horizontal line whose height is the mean of the specific data set, the Spearman correlation coefficient and the base 10 logarithm of the p-value associated to it. The linear fit parameters are also shown as m and b.

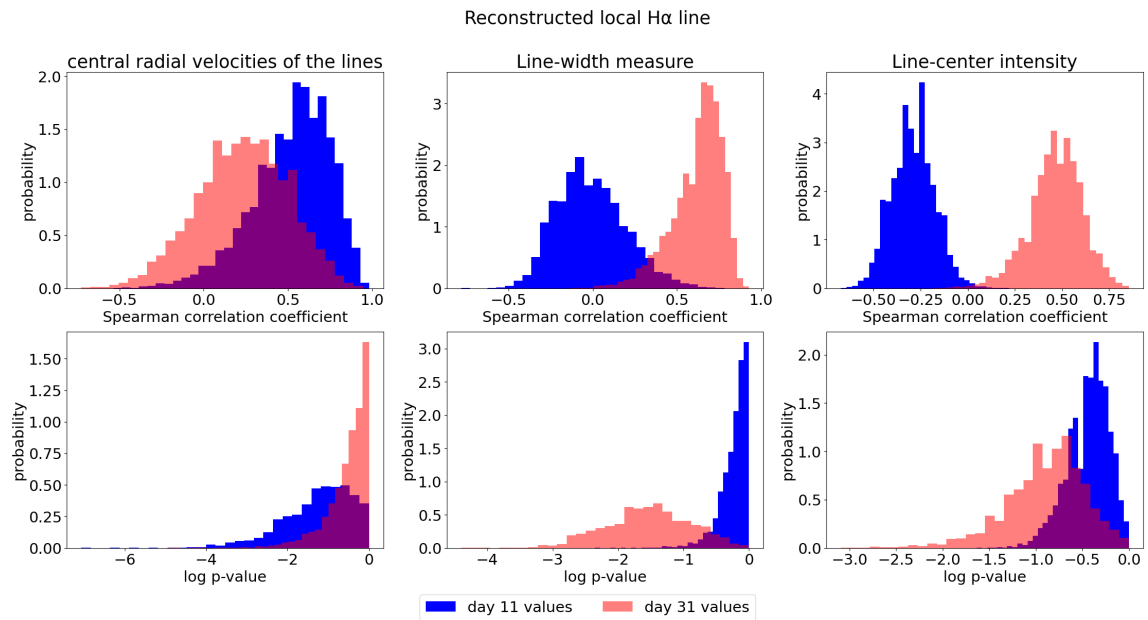


FIGURE 3.23: Plots containing distributions of the spearman coefficient and p-values calculated from 10000 samplings of the data used for each of the 3 linear fits, displayed in the same manner as Figure 3.5.

Chapter 4

Discussion of the data - Center to limb variations

In this chapter, we analyze and discuss multiple sets of results. We will compare these results to models from other sources and provide physical interpretations.

Before discussing specific sets of data, we will mention an anomaly identified in the work of [Cristo, E. et al. \(2024\)](#), which uses the same sets of data. This work focused on the study of the Rossiter-McLaughlin effect (RM), where a computational model was used to fit the RM curve for the same white light CCFs, shown in Figure C.1 of [Cristo, E. et al. \(2024\)](#), which is similar to the right side of Figure 2.10. Upon subtracting the fitted model from the RM curve of the data one anomaly was noticed on the second transit on the 31st of August, a deviation in radial velocities of about 3 m/s right after the ingress was detected, the origin of this anomaly is not confirmed. One possibility, although not the only explanation, as no firm conclusion was reached, was that the planet may have crossed an active region of the star, which was shown to be compatible with a spot of size 0.1% of the stellar disk and a temperature contrast of 660 K.

Looking at our data, one can see the effects of the anomaly most evidently on the white light CCFs which were where the anomaly was first noticed and also the ones with the highest S/N. The effect can be seen in Figure 3.3 for the line-widths versus orbital phases plot and while the values of the second day are overall larger than the ones of the first denoting wider profiles, a noticeable increase in width is seen shortly after ingress on the location of the anomaly. Another unusual behavior is seen in the line-center intensities versus orbital phases plot, where at the ingress the profiles are deeper for the second day

but after the ingress and the rest of transit the profiles of the first day are deeper this could also be due to the mentioned anomaly.

Beyond the white light CCFs it is hard to notice more evident anomalies and one would need to compare with models of the profiles of the selected spectral features in order to evaluate the effect this anomaly has on more specific lines.

4.1 The white light CCF

We now look at each set of data at a time starting with the white light CCFs. As explained previously, these CCFs were obtained from the cross correlation of the spectra with a class K2 mask to give a general representative view of the stellar spectrum as a whole. As such it gives the typical behavior of most lines in the interval that our spectrum covers.

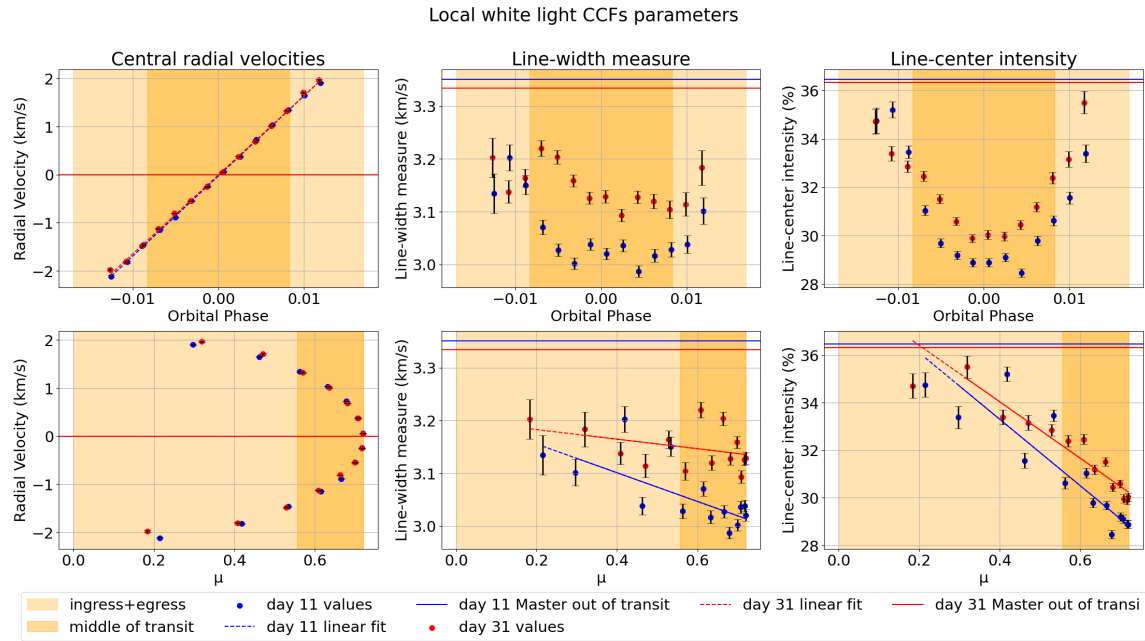


FIGURE 4.1: Figure 3.3 for ease of view of the results of the white light CCFs.

In the previous chapter we presented three parameters that describe the CCF profiles and we now give a more physical discussion, Figure 4.1 is a copy of Figure 3.3 for convenience of reading. Starting with the central radial velocities, the white light CCFs gave a very clear result. In this parameter we observe the radial velocity of a given segment of the stellar surface and the effect of stellar rotation is clearly visible. At the start of the transit the planet covers regions that are blue shifted with negative RVs, as the transit progresses we move to the center of the star where the effect of rotation on our line of sight

cancels out and where the RV crosses a value of zero only for it to start moving towards redshifted regions with positive RVs.

This behavior is nearly linear with our linear fits performing very well, but this is an oversimplification due to two factors. First we know that the orbit of the exoplanet is not exactly perpendicular to the axis of stellar rotation being at $85.465^{+0.020}_{-0.021}$ (Table 2.2), causing it to cover different stellar latitudes. Second the star does not rotate like a rigid body. This was demonstrated by Cegla et al. (2016) and reproduced by Cristo, E. et al. (2024) from data based on the RM effect, with the last reference giving a differential rotation coefficient of $\alpha_B = 0.29^{+0.12}_{-0.13}$. A more accurate value was obtained by Fares et al. (2010) using a different technique of 0.278 ± 0.093 .

This results in more complex behavior than a simple linear relation as not only does the planet cover different latitudes of the star, but the angular rotation velocity varies with latitude. These phenomena are reflected in the residuals in Figure 3.4 where small wiggles are observed and at the start of the transit the linear model underestimates the radial velocities while at the end it overestimates. These curvatures at the limb in the residuals were also reported in Cegla et al. (2016) in Figure 5. In the reference by using a model based on differential rotation the trend of the residuals was removed with only small wiggles remaining. So fitting with a differential rotation model to our data could improve the residuals.

The line-widths we analyzed on the previous chapter gave a trend for the first transit with the lines being narrower at the center of the stellar disk and wider at the limb, with the second transit being more inconclusive. Now with the context of the anomaly observed after ingress at the 31st of August and looking back at Figure 3.3. We observe that outside of the time interval affected by the anomaly and especially at the later part of the transit that the data of the 31st of August resembles that of the 11th.

The general trend is that the typical line profile seems to widen from the center of the stellar transit to the limb. This may be primarily due to the convective blueshift on the stellar surface which shown in Cegla et al. (2018) is capable of widening spectral lines.

Moreover, all in transit local profiles are significantly narrower than the master out-of-transit profile, which represents the spectra of the entire star outside of transit. This is primarily due to the different segments of the stellar surface having different radial velocities and thus when measured all the individual components are overlapped resulting in an integrated spectrum with wider line profiles.

In the previous chapter we looked at the line-center intensities which gave us how deep a profile is in relation to their continuum. For both transits, we observed highly correlated behavior with μ , although a linear approximation may be crude, as the residuals showcase curvature. The slight difference in the days of observation could be due to the anomaly mentioned and possible influence of other stellar activity.

The major conclusion is that the average line profile decreases in depth from the center of the stellar disk towards the limb most likely due to line of sight effects as shown in [Cegla et al. \(2018\)](#).

Furthermore, just like with the line-widths the reason that the master out-of-transit profile has a smaller relative depth is due to the integration of multiple profiles with different radial velocities resulting in both a narrower and shallower CCF.

These results seem to be in agreement with simulations in the literature, although modeling of the radial velocities due to differential geometry would help our analysis.

4.2 Iron CCFs

Next we look at both the weak Fe I lines and strong Fe I lines local CCFs. These sets of lines were chosen to provide a direct comparison with the work of [Dravins et al. \(2018\)](#) with our higher quality dataset. In their work the iron lines were chosen as they form under very high temperatures, higher than any found in the exoplanet. This allows treating the planet and its atmosphere as an opaque disk. Moreover, the iron lines are very numerous and found in the wavelength range of HARPS used in the original work and also covered by ESPRESSO. The iron lines were then divided into two groups based on their intensity.

In Figure 4.2 is compiled the results of [Dravins et al. \(2018\)](#) for their iron CCFs together with our own data points, providing a direct comparison. In their work the data is presented in function of transit epoch which give the order of the points, although different this is essentially proportional to the use of orbital phases.

We start by comparing the line-widths. According to [Cegla et al. \(2018\)](#) the profiles are expected to widen moving from star center to limb, but it is of note that these simulations were done for a star with a surface temperature of 5800 K. In [Dravins et al. \(2018\)](#) simulations were made for three different stellar classes at 3960 K, 5900 K and 6730 K presented in Figure 4.3.

It was then pointed out that although it seems to be the case for Solar type and hotter stars that widths increase from center to limb. It becomes less clear for cooler stars as the

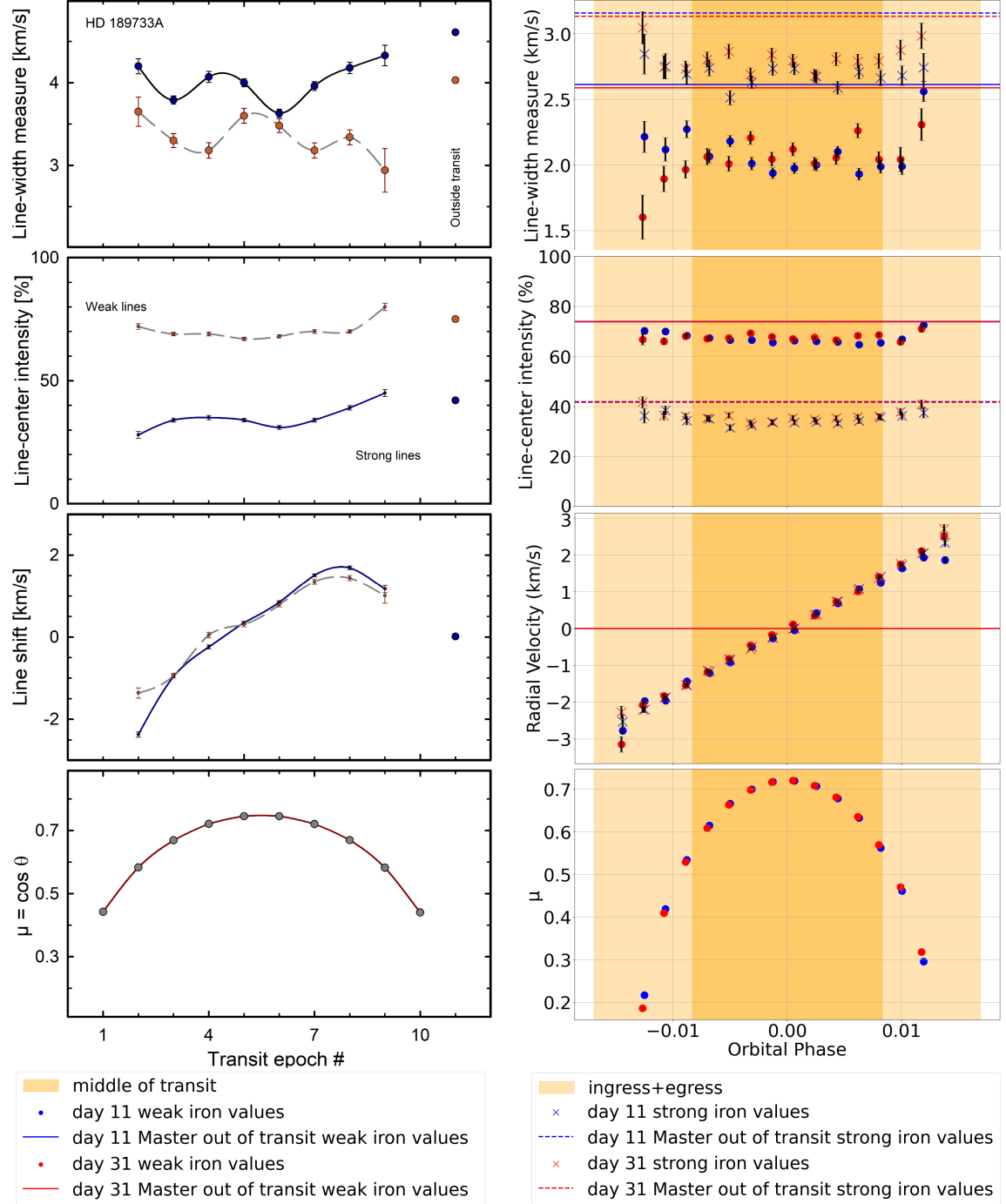


FIGURE 4.2: Direct comparison of the weak iron and strong iron CCFs. Left is Figure 10. of Dravins et al. (2018) while right is our results. The plots from up to bottom contain the line-widths, line-depths, central radial velocities and μ for reference.

contrast of the granulation is smaller, with the simulated 3960 K case having almost no variation. This seems to be the case for their results as no clear trend is visible, for our data this is especially true for the weak iron CCFs as seen in the previous chapter. While for the strong iron there seemed to be a trend, but we decided not consider it significant due to the statistical tests performed. Given that our star with an effective temperature of $4969 \pm 43\text{K}$ (Table 2.2) is between the two simulations of [Dravins et al. \(2018\)](#) with one of them showing a trend in line-widths it raises the possibility that with better data one could detect it. For now, we conclude that no significant trend has been found.

Another detail to point out is that our local CCFs are much narrower compared to the ones in [Dravins et al. \(2018\)](#). This seems to be a side effect of the methodology used. While we worked with our observations individually, in [Dravins et al. \(2018\)](#) their observations were grouped in groups of four and averaged. This resulted in the spatial smearing of the data along the stellar surface, giving then wider profiles. Furthermore, this could also be due to the differences in resolving power, with ESPRESSO having 140 000, while HARPS has 100 000.

The line-center intensities are predicted by [Dravins et al. \(2018\)](#) to decrease from center to limb as the CCFs get less deep with a similar observation being made in [Cegla et al. \(2018\)](#). This behavior does not seem to be clear looking at their plots (Figure 4.2) with strong iron CCFs apparently decreasing in depth along the entire transit. However, from our analysis in the previous chapter we established that the data of the 11th of August for the weak iron CCFs and the data of the strong iron CCFs is in agreement with these models. We then assert that profiles for the iron CCFs decrease in depth from center to limb.

The final parameter is the line-shift or central radial velocities these are the most well understood with their behavior mostly linked by rotation of the stellar surface. It was already predicted from the Rossiter-McLaughlin effect that the radial velocities of the stellar surface would vary in an almost linear fashion from one side to the other. In the data given by [Dravins et al. \(2018\)](#) we do observe an increase in line-shift although deviating significantly from an almost linear behavior, but in contrast our results show an almost perfect linear behavior as one would expect. Part of the difference could be due to the averaging of groups of 4 sets of data done in [Dravins et al. \(2018\)](#), which included regions of slightly different radial velocity which deformed the curve further from a line. In any form these results highlight the quality of our data and what may be possible in the future.

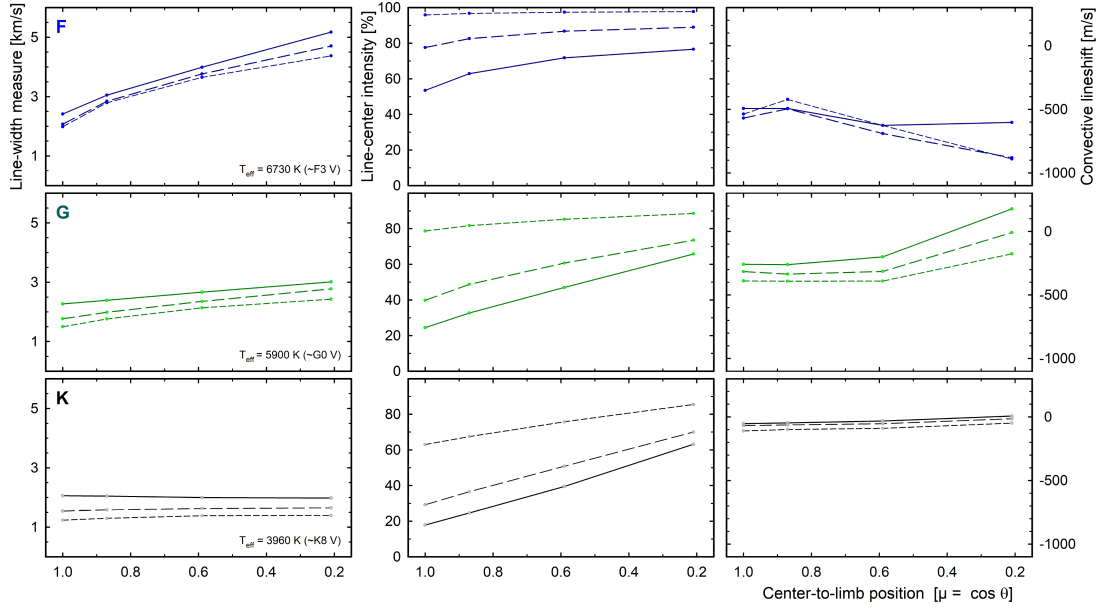


FIGURE 4.3: Figure 11. of [Dravins et al. \(2018\)](#). From their description: Center-to-limb changes of the parameters for differently strong lines computed from the various simulations representing stars of spectral types F, G, and K. Properties of strong lines are shown by solid curves, of intermediate lines by long-dashed, and of weak lines by short-dashed curves. Line-width measures, line-center intensities, and convective line shifts are from symmetric Gaussian fits to the full synthetic profiles and therefore do not resolve any line asymmetries. Convective line shifts are in absolute units but gravitational redshifts are neglected.

Moving beyond the comparisons between plots another matter of discussion is that of convective blueshift. In [Cegla et al. \(2018\)](#) and shown in Figure 4.4 working at a temperature of 5800 K the simulations suggest an increase in blueshift from $\mu = 1$ to $\mu \approx 0.7$ due to convection and then decreasing until around $\mu \approx 0.5$ only to turn into a redshift in the limb. In Figure 4.3 a slight blueshift is given that decreases until the limb for the cooler type with a more abrupt decrease in the limb given for the G class star.

In our work we assumed a linear model for our radial velocities, but as mentioned before this is not accurate due primarily to the slight misalignment of the planet's orbit and differential rotation. These effects need to be taken into account in order to properly retrieve the RVs due to convective blueshift from our data.

Still we can make a basic observation, Figure 4.5 contains the residuals of the linear fits to the radial velocities of the strong iron CCFs (Figure 3.10) which have been put in function of μ . These residuals were intriguing as they presented a clear symmetry along the transit showcasing a possible dependence in μ and a connection to convective blueshift. The closest we get to center of the stellar disk is around $\mu = 0.72$ with a noticeable blueshift, from there it decreases until a redshift around $\mu \approx 0.5$ and then the

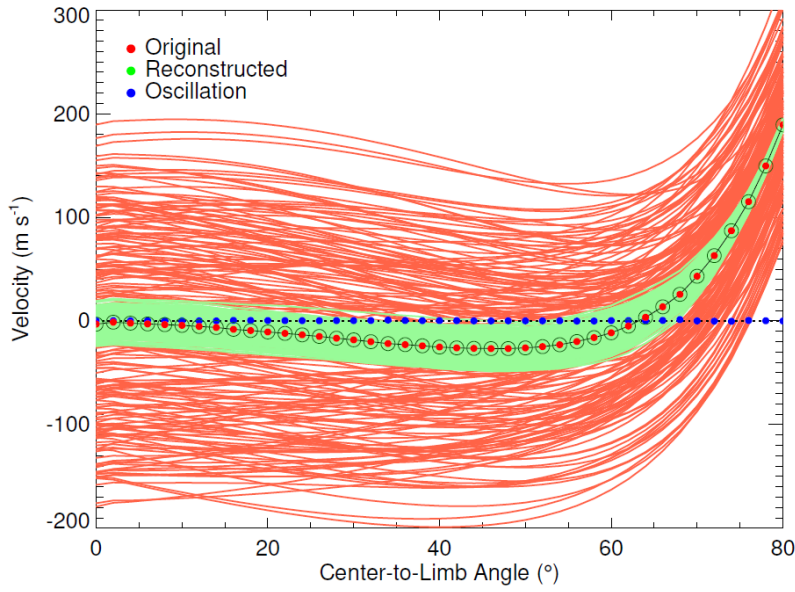


FIGURE 4.4: Figure 2 of [Cegla et al. \(2018\)](#). Total RVs from the snapshots of their simulated profiles (lines) and averages over a time-series (points). Original profile results are in the red points, connected by a solid black line, and the reconstructed granulation-only profiles are in green. Also plotted are the oscillation RVs averaged over the entire series (blue points). Plus, a black dotted line at 0 m/s was added to guide the eye.

blueshift seems to increase again until the limb. This behavior is odd compared to the models in the other works, reminding again that it is possible that the effects of differential rotation are distorting our residuals.

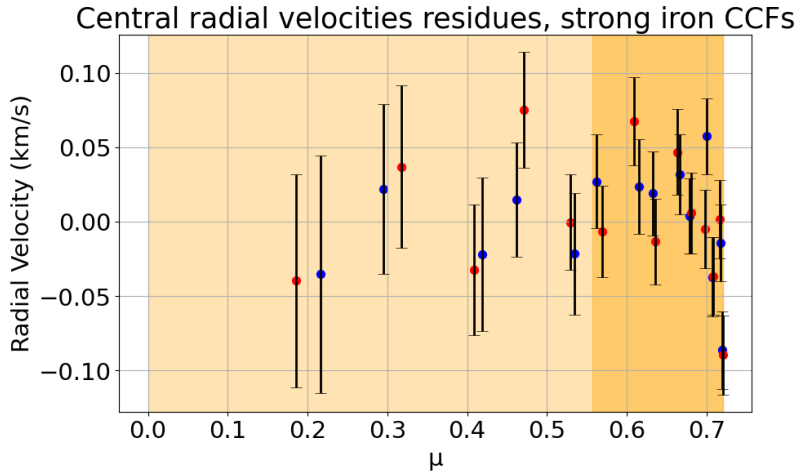


FIGURE 4.5: Plots containing the residuals of the linear fits to the central radial velocities of the strong iron CCFs [3.10](#) in function of μ . The colors follow the same legend as the referenced Figure.

4.3 Sodium Doublet lines

As opposed to the CCFs which were constructed with the goal of replicating and verifying previous results and models, we performed reconstruction of individual spectral lines for the first time. In this section we discuss the results of the reconstructed local sodium doublet lines, but since this is a first time we are unable to compare to other results. We instead turn to the work of [Dethier & Bourrier \(2023\)](#) where simulations of the sodium D2 line were made taking into account stellar rotation and simulated center to limb variations. Their work being done in the context of transmission spectroscopy for HD 209458 b and MASCARA-1 b.

Although the models were only done for the sodium D2 line, the lines of the doublet share a common origin and we found similar results for our data so we will make the discussion for both lines simultaneously.

First we look at the central radial velocities, these are straightforward as they are the result of the lines being Doppler shifted mostly due to the rotational velocity of that section of stellar surface. In the previous chapter we verified as a first order approximation a clear linear trend as expected although much noisier and with higher uncertainties compared to the CCFs as we are working with single lines.

In the residuals of the central RVs there are visible wiggles in an almost sinusoidal shape in Figures 3.15 and 3.17. These are most certain by due to us only using a linear

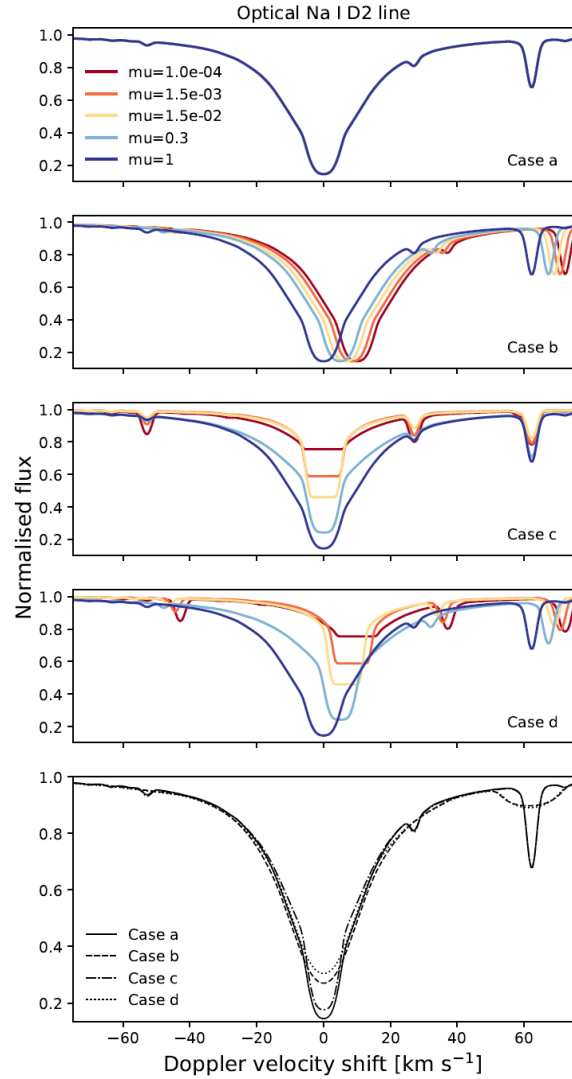


FIGURE 4.6: Figure 2 of [Dethier & Bourrier \(2023\)](#) containing simulated Na I D2 lines at different μ s. The upper plot (case a) represents these lines on a uniform static disk. Case b introduces stellar rotation. Case c the lines are subject to center to limb variations. Case d combines center to limb variations and Doppler shifts due to stellar formation. The bottom plot contains the profiles of the integrated disk for all cases.

model and not accounting for differential rotation. Despite this, some may be due to convective blueshift.

For the line-widths these tell a different story than compared to the CCFs. The data of the 11/08/2021 show a slight increasing trend with μ , but we deemed it not significant in the previous chapter based on the statistical tests performed. However, we determined that for the 31/08/2021 that the data did show a growing trend with μ . This is interesting as it goes against the behavior of the CCFs and from the other models made for iron lines.

Simulations made by [Dethier & Bourrier \(2023\)](#) shed light to this behavior seen in Figure 4.6 (Figure 2 of their work). Moving from center to limb the wings of the sodium lines shrink leaving just the "core" of the line resulting in a smaller line-widths. In addition, in [Cegla et al. \(2018\)](#) where different components of the photosphere are simulated it is pointed that while for granules and non-magnetic intergranular lanes the line-widths increase, the same does not happen for the magnetic intergranular lanes whose spectral features seem to decrease in width, a possible explanation given is that our line of sight becomes inclined to the magnetic flux tubes resulting in a reduction in Zeeman splitting.

Lastly, for the line-center intensities, which give the line depths, our results did not show a meaningful trend as they were too noisy so no conclusions could be taken. Nevertheless, we mention that the simulations in [Dethier & Bourrier \(2023\)](#) and [Cegla et al. \(2018\)](#) predict a decrease in line depth (increase in line-center intensities) with μ .

4.4 H α line

At last we discuss the H α line. This spectral line was selected as another candidate for studying due to being a very distinct feature in the spectrum of the star in the wavelength range of ESPRESSO just like the sodium doublet. Such a prominent feature could then be put through our methodology in order to retrieve the reconstructed local lines.

However, the H α is very sensitive to stellar activity being a well-known activity indicator ([Cincunegui, C. et al. 2007](#)), which had the side effect of cluttering our data with noise and other deviations from readable results. We retrieved some data points despite the impact of stellar activity.

The central radial velocity seems to have the same growing trend for the 11th of August as one would expect, although with very high uncertainties that overshadow any meaningful correlation. For the 31st of August its data has an unusual curvature as seen in the residuals in Figure 3.22 and a possible linear dependence in μ rather than orbital

phases, but we cannot say for sure due to the uncertainties. This potential behavior is quite odd, but we do know from [Cristo, E. et al. \(2024\)](#) that there was an anomaly on the second day so we might be seeing the effects of stronger activity in that day.

As for the other two parameters of line-width and line-center intensities no clear pattern was observed with some of the values being greater than the out-of-transit case. These values are clearly too noisy so no conclusion can be taken from them.

Overall this line shows the limitations of the Doppler shadow method with very noisy spectral lines affected by stellar activity and this is a major barrier to overcome requiring higher S/N data.

Chapter 5

Conclusion

In this work we studied how the spectra of a star changes across its surface using a planet as a probe. The study of the stellar surface is important since one of the major obstacles in the detection of small terrestrial planets is noise from the star that clutters the radial velocity signal. By studying the stellar surface, we can shed light on the physics of stellar surfaces and verify current models to improve future exoplanet detections and characterizations.

For this purpose, we employed a Doppler shadow method based on the work of [Dravins et al. \(2017a\)](#). This technique was based on using a transiting exoplanet as a probe to subtract and retrieve the spectral information of the sections of the stellar surface covered by the exoplanet along its transit chord. Furthermore, we utilized two new sets of high resolution and high S/N data taken by the spectrograph ESPRESSO of two transits of HD 189733 Ab.

Equipped with these sets of data we manipulated them to analyze four features of interest. The first was a white light CCF taken from the cross correlation of the spectra with a stellar spectral type K2 mask to understand the overall behavior of the spectrum. The second feature of interest were the Fe I lines in order to provide a direct comparison with the work of [Dravins et al. \(2018\)](#), where the iron lines were grouped into two types of CCFs for the weak iron and strong iron lines. The third were the sodium doublet lines which are the first time individual lines were reconstructed using this methodology. These two lines were chosen because they are very strong and distinctive in the stellar spectrum, and they have been of significant interest. The final feature analyzed was the $H\alpha$ line. This absorption feature was chosen to study the limitations of the current methodology and the impact of stellar activity in our data.

After applying our methodology, we were left with the localized profiles along the transit chord. We then proceeded to fit our data with a modified Gaussian function also used in [Dravins et al. \(2018\)](#) which proved to provide adequate fits. With the fitting of the profiles we extracted descriptive parameters such as the central radial velocities, line-widths and line-center intensities which describe the profile depths.

Analyzing the white light CCFs we verified the almost linear changes in radial velocity due to stellar rotation and how the residuals of a linear regression suggest differential rotation and/or the signature of convective blueshift when compared to [Cegla et al. \(2016\)](#). We also verified that the first transit had line-widths that decreased from the center of the stellar disk to the limb, after that we also verified that the CCFs get shallower from the center to the limb. These behaviors were consistent with what was expected by models and simulations such as those presented in [Cegla et al. \(2018\)](#). In addition, we suspect to see the effects of an anomaly due to a possible stellar spot, although not known for sure on the data of the second transit as first stated in [Cristo, E. et al. \(2024\)](#).

The iron CCFs allowed for a direct comparison with the work of [Dravins et al. \(2018\)](#). From here we noted the improvements that our data set had over their work, without the need of grouping observations for averaging and increasing S/N. The central RVs showed the same linear behavior with the strong iron CCFs having a clear extra component on the RVs visible on the residuals of the linear fit which we suspect to be due to convective blueshift. In a similar fashion to [Dravins et al. \(2018\)](#) we did not find a statistically relevant trend in the line-widths. In contrast we did see for the first day of observations a noticeable trend in line-center intensities with the depth of the profiles decreasing from star center to limb.

The sodium doublet lines show the potential of this method as with just two spectral lines we were able to retrieve information about center to limb variation. We compared the results to the simulations made in [Dethier & Bourrier \(2023\)](#). We saw the same process with changing central RVs and we observed a different phenomenon to the CCFs with the line-widths. They decreased from center to limb highlighting different behavior that was predicted by the simulations. No significant trend was detected for the line-center intensities.

Finally, for the reconstructed local $H\alpha$ lines not much could be recovered as this line is more strongly affected by stellar activity with only the usual central RVs behavior being seen for data of the first day where the anomaly was not present.

Many notable features were uncovered and analyzed and for the future it would be valuable to accurately model our data with the conditions found in HD 189733 A. The SOAP code provides the ability to simulate differential geometry in the surface of stars giving the opportunity of decoupling stellar differential rotation from convective blueshift and other forms of stellar activity. In the future simulations of spectral lines could be compared to the results of the Doppler shadow method to potentially test existing models of photospheric line formation and stellar activity.

Our results demonstrate the feasibility of the Doppler shadow method in retrieving information about center-to-limb variations in stellar surfaces. Such analysis can already be extended to individual lines, although stellar activity seems to be a limiting factor in the precision of the results. As such multiple observations timed to cover the same area of stellar surface may be necessary to average the data and increase the S/N. Understanding the stellar surface can help us in the future to detect smaller radial velocity signals and decouple the data from the stars and their exoplanets.

Appendix A

Uncertainties

Propagation of uncertainties is a critical step in understanding the significance of data. Many steps had to be taken to propagate our uncertainties from the raw spectra to our results.

The methodology employed to extract the local spectra/CCFs from our data was covered in Chapter 2 and in this appendix we complement it with the respective propagation of uncertainties. The methodologies differ slightly if we are working with the CCFs or single lines.

The general formula for propagation of uncertainties taking into account covariances of a function of the type $y = f(x_1, x_2, \dots, x_N)$ is of the form (Gardner 2003):

$$\sigma_y^2 = \sum_{i=1}^N \left(\frac{\delta f}{\delta x_i} \right)^2 \sigma_{x_i}^2 + \sum_{i=1}^N \sum_{j \neq i=1}^N \frac{\delta f}{\delta x_i} \frac{\delta f}{\delta x_j} \sigma_{x_i, x_j} \quad (\text{A.1})$$

Where σ_y is the standard deviation of y , likewise for the other variables and σ_{x_i, x_j} is the covariance between x_i and x_j .

A.1 CCF uncertainty propagation

We will showcase the process for the CCFs first. In the case of the white light CCFs, these were calculated by the ESPRESSO pipeline from the raw spectra and provided to us the data files that contained information of the radial velocities, flux and respective flux errors. For the iron CCFs we performed the cross correlation using equation 2.1.

The raw spectra also contained information about the flux and their respective errors. In order to propagate these uncertainties when performing cross correlation, we started by sampling the spectra. For each data point of a spectrum we took N samples from a

Gaussian distribution with the mean being the flux value of the data point and standard deviation as the flux error, having done this for all data points of a spectrum we were left with N sampled spectra, with these sampled spectra we calculated their CCF resulting in N sampled CCFs. With the N sampled CCFs of the same original spectrum we calculated their covariance matrix which proved more useful than simple standard deviations. In total 1000 samples were used for each CCF.

The next step was to take the CCFs and fit them with a modified Gaussian function 2.2 using y_0 for normalization and x_0 for barycentric correction. For the fitting we used the `curve_fit` function from the python package `scipy`, which accepts uncertainties as weights and provides the covariance matrix of our fitting parameters, taking the square root of the diagonals we get the standard deviations of the fitting parameters that we use as uncertainties.

When normalizing the white light CCFs by y_0 ($f = \frac{F}{y_0}$) the new uncertainties are given by:

$$\sigma_f = f \sqrt{\left(\frac{\sigma_{y_0}}{y_0}\right)^2 + \left(\frac{\sigma_F}{F}\right)^2} \quad (\text{A.2})$$

Where f is the normalized flux, F is the flux of the CCF prior to normalization and σ represents the uncertainties.

For the other types of CCFs ($f_i = \frac{F_i}{y_0}$) where we have a covariance matrix, the propagation was done with the following formula derived using Eq. A.1:

$$\sigma_{f_i, f_j} = \begin{cases} \frac{F_i F_j}{y_0^2} \left(\left(\frac{\sigma_{y_0}}{y_0} \right)^2 + \frac{\sigma_{F_i, F_j}}{F_i F_j} \right) + \frac{F_i F_j}{y_0^2} & i \neq j \\ \frac{F_i F_j}{y_0^2} \left(\left(\frac{\sigma_{y_0}}{y_0} \right)^2 + \frac{\sigma_{F_i, F_j}}{F_i F_j} \right) & i = j \end{cases} \quad (\text{A.3})$$

Here, σ_{F_i, F_j} is the covariance between the i th and j th CCF flux values. This formula is similar to A.2 and simplifies to it for the $i = j$ case. The extra term comes from the symmetry of the covariance matrix.

The barycentric correction results in shifting the abscissas of the CCFs which we treat as having uncertainties. The uncertainties shown in Figures such as 2.10 are from the fitting parameter x_0 . After barycentric correction the values of radial velocity are calculated as $x_{0,i,corrected} = x_{0,i} - b - m \times \phi_i$ where m and b were taken from the linear fit like the ones shown in the same figure, and ϕ_i is the orbital phase of the given CCF. The uncertainty shown on the left side of the figure was then calculated using:

$$\sigma_{x_{0,i,corrected}} = \sqrt{\sigma_{x_{0,i}}^2 + \sigma_b^2 + \sigma_m^2 \times \phi_i^2} \quad (\text{A.4})$$

The next step was the interpolation of the CCFs to share the same x axis, this was done to calculate the master out-of-transit and perform the subtraction. We used a simple linear interpolation, however the propagation of uncertainties proved to be trickier. At first the following formula was used:

$$\sigma_{f_i} = \sqrt{\left(\frac{rv_{i-1} - rv_i}{rv_{i+1} - rv_{i-1}} + 1\right)^2 \sigma_{f_{i-1}}^2 + \left(\frac{rv_i - rv_{i-1}}{rv_{i+1} - rv_{i-1}}\right)^2 \sigma_{f_{i+1}}^2} \quad (\text{A.5})$$

where rv are the radial velocities of the respective flux values for each of the CCF points. The i index represents the new interpolated point for the chosen rv_i value for the new x axis, the $i - 1$ and $i + 1$ refer to points of the original CCF before and after the interpolated point. This was used for the white light CCF.

However, despite the simplicity of the linear interpolation this formula seems to be an oversimplification that assumes no correlation between points. As such ends up slightly underestimating the uncertainties as seen in the comparison in Figure A.1. We found that a better propagation could be achieved if we took covariances into account, which for our data could be only obtained for the iron CCFs. In the covariance matrix of the iron CCFs calculated from cross correlating sampled spectra subdiagonal terms were found to be on the same order of the diagonal variances (around half) and by including them in error propagation resulted in non-underestimated uncertainties. Including the non-diagonal terms gives the following formula:

$$\begin{aligned} \sigma_{f_i}^2 = & \left(\frac{rv_{i-1} - rv_i}{rv_{i+1} - rv_{i-1}} + 1\right)^2 \sigma_{f_{i-1},f_{i-1}} + \left(\frac{rv_i - rv_{i-1}}{rv_{i+1} - rv_{i-1}}\right)^2 \sigma_{f_{i+1},f_{i+1}} \\ & + 2 \left(\frac{rv_i - rv_{i-1}}{rv_{i+1} - rv_{i-1}} - \left(\frac{rv_{i-1} - rv_i}{rv_{i+1} - rv_{i-1}}\right)^2\right) \sigma_{f_{i-1},f_{i+1}} \end{aligned} \quad (\text{A.6})$$

This formula was used for the interpolation of the iron CCFs while for the remaining data A.5 was used as we did not possess a covariance matrix.

The following step was the averaging of the out-of-transit CCFs to create the master out-of-transit CCF, this was straightforward due to the prior interpolation. Since we took the mean of the respective CCFs and their uncertainties are not exactly equal so we propagated the formula of the mean:

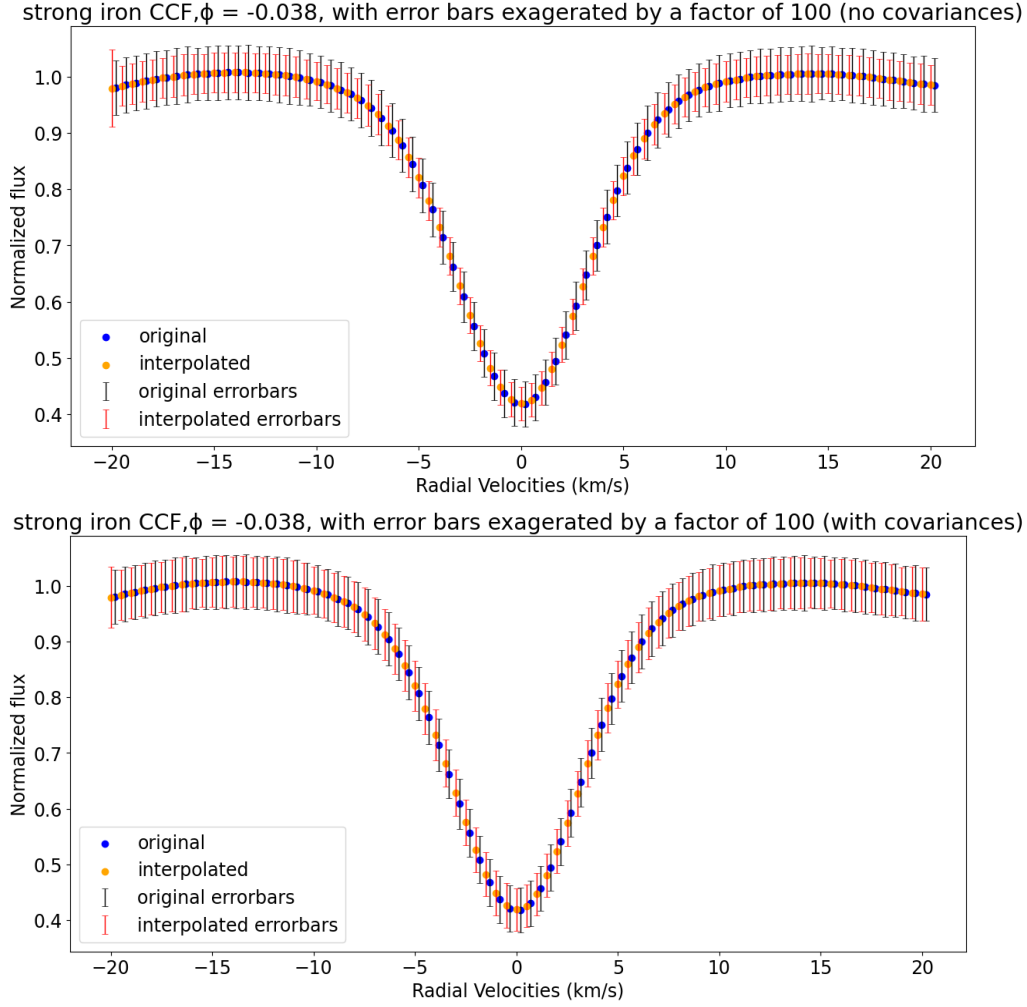


FIGURE A.1: Top: interpolation of a strong iron CCF without taking into account uncertainties using formula A.5. Bottom: same as the top except we consider covariances between points using formula A.6. The blue points are the CCF prior to interpolation, while orange is after interpolation. The black error bars are the uncertainties (square root of the diagonal of the covariance matrix) of the CCFs before interpolation with the red error bars being after interpolation. The uncertainties have been scaled up by a factor of 100 for ease of viewing.

$$\sigma_{f_{master,j}} = \frac{1}{N} \sqrt{\sum_{i=0}^{N-1} \sigma_{f_{i,j}}^2} \quad (\text{A.7})$$

Where the j index refers to the j th point on a given CCF, while the i index refers to a specific out-of-transit CCF with N being the total amount of CCFs.

Finally, we get to the subtraction to arrive at the local CCFs. This was done using formula 2.4, propagating results in the following relation:

$$\sigma_{F_{local}} = \sqrt{\sigma_{f_{Master,out}}^2 + \left(\sigma_{f_{in}} F_{SOAP}\right)^2} \quad (\text{A.8})$$

Next the local CCFs representative of specific portions of the stellar surface were fitted with a modified Gaussian 2.2 enabling us to extract the three descriptive parameters of their profiles. These were the central radial velocity x_0 , the line-widths b and line-center intensities defined as $(1 - \frac{a}{y_0}) * 100\%$. The first two have their uncertainties taken directly from the respective diagonal in the covariance matrix given by the `scipy curve_fit` function, the propagation for the last one was taken as:

$$\sigma_{line-center} = \frac{a}{y_0} \sqrt{\left(\frac{\sigma_a}{a}\right)^2 + \left(\frac{\sigma_{y_0}}{y_0}\right)^2} \times 100\% \quad (\text{A.9})$$

This concludes the propagation of uncertainties for the CCFs.

A.2 Individual lines uncertainty propagation

As discussed in Chapter 2 the analysis for the individual lines is very similar to that of the CCFs, the main differences are in the normalization process at the start, as such we will focus on the differences.

One of the first operations done on the spectra was a global normalization of an entire spectrum using the average value of a wavelength interval between 640 nm and 680 nm. The mean (F_{mean}) of the chosen interval is trivial to calculate and its uncertainty using essentially the same formula as A.7 but with the uncertainties of the flux values on that interval, due to the high resolution of the data the relatively small interval contained around 36000 so the resulting error in percentage was very small and barely effected our data. The uncertainty of this global normalization is given by:

$$\sigma_{f_g} = f_g \sqrt{\left(\frac{\sigma_F}{F}\right)^2 + \left(\frac{\sigma_{F_{mean}}}{F_{mean}}\right)^2} \quad (\text{A.10})$$

However as discussed in Chapter 2 this was not enough as there was a variation in the entire spectrum comparing different observations, as such this was complemented with local continuum linear regressions in order to properly normalize the specified spectral lines. At first we tried the usual approach to propagating uncertainties by taking into account the fitting parameters of the linear fit.

However, this approach resulted in increasing the uncertainties from around 0.5% of the flux values to over 30%. This was obviously incorrect specially given to how the new

flux values are in the same order of magnitude. So instead we opted to rescaling the uncertainties based on the flux values in this manner:

$$\sigma_f = f \frac{\sigma_{f_g}}{f_g} \quad (\text{A.11})$$

Where f represents the locally normalized flux, while f_g represents the globally normalized flux.

After that the process was the same as for the CCFs, the main differences were the line-widths that are expressed in Å and the central radial velocities which have to be converted from the central wavelengths. We also renormalize again to be consistent with the methodology of the CCFs using equation A.2 and interpolating with formula A.5. The conversion from central wavelengths to central RVs ads an additional step, starting with the Doppler shift formula $RV = c \frac{\lambda - \lambda_0}{\lambda_0}$ we get the following propagation:

$$\sigma_{RV} = c \frac{\lambda}{\lambda_0} \sqrt{\left(\frac{\sigma_{\lambda}^2}{\lambda}\right)^2 + \left(\frac{\sigma_{\lambda_0}}{\lambda_0}\right)^2} \quad (\text{A.12})$$

The remaining processes were essentially identical so we conclude the propagation of uncertainties.

Appendix B

Derivation of μ

We showcase a simple derivation of μ , this is a geometric problem relating the projection of the planet's position on the stellar surface. Figures B.1, B.2 and B.3 give an helpful view of the geometry from 3 different perspectives.

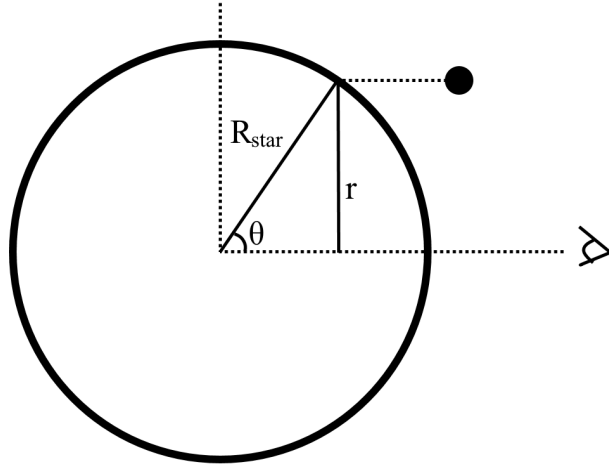


FIGURE B.1: Side view perpendicular to our line of sight showing the planet's projection on the stellar surface.

The definition of μ is as the cosine of the angle θ seen on the left side of Figures B.1, from the definition of cosine we get:

$$\mu = \frac{\sqrt{R_{star}^2 - r^2}}{R_{star}} \quad (B.1)$$

The angle α measured relative to the center of the transit at time t_{epoch} (Figure B.2) can be derived from the planet's orbital phases assuming a circular orbit as $\alpha = \omega(t - t_{epoch})$, where ω is the angular velocity, which for a circular orbit with uniform circular movement

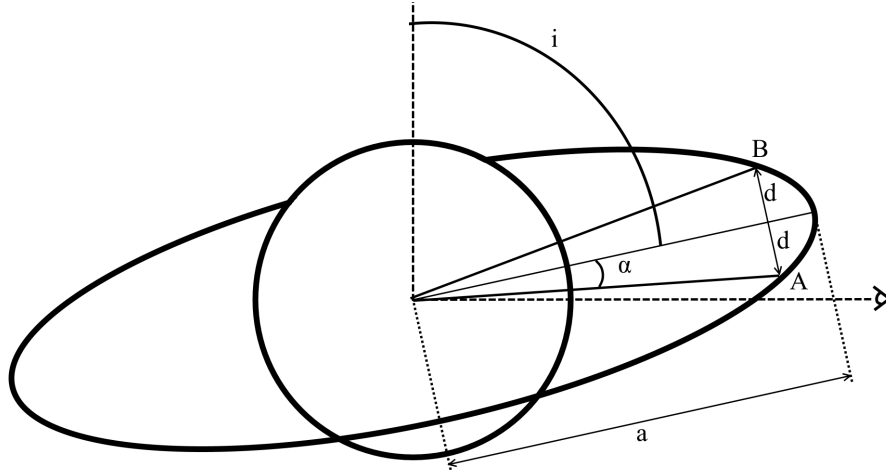


FIGURE B.2: Right side shows a different perspective showing the planet's orbit around a star and the angle α it traces in space.

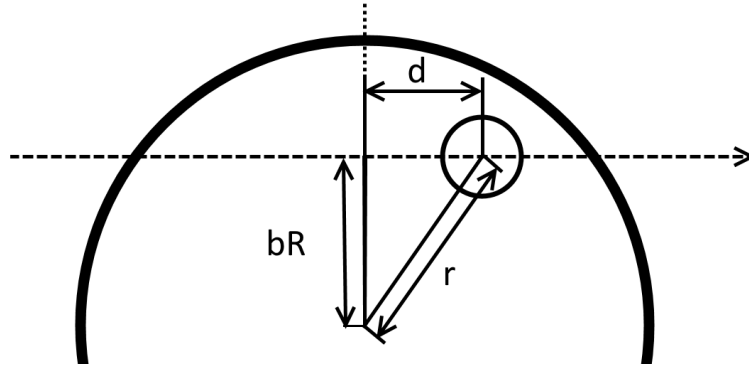


FIGURE B.3: Front view of the star and planet during a transit.

has value $\omega = \frac{2\pi}{P}$. Replacing into the expression for α gives $\alpha = 2\pi \frac{t-t_{epoch}}{P}$, but this now gives the definition of orbital phase in equation 1.5 giving us:

$$\alpha = 2\pi\phi_t \quad (\text{B.2})$$

Here, ϕ_t is the orbital phase of the planet at time t . Looking at the angle α we can define a right triangle with catheti of length d and of the semi-major axis (a), which in this simplified case is the radius of the orbit. From which we get $\sin \alpha = \frac{d}{a}$, equating with B.2 results in $d = a \sin (2\pi\phi_t)$.

Looking at the front view of the star and planet in Figure B.3 we have a right triangle of catheti of length bR_{star} and d and hypotenuse r . Applying Pythagoras' theorem we get that $r = \sqrt{(bR_{star})^2 + d^2}$, substituting gives:

$$r = \sqrt{(bR_{star})^2 + a^2 \sin^2 (2\pi\phi_t)} \quad (\text{B.3})$$

Replacing equation B.3 into equation B.1 gives our final results in the form of:

$$\mu = \frac{\sqrt{R_{star}^2 - (bR_{star})^2 - a^2 \sin^2(2\pi\phi_t)}}{R_{star}} \quad (\text{B.4})$$

Which is the same as equation 1.6.

Appendix C

Statistical tests

To evaluate the behavior of our results and the significance of the trends observed we used some statistical quantities to test our data. In this appendix we give a brief description of the statistical quantities used.

The first was the R^2 also known as the coefficient of determination ([Lewis-Beck 1980](#)). This is a very common statistical quantity used to evaluate how good a fit with a function can be to a set of data. The R^2 can be given by (page 21 of [Lewis-Beck \(1980\)](#)):

$$R^2 = 1 - \frac{ESS}{TSS} \quad (C.1)$$

where ESS is the sum of the square of the differences between the data (y_i) and the fit (f_i):

$$ESS = \sum_{i=0}^{N-1} (y_i - f_i)^2 \quad (C.2)$$

While TSS is the sum of the squares of the differences between the data and their mean ($\bar{y}$)

$$TSS = \sum_{i=0}^{N-1} (y_i - \bar{y})^2 \quad (C.3)$$

A value of 1 indicates that the fit perfectly matches the data, while 0 would be the case of a baseline model that predicts the average value of the sample of data. Negative values can occur specially for cases where the data for example follows a non-linear behavior and tends to indicate that our model performs worse than a model that just predicts the average value.

The Pearson correlation coefficient ([David Freedman 2007](#)) is a quantity that measures the linear correlation of a set of data, for a sample it is often represented as r , for a set of x and y values is calculated as the covariance of x and y divided by the product of the standard deviations of x and standard deviation of y (page 134 of [David Freedman \(2007\)](#)):

$$r = \frac{\text{cov}(x, y)}{(\text{SD of } x) \times (\text{SD of } y)} \quad (\text{C.4})$$

The formula we used for a set of N (x, y) points was then:

$$r_{xy} = \frac{\sum_{i=0}^{N-1} (x_i - \bar{x}) (y_i - \bar{y})}{\sqrt{\sum_{i=0}^{N-1} (x_i - \bar{x})^2} \sqrt{\sum_{i=0}^{N-1} (y_i - \bar{y})^2}} \quad (\text{C.5})$$

This coefficient measures basically how close to a line the data is. A value of 1 would be a perfect line with positive slope, while -1 would be with negative slope, values close to 0 could be caused by a random distribution of points.

The Spearman's rank correlation coefficient ([Kokoska & Zwillinger 2000](#)) evaluates the rank correlation of two variables, that is it measures how well the data is described by a monotonic function. It can be thought as a more general version of the Pearson coefficient, but instead of evaluating the linearity of the data it describes in general monotonic relationships, which is partly what we were interested in our data. For this we used the `scipy.stats.spearmanr` function to calculate this quantity.

Another quantity calculated using the `scipy.stats.spearmanr` function is the p-value. The p-value is the probability of getting results as extreme or greater than the ones measures assuming that the null hypothesis is true, meaning that the lower the values the more likely that the null hypothesis is not true and we are observing something different. In the `scipy` function used, the null hypothesis is that of no ordinal correlation, that is the data is not caused by a monotonic function. This is useful for us as we are trying to identify growing or decreasing behaviors in our results. The cutoff value for what one may consider statistically relevant may depend on what one studies, but it is common to use 0.05 as a limit as given in [Fisher \(1954\)](#).

The final quantity used is the χ^2 ([Taylor 1997](#)). One of the uses of this quantity is for goodness of fit tests, where we can compare different hypotheses. Since our data contains uncertainties we opted to use the following formula:

$$\chi^2 = \sum_{i=0}^{N-1} \frac{(y_i - f_i)^2}{\sigma_i^2} \quad (\text{C.6})$$

Here, y_i is our data, σ_i the uncertainties of our data in the form of standard deviations and f_i is the values predicted by our model. In chapter 3 we opted for two models one being a line of non-zero slope and the other being a horizontal line of height given by the mean of the data. Looking at the formula it is clear that χ^2 will be smaller for the more optimal model.

Appendix D

Telluric correction comparison

Here we give another example of the effect of telluric lines in our data. We compare Figure [2.25](#) with the same analysis done without telluric correction resulting in Figure [D.1](#)

As one can see the effect of telluric contamination is very severe. This is because our methodology involves subtraction and working on the star's referential which involved RV correction. So when performing this step, the telluric lines are misaligned and they do not cancel out upon subtraction resulting in lots of contamination in the subtracted spectra.

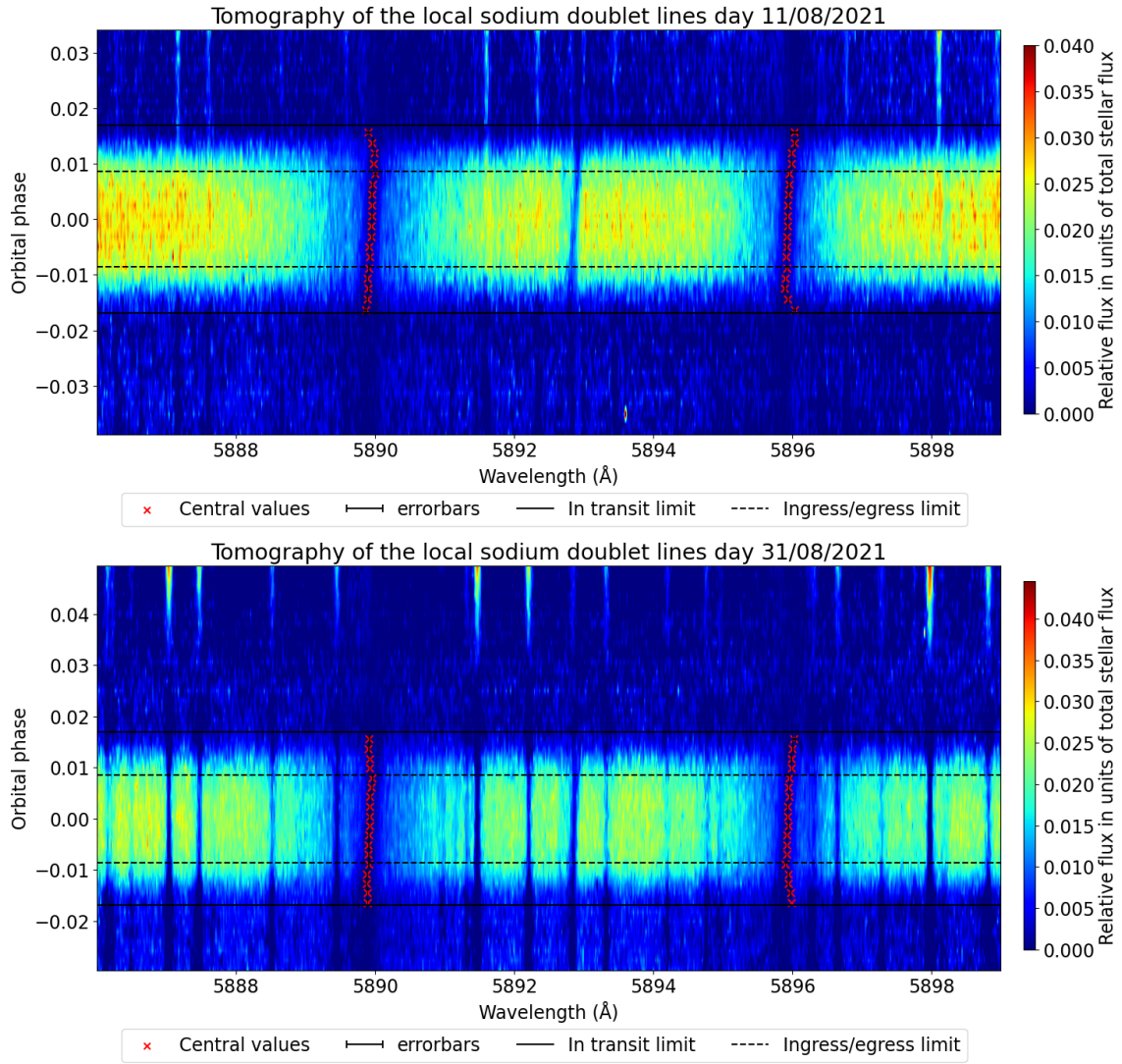


FIGURE D.1: Tomography of the reconstructed sodium doublet lines without telluric correction displayed in the same manner as Figure 2.15.

Appendix E

Iron lines list

In this appendix we include the data for the iron lines used to create the CCFs taken directly from [Dravins et al. \(2018\)](#). According to them the wavelengths and excitation potentials were taken from [Stenflo & Lindegren \(1977\)](#), while the line-center intensities are normalized to their continuum and were calculated from Gaussian fits to the HARPS dataset used in their work and also subject to a convolution with their respective instrument profile.

Wavel. (nm)	Int. (%)	Exc. pot. (eV)	Wavel. (nm)	Int. (%)	Exc. pot. (eV)	Wavel. (nm)	Int. (%)	Exc. pot. (eV)
455.16499	62	3.94	553.92824	79	3.64	616.53641	61	4.14
459.35268	61	3.94	555.26919	91	4.95	622.67403	71	3.88
465.75879	55	2.84	556.88689	86	3.63	631.15050	68	2.83
465.82976	72	3.27	560.76668	83	4.15	631.58164	63	4.07
467.28364	59	1.61	561.13587	86	3.63	638.07483	59	4.19
472.61396	73	3.00	561.86360	55	4.21	638.57206	88	4.73
478.59583	67	4.14	563.58247	65	4.26	639.25429	73	2.28
479.43571	76	2.42	563.66992	74	3.64	643.64102	88	4.19
479.94092	58	3.64	564.66854	90	4.26	649.64738	57	4.79
480.25216	81	4.61	565.14716	79	4.47	657.42325	57	0.99
480.81509	60	3.25	565.23194	72	4.26	658.12143	71	1.48
480.99400	70	3.57	566.13480	75	4.28	660.80301	74	2.28
521.38071	88	3.94	567.76875	90	4.10	662.75488	76	4.55
522.31851	63	3.63	568.02441	86	4.19	664.69355	85	2.61
523.62039	65	4.19	572.08950	83	4.55	670.35720	62	2.76
529.39609	67	4.14	579.39178	66	4.22	671.03213	74	1.48
529.45493	79	3.64	580.77868	89	3.29	673.79830	82	4.56
535.81168	84	3.30	581.19172	86	4.14	673.95243	78	1.56
538.63345	64	4.15	581.48092	75	4.28	674.59626	92	4.07
539.52187	77	4.44	582.78794	84	3.28	674.69508	92	2.61
541.70405	65	4.41	583.51018	84	4.26	675.27107	74	4.64
546.08762	88	3.07	585.22222	64	4.55	679.32656	87	4.07
546.15530	71	4.44	585.87840	85	4.22	680.68491	64	2.73
546.42825	63	4.14	586.11123	90	4.28	682.03730	69	4.64
547.00957	71	4.44	588.12822	84	4.61	682.85976	61	4.64
548.71489	67	4.41	590.24755	85	4.59	683.98340	67	2.56
549.18346	85	4.19	592.96802	65	4.55	684.13450	60	4.61
549.35012	65	4.10	603.40365	90	4.31	684.36606	59	4.55
552.24491	58	4.21	608.27147	59	2.22			

TABLE E.1: Table containing the 86 weak Fe I lines taken directly from [Dravins et al. \(2018\)](#). The weak lines were selected with line-center intensities $\geq 55\%$.

Wavel. (nm)	Int. (%)	Exc. pot. (eV)	Wavel. (nm)	Int. (%)	Exc. pot. (eV)	Wavel. (nm)	Int. (%)	Exc. pot. (eV)
436.59004	37	2.99	525.06527	26	2.20	602.70562	49	4.07
443.25726	39	3.57	525.34693	36	3.28	605.60114	49	4.73
444.54760	33	0.09	536.54063	38	3.57	606.54921	29	2.61
452.34015	47	3.65	536.74755	29	4.41	615.16217	50	2.18
457.42191	48	3.21	536.99702	30	4.37	615.77331	52	4.07
460.20006	24	1.61	537.37136	48	4.47	617.33433	43	2.22
467.88519	23	3.60	537.95796	44	3.69	618.02084	51	2.73
474.15341	29	2.83	538.94866	38	4.41	620.03204	41	2.61
483.95500	39	3.27	543.29525	46	4.44	621.34375	36	2.22
487.43565	64	3.07	544.50502	34	4.39	621.92886	35	2.20
487.58815	40	3.33	547.39076	40	4.15	623.26493	43	3.65
488.54361	35	3.88	552.55472	53	4.23	626.51412	35	2.18
489.28624	50	4.22	554.39399	49	4.22	627.02322	52	2.86
490.93874	42	3.93	563.82675	41	4.22	629.78013	40	2.22
491.80152	51	4.23	566.25233	37	4.18	632.26936	41	2.59
494.63941	26	3.37	570.15527	34	2.56	633.53378	33	2.20
495.01108	34	3.42	570.54677	63	4.30	640.80262	42	3.69
498.32566	31	4.15	571.78379	51	4.28	643.08538	32	2.18
504.42164	34	2.85	573.17666	51	4.26	648.18784	46	2.28
509.07807	35	4.26	575.31287	43	4.26	649.89461	49	0.96
514.17460	26	2.42	577.50849	50	4.22	659.38798	39	2.43
524.24988	31	3.63	591.62535	47	2.45	660.91189	47	2.56
524.37823	45	4.26	595.66997	41	0.86	675.01597	43	2.42
525.02171	29	0.12	600.79656	54	4.65	685.51684	54	4.56

TABLE E.2: Table containing the 72 strong Fe I lines taken directly from [Dravins et al. \(2018\)](#).

Bibliography

- Akinsanmi, B., Barros, S. C. C., Santos, N. C., et al. 2019, Detectability of shape deformation in short-period exoplanets, *Astronomy & Astrophysics*, 621, A117
- Azevedo Silva, T., Demangeon, O. D. S., Santos, N. C., et al. 2022, Detection of barium in the atmospheres of the ultra-hot gas giants WASP-76b and WASP-121b: Together with new detections of Co and Sr+ on WASP-121b, *Astronomy & Astrophysics*, 666, L10
- Bakos, G. Á., Pál, A., Latham, D. W., Noyes, R. W., & Stefanik, R. P. 2006, A Stellar Companion in the HD 189733 System with a Known Transiting Extrasolar Planet, *The Astrophysical Journal*, 641, L57–L60
- Boisse, I., Bonfils, X., & Santos, N. C. 2012, SOAP: A tool for the fast computation of photometry and radial velocity induced by stellar spots, *Astronomy & Astrophysics*, 545, A109
- Bouchy, F., Udry, S., Mayor, M., et al. 2005, ELODIE metallicity-biased search for transiting Hot Jupiters: II. A very hot Jupiter transiting the bright K star HD 189733, *Astronomy & Astrophysics*, 444, L15
- Brion, J., Chakir, A., Charbonnier, J., et al. 1998, Absorption Spectra Measurements for the Ozone Molecule in the 350–830 nm Region, *Journal of Atmospheric Chemistry*, 30, 291
- Cameron, A. C., Bruce, V. A., Miller, G. R. M., Triaud, A. H. M. J., & Queloz, D. 2010, Line-profile tomography of exoplanet transits – I. The Doppler shadow of HD 189733b*, *Monthly Notices of the Royal Astronomical Society*, 403, 151
- Cegla, H. M., Lovis, C., Bourrier, V., et al. 2016, The Rossiter-McLaughlin effect reloaded: Probing the 3D spin-orbit geometry, differential stellar rotation, and the spatially-resolved stellar spectrum of star-planet systems, *Astronomy & Astrophysics*, 588, A127

- Cegla, H. M., Watson, C. A., Shelyag, S., et al. 2018, Stellar Surface Magneto-convection as a Source of Astrophysical Noise. II. Center-to-limb Parameterization of Absorption Line Profiles and Comparison to Observations, *The Astrophysical Journal*, 866, 55
- Charbonneau, D., Brown, T. M., Latham, D. W., & Mayor, M. 2000, Detection of Planetary Transits Across a Sun-like Star, *The Astrophysical Journal*, 529, L45
- Cincunegui, C., Díaz, R. F., & Mauas, P. J. D. 2007, H α and the Ca II H and K lines as activity proxies for late-type stars*, *A&A*, 469, 309
- Cristo, E., Santos, N. C., Demangeon, O., et al. 2022, CaRM: Exploring the chromatic Rossiter-McLaughlin effect: The cases of HD 189733b and WASP-127b, *Astronomy & Astrophysics*, 660, A52
- Cristo, E., Esparza Borges, E., Santos, N. C., et al. 2024, An ESPRESSO view of the HD 189733 system - Broadband transmission spectrum, differential rotation, and system architecture, *A&A*, 682, A28
- David Freedman, Robert Pisani, R. P. 2007, *Statistics*, 4th Edition (W. W. Norton Company), 125–140
- Dethier, W. & Bourrier, V. 2023, Combined analysis of stellar and planetary absorption lines via global forward-transit simulations, *Astronomy & Astrophysics*, 674, A86
- Dravins, D., Gustavsson, M., & Ludwig, H.-G. 2018, Spatially resolved spectroscopy across stellar surfaces: III. Photospheric Fe I lines across HD 189733A (K1 V), *Astronomy & Astrophysics*, 616, A144
- Dravins, D., Ludwig, H.-G., Dahlén, E., & Pazira, H. 2017a, Spatially resolved spectroscopy across stellar surfaces - I. Using exoplanet transits to analyze 3D stellar atmospheres, *A&A*, 605, A90
- Dravins, D., Ludwig, H.-G., Dahlén, E., & Pazira, H. 2017b, Spatially resolved spectroscopy across stellar surfaces: II. High-resolution spectra across HD 209458 (G0 V), *Astronomy & Astrophysics*, 605, A91
- Dumusque, X., Boisse, I., & Santos, N. C. 2014, SOAP 2.0: A TOOL TO ESTIMATE THE PHOTOMETRIC AND RADIAL VELOCITY VARIATIONS INDUCED BY STELLAR SPOTS AND PLAGES, *The Astrophysical Journal*, 796, 132

- Evans, T. M., Pont, F., Sing, D. K., et al. 2013, THE DEEP BLUE COLOR OF HD 189733b: ALBEDO MEASUREMENTS WITH HUBBLE SPACE TELESCOPE /SPACE TELESCOPE IMAGING SPECTROGRAPH AT VISIBLE WAVELENGTHS, *The Astrophysical Journal*, 772, L16
- Fares, R., Donati, J.-F., Moutou, C., et al. 2010, Searching for star-planet interactions within the magnetosphere of HD 189733: Magnetospheric modelling of HD 189733, *Monthly Notices of the Royal Astronomical Society*, 406, 409–419
- Figueira, P., Santos, N. C., Pepe, F., Lovis, C., & Nardetto, N. 2013, Line-profile variations in radial-velocity measurements: Two alternative indicators for planetary searches, *Astronomy & Astrophysics*, 557, A93
- Fisher, R. A. 1954, *Statistical Methods for Research Workers*, 12th edn. (Hafner Publishing Company, NYC), 82
- Gardner, J. L. 2003, Uncertainties in Interpolated Spectral Data., *Journal of research of the National Institute of Standards and Technology*, 108, 69
- Han, E., Wang, S. X., Wright, J. T., et al. 2014, Exoplanet Orbit Database. II. Updates to Exoplanets.org, *PASP*, 126, 827
- Husser, T.-O., Wende-von Berg, S., Dreizler, S., et al. 2013, *Astrophysics* A new extensive library of PHOENIX stellar atmospheres, *A&A*, 553, A6
- Kausch, W., Noll, S., Smette, A., et al. 2015, Molecfit: A general tool for telluric absorption correction: II. Quantitative evaluation on ESO-VLT/X-Shooterspectra, *Astronomy & Astrophysics*, 576, A78
- Knutson, H. A., Charbonneau, D., Allen, L. E., et al. 2007, A map of the day-night contrast of the extrasolar planet HD 189733b, *Nature*, 447, 183
- Kokoska, S. & Zwillinger, D. 2000, *Standard probability and statistics tables and formulae: Student edition* (CRC Press)
- Kramida, A., Yu. Ralchenko, Reader, J., & and NIST ASD Team. 2023, *NIST Atomic Spectra Database* (ver. 5.11), [Online]. Available: <https://physics.nist.gov/asd> [2024, September 13]. National Institute of Standards and Technology, Gaithersburg, MD.

- Lanza, A. F., Malavolta, L., Benatti, S., et al. 2018, The GAPS Programme with HARPS-N at TNG: XVII. Line profile indicators and kernel regression as diagnostics of radial-velocity variations due to stellar activity in solar-like stars, *Astronomy & Astrophysics*, 616, A155
- Léger, A., Grasset, O., Fegley, B., et al. 2011, The extreme physical properties of the CoRoT-7b super-Earth, *Icarus*, 213, 1–11
- Lewis-Beck, M. S. 1980, *Applied regression: An introduction* (Sage publications), 20–25
- Mayor, M. & Queloz, D. 1995, A Jupiter-mass companion to a solar-type star, *Nature*, 378, 355
- Murray, C. D. & Correia, A. C. M. 2011, *Keplerian Orbits and Dynamics of Exoplanets*
- Nardetto, N., Mourard, D., Kervella, P., et al. 2006, High resolution spectroscopy for Cepheids distance determination: I. Line asymmetry, *Astronomy & Astrophysics*, 453, 309–319
- Oshagh, M., Santos, N. C., Boisse, I., et al. 2013, Effect of stellar spots on high-precision transit light-curve, *Astronomy & Astrophysics*, 556, A19
- Paredes, L. A., Henry, T. J., Quinn, S. N., et al. 2021, The Solar Neighborhood XLVIII: Nine Giant Planets Orbiting Nearby K Dwarfs, and the CHIRON Spectrograph’s Radial Velocity Performance, *The Astronomical Journal*, 162, 176
- Parviainen, H. & Aigrain, S. 2015, ldtk: Limb Darkening Toolkit, *MNRAS*, 453, 3821
- Pepe, F., Cristiani, S., Rebolo, R., et al. 2021, ESPRESSO at VLT: On-sky performance and first results, *Astronomy & Astrophysics*, 645, A96
- Queloz, D., Henry, G. W., Sivan, J. P., et al. 2001, No planet for HD 166435, *Astronomy and Astrophysics*, 379, 279–287
- Rosenblatt, F. 1971, A two-color photometric method for detection of extra-solar planetary systems, *Icarus*, 14, 71
- Rosenthal, L. J., Fulton, B. J., Hirsch, L. A., et al. 2021, 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*, 255, 8

- Santerne, A., Díaz, R. F., Almenara, J. M., et al. 2015, PASTIS: Bayesian extrasolar planet validation II. Constraining exoplanet blend scenarios using spectroscopic diagnoses
- Santos, N. C., Barros, S. C., Demangeon, O. D., & Faria, J. P. 2020, Detection and Characterization Methods of Exoplanets
- Seager, S. & Mallen-Ornelas, G. 2003, A Unique Solution of Planet and Star Parameters from an Extrasolar Planet Transit Light Curve, *The Astrophysical Journal*, 585, 1038
- Serrano, L. M., Oshagh, M., Cegla, H. M., et al. 2020, Can we detect the stellar differential rotation of WASP-7 through the Rossiter–McLaughlin observations?, *Monthly Notices of the Royal Astronomical Society*, 493, 5928–5943
- Smette, A., Sana, H., Noll, S., et al. 2015, Molecfit: A general tool for telluric absorption correction: I. Method and application to ESO instruments, *Astronomy & Astrophysics*, 576, A77
- Sousa, S. G., Adibekyan, V., Delgado-Mena, E., et al. 2018, SWEET-Cat updated. New homogenous spectroscopic parameters, *A&A*, 620, A58
- Stenflo, J. O. & Lindegren, L. 1977, Statistical analysis of solar Fe I lines: magnetic line broadening., *A&A*, 59, 367
- Taylor, J. R. 1997, *An introduction to error analysis: The study of uncertainties in physical measurements* (2nd edition) (University Science Books), 265, 266
- Triaud, A. H. M. J., Queloz, D., Bouchy, F., et al. 2009, The Rossiter-McLaughlin effect of CoRoT-3b and HD 189733b, *Astronomy & Astrophysics*, 506, 377
- Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python, *Nature Methods*, 17, 261
- Vugrin, K. W., Swiler, L. P., Roberts, R. M., Stucky-Mack, N. J., & Sullivan, S. P. 2007, Confidence region estimation techniques for nonlinear regression in groundwater flow: Three case studies, *Water Resources Research*, 43 [<https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2005WR004804>]
- Yan, F., Pallé, E., Fosbury, R. A. E., Petr-Gotzens, M. G., & Henning, T. 2017, Effect of the stellar absorption line centre-to-limb variation on exoplanet transmission spectrum observations, *Astronomy & Astrophysics*, 603, A73

Zhao, Y. & Dumusque, X. 2023, SOAP-GPU: Efficient spectral modeling of stellar activity using graphical processing units, *Astronomy & Astrophysics*, 671, A11