Deciphering the atmospheric properties of hot gas giant exoplanets via ground-based transmission spectroscopy

Petros Spyratos

Doctor of Philosophy

School of Chemical and Physical Sciences, Keele University

June 2023

Abstract

Transiting hot gas giant exoplanets are great targets for atmospheric investigations due to their extended atmospheres and their strong obscuring effect on the observed starlight when they pass in front of their host star. Transmission spectroscopy can probe transiting exoplanets at various wavelengths and, hence, reveal the dominant opacity sources in specific regions of their atmospheres. In this thesis, I present ground-based transmission spectroscopy obtained in the optical for three exoplanets and describe the various analysis methods explored in order to produce the characteristic transmission spectrum for each exoplanet.

The first target was the test case WASP-75b for which a rudimentary data reduction and analysis procedure resulted in a transmission spectrum of a low precision. The data set for this target was mainly used for practice purposes and so the analysis is brief.

The next target was the low-density hot Saturn WASP-88b. Here, I examined a polynomial detrending approach and a detrending approach based on Gaussian Processes (GPs) in an effort to isolate the systematic effects from the transit light curves. The resulting transmission spectrum was found to be featureless with atmospheric models indicating the presence of high-altitude haze.

The final target was the hot Jupiter WASP-74b. In this case, a new method that performs a common-mode correction only to the light curves of the target and considers an exponential of airmass during GP detrending was also presented in the spectroscopic analysis. This method resulted in increased transit depth precision and a steep scattering slope in the measured blue-optical transmission spectrum that suggests enhanced haze in the upper atmosphere.

The results for WASP-88b and WASP-74b are representative of an emerging trend of steep scattering slopes observed in an increasing number of exoplanet atmospheres. The physics behind this spectral shape is not yet well-understood but several mechanisms have been proposed.

Acknowledgements

First and foremost, I would like to thank my family because without their support all these 4 years, and especially during the challenging times of the pandemic, I may not have had the mental capacity to complete my research. Their role was vital in keeping me motivated and focused.

Next, I would like to thank my supervisor, John Taylor, for providing guidance, valuable context and constructive comments throughout my research project. I am also very grateful to his family who helped me when I had to isolate during the pandemic.

A huge thank you also goes to my main collaborator, Nikolay Nikolov, who was like a supervisor to me as he aided me in many aspects of my research and helped me gain a deeper understanding of transmission spectroscopy and atmospheric physics.

I am grateful for all the fun moments I had with the PhD students and other members of the astrophysics group at Keele. Here, I would like to give special thanks to my second supervisor, Pierre Maxted, who assisted me in my first observing run and guided me through all the processes involved in the operation of a large telescope, and to Nick Wright, who was the PGR director during the pandemic and was very supportive in all the issues I faced.

Many thanks should also be given to all people who helped my research go forward with their valuable contributions to the work presented in this thesis. Their contribution has been acknowledged in the beginning of Chapters 4 and 5 and in the manuscripts: Spyratos et al. (2021) and Spyratos et al. (2023). More specifically:

- The remote observations and data reduction for WASP-88b, as well as the data reduction for WASP-74b were performed by Nikolay Nikolov (Sections 4.2.2 and 5.2).

- The observations for WASP-75b and the data reduction and analysis of the TESS data for WASP-88b were implemented by John Southworth (Sections 3.2, 4.2.1 and 4.3.1).

- The atmospheric retrievals with AURA were conducted by Savvas Constantinou, with supervision from Nikku Madhusudhan (Sections 4.5.3 and 5.5.3).
- The observations for WASP-74b were done by Elyar Sedaghati (Section 5.2).

In addition, I am also grateful to all ESO staff members who provided on-site support in obtaining the observations.

This work is based on observations collected at the European Organization for Astronomical Research in the Southern Hemisphere under European Southern Observatory programmes 099.C-0301(A), 199.C-0467(D), and 0101.C-0716. It is also based on data collected by the TESS mission. The NTT EFOSC2 and VLT FORS2 data are publicly available on the ESO archive. The TESS data are publicly available on the Mikulski Archive for Space Telescopes (MAST) at the Space Telescope Science Institute (STScI).

This research also relies on the Python programming language and makes heavy use of the science-focused extension libraries NumPy (array manipulation, Harris et al., 2020), SciPy (numerical methods, Virtanen et al., 2020), and AstroPy (FITS file handling, Astropy Collaboration et al., 2013), as well as use of the plotting toolkit Matplotlib (Hunter, 2007).

Contents

List of Figures

List of Tables

1 An introduction to transiting exoplanets and their atmospheres

1.1 Introduction

The year 2022 marked two important events in the field of exoplanetary science. The first one is the 30-year anniversary since the announcement of the first confirmed detections of exoplanets (Wolszczan & Frail, 1992), and the second one is the news that the number of confirmed exoplanets has now officially exceeded the milestone number of 5000 (see e.g. the NASA Exoplanet Archive[1]). These two numbers highlight the dynamic evolution of the field and promise a bright future in terms of exoplanet discoveries.

During this 30-year period many milestones were hit but one of the most important ones came in 2002. At the 10th anniversary of the first confirmed detections, scientists were able to observe the atmosphere of an exoplanet for the first time (Charbonneau et al., 2002). This extraordinary achievement paved the way for a new era where atmospheric characterisation was put at the forefront of exoplanetary science. Over 100 exoplanet atmospheres have now been observed (see e.g. the ExoAtmospheres database[2]) and with the recent launch of the James Webb Space Telescope (JWST) the expectation is that the sample of observed atmospheres will increase substantially. The aim of this thesis is to build upon this legacy by analysing and presenting data from new observations of two exoplanetary atmospheres (see Chapter 4 for a first-time atmospheric exploration of an exoplanet and Chapter 5 for a revisit of an exoplanetary atmosphere).

This chapter will therefore introduce the main topic of this thesis, which is to investigate the transmission spectra of a few selected targets. Starting with a brief history of exoplanet discoveries and a few basic concepts about exoplanets, the goal

[1] `https://exoplanetarchive.ipac.caltech.edu/`
[2] `http://research.iac.es/proyecto/exoatmospheres/table.php`

is to provide a smooth introduction to what is about to follow. The first few sections will then focus on the transit method and the main properties of hot gas giant exoplanets, before we get into the essence of this thesis: the exploration of exoplanetary atmospheres. For this reason, the last few sections will concentrate on how we can observe these atmospheres, what we can observe in them, what we can infer about the environment and what challenges exist in ground- and space-based observations. The chapter will end with a brief overview of my thesis.

1.1.1 A brief history

The era of exoplanet discoveries officially starts with the first confirmed exoplanet detections in 1992. These first detections involved the observation of two nearly Earth-mass exoplanets (what we now call super-Earths) orbiting a rapidly rotating pulsar (Wolszczan & Frail, 1992). This was a ground-breaking discovery revealing that exoplanets may even exist in hostile, post-supernova environments. Unofficially, two more potential exoplanets had been discovered that preceded these detections. The first claim was made by Campbell et al. (1988), who described a substellar companion of planetary nature orbiting a solar-type star. This discovery was later confirmed by Hatzes et al. (2003), who found a minimum mass of $1.7\,\mathrm{M_{Jup}}$ and an orbital period of 906 days. Meanwhile, Latham et al. (1989) presented evidence of the existence of a "brown dwarf" orbiting another solar-type star. The object had a minimum mass of $11\,\mathrm{M_{Jup}}$ and, based on our current understanding of the maximum planet mass limit, this object could be an exoplanet. However, the planetary nature of this object was recently disproved on the basis of astrometric data from Gaia Data Release 3, which showed that the orbital inclination is low and that the secondary component actually has a mass of $0.21 \pm 0.02\,\mathrm{M_\odot}$ (Gaia Collaboration et al., 2022a).

A major breakthrough in exoplanet studies was the first confirmed and indisputable detection of an exoplanet orbiting a Sun-like star (Mayor & Queloz, 1995). The properties of this exoplanet were determined to be unusual at the time as the exoplanet was Jupiter-like with a minimum mass of $0.47\,\mathrm{M_{Jup}}$, but was orbiting ex-

tremely close to its host star having an orbital period of just 4.2 days. This result was in contradiction with the image we have from our own Solar System where rocky planets are close-in and gas giants are further out. Nevertheless, the potential existence and detectability of such planets had been predicted as early as 1952 (Struve, 1952). In the following years, several exoplanet discoveries revealed that such planets are not rare (see e.g. Butler et al., 1997, 1998) and so the name *hot Jupiters* was finally established for this particular group of exoplanets. As we will see in Sections 1.3 and 1.4, some of these peculiar exoplanets have properties that make them excellent targets for atmospheric characterisation.

Exoplanet discoveries are now announced on a daily basis and Figure 1.1 shows the increase in the cumulative total through the years. Some of the most important milestones include the detection of the first transiting exoplanet (Charbonneau et al., 2000; Henry et al., 2000), the discovery of the first multiplanet system (Butler et al., 1999), the detection of the first exoplanet in the habitable zone (Santos et al., 2001), the first directly imaged exoplanet (Chauvin et al., 2004), and the observation of the first Earth-sized exoplanet in the habitable zone with the potential to sustain water in liquid form (Quintana et al., 2014). It should be noted here that although the first detection of a transiting exoplanet was announced in 2000, it took several more years for the first exoplanet discovery to be made using the transit method (Udalski et al., 2002; Konacki et al., 2003, see also Section 1.2). Another important milestone was the detection of an exoplanet orbiting the closest star to our Solar System, Proxima Centauri (Anglada-Escudé et al., 2016). Finally, the announcement in early 2017 of the discovery of seven Earth-sized exoplanets orbiting the M-dwarf TRAPPIST-1, of which three or four may be in the habitable zone, was also a momentous event in the history of the field (Gillon et al., 2017).

Moving on to planetary atmospheres, we trace the development of this field in the last two decades. Undoubtedly, the first ever observation of an exoplanetary atmosphere (Charbonneau et al., 2002) was a historic moment, but other discoveries have also been of great importance. For example, two other highlights include the first observation of an extended atmosphere in an exoplanet that experiences hydrogen

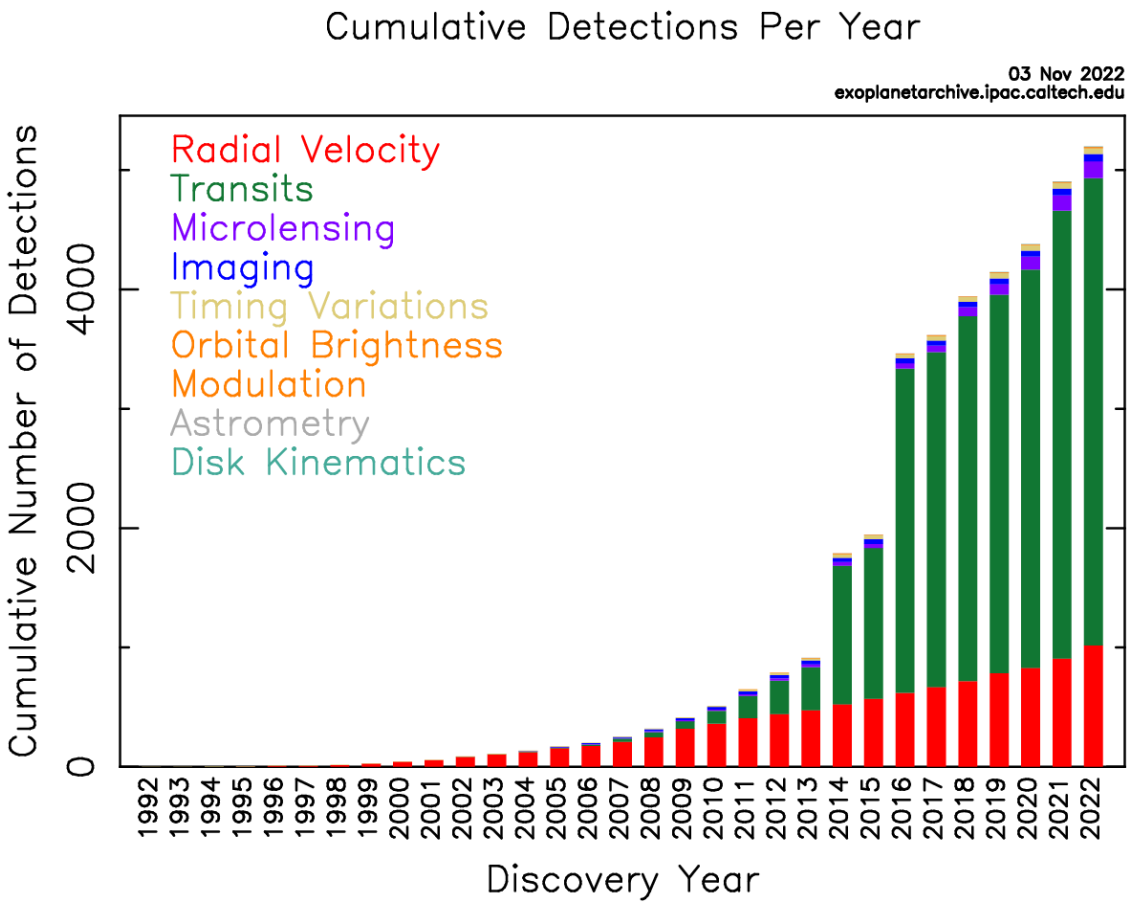

Figure 1.1: Cummulative number of discovered exoplanets binned according to year of detection. Different colours indicate different detection methods. Histogram from the NASA Exoplanet Archive (Accessed on 4/11/2022).

escape (Vidal-Madjar et al., 2003) and the first observation of the atmosphere of an Earth-sized exoplanet (de Wit et al., 2016). Of particular interest is the observation of the atmosphere of the hottest known exoplanet (see e.g. Yan & Henning, 2018; Hoeijmakers et al., 2018) as it gives an idea of the conditions that predominate in extremely irradiated environments. The detection of helium in the atmosphere of a hot Jupiter is another important achievement, as helium is the most common element after hydrogen (Spake et al., 2018).

1.1.2 Definition

Despite the fact that exoplanets have now been discovered by the thousand, the answer to the question "what is an exoplanet?" is still a work in progress. An early definition was given by the Working Group on Extrasolar Planets (WGESP) of the International Astronomical Union (IAU) in 2003 (Boss et al., 2007). This definition consisted of three main points:

1. If the mass of a celestial object is lower than the minimum mass needed for thermonuclear fusion of deuterium to take place, and if the object orbits a star or a stellar remnant, then this object is a *planet*. An extrasolar object should follow the same minimum mass rules[3] as the Solar System objects for it to be considered a planet.

2. If the mass of a celestial object is lower than the limit described in the previous point, and if the object is free-floating in a young star cluster, then this object is a *sub-brown dwarf*.

3. If the mass of a celestial object is higher than the limit described in the first point, but lower than the minimum mass needed for thermonuclear fusion of

[3]See the 2006 IAU definition of a planet:
`https://www.iau.org/static/resolutions/Resolution_GA26-5-6.pdf`

hydrogen to take place (Baraffe et al., 1998), then this object is a *brown dwarf*. Note that in this case the location of the object is unimportant.

According to the above, formation history is irrelevant in the definition of an exoplanet. For celestial objects with solar metal abundances, the minimum limit for thermonuclear fusion of deuterium to occur is set at $13\,\mathrm{M_{Jup}}$ (Spiegel et al., 2011). The above definition is now somewhat outdated, and a newer resolution was recently published (Lecavelier des Etangs & Lissauer, 2022). Based on the update, only the first point changes slightly:

- Planets can also orbit brown dwarfs.
- Planets need to have a mass ratio below ∼4% with the central object.

The mass ratio limit is directly derived from the L4/L5 instability:

$$\frac{M}{M_{\mathrm{central}}} < \frac{2}{25+\sqrt{621}} \approx \frac{1}{25}. \tag{1.1}$$

Here, L4 and L5 are the Langrange points at 60 degrees ahead and behind the planet's orbital path. The L4/L5 instability describes the maximum mass a celestial object can obtain at these points while still remaining in a stable orbit.

As a final note on this matter, it is worth mentioning that large exoplanet catalogues, such as the NASA Exoplanet Archive and the Exoplanet Encyclopedia, tend to use higher upper mass limits of 30 and $60\,\mathrm{M_{Jup}}$ for objects to be considered as exoplanets. This is not done by accident. The lower limit of $30\,\mathrm{M_{Jup}}$ refers to the limit placed by Soter (2006) for objects that have completed their accretion process around a stellar object. On the other hand, the higher limit of $60\,\mathrm{M_{Jup}}$ refers to the clear boundary observed between stars and lower-mass objects in the mass-density diagram presented by Hatzes & Rauer (2015).

1.1.3 Classification

To group the current population of exoplanets into categories, it may be useful to observe their distribution in a mass-period diagram. Figure 1.2 illustrates all current

exoplanets for which the mass and period are known. From this figure, one can easily discern three main clusters of exoplanets: hot Jupiters (upper left corner), warm Neptunes and super-Earths (lower left corner), and standard jovian planets (upper right corner). This distinction is actually over-simplified as the diagram does not reveal information about the temperature of the exoplanets or the type of star they are orbiting but in general it is a good approximation. Furthermore, it depicts a clear bias towards shorter periods for the detection of transiting exoplanets.

The leading proposition is that the exoplanets are divided into three broad categories: gas giant exoplanets, Neptune-like exoplanets and rocky exoplanets (Spiegel et al., 2014). The rocky exoplanets may be further divided into terrestrial and ocean exoplanets. However, in the literature, the prevailing terminology is based on the names of Solar System planets. One parameter that is frequently used as a distinguishing factor is the mass of the exoplanets. For example, Stevens & Gaudi (2013) present a mass-based system where planets are divided into Earths/Super-Earths ($0.1\,\mathrm{M}_\oplus - 10\,\mathrm{M}_\oplus$), Neptunes ($10\,\mathrm{M}_\oplus - 100\,\mathrm{M}_\oplus$), Jupiters ($100\,\mathrm{M}_\oplus - 10^3\,\mathrm{M}_\oplus$), and Super-Jupiters ($3\,\mathrm{M}_{\mathrm{Jup}} - 13\,\mathrm{M}_{\mathrm{Jup}}$). But other names such as Saturn and Venus may also be used depending on the properties of the exoplanet (see e.g. Nikolov et al., 2018b; Kane et al., 2018).

The huge number of transiting exoplanet discoveries gave rise to size-based classification schemes such as the one presented by Borucki et al. (2011). According to this classification scheme, exoplanets are divided into groups based on their radius: Earths ($< 1.25\,\mathrm{R}_\oplus$), Super-Earths ($1.25\,\mathrm{R}_\oplus - 2\,\mathrm{R}_\oplus$), Neptunes ($2\,\mathrm{R}_\oplus - 6\,\mathrm{R}_\oplus$), Jupiters ($6\,\mathrm{R}_\oplus - 15\,\mathrm{R}_\oplus$), and other exoplanets up to twice the size of Jupiter ($15\,\mathrm{R}_\oplus - 22.4\,\mathrm{R}_\oplus$). Combining mass and radius, a somewhat simpler but unconventional categorisation scheme, already mentioned in Section 1.1.2, relies on the properties of the planet mass versus planet density diagram (Hatzes & Rauer, 2015). Under this diagram, exoplanets can be split into low-mass exoplanets and gas giant exoplanets based on a minimum density observed at $\sim 0.3\,\mathrm{M}_{\mathrm{Jup}}$. Furthermore, a peak at $\sim 60\,\mathrm{M}_{\mathrm{Jup}}$ in the mass-density relation indicates that objects above this mass are likely objects of stellar nature. This interpretation reclassifies brown dwarfs as high-mass gas giant exoplanets, but is not

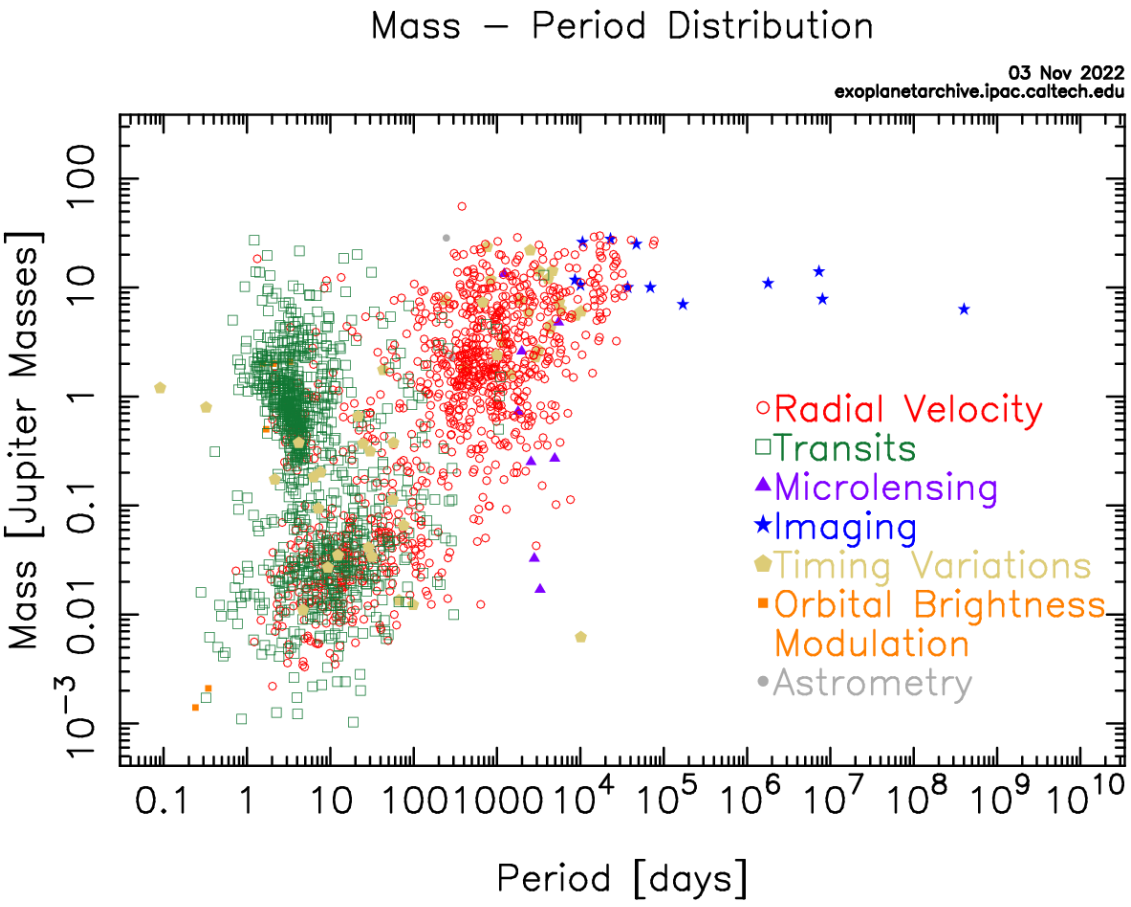

Figure 1.2: Mass-period diagram of detected exoplanets. This is the distribution for exoplanets that have calculated estimates for their period and their mass or minimum mass. There are three clusters: hot Jupiters (upper left corner), warm Neptunes and super-Earths (lower left corner), and standard jovian planets (upper right corner). Plot from the NASA Exoplanet Archive (Accessed on 4/11/2022).

included in the IAU definition presented in Section 1.1.2.

Apart from mass and size, planetary temperature can also be used to distinguish between different groups of exoplanets and specifically it can divide them into hot, warm and cold. This parameter is not so important when describing types of exoplanets, but can play a vital role in the chemical composition of a planetary atmosphere (see Section 1.5). In Chapter 4, I use the term "hot Saturn" to describe the exoplanet WASP-88b, even though "hot Jupiter" would be equally fine. The reason I chose to use this term instead is because the very low-density of this exoplanet and the relatively lower mass compared to Jupiter make it look more Saturn-like.

1.2 Transiting exoplanets

Due to the focus of this thesis on transiting gas giant exoplanetary atmospheres, it is crucial to understand a few basic characteristics of the transit method. This section is, therefore, concentrated on this specific detection technique. Despite this emphasis, it should be noted that other detection techniques have also had some success in the discovery of new exoplanets. As high as 99.5% of all exoplanets have been detected with the transit method ($\sim$76% of all discoveries) and the following three methods: radial velocities (e.g. Mayor & Queloz, 1995), direct imaging (e.g. Chauvin et al., 2004), and gravitational microlensing (e.g. Bond et al., 2004). Another promising technique expected to contribute with exoplanet detections in the coming years is astrometry, with the GAIA mission already revealing over 70 potential exoplanet candidates (Holl et al., 2022). Figure 1.2, presented earlier, also shows the detection technique used for each exoplanet in the distribution.

1.2.1 The transit method

The transit method is based on the concept that the amount of light we receive from an observed stellar object is reduced when an orbiting exoplanet passes in front of it

(Winn, 2010). This happens when the orbit of the exoplanet is aligned with our line of sight, so that the disk of the exoplanet can cover a portion of the stellar disk for a brief period. The obscuration of the star caused by this event, known as *transit*, is visible in the time-series observations of the stellar flux and is seen as a periodic dip in brightness. The phenomenon normally has a relatively short duration compared to the orbital period of the exoplanet. A smaller dip in brightness can also occur when the exoplanet passes behind the star and this is known as *occultation*. Figure 1.3 illustrates the effect of both phenomena on incident flux. The small decrease in observed flux is depicted by the transit light curves and can be used to determine the bulk and orbital properties of the transiting exoplanet.

One defining characteristic of the transit method is the *impact parameter* b. This parameter describes the sky-projected distance between the centres of the two celestial bodies (i.e. the star and the planet) when the conjunction between Earth and the two bodies occurs. Figure 1.4 shows a schematic of the star and the orbiting exoplanet and reveals a relation between the impact parameter and the orbital semi-major axis a and inclination i. This relation is defined as:

$$bR_* = a\cos i \Leftrightarrow b = \frac{a}{R_*}\cos i, \tag{1.2}$$

where R_* is the radius of the star. If the transit is central, then $b = 0$, and, if the centre of the planetary disk touches the edge of the stellar disk, then $b = 1$.

An interesting inference from Equation 1.2 and the two defined boundaries is the probability of transit. According to the above, a transit happens when $b < 1 - \frac{R_\mathrm{p}}{R_*}$, and a partial eclipse[4] happens when $1 - \frac{R_\mathrm{p}}{R_*} < b < 1 + \frac{R_\mathrm{p}}{R_*}$. In these equations, R_p represents the radius of the planet. Replacing impact parameter b with inclination i and rearranging gives:

$$\cos i < \frac{R_* - R_\mathrm{p}}{a} \quad \text{(for a transit)}, \tag{1.3}$$

$$\cos i < \frac{R_* + R_\mathrm{p}}{a} \quad \text{(for a partial eclipse)}. \tag{1.4}$$

[4] *Eclipse* is the general term used to describe the obscuration of a celestial object by another.

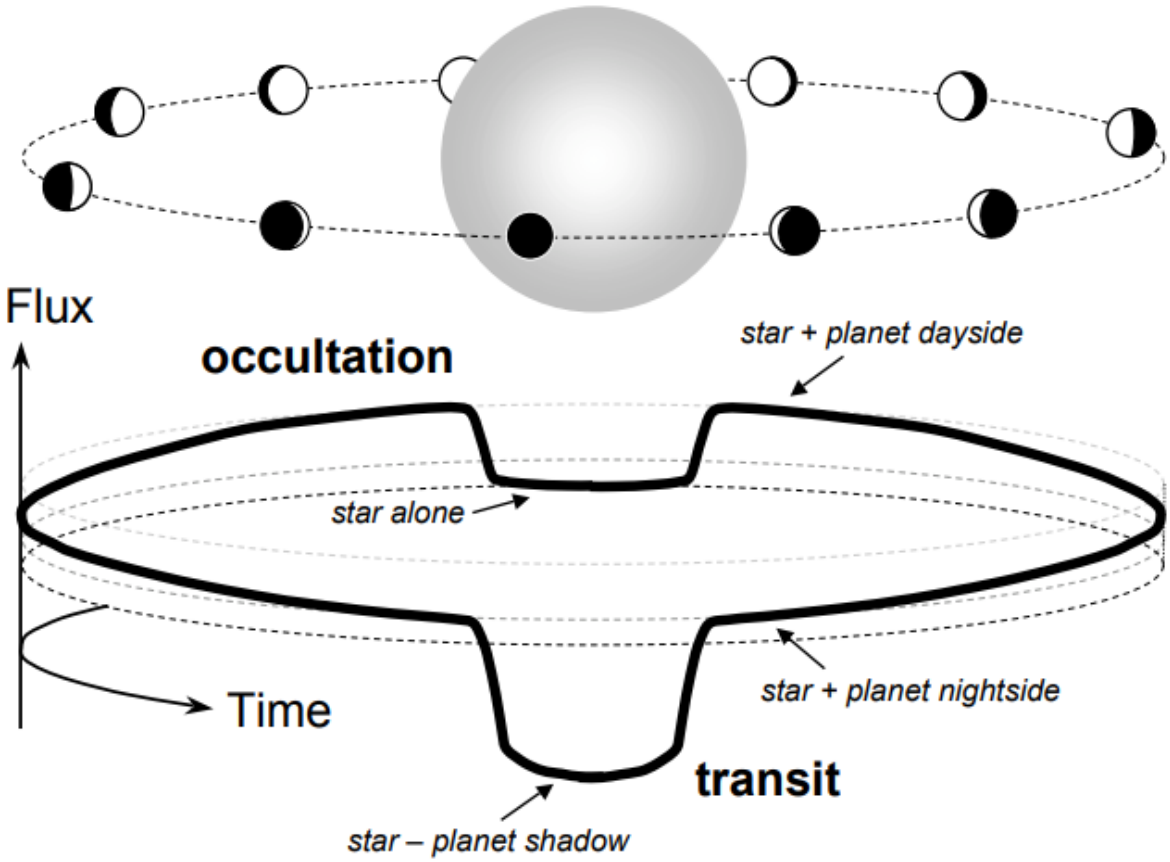

Figure 1.3: A schematic depicting a transiting planet in various orbital positions. The expected full transit light curve is also given below. It can be seen that, during transit and occultation, the passage of the transiting planet causes a drop in observed flux. Image from Winn (2010).

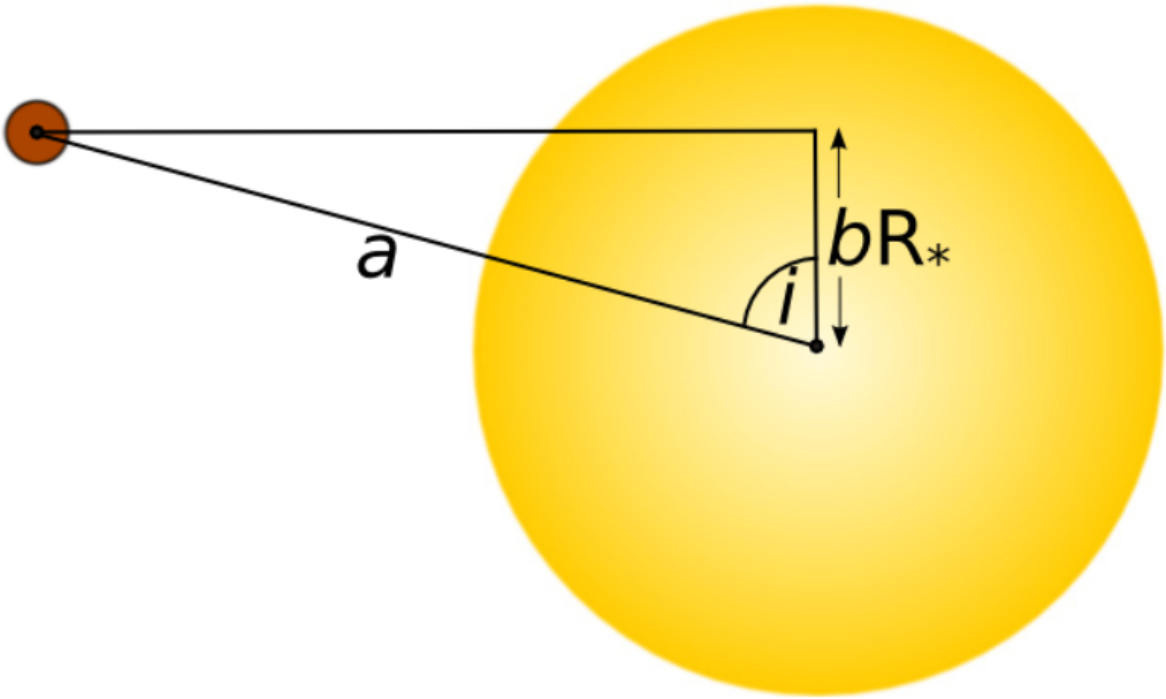

Figure 1.4: The impact parameter b and its relation to the orbital semi-major axis a, the radius of the star R_* and inclination i. Figure from `https://www.paulanthonywilson.com/` (Accessed on 2/11/2022).

From Equations 1.3 and 1.4 it becomes apparent that the probability of transit increases for smaller orbital separations. This is the reason why many transiting systems are found with small orbital separations, and, as a consequence, with small orbital periods (see also Figure 1.2).

1.2.2 Transit light curves

Figure 1.5 shows the effect of the transit on the observed flux. The schematic depicts four contact points (t_I, t_{II}, t_{III} and t_{IV}) on the transit light curve. These indicate different moments during the transit event. More specifically, the steady decrease between points t_I and t_{II} is called *ingress* and describes the beginning of the transit from the point of first contact between the two disks (t_I) till the point where the disk of the planet fully overlaps with the stellar disk for the first time (t_{II}). On the other side, a steady increase can be observed between points t_{III} and t_{IV}. This is called *egress* and depicts the end of the transit, with t_{III} being the last time where the disk of the planet fully overlaps with the stellar disk and t_{IV} being the point where disk overlap ends. Although transits show four contact points, it should be noted that this is not the case for all transiting exoplanets. If the eclipse is partial, then ingress is directly followed by egress and the contact points are reduced to two.

The interval from t_{II} to t_{III}, where the light curve appears flat, gives the duration of totality (t_{tot}). During this interval, the planetary disk is completely encompassed by the stellar disk. The duration of totality is associated with the period P and parameters a, i, R_* and R_p via the following approximation:

$$t_{tot} = t_{III} - t_{II} \approx \frac{P}{\pi}\sqrt{\left(\frac{R_\star - R_\mathrm{p}}{a}\right)^2 - \cos^2 i}. \tag{1.5}$$

Another useful interval is the interval between the first (t_I) and last (t_{IV}) contact points. This interval defines the full duration of the transit (t_{tra}) and is given by a similar approximation:

$$t_{tra} = t_{IV} - t_I \approx \frac{P}{\pi}\sqrt{\left(\frac{R_\star + R_\mathrm{p}}{a}\right)^2 - \cos^2 i}. \tag{1.6}$$

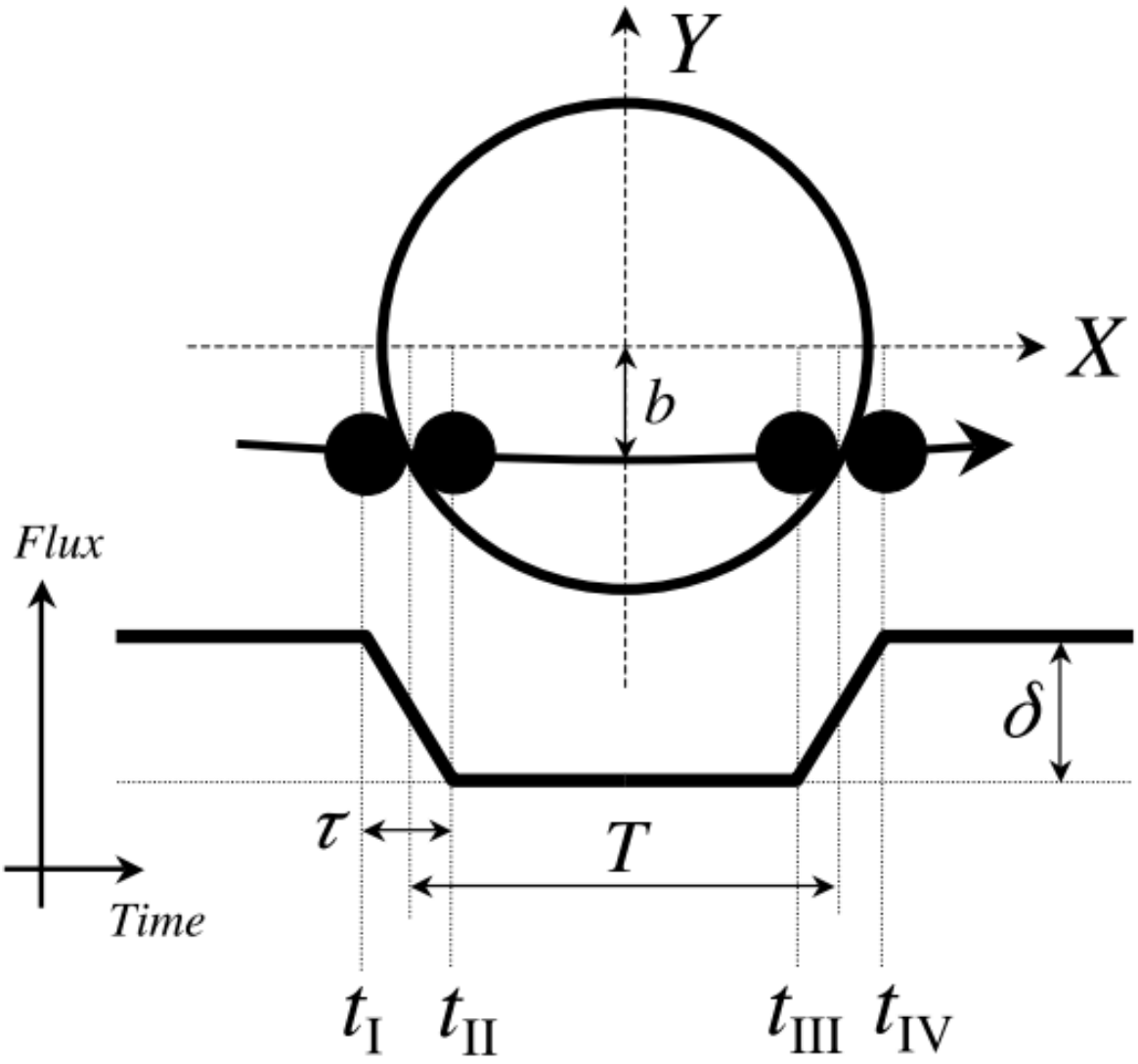

Figure 1.5: A schematic of a transit light curve with all four contact points included (t_I, t_{II}, t_{III} and t_{IV}). Quantities τ and T give the time interval of ingress/egress and the time interval $T = t_{tra} - \tau$, respectively. t_{tra} is the duration of the transit. The transiting exoplanet causes a reduction in flux that is equal to the transit depth δ. Illustration from Winn (2010).

Both approximations are accurate under the assumption that the orbit is circular.

One final but important parameter worth mentioning is the transit depth δ. The parameter is also depicted in Figure 1.5 and expresses the fractional decrease in brightness caused by the transiting planet. The transit depth is connected to the radii of the planet and the star through the following expression:

$$\delta = \left(\frac{R_{\rm p}}{R_*}\right)^2. \tag{1.7}$$

Exact derivation of this parameter from the fractional decrease in brightness is given in Section 1.4.3, where the additional input from the planetary atmosphere is also discussed. The expression is correct for circular orbits when limb darkening and contamination from nearby sources are considered to be negligible. According to the above relation, the transit depth simply is the square of the ratio between the planetary radius and the stellar radius. This means that the transit depth is quite a useful parameter as it can disclose information about the size of an exoplanet.

1.2.3 Limb darkening

Figure 1.5 shows more of an ideal case where contact points II and III are well-defined. In reality, the change from ingress to totality and from totality to egress can be less clear sometimes as the transition can often be smoothed due to a phenomenon known as *limb darkening*. This apparent phenomenon describes the drop in surface brightness as we observe the stellar disk from its centre to its limb.

The reason for the centre-to-limb drop in brightness is the change in the observed stellar photosphere depth with viewing angle and, as a consequence, the change in the observed temperature. More specifically, at the centre of the stellar disc, we attain an optical depth of one at a deeper layer within the stellar atmosphere compared to the limb and so we observe the star at a greater depth where the gas temperature is higher. As our line of sight moves towards the limb, however, our view is restricted to shallower depths where the gas is cooler and so the edge of the stellar disk appears darker. The effect of limb darkening depends on the angle θ between the surface normal, which is a

line perpendicular to the stellar surface, and our line of sight, and is typically defined by the quantity $\mu = \cos\theta$. By definition, μ is equal to 1 at the centre of the stellar disk and 0 at the edge of the disk.

Several limb darkening laws exist in the literature to explain the intensity of light emitted by the surface of the star as a function of μ. The most sophisticated of them is the non-linear, four parameter law first presented by Claret (2000):

$$\frac{I(\mu)}{I(1)} = 1 - \sum_{n=1}^{4} u_n \left(1 - \mu^{\frac{n}{2}}\right). \tag{1.8}$$

Here, $I(\mu)$ is the light intensity of the stellar surface at an angle defined by μ, $I(1)$ is the light intensity at the centre of the stellar disk, and u_n are limb darkening coefficients that are associated with the strength of the dimming. This law offers great accuracy but includes too many parameters and can be slow to compute without providing much difference in the estimated parameters (see also Chapter 4). Furthermore, the added complexity is unnecessary in ground-based observations that may also be affected by the Earth's atmosphere.

In practice, another law with fewer parameters is widely used in ground-based observations to increase computational speed whilst simultaneously keeping results relatively similar. This law is termed as the quadratic limb darkening law (Kopal, 1950) and is expressed as:

$$\frac{I(\mu)}{I(1)} = 1 - u_1(1-\mu) - u_2(1-\mu)^2, \tag{1.9}$$

where u_1 and u_2 are the linear and quadratic limb darkening coefficients, respectively.

Limb darkening plays a pivotal role in transmission spectroscopy as it greatly depends on the gas properties of the stellar surface. This implies that, apart from temperature, wavelength can also affect the limb darkening coefficients and, as a result, the shape of the transit light curves (see e.g. Knutson et al., 2007). Limb darkening tends to be stronger at shorter/bluer wavelengths, increasing the curvature at the bottom of the light curves and causing difficulties in the determination of the second and third contact points. At longer/redder wavelengths, however, the light curves tend

to preserve most of the characteristic trapezoidal shape seen in Figure 1.5. Finally, the values of the limb darkening coefficients can also be influenced by other stellar properties such as surface gravity and metallicity (Magic et al., 2015).

1.3 Hot gas giant exoplanets: a peculiar group

1.3.1 Formation and evolution

The leading theories for the mechanisms behind the formation of hot Jupiters and other close-in, gas giant exoplanets, are two. The first one is the core accretion scenario and the second one is the gravitational instability scenario. In the first scenario, pre-planetary embryos are initially formed via collisions between dust particles in the proto-planetary disk. These embryos tend to become more massive and start to form planetesimals, which accrete more solid material until a rocky proto-planet core is formed. This is followed by a transition period, until a runaway gas accretion mechanism kicks in. This leads to a substantial increase in mass until the formation process finally comes to an end when the rate of accretion slows down (Perri & Cameron, 1974; Pollack et al., 1996; Chabrier et al., 2014). In the second scenario, a gravitational instability in cool, dense areas of the proto-planetary disc causes material to collapse into self-gravitating clumps, which then coalesce to form the exoplanet (Boss, 1997; Durisen et al., 2007).

Hot, gas giant exoplanets may either form in-situ or ex-situ, with ex-situ formation being more plausible. In in-situ formation, hot exoplanets form at their present day distance from the host star (Boley et al., 2016; Batygin et al., 2016). Gravitational instability requires efficient cooling mechanisms and is therefore unable to support the formation of an exoplanet in an environment close to the stellar core (Rafikov, 2005). Core accretion also struggles in-situ due to the smaller amount of material in short orbital paths and the difficulty in attracting material from further out in the disc due to the smaller initial mass (Schlichting, 2014; Johansen & Lambrechts, 2017). Never-

theless, it may still be possible for core accretion to work and for the proto-planetary core to reach the critical mass of $\sim 10\,M_{\oplus}$ required to commence gas accretion and form a gas giant exoplanet (Batygin et al., 2016).

In the case of ex-situ formation, hot Jupiters form further out in the proto-planetary disc and then migrate to their current location through disk or disk-free migration. In disk migration, a gas giant exoplanet starts forming in the outer bounds of the proto-planetary disk and then moves inward before the dissipation of the disk (Goldreich & Tremaine, 1980; Lin & Papaloizou, 1986; Ida & Lin, 2008; Baruteau et al., 2014). This migration mechanism is established on the gravitational interaction between the disc and the migrating exoplanet and the transfer of angular momentum between them. In some cases, migration may be unsuccessful leading to the consumption of the gas giant by the host star (Jackson et al., 2009), but tidal stellar interaction may play a part in the observed survival rate (Heller, 2019). In disk-free/eccentric/tidal migration, a gas giant exoplanet moves inward after the dissolution of the proto-planetary disk and gets excited into a highly elliptical orbit (Rasio & Ford, 1996; Dawson & Johnson, 2018). Reasons for the inward migration include secular interactions between the exoplanet and other celestial bodies (e.g Wu & Murray, 2003; Fabrycky & Tremaine, 2007; Wu & Lithwick, 2011) and planet-planet scattering (e.g Rasio & Ford, 1996; Chatterjee et al., 2008). After the exoplanet obtains a sufficiently elliptical orbit, it starts to lose energy due to tidal forces from the host star, and so its orbit decreases by becoming circular and getting closer to the heavier mass object (Eggleton et al., 1998).

Figure 1.6 depicts the three possible pathways in the formation and evolution of hot gas giant exoplanets based on our current knowledge.

1.3.2 Radius inflation

One key physical parameter that characterises an exoplanet is its radius. An interesting finding from the first transit explorations is that the radius of hot gas giant exoplanets is somewhat larger than expected (see e.g. Charbonneau et al., 2000; Bakos et al.,

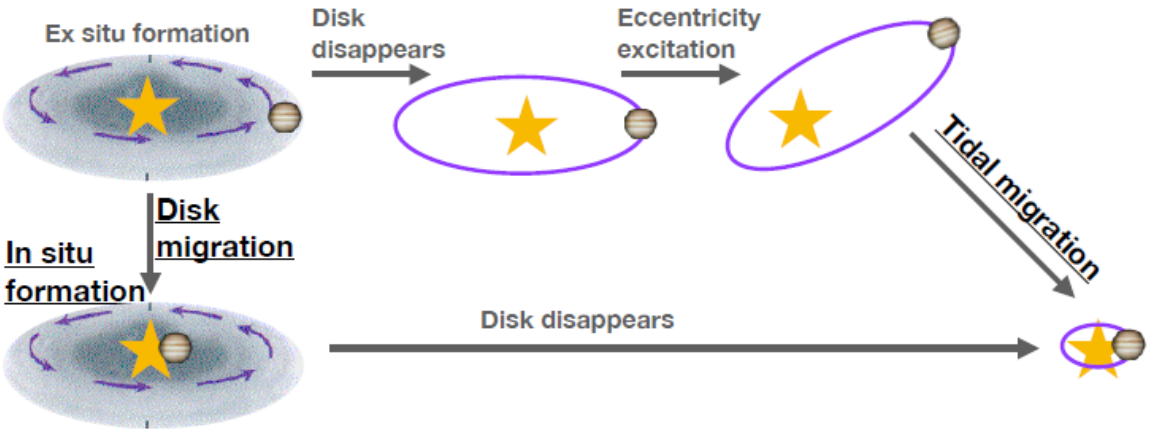

Figure 1.6: The origins of hot gas giant exoplanets based on our current understanding. Three possibilities exist: in situ formation, disk migration or tidal migration. Figure from Fortney et al. (2021).

2007). This characteristic often leads to low mean densities, which in extreme cases can be less than 10% the mean density of Jupiter (see e.g. WASP-17, Anderson et al., 2010, and WASP-127, Lam et al., 2017). This discrepancy in the radius of hot gas giant exoplanets is commonly referred to as radius inflation or the radius anomaly (Burrows et al., 2007) and evolutionary models have been trying to address the issue since the first transit detections (see e.g. Bodenheimer et al., 2001; Guillot & Showman, 2002).

Radius inflation depends on incident flux and the higher the equilibrium temperature of a gas giant exoplanet the likelier it is for the radius to show a positive deviation from theoretical models that are not adjusted for inflation (Laughlin et al., 2011). Figure 1.7 reveals the size of this effect by depicting the departure of the planetary radius from theoretical predictions as a function of incident flux. Interestingly, the increase in the planetary radius seems to come into a halt after 1500 K. This result suggests that the mechanism behind the radius inflation is less efficient at higher temperatures (Thorngren & Fortney, 2018). Furthermore, at temperatures below 1000 K the planetary radius does not seem to experience inflation but may experience deflation. This can be explained by a denser core and an increased metallicity in cooler exoplanets

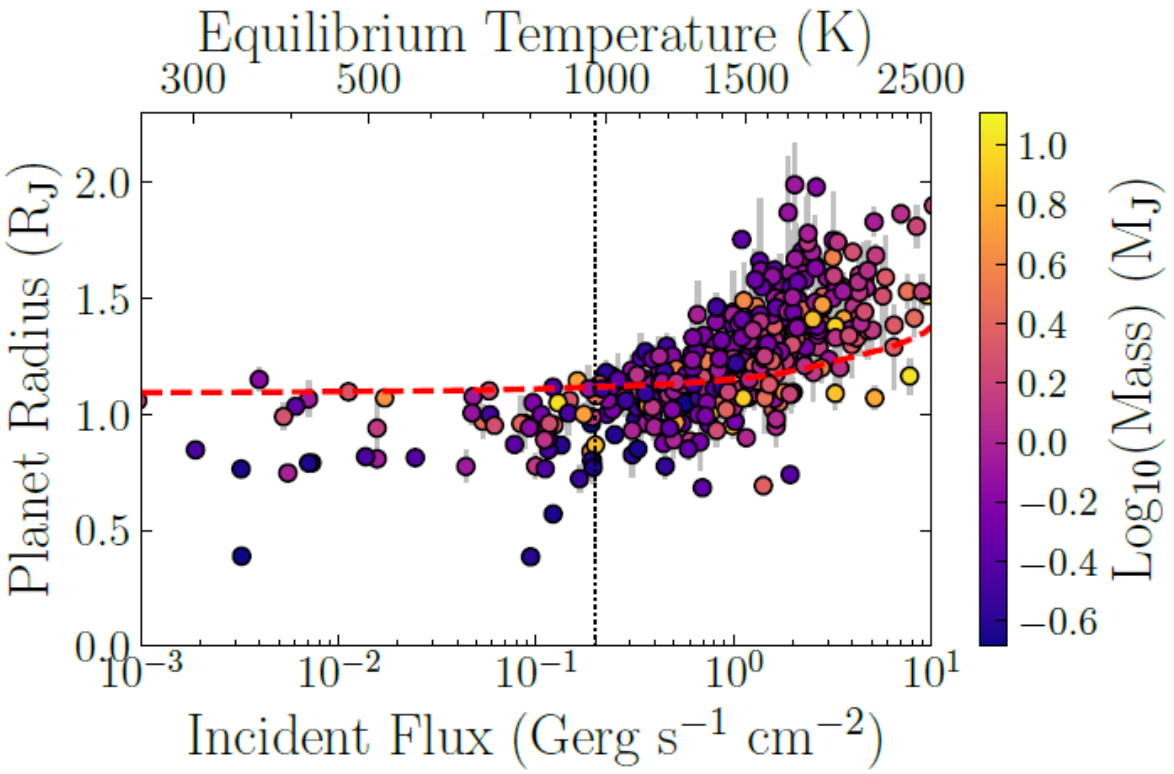

Figure 1.7: The radius anomaly of hot gas giant exoplanets. A diagram of the planet radius as a function of equilibrium temperature (or incident flux). The red dashed line illustrates the predicted radius from an evolutionary model of a Jupiter-mass exoplanet without inflation effects taken into account. The black vertical dotted line indicates the temperature at which the radii start to deviate from the model. Inflation occurs at higher temperatures. Figure from Fortney et al. (2021).

(Burrows et al., 2007).

Proposed solutions for the radius inflation problem can be divided into two main groups. In the first group, internal heating mechanisms use the incident stellar flux or tidal energy to keep the radius inflated. For example, incident stellar flux can be converted into kinetic energy leading to strong circulation patterns that can then produce thermal energy and invoke an extended atmosphere (Guillot & Showman, 2002). A similar mechanism to this is large-scale vertical mixing where kinetic energy is again converted to thermal energy (Youdin & Mitchell, 2010). Stellar irradiation

can also lead to the ionisation of some chemical species. Zonal winds of ionised species can then interact with the magnetic field of the hot gas giant exoplanet leading to the creation of an electric current towards the interior of the exoplanet. As the current moves deeper into the planet it encounters resistance and so electric energy is converted into heat in a phenomenon that is known as Ohmic dissipation (Batygin & Stevenson, 2010). Lastly, the interior of a hot gas giant exoplanet can also be heated through tidal dissipation due to eccentricity tides caused by disk-free migration (Bodenheimer et al., 2001) or due to thermal tides (Arras & Socrates, 2010).

In the second group, mechanisms that delay the cooling of the exoplanet can also keep the radius enlarged. According to planetary evolution models, the temperature in the interior of a hot gas giant exoplanet is connected to the radius, and, as heat reduces after formation, the radius is naturally expected to shrink over time (Arras & Bildsten, 2006). However, in some cases, the rate of cooling and, therefore, the rate of radius contraction, can be altered. This can happen, for instance, when atmospheric opacities are enhanced due to e.g. haze or super-solar metallicities (Burrows et al., 2007), or when there is an increasing gradient of molecular weight with decreasing distance from the planet core that affects convection and reduces heat transport (Chabrier & Baraffe, 2007). Both these mechanisms slow the cooling process and inhibit the reduction in size.

A caveat of the second group of mechanisms is that it cannot satisfactorily explain the observed correlation between insolation and radius inflation. Based on our current understanding of all the aforementioned mechanisms, the most prominent theory is that of Ohmic dissipation as it can also explain the stabilisation of the radius at higher temperatures (Thorngren & Fortney, 2018). Even so, it is very likely that multiple mechanisms play a role as recently suggested by Sarkis et al. (2021).

1.3.3 Occurrence rates

Given the increasing number of exoplanet discoveries, it may be useful to determine the planet occurrence rates, as well as trends in the properties of planet-hosting systems.

Such data can tell us more about the leading mechanisms that define the formation and evolution of exoplanets and can immensely help with the future planning of exoplanet detection surveys.

Current observations show that gas giant exoplanets are frequently found to have high eccentricities (e.g. Carrera et al., 2019). Furthermore, gas giant exoplanets tend to orbit stars with high metal abundances (e.g. Gonzalez, 1997; Fischer & Valenti, 2005). Lower-mass exoplanets, however, do not seem to follow the same relation between host star metal content and planet occurrence rate (e.g. Buchhave et al., 2014) and tend to occur more frequently in densely-packed multiplanet systems (Mayor et al., 2011).

When it comes to occurrence rates, most studies are based on transit and radial velocity surveys, as these two techniques alone account for ∼95% of all exoplanet detections (NASA Exoplanet Archive). According to the works of Cumming et al. (2008) and Mayor et al. (2011), it is estimated that about 10% of Sun-like stars host at least one gas giant exoplanet that orbits with a period of ≤ 2000 d. The percentage is increased for Neptune-like and smaller exoplanets, with up to 50% of Sun-like stars potentially being orbited by such planets (Mayor et al., 2011).

The situation is changed when we reduce the sample to hot Jupiters. Although many early transit and radial velocity studies found a relatively high number of hot Jupiters compared to other planets (see Section 1.1.1), this was mainly a selection effect and more recent radial velocity studies suggest that the occurrence of hot Jupiters is low and around ∼1% for main-sequence stars (e.g. Mayor et al., 2011; Wright et al., 2012). The occurrence rate is found to be slightly lower from transit surveys, at around ∼0.5% (e.g. Fressin et al., 2013; Zhou et al., 2019). This discrepancy can be explained by stellar multiplicity and when correcting for binary systems the percentage increases to the value obtained from radial velocity surveys (e.g. Beleznay & Kunimoto, 2022).

Early analyses found that hotter stars tend to host fewer exoplanets, and that the cooler the stars are, the higher the likelihood is for them to be hosting exoplanets (e.g. Howard et al., 2012). However, according to a more recent study, the distribution seems to peak for G-type stars and is decreasing for cooler and hotter stars in the AFGKM regime (Gan et al., 2023). The decrease looks sharper for cooler M-type

stars.

1.4 Observing exoplanetary atmospheres

There are three main techniques to observe an exoplanetary atmosphere, and all of them are based on methods already used in the detection of exoplanets. These are direct imaging, high-resolution Doppler spectroscopy and transit spectroscopy. Direct imaging is primarily concerned with young, gas giant exoplanets in wide orbits (see e.g. Wang et al., 2021). These planets tend to emit light in the infrared as they cool off while being in the last stages of their formation process. High-resolution Doppler spectroscopy is based on the radial velocity method and relies on the Doppler shifts caused by the motion of the exoplanet (see e.g. Brogi et al., 2016). This technique focuses on time-varying signals as a function of wavelength, and, in that respect, it shares a similarity with transit spectroscopy. Transit spectroscopy, as the name suggests, is based on the transit method and therefore describes the periodic changes in stellar brightness caused by a transiting exoplanet (see e.g. Seager & Sasselov, 2000; Sudarsky et al., 2000). This technique of atmospheric observation is sensitive to large, close-in exoplanets, and just like all aforementioned methods it depends on slight, wavelength-dependent variations in the observed spectrum.

Transit spectroscopy can be further divided into three sub-categories depending on the orbital position of the exoplanet during the time of the observation. These are transmission spectroscopy, occultation (or emission/eclipse) spectroscopy and phase curves. Figure 1.8 illustrates the two main orbital positions at which an exoplanet can be observed. These are in front of the host star (transmission spectroscopy) and behind it (occultation spectroscopy). Phase curves are created when the transiting exoplanet is observed during the whole orbit, including during transit and secondary eclipse.

Occultation spectroscopy is the multi-wavelength observation of an exoplanet as it passes behind its host star. During this event, we observe the atmospheric conditions on the dayside of the exoplanet. The spectrum obtained in this case can be

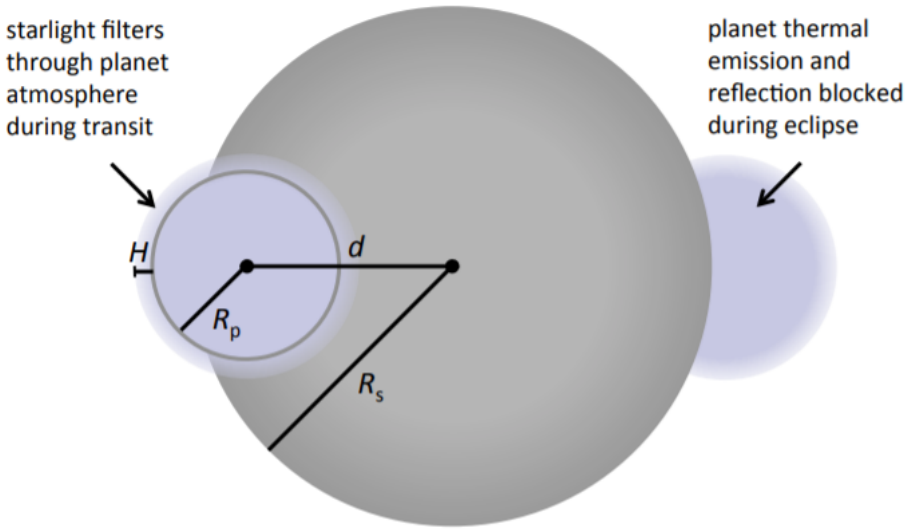

Figure 1.8: A simple sketch of an exoplanet and its atmosphere during the events of transit and eclipse. The exoplanet is observed in transmission during transit and in emission during occultation. The depicted quantities are the atmospheric pressure scale height H, the radius of the exoplanet R_p, the radius of the host star R_*, and the sky separation d between the centres of the two discs. Figure from Kreidberg (2018).

further categorised depending on the spectral region observed, with optical wavelengths being more sensitive to reflected light and infrared wavelengths being more sensitive to thermal emission. The first detections of thermal emission were made by Deming et al. (2005) and Charbonneau et al. (2005), while the first direct detection of reflected light from ground-based facilities was made by Martins et al. (2015). Occultation spectroscopy is a particularly useful technique as it can also indicate the dayside atmospheric temperature of an exoplanet as a function of pressure. An emission (or eclipse) spectrum is therefore quite valuable as it can give information about increasing or decreasing vertical gradients in temperature and, hence, reveal potential mechanisms that control the structure of the atmosphere.

Phase curves are basically the complete transit light curves of an exoplanet's orbit, as they contain transit and secondary eclipse data, as well as out-of-transit data from other orbital positions. In that way, phase curves can observe an exoplanet at all orbital phases. The first phase curve observation was announced by Harrington et al. (2006). The power of this technique lies in the full longitudinal scan of an exoplanet due to it being observed at various stages in its orbital path. This geographical mapping is possible because close-in exoplanets have synchronised their rotational period to their orbital period (Guillot et al., 1996; Rasio et al., 1996). Phase curves are an excellent method to explore wind patterns and other atmospheric dynamics, including cloud formation as a function of longitude, and temperature variations between day and night and between deeper and shallower pressure layers.

Finally, transmission spectroscopy is established on the observation of a transit event as a function of wavelength (see also Seager & Sasselov, 2000). Transmission spectroscopy will be described in more detail in the following paragraphs as it is the technique my research is based on. The reader can learn more about other techniques in recent reviews (e.g. Madhusudhan, 2019, and references therein).

1.4.1 Transmission spectroscopy

Transmission spectroscopy is founded on the notion that the periodic drop in flux caused by a transiting planet varies with wavelength. Most of the decrease in brightness is caused by the bulk of the exoplanet, but part of the light is also filtered through the planetary atmosphere. Chemical species that make up the planetary atmosphere may absorb or scatter incident stellar light and hence block part of it at wavelengths defined by the properties of the atmospheric constituents (Seager & Sasselov, 2000). This means that the atmosphere will appear more opaque in some wavelengths and less so in others. Extensive information about the potential opacity sources is given in Section 1.5.

The minute, wavelength-dependent variations in the intensity of starlight as it interacts with the planetary atmosphere can be measured via two basic techniques. These two techniques are photometry and spectroscopy and are distinguished by the light collection properties of the instruments used. In photometry light is collected using broad multi-colour passbands (see e.g. Mancini et al., 2019), while in spectroscopy light is first refracted through the use of dispersive elements and then grouped backed together into many narrow wavelength channels (see e.g. Nikolov et al., 2018b). It is clear that both techniques are able to explore the full electromagnetic spectrum, but ground-based observations are limited by the Earth's atmosphere due to opacity sources in the long radio wavelengths, in the infrared, and at shorter wavelengths beyond near-ultraviolet (Barmby, 2019).

Photometry often requires much shorter exposure times compared to spectroscopy, unless photometry is performed using the telescope defocusing technique (Southworth et al., 2009). In addition, photometry has a lower spectral resolving power than spectroscopy and is therefore less effective in extracting fine details such as the absorption signatures of atomic elements. In this thesis, I focus my analysis on spectroscopic data obtained from the ground, but I also share some space-based photometry results produced by John Southworth to provide context for the target presented in Chapter 4. In this case, the photometry is not obtained in multiple colour bands and so it cannot

be used for transmission spectroscopy.

One key characteristic of spectroscopic data is that the dispersed light is binned into narrow spectrophotometric channels to construct a sequence of transit light curves as a function of wavelength. These light curves are then fitted separately or simultaneously to obtain the transit depth at each wavelength, which defines the relative change in the apparent size of an exoplanet due to the presence of a planetary atmosphere. The transit depth (or its square root) is a valuable parameter as it can be plotted against wavelength and reveal the desired *transmission spectrum*.

Transmission spectroscopy is credited with the first detection of an exoplanetary atmosphere (Charbonneau et al., 2002) and has been of paramount importance in the field with many observations of alien planetary atmospheres beyond our Solar System, especially in the infrared (see e.g. Edwards et al., 2022). In the following subsections, I will concentrate on the atmospheric pressure scale height and the atmospheric (or transmission) signal, two characteristic properties of atmospheric characterisation via transmission spectroscopy.

1.4.2 Atmospheric pressure scale height

The atmospheric pressure scale height H is a height scale defined between two atmospheric layers and expresses the change in the atmospheric pressure by a factor of e. It is an important parameter in transmission spectroscopy as it affects the amplitude of the observed spectral features. The parameter can be derived by combining the ideal gas equation:

$$PV = Nk_{\mathrm{B}}T_{\mathrm{eq}}, \tag{1.10}$$

and the condition for hydrostatic equilibrium:

$$P = \rho g_{\mathrm{p}} H. \tag{1.11}$$

In these equations, P denotes the absolute pressure, V is the gas volume, N is the number of atmospheric particles, k_B is the Boltzmann constant, T_{eq} is the equilibrium temperature, ρ is the density and g_{p} is the gravity of the planet.

Joining the two equations together and rearranging with respect to H, we obtain:

$$H = \frac{N k_{\mathrm{B}} T_{\mathrm{eq}}}{\rho g_{\mathrm{p}} V}, \tag{1.12}$$

This equation can be simplified by noting that ρV is equal to the total mass M and by replacing the number of particles N with the ratio of the total mass over the mean molecular mass ($\frac{M}{\mu}$). The total mass can then be crossed out, leading to:

$$H = \frac{k_{\mathrm{B}} T_{\mathrm{eq}}}{\mu g_{\mathrm{p}}}. \tag{1.13}$$

The atmospheric pressure scale height is directly proportional to the atmospheric signal (see Section 1.4.3) and as such it is an important indicator of atmospheric detectability.

One notable parameter in Equation 1.13 is the mean molecular mass. This parameter is inversely proportional to the atmospheric pressure scale height and is a useful descriptor of the atmospheric composition. For Jupiter-like exoplanets with a chemical abundance similar to the Sun, the atmospheres are expected to be governed by hydrogen and helium. The volume mixing ratio of He/H_2 should, therefore, be close to the value of 0.158 observed for the jovian atmosphere (Taylor et al., 2004). Such a composition would give a mean molecular mass of ∼2.3 a.m.u., which is more than twice the value of pure hydrogen. For reference, an atmosphere with 100 times the solar metallicity will have a mean molecular mass of ∼4.3 a.m.u. (Drummond et al., 2018).

1.4.3 Atmospheric signal

If the atmospheric pressure scale height of an exoplanet is known, we can then obtain the atmospheric signal or, more precisely, the transit signal due to absorption and/or scattering of light from the atmospheric constituents. This is a measure of the potential strength of spectral lines in the transmission spectrum and is defined as the variation in the transit depth caused by one atmospheric pressure scale height.

The parameter can be determined from the reduction in incident starlight during

transit. This can be expressed as follows:

$$\Delta D = \frac{F_{out} - F_{in}}{F_{out}}, \tag{1.14}$$

where $F_{out} = I\pi R_*^2$ is the out-of-transit flux of the system and $F_{in} = I\pi(R_*^2 - R_{\rm p}^2)$ is the in-transit flux of the system. Here, I refers to the intensity of the incident light, R_* expresses the radius of the star and $R_{\rm p}$ denotes the radius of the planet. If limb darkening is not included, then the intensity is uniform across the stellar disk and can be divided out along with π. Note that the input from the planet is limited to the in-transit flux.

Now, let us assume that a chemical species experiences absorption at a specific wavelength within the planetary atmosphere. This effect will slightly increase the planetary radius at this wavelength in comparison to other wavelengths. The increase will be proportional to $A_H H$, where A_H expresses the number of atmospheric pressure scale heights that get intersected due to the enhanced opacity observed at this wavelength. A_H is needed as the actual opacity observed at a specific wavelength normally deviates from exactly one atmospheric pressure scale height. This deviation depends on the properties of the examined atmosphere. Incorporating $A_H H$ in Equation 1.14, results in:

$$\Delta D = \frac{(R_{\rm p} + A_H H)^2}{R_*^2} = \left(\frac{R_{\rm p}}{R_*}\right)^2 + \frac{2R_{\rm p}A_H H}{R_*^2} + \left(\frac{A_H H}{R_*}\right)^2. \tag{1.15}$$

The relation can be simplified as the atmospheric pressure scale height should be much smaller than the radius of the planet ($H \ll R_p$). We can therefore reasonably assume that $H^2 \approx 0$ and remove the last term from the expression. This leaves two terms:

$$\delta = \left(\frac{R_{\rm p}}{R_*}\right)^2, \tag{1.16}$$

and:

$$\delta_\lambda = \frac{2R_{\rm p}A_H H}{R_*^2}. \tag{1.17}$$

The first term is the familiar fractional expression for the transit depth and the second is the additional wavelength-dependent contribution from the planetary atmosphere.

Early theoretical predictions of clear atmospheres have suggested that the observed variation due to atmospheric absorption can be as high as ten atmospheric pressure scale heights (Brown, 2001). However, recent observations have demonstrated that the effect is less significant, averaging at a little below two atmospheric pressure scale heights (see e.g. Stevenson, 2016; Iyer et al., 2016). There are two key reasons contributing to these lower estimates and discussion as to which one is more prevalent is still on-going. The first one is the influence from clouds, which were found to play a more obscuring role than initially suspected (Barstow et al., 2017), and the second one is that the abundances of chemical species with strong absorption signals may be lower than expected (Madhusudhan et al., 2014a; Pinhas et al., 2019).

By applying the condition that $A_H = 1$, the effect is constrained to one atmospheric pressure scale height and gives the expression for the expected transit signal due to atmospheric absorption and/or scattering effects:

$$\delta_\lambda = \frac{2R_\mathrm{p}H}{R_*^2} = \frac{2R_\mathrm{p}k_\mathrm{B}T_\mathrm{eq}}{\mu g_\mathrm{p} R_*^2}. \tag{1.18}$$

This is a very useful relation as it indicates the potential strength of the atmospheric features in the transmission spectrum and can therefore inform us about whether a target is suitable for atmospheric characterisation.

One interesting consequence from Equation 1.18 is that the transmission spectrum will appear inverted, with stronger absorption lines leading to an increase in the transit depth, which are viewed as peaks in the produced transmission spectrum. Furthermore, the signal is predicted to be stronger for large exoplanets, with hydrogen-dominated atmospheres (small μ), high equilibrium temperatures and low surface gravities, orbiting small stars.

1.5 Deciphering transmission spectra

The transmission spectrum of an exoplanet probes a characteristic region of its atmosphere that takes the form of an annulus in our line of sight (see Figure 1.8). This

means that the information we receive is about the atmospheric conditions at the day-night transition zone. For hot Jupiters, this zone is especially interesting, as they are expected to be tidally locked (Guillot et al., 1996; Rasio et al., 1996). Therefore, one side always faces the host star and receives constant daylight, while the other faces the opposite direction and has never-ending night. Consequently, in the day-night terminator region conditions may change drastically leading to the formation of condensates and a variation in cloud coverage. This assumption also depends on atmospheric dynamics. Highly irradiated exoplanets are often found with temperature variations of several hundred K between day and night (Kataria et al., 2016).

The shape of the transmission spectrum can therefore tell us about the temperature in these regions, the potential presence of high-absorbing chemical species, clouds and/or hazes, the atmospheric phenomena that may occur at these temperatures and the candidate species that can form condensates. However, in certain circumstances, the transmission spectrum not only reveals the properties of the planetary atmosphere, but also properties of the host star. Variations in the stellar flux due to phenomena on the stellar surface may also affect the shape of the transmission spectrum. Such variations can be indicative of stellar activity and can lead to features that resemble similarities to the ones produced by the planetary atmosphere (McCullough et al., 2014). In the following paragraphs, I will present the basic characteristics of measured transmission spectra in more detail.

1.5.1 Atomic and molecular species

When an atmosphere is relatively clear, one can see traces of absorption in the obtained transmission spectrum by observing the peaks in the transit depth at certain wavelengths. These features can be a sign of specific atoms and molecules that constitute the planetary atmosphere. Example spectra from real observations of atomic and molecular absorption features in the optical are shown in Figure 1.9. The possible detection of such species in the observed spectrum was postulated around the same time as the detection of the first transiting exoplanet (Charbonneau et al., 2000; Henry

et al., 2000). Theoretical works by Seager & Sasselov (2000), Brown (2001) and Hubbard et al. (2001), established the atmospheric constituents that would be detectable via this method and revealed that the optical would be dominated by alkali metals and the infrared by water and other species of molecular nature. Furthermore, later works showed that, in the highly irradiated exoplanets, strong molecular absorbers could also be present in the optical (see e.g. Hubeny et al., 2003; Fortney et al., 2008, 2010).

To get a better picture of the elements that can be observed in a planetary atmosphere with this technique, it may be useful to look into the results from observations made in the last two decades. For example, the first element ever detected in an exoplanetary atmosphere is sodium (Na). Charbonneau et al. (2002) were able to isolate its characteristic signature in the atmosphere of HD 209458b from observations of the absorption line doublet at $\sim$589 nm. Interestingly, the strength of the observed line core was lower than expected by a factor of a few, which led to the conclusion that either the concentration of sodium is naturally low in this exoplanet or the presence of clouds at high altitudes has weakened the signal.

Since then, sodium has been detected in numerous other exoplanets (see e.g. Zhou & Bayliss, 2012; Chen et al., 2020; Alam et al., 2021), and so has potassium (K), the other alkali metal observed in the optical ($\sim$770 nm) (see e.g. Sing et al., 2011; Wilson et al., 2015; Sedaghati et al., 2016). Tentative or strong detections of these two alkali metals have now been claimed in about 25% of all the exoplanetary atmospheres observed in transmission[5]. Early discoveries mostly utilised space-borne telescopes such as Hubble (HST/STIS) and Spitzer (see e.g. Sing et al., 2008, 2015; Nikolov et al., 2014), but ground-based detections have been rapidly increasing in recent years (see e.g. Nikolov et al., 2016; Carter et al., 2020; Alderson et al., 2020; Ahrer et al., 2022; Chen et al., 2022).

While Na and K are quite important tracers of atmospheric chemistry and composition, other species can also be observed in the optical. Mineral molecules, such as titanium oxide (TiO) and vanadium oxide (VO), have also been predicted to produce

[5]Based on data from the ExoAtmospheres database.

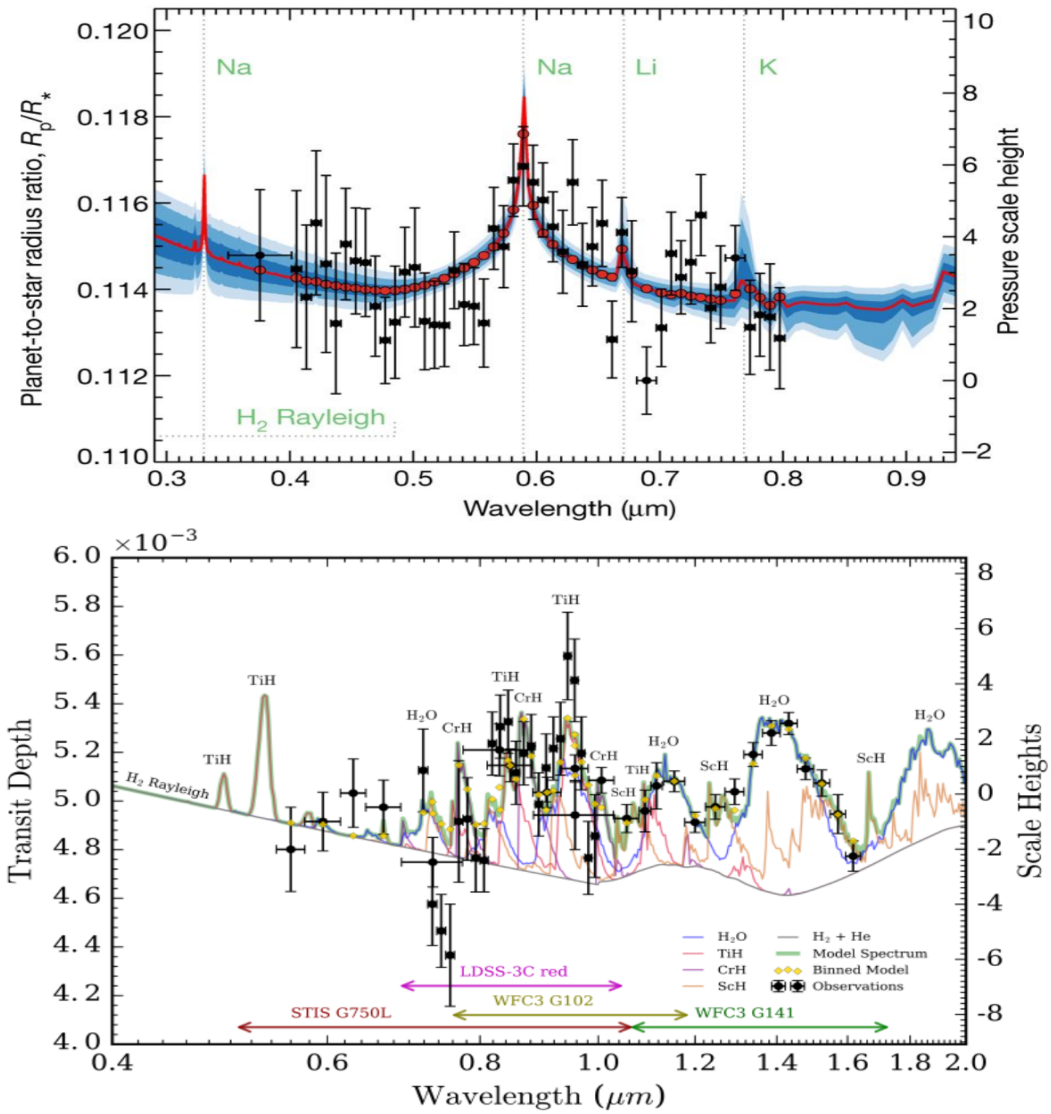

Figure 1.9: Example transmission spectra from observations of moderately hot exoplanets. Top row: Pressure-broadened sodium line along with signs of potassium and lithium in the atmosphere of WASP-96b (Nikolov et al., 2018b). Bottom row: Detection of water, titanium hydride, and possibly other metal hydrides, such as chromium hydride and scandium hydride, in the atmosphere of HAT-P-26b (MacDonald & Madhusudhan, 2019).

strong features in the transmission spectra of some hot and ultra-hot Jupiters (Hubeny et al., 2003; Fortney et al., 2008). Such exoplanets are highly irradiated and as a result metal oxides may be present in gaseous form. There have been several claimed detections of TiO and VO in the last few years (e.g. Sedaghati et al., 2017b; Mancini et al., 2019), but the most promising and reliable result was reported by Chen et al. (2021b), who detected TiO in the atmosphere of the hot Jupiter HAT-P-65b. Furthermore, another noteworthy metal oxide that has been detected in the optical is aluminum oxide (AlO) (von Essen et al., 2019). TiO and VO are potentially strong optical absorbers and can have a huge impact on the structure of a planetary atmosphere (Fortney et al., 2008).

Apart from metal oxides, metal hydrides may also leave traces in the transmission spectra of hot Jupiters and can also affect the planetary temperature profiles as a function of altitude (see e.g. Gandhi & Madhusudhan, 2019). The search for these metals has intensified in recent years with the inclusion of the species in atmospheric retrieval codes. Tentative evidence now exists for the presence of iron hydride (FeH) in the atmosphere of some hot exoplanets (e.g. Evans et al., 2016; Sotzen et al., 2020; von Essen et al., 2020; Skaf et al., 2020). In addition, MacDonald & Madhusudhan (2019) found weak evidence for the presence of scandium hydride (ScH) and chromium hydride (CrH), as well as moderate evidence for the presence of titanium hydride (TiH) in the atmosphere of the hot Neptune HAT-P-26b. Most of these molecules have absorption signatures in the near-infrared, but TiH can also show strong absorption in the optical. Other species predicted, such as sodium hydride (NaH) and magnesium monohydride (MgH), can also have prominent features in the optical (Gandhi & Madhusudhan, 2019).

Taking a closer look at observations in the near-infrared, most of them have so far been made with the Hubble Space Telescope (HST), and specifically the Wide Field Camera (WFC3) (see e.g. Edwards et al., 2022). This camera is focused on the 1.4 μm water (H_2O) band and has been critical in many detections of water in exoplanetary atmospheres (see e.g. Wakeford et al., 2013, 2018; Kreidberg et al., 2014b; Carone et al.,

2021; Wong et al., 2022). According to the ExoAtmospheres database[6], water has now been detected in over 50 exoplanetary atmospheres. The success of this instrument (and STIS in the optical) also led to the production of the first homogeneous population studies (see e.g. Sing et al., 2016; Tsiaras et al., 2018). The successor of HST, the James Webb Space Telescope (JWST), is expected to greatly raise the number of exoplanetary atmospheres observed in the infrared as the observation of such atmospheres is one of its primary missions (Gardner et al., 2006). It has already achieved an impressive milestone with the first-ever detection of CO_2 in an exoplanetary atmosphere using transmission spectroscopy (JWST Transiting Exoplanet Community Early Release Science Team et al., 2023). One interesting outcome so far from atmospheric observations in the infrared is the lack of methane (CH_4) and ammonia (NH_3)[7], molecules that are abundant in some Solar System planets (e.g. Jupiter, Taylor et al., 2004). This is normal as these molecules are expected to be dominant in cooler atmospheres (Burrows et al., 1997).

Returning back to atoms, some spectral lines have been associated with extended exospheres and atmospheric escape due to strong irradiation from the host star. The loss of atmospheric constituents can result in deeper transits at wavelengths where the characteristic absorption lines of hydrogen (H) and helium (He) are expected. For example, Lyman-α absorption due to escaping hydrogen is now routinely observed in the near-ultraviolet region (e.g. Lecavelier Des Etangs et al., 2010; Ehrenreich et al., 2015). But hydrogen can also produce signals in the optical through the Hα Balmer line (see e.g. moderately high resolution results by Jensen et al., 2018). Helium, on the other hand, has been a little more elusive, despite the fact it was predicted quite early (Seager & Sasselov, 2000). Nevertheless, this element was finally confirmed by Spake et al. (2018) in the eroding atmosphere of a Saturn-mass exoplanet.

Atomic absorption in low-resolution spectroscopy is relatively easy to detect as atoms produce clear, single-peaked features. In contrast, molecular absorption is usu-

[6] `http://research.iac.es/proyecto/exoatmospheres/table.php`

[7] This conclusion is drawn from data obtained via transmission spectroscopy and does not take into account other techniques.

ally more complex due to a multitude of spectral lines being combined together. Despite this observation, high resolution spectroscopy actually shows that many atomic lines come in doublets (e.g. Na, K) or triplets (e.g. He) due to the degenerate nature of energy transitions. Furthermore, pressure broadening can also interfere with the width of the atomic lines causing wing-like features. This effect is caused by collisions with other chemical species (e.g. H_2) and is a function of pressure (Seager & Sasselov, 2000; Seager & Deming, 2010).

1.5.2 Clouds and hazes

Planetary atmospheres tend not to be cloud-free. In fact, clouds are quite common in the gaseous envelopes of all types of exoplanets, with most observations to date revealing a significant level of cloudiness (e.g. Sing et al., 2016; Tsiaras et al., 2018; Edwards et al., 2022). Clouds are created when atmospheric constituents form condensates due to phase transitions from their gaseous state to their liquid or solid states (Marley & Robinson, 2015). Such transitions normally occur when atmospheric material encounters cooler temperatures as it moves freely within the atmosphere of a planet. As hot Jupiters are tidally locked (Guillot et al., 1996), they are expected to have one hemisphere that constantly faces daylight and another that constantly faces darkness. Transmission spectroscopy probes the day-night terminator region, which should be affected by temperature gradients and should therefore be an excellent location for the study of cloud formation processes.

Clouds are mostly a grey opacity source for transmission spectra. This means that absorption from clouds reduces the spectral line strength of atoms and molecules in the entire spectrum and is largely a uniform and thus a wavelength-invariant characteristic (Deming et al., 2013). Even so, a wavelength-dependent component may still exist. For example, the presence of high-altitude clouds can also create slopes in the optical and condensate features in the infrared (Pinhas & Madhusudhan, 2017). The slopes in the optical are often described as haze, but haze can also mean condensates that form through photochemical processes or disequilibrium chemistry (Marley et al., 2013).

In some cases, a thick cloud deck can completely mute spectral lines and result in featureless transmission spectra. Flat transmission spectra are frequently found in atmospheric studies of exoplanets (e.g. Gibson et al., 2013a,b, 2017; Mallonn et al., 2015; Kirk et al., 2016; Espinoza et al., 2019; Wilson et al., 2020) and are indicative of the strong influence clouds can have on a transmission spectrum.

Clouds are found across the whole temperature regime, from cool exoplanets (Kreidberg et al., 2014a) to ultra-hot ones (Espinoza et al., 2019), with different condensate species forming at different temperatures. At the lowest temperatures ($<$ 200 K), it is expected that clouds of ammonia and methane will be prevalent, while water should dominate at slightly higher temperatures (Marley & Robinson, 2015). At moderate temperatures (700-1300 K), sulphide species may become the dominant source of opacity at high altitudes where the pressure is lower (Pinhas & Madhusudhan, 2017), but recent theoretical work has questioned this (Gao et al., 2020). The same study also suggests that iron-based compounds may also face difficulties in forming clouds at higher temperatures. At temperatures above 1300 K, silicate species, such as forsterite and enstatite, may start to form condensates (Pinhas & Madhusudhan, 2017). In the hottest atmospheres, refractory species, such as corundum and titanium bearing molecules, are predicted to govern cloud formation (Wakeford et al., 2017).

The prevalence of clouds in many atmospheres has led to several studies introducing cloud metrics in order to determine possible correlations with other parameters such as surface gravity and equilibrium temperature. For example, Sing et al. (2016) and Stevenson (2016) propose that the amplitude of the water feature with respect to the continuum signal in the infrared can be used as an indicator of cloud coverage. Heng (2016) recommend another metric in the optical based on the strength of the Na and K absorption lines, while Tsiaras et al. (2018) suggest a metric based on a correction of a nominal model that includes all observed opacities by a flat (grey) opacity model. Population studies such as these are useful in the establishment of possible trends between the presence of clouds and exoplanet characteristics. Some early studies indicated a trend where increased irradiation is connected to fewer clouds (e.g. Stevenson, 2016; Fu et al., 2017), but other studies failed to establish a definitive

correlation (e.g. Sing et al., 2016; Heng, 2016). A larger, more recent study based on 70 gaseous exoplanets and space-based observations in the infrared only confirmed the weak correlation with temperature found in previous studies and so results on this matter remain somewhat inconclusive (Edwards et al., 2022). Stevenson (2016) also found a relation with surface gravity, but this has been disputed by a larger population study that only found an influence from the radius and not from the mass of an exoplanet (Tsiaras et al., 2018). Interestingly, the work of Edwards et al. (2022) also challenged the idea of the mass-metallicity relation mentioned in Section 1.3.3 and struggled to determine any trends despite analysing the largest sample of exoplanetary atmospheres so far.

All these uncertainties demonstrate the importance of additional observations in both the optical and the infrared range of the spectrum. The broad range is necessary to break any potential degeneracies due to cloud opacity and to place better constraints on the observed spectral features. Population studies based exclusively on observations in the infrared highlight the need for a wider spectral range coverage (e.g. Edwards et al., 2022).

Apart from clouds, hazes can also be present in exoplanetary atmospheres. These tend to have a similar effect on the observed transmission spectra as they also increase atmospheric opacity and weaken the signatures of atoms and molecules. Contrary to most clouds, however, hazes have a clearly defined wavelength-dependent component, which manifests itself in the optical range of the transmission spectrum. More specifically, hazes form a spectral slope that is seen to increase from redder to bluer wavelengths (e.g. Pont et al., 2008). Figure 1.10 illustrates the effect clouds and hazes can have on the observed transmission spectra. Several hot Jupiters to date have been found with very steep blueward slopes in their transmission spectra due to enhanced haze within their planetary atmosphere (e.g. May et al., 2020; Alam et al., 2020; Chen et al., 2021a; Ahrer et al., 2022).

There are a few different mechanisms that can lead to the haze slope seen in the optical. One commonly observed phenomenon is Rayleigh scattering. In this case, starlight is scattered by grains of dust that may exist within the planetary atmosphere.

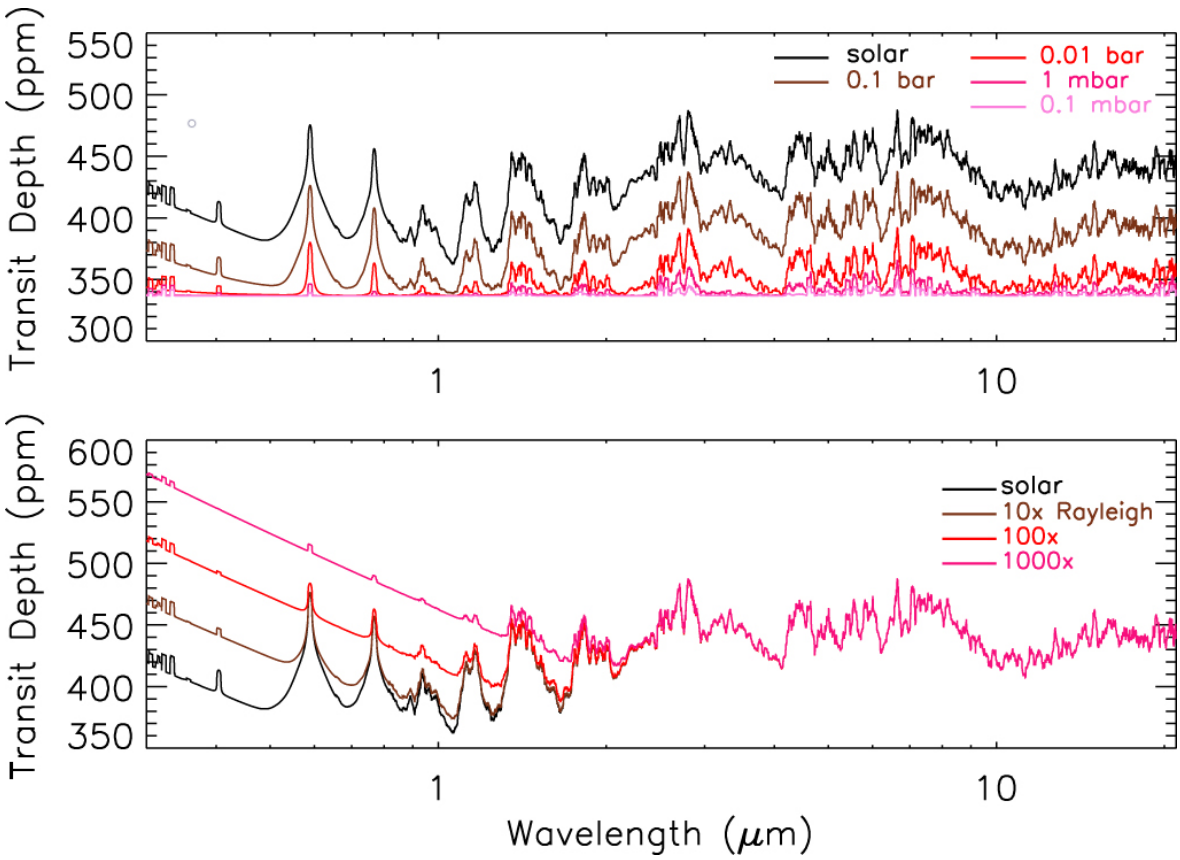

Figure 1.10: Model transmission spectra depicting the effect of clouds and hazes. Top row: The effect of a thick cloud deck at different atmospheric pressure layers. Bottom row: The effect of Rayleigh scattering at different haze enhancement factors. The model spectra are for an exoplanet with solar composition, $T_{\rm iso} = 1000\,{\rm K}$, $g_{\rm p} = 10\,{\rm m\,s^{-2}}$ and $R_{\rm p} = 2\,{\rm R}_{\oplus}$, orbiting a star with $R_{*} = 1\,{\rm R}_{\odot}$. In both rows, the black line represents a clear atmosphere. Figure from Kempton et al. (2017).

A common species that can block starlight in this manner is molecular hydrogen (H_2) (Lecavelier Des Etangs et al., 2008b). For Rayleigh scattering, the cross-section follows a power law that is directly proportional to λ^{-4}. Larger condensate particles that form clouds (e.g. $MgSiO_3$ Lecavelier Des Etangs et al., 2008a) can also create a haze slope. This is known as Mie scattering and theoretical studies have indicated that some condensate species may even induce much steeper scattering slopes than Rayleigh scattering (Pinhas & Madhusudhan, 2017). Another potential mechanism for the creation of slopes in optical transmission spectra is haze formation due to photochemical reactions (Ohno & Kawashima, 2020). If the atmosphere is vigorously mixed, then such reactions can lead to steeper-than-Rayleigh slopes for equilibrium temperatures below 1500 K.

1.5.3 Stellar activity

A transmission spectrum does not only show properties of the planetary atmosphere, but can also reveal characteristics of the observed stellar disk. This can happen if the star is magnetically active, a condition that favours the formation of dark and bright spots on the stellar surface. Magnetic activity is thought to be prevalent in cool stars such as M-dwarfs (Kochukhov, 2021). If stellar activity is strong, then it can mimic absorption and/or scattering features in the obtained transmission spectrum that are otherwise produced by opacity sources within the planetary atmosphere.

One such feature is an increasing scattering slope towards bluer wavelengths. Such a slope can be influenced by haze phenomena, such as Rayleigh scattering (Pont et al., 2008), but can also be a result of stellar contamination caused by unocculted starspots (McCullough et al., 2014). Unocculted starspots are dark spots on the stellar surface that appear away from the path of the transiting exoplanet during the transit event. The spots appear dark because of their lower temperature with respect to the temperature of the photosphere. The same slope can be created by the presence of bright plage areas during the time of transit (Oshagh et al., 2014). If these areas are occulted (i.e. blocked by the passage of the transiting exoplanet), then the transit depth

can also increase towards shorter optical and near-ultraviolet wavelengths.

But stellar activity can also lead to inverted slopes where transit depth decreases towards bluer wavelengths. This is often a sign of unocculted faculae (e.g. Nikolov et al., 2021; Rathcke et al., 2021) and needs to be addressed carefully when analysing transmission spectra with such shapes. Faculae are the opposite of dark spots as they appear bright and have a higher temperature compared to the temperature of the photosphere. A similar slope can also be observed with occulted starspots (Mallonn et al., 2018). It is vital to distinguish between the planetary and stellar nature of the observed signal as this kind of characteristic can also be interpreted by a clear atmosphere in disequilibrium chemistry (Sotzen et al., 2020).

Current forward modelling and atmospheric retrieval codes[8] can model the potential effects of stellar contamination (e.g. Pinhas et al., 2018; Zhang et al., 2019), and I very briefly describe the relevant findings on two of the observed targets in Chapters 4 and 5. However, such analyses usually only present alternative interpretations of obtained transmission spectra and usually other observational means are used to figure out the activity of the host star. For example, spectroscopic measurements of the stellar atmosphere in the ultraviolet can be used to observe the calcium H and K lines where signs of chromospheric emission could be interpreted as stellar activity (Vaughan et al., 1978; Noyes et al., 1984). Another option would be the photometric surveillance of the host star for signs of variability and/or periodicity in the long-term light curves (e.g. García et al., 2010). In Chapters 4 and 5, I take such observations into account in my attempt to interpret the obtained results.

1.6 Ground-based versus space-based observations

This work will focus on ground-based observations, but some space-based results will also be included to aid with the conclusions. For this reason, a comparison between the

[8]More information on these methods will be given in Chapter 2.

two approaches will be useful in order to understand the usefulness of ground-based observations, as well as some of their limitations.

One argument in favour of ground-based observations is the large aperture of the telescopes. This characteristic improves light-gathering power and is important in detecting minute variations in the stellar flux. While the feature is not vital for the detection of transiting exoplanets as smaller telescopes are normally used for that purpose (see e.g. Pollacco et al., 2006; Wheatley et al., 2018), the parameter is quite significant for the detection of a planetary atmosphere. Ground-based spectroscopic surveys of planetary atmospheres often utilise 8-10 m-class telescopes (see e.g. Murgas et al., 2014; Huitson et al., 2017), although 3-4m-class telescopes have also revealed promising results (e.g. Ahrer et al., 2022). Until recently, the only option in space was the 2.4 m Hubble Space Telescope (HST). The newly commissioned James Webb Space Telescope (JWST, Gardner et al., 2006) is a crucial step forward as it has an aperture of 6.5 m, state-of-the-art instrumentation, and a more stable space environment.

Another advantage of ground-based instruments is hardware accessibility. If a problem occurs or something stops working, the telescope and the instruments can be easily accessed and repaired. If fixing them is impossible, then suitable replacements can be arranged. Furthermore, ground-based instruments are often refurbished or repurposed. For example, one of the instruments used in my analyses had been upgraded two years before the observations to correct for systematic effects caused by one of its components (Boffin et al., 2016)[9]. In space, the repairing of flaws is often unfeasible and innovative workarounds are usually the only viable solution (see e.g. Howell et al., 2014).

Not long ago, ground-based telescopes were the only option to obtain continuous transit light curves for transmission spectroscopy as the main space-borne instrument (HST) is in low orbit around the Earth and so space-based observations were fragmentary. However, this has changed with the launch of the JWST, which can spectroscopically observe entire transits without pauses due to its location at the L2 Lagrange

[9]Although, as we will see in the following chapters, substantial systematic effects may still occur

point (Gardner et al., 2006).

The cost for the construction of ground-based telescopes suitable for transmission spectroscopy also tends to be lower compared to space-based counterparts. For instance, one of the largest ground-based telescopes currently able to observe exoplanetary atmospheres, the Gran Telescopio Canarias, had a total development budget of ∼€130,000,000[10]. For comparison, the predecessor of JWST, HST, had a total prelaunch budget of ∼$2,000,000,000[11]. This huge difference in budget explains why there is an increasing interest towards ground-based transmission spectroscopy. At this time, there are about a dozen telescopes capable of doing transmission spectroscopy, but only two of them are space-based.

On the other hand, the Earth's atmosphere is a major limiting factor in ground-based observations. The main reason is the unstable weather conditions. For example, observations can be completely lost or severely limited by cloudy conditions (see also Section 5.2). Furthermore, adverse weather, including windy conditions, can affect the pointing of the telescope and the guidance systems, and this can potentially result in loss of observing time due to displacements or complete disappearance of the target from the field-of-view.

In case of good weather, there is still the contribution from atmospheric extinction. The effect is predominantly caused by absorption and scattering effects due to interaction of starlight with atmospheric atoms and molecules (see also Section 5.3.2.4). The observed reduction in stellar flux is a function of airmass and wavelength. The standard approach to overcome the impact of this phenomenon is to perform differential photometry/spectroscopy by using a suitable nearby comparison star to remove most of the interference. The approach was recently challenged by Panwar et al. (2022), who presented evidence that in some cases extinction may induce completely different light curve profiles between the target and nearby stars.

Extinction can also be influenced by scintillation due to atmospheric turbulence.

[10] https://en.wikipedia.org/wiki/Gran_Telescopio_Canarias

[11] https://www.nasa.gov/sites/default/files/atoms/files/not_yet_imagined_tagged.pdf

The effect is usually more prominent in ground-based photometric surveys of transiting exoplanets orbiting bright stars (Föhring et al., 2019). However, the impact of this source of noise is largely minimised in transmission spectroscopy where exposure times are increased and telescopes with larger apertures are used. This is expected as scintillation is inversely proportional to aperture size and integration time (Dravins et al., 1998, see also Section 3.3).

1.7 Thesis overview

The motivation behind the conduction of this research was to characterise the atmospheres of transiting hot Jupiters that show signs of radius inflation. These targets were deemed to be ideal for observations of their atmosphere via ground-based transmission spectroscopy because of their size, their proximity to their host star, their temperature and their expected atmospheric composition. Theoretical simulations of cloud-free atmospheres carried out by Fortney et al. (2010) predicted the presence of Na and K, as well as the presence of metal oxides, such as TiO and VO, in the optical wavelengths of these exoplanets. Therefore, my research is focused on observations performed in the visible region for hot Jupiters with suitable characteristics that promise an extended atmosphere. To that respect, I analyse data of the inflated hot Jupiters WASP-74b and WASP-88b and present my findings, including the constructed transmission spectra of these exoplanets, in dedicated chapters of this thesis. In addition, I also present a brief analysis of the hot Jupiter WASP-75b. This target was used as a test case for data reduction and analysis processes during the beginning of my research.

Apart from atmospheric characterisation, my research is also heavily concentrated on the different analysis methods applied to obtain the measured transmission spectra. The main reason for the exploration of various analysis techniques is the fact that the ground-based data I used are corrupted by substantial systematic effects and so a lot of effort is needed to disentangle the transit from the background noise. My investigation considers both older and newer techniques in the decorrelation of the noise parameters

and even presents a novel method for the wavelength-specific analysis of WASP-74b.

In Chapter 2, I describe the basic procedures followed during my analysis, including the application of basic decorrelation methods on the transit light curves and the application of synthetic spectra from simulated atmospheres on the observed transmission spectra. I also mention the tools I used towards the final product. Special focus is given on the physics involved during the creation of the atmospheric models. In Chapter 3, I perform an elementary data reduction procedure on a data set obtained for WASP-75b and produce a transmission spectrum with the aid of forward atmospheric models. In Chapters 4 and 5, I present the main body of my research with an exhaustive atmospheric investigation of the hot Jupiters WASP-88b and WASP-74b, respectively. These chapters describe in detail the steps taken during the analysis of the transit light curves and discuss the properties of the obtained transmission spectra. The thesis ends with Chapter 6. In this chapter, I present my conclusions on the methodologies employed for the two bloated hot Jupiters WASP-74b and WASP-88b, put the results into context, give an interpretation of the observed spectral properties and stress the importance of future observations in the understanding of their atmospheres.

2 Methods and tools for transmission spectroscopy

This chapter will concentrate on the methodologies followed in each step of the data investigations and will outline the tools employed throughout the entire procedure. There are three distinct stages: the data reduction process, the transit light curve analysis, and the atmospheric modelling of the obtained transmission spectrum. Starting with Section 2.1, I briefly outline the basic steps taken during the initial reductions. This section is purposely short as more information will be revealed in subsequent chapters. Next, in Section 2.2, I present the methods and tools utilised to model the transit light curves, with special attention given to the decorrelation of the systematic contributions. Finally, in Section 2.3, I describe how simulated atmospheres are constructed and discuss the important parameters considered in the atmospheric models.

2.1 The data reduction process: A short list of the basic steps

The following steps were (roughly) followed to create the transit light curves for the targets investigated in this thesis:

- Conversion of the raw FITS[1] images to easily manipulated data structures.
- Conversion of given time to a time standard based on the barycentre of the solar system.
- Isolation and determination of the spectral aperture of the brightest stars in the field (including the target).
- Removal of the background noise from the spectral signal.

[1]Flexible Image Transport System (Wells et al., 1981)

- Estimation of the error in the measurements.
- Wavelength calibration through the use of a gas lamp with known emission lines.
- Estimation of the total spectral counts for each bright object and separation into wavelength bins.
- Normalisation of the observed fluxes.
- (Optional) Division of the target fluxes by a combination of nearby star fluxes to obtain relative fluxes.
- Application of the above procedure to all images to create the time-series and, hence, the transit light curves.

All these steps are explained in more detail in Chapter 3, where I perform a very basic data reduction procedure for WASP-75b, and in less detail in Chapters 4 and 5.

2.2 Light curve modelling

The modelling of the transit light curves normally consists of two parts. One is the contribution from systematics, which can be instrumental, atmospheric or astrophysical, and the other is the contribution from the actual planetary transit. Normally these two components are evaluated simultaneously during the fitting of the light curves, although in Chapter 3 this is done separately, with noise being modelled out first and then the "clean" transit being modelled afterwards. There are two widely used techniques to decorrelate the nuisance noise component from the useful bit of the light curves, which is the transit itself. One method is the use of baseline parametric functions and the other is the use of Gaussian processes. I make use of both these techniques in my research either separately or in combination.

2.2.1 Parametric baseline functions

The use of parametric functions to decorrelate the systematic effects from the transit light curves is a frequent practice in transmission spectroscopy (see e.g. Gibson et al., 2011; Lendl et al., 2016; Mallonn & Strassmeier, 2016; Nikolov et al., 2016; Lendl et al., 2017). Their use is necessary as the transits are often overshadowed by significant noise contributions in the observed data sets (see e.g. the data sets in Chapters 4 and 5). These functions are either based on time or on auxiliary measurements that are obtained during an observation and may have an impact on the flux measurements. Noise may, therefore, be time-correlated or correlated with physical parameters. Another term often used for these ancillary measurements is "optical state parameters" as they tend to define the state of the optics and the detector. A selection of various optical state parameters are normally considered in the baseline functions and I describe them in the relevant chapters.

The most common parametric functions employed in this work are polynomials. The polynomials of different auxiliary parameters are often combined together in an attempt to create a baseline solution that best matches the observed systematic trends (see details in Section 4.3.2.1). The combinations also explore variations in the polynomial degrees, which are usually of low order. Due to the large number of possible combinations, there needs to be a way to select the best-fit noise model. This is done through the use of two different information criteria.

The first information criterion is the Akaike Information Criterion (AIC, Akaike, 1974). This criterion is expressed as:

$$AIC = \chi^2 + 2k, \tag{2.1}$$

where k represents the number of variables in the fit.

The second criterion is the Bayesian Information criterion (BIC, Schwarz, 1978). This estimator is defined as:

$$BIC = \chi^2 + k\ln(n), \tag{2.2}$$

where n represents the size of the data set.

Both criteria make use of the χ^2 statistic:

$$\chi^2 = \sum_{i=1}^{k} \left(\frac{\mathcal{D}_i - \mathcal{M}_i}{\sigma_i} \right)^2, \tag{2.3}$$

where $\mathcal{D}$ denotes the data, $\mathcal{M}$ symbols the model, and σ is the error of the measurements.

Both these criteria select the best model while trying to eliminate models that under-fit or over-fit the data. As can be seen from Equations 2.1 and 2.2, both criteria penalise models with a higher number of free parameters by inducing an increase in the criterion values. For data sets with hundreds of data points, such as the ones that portray transit light curves, the AIC reveals a preference for more complex models compared to the BIC (see also Section 4.3.2.1).

Apart from polynomials, the BIC parameter is also used in Chapter 4 for the selection of the simulated atmosphere that best describes the observed transmission spectrum.

2.2.2 Gaussian processes

As Gibson et al. (2011) pointed out and as we will see in Chapters 4 and 5, baseline parametric functions are often sensitive to the systematics considered. Furthermore, their flexibility is quite limited and they have difficulties following the behaviour of the data. Such issues may affect the accurate determination of the nuisance parameters and, by extension, the accurate estimation of the transit parameters. To resolve this problem, it may be helpful to consider the transit light curves as *Gaussian processes* (GPs). Under this assumption, the data are described by a collection of stochastic variables, such that every finite subgroup of them is defined by a joint multivariate Gaussian distribution (Rasmussen & Williams, 2006). GPs have received a great deal of attention in recent years and are now increasingly being applied to time-series observations (see e.g. reviews by Roberts et al., 2012; Aigrain & Foreman-Mackey, 2022), including transmission spectroscopy (e.g. Gibson et al., 2012).

In its general form, a GP is defined as (Rasmussen & Williams, 2006):

$$f(x) \sim \mathcal{GP}(\mu(x,\theta), K(x,\phi)) \tag{2.4}$$

Here, x represents the input domain (usually time or auxiliary physical parameters), μ is the mean function that depends on the parameters θ, and K is the covariance matrix that depends on the parameters ϕ. Now, since we work with transit light curves and time-series observations, the mean function can be assumed to be the transit function $T(t,\theta)$, resulting in (Gibson et al., 2012):

$$f(t,x) \sim \mathcal{GP}(T(t,\theta), K(x,\phi)), \tag{2.5}$$

where t is the time domain of the observations and θ are the transit parameters.

Definition 2.5 implies that the joint multivariate Gaussian distribution p of the transit data is given as (Gibson et al., 2012):

$$p(f \mid t, X, \theta, \phi) = \mathcal{N}(T(t,\theta), K(X,\phi)), \tag{2.6}$$

where $\mathcal{N}$ certifies that the distribution is Gaussian, f describes the observations, and X denotes the covariance matrix of the input parameter x considered[2].

Hence, the likelihood function $\mathcal{L}$, described by the covariance matrix K, is defined in matrix form as (Aigrain & Foreman-Mackey, 2022):

$$\mathcal{L}(r \mid t, x, \theta, \phi) = \frac{1}{\sqrt{(2\pi)^N |K|}} exp\left(-\frac{1}{2} r^T K^{-1} r\right), \tag{2.7}$$

where r is a vector of the residuals, determined by:

$$r = f - T(t,\theta), \tag{2.8}$$

r^T is the transpose of this vector, N is the number of flux measurements, and $|K|$, K^{-1} are the determinant and inverse of matrix K.

[2]In Chapters 4 and 5 the noise parameter x is replaced by w_ν, where ν describes the number ($\nu = 1, 2... N$) of auxiliary noise parameters w considered.

In practice, it is computationally faster to consider the logarithmic version of this function (Gibson et al., 2012):

$$\ln \mathcal{L}(r \mid t, x, \theta, \phi) = -\frac{1}{2}\left(r^T K^{-1} r + \ln|K| + N\ln(2\pi)\right). \tag{2.9}$$

In its simplest form, the covariance matrix is constructed in diagonal form under the assumption of pure white noise. The covariance function or *kernel*, in this case, is given by (Aigrain & Foreman-Mackey, 2022):

$$K_{nm} = \delta_{nm}\sigma_n^2, \tag{2.10}$$

where δ_{nm} is the Kronecker delta function and σ_n represents the data point uncertainties. In most situations, however, the data points tend to be correlated with each other, inducing red noise, which can be described by an additional kernel with non-zero off-diagonal matrix elements. The kernel, therefore, transforms to (Aigrain & Foreman-Mackey, 2022):

$$K_{nm} = k(x_n, x_m, \phi) + \delta_{nm}\sigma_n^2, \tag{2.11}$$

where x_n, x_m are the auxiliary noise parameters corresponding to the n^{th} and m^{th} samples, and ϕ are the kernel parameters. The kernel parameters describe properties of the kernel, such as maximum covariance and level of similarity (or correlation) between inputs, and are often referred to as *hyperparameters*.

In transmission spectroscopy, the Matérn class kernel is often selected for the modelling of correlated noise in the presence of substantial external interference and the same is done for the purposes of this thesis. The general form of this kernel class is (Rasmussen & Williams, 2006; Roberts et al., 2012):

$$K_{nm} = k_{\mathrm{M}}(x_n, x_m) = \xi^2 \frac{2^{1-\nu}}{\Gamma(\nu)}\left(\sqrt{2\nu\frac{(x_n - x_m)^2}{\tau^2}}\right)^{\nu} \mathbb{B}_v\left(\sqrt{2\nu\frac{(x_n - x_m)^2}{\tau^2}}\right), \tag{2.12}$$

where ξ is a hyperparameter that defines the height scale or maximum covariance, τ is a hyperparameter that gives the length scale of the input variables, $\Gamma()$ is the Gamma function, $\mathbb{B}()$ is the modified Bessel function of the second kind, and ν is a parameter that regulates the differentiability of the output.

In the special case when ν is half integer ($\nu = p + \frac{1}{2}$), this covariance function can be simplified to a product between an exponential and a polynomial of p^{th} order (Abramowitz & Stegun, 1965; Rasmussen & Williams, 2006):

$$\begin{aligned} K_{nm} &= \xi^2 \exp\left(-\sqrt{(2p+1)\frac{(\Delta x)^2}{\tau^2}}\right) \frac{\Gamma(p+1)}{\Gamma(2p+1)} \sum_{i=0}^{p} \frac{(p+i)!}{i!(p-i)!} \left(2\sqrt{(2p+1)\frac{(\Delta x)^2}{\tau^2}}\right)^{p-i} \\ &= \xi^2 \exp\left(-\sqrt{(2p+1)\frac{(\Delta x)^2}{\tau^2}}\right) \frac{p!}{(2p)!} \sum_{i=0}^{p} \frac{(p+i)!}{i!(p-i)!} \left(2\sqrt{(2p+1)\frac{(\Delta x)^2}{\tau^2}}\right)^{p-i}, \end{aligned} \tag{2.13}$$

where $\Delta x = x_n - x_m$.

For $\nu = \frac{3}{2}$ ($p = 1$), one obtains the Matérn 3/2 kernel (Rasmussen & Williams, 2006; Aigrain & Foreman-Mackey, 2022):

$$K_{nm} = k_{M_{3/2}}(x_n, x_m) = \xi^2 \left(1 + \sqrt{3\frac{(x_n - x_m)^2}{\tau^2}}\right) \exp\left(-\sqrt{3\frac{(x_n - x_m)^2}{\tau^2}}\right), \tag{2.14}$$

which, with the addition of white noise, becomes:

$$K_{nm} = \xi^2 \left(1 + \sqrt{3\frac{(x_n - x_m)^2}{\tau^2}}\right) \exp\left(-\sqrt{3\frac{(x_n - x_m)^2}{\tau^2}}\right) + \delta_{nm}\sigma_n^2. \tag{2.15}$$

Equation 2.15 is the basis for the kernels constructed during the GP analyses in Chapters 4 and 5.

At $\nu \to \infty$ ($p \to \infty$), the degree of differentiability extends to infinity and Equation 2.13 is reduced to (Rasmussen & Williams, 2006; Roberts et al., 2012; Aigrain & Foreman-Mackey, 2022):

$$K_{nm} = k_{SE}(x_n, x_m) = \xi^2 \exp\left(-\frac{(x_n - x_m)^2}{\tau^2}\right). \tag{2.16}$$

This is the squared exponential kernel, another commonly used kernel in transmission spectroscopy, and is also discussed in Chapter 4.

Calculating the GP probabilities for transit light curves is not a speedy task. It requires inversion of 2D matrices and evaluation of their determinants in every step (see Equation 2.9). This is computationally demanding for data sets with hundreds of data

points and normally involves $\mathcal{O}(n^3)$ operations. Ambikasaran et al. (2015), however, developed an $\mathcal{O}(n \log^2 n)$ algorithm for inversion that significantly shortens computational time. The Python GP module `george` (Foreman-Mackey, 2015) employed for the GP regression performed in Chapters 4 and 5 is relatively fast as it incorporates this $\mathcal{O}(n \log^2 n)$ algorithm in the matrix calculations.

2.2.3 Transit modelling

Under the assumption that noise is modelled out using one of the two techniques described earlier, the transit model can then be constructed based on the temporal reduction in flux caused by the transit. Mandel & Agol (2002) provide analytic formulae to model the transit light curves for the cases where stellar limb darkening is either quadratic or follows the non-linear, four-parameter law. During the analysis of all exoplanets presented throughout this thesis, I make use of the `batman` software (BAsic Transit Model cAlculatioN in Python, Kreidberg, 2015). For the quadratic case, which is used most of the time, the package performs model calculations in accordance with Mandel & Agol (2002). In the four-parameter case, however, the analytic algorithms become computationally expensive and so the code resorts to numeric calculations that integrate stellar intensity in 1D radial space by using non-uniform step sizes. This method makes non-linear limb darkening calculations faster, but is still a factor of a few slower than the analytic calculations performed in the quadratic case and is, therefore, not considered in most scenarios.

2.2.4 Light curve fitting algorithms

To obtain the transit parameters and their uncertainties from the transit light curves, it is important to associate the aforementioned light curve models with a suitable fitting algorithm. Two approaches were considered to tackle this matter: a non-linear least-squares minimisation procedure and a Markov Chain Monte Carlo (MCMC) sampling. The first process was exclusively applied to light curves modelled with baseline

parametric functions, whereas the second process was applied to light curves modelled using any of the two systematics realisations. In addition, the two fitting algorithms were also explored in the context of simulated atmospheres (see Section 2.3), with non-linear least squares being applied to a grid of forward models and MCMC being briefly applied to atmospheric retrievals. The latter method, however, was not considered in the final retrieval results as nested sampling was preferred instead (see Section 2.3.2.1).

2.2.4.1 Non-linear least squares minimisation

In least squares minimisation, the goal is to reduce the sum of the (weighted) squared residuals between the model and the observed data. With linear least squares regression the solution can be quickly determined through one set of calculations (see e.g. ordinary least squares). This is normally not the case for non-linear least squares, where the solution is usually determined numerically through an iterative procedure. One commonly used algorithm, also employed throughout this work, is the Levenberg-Marquardt method (Levenberg, 1944; Marquardt, 1963). This algorithm minimises the sum of the squared residuals by providing initial guesses of the model parameters and then progressively updating the model parameters based on a combination of two other algorithms: the Gauss-Newton method and the gradient descent method. In the Gauss-Newton approach, the least squares function is assumed to be locally quadratic and so the new model parameters are obtained through estimation of the minimum value. In the gradient descent approach, the parameters are updated following the direction of the steepest descent (i.e. the direction closer to the minimum). The Levenberg-Marquardt algorithm takes the best from both algorithms by following the gradient descent method when the model values are far from the solution and the Gauss-Newton method when the model values are closer. Non-linear least squares, and the Levenberg-Marquardt method especially, are frequently used in transmission spectroscopy (see e.g. Line et al., 2013; Nikolov et al., 2016; Mancini et al., 2016; Mackebrandt et al., 2017). For the purposes of this thesis, I make use of `lmfit`, the python implementation of the Levenberg-Marquardt algorithm (Newville et al., 2016).

2.2.4.2 MCMC sampling

MCMC sampling is a Bayesian inference method used for the estimation of the model parameters and their uncertainties through calculation of the posterior probability distribution of a given data set. This distribution simply defines the probability of the model given the data and is constructed with the help of some prior information. The method relies on some initial estimates for the model parameters and then explores the multi-dimensional parameter space, defined by the prior limits, by sampling new test estimates in a step-by-step process. Each new estimate is a random proposal based on a given distribution centred on the previous estimate. A widely-used random proposal algorithm is the Metropolis-Hastings algorithm (Metropolis et al., 1953; Hastings, 1970), where this random "walk" of the model parameters is computed from the ratio of the posterior distributions between the new and the old proposals. If the value of this ratio is greater than a number randomly-drawn from a uniform distribution bounded by 0 and 1, then the new proposal is accepted, else it is rejected and the algorithm will continue from the previous estimate. Through this acceptance-rejection process, the model parameters are drawn towards higher-probability space and so get closer to the best fit. Once the chain of random steps (or random walk) has been completed, the joint posteriors of all model parameters can be obtained. One common MCMC approach is the one implemented by the `emcee` module (Foreman-Mackey et al., 2013). Under this approach, the parameter space is examined by an ensemble of walkers instead of one chain or a set of a few independent chains (see more details in Chapter 4). `emcee` is nowadays commonly used in transit light curve fitting (see e.g. Parviainen et al., 2015; Kirk et al., 2017; Mikal-Evans et al., 2021; Panwar et al., 2022) and is also employed for the targets examined in this thesis.

2.3 Simulated atmospheres

Interpretation of the observed transmission spectra is only possible with a good understanding of the processes that govern exoplanetary atmospheres. If the background

theoretical knowledge of the physics and chemistry involved is sufficient, then structural approximations of these atmospheres, called "simulated" or "model" atmospheres can be created. From these models one can build synthetic/theoretical spectra, which can then be compared to the real data. A plethora of custom-made state-of-the-art codes exist in the literature that can do this job and a fairly recent review by (Madhusudhan, 2018) outlines some of them[3]. Most of these atmospheric modelling codes are not publicly available and are mainly maintained by the developers. There is an increasing trend, however, for making atmospheric modelling software open-source and, thus, publicly accessible. Examples of freely available software include Exo-Transmit (Kempton et al., 2017), petitRADTRANS (Mollière et al., 2019), PLATON (Zhang et al., 2019), the ExoCTK generic grid (Goyal et al., 2019a), and TauREx 3 (Al-Refaie et al., 2021).

There are essentially two different approaches in mapping the observed spectrum to a simulated atmosphere: forward modelling and atmospheric retrievals. Forward models are the core of atmospheric modelling as they accept as input given planetary and stellar properties and calculate the structure and composition of a planetary atmosphere based on these assumptions to eventually give an output of the model spectrum. With a grid of forward models the objective is simple: to establish the basic properties of the observed atmosphere without the need to derive the exact details of the atmosphere. In this case, a distinction is often made between clear, cloudy and hazy scenarios and a solution is given with respect to the presence of certain atoms and molecules in the atmosphere (see e.g. Nikolov et al., 2016; Mancini et al., 2019). Atmospheric retrievals are a step further as they couple forward models with parameter estimation techniques to optimise the simulation parameters and get the best-fit between theoretical spectra and observed spectra. Atmospheric retrievals are useful for the determination of metal enrichment parameters and for the estimation of chemical abundances (see e.g. Nikolov et al., 2018b; Sheppard et al., 2021). In Chapter 4, I make

[3]Note that this book review chapter is focused on atmospheric retrievals and so naturally some forward modelling software is excluded.

use of the ExoCTK generic grid and PLATON to create forward models of WASP-88b's atmosphere, while in Chapter 5, I carry out an atmospheric retrieval analysis of WASP-74b's atmosphere using PLATON. Retrievals with AURA (Pinhas et al., 2018) are also included in both chapters to aid the deciphering of the obtained transmission spectra.

Lastly, it is worth mentioning that simulated atmospheres can be explored in multiple dimensions with each additional dimension adding to the computational complexity of the already quite complicated atmospheric models. The two most common realisations of exoplanetary atmospheres are the one-dimensional (1D) plane-parallel models (e.g. Pinhas et al., 2018) and the three-dimensional (3D) general circulation models (e.g. Nixon & Madhusudhan, 2022). Throughout this work atmospheric assumptions are made by making use of 1D vertical temperature profiles and so computations of the simulated atmospheres are relatively fast.

2.3.1 Forward modelling

The simplest forward models are constructed by complying with the rules of radiative-convective equilibrium and chemical equilibrium (e.g. Goyal et al., 2019a). PLATON can also, in principle, explore deviations from equilibrium chemistry through the incorporation of custom-made abundances in the code (Zhang et al., 2019), but this was consciously avoided during modelling, as PLATON is not built for abundance calculations and the retrievals with AURA are more suitable for this purpose. In the paragraphs below, I will concentrate on the basic properties of simulated atmospheres, such as pressure-temperature profiles, chemistry and condensation, and explain their role in deciphering the observed transmission spectra in Chapters 4 and 5. Figure 2.1 illustrates the various processes, temperature profiles and chemistry that can be inferred from an exoplanet's measured transmission spectrum with help from atmospheric models.

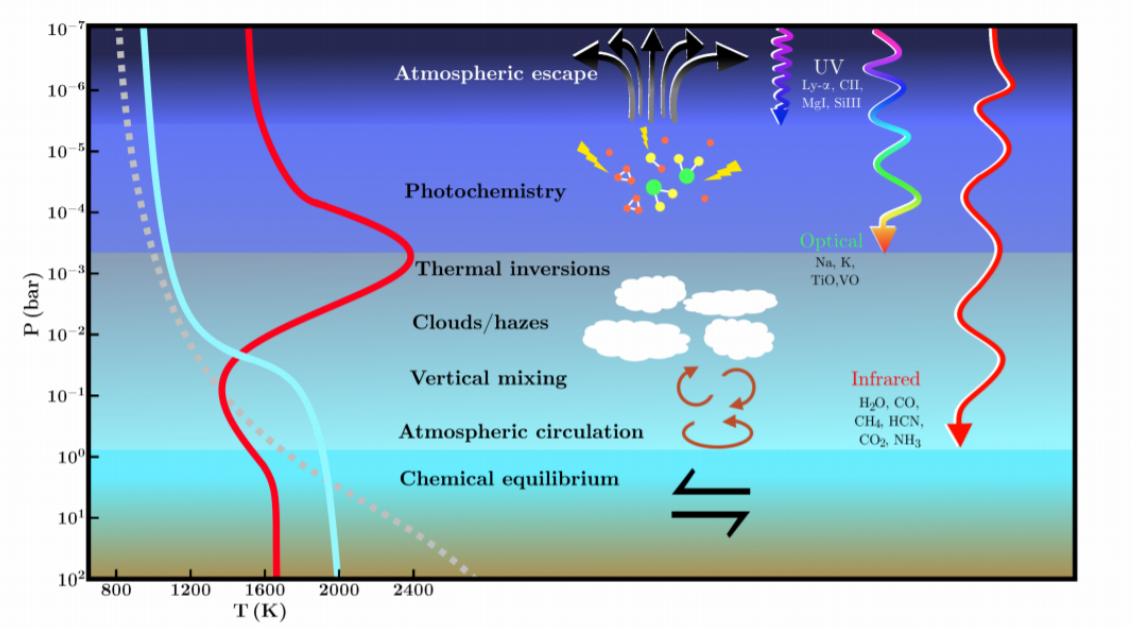

Figure 2.1: A sketch of the structure, chemistry and kinematics that govern exoplanetary atmospheres. The pressure-temperature profiles of a slightly irradiated exoplanet (grey dashed line), a moderately irradiated exoplanet with no thermal inversion (cyan line), and a highly irradiated exoplanet with a thermal inversion (red line) are shown on the left side, whereas the chemical species that can be observed at different depths with the help of ultraviolet, optical and infrared observations are depicted on the right side. The dominant physical and chemical processes at each atmospheric pressure layer are labelled and sketched in the middle. Illustration from Madhusudhan (2019).

2.3.1.1 Pressure-temperature profiles

Pressure-temperature (P-T) profiles are one of the defining characteristics of simulated atmospheres as they describe the temperature structure of exoplanetary atmospheres. Observationally, such profiles can mainly be constructed with the help of an exoplanet's thermal emission spectrum, which observes the dayside brightness temperature (and, hence, the photospheric pressure) of an exoplanet with respect to wavelength. Assumptions about the pressure-temperature profiles can also be made through transmission spectroscopy due to the relation between temperature and atmospheric pressure scale height, the parameter that regulates the strength of absorption lines (see Section 1.4.2). However, transmission spectroscopy probes the day-night transition zone where thermal emission is weaker and so constraints tend to be poor.

So far, observations of pressure-temperature profiles have revealed three regimes of hot Jupiter atmospheres. These are:

- atmospheres that experience no thermal inversion (i.e. temperature decreases monotonically with altitude),
- atmospheres that experience thermal inversion (i.e. temperature increases towards the upper atmosphere), and
- atmospheres that are isothermal (i.e. temperature is roughly the same at all atmospheric pressure layers).

Most exoplanets with equilibrium temperatures below $\sim$2000 K, as well as some ultra-hot Jupiters ($T_{eq} \geqslant 2500$ K), do not show inverted pressure-temperature profiles (e.g. Line et al., 2016). In addition, highly irradiated exoplanets may assume isothermal profiles (e.g. Nikolov et al., 2018a). Only a few ultra-hot Jupiters have been found to date showing a thermal inversion signal in their atmosphere (e.g. Evans et al., 2017).

As thermal inversions have a propensity for occurring in the hottest, most irradiated hot Jupiters, it makes sense that the presence of mineral molecules in the upper atmosphere has been suspected as a potential trigger. Metal oxides, such as

TiO and VO (e.g. Hubeny et al., 2003; Fortney et al., 2008), and metal hydrides, such as NaH and MgH (e.g. Gandhi & Madhusudhan, 2019), can show strong absorption signatures in the optical region of highly irradiated exoplanets and so they have been proposed as candidate species for thermal inversions. Forward models of simulated atmospheres, such as the ones produced by PLATON, tend to include many metal species (e.g. TiO, VO and MgH) and have the ability to explore thermal inversions. However, the exoplanets I investigated, although highly irradiated, tend to have equilibrium temperatures below $\sim$2000 K and so are very unlikely to experience thermal inversions. Therefore, when generating synthetic spectra with PLATON, I stick to the assumption of an isothermal atmosphere.

Planet equilibrium temperature Although technically the equilibrium temperature of the planet is merely an approximation and does not actually define P-T profiles, it may be useful to include it here as it is also largely considered in statistical studies (e.g. Fu et al., 2017; Edwards et al., 2022). Furthermore, the equilibrium temperature is normally used as an initial guess for the P-T profile generation of the simulated atmospheres. For example, in my forward modelling and retrievals with PLATON, I use the equilibrium temperature as input for the construction of the initial isothermal atmosphere.

The equilibrium temperature describes the effective temperature[4] of an exoplanet in the situation where all incident radiation is absorbed by the planetary atmosphere and re-radiated back to space (Seager & Deming, 2010). In this situation the exoplanet is in thermal equilibrium. If internal heating is strong, then the intrinsic luminosity produced will affect the planet's effective temperature and place the planet in a disequilibrium state. In highly irradiated gas giant exoplanets, such as the ones analysed in this thesis, the intrinsic energy from the planet's interior is relatively insignificant and can be safely ignored in most applications. Therefore, the effective temperature

[4]The effective temperature of a celestial object is defined as the temperature of a blackbody with the same size and luminosity as the celestial body itself.

is equal to the equilibrium temperature of the planet and, hence, it is expressed as (Guillot et al., 1996):

$$T_{\rm p,eff} = T_{\rm p,eq} = \left[f\left(1 - A_{\rm B}\right)\right]^{1/4} \left(\frac{R_*}{2a}\right)^{1/2} T_{*,\rm eff}, \tag{2.17}$$

where $T_{*,\rm eff}$ is the effective temperature of the star, R_* is the radius of the star, a is the orbital separation, A_B is the Bond albedo, and f is a day-night heat redistribution parameter equal to 1 for isotropic redistribution and 2 for no redistribution at all.

2.3.1.2 Chemistry and condensation

To create simulated atmospheres one important ingredient is the chemistry involved. Hot Jupiters tend to have similarities with our own Solar System gas giants in that their envelope mostly consists of molecular hydrogen and helium (e.g. Jupiter, Taylor et al., 2004). However, heavy elements and molecules, such as Na, K, H_2O, CO, CO_2, HCN, CH_4, TiO and VO can be great sources of opacity, despite normally being present at much lower concentrations. Due to their higher temperature, hot Jupiters are expected to have a different composition compared to Solar System counterparts, with different chemical species being prominent sources of opacity in the observed transmission spectra (see also Sections 1.5.1 and 1.5.2). Current atmospheric libraries tend to take into account the chemical abundances of the most prominent species for a large range of temperatures. For example, the generic grid (Goyal et al., 2019a) uses opacity contributions from 19 species while PLATON incorporates over 30 species (Zhang et al., 2019, 2020) in the calculation of the forward models.

The atmospheric models used in my work obey the rules of equilibrium chemistry, but disequilibrium chemistry models are also included for completeness. A state of chemical equilibrium means that both products and reactants have reached their final abundances following the minimisation of the Gibbs free energy that defines the system for a given P-T profile (Heng et al., 2016). According to equilibrium chemistry models, carbon is mainly found in carbon monoxide or methane molecules whereas oxygen is mostly contained in water molecules. However, disequilibrium processes can create im-

balances that may lead to an increase in the abundances of less prominent species. The use of free chemistry in atmospheric models takes into account such instabilities and allows for explorations of a wider range of atmospheres. To get a better perspective of the planetary atmospheres, both equilibrium and non-equilibrium chemistry models were examined during the interpretation of the observed transmission spectra in Chapters 4 and 5. More specifically, the chemical equilibrium models were generated with help from the generic grid and PLATON, while situations of disequilibrium chemistry were explored during the abundance retrievals with AURA. Processes that may bring an atmosphere into a state of non-equilibrium chemistry include photochemical reactions in the upper atmosphere and transport-induced quenching due to vertical mixing or circulation patterns deeper within the planetary atmosphere (Moses, 2014, see also Figure 2.1). According to a recent study by Roudier et al. (2021), very hot exoplanets ($T > 1800\,\mathrm{K}$) are expected to be in chemical equilibrium, while exoplanets with temperatures in the range 1200-1800 K are most likely to experience some form of chemical imbalance. The same study also found that disequilibrium chemistry occurs in about 50% of exoplanetary atmospheres.

Clouds are found ubiquitously among exoplanetary atmospheres as we learned in Section 1.5.2. It is, therefore, not surprising that most forward atmospheric models make an attempt at including clouds and hazes as part of the simulated atmospheres. In the generic grid used in Chapter 4 (Goyal et al., 2019a), haze is parameterised by the following equation:

$$\sigma(\lambda) = \alpha_{\mathrm{haze}}\sigma_0(\lambda), \tag{2.18}$$

where $\sigma(\lambda)$ is the total scattering cross-section, $\sigma_0(\lambda)$ is the scattering cross-section without haze enhancement, and α_{haze} is a haze enhancement factor that can be between 1 and 1100. Clouds are also parameterised as follows:

$$\kappa(\lambda)_c = \kappa(\lambda) + \alpha_{\mathrm{cloud}}\kappa_{\mathrm{H_2}}, \tag{2.19}$$

where $\kappa(\lambda)_c$ is given in $\mathrm{cm^2\,g^{-1}}$ and describes the total scattering opacity, $\kappa(\lambda)$ is the nominal Rayleigh scattering opacity, α_{cloud} is a cloudiness factor with values between 0 and 1, and $\kappa_{\mathrm{H_2}}$ is the H_2 scattering opacity at 350 nm. In PLATON, clouds are

parameterised slightly differently by just defining a cloud-top pressure where thick clouds are formed and beyond which no light can penetrate (Zhang et al., 2019). In addition, the haze enhancement factor is more flexible and can obtain any value.

Apart from clouds and hazes, atmospheric models may also consider different condensation schemes. For example, the generic grid allows for two potential condensation mechanisms: local condensation and rainout condensation (Goyal et al., 2019a). In the local condensation scenario, each atmospheric pressure layer is independent and so material that forms condensates at a specific P-T layer is only removed from the abundance recipe for this layer. In the rainout scenario, condensing material gets depleted in the local P-T layer and all P-T layers above it under the premise that condensates sink deeper into the atmosphere. PLATON also considers the rainout condensation scheme, but includes one additional scheme that ignores condensation entirely (Zhang et al., 2020). In this special scenario, all atoms and molecules remain in their gas phase for all layers of the simulated atmosphere. Since rainout condensation is common to both atmospheric modelling codes, it was the preferred option for the simulated atmospheres, but other condensation schemes were also examined briefly.

2.3.1.3 Other important parameters

Metallicity and surface gravity Aside from temperature, two important parameters are metallicity and surface gravity. These two parameters can also affect the strength of the spectral features due to their inverse relation with the atmospheric pressure scale height[5]. Their effect is somewhat similar to the effect of clouds, but clouds tend to shrink the features from below, whereas metallicity and gravity reduce the amplitude from both above and below leading to areas with decreased opacity and areas with increased opacity in the observed transmission spectrum (Goyal et al., 2019a). Also, higher gravity planets are affected by increased pressure broadening in

[5]Increased metallicity leads to an increase in the abundances of heavier chemical species, which, in turn, leads to an increase in the mean molecular mass.

the observed spectral lines due to higher overall pressures[6]. The effect is most prominent for the atomic lines of Na and K.

C/O ratio Due to carbon and oxygen existing in a great number of molecules in exoplanetary atmospheres, atmospheric models often incorporate a parameter called the C/O ratio. This parameter describes the relative abundance between the two elements and is a vital component of exoplanetary architecture. Restrictions on this parameter can mainly be placed through observations in the infrared where carbon and oxygen bearing species such as H_2O, CO and CO_2 can be observed (Madhusudhan, 2018). The most precise constraints are normally achieved through emission spectroscopy, but, in some cases, satisfactory results have also been obtained through transmission spectroscopy (e.g. Kreidberg et al., 2015). The generic grid and PLATON incorporate this parameter in the construction of the forward atmospheric models (see Chapters 4 and 5).

Under solar conditions, the C/O ratio should roughly be equal to 0.5. Any departure from this value could lead to interesting inferences about an exoplanet's formation and evolution history, in a similar way to metallicity. For example, a super-solar C/O favours formation at a large orbital separation and disk-free migration (e.g. Öberg et al., 2011), while a sub-solar C/O ratio may indicate disk migration and accretion of oxygen-rich material during this process (e.g. Madhusudhan et al., 2014b). Madhusudhan (2012) highlighted that for very hot Jupiters, such as the ones examined in my thesis, the transition between carbon-rich ($C/O \geqslant 1$) to oxygen-rich ($C/O \leqslant 1$) atmospheres can result in very steep changes in the molecular volume mixing ratios of the most dominant species. An exception to this rule is CO, which is expected to remain relatively stable.

[6]Particle concentrations are larger at higher pressures and so collision rates increase causing changes to the width of the absorption lines.

Opacities When creating synthetic spectra from simulated atmospheres it is crucial to be able to match them with the observed transmission spectrum. This can be accomplished through accurate computation of the wavelength-dependent opacities for a wide range of P-T profiles. The calculated opacities are usually a mixture of theoretical assumptions, empirical relations and observational data and can be divided into three categories: scattering opacities, line opacities and collisional opacities. Scattering and line opacities are the opacities produced due to scattering (e.g. H_2) and absorption from individual atomic and molecular species, while collisional opacities are the ones induced through collisions between the most dominant atmospheric species (e.g. H_2-H_2 and H_2-He collisions). The absorption cross sections for all chemical species considered in the forward models of Chapters 4 and 5 are generated based on absorption lines obtained from dedicated line lists. More specifically, the three tools used to decipher the observed transmission spectra (namely the generic grid, PLATON and AURA) take into account absorption lines from a variety of sources, including the three most commonly used line list databases: ExoMol (Tennyson et al., 2016), HITRAN (Gordon et al., 2017) and HITEMP (Rothman et al., 2010). At the time of writing, ExoMol offers the most precise data for higher temperatures, while HITRAN provides a more extensive list of molecules and is more sensitive to lower temperatures (Zhang et al., 2020).

2.3.2 Atmospheric retrievals

Atmospheric retrievals are the epitome of atmospheric modelling as they combine forward models with leading-edge statistical tools to fit the observed transmission spectrum and retrieve the optimised parameters for the planetary atmosphere. This method, therefore, also provides uncertainties for the atmospheric parameters. The most common parameter estimation techniques utilised are MCMC (also exploited during light curve fitting) and nested sampling. PLATON is quite flexible in this respect as it can take any of the two Bayesian frameworks into account for the atmospheric model fits (Zhang et al., 2019). Since MCMC sampling here is performed using the

same tools and following a similar methodology to the one described in Section 2.2.4.2, the focus in the next few paragraphs will be on nested sampling. Furthermore, even though MCMC retrievals with PLATON were performed, only nested sampling results will be presented in Chapter 5 as it was the preferred Bayesian inference method for the atmospheric retrievals.

2.3.2.1 Nested Sampling

Nested sampling is another Bayesian inference method, similar to MCMC, but with the basic difference that the sampler starts with a number of randomly produced "live points" drawn from the prior distribution and not from one initial guess (or a number of points in a tight region around an initial guess) to investigate parameter space (Skilling, 2006). The "live points" steadily move towards a parameter space region of higher likelihood as live points with the lowest likelihood are constantly being replaced by new points with greater likelihoods drawn from the same prior distribution. This results in a gradual decrease of the area enclosed by the iso-likelihood contour of the lowest likelihood as old points are substituted by new ones during the optimisation process. The sampling method calculates the Bayesian evidence in every step of the procedure and ends when a pre-determined tolerance (e.g. when changes in the approximation of the evidence start to become small) is attained. This given tolerance guarantees that the samples have converged. The posterior probability distributions are then constructed from all the points sampled during the entire fitting procedure.

Here, nested sampling is utilised in the context of the atmospheric retrievals with PLATON (Chapter 5) and AURA (Chapters 4 and 5). More specifically, the PLATON retrievals perform the sampling through the use of the nested sampling package `dynesty` (Speagle, 2020), while AURA relies on the algorithms of PyMultiNest (Feroz et al., 2009; Buchner et al., 2014). One benefit from performing nested sampling in these retrievals is the efficient estimation of the Bayesian evidence. This parameter is quite useful as its value can be used for model selection purposes. In fact, the logarithmic evidence is used to provide the statistical detection significance between two

simulated atmospheres in the WASP-88b retrieval with AURA (see Section 4.5.3), and in the WASP-74b retrieval with PLATON (see Section 5.5.2). Nested sampling is often the go-to approach in atmospheric retrievals (see e.g. Knutson et al., 2014; MacDonald & Madhusudhan, 2017; Sedaghati et al., 2017b).

3 From raw data to a final product: WASP-75b as a test case

3.1 Introduction

During the first few months of my research, I worked on spectroscopic data obtained for the study of WASP-75b. This was a test target to get a background understanding of transmission spectroscopy and to get acquainted with the processes involved from obtaining the raw data to generating the observed transmission spectrum. In the following paragraphs, I will present what we know about the target from the literature, describe the new observations, concentrate on the data reduction pipeline I created, and show results from the transit light curve analysis conducted. Although some results will be presented here, this Chapter should not be considered as a focused atmospheric characterisation chapter, such as Chapters 4 and 5, but rather as an extension to Chapter 2, with more attention being given to the tools used and the methodology followed instead of the results produced. In a way, it is a preparatory challenge before the analysis of the main targets of this thesis: WASP-88b and WASP-74b.

3.1.1 The WASP-75 system in detail

The WASP-75 system is a planet-hosting system that consists of a Sun-like star and a marginally inflated hot Jupiter (Gómez Maqueo Chew et al., 2013). The planetary system was detected as part of the SuperWASP survey (Pollacco et al., 2006) and is located in the southern celestial hemisphere having a right ascension of 22:49:32.6 and a declination of −10:40:32. The discovery was announced by Gómez Maqueo Chew et al. (2013) and was based on both WASP-South and WASP-North observations. Further constraints on the system parameters were placed using follow-up photometric and spectroscopic observations. The latest information on the planetary parameters was given by Clark et al. (2018) who reported new results based on transit data obtained

during the K2 mission[1], and the most up-to-date stellar parameters were determined by Sousa et al. (2021) and can be found in the SWEET-Cat database[2]. Some of these parameters are presented in Table 3.1. The system is slightly younger than our Solar System with an age of ∼3-4 Gyr (Gómez Maqueo Chew et al., 2013). In addition, the host star has an apparent magnitude of V=11.45 mag and is a borderline F dwarf with a spectral type of F9V. The exoplanet, on the other hand, is a hot Jupiter that does not possess a highly inflated radius and has an atmospheric pressure scale height of less than $\sim$ 450 km (Gómez Maqueo Chew et al., 2013). This value was computed under the assumption of a pure H_2 atmosphere ($\mu \approx 2$) and so the actual atmospheric pressure scale height should be lower. According to the discovery paper, this estimate is too low to get reliable results from most transmission spectroscopy observations and JWST may be the only possible option to observe WASP-75b's atmosphere.

Clark et al. (2018) also revealed a few more valuable properties of the WASP-75 system. First of all, they did not detect any significant transit timing and duration variations in the transit light curves. This finding rules out any large additional exoplanets in close orbits. In addition, they also did not detect additional transiting planets through inspection of the model-subtracted residuals and determined that the phase curves of WASP-75b showed no variations. Interestingly, they found signs of stellar variability in the long-term light curves and identified a periodic signal of 11.2 $\pm$ 1.5 d. However, they did not observe any starspot occultations throughout the data.

From all the above information, there are three points that make the observation of WASP-75b's atmosphere quite challenging. These are the low atmospheric pressure scale height, the high impact parameter (indicative of a transit that occurs close to the stellar edge), and the presence of activity in the host star. Despite the inconvenient properties of this target, I made an attempt to construct a transmission spectrum. This is because the expectation from this exercise was to learn a bit about the steps followed in the characterisation of an exoplanetary atmosphere via transmission spectroscopy

[1]This mission was put forward after the cancelling of the original Kepler mission due to a failure of two reaction wheels (Howell et al., 2014).

[2]`http://sweetcat.iastro.pt/`

Table 3.1: WASP-75 system parameters based on most recent works.

Parameter	Value	Reference
Distance (pc)	295.8 ± 1.6	Bailer-Jones et al. (2021)
Impact parameter $b = a \cos i/\mathrm{R}_*$	0.8926 ± 0.0007	Clark et al. (2018)
Time of mid-transit t_0 ($\mathrm{BJD_{TDB}}$)	2457009.94594 ± 0.00002	Clark et al. (2018)
Orbital period P (d)	2.4842014 ± 0.0000004	Clark et al. (2018)
Effective temperature T_{eff} (K)	6102 ± 27	Sousa et al. (2021)
Metallicity [Fe/H]	0.24 ± 0.02	Sousa et al. (2021)
Stellar density ρ_* ($\rho_\odot$)	0.566 ± 0.003	Clark et al. (2018)
Stellar surface gravity log g_* (c.g.s.)	4.337 ± 0.011	Sousa et al. (2021)
Stellar mass M_* ($\mathrm{M}_\odot$)	1.199 ± 0.009	Sousa et al. (2021)
Stellar radius R_* ($\mathrm{R}_\odot$)	1.271 ± 0.042	Sousa et al. (2021)
Planet/star radius ratio $k = R_\mathrm{p}/R_*$	0.106 ± 0.002	Clark et al. (2018)
Scaled semi-major axis a/R_*	6.38 ± 0.23	Clark et al. (2018)
Semi-major axis a (AU)	0.0377 ± 0.0006	Clark et al. (2018)
Inclination i (degrees)	81.96 ± 0.02	Clark et al. (2018)
Eccentricity e	0 (fixed)	Clark et al. (2018)
Planet mean density ρ_p (ρ_Jup)	0.48 ± 0.02	Clark et al. (2018)
Planet surface gravity log g_p (c.g.s.)	3.16 ± 0.01	Clark et al. (2018)
Planet mass M_p ($\mathrm{M_{Jup}}$)	1.08 ± 0.05	Clark et al. (2018)
Planet radius R_p ($\mathrm{R_{Jup}}$)	1.31 ± 0.02	Clark et al. (2018)
Planet equilibrium temperature T_eq (K)	1688^{+25}_{-26}	Clark et al. (2018)

and not to produce publication quality results. The following sections will therefore outline the whole process followed, starting from the raw data and ending with the final product.

3.2 Observations

The observations for WASP-75b were performed using the New Technology Telescope (NTT) located at the European Southern Observatory (ESO) facility in La Silla, Chile. The NTT[3] is one of the two 3.6 m telescopes located at this site. The other one is the ESO 3.6 m Telescope, which is well-known for the High Accuracy Radial velocity Planet Searcher (HARPS). This instrument has played a crucial role in the discovery of numerous exoplanets via the radial velocity method (e.g. Lovis et al., 2011) and has even made observations of exoplanetary atmospheres via high-resolution Doppler spectroscopy (e.g. Wyttenbach et al., 2015; Seidel et al., 2020). NTT, on the other hand, has not been widely used for the observation of exoplanetary atmospheres. This makes sense as the diameter of this telescope is lower than the typical 8-10 m class telescopes used for transmission spectroscopy from the ground.

WASP-75b was observed as part of the ESO program 099.C-0301(A) led by principal investigator John Southworth. The instrument employed for the one-night observation of WASP-75b was the Faint Object Spectrograph and Camera (EFOSC2). This instrument offers up to six different observing modes, has a $4.1' \times 4.1'$ field of view, and can cover the wavelengths between 305 nm and 1100 nm. The data were recorded on a single Charge-Coupled Device (CCD) chip of size equal to 2048×2048 pixels but the image was visually split in two pieces because of the use of two different amplifiers. For the observation of WASP-75b, the long-slit spectroscopy mode was chosen to allow for the use of long slits with predetermined sizes. The data were binned using the normal mode and a binning of 2×2, which resulted in an image readout and transfer time of $\sim$20 s.

[3]To be precise, NTT has a primary mirror diameter of 3.58 m.

The target was monitored for an entire transit during early August 3rd 2017 (i.e. the observing started after 00:00 UT on August 3rd). A long slit with a width of $15''$ was utilised to record the spectrum on the single CCD chip. During this observation, Grism#11 was used to disperse starlight to a wavelength range between 3380 and 7520 Å. The observation commenced at 04:28 UT and lasted for $\sim$6 hours and 2 min. During this time, WASP-75b was tracked from an airmass of 1.25 to an airmass of 1.06 and then to an airmass of 1.78. The sky was clear throughout the observation and no clouds were reported by the observer. The final product is a data set that consists of 154 exposures with integration times of 120 s. These data were saved into Flexible Image Transport System (FITS) files, which follow ESO's specifications and have detailed headers.

A key distinction of ground-based observations from those obtained in space is the presence of the Earth's atmosphere. This additional source of noise can have a significant impact in ground-based measurements and needs to be dealt with in order to produce reliable results (see also Section 1.6). One established approach to handle line-of-sight variations, caused by e.g. changes in airmass or cloud coverage, is to observe and get the spectrum of a similar star in close proximity to the target (e.g. Nikolov et al., 2021). The small angular separation (typically up to a few arcmin) between the target and the nearby star (also called a reference or comparison star), ensures that the atmospheric path of incident light is similar in length and increases the probability that both objects are affected by similar extinction profiles and other telluric effects. Atmospheric systematics can therefore be largely corrected for by dividing the flux of the target by the flux of a suitable comparison star. The correction is most effective when the comparison star displays a similar spectral energy distribution to the target. In the case of WASP-75, in particular, we are quite fortunate to have a star of similar brightness and spectral distribution at a very close distance. This star, also known as 2MASS J22492971−1039261, has an angular distance of just $1.3'$. The star's spectral trace was simultaneously recorded on the same CCD chip as the target, but was obtained using a different amplifier.

3.3 Data selection and wavelength calibration

One of the first steps in the data reduction process was to check if bias subtraction and flat field corrections could be performed. Bias frames are frames taken with the shutter closed and no exposure time, and are used to correct for the offset in counts induced during readout. Flat fields, on the other hand, are raw images of a light source that have the same slit[4] and exposure time specifications as the science images and are used to correct for variations in pixel-to-pixel sensitivity caused by e.g. detector anomalies or grains of dust trapped in the optical path of the instrument. During the WASP-75b observations, no bias frames were obtained and only a few flat fields were taken. More specifically, the acquired data included one lamp flat and six dome flats. However, the quality of the flat fields was sub-optimal with systematically lower counts in the blue and a huge difference in counts between the blue and the red end of the image. For this reason, flat field corrections were not applied to the raw spectroscopic images of the two stars.

To define the apertures for the spectral profiles of WASP-75 and the comparison star, I first fitted both spectra with Gaussians in the cross-dispersion axis for every ten stacked pixel lines. Stacking ten rows together was useful as it accelerated the aperture determination process. In this way, the spectral peaks in the dispersion axis were approximately determined. These peaks were then fitted with 3^{rd}-degree polynomials so that the aperture of each spectrum could be defined precisely. The procedure was implemented for all images in the time series and the aperture width was determined with respect to the limit where the spectral signal was balanced out by the background noise. The spectral trace of both objects was also shortened in the dispersion axis to remove pixels without photon counts and blue regions where the signal-to-noise ratio was low. The final spectral regions had an aperture radius of 36 pixels and a dispersion length of 900 pixels. One important point to note here is that the identification of each spectrum was done manually on the first image. In retrospect, spectral identification

[4]Assuming that the flats are taken for spectroscopic observations.

could be simplified and automated based on spectral peak detection algorithms. Such a change would have no effect on the final result but would be useful as the code would then require minimal changes when applied to different targets.

There were several sources of noise that could influence the observed spectral signal. One of them is the contribution from the sky background level. This is seen as additional pixel counts in seemingly empty parts of the CCD images (i.e. areas where no visible sources of light are present). These sky background counts could be a result of instrumental and astrophysical effects (e.g. faint unresolved stars), or due to light diffusion (e.g. moonlight, twilight, and human-induced light pollution). Part of the background is also created by the bias level. This type of stray light can be erased by specifying background noise regions on either side of the spectral profiles of both stars, calculating the median count value of both backgrounds, subtracting these values from the spectral signal defined by the aperture in the cross-dispersion axis, and repeating the procedure for all rows in the dispersion axis and for all images in the time-series (see e.g. Nikolov et al., 2016, 2018b). This box summation method was chosen for the removal of background noise and is a useful step as it eliminates the contribution from bias as well. Another option would be to fit a polynomial to both background regions (by masking the region in between). This would be a useful approach in a scenario where a clear slope is visible in the sky background. However, for this specific data set, the fit revealed no significant slope and its application made no difference to the final outcome. This idea was therefore rejected. In the end, background noise was removed through the use of 50-pixel sky region boxes on both sides of the spectral profiles. Example spectra of WASP-75b and the reference star, as well as the contribution from the sky background, are shown in Figure 3.1.

Another potential noise contribution is cosmic rays. Cosmic rays are highly energetic particles of solar, galactic or extra-galactic origin that can occasionally interact with the material of the detector causing a sudden increase in counts that is visible as bright spots or lines with clearly defined edges. The probability of interaction increases with longer exposure times and, in the case of WASP-75, the exposures are moderately long. This means that cosmic rays are frequently present, but most of

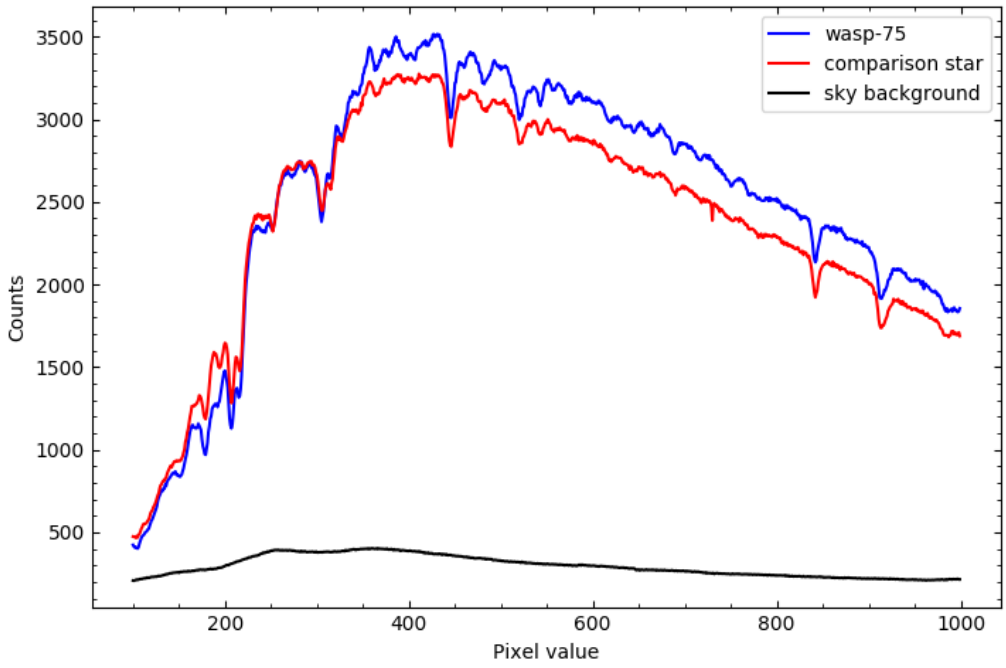

Figure 3.1: Example spectra of WASP-75 (blue line) and the nearby reference star 2MASS J22492971−1039261 (red line). The background contribution is depicted by a black line.

the time they leave their trace far from the spectral profiles. A way to remove them is to use the cosmic ray rejection algorithm developed by van Dokkum (2001), which relies on Laplacian edge detection. The original code is written in IDL, but I used the python implementation `astroscrappy`[5] (McCully & Tewes, 2019). The detection and removal of cosmic rays was performed to the raw images before performing any other data reduction procedures such as defining the spectral apertures and the sky background regions. `astroscrappy` is extremely fast (as the python functions actually call subroutines written in C) but not 100% foolproof and some tweaking with the input parameters may be required in some images. Figure 3.2 shows two images before and after the application of this cosmic ray elimination method.

Other sources of noise considered during data reduction include photon noise, scintillation, and detector artifacts:

- Photon noise (or shot noise) is the main source of error in the background-corrected frames and is a vital component of the error budget in the flux measurements. This form of noise is caused by variations in the emitted stellar photons that reach the detector and leads to pixel-to-pixel fluctuations in the observed counts. The scatter of the data follows a Poisson distribution and the standard deviation is given by the square root of the total counts. Space-based observations tend to reach the photon noise limit (e.g. Spake et al., 2018), while ground-based observations are usually affected by increased scatter (e.g. Bean et al., 2013). Photon noise uncertainties were computed for all data points in the reduced transit light curves of WASP-75b.

- Scintillation is described in Section 1.6 and is a product of the interaction between starlight and the Earth's turbulent atmosphere. The error due to scintillation for the WASP-75b data was also calculated, based on Young's equation (Dravins et al., 1998):

$$\sigma = \frac{0.09 D^{-\frac{2}{3}} A^{\frac{7}{4}} e^{-\frac{h}{H}}}{2t^{\frac{1}{2}}}. \tag{3.1}$$

[5] `https://github.com/astropy/astroscrappy`

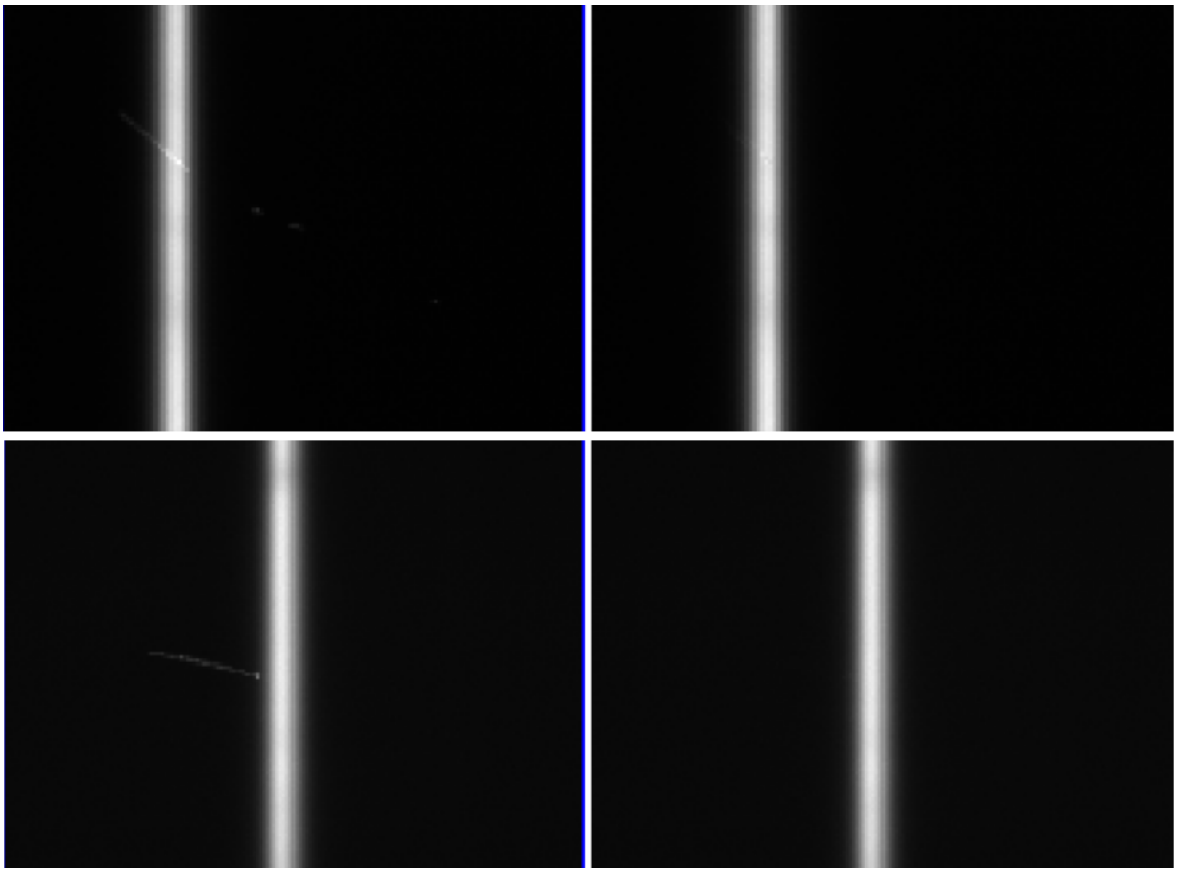

Figure 3.2: The effect of the cosmic ray rejection algorithm on two spectral images. The images on the left are before cosmic ray removal and the images on the right are after.

Here, D defines the diameter (in cm), A denotes the airmass, h gives the altitude of the observer's location (in m), H describes the atmospheric pressure scale height (also in m), and t is the exposure time (in s). Since all measurements were taken using the same exposure time of 120 s, the only variable parameter in Equation 3.1 is airmass. When plugging in all the relevant values, the error due to scintillation was estimated to be quite low and around the order of $\sim 10^{-4}$. Such a low number is expected given the high integration times. These errors were at least an order of magnitude lower than the shot noise uncertainties and so were not taken into account for the remainder of the analysis.

- Detector artifacts can be present in various forms in the raw CCD frames. These tend to be caused by impurities of the CCD camera and can add unwanted noise to the observed spectral signal. Detector artifacts commonly observed in CCD images include black pixels, hot pixels and charge traps. These can be isolated in single pixels, clustered into bigger regions or affect a large part of a column. In addition, a variable illumination due to differences in pixel sensitivity across the full image can induce a more diffused noise feature. Such imperfections can be spotted in the flat field images and this is why these exposures are often used to correct science images. During the determination of the sky background and spectral regions, extra care was taken to avoid dodgy pixels (with help from the flat field exposures). Thankfully, the spectral signatures of the target and the comparison star did not seem to be (strongly) affected by defective pixels of this kind.

After dealing with all the systematics, it was time to construct the white transit light curves. These can be created by summing the cosmic ray- and background-subtracted counts of each spectrum. The summation was first done in the cross-dispersion axis and then in the dispersion axis. The result was two total count values, one for WASP-75 and another for the comparison star, for every image in the time-series. These values were normalised, through division by the mean, and then used to

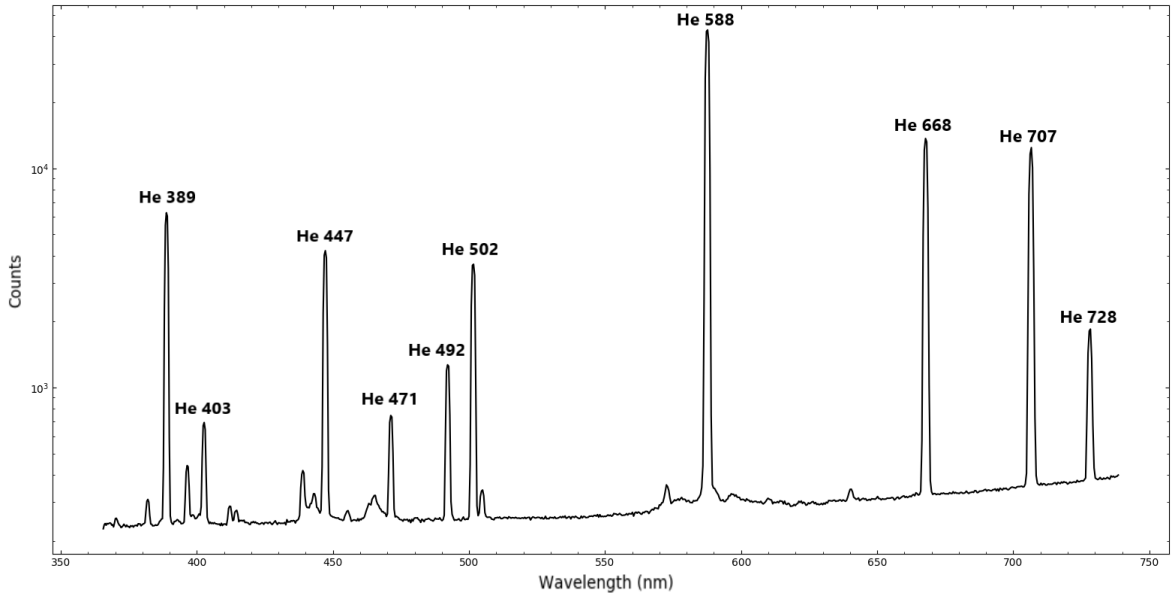

Figure 3.3: Example optical spectrum of helium emission lines. Wavelength values are given for the 10 strongest helium lines.

create relative flux values. As already mentioned in Section 3.2, relative light curves are produced by dividing the light curves of the target by the light curves of the comparison star. This technique creates differential light curves and so any additional conversions (e.g. from counts to electrons) are actually unnecessary as they are cancelled out during division. Differential light curves are useful as they are able to remove atmospheric effects such as extinction to great extent (see also Chapter 5).

To create binned spectrophotometric light curves, which are needed for the production of the transmission spectrum, the light curves need to be calibrated with respect to wavelength. In fact, this was also done for the white light curves, but the necessity of this procedure is more obvious here. For the wavelength calibration, a helium lamp was employed to create an image of the arc spectrum that shows helium emission lines. For this purpose, a slit with a width of $1''$ was utilised. Figure 3.3 shows an example helium spectrum with all observable lines in the wavelength region investigated. The most significant lines are mentioned by their nm value.

The procedure for the wavelength calibration was similar to the approach followed in the determination of the aperture for the spectral profiles of WASP-75 and the comparison star. This means that a combination of Gaussians and polynomials was used to locate the centroids and define the two-dimensional trace of each helium line. An initial calibration was made by applying a polynomial fit to the six most prominent helium lines. Six additional, but weaker, helium lines were also used as "control group" (i.e. their model-subtracted residuals were used to check for deviations from the six-line model fit). Polynomials up to 5^{th}-degree were tested but, in the end, 4^{th}-degree polynomials were preferred. From Figure 3.4 and the depicted example residuals, we see that while 5^{th}-degree polynomials fit the most prominent lines excellently, they do not offer an improvement to the fit with respect to the non-fitted, weaker lines. In addition, the residuals show some scatter from column to column (in the cross-dispersion axis), with the biggest differences being observed for the two weakest lines. In Figure 3.5, the minimum and maximum residual values of each line are displayed, as obtained from 4^{th}-degree polynomial fits. From there, it can be seen that the two weakest lines were characterised by the highest scatter. Therefore, one final calibration was performed using the 10 strongest helium lines, while excluding the other two. The obtained wavelengths were then mapped with respect to the indices of the two stellar spectra to get the calibrated spectra.

After wavelength calibration, the data were split into 12 spectrophotometric bins according to pixel values. This led to a slight, non-linear increase in the size of the bins towards longer wavelengths. This way of separation was possible as the offset between the two spectra was below 1 Å across the whole dispersion range and the average dispersion was ~ 4 Å per double pixel[6]. In retrospect, the bins should have been divided more evenly based on defined wavelength regions. Nevertheless, relative transit light curves were then produced for all spectroscopic channels following the same methodology described for the white light curves.

One final adjustment to the reduced data was the redefinition of the time stamps.

[6]Because a binning of 2×2 was used.

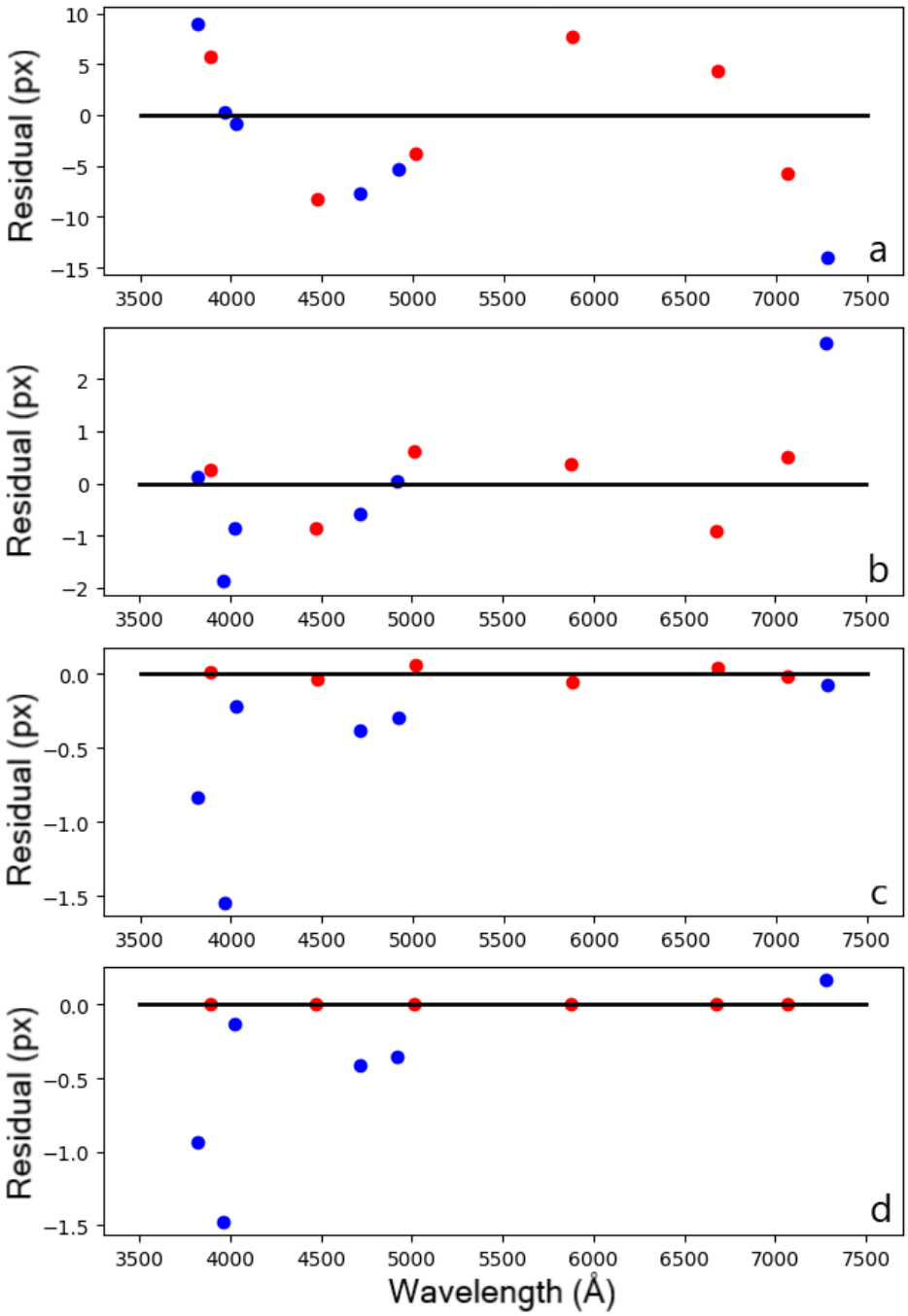

Figure 3.4: Example residuals from all 12 (6+6) lines used during the initial wavelength calibration of WASP-75. Panels a through d show results for fits of polynomial degrees between 2 and 5, respectively. Red data points depict fitted lines, while blue data points show other lines not included in the fit. This result is for an example image.

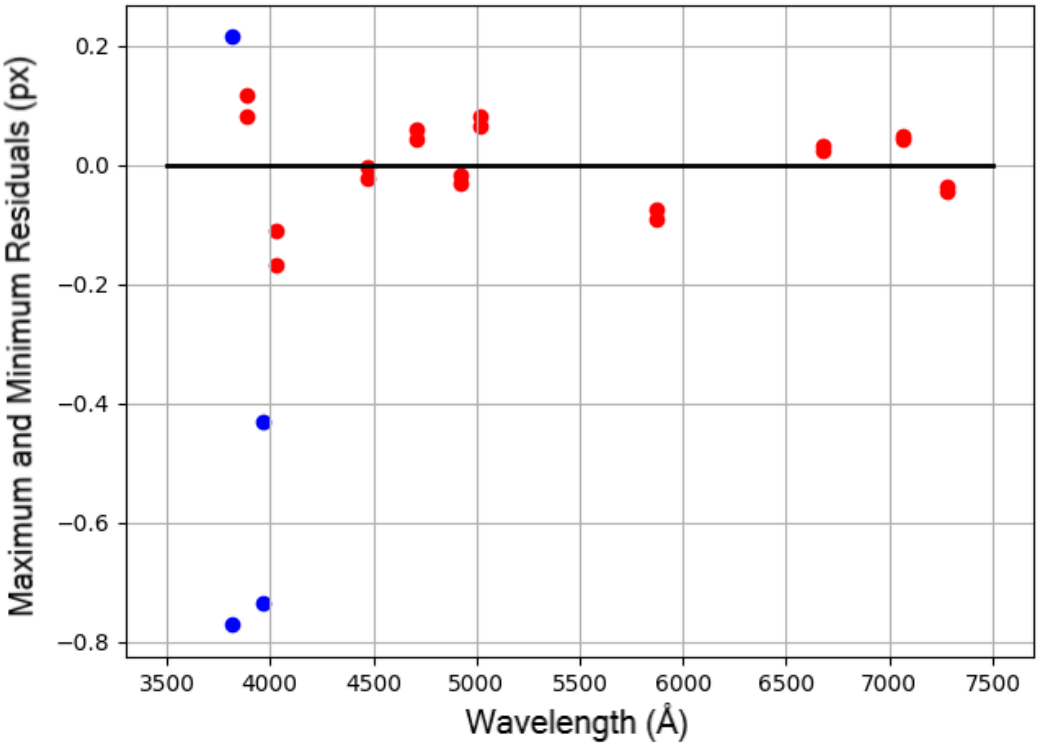

Figure 3.5: Maximum and minimum residual points from a column inspection of all 12 (6+6) lines used during the initial wavelength calibration of WASP-75. The distance is higher for weaker lines, indicating a higher scatter. Strong lines are marked in red, while the two weakest lines are marked in blue. The two weakest lines were ignored during re-calibration. This result is for an example image.

This adjustment was important as having a consistent time standard throughout the literature is considered to be good practice. Eastman et al. (2010) argue that the use of Barycentric Julian Dates is the most practical solution as it provides a high precision to time-series observations (a precision of 1 s or less can be typically achieved these days in certain applications). High-precision measurements are e.g. critical for the accurate determination of an ephemeris and, consequently, the accurate prediction of future transits by a transiting exoplanet. Eastman et al. (2010) state, among other things, that the variety of time standards used in the literature can potentially lead to spurious claims of transit timing variations, and suggest that a conversion from Julian Dates UTC to Barycentric Julian Dates TDB[7] would minimise such timing discrepancies. For this reason, they created an IDL pipeline that performs all the timing corrections required for such a conversion. In my work, I use the python adaptation of this pipeline `barycorrpy` (Kanodia & Wright, 2018). For the WASP-75b observations, in particular, I applied this code to the entire time-series by converting Modified Julian Dates to Barycentric Julian Dates.

3.4 Analysis

The relative white flux produced during data reduction is characterised by a slope that goes upwards with time. This slope is mostly a result of additional instrumental and atmospheric systematics and needs to be removed or separated in order to fit the transit model. A range of time polynomial degrees was explored, with 3rd-degree polynomials providing the best fit. Just like during wavelength calibration, higher-degree polynomials offered little improvement to the fit and were, therefore, rejected for the systematic effects correction. The relative white light curves were then divided by the polynomial of time to obtain the rescaled transit light curves, which were properly aligned with the fitted transit models. In hindsight, a better solution for the removal of the noise

[7]TDB: Barycentric Dynamical Time.

would be to create a two-component function that includes both the polynomial of time and the transit model and to perform a simultaneous fit of the noise and the transit parameters.

The corrected transit light curves were fitted with transit models produced by the open source python library `batman` (Kreidberg, 2015). This package was chosen as it offers a variety of stellar limb darkening laws and includes algorithms that compute the transit models fast and accurately. `batman` was, therefore, used in the modelling of both the white and binned (spectroscopic) light curves. Due to the prevalent use of the quadratic limb darkening law in ground-based transmission spectroscopy observations (e.g. Nikolov et al., 2016, 2018b; Carter et al., 2020; Chen et al., 2022), this law was preferred over others during parameterisation. In this situation, `batman` precisely follows the analytic formalism of Mandel & Agol (2002) in the modelling of the transit light curves.

The median transit parameters and their standard deviation were then inferred from the transit models by performing an MCMC maximum likelihood estimation using the python module `emcee`[8]. For the white transit light curves, the orbital period (P), the star-to-planet radius ratio (R_p/R_*), the orbital separation to star radius ratio (a/R_*), inclination (i), and eccentricity (e) were initially set to their values determined by Clark et al. (2018). In addition, the expected time for the transit mid-point (t_0) was set to zero for fitting purposes. Parameters t_0, R_p/R_*, a/R_* and i were allowed to vary freely under uniform priors, while P and e were fixed. Throughout the analysis, the orbit of WASP-75b was assumed to be circular, in line with Clark et al. (2018). An additional noise parameter σ_x was also included to adjust the fit for unaccounted for noise contributions.

The values for the quadratic limb darkening coefficients were acquired from tables produced by Claret & Hauschildt (2003) according to theoretical models. These tables included computed values of the quadratic limb darkening coefficients based on a grid of stellar metallicities, stellar effective temperatures, and stellar surface gravities, and

[8]More details about the `emcee` estimation method are given in Chapter 4.

a selection of commonly used filters. To get the wavelength-dependent limb darkening coefficients of WASP-75, the Johnson U, B, V, R and I photometric bands were used and an interpolation was performed on the given set of grid values. These values were then mapped against the central wavelengths of each passband and were fitted with a polynomial to obtain an estimate of the quadratic limb darkening coefficients for each spectroscopic channel. For the white flux, the solution was simpler as only the V-band was used for the computation of the initial quadratic coefficients. Parameterisations with zero, one or two free limb darkening coefficients were examined and Gaussian priors were applied to both coefficients due to a high correlation between the two parameters.

A separate analysis was performed for the spectroscopic light curves. The analysis was independent, meaning that no prior assumptions were made for the transit parameters based on the white light curve analysis. Therefore, the free parameters were similar to the white-light analysis, with the important distinction that parameters t_0, a/R_*, i and σ_x were the same for all spectroscopic light curves, while the transmission spectrum parameter R_p/R_* was different for each spectroscopic light curve[9]. All other parameters were kept fixed, with the limb darkening coefficients being estimated following the same principles described earlier. In the same way as R_p/R_*, the limb darkening coefficients u_1 and u_2 have a wavelength dependency and so they were different for each light curve.

During both analyses, an outlier removal algorithm was also implemented. The applied function was able to spot deviating data points according to the root mean square error (RMSE) of the data points. Residuals that were the furthest from the best-fit transit model were discarded in a second fit. Only extreme data points that deviated by a factor of a few were removed.

The undetrended and detrended white light curves, as well as the undetrended and detrended spectroscopic light curves, are depicted in Figures 3.6, 3.7, 3.8 and 3.9, respectively. The final light curves were produced after several reductions were applied,

[9]Essentially, this was a global fit. A variation of it is described in Chapter 5.

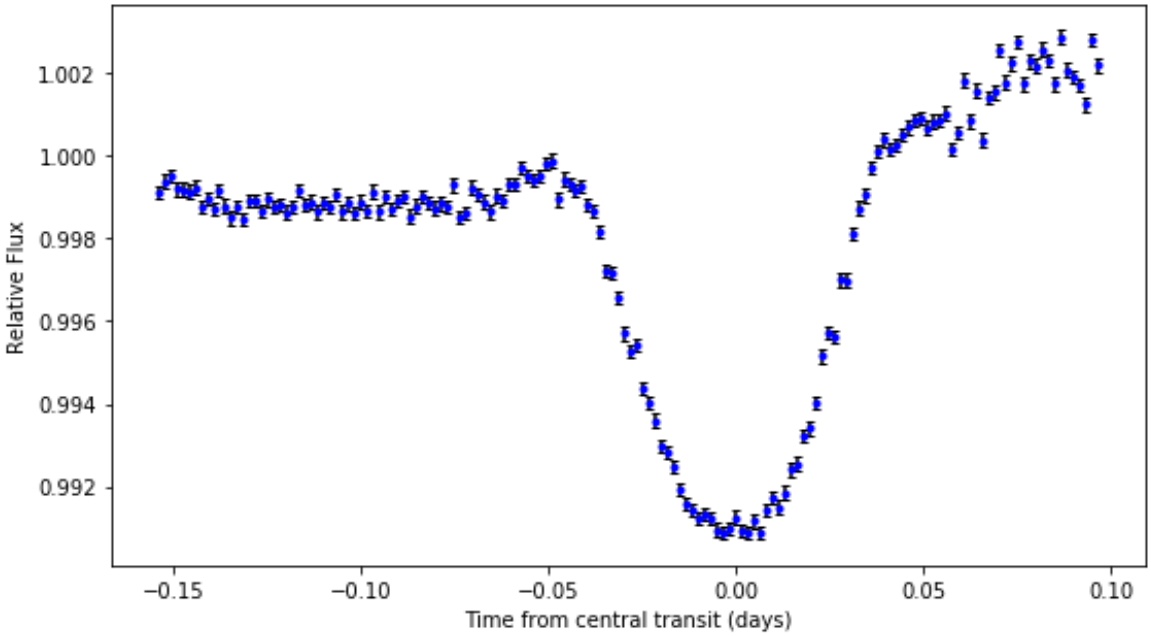

Figure 3.6: Undetrended white light curves of WASP-75b (blue data points). The data points are depicted with their 1σ uncertainties.

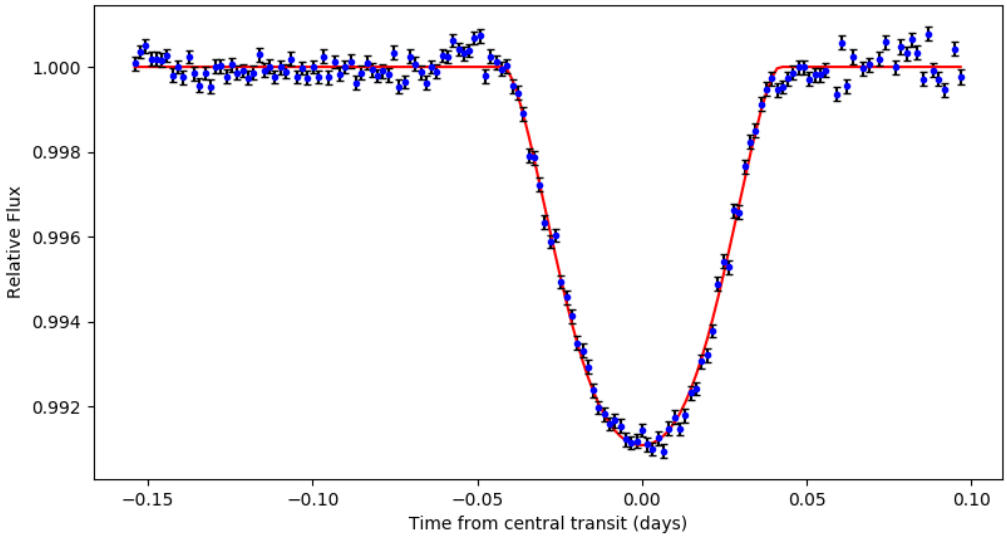

Figure 3.7: Detrended white light curves of WASP-75b (blue data points) and median transit models (red lines). The data points are depicted with their 1σ uncertainties.

Table 3.2: White light curve parameters of WASP-75b.

Parameter	Value
P (d)	2.484201 (fixed)
e	0 (fixed)
t_0 ($\mathrm{BJD_{TDB}}$)	2457968.84767 ± 0.00017
R_p/R_*	$0.1058^{+0.0032}_{-0.0017}$
a/R_*	$6.19^{+0.11}_{-0.13}$
i (degrees)	$81.66^{+0.21}_{-0.26}$
$u1$	0.495 (fixed)
$u2$	0.202 (fixed)
σ_x	0.00024 ± 0.00002

including cosmic ray removal, outlier removal and removal of time-related systematics (see also Section 3.3). Furthermore, the final model fits relied on fixed limb darkening coefficients. This is because the fit resulted in very poor constraints on the R_p/R_* values when the coefficients were considered as free parameters. Such a result is to be expected as Müller et al. (2013) demonstrated that, when the impact parameter is $\geqslant$0.8, then the limb darkening coefficients measured from transit light curve fits are not reliable.

The final system parameters from the white and spectroscopic fits of WASP-75b are shown in Tables 3.2 and 3.3, respectively. As can be seen from these tables, the highly correlated parameters a/R_* and i were found to be slightly lower than the high-precision results reported by (Clark et al., 2018). Overall, there was a good agreement between the white-light and the spectroscopic-light estimates, with most differences being below 1σ. Only the time of mid-transit from the white-light fits was found to deviate by $\sim 1.5\sigma$ from the spectroscopic fits. Again, a better treatment of the systematics would have given more confidence about the truth of these values.

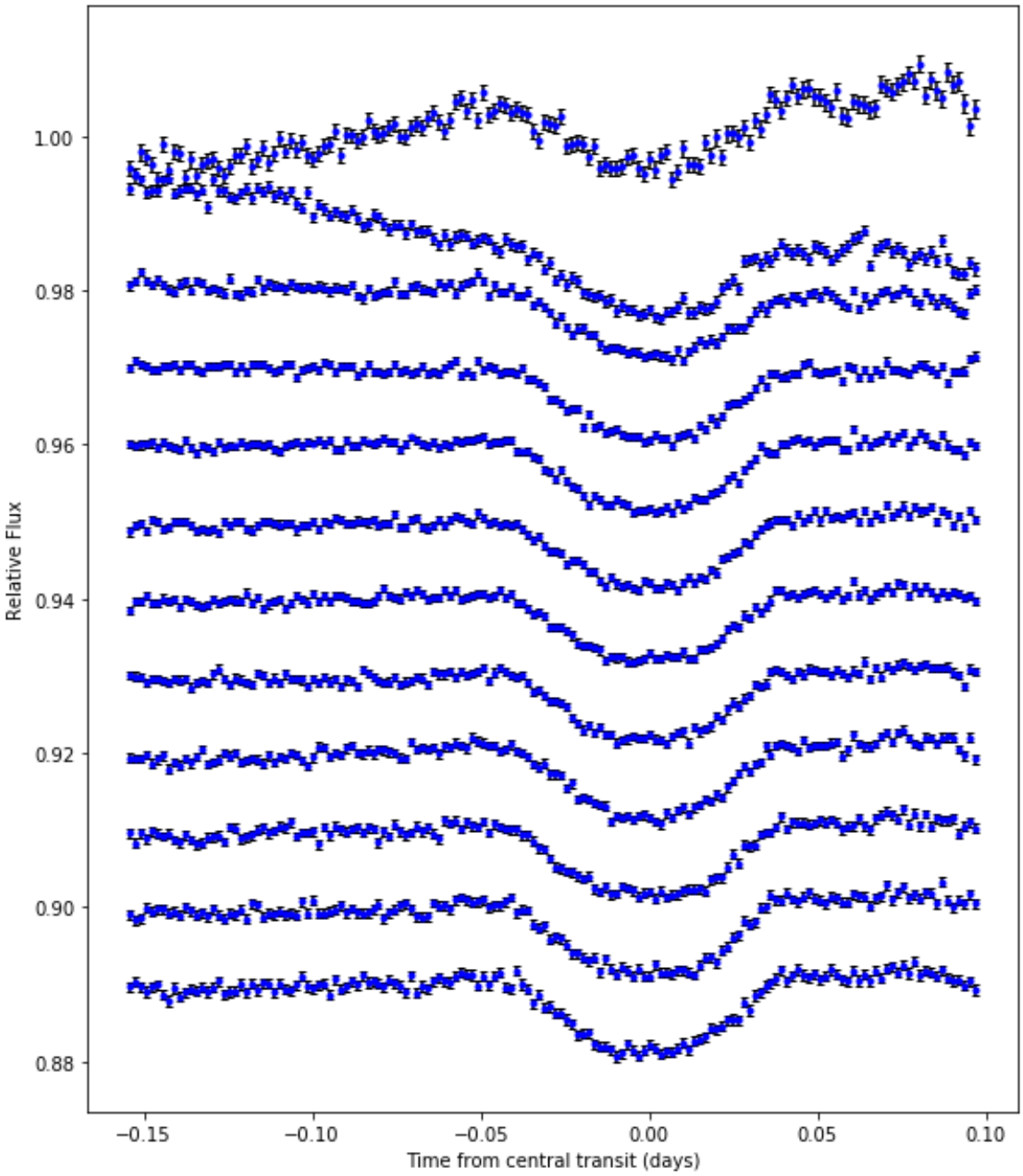

Figure 3.8: Undetrended spectroscopic light curves of WASP-75b (blue data points). The data points are shown with their spectrophotometric uncertainties.

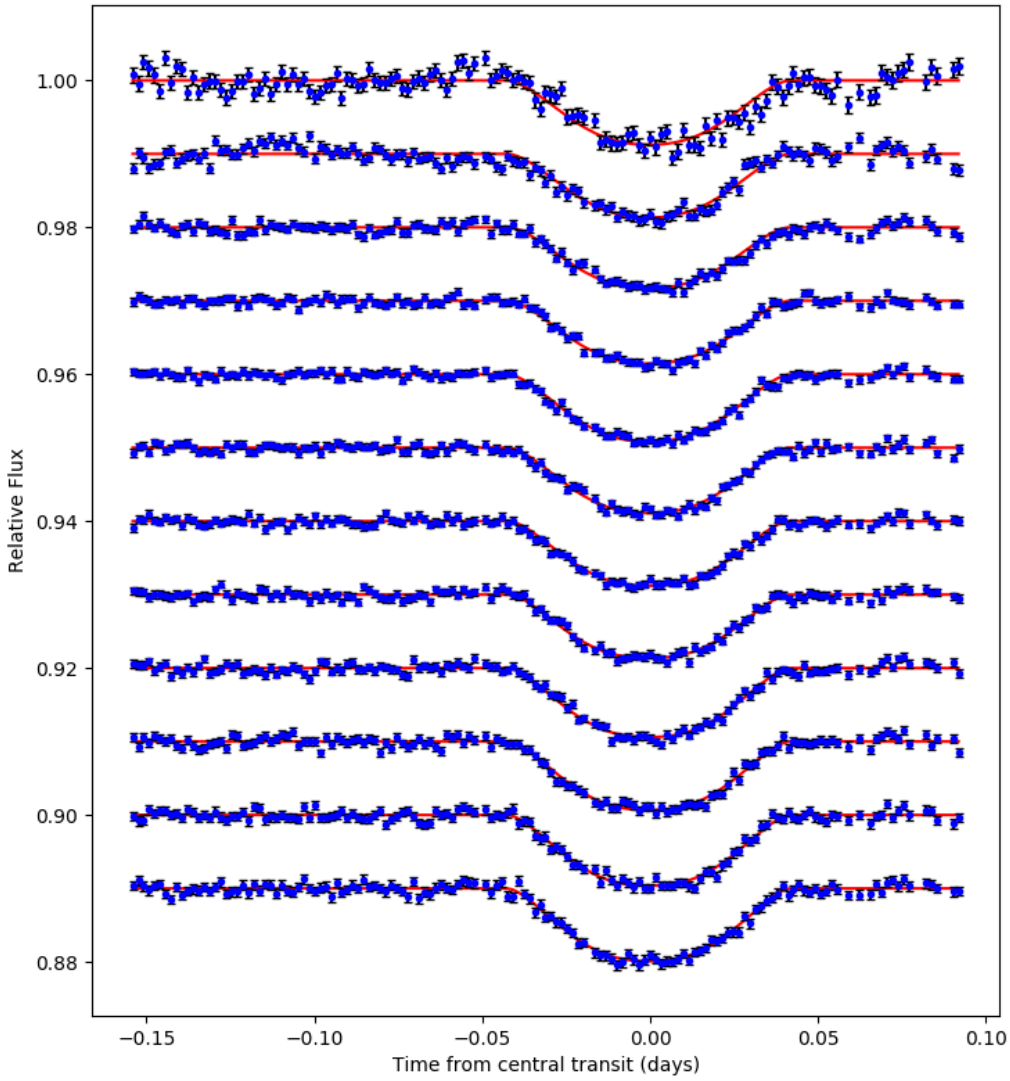

Figure 3.9: Detrended spectroscopic light curves of WASP-75b (blue data points) and median transit models (red lines). The data points are shown with their spectrophotometric uncertainties.

Table 3.3: Shared and binned parameters from the spectroscopic light curve fits of WASP-75b.

Shared parameters	
Parameter	Value
t_0 ($\mathrm{BJD_{TDB}}$)	$2457968.84742 \pm 0.00010$
a/R_*	$6.23^{+0.06}_{-0.07}$
i (degrees)	$81.74^{+0.11}_{-0.14}$
σ_x	0.00023 ± 0.00003

Binned parameters			
Wavelength Range (Å)	R_p/R_*	u_1	u_2
$3655 - 3919$	$0.10764^{+0.00199}_{-0.00178}$	0.689	0.086
$3919 - 4198$	$0.10715^{+0.00174}_{-0.00143}$	0.663	0.130
$4198 - 4490$	$0.10409^{+0.00157}_{-0.00129}$	0.629	0.162
$4490 - 4792$	$0.10529^{+0.00154}_{-0.00121}$	0.592	0.183
$4792 - 5102$	$0.10853^{+0.00159}_{-0.00124}$	0.553	0.196
$5102 - 5419$	$0.10571^{+0.00151}_{-0.00118}$	0.515	0.201
$5419 - 5741$	$0.10416^{+0.00145}_{-0.00114}$	0.481	0.202
$5741 - 6066$	$0.10220^{+0.00135}_{-0.00108}$	0.453	0.200
$6066 - 6394$	$0.10646^{+0.00144}_{-0.00111}$	0.429	0.197
$6394 - 6724$	$0.10558^{+0.00146}_{-0.00118}$	0.411	0.195
$6724 - 7055$	$0.10703^{+0.00144}_{-0.00117}$	0.396	0.194
$7055 - 7388$	$0.10831^{+0.00151}_{-0.00121}$	0.382	0.195

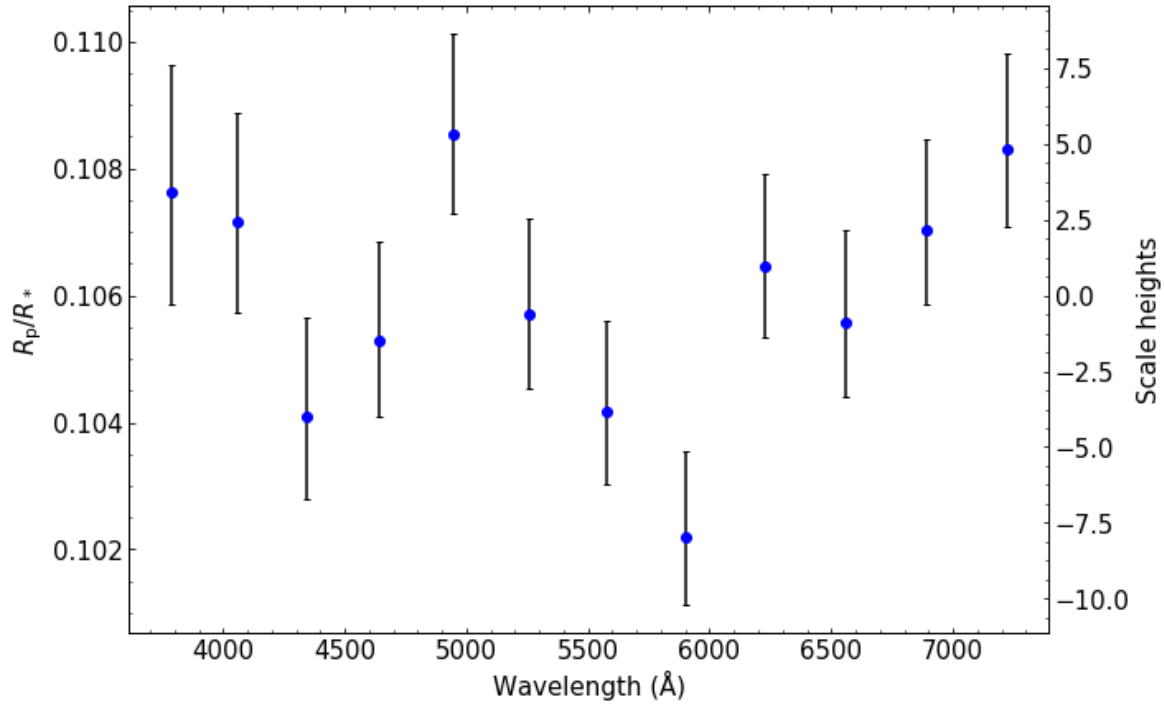

Figure 3.10: The transmission spectrum for the test case WASP-75b.

3.5 Transmission Spectrum

From the 12 fitted spectroscopic light curves, planet-to-star radius ratios were obtained and a transmission spectrum was created. Since the WASP-75 data set was mainly used for exploratory purposes in order to familiarise myself with data reduction procedures and to understand the analysis steps required up to the production of a transmission spectrum, the investigation did not concentrate on improving the results or on deciphering WASP-75b's atmosphere. This was influenced by the fact that the properties of WASP-75 and WASP-75b were less than ideal for observations with a 3.6 m telescope. Still, a variety of adjustments were implemented to examine their effect on the final outcome, including changes in the fitted parameters and changes in the noise parameters considered for the error budget. The complete transmission spectrum of WASP-75b is displayed in Figure 3.10.

The transmission spectrum appears to be slightly wavy but has no discernible features of atomic elements. Given that the errors are only a few scale heights, some information about the properties of WASP-75b's atmosphere can be inferred. For the expected sodium location, in particular, at ∼5890 Å, the spectrum gives a low value, something that could indicate a lack of sodium for WASP-75b. The ∼5900 Å data point is consistently found to show no peak in all realisations of the problem, reinforcing the initial assessment of a non-detection of Na. In addition, there is no significant slope indicative of haze in the blue. While all signs hint at a cloudy atmosphere, there are some uncertainties in the consideration of the systematic effects that make the R_p/R_* values somewhat questionable. Therefore, results from atmospheric models will not be presented here. Such analyses, however, will be provided for the more suitable targets in Chapters 4 and 5.

3.6 Conclusion

WASP-75b is a marginally inflated hot Jupiter with properties that make its atmospheric characterisation a challenging prospect. Nevertheless, telescope time was given to obtain some ground-based spectroscopic measurements using NTT's instrument EFOSC2. WASP-75b was observed for one night, during which an entire transit was recorded for the wavelengths between 3600 and 7400 Å using grating #11. A very basic data reduction and analysis pipeline was created by me in an attempt to process the raw data and decipher WASP-75b's atmosphere from the measured transmission spectrum. The analysis showed tentative evidence for the presence of clouds in WASP-75b's atmosphere, but the result is doubtful considering the various possible mechanisms, presented throughout this chapter, that could introduce additional error. This is the reason why these results will not be discussed in the concluding Chapter 6.

Although current observations of WASP-75b may not be able to reliably uncover the atmospheric composition of this exoplanet, the situation could change in the near future. Large telescopes, specifically designed to probe exoplanetary atmospheres

with cutting-edge instrumentation, are currently or will soon be in operation, offering the opportunity for higher-precision measurements that will enable the observation of WASP-75b's atmosphere with greater confidence (see Section 6.4 in Chapter 6 for more details).

4 The optical transmission spectrum of the low-density hot Saturn WASP-88b

The observations and most of the analysis presented in this chapter were published in the Monthly Notices of the Royal Astronomical Society (Spyratos et al., 2021). The VLT FORS2 analysis as well as the forward modelling with Goyal's grid and PLATON was implemented by me. The initial data reduction was performed by Nikolay K. Nikolov and the TESS analysis by John Southworth. The retrieval analysis with AURA was done by Savvas Constantinou with supervision from Nikku Madhusudhan. Other co-authors aided with the preparation of the proposal for this target and the observations. The parts that were implemented by others are included in the presentation here as they provide context for my own work.

4.1 Introduction

In this chapter, I will present analysis and results for the planet-hosting system WASP-88. The system was recently observed during transit using both space-based and ground-based instruments. The primary focus of the chapter will be on the ground-based observations as they explore the variation in the radius of the exoplanet as a function of wavelength and so probe the exoplanetary atmosphere at the limbs of the planet. Hence, the introduction will concentrate on the motivation for the spectroscopic investigation of WASP-88b and will outline the scientific objectives that this study aims to achieve with aid from previous and other parallel works. Initial information about the system will also be given based on the limited number of publications prior to this study.

WASP-88b is defined by a special set of characteristics that make it an ideal target for atmospheric exploration via transmission spectroscopy (Delrez et al., 2014). First of all, it is a transiting gas giant exoplanet orbiting very close to its host star. This

characteristic indicates that WASP-88b is easily observable producing deep transits in the observed light curves. Furthermore, the small distance from its host star implies that the exoplanet is tidally locked and experiences high levels of irradiation from its stellar companion. These features contribute to a very hot planetary atmosphere with a high equilibrium temperature. Finally yet importantly, this particular target has a very low density, with an almost Saturn-like mass and a highly inflated radius that is nearly two times larger than the radius of Jupiter. The combination of all these properties suggest that WASP-88b is likely surrounded by an extended atmosphere and is, thus, a suitable target for atmospheric characterisation.

This exoplanet was therefore selected as part of a large ground-based survey that aims to explore the diverse properties of gas giant atmospheres (Nikolov et al., 2018b). This survey is targeted at a select list of exoplanets with low mean densities and strong signals in the optical that orbit reasonably bright stars. The purpose of this survey is not only to understand the physics behind the dominant processes within the planetary atmospheres but also to create a large enough sample of atmospherically investigated exoplanets, so that the first population studies can be performed using a statistically significant sample. At the moment, only a few population studies have included exoplanetary atmospheres in their statistics and most of them examine relatively small samples (e.g. Sing et al., 2016; Iyer et al., 2016; Barstow et al., 2017; Tsiaras et al., 2018; Fisher & Heng, 2018; Pinhas et al., 2019; Mansfield et al., 2021; Changeat et al., 2022), with the largest of them examining 70 exoplanets (Edwards et al., 2022). This survey, in conjunction with previous results and other independent work, will create a fairly wide group of exoplanets observed in transmission and will immensely aid in the determination of possible correlations between atmospheric properties and formation/evolution histories, as well as provide links to bulk and orbital characteristics.

Another characteristic of this study is that it relies on transmission spectroscopy observations from the ground. In fact, it is part of one of the largest comparative studies currently conducted from the ground and is therefore a good test for the performance of ground-based facilities. Moreover, it focuses on an exoplanet that has no solar system counterparts and so it is expected that the atmospheric conditions will

be vastly different from the ones found on the solar system gas giants Jupiter and Saturn. This means there is a great opportunity to expand our knowledge of atmospheric physics beyond the cooler conditions of the outer planets in our solar system and to observe the atmospheric phenomena that occur in very hot environments. Furthermore, the spectroscopic observation of WASP-88b during transit will generate a legacy transmission spectrum for the ever-growing database of exoplanetary atmospheres and will provide highly complementary optical data for the newly launched James Webb Space Telescope (JWST, Gardner et al., 2006). This state-of-the-art telescope is sensitive to the near-infrared and will therefore greatly benefit from ground-based observations in the optical in a joint effort to obtain the full picture of the planetary atmosphere.

From the current sample of observed exoplanetary atmospheres (see also the ExoAtmospheres database[1]), only a few dozen of them have been examined at low-resolution in the full optical range. It is therefore hard to draw conclusions about trends in the transmission spectra and, hence, about tendencies in atmospheric compositions. One important observation so far is that clouds and hazes appear to be ubiquitous amongst all types of atmospheres (see e.g. Sing et al., 2016; Tsiaras et al., 2018). It will, thus, be interesting to see how prominent clouds and/or hazes are in the atmosphere of WASP-88b. Furthermore, Tsiaras et al. (2018), based on a sample of 30 hot gaseous exoplanets, found that the planetary radius plays a more significant role than the surface gravity in the potential detection (or cloudiness) of an exoplanetary atmosphere, but this needs to be confirmed with additional observations. Given its highly inflated radius, WASP-88b offers a unique chance to test this hypothesis.

In the following subsection, we will take a closer look at the properties of the WASP-88 system, so we can have a better idea of the target and its environment before we move on to the presentation of the observations, analysis and results.

[1] `http://research.iac.es/proyecto/exoatmospheres/table.php`

4.1.1 The WASP-88 system in detail

The WASP-88 system, with a right ascension of 20:38:02.70 and a declination of −48:27:43.2, is primarily observable from telescopes in the southern hemisphere. The exoplanet was discovered by Delrez et al. (2014) who detected the transit using data from the SAAO (South African Astronomical Observatory) facility of the SuperWASP survey (Pollacco et al., 2006) and performed follow-up photometry using the 1.2 m Euler-Swiss telescope (Lendl et al., 2012) and the 0.6 m TRAPPIST robotic telescope (TRAnsiting Planets and PlanetesImals Small Telescope Jehin et al., 2011)[2]. They also performed spectroscopic observations using the CORALIE spectrograph, also mounted on the Euler-Swiss telescope, to obtain radial velocities and estimate the stellar properties. The stellar parameters were later revised by the SWEET-Cat project[3] (Andreasen et al., 2017; Sousa et al., 2021). The discovery report revealed a brightness of $V = 11.4$ and categorised the host star as a main-sequence dwarf of spectral type F6. The star's core is likely at a medium or an advanced stage of the hydrogen burning process given that the age of the system was determined to be between 1.8 and 5.3 Gyr. Meanwhile, the transiting exoplanet is a hot Saturn with a very low density and a large radius that promise an extended atmosphere suitable for atmospheric observation via transmission spectroscopy. In fact, after taking a mean molecular weight of 2.3 a.m.u. into account, the pressure scale height was calculated to be approximately equal to 1300 km. This size corresponds to an expected variation in the transit depth with respect to one pressure scale height of $\Delta\delta = 2HR_{\rm p}/R_*^2 \approx 170$ ppm (Winn, 2010), which means that the atmospheric signal should be detectable from ground-based spectrographs. The basic properties of the host star and the orbiting exoplanet are summarised in Table 4.1.

One thing to consider when evaluating the shape of the transit light curves and the resulting transmission spectrum is stellar activity. In the case of WASP-88, brightness variations on the stellar surface that are indicative of dark or bright spots formed

[2]The WASP-South telescope used for the transit detections is located in Sutherland, South Africa, while both follow-up observation sites are located in La Silla, Chile.

[3]`http://sweetcat.iastro.pt/`

Table 4.1: The WASP-88 system properties as determined by others.

Parameter	Value	Reference
Distance (pc)	548.8 ± 6.4	Bailer-Jones et al. (2021)
Impact parameter $b = a \cos i/\mathrm{R}_*$	0.23 ± 0.15	Delrez et al. (2014)
Time of mid-transit t_0 ($\mathrm{HJD_{TDB}}$)	$2456474.73154^{+0.00085}_{-0.00086}$	Delrez et al. (2014)
Orbital period P (d)	4.954000 ± 0.000019	Delrez et al. (2014)
Effective temperature $T_{\rm eff}$ (K)	6428 ± 59	Sousa et al. (2021)
Metallicity [Fe/H]	0.06 ± 0.04	Sousa et al. (2021)
Stellar density ρ_* ($\rho_\odot$)	$0.16^{+0.02}_{-0.03}$	Delrez et al. (2014)
Stellar surface gravity log g_* (c.g.s.)	4.025 ± 0.020	Sousa et al. (2021)
Stellar mass M_* ($\mathrm{M}_\odot$)	1.428 ± 0.014	Sousa et al. (2021)
Stellar radius R_* ($\mathrm{R}_\odot$)	1.972 ± 0.065	Sousa et al. (2021)
Planet/star radius ratio $k = R_{\rm p}/R_*$	0.084 ± 0.002	Delrez et al. (2014)
Scaled semi-major axis a/R_*	$6.64^{+0.17}_{-0.34}$	Delrez et al. (2014)
Semi-major axis a (AU)	$0.06431^{+0.0006}_{-0.00062}$	Delrez et al. (2014)
Inclination i (degrees)	$88.0^{+1.4}_{-1.5}$	Delrez et al. (2014)
Eccentricity e	0 (fixed)	Delrez et al. (2014)
Planet mean density $\rho_{\rm p}$ ($\rho_{\rm Jup}$)	0.11 ± 0.03	Delrez et al. (2014)
Planet surface gravity log $g_{\rm p}$ (c.g.s.)	$2.67^{+0.07}_{-0.08}$	Delrez et al. (2014)
Planet mass $M_{\rm p}$ ($\mathrm{M_{Jup}}$)	0.56 ± 0.08	Delrez et al. (2014)
Planet radius $R_{\rm p}$ ($\mathrm{R_{Jup}}$)	$1.70^{+0.13}_{-0.07}$	Delrez et al. (2014)
Planet equilibrium temperature $T_{\rm eq}$ (K)	1772^{+54}_{-45}	Delrez et al. (2014)
Irradiation (erg s^{-1} cm^{-2})	$2.2^{+0.5}_{-0.3}$ 10^9	Delrez et al. (2014)

by strong magnetic fields are highly unlikely to occur. One reason is the very high temperature of the target star (∼6450 K), which does not favour the formation of starspots and/or faculae. Another reason is that the long-term light curves from the WASP photometry show no signs of rotational modulation down to at least 1 mmag (Delrez et al., 2014). Furthermore, a spectrum of the calcium H and K lines that was obtained for an independent program, which currently analyses stellar activity trends for a large sample of exoplanets, revealed a deep line core and, thus, no signatures of chromospheric emission (Southworth, unpublished). The deep line core is consistent with the photometry for this type of star and suggests a lack of stellar activity for WASP-88.

A nearby star used as a reference for relative spectrophotometry was also checked for variability in the long-term light curves acquired from the WASP data. The investigation found a very weak and statistically insignificant periodic signal at the 9 d period. The strength of the signal is negligible considering that the transmission spectrum relies on the much smaller signal from the wavelength-dependent component. Moreover, only the disc-integrated light of the reference star is used and so any brightness variations from e.g. starspots would be unimportant compared to those of WASP-88. Given that the systematics correction also removes traces of variability, the influence of the reference star light curves on the transmission spectrum should be inconsequential.

Apart from that comparison star, WASP-88 also has a much closer companion at an angular separation of only $3.350 \pm 0.015''$ (Bohn et al., 2020). The distance is so small that the star will be enclosed by the point spread function of the target in our observations and so the star will be unresolved. Contamination from nearby stars can also affect the shape of a transmission spectrum, as evidenced in the emblematic case of WASP-103b (Southworth & Evans, 2016; Lendl et al., 2017; Wilson et al., 2020). High-contrast images obtained using the SPHERE instrument on the Very Large Telescope (VLT) showed that the star is fainter than WASP-88 in the K-band by 7.60 ± 0.53 mag (Bohn et al., 2020). This is a significant margin and suggests that the impact of this companion candidate on the transmission spectrum of WASP-88b will be negligible (Southworth et al., 2020).

At the time of writing, there have been no other investigations of WASP-88b's atmosphere. This study presents the first ever transmission spectroscopy results for this transiting and low-gravity exoplanet and probes WASP-88b's atmosphere for the first time, based on observations made on the Earth's surface.

4.2 Data acquisition and reduction

WASP-88b was observed from both the ground and space, but transmission spectroscopy was only performed on the ground-based data sets. This is mainly because the space-based data were obtained through broadband photometry, while the ground-based data sets were taken using multi-object spectroscopy. This key difference between the two methods means that the space-based data cover a broad wavelength range that spans from the optical (orange) to the near-infrared (up to 1 micron), whereas the observations from the ground could easily be sliced into narrow spectrophotometric channels so that any transit parameter estimates can be analysed with respect to wavelength.

Nonetheless, the space-based observations were quite important as they presented a fairly continuous light curve (i.e. with almost no gaps) and they included more transits. This made them ideal for a system characterisation and so they could provide a much needed update to the WASP-88 and WASP-88b values reported in Delrez et al. (2014). Furthermore, the additional transits promised a revised ephemeris and, hence, a refined orbital period for the exoplanet. In addition, the relatively longer duration of the out-of-transit data meant that these light curves were excellent for a short-term photometric investigation. Such an analysis could potentially show variability in the flux of the star and even periodicity, something that could suggest the presence of another celestial companion. Checking for variability in the stellar flux is vital, as it could reveal various phenomena that occur in the stellar atmosphere and indicate whether the star has strong magnetic fields that drive it to an increased atmospheric activity.

4.2.1 TESS photometry

The Transiting Exoplanet Survey Satellite (TESS, Ricker et al., 2015) is a space telescope launched in 2018 by the National Aeronautics Space Administration (NASA) with the primary mission to discover and observe exoplanets that are transiting nearby bright stars. TESS has a 10.5 cm entrance pupil diameter and is a near all-sky survey, currently[4] in its fifth cycle exploring sector 60. The light collection instrument consists of four CCD cameras that produce a combined image with a wide field of view totalling to $24° \times 96°$ (this is also known as one sector). The four CCD chips in each camera are separated by 2 mm and effectively create an image of the sky that is 4096×4096 pixels in size. The bandpass is quite broad (600 to 1000 nm) and (roughly) corresponds to the Cousins I-band. This bandpass is sensitive to smaller exoplanets orbiting M dwarfs as it examines the red and near-infrared region of the spectrum.

The WASP-88 system contains a gas giant planet transiting a bright F-type star, and the system parameters can be inferred from the TESS data with high precision. The system was observed in July 2020 as part of sector 27 and the data were collected in short cadence. A total of five transits were observed, four complete and one partial, in a period of 24.4 days. Data collection was paused in the middle of the observation to transmit the massive amount of data back to Earth. This did not affect any of the transits. The partial transit obtained during the end of the observation was rejected from subsequent analysis as it was of little use in the deduction of the transit parameters. The full light curve includes 16 511 data points altogether that were sampled at a rate of one frame per 120.1 s.

When transmitted back to Earth, the data are saved on the Mikulski Archive for Space Telescopes (MAST[5]) at the Space Telescope Science Institute (STScI). An initial data reduction procedure is then performed by an automated science pipeline (Jenkins et al., 2016). This pipeline invokes several modules, including a calibration module (for e.g. bias, dark and flat field corrections, cosmic-ray removal, pixel-to-pixel sensitivity

[4] at thesis original submission date

[5] `https://mast.stsci.edu/portal/Mashup/Clients/Mast/Portal.html`

corrections), an aperture determination module, a sky background elimination module, and a Pre-search Data Conditioning (PDC) module that corrects for some additional systematics.

The reduced WASP-88b data were accessed through this portal and the PDC Simple Aperture Photometry (PDCSAP) fluxes were extracted by demanding that the quality flag is set to zero to minimise contamination from stray light. These fluxes were produced by the TESS pipeline after the removal of common instrumental systematic trends from the time-series. For the transit analysis, most of the data points were removed, leaving only data that sufficiently explore the entire transits. The cutoff was placed at 1.5 transit durations from either side of the transit midpoints. The remaining 1998 data points were normalised to unit flux by simply fitting a straight line to the out-of-transit data on either side of each transit and then by dividing the observed flux by the straight-line fit. Finally, the out-of-transit light curve was used to determine whether any additional periodic signals could be detected and to find out whether any features suggestive of stellar activity are visible.

The out-of-transit aperture photometry light curve investigation showed no significant variations in the time series. More specifically, a computed periodogram only displayed a maximum amplitude of 0.3 mmag for a peak signal at 0.172 cycles per day. The signal-to-noise ratio was $\sim$3.6 suggesting a negligible detection. This is below the commonly-used signal-to-noise ratio limit of 4.0 (Breger et al., 1993). The observed frequency is not consistent with the expected rotation period that is calculated from its radius and spectroscopic rotational velocity ($v \sin i$), which gives a signal at a much lower frequency. Nevertheless, the estimated frequency is still insignificant and the absence of any periodicities can be safely assumed from the TESS data. This result solidifies our conclusion of a predominantly quiet stellar atmosphere.

4.2.2 VLT FORS2 spectroscopy

The ground-based spectroscopic observations were performed using the Very Large Telescope (VLT), which is located at the European Southern Observatory (ESO) in

Cerro Paranal, Chile. Although it is often simply referred to as the VLT, it actually consists of four large, 8.2 m diameter Unit Telescopes and four additional, movable 1.8 m diameter Auxiliary Telescopes. Altogether the VLT forms one of the most advanced arrays of telescopes with a capability to observe from the ultra-violet to the mid-infrared wavelength region. For the purpose of this work, only Unit Telescope 1 (UT1, or Antu) was utilised to spectroscopically examine the WASP-88 system.

To observe the hot Jupiter system, the FOcal Reducer and low dispersion Spectrograph 2 (FORS2, Appenzeller et al., 1998; Boffin et al., 2016) was employed. This instrument is mounted at the Cassegrain focus of UT1 and is widely used for low-resolution transmission spectroscopy in the optical. Several exoplanetary atmospheres have already been observed and characterised with FORS2 (Bean et al., 2010, 2011; Sedaghati et al., 2015, 2016, 2017a,b; Lendl et al., 2016; Nikolov et al., 2016, 2018b, 2021; Gibson et al., 2017; Carter et al., 2020; Wilson et al., 2020), revealing a variety of atomic and molecular species, scattering slopes and grey opacities, as well as some stellar activity effects. This was the instrument of choice for the large ground-based optical transmission spectroscopy survey that includes WASP-88b.

FORS2 has four different observing modes, with multi-object spectroscopy (MXU mode) being selected as the mode of choice because it offers great flexibility in the number, size and shape of the slits (>400 combinations) through custom built masks. For comparison, long-slit spectroscopy only provides 9 long-slits with fixed widths. Depending on the applied grisms and filters, FORS2 can cover a wavelength range from 330 nm to 1100 nm, which corresponds to the ultraviolet, optical and near-infrared parts of the spectrum. The FORS2 field of view in standard resolution is $6.8' \times 6.8'$ giving a pixel scale of $0.25''$/pixel. From the two detectors available, only the one sensitive to the red wavelengths (MIT) was used. This detector is comprised of two CCD chips that have a size of 2048×4096 pixels each.

WASP-88b was observed in two separate nights as part of program 199.C-0467, which was led by principal investigator Nikolay K. Nikolov. The data were collected in 2017 during the nights of August 19th and 24th, and resulted in an output of two complete WASP-88b transits. The target and a suitably bright comparison star were

simultaneously examined in MXU mode, with each spectrum recorded on a different CCD plate. The comparison star is positioned at an angular separation of 2.7′ and is known as 2MASS J20381555−4829215. A custom-made mask with broad slits of size 22″×132″ was created to diminish differential slit light losses caused by variations in seeing. In addition, a binning of 2×2 was selected to increase the amount of data by reducing the readout time of the duty cycle to ∼30 s. Exposure times were variable throughout the observations due to fluctuations in flux, but this was not an issue for the FORS2 CCD chips as they are linear up to saturation.

During the first night of the campaign, the dispersive element GRIS600B (hereafter blue grism or 600B) was utilised to successfully obtain the first entire transit. This blue grism examined the wavelength region between 3300 and 6200 Å. The seeing was above 1″ (0.82″-1.8″) throughout most of the night, owing to some low level turbulence in the Earth's atmosphere. Despite that, the sky remained clear during the whole observation. The target was tracked from an airmass of 1.38 to an airmass of 1.09 and then to an airmass of 2.21. The final data set (hereafter blue) consists of 206 valid exposures that have integration times of 30, 100 or 120 s. The full observation lasted ∼8 hours and 12 min.

A second complete transit was recorded using a different dispersive element. This time GRIS600RI (hereafter red grism or 600RI) was used in conjunction with blocking filter GG435 to effectively cover the spectral range from 5200 to 8400 Å. The use of the blocking filter was necessary to isolate the first spectral order. Observing conditions were slightly different this night, with seeing staying below 1″ most of the time (0.4″-1.94″). Furthermore, the sky was mostly clear during the observation, and the weather improved in the second part resulting in photometric conditions for the early morning of August 25th. As the target ascended and then descended in the night sky, the airmass also changed, starting from 1.39, reaching a minimum of 1.09, and then ending at 2.22. This data set (hereafter red) contains 494 valid data points with exposure times between 25 and 80 s. Both WASP-88 and the comparison star were followed for

a period of ∼8 hours and 19 min and the two data sets were saved on the ESO archive[6]. The observations are now accessible by anyone as the proprietary period is long over.

A key step before analysis was to reduce the data by identifying and isolating the spectral signal of the target and the reference star from the FITS files produced by the FORS2 software. For this, a well-tested, custom built pipeline, written in IDL, was employed. The pipeline has been extensively used in other FORS2 works (e.g. Nikolov et al., 2018b; Carter et al., 2020) and has been shown to be capable of generating high-precision results that are consistent with space-based observations (e.g. Nikolov et al., 2016; Gibson et al., 2017). A bias frame subtraction and a flat field correction to the raw images were also considered, but were not included in the analysis as they offered no meaningful improvements to the obtained data.

The extraction of the aperture of each spectrum was performed separately by making use of the APALL function, which is part of the IRAF package[7]. This function identifies the dispersed light from the stars and carries out a simple, column by column, box summation. An aperture width of 21 pixels was selected for both spectral traces as it minimised the out-of-transit scatter. In addition, background noise regions were also specified on both sides of each spectrum at a location 30 to 80 pixels away from the maximum spectral count value. The computed median count values of the regions were then computed and subtracted from the spectral signal. This operation was useful as it also eliminated any residual bias. An example spectrum of WASP-88 and the reference star can be seen in Figure 4.1. For simplicity reasons, the spectra in the figure are normalised with respect to maximum counts.

After acquiring the spectra, it was important to calibrate them as a function of wavelength. This action was performed with the help of an emission lamp at the end of each observation. Again, a customised mask was devised in MXU mode, but the slit width was much shorter now ($1''$) compared to the one used for the stellar spectra. A wavelength solution was then established for each spectral aperture in two steps.

[6] `http://archive.eso.org/eso/eso_archive_main.html`
[7] `https://iraf-community.github.io/`

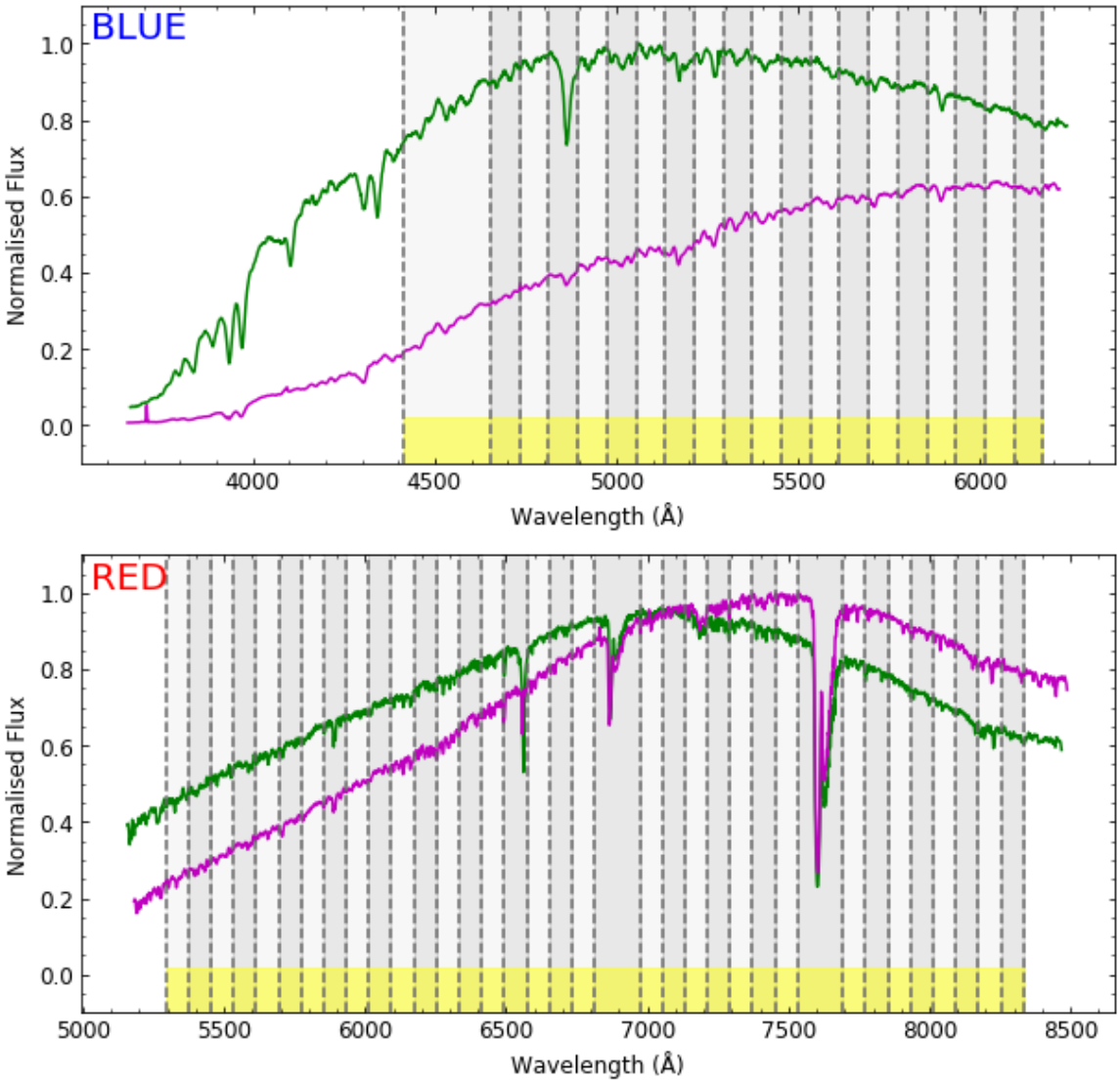

Figure 4.1: Normalised example spectra of WASP-88 (green line) and a selected reference star (magenta line). The spectra from the FORS2 grism 600B are shown on the left, whereas the spectra from grism 600RI are displayed on the right. The range of the spectroscopic bins is illustrated by grey bands, while the range of the broad white-light channels is indicated in yellow.

First, the most conspicuous emission lines were determined by obtaining the peak position from Gaussian fits in the cross-dispersion axis and then low-order Chebyshev polynomials were applied to these peaks in the dispersion axis. Any minor sub-pixel drifts in the spectral traces were also accounted for by mapping the extracted spectra against a Doppler-corrected rest frame.

The pipeline generated a great deal of information, but, most importantly, it produced spectrophotometric light curves for both the combined (white) light and the binned (spectroscopic) light by summing the flux for each pre-determined bandpass. The end product contained separate light curves for each star and their relative fluxes, together with the appropriate spectrophotometric errors. These errors were estimated under the assumption that the data can be fully described by pure photon noise. Furthermore, the pipeline included parameter arrays used for the wavelength calibration, as well as some optical state parameters (or decorrelation parameters). These auxiliary variables consisted of time, airmass, positional changes of the spectra in the x and y axes, variations in the sky background noise, the rotator angle and its rate of change, the full width at half maximum, ambient pressure and ambient temperature. Alternatively, all these parameters could also be readily obtained from the headers of the FITS files.

One final but useful addition to this processed data package were the limb darkening coefficients. Two groups of limb darkening coefficients were computed, one that takes into account the quadratic limb darkening law and another that considers the non-linear four-parameter law. These coefficients represent theoretical values calculated by making use of the Stagger-grid (Magic et al., 2015). This is a grid of 3D model stellar atmospheres that can be used to construct synthetic spectra of relatively cool stars such as WASP-88. With a large range of stellar effective temperatures, stellar surface gravities, and stellar metal abundances, one can compute the limb darkening coefficients at any given passband. By taking advantage of this flexibility, limb darkening coefficients were obtained not only for the white-light time series, but also for the narrow spectroscopic channels.

4.3 Analysis

4.3.1 TESS

The discovery paper of WASP-88b (Delrez et al., 2014) was published in 2014, but since then no other transit observations of the system made their way to the literature. In addition, the system analysis in that paper was based on five transit light curves, with three of them showing incomplete transits. The new data from TESS offered the opportunity to revise the stellar and planetary parameters and to improve the orbital ephemeris. To model the TESS light curves, the FORTRAN-based code JKTEBOP (Southworth, 2013) was employed and a similar procedure to the one described in the *Homogeneous Studies* project (Southworth, 2012, and references therein) was followed. To ease calculations, the data were sorted according to orbital phase and were then binned in successive groups of five data points. This little trick made the data more compact by generating a total of 400 data points with an effective time resolution of 150 s. The unbinned data were then only used to estimate the orbital ephemeris.

The free parameters included in the JKTEBOP fit were the sum of the fractional radii ($r_{\rm p} + r_*$, with $r_{\rm p} = \frac{R_{\rm p}}{a}$ and $r_* = \frac{R_*}{a}$, and where $R_{\rm p}$ is the radius of the planet, R_* is the stellar radius, and a is the semi-major axis of the planetary orbit.), the ratio of the radii ($k = \frac{r_{\rm p}}{r_*}$), the inclination ($i$), the out-of-transit flux level and the phase of mid-transit. Eccentricity was fixed to zero, under the assumption of a circular orbit, in agreement with the findings of Delrez et al. (2014). Limb darkening was explored via four different laws: quadratic, square-root, logarithmic and cubic (Southworth, 2008), and two fits were implemented for each one of them. More specifically, there was one version where both coefficients were fixed and another where one coefficient was free and the other fixed. The initial values for the limb darkening coefficients were set to the theoretical values obtained from Claret (2017).

All fitted parameters were determined using the Monte Carlo technique together with residual-permutation formulae (Southworth, 2008). The uncertainties produced by this method have been proven to be comparable with results from other forms of

MCMC analyses (Maxted et al., 2020). Any differences between the products of the four limb darkening laws used here were also taken into account by adding the limb darkening errors in quadrature to the larger uncertainties from the Monte Carlo and the residual-permutation analysis. The best-fit transit model along with its residuals is displayed in Figure 4.2, and is plotted for both the binned and the unbinned data points from TESS.

The physical quantities of the transiting system were finally estimated using the orbital period, $r_{\rm p}$, r_* and i from the JKTEBOP investigation, the stellar velocity amplitude ($K_* = 53.4^{+6.8}_{-6.6}\,{\rm m\,s^{-1}}$) from Delrez et al. (2014), and the stellar effective temperature ($T_{\rm eff}$) and metallicity [Fe/H] from Andreasen et al. (2017). Given these initial values, the planetary velocity amplitude ($K_{\rm p}$) was determined as the one that provided the best match between the calculated $T_{\rm eff}$ and R_*, and five different sets of theoretical predictions from stellar evolutionary models (see e.g. Southworth, 2010). To obtain errors for all physical properties, a perturbation analysis was performed that propagated the uncertainties from the initial parameters to the final physical outcomes.

Apart from random errors, systematic contributions were also considered. These were estimated from the standard deviation of each output quantity over the theoretical expectations evaluated for the same assortment of stellar simulations mentioned in the previous paragraph. A significant systematic uncertainty was found for the stellar mass as a result of complications in associating the measured $T_{\rm eff}$ with the stellar density calculated from the TESS time series. The final physical parameters adequately reflect the values given by Delrez et al. (2014) with a somewhat lower stellar radius and, therefore, a lower planetary radius. A full list of the light curve parameters and the physical properties for the WASP-88 system is shown in Table 4.2.

Contamination from the faint nearby star (Bohn et al., 2020, see also Section 4.1) was also investigated but was deemed to be negligible, despite the fact its point spread function (PSF) was encompassed by the PSF of WASP-88. Southworth et al. (2020) demonstrated that fainter stars with a magnitude difference greater than 3 mag had an insignificant impact on the flux of the bright star for analyses such as this one. When implementing the calculations presented in Southworth et al. (2020), it became

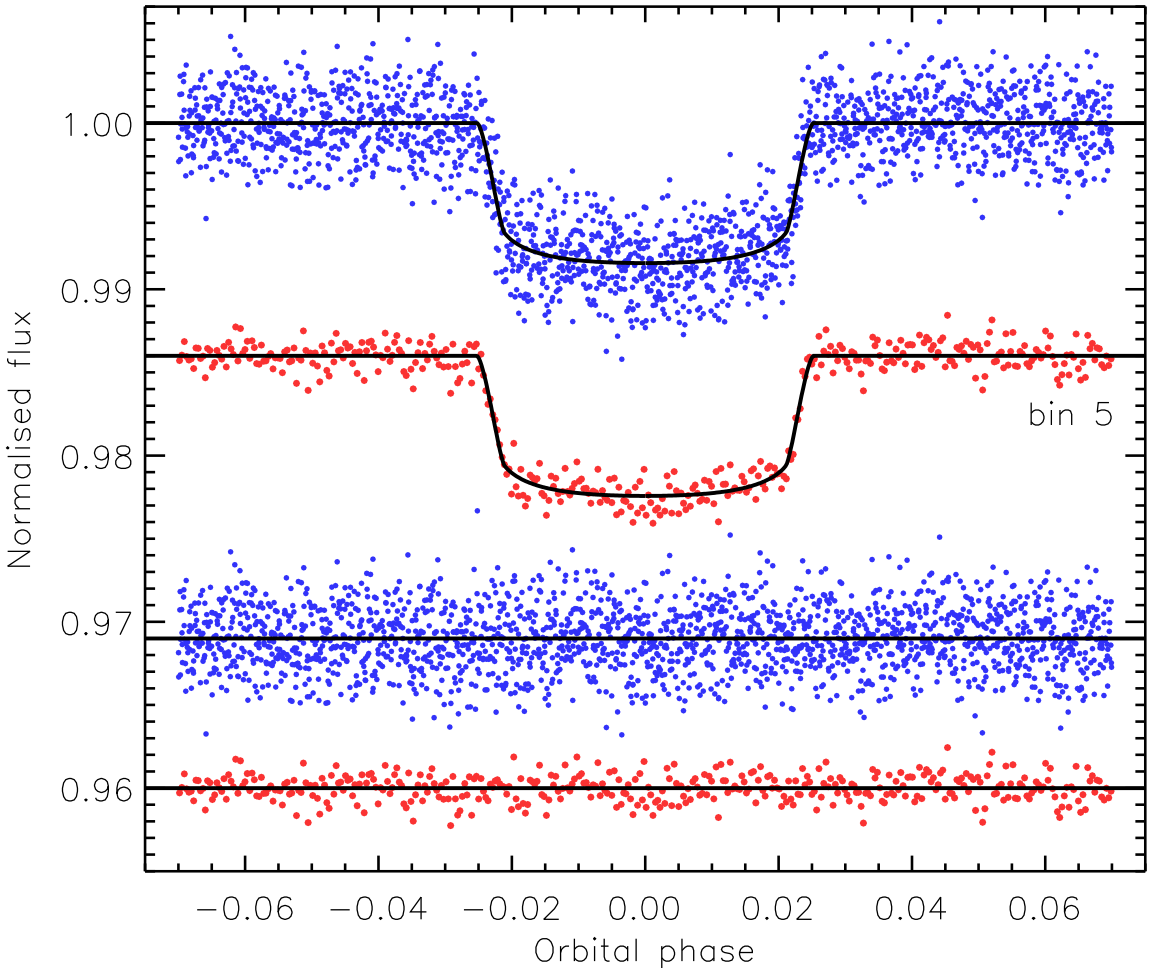

Figure 4.2: TESS transit light curve of WASP-88b, with blue dots indicating the unbinned data used to compute the orbital ephemeris and red dots indicating the phase-binned data used to establish the physical properties of the system. The best-fit transit model from JKTEBOP is illustrated in black solid lines. The fit residuals are also presented below the transit light curve for clarity. All data are plotted with respect to orbital phase.

Table 4.2: The WASP-88 system properties as determined after examination of the TESS light curve.

Parameter	Value
Light curve parameters	
Orbital period (d)	4.9540045 ± 0.0000020
Time of transit (BJD/TDB)	$2456474.73154 \pm 0.00087$
$r_{\rm p} + r_*$	$0.1577^{+0.0099}_{-0.0025}$
$r_{\rm p}$	$0.01261^{+0.00091}_{-0.00024}$
r_*	$0.1451^{+0.0088}_{-0.0024}$
k	$0.0869^{+0.0010}_{-0.0012}$
i (degrees)	$89.7^{+0.3}_{-2.6}$
Physical properties	
$T_{\rm eff}$ (K)	6450 ± 61
[Fe/H]	0.03 ± 0.04
Stellar mass ($M_\odot$)	$1.288^{+0.029\ +0.020}_{-0.020\ -0.025}$
Stellar radius ($R_\odot$)	$1.93^{+0.12\ +0.01}_{-0.03\ -0.01}$
Stellar log g (c.g.s.)	$3.976^{+0.015\ +0.002}_{-0.051\ -0.003}$
Stellar density ($\rho_\odot$)	$0.179^{+0.009}_{-0.029}$
System age (Gyr)	$1.7^{+0.2\ +0.2}_{-0.4\ -0.4}$
Planet mass ($M_{\rm Jup}$)	$0.520^{+0.067\ +0.005}_{-0.066\ -0.007}$
Planet radius ($R_{\rm Jup}$)	$1.63^{+0.12\ +0.01}_{-0.03\ -0.01}$
Planet surface gravity (m s^{-2})	$4.84^{+0.64}_{-0.88}$
Planet mean density ($\rho_{\rm Jup}$)	$0.112^{+0.016}_{-0.026}$
Planet equilibrium temperature (K)	1737^{+54}_{-15}
Semimajor axis (au)	$0.06189^{+0.00048\ +0.00031}_{-0.00032\ -0.00040}$

apparent that the fractional contribution of this stellar object was only $(9.1 \pm 4.1) \times 10^{-6}$. This result is well below the error produced by the TESS photometry.

4.3.2 VLT FORS2

In accordance with most spectroscopic analyses of this kind, WASP-88b was investigated using both the full (white) spectrum and a set of narrow channels that were binned as a function of wavelength. This was done for both the blue and red data sets.

In particular, one white and 20 spectroscopic light curves were constructed for the blue, and one white and 36 spectroscopic light curves were produced for the red data set. The two time series cover individual wavelength ranges of 4413 to 6173 Å (blue), and 5293 to 8333 Å (red), for a combined coverage between 4413 and 8333 Å. Most spectroscopic channels had bin sizes of 80 Å, with a few having double or triple that size to improve the signal. This bin distribution closely follows the passbands used for other FORS2 targets (e.g. Nikolov et al., 2018b; Carter et al., 2020). Data below 4413 Å were not included in the final cut due to low signal to noise ratio. Relative flux light curves were then created by dividing the target light by the light of the reference star. This was necessary to minimise atmospheric effects, such as changes in the extinction levels due to variations in airmass or corruption from telluric lines.

Two main techniques were considered to disentangle the noise from the transit light curves and to analyse the system. The first one involved the use of various baseline polynomial functions and the second one was based on Gaussian Processes. Both techniques have been introduced in Chapter 2 and are used to decorrelate systematic errors that naturally occur in the course of any type of observation. For ground-based data, in particular, the most common noise contributors are the Earth's atmosphere and instrumental effects. A third technique that combines the two methods was also taken into account. Results from all these methods were compared together in the end to draw useful conclusions about the strengths and weaknesses of each one of them.

In all cases, the investigation was split in two parts. First, the white spectrophotometric light curves were analysed to gather the initial knowledge for the planetary parameters and then a wavelength-dependent examination was performed to produce the transmission spectrum. In the first case, the initial fits consisted of polynomials of airmass but this was later expanded to also take various other physical parameters into consideration. Noise was therefore fully evaluated under linear regression and the planetary quantities were estimated following a deterministic approach.

In the second case, both the white-light time series and the spectroscopic time series were assessed under a Gaussian Process framework by incorporating noise in the kernel function of the non-parametric solution. Again, time was initially examined as

a regressor but the impact of other physical parameters was explored too. To aid the fit towards the optimised solution, parametric baseline functions were also studied and were included as part of the mean function that typically consists of the analytically computed transit model. Systematics were therefore mostly described through non-linear regression methods and the best-fit transit model was calculated by adhering to a stochastic approach. Linear regression was only involved during common-mode correction and during the inclusion of the polynomials.

In the last case, a middle ground solution was probed to satisfactorily cover all worthwhile possibilities towards an outcome that best explains the properties of WASP-88b. This solution comprised of a white light curve analysis based on Gaussian Processes and a spectroscopic analysis based on linear baseline functions. This alternative, part-stochastic-part-deterministic, approach was important in determining whether a combined approach could additionally provide results that make sense.

4.3.2.1 Baseline polynomial analysis

White light curves In many ways, this analysis reflects the rudimentary investigation performed for the case study WASP-75b. To simulate the transit light curves, the same modelling package (`batman`, Kreidberg, 2015) was utilised, and the same two-parameter, quadratic limb darkening law was applied (Kopal, 1950). These two conditions together resulted in very fast model evaluations each time the transit function was called. Moreover, parametric polynomial functions were employed to filter out any interference, something that is similar to the noise decorrelation method implemented for WASP-75b using a polynomial of time. A key difference with the WASP-75 system analysis is that, in this case, the optimised solution was computed by carrying out non-linear least squares fits instead of performing a much more time consuming Markov Chain Monte Carlo sampling process.

The most important change from that analysis, however, was that the transit model and the systematics contributions were fitted simultaneously in one step in lieu of the two-step fitting strategy followed for WASP-75b. This simplified the overall pro-

cedure and reduced computational time. The simultaneous fit of both the transit and the decorrelation parameters was achieved by constructing a two-component function of the form:

$$P(w) \times T(t, \theta) = \prod_{\nu=1}^{N} \left(\sum_{\mu=0}^{M} a_\mu w_\nu^\mu \right) \times T(t, \theta) \tag{4.1}$$

where P represents a single polynomial or a product of polynomials depending on the number ($\nu = 1, 2... \ N$) of auxiliary variables w considered, a_μ are the coefficients for each polynomial of a certain degree ($\mu = 1, 2... \ M$) used, and T is the transit model comprised of the transit parameters θ and time t.

This composite model was devised to include a range of temporal and physical systematic parameters, such as time, airmass, sky background, FWHM, horizontal and vertical spectrum profile drifts, the angle of the rotator and its rate of change, precipitable water vapour, and ambient atmospheric pressure and temperature. Figure 4.3 illustrates the temporal trends in some of the most important optical state parameters. Every single variable used was inspected either on its own or in conjunction with other variables. To cover all possible combinations of external variables, a combinations algorithm was implemented that took into account both the variables and their respective polynomial degrees. For this, it was necessary to apply a condition as the number of unique combinations was practically infinite. Ultimately, only polynomials without cross terms and up to second degree were considered, resulting in a total of 59,048 unique combinations for each relative light curve.

The very large number of models meant that a typical MCMC analysis would likely last for weeks, if not months or years. Instead, a non-linear least squares algorithm was selected to accelerate model optimisation. This implies that the transit quantities and polynomial coefficients were recovered concurrently by following the Levenberg-Marquardt least squares formalism. This method works by performing a series of non-linear least squares fits to minimise a vector of residuals specified by a given objective function. These residuals were estimated from the difference between the observed data and the model, and had weights that were defined by the spectrophotometric uncertainties of the data. This minimisation technique led to an extremely

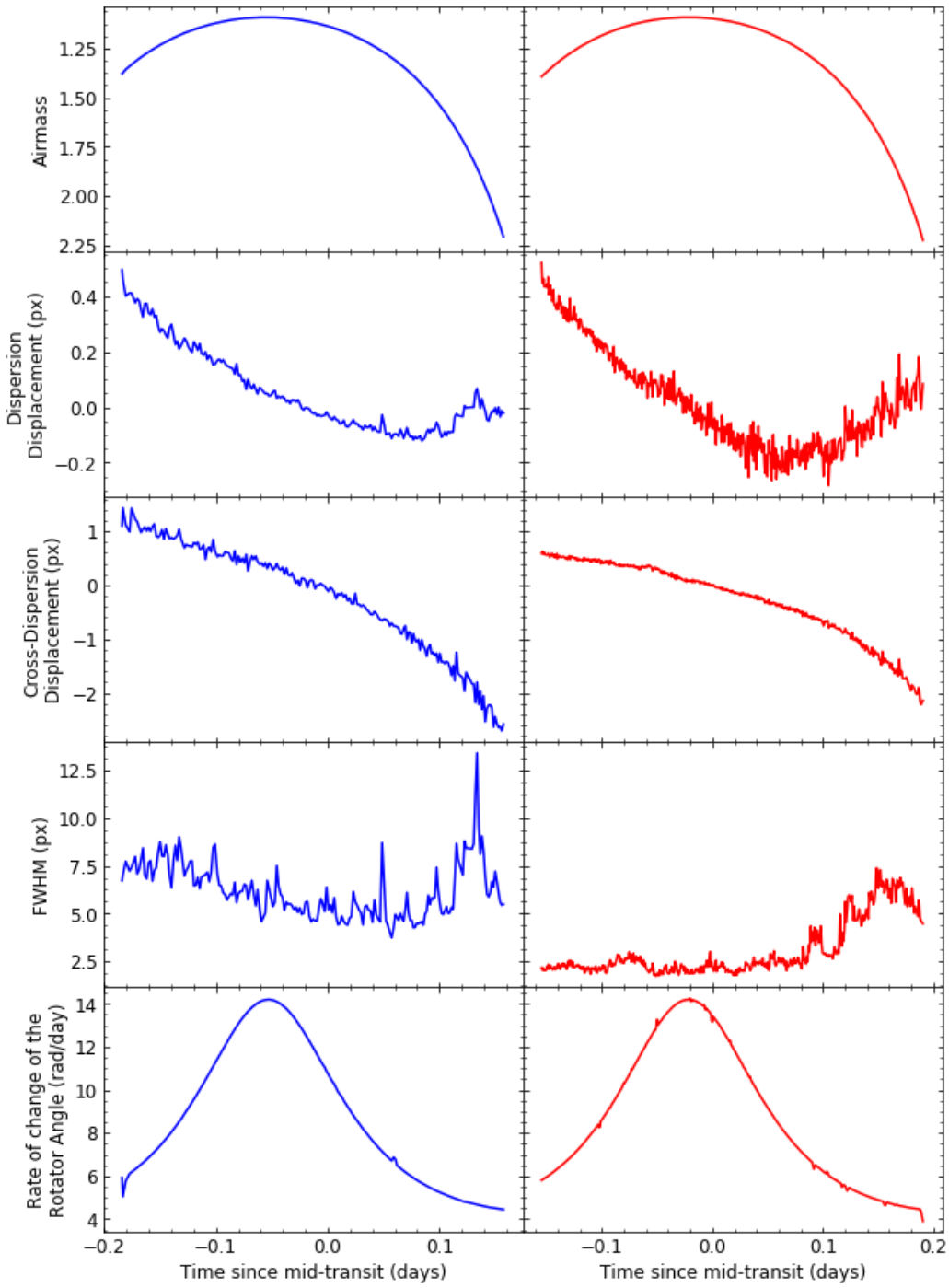

Figure 4.3: Detrending variables as a function of time for the blue (left) and red (right) data sets. From top to bottom, the figure includes airmass, displacements in the x and y axis, FWHM, and the angular velocity of the rotator.

fast convergence towards the best solution and results for the $\sim 6 \times 10^4$ models were produced within 1 or 2 days. All the computed parameters were optimised by making use of the versatile, curve fitting, python software `lmfit` (Newville et al., 2016).

Given all these models, it was vital to implement a model selection strategy to obtain the best model fit for the observed transit light curves. For this purpose, both the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) were employed to determine the best model configuration (see Chapter 2 and Liddle, 2007, for a detailed decription of these two criteria). According to each criterion, the best model is established as the one that minimises the value of the criterion. During this analysis, the two selection criteria displayed similar model preferences, with BIC favouring slightly simpler models and AIC promoting models with more complex systematics. This is because BIC tends to penalise the number of free parameters more severely. In the end, I opted to use the AIC values as they provided a more conservative solution with larger error bars. This behaviour is in agreement with findings from transit light curve investigations of other targets (e.g Gibson, 2014; Nikolov et al., 2016).

The AIC values revealed that the best-fit systematics model for the blue data set is given by the product of six polynomials (time, airmass, shifts in the x and y axes, rate of change of the rotator angle, and FWHM) and that the best-fit model for the red data set is defined by the product of five polynomials (time, shifts in the x and y axes, rate of change of the rotator angle, and FWHM). In both cases, most systematic parameters are described by a 2nd-degree polynomial, while airmass and FWHM are described by a 1st-degree polynomial. Interestingly, there is a large set of functions with similar AIC values for the blue data set, indicating that the effect of some physical parameters on the shape of the light curves is likely less significant. This is not the case for the red data set where the preferred systematics model clearly stands out, having an AIC value that is a little lower than the AIC value of the second-best model and significantly lower than the AIC values of all other models.

For the transit parameters, the initial setup was the same as for WASP-75b, with the time of mid-transit, the semi-major axis to stellar radius ratio, the inclination,

the ratio of the radii and one limb darkening coefficient left to be free and the other transit parameters kept fixed. All starting parameter values were taken from Delrez et al. (2014), with the midpoint of the transit initially placed to the expected time (in BJD) from the ephemeris reported in that paper. Eccentricity was again fixed to a circular orbit. For the limb darkening coefficients, three possibilities were explored: the one already described with one linear coefficient free, another one with two coefficients included in the fit, and a last one with both coefficients fixed. However, it soon emerged from the fits that allowing for a free limb darkening coefficient gave unrealistic results, with the limb darkening coefficient settling at much lower values than expected. The situation was worse for two free limb darkening coefficients. It was, therefore, decided to keep both coefficients fixed during the white-light fits.

The white light curves were fitted twice with each fit consisting of two iterations. The two fits were established due to changes in the free parameters. In the first fit, the time of mid-transit, the planet-to-star radius ratio, inclination and the semi-major axis to stellar radius ratio were allowed to vary freely. In the second fit, only R_p/R_* was allowed to vary freely from the transit parameters. This is because, in the second fit, the time of the mid-transit was fixed to the solution from the first fit, and the semi-major axis to stellar radius ratio, as well as the inclination, were fixed to their weighted mean values from the first fits. The inclusion of two iterations was important in order to determine outliers in between. Residuals that departed from the best fit by at least three times the standard deviation were identified and removed before the second iteration. The two-fits procedure reflects the white light curve methodology described in Nikolov et al. (2016).

Figure 4.4 reveals the best-fit light curve models using baseline polynomial functions to decorrelate systematic errors. Interestingly, the common mode (as seen from the systematics line fit in the plots) shows a light curve peak at the beginning of the transit for the blue data set and at the end of the transit for the red data set. This observed curvature, to great extent, reflects the behaviour of airmass in the Earth's atmosphere. However, other effects (e.g. instrumental) may also be at play, especially in the red data where the peak tends to arrive when the target is already descending,

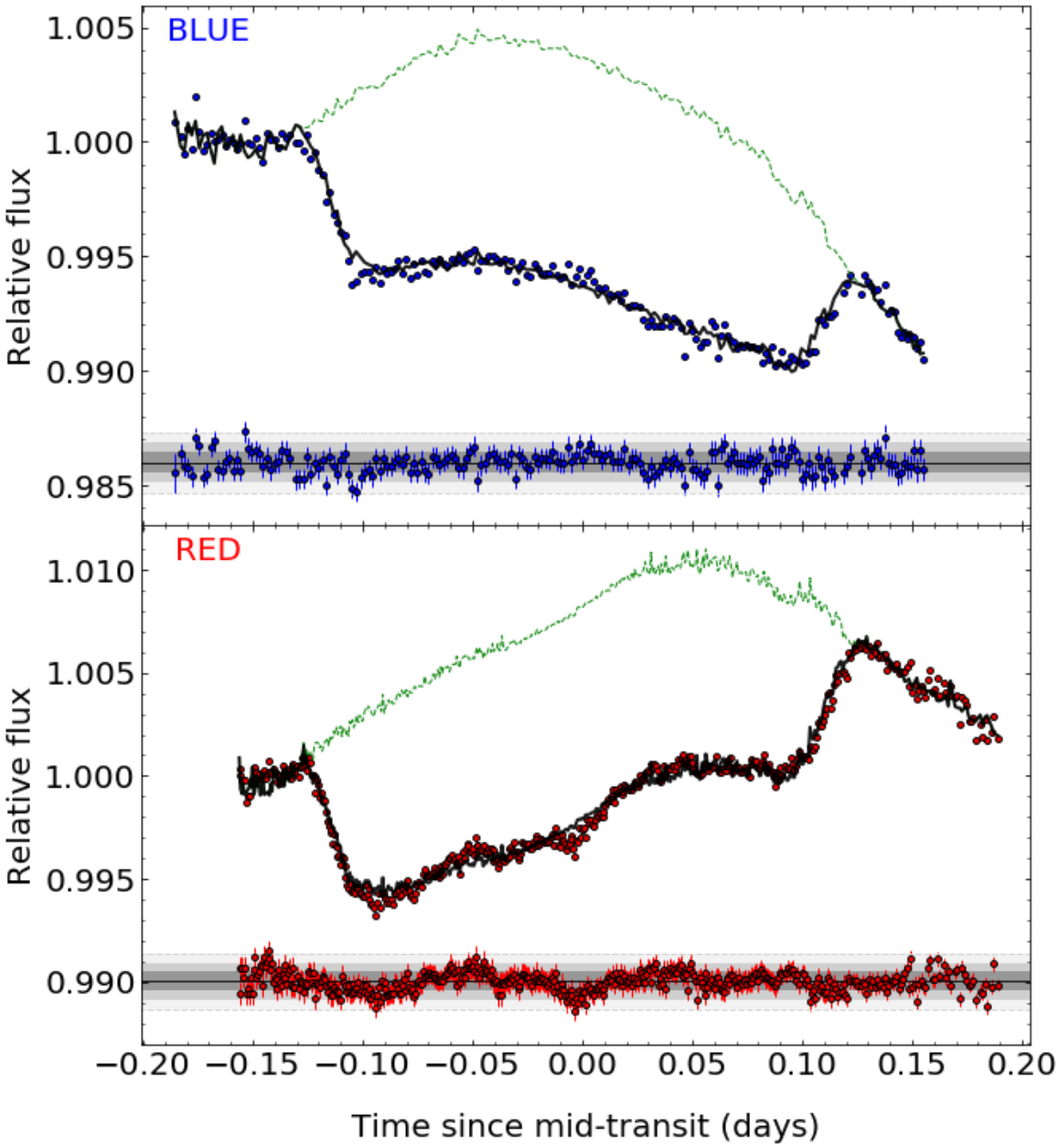

Figure 4.4: The white transit light curves of WASP-88b using baseline polynomial fits. Raw data, depicted by dark blue (top) and dark red dots (bottom), with the best-fit composite model (black line) and the noise model (green dashed line) that shows the common-mode trend. The residuals and their modified shot noise unsertainties are plotted at the bottom to better visualise their behaviour. Grey areas indicate the 1σ, 2σ and 3σ uncertainty regions, from darker to lighter shades.

but also in the blue data where a steeper increase is observed right at the start of the ingress. Part of these observed trends can also be attributed to the fitting algorithms preferred through the model selection procedure. And, indeed, as we will see later on, the influence from the fitting formula is high and sometimes it gives very dubious values for the common-mode corrected transmission spectrum parameters.

Another striking observation is the behaviour of the residuals, which appear to be wavy. This is especially true for the red data set where an almost sinusoidal-like feature is clearly visible in the graph of the light curve. Given these characteristics, it is obvious that the parametric solution had difficulties in establishing a good fit for the two time series, and particularly so for the red case. These effects suggest that the findings from the polynomial decorrelation method are immensely corrupted by unaccounted for systematics. Furthermore, this heavy contamination from unknown factors is at odds with the high precision that comes with the results obtained from this method. This is a major weakness of the fitting algorithms that are exclusively based on baseline polynomial functions.

Spectroscopic light curves Part of the white light curve systematics is common to all spectroscopic light curves and is, consequently, not dependent on wavelength. This common noise can be accounted for in the spectroscopic fits by the application of a common-mode correction. More specifically, this common mode was obtained by dividing the white-light relative fluxes by the best-fit transit models. The common mode factors were then removed from the spectroscopic time series by dividing the spectroscopic relative fluxes by these factors. A case without common-mode correction was also evaluated but was deemed to be unhelpful as it resulted in spectroscopic fits with unnatural limb darkening coefficients.

Despite the fact that the white light curve fits for the red data set were problematic, I opted to perform this analysis on both time series. Noise here was explored in a similar fashion to the white-light analysis, with multiple polynomial combinations being considered and assessed by the AIC rule. To speed up the process, models that

included the angle of the rotator (except its rate of change), the sky background, water vapour, and the ambient temperature and pressure were excluded from the fits. These optical state parameters were found to produce the worst results in the white-light time series investigation and were, therefore, not expected to perform better in the wavelength specific fits. Furthermore, the AIC value indicated simpler systematics models compared to the white light curves. The initial values for the transit parameters were taken from the output in the white-light analysis and were held fixed, except for the linear limb darkening coefficient u_1 and the transmission spectrum parameter R_p/R_*, which were allowed to vary.

The wavelength-dependent light curves also included two iterations and obeyed the same outlier rejection requirement mentioned in the white light curve analysis. Furthermore, the final data point uncertainties in both the white and spectroscopic light curves were re-evaluated. This is because the spectrophotometric uncertainties on their own only consider photon noise contributions and exclude other potential noise sources. This means the initial data point errors are likely underestimated. To correct for this, the uncertainties were multiplied by a constant factor $\sqrt{\chi^2_\nu}$, where χ^2_ν is the reduced χ^2 of the model fit.

Spectroscopic light curves were produced for both data sets, after the application of a common-mode correction, and are shown in Figures 4.5 and 4.6. These figures follow the whole approach, from the raw light curves to the detrended transits, and also display the residuals with their spectrophotometric uncertainties. It can be observed that the inadequate common mode generated from the white light curve analysis (see Section 4.3.2.1, White light curves) greatly removes the wavelike pattern seen in the raw light curves of the red data set. Although this is a desired outcome, the common-mode correction here does not take into account additional unknown systematics and may lead to unreliable values for the transit parameters.

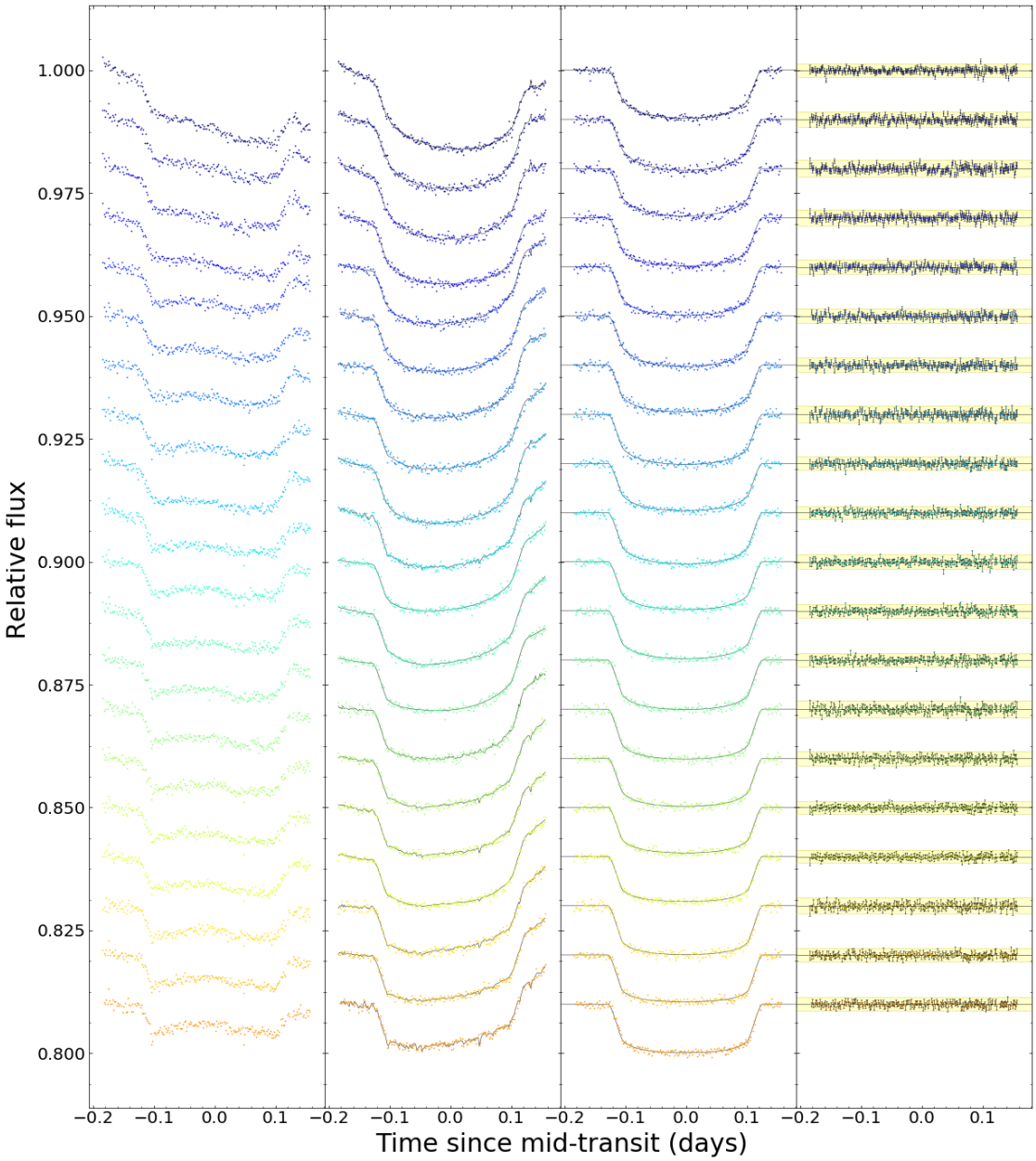

Figure 4.5: Spectroscopic light curves for the blue data set of WASP-88b using the polynomial decorrelation technique. Shorter wavelengths are indicated by violet dots, whereas longer wavelengths are represented by brown dots. Left column: Raw transit light curves. Middle left column: Common-mode corrected light curves with model fits from the two-component function. Middle right column: Decorrelated light curves and best-fit transit model. Right column: Light curve residuals from the best-fit model and their uncertainties.

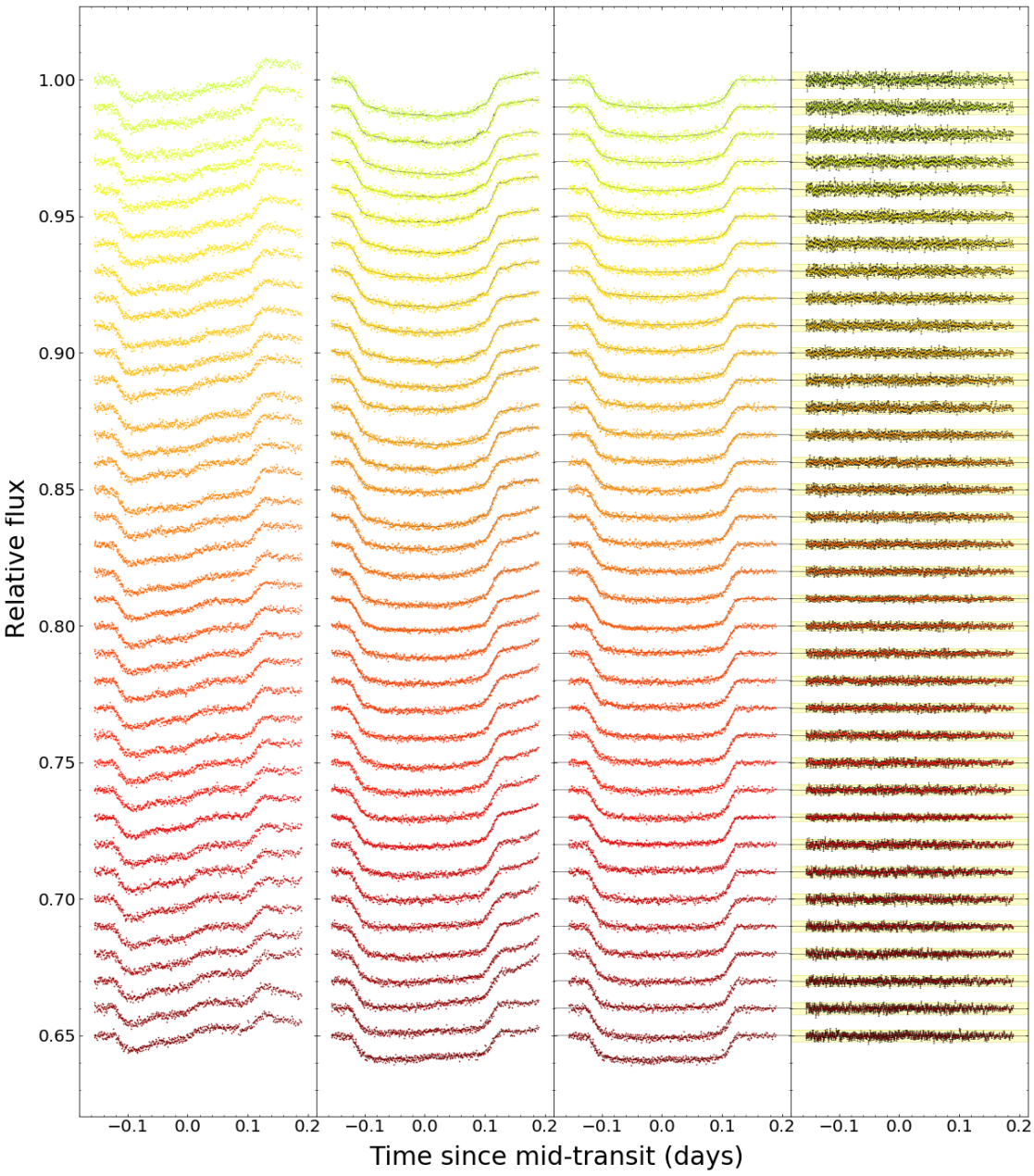

Figure 4.6: Same as Figure 4.5 but for the red data set. Longer wavelengths are depicted in dark red and shorter in light green.

4.3.2.2 GP analysis

White light curves Despite the simplicity and precision of the linear baseline models, the results were found to be characterised by unaccounted for noise contributions. To solve this problem, I changed my approach and followed a more flexible method by treating the time series as Gaussian Processes (GP). This was achieved by taking advantage of the speed and adaptability provided by the Python GP library `george` (Ambikasaran et al., 2015; Foreman-Mackey, 2015). Under this approach, the data points were defined by a multivariate normal probability distribution that consisted of a deterministic realisation of the transit component and a stochastic understanding of the systematics behaviour. This probability distribution p is given by (Gibson et al., 2012; Gibson, 2014):

$$p(\mathcal{D} \mid \theta, \phi) = \mathcal{N}(f \mid T(t, \theta), K), \tag{4.2}$$

where $\mathcal{N}$ portrays the multivariate normal distribution describing the data $\mathcal{D}$, f symbolises the relative flux measurements, T represents the mean function expressed by time t and the transit model parameters θ, and K is the covariance matrix containing all noise information declared by the hyperparameters ϕ. Time especially, as specified here, was obtained from the mid exposure times that were properly converted from Modified Julian Dates (MJD) to Barycentric Julian Dates (BJD) by making use of the python time-conversion module `barycorrpy` (Kanodia & Wright, 2018).

A variety of kernel functions were considered to model interference from external factors in the transit light curves. The most promising kernel configurations that have convincingly proven their effectiveness were the squared exponential kernel and the Matérn 3/2 kernel. We ultimately selected the latter option as the contribution from noise in these particular data sets was substantial. The Matérn 3/2 kernel is quite efficient at following the incoherent nature of light curves affected by a considerable amount of unknown systematics. This is due to its capacity to be differentiated a finite number of times. This property is what makes this choice superior to the smoother and infinitely differentiable squared exponential kernel. Moreover, for data sets that exhibit time-correlated noise, the Matérn 3/2 kernel was found to overshadow other

kernels, achieving better results (Gibson et al., 2013a). As will be explained later, both data sets display some level of time-dependence in the shaping of the light curves.

On the basis that the Matérn 3/2 kernel is used, the covariance function is then formulated as:

$$K_{nm} = \xi^2 \left(1 + \sqrt{3} D_{nm}\right) \exp(-\sqrt{3} D_{nm}) + \delta_{nm} \left(\sigma_p \beta\right)^2 . \tag{4.3}$$

In this slightly modified version of the equation presented in Chapter 2 (see Equation 2.15), ξ determines the correlation amplitude or height scale, δ_{nm} defines the Kronecker delta, σ_p represents the spectrophotometric uncertainties calculated under the expectation of pure photon shot noise, and D_{nm} is regulated by:

$$D_{nm} = \sqrt{\sum_{\nu=1}^{N} \left(\frac{\left(\hat{w}_{\nu,n} - \hat{w}_{\nu,m}\right)^2}{\tau_{w_\nu}^2} \right)}, \tag{4.4}$$

where τ_{w_ν} describe the length scale parameters for a set of auxiliary noise parameters ($\hat{w}_\nu$) examined. The hats, in this instance, signify that the systematic parameters are put on the same scale by subtraction of the mean and division by the standard deviation. Furthermore, parameter β is a multiplicative factor included in the fit so that the underestimated shot noise uncertainties are rescaled towards more sensible values. This parameter is constant across all flux measurements of a given light curve.

Again, an assortment of instrumental and atmospheric noise parameters were taken into account for the two data sets. This time, most systematic contributions were evaluated, either individually or in combination, as part of the kernel function. The best systematics models were then selected as the ones that lead to low residual scatter. Based on this approach, it was established that each white-light time series was best explained by a distinct systematics model within the GP kernel. More specifically, for the blue light curve, a kernel that includes only time as a detrending element was favoured. On the contrary, the red data set displayed more complex systematics, and, thus, had to be fitted with a kernel function that combines positional drifts in the dispersion (x) and cross-dispersion (y) axes, as well as the rate of change of the rotator angle (z). This meant that, in this situation, three systematic parameters were used:

$\hat{w}_\nu = (\hat{w}_1, \hat{w}_2, \hat{w}_3) = (\hat{x}, \hat{y}, \hat{z})$. The use of z as a detrending factor in the red data is also supported by the fact that the rotator angle is sensitive to the non-uniform spatial transmission of the longitudinal atmospheric dispersion corrector (LADC). This could introduce temporal discrepancies in the data regardless of some recent adjustments (Boffin et al., 2016). Considering that the rate of change of the rotator angle is a function of time, then both data sets were found to exhibit time-dependent interference.

Linear trends defined by functions of time or airmass were also explored and assessed for their influence on the fit. However, it was determined that such parametric functions offered no improvement to the fit and generated a similar output to simpler forms of the mean function. For this reason, any baseline models of this kind were not included in the fits for the remainder of this analysis. In addition, it is worth noting here that despite the use of a different noise configuration for each data set, analysis of the transmission spectra in a later stage revealed a consistency in the results between different models. This further reassured us that the choice of systematics model had little effect on the common-mode correction.

In the same way as for the baseline polynomial analysis, the transit models for the relative light curves were built using `batman` (Kreidberg, 2015). The very reliable quadratic limb darkening law was also employed to bring in computational speed and balance out the demanding kernel calculations of the GPs. Nevertheless, we found that model arrangements with one or two free limb darkening coefficients led to poor constraints on their estimates and resulted in similar patterns to the ones observed in the baseline model fits. Therefore, both limb darkening coefficients were maintained fixed to their theoretical values obtained from the Stagger-grid (Magic et al., 2015). These values were retrieved using the effective temperature, surface gravity and metal abundance reported in Delrez et al. (2014). The more complex, non-linear, four-parameter limb darkening law (Claret, 2000) was also examined briefly to verify the effectiveness of the two-parameter law. This limb darkening law reliability check resulted in very good agreement between the two methods, with results indicating a match below $1\,\sigma$ for all measurements.

To appropriately model the dip in flux caused by the transit of WASP-88b, the

transit parameters $\theta = (t_0, R_\mathrm{p}/R_*, a/R_*, i)$ were permitted to vary freely in the fit. In addition, a group of hyperparameters $\phi = (\xi, \tau_{w_\nu})$ were utilised to assist with the handling of systematic effects. These noise parameters, along with the additional white noise term β, were also instructed to be variable during the fitting process. Initial assumptions for R_p/R_*, a/R_* and i were obtained from the values presented in Table 4.2, and the time of mid-transit was placed to the predicted value from the ephemeris given in the same table. Again, following the premise of a circular orbit, eccentricity was fixed to zero, while the orbital period was fixed to the computed value from the TESS data. We applied broad uniforms priors for β and the transit parameters, and wide log-uniform priors for the hyperparameters. Using log-space is the natural parameterisation for the hyperparameters, given that all these quantities are strictly positive.

The best-fit transit and noise parameters were then calculated by implementing a maximum likelihood optimisation procedure using a variation of the Markov-Chain Monte Carlo (MCMC) technique. For this purpose, we made use of the affine invariant ensemble sampler (Goodman & Weare, 2010) that is provided with the cutting-edge python MCMC package `emcee` (Foreman-Mackey et al., 2013). This approach was favoured over conventional MCMC methods as it examines parameter space fast and efficiently. More specifically, it creates an ensemble of MCMC chains that run simultaneously and are known as walkers. Each member of this group of walkers progressively moves towards higher likelihoods in a defined probability distribution through random steps and under the linear influence from other members of the ensemble. Parameter space is therefore completely covered in a decent amount of time and samples converge based on the exchange of information between the walkers.

In my investigation, I adopted a set of 150 walkers, which were more than enough for a parameter dimension of less than 10. The best-fit result was then established in a two stage fitting process that involved two iterations per fit. Each iteration consisted of three phases, two burn-in and one production, that were essential in the accurate recovery of the parameter uncertainties. In phase one, the walkers were set at random positions close to the TESS estimates for the transit parameters and close to arbitrary

quantities based on results from other similar targets for the hyperparameters. In phase two, the walkers were re-set to a compact space around the location of the best walker determined from the best (maximum) log-likelihood. This phase was important for the fitting formula as it facilitated a faster convergence for the members of the ensemble in the direction of the values that best explain the data. Both these burn-in phases were conducted for chains with a step size of 500 steps. The final production phase was then executed in 5000 steps to create thick posterior distributions that fully show the probability density of each parameter. A second iteration was performed with the rescaled light curve uncertainties using the estimated parameter β computed in the first iteration. The median parameter values were then identified from the marginalised posterior distribution of this iteration.

Following the logic of Nikolov et al. (2016), a second fit was carried out with a slightly changed parameterisation and fewer free parameters. In particular, after calculating the weighted mean of a/R_* and i, both parameters were kept fixed to these values in the second fit. The time of mid-transit was also held fixed to the outcome from the first fit. The only free transit parameter in this configuration was R_p/R_* to give the variation in the depth of the transit with respect to wavelength. In addition, a data removal condition was applied to the first fit in a similar fashion to the baseline polynomial analysis. Data points that were more than 3σ away from the model GP fit were not included in the second fit.

The median parameter estimates from the posterior distributions of the white light curve fits are presented in Table 4.3 and the full distributions from the first fit of the blue and red data sets are shown in Figures 4.7 and 4.8. The white light curves, in particular, are illustrated in Figure 4.9. This figure shows the best-fit GP model for the relative fluxes of the blue and red data sets, as well as the noise model together with the calculated residual errors.

Table 4.3: Transit parameters from the GP fit on the white light curves of WASP-88b.

Parameter	Value	Prior
P (d)	4.9540045 (fixed)	
e	0 (fixed)	
blue (1st fit)		
t_0 ($\mathrm{BJD_{TDB}}$)	$2457985.70345^{+0.00070}_{-0.00069}$	$\mathcal{U}$(-0.01,0.01)*
R_p/R_*	$0.0885^{+0.0046}_{-0.0044}$	$\mathcal{U}$(0.03,0.15)
a/R_*	$6.22^{+0.24}_{-0.22}$	$\mathcal{U}$(4,9)
i (degrees)	$86.07^{+0.87}_{-0.73}$	$\mathcal{U}$(80,90)
u_1	0.327 (fixed)	
u_2	0.363 (fixed)	
$\ln \alpha$	$-8.5^{+2.0}_{-1.4}$	$\mathcal{U}$(-20,15)
$\ln \tau_\mathrm{t}$	$2.9^{+1.5}_{-1.1}$	$\mathcal{U}$(-15,15)
β	$4.31^{+0.23}_{-0.21}$	$\mathcal{U}$(0,10)
red (1st fit)		
t_0 ($\mathrm{BJD_{TDB}}$)	$2457990.65735 \pm 0.00045$	$\mathcal{U}$(-0.01,0.01)*
R_p/R_*	$0.0868^{+0.0032}_{-0.0035}$	$\mathcal{U}$(0.03,0.15)
a/R_*	$6.62^{+0.16}_{-0.17}$	$\mathcal{U}$(4,9)
i (degrees)	$87.68^{+0.98}_{-0.74}$	$\mathcal{U}$(80,90)
u_1	0.193 (fixed)	
u_2	0.364 (fixed)	
$\ln \alpha$	$-10.33^{+1.12}_{-0.78}$	$\mathcal{U}$(-20,15)
$\ln \tau_\mathrm{x}$	$6.2^{+1.3}_{-1.1}$	$\mathcal{U}$(-15,15)
$\ln \tau_\mathrm{y}$	$1.14^{+0.81}_{-0.60}$	$\mathcal{U}$(-15,15)
$\ln \tau_\mathrm{z}$	$2.64^{+0.95}_{-0.77}$	$\mathcal{U}$(-15,15)
β	3.52 ± 0.13	$\mathcal{U}$(0,10)

Table continues on the next page...

Table 4.3: Continued from previous page.

Parameter	Value	Prior
Weighted mean:		
a/R_*	6.49 ± 0.14	
i (degrees)	$87.07^{+0.59}_{-0.50}$	
blue (2nd fit)		
R_p/R_*	$0.0882^{+0.0037}_{-0.0036}$	$\mathcal{U}$(0.03,0.15)
$\ln \alpha$	$-8.5^{+2.1}_{-1.4}$	$\mathcal{U}$(-20,15)
$\ln \tau_\mathrm{t}$	$2.9^{+1.5}_{-1.1}$	$\mathcal{U}$(-15,15)
β	$0.83^{+0.05}_{-0.04}$	$\mathcal{U}$(0,10)
red (2nd fit)		
R_p/R_*	$0.0858^{+0.0031}_{-0.0033}$	$\mathcal{U}$(0.03,0.15)
$\ln \alpha$	$-10.56^{+0.96}_{-0.68}$	$\mathcal{U}$(-20,15)
$\ln \tau_\mathrm{x}$	$5.9^{+1.2}_{-1.0}$	$\mathcal{U}$(-15,15)
$\ln \tau_\mathrm{y}$	$0.81^{+0.70}_{-0.55}$	$\mathcal{U}$(-15,15)
$\ln \tau_\mathrm{z}$	$2.17^{+0.88}_{-0.73}$	$\mathcal{U}$(-15,15)
β	0.88 ± 0.03	$\mathcal{U}$(0,10)

*Mid-transit times are computed based on the ephemeris from Table 4.2. The expected values are set to 0 for the prior assumptions.

Contrary to the baseline polynomial fitting method, the flexible GP model fits seem to follow the data quite well leading to lower residual scatter, as expected. Any unaccounted for systematics in the form of instrumental or atmospheric effects that make up common noise are significantly reduced through stochastic means. This is also true for the unaccounted residual errors caused by fixing the limb darkening coefficients to their theoretical values. This means that the GP model is found to be capable of correcting not only noise factors but also possible parameterisation effects. For this reason, the GP approach is quite powerful as a fitting technique and can help extract useful information even if the data sets are highly contaminated by unknown noise contributors. This is why GPs are the go-to method for particularly demanding data sets, and, indeed, this significant improvement in the fit is what made us keep this

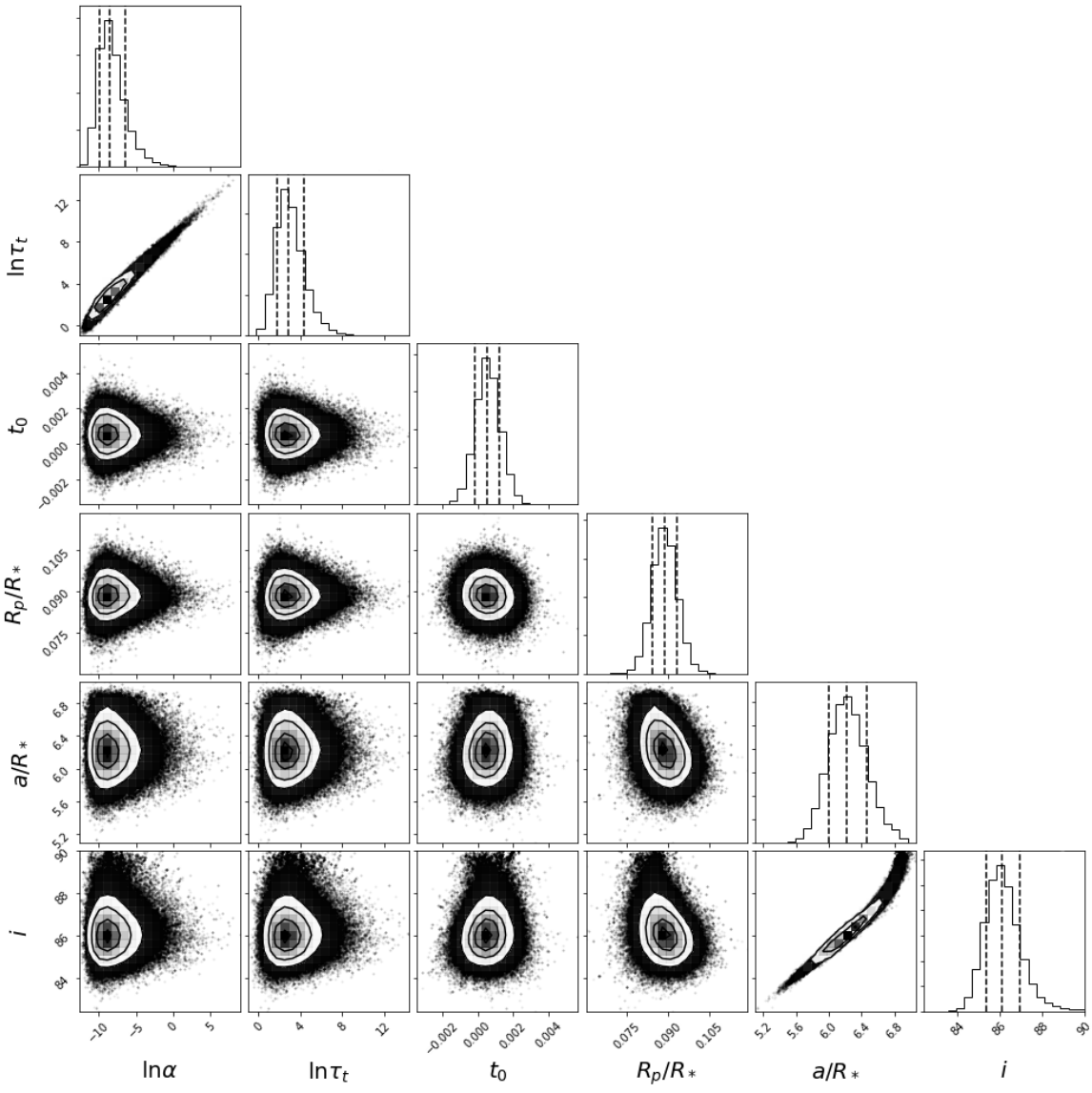

Figure 4.7: Marginalised posterior distribution from the initial fit on the blue data.

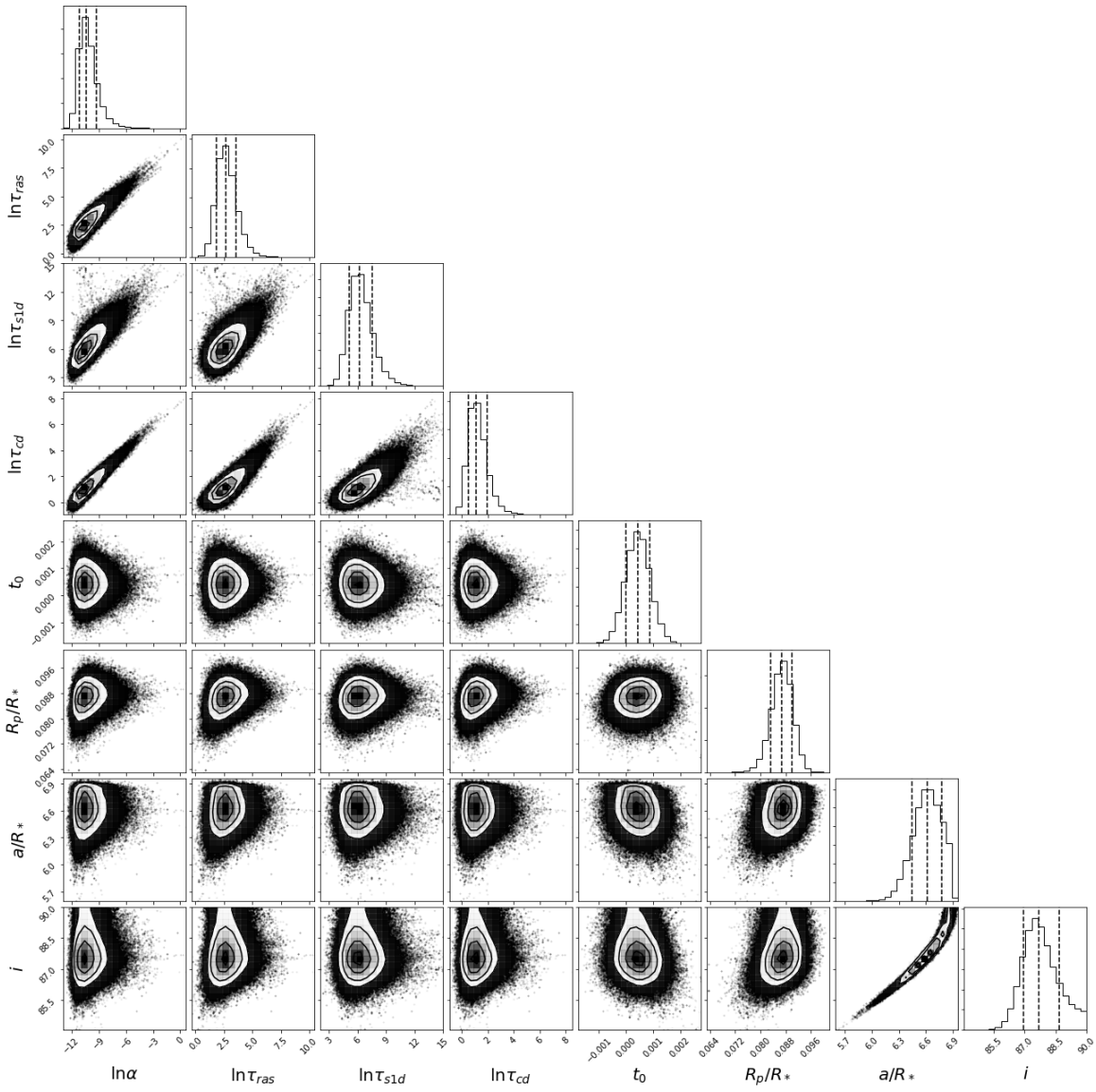

Figure 4.8: Marginalised posterior distribution from the initial fit on the red data.

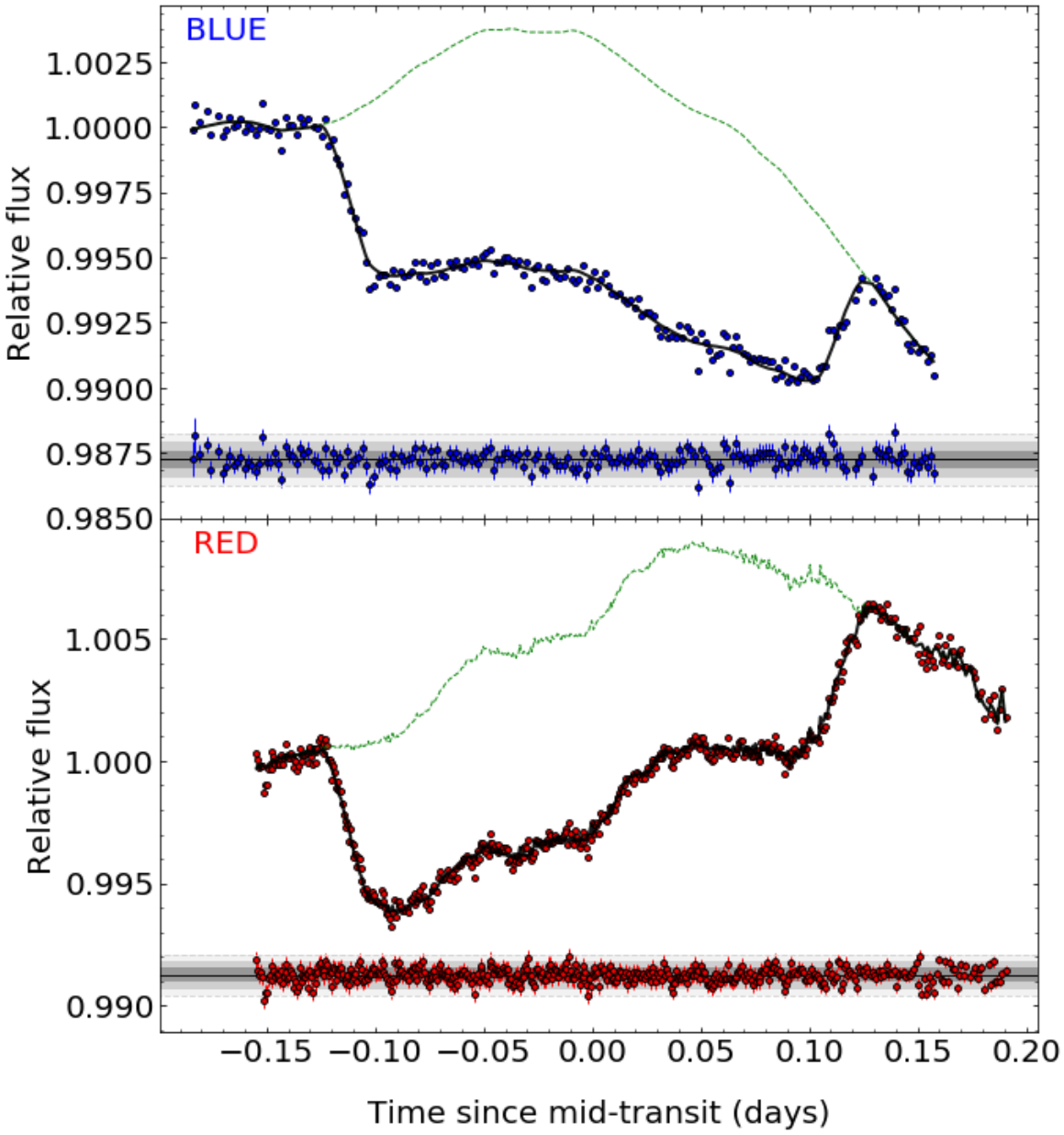

Figure 4.9: The white light curves of WASP-88b from the GP analysis. Data are depicted in dark blue for the blue data (top) and red for the red data (bottom). The GP model is represented by a black line and the systematics model is indicated by a green dashed line. The residuals are plotted below the transit light curves and their error bars show the modified spectrophotometric uncertainties. The transparent regions, from darker to lighter shades of grey, indicate the 1σ, 2σ and 3σ uncertainty of the residuals.

analysis for use in any subsequent atmospheric modelling.

Another characteristic of the GPs, that is more evident from Table 4.3 and the corner plots, is the relatively large error bars for the planet-specific parameters, which are found to be about two to three times larger than their counterparts obtained through polynomial fitting. This indicates that the error estimates from this method are more conservative and take into account plausible contributions from unknown systematics. This is actually an improvement from the baseline models that gave fairly underestimated errors for the rather poor generated fits. These results further demonstrate the reliability of the GP method.

Spectroscopic light curves The GP fit in the white light curve analysis resulted in a more flexible fit compared to the polynomial decorrelation analysis, creating a cleaner (i.e. with less unaccounted for systematics) common mode (see Figure 4.9). This common trend was again derived through division of the raw white transit light curves by the median transit model. The only difference this time was that the best-fit transit parameters were likely closer to reality as they were evaluated with the aid of a GP framework. We used this trend and applied a common mode correction to the binned relative fluxes by dividing them by this white light behaviour. Despite this correction, a wavelength-dependent noise component still affected the shape of the spectroscopic light curves producing minor trends in the form of upward or downward slopes.

Much of the spectroscopic analysis reflected procedures followed in the white light curve analysis as we relied again on the power of the GPs to de-trend the wavelength-dependent noise segments from the spectroscopic relative fluxes. One key difference was the use of time only as a regressor for both the blue and red data sets. The rationale behind the use of a GP kernel of time for the spectroscopic fits was that any influence from physical systematic parameters was extracted and removed during the common mode correction step. In that respect, any remaining impact from external systematic elements could be the result of temporal effects. In any case, more complex

kernel parameterisations that include physical parameters were also investigated but the results were deemed to be unreliable due to the 'untamed' nature of the limb darkening coefficients. In contrast, time was found to produce relatively good fits for any leftover residuals, confirming that our reasoning was sound.

To obtain the optimised values for the kernel and transit parameters, a similar MCMC sampling process to the white-light time series analysis was followed. This means that the binned light curves were also fitted in two iterations that consisted of three phases each. Again, an ensemble of 150 walkers was employed to explore the parameter space through two burn-in phases of 500 steps each, and one production phase of 1000 steps. This reduction in step size to only double the amount of the burn-in phases for the final phase was strategically put in place to accelerate the time-consuming GP analysis to an execution time of about two days. Nevertheless, an analysis that included a higher number of steps for the production phase was also performed to see whether the results are any different then. I found that the outcomes were highly similar, despite the tripling of the execution time. This is naturally expected as convergence is attained pretty fast (taking also into account the two burn-in phases).

The parameters for each spectroscopic light curve were set up in accordance with the findings from the white light curve analysis. This implies that the transit parameters t_0, a/R_*, and i were kept fixed to the estimates from the white light investigation. The stellar-to-planetary radius ratio R_p/R_*, however, varies with wavelength and was, therefore, incorporated in the fit as a free parameter. Again, all possible combinations of free and fixed limb darkening coefficients from the quadratic law were examined with respect to the obtained transmission spectrum. Ultimately, we proceeded with the case where the linear limb darkening coefficient is left free and the quadratic limb darkening coefficient is held fixed. This limb darkening configuration produced the best results in the spectroscopic light curves. In addition, all noise contributors (i.e. the two hyperparameters and parameter β) were allowed to be free, to get the optimal transit characteristics of WASP-88b.

An adaptation without common mode correction was also tested in a repetition of

the baseline polynomial paradigm. This revealed a similar pattern in the response of the limb darkening coefficients, which again had the tendency to settle at considerably lower values than their theoretical predictions. The cause of this behaviour is not exactly known but a degeneracy between noise factors and limb darkening was suspected as a possible culprit. Once again, a second iteration was needed to eliminate residual effects caused by deviating data points. The same 3σ rule was applied for each spectroscopic light curve and any outliers were discarded from the rest of the fitting process in order to arrive at the final solution.

The whole step-by-step procedure for the spectroscopic light curves is depicted in Figures 4.10 and 4.11. The methodology here is essentially very similar to the one followed for the polynomial fits with the only key difference being the handling of the noise, which is isolated stochastically. The final product contains the wavelength-dependent transmission spectrum parameters $R_{\rm p}/R_*$ and u_1, and is presented in Table 4.4. Note that, again, the uncertainties are about three times larger compared to the simple linear decorrelation case.

Table 4.4: Transmission spectrum of WASP-88b from the GP analysis of the spectroscopic light curves.

Wavelength Range (Å)	$R_{\rm p}/R_*$	u_1	u_2
$4413-4653$	$0.0916^{+0.0041}_{-0.0046}$	$0.392^{+0.078}_{-0.085}$	0.336
$4653-4733$	$0.0939^{+0.0082}_{-0.0040}$	$0.422^{+0.060}_{-0.059}$	0.355
$4733-4813$	$0.1009^{+0.0039}_{-0.0079}$	$0.431^{+0.069}_{-0.075}$	0.347
$4813-4893$	$0.0885^{+0.0044}_{-0.0039}$	$0.264^{+0.073}_{-0.081}$	0.416
$4893-4973$	$0.0880^{+0.0038}_{-0.0040}$	$0.348^{+0.077}_{-0.076}$	0.374
$4973-5053$	$0.0873^{+0.0035}_{-0.0038}$	$0.449^{+0.063}_{-0.066}$	0.340
$5053-5133$	$0.0873^{+0.0036}_{-0.0042}$	$0.281^{+0.075}_{-0.093}$	0.358
$5133-5213$	$0.0855^{+0.0040}_{-0.0047}$	$0.333^{+0.087}_{-0.095}$	0.353
$5213-5293$	$0.0904^{+0.0040}_{-0.0045}$	$0.336^{+0.078}_{-0.084}$	0.371
$5293-5373$	$0.0886^{+0.0028}_{-0.0029}$	$0.255^{+0.062}_{-0.068}$	0.365
$5373-5453$	$0.0892^{+0.0028}_{-0.0029}$	$0.170^{+0.069}_{-0.074}$	0.400
$5453-5533$	$0.0893^{+0.0028}_{-0.0028}$	$0.268^{+0.057}_{-0.060}$	0.361

Table continues on the next page...

Table 4.4: Continued from previous page.

Wavelength Range (Å)	$R_{\rm p}/R_*$	u_1	u_2
$5533 - 5613$	$0.0904^{+0.0021}_{-0.0022}$	$0.260^{+0.048}_{-0.052}$	0.368
$5613 - 5693$	$0.0890^{+0.0025}_{-0.0027}$	$0.183^{+0.058}_{-0.066}$	0.369
$5693 - 5773$	$0.0847^{+0.0029}_{-0.0031}$	$0.203^{+0.064}_{-0.068}$	0.376
$5773 - 5853$	$0.0878^{+0.0023}_{-0.0024}$	$0.242^{+0.050}_{-0.052}$	0.358
$5853 - 5933$	$0.0867^{+0.0023}_{-0.0025}$	$0.247^{+0.054}_{-0.058}$	0.380
$5933 - 6013$	$0.0861^{+0.0022}_{-0.0023}$	$0.198^{+0.055}_{-0.061}$	0.377
$6013 - 6093$	$0.0874^{+0.0024}_{-0.0025}$	$0.126^{+0.061}_{-0.066}$	0.374
$6093 - 6173$	$0.0864^{+0.0025}_{-0.0024}$	$0.165^{+0.058}_{-0.064}$	0.358
$6173 - 6253$	$0.0874^{+0.0028}_{-0.0027}$	$0.205^{+0.065}_{-0.069}$	0.378
$6253 - 6333$	$0.0914^{+0.0056}_{-0.0053}$	$0.256^{+0.087}_{-0.095}$	0.364
$6333 - 6413$	$0.0862^{+0.0036}_{-0.0034}$	$0.070^{+0.091}_{-0.108}$	0.375
$6413 - 6493$	$0.0875^{+0.0029}_{-0.0026}$	$0.299^{+0.062}_{-0.061}$	0.379
$6493 - 6573$	$0.0887^{+0.0066}_{-0.0047}$	$-0.074^{+0.143}_{-0.137}$	0.391
$6573 - 6653$	$0.0866^{+0.0034}_{-0.0030}$	$0.222^{+0.067}_{-0.074}$	0.390
$6653 - 6733$	$0.0850^{+0.0031}_{-0.0029}$	$0.207^{+0.069}_{-0.075}$	0.369
$6733 - 6813$	$0.0873^{+0.0034}_{-0.0029}$	$0.170^{+0.071}_{-0.077}$	0.361
$6813 - 6973$	$0.0873^{+0.0028}_{-0.0026}$	$0.267^{+0.058}_{-0.058}$	0.363
$6973 - 7053$	$0.0867^{+0.0033}_{-0.0029}$	$0.236^{+0.071}_{-0.071}$	0.363
$7053 - 7133$	$0.0862^{+0.0033}_{-0.0031}$	$0.176^{+0.077}_{-0.082}$	0.361
$7133 - 7213$	$0.0844^{+0.0029}_{-0.0030}$	$0.172^{+0.074}_{-0.083}$	0.351
$7213 - 7293$	$0.0852^{+0.0029}_{-0.0030}$	$0.131^{+0.077}_{-0.091}$	0.360
$7293 - 7373$	$0.0835^{+0.0030}_{-0.0031}$	$0.131^{+0.081}_{-0.092}$	0.361
$7373 - 7453$	$0.0864^{+0.0030}_{-0.0029}$	$0.161^{+0.070}_{-0.078}$	0.360
$7453 - 7533$	$0.0843^{+0.0030}_{-0.0032}$	$0.212^{+0.080}_{-0.086}$	0.357
$7533 - 7693$	$0.0852^{+0.0028}_{-0.0027}$	$0.150^{+0.066}_{-0.071}$	0.361
$7693 - 7773$	$0.0863^{+0.0039}_{-0.0038}$	$0.178^{+0.086}_{-0.099}$	0.351
$7773 - 7853$	$0.0856^{+0.0047}_{-0.0051}$	$0.230^{+0.102}_{-0.116}$	0.362
$7853 - 7933$	$0.0877^{+0.0045}_{-0.0048}$	$0.383^{+0.084}_{-0.087}$	0.355
$7933 - 8013$	$0.0854^{+0.0046}_{-0.0044}$	$0.241^{+0.101}_{-0.109}$	0.352
$8013 - 8093$	$0.0828^{+0.0046}_{-0.0047}$	$0.396^{+0.080}_{-0.088}$	0.348
$8093 - 8173$	$0.0875^{+0.0043}_{-0.0041}$	$0.257^{+0.086}_{-0.096}$	0.353
$8173 - 8253$	$0.0848^{+0.0018}_{-0.0019}$	$0.182^{+0.069}_{-0.059}$	0.349
$8253 - 8333$	$0.0792^{+0.0030}_{-0.0034}$	$0.224^{+0.084}_{-0.097}$	0.349

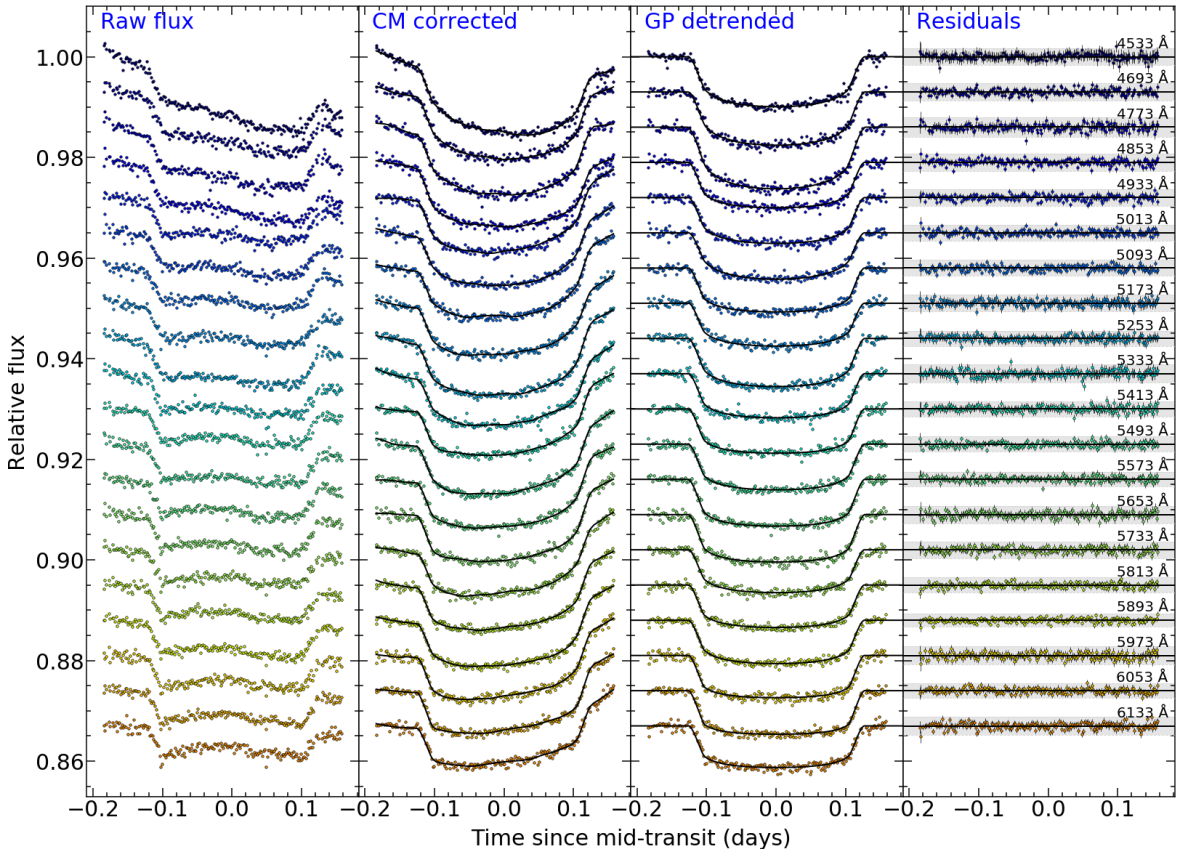

Figure 4.10: Spectroscopic light curves from the GP analysis of the blue WASP-88b data set. Shorter wavelengths are indicated by violet dots, whereas longer wavelengths are represented by brown dots. First column: Raw transit light curves. Second column: Common-mode corrected light curves with GP model fits. Third column: Decorrelated light curves and best-fit transit model. Fourth column: Light curve residuals from the best-fit model and their 1σ uncertainties.

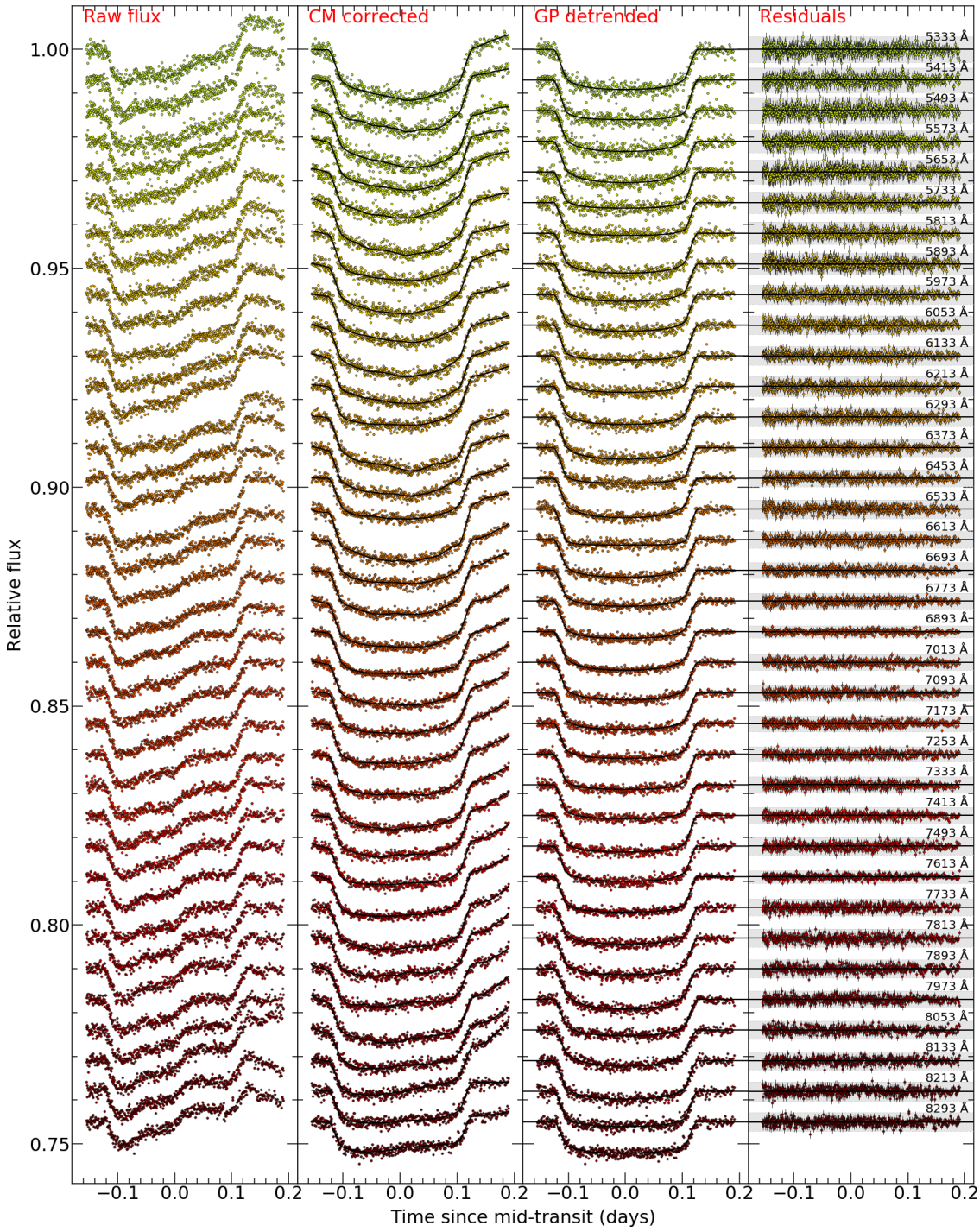

Figure 4.11: Same as Figure 4.10, but for the red WASP-88b data set. Shorter wavelengths are depicted in light green and longer in dark red.

One interesting observation is that the spectroscopic light curves exhibit less systematic effects compared to the white light curves, but only after common mode correction. The raw, red spectroscopic light curves still show the wavelike pattern observed during the white-light analysis. This light curve trend due to unknown interference is common for all spectroscopic bins and demonstrates that the application of a common mode correction is a justified way to analyse spectroscopic data.

4.3.2.3 Combined analysis

This analysis is similar to the one performed in Nikolov et al. (2018b), where the relatively clear atmosphere made it possible to obtain an absolute sodium abundance for a slightly cooler exoplanet. This was the first time an absolute sodium abundance was estimated for the atmosphere of an exoplanet. These observations were also made from the ground and used the same instrument employed here (FORS2). The idea in this case is that the more severe systematics can be modeled out during the white light curve analysis, whereas the less significant systematics in the common-mode corrected spectroscopic light curves may simply be evaluated using conventional methods. This implies that noise is handled stochastically under a GP framework in the white-light analysis and deterministically using baseline parametric functions in the spectroscopic analysis. Interestingly, this approach was considered before the fully GP approach presented earlier.

For the white light curve analysis, the results from the most credible solution described in Section 4.3.2.2 were used. This means that the GP model for the blue data set was merely defined by a regressor of time whilst the GP model for the red data set was determined from a more complex function that involved horizontal and vertical displacements, as well as the rate of change of the rotator angle. This configuration produced the most reliable fits with conservative errors but without a significant loss in precision. It was therefore the ideal frame for a spectrophotometric investigation that explored common mode corrected light curves.

Based on the GP computed correction from the white-light analysis, the transit

parameters $R_{\rm p}/R_*$ and u_1 were then estimated as a function of wavelength using the same narrow spectroscopic channels mentioned previously. Many possible parametric fit combinations were considered and only the worst noise factors as determined in Section 4.3.2.1 were excluded (e.g. pressure, temperature, background etc). These combinations were explored under the Levenberg-Marquardt least squares framework and the best-fit models for the spectroscopic light curves were eventually selected based on the AIC statistic.

In the same way as in the other analyses of WASP-88b, outliers were spotted and removed in between fits following the now familiar 3σ condition in the residual trend. This was an important step in every procedure because we found that even a single outlier can have a tremendous effect on the outcome of the fit. Nevertheless, the acquired transit parameters did heavily rely on the parametric noise function applied, which eventually made this technique somewhat questionable for this particular target. This is not the case with the GP model fits that show a higher consistency between different models. For this reason, the GP approach was eventually selected as the most favourable description of the transiting system. In any case, the best-fit transmission spectrum from this combined approach is compared against other solutions in Section 4.4.

4.4 Transmission spectrum

Once the analysis of the system from all three methodologies was completed, it was time to construct the transmission spectrum for both the blue and red data sets based on the R_p/R_* values determined from the spectrophotometric channels. We produced transmission spectra for all techniques, including some additional sub-cases to check the reliability of our results. The comparisons revealed the superiority of the fully GP approach, and so this section will focus on the characteristic properties of the transmission spectrum created using this method. The end result will be placed in the context of the transmission spectra obtained through other approaches to address any

doubts about the success of this method.

A careful examination of the planet-to-star radii ratios for each individual data set found a mean R_p/R_* value of 0.0884 ± 0.0035 for the blue data set and a mean R_p/R_* value of 0.0865 ± 0.0024 for the red data set. These values are in excellent agreement with the photometry estimate from TESS (see Table 4.2) and slightly higher compared to the result reported in the discovery paper (Delrez et al., 2014). A minor offset of magnitude $\Delta(R_\mathrm{p}/R_*) = 0.0006 \pm 0.0026$ was also detected in the overlapping region between the two observations. The offset is expected based on the outcomes from the white-light analysis and the subsequent common-mode correction. It is, however, marginal, demonstrating the great compatibility of using stochastic processes throughout the whole fitting procedure from the white-light analysis to the spectroscopic analysis. Another noticeable point is the large size of the errors for the transmission spectrum quantities. Due to the use of the more flexible and conservative GP fits, systematics from unknown physical sources are also taken into consideration adding extra weight to the uncertainties. This effect also arises from the model setup in the white light curve analysis where limb darkening coefficients were held fixed. Any remaining residuals from the fixed parameters there are propagated to the common-mode corrected spectroscopic light curves and assessed as unidentified systematics by the GP model.

Despite the fact that the offset is minuscule, a decision was made to apply a vertical displacement in one of the two data sets. Motivated by the larger atmospheric extinction effects in the blue data, we implemented this correction on that specific data set. The impact on the transmission spectrum from this adjustment was insignificant and a comparison with an uncorrected version showed an extremely low difference of 0.0003 on the average level of the spectra. Nevertheless, the offset was applied on the blue radii ratios and a combined transmission spectrum was produced using the weighted mean R_p/R_* estimates from the overlapping wavelength region of the two observations. The resulting transmission spectrum, as shown in Figure 4.12, appears to be relatively flat and highly affected by strong opacity sources in the upper atmosphere that limit the capacity to probe deeper. A straight, horizontal line was used to assess the level of grey opacity and, hence, the potential influence from a thick cloud deck.

The horizontal fit employs the Bayesian information criterion (BIC) (Schwarz, 1978) to stimulate meaningful statistical comparisons with other types of atmospheres. This model selection criterion is useful in our case as it penalises models with larger free parameter dimensions in an attempt to avoid potential over-fitting effects caused by the large uncertainties. The application of this grey opacity fit resulted in a BIC value of 38.0 for an atmosphere that completely blocks transmitted light. Figure 4.12 shows the steps followed until the generation of the combined transmission spectrum.

All data points were found to agree well with the horizontal line and any deviations were up to 2σ from this mark. Despite the featureless appearance of the transmission spectrum and the absence of the characteristic signatures of the sodium and potassium lines at ∼5890 and ∼7700 Å, respectively, the shape of the transmission spectrum hinted at a slope that goes upward towards bluer wavelengths. Such as a slope is often indicative of scattering phenomena from small particles in the upper atmosphere and so we planned to investigate this further.

In accordance with the reasoning outlined by Lecavelier Des Etangs et al. (2008a), atmospheric opacity due to scattering can be described by a cross-section of the form $\sigma = \sigma_0(\lambda/\lambda_0)^\gamma$. This function varies with respect to wavelength and is defined by index γ that describes the scattering slope. γ can be estimated from the slope $\frac{\mathrm{d}(R_\mathrm{p}/R_*)}{\mathrm{d}\ln\lambda}$ of the transmission spectrum as follows:

$$\gamma = \frac{R_*}{H}\frac{\mathrm{d}\left(R_\mathrm{p}/R_*\right)}{\mathrm{d}\ln\lambda}. \tag{4.5}$$

Here, H symbolises the atmospheric pressure scale height and is given by:

$$H = \frac{k_\mathrm{B}T_\mathrm{eq}}{\mu g_\mathrm{p}}, \tag{4.6}$$

where k_B is the Boltzmann constant, T_eq is the equilibrium temperature of the planet, μ is the mean molecular mass of the atmospheric constituents and g_p is the surface gravity of the planet (see also Section 1.4.2 for an exact derivation of this parameter).

A scattering line fit with two free parameters (slope, intercept) was performed on the whole spectrum in order to get the slope $\frac{\mathrm{d}(R_\mathrm{p}/R_*)}{\mathrm{d}\ln\lambda}$, which was then used to obtain index γ from Equation 4.5. The fit revealed a BIC value of 25.8 and is an improvement

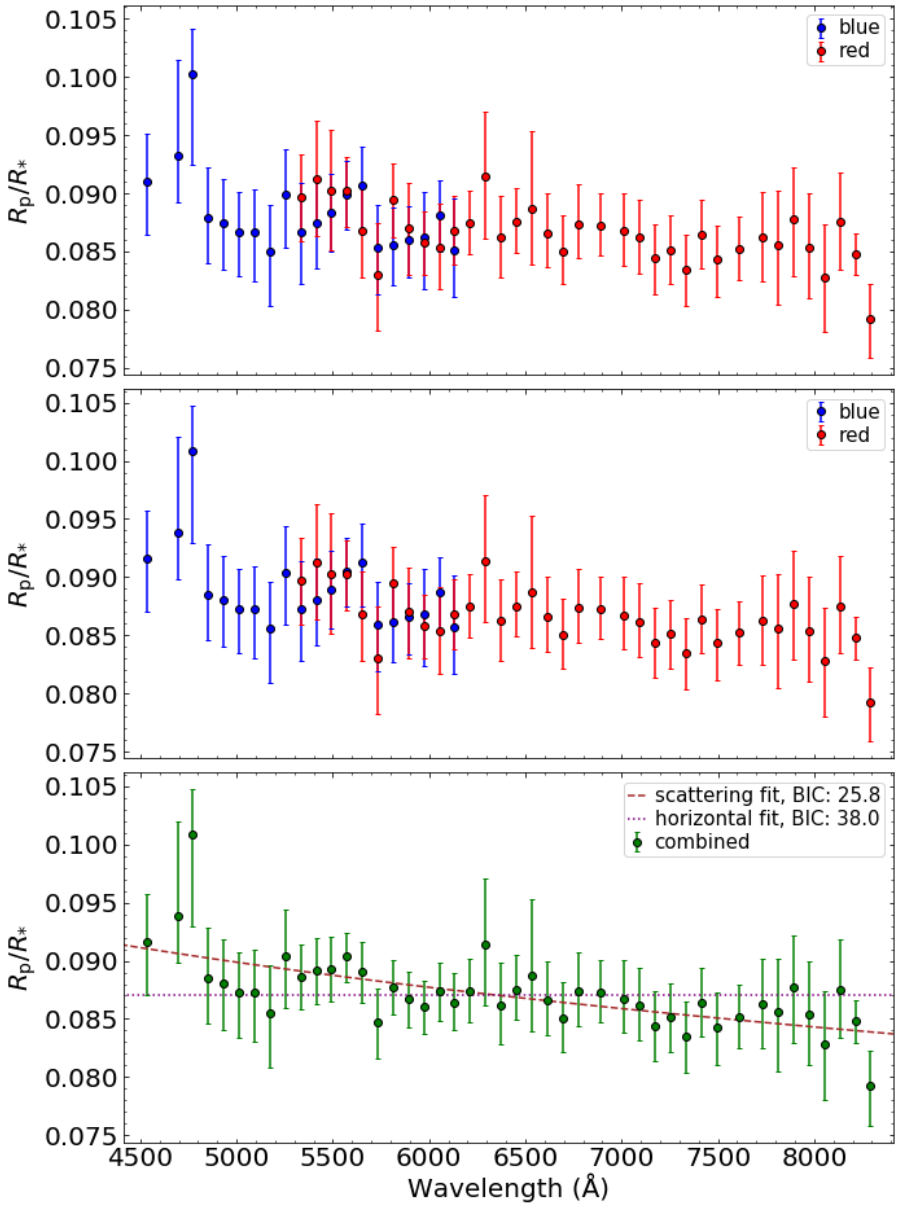

Figure 4.12: The transmission spectrum of WASP-88b from the GP analysis. Top row: Uncorrected transmission spectra from the two separate data sets. Middle row: After application of a very small offset to the blue data. Bottom row: The combined result after computing the weighted mean R_p/R_* in the overlapping wavelength region. Green circles indicate the combined data set, while blue and red circles depict the two separate data sets. In the bottom row, a grey opacity fit (dotted magenta line) and a Rayleigh slope fit (brown dashed line) to the data are also included.

compared to the horizontal fit as the BIC value here is significantly lower. In fact, the scattering fit is strongly favoured giving $\Delta\text{BIC} = 12.2$. By plugging in the value of the slope and the values for the stellar radius, planet surface gravity and planet equilibrium temperature from Table 4.2 in Equation 4.5, the index was calculated as $\gamma = -12.3^{+2.7}_{-3.0}$. This number indicates significant scattering in the atmosphere that is about three times stronger than Rayleigh scattering ($\gamma = -4$). This level of excess scattering is not unheard of and is now increasingly being observed in the atmospheres of hot gaseous exoplanets (e.g. Alam et al., 2020; Chen et al., 2021a).

A transmission spectrum based exclusively on baseline polynomials fits was also created, despite the poor fits on the red data during the white-light analysis. The shape of this transmission spectrum was found to be similar to the one obtained from the combined polynomial/GP analysis but was significantly displaced compared to that one. This result is to be expected as the common-mode was quite different between the white-light polynomial analysis and the white-light GP analysis. Both these transmission spectra were produced following the AIC model selection criterion. Figure 4.13 shows the transmission spectra from all three approaches.

Another parametric fit relying on a second degree polynomial of time in the spectroscopic analysis of the combined approach was also performed. Figure 4.14 illustrates how this transmission spectrum stands against the preferred GP model. It can be quickly observed that the parametric fit greatly reduces the uncertainties in R_p/R_*. However, another more important observation is that the behaviour of the transmission spectrum from the time polynomials is completely different to the behaviour of the spectrum from the best-fit AIC polynomials. Moreover, a parametric function of time induces a much steeper slope in the transmission spectrum. All these inconsistencies are signs that the parametric approach in both the white and spectroscopic analysis is rather unreliable.

Upon request by the referee during paper submission process, we also performed a spectroscopic analysis based on a GP analysis using FWHM as an additional regressor in the white-light analysis of the red data set. The white-light analysis had revealed that the GP model without FWHM was a better fit to the observed light curves and so

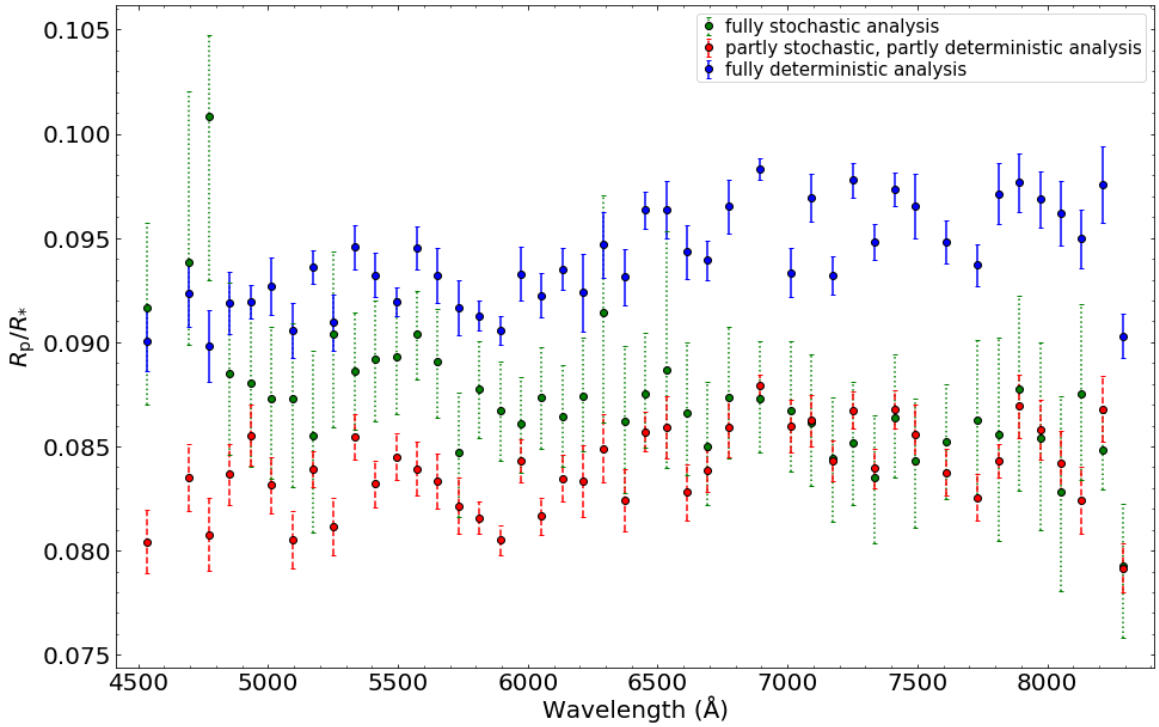

Figure 4.13: The transmission spectrum of WASP-88b from the GP analysis compared to the polynomial baseline model approach and the combined polynomial/GP approach. Note the increased R_p/R_* found in the fully deterministic case.

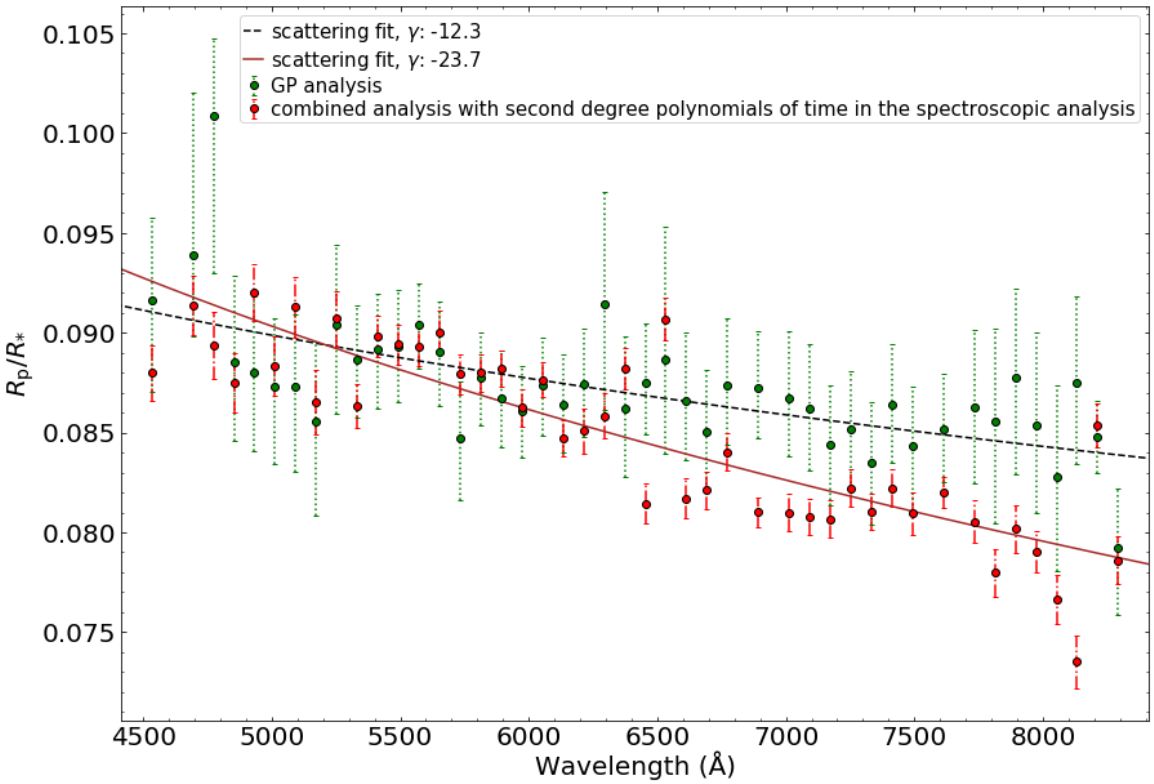

Figure 4.14: The transmission spectrum of WASP-88b from the GP analysis compared to results from the combined polynomial/GP approach for the special case where a second degree polynomial of time is used for all spectroscopic light curves. Note the combination of a very steep slope and small error bars in the combined polynomial/GP approach suggesting potentially unaccounted for systematics.

a fit that takes FWHM into account was not considered in the spectroscopic analysis. Nevertheless, due to the referee's useful comment, a transmission spectrum was also generated for this particular case. Figure 4.15 shows how this result compares with our favoured model. The similarities are remarkable with all data points agreeing at the 1σ level. The version that includes FWHM in the red fit finds a slightly less steep slope, but it is still quite significant. This excellent agreement between the two results demonstrates the relative stability of the GP result with changing systematic assumptions and provides greater confidence about the validity of the measured transmission spectrum.

4.5 Forward modelling and retrieval results

A more comprehensive analysis of the resulting transmission spectrum also took place by comparing the observed data with synthetic spectra from simulated atmospheres. This additional analysis was important in our effort to better understand the enhanced slope and the lack of atomic and molecular features in the transmission spectrum. For this purpose, we employed both forward modelling and retrieval methods using three different software packages to generate the best-fit model atmospheres.

4.5.1 Generic Grid

Our first and simplest option was to use the extensive grid of distinct model atmospheres created by Goyal et al. (2018) and based on the 1D plane-parallel radiative-convective equilibrium library ATMO (Amundsen et al., 2014; Tremblin et al., 2015, 2016). The grid itself is constantly upgraded (Goyal et al., 2019a,b, 2020) and has been very useful in the observation of other hot Jupiter atmospheres such as WASP-6b (Carter et al., 2020) and WASP-103b (Wilson et al., 2020). In this study, we make use of the generic grid[8] assuming an isothermal structure for the atmosphere and condi-

[8]This grid is publicly accessible on `https://exoctk.stsci.edu/generic`.

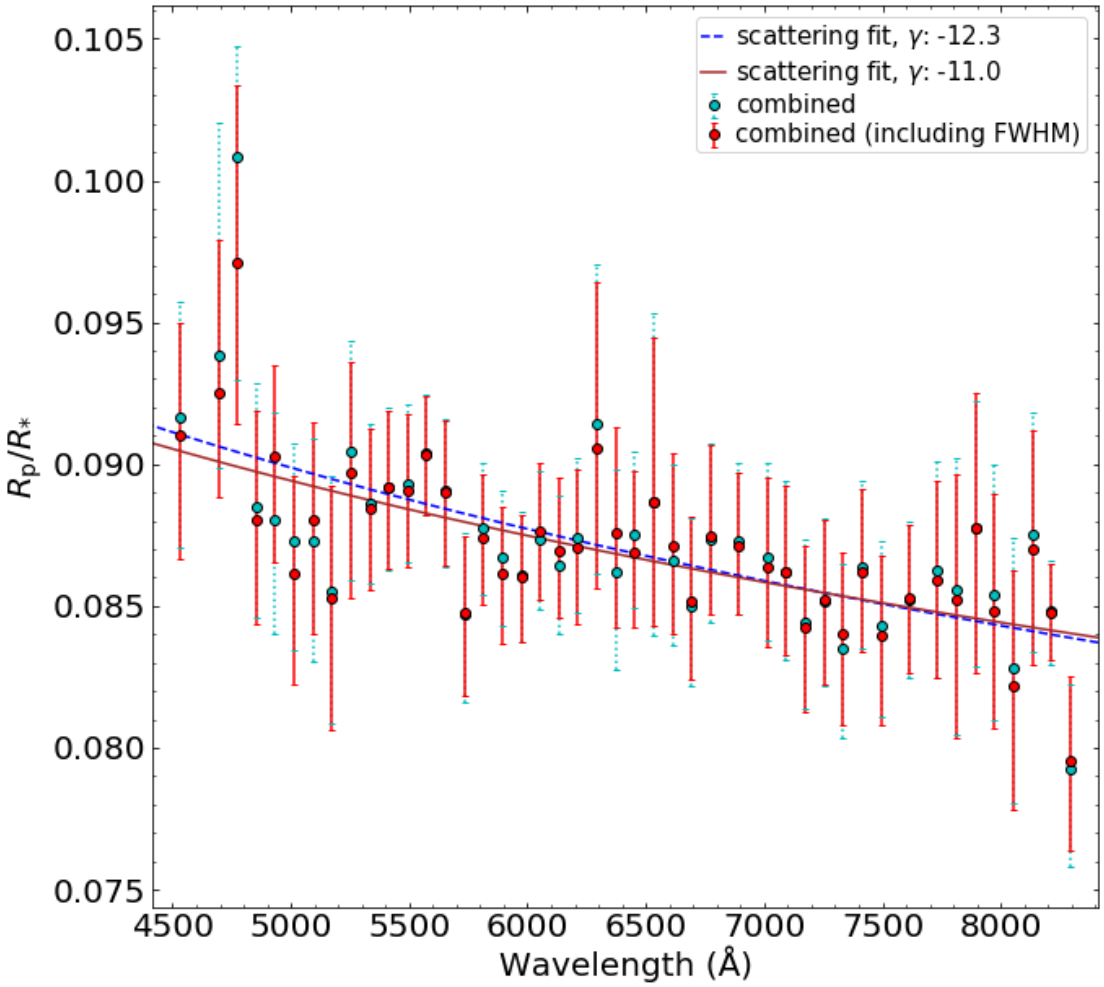

Figure 4.15: The transmission spectrum of WASP-88b from the GP analysis (cyan circles with dotted error bars) compared to a variation that includes FWHM in the white-light GP analysis (red circles with solid error bars). The original GP fit (solid brown line) agrees astoundingly well with the FWHM solution (blue dashed line) with only a minor difference in the steepness of the scattering slope.

tions that are governed by equilibrium chemistry (Goyal et al., 2019a). This version contains 21 opacity sources, with 19 coming from atomic and molecular species and another two coming from collision induced absorption from the atoms that make up the characteristic H/He envelopes of hot gaseous atmospheres.

To simplify the process and produce the most suitable fits, a few assumptions were made about the stellar parameters, the bulk properties of the planet and the physics that contribute to the observed atmospheric opacity. While the planetary and stellar radii were taken from Table 4.2 and remained the same, the generic grid offers only a selection of planetary surface gravities and equilibrium temperatures to explore. Given that the surface gravity of the planet was calculated at 4.84 m s^{-2} (see Table 4.2), the nearest value in the grid was 5 m s^{-2}. This value is well within 1σ from our measurements, making it a fair assumption, and so it was selected for the forward models. The planetary temperature, however, was fully examined in steps of 100 K, starting from 800 K and going up to 2000 K. Apart from the planet parameters, we made two further assumptions about the stellar atmosphere based on the derived quantities reported in (Delrez et al., 2014; Andreasen et al., 2017) and Table 4.2. These estimates indicate solar values for the metal abundance [Fe/H] and the C/O ratio, a result that was incorporated in the atmospheric models used for the creation of the synthetic spectra. One final assumption made was that material forming condensates at a specific pressure and temperature can be removed from the local pressure layer and all layers above it under the hypothesis that droplets will sink deeper into the atmosphere allowing the upper atmosphere to be free of condensing chemical species.

Three distinct versions of simulated atmospheres were explored in order to establish the best-fit scenario. These are an atmosphere with significantly enhanced scattering towards bluer wavelengths by a factor of 1100, an atmosphere with a thick cloud deck simulated by a cloudiness factor of 1, and a cloud-free atmosphere (see also Goyal et al., 2019a, for more information on the haze and cloudiness factors). The best-fit was then determined by calculating the mean model transit depth for every spectroscopic bin and following a least squares algorithm. The only free parameter of the fit was the vertical offset of the computed model spectrum from the observed

transmission spectrum. The fit was implemented for each atmospheric scenario and the best simulated atmosphere was selected according to the lowest BIC rule. This criterion indicated a hazy atmosphere for the exoplanet and a limb temperature of 1800 K as the best match to the obtained transmission spectrum. This scenario was found to dominate over the clear and cloudy atmosphere scenarios and a comparison of the BIC quantities revealed ΔBIC values of 6.4 and 6.6 for the two simulated atmospheres respectively. It should be noted here that these moderate differences are between simulated atmospheres of the same temperature.

4.5.2 PLATON

A second option with broad applications due to the wide available range provided in parameters such as temperature, metal abundance and carbon-to-oxygen ratio is the forward modelling and retrieval toolbox PLATON (PLanetary Atmospheric Tool for Observer Noobs, Zhang et al., 2019, 2020). In our case, this python package is used to perform forward modelling and produce theoretical spectra that closely match the observed data for the atmosphere of WASP-88b. To allow useful comparisons with the results from Section 4.5.1 we made a few identical assumptions about the properties of the atmosphere. Therefore, under this analysis, WASP-88b is also governed by an atmosphere that obeys equilibrium chemistry and isothermal temperature-pressure profiles. In addition, condensation affects the atmosphere in the same way by depleting condensing material from the local layer and all other layers at higher altitudes. One notable difference is the increased number of included opacities here as PLATON takes absorption from a total of 34 chemical species into account. However, it should be added that most of the molecules will have a negligible effect on the simulated transmission spectrum and that any discrepancies between the two methods are mainly a result of variations in the opacity tables used for the most influential chemical species. Another important distinction is the management of the cloud contribution on the observed spectrum. PLATON uses cloud-top pressure, which indicates the pressure layer at which light is effectively blocked due to the formation of a thick cloud deck, as an

equivalent parameter to the cloudiness factor described in Goyal et al. (2019a). Because of the very broad supported range in cloud and scattering parameters, PLATON is an excellent tool for investigations of extreme atmospheric phenomena at higher altitudes.

Emulating the approach taken for the generic grid (Goyal et al., 2019a), we study the gaseous envelope of WASP-88b assuming the planet radius and stellar radius presented in Table 4.2. Instead of the planet surface gravity, we consider its mass from Table 4.2 and make the further assumption that the star's composition reflects solar values. Therefore, a solar C/O ratio is applied, with the values for the stellar effective temperature and the solar metallicity being taken from Andreasen et al. (2017). The best-fit theoretical transmission spectrum is then computed through interpolation based on the same temperature range for the planet outlined in Section 4.5.1. One key difference here is the way the cloud deck is represented and so a divergence from the corresponding generic grid result is to be expected. In particular, the impact of a thick high-altitude grey opacity source is simulated by a cloud-top pressure of 1 Pa. Haze, on the other hand, is modelled in a similar way to the generic grid approach by placing the scattering factor to a value of 1100. The best-fit scenario is then determined again through a least squares fitting process and is selected as the model that gives the smallest BIC value. The results from PLATON are found to be quite similar to those obtained from the generic grid and favour a hazy sky for WASP-88b with a temperature of 1400 K. Clear and cloudy skies perform slightly worse, giving ΔBIC values of 7.5 and 6.8, respectively. Despite the lower limb temperature estimated here, the result, in conjunction with the outcome from Section 4.5.1, confirms the previous conclusion of enhanced scattering in the atmosphere of WASP-88b and reveals a relatively flat transmission spectrum.

The measured transmission spectrum, as well as the best-fit forward simulations of clear, cloudy and hazy atmospheres from PLATON (Zhang et al., 2019) and the generic grid (Goyal et al., 2019a) are depicted in Figure 4.16. The impact of stellar activity on the transmission spectrum was also evaluated in an attempt to make the analysis more thorough. Using very wide ranges for the temperatures and stellar surface fractional coverages of the magnetically-induced starspots, the forward mod-

els were found to produce an excellent fit to the transmission spectrum. The result, however, was not to be trusted as the outcome merely advocated the presence of widespread spot activity covering most of the stellar disk. This solution was deemed to be unphysical given our knowledge of the system (see Section 4.1.1) and was therefore not considered as proof for WASP-88's activity. For this reason, such a fit was not included in Figure 4.16.

4.5.3 AURA

A retrieval analysis of the observed transmission spectrum was ultimately also performed by taking advantage of the most recent update of the AURA retrieval software (Pinhas et al., 2018; Welbanks & Madhusudhan, 2019). This code pairs forward modelling and multi-modal nested sampling, a feature that is also offered with PLATON. However, AURA is slightly more versatile, providing the opportunity to explore elemental abundances in addition to atmospheric types. For the nested sampling procedure, the PyMultiNest package is employed (Buchner et al., 2014). This python implementation of MultiNest (Feroz et al., 2009) not only retrieves parameters, but also compares models using Bayesian statistics. For the forward models, AURA performs a line-by-line radiative transfer computation and handles the day-night transition zone at the limbs of the exoplanet as plane-parallel. Furthermore, the code assumes that the atmosphere is in hydrostatic equilibrium and that atomic and molecular species follow a uniform distribution as a function of altitude.

For WASP-88, a variety of chemical species were considered but the most prominent opacity sources were H_2O, Na and K, as well as the H_2 pressure broadening of the Na and K cores (Rothman et al., 2010; Welbanks et al., 2019). In addition, opacity contributions from H_2-H_2 and H_2-He collision-induced absorption were also taken into account. Apart from a clear atmosphere, the effects of clouds and hazes were also investigated, with clouds being simulated as grey opacity and hazes being included in the function for the cross-section $\sigma = a\sigma_0(\lambda/\lambda_0)^\gamma$ that gives the opacity due to scattering above the cloud deck. In this function, σ_0 is equal to 5.31×10^{-27} cm^2 and yields the

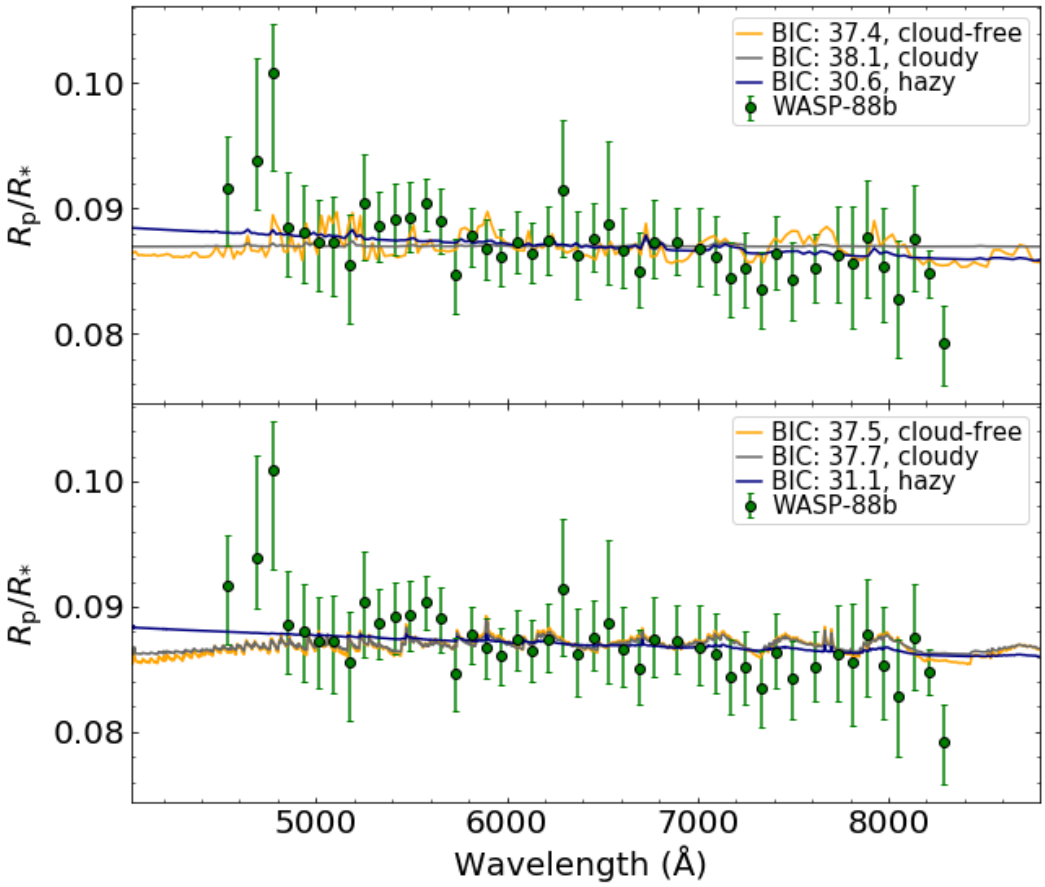

Figure 4.16: Transmission spectrum of WASP-88b and simulated atmosphere fits based on forward models from PLATON (top, Zhang et al., 2019) and the generic grid (bottom, Goyal et al., 2019a). Three distinct atmospheric scenarios are contemplated (clear, cloudy and hazy) with the synthetic spectra of simulated atmospheres showing a slight preference for the hazy scenario, even though both cloudy and cloud-free atmospheres generate reasonable fits.

H_2 Rayleigh scattering cross-section at $\lambda_0 = 350$ nm. a and γ are free variables with γ providing the value for the scattering slope. The temperature-pressure profile at the day-night terminator region is characterised by six parameters following the recipe outlined in Madhusudhan & Seager (2009). Altogether, in the best-fit scenario, the atmosphere of WASP-88b is described by a total of 14 free variables, of which six express the temperature-pressure profile, four explain the contribution from clouds and hazes, as well as their fractional coverage, three give the volume mixing ratios of H_2O, Na and K, and one provides the reference pressure for the estimated radius of the exoplanet.

Broad uniform priors were applied to five atmospheric parameters, whereas log-uniform priors were imposed on the nine remaining parameters. More particularly, a uniform prior in the range between -20 and 2 was used for the scattering slope γ, and in the range between 0 and 1 for the cloud and haze coverage fraction. The temperature at the top of the gaseous envelope was limited in the range between 800 and 2000 K to eliminate unreal atmospheric parameter values whilst simultaneously allowing for the examination of a large range of temperature-pressure profiles. Parameters a_1 and a_2 of the temperature-pressure profiles were also restricted uniformly in the range between 0 and 2. At the same time, relative abundances for H_2O, Na and K, followed a log-uniform distribution spanning from 10^{-12} up to 10^{-1}. Pressure parameters P_1, P_2, P_c and P_{ref} had log-uniform priors covering the range from 10^{-6} to 10^2, while P_3 was more constrained in the range between 10^{-2} and 10^2. Finally, haze parameter a had a log-uniform prior ranging from 10^{-4} to 10^{10}.

Just as in the forward models with the generic grid and PLATON, the retrievals recovered a featureless transmission spectrum for WASP-88b with a steep scattering slope that can best be interpreted by the existence of significant high-altitude haze in WASP-88b's atmosphere. Specifically, the scattering slope was found to be $\gamma = -14.42^{+4.88}_{-3.64}$, while the scattering factor was retrieved as $\log(a) = 5.53^{+1.56}_{-1.91}$. In addition, the abundances for the key atmospheric constituents H_2O, Na and K, could not be constrained. The median transmission spectrum, along with the 1σ and 2σ uncertainty regions, is illustrated in Figure 4.17, and the marginalised posterior distributions of the 14 free parameters are presented in Figure 4.18.

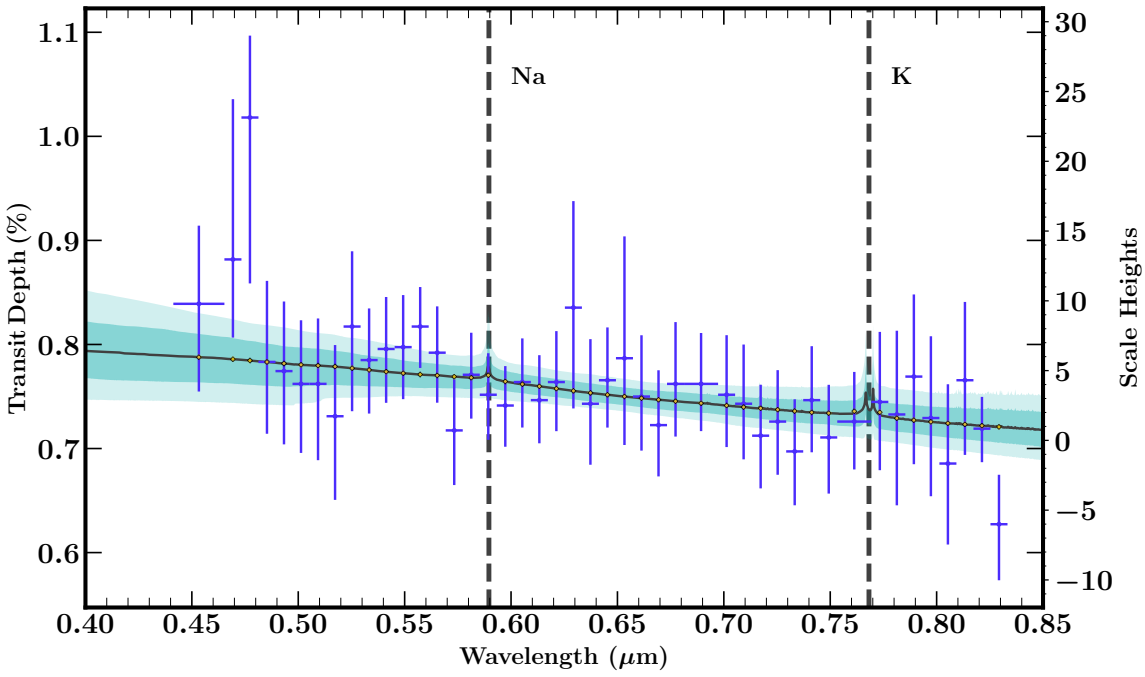

Figure 4.17: Retrieved transmission spectrum of WASP-88b with 1σ (dark turquoise) and 2σ (light turquoise) contours. The expected position of the Na and K cores is depicted by grey vertical dashed lines.

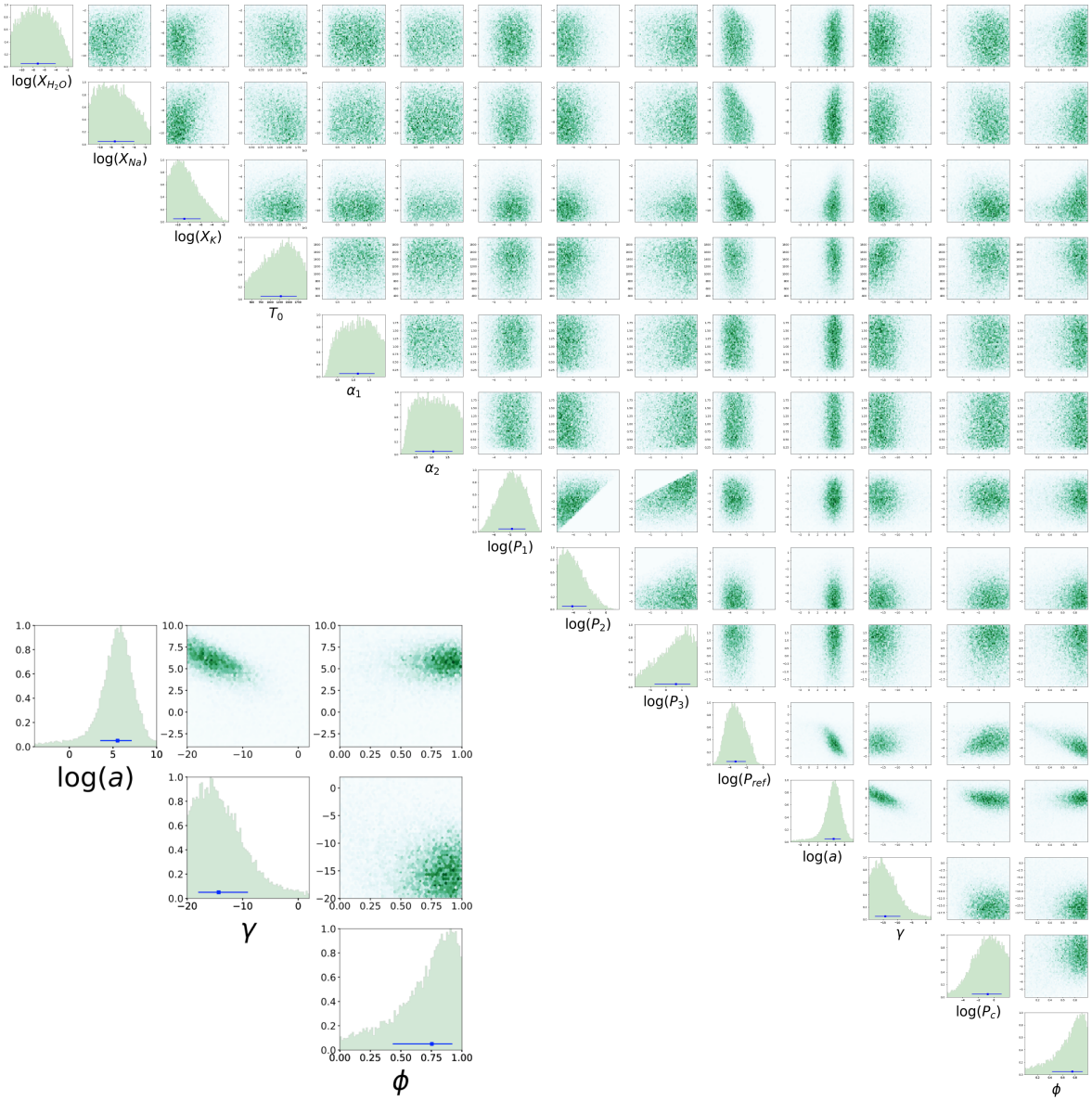

Figure 4.18: Results from the WASP-88b retrieval analysis using AURA. The marginalised posterior distributions of the atmospheric parameters from the preferred model are shown on the upper right corner together with their histograms. The abundances of H_2O, Na and K are unconstrained, whereas the haze parameters a and γ are constrained. On the bottom left, we zoom in on the two haze parameters and the fractional cloud and haze coverage ϕ for clarity reasons.

A second atmospheric retrieval was also executed to measure the significance of this detection using Bayesian model selection. The second retrieval assumed the high-altitude presence of clouds and ignored haze by fixing the two haze parameters to the values expected from H_2 Rayleigh scattering. Therefore, by setting the two haze parameters at $a = 1$ and $\gamma = -4$, this model effectively considered a cloudy atmosphere and resulted in a flat transmission spectrum for the observed data. The model comparison yielded a marginal preference for a hazy sky, revealing a detection significance of 2.5 σ over a sky obscured by clouds. This finding is very similar to the forward model-generated results in Sections 4.5.1 and 4.5.2, suggesting that even though enhanced haze due to scattering is very probable, a sky filled with high-altitude clouds cannot not be decisively ruled out.

Other chemical species predicted to be abundant at high temperatures, such as TiO, VO and AlO, had no significant impact on the obtained transmission spectrum and were thus excluded from the best-fit retrieval. These metal oxides could potentially show strong absorption signals in the optical, something that cannot be seen here. Additional retrievals checking the presence of various metal hydrides were also examined but had a negligible effect on the observed transmission spectrum. The volume mixing ratios of the three chemical species considered in the most favoured retrieval analysis were hugely unrestricted, with H_2O and Na reaching 2σ upper-limits of $\sim 10^{-2}$, and K approaching a 2σ upper abundance limit of $\sim 10^{-4}$. At the end, another retrieval was conducted to explore the prospect of stellar activity affecting our data. The AURA package includes subroutines that can investigate the likelihood of stellar heterogeneity producing the enhanced slope seen in the data, including models that simulate the behaviour of starspots and faculae Pinhas et al. (2018). The code was recently tested on the atmosphere of WASP-110b showing the presence of faculae in the transmission spectrum of this exoplanet (Nikolov et al., 2021). No evidence was found to support the idea of a heterogeneous stellar disc influencing our results and so we conclude that the star is likely inactive.

4.6 Conclusion

Observations in the optical and the near-infrared enabled the detection of the atmosphere of WASP-88b for the first time. This hot and low-gravity transiting exoplanet was observed from both ground and space and an in-depth analysis uncovered the properties of the system as well as details about the structure of its atmosphere. More specifically, space-based surveillance with TESS provided a refined ephemeris and revised estimates for the system parameters, which were found to be in good agreement with the values reported in Delrez et al. (2014). The updated system parameters give a slightly increased planet-to-star radius ratio as a result of lower stellar and planetary radii. The result is supported by ground-based transmission spectroscopy with VLT FORS2 that probes the optical regime in the wavelength range between 4413 and 8333 Å. The measured transmission spectrum reveals no absorption signatures from known elements with signals in the optical and a steep scattering slope towards bluer wavelengths. The featureless shape of the transmission spectrum was evaluated with forward modelling and retrieval methods and the verdict was unanimous: WASP-88b has an atmosphere that is likely obscured by its constituents, producing enhanced haze at high-altitudes. This scenario is tentatively preferred over other scenarios and additional observations are required to reject the cloudy and cloud-free hypotheses.

Furthermore, there are strong indications that WASP-88 is inactive and thus stellar activity cannot be the cause behind the observed slope in the transmission spectrum. Photometric monitoring of the host star and a stellar spectrum of the Calcium H and K lines show no evidence of strong photospheric and chromospheric activity indicating that WASP-88 is in a quiet phase of its evolution. This conclusion is endorsed from retrievals with AURA that also do not detect any signs of increased stellar activity. It is therefore relatively safe to attribute the steep slope to intrinsic phenomena within the planetary atmosphere and not to the external influence from the host star.

In Chapter 6, we will briefly take a deeper look into WASP-88b in an attempt to provide an interpretation for the measured shape of the transmission spectrum and to

understand the physical phenomena that may occur within its atmosphere. There we will also discuss the importance of additional observations in getting a better idea of the planetary atmosphere by exploring a broader wavelength range. Only a complete picture of the atmosphere will eliminate doubts and decisively show what WASP-88b's atmosphere is made of.

5 The blue-optical transmission spectrum of the hot Jupiter WASP-74b

The observations and most of the analysis presented in this chapter were submitted for publication in the Monthly Notices of the Royal Astronomical Society (Spyratos et al., 2023). The VLT FORS2 analyses as well as the atmospheric retrievals with PLATON were performed by me. Others helped by implementing the observations (Elyar Sedaghati), the initial reductions (Nikolay K. Nikolov), and the additional atmospheric retrievals with AURA (Savvas Constantinou). John Southworth provided useful comments throughout the project as my supervisor. The contribution from others is included for context.

5.1 Introduction

The WASP-74 system is the second planet-hosting system I investigated as part of my thesis. This time the system was only observed from the ground with the clear objective to study its atmosphere at short optical wavelengths. Contrary to WASP-88b's atmosphere, which was observed for the first time, WASP-74b's atmosphere had already been examined by other studies during both transit and occultation. This means that the properties of the system are known to greater precision and, thus, a ground-based exploration would be based on stronger foundations. In this chapter, I will focus on the various ways the new ground-based transit observations were analysed, introduce successful solutions and unravel the characteristics of the obtained WASP-74b transmission spectrum. Before doing that, however, I will present the motivation behind these observations, which differs compared to the WASP-88b investigation, and outline the properties of the system as determined from previous studies.

In many ways, WASP-74b is quite similar to WASP-88b, as it is orbiting close to its host star and has an inflated radius coupled with a hot atmosphere (Hellier et al.,

2015). These are some of the main characteristics that inspired subsequent observations of this gas giant exoplanet in an attempt to detect its atmosphere. One of the first atmospheric studies of WASP-74b was by Mancini et al. (2019), who performed transmission spectroscopy on photometric data obtained from the ground. Their analysis was supported with space-based spectroscopic data from HST (Tsiaras et al., 2018). While their measurements revealed fairly normal transit light curves in the red-optical, they also unexpectedly found some anomalous readings in the U-band, which explores the near-ultraviolet region. More specifically, they detected a deeper transit that was almost twice as deep as the transits in redder passbands, and a longer transit duration by about 0.4 hours. Such a finding could be indicative of an escaping atmosphere caused by strong irradiation from its stellar companion (see e.g. Allan & Vidotto, 2019, for theoretical predictions about the magnitude of this effect in the Ly-α and H-α absorption regions for a wide range of simulated exoplanets).

The phenomenon of atmospheric evaporation is not rare among hot exoplanets and a few studies have already detected such signatures in the entire regime between Neptune-sized and Jupiter-sized exoplanets (e.g. Vidal-Madjar et al., 2003, 2004; Lecavelier Des Etangs et al., 2010; Kulow et al., 2014; Lavie et al., 2017; Bourrier et al., 2018; Spake et al., 2018; dos Santos et al., 2019; Ninan et al., 2020). One notable example is the observation of a comet-like hydrogen tail in the evaporating atmosphere of GJ 436b (Ehrenreich et al., 2015). The study of this hot Neptune reported a longer transit duration by nearly 5 hours in the near-ultraviolet and a five-fold increase in the transit depth compared to the corresponding value in the optical. According to Ehrenreich et al. (2015), this effect translates into a mass loss rate of 10^8 to $10^9\,\mathrm{g\,s^{-1}}$. This mass loss rate may seem large at first, but is insignificant compared to the exoplanet's lifetime.

The observation of a longer and deeper transit in the U-band for WASP-74b was therefore the stimulus for a spectroscopic investigation of the blue-optical region in an effort to study the potential signature of mass loss. This was the first time such a signal was observed in a spectral region that is so close to the bluest optical wavelengths. If this distinctive transit signature was confirmed in the blue, it would

have major implications on the computation of the mass loss rate and heating efficiency (e.g. Ehrenreich & Désert, 2011), as well as on the evolution history of this exoplanet (e.g. Kurokawa & Nakamoto, 2014; Jin et al., 2014).

Follow-up observations from ground and space, however, failed to detect a significant variation in the transit depth with respect to wavelength that would be indicative of atmospheric escape (Luque et al., 2020; Fu et al., 2021). The same studies also reported contradictory results about the atmospheric composition of WASP-74b. These results not only disagreed between themselves but also disputed the findings by Mancini et al. (2019). Therefore, the exact properties of WASP-74b's atmosphere are still unclear and this predicament highlights the importance of this additional study.

In terms of whether WASP-74b experiences atmospheric evaporation or not, the blue-optical transmission spectrum will be able to provide answers for both situations. More specifically, in the scenario where the shape of the transit changes drastically towards bluer regions due to mass loss, the transmission spectrum will integrate these effects through the inclusion of particle simulations in the model fits (e.g. Ehrenreich & Désert, 2011; Ehrenreich et al., 2015). On the other hand, in the event where the imprints of mass loss cannot be observed in the transit light curves, the transmission spectrum can be examined following the standard techniques already described in previous chapters of this thesis.

Regardless of this investigation's outcome, WASP-74b will also provide a legacy transmission spectrum of tremendous significance to population studies. Together with the atmosphere of WASP-88b and the atmospheres of other exoplanets already observed or with plans to be observed in transmission, WASP-74b's atmosphere will help illuminate the processes that regulate the structure and composition of hot gas giant exoplanetary atmospheres. Part of this combined effort will also allow for a better understanding of the mechanisms that form clouds and hazes in the atmospheres of hot gaseous planets. So far, the first population studies based on small samples have shown that a degree of cloudiness can be observed ubiquitously among exoplanetary atmospheres (Sing et al., 2016; Pinhas et al., 2019).

The following subsection will give more details about the WASP-74 system char-

acteristics and will discuss the findings from previous atmospheric studies. This step is important before we proceed to the data analysis and results from this study.

5.1.1 The WASP-74 system in detail

The WASP-74 system has a right ascension of 20:18:09.32 and a declination of −01:04:32.6 and was observed from the WASP-South observatory in Sutherland, South Africa, as part of the SuperWASP survey (Pollacco et al., 2006). The discovery of the transiting exoplanet was announced by Hellier et al. (2015), who also carried out a follow-up investigation using the La Silla-based 0.6 m TRAPPIST telescope (Jehin et al., 2011) and the 1.2 m Euler-Swiss telescope (Lendl et al., 2012). The report also included one additional transit acquired using the RISE imager mounted on the 2.0 m Liverpool Telescope (Steele et al., 2008). In addition, the CORALIE spectrograph was once again employed to calculate the radial velocities and compute the stellar parameters. These stellar quantities were recently updated by Sousa et al. (2021) and were included in the SWEET-Cat catalogue of planet-hosting stars[1]. Furthermore, the target star was part of Gaia Data Release 3 and so a few system parameters were determined to a very high precision (Gaia Collaboration et al., 2016, 2022b; Babusiaux et al., 2022). Other studies also provided revised solutions for a wide range of orbital, planetary and stellar characteristics (Mancini et al., 2019; Garhart et al., 2020; Luque et al., 2020; Ivshina & Winn, 2022; Baştürk et al., 2022). Table 5.1 gives a summary of the derived light curve parameters and physical properties of the system. The host star is a main-sequence dwarf of spectral type F9 with a brightness of $V = 9.7$. As can be seen from Table 5.1, WASP-74 is about the same age as WASP-88, indicating that hydrogen is still burning in its core. At the same time, the orbiting exoplanet is a close-in, hot Jupiter with a moderately bloated radius that gives a pressure scale height of 670 km. This value roughly translates to an atmospheric signal of $\Delta\delta = 2HR_{\rm p}/R_*^2 \approx 120$ ppm (Winn, 2010), which is slightly lower compared to WASP-88b.

[1] `http://sweetcat.iastro.pt/`

Table 5.1: The WASP-74 system characteristics based on other works.

Parameter	Value	Reference
Distance (pc)	148.75 ± 0.39	Bailer-Jones et al. (2021)
Impact parameter $b = a \cos i/\mathrm{R}_*$	$0.8460^{+0.0031}_{-0.0033}$	Baştürk et al. (2022)
Time of mid-transit t_0 ($\mathrm{BJD_{TDB}}$)	2457173.871443 ± 0.000080	Baştürk et al. (2022)
Orbital period P (d)	$2.13775132^{+0.00000053}_{-0.00000055}$	Baştürk et al. (2022)
Effective temperature T_{eff} (K)	6006 ± 23	Sousa et al. (2021)
Metallicity [Fe/H]	0.44 ± 0.02	Sousa et al. (2021)
System Age (Gyr)	$3.43^{+1.00}_{-0.82}$	Baştürk et al. (2022)
Stellar density ρ_* ($\rho_\odot$)	0.472 ± 0.011	Baştürk et al. (2022)
Stellar surface gravity log g_* (c.g.s.)	$4.1608^{+0.0088}_{-0.0091}$	Baştürk et al. (2022)
Stellar mass M_* ($\mathrm{M}_\odot$)	$1.316^{+0.052}_{-0.053}$	Baştürk et al. (2022)
Stellar radius R_* ($\mathrm{R}_\odot$)	$1.578^{+0.024}_{-0.025}$	Baştürk et al. (2022)
Planet/star radius ratio $k = R_\mathrm{p}/R_*$	0.09161 ± 0.00028	Baştürk et al. (2022)
Scaled semi-major axis a/R_*	$4.849^{+0.038}_{-0.037}$	Baştürk et al. (2022)
Semi-major axis a (AU)	$0.03559^{+0.00046}_{-0.00049}$	Baştürk et al. (2022)
Inclination i (degrees)	$79.95^{+0.12}_{-0.11}$	Baştürk et al. (2022)
Eccentricity e	0 (fixed)	Hellier et al. (2015)
Planet mean density ρ_p (ρ_Jup)	0.25 ± 0.02	Hellier et al. (2015)
Planet surface gravity log g_p (c.g.s.)	2.95 ± 0.02	Hellier et al. (2015)
Planet mass M_p ($\mathrm{M_{Jup}}$)	$0.877^{+0.037}_{-0.038}$	Baştürk et al. (2022)
Planet radius R_p ($\mathrm{R_{Jup}}$)	$1.407^{+0.023}_{-0.024}$	Baştürk et al. (2022)
Planet equilibrium temperature T_eq (K)	1947 ± 16	Baştürk et al. (2022)

The transit ephemeris was estimated and refined by multiple studies (Hellier et al., 2015; Mancini et al., 2019; Garhart et al., 2020; Luque et al., 2020; Ivshina & Winn, 2022; Başt ürk et al., 2022). In Table 5.1, I present the ephemeris from Başt ürk et al. (2022) as they provide the most recent calculation on the orbital period of WASP-74b based on 42 transit light curves. The same study also explored the system for transit timing variations using newly obtained data, as well as data from published literature and open databases. They concluded that the time of mid-transit is unaffected by systematic trends such as periodic or quadratic variations. This outcome suggests that WASP-74b is less likely to have a decaying orbit.

Interestingly, all studies after the initial discovery (Hellier et al., 2015) found slightly lower values for the transit depth. For example, Garhart et al. (2020), who observed the system during transit and eclipse using the Spitzer 3.6 and 4.5 μm pass-bands, found a transit depth difference of about 5%. They attributed this discrepancy to the nearly grazing transit and the effect this may have on the limb darkening coef-ficients. As one may notice in Table 5.1, the measured value for the impact parameter is quite high ($\sim$0.85).

The first analysis of the atmosphere was performed by Tsiaras et al. (2018) as part of their population study that targeted the near-infrared H_2O feature at 1.4 μm. Their space-based spectroscopic analysis revealed an effectively flat transmission spectrum indicating that the exoplanetary atmosphere is either water-depleted or obscured by clouds. More importantly, their newly created detection significance metric, dubbed the Atmospheric Detectability Index, returned a value of 0, suggesting that the atmosphere of WASP-74 was not detected due to clouds. However, the same study also notes that, despite this result, it may still be possible to constrain the atmospheric parameters to a great degree with additional observations and to possibly detect the presence of water.

A simultaneous system characterisation and atmospheric investigation was then performed by Mancini et al. (2019). They found smaller values for the planetary and stellar radii, and for the planetary and stellar masses, compared to Hellier et al. (2015). Their transmission spectrum was produced based on photometry from the ground and

included the re-analysed HST near-infrared observations from Tsiaras et al. (2018). Apart from the already discussed anomalous findings in the U-band, they also detected significant absorption in the optical bands. Forward models showed that the inverted U shape of the transmission spectrum could best be described by the presence of strong metal absorbers in the atmosphere, such as TiO and VO.

The WASP-74 system was then again observed from the ground by Luque et al. (2020). Just as Mancini et al. (2019), Luque et al. (2020) were able to obtain additional high-resolution spectra to get the stellar parameters and fully characterise the system. The obliquity of the exoplanet was also explored via the Rossiter-McLaughlin effect on the radial velocity data. This inspection revealed an aligned orbit for WASP-74b. At the same time, the transmission spectrum was constructed based on a re-analysis of the published data and the inclusion of new, ground-based photometry. The transmission spectrum was found to display a significant slope towards bluer wavelengths. A linear scattering fit showed that the slope is about five times steeper than Rayleigh scattering, something that indicates additional scattering phenomena within the planetary atmosphere leading to enhanced haze.

Another transmission spectrum was later generated by Fu et al. (2021), who combined space-based photometry and spectroscopy. This study managed to observe the exoplanet in both transmission and emission. Their results could not reproduce earlier findings in the optical (Mancini et al., 2019; Luque et al., 2020), as well as the highly increased planet-to-star radius ratio found in the U-band (Mancini et al., 2019). The transmission spectrum exhibited an overall featureless shape revealing a largely cloudy atmosphere, in agreement with previous space-based results (Tsiaras et al., 2018). Furthermore, a strong Rayleigh scattering signal that stretches from the near-ultraviolet up to the near-infrared was also established.

High-resolution transmission spectroscopy by Lira-Barria et al. (2022) was used to scrutinise the atmosphere of WASP-74b in a search for various elements. Their study was only able to report the tentative detection of aluminium after inspecting the four strongest observable lines of this element in the optical. Only one of them was found to produce a strong signal, giving a detection significance of 5.6σ. The results from

this study are further evidence that the atmosphere of WASP-74b is highly concealed by processes that, so far, are not well-understood.

Finally, a number of population studies have also performed independent analyses of the space-based observations reported by Tsiaras et al. (2018) and Fu et al. (2021). These studies performed atmospheric retrievals on either the transit light curves (Fisher & Heng, 2018) or the eclipse light curves (Mansfield et al., 2021; Changeat et al., 2022). Their atmospheric simulations largely reproduced the results from the original space-based investigations and showed that the observations from space favour a cloudy sky for WASP-74b.

In view of the fact that some previous observations found increased absorption in the shorter optical wavelengths, it should be emphasized that WASP-74 shows no signs of magnetic activity. Long term photometric surveillance of the stellar light curves obtained with WASP found no rotational modulation above 0.7 mmag (Hellier et al., 2015). Similar findings were presented by Fu et al. (2021), who monitored the star photometrically from June 2018 to June 2021 using a 14-inch Celestron telescope based at Fairborn Observatory in Arizona, USA. Their work found no significant periodicities and largely invariable light curves, confirming that the star is inactive. Furthermore, a work in progress that aims to map stellar activity in a high number of planet-hosting stars through spectroscopic observations of their chromosphere also returned a negative result for WASP-74 (Southworth, unpublished). Specifically, the study found that the calcium H and K lines showed no traces of chromospheric emission, in support of the previous photometric results that hinted at a magnetically quiet star. All these pieces of data demonstrate the inactivity of WASP-74, something that is shared in common with WASP-88. Therefore, the shape of the transmission spectrum should not be affected by the atmosphere of the host star.

Despite the many attempts at deciphering the atmospheric properties of WASP-74b, the current findings are controversial with most results pointing to a high fractional coverage of cloud/haze in the upper atmosphere. Furthermore, none of these investigations includes transit spectroscopy observations from the ground. This study will produce the first transmission spectrum based on such observations and will, hence,

give a fresh and detailed insight into the enigmatic atmosphere of WASP-74b.

5.2 Observations and Data Reduction

The observations for this target were performed as part of ESO programme 0101.C-0716 with John Southworth being the principal investigator. A total of three nights were allocated to our research team to observe the WASP-74 system. Just like the observations for the WASP-88 system, the data here were obtained by taking advantage of the ESO telescope facilities at Paranal, in Chile. Again, the Antu Telescope from the VLT array was utilised to perform spectroscopy with the FORS2 instrument. The data were then saved on the ESO archive, where they are freely available now.

Due to the slightly different purpose of this investigation compared to the one performed for WASP-88b, the data collected for this target covered a narrower region of the electromagnetic spectrum. This smaller spectral region was explored by employing the dispersive element GRIS600B on all three nights. This grating is sensitive to the blue, having a central wavelength of 465 nm. Furthermore, it can contain the wavelengths between 330 and 621 nm, with a dispersion of 50 Å/mm. A complete transit was observed in all three nights, so, in theory, the observations were deemed to be successful. In practice, the atmospheric conditions during the observations, as well as instrumental problems, made the analysis of this target quite challenging.

The target and three additional reference stars were monitored throughout the nights and their spectra were recorded on CCD cameras using the multi-object spectroscopy mode (MXU) of FORS2. In particular, the spectrum of the brightest reference star was collected on one CCD chip, and the target with the fainter reference stars were collected on the other CCD chip. Contrary to the WASP-88 observations, the CCDs here were part of the blue detector (EEV) of FORS2, which is expected to work better at shorter optical wavelengths. To observe the spectra of the stars, a customised mask with moderately wide slits ($22'' \times 22''$) was created. In addition, the operating cycle was set to a binning of 2×2 and the readout time to $\sim$30 s. This was done in order to

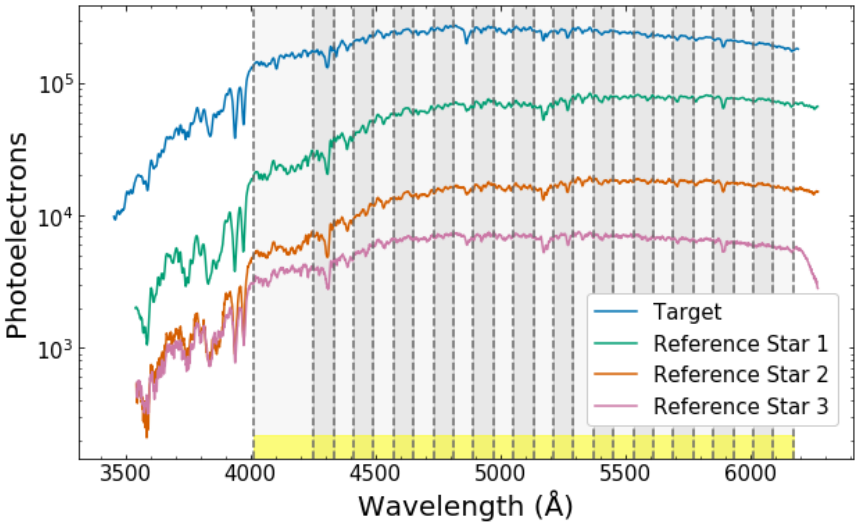

Figure 5.1: Example spectra of WASP-74 (blue line) and three comparison stars (green, brown, and magenta lines). Yellow indicates the wavelength region used for the white light curves, while the shaded grey vertical bands indicate the wavelength regions used for the spectroscopic light curves.

optimise the observations by collecting the highest amount of data in the least amount of time. Figure 5.1 shows example spectra of the target and the three reference stars.

The first transit observation was made on May 19th 2018. These data contain one full transit, but are highly contaminated by cloudy conditions. The weather was less than ideal, with a good amount of clouds at higher altitudes affecting most of the out-of-transit data and also part of the transit. Seeing remained low and fluctuated between 0.3″ and 1.5″. The target and the three reference stars were followed from an airmass of 2.41 upward to an airmass of 1.09 and then downward to an airmass of 1.17. A total of 400 data points were accumulated with integration times between 20 and 45 s. This resulted in an observation that lasted 5 hours and 38 min.

Another entire transit was observed on June 20th 2018. This data set was acquired during improved weather conditions (sky was mostly clear) and the transit light curve was less corrupted by telluric effects. Nevertheless, the humidity was relatively high for the desert location (15-20%) and the integrated water vapour was almost three

times higher compared to the other two observations. Seeing was above 1″ at the beginning of the observation but improved as the night went on (range: 0.39″-1.67″). WASP-74 and the other stars were tracked from an airmass of 1.48 to an airmass of 1.09, and then to a final airmass of 1.50. This led to a data set consisting of 405 data points and assembled over a period of 5 hours and 45 min. These data points have exposure times between 16 and 70 s.

The final complete transit of WASP-74b was recorded during the last night of observations on August 19th 2018. This data set was characterised by a low precipitable water vapour ($\sim$ 1.5), very low humidity (<6%) and a seeing level that stayed below 1″ most of the night (0.3″-1.1″). This was indicative of a smaller amount of turbulence in the Earth's atmosphere. Still, the observer noted some occasional thin cirrus clouds passing over the field during the data acquisition process. The same four stars were monitored during their ascent from an airmass of 1.36 to an airmass of 1.09, and then during their descent to an airmass of 1.47. In total, 379 exposures were obtained for a narrow range of integration times (20-35 s). The observation came to a conclusion after 5 hours and 16 min of data collection.

Figure 5.2 shows the raw fluxes of the target for each data set. As can be seen from the first panel, the May data set displays the most severe systematics with an overall very poor quality of light curve behaviour. Thick and/or thin cirrus clouds have caused a great deal of corruption to the transit making fits quite demanding and any results unreliable. This time-series was therefore not included in most of the subsequent analyses. In addition, a small, 7 min gap in the data of the second observation can also be identified in panel two. This is likely caused by instrumental effects such as a guiding failure or a minor software crash. All data after this gap were removed as they were out-of-transit and the data set already contained enough information from before, during and after the planetary transit. A single outlier in the last data set was also discarded from the rest of the investigation. This effectively reduced the valid exposures to 289, 333 and 378 for each data set respectively.

In the same way as for the WASP-88 system, the customised IDL pipeline from our collaborator Nikolay K. Nikolov produced an assortment of decorrelation param-

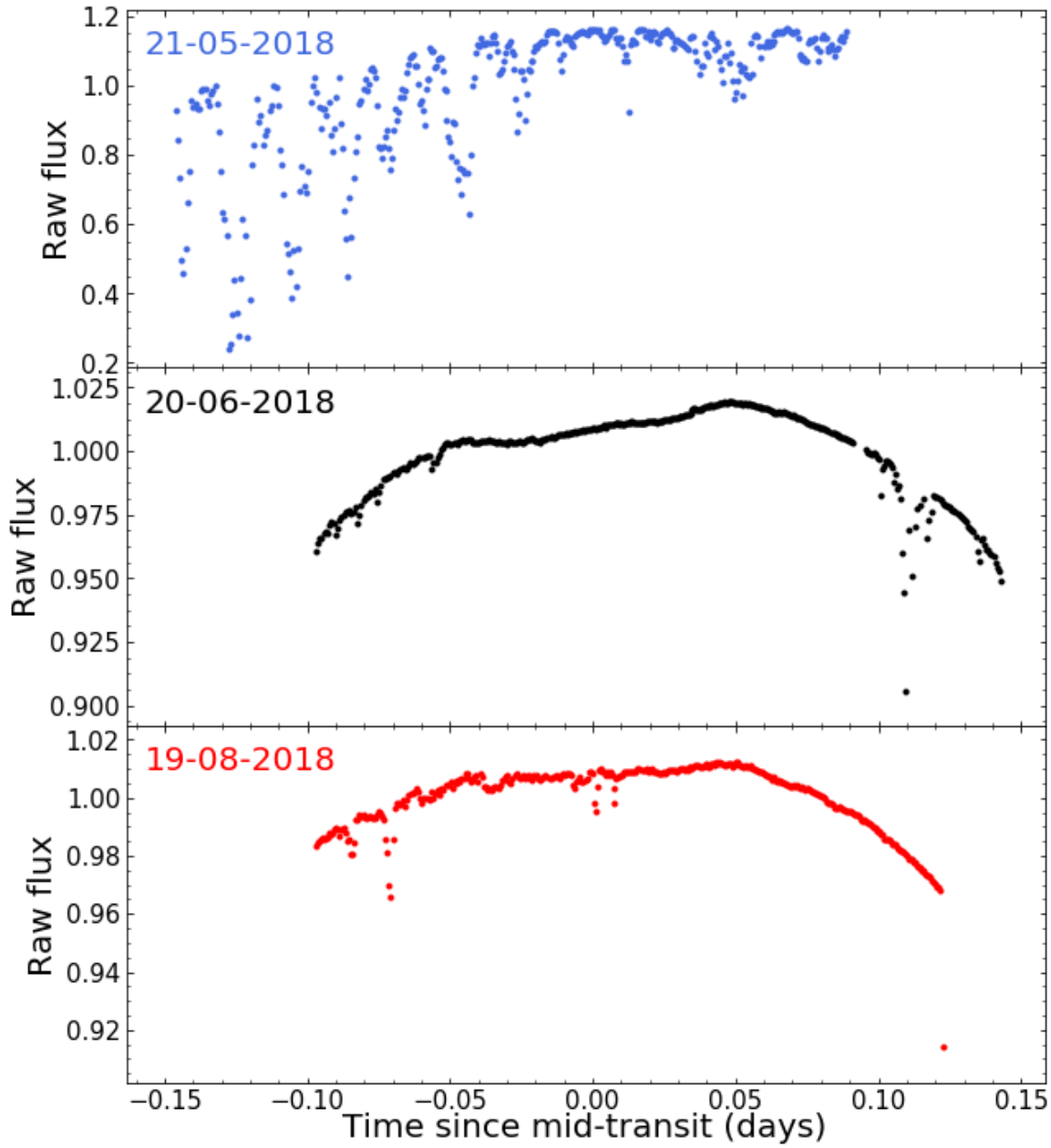

Figure 5.2: The white flux of WASP-74 from all three observations. The first data set (sky blue dots, top) reveals heavy corruption from thin and thick cirrus clouds, the second data set (black dots, middle) displays cleaner data obtained under better weather conditions, and the third data set (red dots, bottom) shows a thin cirrus cloud crossing around the transit midpoint. A moderate cloud crosses the field after the 7-minute gap in the second observation, but this is inconsequential as these data were removed. The outlier at the end of the third observation was also removed.

eters and limb darkening coefficients. In addition, relative fluxes were constructed between the target and the three reference stars within the field of view[2]. The main difference with WASP-88 was that now there were more than one comparison stars available. This resulted in a set of relative flux time series between the target and each comparison star individually or in combination (with and without weights).

To get the best relative flux time series, a systematic analysis was implemented. Unfortunately, it was soon discovered that, during the June observation, the second CCD chip failed to collect data properly. Figure 5.3 illustrates the massive impact this had on the brightest comparison star (2MASS J20181210−0107143) in both the raw and the relative light curve, where the transit effectively disappeared in all the noise. Due to this problem, the brightest comparison star was excluded from the rest of the study. The out-of-transit examination of the relative fluxes eventually showed that a relative flux light curve that consists of the target and the second brightest reference star (2MASS J20180844−0102001) generates the smallest amount of scatter. This reference star was therefore selected to correct for any atmospheric phenomena in subsequent relative flux analyses.

To isolate the spectra from the background, the procedure was essentially similar to the one followed for WASP-88. The APALL subroutine from the IRAF software was again employed to obtain the spectral apertures through summation. Then the rest of the data reduction was carried out using the custom-made IDL code first presented in Chapter 4. An aperture radius of 15 pixels was defined for all time-series as the radius that decreases the out-of-transit scatter the most. In addition, sky background regions 25 to 30 pixels away from the peak of the spectral profiles were specified on each side of the spectra. Then, the median value of this background noise was estimated and subtracted from the integrated spectral apertures.

For the wavelength solution, the process was nearly identical to the one described for the WASP-88 system. Once more an emission lamp was used to extract

[2]The 2MASS identifiers for these stars are 2MASS J20181210−0107143, 2MASS J20180844−0102001 and 2MASS J20180810−0101305.

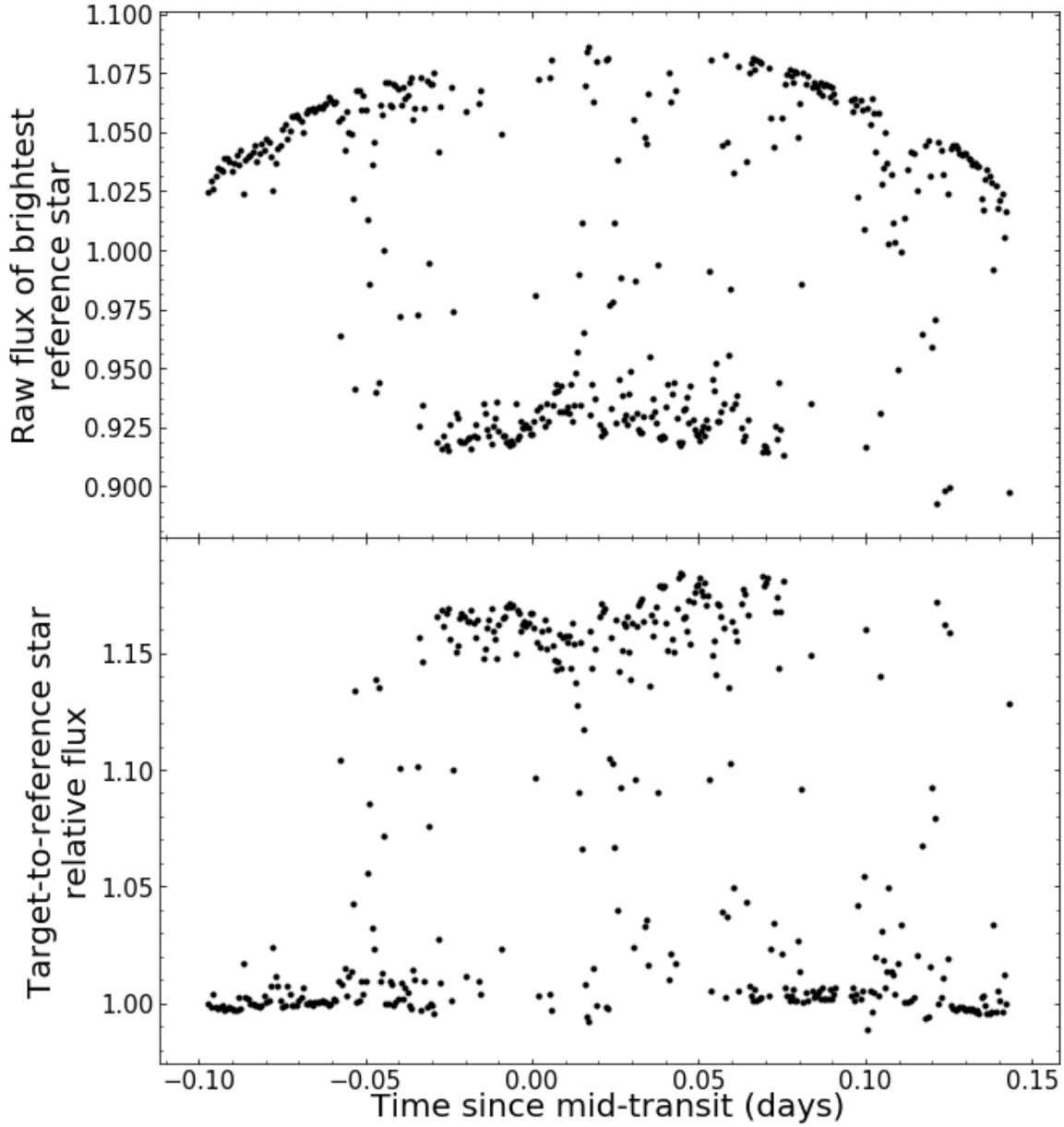

Figure 5.3: The comparison star flux from June 20th 2018. During this observation the second CCD chip failed to collect data properly and as a result the recording consists of data points that are scattered all around. This is evident in both the raw white flux of the brightest comparison star (top), as well as in the target-to-reference star relative white flux (bottom). The cause of this systematic effect is unknown.

several spectral lines using a narrow slit. Then, to get the central wavelengths of the most prominent lines, the spectra were fitted with Gaussians and low-order Chebyshev polynomials in the horizontal and vertical axes of the images. The end product was then compared against a reference rest frame and corrected for at the sub-pixel level. The only difference versus WASP-88 was that now calibration had to be implemented for two additional faint stars. This was actually fairly trivial as their position could easily be identified in the image.

As a final step in the data reduction process, we produced differential spectrophotometric light curves for each data set spanning over a wavelength region between 4013 and 6173 Å. More specifically, we created 25 narrow spectroscopic bins of 80 Å in width to explore the planet-to-star radius ratio as a function of wavelength. These spectroscopic channels are in accordance with the grouping criteria followed in previous FORS2 observations using the same pipeline (Nikolov et al., 2018b). We adopted a wider width of 240 Å for the wavelengths between 4013 and 4253 Å to compensate for the lower signal-to-noise ratio at these wavelengths. We then obtained the differential light curve for each bin by dividing the summed flux of the target by the summed flux of the second brightest reference star. We also constructed differential light curves for each data set from the combined light of the entire wavelength range investigated. These "white" light curves are important for the identification of the common mode factors that will need to be removed during the spectroscopic analysis (see Section 5.3).

Upon inspection of the spectroscopic light curves, a deeper transit was not observed for any of the light curves. This observation is a first sign that the atmosphere of WASP-74b does not experience strong evaporation and that the U-band data presented in Mancini et al. (2019) are likely a result of unaccounted for systematics. Therefore, a decision was made to analyse the data according to procedures that do not take mass loss into account. This approach reflects all previous atmospheric investigations of WASP-74b.

5.3 Data Analysis

In the same way as in the WASP-88b analysis, the approach here was two-fold. First, we analysed the transit from the combined (white) light time series and then we used the computed common mode to estimate the transmission spectrum from the spectrophotometric light curves. In the case of WASP-74b, however, the use of fits relying on baseline polynomial decorrelation of the systematics contributions will not be presented. Such an investigation, as well as its combined version with GPs, was explored but was deemed to be redundant and not useful for our purposes. Instead, the analysis section will focus on the more reliable and conservative approach involving GPs.

The original GP investigation of the spectroscopic time series revealed enlarged uncertainties for the transit parameters mostly due to the high residual scatter of the spectroscopic light curves. The spectroscopic analysis was therefore split into several parts as we examined a sequence of different techniques in our aim to improve the precision of the measurements. These techniques involved a global approach of the relative spectrophotometric light curves, as well as solutions that do not require the use of a comparison star, including one novel approach. These methods were found to either have no significant effect on the produced transmission spectrum or provided varying levels of improvement.

Contrary to WASP-88b, WASP-74b is relatively well-studied with its atmosphere being observed from both the ground and space. Therefore, WASP-74b is a good target for atmospheric comparisons between different data sets. Despite the great data availability, however, the analysis concentrated on the VLT data (from this study) and the broad transmission spectrum obtained from space (Fu et al., 2021) due to its higher number of data points, its broader coverage and its increased precision. The aim was to add confidence to the ground-based observations whilst also keeping a GP model setup that allows for direct comparisons with the published space-based results.

Due to the implementation of a novel method in the spectroscopic analysis, it was important to check the reliability of this method by applying the fits to a known published data set. This would give a certain degree of confidence about the robustness

of the result and would add value to the method if this technique could reproduce the published outcome within a certain margin of error. For this purpose, WASP-39b was chosen as the ideal target due to the tentative to moderate evidence for the presence of sodium in its atmosphere and the additional weak evidence for potassium. Both alkali metals show signatures in the wavelength range covered by VLT FORS2 (Nikolov et al., 2016). As we will see further on, the new method relies on a few innovations that differentiate it from the published analysis quite considerably and so the implications for ground-based transmission spectroscopy would be significant in the event of a very good agreement being established. A small part of Section 5.4 will therefore be dedicated to WASP-39b.

5.3.1 White light curves

The white light curve analysis focused on the June and August data sets, but the May data set was also inspected briefly. Due to adverse weather affecting the May observations more severely, this data set did not make the final cut into the submitted paper. Nevertheless, white and spectroscopic light curve fits were performed and a result will be shown in Section 5.4 to observe how thick clouds can affect the shape and the transmission spectrum.

Unlike May, the June and August observations were less significantly affected by the conditions in the Earth's atmosphere during the time of the observations. The white light curves show a relatively higher stability in the in-and-out-of-transit flux, suggesting a more transparent atmosphere and, hence, better observing conditions. Despite this improvement, however, interference from noise factors was still substantial, something that inspired us to rely, once again, on stochastic methods for the white light curve analysis. This means that noise contributions were handled by GP fits and that the system parameters were obtained following a similar procedure to the one outlined in Gibson et al. (2012) and in Chapter 4[3].

[3]For readability purposes, some of the steps and definitions will be repeated here, but most of the

Under the GP approach, transit light curves are described by a multivariate Gaussian distribution composed of two individual components. The first component models the deterministic input from the transit and the second component evaluates noise contributions following a stochastic formulation. For WASP-74b, in particular, the noise component is expressed by a Matérn 3/2 kernel. This kernel is used throughout the whole study, as WASP-74b, in a similar fashion to WASP-88b, displays time-dependent systematics. Gibson et al. (2013a) empirically demonstrated that the Matérn 3/2 kernel greatly outperforms other kernels in situations such as this one. Again, the Python GP package `george` (Ambikasaran et al., 2015; Foreman-Mackey, 2015) is employed to detach systematic errors from the actual transit.

Given this setup, the covariance function K_{nm} is then described by:

$$K_{nm} = \xi^2 \left(1 + \sqrt{3 \sum_{\nu=1}^{N} \left(\frac{\Delta \hat{w}_\nu}{\tau_{w_\nu}}\right)^2}\right) \exp\left(-\sqrt{3 \sum_{\nu=1}^{N} \left(\frac{\Delta \hat{w}_\nu}{\tau_{w_\nu}}\right)^2}\right) + \delta_{nm} \left(\sigma_p \beta\right)^2 . \quad (5.1)$$

In this equation, τ_{w_ν} are the characteristic length scale parameters for the optical state parameters $\hat{w}_\nu$, ξ is the characteristic height scale, and δ_{nm} is the Kronecker delta. In addition, σ_p are the shot noise uncertainties of the spectroscopic measurements and β is a multiplicative factor that rescales the computed residual errors to more reasonable values. Parameter β is assumed to be constant for all measurements. This equation is practically the same as the combination of Equations 4.3 and 4.4 presented in Chapter 4, but is written in a slightly shortened, more compact, form. This is because the difference between the kernel parameters of the auxiliary variables is given as:

$$\Delta \hat{w}_\nu = \Delta \hat{w}_{\nu,nm} = \hat{w}_{\nu,n} - \hat{w}_{\nu,m}, \quad (5.2)$$

where the circumflex denotes that the variables are standardised.

Meanwhile, the mean function μ is represented by a two-part expression that consists of a polynomial of time ($P(t)$) multiplied by the transit model ($T(t, \theta)$). This

analysis should be different as WASP-74b is simply a different target.

polynomial was included in the fit to assist with the description of the observed linear trend. The function is, therefore, defined as:

$$\mu\,(t,\theta) = P(t) \times T(t,\theta) = [c_0 + c_1 t] \times T(t,\theta), \tag{5.3}$$

where c_0 and c_1 are coefficients that explain the linear shape of the light curve. Again, θ are the transit parameters and t refers to the central times of the exposures. These times are converted from MJD to BJD with support from the time-conversion package `barycorrpy` (Kanodia & Wright, 2018). The transit models are once again constructed by taking advantage of the fast and well-tested `batman` modelling toolbox (Kreidberg, 2015), and by making the assumption that the transit light curves are well-described by the quadratic limb darkening law (Kopal, 1950).

Various optical state parameters were considered as potential noise contributors and were, hence, incorporated in the kernel function as GP regressors. Factors that could have caused additional systematic errors include, but are not limited to, the rate of change of the rotator angle, spectral drifts in the dispersion and cross-dispersion axes, FWHM and time. Figure 5.4 depicts the systematic tendencies in some of the most important ancillary parameters. These parameters were inserted as inputs in the GP kernel and were assessed independently, as well as in conjunction with other parameters. The final solution was determined to be a systematics description based on the rate of change of the rotator angle (x) for the June data set, and on the combination of the rate of change of the rotator angle and changes in the position of the spectra with respect to the cross-dispersion axis (y) for the August data set. This parameterisation leads to well-fitted light curves and a low scatter for the residuals and was, therefore, selected as the preferred choice. Other parameterisations were also checked and were found to result in very consistent transmission spectra. From this outcome, it was concluded that the GP kernel inputs in the white-light analysis were of little consequence to the final, wavelength-dependent $R_{\rm p}/R_*$ values.

To estimate the transit and noise parameters and get their uncertainties, it was important to couple the GP model with an MCMC sampling method. For this purpose, we relied one more time on the `emcee` library (Foreman-Mackey et al., 2013). This

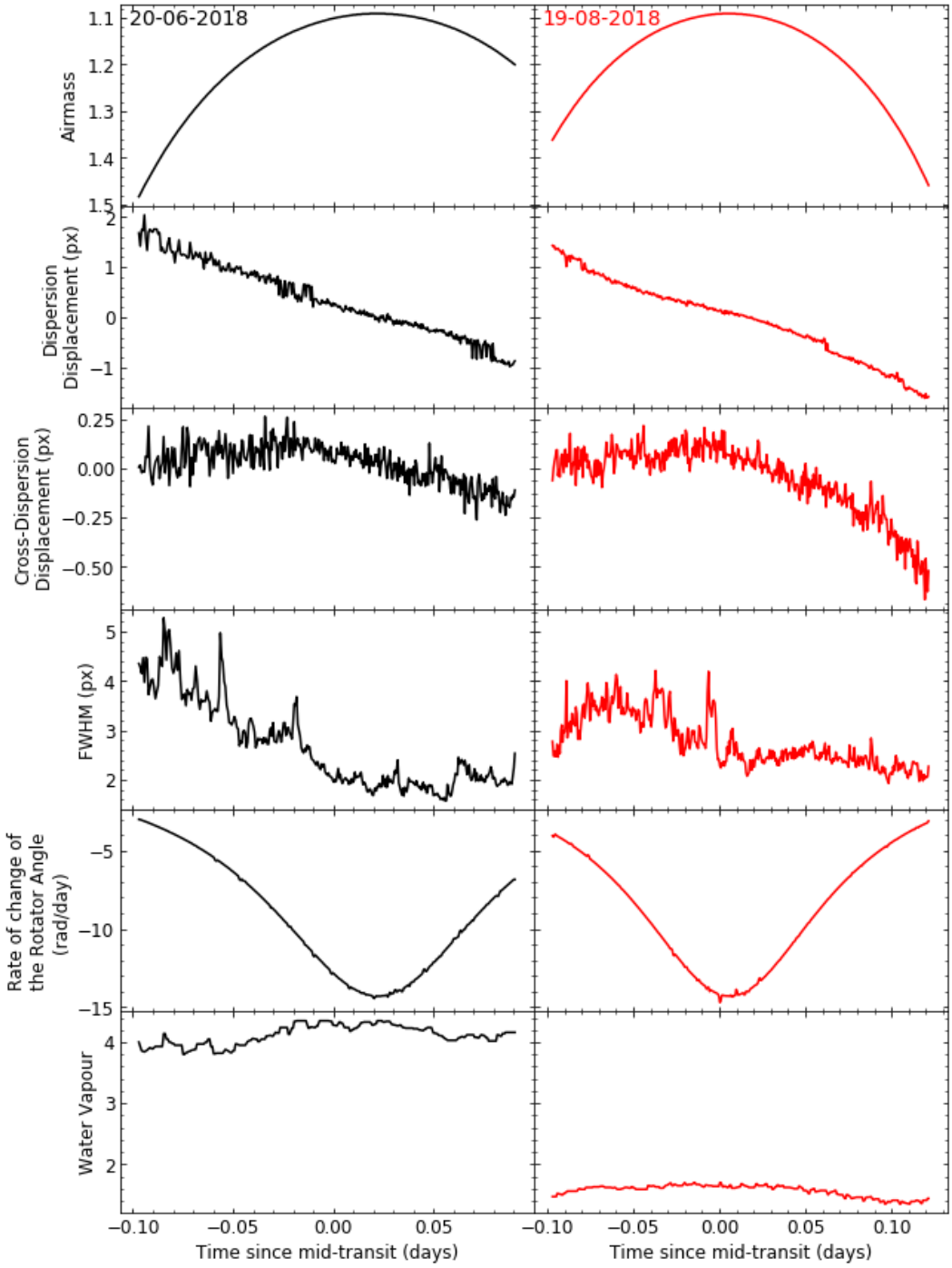

Figure 5.4: Optical state parameters for the two data sets that were fully investigated. Starting from the top, the parameters are airmass, displacements in the horizontal and vertical axes, FWHM, rotator angular velocity, and precipitable water vapour.

powerful package adds speed to the calculations without consuming a lot of memory and computes the given set of parameters by probing the parameter space through an ensemble of walkers that run in parallel and steadily progress to higher probability regions. Contrary to WASP-88b, where two fits were performed, the white light curves of WASP-74b were only fitted once. This change was necessary as both the semi-major axis to stellar radius ratio (a/R_*) and the inclination (i) were fixed to the values presented in Mancini et al. (2019) from the beginning of the fitting process. This adjustment allowed comparisons with results from the space-based analysis by Fu et al. (2021), who also made the same initial assumptions about the system properties. This single fit comprised of two iterations that were split in three parts: two burn-in phases of 500 steps and one production phase of 2000 steps. A group of 150 walkers was used in each case to efficiently explore the parameter space. In between the two burn-in phases, the walkers were re-initialised to a thin region around the walker with the highest probability. This measure stimulated a faster convergence in the second burn-in phase.

Apart from a/R_* and i, the initial planet-to-star radius ratio (R_p/R_*), the period (P) and the transit midpoint time (t_0) were also based on the values reported in Mancini et al. (2019). For t_0, especially, we projected forward to the expected value in BJD from the ephemeris computed therein. Eccentricity was fixed to 0, under the assumption of a circular orbit, in accordance with the findings from the discovery paper (Hellier et al., 2015). The linear and quadratic limb darkening coefficients were computed in the same way as for WASP-88, by employing the Stagger-grid (Magic et al., 2015) and getting the theoretical predictions on the basis of the stellar values for the temperature, gravity and metallicity provided in Hellier et al. (2015). In total, 8 to 9 parameters were let to vary freely in the fit, depending on the data set. These were the three transit parameters (t_0, R_p/R_*, u_1), the two or three kernel parameters (ξ, τ_{w_x}, τ_{w_y}), the two time polynomial coefficients (c_0, c_1), and the multiplicative factor (β). While u_1 was allowed to vary freely, the quadratic limb darkening coefficient (u_2) was fixed. Finally, the kernel parameters were broadly constrained using log-uniform priors, whereas the remaining parameters were estimated using uniform priors.

A second iteration was performed to re-evaluate the fit after outlier removal. To identify the outliers, the same 3σ rule used in the WASP-88b analysis was employed. The new results were found to be largely consistent with those from the first iteration. The white transit light curves along with the median GP fit and the systematics model from the second iteration are displayed in Figure 5.5. The noise and system parameters from the marginalised posterior are presented in Table 5.2.

As can be seen from the light curve fits, the GP model successfully follows the shape of the light curves despite some discontinuities. These discontinuities probably add some weight to the uncertainties but the effect is not significant in the white-light analysis. The linear slope that can be clearly observed in the August data set prompted us to incorporate a linear baseline function of time in the fit. The inclusion of such a relation in the mean function aided the pace of the walkers towards the maximised likelihood.

5.3.2 Spectroscopic light curves

The spectroscopic analysis was quite thorough, as already mentioned in the beginning of Section 5.3. A total of five different main approaches were considered to analyse the transit light curves. These methods are discussed below and are summarised in Table 5.3. The aim was to determine a solution that gives a precise description of the atmosphere while successfully eradicating interference from atmospheric or instrumental effects.

The first approach was similar to the one described in Section 4.3.2.2 for the GP analysis of the spectroscopic light curves of WASP-88b. According to that procedure, the analysis was performed on the relative spectrophotometric light curves, i.e. on the target flux divided by the reference star flux of the narrow binned channels. These light curves were then common-mode corrected using the common-mode obtained from the relative white flux, before being assessed with a GP fit described by a kernel function of time. This approach is commonly used in GP analyses of the spectroscopic light curves (e.g Gibson et al., 2017; Wilson et al., 2020) and so from hereafter it will be

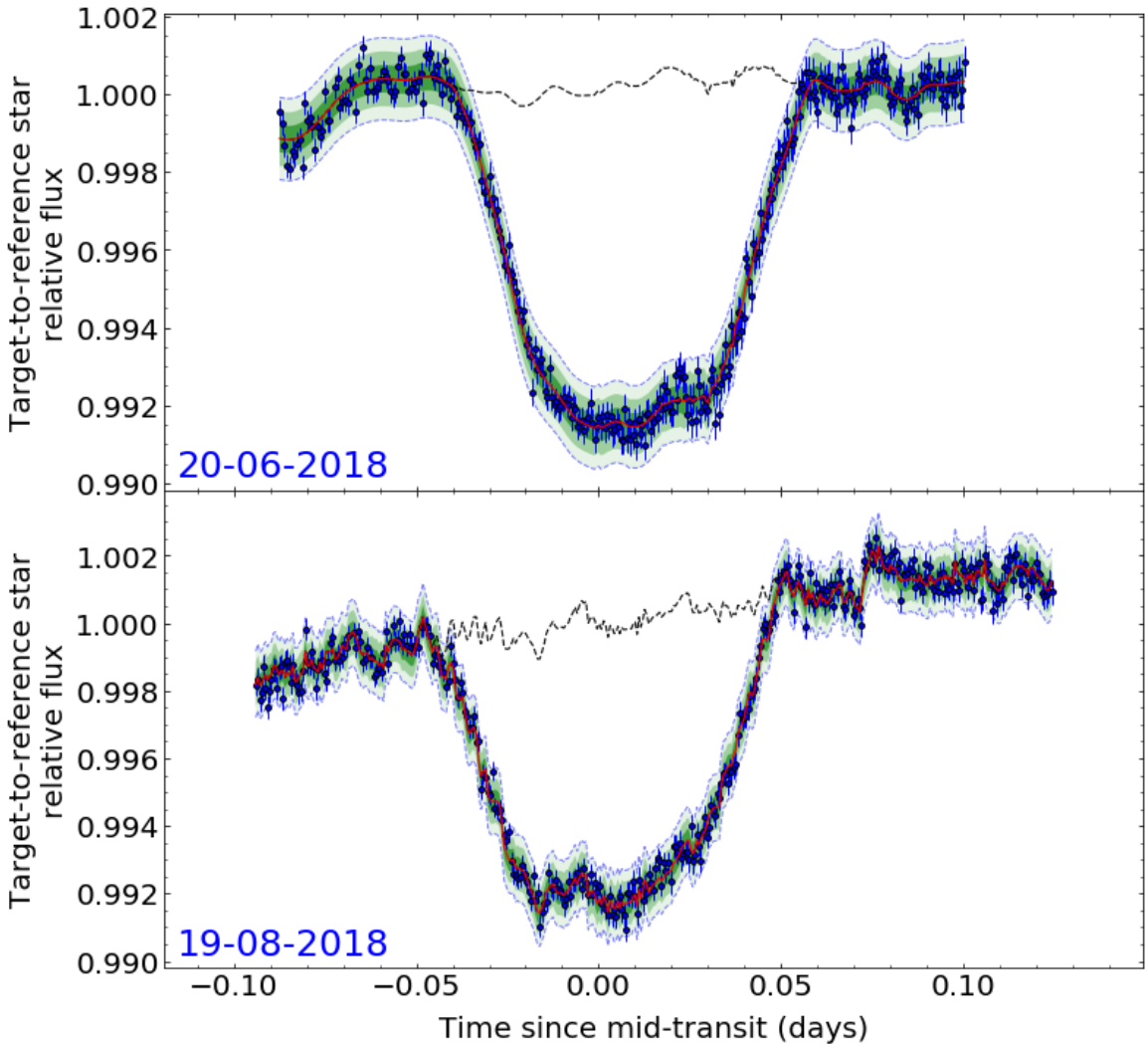

Figure 5.5: Transit light curves from the white-light GP analysis of WASP-74b. The data for the best two observations are depicted by blue circles, along with their 1σ modified spectrophotometric uncertainties (vertical error bars). The red line shows the median model from the GP fit and the black dashed line indicates the contribution from noise. The green regions, from darker to lighter shades, reveal the standard deviation of the residuals up to a factor of three.

Table 5.2: System and systematics parameters from the white-light GP analysis of WASP−74b.

Parameter	Value	Prior
Period (d)	2.1377445 (fixed)	
Eccentricity	0 (fixed)	
a/R_*	4.82 (fixed)	
i (degrees)	79.86 (fixed)	
u_2	0.29 (fixed)	
20-06-2018		
$\ln \alpha$	$-15.41^{+0.57}_{-0.46}$	$\mathcal{U}$(-20,15)
$\ln \tau_{\rm x}$	$-3.54^{+0.41}_{-0.49}$	$\mathcal{U}$(-15,15)
c_0	$1.00008^{+0.00016}_{-0.00017}$	$\mathcal{U}$(0.9,1.1)
c_1	$0.000087^{+0.000084}_{-0.000088}$	$\mathcal{U}$(-0.1,0.1)
t_0 ($\rm BJD_{TDB}$)	2458289.77937 ± 0.00038	$\mathcal{U}$(-0.01,0.01)*
$R_{\rm p}/R_*$	$0.09883^{+0.00072}_{-0.00074}$	$\mathcal{U}$(0.03,0.15)
u_1	$0.420^{+0.040}_{-0.044}$	$\mathcal{U}$(0,1)
β	$1.546^{+0.077}_{-0.091}$	$\mathcal{U}$(0,10)
19-08-2018		
$\ln \alpha$	-15.81 ± 0.21	$\mathcal{U}$(-20,15)
$\ln \tau_{\rm x}$	$-6.21^{+0.50}_{-0.31}$	$\mathcal{U}$(-15,15)
$\ln \tau_{\rm y}$	$-0.49^{+0.55}_{-0.43}$	$\mathcal{U}$(-15,15)
c_0	$1.000265^{+0.000070}_{-0.000074}$	$\mathcal{U}$(0.9,1.1)
c_1	0.000897 ± 0.000046	$\mathcal{U}$(-0.1,0.1)
t_0 ($\rm BJD_{TDB}$)	2458349.63287 ± 0.00023	$\mathcal{U}$(-0.01,0.01)*
$R_{\rm p}/R_*$	0.0951 ± 0.0013	$\mathcal{U}$(0.03,0.15)
u_1	$0.411^{+0.043}_{-0.046}$	$\mathcal{U}$(0,1)
β	$1.410^{+0.112}_{-0.096}$	$\mathcal{U}$(0,10)

*The prior range for the transit midpoint time was based on the prediction from the ephemeris presented in Mancini et al. (2019).

Table 5.3: A list of the methodologies followed in the spectroscopic analysis of WASP-74b. The abbreviations T and E signify the transit model and the exponential function of airmass, respectively.

Method	Light curve	Global fit	CM[a] correction	GP regressors	Mean function	$\bar{\sigma_\delta}$ (ppm)[b]
Classic approach	$\frac{\text{Target}}{\text{Comparison Star}}$	No	$\frac{\text{Relative SLCs}^{c}}{\text{Relative CM}}$	Time	T	719
Global approach	$\frac{\text{Target}}{\text{Comparison Star}}$	Yes	$\frac{\text{Relative SLCs}}{\text{Relative CM}}$	Time	T	–
Panwar et al. (2022) approach	Target	No	–	Time Raw CM	T	533
Panwar et al. (2022) modified approach	Target	No	–	Time Raw CM	T × E	412
New approach	Target	No	$\frac{\text{Raw SLCs}}{\text{Raw CM}}$	Time	T × E	211

[a]CM: common mode
[b]$\bar{\sigma_\delta}$: average error in the transit depth $\delta = (R_p/R_*)^2$ given in parts-per-million (ppm)
[c]SLCs: spectroscopic light curves

referred to as "the classic approach".

Due to insufficient constraints on the transit parameters, other options were also investigated. One of them is the so-called "global approach". Based on this method, the spectroscopic transit light curves were not analysed individually but simultaneously using a global fit where information is shared between all the spectroscopic light curves. This led to large multi-variate distributions where some variables were common for all light curves and others such as $R_{\rm p}/R_*$ were different for each light curve.

Another method that was explored isolated the raw spectroscopic light curves of the target and performed the GP analysis only on them. This approach effectively removed the need to use a comparison star, a technique normally implemented on the light curves to correct for telluric effects. This method was developed by Panwar et al. (2022) and also did not involve the classic linear way of correcting the light curves by the common mode. Instead, the common mode was incorporated as a regressor in the GP kernel. Therefore, the data were fully treated stochastically and the analysis completely relied on non-linear methods to decorrelate systematic errors.

Upon evaluation of the results from this method, a modified version was also examined. The modified example was still quite similar in its implementation and only differed from the Panwar et al. (2022) method in that the mean function was described differently. According to the modified approach, a parametric function of airmass was also included in the description of the transit. This adjustment was carried out in light of the exponential shape seen in the raw, out-of-transit, spectroscopic light curves.

And, finally yet importantly, a novel method first presented in the submitted WASP-74b manuscript will also be discussed. The new method is similar to the Panwar et al. (2022) approach in the sense that the spectroscopic analysis only considers the light curves from the target. The common mode, however, is corrected for linearly. More specifically, the common mode is determined by dividing the raw target flux of the white light curve by the MCMC-established transit model. Then, the raw target fluxes of the spectroscopic light curves is divided by this common mode to obtain the canvas for the GP fit. The final fit is computed by taking into account a kernel function

of time and a mean function comprised of the product between an airmass exponential and a transit model.

5.3.2.1 The classic approach: target-to-reference star relative flux

The classic approach described here is practically the same as the preferred analysis for WASP-88b. The investigation focuses on the relative (target-to-comparison star) spectroscopic light curves and performs a common-mode correction in accordance with other recent FORS2 investigations (e.g. Nikolov et al., 2016, 2018b; Carter et al., 2020; Wilson et al., 2020). The wavelength-independent trend (or common mode) is established by dividing the relative white light curve of each data set by the median transit model obtained from the GP fit. The relative spectroscopic light curves are then subsequently divided by this common systematics behaviour to eliminate any influence from external factors.

After common-mode correction, the spectroscopic GP analysis reflects the procedure followed for the white light curves in Section 5.3.1. This means that the transit light curves were treated as GPs, where the transit is part of the mean function and noise is part of the kernel function. In contrast to the white-light analysis, the kernel function here was simply defined by a regressor of time, under the assumption that any physical systematic effects were corrected with the application of the common-mode correction. The fitting process was then very similar to the one described in Section 5.3.1 and the MCMC chains were run for the same amount of steps and walkers.

Some changes compared to the white-light analysis include the fewer free parameters in the fit and the omission of the parametric function in the mean function. More specifically, the only parameters that were allowed to vary freely now were the two transit parameters R_p/R_* and u_1, the two kernel parameters $\ln\alpha$ and $\ln\tau_\mathrm{t}$, and the residual scaling factor β. Both free and fixed parameters were initially set to the values obtained from the white-light analysis. Moreover, time was already accounted for in the GP kernel and so it was deemed unnecessary to aid the fit with a polyno-

mial of time. Finally, two outliers were removed after the first iteration and a second iteration was implemented to derive the final system parameters and build the transmission spectrum. Figures 5.6 and 5.7 illustrate the relative spectroscopic light curves generated by this approach, along with the median GP fits and their residuals.

5.3.2.2 The global approach: simultaneous fit of all relative light curves

The use of this method was inspired by Ahrer et al. (2022), who utilised this method to achieve a higher precision in the transmission spectrum of WASP-94Ab. They initially used the classic method and observed an average transit depth uncertainty of 181 ppm for wavelength bins of ∼200Å. When they implemented a global approach, however, they managed to reduce the size of the transit depth uncertainties to ∼128 ppm. This result is equivalent to a reduction of ∼30% in the size of the errors. The new uncertainties enabled improved constraints for the observed sodium feature and made a steep scattering slope clearly visible. Interestingly, Ahrer et al. (2022) employed the NTT EFOSC2 instrument for their measurements, which is the same instrument used for the WASP-75b observations (see Chapter 3).

In order to use a global model in the spectroscopic analysis it is necessary to assume that part of the systematics is common to all spectroscopic light curves. Upon inspection of the light curves, we found that the characteristic lengthscales of time showed a similar behaviour from light curve to light curve and an agreement was determined at the 1σ level. Therefore, this parameter was selected to be common for all spectroscopic bins. In their WASP-94Ab analysis, Ahrer et al. (2022) did the same with the lengthscale parameters, adding further confidence to our approach. Figure 5.8 illustrates the behaviour of the logarithmic lengthscale values as a function of wavelength.

The fitting procedure in the global approach was similar to the one from the classic method with two iterations consisting of two burn-in phases and one production phase. Nonetheless, the number of steps and the number of walkers were different to accommodate for the increased number of free variables. More precisely, the steps rose

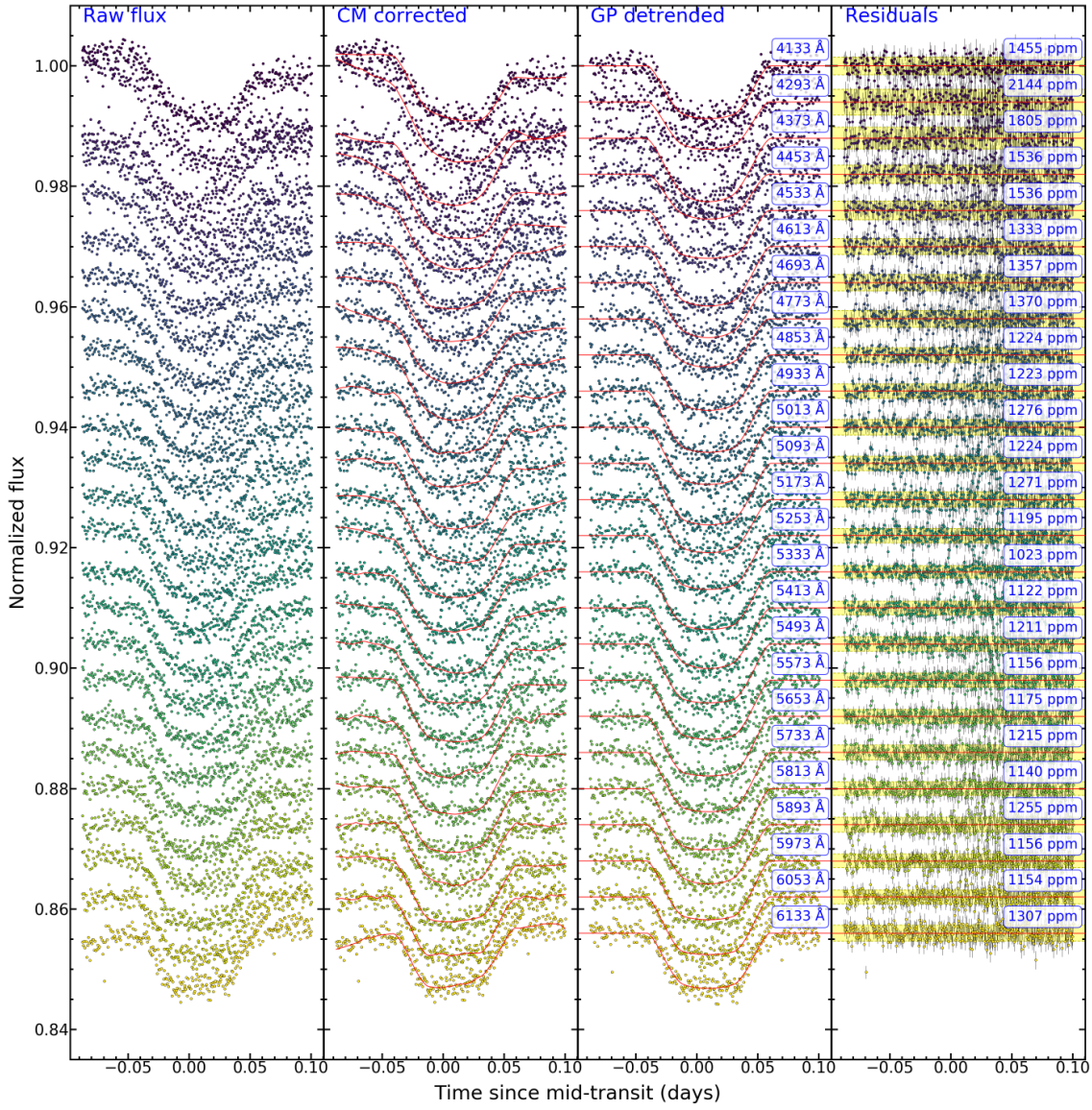

Figure 5.6: Spectroscopic light curves of WASP-74b from June 20th 2018 as obtained through the classic approach. Longer wavelengths are shown by yellow dots whereas shorter wavelengths are represented by dark blue dots. First column: Relative (target-to-comparison star) transit light curves. Second column: Common-mode corrected light curves fitted with GP models (red lines). Third column: Detrended light curves overplotted with median transit models (red lines). Fourth column: Residuals from the median model fits, 1σ residual regions (yellow areas), and rescaled spectrophotometric uncertainties (vertical error bars). The transit light curves are offset for clarity.

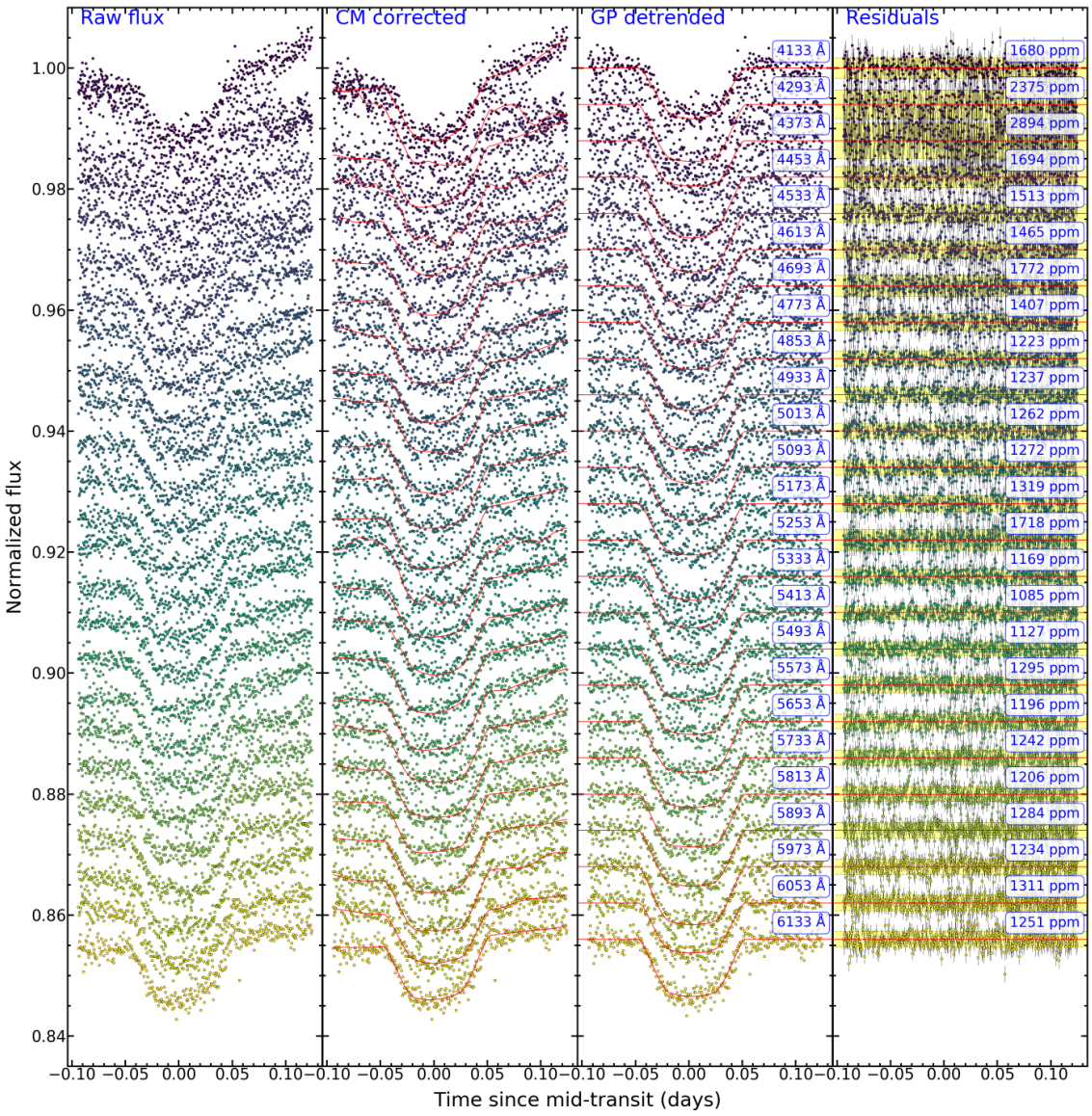

Figure 5.7: Same as Figure 5.6, but for the spectroscopic light curves of WASP-74b from August 19th 2018.

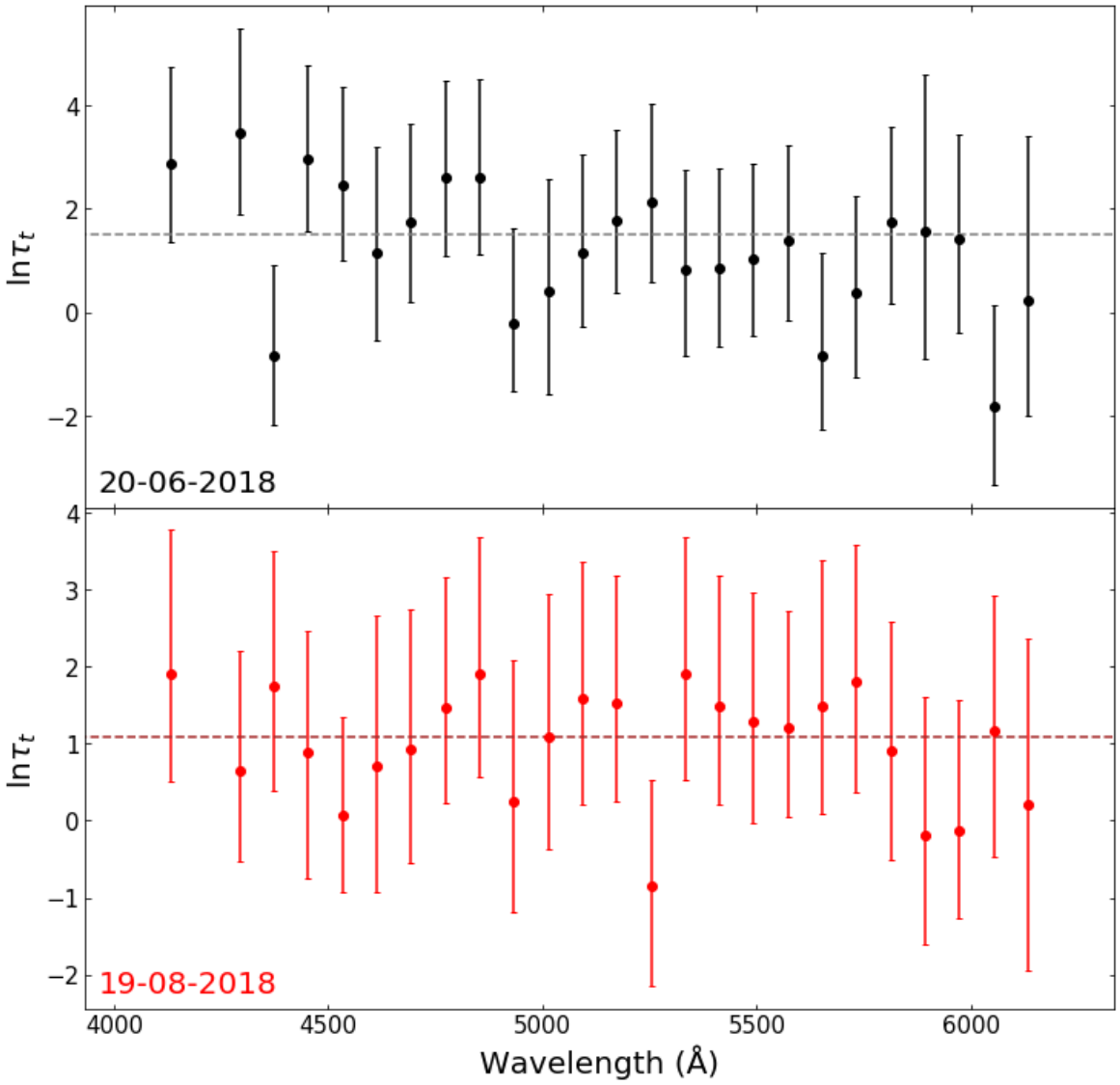

Figure 5.8: The lengthscale of time as a function of wavelength from the classic method. The data set from June 20th 2018 (black data, top) shows a slightly higher scatter compared to the data set from August 19th 2018 (red data, bottom). The discrepancy, however, is minor, and an agreement with a horizontal line fit of less than 1σ was established for both data sets.

to 1000 for the burn-in phases and to 5000 for the production phase, while the walkers increased to 600. The higher number of steps was crucial to ensure convergence after the exploration of a larger parameter space.

All in all, 77 free parameters were taken into account during the global GP fit. These were the 25 planet-to-star radii ratios (R_p/R_*), the 25 linear limb darkening coefficients (u_1) and the 25 height scale parameters ($\ln \alpha$), as well as the shared lengthscale of time ($\ln \tau_\mathrm{t}$) and the shared multiplicative factor (β). The number 25 here indicates that there was one distinct variable per spectroscopic light curve. The remaining transit variables were fixed to the estimates from the white-light analysis, while for the free variables the white light estimates were used as initial guesses. The data were then filtered through a 3σ outlier rejection algorithm before revealing the final quantities of the transmission spectrum in a second iteration. The spectroscopic light curve fits from the global approach can be seen in Figures 5.9 and 5.10.

This method resulted in a transmission spectrum that is in very good agreement with the outcome from the classic method giving a match that is well below 1σ. It did not, however, provide an improvement to the R_p/R_* uncertainties, resulting in an increase of 15%. The computational effort in order to get this result was high and each variation in the setup of the framework required about one-to-two weeks of calculations. The reason for the increased amount of time was the combination of a larger parameter dimension (and, hence, a higher number of walkers) and an increased number of steps. The approach was, therefore, computationally demanding, leading to a search for other alternatives.

5.3.2.3 The Panwar et al. (2022) approach: fit on the raw flux of the target using a GP kernel of time and common mode

Relative, target-to-comparison star, light curves are sometimes characterised by a substantially increased scatter compared to raw light curves. This happens as additional scatter from the light curves of the comparison star is introduced to the light curves of the target during the production of the relative light curves. This effect is especially

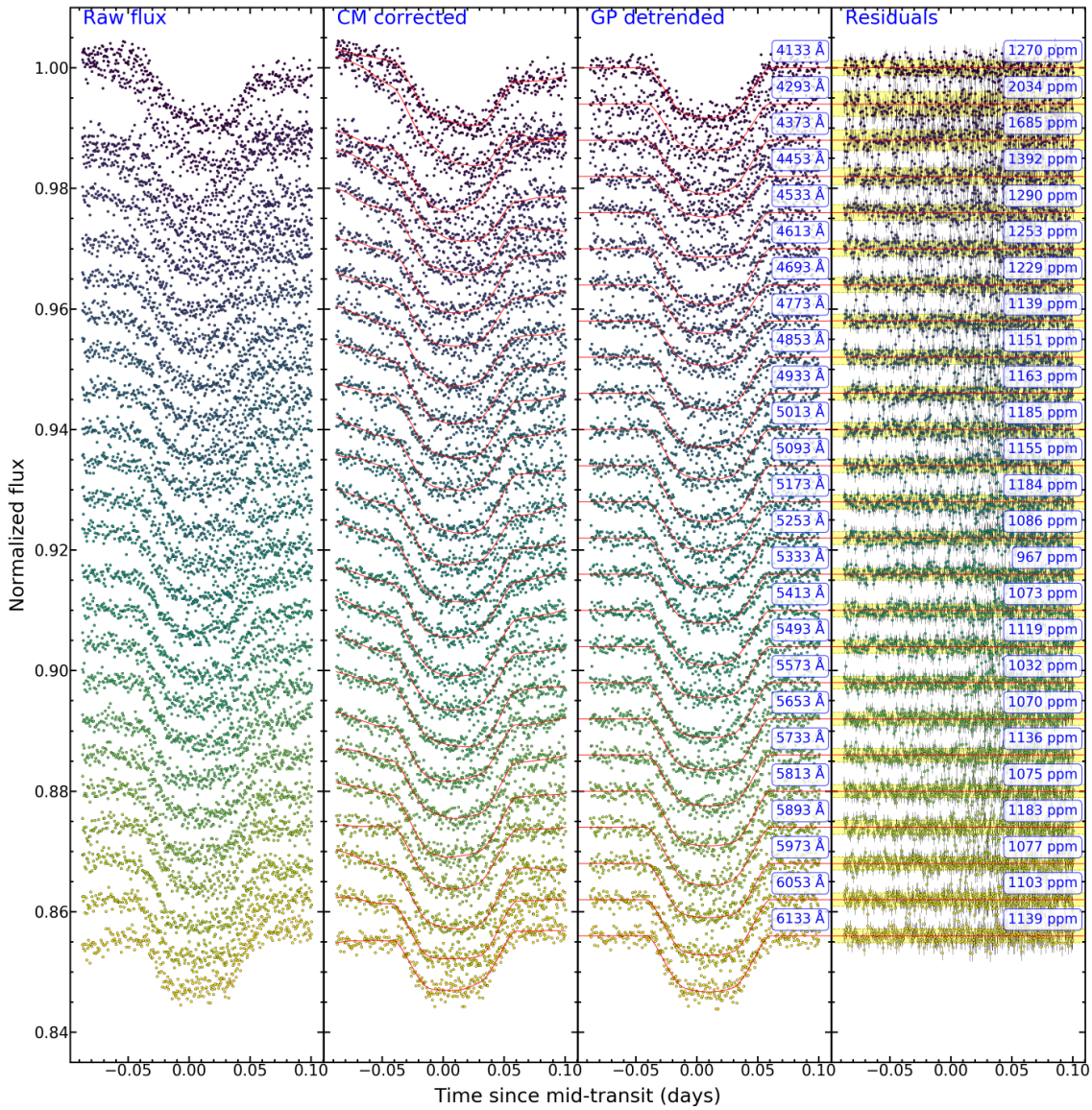

Figure 5.9: Spectroscopic light curves of WASP-74b from June 20th 2018 as obtained through the global approach. Longer wavelengths are shown by yellow dots whereas shorter wavelengths are represented by dark blue dots. First column: Relative (target-to-comparison star) transit light curves. Second column: Common-mode corrected light curves fitted with models obtained from a global GP fit (red lines). Third column: Detrended light curves overplotted with median transit models (red lines). Fourth column: Residuals from the median model fits, 1σ residual regions (yellow areas), and rescaled spectrophotometric uncertainties (vertical error bars). The transit light curves are offset for clarity.

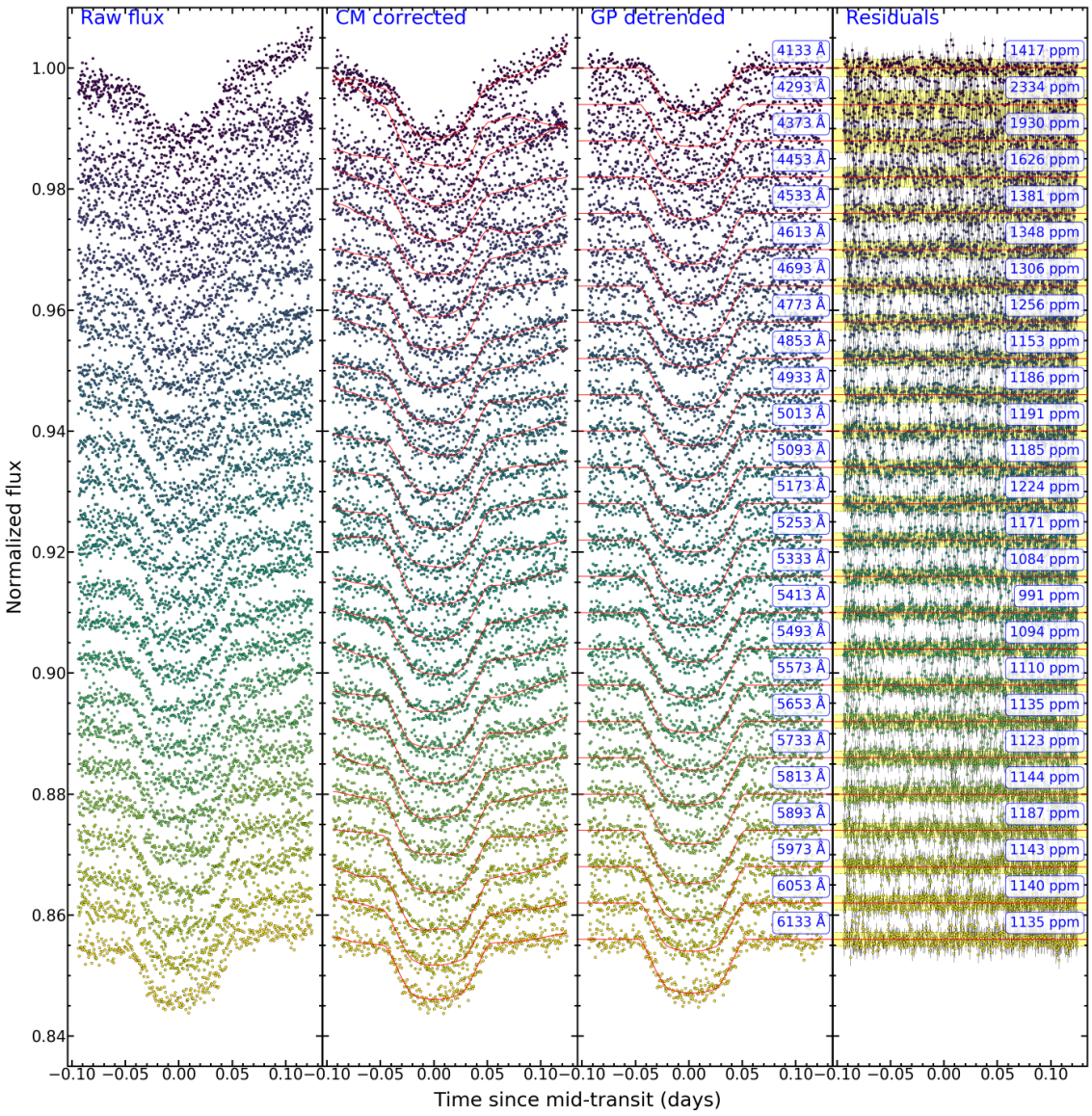

Figure 5.10: Same as Figure 5.9, but for the spectroscopic light curves of WASP-74b from August 19th 2018.

significant when much fainter comparison stars are used. This is the case for WASP-74b and the impact of this effect can be clearly seen in the relative, common-mode corrected, spectroscopic light curves (see Figures 5.6, 5.7, 5.9 and 5.10). As a consequence of this process, systematic errors are increased and this can affect the values and uncertainties of the transit parameters. Here, the issue is particularly evident in the classically-produced transmission spectrum of WASP-74b, where the average transit depth uncertainty is determined to be about five times larger than the corresponding uncertainty from HST. Other factors such as atmospheric conditions may also have contributed to this much lower precision but the fact remains that the raw spectroscopic light curves show a nearly five times lower scatter compared to the relative counterparts.

The solution to this problem was to apply the GP fits on the raw spectroscopic light curves of the target and to completely disregard the spectroscopic light curves of the reference star. Such an approach had already been implemented by Panwar et al. (2022) to produce the transmission spectrum of the warm Neptune HAT-P-26b. They achieved a reduction of $\sim$25% in the transit depth uncertainties compared to the classic approach, a promising improvement that motivated the further trial of this method in the study of WASP-74b. For the record, the transmission spectrum of that exoplanet was found to be flat, indicating a cloudy atmosphere.

The Panwar et al. (2022) approach, apart from focusing exclusively on target light curves, offered one additional innovation. The common mode from the white-light analysis was decorrelated from the spectroscopic light curves not through the normal method of division ("common-mode correction") but through incorporation in the GP kernel. This means that the common mode was assessed stochastically during the fit as a regressor alongside time. In other words, common-mode correction was skipped entirely and a fit that comprised of a GP kernel with two regressors (time and common-mode) was applied on the raw spectroscopic light curves of the target.

Following the recipe described by Panwar et al. (2022), we performed a similar fit on the raw spectroscopic fluxes of WASP-74. Much of the approach here reflected the procedure in the classic approach as, again, the MCMC sampling process involved two

burn-in stages and one production stage that consisted of 500 and 2000 steps respectively. The same five parameters were let to vary freely in the fit, while an additional sixth parameter representing common-mode was also permitted to vary freely. The remaining parameters were once again fixed to the values from the white light investigation. Figures 5.11 and 5.12 depict the steps that were followed during this approach and the GP fits that were applied to the raw spectroscopic light curves of the target.

This method led to a non-negligible improvement in the resulting transmission spectrum as it provided a higher precision for the measured $R_{\rm p}/R_*$ values (see Table 5.3). In particular, an average reduction in the transit depth uncertainty of $\sim$26% was established, which, remarkably, is very similar to the result found for HAT-P-26b by Panwar et al. (2022). This outcome highlights the benefits of applying the fit only on the spectroscopic data of the target in cases where substantial light curve systematics are observed.

5.3.2.4 The modified Panwar et al. (2022) approach: taking extinction into account

Incident starlight passes through many layers of the atmosphere before reaching the observing facility on the ground. Part of this light experiences absorption and/or scattering as it engages the atoms and molecules that make up Earth's atmosphere and, therefore, never reaches the surface at the point of observation. Such interactions lead to a reduction in the observed starlight from the ground facility, a phenomenon that is known as atmospheric extinction. The effect is quite evident in the raw light curves where a curvature with a characteristic exponential shape can be recognised.

The properties of this shape are formulated in the Beer-Lambert Law that describes the physical conditions under radiative transfer:

$$F(\lambda) = F_0(\lambda) e^{-\frac{\tau(\lambda)}{\cos z}}, \tag{5.4}$$

where $F(\lambda)$ is the observed flux, $F_0(\lambda)$ is the flux before entry in the Earth's atmosphere, $\tau(\lambda)$ is a function that defines the degree of attenuation or extinction within

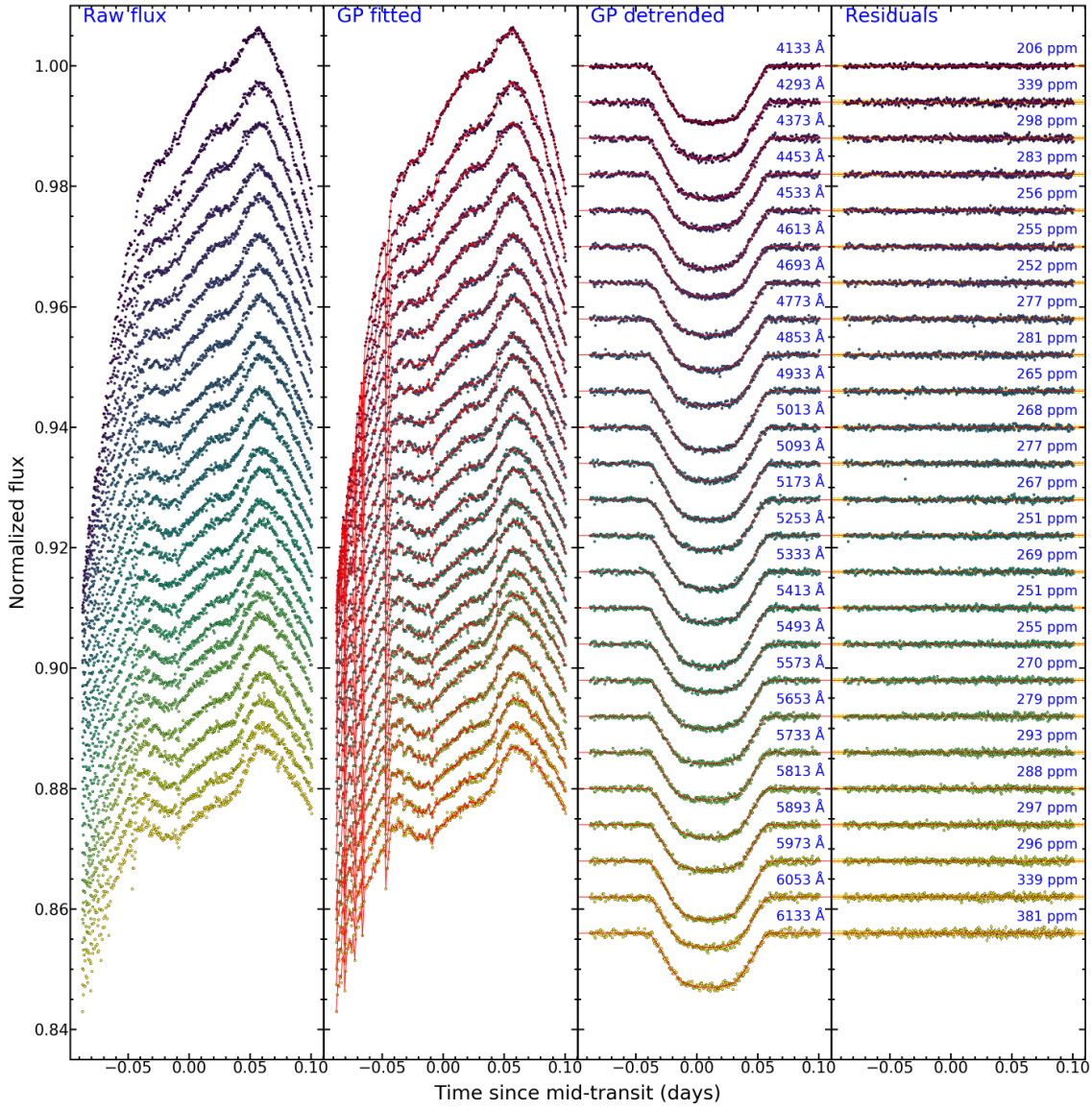

Figure 5.11: Spectroscopic light curves of WASP-74b from June 20th 2018 as obtained through the Panwar et al. (2022) approach. Longer wavelengths are shown by yellow dots whereas shorter wavelengths are represented by dark blue dots. First column: Raw transit light curves of the target. Second column: Raw light curves fitted with GP models (red lines). Third column: Detrended light curves overplotted with median transit models (red lines). Fourth column: Residuals from the median model fits, 1σ residual regions (yellow areas), and rescaled spectrophotometric uncertainties (vertical error bars). The transit light curves are offset for clarity.

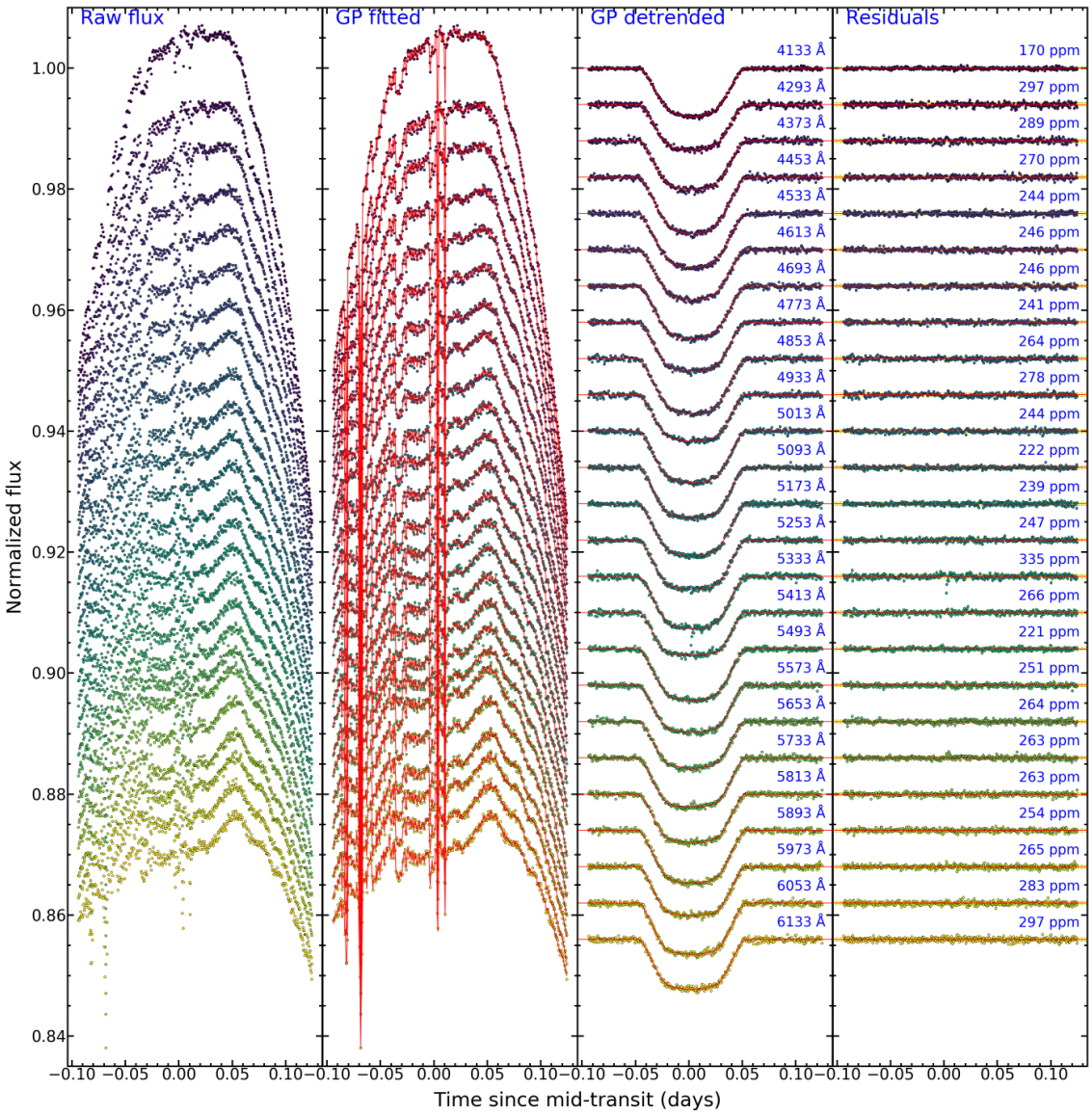

Figure 5.12: Same as Figure 5.11, but for the spectroscopic light curves of WASP-74b from August 19th 2018.

the atmosphere, and $\frac{1}{\cos z} = \sec z$ is determined by the effective zenith distance z and is equivalent to the airmass A. λ here indicates that the parameters are a function of wavelength and, thus, a variation in the curvature of the spectroscopic light curves is to be expected.

The formula for $\tau(\lambda)$ can be quite complex as it heavily depends on the atmospheric conditions during the time of the observation. Typically, $\tau(\lambda)$ is defined by the sum of several scattering and absorption coefficients that explain the attenuation effect as light interacts with atmospheric constituents. However, the impact of each atmospheric phenomenon on the observed light curves may vary and is subject to the composition of the Earth's atmosphere above the ground-based facility. Thankfully, for the purpose of this study, exact knowledge of the degree of contribution from each source of scattering and absorption is not needed and extinction can be assessed as a whole by one single parameter. Furthermore, it should be noted that, although instrumental effects can also influence the shape of the raw light curves, their involvement here tends to be minimal compared to atmospheric extinction.

The use of exponential functions of airmass to correct for telluric effects is not something new in transmission spectroscopy, but the applications so far have been limited to observations made in high resolution. For example, Astudillo-Defru & Rojo (2013) were able to determine the presence of calcium in the atmosphere of HD 209458b for the first time by performing such telluric corrections on their high resolution spectra. A similar approach was taken by Wyttenbach et al. (2015) in a successful attempt to detect sodium in the atmosphere of HD 189733b. More impressively, a parametric function of airmass was used even earlier by Vidal-Madjar et al. (2010) to observe the Earth's optical transmission spectrum. Vidal-Madjar et al. (2010) managed to do this through lunar eclipse observations and were able to detect prominent atmospheric species such as sodium, nitrogen, and ozone, as well as the characteristic signatures of oxygen.

Given the sizeable effect of extinction on the raw spectroscopic light curves of the target, it was deemed necessary to adjust the GP model and incorporate a more general version of the Beer-Lambert law. Therefore, the mean function was modified

accordingly to include a parametric function of airmass. The new, modified relation is expressed as follows:

$$\mu\left(A,t,\theta\right) = E(A) \times T(t,\theta) = \left[\mathrm{a_0 e^{-a_1} } A\right] \times T(t,\theta). \tag{5.5}$$

In this modified form of the mean function, $F_0(\lambda)$, $\tau(\lambda)$ and $\frac{1}{\cos z}$ are replaced by the two parameters $\mathrm{a_0}$ and $\mathrm{a_1}$ that explain the exponential curvature, and parameter A that represents airmass. Again, t symbolises the central exposure times in BJD and θ are the transit parameters that construct the transit model T. E signifies that the first part of the function takes the form of an exponential. Both $\mathrm{a_0}$ and $\mathrm{a_1}$ are free parameters in the fit with $\mathrm{a_0}$ settling at values close to unity, as the spectroscopic flux is normalised, and $\mathrm{a_1}$, the attenuation coefficient, being positive in the bluer wavelengths and negative in the redder wavelengths.

Apart from the adjusted mean function, this approach included one additional stage before the GP fitting procedure. This complementary stage entailed the application of a Levenberg-Marquardt fit on the out-of-transit flux of each spectroscopic bin. The implementation of this extra measure was essential in order to calculate the initial values for the coefficients $\mathrm{a_0}$ and $\mathrm{a_1}$. The use of model-generated and, therefore, more accurate initial guesses for the parameters of the mean function was found to stimulate a faster convergence towards the optimised description of the raw, transit light curves. For the non-linear Levenberg-Marquardt least squares fits, the curve fitting module `lmfit` (Newville et al., 2016) was once again utilised. During the fits, five to six outliers generated by the presence of clouds were also identified and discarded in every spectroscopic light curve of the August data set.

The new guesses of the exponential coefficients were then used as starting points for the modified GP fits applied on the complete transit light curves. The modified GP models now consisted of a simple kernel function of time and a more complicated mean function defined by the product between the airmass exponential and the transit model. The spectroscopic light curves produced by this approach and their residuals are shown in Figures 5.13 and 5.14. This realisation of the problem resulted in finer constraints for the transmission spectrum parameters. The effect was more evident

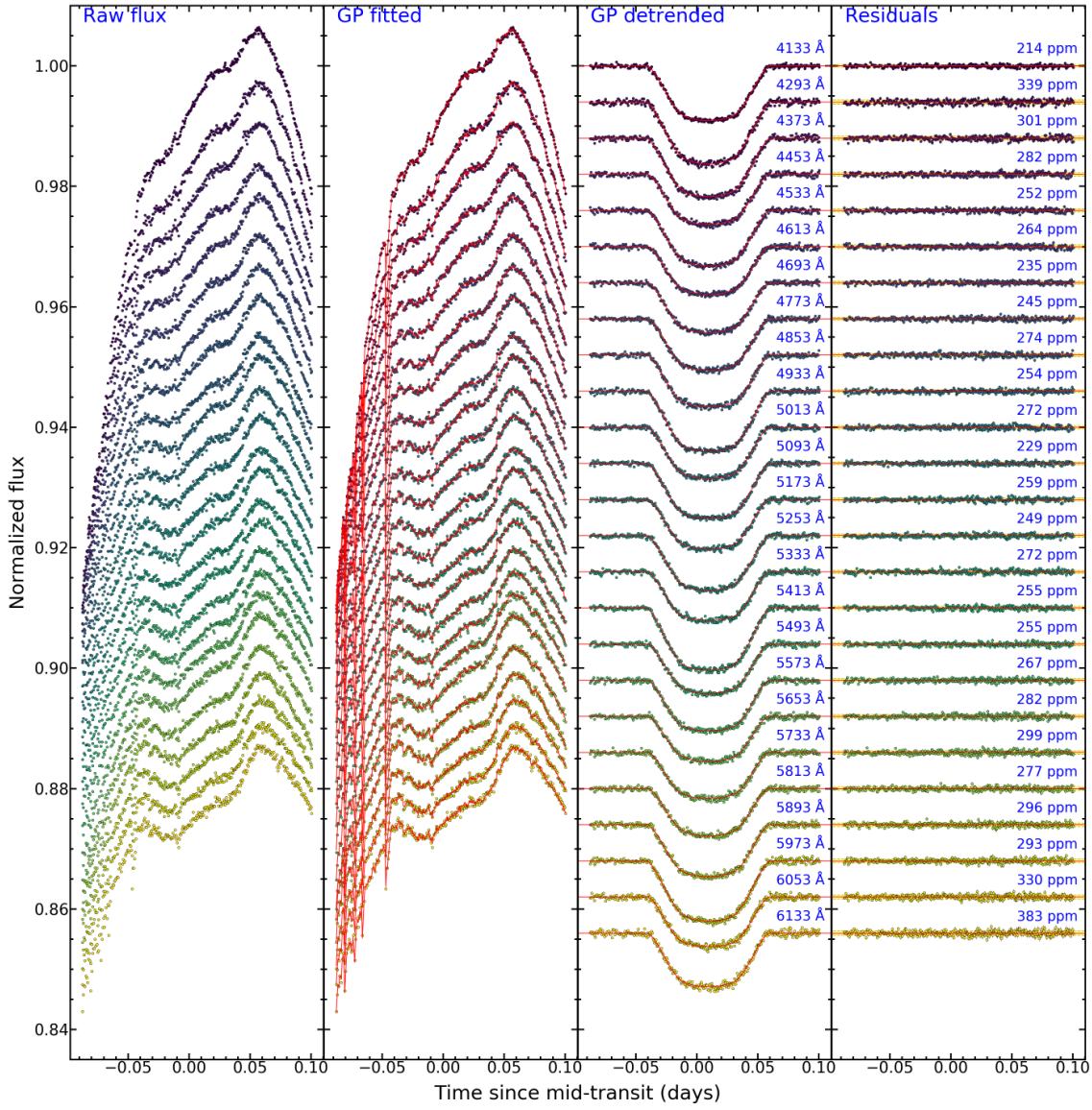

Figure 5.13: Spectroscopic light curves of WASP-74b from June 20th 2018 as obtained through the modified Panwar et al. (2022) approach. Longer wavelengths are shown by yellow dots whereas shorter wavelengths are represented by dark blue dots. First column: Raw transit light curves of the target. Second column: Raw light curves fitted with GP models (red lines). Third column: Detrended light curves overplotted with median transit models (red lines). Fourth column: Residuals from the median model fits, 1σ residual regions (yellow areas), and rescaled spectrophotometric uncertainties (vertical error bars). The transit light curves are offset for clarity.

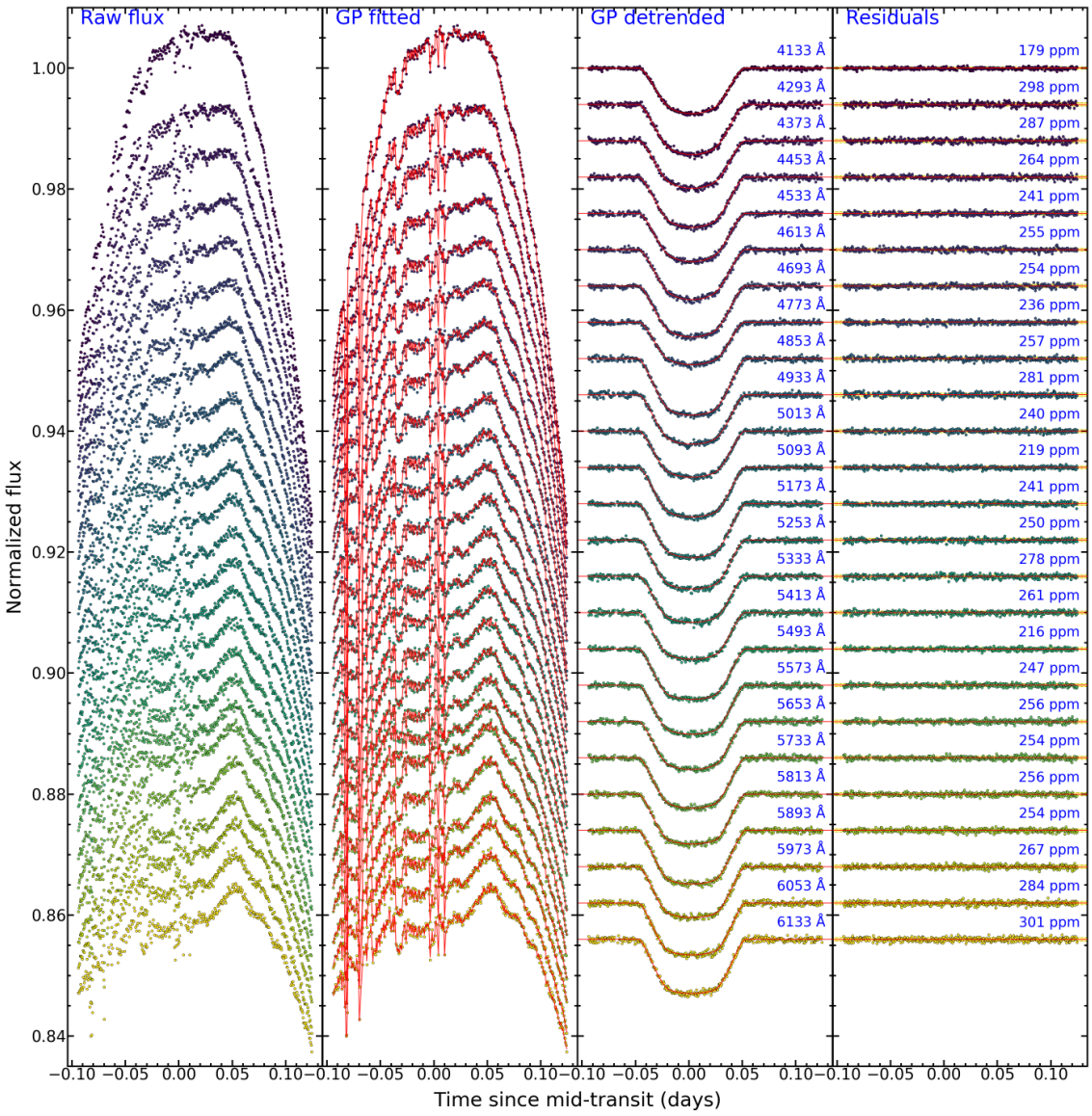

Figure 5.14: Same as Figure 5.13, but for the spectroscopic light curves of WASP-74b from August 19th 2018.

at the two ends of the transmission spectrum where spectroscopic light curves were inclined to experience a higher amount of scatter. On average, the uncertainties in $R_{\rm p}/R_*$ underwent a considerable reduction of more than 40% in size compared to the classic approach. This result suggests that the addition of an exponential function of airmass can further improve the fitting methodology developed by Panwar et al. (2022).

5.3.2.5 The new approach: common-mode correction on the raw flux of the target

The modified Panwar et al. (2022) approach was a success considering that the precision was not only increased compared to the classic method but also compared to the original Panwar et al. (2022) method. It also demonstrated the effectiveness of a parametric function of airmass. Nevertheless, the errors were still more than twice higher compared to the published space-based results by (Fu et al., 2021). This weakened the usefulness of the blue-optical data sets given that they explore a much narrower spectral range. It also made blue-optical atmospheric retrievals less powerful in placing constraints on the obtained transmission spectrum.

Therefore, one more approach was followed in an attempt to reduce the uncertainties even further. This new method was based on the classic common-mode correction method but differed from the classic approach in that the correction was applied on the raw spectroscopic light curves of the target. Furthermore, the model fit was reinforced by an exponential function of airmass in a similar way to the modified Panwar et al. (2022) approach. In fact, these two methods are equivalent with the only difference being that the common mode was removed through a linear operation here instead of being identified and isolated through a stochastic process. According to this technique, the raw white flux of the target was divided by the optimised transit model from the relative (target-to-reference star) white light analysis. Then, the raw spectroscopic fluxes of the target were divided by this common trend to obtain the common-mode corrected light curves. The procedure was performed on both the June and August data sets before we progressed to the GP fitting process.

Since the fitting process was similar to the methodology detailed in Section 5.3.2.4, it involved one initial Levenberg-Marquardt fit to the out-of-transit data. This step was crucial in the determination of the initial estimates for the normalisation parameter a_0 and the attenuation coefficient a_1, and resulted in the removal of three outliers. The initial estimates for the transit and kernel parameters were taken from the white light curve results, while the initial limb darkening coefficients for each binned light curve were obtained one more time from theoretical predictions based on the Stagger-grid (Magic et al., 2015). Altogether, a total of seven free parameters were included in the fit, including the transit parameters R_p/R_* and u_1, and the systematic parameters β, a_0, a_1, $\ln \alpha$ and $\ln \tau_t$.

This novel method produced the most impressive results with an average uncertainty reduction of almost 50% compared to the modified Panwar et al. (2022) approach. This translates to a staggering decrease of over 70% in the errors from the classic approach. The new spectroscopic light curves were also characterised by a much lower scatter compared to the classic method with ~ 275 ppm for the June data set and ~ 258 ppm for the August data set. This is only $\sim$1.25 times higher than the photon noise limit. For comparison, the spectroscopic light curves from the classic methodology had a scatter that was $\sim$6 times higher than the photon noise limit. Figures 5.15 and 5.16 show in succession the raw spectroscopic light curves of the target, followed by the common-mode corrected and the detrended light curves, and then by the residuals and their uncertainties. The GP fit results for R_p/R_*, u_1 and u_2 are displayed in Table 5.4.

An interesting observation from the second panel of the spectroscopic light curve plots is the effect of the linear common-mode correction on the curvature of the raw spectroscopic light curves. More specifically, the light curves tend to assume a convex shape at the top and a concave shape at the bottom. The reason for this seemingly odd behaviour is the nature of the common-mode corrected light curves, which are in fact a ratio of exponentials with two distinct attenuation indexes. In particular, it is the ratio between the extinction-contaminated white flux and the also contaminated raw spectroscopic fluxes. However, the effect of extinction is not monochromatic and

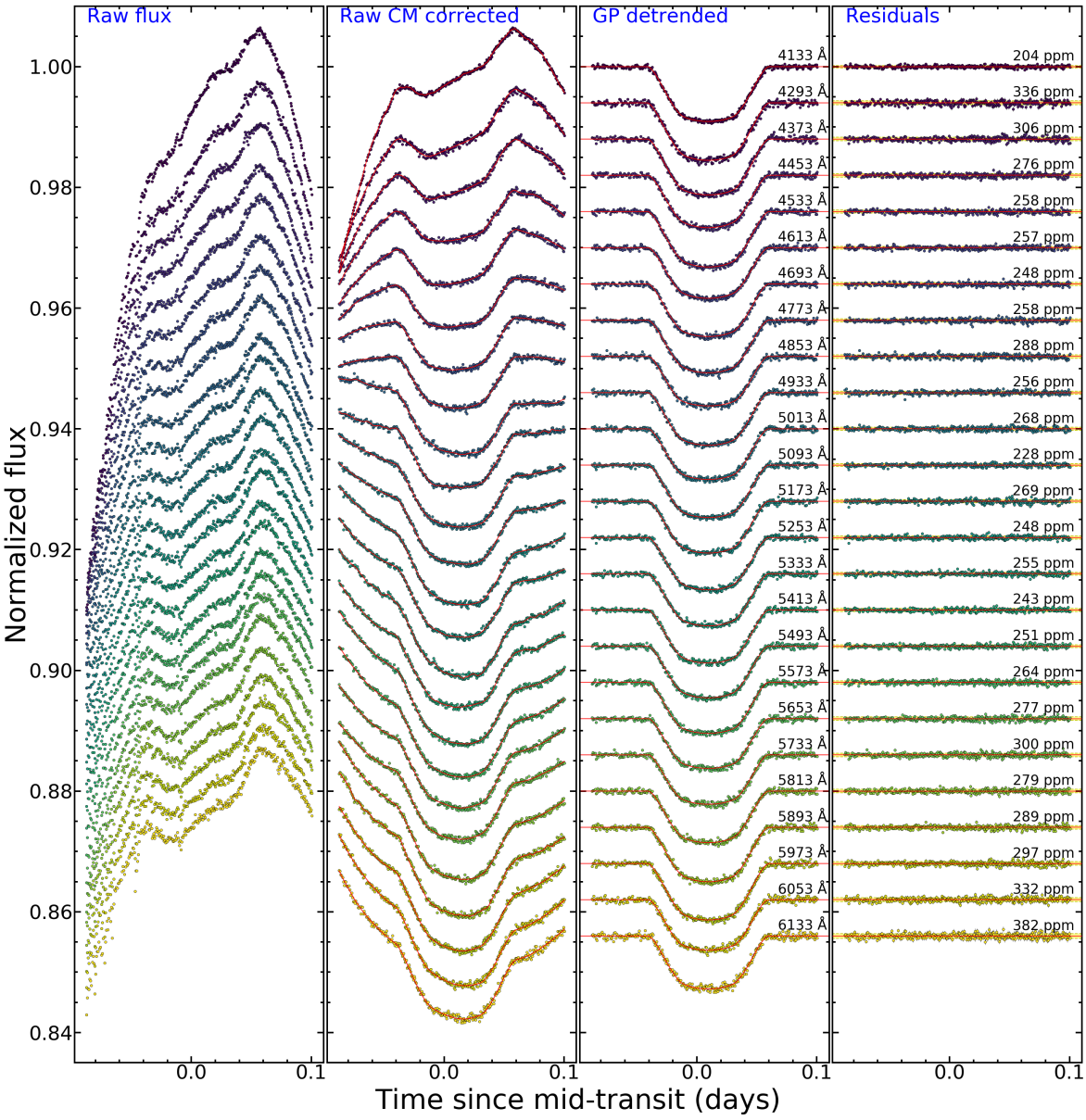

Figure 5.15: Spectroscopic light curves of WASP-74b from June 20th 2018 as obtained through the new approach. Longer wavelengths are shown by yellow dots whereas shorter wavelengths are represented by dark blue dots. First column: Raw transit light curves of the target. Second column: Common-mode corrected light curves fitted with GP models (red lines). Third column: Detrended light curves overplotted with median transit models (red lines). Fourth column: Residuals from the median model fits, 1σ residual regions (yellow areas), and rescaled spectrophotometric uncertainties (vertical error bars). The transit light curves are offset for clarity.

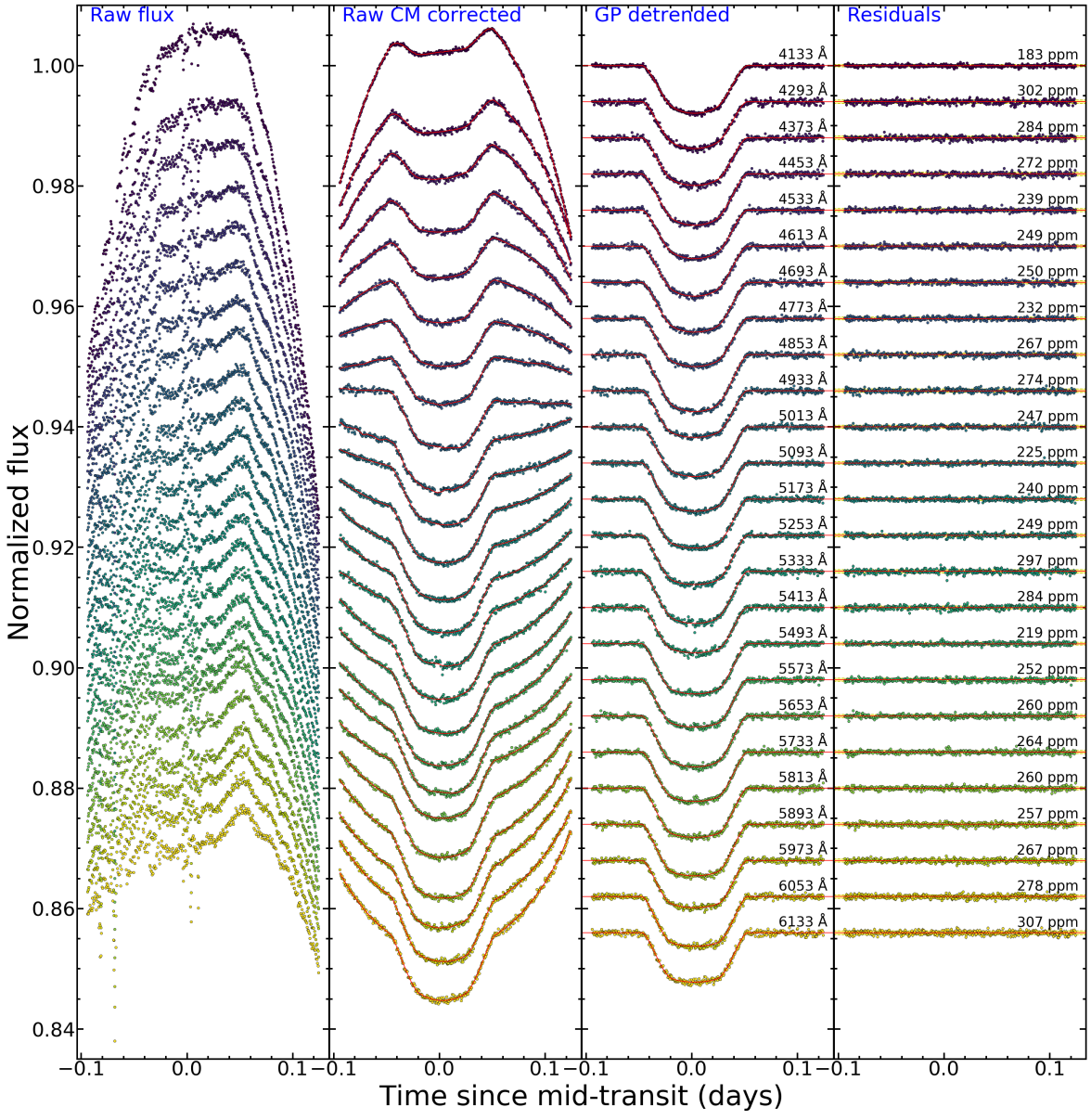

Figure 5.16: Same as Figure 5.15, but for the spectroscopic light curves of WASP-74b from August 19th 2018.

Table 5.4: Transmission spectrum parameters for WASP-74b from the new approach.

Wavelength Range (Å)	$R_{\rm p}/R_*$	u_1	u_2
$4013-4253$	$0.09650^{+0.00136}_{-0.00138}$	$0.641^{+0.021}_{-0.021}$	0.195
$4253-4333$	$0.09705^{+0.00164}_{-0.00144}$	$0.627^{+0.020}_{-0.021}$	0.156
$4333-4413$	$0.09610^{+0.00127}_{-0.00133}$	$0.487^{+0.040}_{-0.044}$	0.240
$4413-4493$	$0.09838^{+0.00162}_{-0.00138}$	$0.549^{+0.027}_{-0.027}$	0.259
$4493-4573$	$0.09685^{+0.00089}_{-0.00079}$	$0.511^{+0.020}_{-0.022}$	0.262
$4573-4653$	$0.09737^{+0.00091}_{-0.00093}$	$0.555^{+0.019}_{-0.020}$	0.250
$4653-4733$	$0.09646^{+0.00086}_{-0.00070}$	$0.488^{+0.015}_{-0.014}$	0.259
$4733-4813$	$0.09513^{+0.00086}_{-0.00099}$	$0.469^{+0.022}_{-0.024}$	0.273
$4813-4893$	$0.09742^{+0.00273}_{-0.00267}$	$0.364^{+0.062}_{-0.066}$	0.356
$4893-4973$	$0.09469^{+0.00076}_{-0.00073}$	$0.397^{+0.033}_{-0.034}$	0.312
$4973-5053$	$0.09390^{+0.00059}_{-0.00062}$	$0.417^{+0.017}_{-0.016}$	0.274
$5053-5133$	$0.09453^{+0.00080}_{-0.00080}$	$0.394^{+0.016}_{-0.018}$	0.285
$5133-5213$	$0.09457^{+0.00083}_{-0.00080}$	$0.408^{+0.016}_{-0.015}$	0.287
$5213-5293$	$0.09482^{+0.00089}_{-0.00097}$	$0.349^{+0.020}_{-0.021}$	0.316
$5293-5373$	$0.09555^{+0.00129}_{-0.00126}$	$0.401^{+0.047}_{-0.049}$	0.312
$5373-5453$	$0.09168^{+0.00125}_{-0.00135}$	$0.258^{+0.041}_{-0.048}$	0.356
$5453-5533$	$0.09495^{+0.00060}_{-0.00051}$	$0.303^{+0.017}_{-0.017}$	0.311
$5533-5613$	$0.09233^{+0.00110}_{-0.00118}$	$0.293^{+0.026}_{-0.028}$	0.321
$5613-5693$	$0.09462^{+0.00162}_{-0.00186}$	$0.357^{+0.041}_{-0.041}$	0.319
$5693-5773$	$0.09371^{+0.00090}_{-0.00101}$	$0.275^{+0.028}_{-0.029}$	0.339
$5773-5853$	$0.09362^{+0.00089}_{-0.00094}$	$0.288^{+0.026}_{-0.026}$	0.311
$5853-5933$	$0.09608^{+0.00090}_{-0.00079}$	$0.256^{+0.040}_{-0.034}$	0.340
$5933-6013$	$0.09358^{+0.00117}_{-0.00119}$	$0.270^{+0.033}_{-0.033}$	0.344
$6013-6093$	$0.09363^{+0.00073}_{-0.00076}$	$0.214^{+0.025}_{-0.025}$	0.340
$6093-6173$	$0.09384^{+0.00133}_{-0.00107}$	$0.248^{+0.026}_{-0.027}$	0.320

varies with wavelength with the white flux roughly resembling the average effect of this phenomenon. Bluer light curves are affected more strongly by extinction and so they tend to keep the natural convex form as the common mode flux is unable to change their shape. Meanwhile, redder light curves are less affected by scattering and absorption in the atmosphere and as a result common mode flux prevails over them inverting their form. In the transition zone between shorter and longer wavelengths the light curves are flattened as the difference in the extinction indexes is nearly nullified. To ease calculations, we transform this ratio of exponentials into one exponential and so the ratio is converted into a difference between indexes. This difference can then effectively be expressed by one free coefficient in the fit that defines the curvature.

5.4 Blue-optical transmission spectrum

Figure 5.17 illustrates the blue-optical trend of the relative planet-to-star radius ratio with respect to wavelength according to four of the five spectroscopic methodologies described in Section 5.3.2 (the global approach is excluded here due to its great similarity with the classic approach and will, therefore, be shown in a separate graph). The produced transmission spectra of WASP-74 displayed a consistent offset between the June and August data sets in all approaches considered. Such an offset is to be anticipated given that the white light curve analysis resulted in slightly different values for the median $R_{\rm p}/R_*$ for each data set and so the transit models used in the common-mode corrections were also somewhat dissimilar. For the favoured new approach outlined in Section 5.3.2.5, this minor offset was determined to be equal to $\Delta R_{\rm p}/R_* = 0.0033 \pm 0.0037$. A quick check on all other approaches revealed a practically identical result with the only variable quantity being the uncertainty of the measurement. The reason for the offset may lie on differences between the observing conditions during the two nights. The data from August are likely more influenced by thin clouds while the data from June are likely more affected by increased precipitable water vapour (see Figure 5.4 for the variability in precipitable water vapour). Ultimately, the offset

was applied on the June data set, with the first column of Figure 5.17 depicting the corrected spectrum for this data set. The other column shows the combined transmission spectrum obtained through calculation of the weighted mean $R_{\rm p}/R_*$ between the two observations. The effect of the offset on the weighted mean is not significant and similar conclusions can be drawn in the case where the offset is not considered.

The combined VLT spectrum is plotted against the corresponding result from the space-based observations using HST STIS (Fu et al., 2021). It can be easily observed that the VLT data based on the new approach agree quite well with the HST transmission spectrum. Only a closer inspection shows that the VLT data on average are slightly higher at the short end and lower at the long end of the blue-optical spectrum. Nevertheless, the significance of these differences is below 1σ. Furthermore, it is worth noting here that all limb darkening coefficients were fixed in the HST analysis whereas in the VLT analysis one limb darkening coefficient was allowed to vary freely during the fit. This change was necessary in order to correct for the potential impact of the Earth's atmosphere on the theoretical limb darkening values. The combined VLT transmission spectrum from the new method is then also compared with all published results in Figure 5.18.

In Section 5.2 we looked at the spectroscopic transit light curves and deduced that atmospheric escape is not evident from the data. This observation prompted us to explore the light curves through the usual methods without taking mass loss into account. Since we now also have an idea of the transmission spectrum, we can confirm our early conclusion that WASP-74b does not show an intense absorption signature, characteristic of atmospheric evaporation, in the blue. All transit depth values are found to be much lower compared to the U-band measurements from Mancini et al. (2019) and projections into the near-ultraviolet range reveal a mismatch between the VLT data and the U-band estimates. This disagreement suggests that the U-band results are likely a product of instrumental errors. Moreover, the VLT results are supported with the observations from Başt ürk et al. (2022), who did not detect transit timing variations. Such a detection could imply, among other things, a deteriorating orbit for WASP-74b. Such orbits are linked to escaping atmospheres for close-in exoplanets and

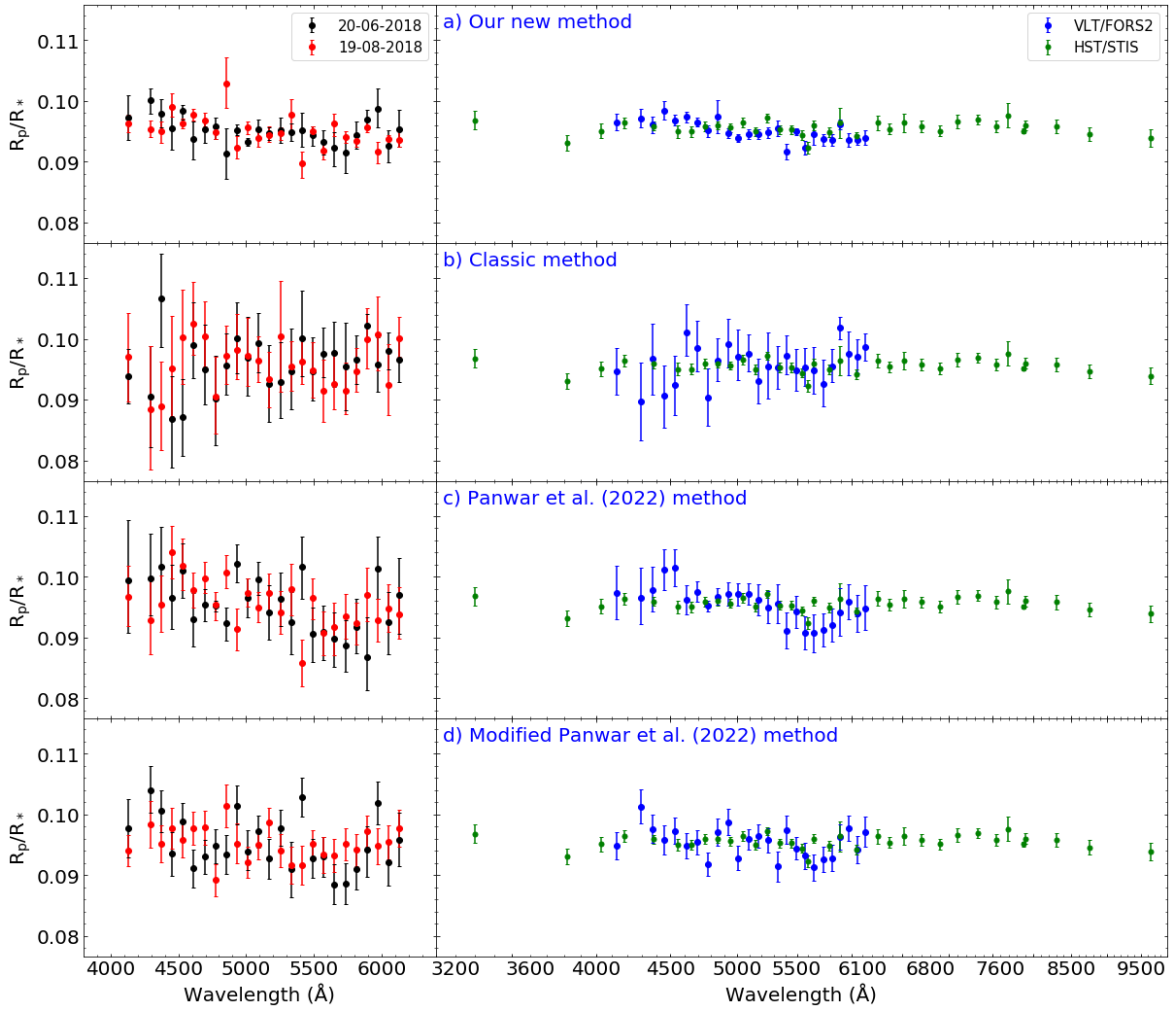

Figure 5.17: The transmission spectrum of WASP-74b according to four out of five different methodologies considered. On the left, the two individual data sets are shown, with black circles indicating the June data set and red circles indicating the August data set. A vertical displacement is administered to the transmission spectrum from June to place it closer to the result from August. On the right, the combined VLT/FORS2 spectrum (blue dots) is depicted together with the HST/STIS spectrum (green dots). Starting from the first row, the figure shows the transmission spectra from a) the new approach, b) the classic approach, c) the Panwar et al. (2022) approach, and d) the modified Panwar et al. (2022) approach.

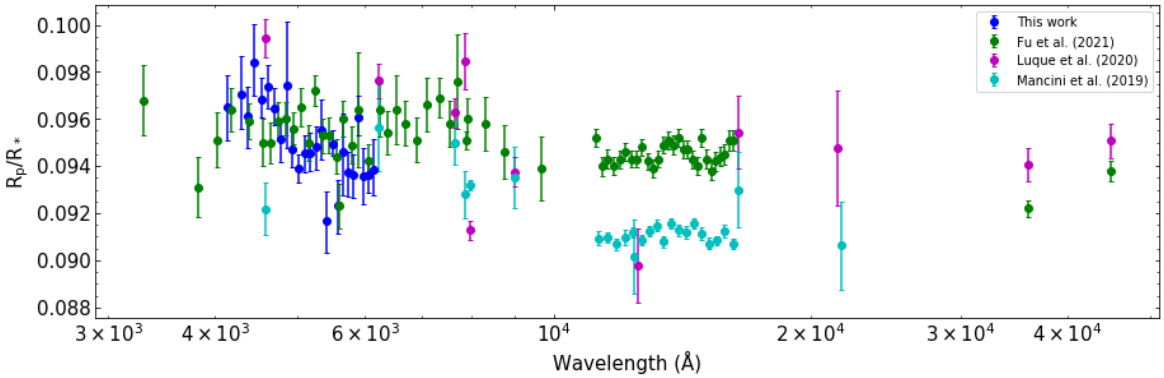

Figure 5.18: The WASP-74b transmission spectrum from this study (dark blue data points) plotted against results from Fu et al. (2021) (green data points), Luque et al. (2020) (magenta data points), and Mancini et al. (2019) (cyan data points).

could, therefore, leave marks in the ultraviolet and possibly in the blue-optical domain of the electromagnetic spectrum.

Going back to the behaviour of the VLT spectrum with respect to the HST one, we noticed that the VLT spectrum is slightly inclined compared to the HST data. In fact, if we take a closer look at the VLT data in Figure 5.18, we can clearly observe a slope that increases towards bluer wavelengths. The scattering slope seems to be quite sharp and is analogous to the slope determined photometrically by Luque et al. (2020). In contrast to Mancini et al. (2019), however, the strong absorption signatures from metal oxides such as TiO and VO cannot be seen. This finding suggests that the atmosphere of WASP-74b likely does not contain a high abundance of molecular species that absorb strongly in the optical.

The enhanced scattering slope of the VLT data motivated the application of a scattering fit to determine the steepness of the inclination. The fit is essentially the same as the one described for the WASP-88b data (see Section 4.4) and is based on the formalism of Lecavelier Des Etangs et al. (2008a). From the scattering fit, one can obtain parameter γ that defines the gradient of the scattering slope and is proportional to the atmospheric scale height H. If Rayleigh scattering is assumed, then parameter

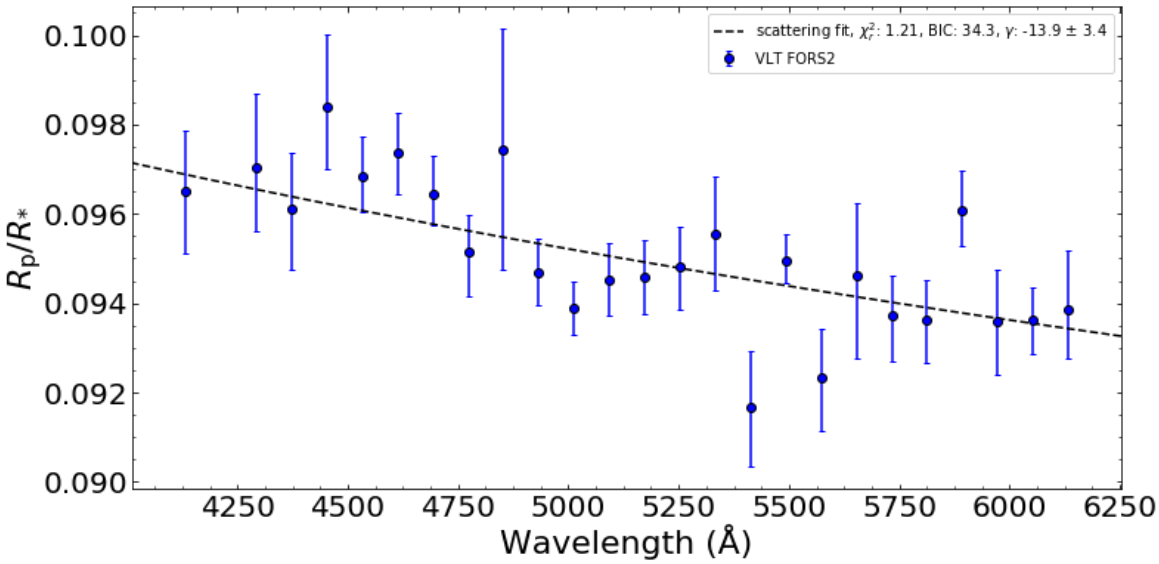

Figure 5.19: The constructed transmission spectrum of WASP74b from the new method (blue data points) along with a scattering fit based on the formulation of Lecavelier Des Etangs et al. (2008a) (black dashed line). The steepness of the slope suggests enhanced haze in the atmosphere of WASP-74b.

γ is equal to -4. To estimate γ, the scattering fit was applied on the entire blue-optical transmission spectrum and included two free parameters: slope and intercept. Figure 5.19 shows the scattering fit on the transmission spectrum obtained from the new approach. The fit revealed a steep slope equal to $\gamma = -13.9 \pm 3.4$ and yielded a χ_r^2 value of 1.21 or equivalently a BIC value of 34.3. This result indicates that the slope in the transmission spectrum of WASP-74b is nearly the same but slightly steeper than the one found for WASP-88b. According to these observations, WASP-74b also joins the growing sample of hot gas giant exoplanets that demonstrate enhanced scattering in their atmosphere (see e.g. Alderson et al., 2020; Ahrer et al., 2022).

Figure 5.20 illustrates the effect of the global approach on the observed transmission spectrum in comparison with the outcome from the classic method. The shape of the spectrum from the global approach remains similar to the corresponding shape from the classic approach, while the precision in the error bars is decreased. This ob-

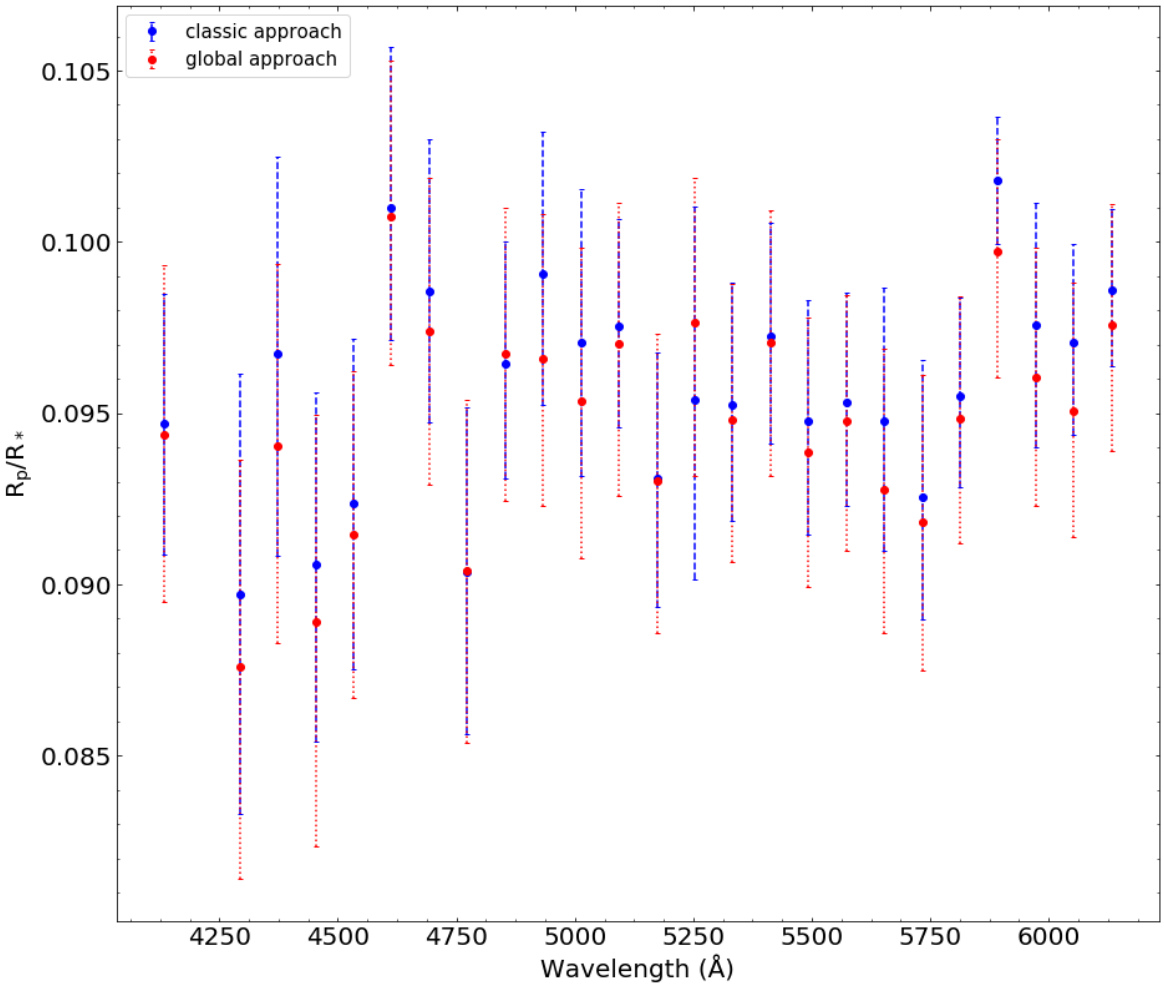

Figure 5.20: Comparison between the transmission spectra from the classic and the global approach. The blue data points with dashed error bars show the result from the classic method, while the red data points with dotted error bars show the result from the global method.

servation suggests that while the global approach may be a useful alternative, it may not always provide the greatest precision.

In Figure 5.21, a transmission spectrum from the May 21st data set is also included. This additional transmission spectrum is used to create a combined transmission spectrum from all three data sets after application of an offset and computation of their weighted mean. The original May 21st data set exhibits a much higher average for the $R_{\rm p}/R_*$ values, giving a large vertical displacement of $\Delta R_{\rm p}/R_* = 0.0071 \pm 0.0038$. In addition, many error bars are unrealistically low for the level of cloudiness during this observation. One reason for the lower error bars may be the insufficient detection and removal of all outliers. This means that the remaining outliers could be influencing the fit towards unrealistic values as the GP model tries to flexibly fit the data. The outcome was found to be very sensitive to the systematics considered and, hence, this data set was not included in the main analysis.

To complement the classic analysis, solutions founded on baseline polynomials were also explored. These solutions were obtained through similar approaches to the ones described in Sections 4.3.2.1 and 4.3.2.3 for WASP-88b. Figure 5.22 presents the resulting transmission spectra from the fully deterministic and partially deterministic approaches in comparison to the result from the classic approach. As expected, the precision is somewhat higher when using parametric functions to decorrelate noise in the spectroscopic analysis. Nevertheless, these solutions were discarded as the VLT FORS2 light curves of WASP-74 showed increased systematic effects that could not be addressed effectively by polynomial functions alone.

A few further tests were performed to check the consistency of the new method with respect to different systematic parameterisations and variations in the number of free transit parameters in the white light curve analysis. Figure 5.23 demonstrates the reliability of the new method and showcases the relative stability of the features for two of these cases. In the first case, the white flux systematics are assumed to be the same as in the main analysis, while the free transit parameters were five as a/R_* and i were also allowed to vary freely in the initial white-light fit. This parameterisation led to a slope that closely matches the finding from the approach of Section 5.3.2.5. A

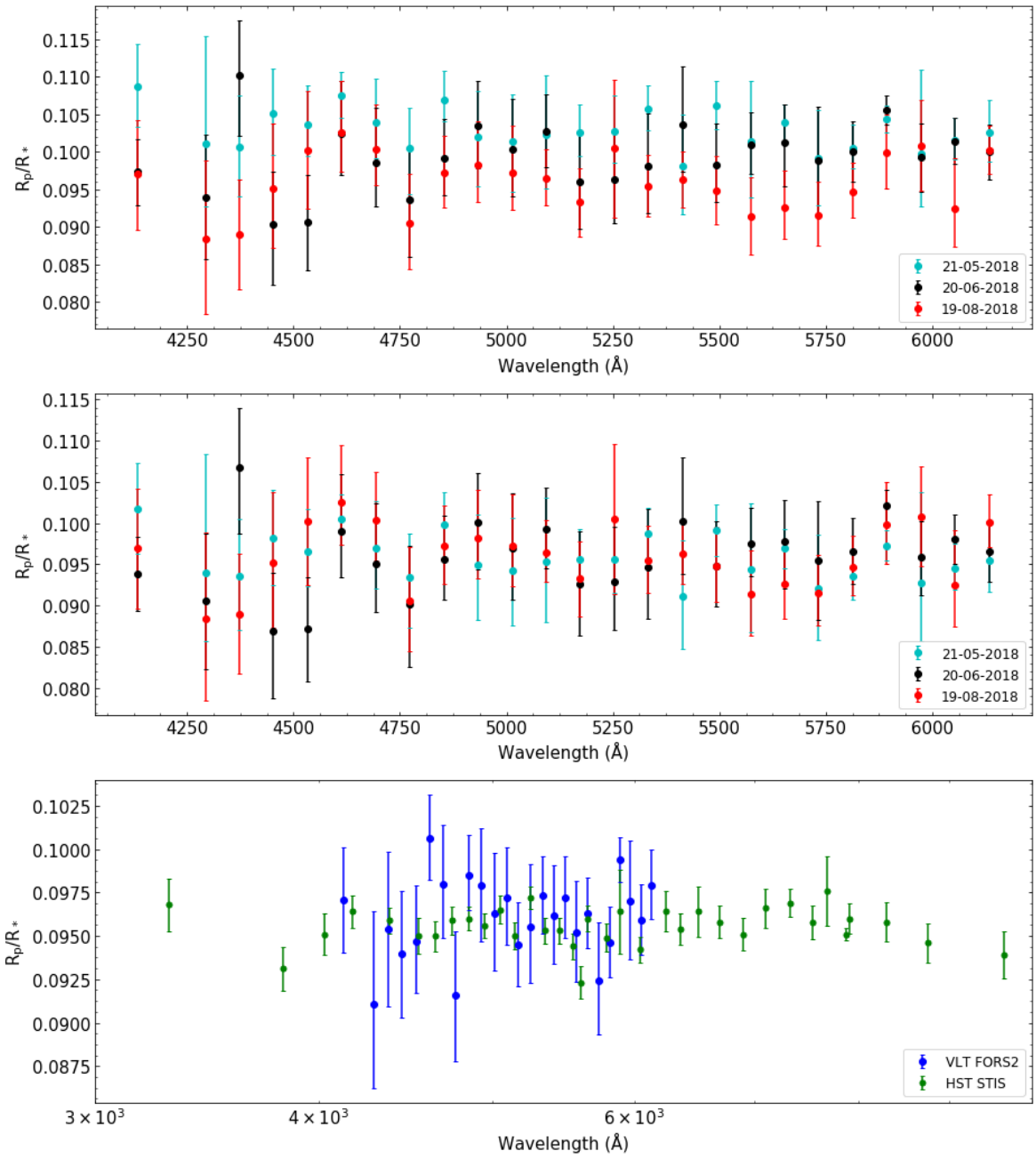

Figure 5.21: Transmission spectrum of WASP-74b from all three data sets. The first two rows show the separate data sets, with cyan depicting the May observations, black depicting the June observations, and red depicting the August observations. In the first row, the initial result is shown, whereas, in the second row, an offset is applied on the May and June data sets. On the last row, the combined transmission spectrum is plotted against the result from HST.

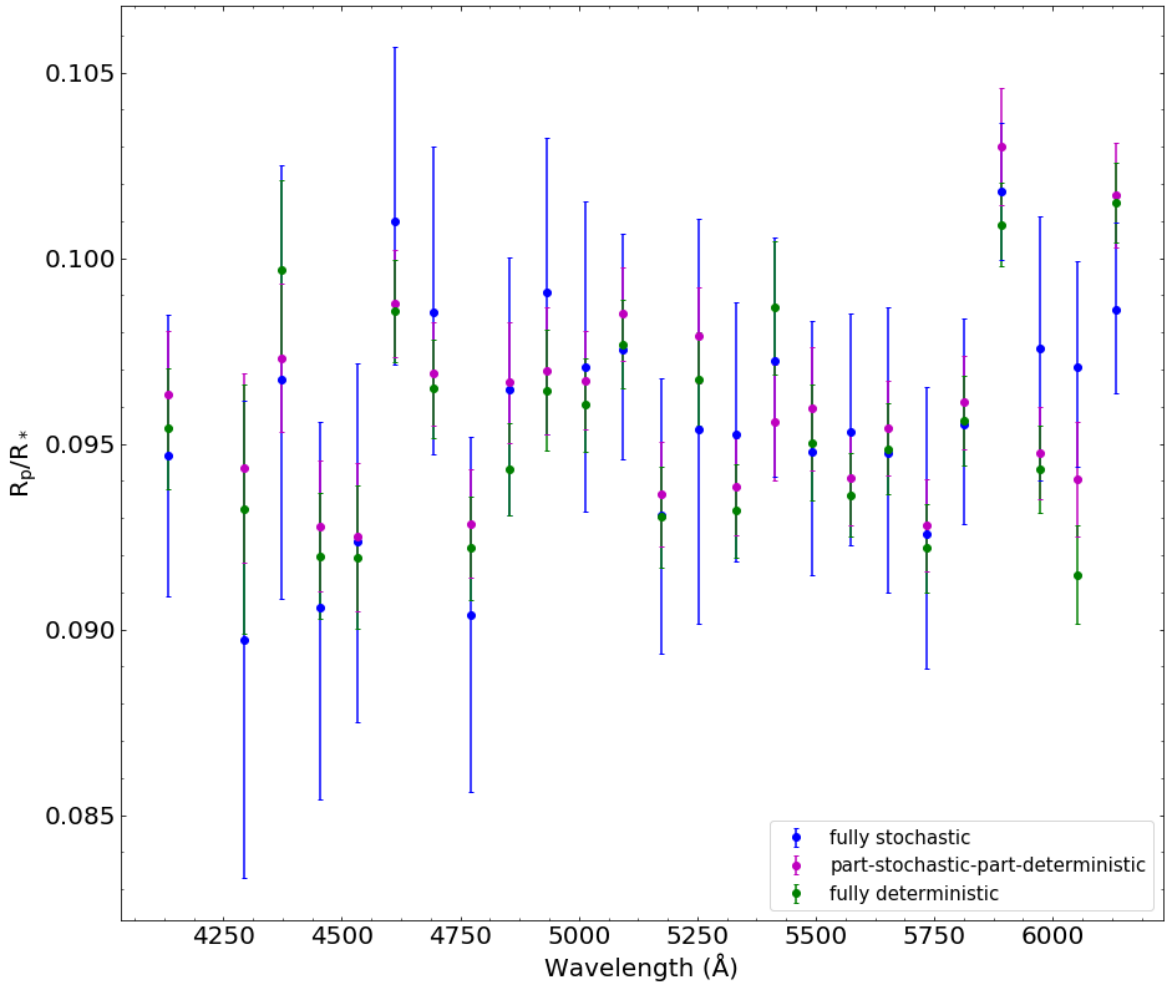

Figure 5.22: The transmission spectrum of WASP-74b as obtained from the fully stochastic approach (blue), the part-deterministic-part-stochastic approach (magenta), and the fully deterministic approach (green).

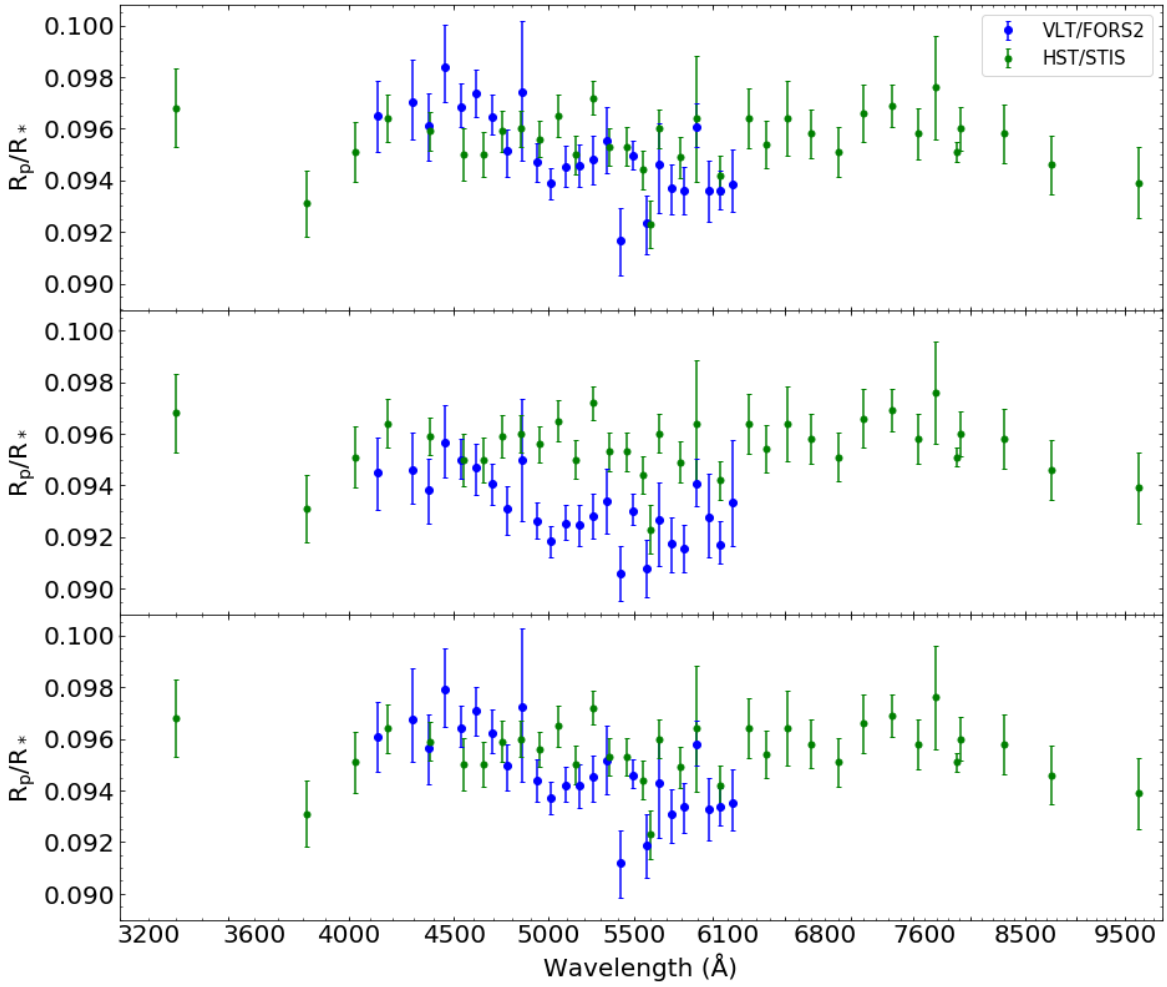

Figure 5.23: The transmission spectrum of WASP-74b based on different transit and systematic configurations. Each VLT configuration from the new method is shown in blue and a comparison is been made with the published HST data (green). An offset is applied on the June data set in all cases. Top row: The favoured transmission spectrum from the new method. Middle row: The transmission spectrum from a configuration where a/R_* and i are free in the white-light analysis. Bottom row: The transmission spectrum from a configuration where GP regressors of time are used in the white-light analysis.

small offset can also be observed in comparison to the favoured methodology, but this is a natural outcome given that the a/R_* and i assumptions were different leading to a minor deviation in the transit models used for the common-mode correction. In the second case, the physical regressors of the rotator angular velocity and the positional displacement in the cross-dispersion axis are replaced with regressors of time in the white-light GP analysis. Here, the free transit parameters are the same as in the preferred model. This solution resulted in a more impressive agreement with the best-fit model revealing a nearly identical transmission spectrum. These findings indicate that the new method is robust against systematic effects and increase our confidence in the validity of the new approach.

5.4.1 Testing the reliability of the new approach with a toy target

Reproducible results are key to scientific progress. Seeing that the novel method produced precise results for WASP-74b, it was important to check whether this method also works well for other targets. This would reassure us that the approach is sound and would give us a higher certainty about the accuracy of the results from this method[4]. To this effect, an exoplanet with a relatively clear atmosphere and a published ground-based transmission spectrum in the optical was needed to investigate the impact of the new method.

WASP-39b was the ideal target for this purpose as it was also observed using VLT FORS2. The exoplanet was discovered by Faedi et al. (2011) and is an inflated hot Saturn with a somewhat lower equilibrium temperature compared to WASP-74b and WASP-88b ($T_{\rm eq} = 1116\,{\rm K}$). This means that the alkali metals Na and K should leave prominent signals in the optical under the assumption of a clear atmosphere (Fortney et al., 2008, 2010). Indeed, subsequent observations of the atmosphere revealed the

[4]Note that an agreement between results is not necessarily a sign of accuracy. Both results can be similar and precise but inaccurate. An agreement, however, will give us greater confidence that the result is closer to the truth.

presence of alkali metals (Sing et al., 2016; Fischer et al., 2016; Nikolov et al., 2016), as well as water in the infrared (Wakeford et al., 2018; Tsiaras et al., 2018; Kirk et al., 2019). The features in the optical are less strong than expected in the ground-based observations and may indicate a thin haze/cloud layer higher in the atmosphere or limitations due to the Earth's atmosphere. More recently, carbon dioxide was also detected in the infrared using JWST (JWST Transiting Exoplanet Community Early Release Science Team et al., 2023), making it the first exoplanet with such a detection through this technique. The target was shortly after scrutinised further by JWST, confirming the earlier detection of CO_2, establishing the presence of H_2O, Na, K and non-grey clouds, and revealing the existence of CO, as well as the existence of photochemically produced SO_2 (Rustamkulov et al., 2023; Alderson et al., 2023; Ahrer et al., 2023; Tsai et al., 2023; Feinstein et al., 2023). The paper to note from this short introduction is the one from Nikolov et al. (2016) as their VLT FORS2 data were the canvas on which the new method was tested to produce a new transmission spectrum.

Instead of starting the analysis from the raw data provided by the ESO archive, I made use of the processed data provided by my collaborator Nikolay K. Nikolov. These reduced data included the raw spectroscopic light curves of the target based on the exact calibrations made by Nikolov et al. (2016). A white light curve analysis was not implemented in this case as all values for the transit model used in the common-mode correction were acquired from the published tables (see Nikolov et al., 2016). The rationale for the use of these data instead of carrying out all reductions from scratch is to ensure that the analysis will be closer to the analysis performed by Nikolov et al. (2016). Nevertheless, it should be noted that this analysis is not applied on the target-to-reference star light curves and so results will naturally be slightly different. After following the precepts of the novel method outlined in Section 5.3.2.5, an updated transmission spectrum was generated.

Figure 5.24 displays the result from this analysis in comparison to the transmission spectrum reported in Nikolov et al. (2016). For simplicity, data in the overlapping regions have been merged based on the weighted mean $R_{\rm p}/R_*$ from the two data sets (i.e. blue and red, or, equivalently, GRIS600B and GRIS600RI). A small offset of

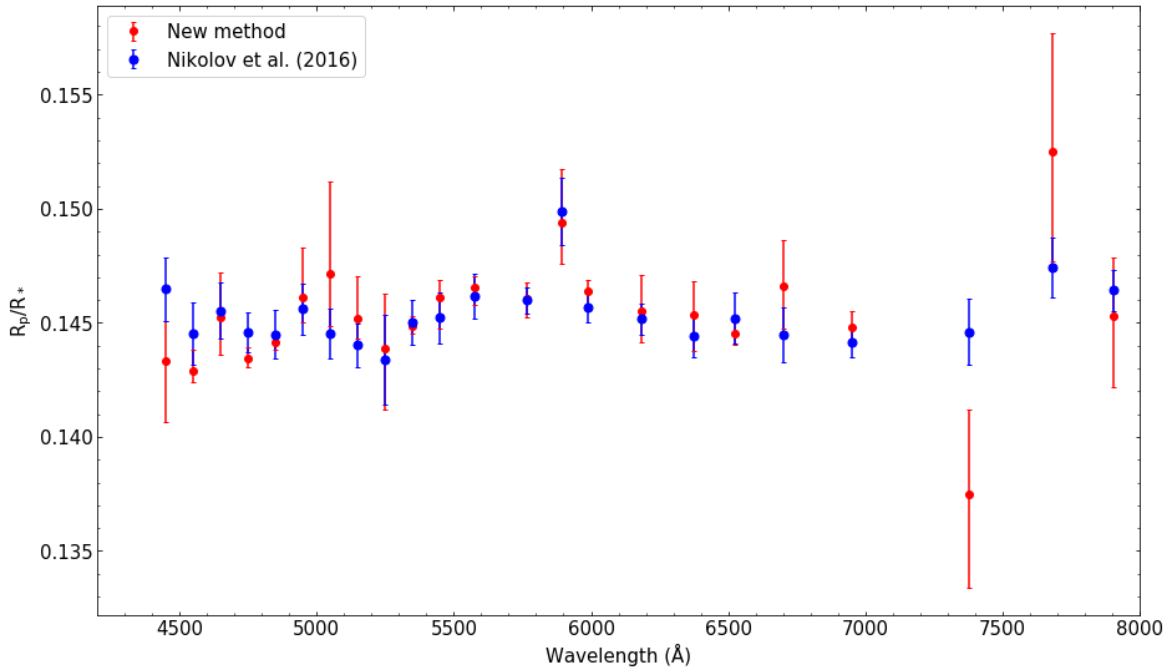

Figure 5.24: Application of the new method on VLT FORS2 data of WASP-39b. The resulting transmission spectrum (red data) compared to published results (blue data). After consideration of a small offset the agreement is quite good. The deviation in the antepenultimate data point is due to the linear characteristics of that particular spectroscopic light curve.

$\Delta R_{\rm p}/R_* = 0.0014 \pm 0.0022$ was applied on the data from the new method. Such an offset is to be expected given the different methodologies followed. One can quickly observe that there is a very good agreement between the two outcomes and that the weak to moderate features of the alkali metals are visible in both spectra. The alkali metals are resolved at 2.5σ (moderate, for Na) and 1.6σ (weak, for K). The significance is slightly lower compared to the findings from Nikolov et al. (2016) (3.2σ and 1.7σ, respectively). There are two reasons for this minor discrepancy: one is the approach taken in Nikolov et al. (2016) and two is the shape of the spectroscopic light curves in some cases. On the first matter, it should be clarified that Nikolov et al. (2016) disentangled the noise from the transit using baseline polynomial functions instead of GPs. As already mentioned previously, such an approach tends to reduce errors significantly (but may not account for all systematics). The second reason is that the form of the spectroscopic light curves approaches linearity when the difference in the exponential indexes is close to zero (see e.g. the ninth light curve in the second column of Figure 5.16). This can lead to a departure of the estimated value from the actual value (because the exponential function struggles to follow the linear pattern), a behaviour that the GP tries to correct by increasing the uncertainties. Contrary to the WASP-74b investigation, the analysis for this target was not extensive and additional corrections could potentially result in improved results for the transmission spectrum of WASP-39b. This is because a deeper inspection of the atmosphere of WASP-39b was not the purpose of this study and so it was not pursued.

5.5 Discussion

5.5.1 Advantages and disadvantages of the new method

The new methodology detailed in Section 5.3.2.5 is the combined product of several different approaches developed through the years to model data for exoplanets observed in transmission. At the same time, it is characterised by some novelties that clearly

distinguish it from all previous analyses. More specifically, the new method is based on the widely used practice of common-mode correction in ground-based data, the use of the target flux only during the fitting process (a traditional space-based approach that was recently also employed for ground-based data, see Panwar et al., 2022), and the inclusion of an exponential curve to account for extinction (a method commonly applied for telluric corrections in high-resolution transmission spectroscopy). There are two key differences, however, from all prior investigations. One, the common-mode correction is applied on the raw light curves instead of the relative, target-to-reference star, light curves, and, two, extinction is taken into account as part of an exponential function of airmass in conjunction with other systematic effects through a Gaussian process. Both this kind of common-mode correction and the inclusion of an exponential of airmass to model the raw spectroscopic fluxes of the target are, to my knowledge, implemented for the first time in low-resolution transmission spectroscopy. It is, therefore, imperative to discuss the arguments for and against the use of this method.

Starting from the advantages, one obvious benefit from this approach is the non-inclusion of the flux from the comparison star in the spectroscopic analysis. This means the fits are only applied on the target light curves. Such an approach is not burdened by the additional scatter introduced by the flux of the comparison star, a factor that can reduce the precision of the measurements. In that regard, this technique is similar to Panwar et al. (2022), who also ignored the comparison star in their fits. As already mentioned in Section 5.3.2.5, the residual scatter is only $\sim$1.25 times higher than the photon noise limit, and the same is found for the methodologies that rely on the Panwar et al. (2022) prescription.

Furthermore, the new method takes into account noise contributions from the extinction of light within the Earth's atmosphere. This is achieved through the inclusion of an exponential function of airmass in the mean function of the spectroscopic GP fits. Such an exponential function has been used to model extinction in target-to-reference star light curves in the past (see e.g. Sedaghati et al., 2015), but this is probably the first time an exponential is applied to the spectroscopic light curves of the target alone. This idea greatly aids the fit, resulting in reduced uncertainties for the resulting trans-

mission spectrum. The improvement is evident not only in the new approach, but also in the modified Panwar et al. (2022) approach. The shape of the transmission spectrum is similar in both cases, as well as in comparison with the unmodified Panwar et al. (2022) approach, and reveals a steep scattering slope for fits applied on the spectroscopic light curves of the target. This agreement between methods with key differences is further evidence that the new method has merit.

The most notable advantage of this new method, however, is the important increase in precision, with uncertainties being three and a half times lower compared to the classic method, two and a half times lower compared to the Panwar et al. (2022) method, and two times lower compared to the modified Panwar et al. (2022) method. The precision is, thus, remarkably good and only slightly lower with respect to the space-based results from HST. The result is more impressive when factoring in that the space-based values were estimated using baseline parametric functions whereas here we followed the more cautious route of GPs. A reason for this excellent performance is the smoother form of the common-mode corrected spectroscopic light curves in the new approach. In contrast, the Panwar et al. (2022) method is applied on the raw spectroscopic light curves, which tend to be characterised by discontinuities due to cloud crossings or other systematics effects. As a result, the GP fits, in an attempt to evaluate the systematic errors with time and common mode as regressors, struggle more in modelling the behaviour of the data. Meanwhile, in the new method, the linear procedure of common-mode correction efficiently removes most of the interference caused by aerosols due to the homogeneous nature of the thin clouds sporadically affecting the data. There is, therefore, no need for these disruptions to be considered as unaccounted systematics by the GP fits.

Although the new method is a great alternative to standard approaches, there are some pitfalls that need to be taken into account while making use of this method. First and foremost, the new method can effectively be applied on light curves that exhibit a strong correlation with extinction. This implies that the method is limited to ground-based observations, which show a characteristic exponential curvature due to attenuation phenomena within the Earth's atmosphere. The effect is wavelength-

dependent, a function of airmass and visible in the raw light curves of the target, as well as in the raw light curves of other stars in the field of view. Preliminary findings indicate that such a method is less efficient in light curves that present linear characteristics, a logical outcome considering the curvy nature of the exponential function.

Moreover, this new approach can be inhibited by the use of the common-mode correction itself. In particular, uncorrected for light curve systematics can be propagated during common-mode correction from one light curve to another causing a possibly dangerous "domino effect". For instance, a cosmic ray affecting certain wavelength regions of the stellar spectrum can be accidentally incorporated in the common mode, introducing additional noise in other wavelength channels. Another example would be detector impurities that can make some recording areas more sensitive to light than others. This effect can lead to increased scatter in parts of the CCD chips that can be included in the description of the white noise in the common mode and can thus be disseminated to other wavelength bins. In the occasion of very strong systematics, the effects of this source of noise on the spectroscopic light curves can be significant. The inclusion of an outlier rejection algorithm is quite useful as it can adjust the time-series for noise contributions from cosmic rays but detector cosmetics can be more difficult to treat as they tend to affect data temporally. In the case of cosmic rays, an alternate option would be the application of cosmic ray removal algorithms (see e.g. van Dokkum, 2001; McCully & Tewes, 2019; Zhang & Bloom, 2020).

The methodology explored in this study is founded on the use of GP functions, a framework that has received a lot of attention in the last decade, and in the field of transmission spectroscopy specifically, due to its relevant success in explaining data that are affected by unknown parameters and are otherwise difficult to decipher (see e.g. Gibson et al., 2012). It is the successor of the traditional noise decorrelation method that employs baseline parametric functions to describe systematic trends in the data (see e.g. Gilliland & Arribas, 2003). Due to the rigidity of the traditional method and its inability to successfully account for telluric and other unknown systematic effects, its direct use on the light curves of planet-hosting stars was generally avoided. Instead, the standard approach was to correct for atmospheric effects by dividing the target flux

by the flux of a nearby comparison star, under the assumption that the atmospheric material in the path of the starlight is similar for objects with small angular separation. The inflexibility of the parametric approach in the light curves of the target can also be seen in the Levenberg-Marquardt fitting process described in Section 5.3.2.4. Some fits there were unable to effectively follow the shape of the out-of-transit data and this behaviour led to limiting such use for the computation of the initial systematic parameters. The introduction of GPs, however, changed this perspective, as GPs are more flexible and can, therefore, account for unknown systematics introduced by the Earth's atmosphere. The new method, as well as the methods based on the formalism of Panwar et al. (2022), rely on this concept to successfully produce a valid transmission spectrum.

Finally, it is worth emphasizing here that the refinement seen in the uncertainties of the planet-to-star radii ratios stems from the application of the new method on blue-optical observations for a certain target and the increase in precision may vary for other targets and for observations at different light frequencies. Despite some doubts in the degree of success, this novel technique can be a useful asset in the analysis of ground-based transmission spectroscopy data and can greatly aid in the understanding of the atmospheres of transiting exoplanets by providing tighter constraints.

5.5.2 Atmospheric retrievals with PLATON

Similarly to the WASP-88b analysis, the PLATON library (Zhang et al., 2019, 2020) was employed to model the transmission spectrum and to figure out what theoretical models predict about the atmosphere of WASP-74b. In contrast to the analysis presented in Chapter 4, details about WASP-74b's atmosphere were obtained through atmospheric retrievals and not through a set of forward models. This feature was selected this time because previous works on WASP-74b (Luque et al., 2020; Fu et al., 2021) also utilise the retrieval capabilities of PLATON. Again, the atmosphere is assumed to be isothermal throughout all pressure layers and the atmospheric composition is defined by equilibrium chemistry. In addition, the formation of condensate species at

Table 5.5: Atmospheric parameters for WASP-74b from the retrievals using PLATON.

Parameter	Prior	VLT	VLT, HST and Spitzer
$R_{\rm p}$ ($R_{\rm Jup}$)	$\mathcal{U}(0.5R_{\rm p}, 1.5R_{\rm p})$	1.33 ± 0.03	1.34 ± 0.03
$T_{\rm p}$ (K)	$\mathcal{U}$(950, 1900)	1602^{+193}_{-312}	1310^{+260}_{-228}
β	$\mathcal{U}$(0.1, 10)	$1.23^{+0.21}_{-0.17}$	1.36 ± 0.10
$\log f_{\rm scatter}$	$\mathcal{U}$(-4, 10)	$1.1^{+2.3}_{-2.1}$	$7.4^{+1.2}_{-1.8}$
$\log(Z/Z_{\odot})$	$\mathcal{U}$(-1, 3)	$0.30^{+0.95}_{-0.83}$	$1.18^{+1.06}_{-1.37}$
C/O	$\mathcal{U}$(0.05, 2)	$0.94^{+0.68}_{-0.56}$	$0.97^{+0.69}_{-0.58}$
$\log P_{\rm cloudtop}$	$\mathcal{U}$(-3.99, 8)	$4.0^{+2.6}_{-3.2}$	$3.7^{+2.7}_{-2.8}$
scatter slope	$\mathcal{U}$(0, 20)	$15.9^{+2.7}_{-4.4}$	$3.2^{+3.5}_{-1.1}$

specific layers of the atmosphere is taken into account through removal of the species from the expected pressure layer and all layers over it.

Two main retrieval versions were considered, but an extensive exploration of sub-cases was performed in order to establish the best-fit scenario for WASP-74b. The first version involved a retrieval on the combined VLT data set from the new method and the second version involved an atmospheric analysis of a more extended data set that included the VLT FORS2 observations, the HST STIS and WFC3 observations, and the Spitzer observations on the 3.6 and 4.5 μm passbands. A total of eight parameters were allowed to vary freely in the fit: the radius of the planet, the temperature at the limb of the planet, a scaling parameter for the uncertainties, the two haze parameters (scattering factor and slope), the metallicity, the C/O ratio and a nebulosity parameter defined by the cloud-top pressure. Log-uniform priors were used for the cloud-top pressure, the scattering factor and the metallicity, while uniform priors were used for all remaining parameters. The atmospheric simulations were coupled with a multimodal nested sampling algorithm (Speagle, 2020) consisting of 1000 live points. The best-fit atmospheric scenarios from these two cases are presented in Table 5.5, with their best-fit synthetic spectra illustrated in Figure 5.25. A comparative posterior distributions plot for these two atmospheric retrievals was also created to get a better idea of the trends in the fitted parameters. These results are shown in Figure 5.26.

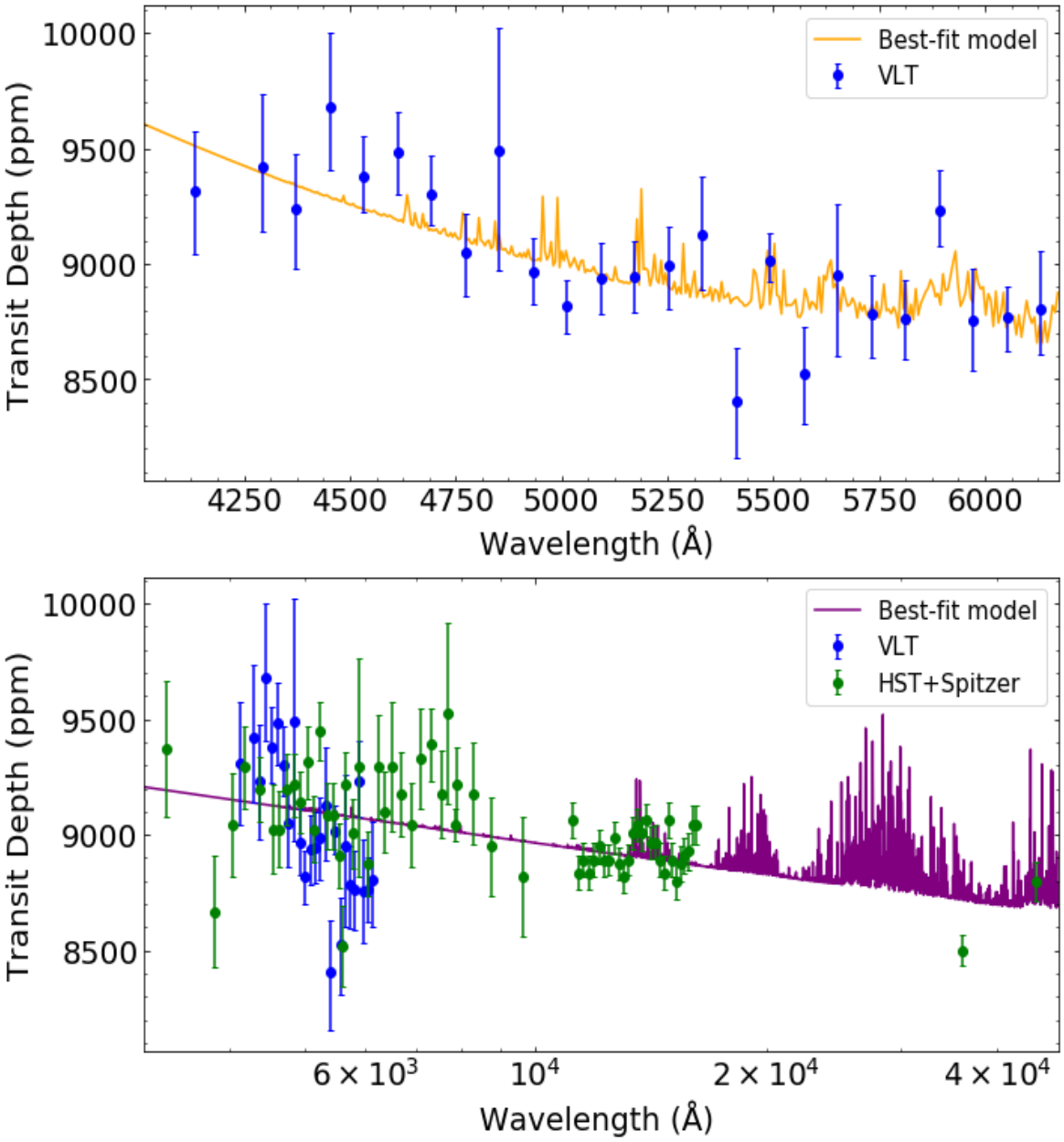

Figure 5.25: The transmission spectrum of WASP-74b compared to best-fit models from PLATON. Blue data indicate the VLT observations, while green data indicate the HST and Spitzer observations. Top: Best-fit retrieval (orange) on the transmission spectrum from the VLT specific analysis. Bottom: Best-fit retrieval (purple) on the transmission spectrum from the combined VLT, HST and Spitzer analysis.

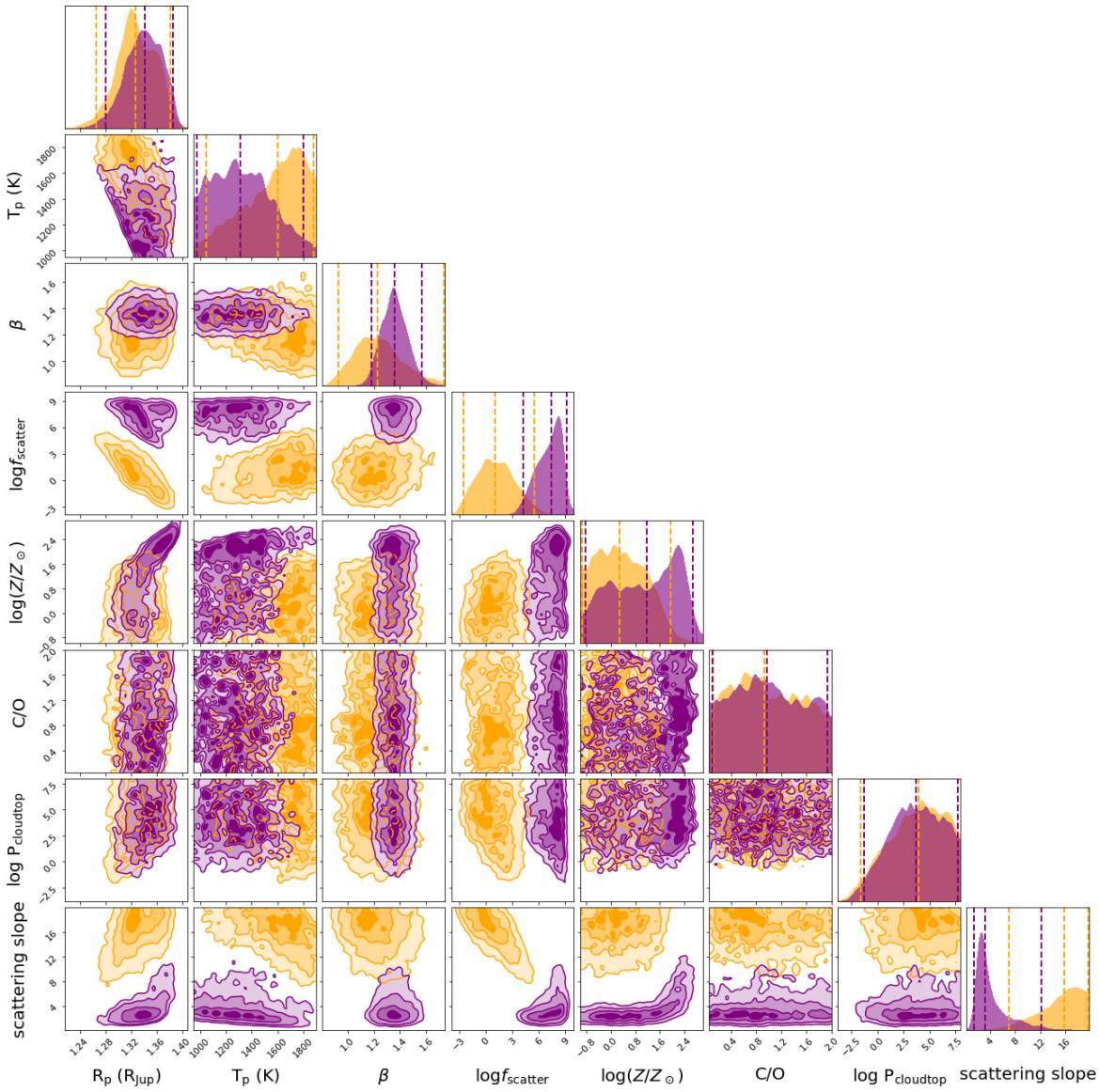

Figure 5.26: Posterior distributions of WASP-74b from the atmospheric retrievals using PLATON. Orange contours describe the retrieval result for the VLT observations, whereas purple contours describe the retrieval result for the combined VLT, HST and Spitzer observations.

From the retrieval analysis on the VLT data, a scattering slope of $15.9^{+2.7}_{-4.4}$ was determined, in excellent agreement with the scattering slope fit performed in Section 5.4. Furthermore, the distribution of the slope was found to be slightly skewed towards the top limit, a result also found by Luque et al. (2020). The log scattering factor was relatively low at $1.1^{+2.3}_{-2.1}$ and the temperature at the limb (1602^{+193}_{-312} K) was close to the equilibrium temperature of the planet ($\sim$1900 K). The value of the scattering slope is high, suggesting increased haze within the atmosphere of WASP-74b.

The picture is quite different when the published HST and Spitzer data are taken into account (Fu et al., 2021). In that case, the atmospheric parameters tend to assume values that are close to the values reported for the combined HST and Spitzer transmission spectrum (Fu et al., 2021). Two notable but minor discrepancies are the slightly increased estimate for the log scattering factor ($7.4^{+1.2}_{-1.8}$) and the slightly steeper scattering slope ($3.2^{+3.5}_{-1.1}$). Despite the rise in the values of the scattering parameters, the slope still agrees with the assumption of Rayleigh scattering and both estimates match the results from Fu et al. (2021) to better than 1σ. This outcome is normal considering the somewhat higher uncertainties of the VLT data set, the steeper slope indicated by the ground-based data and the narrower bandwidth of the VLT observations. For reference, the spectral range of the combined ground- and space-based spectrum is about 20 times larger compared to the range of the VLT spectrum alone. This difference highlights the wide spectral range of the combined HST and Spitzer data and explains the certain bias of this retrieval towards the values retrieved by Fu et al. (2021). It should be noted here that, while most of the parameters are consistent with the quantities reported in Fu et al. (2021), metallicity is found to be higher than the result from the space-based retrieval. Such a deviation is expected given the narrow metallicity prior placed by Fu et al. (2021) and the much wider prior used in this study.

To check the statistical significance of these findings, both retrievals were compared to two other atmospheric simulations and were evaluated through their Bayesian evidences. The two additional simulations were based on the assumption of fixed Rayleigh scattering and were split into a scenario where cloudiness is assessed as a free parameter in the fit and a scenario where clouds are practically minimised by fixing the

cloud-top pressure to infinity. For the VLT specific retrieval, the logarithmic difference in the Bayesian evidence $\Delta\log\mathcal{Z}$ was determined at 4.49 ± 0.24 for the first scenario and at 5.34 ± 0.26 for the second scenario. The statistical significance was higher in the combined ground- and space-based data, with Bayesian evidences revealing values of 8.73 ± 0.23 and 9.78 ± 0.24 for the two scenarios, respectively. A difference greater than five in the logarithmic evidences corresponds to a statistical significance greater than 3σ, indicating strong to very strong evidence in support of the initial retrievals (see e.g. Jeffreys, 1961; Kass & Raftery, 1995). Only the comparison between the unrestricted VLT retrieval and the fixed Rayleigh VLT retrieval showed a lower $\Delta\log\mathcal{Z}$ of 4.49, which in reality is also quite close to the 3σ mark, suggesting that the basic VLT retrieval is also strongly favoured over the solution with fixed Rayleigh scattering. Finally, the combined VLT, HST and Spitzer data set was also compared against a cloudy model with fixed Rayleigh scattering and a fixed cloud-top pressure of 0.001 Pa. This retrieval resulted in the highest $\Delta\log\mathcal{Z}$ difference ($\Delta\log\mathcal{Z} = 10.65 \pm 0.22$), demonstrating that the hazy model is a much better fit to the combined data than the cloudy model.

A combined retrieval with an offset to the VLT FORS2 data set was also examined briefly but the results were quite similar to before as the offset was found to be small (132^{+53}_{-56} ppm).

5.5.3 Atmospheric retrievals with AURA

Additional retrievals with AURA were also performed on the combined ground- and space-based data set for completeness. This retrieval framework, first presented in Pinhas et al. (2018), allows for free chemistry retrievals and, hence, abundance calculations (see also Chapter 4 for more information). This distinctive property compared to PLATON provided a different perspective on WASP-74b's atmosphere and helped in creating a better picture of the atmospheric structure and composition.

Contrary to Chapter 4, where the main retrieval analysis with AURA focused on the most prominent chemical species, here the primary retrieval was based on over a

dozen species. The increased precision of the WASP-74b data set, enabled some broad constraints on the abundances of some species rendering multi-species retrievals more informative. More specifically, opacity contributions from the following species were considered: CH_4 (Yurchenko & Tennyson, 2014), NH_3 (Yurchenko et al., 2011), HCN (Barber et al., 2014), Na and K (Welbanks et al., 2019), as well as H_2O, CO and CO_2 (Rothman et al., 2010). Due to the high equilibrium temperature of the exoplanet, heavier species were also included in the simulations. These were the metal oxides TiO (McKemmish et al., 2019) and VO (McKemmish et al., 2016), and the metal hydrides FeH, CrH and TiH (Bauschlicher et al., 2001; Dulick et al., 2003; Burrows et al., 2005; Tennyson et al., 2016). In addition, the impact of AlO (Patrascu et al., 2015) was also investigated in a separate retrieval.

The main retrieval analysis resulted in a small vertical offset (similar to the one found using PLATON) for the VLT FORS2 data set and revealed a fit with an enhanced scattering slope and muted overall absorption signatures. Figure 5.27 shows the AURA retrieval fit that indicates the presence of haze in the atmosphere of WASP-74b. The haze parameters in this case were determined at $\gamma = -7.2^{+2.5}_{-1.9}$ and $\log(a) = 7.2^{+1.6}_{-2.3}$. In addition, the retrieval found that the fractional coverage of clouds and hazes in the day-night transition region is equal to $\phi = 0.6^{+0.2}_{-0.2}$. The posterior distributions of the retrieved chemical abundances are depicted in Figure 5.28. From there, it can be seen that the volume mixing ratios of most species are very broadly constrained with the best-defined peaked distributions being observed for H_2O and K. The 2σ upper limits for the abundances of these two species were determined at $10^{-2.5}$ and $10^{-2.9}$, respectively. Interestingly, CO also shows a peak towards the right end of the distribution, resulting in a 2σ upper limit of $10^{-1.8}$. This distribution, however, is largely unconstrained towards the lower end.

In the retrieval that incorporated AlO in the atmospheric simulations, the scattering slope was found to be slightly increased compared to the main retrieval, giving $\gamma = -13.7^{+6.1}_{-4.2}$. At the same time, the haze enhancement factor and the cloud and haze coverage parameter remained similar agreeing with the previous AURA estimates at the 1σ level ($\log(a) = 5.5^{+1.8}_{-1.3}$ and $\phi = 0.6^{+0.2}_{-0.2}$, respectively). Surprisingly, the vol-

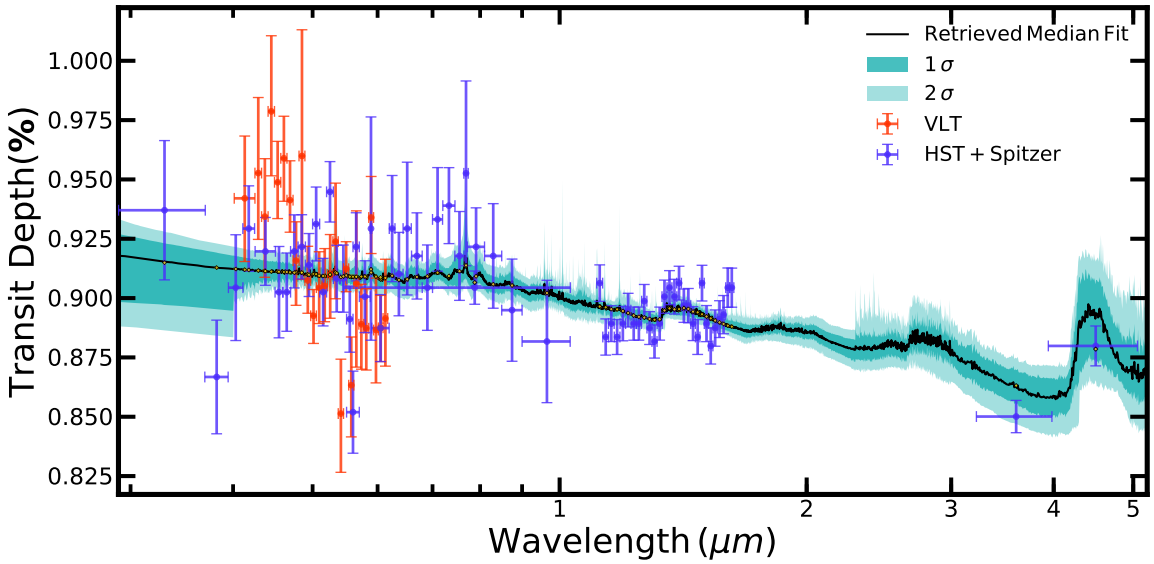

Figure 5.27: Transmission spectrum of WASP-74b and model spectral fit from the retrieval with AURA. The median fit is indicated by a black line, while the 1σ and 2σ uncertainty regions are indicated by darker and lighter shades of turquoise.

ume mixing ratio of AlO was found to be around $\sim 10^{-3.5}$. This value is $\sim$5 orders of magnitude higher than equilibrium estimates for AlO in the temperature range of WASP-74b. The result is influenced by multiple weak absorption features in the optical that are partially hidden by the strong effect of haze. Nevertheless, it is highly unlikely that the abundance of AlO is so high and it is more likely that the outcome is affected by minute noise contributions.

A brief exploration of stellar activity found that the presence of starspots could also explain the observed slope[5]. In this case, the spot coverage would have to be small leading to weaker abundance constraints on most retrieved species. An exception here is the precise result for the volume mixing ratio of VO, but the simultaneous weakening of all other abundance constraints make this outcome quite unrealistic. The fact that long-term photometry and ultraviolet spectroscopy show no evidence of stellar activity

[5]The same was found from a retrieval using PLATON.

Figure 5.28: The AURA posterior distributions of the retrieved volume mixing ratios from the main retrieval analysis of the combined VLT, HST and Spitzer observations.

(see Section 5.1.1) indicates the implausibility of this scenario and strengthens the assumption that WASP-74 is quiet.

5.6 Conclusion

In Chapter 4, I reported results for WASP-88b and presented the first transmission spectrum of this low-density exoplanet. Contrary to WASP-88b, WASP-74b is relatively well-studied and a few explorations of its atmosphere had already been performed prior to this study revealing a variety of atmospheric scenarios. More specifically, observations in the infrared by Tsiaras et al. (2018) indicated the presence of clouds or a lack of water at high altitudes, while combined optical and near-infrared observations by Mancini et al. (2019) determined strong absorption from mineral molecules such as TiO and VO. Follow-up observations by Luque et al. (2020) detected a steep slope in the optical and hinted at hazy conditions, but Fu et al. (2021) found a less significant slope that corresponds to Rayleigh scattering and affects wavelengths up to the near-infrared. A high-resolution search for metals detected a single aluminium line at greater than 5σ statistical significance, but failed to confirm the presence of additional lines, including the strongest one at 5558 Å (Lira-Barria et al., 2022).

In this study, the VLT FORS2 instrument was once again utilised to obtain ground-based spectroscopic observations of a planet-hosting star. This time, WASP-74 and three comparison stars were observed in the blue-optical wavelength region by employing the dispersive element GRIS600B. The stars were monitored for three nights, but one night (May 21st 2018) was later removed from the analysis due to thick cirrus clouds disrupting most of the observations. Furthermore, the brightest comparison star was also excluded from the analysis due to issues with the second CCD chip (see Section 5.2). The aim of this study was three-fold:

1. to assess whether the anomalous readings found by Mancini et al. (2019) in the U-band could be supported with new, spectroscopic observations in the blue-optical,

2. to probe the atmosphere of WASP-74b through transit spectroscopy based on spectroscopic data from the ground (which is a first time for WASP-74b), and

3. to determine effective solutions for the decorrelation of systematic errors.

Starting from the last point, a sequence of approaches were explored in order to minimise residual scatter and increase the precision of the transmission spectrum parameters. This exploration led to the development of a new method that avoids the use of a comparison star, resembles characteristics of a previously established method (Panwar et al., 2022), and takes atmospheric extinction into account. The foundation of this method lies on two key novelties:

1. the usual step of "common-mode correction" is performed, but only on the raw spectroscopic fluxes of the target, and

2. the effects of atmospheric attenuation are considered in the fits through the inclusion of an airmass exponential in the mean function.

The new method greatly reduced transit depth uncertainties and preserved the shape of the transmission spectrum as seen in the Panwar et al. (2022) analysis.

Using the two remaining data sets, a transmission spectrum was produced from the combined weighted mean. The transmission spectrum was then mapped against synthetic spectra from atmospheric retrievals using PLATON, revealing a sharp scattering slope towards the blue end of the VLT specific data set. This outcome is in agreement with previous ground-based observations by Luque et al. (2020) and is also confirmed through all approaches based on the Panwar et al. (2022) method. The situation is somewhat changed when considering the combined transmission spectrum from VLT, HST and Spitzer. In that case, the transmission spectrum only shows a Rayleigh scattering slope. This outcome is normal as the space-based data cover a wider wavelength range and so they prevail over the VLT data. Nonetheless, the haze parameters were found to be slightly higher compared to Fu et al. (2021), but were still within 1σ from the values reported therein.

Apart from PLATON, AURA was also employed to carry out atmospheric retrievals of the combined VLT, HST and Spitzer data set. These free chemistry retrievals allowed for a broader exploration of WASP-74b's atmosphere and resulted in wide constraints for the abundances of several chemical species, including water and some molecular metals. The retrievals with AURA also indicated a steep scattering slope, strengthening the hypothesis that WASP-74b's atmosphere is characterised by haze. In addition, one of the retrievals suggested a high concentration of aluminum oxide, but the solution was determined to be somewhat unrealistic as the mixing ratio of aluminum oxide was orders of magnitude higher than equilibrium values. Therefore, part of the observed features could be attributed to small noise effects.

All transit light curve investigations showed no evidence of extreme absorption in the form of much deeper transits in the blue. Furthermore, the steep slope found in some scenarios is gradual and smooth and cannot match the U-band observations by Mancini et al. (2019). This result indicates the outlying nature of the U-band data and suggests that the atmosphere of WASP-74b is likely not escaping at a significant mass loss rate. The U-band data were highly affected by sub-optimal atmospheric conditions during the observations, something that could have introduced systematic errors. This is noted by Mancini et al. (2019), who attribute the U-band anomalies to photometry systematics.

Due to the similarities with WASP-88b (e.g. enhanced haze in the atmosphere), a more detailed interpretation of the WASP-74b transmission spectrum will be presented in Chapter 6. Unlike the usual full optical range used in transmission spectroscopy, the WASP-74b VLT data presented in this study only explore the blue-optical region. Additional ground- and space-based spectroscopic observations in the red-optical and near-infrared will therefore be crucial in the understanding of the underlying physical phenomena that shape WASP-74b's atmosphere.

6 Summary and conclusions

Some of the discussion presented in this chapter was also included in the works of Spyratos et al. (2021) and Spyratos et al. (2023).

6.1 A brief evaluation of the methodologies considered

6.1.1 Parametric versus non-parametric approach

In the WASP-88b analysis, one of the key aspects of the investigation was to determine a fit that best describes the data by properly addressing the white and red noise contributions to the transit light curves. Therefore, a comparison was made between the two most commonly-used noise decorrelation techniques. The first one was the deterministic method, which depends on baseline parametric functions of various optical state parameters (Gilliland & Arribas, 2003), and the second one was the stochastic method of Gaussian Processes (Gibson et al., 2012), where noise is handled non-parametrically through a kernel function that defines the random effect of the optical state parameters. Both techniques were qualitatively assessed with respect to their impact on the transit light curves, their effectiveness in eliminating interference from auxiliary parameters and their reliability in the produced uncertainties.

Before I briefly describe the results from this comparison, it is worth pointing out that these findings are based on data sets that are profoundly affected by systematics that could not be entirely corrected through the normal set of optical state parameters. For example, in the red WASP-88b data, the source of the wavelike interference is unknown, whereas in the WASP-74b data, the impact of e.g. clouds cannot be accounted for by the given set of optical state parameters. Therefore, the findings presented here reveal tendencies in data that are corrupted by unaccounted for systematics, and so results may be different in situations where the systematic contributions

are well-understood and/or the data are cleaner.

One immediate conclusion from this comparison is that baseline parametric functions on their own are unable to follow trends in the data when the given set of optical state parameters are not fully responsible for the observed trends. The almost sinusoidal pattern seen in the residuals of the red data set of WASP-88b is a clear indication of the difficulties faced by the fitting algorithms. Furthermore, such an observation is not only limited in the WASP-88b data as challenges in the modelling of the light curves are also observed in the raw spectroscopic light curve fits of the out-of-transit WASP-74b data. In some cases, the parametric solution tends to go above or below the data making the fit unreliable. One could try to use high-degree parametric functions to fix this issue, but this may lead to potential over-fitting and, thus, does not actually solve the problem.

When a fit cannot successfully follow discontinuities in the data, then the obtained transit parameters tend to be unrealistic. A characteristic effect observed in such cases is the large scatter of the $R_{\rm p}/R_*$ parameter in the transmission spectrum. In some pathological cases, the scatter can be almost 10 times the size of the error bars. In other cases, very steep slopes may be induced (see e.g. Figure 4.14). These effects are normally the result of unaccounted for interference caused by temporal changes in the properties of the observing instrument or by variations in the Earth's atmosphere. Extra care should therefore be taken when applying the parametric technique in the decorrelation of systematics to avoid potential loss of valuable transit information.

One final remark on the parametric fits is the relatively low uncertainties and thus the very precise nature of the measurements. While this is often a desirable effect, the parametric functions can in some cases lead to results that deviate from the truth because of an ineffective evaluation of the systematics. This finding suggests that error bars may be prone to underestimation when noise factors are not well-established. Hence, it is necessary to fully understand all noise contributions when applying this method to transit light curves.

Contrary to the parametric functions, the stochastic nature of the GPs ensures a great degree of flexibility in the fit, resulting in an enhanced agreement between

the data and the model. More importantly, GPs are able to incorporate contributions from unknown systematics within the fit and are also able to assess their influence in the estimation of the uncertainties. In other words, in cases where the impact of unknown noise contributions is high, the errors are properly adjusted to account for any discontinuities in the data and are therefore increased. The lower precision of the measurements may seem discouraging at first, but the conservative errors give greater confidence about the genuineness of the result and are clearly a sign that underestimation effects are not a characteristic of this method. In addition, GPs were found to be less sensitive to the systematic model considered compared to the parametric approach and led to a greater consistency between results.

However, even the preferred approach of GPs comes with a weakness that needs to be taken into account during the fitting process. The power of the GPs may lie in their flexibility, but this can turn into a disadvantage if the prior assumptions are not well-defined. Consequently, it is important for the initial guesses to be as close to reality as possible. Moreover, the ranges for the prior distributions should be based on logical assumptions of the system and should occasionally exclude parameter space that may lead to unphysical values for the free parameters. Once the prior information is set up correctly, the GP approach can produce a reliable solution and eliminate any spurious features that may be generated by a fit that is too flexible.

6.1.2 Global approach

Two global approaches were considered in my research: one that was applied to the detrended spectroscopic light curves of WASP-75b and another that was based on a GP framework and was applied to the spectroscopic light curves of WASP-74b. The WASP-75b global approach was a simple fit of transit models applied to only 12 spectroscopic light curves consisting of relatively few data points (154). This fit did not take any noise contributions into account as noise was removed beforehand. For all these reasons, the WASP-75b fits were fast to compute but the quality of the data did not allow for a good precision despite the large size of the spectroscopic bins.

The WASP-74b global approach, on the other hand, has been shown to increase the precision of the measured parameters (see e.g. Ahrer et al., 2022). For the method to work, the length scale parameters τ from all light curves need to have similar values as τ is one of the shared parameters in the global fit. The WASP-74b classic approach revealed that τ is indeed quite stable showing only some minor variation from light curve to light curve (see Figure 5.8). This global approach led to a result that was in good agreement with the outcome from the WASP-74b classic approach, but reduced the precision in the transit depth by $\sim$15%. In addition, due to the increase in the number of free parameters and the MCMC chains, the calculations were found to be quite time consuming. For reference, the fitting process had a duration that was about five times higher than other approaches. For this reason, the approach was only briefly explored and a thorough investigation was not pursued.

6.1.3 Target-based approaches

During the spectroscopic analysis of WASP-74b three target-based approaches were considered. The first one is the approach developed by Panwar et al. (2022) and involved the stochastic treatment of both the common-mode and the wavelength-dependent noise component, the second one is a modified version of this approach and included an additional parametric function of airmass in the handling of the noise, and the third one is a novel approach first presented by me and involves a slightly-modified common-mode correction approach where the systematics correction is exclusively applied on the raw light curves of the target. The last method also benefits from the inclusion of an airmass exponential. These three techniques provided increased precision in the produced transmission spectrum compared to the classic method and so offer a viable alternative solution in cases where no suitable comparison stars are available (due to faintness, spectral type differences, variable extinction or detector failures).

One principal conclusion from all the analyses that incorporated an exponential of airmass in the mean function during the fitting process was that it improved precision without significantly altering the result compared to the simpler fit where no

exponential was used (see last two rows in Figure 5.17). The role of the initial out-of-transit Levenberg-Marquardt fit was also vital as it gave a hint about the curvature of the exponential. The improvement in precision was found to be significant and so this work favours the use of an exponential in ground-based observations that rely on the target flux only. The form of the exponential can be simple (like in this work) or more sophisticated (in cases where the effects of extinction are relatively well-understood).

Although the benefits and drawbacks of the new method were thoroughly analysed in Section 5.5.1, there is still room for improvement. In particular, the slightly erratic behaviour of the exponential in the transition zone where the light curves tend to be linear needs to be addressed. One immediate solution would be to replace the exponential of airmass with a linear function of time in those cases. However, for this to happen we would also need to devise a reliable test that successfully determines the cases where the linear fit outperforms the exponential. A better understanding of extinction and other systematics could also be useful as it could lead to more sophisticated models for the fitting process. In other words, if the contribution from an extinction-related physical parameter is known then its value can be incorporated in the index of the exponential function. This will probably make the exponential function more complex but at the same time we will not have to guess the origin of some of the systematics.

6.2 Results and context

In this thesis, I presented ground-based transmission spectroscopy results for a test target (Chapter 3) and two other more promising targets that were selected as part of a survey (Chapter 4) or as part of follow-up observations (Chapter 5). In the last two cases, the presence of aerosols at high altitudes was determined to be a significant obscuring factor in the observed transmission spectra. These findings confirm previous conclusions from population studies (Sing et al., 2016; Tsiaras et al., 2018; Edwards et al., 2022) that clouds and hazes are found ubiquitously among exoplanetary atmo-

spheres and demonstrate the important contribution of ground-based observations in deciphering the properties of planetary atmospheres beyond our Solar System.

The two main targets investigated here were the low-density hot Saturn WASP-88b (see Chapter 4) and the well-studied hot Jupiter WASP-74b (see Chapter 4). For WASP-88b, the obtained transmission spectrum was marked by the absence of any features, and a steep scattering slope that goes down towards longer wavelengths. Grids of forward atmospheric models were employed to distinguish between clear, cloudy and hazy atmospheres, and a model selection procedure using the BIC metric revealed a moderate to strong preference for haze in the atmosphere of WASP-88b. An independent atmospheric retrieval analysis by one of my collaborators also indicated the presence of haze and showed that the hazy scenario is slightly favoured over the cloudy scenario at 2.5σ significance. Overall, the slope found for WASP-88b is likely produced by phenomena that occur within the planetary atmosphere and is not a result of stellar conditions. This conclusion is supported by both atmospheric simulations and additional photometric and spectroscopic observations.

In the case of WASP-74b, the precision is higher as a result of the new approach outlined in Section 5.3.2.5, but the spectrum is similar and an enhanced slope going upward towards the blue is once again observed. One notable difference is that the VLT FORS2 spectrum in this instance has a smaller wavelength coverage compared to the WASP-88b observations. For this reason, the computations for the simulated atmospheres were also assisted by published space-based data obtained using HST and Spitzer (Fu et al., 2021). Several atmospheric retrievals were conducted both by me and a collaborator and all of them demonstrated the existence of haze in the atmosphere of WASP-74b. Depending on model assumptions the magnitude of the scattering slope may vary but the spectral gradient is there in all situations. Broad abundances were retrieved for several species, including water, potassium and carbon monoxide. One surprising result was the high abundance of AlO, which was mainly attributed to minor unaccounted for interference due to the small size of the spectral features and the partial masking from haze. In addition, the slope was also well-fitted by atmospheric retrievals that incorporate stellar effects. Given that this well-studied

system shows no signs of stellar activity through the usual observational means, it is highly unlikely that the spectrum is strongly affected by photospheric phenomena on the host star.

In Chapter 1, a potential anti-correlation between cloud and haze coverage and planetary temperature was briefly mentioned. The two exoplanets I examined here have relatively high equilibrium temperatures but do not confirm this cloud-temperature relation highlighted in some population studies (e.g. Stevenson, 2016; Fu et al., 2017). Furthermore, my findings are also in part or full disagreement with the temperature and surface gravity thresholds described in Stevenson (2016). As stated in that study, the likelihood for clouds is lower for exoplanets with equilibrium temperatures above 700 K and logarithmic surface gravities above 2.8. WASP-74b, with a temperature of 1947 K and a log surface gravity of 2.95 hovers above both thresholds, and WASP-88b satisfies the surface gravity limit ($\log\, g_{\rm p} = 2.68$) but contradicts the first criterion ($T_{\rm eq} = 1737\,{\rm K}$). Furthermore, both exoplanets do not follow the radius trend outlined in Tsiaras et al. (2018). Despite these conflicts, however, the exoplanets investigated in this thesis are not some odd outliers as high-altitude clouds and hazes are now increasingly being observed in atmospheres with a high level of irradiation (e.g. Espinoza et al., 2019; Alam et al., 2020; Chen et al., 2021a).

Both WASP-88b and WASP-74b have equilibrium temperatures that are in the temperature range between 1500 K and 2000 K. This puts them in a transition region where spectral features from alkali metals are being replaced by features from heavier metals, such as TiO and VO, when conditions are clear (Fortney et al., 2008). It would, therefore, be interesting to find out how WASP-88b and WASP-74b compare against exoplanets with similar characteristics in terms of temperature and surface gravity.

Some analogous exoplanets to WASP-88b with published transmission spectroscopy results in the optical are WASP-17b (Pinhas et al., 2019), WASP-31b (McGruder et al., 2020), HAT-P-32Ab (Alam et al., 2020) and HAT-P-41b (Wakeford et al., 2020; Sheppard et al., 2021). Most of them demonstrate partially or entirely muted features indicating the presence of clouds and/or haze. WASP-17b and HAT-P-41b seem to be the less cloudy exoplanets of this sample, with WASP-17b showing signs

of Na (Sing et al., 2016) and K (Sedaghati et al., 2016) absorption, and HAT-P-41b displaying features from strong molecular absorbers (Sheppard et al., 2021). Initial observations also detected the presence of K in the atmosphere of WASP-31b (Sing et al., 2015), but this result was firmly refuted afterwards through additional low- and high-resolution transmission spectroscopy (Gibson et al., 2017, 2019; McGruder et al., 2020). HAT-P-32Ab (Mallonn & Strassmeier, 2016; Alam et al., 2020) and WASP-31b (McGruder et al., 2020) are quite similar to WASP-88b as they also exhibit a scattering slope in the optical going upward with decreasing wavelength. This demonstrates that low-gravity exoplanets are not necessarily cloud-free. Even so, WASP-88b has one distinctive quality compared to other planets in this small sample: its scattering slope is at least one and half times stronger.

On the other hand, fewer exoplanets with similar properties to WASP-74b have been observed in transmission, and specifically at low-resolution and at optical to near-infrared wavelengths. The two closest matches to WASP-74b that I found in the literature were CoRoT-1b (Southworth, 2011) and WASP-79b (Brown et al., 2017). From these two exoplanets, CoRoT-1b also displays characteristics of a non-clear atmosphere, with all attempts at deciphering its transmission spectrum in the red-optical and the near-infrared pointing at a lack of absorption features and widespread clouds (Schlawin et al., 2014; Ranjan et al., 2014; Glidic et al., 2022). WASP-79b, however, is quite special as it shows an inverted slope towards shorter wavelengths. Such a slope could suggest an active host and, in particular, it could indicate the presence of faculae on the stellar disk that were unocculted by the transiting planet (Rathcke et al., 2021). Despite this peculiar slope, WASP-79b also presents strong evidence for water and weak to moderate evidence for H^- and FeH (Sotzen et al., 2020; Skaf et al., 2020; Rathcke et al., 2021). WASP-74b is, therefore, the only exoplanet here displaying a significant scattering slope suggestive of a hazy atmosphere.

6.3 Interpretation

The sharp scattering slopes observed for both WASP-88b and WASP-74b are hard to explain. The most common scattering slope seen in exoplanetary atmospheres is the one produced due to Rayleigh scattering (e.g. Lecavelier Des Etangs et al., 2008a; Kirk et al., 2017), but, as discussed in Chapter 1, the index in the Rayleigh power law is proportional to -4 and the findings from this work suggest likely lower values. Despite this discrepancy, these findings are not unusual and extreme scattering slopes in the blue are now increasingly being found in the transmission spectra of various exoplanets (e.g. May et al., 2020; Alam et al., 2020; Chen et al., 2021a; Ahrer et al., 2022).

One hypothesis that may explain the observed slope in both exoplanets is stellar activity. In this scenario, as the transiting exoplanets pass in front of their host star, phenomena that occur on the stellar atmosphere may disrupt the homogeneous luminosity from the stellar disk. The resultant heterogeneity could present itself in the form of starspots and faculae that form due to increased magnetic activity. If the starspots are unocculted or if the faculae are occulted, then the transmission spectra could exhibit steep "scattering" slopes that rise towards bluer wavelengths (McCullough et al., 2014; Oshagh et al., 2014). In the cases of WASP-88b and WASP-74b, however, this scenario is weak. Both exoplanets were monitored photometrically and the long-term light curves were found to be clean from significant periodic signals down to 1 mmag (Delrez et al., 2014; Hellier et al., 2015; Fu et al., 2021). In addition, an inspection of the calcium H and K lines for chromospheric emission in the stellar spectra came back negative in both cases (Southworth, unpublished). Atmospheric retrievals, on the other hand, provided mixed results revealing no stellar effects for WASP-88 and advocating the plausible presence of starspots for WASP-74. Given the rest of the evidence, however, it is highly implausible that either of the host stars is active.

The observed slopes are, therefore, more likely to be a product of haze processes within the planetary atmospheres. One such trigger could be the prevalence of photochemical reactions in the upper atmosphere (Kawashima & Ikoma, 2019; Ohno & Kawashima, 2020). In this case, two conditions need to be met: the eddy diffusion

has to be high and combined with a modest haze mass flux (these two parameters define particle mixing and mass flow within the atmosphere), and the equilibrium temperature has to be between 1000 and 1500 K. Both WASP-88b and WASP-74b fail to meet the second criterion, and additional observations are required to explore whether sufficient vertical mixing occurs within their atmospheres.

The formation of mineral condensates can also produce enhanced haze slopes in the optical. This scenario is more convincing for the steep gradient that characterises the transmission spectra of WASP-88b and WASP-74b. Given their equilibrium temperature (and even some of the retrieved temperatures), heavy silicate molecules such as forsterite and enstatite may form mineral clouds at high altitudes (Pinhas & Madhusudhan, 2017). In addition, contributions from alumina condensates are also expected at these temperatures. In both situations, however, the gradient is moderate and so may not be high enough. A potential solution could be the presence of manganese sulphide condensates in the atmosphere, but the temperatures in the day-night transition zone need to be somewhat lower than the estimated equilibrium temperatures (Pinhas & Madhusudhan, 2017). Manganese sulphide can generate extreme slopes with gradients that can go below -10. Nonetheless, some doubts were raised recently about the actual potential of sulphide species to form haze. Gao et al. (2020) argue that low nucleation energies may hamper the formation of sulphide clouds considerably, and the same conclusion was inferred for iron. Their results reinforce the idea that silicate aerosols play a dominant role in condensation processes for exoplanets with the properties of WASP-88b and WASP-74b.

Extremely steep scattering slopes can also be induced by circumplanetary dust rings (Ohno et al., 2022). These slopes could be significantly steeper than the ones found for WASP-88b and WASP-74b. According to Ohno et al. (2022), these dust rings need to be optically thin and above an optical thickness threshold of $\tau_{\rm ring} = 10^{-3}$, where $\tau_{\rm ring}$ is the line-of-sight attenuation of light due to the circumplanetary dust ring. If the dust rings are too thick, then this will flatten the observed transmission spectra (Ohno & Fortney, 2022). The lifespan of dust rings is predicted to be short for close-in exoplanets (e.g. Schlichting & Chang, 2011), so there needs to be a mechanism, such

as interplanetary dust capture (e.g. Mitchell et al., 2005) or leftover dust from impacts of interplanetary micrometeroids and natural satellites (e.g. Krivov et al., 2002), that replenishes the dust material around the planet. Such mechanisms are expected to be more common for planetary systems that are much younger than the WASP-88 and WASP-74 systems and so the chances for dust ring effects are quite low.

Apart from processes that may create slopes in the observed transmission spectra, it is important to stress here that all of the atmospheric modelling was conducted using 1D simulated atmospheres. The reason for this choice is simple: 3D general circulation models take way too much time to compute. While 1D realisations can provide useful conclusions about planetary atmospheres, they may lose some information due to 3D effects. For example, one prediction from 3D models is a day-night temperature gradient in hot, close-in exoplanets (see e.g. Showman & Guillot, 2002). The effect seems to be correlated with temperature as hotter exoplanets may experience larger day-night temperature contrasts (e.g. Komacek & Showman, 2016; Komacek et al., 2017). Since 1D models can only explore vertical gradients, they may fail to provide a proper temperature profile at the terminator region. Indeed, Caldas et al. (2019) found that terminator temperatures from 1D models may be biased towards dayside values. Furthermore, MacDonald et al. (2020) showed that the systematically lower terminator temperatures found from retrievals can be attributed to 1D modelling effects, which may be unable to distinguish between morning and evening terminator conditions. These hypotheses have recently been disputed, however, by Welbanks & Madhusudhan (2022) who argued that such biases likely stem from initial parameter assumptions and are not due to intrinsic limitations of the 1D simulated atmospheres (except maybe in extreme cases where the data quality is high and the terminator inhomogeneities are severe). Aside from temperature, other 3D effects that may be captured by 3D simulations and not by 1D frameworks are kinetic cloud formation (e.g. Lines et al., 2018) and chemical kinetics (e.g. Drummond et al., 2020). Some studies here have suggested that 1D simulations can result in inconsistencies between input assumptions and inferred chemical abundances (e.g. Caldas et al., 2019; Pluriel et al., 2020), but this has also been contested by Welbanks & Madhusudhan (2022).

Overall, it seems that, while 3D frameworks do offer improvements to the inferred atmospheric properties, the 1D simulated atmospheres employed throughout this work are already quite sufficient in generating reliable results.

Lastly, although there has been a meticulous effort to remove most systematic effects, it is still possible that small noise features may still remain and, hence, affect some of the conclusions drawn. In particular, the AlO abundance found for WASP-74b's atmosphere is unexpectedly high and is, therefore, more likely a result of minor unaccounted for noise effects. This conclusion is supported by the fact that the AlO result here is also one order of magnitude higher than the super-solar abundance estimate calculated by von Essen et al. (2019) for the ultra-hot Jupiter WASP-33b. Even so, however, the presence of AlO in the atmosphere of WASP-74b cannot be rejected completely given the tentative detection of aluminium by Lira-Barria et al. (2022). Ultimately, even if the AlO feature is real, other minor systematic effects may still be present in the measured transmission spectrum.

6.4 Future work and outlook

The field of exoplanet atmospheres has had many successes in recent years but is still in its early stages. For example, atmospheric retrievals were first applied to exoplanets just a little over a decade ago (e.g. Madhusudhan & Seager, 2009). It will, therefore, be quite interesting to see the development of the field, and transmission spectroscopy in particular, in the next few years. Advancements in technology will be essential, but in this section the focus will be on the two directions that naturally arise from the work presented in this thesis. These are the necessity for additional observations (especially in the infrared) and the potential application of the method developed in Chapter 5 (Section 5.3.2.5) on other targets.

Firstly, all the data presented here cover a limited range of the transmission spectrum. Specifically, all analysed ground-based data sets probe the optical, with the data for the test target WASP-75b slightly expanding towards the near-ultraviolet

and the data for WASP-88b narrowly exploring the beginning of the near-infrared. The range is especially tight for WASP-74b where the VLT data only cover the blue-optical. The confined range makes sense given the blocking effect of the Earth's atmosphere in most of the near-infrared and near-ultraviolet. Nevertheless, it would be quite beneficial to extend this range in both directions, especially into the infrared where many chemical species are expected to produce prominent features (e.g. Wakeford et al., 2013; JWST Transiting Exoplanet Community Early Release Science Team et al., 2023; Rustamkulov et al., 2023; Alderson et al., 2023). And while WASP-74b has already been examined in the near-infrared from space (Tsiaras et al., 2018; Fu et al., 2021), a better understanding of the atmospheres of the two main thesis targets (WASP-88b and WASP-74b) requires additional observations in that wavelength regime. With the presence and development of new ground- and space-based telescopes, the prospects are tremendous in the coming decades for a complete atmospheric characterisation of these two exoplanets.

The expectations are also high for the atmospheric characterisation of many other old and newly discovered exoplanets. In that direction, it will be crucial to develop robust and reliable analysis methods that are able to reduce the impact of noise without being affected by increased uncertainty. The novel method detailed in Section 5.3.2.5 is a step forward as it opens up a new avenue in noise decorrelation techniques and can aid in the understanding of exoplanetary atmospheres by generating precise results. This new method does have shortcomings (see Sections 5.5.1 and 6.1.3), but can be improved upon and become a valuable option in ground-based transmission spectroscopy.

The recent launch of the James Webb Space Telescope (JWST, Gardner et al., 2006) and the initiation of cycle 1 of its science operations was a major milestone in the field. It is anticipated that this telescope will observe exoplanetary atmospheres in unprecedented detail in the red-optical, near-infrared and mid-infrared wavelength regions. Its instrumentation includes the Near Infrared Spectrograph (NIRSpec) and the Near Infrared Imager and Slitless Spectrograph (NIRISS) in the red-optical and near-infrared, and the Mid-Infrared Instrument (MIRI) in the mid-infrared. All of these instruments are well-suited for the atmospheric characterisation of exoplanets.

Interestingly, the mission has already reached or even exceeded expectations, with multiple discoveries having been announced so far, including the detection of several chemical species in the atmosphere of WASP-39b (JWST Transiting Exoplanet Community Early Release Science Team et al., 2023; Rustamkulov et al., 2023; Alderson et al., 2023; Ahrer et al., 2023; Tsai et al., 2023; Feinstein et al., 2023). Such studies are of great importance as they allow for investigations of atmospheric compositions and other atmospheric properties, including metallicity trends (e.g. JWST Transiting Exoplanet Community Early Release Science Team et al., 2023) and photochemical phenomena (e.g. Tsai et al., 2023).

Another space telescope currently under development is the Atmospheric Remote-sensing Infrared Exoplanet Large-survey (ARIEL, Tinetti et al., 2018). ARIEL will have an oval-shaped primary mirror of $\sim$1 m in diameter and will arguably be the most specialised space-borne telescope in the field with its exclusive focus on the observation of known exoplanetary atmospheres. The plan is to perform one of the largest statistical surveys of warm and hot exoplanetary atmospheres ($\sim$1000 atmospheres), with the intention to establish potential relations between observed chemistry, stellar environment, and planetary formation and evolution histories. ARIEL is expected to get first light in the late 2020s or early 2030s and will explore a region in the infrared where many major molecular species could show prominent absorption features.

Significant progress is also made from the ground, with three extremely large, state-of-the-art telescopes currently being under construction. The Extremely Large Telescope (ELT), the Thirty Meter Telescope (TMT) and the Giant Magellan Telescope (GMT) will have primary mirrors with diameters that will be more than twice the size of present-day telescope diameters (39.3, 30 and 24.5 m, respectively), and will have instruments suitable for transmission spectroscopy in the optical and near-infrared. The enormous light-gathering power of these telescopes will enable the detection of the atmospheres of even terrestrial exoplanets and will provide insights into the little-known characteristics of these rocky worlds (see e.g. Lopez-Morales et al., 2019).

All these new telescopes will largely complement and improve upon the legacy left by predecessors, such as HST, NTT, and VLT, and will undoubtedly lead to a refined

perception of the exoplanets investigated in this thesis. At the same time, the results from this work will perhaps motivate the implementation of new analysis techniques in transmission spectroscopy and will have an impact on future atmospheric observations of these hot Jupiters. This research, by deciphering the atmospheric properties of hot gas giant exoplanets, provides a valuable contribution to the rapidly evolving field of exoplanetary atmospheres and illustrates the important role of transmission spectroscopy in increasing our knowledge of the dominant atmospheric characteristics in other planetary systems.

Publications

The list of publications is divided into two groups to clearly mark the ones led by me and the others where I am included as a co-author. In the latter ones, I contributed by performing some of the observations. All publications are in peer-reviewed journals.

Led

- Spyratos P., Nikolov N. K., Southworth J., et al., 2021, MNRAS, 506, 2853
- Spyratos P., Nikolov N. K., Constantinou S., et al., 2023, MNRAS, 521, 2163

Co-authored

- Southworth J., Dominik M., Jørgensen U. G., et al., 2019, MNRAS, 490, 4230
- Herald A., Udalski A., Bozza V., et al., 2022, A&A, 663, A100
- Southworth J., Barker A. J., Hinse T. C., et al., 2022, MNRAS, 515, 3212

Bibliography

Abramowitz M., Stegun I. A., 1965, Handbook of mathematical functions with formulas, graphs, and mathematical tables

Ahrer E., Wheatley P. J., Kirk J., Gandhi S., King G. W., Louden T., 2022, MNRAS, 510, 4857

Ahrer E.-M., et al., 2023, Nature, 614, 653

Aigrain S., Foreman-Mackey D., 2022, arXiv e-prints, p. arXiv:2209.08940

Akaike H., 1974, IEEE Transactions on Automatic Control, 19, 716

Al-Refaie A. F., Changeat Q., Waldmann I. P., Tinetti G., 2021, ApJ, 917, 37

Alam M. K., et al., 2020, AJ, 160, 51

Alam M. K., et al., 2021, ApJL, 906, L10

Alderson L., et al., 2020, MNRAS, 497, 5182

Alderson L., et al., 2023, Nature, 614, 664

Allan A., Vidotto A. A., 2019, MNRAS, 490, 3760

Ambikasaran S., Foreman-Mackey D., Greengard L., Hogg D. W., O'Neil M., 2015, IEEE Transactions on Pattern Analysis and Machine Intelligence, 38

Amundsen D. S., Baraffe I., Tremblin P., Manners J., Hayek W., Mayne N. J., Acreman D. M., 2014, A&A, 564, A59

Anderson D. R., et al., 2010, ApJ, 709, 159

Andreasen D. T., et al., 2017, A&A, 600, A69

Anglada-Escudé G., et al., 2016, Nature, 536, 437

Appenzeller I., et al., 1998, The Messenger, 94, 1

Arras P., Bildsten L., 2006, ApJ, 650, 394

Arras P., Socrates A., 2010, ApJ, 714, 1

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

Astudillo-Defru N., Rojo P., 2013, A&A, 557, A56

Babusiaux C., et al., 2022, arXiv e-prints, p. arXiv:2206.05989

Başẗurk Ö., et al., 2022, MNRAS, 512, 2062

Bailer-Jones C. A. L., Rybizki J., Fouesneau M., Demleitner M., Andrae R., 2021, AJ, 161, 147

Bakos G. Á., et al., 2007, ApJ, 656, 552

Baraffe I., Chabrier G., Allard F., Hauschildt P. H., 1998, A&A, 337, 403

Barber R. J., Strange J. K., Hill C., Polyansky O. L., Mellau G. C., Yurchenko S. N., Tennyson J., 2014, MNRAS, 437, 1828

Barmby P., 2019, The Open Journal of Astrophysics, 2, 2

Barstow J. K., Aigrain S., Irwin P. G. J., Sing D. K., 2017, ApJ, 834, 50

Baruteau C., et al., 2014, in Beuther H., Klessen R. S., Dullemond C. P., Henning T., eds, Protostars and Planets VI. p. 667 (arXiv:1312.4293), doi:10.2458/azu˙uapress˙9780816531240-ch029

Batygin K., Stevenson D. J., 2010, ApJL, 714, L238

Batygin K., Bodenheimer P. H., Laughlin G. P., 2016, ApJ, 829, 114

Bauschlicher C. W., Ram R. S., Bernath P. F., Parsons C. G., Galehouse D., 2001, JChPh, 115, 1312

Bean J. L., Miller-Ricci Kempton E., Homeier D., 2010, Nature, 468, 669

Bean J. L., et al., 2011, ApJ, 743, 92

Bean J. L., Désert J.-M., Seifahrt A., Madhusudhan N., Chilingarian I., Homeier D., Szentgyorgyi A., 2013, ApJ, 771, 108

Beleznay M., Kunimoto M., 2022, MNRAS, 516, 75

Bodenheimer P., Lin D. N. C., Mardling R. A., 2001, ApJ, 548, 466

Boffin H. M. J., et al., 2016, in Evans C. J., Simard L., Takami H., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 9908, Ground-based and Airborne Instrumentation for Astronomy VI. p. 99082B (arXiv:1607.07237), doi:10.1117/12.2232094

Bohn A. J., Southworth J., Ginski C., Kenworthy M. A., Maxted P. F. L., Evans D. F., 2020, A&A, 635, A73

Boley A. C., Granados Contreras A. P., Gladman B., 2016, ApJL, 817, L17

Bond I. A., et al., 2004, ApJL, 606, L155

Borucki W. J., et al., 2011, ApJ, 736, 19

Boss A. P., 1997, Science, 276, 1836

Boss A. P., et al., 2007, Transactions of the International Astronomical Union, Series A, 26A, 183

Bourrier V., et al., 2018, A&A, 620, A147

Breger M., et al., 1993, A&A, 271, 482

Brogi M., de Kok R. J., Albrecht S., Snellen I. A. G., Birkby J. L., Schwarz H., 2016, ApJ, 817, 106

Brown T. M., 2001, ApJ, 553, 1006

Brown D. J. A., et al., 2017, MNRAS, 464, 810

Buchhave L. A., et al., 2014, Nature, 509, 593

Buchner J., et al., 2014, A&A, 564, A125

Burrows A., et al., 1997, ApJ, 491, 856

Burrows A., Dulick M., Bauschlicher C. W. J., Bernath P. F., Ram R. S., Sharp C. M., Milsom J. A., 2005, ApJ, 624, 988

Burrows A., Hubeny I., Budaj J., Hubbard W. B., 2007, ApJ, 661, 502

Butler R. P., Marcy G. W., Williams E., Hauser H., Shirts P., 1997, ApJL, 474, L115

Butler R. P., Marcy G. W., Vogt S. S., Apps K., 1998, PASP, 110, 1389

Butler R. P., Marcy G. W., Fischer D. A., Brown T. M., Contos A. R., Korzennik S. G., Nisenson P., Noyes R. W., 1999, ApJ, 526, 916

Caldas A., Leconte J., Selsis F., Waldmann I. P., Bordé P., Rocchetto M., Charnay B., 2019, A&A, 623, A161

Campbell B., Walker G. A. H., Yang S., 1988, ApJ, 331, 902

Carone L., et al., 2021, A&A, 646, A168

Carrera D., Raymond S. N., Davies M. B., 2019, A&A, 629, L7

Carter A. L., et al., 2020, MNRAS, 494, 5449

Chabrier G., Baraffe I., 2007, ApJL, 661, L81

Chabrier G., Johansen A., Janson M., Rafikov R., 2014, in Beuther H., Klessen R. S., Dullemond C. P., Henning T., eds, Protostars and Planets VI. p. 619 (arXiv:1401.7559), doi:10.2458/azu˙uapress˙9780816531240-ch027

Changeat Q., et al., 2022, ApJS, 260, 3

Charbonneau D., Brown T. M., Latham D. W., Mayor M., 2000, ApJL, 529, L45

Charbonneau D., Brown T. M., Noyes R. W., Gilliland R. L., 2002, ApJ, 568, 377

Charbonneau D., et al., 2005, ApJ, 626, 523

Chatterjee S., Ford E. B., Matsumura S., Rasio F. A., 2008, ApJ, 686, 580

Chauvin G., Lagrange A. M., Dumas C., Zuckerman B., Mouillet D., Song I., Beuzit J. L., Lowrance P., 2004, A&A, 425, L29

Chen G., Casasayas-Barris N., Pallé E., Welbanks L., Madhusudhan N., Luque R., Murgas F., 2020, A&A, 642, A54

Chen G., et al., 2021a, MNRAS, 500, 5420

Chen G., Pallé E., Parviainen H., Murgas F., Yan F., 2021b, ApJL, 913, L16

Chen G., Wang H., van Boekel R., Pallé E., 2022, AJ, 164, 173

Claret A., 2000, A&A, 363, 1081

Claret A., 2017, A&A, 600, A30

Claret A., Hauschildt P. H., 2003, A&A, 412, 241

Clark B. J. M., Anderson D. R., Hellier C., Turner O. D., Močnik T., 2018, PASP, 130, 034401

Cumming A., Butler R. P., Marcy G. W., Vogt S. S., Wright J. T., Fischer D. A., 2008, PASP, 120, 531

Dawson R. I., Johnson J. A., 2018, ARA&A, 56, 175

Delrez L., et al., 2014, A&A, 563, A143

Deming D., Seager S., Richardson L. J., Harrington J., 2005, Nature, 434, 740

Deming D., et al., 2013, ApJ, 774, 95

Dravins D., Lindegren L., Mezey E., Young A. T., 1998, PASP, 110, 610

Drummond B., Mayne N. J., Baraffe I., Tremblin P., Manners J., Amundsen D. S., Goyal J., Acreman D., 2018, A&A, 612, A105

Drummond B., et al., 2020, A&A, 636, A68

Dulick M., Bauschlicher C. W. J., Burrows A., Sharp C. M., Ram R. S., Bernath P., 2003, ApJ, 594, 651

Durisen R. H., Boss A. P., Mayer L., Nelson A. F., Quinn T., Rice W. K. M., 2007, in Reipurth B., Jewitt D., Keil K., eds, Protostars and Planets V. p. 607 (arXiv:astro-ph/0603179)

Eastman J., Siverd R., Gaudi B. S., 2010, PASP, 122, 935

Edwards B., et al., 2022, arXiv e-prints, p. arXiv:2211.00649

Eggleton P. P., Kiseleva L. G., Hut P., 1998, ApJ, 499, 853

Ehrenreich D., Désert J. M., 2011, A&A, 529, A136

Ehrenreich D., et al., 2015, Nature, 522, 459

Espinoza N., et al., 2019, MNRAS, 482, 2065

Evans T. M., et al., 2016, ApJL, 822, L4

Evans T. M., et al., 2017, Nature, 548, 58

Fabrycky D., Tremaine S., 2007, ApJ, 669, 1298

Faedi F., et al., 2011, A&A, 531, A40

Feinstein A. D., et al., 2023, Nature, 614, 670

Feroz F., Hobson M. P., Bridges M., 2009, MNRAS, 398, 1601

Fischer D. A., Valenti J., 2005, ApJ, 622, 1102

Fischer P. D., et al., 2016, ApJ, 827, 19

Fisher C., Heng K., 2018, MNRAS, 481, 4698

Föhring D., Wilson R. W., Osborn J., Dhillon V. S., 2019, MNRAS, 489, 5098

Foreman-Mackey D., 2015, George: Gaussian Process regression (ascl:1511.015)

Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306

Fortney J. J., Lodders K., Marley M. S., Freedman R. S., 2008, ApJ, 678, 1419

Fortney J. J., Shabram M., Showman A. P., Lian Y., Freedman R. S., Marley M. S., Lewis N. K., 2010, ApJ, 709, 1396

Fortney J. J., Dawson R. I., Komacek T. D., 2021, Journal of Geophysical Research (Planets), 126, e06629

Fressin F., et al., 2013, ApJ, 766, 81

Fu G., Deming D., Knutson H., Madhusudhan N., Mandell A., Fraine J., 2017, ApJL, 847, L22

Fu G., et al., 2021, AJ, 162, 271

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

Gaia Collaboration et al., 2022a, arXiv e-prints, p. arXiv:2206.05595

Gaia Collaboration et al., 2022b, arXiv e-prints, p. arXiv:2208.00211

Gan T., et al., 2023, AJ, 165, 17

Gandhi S., Madhusudhan N., 2019, MNRAS, 485, 5817

Gao P., et al., 2020, Nature Astronomy, 4, 951

García R. A., Mathur S., Salabert D., Ballot J., Régulo C., Metcalfe T. S., Baglin A., 2010, Science, 329, 1032

Gardner J. P., et al., 2006, SSRv, 123, 485

Garhart E., et al., 2020, AJ, 159, 137

Gibson N. P., 2014, MNRAS, 445, 3401

Gibson N. P., Pont F., Aigrain S., 2011, MNRAS, 411, 2199

Gibson N. P., Aigrain S., Roberts S., Evans T. M., Osborne M., Pont F., 2012, MNRAS, 419, 2683

Gibson N. P., Aigrain S., Barstow J. K., Evans T. M., Fletcher L. N., Irwin P. G. J., 2013a, MNRAS, 428, 3680

Gibson N. P., Aigrain S., Barstow J. K., Evans T. M., Fletcher L. N., Irwin P. G. J., 2013b, MNRAS, 436, 2974

Gibson N. P., Nikolov N., Sing D. K., Barstow J. K., Evans T. M., Kataria T., Wilson P. A., 2017, MNRAS, 467, 4591

Gibson N. P., de Mooij E. J. W., Evans T. M., Merritt S., Nikolov N., Sing D. K., Watson C., 2019, MNRAS, 482, 606

Gilliland R. L., Arribas S., 2003, High Signal-to-Noise Differential NICMOS Spectrophotometry, Instrument Science Report NICMOS 2003-001, 14 pages

Gillon M., et al., 2017, Nature, 542, 456

Glidic K., Schlawin E., Wiser L., Zhou Y., Deming D., Line M., 2022, AJ, 164, 19

Goldreich P., Tremaine S., 1980, ApJ, 241, 425

Gómez Maqueo Chew Y., et al., 2013, A&A, 559, A36

Gonzalez G., 1997, MNRAS, 285, 403

Goodman J., Weare J., 2010, Communications in Applied Mathematics and Computational Science, 5, 65

Gordon I. E., et al., 2017, JQSRT, 203, 3

Goyal J. M., et al., 2018, MNRAS, 474, 5158

Goyal J. M., Wakeford H. R., Mayne N. J., Lewis N. K., Drummond B., Sing D. K., 2019a, MNRAS, 482, 4503

Goyal J. M., et al., 2019b, MNRAS, 486, 783

Goyal J. M., et al., 2020, MNRAS, 498, 4680

Guillot T., Showman A. P., 2002, A&A, 385, 156

Guillot T., Burrows A., Hubbard W. B., Lunine J. I., Saumon D., 1996, ApJL, 459, L35

Harrington J., Hansen B. M., Luszcz S. H., Seager S., Deming D., Menou K., Cho J. Y. K., Richardson L. J., 2006, Science, 314, 623

Harris C. R., et al., 2020, Nature, 585, 357

Hastings W. K., 1970, Biometrika, 57, 97

Hatzes A. P., Rauer H., 2015, ApJL, 810, L25

Hatzes A. P., Cochran W. D., Endl M., McArthur B., Paulson D. B., Walker G. A. H., Campbell B., Yang S., 2003, ApJ, 599, 1383

Heller R., 2019, A&A, 628, A42

Hellier C., et al., 2015, AJ, 150, 18

Heng K., 2016, ApJL, 826, L16

Heng K., Lyons J. R., Tsai S.-M., 2016, ApJ, 816, 96

Henry G. W., Marcy G. W., Butler R. P., Vogt S. S., 2000, ApJL, 529, L41

Hoeijmakers H. J., et al., 2018, Nature, 560, 453

Holl B., et al., 2022, arXiv e-prints, p. arXiv:2206.05439

Howard A. W., et al., 2012, ApJS, 201, 15

Howell S. B., et al., 2014, PASP, 126, 398

Hubbard W. B., Fortney J. J., Lunine J. I., Burrows A., Sudarsky D., Pinto P., 2001, ApJ, 560, 413

Hubeny I., Burrows A., Sudarsky D., 2003, ApJ, 594, 1011

Huitson C. M., Désert J. M., Bean J. L., Fortney J. J., Stevenson K. B., Bergmann M., 2017, AJ, 154, 95

Hunter J. D., 2007, Computing in Science and Engineering, 9, 90

Ida S., Lin D. N. C., 2008, ApJ, 673, 487

Ivshina E. S., Winn J. N., 2022, ApJS, 259, 62

Iyer A. R., Swain M. R., Zellem R. T., Line M. R., Roudier G., Rocha G., Livingston J. H., 2016, ApJ, 823, 109

JWST Transiting Exoplanet Community Early Release Science Team et al., 2023, Nature, 614, 649

Jackson B., Barnes R., Greenberg R., 2009, ApJ, 698, 1357

Jeffreys H., 1961, Theory of Probability, third edn. Oxford University Press

Jehin E., et al., 2011, The Messenger, 145, 2

Jenkins J. M., et al., 2016, in Proc. SPIE. p. 99133E

Jensen A. G., Cauley P. W., Redfield S., Cochran W. D., Endl M., 2018, AJ, 156, 154

Jin S., Mordasini C., Parmentier V., van Boekel R., Henning T., Ji J., 2014, ApJ, 795, 65

Johansen A., Lambrechts M., 2017, Annual Review of Earth and Planetary Sciences, 45, 359

Kane S. R., Ceja A. Y., Way M. J., Quintana E. V., 2018, ApJ, 869, 46

Kanodia S., Wright J., 2018, Research Notes of the American Astronomical Society, 2, 4

Kass R. E., Raftery A. E., 1995, Journal of the American Statistical Association, 90, 773

Kataria T., Sing D. K., Lewis N. K., Visscher C., Showman A. P., Fortney J. J., Marley M. S., 2016, ApJ, 821, 9

Kawashima Y., Ikoma M., 2019, ApJ, 877, 109

Kempton E. M.-R., Lupu R., Owusu-Asare A., Slough P., Cale B., 2017, PASP, 129, 044402

Kirk J., Wheatley P. J., Louden T., Littlefair S. P., Copperwheat C. M., Armstrong D. J., Marsh T. R., Dhillon V. S., 2016, MNRAS, 463, 2922

Kirk J., Wheatley P. J., Louden T., Doyle A. P., Skillen I., McCormac J., Irwin P. G. J., Karjalainen R., 2017, MNRAS, 468, 3907

Kirk J., López-Morales M., Wheatley P. J., Weaver I. C., Skillen I., Louden T., McCormac J., Espinoza N., 2019, AJ, 158, 144

Knutson H. A., Charbonneau D., Noyes R. W., Brown T. M., Gilliland R. L., 2007, ApJ, 655, 564

Knutson H. A., Benneke B., Deming D., Homeier D., 2014, Nature, 505, 66

Kochukhov O., 2021, A&A Rv, 29, 1

Komacek T. D., Showman A. P., 2016, ApJ, 821, 16

Komacek T. D., Showman A. P., Tan X., 2017, ApJ, 835, 198

Konacki M., Torres G., Jha S., Sasselov D. D., 2003, Nature, 421, 507

Kopal Z., 1950, Harvard College Observatory Circular, 454, 1

Kreidberg L., 2015, PASP, 127, 1161

Kreidberg L., 2018, Exoplanet Atmosphere Measurements from Transmission Spectroscopy and Other Planet Star Combined Light Observations. p. 100, doi:10.1007/978-3-319-55333-7_100

Kreidberg L., et al., 2014a, Nature, 505, 69

Kreidberg L., et al., 2014b, ApJL, 793, L27

Kreidberg L., et al., 2015, ApJ, 814, 66

Krivov A. V., Krüger H., Grün E., Thiessenhusen K.-U., Hamilton D. P., 2002, Journal of Geophysical Research (Planets), 107, 5002

Kulow J. R., France K., Linsky J., Loyd R. O. P., 2014, ApJ, 786, 132

Kurokawa H., Nakamoto T., 2014, ApJ, 783, 54

Lam K. W. F., et al., 2017, A&A, 599, A3

Latham D. W., Mazeh T., Stefanik R. P., Mayor M., Burki G., 1989, Nature, 339, 38

Laughlin G., Crismani M., Adams F. C., 2011, ApJL, 729, L7

Lavie B., et al., 2017, A&A, 605, L7

Lecavelier Des Etangs A., Pont F., Vidal-Madjar A., Sing D., 2008a, A&A, 481, L83

Lecavelier Des Etangs A., Vidal-Madjar A., Désert J. M., Sing D., 2008b, A&A, 485, 865

Lecavelier Des Etangs A., et al., 2010, A&A, 514, A72

Lecavelier des Etangs A., Lissauer J. J., 2022, NewAR, 94, 101641

Lendl M., et al., 2012, A&A, 544, A72

Lendl M., et al., 2016, A&A, 587, A67

Lendl M., Cubillos P. E., Hagelberg J., Müller A., Juvan I., Fossati L., 2017, A&A, 606, A18

Levenberg K., 1944, Quarterly of Applied Mathematics, 2, 164

Liddle A. R., 2007, MNRAS, 377, L74

Lin D. N. C., Papaloizou J., 1986, ApJ, 309, 846

Line M. R., Knutson H., Deming D., Wilkins A., Desert J.-M., 2013, ApJ, 778, 183

Line M. R., et al., 2016, AJ, 152, 203

Lines S., et al., 2018, MNRAS, 481, 194

Lira-Barria A., Rojo P. M., Mendez R. A., 2022, A&A, 657, A36

Lopez-Morales M., et al., 2019, BAAS, 51, 162

Lovis C., et al., 2011, A&A, 528, A112

Luque R., et al., 2020, A&A, 642, A50

MacDonald R. J., Madhusudhan N., 2017, MNRAS, 469, 1979

MacDonald R. J., Madhusudhan N., 2019, MNRAS, 486, 1292

MacDonald R. J., Goyal J. M., Lewis N. K., 2020, ApJL, 893, L43

Mackebrandt F., et al., 2017, A&A, 608, A26

Madhusudhan N., 2012, ApJ, 758, 36

Madhusudhan N., 2018, Atmospheric Retrieval of Exoplanets. p. 104, doi:10.1007/978-3-319-55333-7_104

Madhusudhan N., 2019, ARA&A, 57, 617

Madhusudhan N., Seager S., 2009, ApJ, 707, 24

Madhusudhan N., Crouzet N., McCullough P. R., Deming D., Hedges C., 2014a, ApJL, 791, L9

Madhusudhan N., Amin M. A., Kennedy G. M., 2014b, ApJL, 794, L12

Magic Z., Chiavassa A., Collet R., Asplund M., 2015, A&A, 573, A90

Mallonn M., Strassmeier K. G., 2016, A&A, 590, A100

Mallonn M., et al., 2015, A&A, 580, A60

Mallonn M., et al., 2018, A&A, 614, A35

Mancini L., Giordano M., Mollière P., Southworth J., Brahm R., Ciceri S., Henning T., 2016, MNRAS, 461, 1053

Mancini L., et al., 2019, MNRAS, 485, 5168

Mandel K., Agol E., 2002, ApJL, 580, L171

Mansfield M., et al., 2021, Nature Astronomy, 5, 1224

Marley M. S., Robinson T. D., 2015, ARA&A, 53, 279

Marley M. S., Ackerman A. S., Cuzzi J. N., Kitzmann D., 2013, in Mackwell S. J., Simon-Miller A. A., Harder J. W., Bullock M. A., eds, , Comparative Climatology of Terrestrial Planets. pp 367–392, doi:10.2458/azu˙uapress˙9780816530595-ch015

Marquardt D. W., 1963, Journal of the Society for Industrial and Applied Mathematics, 11, 431

Martins J. H. C., et al., 2015, A&A, 576, A134

Maxted P. F. L., et al., 2020, MNRAS, 498, 332

May E. M., Gardner T., Rauscher E., Monnier J. D., 2020, AJ, 159, 7

Mayor M., Queloz D., 1995, Nature, 378, 355

Mayor M., et al., 2011, arXiv e-prints, p. arXiv:1109.2497

McCullough P. R., Crouzet N., Deming D., Madhusudhan N., 2014, ApJ, 791, 55

McCully C., Tewes M., 2019, Astro-SCRAPPY: Speedy Cosmic Ray Annihilation Package in Python, Astrophysics Source Code Library, record ascl:1907.032 (ascl:1907.032)

McGruder C. D., et al., 2020, AJ, 160, 230

McKemmish L. K., Yurchenko S. N., Tennyson J., 2016, MNRAS, 463, 771

McKemmish L. K., Masseron T., Hoeijmakers H. J., Pérez-Mesa V., Grimm S. L., Yurchenko S. N., Tennyson J., 2019, MNRAS, 488, 2836

Metropolis N., Rosenbluth A. W., Rosenbluth M. N., Teller A. H., Teller E., 1953, JChPh, 21, 1087

Mikal-Evans T., et al., 2021, AJ, 161, 18

Mitchell C. J., Colwell J. E., HoráNyi M., 2005, Journal of Geophysical Research (Space Physics), 110, A09218

Mollière P., Wardenier J. P., van Boekel R., Henning T., Molaverdikhani K., Snellen I. A. G., 2019, A&A, 627, A67

Moses J. I., 2014, Philosophical Transactions of the Royal Society of London Series A, 372, 20130073

Müller H. M., Huber K. F., Czesla S., Wolter U., Schmitt J. H. M. M., 2013, A&A, 560, A112

Murgas F., Pallé E., Zapatero Osorio M. R., Nortmann L., Hoyer S., Cabrera-Lavers A., 2014, A&A, 563, A41

Newville M., Stensitzki T., Allen D. B., Rawlik M., Ingargiola A., Nelson A., 2016, Lmfit: Non-Linear Least-Square Minimization and Curve-Fitting for Python, Astrophysics Source Code Library (ascl:1606.014)

Nikolov N., et al., 2014, MNRAS, 437, 46

Nikolov N., Sing D. K., Gibson N. P., Fortney J. J., Evans T. M., Barstow J. K., Kataria T., Wilson P. A., 2016, ApJ, 832, 191

Nikolov N., et al., 2018a, MNRAS, 474, 1705

Nikolov N., et al., 2018b, Nature, 557, 526

Nikolov N., et al., 2021, AJ, 162, 88

Ninan J. P., et al., 2020, ApJ, 894, 97

Nixon M. C., Madhusudhan N., 2022, ApJ, 935, 73

Noyes R. W., Hartmann L. W., Baliunas S. L., Duncan D. K., Vaughan A. H., 1984, ApJ, 279, 763

Öberg K. I., Murray-Clay R., Bergin E. A., 2011, ApJL, 743, L16

Ohno K., Fortney J. J., 2022, ApJ, 930, 50

Ohno K., Kawashima Y., 2020, ApJL, 895, L47

Ohno K., Thao P. C., Mann A. W., Fortney J. J., 2022, ApJL, 940, L30

Oshagh M., Santos N. C., Ehrenreich D., Haghighipour N., Figueira P., Santerne A., Montalto M., 2014, A&A, 568, A99

Panwar V., Désert J.-M., Todorov K. O., Bean J. L., Stevenson K. B., Huitson C. M., Fortney J. J., Bergmann M., 2022, MNRAS, 510, 3236

Parviainen H., Aigrain S., Thatte N., Barstow J. K., Evans T. M., Gibson N., 2015, MNRAS, 453, 3875

Patrascu A. T., Yurchenko S. N., Tennyson J., 2015, MNRAS, 449, 3613

Perri F., Cameron A. G. W., 1974, Icarus, 22, 416

Pinhas A., Madhusudhan N., 2017, MNRAS, 471, 4355

Pinhas A., Rackham B. V., Madhusudhan N., Apai D., 2018, MNRAS, 480, 5314

Pinhas A., Madhusudhan N., Gandhi S., MacDonald R., 2019, MNRAS, 482, 1485

Pluriel W., Zingales T., Leconte J., Parmentier V., 2020, A&A, 636, A66

Pollacco D. L., et al., 2006, PASP, 118, 1407

Pollack J. B., Hubickyj O., Bodenheimer P., Lissauer J. J., Podolak M., Greenzweig Y., 1996, Icarus, 124, 62

Pont F., Knutson H., Gilliland R. L., Moutou C., Charbonneau D., 2008, MNRAS, 385, 109

Quintana E. V., et al., 2014, Science, 344, 277

Rafikov R. R., 2005, ApJL, 621, L69

Ranjan S., Charbonneau D., Désert J.-M., Madhusudhan N., Deming D., Wilkins A., Mandell A. M., 2014, ApJ, 785, 148

Rasio F. A., Ford E. B., 1996, Science, 274, 954

Rasio F. A., Tout C. A., Lubow S. H., Livio M., 1996, ApJ, 470, 1187

Rasmussen C. E., Williams C. K. I., 2006, Gaussian Processes for Machine Learning

Rathcke A. D., et al., 2021, AJ, 162, 138

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

Roberts S., Osborne M., Ebden M., Reece S., Gibson N., Aigrain S., 2012, Philosophical Transactions of the Royal Society of London Series A, 371, 20110550

Rothman L. S., et al., 2010, JQSRT, 111, 2139

Roudier G. M., Swain M. R., Gudipati M. S., West R. A., Estrela R., Zellem R. T., 2021, AJ, 162, 37

Rustamkulov Z., et al., 2023, Nature, 614, 659

Santos N. C., Mayor M., Naef D., Pepe F., Queloz D., Udry S., Burnet M., 2001, A&A, 379, 999

Sarkis P., Mordasini C., Henning T., Marleau G. D., Mollière P., 2021, A&A, 645, A79

Schlawin E., Zhao M., Teske J. K., Herter T., 2014, ApJ, 783, 5

Schlichting H. E., 2014, ApJL, 795, L15

Schlichting H. E., Chang P., 2011, ApJ, 734, 117

Schwarz G., 1978, Annals of Statistics, 6, 461

Seager S., Deming D., 2010, ARA&A, 48, 631

Seager S., Sasselov D. D., 2000, ApJ, 537, 916

Sedaghati E., Boffin H. M. J., Csizmadia S., Gibson N., Kabath P., Mallonn M., Van den Ancker M. E., 2015, A&A, 576, L11

Sedaghati E., et al., 2016, A&A, 596, A47

Sedaghati E., Boffin H. M. J., Delrez L., Gillon M., Csizmadia S., Smith A. M. S., Rauer H., 2017a, MNRAS, 468, 3123

Sedaghati E., et al., 2017b, Nature, 549, 238

Seidel J. V., et al., 2020, A&A, 643, A45

Sheppard K. B., et al., 2021, AJ, 161, 51

Showman A. P., Guillot T., 2002, A&A, 385, 166

Sing D. K., Vidal-Madjar A., Désert J. M., Lecavelier des Etangs A., Ballester G., 2008, ApJ, 686, 658

Sing D. K., et al., 2011, A&A, 527, A73

Sing D. K., et al., 2015, MNRAS, 446, 2428

Sing D. K., et al., 2016, Nature, 529, 59

Skaf N., et al., 2020, AJ, 160, 109

Skilling J., 2006, Bayesian Analysis, 1, 833

Soter S., 2006, AJ, 132, 2513

Sotzen K. S., et al., 2020, AJ, 159, 5

Sousa S. G., et al., 2021, A&A, 656, A53

Southworth J., 2008, MNRAS, 386, 1644

Southworth J., 2010, MNRAS, 408, 1689

Southworth J., 2011, MNRAS, 417, 2166

Southworth J., 2012, MNRAS, 426, 1291

Southworth J., 2013, A&A, 557, A119

Southworth J., unpublished

Southworth J., Evans D. F., 2016, MNRAS, 463, 37

Southworth J., et al., 2009, MNRAS, 396, 1023

Southworth J., Bohn A. J., Kenworthy M. A., Ginski C., Mancini L., 2020, A&A, 635, A74

Spake J. J., et al., 2018, Nature, 557, 68

Speagle J. S., 2020, MNRAS, 493, 3132

Spiegel D. S., Burrows A., Milsom J. A., 2011, ApJ, 727, 57

Spiegel D. S., Fortney J. J., Sotin C., 2014, Proceedings of the National Academy of Science, 111, 12622

Spyratos P., et al., 2021, MNRAS, 506, 2853

Spyratos P., Nikolov N. K., Constantinou S., Southworth J., Madhusudhan N., Sedaghati E., Ehrenreich D., Mancini L., 2023, MNRAS, 521, 2163

Steele I. A., Bates S. D., Gibson N., Keenan F., Meaburn J., Mottram C. J., Pollacco D., Todd I., 2008, in McLean I. S., Casali M. M., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 7014, Ground-based and Airborne Instrumentation for Astronomy II. p. 70146J (arXiv:0809.3351), doi:10.1117/12.787889

Stevens D. J., Gaudi B. S., 2013, PASP, 125, 933

Stevenson K. B., 2016, ApJL, 817, L16

Struve O., 1952, The Observatory, 72, 199

Sudarsky D., Burrows A., Pinto P., 2000, ApJ, 538, 885

Taylor F. W., Atreya S. K., Encrenaz T., Hunten D. M., Irwin P. G. J., Owen T. C., 2004, in Bagenal F., Dowling T. E., McKinnon W. B., eds, , Vol. 1, Jupiter. The Planet, Satellites and Magnetosphere. pp 59–78

Tennyson J., et al., 2016, Journal of Molecular Spectroscopy, 327, 73

Thorngren D. P., Fortney J. J., 2018, AJ, 155, 214

Tinetti G., et al., 2018, Experimental Astronomy, 46, 135

Tremblin P., Amundsen D. S., Mourier P., Baraffe I., Chabrier G., Drummond B., Homeier D., Venot O., 2015, ApJL, 804, L17

Tremblin P., Amundsen D. S., Chabrier G., Baraffe I., Drummond B., Hinkley S., Mourier P., Venot O., 2016, ApJL, 817, L19

Tsai S.-M., et al., 2023, Nature, 617, 483

Tsiaras A., et al., 2018, AJ, 155, 156

Udalski A., Zebrun K., Szymanski M., Kubiak M., Soszynski I., Szewczyk O., Wyrzykowski L., Pietrzynski G., 2002, AcA, 52, 115

Vaughan A. H., Preston G. W., Wilson O. C., 1978, PASP, 90, 267

Vidal-Madjar A., Lecavelier des Etangs A., Désert J. M., Ballester G. E., Ferlet R., Hébrard G., Mayor M., 2003, Nature, 422, 143

Vidal-Madjar A., et al., 2004, ApJL, 604, L69

Vidal-Madjar A., et al., 2010, A&A, 523, A57

Virtanen P., et al., 2020, Nature Methods, 17, 261

Wakeford H. R., et al., 2013, MNRAS, 435, 3481

Wakeford H. R., Visscher C., Lewis N. K., Kataria T., Marley M. S., Fortney J. J., Mandell A. M., 2017, MNRAS, 464, 4247

Wakeford H. R., et al., 2018, AJ, 155, 29

Wakeford H. R., et al., 2020, AJ, 159, 204

Wang J. J., et al., 2021, AJ, 162, 148

Welbanks L., Madhusudhan N., 2019, AJ, 157, 206

Welbanks L., Madhusudhan N., 2022, ApJ, 933, 79

Welbanks L., Madhusudhan N., Allard N. F., Hubeny I., Spiegelman F., Leininger T., 2019, ApJL, 887, L20

Wells D. C., Greisen E. W., Harten R. H., 1981, A&AS, 44, 363

Wheatley P. J., et al., 2018, MNRAS, 475, 4476

Wilson P. A., et al., 2015, MNRAS, 450, 192

Wilson J., et al., 2020, MNRAS, 497, 5155

Winn J. N., 2010, arXiv e-prints, p. arXiv:1001.2010

Wolszczan A., Frail D. A., 1992, Nature, 355, 145

Wong I., et al., 2022, AJ, 164, 30

Wright J. T., Marcy G. W., Howard A. W., Johnson J. A., Morton T. D., Fischer D. A., 2012, ApJ, 753, 160

Wu Y., Lithwick Y., 2011, ApJ, 735, 109

Wu Y., Murray N., 2003, ApJ, 589, 605

Wyttenbach A., Ehrenreich D., Lovis C., Udry S., Pepe F., 2015, A&A, 577, A62

Yan F., Henning T., 2018, Nature Astronomy, 2, 714

Youdin A. N., Mitchell J. L., 2010, ApJ, 721, 1113

Yurchenko S. N., Tennyson J., 2014, MNRAS, 440, 1649

Yurchenko S. N., Barber R. J., Tennyson J., 2011, MNRAS, 413, 1828

Zhang K., Bloom J. S., 2020, ApJ, 889, 24

Zhang M., Chachan Y., Kempton E. M. R., Knutson H. A., 2019, PASP, 131, 034501

Zhang M., Chachan Y., Kempton E. M. R., Knutson H. A., Chang W. H., 2020, ApJ, 899, 27

Zhou G., Bayliss D. D. R., 2012, MNRAS, 426, 2483

Zhou G., et al., 2019, AJ, 158, 141

de Wit J., et al., 2016, Nature, 537, 69

dos Santos L. A., et al., 2019, A&A, 629, A47

van Dokkum P. G., 2001, PASP, 113, 1420

von Essen C., Mallonn M., Welbanks L., Madhusudhan N., Pinhas A., Bouy H., Weis Hansen P., 2019, A&A, 622, A71

von Essen C., Mallonn M., Hermansen S., Nixon M. C., Madhusudhan N., Kjeldsen H., Tautvaišienė G., 2020, A&A, 637, A76