Characterisation of Exoplanetary Atmospheres and Planetary Systems

Marco Rocchetto

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

of

University College London.

Department of Physics & Astronomy

University College London

January 19, 2017

I, Marco Rocchetto, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.

Abstract

The discovery of over three thousand exoplanets in the past two decades has unveiled a large and diverse population, far exceeding the diversity seen in our own Solar System. Today, research efforts need to shift from the discovery to the characterisation of exoplanetary systems, and this thesis aims to be a further step in this direction. Two different techniques are investigated to chemically characterise exoplanetary systems: atmospheric retrievals and metal-polluted white dwarfs.

The study of exoplanetary atmospheres through their spectra offers a very promising way to understand not only the chemistry of exoplanets, but also their atmospheric dynamics, formation and evolution history. As part of this thesis, a novel retrieval tool, called TauREx, was developed to interpret exoplanetary spectra. Spectral models were created and benchmarked with existing models, and a state-of-the-art database of absorption cross sections was also developed. The uncertainties in these models, and their propagation in the retrieval stage, were analysed in detail. These methods were used to investigate the retrievability of the carbon-to-oxygen ratio in simulated exoplanet spectra, and to interpret the atmospheres of two exoplanets, HD209458 b and 55 Cnc e. Lastly, these models were used to study the effects of stellar flares on the chemistry and spectra of typical exoplanets.

Complementary to the observations of exoplanetary atmospheres, metal polluted white dwarfs are today a unique laboratory to infer the chemical composition of terrestrial exoplanets, and to study evolved planetary systems. It has become clear that the metals seen at a fraction of white dwarfs result from accreted circumstellar dust, originating from the tidal disruption of rocky planetesimals. Through the analysis of these stars, it is possible to infer the composition of terrestrial planetesimals, as their photospheres, in principle, mirror the composition of the accreted material, in turn providing clues on the nature of rocky planetary bodies. In this thesis I will discuss this technique, and present a recent survey that has unambiguously determined the fraction of detectable planetary debris at white dwarfs.

Acknowledgements

There are many people that made this thesis possible and deserve special gratitude.

First and foremost I would like to thank my parents, my sister and my grandmother, who have supported me not only throughout this journey, but for my entire life, always allowing and encouraging me with love to pursue my interests.

I would like to express my gratitude and appreciation to my supervisor, Giovanna Tinetti. Working with you has been a privilege and a pleasure: I could have not found a better mentor and I will be forever grateful to you.

To Ingo Waldmann, my second supervisor, and a true friend. Most of this thesis is the result of a special synergy that made these three years a real adventure, and amongst the best of my life. I am sure that our friendship will continue for years to come.

To Steve Fossey, my master supervisor, and a mentor since I began my studies at UCL as a kid. Your passion and dedication, and the continuous support throughout these years, have been always inspiring to me.

My deepest thanks goes also to my close collaborators: Olivia Venot, Michael Line, Pierre-Olivier Lagage, and Carolina Bergfors. Working with you has been a pleasure.

I also thank my two examiners, Dr. Daphne Stam and Prof. Mike Barlow for taking the time to read this thesis and discuss it during the Viva.

Lastly, special gratitude goes to my friends, and especially to my housemate, who bore me during the past three years. You guys are the people that make my life special. And to my dog, who has kept me company for infinite hours!

Contents

List of Figures

List of Tables

Chapter 1

Introduction

> *Man must rise above the Earth – to the top of the atmosphere and beyond – for only thus will he fully understand the world in which he lives.*
>
> Socrates, 470 – 399 BC

This quote by Socrates, a Greek philosopher credited as one of the founders of western philosophy, best represents one of the major drivers that has always guided the scientific interest of many astronomers, including myself. Similarly to the early explorers sailing across the oceans, scientific exploration brings humankind into a voyage that will help them understand not only the physical nature of the universe, but our place as humans in it. In this context, the search for worlds similar to our own, emphasised by the question "are we alone?", is probably one of the most profound voyage that astronomy has ever taken. Countless philosophers and astronomers have tried to answer this question, which remained in the realm of scientific and popular speculation until about 25 years ago, when the first evidence of the existence of a world outside our own solar system was confirmed.

After centuries of conjecturing, it is only at the end of the 1980's that technological advances led to the first hints of planetary mass objects orbiting around other stars (Campbell et al., 1988; Latham et al., 1989), with the first unambiguous detection of a planetary system discovered around a pulsar occurring in 1992 (Wolszczan & Frail, 1992). We then need to wait three more years to have the first evidence of an exoplanet orbiting around a main-sequence star (Mayor & Queloz, 1995). From the Observatoire de Haute-Provence, Michel Mayor and Didier Queloz observed the tiny wobble of the star 51 Peg caused by the presence of a Jupiter-sized object orbiting at just 0.05 AU from its parent star. This discovery marks the beginning of an intense search for extrasolar planets, which quickly led to the rapid discovery of new planets (the total count today stands at 3375, with 2416 un-

confirmed candidates[1]), leading to a new understanding of planetary systems. With the first results from the Kepler space telescope (e.g. Borucki et al., 2011), a space mission designed to detect thousands of new exoplanets, it has been realised that our solar system is certainly not representative of the incredible diversity of the vast exoplanet population. Today, some questions about planetary science have found partial answers, while many others still lack proper explanations. Amongst others, planetary scientists pose the following questions:

- What are planets made of? What is the chemical make up of their atmosphere?
- Why do we observe such a diversity of exoplanets?
- How do planets form and evolve with age?
- How do planetary properties correlate with factors such as stellar metallicity and stellar type?
- What is the influence on the formation mechanism of the environment in which the planets form?

Different approaches exist to answer these questions. Firstly, large discovery surveys, finding thousands of exoplanets, are able to map statistically the exoplanet population, providing key information about the bulk properties of exoplanets, and the frequency of planets as a function of various external factors. Secondly, detailed characterisation of known exoplanets, through analyses of their atmospheres for instance, is paramount to have a more detailed picture, helping to explain why we observe such a diversity. This thesis will focus on this second aspect, and will discuss at length the methods used to chemically characterise planetary systems.

In the following sections I will firstly discuss the main detection methods that have been used over the past few decades to detect new exoplanets. Then, I will present some of the main results that these discoveries have led to. Lastly, I will discuss how we can shift our attention from the mere discovery of new planets, to the characterisation of existing systems. In the last section, I will give a brief overview of the thesis outline. Please note that a glossary of mathematical symbols used throughout this and the following chapters is presented at the end of the thesis.

[1]Checked on 01/09/2016, NASA Exoplanet Archive (`http://exoplanetarchive.ipac.caltech.edu`)

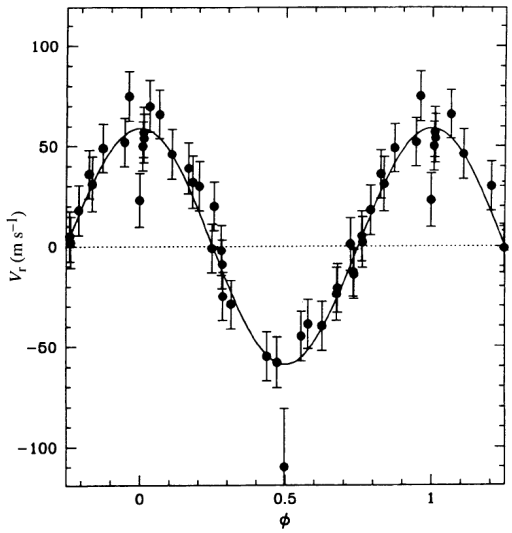

Figure 1.1: Radial velocity curve of 51 Peg (Mayor & Queloz, 1995)

1.1 Detection methods

There is a wealth of methods that, since the discovery of the first exoplanet in 1995, have been developed and refined to detect new exoplanets (see Perryman, 2014, Figure 1.1). These methods can be divided into two broad categories: direct and indirect ones. The vast majority of exoplanets have been detected using indirect methods, relying on the detection of the effects that the orbiting planet has on the host star or other sources. These include the radial velocity method, the transit method, astrometry and gravitational microlensing. Other indirect methods include timing variations, in which the orbiting planet causes some variations on some periodic phenomena on the star or other companion planets. These, amongst others, include pulsar timing and transit timing variations.

Only a tiny fraction of exoplanets has been detected directly, as the extreme contrast between a planet and its parent star makes it incredibly difficult to spatially resolve exoplanets. Nevertheless, in recent years, several dedicated imaging instruments have been developed, which, together with the creation of sophisticated post-processing algorithms, have allowed the discovery and spectroscopic characterisation of a few exoplanets.

1.1.1 The radial velocity and astrometry methods

These methods rely on the fact that, in a planet-star orbiting system, both the planet and the star orbit around the common system barycentre, or centre of mass. This causes the star to undergo a reflex motion that can be detected from a distant observer. There are two potential observable properties of the star that can be measured: a periodic change in the

radial velocity of the star, observable via the periodic shift of the star spectrum, and a change in the astrometric position of the star on the sky. The amplitude of these two effects will in turn depend on the inclination of the orbit with respect to the line of sight. In this section I will mainly focus on the radial velocity method, which has essentially driven the field for more than two decades, and I will only briefly discuss the astrometry method.

Radial Velocity

The radial velocity (RV) method symbolically represents the onset of the field of exoplanets as an observational physical science. After several years of search and speculation (Walker et al., 1995; Cochran & Hatzes, 1994; McMillan et al., 1994; Marcy & Butler, 1992), the first Jupiter-mass companion to a Sun-like star was eventually discovered using this method (Mayor & Queloz, 1995). Figure 1.1 shows the radial velocity curve of the star 51 Peg: a lot of information can be obtained from this curve. Using simple formalisms – Kepler's third law and the conservation of momentum – from a radial velocity curve we can easily determine the minimum mass of the planet, its period, eccentricity and its semi-major axis.

For a circular orbit, the semi-major axis a is related to the period P and the mass of the star M_* by

$$a^3 = \frac{GM_*}{4\pi^2}P^2, \tag{1.1.1}$$

where G is the gravitational constant. From the conservation of momentum, we also have

$$M_p v_p = M_* v_*, \tag{1.1.2}$$

where M_p is the planet mass and v_p and v_* are the planet and star orbital tangential velocities respectively. The radial velocity semi-amplitude is then given by

$$K_* = v_* \sin i = \frac{M_p v_p \sin i}{M_*}, \tag{1.1.3}$$

where i is the orbital inclination. Finally, as $v_p = 2\pi a/P$, we have

$$K_* = v_* \sin i = \frac{2\pi a M_p \sin i}{M_* P}. \tag{1.1.4}$$

This can be rearranged to give the mass of the planet, as a factor of the orbital inclination i:

$$M_p \sin i = \frac{K_* M_* P}{2\pi a}. \tag{1.1.5}$$

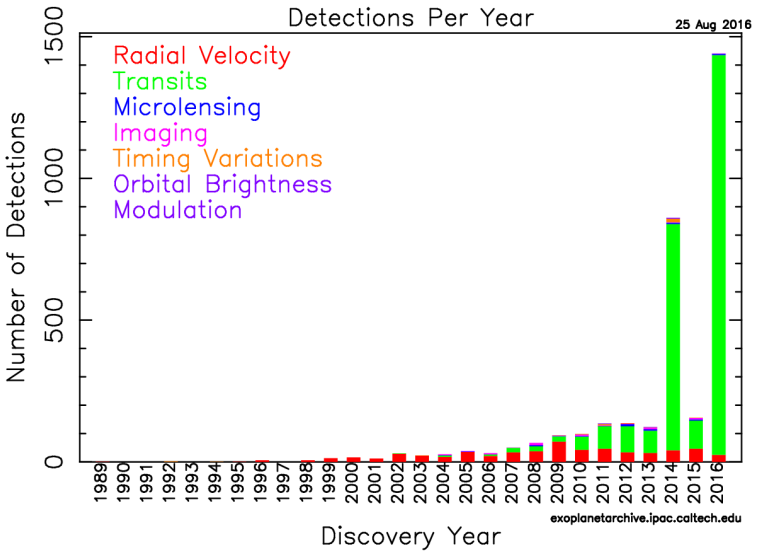

Figure 1.2: Histogram showing the total number of confirmed exoplanet discoveries by detection method through 2015 from the NASA Exoplanet Archive.

The velocity semi-amplitude, K_*, and the period, P, can be directly measured from the radial velocity curve, the semi-major axis can be computed using Equation 1.1.1, while the stellar mass, M_*, can be obtained using stellar models of the planet's host.

The radial velocity method alone, however, cannot unambiguously confirm the presence of a planetary mass companion. Due to the degeneracy between M_p and $\sin i$, small values of K_* could equally suggest either a planet, or an object with much higher mass but smaller orbital inclination. Such a degeneracy can be broken in several ways. Methods to provide additional information leading to an unambiguous determination of the planet mass include constraints on the orbital inclination from photometric transits (e.g. Charbonneau et al., 2000), spectroscopic line profiles analysis (e.g. Hekker & Aerts, 2010) or statistical deconvolution (e.g. Jorissen et al., 2001).

The idea of detecting radial velocity signals from planetary companions spans many decades (Struve, 1952), but it is only with the advances in observational techniques at the end of the last century (e.g. ELODIE spectrograph, Baranne et al., 1996), that the very tiny changes in radial velocity expected from planet-like companions (tens of m/s for the most massive and close-in planets), could be detected.

From the first unambiguous detection of a planet orbiting a main sequence star in 1995, the radial velocity method has been the primary way of detecting planets until about 2011 (see Figure 1.2), when the first results from Kepler, a major space-based transit survey, were published. Figure 1.3 shows the planets discovered using this technique as a function

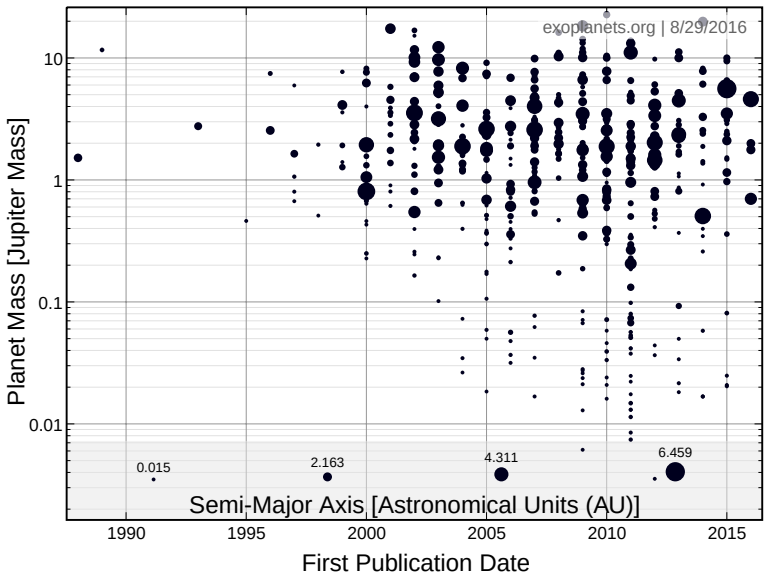

Figure 1.3: Planets discovered by radial velocity measurements according to mass (y-axis) and year of discovery (x-axis). The circle sizes are proportional to the semi-major axis. Data and plot are from `exoplanets.org`, 29/08/2016.

of planet mass and semi-major axis. It can be seen that the lower mass bound decreases exponentially, reflecting the increasing accuracy in radial velocity measurements.

The two most successful high-performance instruments available today are HARPS at the ESO 3.6-m telescope at La Silla (Chile) and HIRES at the Keck I 10-m telescope (Hawaii/US). Both instruments are echelle spectrographs: HIRES (High Resolution Echelle Spectrometer) has a resolution of $R = 80\,000$ (Vogt et al., 1994), with radial velocity precision up to $1\,\mathrm{m\,s^{-1}}$; HARPS (High Accuracy Radial Velocity Planet Searcher), has a resolution of $R = 115\,000$ and was designed to achieve accuracies of around $1\,\mathrm{m\,s^{-1}}$ (Mayor et al., 2003). One of the most recent and interesting discoveries made by HARPS is the detection of a terrestrial planet orbiting within the habitable zone of Proxima Centauri, the closest star to the Sun (Anglada-Escudé et al., 2016).

Nowadays significant effort is invested in the development of spectrographs that can reach sensitivity below $1\,\mathrm{m\,s^{-1}}$, allowing the potential detection of Earth-like exoplanets around solar type stars. Future instruments include ESPRESSO, a super-stable high resolution spectrograph whose main objectives include the measurement of high precision radial velocities of solar type stars for search for rocky planets. ESPRESSO (Pepe et al., 2010), is the successor of several echelle spectrometers (e.g. Elodie, Coralie, HARPS), and will improve on their best properties. It will have enhanced stability and increased resolution

($R = 225\,000$), reaching accuracies of less than $10\,\mathrm{cm\,s^{-1}}$ (Mégevand et al., 2010). It will be installed at ESO/VLT at Paranal, and can operate using the combined light of all four 8.2 m Unit Telescopes (UT). Building on techniques developed for HARPS, in to order to ensure stability and accuracy of better than 10^{-12}, ESPRESSO will integrate a laser frequency comb (Steinmetz et al., 2008; Pepe et al., 2010).

Further developments and enhanced accuracy in the instrumentation will certainly make the radial velocity method an invaluable resource to search for new exoplanets for many decades to come.

Astrometry

The method based on astrometric measurements had much lower success in finding new exoplanets, mainly due to the lack of sufficient astrometric accuracy. Astrometry is closely related to the radial velocity method, as both aim to detect the same periodic gravitational perturbation of an orbiting planet on the host star. Repeated high-accuracy astrometric measurements of a planet-hosting star can measure the transverse component of the displacement of the star due to the orbiting planet. However, the best accuracies of about 1 mas, achieved with Hipparcos and the Fine Guidance Sensors on-board the Hubble Space Telescope (HST), have only began to reach the regimes where these periodic displacements of known systems can be detected (Reffert & Quirrenbach, 2011; Sahlmann et al., 2011; McArthur et al., 2010). As of today, there is only one confirmed detection of a massive $28 \pm 2\ M_J$, long period ($P = 246.4 \pm 1.4$ days) object using this method (Sahlmann et al., 2013). This exoplanet was observed using FORS2/VLT optical imaging for high precision astrometry, reaching an accuracy of two tenths of a milli-arcsecond over two years.

The picture will change with the Gaia space-telescope, launched at the end of 2013, which will monitor the astrometric positions of billion of stars with an accuracy of 20–25 μas. Gaia is expected to discover thousands of exoplanets astrometrically (Dzigan & Zucker, 2012; Sozzetti et al., 2014; Perryman et al., 2014).

1.1.2 The transit method

A transit is produced when two or more orbiting objects are seen nearly edge-on. In this case the two objects periodically eclipse each other. Transits have always been of great importance in astronomy, leading for example to the confirmation of Einstein's prediction of the light deflection due to gravity (Poor, 1930). When the orbital inclination of a planet-star system relative to the normal of the line-of-sight is close to 90 degrees, the planet is

expected to *transit* in front of its host star. The *indirect* observable effect is a temporary and periodic dimming of the star brightness, proportional to the square of the ratio of the planet and star radii. The first transiting extrasolar planet independently observed by two groups with this method is HD 209458b (Charbonneau et al., 2000; Henry et al., 2000), a Jupiter-sized planet with a period of ≈ 3.5 days, originally discovered using the radial velocity method (Mazeh et al., 2000).

Transiting exoplanets need to be confirmed using radial velocity measurements to obtain a tight constrain on their mass, and to exclude alternative solutions, such as eclipsing binaries or blended background stars. The two methods are therefore complementary. The amount of information that can be derived from a transit light curve *and* radial velocity curve goes far beyond the physical parameters obtainable with each individual method alone. When the two methods are combined, the planet and its orbit can be fully characterised: while the light curve provides the planet radius and the orbit inclination, the radial velocity curve gives the mass of the planet, providing in turn the planet density, hence giving a first hint of its composition. Transiting planets are also accessible to spectroscopic measurements, giving valuable clues about their atmospheric composition (see Section 1.3.1).

Before presenting the relevant equations, it is useful to clarify some terminology to avoid confusion. An *eclipse* is the obscuration of one object by another. When the relative sizes of the objects differ significantly, the passage of the small body in front of the bigger one is called a *transit* or *primary eclipse*, while the passage of the smaller body behind the larger one is called *occultation* or *secondary eclipse* (see Figure 1.4). When the two bodies partially overlap each others, the eclipse is referred to as *grazing*.

From a transit light curve it is possible to find a unique solution for the stellar mass, M_*, stellar radius, R_*, companion radius, R_p, orbital semi-major axis, a and orbital inclination ,i, under the assumptions that the stellar mass-radius relation is known and that the period, P, can be derived from consecutive observations of the light curve, or through RV data (Seager & Mallén-Ornelas, 2003a). In addition, the light needs to come from a single star, rather than from two or more blended stars. Then, the various parameters can be derived as follows[2]:

[2]Note that, for simplicity, in the following expressions we assume a circular orbit. For elliptical orbits see Kipping (2008).

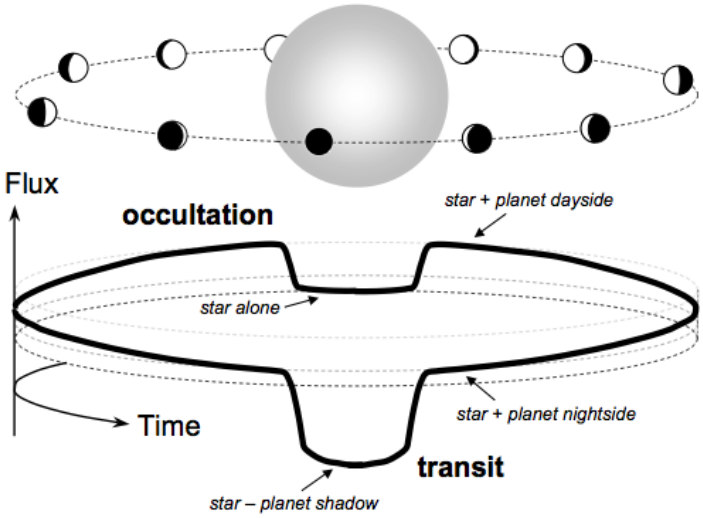

Figure 1.4: Illustration of transits and occultations (Winn, 2010).

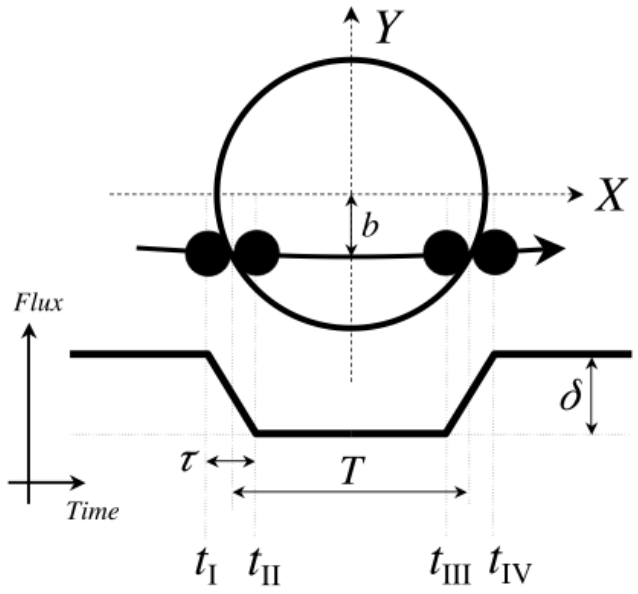

Figure 1.5: Geometry of a primary transit light curve (Winn, 2010).

The ratio R_p/R_* is given by

$$\frac{R_p}{R_*} = \sqrt{\Delta F} = \sqrt{\frac{F_{\text{no transit}} - F_{\text{transit}}}{F_{\text{no transit}}}}, \tag{1.1.6}$$

where F_{transit} and $F_{\text{no transit}}$ are the measured relative star fluxes when the planet is in transit and out of transit, respectively. This simple expression allows us to see that the sought effect is relatively small: for a Jupiter-like planet orbiting around a Sun-like star, the expected drop in star brightness is about 1%, or 0.01 mag. Such a change in star magnitude is easily detectable even using small "backyard" telescopes[3]. However, for a smaller, Earth-like planet, orbiting a Sun-like star, drops of only 10^{-4} are expected, requiring far greater photometric accuracy than that achievable from the ground.

Various other parameters can be obtained from a transit light curve. The impact parameter, b, is given by

$$b = \left[\frac{\left(1 - \sqrt{\Delta F}\right)^2 - (t_F/t_T)^2 \left(1 + \sqrt{\Delta F}\right)^2}{1 - (t_F/t_T)^2} \right]^{1/2}, \tag{1.1.7}$$

where $t_T = t_{\text{IV}} - t_{\text{I}}$ and $t_F = tt_{\text{III}} - t_{\text{II}}$ (see Figure 1.5). The ratio a/R_* is

$$\frac{a}{R_*} = \frac{2P}{\pi} \frac{\Delta F^{1/4}}{\left({t_T}^2 - {t_F}^2\right)^{1/2}}, \tag{1.1.8}$$

the stellar density, ρ_*, is

$$\rho_* = \frac{32}{G\pi} P \frac{\Delta F^{3/4}}{\left({t_T}^2 - {t_F}^2\right)^{3/2}}, \tag{1.1.9}$$

where G is the gravitational constant. Combining Equations 1.1.8 and 1.1.9, the stellar density is also given by

$$\rho_* = \frac{3\pi}{GP^2} \left(\frac{a}{R_*}\right)^3. \tag{1.1.10}$$

From Kepler's third law, assuming $M_p \ll M_*$, the semi-major axis, a, is

$$a = \left(\frac{P^2 G M_*}{4\pi^2}\right)^{1/3}, \tag{1.1.11}$$

[3]The Exoplanet Transit Database (`http://var2.astro.cz/ETD/`) contains thousands of transit light curves obtained with all sorts of amateur equipment, and several peer-reviewed papers have used data from this resource (see Poddaný et al., 2011, and references therein)

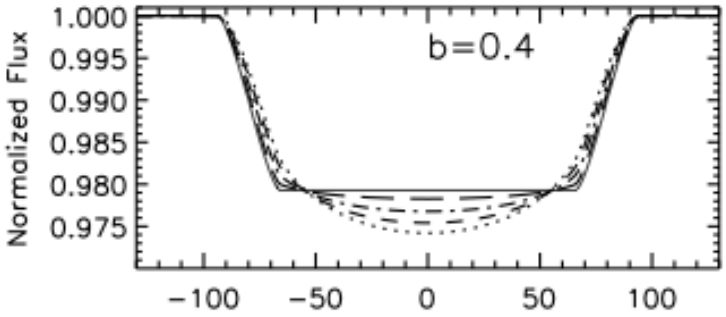

Figure 1.6: The effect of limb darkening on transit light curves. The solid curve shows a transit with limb darkening neglected. The other light curves have solar limb darkening at wavelengths 3, 0.8, 0.55 and 0.45 μm (Seager & Mallén-Ornelas, 2003b)

and, based on the definition of impact parameter (Equation 1.1.7), the orbital inclination i is

$$i = \cos^{-1}\left(b\frac{R_*}{a}\right). \tag{1.1.12}$$

Southworth (2010) also showed that it is possible to derive the planetary surface gravity $g_{p,\mathrm{surf}}$ from the radial velocity curve of its parent star and the analysis of its transit light curve:

$$g_{p,\mathrm{surf}} = \frac{2\pi}{P}\frac{K_*}{(R_p/a)^2\sin i}, \tag{1.1.13}$$

where K_* is the radial velocity semi-amplitude of the star,

A complete derivation of these equations is given by Seager & Mallén-Ornelas (2003b) for circular orbits, expanded by Kipping (2008) for elliptical orbits.

So far we have ignored the fact that real stellar disks are brighter in the middle and fainter at the edge, a phenomenon called *limb darkening*. Limb darkening causes a change in the depth of the light curve ΔF as a function of impact parameter, making the flat bottom rounder and causing a blur in the boundary between ingress/egress and the flat bottom (see Figure 1.6). Many limb-darkening laws that attempt to model this effect have been proposed. The most commonly used in the analysis of transiting exoplanets are the quadratic law (Kopal, 1950):

$$\frac{I_\mu}{I_0} = 1 - u_1(1-\mu) - u_2(1-\mu)^2, \tag{1.1.14}$$

and the Claret four-parameter (non linear) law (Claret, 2000):

$$\frac{I_\mu}{I_0} = 1 - u_1(1-\mu^{1/2}) - u_2(1-\mu) - u_3(1-\mu^{3/2}) - u_4(1-\mu^2), \tag{1.1.15}$$

where I_0 is the specific intensity at the centre of the disk, $\mu = \cos\theta$, θ being the angle between the line of sight and the emergent intensity, u_{1-4} are the *limb darkening coefficients*, which can be derived from numerical models.

Although a unique solution still exists when limb darkening is taken into account, the equations describing the transit light curve are very complicated and a simple analytical solution no longer exists. Mandel & Agol (2002) developed the exact analytic solution for the eclipse of a star described by quadratic or nonlinear limb-darkening laws. This publicly available algorithm[4] is today commonly used to model exoplanet transit light curves. I refer the interested reader to Haswell (2010) (pp. 103-112) for a full description of the algorithm.

After the first discovery of a transiting exoplanet (Charbonneau et al., 2000; Henry et al., 2000), planets detected from radial velocity measurements were soon closely followed-up to look for possible transits (Kane, 2007; Kane et al., 2009). These early discoveries were followed by an increasing number of detections coming from "blind" wide-field photometric surveys, where thousands of stars were continuously monitored looking for potential transit signatures. Ground based surveys aim at detecting $\sim$1% variations in the light curves, enabling the discovery of Jupiter-sized planets around bright Sun-like stars, or, alternatively, super-Earth planets around M-dwarfs. Surveys from space allow the detection of much smaller planets, down to $\sim 1\ M_{\oplus}$, thanks to the enhanced photometric precision obtainable from space. For unambiguous discoveries, however, such detections need to be confirmed by ground-based precision radial velocity measurements using high-performance instruments such as HIRES and HARPS. For the smallest planets, or for planets orbiting very active stars, this is not always achievable with the current instrumentation, and future facilities such as ESPRESSO at the ESO/VLT or larger telescopes such as the European Extremely Large Telescope, together with improved stellar variability modelling, will be needed to confirm these planets.

The most successful ground based transit surveys are the HAT and SuperWASP surveys. The Hungarian Automated Telescope is divided into its northern hemisphere version (HATNet, Bakos et al., 2002, 2004), which saw first light in 2003, and its southern hemisphere counterpart (HATSouth, Bakos et al., 2013), with first light in 2009. HATNet is a network of seven small telescopes distributed in Arizona and Hawaii (US). HATSouth consists of six telescopes distributed over South America, South Africa and Australia. Both

[4] `http://faculty.washington.edu/agol/transit.html`

surveys together have discovered 93 planets as of July 2016[5]. The SuperWASP network (Pollacco et al., 2006) is a UK led program consisting of two wide-field camera arrays in the northern (La Palma, Spain) and southern (Sutherland, South Africa) hemispheres. It is similar in design and scope to the HAT survey. SuperWASP is so far the most successful ground based survey, with 130 planets discovered as of today[6]. Other ground based surveys include OGLE (Udalski et al., 2002, 7 hot Jupiters discovered), TrES (O'Donovan et al., 2006; Alonso et al., 2007, 5 hot Jupiters), XO (McCullough et al., 2005, 6 hot Jupiters), and MEarth (Nutzman & Charbonneau, 2008, 2 super-Earths).

Space based dedicated transit surveys have the obvious advantage of providing increased sensitivity to much smaller planets. The CoRoT satellite, a French-led mission with several European partners, was the first of its kind, and saw first light in 2006. Its main objectives were asteroseismology and exoplanet detection. It consisted of a 0.27 m diameter telescope with a field of view of 2.7×3.0 square degrees (Auvergne et al., 2009), and led to the detection of 32 confirmed planets. The most notable discovery is probably CoRoT-7b (Queloz et al., 2009), a super-Earth with $M_p = 8.0 \pm 1.2\,M_\oplus$ (Ferraz-Mello et al., 2011), the first extrasolar potential terrestrial planet to be found.

Undoubtedly the king of all transit surveys is NASA's Kepler satellite (Borucki et al., 2010; Koch et al., 2010), launched in 2009 into an Earth-trailing heliocentric orbit. It was designed to discover Earth-size exoplanets near or in the habitable zone[7], and to provide definite statistics on the occurrence rate of such planets around main-sequence stars in the Milky Way. Kepler is a 0.95 m aperture telescope, covering 115 square degrees of sky. It has monitored 150 000 main sequence stars between V = 8–15 mag for over three years, reaching a photometric precision of 25 ppm (Gilliland et al., 2011). The core mission reached an end when two of the four reaction wheels failed in 2013[8]. However, this technical failure did not lead to a decommissioning of the telescope. At the end of 2013, a new ingenious plan, named K2, was presented by NASA: using only two reaction wheels, and the radiation pressure from the photons of sunlight, the spacecraft can be stably balanced, achieving a photometric precision of 80 ppm over 6 hr (Howell et al., 2014). Such position can be maintained for about 83 days, until the spacecraft needs to be rotated to avoid sunlight con-

[5] `http://hatnet.org/planets/` and `http://hatsouth.org/planets/`

[6] `http://exoplanet.eu/catalog/`, checked on 30/08/2016

[7] The habitable zone is defined as the range of orbits around a star within which a planetary surface can support liquid water given sufficient atmospheric pressure

[8] Kepler Mission Manager Update, May 21, 2013 (`http://www.nasa.gov/mission_pages/kepler/news/keplerm-20130521.html`

tamination. This has clearly changed the mission objectives, but K2 still provides data to study and discover more exoplanets, and it is expected to operate in this mode until 2019. Kepler has so far found 2330 planets, with 2416 additional unconfirmed candidates, while K2 has found 129 confirmed planets, with 329 additional unconfirmed candidates[9].

The future of space based transit searches is bright. There are two major planned dedicated missions in the pipeline: PLATO 2.0 and TESS. PLATO 2.0 (Rauer et al., 2014) is an ESA-led mission with a launch opportunity in 2022–2024. It consists of 34 small aperture telescopes combined to provide a total field-of-view of 2232 square degrees and a large photometric magnitude range (V = 4–16 mag). Its focus is on the detection of Earth-size transiting planets whose mass can be determined from ground-based radial-velocity follow-up, overcoming one of the major limitation of the Kepler planet candidates, which are often too faint to be followed-up from the ground. TESS (Ricker et al., 2014) is a NASA led mission scheduled for launch in 2017, which will monitor 200 000 main sequence bright dwarf stars, allowing the follow-up of these targets from the ground. One of the main goals of TESS is to provide targets for further spectroscopic characterisation by the JWST (Deming et al., 2009).

1.1.3 Other detection methods

Gravitational microlensing

Gravitational microlensing occurs when the gravitational field of a star acts like a lens, enhancing the light coming from a distant background star in our galaxy (Paczynski, 1986). In the context of exoplanet detection, the star-planet system acts as a multiple lens, while a distant star coming from the Galaxy acts as the source (for a review see Sackett, 2004; Wambsganss, 2004; Gould, 2005; Gaudi, 2008). The enhanced magnification of the source due to the changing alignment geometry of the observer, planetary host, and background source, can lead to changes in brightness over time scales of days to weeks. These changes can be detected by monitoring continuously the light curve as the alignment of the lensing system changes. As for transiting planets, continuous monitoring of a large number of distant stars is needed. However, unlike transit events, microlensing events are once-and-only events. This method is most efficient for planets between the Earth and the centre of the galaxy, as the galactic bulge provides a larger number of relatively fast-moving background

[9]Checked on 01/09/2016, NASA Exoplanet Archive (`http://exoplanetarchive.ipac.caltech.edu`)

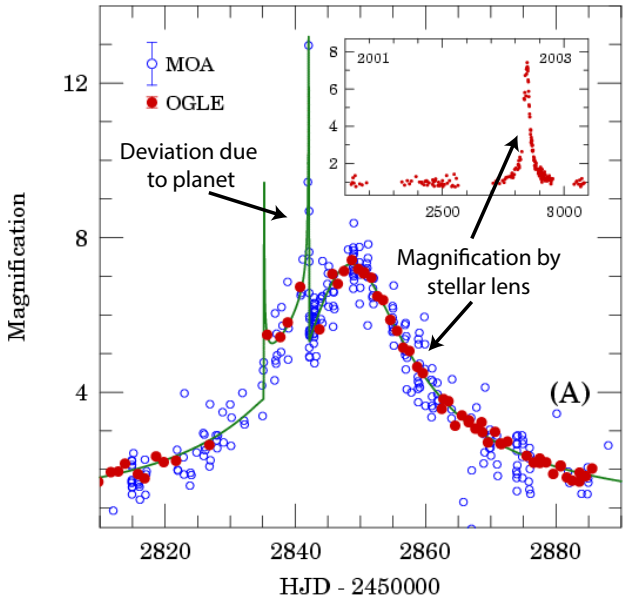

Figure 1.7: The first microlensing planet system, OGLE-2003-BLG-235. The red dots and blue dots show data from OGLE and MOA respectively. The two spikes seen at days 2835 and 2842 correspond to the caustic crossing (entry and exit) of the planet.

sources.

Despite some obvious disadvantages of this method, such as non-repeatability and small probability of detection, there are number of advantages in using this technique. Firstly, there is no bias for nearby stars and for planets around certain type of stars. Almost all conventional planet detection methods concentrate their efforts on nearby stars and on specific stellar types. There is also no bias for planets with large masses, unlike conventional techniques that are significantly more sensitive to massive planets. This method has also the advantage to detect large-semi major axes, long-period planets with an instantaneous detection (see Figure 1.9, showing that all planets with orbits above 10 AU have been detected by microlensing events). The detection of such planets using the radial velocity or transit methods would instead require decades. Lastly, microlensing events can provide an independent statistics of galactic population of planets (e.g. Cassan et al., 2012).

The current effort to detect microlensing events operate in a two-step mode (Gould & Loeb, 1992; Han & Kim, 2001; Han, 2007). A wide-field survey telescope observes the early stages of a microlensing event using coarse time sampling. Once an event is detected, an alert is issued, and an array of smaller follow-up telescopes distributed across the Earth is used to provide high-precision photometry. Currently, there are two monitoring

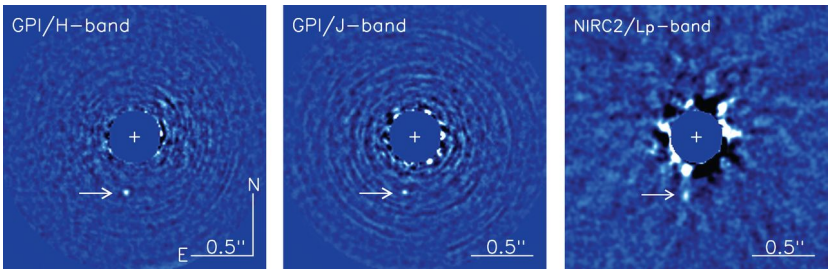

Figure 1.8: *H*, *J*, and *Lp*-band images of 51 Eridani b (shown with an arrow) obtained with GPI (Macintosh et al., 2015).

teams, the MOA (Bond et al., 2001) and OGLE (Udalski et al., 2002) collaborations, and three monitoring groups, PLANET/RoboNet (Beaulieu et al., 2007), MicroFUN (Yoo et al., 2004) and MiNDSTEp (Dominik et al., 2010), pursing follow-up. As of today, there are 16 exoplanets discovered using the microlensing technique[10]. The first confirmed exoplanet detected using such method is OGLE-2003-BLG-235 (Bond et al., 2004, Figure 1.7).

Direct imaging

The direct imaging is technically the only *direct* method available to detect exoplanets. It implies the detection of a point source image of the exoplanet, coming either from the reflected light from the parent star in the visible, or through its thermal emission in the infrared. It is worth mentioning that what we observe is just a point source: obtaining resolved spatial imaging of an exoplanet surface (e.g. Bender & Stebbins, 1996) will remain in the realm of science fiction for several decades.

One of the major challenges in the direct imaging of exoplanets is the extreme contrast ratio between the planet and host star. The ratio of the planet to stellar brightness is in fact very small: for the Jupiter-Sun system, it is $\sim 10^{-9}$ in the near infrared, while for typical exoplanets this value is expected to range from $\sim 10^{-5}$ in the infrared, to $\sim 10^{-10}$ in the visible (Perryman, 2014). Another important aspect is the angular separation: typical directly imaged exoplanets lie within 0.1 to 0.5 arcsec from their host, and are therefore hidden below the point spread functions (PSF) of typical ground based observations. Even if adaptive optics is used, or if observations are performed from space, diffracted light from the telescope and scattered light from aberrations result in instrumental 'speckles' that are often difficult to disentangle.

In recent years dedicated ground-based instruments have been developed to perform targeted imaging surveys of exoplanets. SPHERE (Spectro-Polarimetric High Contrast

[10] `exoplanets.org`, 31/08/2016

Exoplanet Research, Beuzit et al., 2008) is a second-generation instrument installed at ESO/VLT (Chile) that achieved first light in 2014. It includes an extreme adaptive optics system, with a large actuator deformable mirror and a wavefront sensor. It is designed to detect giant planets at contrasts of 10^{-6} at J = 6 mag, and at 0.1–3 arcsec from the host star.

A similar instrument, installed at the Gemini Observatory in Chile, is GPI (Gemini Planet Imager, Macintosh et al., 2008), which saw first light in 2013. It shares many design features with SPHERE, including an extreme adaptive optics system with 1800 actuators. It is designed to detect planets with a contrast of 10^{-7} at I = 6 mag, with an inner working angle of only 0.13 arcsec.

The first exoplanet discovered with SPHERE is HD 131399Ab (Wagner et al., 2016), a $4 \pm 1 M_J$ exoplanet orbiting in a triple-star system, making this planet dynamically unlike any other known system, while 51 Eridani b (Macintosh et al., 2015) (Figure 1.8) is the first exoplanet discovered with GPI. The integral field spectrographs included in both instruments enabled the acquisition of low resolution emission spectra of their atmospheres: both spectra show strong features of methane and water.

1.2 The current picture

Amongst the exciting discoveries of recent years, one of the most striking is that, unlike in our solar system, there appears to be a continuous distribution of planetary radii (Batalha et al., 2013), ranging from sub-Earths to super-Jupiters: no transition is observed between terrestrial planets to gas and icy giants. Orbital characteristics also seem not to follow the solar system model: several highly eccentric planets have been found (e.g. HD 80606b), or planets orbiting two stars (e.g. Kepler-34b, Kepler-35b, Kepler-38b), that have no solar system analogue. Detected planets also range from exotic planets such as the extremely hot and possibly rocky CoRoT-7b, Kepler-10b and Kepler-78b (Rouan et al., 2011; Grasset et al., 2009), to so-called ocean planets with density in between silicate and gaseous planets (Léger et al., 2004; Sotin et al., 2007).

Figure 1.9 shows the status of the confirmed exoplanet detections whose mass and semi-major axis have been measured, using different discovery methods. We can see that the Jovian planets are well represented up to a few tens of AU, while, for intermediate mass planets (0.01–1 M_J) planets up to a few AU have been detected. For masses below 0.01 M_J ($\approx 3 M_\oplus$), only planets within 1 AU have been discovered. The lack of discoveries of wide

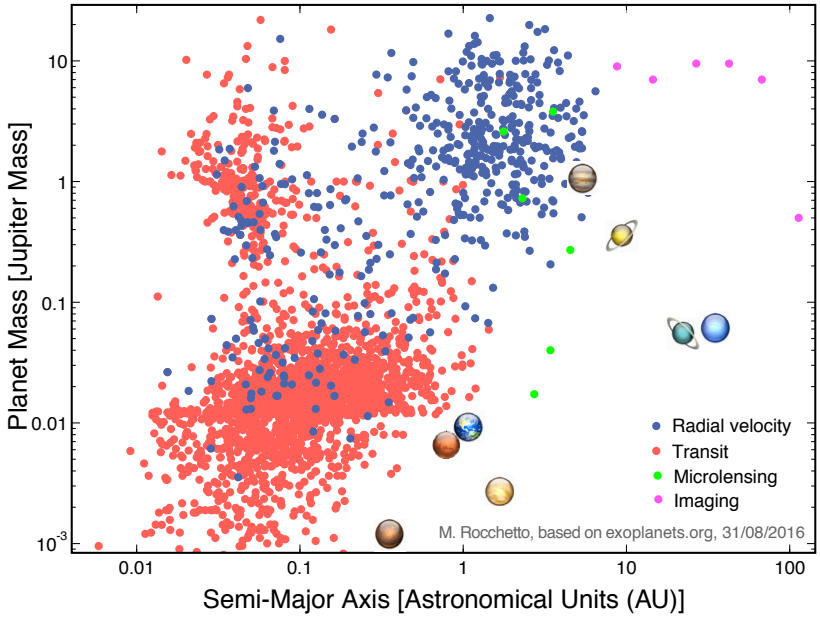

Figure 1.9: Planet discovered as of 31 August 2016 using radial velocity (blue dots), transit (red dots), microlensing (green dots), and direct imaging (purple dots), according to mass (y-axis) and semi-major axis (x-axis).

orbit, low-mass planets is, however, mainly due to limitation in current discovery methods. Transit and RV detections beyond 1 AU require years of continuous monitoring, something that PLATO will attempt to do for transiting systems. It is also important to stress here that most of the RV planets do not have measured radii, hence do not provide clues about their mean density. From this figure it is clear that much effort is still needed to access the region of smaller and wider orbit planets.

Accurate and complete planet population statistics helps to constraint planet formation models. By comparing simulations of dynamical evolutions with the observed distributions of masses and locations in planetary systems, population statistics can efficiently inform theories of planetary formation. For example, in the context of giant planet formation, models need to identify how a sufficiently massive accreting core can form, which can in turn accumulate the necessary amount of gas before the protoplanetary gas disk dissipates. It is with the orbital characterisation and general structure of a large sample of exoplanetary systems that models can therefore be tested.

In the case of gas giant planets, current models favour two potential scenarios, initially developed for the solar system: core-accretion, and gravitational disk instability. The *core-accretion scenario* (Alibert et al., 2004, 2005; Hubickyj et al., 2005; Fortier et al., 2007; Guilera et al., 2011; Fortier et al., 2013), can be seen as a two stage process. In the first

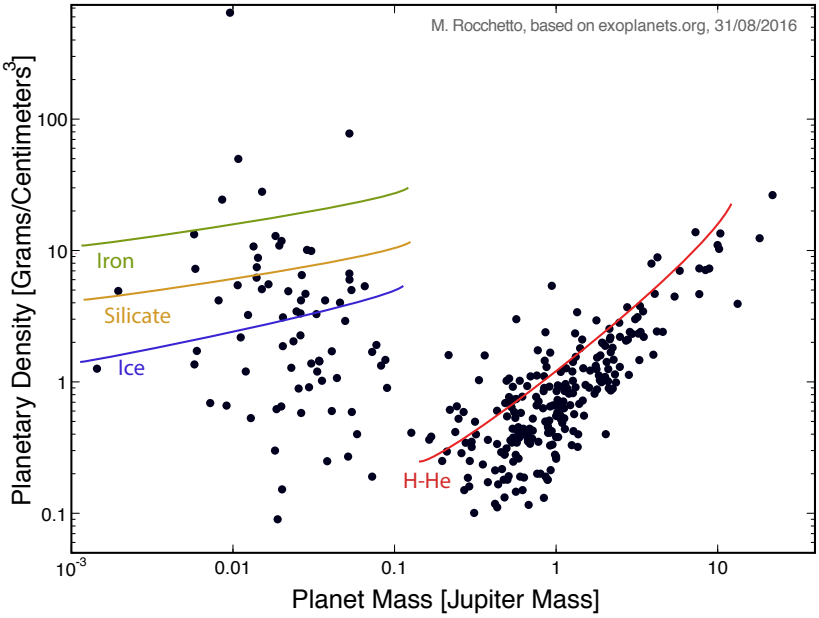

Figure 1.10: Planet density as a function of planet mass. Density lines are shown for different bulk compositions. These lines were taken from Rauer et al. (2014) (Figure 4), who derived them following Wagner et al. (2012).

stage, a massive planet or core (5–10 $M_\oplus$) is formed by the accretion of planetesimals. This is followed by a second stage of rapid accretion of gas onto the core, with continued accretion of planetesimals. Only cores with sufficient mass (the "critical core mass" or "cross over mass") can lead to large-scale gas accretion, and these can only form under certain conditions within the protoplanetary disk. Moreover, the disk dissipation timescales impose a strict timescale (5-10 Myr) over which this process can occur. A second scenario is the *disk instability scenario* (Boss, 1997; Mayer et al., 2005; Boley et al., 2010; Helled & Bodenheimer, 2011; Vazan & Helled, 2012), where gravitational instability in a cold planetary disk results in the formation of a gas giant clump. This clump then contracts as it cools down, eventually accreting planetesimals and forming a planetary core.

The core accretion scenario can be tested by high density, low-mass rocky exoplanets, as these planets define the critical mass over which accretion of the surrounding gas starts. Figure 1.10 shows the planet density as a function of planet mass for all confirmed exoplanets with measured radii and masses. The coloured lines show different bulk compositions. By following the silicate composition line with increasing mass, we can see that there are no planets detected above about 0.03 M_J (or $\sim 10 M_\oplus$). This favours the core accretion scenario, as high-mass silicate planets would quickly accrete significant envelopes, ending up as high mass but lower density planets (Hori & Ikoma, 2011). However, it can also be

appreciated that Figure 1.10 is sparsely populated in the low-mass regime: increasing the number of detections of low-mass planets with measured radii is therefore paramount to constrain such models.

Although detailed statistics on planet population are still a matter of debate, there is general agreement that rocky planets are ubiquitous around solar-type stars. This has been independently confirmed by radial velocity measurements (Bonfils et al., 2013; Mayor et al., 2011; Udry & Santos, 2007), microlensing discoveries (Cassan et al., 2012), and the Kepler sample (Dong & Zhu, 2013; Fressin et al., 2013; Howard et al., 2012; Petigura et al., 2013). This finding supports again the core accretion formation model, which predicts rocky planets to outnumber Jupiter and Neptune-like planets (Ida & Lin, 2004; Mordasini et al., 2009, 2012).

In this section I briefly explored how current and future *discovery* surveys are crucial for the understanding of planetary systems and for the development of formation and evolution models. While here I only gave an example of how current discoveries have driven planetary system models, I refer the interested reader to more comprehensive reviews (Mayor et al., 2014; Lissauer et al., 2014; Hatzes, 2014).

1.3 From discovery to characterisation

Despite the important discoveries that planet population statistics have yield so far, there is still large uncertainty about the detailed composition and structure of exoplanets. For instance, the density derived from mass and radius, despite giving some clues about the properties of an exoplanet, cannot uniquely constrain its bulk composition, due to the degeneracies between models, and the relative uncertainties of the measurements (e.g. Adams et al., 2008; Valencia et al., 2007, 2013). Today, it is therefore necessary to shift part of the attention to the characterisation of existing systems. I will introduce how this can be achieved in the next two sections using two techniques: spectroscopy of exoplanetary atmospheres, and the study of metal-polluted white dwarfs.

1.3.1 Spectroscopy of exoplanets

The study of exoplanetary atmospheres represents one of the most immediate and direct ways to characterise exoplanets, and can potentially provide answers to some important questions. Amongst others, knowledge of the atmospheric structure and composition of an exoplanet could provide important clues about the following issues:

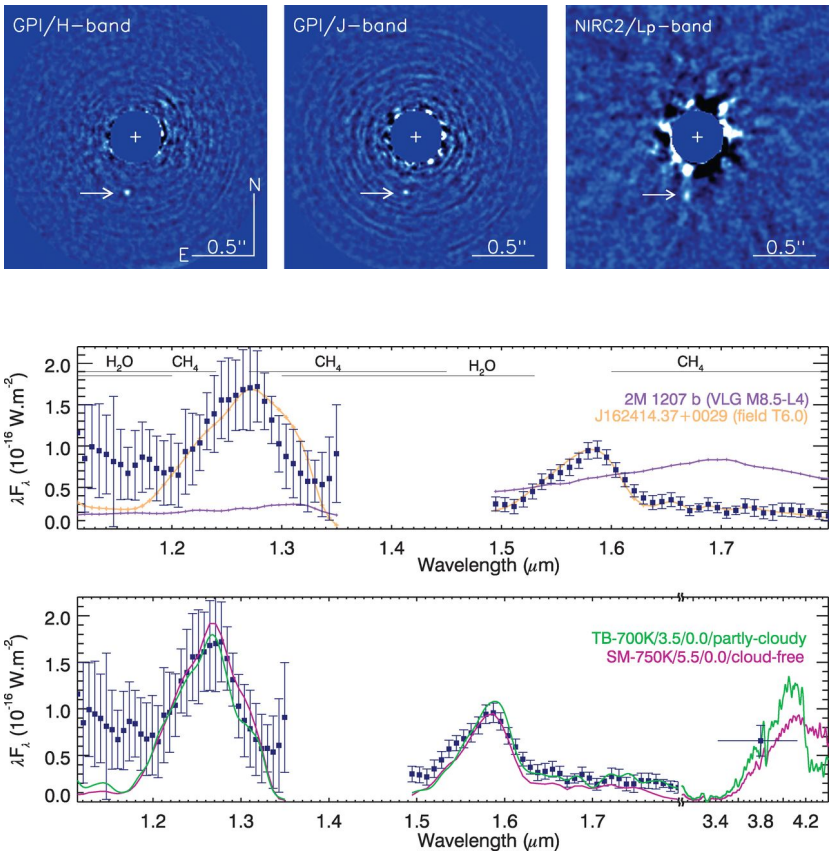

Figure 1.11: *J* and *H*-band spectra for 51 Eridani b from GPI data (Macintosh et al., 2015).

- What is the relationship between bulk and atmospheric composition?
- How are heavy elements distributed in gaseous planets? Are they retained inside the core?
- Do planets around active stars keep their atmospheres? This is an important question for habitability.
- Many hot Jupiters are found to have a radius larger than expected (Bodenheimer et al., 2001; Guillot et al., 2006). Why are these gaseous planets *inflated*?

There are mainly three techniques to obtain the spectrum of an exoplanet:

Spectroscopy of directly imaged planets

Dedicated instruments such as SPHERE and GPI include a low resolution integral field unit spectrograph, making it possible to observe a low resolution emission spectrum of the imaged exoplanet (e.g. Macintosh et al., 2015; Wagner et al., 2016). Figure 1.11 shows the images and spectra of 55 Eridani b obtained with GPI (Macintosh et al., 2015), showing

strong features of methane and water. It is worth noting, however, that determining the detailed atmospheric temperature structure and composition with such spectra is somewhat difficult, as their mass and radius are often unknown, therefore leading to significant model degeneracies.

Directly imaged planets are also well suited to detect polarisation signals (Stam et al., 2006; Marley & Sengupta, 2011; de Kok et al., 2011), which can help to constrain the presence of clouds and atmospheric inhomogeneities, or the flatting due to the planet's rotation rate. However, detecting polarised signals will be very challenging, as sensitivities of 10^{-6} need to be reached.

High dispersion spectroscopy

By observing consecutive high-resolution ($R \sim 100\,000$) spectra of an exoplanet as it orbits around the star, it is possible to detect the changing Doppler shift of the planet signal. This is due to the varying radial component of the planet orbital velocity, which is of the order of ~ 100 km s^{-1} for hot Jupiters. This ground-based technique has led to the detection of carbon monoxide and water vapour in the atmosphere of a few transiting, and non-transiting, exoplanets (Snellen et al., 2008; Brogi et al., 2012, 2013; Rodler et al., 2012; de Kok et al., 2013; Brogi et al., 2014).

Transit spectroscopy

By observing the planetary transit at multiple wavelengths using a mid-resolution spectrograph, it is possible to trace the absorption signature of broad roto-vibrational transitions of gases present in the atmosphere. This technique can probe both the transmission and emission spectra, by observing the planet's primary and secondary transits respectively. Technical details about this method are discussed at length in Chapter 2.

In recent years, several tens of exoplanet spectra have been observed using the transit spectroscopy technique. The bulk of these observations have been obtained with HST (e.g. Kreidberg et al., 2014; Sing et al., 2016; Tsiaras et al., 2016), Spitzer Space Telescope (e.g. Stevenson et al., 2010; Knutson et al., 2011; Deming et al., 2011; Todorov et al., 2013), and other ground based facilities (Waldmann et al., 2012; Bean et al., 2013; Zellem et al., 2014). These observations have led to the detection of a few ionic, atomic and molecular species, and to the constraining of the planet's thermal structure.

Although the current picture is limited by the restricted wavelength range offered by current instruments, it is clear that hot Jupiters are dominated by the signature of water

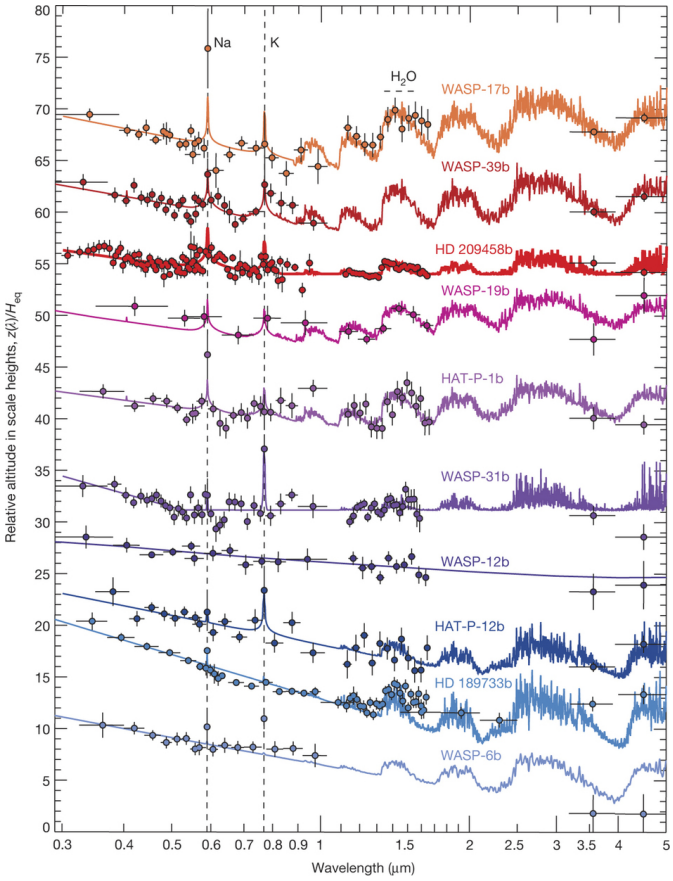

Figure 1.12: HST/Spitzer transmission spectra of ten hot-Jupiters (Sing et al., 2016). Solid coloured lines show atmospheric models. The spectra have been offset for clarity. Planets with clear atmospheres are shown at the top, and exhibit strong alkali and H_2O absorption. Hazy and cloudy planets are shown at the bottom, and exhibit strong optical scattering slope, but no alkali lines or H_2O absorption.

vapour (e.g. Iyer et al., 2016). Tentative detections of other molecules such as carbon monoxide and methane have been reported (Swain et al., 2008; Desert et al., 2009; Swain et al., 2009; Tinetti et al., 2010), but it is only with higher signal-to-noise and broader wavelength range spectra that the signature of these molecules can be confirmed with confidence. Interestingly, recent findings show that there is a continuous from clear to cloudy hot-Jupiter exoplanets, suggesting no evidence of water depletion in the protoplanetary disk at the planet's formation location (Sing et al., 2016, see Figure 1.12). The low-amplitude features observed in many hot Jupiters can in fact be confidently attributed to clouds rather than to severe depletion of atmospheric H_2O relative to solar values. In the latter case, the water depletion would be a result of planet formation, as water vapour is expected to be well mixed in hot-Jupiter atmospheres, such that the measured abundances would be consistent with primordial values. As depletion of H_2O can occur only beyond the snow line (where

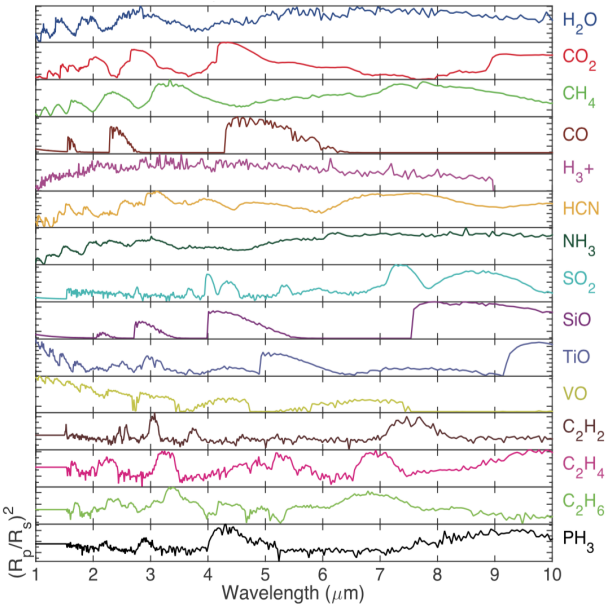

Figure 1.13: Molecular signatures in the 1–10 μm range at the spectral resolving power of $R = 100$ (courtesy of I. Waldmann).

water is found as solid ice), a hot-Jupiter with less-than-expected H_2O would imply that the planet formed beyond the snow line (i.e. at large orbital distances), and avoided accretion of icy planetesimals (Sing et al., 2016). Such scenario has been proposed for Jupiter (Lodders, 2004; Mousis et al., 2012).

The available data also show that clouds and hazes are present in hot Jupiters and warm Neptunes. High altitude clouds have been suggested to explain the weaker-than-expected absorption features of water of several hot Jupiters (Sing et al., 2016), and the flatness of the transmission spectra of warm Neptunes, such as GJ 436b and GJ 3470b (Stevenson et al., 2010; Knutson et al., 2011; Fukui et al., 2013; Ehrenreich et al., 2014). The data available for the only super-Earth extensively studied, GJ 1214b, suggests the presence of either a metal-rich or cloudy atmosphere (Bean et al., 2010; Berta et al., 2012; Kreidberg et al., 2014).

Despite the early successes of this technique, it is clear that we are currently limited by the wavelength coverage and signal-to-noise of these observations. Most of these spectra have been obtained using the Wide Field Camera 3 (WFC3) onboard the HST, covering the spectral range 1.1–1.7 μm. Although some molecules have some strong features in this range (such as H_2O or HCN, or metal oxides such as TiO and VO, see Figure 1.13), it is at longer wavelengths that most molecular roto-vibrational transitions occur. The Spitzer

Space Telescope has given some insight in this long-wavelength regime, but the data are mostly represented by photometric measurements. Significant advances of this field can therefore only occur with higher quality, broader wavelength spectroscopic observations, such as those expected from the James Webb Space Telescope (JWST) (Beichman et al., 2014; Cowan et al., 2015; Batalha et al., 2015; Greene et al., 2016; Barstow et al., 2015; Barstow & Irwin, 2016). One of the four major goals of this NASA/ESA major general-purpose observatory, scheduled for launch in October 2018, is in fact to study and characterise exoplanetary atmospheres using the transit method[11].

However, it is with a dedicated mission, targeting hundreds of exoplanets, that the physical understanding of exoplanetary atmospheres can progress swiftly. The dedicated mission EChO (Tinetti et al., 2015) was a space telescope proposed as part of the Cosmic Vision roadmap of ESA, which competed with four other missions for the M3 call in the program, but was eventually rejected in favour of PLATO 2.0[12]. From the experience gained with EChO, a second dedicated mission called ARIEL, with similar objectives to EChO, was proposed. ARIEL has been selected together with two other missions for the M4 call, and final selection will happen in 2017.

1.3.2 Rocky planetary bodies around metal-polluted white dwarfs

The chemical characterisation of exoplanets is mostly dominated by the study of their atmospheric spectra. Such spectra, however, provide little to no information about the internal composition of the planet. Although some knowledge about the bulk composition can be inferred from the planetary mass and radius, a detailed picture of their internal chemistry is beyond the capabilities of common methods.

In this scenario, white dwarfs represent a unique tool to infer the detailed composition of terrestrial exoplanets. The metal pollution that is often seen at these stars (Zuckerman et al., 2003, see also Figure 1.14) is believed to come from the external accretion of disrupted rocky planetesimals, once part of ancient planetary systems. Due to the high surface gravity and negligible radiative forces, heavy elements sink on relatively short timescales within the atmospheres of cool white dwarfs if compared to the evolution timescales (Paquette et al., 1986; Fontaine & Michaud, 1979). Therefore, the metals seen in their atmospheres effectively mirror the composition of the accreted material. By studying the metal contamination of such stars, it is therefore possible to infer the chemical composition of the accreted

[11] http://jwst.nasa.gov/science.html
[12] http://sci.esa.int/plato/53707-esa-selects-planet-hunting-plato-mission/

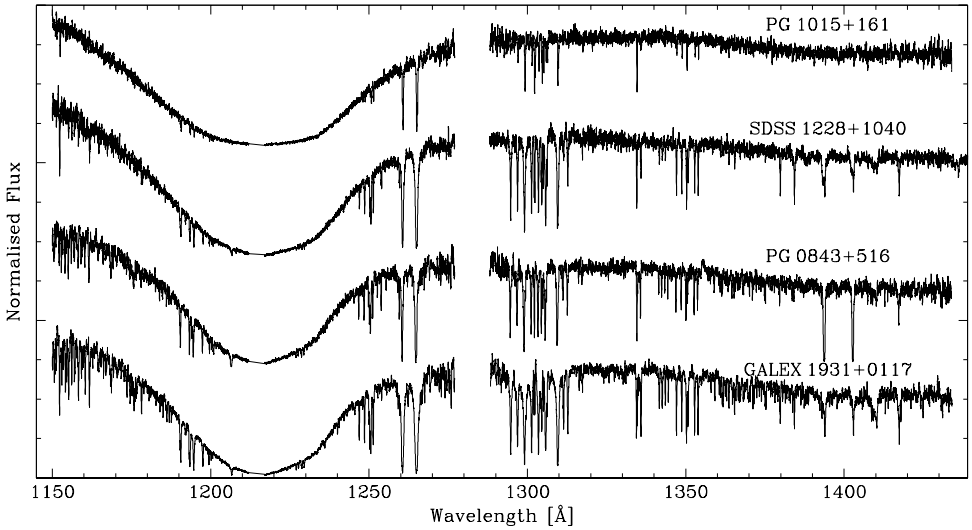

Figure 1.14: HST/COS spectra of four DA white dwarfs known to have circumstellar debris. The large absorption line centred at 1215 Å is the Lyman alpha absorption coming from the white dwarf atmosphere. The additional absorption lines come from the accreted circumstellar debris, and correspond to C, O, Al, Si, P, S, Cr, Fe and Ni. (Gänsicke et al., 2012).

material, which originates from disrupted rocky terrestrial planets.

In the last chapter of this thesis I will discuss this method in more detail, and present a recent Spitzer survey aimed at determining the fraction of white dwarfs with circumstellar debris.

1.4 Thesis outline

In **Chapter 2** I introduce the use of retrieval techniques to interpret the observed spectra of exoplanetary atmospheres, with particular attention given to the use of Bayesian methods. I also present TauREx, a novel spectral retrieval algorithm: I discuss the implemented forward models, and show a benchmark of these models with other equivalent codes published in the literature.

Central to the forward models included in TauREx are the absorption cross sections of different species contained in the atmosphere. In **Chapter 3**, I discuss how absorption cross sections are generated, and I present a new state-of-the-art database of molecular cross sections for the study of H/He dominated atmospheres, created from the most complete line-list databases and up-to-date line broadening coefficients available today. The uncertainties in the generation of these cross sections, and their effects on the final transmission and emission model spectra, is then discussed in **Chapter 4**.

In **Chapter 5**, I investigate the retrievability of the carbon-to-oxygen ratio in exoplanetary atmospheres, and explore the biases that common approximations in atmospheric retrievals lead to. In particular, I study how the use of isothermal profiles and constant-with-altitude abundance profiles in the forward models affect the interpretation of more complex atmospheres.

Chapter 6 presents the application of TauREx to the observed spectra of two exoplanets: a hot Jupiter (HD 209458b) and a super-Earth (55 Cancri e). I show that the spectrum of HD 209458b contains water vapour and possibly clouds, while the modulation seen in the spectrum of 55 Cancri e is likely due to a hydrogen-dominated atmosphere. This latter result represents the first likely detection of an atmosphere around a super-Earth.

In **Chapter 7**, I present a study on the effects that stellar flares have on the atmospheric composition of exoplanets and on their spectra.

Finally, in **Chapter 8**, I shift the focus from the spectroscopy of planetary atmospheres, to a different technique used to chemically characterise planetary bodies: the study of metal-polluted white dwarfs. In particular, I will present the first unbiased survey aimed at determining the fraction of detectable planetary debris at relatively young white dwarfs.

Chapter 2

Spectral retrieval of exoplanetary atmospheres

After more than 20 years of detections, and with over three thousand exoplanets discovered, it has now become clear that we need to shift our attention from the discovery to the characterisation of extrasolar planets. Amongst the various techniques to characterise exoplanets discussed in the introduction, the study of their spectra certainly represents one of the most promising ways to increase our understanding of planetary systems. Atmospheric retrieval techniques are nowadays commonly used to interpret the observed data, and with the imminent launch of JWST and of possible dedicated missions, which will provide spectra in a much broader wavelength range and with higher signal-to-noise, this technique will be even more fundamental to characterise exoplanetary atmospheres.

In this chapter I will discuss the use of retrieval methods to interpret the transmission and emission spectra of exoplanets. In particular, I will present TauREx, a novel inverse retrieval code for exoplanetary atmospheres.

2.1 Introduction to retrieval theory

Remote or indirect measurements are today widely used to acquire information about physical phenomena without making direct contact with the object. Remote sensing is used in a variety of fields, including geography, meteorology, medical sciences and earth sciences, and it is often the most efficient – if not the only – way to study certain phenomena, which would be otherwise unaccessible from direct measurements. When a remote measurement is made, the quantity measured is usually a complicated function of the parameter that is actually required. The inversion of this complicated function, and the estimation of the underlying parameters, can then lead to a wealth of information about the phenomenon being

investigated. This problem is commonly known as *retrieval* or *inverse problem*.

Indirect measurements have often many advantages compared to direct sounding. For instance, medical tomography, in which virtual slices of specific areas of the human body are obtained using a X-ray scanner, has revolutionised the way in which doctors reach a correct diagnosis, thanks to its non-invasive nature. In the field of astrophysics, remote sounding finds an incredible large number of applications, as most of the times the phenomenon or the object being investigated is too far away to be reached by in-situ probes. In the context of atmospheric studies, remote sounding has been essential to increase our understanding of the Earth's atmosphere, of the solar system planets' atmospheres, and, today, of exoplanetary atmospheres. Single measurements obtained with space or ground-based telescopes can in fact provide significant insight about the thermal or compositional structure of a planetary atmosphere.

The retrieval of atmospheric properties from remote sounding has been initially pioneered for Earth studies. With the development of the first meteorological satellites, methods to determine the state of the Earth's atmosphere from space were pioneered in the late 1950s (Kaplan, 1959; King, 1956; Greenfield & Kellogg, 1960; Wark, 1961; Yamamoto, 1961), and were further developed with the advent of dedicated sounders (Chahine, 1968; Wark & Hilleary, 1969; Wark, 1970; Hanel & Conrath, 1970; Conrath et al., 1970; Conrath, 1972). Indirect measurements have also been widely used to study solar system planets (e.g. Larson, 1980; Combes & Encrenaz, 1979; de Bergh et al., 1990; Jakosky & Haberle, 1992; Lellouch et al., 2000) and their rings and satellites (e.g. Smith et al., 1979; Clark & McCord, 1980b,a). With the discovery of the first extrasolar planets in the mid-1990s, part of the attention has begun to shift towards the remote sounding of exoplanetary atmospheres, and the same retrieval techniques developed for the solar system planets are today applied to interpret exoplanetary spectra (e.g. Terrile et al., 2005; Irwin et al., 2008; Madhusudhan & Seager, 2009; Benneke & Seager, 2012, 2013; Line et al., 2013; Waldmann et al., 2015b).

2.1.1 Technical details

The *retrieval problem*, also known as the *inverse problem*, consists in determining the state vector of a model given a data vector. In this case, the state vector can be a set of temperatures and molecular abundances at different altitudes in the atmosphere, and the data vector an emission or transmission spectrum of said atmosphere.

Here I consider a set of measurements described by a vector, $\mathbf{y}$, the measurement or

data vector, and a set of physical parameters describing the state of the atmosphere represented by a state vector, $\mathbf{x}$. These two vectors can be related by:

$$\mathbf{y} = \mathcal{M}(\mathbf{x}) + \varepsilon, \tag{2.1.1}$$

where ε is the error vector and $\mathcal{M}(\mathbf{x})$ is defined as the *forward model*. The forward model is a function representing our understanding of the physics of the measurements. As such, it is often an approximation of the real underlying physics.

Assuming that the relationship between $\mathbf{x}$ and $\mathbf{y}$ is linear, and that $\mathcal{M}(\mathbf{x})$ is linear within the error bounds in the retrieval, we can write:

$$\mathbf{y} - \mathcal{M}(\mathbf{x}_0) = \frac{\delta \mathcal{M}(x)}{\delta \mathbf{x}}(\mathbf{x} - \mathbf{x}_0) + \varepsilon = \mathbf{J}(\mathbf{x} - \mathbf{x}_0) + \varepsilon, \tag{2.1.2}$$

where $\mathbf{J}$ is *weighting function matrix*, or *Jacobian matrix*. The vector $\mathbf{x}_0$ is the *a priori* state vector.

However, it is often hard to find a linear relationship between the state and data vectors, as linearisation depends on underlying processes that are usually unknown. Linearisation can only be readily obtained if the atmospheric state and underlying physics are well known a priori, and in the presence of extremely high signal to noise data. This is sometimes possible in the presence of atmospheric observations of solar system planets, where probes can obtain in-situ measurements. But in the presence of low resolution, lower signal-to-noise spectra of exoplanets, such formalism is not valid. In this case, *Bayesian inference* is a better statistical tool allowing us to constraint the atmospheric state by taking into account the large uncertainties of the measurements, while also enabling us to impose very little prior knowledge on the atmospheric state. This technique is described in the following sections.

2.1.2 Bayesian inference

The Bayesian method provides a consistent approach to parameter estimation. The Bayesian approach begins with the construction of a model that is believed to be adequate to describe the physical phenomenon under study. Then, a "prior distribution" is formulated, which should capture our understanding and knowledge of the system before the measurement is actually made. Lastly, after observing the data, Bayes theorem is used to obtain the posterior distributions of the unknown underlying parameters, which take into account both

the prior and the measurements. From these posteriors, the probability distributions of the model parameters can be evaluated. In the following, I will give a brief technical overview of Bayesian inference and its applications in the context of retrieval theory. For a more in-depth discussion of Bayesian statistics, I refer the interested reader to the relevant books (e.g. Gelman et al., 2013; Bolstad, 2007). See also Trotta (2008) for a review of Bayesian methods in cosmology and astrophysics.

Bayes theorem tells us how the probability density function of the data vector, $\mathbf{y}$, relates to the probability density function of the state vector, $\mathbf{x}$, given a forward model, $\mathcal{M}$. It can be written as:

$$\mathcal{P}(\mathbf{x}|\mathbf{y},\mathcal{M}) = \frac{\mathcal{P}(\mathbf{y}|\mathbf{x},\mathcal{M})\mathcal{P}(\mathbf{x},\mathcal{M})}{\mathcal{P}(\mathbf{y}|\mathcal{M})}, \tag{2.1.3}$$

where $\mathcal{P}(\mathbf{x}|\mathbf{y},\mathcal{M})$ is the posterior distribution of the state when the measurement is given, $\mathcal{P}(\mathbf{x},\mathcal{M})$ is the prior distribution representing our knowledge of the state before the measurement is made, and $\mathcal{P}(\mathbf{y}|\mathbf{x},\mathcal{M})$ is the likelihood function. The denominator on the right hand side of Equation 2.1.3 can be obtained by integrating $\mathcal{P}(\mathbf{y}|\mathbf{x},\mathcal{M})\ \mathcal{P}(\mathbf{x},\mathcal{M})$ over all states, and is also called Bayesian evidence ($\mathscr{E}$).

Assuming a normal distribution for the error vector $\boldsymbol{\varepsilon}$, and no correlation between measurements, the likelihood function can be written as

$$\mathcal{P}(\mathbf{y}|\mathbf{x},\mathcal{M}) = \prod_{k=0}^{N} \frac{1}{\varepsilon_k\sqrt{2\pi}} \exp\left[-\frac{(x_k - y_k)^2}{2\varepsilon_k^2}\right]. \tag{2.1.4}$$

Different techniques can be used to iteratively solve Equation 2.1.3. In the study of extrasolar planets' atmospheres common methods include Markov Chain Monte Carlo, Nested Sampling and Optimal Estimation.

2.1.3 Markov Chain Monte Carlo

Markov Chain Monte Carlo (MCMC) techniques are a set of algorithms for sampling from a probability distribution by constructing a *Markov chain* that has the desired distribution as its equilibrium distribution. The state of the chain after a large number of steps is then used as a sample of the desired distribution.

Monte Carlo methods were developed in the 1940s and 1950s (Turing, 1950; Kahn & Harris, 1951; Barricelli, 1954; McKean, 1967; McK, 1966) and played a central role in the simulations required for the development of the hydrogen bomb, part of the Manatthan

Project (Anderson, 1986; Sabelfeld, 2004). They allow us to estimate the value of integrals by drawing random samples from posterior distributions.

Consider the expected value of the state vector $\mathbf{x}$ with probability distribution $f(\mathbf{x})$:

$$\exp(\mathbf{x}) = \int \mathbf{x} f(\mathbf{x}) d\mathbf{x} \tag{2.1.5}$$

It is possible to approximate the unknown integral by producing a reasonable number of random vectors $\mathbf{x}_i$. This technique is known as *Monte Carlo integration*. In practice, Monte Carlo methods allow us to approximate the values of integrals against probability density functions with finite sums over N_{samp} samples:

$$\mathcal{I} = \int h(\mathbf{x}) f(h(\mathbf{x})) dh(\mathbf{x}) \quad \rightarrow \quad \hat{\mathcal{I}} = \frac{1}{N_{\text{samp}}} \sum_{i}^{N_{\text{samp}}} h(\mathbf{x}_i) f(h(\mathbf{x}_i)) \quad \text{(for large } N_{\text{samp}}\text{)}, \tag{2.1.6}$$

where $h(\mathbf{x}_i)$ is a sample drawn from f. In Bayesian inference, we can replace $f(h(\mathbf{x}))$ with a posterior distribution, $\mathcal{P}(\mathbf{x}|\mathbf{y},\mathcal{M})$, and make $h(\mathbf{x})$ a function of the unknown parameter. Then, the resulting expected value is:

$$\exp(h(\mathbf{x})|\mathbf{y}) = \int \mathcal{P}(\mathbf{x}|\mathbf{y},\mathcal{M}) h(\mathbf{x}) d\mathbf{x} \tag{2.1.7}$$

$$\approx \frac{1}{N_{\text{samp}}} \sum_{i=1}^{N_{\text{samp}}} h(\mathbf{x}), \tag{2.1.8}$$

and the sample variance is:

$$\text{var}(\hat{\mathcal{I}}) = \frac{1}{N_{\text{samp}}(N_{\text{samp}}-1)} \sum_{i=1}^{N_{\text{samp}}} (\mathbf{x}_i - \hat{\mathcal{I}})^2. \tag{2.1.9}$$

One of the major problems of the Monte Carlo integration method described above is that we need to know how to generate efficiently the samples $\mathbf{x}_i$. This can be readily achieved using a *Markov chain*.

Markov chains are *stochastic processes* that undergo transitions from one state to another on a state space, in which the probability distribution of the future state depends only on the current state, but not on past states (see e.g. Brooks et al., 2011; Gamerman & Lopes, 2006). This property is called *Markovian property* (Markov, 1954; Feller, 1971), and can be thought as a property of "mild non-independence", as the chain only exhibits a one-step

dependence.

A sequence of random variables $X_0, X_1, \ldots$ taking values in the likelihood space $1, 2, \ldots$ is called a *Markov chain* if, for all $n \geq 0$, we have:

$$\mathcal{P}(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) = \mathcal{P}(X_{n+1} = j | X_n = 1) \tag{2.1.10}$$

where $\mathcal{P}(X_{n+1} = j | X_n = 1)$ is called the *transition probability* or *transition kernel* from state i to state j.

In the context of Markov Chain Monte Carlo (MCMC) techniques, we simulate Markov chains in which the stationary distributions are the distributions that we want to simulate. In other words, a Markov chain is simulated for which some function, such as the joint distribution of the parameter of some model, is the limiting distribution when the Markov chain has run for a sufficiently long time.

Various MCMC algorithms exist in the literature, such as the Metropolis-Hastings algorithm (Metropolis et al., 1953; Hastings, 1970) and Gibbs sampling (Geman & Geman, 1984; Gelfand & Smith, 1990). In the following I only describe the first method.

2.1.3.1 The Metropolis-Hastings algorithm

Let $f(\mathbf{x})$ be the "target distribution", i.e. a function proportional to the desired probability distribution, $\mathbf{x}_j$ be a current value, and $q(\mathbf{x}|\mathbf{x}_j)$ be a "proposal distribution" that may depend on the current value $\mathbf{x}_j$. The algorithm is as follow:

- Sample a proposed value, $\mathbf{x}_*$, from the proposal distribution $q(\mathbf{x}|\mathbf{x}_j)$
- Calculate the acceptance probability:

$$\mathcal{P}(\mathbf{x}_j, \mathbf{x}_*) = \min\left[1, \frac{f(\mathbf{x}_*)}{f(\mathbf{x}_j)} \frac{q(\mathbf{x}_j|x_*)}{q(\mathbf{x}_*|\mathbf{x}_j)}\right] \tag{2.1.11}$$

- If the acceptance probability is unity, accept the sample and set $\mathbf{x}_{j+1} = \mathbf{x}_*$. Otherwise accept the sample with probability $\mathcal{P}(\mathbf{x}_j, \mathbf{x}_*)$, or reject the sample and set $\mathbf{x}_{j+1} = \mathbf{x}_j$

The Metropolis-Hastings algorithm uses the original transition probabilities to propose the next step, then accepts the proposal with a given probability. This algorithm represents an extremely general way to construct a Markov chain with the desired distribution, as no assumptions are made about the starting distributions.

Different improvements of the Metropolis-Hastings algorithm exist in the literature, such as the Delayed-rejection Adaptive-MCMC (DRAM, Haario et al., 2006). Compared to a more classical Metropolis-Hastings sampler, the DRAM algorithm implements a Delayed-rejection algorithm (Mira, 2001) and an adaptive proposal distribution (Adaptive Metropolis Haario et al., 2001). In order to improve the efficiency of the chain, in the Adaptive Metropolis the covariance matrix of the proposal distribution, assumed to be Gaussian, is tuned to the shape and size of the target distribution using the previous chain. The problem of this method is that, as accepted values now depend on the history of the chain, the resulting chain is not Markovian[1]. However, it can be shown that ergodicity[2] still applies (Haario et al., 2006). The Delayed-rejection mechanisms allows for a second proposal attempt to be made which is dependent on the previous chain as well as the previously rejected proposals. This delayed rejection can be repeated for a fixed or random number of stages. When these two methods are combined, we use different proposals using the Delayed-rejection mechanism, and adapt them using the Adaptive Metropolis method. It can be shown that this technique can improve the efficiency of the MCMC algorithm in high dimensional likelihood spaces.

2.1.4 Nested Sampling and MULTINEST

Current MCMC algorithms concentrate on sampling the posterior distributions, and do not calculate the Bayesian evidence (the denominator on right hand side of Equation 2.1.3). In Nested Sampling (Skilling, 2004, 2006; Sivia, 2011) the method is somewhat reversed: the evidence is accurately computed, leaving posterior samples as a byproduct. Amongst the various advantages of this method, the calculation of the evidence enables us to discriminate between different models (see next section).

Standard MCMC methods, such as the Metropolis-Hastings algorithm described above, can lead to problems in sampling efficiently from a multi-modal posterior distributions or one with large and curved degeneracies between parameters (Feroz et al., 2010b). In addition, MCMC usually requires tuning of the proposal distribution to sample efficiently, and convergence can be problematic. The Nested Sampling algorithm and the MULTINEST implementation (Feroz et al., 2009b) aim at solving both problems, while also providing

[1] A stochastic process is *Markovian* if the conditional probability distribution of future states of the process depends only upon the present state, not on past states.

[2] A Markov chain is *ergodic* if there is a positive probability to pass from any state to any other state in one step.

means of calculating the Bayesian evidence, something that requires additional steps when using the standard MCMC method (see e.g. Friel & Wyse, 2012). MULTINEST is optimised for highly degenerate (e.g. curved) and multimodal posteriors, and has found extensive applications in the astrophysical context (Skilling, 2004; Feroz et al., 2009a, 2010a, 2011b,a; Bridges et al., 2009; Graff et al., 2012; White & Feroz, 2010; Karpenka et al., 2013; Strege et al., 2013), including the study of exoplanetary atmospheres (e.g. Benneke & Seager, 2012; Waldmann et al., 2015b). In the following I will briefly describe the Nested Sampling algorithm and the MULTINEST implementation, while I refer the reader to (Skilling, 2006; Feroz & Hobson, 2008; Feroz et al., 2009b) for a more complete discussion.

Let us first remind that the evidence for a set of parameters $\mathbf{x}$ is:

$$\mathscr{E} = \int \mathcal{P}(\mathbf{y}|\mathbf{x},\mathcal{M})\mathcal{P}(\mathbf{x},\mathcal{M})d\mathbf{x}, \tag{2.1.12}$$

where $\mathscr{E} = \mathcal{P}(\mathbf{y}|\mathcal{M})$. In Nested Sampling, this multidimensional integral is transformed into a one-dimensional integral, exploiting the relation between the likelihood and prior volume. Following Feroz & Hobson (2008), the prior volume is defined by

$$\mathscr{X}(\xi) = \int_{\mathcal{P}(\mathbf{y}|\mathbf{x},\mathcal{M})>\xi} \mathcal{P}(\mathbf{x},\mathcal{M})d\mathbf{x}, \tag{2.1.13}$$

where the integral extends over all regions for which the likelihood function is contained within the iso-likelihood contour $\mathcal{P}(\mathbf{y}|\mathbf{x},\mathcal{M}) = \xi$. Assuming that the inverse of this equation is monotonically decreasing function of $\mathscr{X}$, Equation 2.1.12 (the evidence) can be written as:

$$\mathcal{P}(\mathbf{y}|\mathcal{M}) = \int_0^1 \mathcal{P}(\mathscr{X})d\mathscr{X} \tag{2.1.14}$$

$$= \mathscr{E}. \tag{2.1.15}$$

The evidence can then be approximated numerically using standard quadrature methods. If the likelihood values $\mathcal{P}(\mathscr{X}_j)$ can be evaluated at a sequence of decreasing values $0 < X_M < \cdots < X_2 < X_1 < X_0 = 1$, we have:

$$\mathcal{P}(\mathbf{y}|\mathcal{M}) = \sum_{i=1}^{M} \mathcal{P}(\mathscr{X}_i)w_i, \tag{2.1.16}$$

where the weights w_i are obtained using the simple trapezium rule:

$$w_i = \frac{1}{2}(X_{i-1} - X_{i+1}). \tag{2.1.17}$$

The summation in Equation 2.1.16 is performed by drawing $N_{\rm live}$ "live" random samples from the full prior $\mathcal{P}(\mathbf{x}, \mathcal{M})$, so that the initial prior volume $\mathscr{X}_0$ is unity. These samples are then sorted in order of their likelihood value and the smallest likelihood, $\mathcal{P}(\mathscr{X}_0)$, is removed from the live set. This sample, corresponding to the case with the worst fit to the data, is then replaced by a point drawn from the prior distribution, provided that its likelihood is larger than $\mathcal{P}(\mathscr{X}_0)$. The new prior volume contained within this iso-likelihood contour will be a random variable given by $X_1 = t_1 X_0$, where t_1 (i.e. the "shrinkage" ratio) has a distribution $\mathcal{P}(t) = N_{\rm live} t^{N_{\rm live}-1}$. This is the probability distribution for the largest of $N_{\rm live}$ samples drawn uniformly from the interval [0, 1], and represents the shrinkage in prior volume between consecutive likelihood contours.

At each subsequent iteration i, the procedure is repeated: the point with lowest likelihood $\mathcal{P}(\mathscr{X}_i)$ in the live set is excluded and replaced with a sample with $\mathcal{P}(\mathscr{X}) > \mathcal{P}(\mathscr{X}_i)$, and the corresponding prior volume $\mathscr{X}_i = t_i \mathscr{X}_{i-1}$ is further shrunk. As the prior volume reduces in size, the algorithm travels through "nested shells" of iso-likelihood contours, identifying the regions of highest likelihoods. The algorithm then stops when the product of the remaining prior volume and maximum-likelihood value does not change by less than a specified tolerance (usually 0.1 in log-evidence).

Posteriors can be easily generated once the evidence is computed, using the full set of discarded points (i.e. the points with the lowest-likelihood value at each iteration i). Each of these points is simply assigned a weight:

$$w_{\text{discarded},i} = \frac{\mathcal{P}(\mathbf{y}|\mathbf{x}, \mathcal{M}) w_i}{\mathcal{P}(\mathbf{y}|\mathcal{M})}. \tag{2.1.18}$$

These samples can then be used to infer the mean, standard deviations, covariances etc., of the retrieved parameters or to compute marginalised posterior distributions.

One of the major limitation of this approach is the efficient draw of random samples with the constraint $\mathcal{P}(\mathscr{X}) > \mathcal{P}(\mathscr{X}_i)$. As the algorithm progresses, if one repeatedly draws random samples from the prior distribution, the probability of randomly getting one that is within the iso-likelihood contour progressively decreases, and the efficiency of the method

decreases exponentially as $\mathcal{P}(\mathscr{X}_i)$ increases. Several improvements to this method have been developed (Mukherjee et al., 2006; Shaw et al., 2007). TauREx implements the Multimodal Nested Sampling (MULTINEST), as described in Feroz et al. (2009b).

In the MULTINEST algorithm, at each iteration the set of live points is redistributed into an optimised number of clusters, encompassed by overlapping ellipsoids. These clusters are constructed such that the union of all ellipsoids includes all samples. A higher acceptance rate for the new random samples is obtained by drawing the new samples from within the ellipsoids, as these points have a high likelihood with higher probability. As the algorithm progresses, the live points become clustered in regions of high likelihood encompassed by ellipsoids.

2.1.5 Model selection

Bayesian model comparison offers a way to select the best model, i.e. the one that has the best balance between quality of the fit and model complexity, in the presence of several competing models. Applying the principle of Occam's razor, added complexity in a model should be avoided whenever a simpler model gives an adequate fit to the observations. In the Bayesian framework, a more complicated model will have a higher evidence only if it fits the data significantly better than a simpler theory (Liddle, 2004; Trotta, 2007). In the case of fitting spectra of exoplanets, this principle can be adapted to evaluate the detection confidence of a given atmospheric constituent.

The main tool for model comparison in a Bayesian framework is the Bayesian evidence. Its definition (Equation 2.1.3) and evaluation through Nested Sampling techniques have already been shown in the previous section. When comparing two models, $\mathcal{M}_0$ and $\mathcal{M}_1$, we can define the ratio of the models' probabilities, or *Bayes factor*:

$$\mathcal{B}_{21} = \frac{\mathcal{P}(\mathcal{M}_2|\mathbf{y})}{\mathcal{P}(\mathcal{M}_1|\mathbf{y})} = \frac{\mathcal{P}(\mathcal{M}_2)}{\mathcal{P}(\mathcal{M}_1)} \frac{\mathcal{P}(\mathbf{y}|\mathcal{M}_2)}{\mathcal{P}(\mathbf{y}|\mathcal{M}_1)} = \frac{\mathcal{P}(\mathcal{M}_2)}{\mathcal{P}(\mathcal{M}_1)} \frac{\mathscr{E}_2}{\mathscr{E}_1}. \tag{2.1.19}$$

Assuming the model priors to be identical ($\mathcal{P}(\mathcal{M}_2) = \mathcal{P}(\mathcal{M}_1)$), we can reduce equation 2.1.21 to

$$\mathcal{B}_{21} = \frac{\mathscr{E}_2}{\mathscr{E}_1} = \frac{\mathcal{P}(\mathbf{y}|\mathcal{M}_2)}{\mathcal{P}(\mathbf{y}|\mathcal{M}_1)} \tag{2.1.20}$$

which is the ratio of the models' evidences. A value of $\mathcal{B}$ greater (less) of unity represents an increase (decrease) in support in favour of model 2 versus 1. Using the Jeffrey's scale (Jeffreys, 1998), an empirically calibrated scale shown in Table 2.1, it is possible to determine

Table 2.1: Jeffreys' scale and translation to frequentist significance values in favour of a more complex model (Adapted from Trotta, 2008).

$\lvert\ln \mathcal{B}_{21}\rvert$	Probability	"sigma"	Interpretation
< 1.0	< 0.750	$< 2\sigma$	Inconclusive
1.0	0.750	2.1σ	Weak evidence
2.5	0.923	2.7σ	Moderate evidence
5.0	0.993	3.6σ	Strong evidence

qualitatively whether a more complex model is favoured strongly, weakly or inconclusively.

Sellke et al. (2001) provides an expression to translate the Bayes factor to the frequentist measures of confidence (p-value = $p_{\rm val}$ and sigma significance n_σ):

$$\mathcal{B}_{21} \leq -\frac{1}{e \times p_{\rm val} \times \ln(p_{\rm val})}, \tag{2.1.21}$$

$$p_{\rm val} = 1 - {\rm erf}\left(\frac{n_\sigma}{\sqrt{2}}\right), \tag{2.1.22}$$

where erf is the error function. Values for a representative samples of $\mathcal{B}_{21}$ are given in Table 2.1.

In the context of atmospheric retrievals, model selection can be used to estimate the detection significance of a given molecular constituent in a Bayesian framework. In this case, multiple retrievals including (and excluding) a range of molecules are run, and their Bayesian evidence is compared to determine the best model. The Bayes factor between these models can then be used to estimate the preference for adding, or removing, a given molecule, following the same scheme given in Table 2.1.

2.2 Spectral retrieval of exoplanetary atmospheres

The study of exoplanetary atmospheres is a very young field, which has shown rapid changes over the past decade. In the early days of exoplanet spectroscopy, determination of the compositions and temperatures were commonly obtained using self-consistent, ab-initio one-dimensional models (Burrows et al., 2005, 2007; Fortney et al., 2005). These models have been progressively improved over the past decade, and today there are several 3D general circulation models trying to model the atmospheric dynamics and compositions (Showman & Guillot, 2002; Showman et al., 2008), while also explaining the observed spectra (Cooper & Showman, 2005; Showman et al., 2009; Parmentier et al., 2013). However, due to the complexity of the models, this "ab-initio approach" can only explore a few

potential solutions, and it is therefore impossible to characterise the uncertainties and the degeneracies of the model parameters. Moreover, self-consistent models usually assume some constraints on the chemistry, such as chemical equilibrium, that might not fully represent the true state of the atmosphere. In this case, the results and conclusions are potentially biased by the assumed model, instead of being driven by the observations.

A different approach, the "retrieval approach" or "data-driven approach", is to use simpler but less constraining models, and use the observations to identify the possible range of solutions, exploring a much larger parameter space. Such method is not, however, bias free, as the numerous approximations that needs to be done to speed up the computation of spectra do not fully grasp the complexity of real atmospheres. Nevertheless, this approach is today very important for the interpretation of exoplanetary spectra.

Spectral retrieval theory, which I briefly discussed in the previous sections, was originally pioneered in solar system studies (e.g. Adler-Golden et al., 1998; Snell et al., 1995), but quickly found applications when the first exoplanet spectra were obtained (e.g. Terrile et al., 2006; Irwin et al., 2008; Lee et al., 2011). However, the first attempts to use retrieval techniques on exoplanet spectra were limited by the lack of large-scale computing power, and could not make full use of the Bayesian methodologies described above. It is thanks to the increasing availability of cheap computing power that the calculation of a large number of models, needed in a Bayesian framework to explore the parameter space, could be obtained. Nowadays, the retrieval of exoplanetary atmospheres is obtained using the MCMC or Nested Sampling approaches (e.g. Madhusudhan et al., 2011b; Benneke & Seager, 2012; Line et al., 2013; Benneke & Seager, 2013; Waldmann et al., 2015b; Line & Parmentier, 2016). Such tools are in fact ideal for the exploration of highly degenerate parameter spaces, such as those used to model exoplanet spectra.

In what follows, I will present TauREx, a retrieval code for exoplanetary transmission and emission spectra, which I developed together with Ingo Waldmann from mid-2014. A major part of the results presented in this thesis have been obtained thanks to this novel algorithm, which has already found several applications in the context of exoplanetary atmosphere studies (Tsiaras et al., 2015, 2016; Venot et al., 2016; Rocchetto et al., 2016).

2.2.1 TauREx: A novel retrieval code for exoplanetary atmospheres

TauREx (Tau Retrieval for Exoplanets) is a radiative transfer fully Bayesian retrieval framework that can be used to model and interpret exoplanetary atmosphere spectra. Its first def-

inition and implementation were presented in Waldmann et al. (2015a,b), while subsequent optimisations and improvements were included in a second major release, presented in this thesis.

TauREx has been developed from scratch in object oriented Python and C++, and it is highly scalable, thanks to its modular structure. The main features of TauREx are:

- highly optimised transmission and emission forward models, allowing the computation of several thousands of low resolution spectra in a short timeframe, using both the cross section sampling method and the *k*-distribution technique.
- the ability to run line-by-line forward models.
- the use of MCMC and Nested Sampling techniques to sample the Bayesian likelihood space.
- the ability to include a large number of parameters in the retrieval model thanks to optimal code parallelisation and scalability to cluster computing.
- iterative Bayesian parameter and model selection using the global Bayesian evidence.
- the use of state-of-the-art molecular opacity cross sections.

Non parametric prior constraints based on pattern recognition software, and prior selection through the use of a deep-believe neural network, are also included in TauREx. These modules are described in Waldmann et al. 2015a and Waldmann 2016 and I refer the interested reader to these publications.

In the following sections I will describe the forward models implemented in TauREx for emission and transmission spectroscopy. Then, I discuss how the atmosphere is parametrised, what are the different approaches used to model the atmospheric temperature profile, and how the different opacity sources are implemented.

2.3 Forward models

Transiting exoplanets offer a unique opportunity to observe transmission and emission spectra of exoplanets, thanks to the particular geometry of a transit.

The radiative transfer models included in TauREx are one dimensional and assume a plane-parallel atmosphere. They take into account molecular absorption, Rayleigh scattering, and collision-induced absorption, but ignore other effects such as larger particle scattering (like haze and cloud particles), forward-scattering, Raman scattering and polarisation

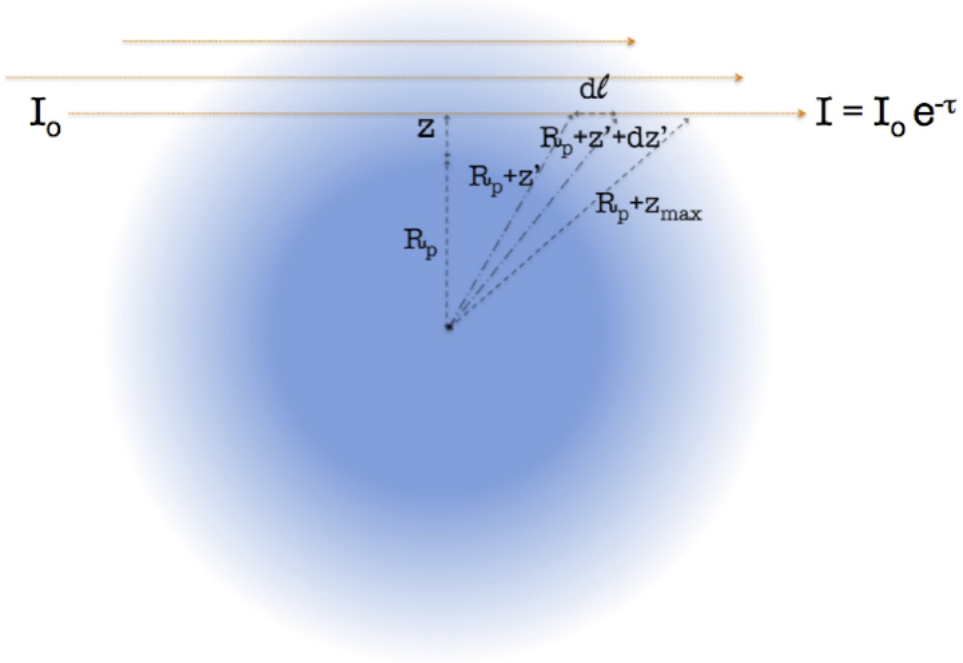

Figure 2.1: Geometry of a primary transit observation, illustrating the paths of the stellar photons filtered through the planetary atmosphere (Tinetti et al., 2012)

effects. These ignored effects and simplifications can potentially introduce significant differences and biases (e.g. de Kok et al., 2011; de Kok & Stam, 2012; Line & Parmentier, 2016), and future work will therefore be needed to address these issues in detail.

In the following sections I will describe the transmission and emission forward models implemented in TauREx.

2.3.1 Transmission spectroscopy

When some radiation goes through a medium along a path ds, its intensity I_λ is weakened by absorption and scattering by the material by an amount equal to

$$dI_\lambda = -I_\lambda \sigma_\lambda \rho_N ds \tag{2.3.1}$$

where ρ_N is the number density of the medium and σ_λ is the absorption or extinction cross section in $\mathrm{m}^2\mathrm{mol}^{-1}$ for radiation of wavelength λ. This can be rearranged to give

$$\frac{dI_\lambda}{I_\lambda} = -\sigma_\lambda \rho_N ds. \tag{2.3.2}$$

We then define the optical path length o, and the optical depth τ_λ at wavelength λ:

$$o = \int \rho_N ds \qquad \text{and} \qquad \tau_\lambda = \sigma_\lambda \int \rho_N ds = \sigma_\lambda o, \tag{2.3.3}$$

respectively. It follows that the intensity at wavelength λ at a point at altitude z in the atmosphere is given by integrating along the vertical path from the top of the atmosphere:

$$I_\lambda(z) = I_\lambda(0) \exp\left(- \int_z^{z_\infty} \sigma_\lambda(z') \rho_N(z') dz' \right) \tag{2.3.4}$$

which is the monochromatic intensity at altitude z, as given by the Beer-Bouguer-Lambert Law:

$$I_\lambda(z) = I_\lambda(0) e^{-\tau_\lambda(z)}. \tag{2.3.5}$$

In the case of transmission through the terminator region of the planetary atmosphere, the optical path of the stellar photons is shown in Figure 2.1. The optical depth at an altitude z along a path is given by

$$\tau_\lambda(z) = \sum_{i=1}^{N_{\text{gas}}} \left(\int_0^{l(z)} \sigma_{i,\lambda}(z') \chi_i(z') \rho_N(z') dl \right) \tag{2.3.6}$$

where N_{gas} is the total number of molecules in the atmosphere, χ_i is the mixing ratio of molecule i, $\sigma_{i,\lambda}$ is the absorption cross section of molecule i at wavelength λ, and ρ_N is the number density in m^{-3}. Note that the path l is parallel to the line of sight, and that the path element dl is given by (see Figure 2.1):

$$dl = 2 \left(\sqrt{(R_p + z + dz')^2 - (R_p + z')^2} - \sqrt{(R_p + z)^2 - (R_p + z')^2} \right) \tag{2.3.7}$$

where R_p (the planet radius), z, z' and dz are defined in Figure 2.1.

Equation 2.3.6 can be finally converted to an equivalent area, $A(\lambda)$, by integrating along all the viewing paths

$$A(\lambda) = 2 \int_0^{z_{\max}} (R_p + z) \left(1 - e^{-\tau(\lambda, z)} \right) dz. \tag{2.3.8}$$

If we assume a stratified atmosphere with N_{layers} layers and N_{gas} molecules, we can replace

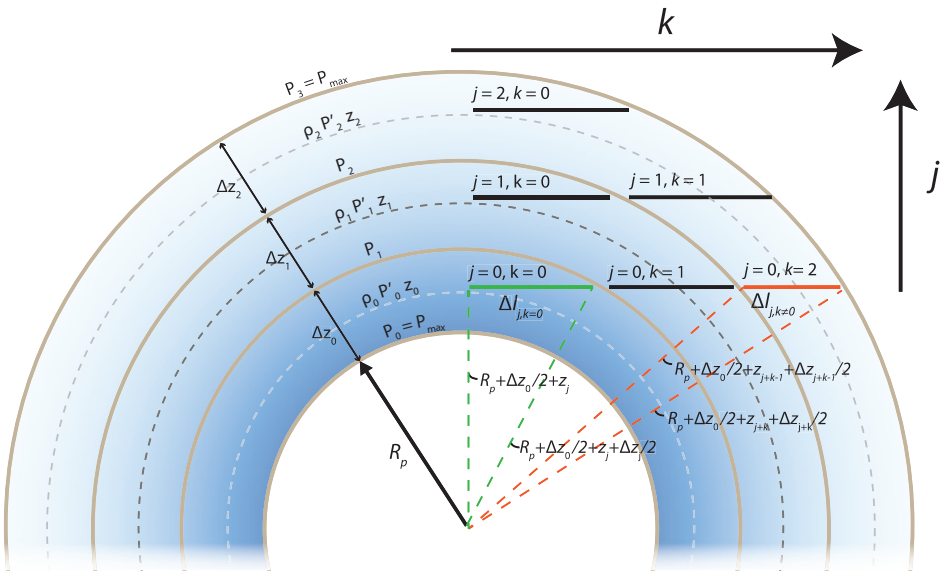

Figure 2.2: Geometry of a primary transit observation for the numerical integration shown in Equation 2.3.9 and 2.3.10, showing how the path lengths (Δl) are calculated. The solid circles show the pressure boundaries of each layer i (P_i and P_{i+1}), while the dashed lines shows the mid-layer where the density (ρ_{Ni}), pressure (P'_i) and altitude (z_i) of layer i are evaluated. The pressure P'_i of the mid layer is calculated by taking the log-average of the pressure boundaries (P_i and P_{i+1}). $P_{\rm max}$ and $P_{\rm min}$ are the maximum and minimum pressure, while Δz_i is the change in altitude at layer i. See also section 2.6 for additional details on this parmetrisation.

this integral over z and l with this expression:

$$A(\lambda) \approx 2 \sum_{j=0}^{N_{\rm layers}} \left((R_p + z_j) \left\{ 1 - \exp\left[- \sum_{k=0}^{N_{\rm layers}-j} \left(\sum_{t}^{N_{\rm gas}} \sigma_{t,j+k}(\lambda) \cdot \chi_{t,j+k} \cdot \rho_{N\,j+k} \cdot 2\Delta l_{j+k} \right) \right] \right\} \Delta z_j \right), \tag{2.3.9}$$

where the path length $\Delta_{lk,j}$ is given by

$$\begin{aligned} \Delta l_{j,k\neq 0} = & \sqrt{\left(R_p + \frac{\Delta z_0}{2} + z_{j+k} + \frac{\Delta z_{j+k}}{2} \right)^2 - \left(R_p + \frac{\Delta z_0}{2} + z_j \right)^2} - \\ & - \sqrt{\left(R_p + \frac{\Delta z_0}{2} + z_{j+k-1} + \frac{\Delta z_{j+k-1}}{2} \right)^2 - \left(R_p + \frac{\Delta z_0}{2} + z_j \right)^2}, \\ \Delta l_{j,k=0} = & \sqrt{\left(R_p + \frac{\Delta z_0}{2} + z_j + \frac{\Delta z_j}{2} \right)^2 - \left(R_p + \frac{\Delta z_0}{2} + z_j \right)^2}, \end{aligned} \tag{2.3.10}$$

and the change in altitude Δz_i is given by:

$$\Delta z_i = -H_i \cdot \log\left(P_{i+1}/P_i\right) \tag{2.3.11}$$

where H_i is the scale height of layer i, and P_i and P_{i+1} are its pressure boundaries. Note also that each path length Δl is multiplied by two in Equation 2.3.9 as the geometry is symmetric. The different variables used in Equation 2.3.9 and 2.3.10 are shown schematically in Figure 2.2.

Rayleigh scattering and collision-induced absorption for H_2–H_2 and H_2–He pairs are also taken into account. These two sources of opacity are simply included in the computation of the optical depth for each path length, using pre-computed cross sections and the relative abundances of each species. Collision induced absorption cross sections are taken from HITRAN (Richard et al., 2012), and from Borysow et al. (2001) and Borysow (2002), while the Rayleigh scattering cross sections are pre-computed using the following expression (Van de Hulst & Twersky, 1957; McCartney, 1976; Bodhaine et al., 1999):

$$\sigma_i(\lambda) = \frac{24\pi^3(n_{R_i}^2 - 1)^2 F_{\mathrm{King}i,\lambda}}{{N_{\mathrm{air}}}^2(n_{R_i}^2 + 2)^2} \times \frac{1}{\lambda^4} \tag{2.3.12}$$

where λ is the wavelength, n_{Ri} is the wavelength dependent refractive index of the i-th species, N_{air} is the molecular number density for standard air (otherwise known as the Loschmidt constant, $N_{\mathrm{air}} = 2.6867805 \times 10^{25}$ m^{-3}), $F_{\mathrm{King}i,\lambda}$ is the King's factor, or depolarisation term, which describes the effect of molecular anisotropy. The refractive indexes and King's factor are taken from a variety of sources (Peck & Khanna, 1966; Mansfield & Peck, 1969; Peck & Reeder, 1972; Bideau-Mehu et al., 1973; Bates, 1984; Zhang et al., 2008; Sneep & Ubachs, 2005; Cox, 2015).

Note that larger particle scattering, described by Mie theory (Van de Hulst, 1957; De Rooij & van der Stap, 1984), is not taken into account; future work to include this source of absorption is therefore needed.

When the equivalent area $A(\lambda)$ is computed, the total transit depth as a function of wavelength is simply given by

$$D(\lambda) = \frac{{R_p}^2 + A(\lambda)}{{R_*}^2} \tag{2.3.13}$$

where R_p and R_* are the planet and star radius respectively.

2.3.2 Emission spectroscopy

A vast literature spanning several decades exists about radiative transfer through the atmosphere (e.g. Chandrasekar, 1960; Goody & Yung, 1995; Liou, K N, 2002). Here I will only

give an overview of the emission forward model included in TauREx, also discussed in Waldmann et al. (2015a).

In the case of non-scattering atmospheres in local thermodynamic equilibrium, the equations describing how thermal radiation diffuses in the atmosphere is given by the Schwartzschild equation:

$$\mu \frac{dI_\lambda(\tau,\mu)}{d\tau} = I_\lambda(\tau,\mu) - B_\lambda(T), \tag{2.3.14}$$

where I_λ is the intensity per wavelength, λ, $B_\lambda(T)$ is the Planck function at temperature T, $\mu = \cos\phi$ is the upwards inclination, τ is the overall optical depth, given as function of altitude (z) by

$$\tau_\lambda(z) = \sum_{i=1}^{N_{\text{gas}}} \tau_{\lambda,i}(z), \tag{2.3.15}$$

where $\tau_{\lambda,i}$ denotes the optical depth per absorbing species, i, given by

$$\tau_{\lambda,i} = \int_z^{z_\infty} \sigma_{\lambda,i}(z')\chi_i(z')\rho_N(z')dz'. \tag{2.3.16}$$

Here $\sigma_{\lambda,i}$ is the absorption cross section, χ_i the mixing ratio and ρ_N the number density. We can now express the upwards welling radiance as

$$I_\lambda(\tau,\mu) = I_\lambda(\tau_s)e^{-(\tau_s-\tau)/\mu} + \int_\tau^{\tau_s} B_\lambda(T_{\tau'})e^{-(\tau'-\tau)/\mu}\frac{d\tau'}{\mu}, \tag{2.3.17}$$

where the first right-hand-side term is the radiation at the planetary surface (or defined surface pressure for gaseous planets), and the second term denotes the integrated emission contributions for individual plane-parallel layers. The monochromatic transmittance and its derivative (weighting function) can be defined as

$$\Im_\lambda(\tau/\mu) = e^{-\tau/\mu}, \quad \frac{\partial \Im_\lambda(\tau/\mu)}{\partial \tau} = -\frac{1}{\mu}e^{-\tau/\mu}. \tag{2.3.18}$$

Hence, we can express the total integrated radiation at the top of the atmosphere ($\tau = 0$, $z = \infty$) as

$$I_{\lambda,0} = I_\lambda(\tau=0) = B_\lambda(T_{\text{surf}})e^{-\tau_{\text{surf}}/\mu} + \int_{\tau_{\text{surf}}}^0 B_\lambda(T_\tau)\frac{\partial \Im_\lambda(\tau/\mu)}{\partial \tau}d\tau \tag{2.3.19}$$

where τ_s and T_{surf} are the optical depth and temperature at the planetary surface. Note that one has also to integrate over the zenith angle (μ) to obtain the *diffuse transmittance* (see Liou, K N, 2002, pg. 125):

$$\Im^f_\lambda(\tau) = 2\int_0^1 \Im(\tau/\mu)\mu d\mu. \tag{2.3.20}$$

Assuming a stratified atmosphere with N_{layers} layers and molecules, we can replace these integrals with discrete summations:

$$\begin{aligned} I_{\lambda,0} =& 2\pi \sum_j^{N_{\text{layers}}} \sum_g^{N_{\text{quad}}} B(T_{\text{surf}}) \left[e^{\left(-\sum_{k=0}^{N_{\text{layers}}} \sum_t^{N_{\text{gas}}} \sigma_{t,k}\ \rho_{Nk}\ dz_k\right)} \right] \mu_g w_g + \\ &+ 2\pi \sum_j^{N_{\text{layers}}-1} \sum_g^{N_{\text{quad}}} B(T_j) \left[e^{\left(-\sum_{k=j+1}^{N_{\text{layers}}} \sum_t^{N_{\text{gas}}} \sigma_{t,k}\ \rho_{Nk}\ dz_k\right)} - e^{\left(-\sum_{k=j}^{N_{\text{layers}}} \sum_t^{N_{\text{gas}}} \sigma_{t,k}\ \rho_{Nk}\ dz_k\right)} \right] \mu_g w_g, \end{aligned} \tag{2.3.21}$$

where the first line of the equation describes the emission from the surface, while the second line the emission from the rest of the atmosphere. Note here that the integration is over the zenith angle is achieved using the Gaussian quadrature method, with N_{quad} points: μ_g represents the g-th quadrature point, and w_g the corresponding weight. We generally adopt N_{quad} = 4, as it was found to be a good compromise between computational speed and accuracy.

The final exoplanetary emission spectrum is then given by

$$\left(\frac{F_p}{F_*}\right)_\lambda = \frac{I_\lambda(\tau=0)}{I_{*,\lambda}} \left(\frac{R_p}{R_*}\right)^2, \tag{2.3.22}$$

where I_* is the stellar intensity.

2.4 Benchmark of forward models with CHIMERA and NEMESIS

In order to validate the forward models described in Section 2.3, I compared spectra generated with TauREx with the equivalent spectral models generated with other codes published in the literature, specifically NEMESIS (Irwin et al., 2008) and CHIMERA (Line et al., 2013). This comparison work was possible thanks to the collaboration of M. Line (CHIMERA) and J. Barstow (NEMESIS).

Significant work was done to achieve an acceptable level of agreement between the

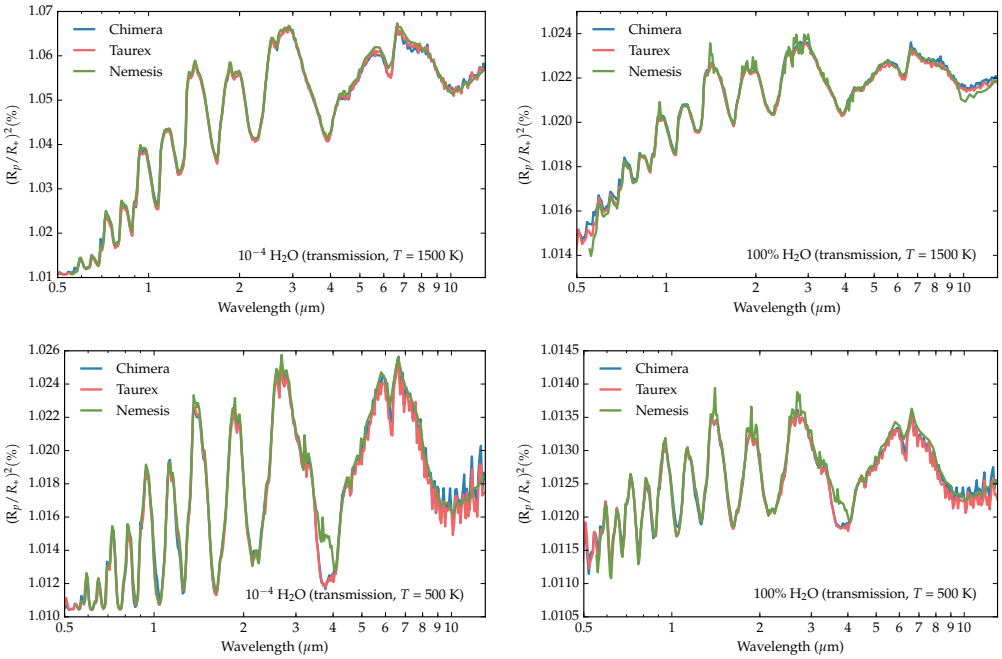

Figure 2.3: Comparison between four different transmission spectral models for an atmosphere with 10^{-4} H_2O and 100% H_2O , and with isothermal temperatures of 500 and 1500 K, obtained using TauREx, CHIMERA (Line et al., 2013) and NEMESIS (Irwin et al., 2008). The simulated planet has $M_p = 1\,R_J$, $R_p = 1\,R_J$, and it is orbiting a star with $R_* = 1\,R_\odot$. Note that while the spectra generated with TauREx and CHIMERA have the same resolution ($R = 100$), the spectrum generated with NEMESIS has lower resolution at longer wavelengths.

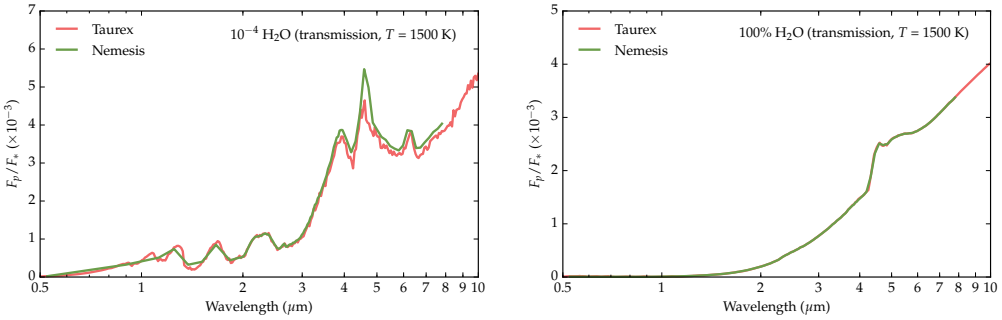

Figure 2.4: Comparison between two different emission spectral models for an atmosphere with 10^{-4} H_2O and 100% H_2O, and with a TP profile with high altitude temperature of 2200 K and low altitude temperature of 1260 K, obtained using TauREx and NEMESIS (Irwin et al., 2008). The simulated planet is the same as the one shown in Figure 2.3. Note that the resolution of the NEMESIS spectrum is significantly lower than the resolution of the TauREx spectrum. This is due to differences in the native resolution of the absorption cross sections used to generate the models.

models. Different sources of errors and uncertainties in the computation of the forward models were investigated and assessed in detail. The major differences between these forward models found at the beginning of the benchmark were found to be mainly due to issues related to the generation of the pressure-altitude grid. Residual differences are attributed to the use of different absorption cross sections. In summary, the strongest differences were caused by:

- Assumption of a constant gravity or average scale height throughout the atmosphere. It is important to take into account the varying gravity and scale height as a function of altitude. Approximating it to a constant value leads to large errors.

- Difference in the definition of constants. This included obvious things like the definition of the Jupiter radius (equatorial, mean, polar radius and so on), used to define the radius of the planet, but also the number of significant figures used in the Boltzmann constant (k_B), which was rather surprising. Small differences in k_B are in fact amplified exponentially in the computation of the altitude grid (see Section 2.6).

- Number of layers (N_{layers}) used to compute the transmittance or emission flux. A sufficient number of layers (at least 75) is needed to approximate the integral in Equations 2.3.13 and 2.3.19 with sufficient accuracy.

- Use of cross section sampling compared to the use of the k-distribution method. This issue will be assessed in detail in Chapter 4.

The successful completion of the benchmark between different models led to general agreement between forward models. Figure 2.3 and Figure 2.4 show the transmission and emission forward models generated with the different algorithms (TauREx, CHIMERA and NEMESIS). Residual differences are mainly due to different opacity cross sections used, an issue discussed in Chapter 4.

2.5 Line-by-line integration and cross section sampling

The forward models described in the previous depend on the absorption cross section for a given species. In turn, these directly depend on the absorption lines of the gas probed and their profiles. Absorption cross sections are usually precomputed from large line list databases for different pressures and temperatures (e.g. Sharp & Burrows, 2007; Hill et al., 2013), taking into account broadening effects such as thermal and pressure broadening

within the medium. As part of this thesis, a large database of high resolution cross sections was computed for the study of planetary atmospheres. This set, described in the next chapter, is made publicly available[3].

Computing a spectrum using a high resolution cross section, where all the line profiles are well sampled, is equivalent to a *line-by-line* (LBL) integration. This is a computationally expensive task, which I will describe in detail in the next chapter. If we simulate a transmission spectrum of an atmosphere with 100 layers containing only one gas between 1 and 10 μm using LBL integration, we will need to compute about 100 billion optical depths. This takes hours in an average computer, and is therefore computationally unfeasible for a retrieval, in which several thousands of models need to be evaluated. There are different ways to overcome this problem. Firstly, it is possible to use minimisation techniques such as Optimal Estimation (Rodgers, 1976) to converge to a solution much more quickly than with sampling techniques such as MCMC and Nested Sampling. This method, however, is unable to grasp the degeneracies of the model parameters, or to find multimodal solutions, and it is therefore sub-optimal for the interpretation of exoplanetary spectra. Alternatively, one can simply decide to sample the cross section at much lower resolution. This method gives generally good results when the sampling resolution is about two order of magnitudes higher than the resolution of the final spectrum (see Chapter 4, Section 3), and it is often adopted in the retrieval of low resolution exoplanet spectra. Another, more accurate, and faster method is to use the correlated k-distribution method, a technique, described in the next section, that allows a much faster computation of the spectral transmittance, based on the grouping of spectral transmittances according to the absorption coefficient. However, cross section sampling and the k-distribution method bring some uncertainties and generates biases in atmospheric retrievals. These will be discussed and assessed in Chapter 4.

TauREx implements all these methods to compute low and high resolution spectra. It can generate line-by-line forward model spectra using precomputed cross sections at their native resolution ($R \sim 10^6$). However, due to the computational time needed to generate such high resolution spectra, this method is never used for spectral retrieval, but only for individual forward model generation. TauREx can also generate spectra from sampled lower resolution cross sections. Lastly, it can use pre-tabulated k-distributions at different resolutions. The generation of absorption cross sections and k-distributions tables is discussed in

[3]The database of cross sections is available at `http://www.taurex.online`. Please note that a registration is required to access the database.

the next chapter.

2.5.1 The k-distribution method

The k-distribution method lies its foundation on the fact that the spectral transmittance is independent of the ordering of the absorption coefficient k_ν, or equivalently the absorption cross section σ_ν, within a given spectral bin through a homogeneous layer. The method is well described in the literature (e.g. Lacis & Oinas, 1991; Liou, K N, 2002; Pierrehumbert, 2010), and is extensively used for radiative transfer calculations in the context of planetary atmospheres (e.g. Irwin et al., 2008; Showman et al., 2009; Fortney et al., 2010; Lee et al., 2011; Grimm & Heng, 2015). Here I will give a brief overview of its fundamentals.

Firstly, note that for consistency with the literature I will make use of the mass absorption coefficient, k_ν , rather than the absorption cross section, σ_ν. The mass absorption coefficient k_ν (in units of g cm^{-1}) is related to the absorption cross section σ_ν (in units of cm^{-2} mol^{-1}) by

$$k_\nu = \frac{\sigma_\nu \times N_A}{M_{\rm mol}}, \tag{2.5.1}$$

where N_A is the Avogadro number, and $M_{\rm mol}$ is the molar mass of the given species.

In the k-distribution method the integration of the transmittance over wavenumber space is replaced by an integration over k space. The transmittance within a bin interval $\Delta\nu$ in a single homogeneous layer at altitude z and height dz, containing one gas, can be expressed as:

$$\mathcal{T}_{\Delta\nu}(z) = \frac{1}{\Delta\nu}\int_{\Delta\nu} e^{-k_\nu \rho(z)\chi(z)dz} d\nu = \int_0^\infty e^{-k_\nu \rho(z)\chi(z)dz} f(k)dk \tag{2.5.2}$$

where ρ is the mass density, χ is the mixing ratio of the gas, and $f(k)$ is the normalised probability distribution function for $k_\nu(z)$ in the interval $\Delta\nu$, such that $\int_0^\infty f(k)dk = 1$. Here we assume k to be the mass absorption coefficient at altitude z of the absorbing gas.

A cumulative probability distribution function $g(k)$ can now be defined:

$$g(k) = \int_0^k f(k)dk. \tag{2.5.3}$$

As $g(k)$ is by definition a single-valued monotonically increasing function, it has a unique, smooth and monotonically increasing inverse, $k(g)$, known as the k-distribution function.

Equation 2.5.2 can therefore be rewritten as

$$\mathcal{T}_{\Delta\nu}(z) = \int_0^1 e^{-k(g)\rho(z)\chi(z)dz} dg. \tag{2.5.4}$$

Because $k(g)$ is a smooth function, this integral can be approximated numerically with far fewer steps than the integration in frequency space. Various method exists to integrate numerically such a smooth function. The most commonly used is the Gaussian quadrature method:

$$\mathcal{T}_{\Delta\nu}(z) \approx \sum_{i=0}^{N_{\text{quad}}} e^{-k_i(g)\rho(z)\chi(z)} w_i. \tag{2.5.5}$$

where k_i is the value of the k distribution at the i-th quadrature point, and w_i is the corresponding weight. A Gaussian quadrature integration with N_{quad} points can yield an exact result for polynomials of degrees $2N_{\text{quad}} - 1$. About 20 gaussian quadrature points ($N_{\text{quad}} = 20$) are generally sufficient to approximate the integral in Equation 2.5.2 with high accuracy. An example of $f(k)$ and $k(g)$ for the absorption cross section of H_2O is shown in Figure 2.5.

This method works also in the presence of multiple absorbers. When two gases are present, and when the integration is done through the entire atmosphere, the integral in Equation 2.5.4 becomes:

$$\mathcal{T}_{\Delta\nu}(z) = \int_0^1 \int_0^1 e^{-\int [k(g_1)\chi_1 + k(g_2)\chi_2]\rho dz} dg_1 dg_2, \tag{2.5.6}$$

where χ_1 and χ_2 are the mixing ratios of the two gases and ρ is the mass density. This can be approximated by a double sum:

$$\mathcal{T}_{\Delta\nu}(z) = \sum_{i=0}^{N_{\text{gas}}} \sum_{j=0}^{N_{\text{gas}}} e^{-\int [k_i\chi_1 + k_j\chi_2]\rho dz} w_i w_j, \tag{2.5.7}$$

and rearranged to give

$$\mathcal{T}_{\Delta\nu}(z) = \sum_{i=0}^{N_{\text{gas}}} e^{-\int k_i\chi_1\rho dz} w_i \times \sum_{i=0}^{N_{\text{gas}}} e^{-\int k_i\chi_2\rho dz} w_i. \tag{2.5.8}$$

By doing this, we are effectively multiplying the transmittances of the two gases together, therefore loosing the non-gray information within the k-distribution. This can be generalised

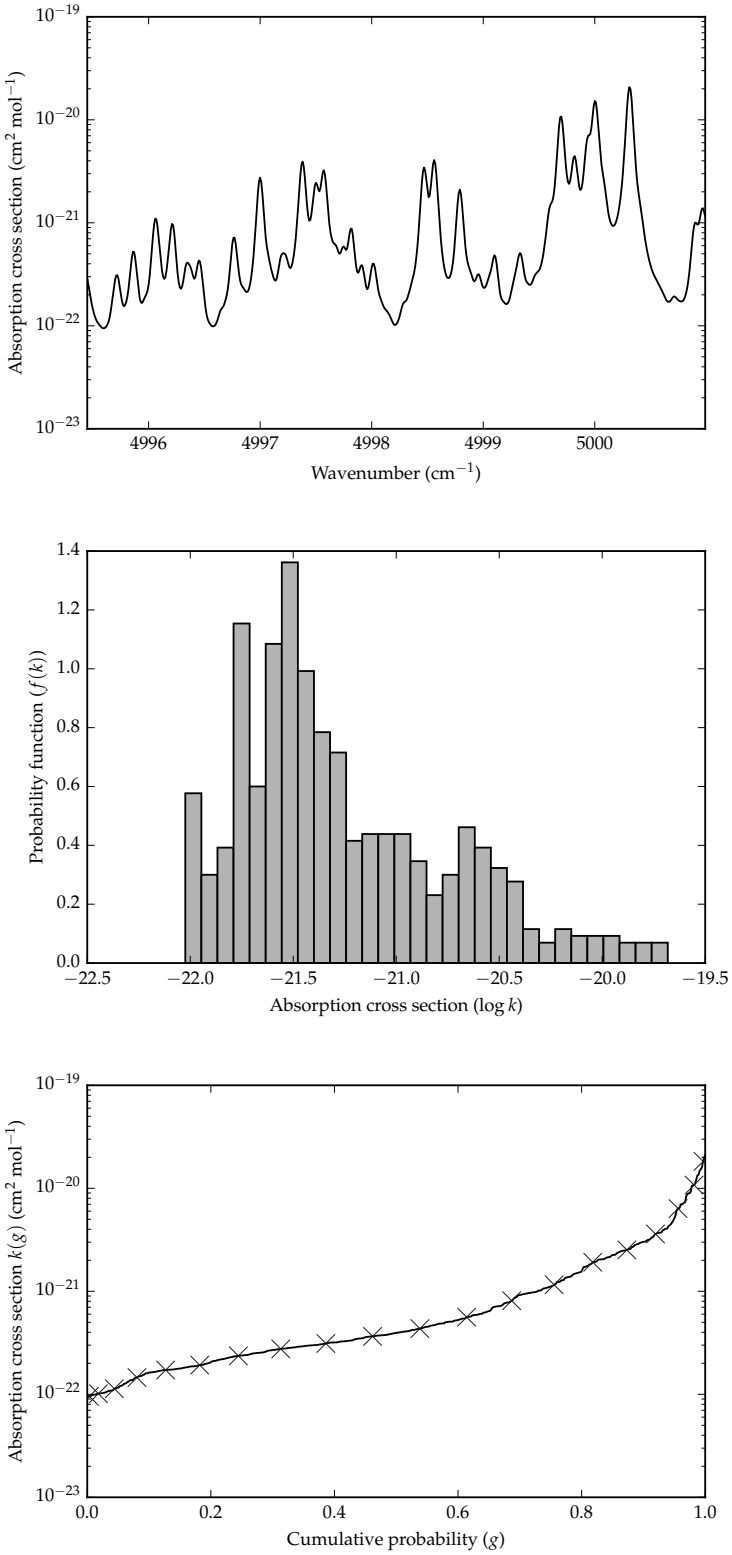

Figure 2.5: Top: absorption cross section as a function of wavenumber for H_2O at $T = 1500$ K and $P = 1$ bar, between 1619 and 1635 cm^{-1}. Middle: The probability function $f(k)$ of the absorption cross section shown in the top plot. Bottom: Cumulative probability function for $f(k)$ shown in the middle plot. The 'X' symbols shows the sampled Gaussian quadrature points for $N_{\text{quad}} = 20$.

to N_{gas} gases:

$$\mathcal{T}_{\Delta\nu}(z) = \prod_{j}^{N_{\text{gas}}} \left(\sum_{i=0}^{N_{\text{quad}}} e^{-\int k_{j,i} X_1 \rho dz} w_{j,i} \right). \tag{2.5.9}$$

When the k-distribution method is applied to transmission spectroscopy the numerical computation of the equivalent atmospheric depth (Equation 2.3.9) for the n-th wavelength bin becomes:

$$A(\lambda_n) \approx \sum_{j=0}^{N_{\text{layers}}} \left((R_p + z_j) \left[1 - \prod_{t}^{N_{\text{gas}}} \exp\left(- \sum_{i}^{N_{\text{quad}}} \sum_{k=0}^{N_{\text{layers}}-j} k_{n,t,i,k+j}\ \chi_{t,k+j}\ \rho_{k+j}\ \Delta l_{k,j} \right) 2 w_{n,i} \right] \right) \tag{2.5.10}$$

where N_{layers} is the number of layers, N_{gas} is the number of gases, N_{quad} is the number of quadrature points of the k-distribution; $k_{n,t,i,k+j}$ is the k-coefficient for the n-th wavelength bin, for gas t, for the quadrature point i and atmospheric layer $k+j$, $w_{n,i}$ is the weight of the i-th quadrature point for gas i and the n-th spectral bin, and $\Delta l_{k,j}$ is still defined in Equation 2.3.10. The different variables are otherwise identical to those in Equation 2.3.9 and are explained schematically in Figure 2.2.

The numerical expression in the case of emission for multiple gases is far more complicated and is not included here for convenience. We refer the interested reader to the source code of TauREx once it will become open source[4], where the relevant algorithm is fully commented.

2.5.2 The *correlated* k-distribution method

The k-distribution method assumes that the absorption coefficients are strictly correlated throughout the atmosphere so that integration in ν space can be readily replaced by integration in g space. However, in real exoplanet atmospheres, the line profiles for a given gas vary throughout different atmospheric layers as the temperature and pressure vary. In such cases, Equation 2.5.4 might not be exact. This is because, as the number of lines and their profiles vary within pressure levels, the order in which all the lines are rearranged to create the cumulative probability distribution is not necessarily the same. This, in turn, implies that the ν and g spaces are not exactly identical.

However, this lack of perfect correlation is generally assumed to lead to small and negligible uncertainties when particle scattering plays a minor contribution. When the k-

[4]The current plan is to make the code available on Github over the course of 2017

distribution method is used in atmospheric studies, it is therefore referred to as the "correlated" k-distribution approximation. In other words, it is assumed that there is a strict correlation between ν and g integrations even if the transmittance is evaluated over several layers, in which the absorption coefficients for a given spectral bin vary. The uncertainties caused by this approximation will be explored in Chapter 4.

2.6 Atmospheric parametrisation

From the previous two sections it is clear that both the transmission and emission spectra can probe a range of pressures in the atmosphere. In its simplest form, the main objective of a retrieval model is to obtain as much information as possible about the thermal structure of the atmosphere (namely the "temperature-pressure" or TP profile), the mixing ratio profiles of the various molecular constituents, and the presence of clouds and hazes. However, the finite number of data points in the relatively low resolution spectra of extrasolar planets forces us to make a number of assumptions. For instance, transmission spectra are much less sensitive to thermal gradients than emission spectra, hence most models assume an isothermal profile.

In TauREx, the integrals in Equation 2.3.13 and 2.3.19 defining the equivalent atmospheric depth in transmission and the total integrated radiation in emission, are converted into simple summations over a discrete number of layers N_{layers}. These summations in turn require the definition of the altitude, z, and the number density, ρ_N, of each layer. In the following I will describe how these two grids are derived.

Firstly, the pressure profile is sampled uniformly in logarithmic space between a minimum pressure P_{min} and maximum pressure P_{max} over a fixed number of layers N_{layers}. The mid-point pressure is then defined between each layer boundaries. This pressure is approximated to the geometric average (or log-average) of the pressure boundaries of each layer. The temperature T_i, number density $\rho_{N,i}$ and altitude z_i for each layer i and the mixing ratio $\chi_{m,i}$ for each molecule m and layer i are then defined. The number density at each atmospheric level i is computed *assuming hydrostatic equilibrium and the ideal gas law*, and is given by

$$\rho_{N,i} = \frac{P_i}{k_B T_i} \tag{2.6.1}$$

where k_B is the Boltzmann constant. The scale height of each atmospheric level i is then

given by

$$H_i = \frac{k_B T_i}{\mu_i g_i}, \tag{2.6.2}$$

where μ is the mean molecular weight and g is the gravity. The mean molecular weight is given by

$$\mu_i = \sum_{n}^{N_{\text{gas}}} \chi_{n,i} m_{\text{mol},n}, \tag{2.6.3}$$

where N_{gas} is the total number of molecules in the atmosphere, $\chi_{n,i}$ is the mixing ratio of molecule n at layer i and $m_{\text{mol},n}$ is the molecular weight of species n. The gravity is given by:

$$g_i = \frac{GM_p}{(R_p + z_i)^2}, \tag{2.6.4}$$

where z_i is the altitude at the i-th layer, R_p and M_p are the radius and mass of the planet, and G is the gravitational constant.

In order to calculate the altitude at each pressure level we need to compute the scale height, which requires the computation of the gravity, which in turns requires the computation of the altitude. This can only be achieved with an iterative process. In TauREx, we firstly compute the gravity and scale height of the zeroth layer. Then, the change in altitude of the next layer is simply given by

$$\Delta z(n) = -H_{n-1} \log\left(P_n / P_{n-1}\right), \tag{2.6.5}$$

where P is the mid-layer pressure as defined above. With the value of Δz for the zero-th layer, the gravity and scale height at the top of the zeroth layer (and, equivalently, bottom of the first layer) can be computed. This computation is then repeated iteratively until the top of the atmosphere is reached[5]. Note that the height of the zeroth layer, defined as z_0, is located at $\Delta z(0)/2$ above the surface defined by the radius R_p.

The model so far described allows us to generate transmission and emission spectra with arbitrary temperature and mixing ratio profiles. However, in the context of atmospheric sounding of extrasolar planets, the number of parameters in the current parametrisation is

[5]The top of the atmosphere is defined at a given pressure, usually 10^{-6} mbar. At such pressures, the optical depth is assumed to be zero.

quite substantial, making any retrieval attempt ill-posed. For each atmospheric layer we have in fact individual parameters for the temperature and the mixing ratio of each atmospheric constituent. While a great number of levels ($\sim$ 100) are needed to approximate correctly the integral in Equation 2.3.13 and 2.3.19, the signal-to-noise and resolution of current data do not allow us to constrain the temperature and mixing ratios of each individual atmospheric layer. Further approximations are therefore adopted. Firstly and most simply, in both emission and transmission spectroscopy, constant-with-altitude mixing ratio profiles are usually assumed (e.g. Benneke & Seager, 2012; Line et al., 2013; de Wit & Seager, 2013; Waldmann et al., 2015b). This reduces the parameter space considerably to just one parameter for each atmospheric constituent. A constant-with-altitude temperature profile is also usually assumed in the case of transmission. While this is a widely adopted approximation that reduces the parameter space considerably, in Chapter 5, I will show that it can lead to strong biases, and I will propose alternative solutions. For the case of emission spectroscopy, where the spectra are highly non-linearly dependent on temperature gradients, further parameterisations of the TP profiles are adopted. These are described in the next section.

2.6.1 Temperature-pressure profiles

The determination of the vertical atmospheric temperature profile is one of the key challenges in the retrieval of atmospheric emission spectra. Typically two approaches exist in the retrieval of the TP-profile: 1) Layer-by-layer retrieval; 2) Analytic parameterisation. In TauREx we include both approaches, together with a "hybrid" method. We also include an isothermal profile, and a "3-point" and "4-point" profile.

2.6.1.1 Layer-by-layer

In the layer-by-layer method the temperature of each atmospheric layer is fitted independently. However, this method can converge to a solution only in the presence of extremely high signal-to-noise and broad wavelength coverage observations, such as those obtained for the Earth's atmosphere and solar system planets (Rodgers, 1976; Hanel et al., 2003). In a low signal-to-noise and resolution scenario, such as that expected from the observations of exoplanets, this method has extremely poor convergence properties. A common solution to this problem is to impose a "regularisation" of the temperature (e.g. Irwin et al., 2008), based on the fact the adjacent atmospheric layers should exhibit some correlation in temperature.

2.6.1.2 Parametric model

Another approach consists of parametrising the temperature profile based on some analytic models. This method attempts to model the underlying physics of the thermal structure, while minimising the number of parameters. However, despite the clear advantage of reducing the parameter space compared to a layer-by-layer approach, the solution will always be constrained within the bounds of the model assumed.

Several analytical models exist in the literature, ranging from radiative-convective approximations to global circulation models (Liou, K N, 2002; Hubeny et al., 2003; Burrows et al., 2008; Hansen, 2008; Madhusudhan & Seager, 2009; Showman et al., 2009; Guillot, 2010; Pierrehumbert, 2010; Robinson & Catling, 2012; Heng & Workman, 2014; Heng et al., 2014). Following Guillot (2010), the mean global temperature profile, as a function of the thermal optical depth (see Equation 2.6.8), for a simple radiative downstream-upstream approximation can be expressed as

$$T^4(\tau_{\rm th}) = \frac{3{T_{\rm int}}^4}{4}\left(\frac{2}{3}+\tau_{\rm th}\right)+\frac{3{T_{\rm irr}}^4}{4}\zeta_{\gamma_1}(\tau_{\rm th}) \tag{2.6.6}$$

where $T_{\rm int}$ is the planet internal heat flux, $T_{\rm irr}$ the stellar flux at the top of the atmosphere and

$$\zeta_{\gamma_i}(\tau_{\rm th}) = \frac{2}{3}+\frac{2}{3\gamma_i}\left[1+\left(\frac{\gamma_i\tau_{\rm th}}{2}-1\right)e^{-\gamma_i\tau_{\rm th}}\right]+\frac{2\gamma_i}{3}\left(1-\frac{{\tau_{\rm th}}^2}{2}\right)E_2(\gamma_i\tau_{\rm th}) \tag{2.6.7}$$

where $\gamma_1 = \kappa_V/\kappa_{IR}$ is the ratio of mean opacities in the optical (κ_V) and infra-red (κ_{IR}) and E_2 is the second-order exponential integral.

We note that similar parameterisations exist in the literature (e.g. Robinson & Catling, 2012). We also include the variation by Line et al. (2013) and Parmentier et al. (2015) including two optical opacity sources κ_{V1} and κ_{V2} and a weighting factor between optical opacities (left as free parameter), $\alpha_{\rm V}$,

$$T^4(\tau_{\rm th}) = \frac{3{T_{\rm int}}^4}{4}\left(\frac{2}{3}+\tau_{\rm th}\right)+\frac{3{T_{\rm irr}}^4}{4}(1-\alpha_{\rm V})\zeta_{\gamma_1}(\tau_{\rm th})+\frac{3{T_{\rm irr}}^4}{4}\alpha_{\rm V}\zeta_{\gamma_2}(\tau).$$

This parametrisation with two visible streams allows more freedom for a temperature inversion (Line et al., 2013).

The temperature as function of opacity $\tau_{\rm th}$ can be mapped to a pressure grid by assuming the following relation

$$\tau_{\rm th} = \frac{\kappa_{IR} P}{g}. \tag{2.6.8}$$

2.6.1.3 Other TP-profiles

In addition to the above TP-profiles, we include an isothermal profile as well as a 3-point and "N-point" profile. The 3-point profile is purely geometric and keeps the top of atmosphere temperature the tropopause temperature and pressure, T_1, P_1, and the surface (or 10 bar pressure) temperature T_{10bar} as free variables. The temperature profile is then linearly interpolated in $\ln(P)$. The N-point profile add extra "nodes" to the parametrised profile.

2.6.1.4 The hybrid approach of TauREx

TauREx implements a "hybrid" approach, combining the parametrised method to a layer-by-layer approach. This method consist of two retrievals, or "stages". In the first stage a parametric model retrieval is computed, then, in the second stage, the retrieval solution is used to guide a layer-by-layer model, thus relaxing the parametric model constraint of the first stage retrieval.

After the first stage retrieval, the error on the sampled parametric model parameters is converted to a one sigma lower and upper temperature bounds for each atmospheric layer. These bounds are obtained by computing the temperature profiles corresponding to all sampled parametric models in the Bayesian retrieval, and then calculating the standard deviation of the temperature value of each atmospheric layer. The following matrix is then calculated:

$$D^2_{i,j} = |\hat{T}_i - \hat{T}_j|^2 + (\sigma_i + \sigma_j)^2, \tag{2.6.9}$$

where i and j are subsequent atmospheric layers, $\hat{T}$ is the maximum likelihood temperature estimator of the parametric fit, and σ_i is the error in temperature for each layer calculated as described above. This matrix is then normalised in terms of the minimal and maximal temperature variations found in the TP-profile:

$$C_{i,j} = 1 - \frac{D^2_{i,j} - \text{argmin}(D)}{\text{argmax}(D)} \tag{2.6.10}$$

which can be thought as a temperature correlation matrix with layers most similar in tem-

perature featuring the highest correlation.

A second correlation matrix is then defined, imposing an exponential correlation length across pressure levels (Rodgers, 1976):

$$S_{i,j} = (S_{i,i}S_{j,j})^{1/2} \exp\left(-\frac{|\ln(P_i/P_j)|}{\mathcal{C}}\right), \tag{2.6.11}$$

where $\mathcal{C}$ is a correlation length in terms of atmospheric scale heights. Larger values of $\mathcal{C}$ correspond to a stronger smoothing in the TP-profile, and might be preferable in low signal-to-noise, low resolution observations. This correlation matrix is the same used by Irwin et al. (2008). A hybrid correlation matrix is then built combining the previous two matrixes:

$$Q_{i,j}(\alpha) = C_{i,j} + (1-\alpha)S_{i,j}, \tag{2.6.12}$$

where α is a scaling factor ranging from 0 to 1. In the second stage retrieval, the parameter α is set as a free parameter. This allows us to dynamically relax the parametric model solution from a model-constrained solution (i.e. stage 1) to an unconstrained solution.

Lastly, in order to optimise the number of free parameters, the sparsity of the temperature profile solution of the first stage is used to compute a nonlinear sampling of the profile used in the second stage. This compression algorithm uses the correlation matrix C to only retain layers corresponding to a change greater than 2% in the temperature gradient with respect to the previously retained layer. Whenever no change in thermal gradient is detected for > 10 layers, a new layer is included nevertheless. With this approach a 100-layers atmospheric model can typically be reduced to 15–25 free parameters. Such number of free parameters is still very high, therefore requiring high signal-to-noise and high resolution spectra.

2.6.2 Compositional data analysis

The retrieval of atmospheric composition from exoplanetary spectra should be treated in the context of compositional data analysis (Aitchison, 1986). In other words, the absolute abundance of each species cannot in principle be considered independently, as they are part of a whole. Hence, when the bulk composition of the atmosphere cannot be assumed a priori, it becomes important to preserve the unit sum constraint of all abundances.

The mixing ratios of all species must satisfy the following constraints:

$$0 < \chi_i < 1 \tag{2.6.13}$$

$$\sum_{i}^{N_{\rm gas}} \chi_i = 1 \tag{2.6.14}$$

where χ_i is the mixing ratio of the i-th gas and $N_{\rm gas}$ is the total number of absorbers and inactive gases.

This is not a problem for hot Jupiter atmospheres. In this case the bulk composition of the atmosphere is made by hydrogen and helium, and the fitted gases are only trace gases, making a very small contribution to the whole atmosphere. The unit sum constraint is easily preserved, by summing all the trace gases, and assuming that the remainder of the atmosphere is made by a mixture of hydrogen and helium. The ratio of hydrogen to helium can be assumed a priori (e.g. solar), or independently fitted. However, in the case of high mean molecular weights atmospheres, such as for super-Earth or Earth-like atmospheres, absorbing gases such as H_2O, CO, CO_2, etc., can represent a significant fraction of the whole. We cannot therefore let each gas abundance vary independently between 0 and 1, as Equation 2.6.14 would not necessarily be satisfied. In such case, the unit sum constraint can only be preserved by employing appropriate transformations.

As shown by Aitchison (1994), the positive and unit-sum constraints preclude the application of standard statistical techniques. The properties of compositional data are indeed peculiar, as they represent parts of a whole, and, as such, they carry only relative information. Variables are not free to vary independently, as shown in their variance-covariance structure (see Aitchison, 1986, Chapter 3) and at least one covariance is negative, implying spurious correlations (Chayes, 1971). For an introduction on the subject, see Pawlowsky-Glahn & Egozcue (2006).

The appropriate sample space for the study of compositions is the N dimensional unit *simplex*:

$$\begin{aligned} \mathcal{S}^N = \big\{ &(X_1,\ldots,X_N) : X_i > 0\,(i = 1,\ldots,N)\,, \\ &X_1 + \cdots + X_N = 1 \big\}. \end{aligned} \tag{2.6.15}$$

Aitchison (1982) introduced the idea of log-ratio representation of compositional data,

and in the TauREx we make use of the centered-log-ratio (clr) transformation, defined as:

$$\mathscr{Y} = \mathrm{clr}(\chi) = \left[\ln \frac{\chi_1}{g(\mathscr{Y})}, \ldots, \ln \frac{\chi_{N_{\mathrm{gas}}}}{g(\mathscr{Y})}\right], \tag{2.6.16}$$

$$\chi = \mathrm{clr}^{-1}(\mathscr{Y}) = \mathscr{C}\left[e^{\mathscr{Y}_1}, e^{\mathscr{Y}_2}, \ldots, e^{\mathscr{Y}_{N_{\mathrm{gas}}}}\right], \tag{2.6.17}$$

where $g(\mathscr{Y}) = \sqrt[N_{\mathrm{gas}}]{\chi_1 \cdots \chi_N}$ and $\mathscr{C}$ is the closure operation. Note also that $\sum_i^{N_{\mathrm{gas}}} \mathscr{Y}_n = 0$.

In TauREx, the variables $\left\{\mathscr{Y}_1, \ldots, \mathscr{Y}_{N_{\mathrm{gas}}-1}\right\}$ are set as free parameters, while $\mathscr{Y}_{N_{\mathrm{gas}}}$ is derived ($\mathscr{Y}_{N_{\mathrm{gas}}} = -\sum_i^{N_{\mathrm{gas}}-1} \mathscr{Y}_n$). Note that this effectively allows us to reduce the parameter space by one. A relatively higher value of $\mathscr{Y}_i$ means that component i is more abundant with respect to the other components, and vice-versa for relatively smaller values.

2.6.3 Chemically-consistent approach

So far I have described a model in which the atmospheric composition is not constrained by any pre-determined chemistry and is allowed to vary freely. This is certainly the most advisable approach in a data-driven context and when the underlying physics of the atmosphere is unknown a priori.

However, it is also interesting to investigate the retrievability of the atmospheric state using chemical models that do make some assumptions on the atmospheric chemistry. In TauREx, we have implemented a chemical scheme that assumes chemical equilibrium. Chemical equilibrium can be considered as a fair approximation for hot Jupiters with temperatures higher than ~ 1500 K (Venot et al., 2012; Moses et al., 2013a) as, at such high temperatures, the chemical reaction timescales are much shorter than the dynamical timescales.

The chemical equilibrium calculations are performed using the Atmospheric Chemical Equilibrium (ACE) program, an adaptation of the code developed by Agúndez et al. (2012). It determines the chemical equilibrium of a system by minimizing the Gibbs energy following the algorithm of Gordon & Mcbride (1994), and uses the thermochemical data described in Venot et al. (2012). The parameters of this model are the C to O ratio and the metallicity, while the He and N fractions are fixed to solar values. When this chemically-consistent approach is used, the mixing ratios of the individual species, set as free parameter in the more typical approach, are replaced by the C to O ratio and the metallicity.

Chapter 3

Molecular absorption cross sections

A key input to the emission and transmission spectra forward models described in the previous chapter is the absorption cross section for each absorbing molecule. Absorption cross sections are a measure of the probability of a molecule to absorb electromagnetic radiation, due to scattering, collisions, or excitation of quantised states, and it is usually expressed in units of area as a function of wavelength (or, equivalently, wavenumber). In the context of atmospheric models, the main phenomena leading to the observed opacity include Rayleigh and Mie scattering, collision induced absorption, and molecular, atomic and ionic absorption. In this chapter, I will discuss the computation of molecular absorption cross sections, originating from the excitation of roto-vibrational transitions.

Estimating the cross sections of an absorbing species is a computationally expensive task. It involves the tabulation of transitions (or "lines") due to excitation of energy states, and the creation of absorption profiles for each absorption line. Such profiles, due to local broadening effects, strongly depend on the local temperature and pressure. For this reason, cross sections are usually precomputed for a range of pressures and temperatures, and then interpolated to the required grid when computing spectral models.

In the following sections, I will describe how cross sections can be computed from transitions databases (or "line lists"). I will also present a precomputed set of cross sections for a range of molecules important for modelling exoplanetary atmospheres, created using the best and most up-to-date inputs available today. This set is made publicly available to the community[1], and represents today one of the largest and most complete databases of cross sections readily available.

[1] `www.taurex.online`

3.1 Molecular line lists

Molecular opacities are usually precomputed in the form of absorption coefficients (or equivalently absorption cross sections) as a function of wavenumber from *molecular line lists*. These line-lists are organised in the form of catalogues of transitions, and contain part of the information needed to compute cross sections, specifically the spectral line positions and intensities. In turn, the accuracy of the cross sections will mainly depend on the completeness of these databases. Large databases of molecular line lists exist for a wide range of molecules (e.g. HITRAN, Rothman et al., 2013), but mainly for Earth-like temperatures, where experimental data are widely available. However, in the study of exoplanetary atmospheres much higher temperatures are usually probed (600–3000 K), and line lists computed for Earth-like temperatures are therefore inappropriate. At these high temperatures several additional vibrational levels are in fact populated, resulting in several millions of additional spectral lines, which are simply missing from these lower-temperature databases.

In recent years, there has been intense work in the computation of more complete molecular line lists valid for higher temperature regimes. For instance, the HITEMP database (Rothman et al., 2010) contains experimental and theoretical high-temperature line lists for molecules such as CO, CO_2, NO and OH. More recently, the ExoMol project (Tennyson & Yurchenko, 2012) has generated theoretical high-temperature line lists for a large number of molecules, many of which are relevant to the study of exoplanetary atmospheres (e.g. CH_4, NH_3, H_2O, HCN, VO, PH_3, SO_2, H_2S and SiO).

Usually line list databases contain Einstein coefficients (A_E), measuring the probability of spontaneous emission for a given transition between states, together with the degeneracies and energies for each state. With these information, line intensities can easily be calculated. The intensity of a line j resulting from a transition between two states is given by:

$$S_j(T) = \frac{A_{Ej} g_{\mathrm{up}\,j}}{8\pi c \nu_j^2 Q(T)} \exp\left(\frac{-hcE_{\mathrm{low}\,j}}{k_B T}\right)\left[1 - \exp\left(\frac{-hc\nu_j}{k_B T}\right)\right], \tag{3.1.1}$$

where ν_j is the line position, $g_{\mathrm{up}\,j}$ and $E_{\mathrm{low}\,j}$ are the upper-state degeneracy and lower-state energy respectively, A_{Ej} is the Einstein coefficient for the transition, k_B is the Boltzmann constant, c is the speed of light, h is the Planck constant, and $Q(T)$ is called the *partition function*, given by

$$Q(T) = \sum_j g_{\mathrm{up}\,j} e^{-E_{\mathrm{low}\,j}/k_B T}. \tag{3.1.2}$$

3.1.1 Line broadening

In any real medium, spectral lines extend over a range of wavelengths, and they are said to be "broadened". There are mainly three form of broadening due to local effects: thermal (or doppler) broadening, pressure broadening, and natural broadening. Only doppler and pressure broadening play a significant role when modelling planetary atmospheres, while natural broadening is often considered negligible.

Natural broadening is caused by the uncertainty principle that relates the lifetime of an excited state with the uncertainty of its energy. This effect results in a Lorentzian profile. Doppler, or thermal, broadening depends on the line-of-sight velocity distributions of the atoms or molecules present in the atmosphere, and is caused by the distribution of red- and blue-shifts due to the Doppler effect. A higher local temperature will produce larger Doppler broadening. Thermal broadening results in a Gaussian profile. Lastly, pressure broadening is caused by the collision of other particles with the emitting particle. When an emitting particle collides with another particle, the emitting process is interrupted, hence decreasing the process characteristic time, and increasing the uncertainty in the energy emitted. This effect depends on the pressure and temperature of the gas, and results in a Lorentzian profile (as for natural broadening).

Other, non-local, sources of broadening might be present when studying exoplanetary atmospheres. In close-in hot Jupiters the presence of strong winds or fast rotation can cause further broadening (Spiegel et al., 2007; Miller-Ricci Kempton & Rauscher, 2012; Showman et al., 2012; Snellen et al., 2014). Broadening can also occur in the presence of strong magnetic fields (Kislyakova et al., 2014). Such non-local effects are not taken into account here, as they are very dependent on the nature of each system.

When both pressure and doppler broadening act on the same line, the resulting profile is a Voigt profile, the convolution of a Lorentzian and a Doppler profile. These three profiles are shown in Figure 3.1.

The Doppler profile as a function of wavenumber ν is given by

$$f_G(\nu - \nu_0) = \frac{1}{\gamma_{G,\nu_0}\sqrt{\pi}} \exp\left(-\frac{(\nu - \nu_0)^2}{\gamma_{G,\nu_0}^2}\right), \tag{3.1.3}$$

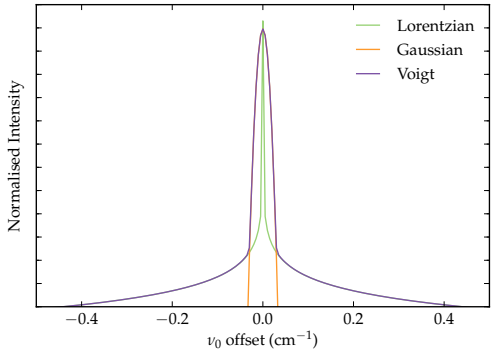

Figure 3.1: Lorentzian profile (green line), Gaussian profile (orange line) and their corresponding convolution (dashed line), a Voigt profile, for an arbitrary H_2O transition at 10,000 cm^{-1}, assuming a temperature of 1500 K and a pressure of 1 bar.

where ν_0 is the line centre position in wavenumber, γ_G is the Doppler width, given by:

$$\gamma_{G,\nu_0} = \sqrt{\frac{2k_B T}{m_{\text{mol}}}} \frac{\nu_0}{c}, \tag{3.1.4}$$

where k_B is the Boltzmann constant, T is the temperature, m_{mol} is the mass of the molecule and c is the speed of light. The Lorentzian broadening profile for pressure P is given by:

$$f_L(\nu - \nu_0) = \frac{1}{\pi} \frac{\gamma_{L,\nu_0}}{(\nu - \nu_0)^2 + \gamma_{L,\nu_0}^2}, \tag{3.1.5}$$

where γ_L is the Lorentzian half-width at half-maximum (HWHM), which can be approximated using the following relation (Sharp & Burrows, 2007):

$$\gamma_{L,\nu_0} = \left(\frac{T_{\text{ref}}}{T}\right)^{n_T} \frac{P}{P_{\text{ref}}} \sum_{i}^{N_{\text{broad}}} \gamma_{L,i}(\nu_0)\chi_i, \tag{3.1.6}$$

where N_{broad} is the number of broadening agents, $\gamma_{L,i}(\nu_0)$ is the Lorentzian HWHM of broadener i at the reference temperature and pressure at which $\gamma_{L,b}(\nu_0)$ is evaluated, n_T is the temperature index, T and P are the temperature and partial pressure of the gas respectively, χ_i is the mixing ratio of the broadening agent i. Note that spectral lines can be pressure-broadened by several broadening agents, i.e. gases, including from the same gas producing the absorption lines (self-broadening). In hot Jupiters atmospheres, for example, pressure-broadening is mainly due to hydrogen and helium which dominates the bulk atmosphere, while other broadeners are usually ignored. However, self-broadening, and broadening from other molecules, may become important in high mean molecular weight

atmospheres such as those expected in super-Earths.

Compared to Gaussian profiles, which are relatively straight forward to compute, Lorentzian profiles present more caveats, as the HWHM, $\gamma_{L,b}(\nu_0)$, of a specific broadener for a given line needs to be measured experimentally in most cases, and data are usually unavailable or scarce for the range and temperatures probed in exoplanetary atmospheres. Therefore, extrapolation to different temperature and pressure regimes is obtained using Equation 3.1.6.

Another important factor is to determine the extension of each Voigt profile. Real profiles do not extend indefinitely as the Lorentzian profiles, but become sub-Lorentzian at a certain distance from the line centre (Allard et al., 2005; Freedman et al., 2008; Allard et al., 2012). Again, due to the lack of experimental data, the sub-Lorentzian profile of broadened lines is relatively unknown, and different *wing cutoffs* are therefore assumed, often without experimental justification.

In the presence of both pressure and doppler broadening, the resulting profile is the convolution of the two line shapes, which is also known as the *Voigt profile*:

$$f_V(\nu-\nu_0) = \int_{-\infty}^{\infty} f_L(\nu'-\nu_0) f_G(\nu-\nu') d\nu' \tag{3.1.7}$$

$$= \frac{1}{\pi^{3/2}} \frac{\gamma_{L,\nu_0}}{\gamma_{G,\nu_0}} \int_{-\infty}^{\infty} \frac{1}{(\nu'-\nu_0)^2+\gamma_{L,\nu_0}^2} \exp\left[\frac{-(\nu-\nu')^2}{\gamma_{G,\nu_0}^2}\right] d\nu'. \tag{3.1.8}$$

The integral in Equation 3.1.7 cannot be solved analytically, and numerical methods need to be used (e.g. Humlícek, 1982; Thompson et al., 1987; Schreier, 1992; Thompson, 1993; Kuntz, 1997). Figure 3.1 shows the Lorentzian and Gaussian profiles together with their convolution.

3.2 A database of cross sections for planetary atmospheres

Line intensities and positions provided by line list databases, and broadening profiles, provided by experiments or semi-empirical laws, allow the computation of absorption cross sections.

The contribution at a given pressure P, temperature T and for a given line j is given by

$$\sigma_{j,P,T}(\nu) = S_j(T) \frac{f_V(\nu)}{\int_{\nu_j-\Delta\nu_j/2}^{\nu_j+\Delta\nu_j/2} f_V(\nu) d\nu} \tag{3.2.1}$$

Table 3.1: Molecular line lists from which cross sections and k-distributions tables have been created.

Molecule	Line list	Coverage cm^{-1}	T_{max} K	Number of lines	Reference(s)
CH_4	YT10to10/ExoMol	0 – 12000	1500	9 819 605 160	Yurchenko et al. (2013)
NH_3	BYTe/ExoMol	0 – 12000	1500	1 138 323 351	Yurchenko et al. (2011)
H_2O	BT2/ExoMol	0 – 30000	3000	505 806 202	Barber et al. (2006)
HCN	Harris/ExoMol	0 – 12000	3000	168 110	Harris et al. (2006), Barber et al. (2014a)
CO	HITEMP	3 – 8465	–	113 631	Rothman et al. (2010)

where $S_j(T)$ is the line intensity (Equation 3.1.1), ν_j is the line position, $f_V(\nu)$ is the Voigt profile for pressure P and temperature T, and $\Delta\nu_j$ is the wing cutoff for line j. The integral in the denominator is used to renormalise the profile after the wing cutoff is applied.

The contribution of the cross section within a bin ν, $\sigma(\nu)$, is then the sum over the contributions of individual lines:

$$\sigma_{P,T}(\nu) = \sum_j \sigma_{j,P,T}(\nu). \tag{3.2.2}$$

As part of this work, a complete and consistent set of high resolution cross sections was created for several molecules, in a range of temperatures and pressures, from the most complete line lists available and using the most up-to-date broadening coefficients found in the literature. These cross sections form the foundation on which atmospheric exoplanet spectra are modelled within TauREx, but can also be used for other applications, such as for modelling brown dwarf atmospheres (Line et al., 2015). Table 3.1 shows all the molecules for which cross sections have been computed, with their relative line list sources, number of lines, spectral and temperature coverage. Note that CO_2 is not included here, despite being one of the most important molecules in exoplanetary spectra. The cross section for this molecule will be computed in the very near future.

This computationally expensive task has been made possible thanks to codes developed as part of the ExoMol project[2] (Al-Refaie et al, in preparation) and to large scale cluster computing. In order to create this set of cross sections, the COBWEB and LEGION[3] high performance computing (HPC) platforms were used. COBWEB is a 148 CPUs cluster exclusively available to our group, and LEGION is the UCL's general-use cluster. This set

[2] https://github.com/ahmed-f-alrefaie/CEXSY
[3] https://www.ucl.ac.uk/research-it-services/research-computing

Table 3.2: Temperatures and pressures for which cross sections and k-distribution tables for the molecules shown in Table 3.1 have been created.

T (K)	100	200	300	400	500	600
	700	800	900	1000	1100	1200
	1300	1400	1500	1600	1700	1800
	1900	2000	2200	2400	2600	2800
	3000	3200	3400			
P (bar)	1×10^{-5}	2.1544×10^{-5}	4.6416×10^{-5}	1×10^{-4}	2.1544×10^{-4}	4.64×10^{-4}
	1×10^{-3}	2.1544×10^{-3}	4.6416×10^{-3}	1×10^{-2}	2.1544×10^{-2}	4.6416×10^{-2}
	1×10^{-1}	2.1544×10^{-1}	4.6416×10^{-1}	1	2.1544	4.6416
	10	21.544	46.416	100		

of cross sections is publicly available through the TauREx website[4].

3.2.1 Temperature and pressure grids

Each cross section was computed for 22 pressures and 27 temperatures, for a total 594 temperature-pressure (TP) nodes. The 22 pressures were determined keeping a constant sampling in log space, while the temperatures were determined arbitrarily. The grid is shown in Table 3.2 and has been used for all molecules. Cross sections for intermediate temperatures and pressures can be generated by interpolating existing TP nodes. In the future, additional temperatures and/or pressures will be likely added.

When data is available, the maximum wavenumber grid used extends between 200 and 30 000 cm^{-1}, or between 0.33 and 50.0 μm, otherwise the range is reduced to match the spectral coverage of the line list. The lower wavenumber range was cut to 200 cm^{-1} mainly for computational reasons, as the required sampling increases for smaller wavenumbers, but also because current and future instruments are not expected to probe wavelength ranges above 50 μm.

3.2.2 Spectral resolution

In order to achieve optimal sampling of all line profiles, while also optimising the computational time required to generate the cross sections, an adaptive sampling grid was built for each molecule, at each temperature and pressure. This adaptive grid was computed such that the Voigt profiles are well sampled across the entire spectral range.

In order to understand how the Voigt profile width changes as a function of wavenumber, temperature and pressure, we can approximate the Voigt width with this expression (Olivero & Longbothum, 1977):

[4] `www.taurex.online`

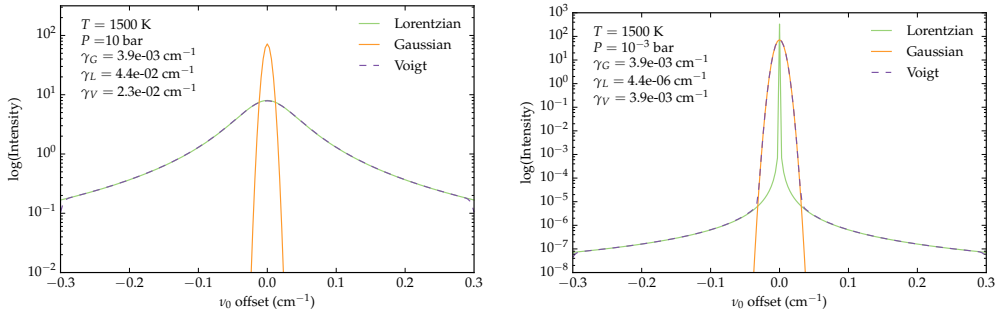

Figure 3.2: Lorentzian, Gaussian and the corresponding Voigt profiles for a line centred at 10,000 cm^{-1} with arbitrary intensity, for $T = 1000$ K, in the high (left plot) and low (right plot) pressure regimes.

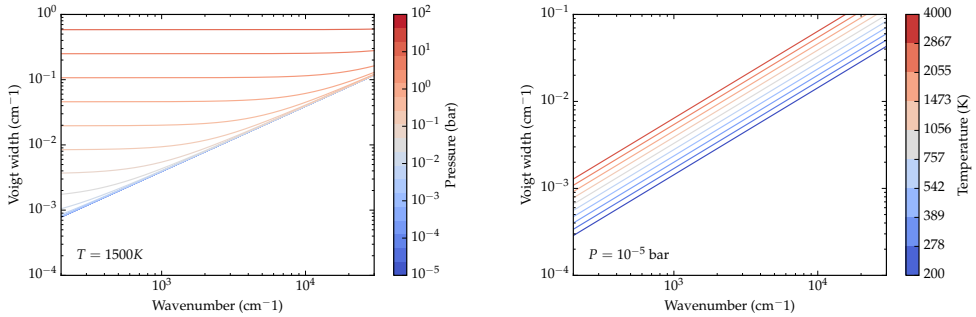

Figure 3.3: Wavelength-dependent Voigt profile width for different pressures assuming a constant temperature of 1500 K (left) and for different pressures assuming a constant pressure of 1×10^{-5} bar (right).

$$\gamma_V \approx 0.5346\,\gamma_L + \sqrt{0.2166\,{\gamma_L}^2 + {\gamma_G}^2} \tag{3.2.3}$$

where γ_L and γ_G are the Lorentzian HWHM and the Doppler width respectively, shown in Equations 3.1.6 and 3.1.5.

Figure 3.2 shows the Voigt profiles in a low and high pressure regime. We can see that while at low pressure the Lorentzian width is much smaller than the Doppler width, in the high pressure regime the Voigt width is dominated by the Lorentzian width. It is also instructive to see how the Voigt width changes as a function of three variables: pressures, temperature, and wavenumber. Figure 3.3 shows the Voigt width, γ_V, for lines with positions between 200 and 30 000 cm^{-1} for a range of pressures and temperatures. Here, for the calculation of the Lorentzian width, only one broadener was considered, with $\gamma_L = 0.01$ cm^{-1} and $n_T = 0.5$ (but note that usually γ_L also changes as a function of different quantum numbers, in particular J). It can be seen that for low pressures (right plot), the Voigt

width is proportional to the wavenumber at all temperatures, signifying that the Voigt width is always dominated by the Doppler width (see Equation 3.1.5). Conversely, in the high pressure regimes (left plot) the Voigt width is constant in wavenumber, meaning that this time the profile is dominated by the Lorentzian width (see Equation 3.1.6). Lastly, at some combinations of pressures and temperatures, the Voigt width is dominated by the Doppler profile and Lorentzian profiles at large and low wavenumbers respectively.

The cross sections generated as part of this work have been computed for the modelling of hydrogen dominated atmospheres, and therefore consider H_2 and He as the only broadening agents. The optimal sampling for each cross section, at each temperature, pressure and spectral range was therefore computed assuming the Voigt profiles widths resulting from H_2/He broadening. Note that the cross section code used here requires a fixed grid spacing for each computation. Each cross section was therefore sub-divided into several spectral ranges, and an average width was then computed for each sub-range. In addition, as the code can more efficiently calculate several TP nodes simultaneously, up to three temperatures for a given pressure were combined, and the corresponding Voigt widths averaged. It was decided to sample each Voigt width with four points. This was found to be a good compromise between good sampling and computational requirements. The value of γ_L and n_T assumed to calculate the Voigt widths were determined by averaging the J-dependent γ_L and n_T (when available) for hydrogen only (i.e. ignoring helium). We consider this a good approximation for the purpose of determining an optimal sampling. Examples of sampling grids for specific molecules for different pressures and temperatures can be seen in the next sections (e.g. Figure 3.5), where the generation of cross sections for individual molecules is discussed.

3.2.3 The cross section of H_2O

Cross sections for the main isotopologue of water ($^1H_2{}^{16}O$) were created from the BT2 line list (Barber et al., 2006) as provided by ExoMol. Other line lists for water exist (e.g. Partridge & Schwenke, 1997), but the ExoMol version was found to be the most complete one available. This line list is today the *de-facto* standard for modelling water spectra, especially at high temperatures, and it has been used to model succesfully one of the first water spectra of an exoplanet (Tinetti et al., 2007). This line list contains over half a billion transitions between 0 and 30 000 cm^{-1}, compared to the 80 000 water lines known experimentally (HITRAN, Rothman et al., 2013), and it is valid for temperatures up to 3000 K.

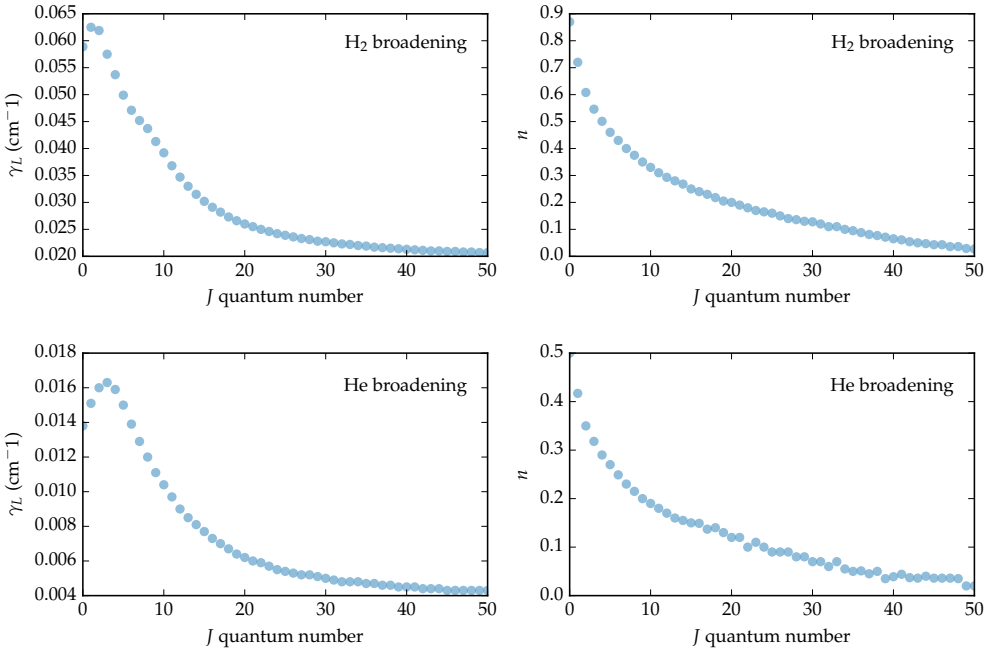

Figure 3.4: Lorentzian width γ_L and temperature index n_T as a function of the angular momentum quantum number J for H_2O lines

Pressure broadening coefficients for H_2 and He for a large number of quantum numbers are available (Solodov & Starikov, 2008; Voronin et al., 2010; Petrova et al., 2013, 2016, Barton et al. in prep.). Here only the dependence with the total angular momentum quantum number J is considered. The broadening parameters are taken from the `.broad` files for both H_2 and He as provided by ExoMol (Tennyson et al., 2016), and assume that each line at all pressures and temperatures is broadened by a mixture of hydrogen and helium. Figure 3.4 shows the Lorentzian width γ_L and the temperature dependent index n_T as a function of J used to compute the Voigt profiles.

The wavenumber grid is divided into 10 sub-ranges, with uniform distance in log space. The sampling for each wavenumber sub-range, temperature and pressure is defined as described in the previous section. In Figure 3.5 the grid spacing is shown for a low and high pressure regime, while in Figure 3.6 it can be appreciated that sufficient sampling of each Voigt profile is obtained in different conditions.

3.2.4 The cross section of CH_4

The Y10to10 line list from ExoMol (Yurchenko et al., 2013) was used to create the cross sections for CH_4. Computing a large number of high resolution cross sections for this molecule represents a tougher challenge compared to water, as the number of transitions

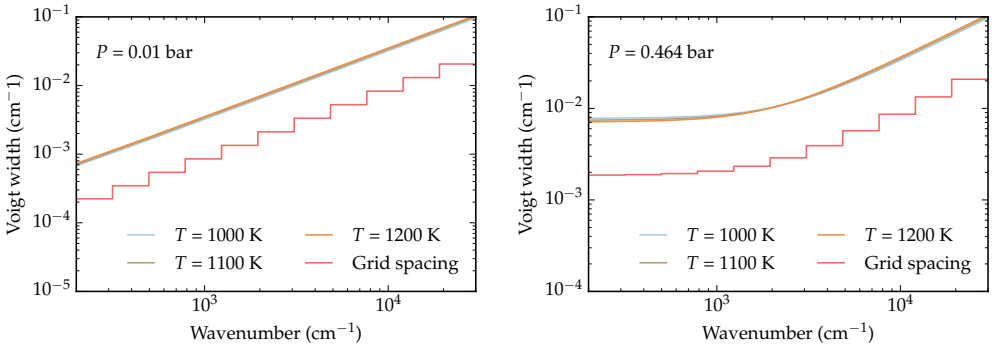

Figure 3.5: Adaptive grid spacing as a function of wavenumber used to create the cross sections of H_2O for a low and high pressure regime (P = 0.01 and 0.465 bar respectively), for a set of three temperatures (T = 1000, 1100 and 1200 K). The grid spacing is calculated as described in the text.

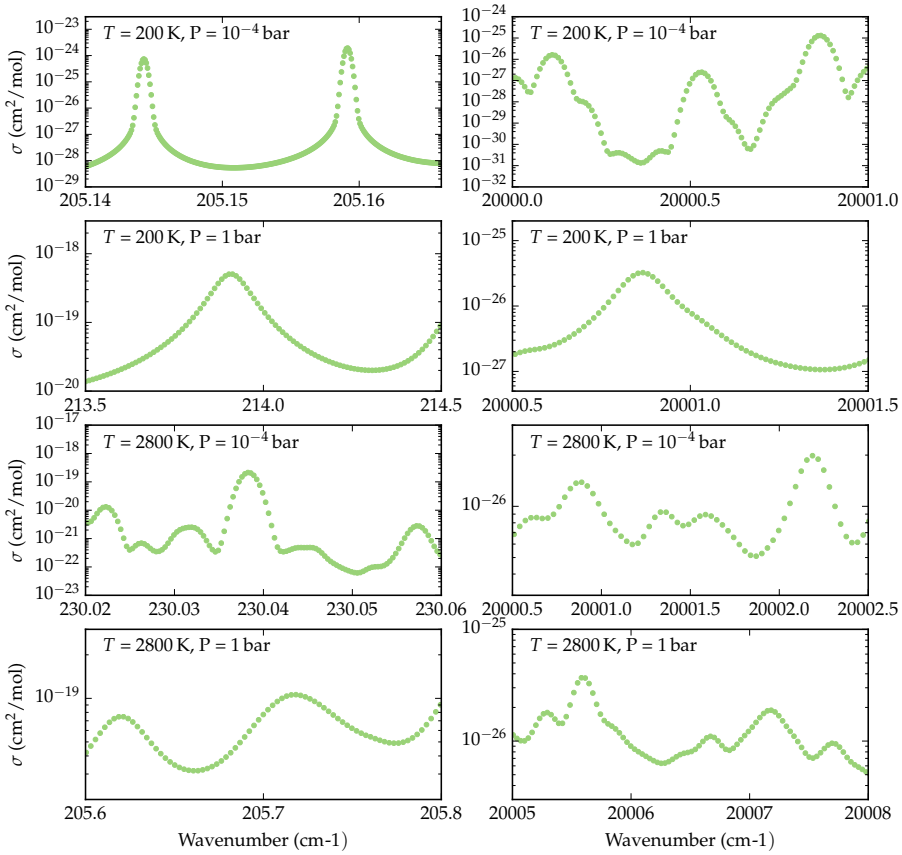

Figure 3.6: Sampling of the cross sections of H_2O for low and high wavenumber ranges ($\sim$200 and $\sim$ 20000 cm^{-1}), low and high temperatures (200 and 2800 K), and low and high pressure (10^{-3} and 1 bar). It can be seen that line profiles are well sampled in all cases.

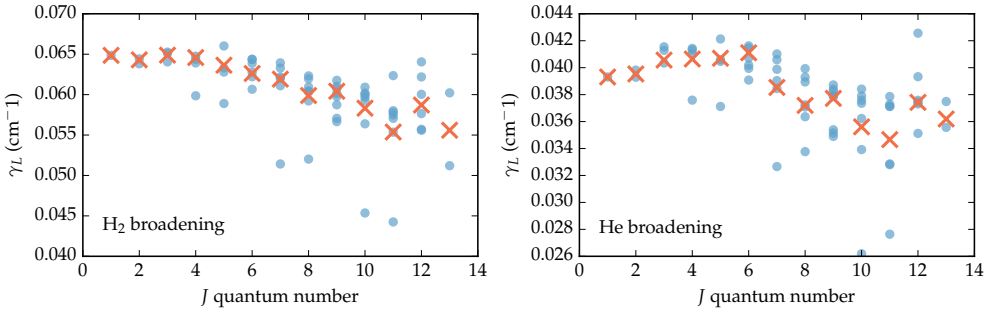

Figure 3.7: Lorentzian width γ_L for CH_4 as a function of the angular momentum quantum number J (Pine et al. 1992). The orange crosses show the weighted average of the Lorentzian widths available for a given J.

exceeds 10 billions. This is about 20 times more transitions than those found in the BT2 line list.

Transitions for Y10to10 are provided in separate `.trans` files (Tennyson et al., 2016) in ranges of 100 cm^{-1}, from 0 to 12000 cm^{-1}. Given the large number of transitions, in some spectral regions the size of each single uncompressed `.trans` file is about 8 GB, with the entire line list occupying about 350 GB in its uncompressed format. In order to compute the cross sections, the wavenumber grid was therefore divided into 100 cm^{-1} sub-ranges, matching the ranges provided by the `.trans` files. In this way, taking into account the wing contributions from lines outside each sub-range, only three `.trans` files had to be considered for each wavenumber sub-range, optimising the storage.

Pressure broadened coefficients for H_2 and He are scarcely available. Various sources were considered (Fox et al., 1988; Varanasi & Chudamani, 1990; Pine, 1992; Margolis, 1993; Strong et al., 1993; Grigoriev et al., 2001; Gabard et al., 2004), but only data from Pine (1992) were used for the Lorentzian widths of hydrogen and helium, as they are explicitly given with a J dependence. As different Lorentzian widths are available for a given J, corresponding to different transitions, a weighted average value was used. Figure 3.7 shows the Lorentzian widths used for J between 0 and 15. When data was not available for a given J, an average value was used.

The temperature index n_T for H_2 was taken from Margolis (1993), but no J dependence was available, so an average value is used ($n_T = 0.44$). This value compares well with the value given by Varanasi & Chudamani (1990) ($n_T = 0.46$). For He the temperature index is determined from Varanasi & Chudamani (1990) ($n_T = 0.26$), and again no J dependence is considered.

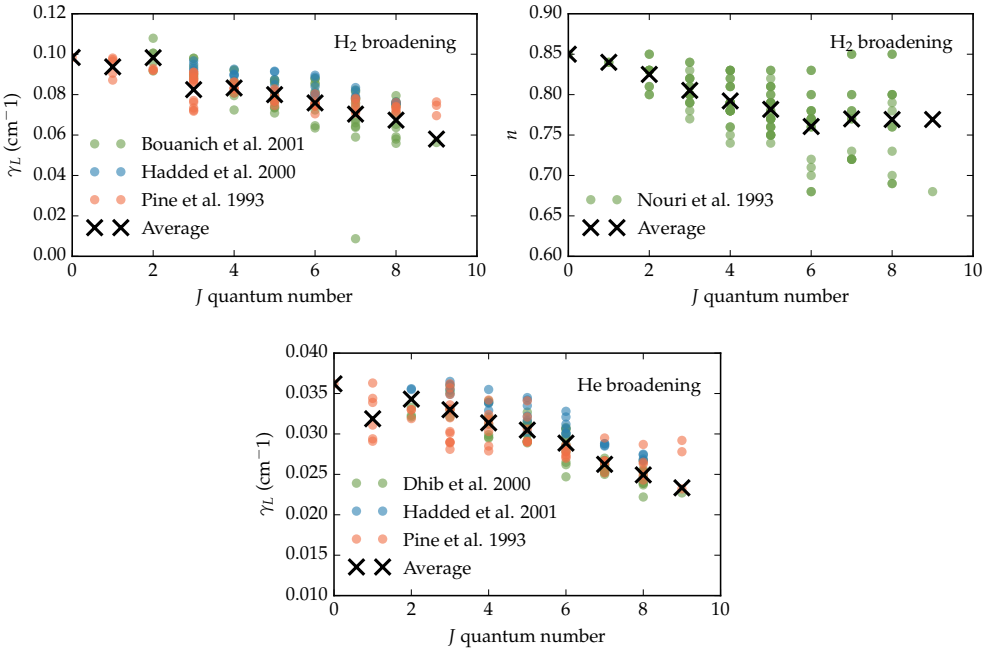

Figure 3.8: Lorentzian width γ_L and temperature index n_T as a function of the angular momentum quantum number J for NH_3 lines broadened by H_2 (top) and He (bottom). Different colours correspond to data extracted from difference sources, as shown by the legend. The black crosses show the weighted average of coefficients corresponding to the same J value.

3.2.5 The cross section of NH_3

The most comprehensive line list valid for temperatures up to 1500 K is provided by ExoMol (BYTe, Yurchenko et al., 2011). It contains over one billion transitions in the spectral range 0–12,000 cm^{-1}, built from 1373,897 energy levels with J values less than 36.

Pressure broadening coefficients for H_2 and He were obtained and combined using multiple sources. The J dependent Lorentzian width for H_2 broadening was determined using data from Hadded et al. (2001), Bouanich et al. (2001) and Pine et al. (1993). The temperature index for H_2 was extrapolated from Nouri et al. (2004). When multiple coefficients for a given J were available, a weighted average for each J value was used. In the case of He-induced broadening, the Lorentzian widths were determined using experimental data from Dhib et al. (2000), Hadded et al. (2001) and Pine et al. (1993) (see Figure 3.8). There is no comprehensive data available for the temperature index of He induced broadening, hence a value of 0.5 was assumed for all transitions.

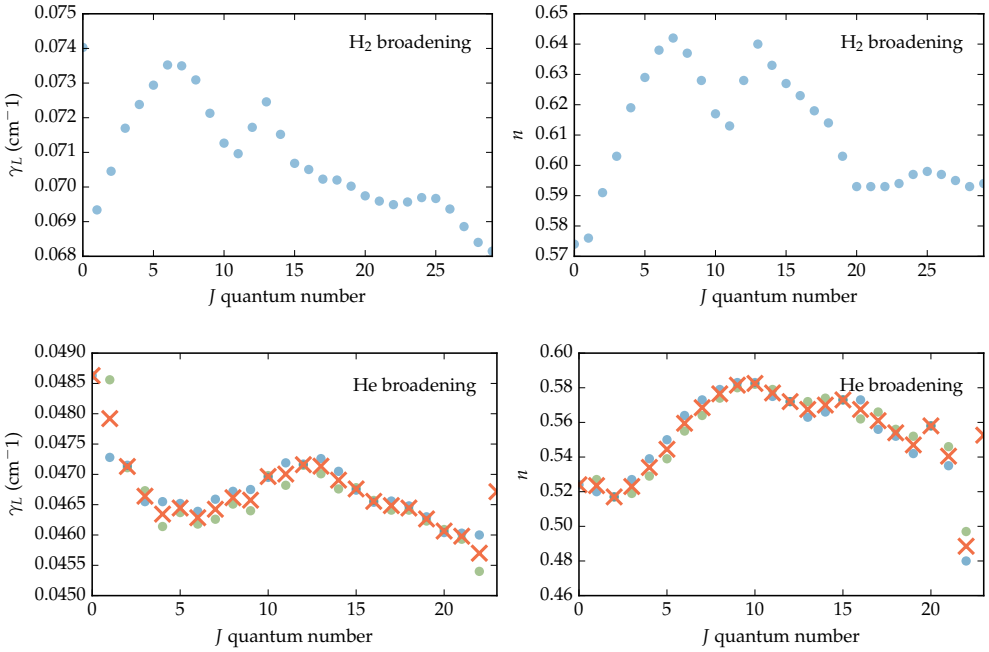

Figure 3.9: Lorentzian width γ_L and temperature index n_T as a function of the angular momentum quantum number J for CO lines broadened by H_2 (top) and He (bottom). In the bottom plots, the blue and green circles represent the coefficients for the R and P branch respectively, while the orange crosses represent the average between the two.

3.2.6 The cross section of CO

The HITEMP database (Rothman et al., 2010) provides a line list for CO, containing 113,631 transitions up to $J = 149$ in the spectral range 3–8465 cm^{-1}, sourced from the line list provided by Goorvitch (1994). The line intensities are converted from the standard temperature of 296 K to other temperatures using the partition function as recommended by Rothman et al. 2010. The partition functions are available through HITRAN[5] up to $T = 3000$ K. The value of the partition function at higher temperatures was extrapolated using a quadratic spline interpolation.

Broadening coefficients for H_2 and He were obtained from a variety of sources. For H_2, data were obtained from Faure et al. (2013), providing the Lorentzian width γ_L and temperature index n_T for values of J up to 29 (Figure 3.9, top). An average value was used for larger Js ($\gamma_L = 0.0708$ cm^{-1} and $n_T = 0.612$). For He the broadening coefficients for Js up to 23 were obtained from Mantz et al. (2005). The values used correspond to the average value for the P and R branches, as shown in Figure 3.9 bottom.

[5]The partition functions are available at `http://hitran.iao.ru/partfun`

3.3 *k*-distribution tables

The correlated *k*-distribution method was presented in Chapter 2 as a technique to decrease the computational time needed to model low resolution emission and transmission spectra. One can pre-compute *k*-distributions for low resolution spectral grids and then efficiently compute transmission and emission spectra.

Once high resolution cross sections are computed, it is trivial to compute *k*-distribution tables. As part of this work, *k* distribution tables were calculated for a range of spectral resolutions, for all the molecules shown in Table 3.1 and for all the temperatures and pressures shown in Table 3.2. The full set of *k*-tables is available online[6]`www.taurex.online`.

[6](

Chapter 4

Uncertainties in absorption cross sections and their effects on atmospheric retrievals

The atmospheric models used for modelling exoplanetary atmospheres contain several approximations, and have inherent uncertainties that are often difficult to assess. These include simplifications of the geometry, such as using 1D models instead of 3D models, assumptions about the uniformity of the atmosphere probed by a transmission or emission spectrum, lack of complete line lists or accurate broadening coefficients when computing molecular cross sections. This represents a major issue, especially when these models are used as forward models in atmospheric retrievals. Assessing the induced biases is therefore important.

In this chapter, I will investigate some of the biases linked to the generation and use of molecular cross sections. These biases are inherent to all atmospheric models: ultimately, even the most advanced general circulation model, taking into account 3D dynamical effects, will be limited by the uncertainties in the atmospheric opacities used. It is therefore useful to firstly assess these potential sources of error, such that further errors introduced by different approximations can be compared with these, often unavoidable, systematics.

There are two kinds of uncertainties when dealing with absorption cross sections. There are physical sources of error, such as the completeness of the line list, and knowledge of the correct line profiles, including broadening parameters, line profile wing cut offs and line positions. Then there are errors introduced by using techniques to reduce the computational time needed to deal with the highest resolution cross sections, such as cross section sampling and the k-distribution method. In the next sections, some of these uncertainties will be investigated and propagated to simulated spectra of typical exoplanets, in the resolutions expected by future instruments such as JWST, so that their effects on atmo-

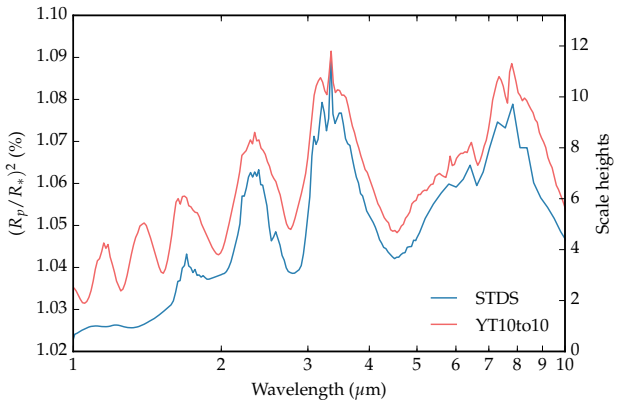

Figure 4.1: Transmission spectrum of a hot Jupiter orbiting around a $1\,R_\odot$ star, with radius $R_p = 1\,R_J$, mass $M_p = 1\,M_J$, isothermal TP profile with $T = 1500\,\mathrm{K}$ and with 10^{-4} of CH_4. The two spectra were obtained using the YT10to10 and STDS line lists (orange and blue spectra respectively). The secondary y-axis indicates the atmospheric scale height. One scale height corresponds to $\approx 209\,\mathrm{km}$, or equivalently 61 ppm (parts per million) in units of $(R_p/R_*)^2$.

spheric retrievals can be better understood. For the most significant cases, I will assess the biases that these uncertainties cause in atmospheric retrievals.

4.1 Completeness of line lists

The uncertainties related to the lack of completeness of a line list for a given temperature are an important factor to assess. This is especially true for the hottest temperatures, where experimental data are often unavailable, and a combination of experimental and analytical methods need to be used to generate line lists. At these high temperatures, an increasing number of energy levels is populated, so that the number of additional transitions often increases dramatically. This causes the emergence of new spectral features and the strengthening of bands also seen at lower temperatures. Incomplete line lists would miss these features, and would therefore be a major source of error in spectral modelling. While it is difficult to assess accurately these uncertainties in most cases, as experimental data are often lacking, it is instructive to see the sort of effects that the use of an incomplete line list has in the interpretation of a typical exoplanet spectrum.

I take CH_4 as an example, and compare transmission spectra obtained using two line lists. One is the YT10to10 line list from Exomol (Yurchenko et al., 2013), containing about 10 billion transitions, and considered to be valid up to 1500 K, in the 0–12000 cm^{-1} range. The second one is obtained from the Spherical Top Data System (STDS; Wenger & Champion, 1998), and contains only about 9 million transitions in the 0–6500 cm^{-1} range. The

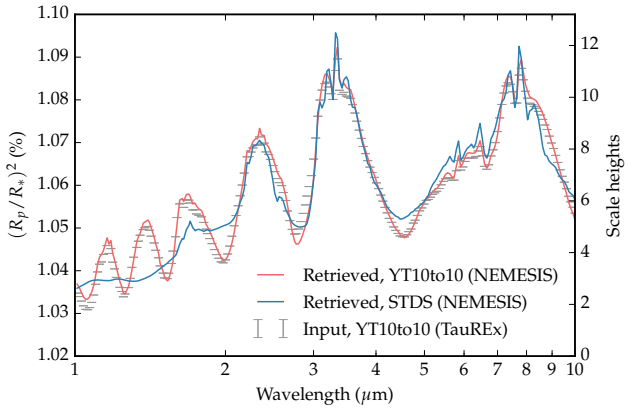

Figure 4.2: Best fit spectra of the YT10to10 spectrum shown in Figure 4.1 retrieved using NEMESIS with the STDS and YT10to10 line lists (orange and blue spectra respectively). The input spectrum is shown in grey, with error bars corresponding to the assumed noise of 20 ppm.

two spectra, shown in Figure 4.1, were calculated assuming a H_2/He dominated hot Jupiter atmosphere ($R_p = 1\,R_J$, $M_p = 1\,M_J$), with an isothermal TP profile with T = 1500 K, containing 10^{-4} of CH_4. It can be seen that, although the shapes of the absorption features are similar, the difference between the two transmission spectra is still significant. This is not surprising, given that the YT10to10 contains 1000 times more transitions than the STDS line list. The most evident effect is that the STDS spectrum, with respect to the YT10to10 spectrum, misses bands below 1.5 μm and underestimates the absorption by about 150 ppm between 1.5 and 3 μm, and by about 100 ppm at longer wavelengths. The use of the STDS line list at these temperature would therefore lead to strong biases in the retrieved atmospheric parameters.

In order to assess these biases, the YT10to10 spectrum generated with TauREx was retrieved using the NEMESIS retrieval tool (Irwin et al., 2008) using k-distribution tables obtained from both the YT10to10 and STDS line lists. NEMESIS was used instead of TauREx to retrieve the spectrum because the former can use CH_4 opacities generated from both STDS and YT10to10, while the latter can only use YT10to10 cross sections. Note that a benchmark between TauREx and NEMESIS was presented in Chapter 2, Section 4,. The input spectrum had a resolution of $R = 100$ and 20 ppm noise was assumed for each data point. The best fit spectra are shown in Figure 4.2. We can see that while the retrieval using the YT10to10 line list can fit the spectrum well, as expected the fit using the STDS line list is quite poor. The retrieved CH_4 abundance using the STDS data is $4.5 \pm 0.4 \times 10^{-5}$, and 13.4 sigma below the input value of 10^{-4}. The retrieved radius is $R_p = 1.0049 \pm 0.0001\,R_J$,

almost 50 sigma away from the input value of 1 R_J. This also shows the retrieved error bar is significantly underestimated. Surprisingly, despite the strong differences between the input forward model and the retrieved model, the retrieved temperature is in agreement with the input value of 1500 K ($T = 1550 \pm 55$ K), but this could likely be due to the prior imposed to this parameter. The second retrieval using the YT10to10 line list produces the expected constraints: the temperature is $T = 1500 \pm 60$ K, the CH_4 abundance is $1.0 \pm 0.2 \times 10^{-4}$ and the radius is $R = 1.000 \pm 0.004 R_J$, in agreement with the true state.

While only one example was used to assess the potential problems of using incomplete line lists, similar effects are expected with other molecules and in other scenarios, such as when interpreting emission spectra. Although it is difficult to characterise the uncertainties due to the incompleteness of a line list, in this section I showed that the potential effects are significant. It is therefore important to always use complete and valid line lists when interpreting the spectra of exoplanets, in particular when detailed abundances are retrieved. When complete line lists are not available for the temperatures probed, much care is needed in interpreting atmospheric spectra.

4.2 Line profiles

In the following two sections I will investigate the effects that uncertainties in the line profiles have on high resolution cross sections and lower resolution exoplanet spectra. Two sources of error will be investigated: the uncertainties on the broadening parameters (namely uncertainties on the Lorentzian width of each line), and the choice of wing cutoff. It is noted that a previous work has already investigated the role of these uncertainties in cross sections (Hedges & Madhusudhan, 2016), but a full assessment of how these propagate in typical exoplanet spectra was not explored. Ultimately, this is the most interesting question: *are the propagated uncertainties in exoplanet spectral models relevant compared to other systematics?*

The assessment will proceed in two stages, in both cases. Firstly, following Hedges & Madhusudhan (2016), differences in the high resolution cross sections computed using different assumptions on the line profiles will be quantified using the relative median difference, defined as:

$$\bar{\Delta\sigma} = \text{median}\left\{\left|\frac{\sigma_2(\hat{\nu}) - \sigma_1(\hat{\nu})}{\sigma_2(\hat{\nu})}\right|\right\} \qquad \text{for} \quad \hat{\nu} = \{\nu_\text{i} \quad ... \quad \nu_\text{f}\} \tag{4.2.1}$$

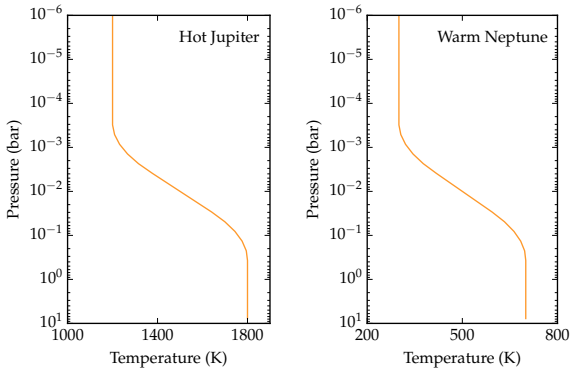

Figure 4.3: Temperature-pressure profiles used for the computation of the emission spectra of the hot Jupiter (left) and warm Neptune (right).

where σ_2 and σ_1 are the two cross sections, and $\{\nu_i ... \nu_f\}$ is the spectral interval.

Secondly, these differences will be propagated in the transmission and emission spectra of two transiting exoplanets: a typical hot Jupiter and a typical warm Neptune. The properties of the hot Jupiter resemble the exoplanet HD 189733b (Bouchy et al., 2005). The planet has radius $R_p = 1.138\,R_J$, mass $M_p = 1.144\,M_J$, and orbits around a star with radius $R_* = 0.756\,R_\odot$ and temperature $T_* = 5,000$ K. The atmosphere is assumed to be hydrogen and helium dominated (the mean molecular weight is 2.3 amu), and to be composed of 10^{-4} of H_2O, which is roughly consistent with the abundance of H_2O expected from chemical equilibrium (Venot et al., 2012; Moses, 2013). For simplicity, however, I do not consider other molecules expected from chemical equilibrium, such as CO or CO_2. The transmission spectrum assumes an isothermal profile ($T = 1500$ K), while in emission the TP profile varies from 1300 to 1800 K, as shown in Figure 4.3. The typical warm Neptune resembles the properties of GJ 436 b (Butler et al., 2004). The planet has radius $R_p = 0.376\,R_J$, mass $M_p = 0.073\,M_J$, and orbits around a star with radius $R_* = 0.464\,R_\odot$ and temperature $T_* = 3000$ K. In transmission, an isothermal profile with $T = 500$ K was assumed, while in emission the TP profile varies from 300 to 700 K (Figure 4.3). The other atmospheric properties of this smaller planet are the same as for the hot Jupiter.

The transmission and emission spectra of these exoplanets are shown in Figure 4.4 and Figure 4.5, binned at a resolving power of 100 for clarity. These spectra were modelled using high resolution cross sections ($R \sim 10^6$) computed assuming the best set of broadening parameters available, and an adaptive wing cutoff of 500 Voigt widths ($500\gamma_V$), as described in the previous chapter. The integrated spectral transmittance and the thermal emission

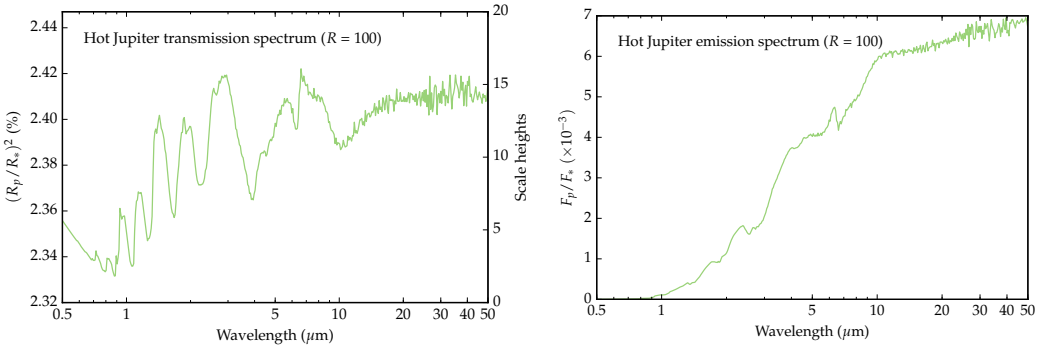

Figure 4.4: Transmission (left) and emission (right) spectra at a resolving power of 100 of a typical hot Jupiter. The secondary y-axis in the transmission plot indicates the atmospheric scale height. One scale height corresponds to ≈ 110 km, or equivalently 64 ppm in units of $(R_p/R_*)^2$.

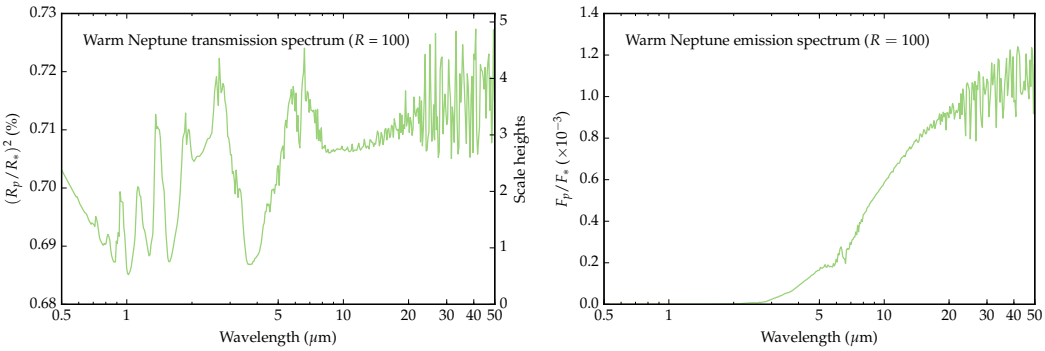

Figure 4.5: Transmission (left) and emission (right) spectra at a resolving power of 100 of a typical warm Neptune. The secondary y-axis in the transmission plot indicates the atmospheric scale height. One scale height corresponds to ≈ 190 km, or equivalently 97 ppm in units of $(R_p/R_*)^2$.

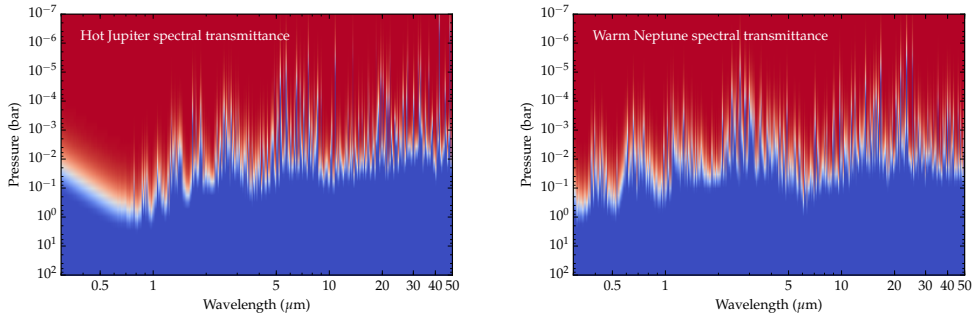

Figure 4.6: Integrated spectral transmittance as a function of pressure for the transmission spectra of the hot Jupiter and warm Neptunes shown in Figures 4.4 and 4.5. The transmittance is integrated over the path parallel to the line of sight. Red represents a fully transparent medium, where the transmittance is unity, while blue represents a fully opaque medium, where the transmittance is zero. These plots allow us to see the pressures (and therefore the temperatures and scale heights) probed at different wavelengths. We can see that most absorption occurs between 1 and 10^{-4} bar in both cases, with increasingly deeper layers probed by shorter wavelengths.

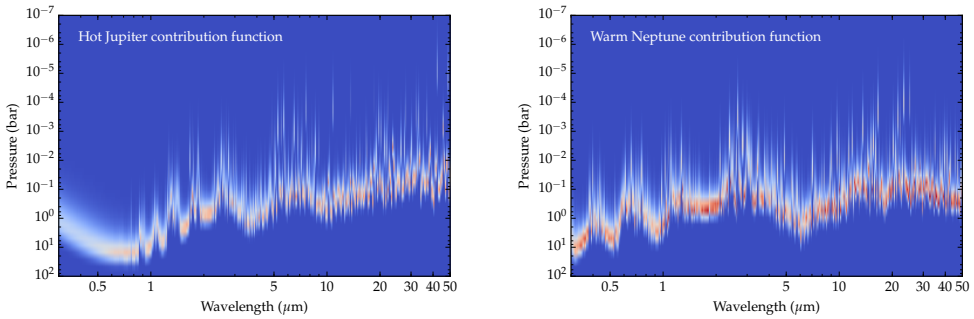

Figure 4.7: Thermal emission contribution function for the emission spectra of the hot Jupiter and warm Neptunes shown in Figures 4.4 and 4.5. This plot shows at which pressures the emission is coming from as a function of wavelength. Red corresponds to a maximum of the thermal emission weighting function, where $\tau = 1$, and blue represents no emission. We can see that most emission occurs between 50 and 0.01 bar in both cases, with increasingly deeper layers probed by shorter wavelengths.

contribution function, for the transmission and emission spectra respectively, are shown in Figures 4.6 and 4.7. These plots allows us to see from which atmospheric layers most of the absorption or emission comes from. We can see that transmission spectra probe pressures from 1 to 10^{-4} bar, emission spectra probe higher pressures from more than 10 bar to 10^{-2} bar.

In what follows, I will compare the differences between these spectra with those computed using cross sections that assume slightly different line profiles. This will be done at several resolving powers, ranging from 50 to 1000. These extend from what it is achievable today (mainly with HST), to what will be possible with the instruments on-board JWST.

The differences between these spectra will be given in parts per million (ppm), as they refer to a change in transit depth, which is the main observable when interpreting spectra of transiting exoplanets. As explained in Chapter 2, the transmission and emission spectra are expressed in terms of change in transit depth of the primary and secondary transits respectively, as a function of wavelength. Therefore, any difference in the spectra will translate into a difference in transit depth. I will refer to a *detectable difference* when the change in transit depth is greater than the sensitivity of current and future instruments. For reference, the best error bar in transit depth achieved with WFC3/HST (covering the range 1.1–1.7 μm, and with a resolving power of ≈ 40) is 22 ppm for a single transit (Tsiaras et al., 2016), while the predicted sensitivity of JWST instruments varies from 20 ppm in the lowest wavelength range, to about 40 ppm in the longest wavelength range (Greene et al., 2016; Beichman et al., 2014), for a resolving power of about 100.

It is also useful to convert a change in $(R_p/R_*)^2$ in units of scale heights. One scale height corresponds to a change in transit depth equal to:

$$\mathcal{H} = \left(\frac{R_p + H}{R_*}\right)^2 - \left(\frac{R_p}{R_*}\right)^2 \tag{4.2.2}$$

where R_p and R_* are the planet and star radii respectively, and H is the scale height.

Note that the average scale height of the hot Jupiter and warm Neptune modelled here are 110 and 190 km respectively, or, in units of transit depth, 65 and 97 ppm respectively. Instead of expressing changes in transit depth in absolute terms, we can therefore "normalise" them and express them in units of scale heights. This is especially useful when considering transmission spectra, as the scale height determines the total amplitude of the spectrum, and therefore directly influences the signal-to-noise of an observed spectrum[1].

4.2.1 Broadening parameters

In the previous chapter we saw how the shape of the Lorentzian profile of each transition line depends on the half width at half maximum (HWHM, or γ_L) of the profile. The Lorentzian width, in turn, depends in a complex way on the temperature and pressure of the gas, and the specific broadening agent. Ideally, these parameters should be measured experimentally, but most times these are measured in Earth-like conditions, so that widths for different pressures and temperatures need to be extrapolated, using some simple relationships. As a reminder, the relationship most commonly used, and also used in this work, is (Thomas & Stamnes, 2002; Sharp & Burrows, 2007):

$$\gamma_{L,\nu_0} = \left(\frac{T_{\text{ref}}}{T}\right)^{n_T} \frac{P}{P_{\text{ref}}} \sum_{i}^{N_{\text{broad}}} \gamma_{L,i}(\nu_0)\chi_i, \tag{4.2.3}$$

where $\gamma_{L,i}(\nu_0)$ is the Lorentzian HWHM of broadener i at the reference temperature and pressure at which $\gamma_{L,b}(\nu_0)$ is evaluated, n_T is the temperature index, T and P are the temperature and partial pressure of the gas respectively, and χ_i is the mixing ratio of the broadening agent i.

The major source of error to investigate here is the effect that inaccurate broadening parameters propagates in low and high resolution exoplanet spectra at difference temperatures. Hedges & Madhusudhan (2016) found significant differences in the high resolution

[1]For a given error in transit depth, the signal-to-noise of a spectral feature will increase if the scale height increases. Here the signal-to-noise is assumed to be the ratio between the total amplitude of the spectral feature, divided by the average error in each spectral datapoint.

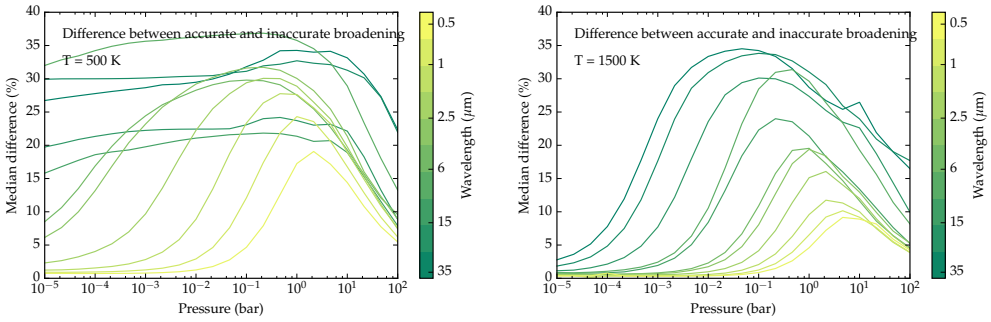

Figure 4.8: Median percentage difference between cross sections at 500 K (left) and 1500 K (right). One set of cross sections is computed with accurate broadening parameters, the other one with averaged parameters. The different colour lines show the median difference for 11 spectral ranges as shown by the colour bar, for pressures ranging from $10^{-}5$ to 100 bar.

cross sections, especially at low temperatures and high pressures.

In order to simulate the effect of inaccurate broadening parameters, a second set of cross sections was generated for H_2O, assuming "average" broadening parameters. This set was generated as explained in the previous chapter, but assuming that the only broadening agent is hydrogen, and that the broadening parameters are "uncertain". Uncertain parameters were generated assuming no J-dependence on the Lorentzian width (a simple average over J was assumed, $\gamma_L = 0.03\,\mathrm{cm}^{-1}$). A temperature index $\bar{n} = 0.5$ was also assumed for all transitions. Finally, line-by-line transmission and emission spectra for the typical hot Jupiter and warm Neptune were generated using this second set of high resolution cross sections.

Before looking at the changes in the spectra it is instructive to see at which pressures and spectral ranges the difference between these two sets of cross sections is greater. I consider two temperatures (500 and 1500 K) and 11 different spectral ranges, from 0.3 to 50 μm, determined assuming constant width in log space. The median differences, as defined in Equation 4.2.1, are shown in Figure 4.8. It can be seen that changes of up to 35% are observed for both temperatures, in agreement with Hedges & Madhusudhan (2016), but only at the longest wavelengths. In the shortest wavelengths the differences are closer to 10% for $T = 1500$ K and 20% for $T = 500$ K. In addition, it can be seen that the peak of the difference shifts towards lower pressures at longer wavelengths. Interestingly, while at 1500 K the median difference drops at lower pressure at all wavelengths, at 500 K, for wavelengths larger than about 6 μm, the difference remains constant at decreasing pressures. This is because even at the lowest pressures, at low temperatures and small wavenumbers

Table 4.1: Maximum differences between spectra binned at a resolving power of 100 computed using cross sections with accurate and inaccurate broadening parameters for different spectral ranges. The differences are given in parts per millions (ppm), and in units of scale heights H (in brackets, see Equation 4.2.2).

	Transmission		Emission	
Spectral range	Hot Jupiter	Warm Neptune	Hot Jupiter	Warm Neptune
[μm]	[ppm] ([H])	[ppm] ([H])	[ppm] ([H])	[ppm] ([H])
$1-5$	10 (0.2)	15 (0.2)	14 (0.2)	1 (0.01)
$5-10$	13 (0.2)	17 (0.2)	50 (0.8)	15 (0.2)
>10	27 (0.4)	14 (0.1)	100 (1.6)	35 (0.4)

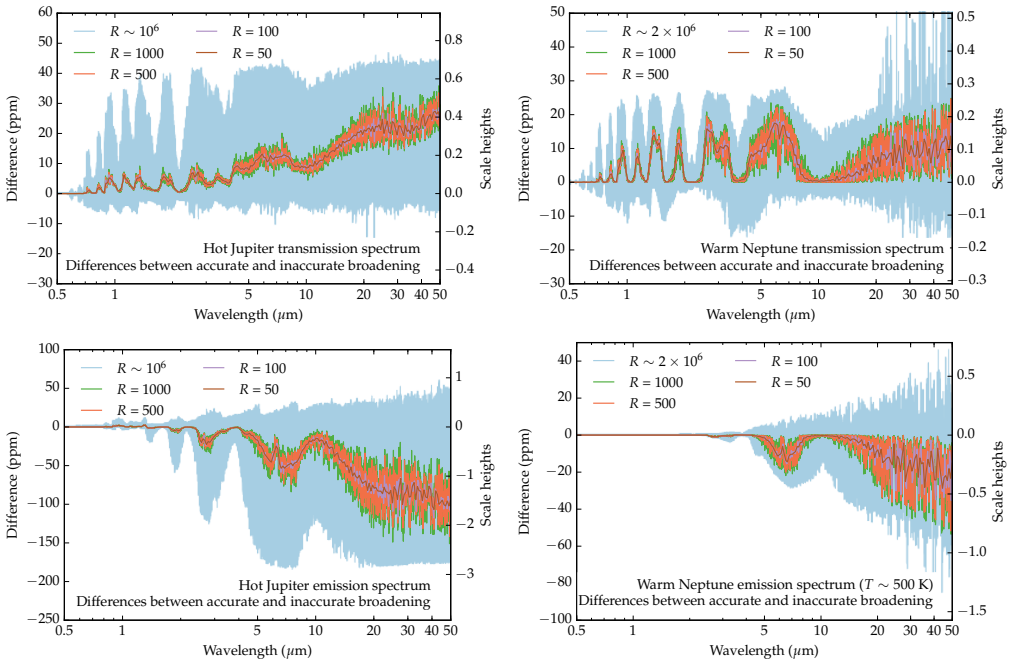

Figure 4.9: Differences between transmission and emission spectra computed using cross sections with accurate and average broadening parameters, binned at several resolving powers from 50 to 1000. The line-by-line difference, corresponding to a resolving power of about 10^6, is also shown in light blue.

(i.e. large wavelenghts) the Voigt width is dominated by the Lorentzian profile (see Figure 3.3 in the previous chapter), so that uncertainties in γ_L will have more impact on the final cross sections. On the other hand, at very low pressures, and at high temperatures and/or high wavenumbers, the Voigt profile is dominated by the Gaussian profile, so that uncertainties in γ_L will have little to no effect on the final cross section.

It is now interesting to analyse how these differences and uncertainties propagate on the modelling of typical exoplanet spectra. Figure 4.9 shows the differences at various resolving powers between the transmission and emission spectra for the typical hot Jupiter and warm Neptune computed with both set of cross sections. Difference of up to 200 ppm,

or, equivalently, 3 scale heights, are observed, but only at the highest resolving power ($R \sim 10^6$). Table 4.1 shows the maximum differences for $R = 100$ and various spectral ranges. The most salient things to notice are:

- Most of the changes are observed at the peak of the absorption features in the spectra
- The differences decrease significantly at lower resolving power, as negative differences are offset by positive differences within each bin. However, once a resolving power of about 100 is reached, a further decrease in resolving power does not reduce these differences significantly.
- Differences increase at larger wavelengths in all cases. This can be easily explained as the absolute differences in the high resolution cross sections increase at larger wavelenghts (Figure 4.8).
- In the case of transmission, similar differences in scale heights are seen between the hot Jupiter and warm Neptune, with slightly larger absolute differences seen for the warm Neptune. For wavelengths larger than 10 μm, a larger difference is seen for the hot Jupiter spectra, likely due to the stronger absorption feature of water at higher temperatures.
- Emission spectra are more affected than transmission spectra, especially for the hot Jupiter. This is because at higher pressures the absolute differences in the cross sections are larger (Figure 4.8), and emission spectra probe higher pressures than transmission spectra (Figure 4.6 and Figure 4.7).
- Significantly larger differences are seen in the hot Jupiter than in the warm Neptune emission spectra. This is mainly because there is significantly more emission of radiation from the hot Jupiter compared to the warm Neptune, so that differences are enhanced.
- Taking as a reference the future systematics expected by JWST instruments for a resolving power of 100 (Table 4.1): Between 1–5 μm, the differences are always below 15 ppm, and therefore slightly less than the expected instrumental systematics ($\approx$ 20 ppm). Similarly, between 5–10 μm, differences are below 17 ppm, and therefore significantly less than the expected instrumental systematics ($\approx$ 40 ppm), apart

from the emission spectrum of the hot Jupiter, where differences of up to 50 ppm are seen.

Although not as dramatic as the differences caused by the use of incomplete line lists, in this section we have seen that the differences caused by the uncertainty in broadening parameters can be relatively strong, and at times comparable with the sort of systematics expected by future instruments.

Lastly, it is worth mentioning that I have only considered the case of water. Uncertainties of broadening parameters of other molecules might lead to different changes in the spectra, but similar differences are still expected. It is also important to note that these biases arise from simply averaging broadening parameters for the appropriate broadening agents. The use of "wrong" broadening agents might have significantly stronger effects. For instance, using air-broadening parameters for the study of H_2/He dominated atmospheres could potentially cause significantly larger biases than those seen here.

4.2.2 Wing cut-offs

Voigt profiles in principle extend to infinity. Real profiles, however, become sublorentzian at a certain distance from the line centre, which is somewhere in between 1 and 30 cm^{-1} (Edwards & Strow, 1991; Birnbaum, 1979). These values are however highly uncertain. The most common approach to solve this problem is to apply an arbitrary wing cutoff at a certain distance from the line centre. The distance can be either fixed, and values between 10 and 100 cm^{-1} are often adopted (Sharp & Burrows, 2007), or can be a multiple of the Lorentzian width (Grimm & Heng, 2015) or of the Voigt width (Hedges & Madhusudhan, 2016). These authors noted that the choice of wing cutoffs can significantly impact the computation of molecular opacities, however an accurate assessment of this effect on transmission and emission spectra of typical exoplanets has not been fully explored yet, but attempts have been made. Grimm & Heng (2015) for example, claim that the line-wing cutoff is the largest source of error. In their work fractional differences of $\sim$10% are seen when using a 500 Lorentzian widths cutoff compared to a fixed cutoff of 100 cm^{-1}, and imply that a similar difference should be present for the computed flux or synthetic spectrum. However, this might not be true. As we saw in the previous section, relative differences of 10–35% in the cross sections caused negligible differences in the transmission and emission spectra of a hot Jupiter, and still relatively small differences in the spectra of a warm Neptune. Here I will try to assess the effect that the choice of wing cutoffs has on transmission

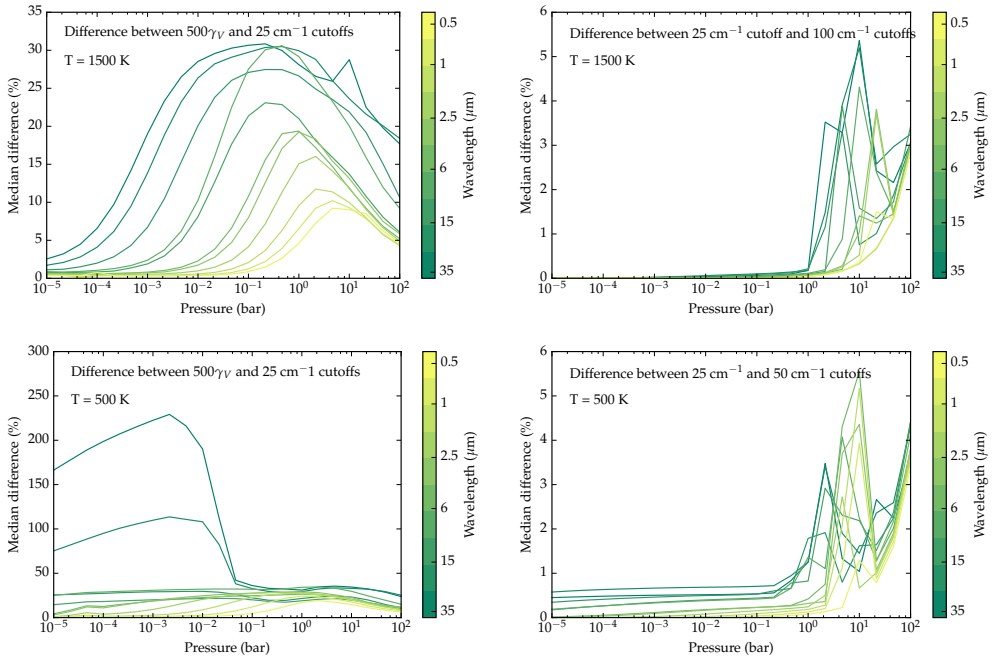

Figure 4.10: Median percentage difference between cross sections computed using an adaptive wing cutoff (500 Voight widths) and two fixed cutoffs (25 and 100 cm^{-1}) at 500 and 1500 K. The different colour lines show the median difference for 11 spectral ranges as shown by the colour bar, for pressures ranging from 10^-5 to 100 bar.

and emission spectra, fully propagating these uncertainties on the final spectra.

In order to assess the uncertainties caused by these different methods several sets of cross sections were computed using different cutoff schemes: 500 Voigt widths cutoff, and 25 cm^{-1}, 50 cm^{-1} and 100 cm^{-1} fixed cutoffs. As before, I used H_2O as an example, using the complete BT2 line list. The assessment of the differences follows the same procedure outlined in the previous section.

Firstly, it is instructive to directly compare the cross sections using the median percentage difference. Figure 4.10 shows the relative difference between the cross sections computed assuming different cutoff schemes, at different pressures and different wavelengths, for $T = 500$ and 1500 K. The strongest differences are seen between the cross sections computed assuming 500 Voigt widths cutoffs and a fixed cutoff (25 cm^{-1}), especially at lower pressures. Stronger differences are seen at the lower temperature ($T = 500$ K), especially at longer wavelengths, where differences of 100 and 200% are seen. These differences almost disappear when cross sections computed using two different *fixed* cutoffs are compared. The right plots show the relative difference at different pressure and wavelengths between cross sections that assume a 25 and 100 cm^{-1} cutoff. Differences of just 3 to 4% are seen

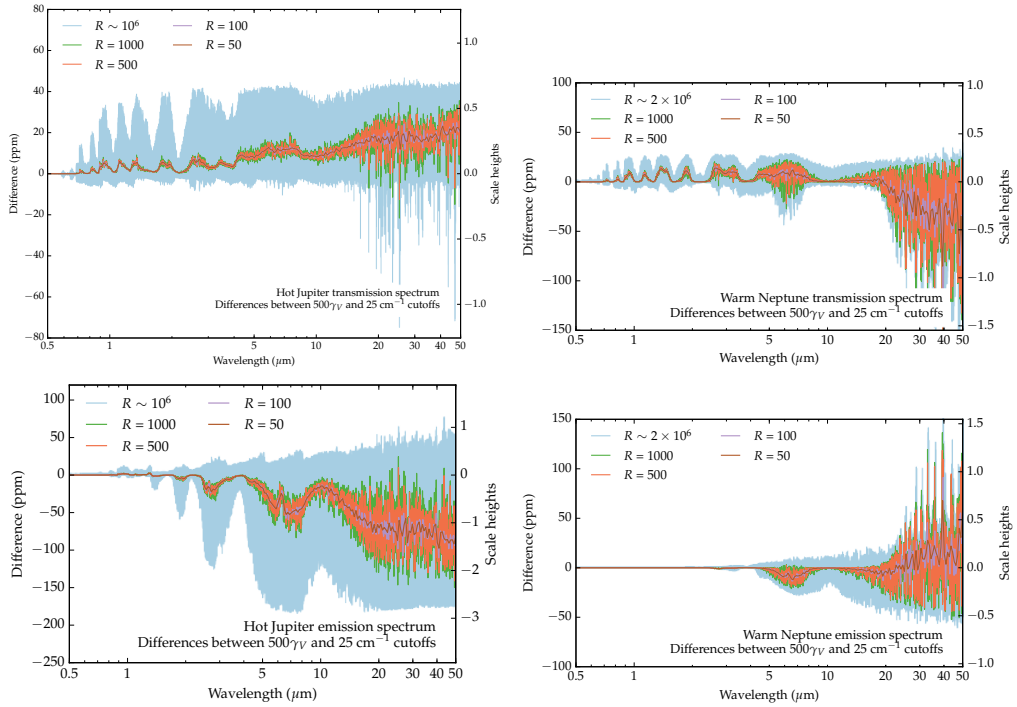

Figure 4.11: Differences between transmission and emission spectra computed using cross sections that assume a 500 Voigt widths ($500\gamma_L$) and 25 cm^{-1} cutoff in the Voigt profile, binned at several resolving powers from 50 to 1000. The line-by-line difference, corresponding to a resolving power of about 10^6, is also shown in light blue.

Table 4.2: Maximum differences for different spectral ranges between spectra binned at a resolving power of 100 computed using cross sections that assume a 500 Voigt widths ($500\gamma_L$) and 25 cm^{-1} cutoff in the Voigt profile. The differences are given in parts per millions (ppm), and in units of scale heights H (in brackets, see Equation 4.2.2).

	Transmission		Emission	
Spectral range	Hot Jupiter	Warm Neptune	Hot Jupiter	Warm Neptune
[μm]	[ppm] ([H])	[ppm] ([H])	[ppm] ([H])	[ppm] ([H])
$1-5$	10 (0.2)	15 (0.2)	25 (0.4)	2 (0.03)
$5-10$	13 (0.2)	15 (0.2)	15 (0.2)	14 (0.2)
>10	25 (0.4)	50 (0.8)	100 (1.6)	60 (1.0)

above 1 bar at both temperatures, while differences smaller than 0.5% are seen at lower pressures. This indicates that while the choice of different fixed cutoffs (between 25 and 100 cm^{-1} for example) causes almost no differences, the choice of an adaptive cutoff (in this case 500 Voigt widths) causes significant differences compared to a fixed cutoff.

The differences of the transmission and emission spectra at different resolving powers, computed using 500 Voigt widths and 25 cm^{-1} cutoffs are shown in Figure 4.11. Table 4.2 shows the maximum differences for $R = 100$ and various spectral ranges. Interestingly, the differences are very similar to what we saw in the previous section. This is not surprising, given that the differences in the high resolution cross sections are also very similar to those

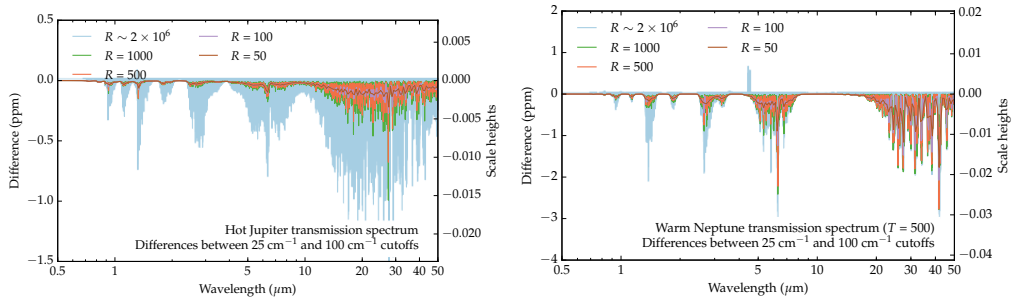

Figure 4.12: Difference between transmission and emission spectra computed using cross sections that assume a 25 and 100 cm^{-1} cutoff.

seen before (compare Figure 4.10 and Figure 4.8). The same list of salient points detailed above is therefore also valid here. The only difference is for the warm Neptune, where at longer wavelengths differences are now slightly stronger. This is because, as noted above, at $T = 500$ K, at wavelength larger than about 15 μm, differences of up to 250% are seen in the high resolution cross sections.

Lastly, Figure 4.12 shows the changes in the spectra using the 25 and 100 cm^{-1} cutoffs. Negligible differences are found: in the case of the hot Jupiter, these are less than a few ppm both in transmission and emission.

In this section we have seen that while adopting different *fixed* line profile cutoffs (25 or 100 cm^{-1}) has little to no effect in the final spectra, the choice of an adaptive cutoff (e.g. 500 Voigt widths) compared to a fixed cutoff (25 cm^{-1}) has some effects in the final spectra. It is noted however that this does not indicate that one method is better than the other, as there is still large uncertainty about where the line cutoff actually occurs (Edwards & Strow, 1991). These differences are very similar to those seen when uncertain broadening parameters were used, and are of the order of 10 to 20 ppm. This is certainly below the $\sim$10% change predicted by Grimm & Heng (2015). Nevertheless, the propagated uncertainties still contribute to the total systematic error of our spectral models.

4.3 Cross section sampling

Computation of line-by-line spectra requires an incredible amount of computing time and large amounts of memory. TauREx can run forward models for transmission and emission spectra using multiple CPUs, but, even when 24 CPUs are used simultaneously, the required time is of the order of a few minutes. This makes line-by-line models impractical for spectral retrievals, especially when Bayesian frameworks are used. Memory requirements are

also a problem, especially when multiple temperatures need to be used. Each high resolution cross section is about 1 to 2 GB in size for each temperature. When multiple molecules need to be considered, and when multiple temperatures are used, the total amount of memory required often exceeds several tens of GB. This makes it often impractical, if not impossible, to simply generate even single spectra.

Cross section sampling is a very simple method allowing us to optimise the generation of low resolution spectra. The cross sections are simply sampled at much lower resolutions, and these are then used to compute transmission or emission spectra, which are eventually binned down[2] to the final resolution. Despite being quite simple, this method is effective when the final resolution of the spectra are relatively low compared to the resolution of the sampled cross sections. Several authors in the literature use cross section sampling (Line et al., 2013; Waldmann et al., 2015b; Benneke & Seager, 2012; Sharp & Burrows, 2007), compared to other techniques such as the *k*-distribution method. However, little work has been done to assess the impact of this method in the final spectra, and to understand how this approximation can bias spectral retrievals.

In this section this issue is analysed in detail, for both transmission and emission spectra, using simulated spectra for the same typical hot Jupiter and warm Neptune used in the previous sections. The impact of cross section sampling is assessed in the following way: firstly, the line-by-line spectra are binned down to several resolving powers ranging from 10 to 1,000 to obtain the final low resolution spectra. Note that the line-by-line spectra have a native resolving power of the order of 10^6. These low resolution spectra are considered to be the "ground truth", and all spectra generated at the same resolving powers but using sampled cross sections will be compared to these ones. High resolution cross sections are then sampled at several lower resolving powers, ranging from 1,000 to 100,000. Note that the cross sections are not binned down to this new resolutions, but *sampled*. In other words, they are interpolated to a new wavenumber grid with constant resolving power (i.e. constant $\nu/d\nu$), which has much fewer points than the original grid. Transmission and emission spectra are then computed using these sampled cross sections, and binned down the same resolutions of the "ground truth" spectra. The aim is to find the minimum sampling resolution required such that the final binned spectra have negligible differences compared to the "ground truth" spectra.

[2]The binned spectrum is calculated by taking the average of each spectral bin

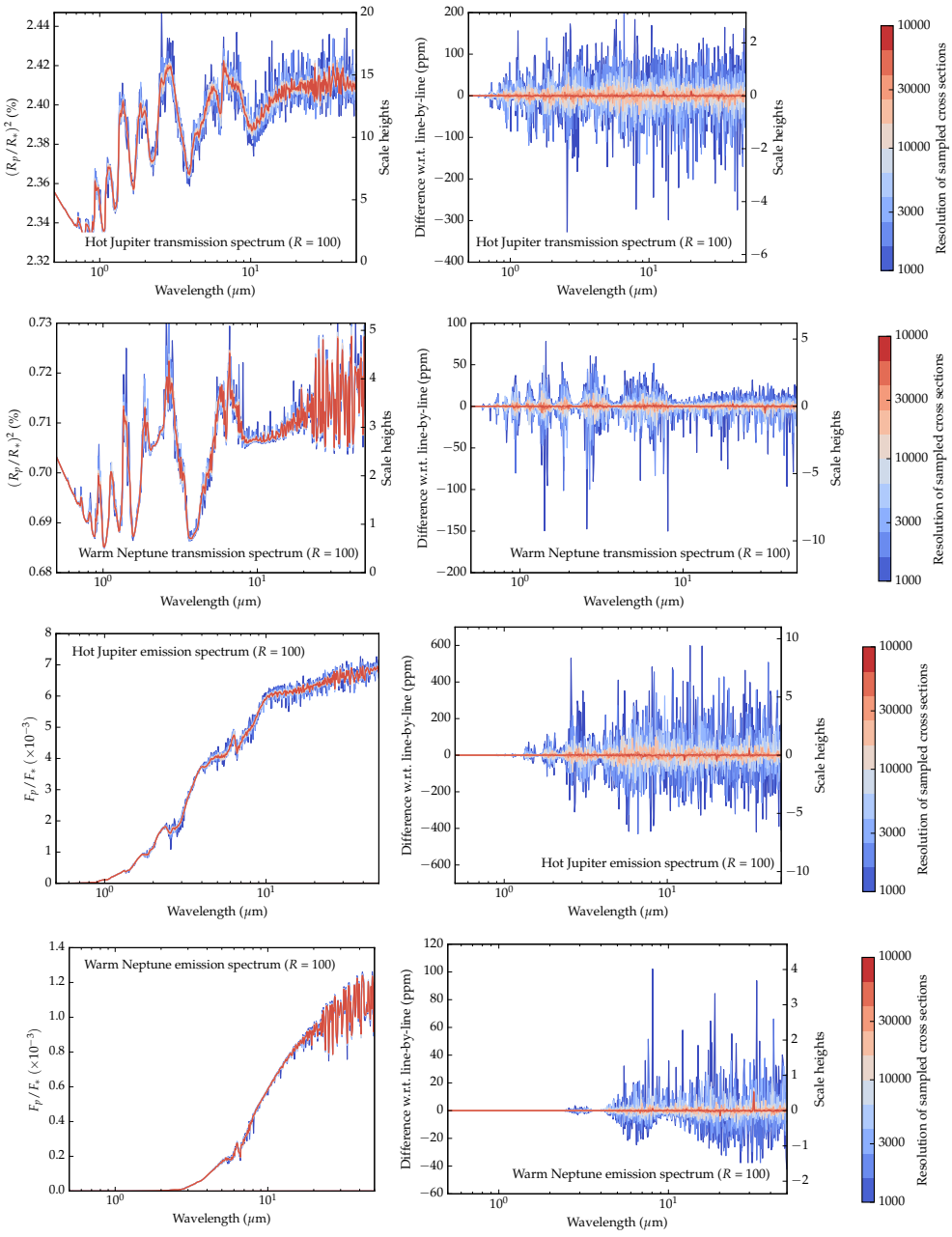

Figure 4.13: Left: transmission and emission spectra of a typical hot Jupiter and warm Neptune with a final resolving power of $R = 100$, computed with cross sections sampled at different resolving powers between 1,000 and 100,000, as shown by the colour bar on the left hand side. The darkest red line shows the line-by-line spectrum binned at $R = 100$. Right: differences of the spectra shown on the left with respect to a line-by-line spectrum binned at the same resolution.

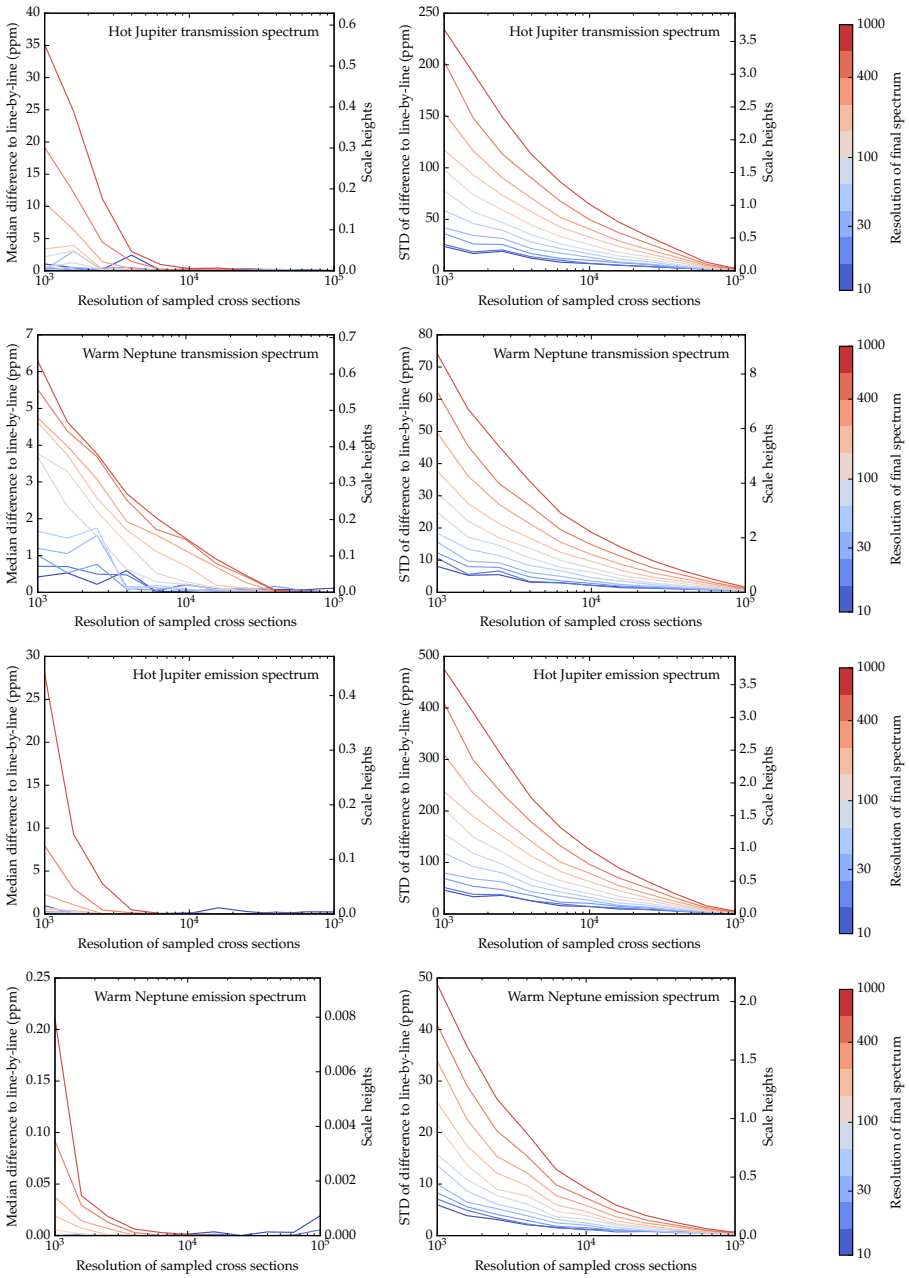

Figure 4.14: Median (left plots) and standard deviation (right plots) of the difference between a line-by-line spectrum binned at final resolving powers from 10 to 1000 (as shown by the colour bar), with respect to the same spectra obtained using sampled cross sections (with resolving powers shown in the x-axis), binned to the same final resolutions. The spectra considered are the emission and transmission spectra of a typical warm Neptune and hot Jupiter

Figure 4.13 shows the effects that different sampling resolutions have on typical transmission and emission spectra binned at a resolving power of 100. It can be seen that, while extremely large differences (150 to 600 ppm) are seen when very low sampling resolutions are used (i.e. $R = 1000$), relatively low scatter is seen when sampling resolving power of 10,000 and above are used. It is also important to note that, while the differences seen in the previous sections were not symmetric and varied a lot as a function of wavelength, the scatter produced due to cross section sampling is highly symmetric, and relatively uniform with wavelength, especially in transmission. It is therefore useful to assess these differences with two quantities: the standard deviation and median of the differences across the entire spectral range.

Figure 4.14 show these two quantities plotted as a function of resolution of the sampled cross sections, and as a function of the resolution of the final spectrum. We can see that the median is about one order of magnitude less than the standard deviation, and is much less affected by the sampling resolution than the standard deviation. For all spectral resolutions (i.e. the resolution of the final spectrum), the median drops very quickly to less than a couple of ppm for sampling resolving powers above a $10,000$. This is because the differences are uniformly scattered around zero, so that the differences within each spectral bin are offset. On the other hand, the standard deviation of decreases more uniformly, approaching < 1 ppm at sampling resolving powers of 10^5 for all spectral resolutions.

Interestingly, the warm Neptune spectra are less affected than the hot Jupiter spectra. In the case of transmission, this is because large parts of the water lines are hidden below the continuum from collision induced absorption of H_2–H_2 and H_2–He pairs, while in the case of emission it is simply because the emitted radiation in the warm Neptune atmosphere is much weaker than in the hot Jupiter atmosphere. Moreover, for the hot Jupiter the scatter seems worse in the case of emission than transmission, while, for the warm Neptune, the opposite is true. Direct comparison of emission and transmission is however difficult, as the amount of absorption seen in emission, and therefore the degree of scatter, relates in a highly non-linear way to the temperature-pressure profile. Nevertheless, we note that the potential negative effects of cross section sampling can be reasonably high even in emission, where the scatter would in principle expected to be lower, as the pressures probed are higher, and the line profiles broader (i.e. “dampened”).

For a typical resolving power of 100, using cross sections with resolving power of

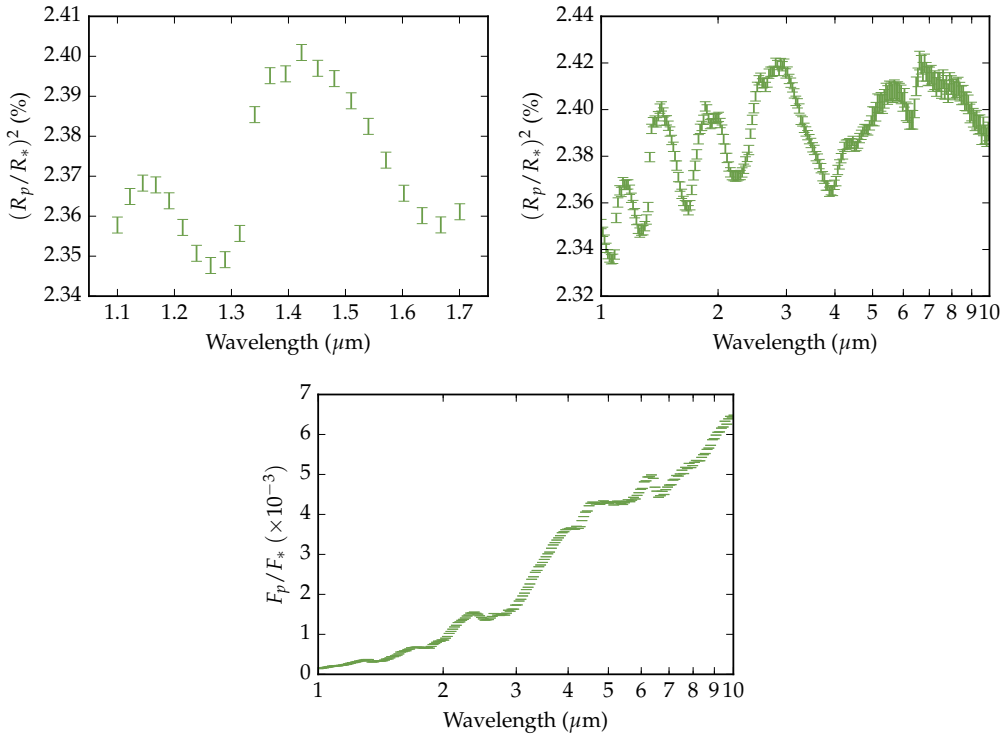

Figure 4.15: Simulated transmission and emission spectra of the typical hot Jupiter. From top left, clockwise: WFC3 transmission spectrum; JWST transmission spectrum; JWST emission spectrum.

10,000 is often considered to be a good compromise between accuracy and computational time. At this sampling resolution, while the median difference is always below 1 ppm, the average standard deviation is relatively large. In the case of transmission spectra, the standard deviation is 25 and 6 ppm for the hot Jupiter and warm Neptune respectively. In the case of emission spectra, it is 40 ppm and 3 ppm for the hot Jupiter and warm Neptune spectra respectively. Note that these values are relatively high compared to the typical systematic expected from future instruments. For lower resolutions, such as those from WFC3/HST, the same standard deviations drop to 16, 5, 30 and 2 ppm respectively.

Given these relatively large differences, and the fact that several spectral retrievals presented in the literature use this sort of approximation (e.g. Sharp & Burrows, 2007; Line et al., 2013; Waldmann et al., 2015b), it is useful to investigate how retrievals are affected by cross section sampling. Given that the observed differences are uniformly scattered around zero, the χ^2 of a spectral fit is expected be only weakly affected by these differences, such that the biases in atmospheric retrievals are also expected to be relatively weak.

For simplicity, I only consider the transmission and emission spectra of the typical hot

Jupiter and simulate synthetic observations in two spectral ranges, corresponding to what is achievable with WFC/HST and JWST. Spectra simulated for WFC/HST cover the spectral range 1.1–1.7 μm, with $R = 50$ and 20 ppm error bars; spectra simulated for JWST cover the range 1–10 μm, with $R = 100$ and 20 ppm error bars between 1–5 μm and 40 ppm between 5–10 μm. The same error bars were assumed both in the transmission and emission spectra. The emission spectrum was only simulated and retrieved for the JWST spectral range. Note that the large spectral range of the JWST synthetic spectrum can only be obtained by combining multiple observations with different instruments on-board JWST. The synthetic observations are shown in Figure 4.15. Note also that error bars have simply been "stuck" on top of the data points, and no Gaussian noise was added. This was done to remove the effect that a specific noise instance (or noise draw) has on the spectrum retrieval. These spectra are computed from the same line-by-line spectra generated in the previous sections.

The effect of cross section sampling is then investigated by retrieving these spectra with TauREx, using three sampling resolving powers ($R = 5000, 1000, 15000$). Figure 4.16 shows the posterior distributions of the retrieved atmospheric parameters for both spectra seen in transmission, using these different cross sections. It can be seen that for the WFC3 spectrum very little differences are seen between different retrievals. In all cases, the atmospheric parameters are within one sigma of the input value (shown with blue vertical lines). The posterior distributions are also well centred at the input value. The only small difference is seen for the retrieval of the temperature using the $R = 10,000$ cross section, where the retrieved temperature is slightly higher than the other two retrieved values, but always within one sigma of the true state.

Slightly larger differences are seen in the JWST spectrum. In this case the retrieved abundances of temperature and radius are more affected by low cross section sampling. The retrieved abundance of water, however, seems unaffected. The worse case is clearly for $R = 5,000$, where the retrieved values of the temperature and radius are 3.1 and 3.4 sigma away from the true state respectively. Better results are seen with $R = 10,000$: the same retrieved values are 1.9 and 1.6 sigma away respectively. Lastly, for $R = 15,000$, the retrieved state is always within one sigma of the true state.

Stronger differences are seen in the retrieval of the JWST emission spectrum. The retrieved water abundance is shown in Figure 4.17. It can be seen that, while for $R = 10,000$ and 15,000 the retrieved state is within one sigma of the true state, in the case of $R = 5,000$

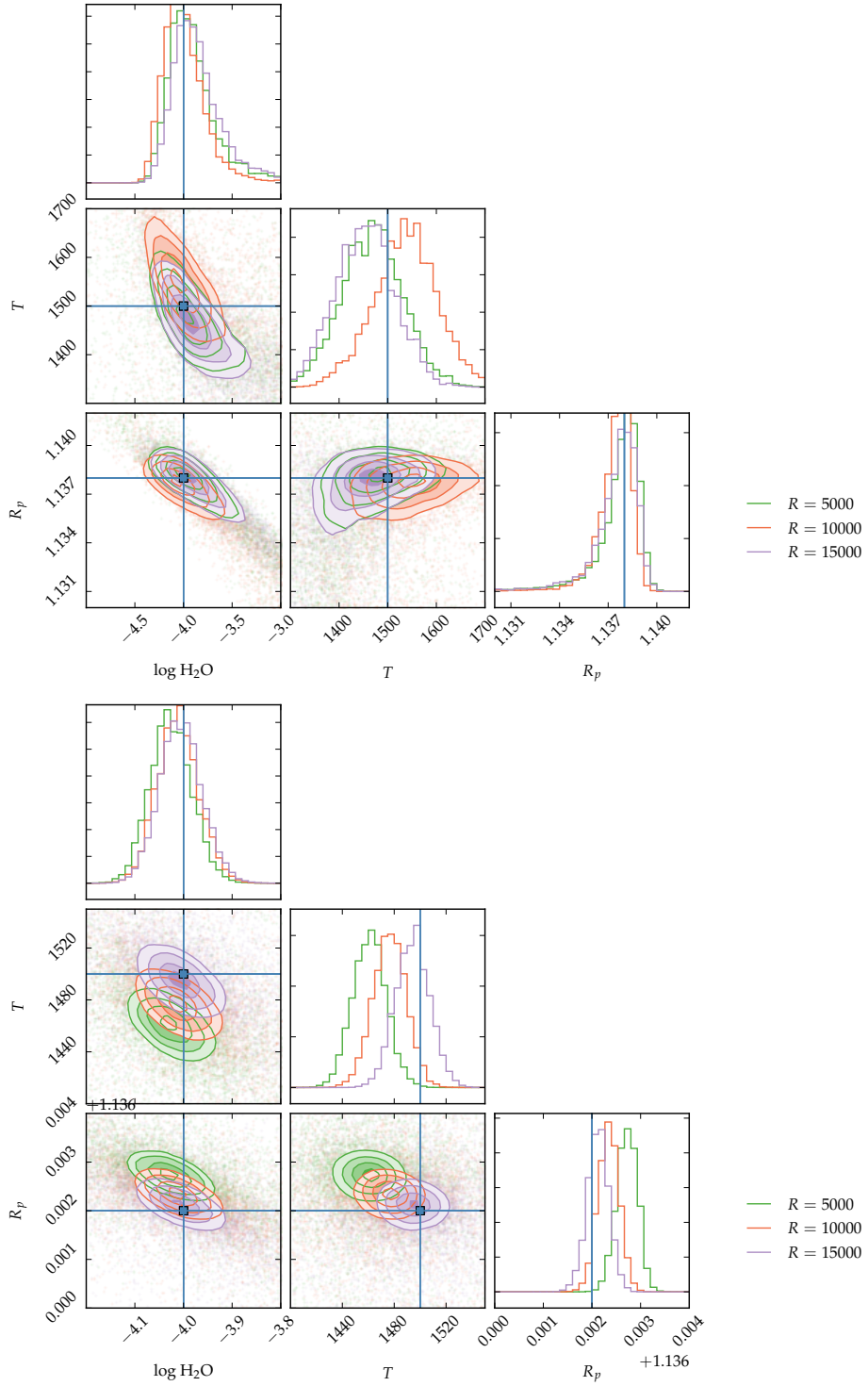

Figure 4.16: Posterior distributions of the retrieved of the H_2O abundance, temperature and radius for the WFC3 (top-left plot) and JWST (top-right plot) transmission spectra shown in Figure 4.15. The different colours correspond to separate retrievals using cross sections sampled at different resolutions, as shown by the legend. The true state, or input values for each parameter, are shown with blue lines.

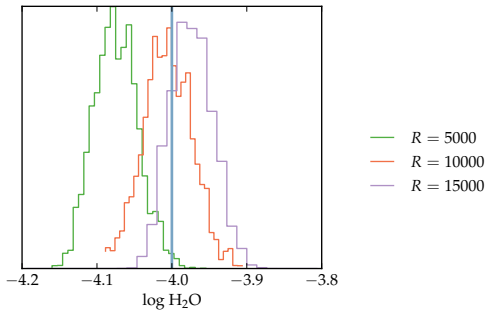

Figure 4.17: Posterior distributions of the retrieved of the H_2O abundance for the JWST emission spectrum shown in Figure 4.15. The different colours correspond to separate retrievals using cross sections sampled at different resolutions, as shown by the legend. The true state, or input values for each parameter, is shown with a blue line.

we are about 2 sigma away. This is not surprising, given that the strongest differences caused by cross section sampling are seen for the emission case of the hot Jupiter (see Figure 4.13).

From these results, it is clear that while a cross section sampling resolving power of $10,000$ is sufficient to model WFC3 spectra, it is inadequate to model larger wavelength range spectra, such as those expected by JWST. In this case, higher resolving powers ($R \sim 15,000$) should instead be used.

4.4 Correlated-*k* approximation

We already saw that the *k*-distribution method allows to reduce the computational time required to model a spectrum. This method takes advantage of the fact that the spectral transmittance is independent of the ordering of the absorption coefficients. By integrating the spectral transmittance over the cumulative distribution function of the absorption coefficients within each spectral bin, we can significantly decrease the number of sampling steps, as this smooth function can be sampled with far fewer points. This method assumes that the line strength distributions within each spectral bin are vertically correlated across the atmosphere. In real atmospheres, however, this correlation is not strictly maintained. In this section I will explore how this assumption affects the observation of the typical exoplanets already studied in the previous pages.

The *correlated k*-distribution method assumes that, for each molecule, the absorption coefficients within a given bin are *correlated* in frequency space across all atmospheric layers. This is because, in order for the method to be technically correct, the way in which the *k* values are reordered within each bin to create the cumulative distribution function should be the same at all atmospheric levels. It is easy to see that in the case of a single absorp-

tion line a one-to-one correlation between the ordered absorption coefficients is maintained, even when the line profile changes across the atmosphere due to pressure and temperature inhomogeneousity. The strongest absorption will always occur at the line-centre at all pressure levels. However, when multiple lines are present, the correlation is not automatically maintained. In this case, the *k*-distribution method is assumed to be a good approximation only in the *weak-* and *strong-line limits*. In the *weak-line limit* the path length and/or absorption coefficient is small, so that the transmittance can be expressed by an exponential function without integration over ν. This is also also referred to as the *grey approximation*. In the strong-line limit, the assumption is that the width of the profiles is small compared to the line spacing, so that the absorption of radiation through the wing regions dominates, allowing the transformation of a non-homogeneous path into a homogenous one. This is also referred to as the *scaling approximation*. See Liou, K N (2002) for details and full derivations of these two approximations.

According to these two approximations, in the case of real atmospheres correlation might be well maintained in the peaks and troughs of the absorption lines, but this strict correlation is not likely maintained over the entire range of absorption coefficients' strengths. The emergence of new spectral lines, or large shifts in line strength within hot bands, can potentially produce spectrally uncorrelated changes in *k* distribution (see also Lacis & Oinas, 1991). The easiest way to assess the induced errors is to simulate typical exoplanet spectra using the *k*-distribution method, and compare them with the same spectra computed with a line-by-line integration, and binned down at the same resolution.

To begin with, it is firstly useful to see the typical correlation errors caused by the lack of strict correlation between atmospheric layers. Two correlations are explored: one between two layers with different pressures (0.1 and 1 bar) but equal temperature (T = 1500 K), and one between two layers with different temperatures (1000 and 1500 K) but equal pressures (P = 1 bar). In both cases I consider the absorption cross section of water between 4996 and 5001 cm^{-1}. Following Lacis & Oinas (1991), in both cases the absorption cross section of the second layer is recomputed by mapping the absorption cross section of the first layer via their respective *k* distributions. This is done by taking the absorption coefficient at a frequency ν of the cross section of the first layer, and computing its corresponding *g* value in the *k*-distribution. Then, the "mapped" cross section value of the second layer at the same frequency ν is calculated by taking the value of *k* from the *k*-distribution

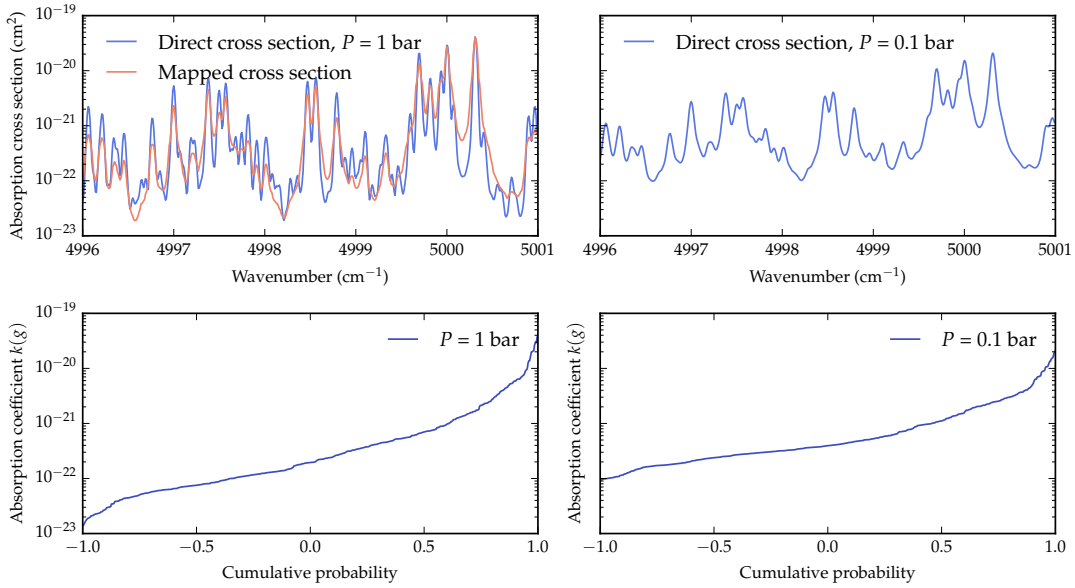

Figure 4.18: Spectral correlation for transmission across two pressure layers, with pressures of 1 and 0.1 bar. Top plots: the blue lines show the absorption cross section of H_2O between 4996 and 5001 cm^{-1} for $T = 1500$, and $P = 1$ bar (left) and $P = 0.1$ bar (right). Middle plots: the blue lines show the cumulative distribution $g(k)$ for the two cross sections shown in the top panels. The orange line in the first plot is the mapping of the cross section for $P = 0.1$ bar onto the $P = 1$ bar cross section, using the k distributions for both pressures.

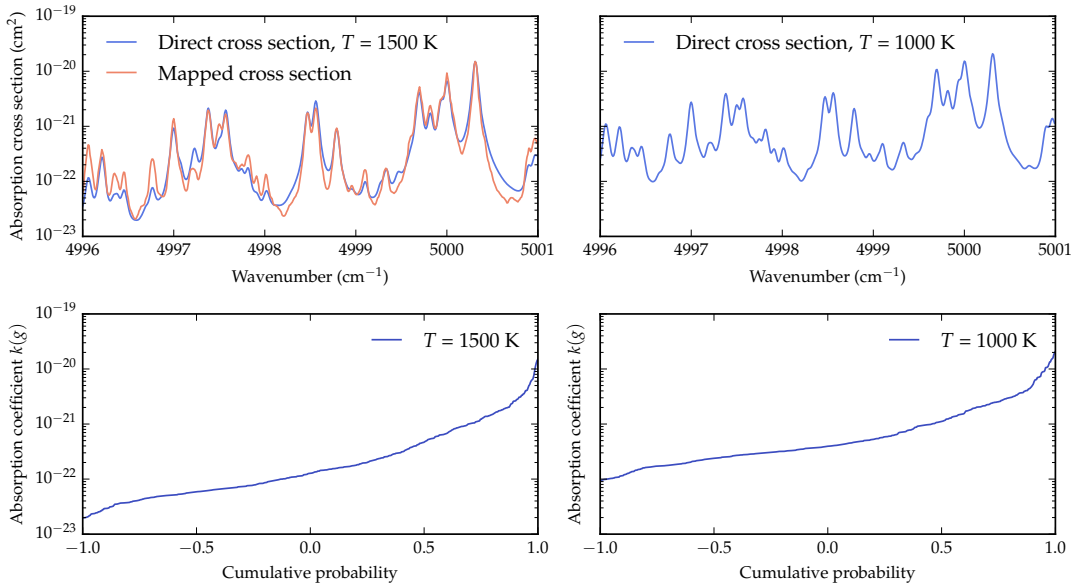

Figure 4.19: Spectral correlation for transmission across two temperature layers, with temperatures 1000 K and 15,000 K. Caption as in Figure 4.18

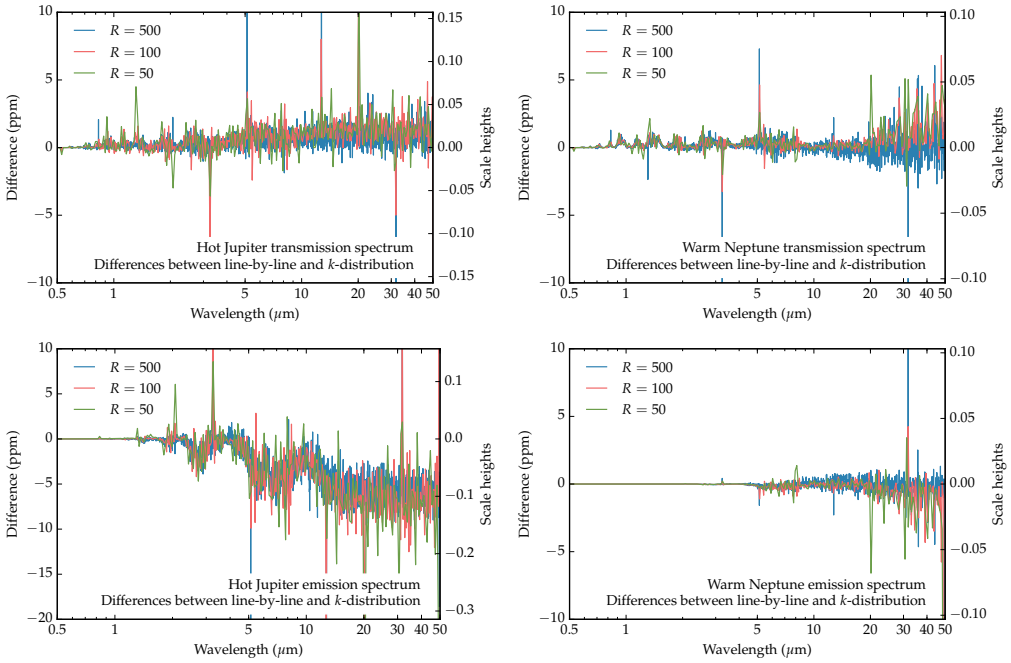

Figure 4.20: Differences in ppm between spectra computed using line-by-line integration and k-distribution tables, for different resolving powers ($R = 50, 100, 500$). The differences arise from both the "correlated" k-approximation, and the quadrature sampling of each k-distribution. Transmission and emission spectra of the typical hot Jupiter and warm Neptune are considered.

of the second layer that shares the same g value computed before.

Figure 4.18 and Figure 4.19 show the direct cross sections for the two pressures (or temperatures) together with their k-distributions, and the mapped cross section (in orange). Under strict correlation, the mapped and direct cross sections should match exactly. It can be seen that, although small differences are present due to the lack of perfect correlation, relatively small differences are seen, especially considering that extreme pressure and temperatures are considered.

While it is instructive to see these plots to understand where the main source of error in the k-distribution method comes from, it is more interesting to assess how this approximation propagates on transmission and emission spectra of exoplanet atmospheres. Following the same method outlined in the previous sections, spectra computed with k tables are directly compared with spectra computed with line-by-line integration. The typical hot Jupiter and warm Neptune, seen in transmission and emission, were taken into account. Three typical resolving powers were used: 50, 100, and 500.

The resulting differences are shown in Figure 4.20. Transmission spectra computed using k-coefficients and those computed using LBL integration have maximum differences

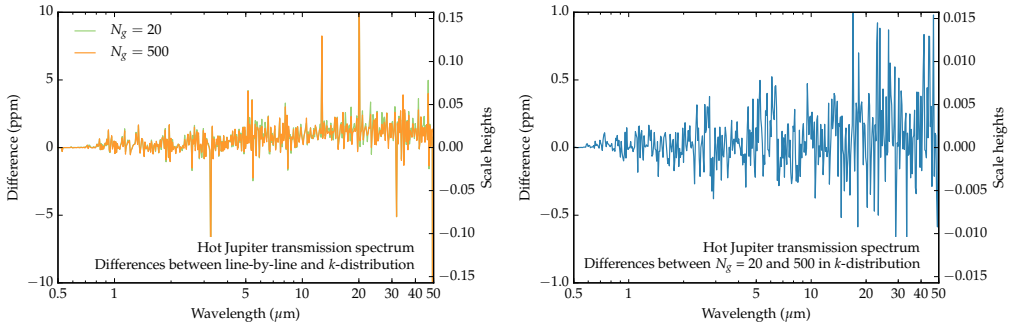

Figure 4.21: Left: difference between spectra with resolving power of 100 computed with k-distributions adopting 20 and 500 Gaussian quadrature points (green and orange lines respectively) and line-by-line spectra binned at the same resolution. Right: Difference between the two spectra computed using the k-distribution method adopting 20 and 500 quadrature points.

of only a few ppm at all resolutions, in both the hot Jupiter and warm Neptune spectra. Interestingly, a change in resolution does not affect the scatter, indicating that the magnitude of the difference is independent of the resolution. In addition, as we saw for the cross section sampling case, the scatter is quite symmetric, and it is slightly lower at smaller than at larger wavelengths. In all cases, the median of the difference is less than 1 ppm, suggesting that the effect on spectral retrievals will be insignificant.

Slightly larger differences are seen in the emission spectrum of the hot Jupiter. In this case, the scatter is not symmetric, and tends to increase at larger wavelengths, where differences of up to 15 ppm are seen. This, however, is not seen in the emission spectrum of the warm Neptune. Such differences are likely due to the approximation of the spectrum of the star within the spectral bin: TauREx takes the face value of the star spectrum at the bin centre wavelength, while an integrated average within the bin should be computed. The induced errors are, however, significantly smaller than the expected systematics in a real astrophysical scenarios, and are, in all cases, unlikely to bias spectral retrievals.

In order to confirm that the cause of this difference is the lack of perfect correlation of absorption coefficients across the atmosphere, I investigated the effect of the choice of the number of Gaussian quadrature points, N_{quad}, used to sample each k-distribution. The k-distribution computed as part of this thesis, and used by TauREx, assume 20 gaussian quadrature points. This is often considered to be a good compromise between computational time and uncertainties (e.g. Irwin et al., 2008). The transmission spectrum computed with this k-tables was compared with one computed with a k-table that uses 500 quadrature points. Both spectra were then compared to the line-by-line integration. Figure 4.21

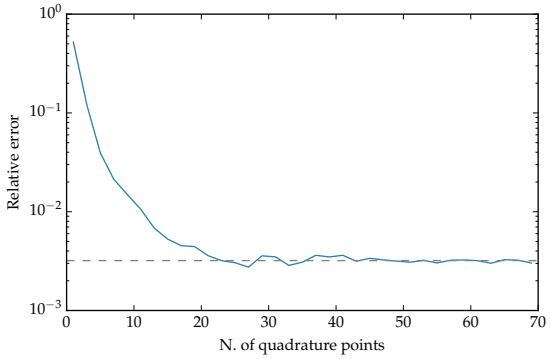

Figure 4.22: Relative error in the calculation of the transmittance using the k-distribution method, as a function of the number of quadrature points used to sample the k-distribution of each bin.

shows the differences between these two spectra. The plot on the left shows the differences between these spectra and the line-by-line spectra, while the plot on the right shows the differences between the spectra computed using $N_{\text{quad}} = 20$ and 500. We can clearly see that the difference between 20 and 500 quadrature points is negligible (less than about 0.3 ppm) compared to the difference with respect to the line-by-line spectra (less than 4 ppm). This is also confirmed by looking at Figure 4.22, showing the relative error in the calculation of the transmittance as a function of the number of quadrature points. After about 20 points, the relative error plateaus at about 0.3 %.

In this section the uncertainties induced by the k-distribution method were assessed and propagated into the modelling of exoplanetary atmospheres. It was shown that the assumption of correlation between atmospheric layers is the main source of uncertainty, and it is therefore unavoidable, while the use of a larger number of quadrature points used to sample each k-distribution plays a relatively small role. Differences, however, were found to be small in all cases (less than 10 ppm even at the larger wavelengths), and much smaller than the expected systematic errors of current and future observations. These uncertainties are therefore highly unlikely to affect and bias atmospheric retrievals in any way.

These results also showed that the k-distribution method gives significantly better results compared to cross section sampling. It should therefore be the preferred method to model, and retrieve, exoplanetary spectra.

4.5 Summary and conclusions

In this chapter I investigated the effects that different approximations in the computation of absorption cross sections have in the final spectra. When these uncertainties were higher

than the expected sensitivity of current or future observations, I investigated their effect on atmospheric retrievals. Five effects were investigated: the use of incomplete line lists; the use of inaccurate broadening parameters; the different choice of wing cutoff in the Voigt profile; the use of cross section sampling; the use of the *k*-distribution method. In summary, I found that:

- The strongest differences are seen when incomplete line lists are used. I compared the 1500 K CH_4 spectrum computed with the STDS line list (the "incomplete" line list, with $\sim$ 9 million transitions) and the YT10to10 line list (the "complete" line list, with $\sim$ 10 billion transitions) and found significant differences of 100–150 ppm across the entire spectrum. The TY10to10 spectrum shows greater absorption, and includes more bands in the short wavelength range. The effects of using the incomplete line list to retrieve a spectrum generated with the more complete one are substantial: the abundance is underestimated and the radius overestimated. The retrieved uncertainties are also significantly underestimated.

- The effects of using inaccurate, or "average", broadening parameters and the use of different Voigt width cutoffs to compute the absorption cross sections, lead to similar differences in the final emission and transmission spectra. These differences are usually small (a few tens of ppm) for the range of resolving powers probed by current and future observations ($R \sim 100$), and always smaller than the expected uncertainties. They are therefore unlikely to cause noticeable biases in the atmospheric retrieval.

- For typical resolving powers of $R \sim 100$, the use of sampled cross section to compute transmission and emission spectra was found to lead to significant differences with respect to a line-by-line spectrum, especially when the sampling resolving power is below 10,000. These differences are however highly symmetric across the entire spectral range, and have therefore a relatively small impact on the atmospheric retrieval. I found that while a sampling resolving power of 10,000 was sufficient to retrieve WFC3/HST spectra, a resolving power of at least 15,000 is needed to model JWST spectra.

- The *k*-distribution method was found to outperform the cross section sampling both in terms of accuracy and computational speed. The correlated-*k* approximation was found to lead to differences of only a few ppm with respect to a line-by-line spectrum.

These differences will therefore not bias atmospheric retrievals. I also found that the use of 20 Gaussian quadrature points is sufficient to achieve good accuracy. As such, the correlated k-method should be the technique of choice to model and retrieve exoplanet spectra.

Chapter 5

Exploring biases in the retrieval of the C/O in hot Jupiter atmospheres

With a scheduled launch in October 2018, the James Webb Space Telescope (JWST) is expected to revolutionise the field of atmospheric characterisation of exoplanets. The broad wavelength coverage and high sensitivity of its instruments will allow us to extract far more information from exoplanet spectra than what has been possible with current observations. In this chapter, I will investigate whether current retrieval methods will still be valid in the era of JWST, exploring common approximations used when retrieving transmission spectra of hot Jupiters, with particular emphasis on the retrieval of the carbon-to-oxygen ratio (C/O). I use 1D photochemical models to simulate a typical hot Jupiter cloud-free atmospheres and generate synthetic observations for a range of C/O. Then, I retrieve these spectra using TauREx, adopting two methodologies: one assuming an isothermal atmosphere, and one assuming a parametrised temperature profile. Both methods assume constant-with-altitude abundances. I found that we can easily differentiate between C/O $>$ 1, C/O = 1, and C/O $<$ 1, as the atmospheric chemistry and resulting spectra change drastically at the C/O = 1 threshold, while determining tighter constraints is more difficult. I also found that the isothermal approximation biases the retrieved parameters considerably, overestimating the abundances by about one order of magnitude. On the contrary, the use of a parametrised temperature profile enables the retrieval of atmospheric abundances within 1 sigma in almost all cases.

Interestingly, I also found that using the parametrised temperature profile I could place tight constraints on the temperature structure. This opens the possibility to characterise the temperature profile of the terminator region of hot Jupiters using transmission spectra.

Lastly, I found that assuming a constant-with-altitude mixing ratio profile is a good approximation for most of the atmospheres under study.

The methods and results presented in this chapter are published in Rocchetto et al. 2016 and are here reproduced with permission from the publisher.

5.1 Introduction

With the imminent launch of the JWST, it has become fundamental to assess whether the current methods used to interpret exoplanetary spectra will still be valid when higher quality datasets will be available. In here I aim to answer, in part, this question, exploring the biases induced by common assumptions used in atmospheric retrievals, with particular focus on the retrieval of the atmospheric C/O.

One of the major limitations of current observations is the limited wavelength coverage. The best quality datasets, which led to the confident detection of water vapour in several hot Jupiters and warm Neptunes, have been mainly obtained with the Wide Field Camera 3 onboard HST, covering the spectral range 1.1–1.7 μm. Nevertheless, it is at longer wavelengths that most roto-vibrational transitions of molecular species occur. While the Spitzer Space Telescope has given some insight into the longer wavelength regime to several tens of close-in hot Jupiters, the data have relatively large uncertainties, and they are mostly only photometric measurements. Significant advances in the field of atmospheric characterisation can therefore only happen if high quality observations extending to the longer wavelength regime are obtained.

In this scenario, the James Webb Space Telescope (JWST) will undoubtedly revolutionise the field of exoplanetary atmospheres, addressing two major problems affecting current observations: wavelength coverage and instrument sensitivity. With a scheduled launch for 2018 October, the large spectral coverage (0.7–28 μm) covered by its multiple instruments, combined with high sensitivity and high degree of instrumental characterisation and calibration, will ensure a significant advance in atmospheric characterization (Beichman et al., 2014; Cowan et al., 2015; Batalha et al., 2015; Greene et al., 2016; Barstow et al., 2015; Barstow & Irwin, 2016).

As we saw in the previous chapters, atmospheric spectra of transiting exoplanets in a broad spectral range will enable us to constrain the abundances of different molecular species, the temperature structure of the atmosphere, and the presence or absence of clouds and hazes. In the case of warm H/He dominated atmospheres one of the key elemental

ratio that we aim to constrain is the carbon-to-oxygen ratio (C/O). Such measurements will enable us to distinguish between different formation and migration scenarios, so far poorly constrained (Öberg et al., 2011; Madhusudhan et al., 2011b; Ali-Dib et al., 2014; Thiabaud et al., 2015). While transmission and emission spectra do not provide direct constraints on the elemental abundances, the measurement of the absolute abundances of O-bearing and C-bearing molecules will provide some constraints on the C/O ratio. In particular, the excess carbon and oxygen not locked in CO will form either oxygen-bearing molecules such as H_2O in atmospheres with C/O < 1, or, in atmospheres with C/O > 1, carbon-rich species such as HCN, C_2H_2 and CH_4 (Madhusudhan, 2012; Moses et al., 2013a,b; Venot et al., 2015). Determining the atmospheric abundances of these gases in hot Jupiters with high accuracy is therefore paramount and JWST will give us direct access to absorption features of these molecules both in emission and transmission.

Determining the absolute abundances of atmospheric gases from atmospheric spectra requires the use of retrieval methods, as discussed in Chapter 2. Atmospheric retrieval techniques are now commonly used to infer the properties of exoplanetary atmospheres, including molecular abundances and temperature profiles (Madhusudhan & Seager, 2009; Madhusudhan et al., 2011a; Benneke & Seager, 2012, 2013; Lee et al., 2011; Line et al., 2012, 2013, 2014; Irwin et al., 2008; de Wit & Seager, 2013; Waldmann et al., 2015a,b). These tools enable us to fully map the likelihood space of atmospheric models, and to place upper limits and constraints on the abundances of molecules and temperature profiles.

The lack of high signal-to-noise and broad wavelength coverage observations have however led current retrievals and forward models to make several assumptions and approximations to reduce the parameter space. The forward model included in most retrieval methods is a 1D radiative transfer model (see Chapter 2 and e.g. Brown, 2001; Liou, K N, 2002; Seager, 2011; Tinetti et al., 2012; Hollis et al., 2013), implementing opacity cross sections for the major molecular absorbers, Rayleigh scattering and collision induced absorption. Transmission spectra are usually retrieved assuming constant-with-altitude temperature and molecular abundances. This might be a fair approximation when probing narrow wavelength ranges, but can lead to significant biases when larger wavelength ranges are probed. *One of the pressing questions we are facing today is whether these assumptions will still be valid in the era of JWST.*

In this chapter, I aim to address these issues. I study the biases and degeneracies of

atmospheric retrievals of high quality, broad wavelength range transmission spectra of hot Jupiters, such as those that will be obtained with instruments onboard JWST. I apply and compare different retrieval approaches to synthetic observations for a range of hot atmospheres with different C/O computed using photochemical models, and study the biases of common assumptions used in today's retrievals.

This study aims at answering the following questions:

a) Are our retrieval approaches and forward models appropriate for the high signal-to-noise, and broader wavelength range spectra expected from future facilities such as JWST?

b) Can we confidently retrieve absolute molecular abundances and infer the C/O ratio?

In Section 6.2 I describe the chemical and radiative transfer models used to generate the synthetic transmission spectra. I also present the JWST synthetic observations, and describe the two retrieval approaches used to interpret these synthetic observations. In Section 6.3 I describe qualitatively the simulated transmission spectra and present the results of the retrievals. In Section 6.4 I discuss the results, and in Section 6.5 I summarise the main conclusions of this study.

5.1.1 Acknowledgments

This work, published in Rocchetto et al. 2016, has been possible thanks to the additional contributions of other authors (I.P. Waldmann, O. Venot, P.O. Lagage and G. Tinetti). In particular, O.Venot computed the chemical models, P.O. Lagage calculated the noise in the JWST observations, I.P. Waldmann assisted with the atmospheric retrieval and the interpretation, and G. Tinetti gave invaluable help in setting the objectives of this study.

5.2 Method

5.2.1 Chemical models

The 1D atmospheric chemical models were generated using the photochemical model developed for hot atmospheres (Venot et al., 2012, and references therein). These models have been used to study exoplanets (Venot et al., 2014; Agúndez et al., 2014; Venot et al., 2015; Venot & Agúndez, 2015; Tsiaras et al., 2016) as well as solar system giant planets (Cavalié et al., 2014; Mousis et al., 2014). The chemical scheme has been developed with combustion specialists and validated in a wide range of pressures (0.001–100 bar) and

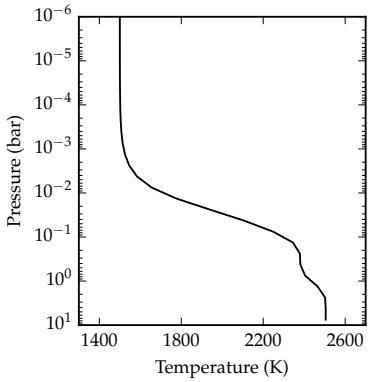

Figure 5.1: Temperature-pressure profile used for the atmospheres under study.

temperatures (300–2500 K), making this model one of the currently most reliable chemical schemes (Battin-Leclerc et al., 2006; Bounaceur et al., 2007; Anderlohr et al., 2010; Wang et al., 2010). Venot et al. (2015) showed that the use of more complete chemical models, including species with up to six carbon atoms, has little effect on the synthetic spectra. We therefore used the simpler, and computationally faster, scheme which includes species with up to four carbon atoms and is able to model the kinetic behaviour of species with up to two carbon atoms. This scheme includes 105 neutral species and 960 reactions (and their reverse reactions). We used a constant diffusion coefficient, $K_{zz} = 10^8$ cm^{2}s^{-1} due to the uncertainties on the vertical mixing acting in exoplanet atmospheres. A similar value has been often used in the literature (Lewis et al., 2010; Moses et al., 2011; Line et al., 2011; Venot et al., 2013). We note that, although this value might be too high (Parmentier et al., 2013), its effect on the final spectra is small compared to other factors such as metallicity and temperature profile.

We used a temperature-pressure (TP) profile with a high-altitude temperature of 1500 K. The vertical profile is the same as the one used in Venot et al. (2015). It was computed using the analytical model one of Parmentier & Guillot (2014), using coefficients from Parmentier et al. (2015) and the opacities from Valencia et al. (2013). The profile, shown in Figure 5.1, was obtained by setting the irradiation temperature to 2300 K and the internal temperature $T_{\rm int} = 100$ K. We assumed a planet with $R_p = 1.162\,R_J$ and $M_p = 1.138\,M_J$.

We computed chemical models for an atmosphere of solar metallicity with C/O of 0.5, 0.7, 0.9, 1.0, 1.1, 1.3, and 1.5.

Table 5.1: JWST instrument modes

Instrument	Mode	Wavelength range (μm)
NIRISS	SOSS/GR700XD	1.0–2.5 μm
NIRCam	LW grism/F322W2	2.5–3.9 μm
NIRCam	LW grism/F444W	3.9–5.0 μm
MIRI	slitless/LRS prism	5.0–10.0 μm

5.2.2 Synthetic high resolution transmission spectra

High resolution ($R \approx 10000$) synthetic transmission spectra were computed using the forward models included in TauREx. This forward model is based on a 1D radiative transfer model, and was described in Chapter 3. The temperature profile used is the same as the one used for the computation of the photochemical models (Figure 5.1). We include a precise computation of the pressure-altitude profile, and take into account the effect of gravity, temperature and mean molecular weight in the computation of the scale height in each of the 100 atmospheric layers included in the model. We compute the pressure grid from 10^{-4} to 10 bar, and define the 10 bar pressure radius to be $R_p = 1.162 R_J$. The mass is set to $M_p = 1.138 M_J$. Amongst the 105 molecules considered in the photochemical model we only consider the following seven molecules in the computation of the opacity in the synthetic spectra: C_2H_2, CH_4, CO, CO_2, H_2O, HCN and NH_3. We found that amongst the complete set of 105 molecules contained in the chemical model, these are the most abundant ones in all cases and will therefore dominate the spectral modulation. The wavelength dependent cross sections for these absorbing molecules were computed using line lists from ExoMol (Barber et al., 2006; Harris et al., 2006; Yurchenko et al., 2011; Tennyson & Yurchenko, 2012; Yurchenko et al., 2013; Barber et al., 2014a), HITRAN (Rothman et al., 2013) and HITEMP (Rothman et al., 2010). Note that the mean molecular weight of each atmospheric layer is coupled to the mixing ratio of all 105 molecules. We included additional opacity from Rayleigh scattering of H_2 and from collision induced absorption of He and H_2–H_2 and H_2–He pairs (Richard et al., 2012).

5.2.3 JWST spectra

We simulated spectra for the Near-InfraRed Imager and Slitless Spectrograph (NIRISS) in Single-Object Slitless Spectroscopy (SOSS) mode using the GR700XD optics (Doyon et al., 2012). We applied a lower wavelength cutoff at 1 μm to avoid saturation and a long wavelength cutoff at 2.5 μm to avoid spectral contamination (Greene et al., 2016). We then used the Near Infrared Camera (NIRCam) using the long wavelength (LW) channel and the

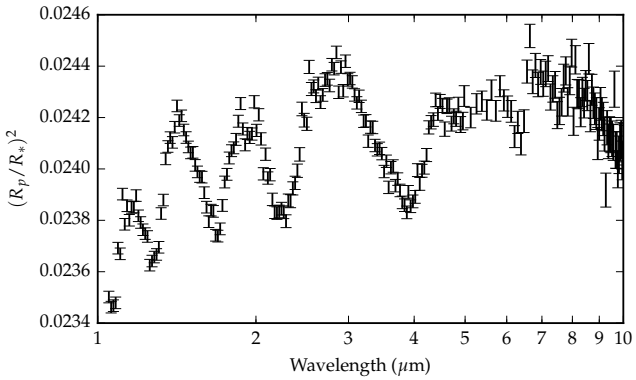

Figure 5.2: Simulated JWST observation for C/O = 0.5. The spectrum was obtained combining four separate synthetic observations obtained with NIRISS, NIRCam and MIRI to cover the 1–10 μm spectral range. This spectrum would therefore require observing a total of four transits.

F322W2 and F444W filters, covering the 2.5–3.9 and 3.9–5.0 μm spectral ranges respectively (Greene et al., 2007). An alternative could be the use of the Near Infra Red SPECtrometer (NIRSPEC) in its high resolution mode with the 2 instrumental configurations: F170LP/G235H (1.7–3.1 μm] and F290LP/G395H (2.9–5.2 μm) (Ferruit et al., 2014). Finally, we use the Mid Infrared Instrument (MIRI) to cover the 5.0–10.0 μm wavelength range. We use MIRI in slitless mode, using the Low Resolution Spectrometer (LRS) and we apply a long-wavelength cutoff of 10 μm due to the degrading S/N at longer wavelengths (Kendrew et al., 2015). Each observation covering the full wavelength range 1–10 μm will therefore require four separate observations. We have considered a one hour effective integration time during the transit and the same amount of time on the star alone. For each mode, the same amount of time was used. Table 5.2.3 summarises the instrument modes considered in this study.

The noise in the spectra was calculated taking into account the star photon noise, the zodiacal and telescope background noise (integrated over the entire band pass of the spectrometer for the slitless mode), the detector dark current and noise. We assumed a star similar to HD189733. The star spectrum used was generated using the PHOENIX atmosphere star code (Husser et al., 2013). For NIRISS and NIRCAM, we have binned the spectra to a constant spectral resolution of $R = 100$. For such a bright star we realised that we are in fact very close to the limitation from systematics of the JWST. Such systematics are difficult to assess but we can reasonably assume that they will be lower than HST. Given the latest performances achieved with HST (e.g. Tsiaras et al., 2016), we can anticipate that

Table 5.2: Free parameters of the two retrieval approaches used in this study. The TP-ISO approach refers to the retrieval using an isothermal TP profile, while the TP-PARAM refers to the retrieval using a parametrised TP profile.

Approach	Parameter	Prior	Description
TP-ISO	log H_2O, log CO,	$-12\ldots1$	Molecular abundances
(10 free parameters)	log CO_2, log CH_4,		
	log NH_3, log HCN,		
	log C_2H_2		
	$T_{\rm iso}$ [K]	$1300\ldots2600$	Isothermal temperature
	R_p [R_J]	$1.05\ldots1.28$	Planetary radius at 10 bar
	$\log(P_{\rm top}\,[\rm Pa])$	$0\ldots6$	Cloud top pressure
TP-PARAM	log H_2O, log CO,	$-12\ldots1$	Molecular abundances
(14 free parameters)	log CO_2, log CH_4,		
	log NH_3, log HCN,		
	log C_2H_2		
	$T_{\rm irr}$ [K]	$1300\ldots2600$	Stellar flux at the TOA
	$\log\kappa_{IR}$	$-4\ldots1$	Mean infrared opacity
	$\log\kappa_{V1}$, $\log\kappa_{V2}$	$-4\ldots1$	Optical opacity sources
	$\alpha_{\rm V}$	$0\ldots1$	Weighting factor for κ_{V1} and κ_{V2}
	R_p [R_J]	$1.05\ldots1.28$	Planetary radius at 10 bar
	$\log(P_{\rm top}\,[\rm Pa])$	$0\ldots6$	Cloud top pressure

the systematics for NIRISS and NIRCAM will be better than about 20 ppm. For MIRI, Greene et al. (2016) adopted a value of 50 ppm, and Beichman et al. (2014) took a value of 30 ppm; we have adopted an intermediate value of 40 ppm. An example of a final spectrum is shown in Figure 5.2.3.

5.2.4 Atmospheric Retrieval

The analysis and interpretation of the simulated observed spectra was carried out using TauREx. Two retrieval approaches were used as part of the current study. Both approaches did not assume any prior knowledge on the chemistry, i.e. the absolute abundance of all gases taken into account is fitted independently. The only difference between the two approaches is in the parametrisation of the temperature profile:

- In the first case we assumed an isothermal TP profile. We will refer to this method as "TP-ISO". This approach is the most commonly used when fitting transmission spectra (Line et al., 2012; Benneke & Seager, 2012; Irwin et al., 2008), and includes a parametrisation of the atmosphere assuming constant-with-altitude mixing ratio and temperature profiles. Crucially, it does not assume any prior on the chemistry of the atmosphere. The free parameters of the retrieval were the absolute abundance of each atmospheric constituent taken into account, the isothermal temperature, the cloud parameters and the 10 bar pressure radius. The mean molecular weight is coupled

to the fitted composition, and we assumed the bulk atmosphere to be formed by a mixture of hydrogen and helium, whose ratio is fixed to solar value (85% H_2 and 15% He). We assumed uniform priors in log space for the absolute abundances, ranging from 10^{-12} to 1. We assumed uniform priors for the temperature (1300–2500) K and for the 10 bar radius (1.05–1.28 $R_{\rm Jup}$). The prior width of the 10 bar radius was determined by assuming a relative uncertainty on R_p of 20% ($R_p = 1.162\ R_J$). Lastly, we fitted the cloud top pressure with a uniform prior in log space (10^{-5}–10 bar). This parametrisation resulted in 10 free variables.

- In the second case, we assumed a more complex TP profile described by five separate parameters. We will refer to this method as "TP-PARAM". Since the temperature profile of the atmospheres under study is highly non-isothermal for pressures greater than 1 mbar (see Figure 5.1), fitting an isothermal profile might lead to biases. We therefore investigate the effectiveness of fitting a more complex profile using this second method. We used the parametrisation of Guillot (2010) modified by Line et al. (2013) and Parmentier & Guillot (2014). There are five parameters that define the temperature profile: one related to the planet internal heat flux ($T_{\rm int}$), and one to the stellar irradiation flux ($T_{\rm irr}$); then there are the opacities in the optical and infrared (κ_{V1}, κ_{V2}), and a weighting factor between optical opacities ($\alpha_{\rm V}$). For a full description of this model we refer the reader to Section 3.1 in Line et al. (2013). These five parameters replace the single parameter used for the isothermal profile in the first method. This model only differs from the first one for the type of TP profile used. This parametrisation resulted in 14 free variables.

The parametrised profile described above is commonly used in the retrieval of emission spectra, where the spectral features are more sensitive to temperature gradients than in transmission. It has received little attention in the retrieval of transmission spectra, as it is assumed that transmission spectra are much less sensitive to temperature gradients, and therefore isothermal profiles, thought to represent the "average" atmospheric temperature, have always been used. Previous studies have addressed the potential bias of the isothermal assumption (Barstow et al., 2013), and found that some information on the temperature profile could be retrieved in transmission only in the highest signal-to-noise and broad wavelength coverage cases.

We used these two approaches to interpret the synthetic JWST observations in a range

of C/O. In all cases we used the MultiNest sampling algorithm (Feroz & Hobson, 2008) to finely sample the parameter space and obtain the posterior distributions of the model parameters. We chose this method instead of a more classical MCMC, as MultiNest can better map the likelihood of highly degenerate parameter spaces. Table 5.2 summarises the free parameters and the corresponding prior widths used in the two retrieval methods.

5.3 Results

5.3.1 Chemical models and transmission spectra for different C/O

Figure 5.3 shows the vertical abundance profiles of seven molecules for all C/O ratios considered in this study, and Figure 5.4 shows the synthetic transmission spectra and contributions of the major opacity sources for the same C/O values. It can be clearly seen that the chemistry and the resulting spectra change significantly between C/O < 1, C/O = 1 and C/O > 1.

Firstly, we note that while the transmission spectra of an oxygen rich atmosphere are dominated almost entirely by H_2O, with additional features from CO at 4.6 μm and from CO_2 at 4.3 μm, a carbon rich atmosphere is dominated by HCN and CH_4, with additional features from CO at 4.6 μm and C_2H_2 at 1.7, 3.0 and 7.5 μm. At the C/O = 1.0 threshold the transmission spectrum is dominated by H_2O and HCN, and exhibits strong features of CO at 2.3 and 4.6 μm. Weak features from CH_4 are also seen at 3.4 and 7.6 μm. Tight constraints on the abundances of all these molecules is therefore paramount to constrain the chemistry and C/O of these atmospheres.

Between C/O = 0.5 and 0.9 we see a gradual decrease in the molar fractions of H_2O and CO_2, and a slight increase in the CH_4, HCN and C_2H_2 abundance, while CO remains relatively constant. The resulting transmission spectra in this C/O range show the progressive decrease in the absorption of H_2O (which remains the dominant absorber across this C/O range) and the resulting emergence of CO, while all the other molecules remain hidden. It is only at C/O > 1.0 that HCN is sufficiently abundant to be clearly seen in the transmission spectrum (see Figure 5.4). We note that at this threshold we see the minimum average absorption from active gases across most of the spectrum, so that in some regions we can also see the emergence of the collision induced absorption from H_2–H_2 and H_2–He pairs.

At C/O = 1.1 the H_2O and CO_2 content drastically drops, while the abundances of CH_4, HCN and C_2H_2 increase significantly. The corresponding transmission spectra show fea-

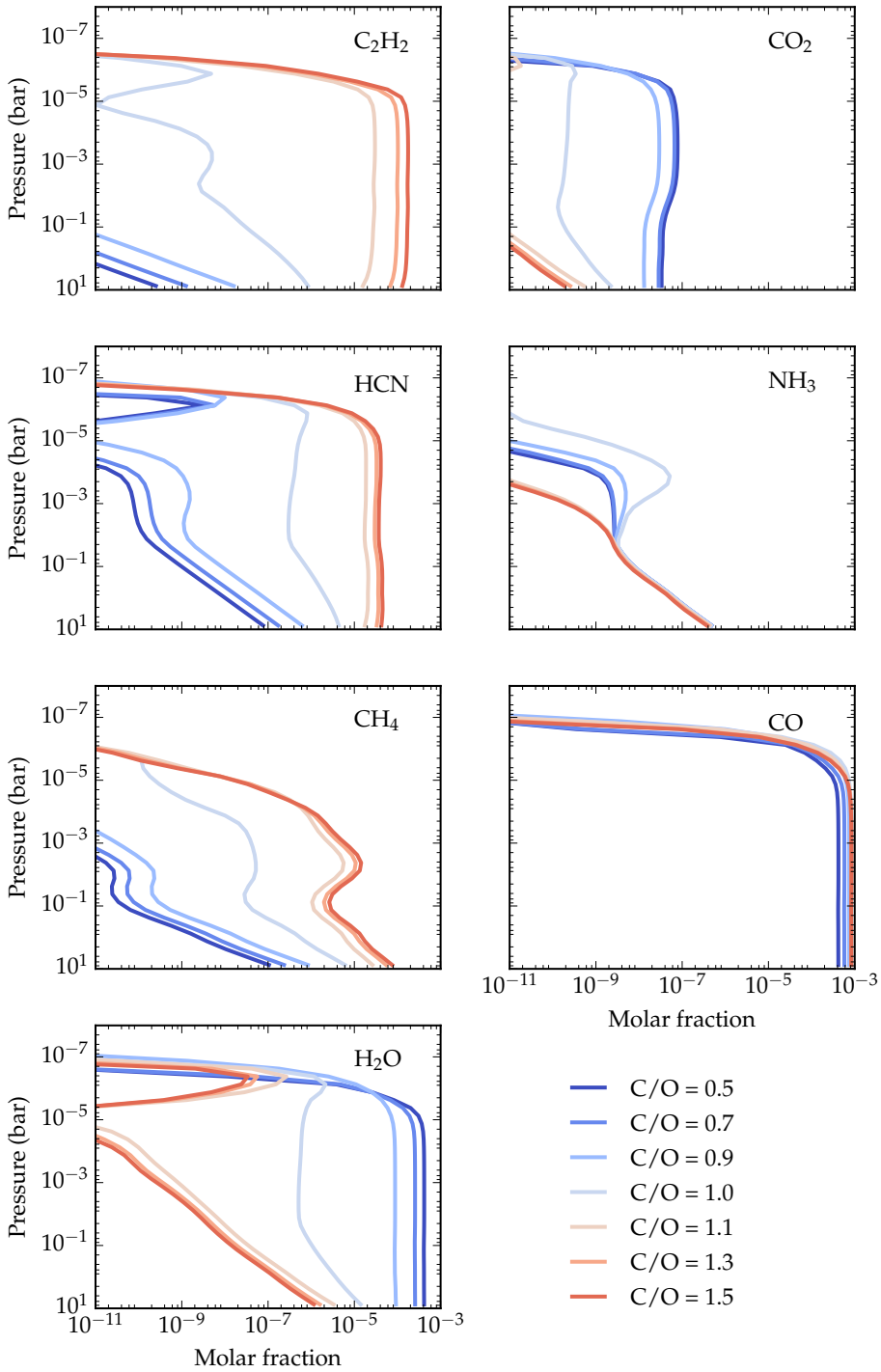

Figure 5.3: Vertical abundance profiles for different molecules for a range of C/O. The different coloured lines show the molar fraction profiles at different C/O, as shown by the legend.

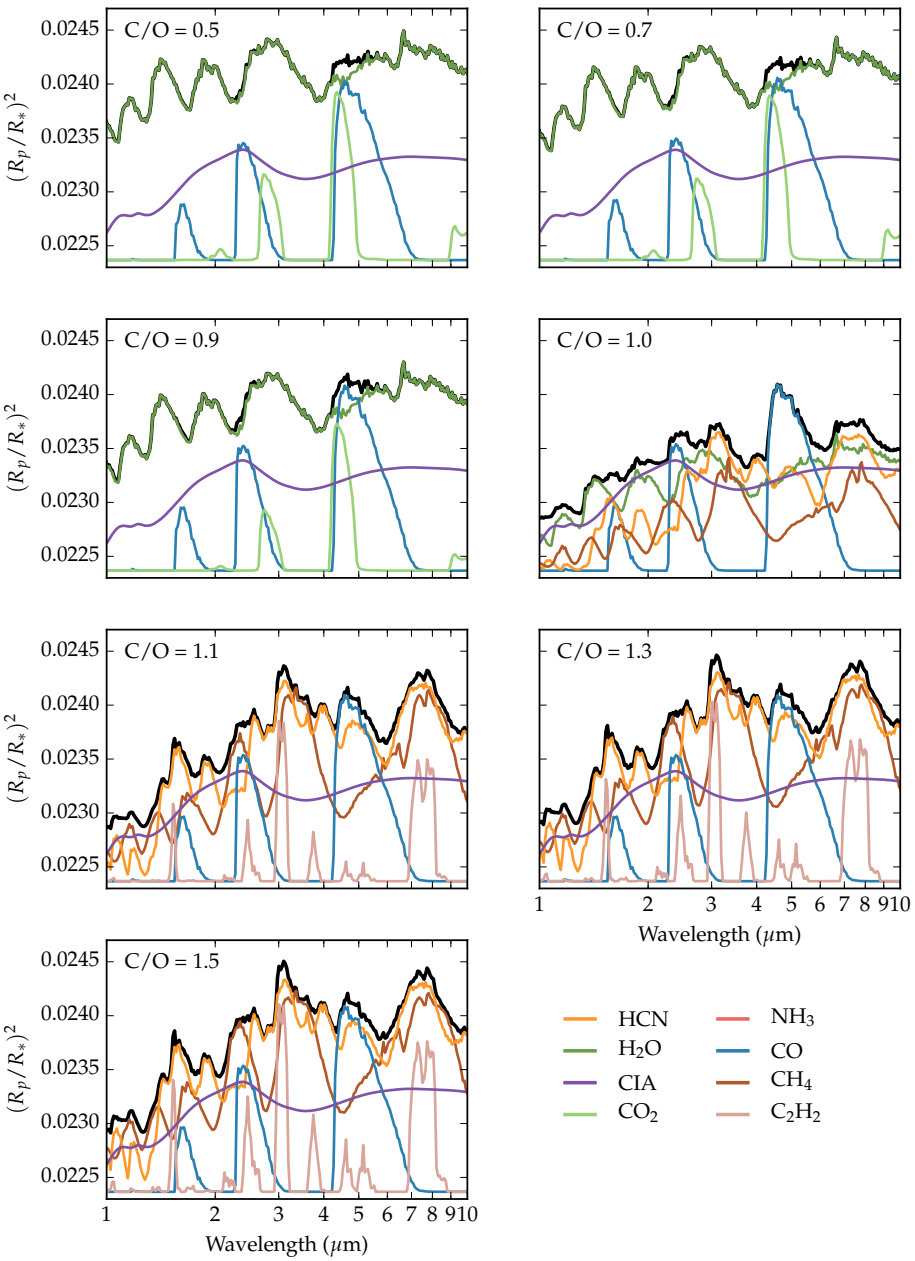

Figure 5.4: Synthetic transmission spectra (black lines) and contributions of the major opacity sources (coloured lines, see legend) for the atmospheres whose chemistry is shown in Figure 5.3, for different C/O values. The opacity sources include the seven molecules considered in this study, and the collision induced absorption (CIA) from H_2–H_2 and H_2–He pairs. Note that for each plot we only show the major opacity contributors to the spectrum, and we hide the molecules that do not significantly contribute to the transmission spectrum features.

tures of CH_4, HCN, CO, and C_2H_2. At progressively higher C/O ratios we see the increase in abundance of CH_4, HCN, and C_2H_2, and the progressive decrease of CO abundance. However, we note that the resulting spectra are very similar to each other. The only differences in the spectra are the weakening of CO at 4.6 μm and the strengthening of C_2H_2 at 3 and 7.5 μm.

Finally, we note that C_2H_2 might actually have additional and much stronger features than those seen here. This is because the line list used for this molecule comes from HITRAN and has been computed experimentally at Earth-like temperatures. It is therefore sub-optimal to use this line list for such high temperatures (> 1500 K). As an appropriate hot line list would include many more transitions resulting from the population of higher vibrational levels, additional spectral features (i.e. "hot bands") are expected, together with the strengthening of the features that can already be seen at lower temperatures. Such a list is under development at ExoMol[1] (private communication).

5.3.2 Retrieval of temperature profiles

Figure 5.5 shows the retrieved temperature profiles using the two approaches for all C/O values. It can be clearly seen that in most cases the retrieved TP profile is within 1 sigma of the input profile using the TP-PARAM method, while using the TP-ISO method the input profile is almost entirely outside the 1 sigma retrieved error bars.

For C/O < 1 (first three plots), it can be seen how the TP-PARAM method fits both the upper atmosphere temperature and the lower-altitude part of the atmosphere. We found that the upper atmospheric temperature could be well fitted within about 1 to 3 sigma using the parametrised TP profile in all cases. The high-altitude temperature was found to be $T = 1502 \pm 66$ K, $T = 1425 \pm 27$ and $T = 1433 \pm 50$ K for C/O = 0.5, 0.7 and 0.9 respectively. Using the TP-ISO method the retrieved temperatures for the same C/O values were $T = 1572 \pm 14$ K, $T = 1610 \pm 17$ K and $T = 1716 \pm 24$ K, respectively. In all cases, the input profile has a high altitude temperature of 1500 K.

From these plots we can also appreciate that the non-isothermal part of the profile could be fitted within one sigma for C/O = 0.5, 0.7 and 0.9. Interestingly, we also note that for C/O < 1 the constraint of the low-altitude temperature ($P > 10^{-3}$ bar) improves for higher C/O, while the fit of the high-altitude part of the profile ($P > 10^{-3}$ bar) improves for lower C/O.

[1] http://www.exomol.com

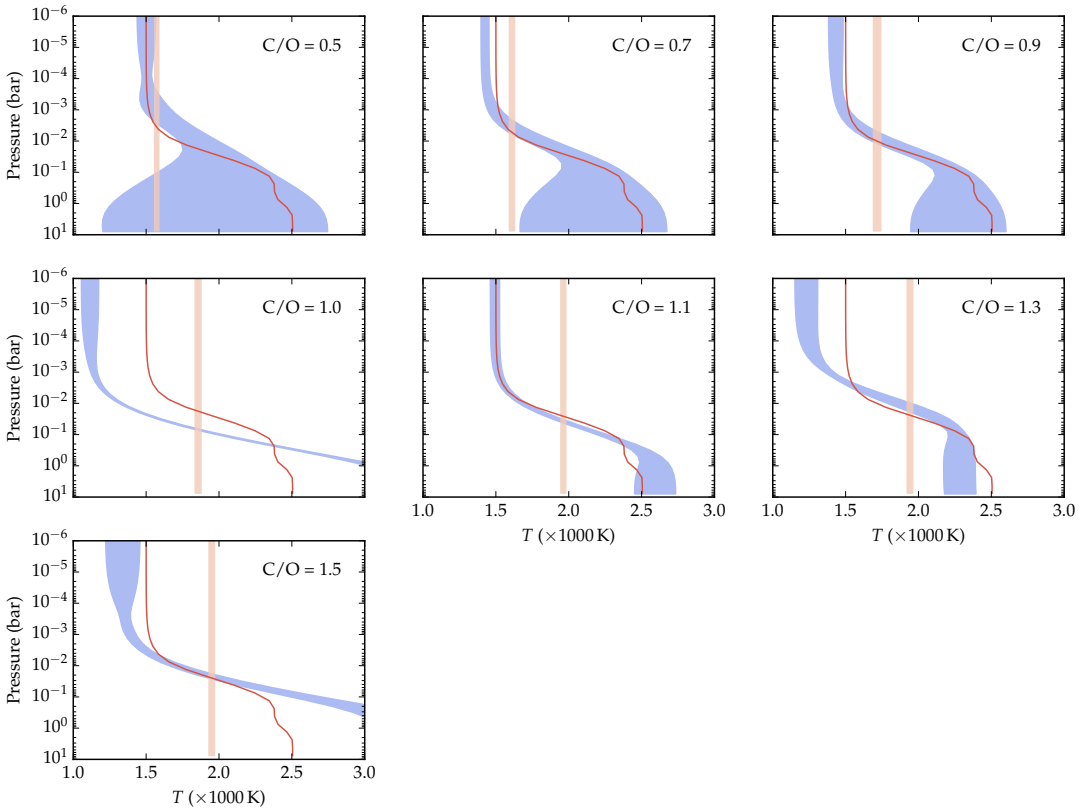

Figure 5.5: Retrieved temperature profiles for the approach with isothermal profile (pink) and parametrised profile (blue) for different C/O . The red line shows the input profile. The shaded areas show the 1 sigma confidence level.

The last three plots in Figure 5.5 show the retrieved temperature profiles using both approaches for C/O $>$ 1. We can see that the TP-ISO approach retrieves a temperature of $\approx$ 2000 K, with an uncertainty of $\approx$ 20 K in all cases. Using the parametrised approach we could fit the high-altitude temperature within about 1 sigma for C/O = 1.1 and 1.5, and within 3.4 sigma for C/O = 1.3. We also note that while the low-altitude part of the TP profile for C/O = 1.1 and 1.3 is well constrained within about 1 sigma, for C/O = 1.5 the fit is poor for pressures higher than 0.1 bar.

For C/O = 1 we note that the TP profile is poorly retrieved, with the TP-PARAM method giving slightly better results. In both cases however the input profile cannot be retrieved within several sigma: the retrieved upper atmosphere temperature is 6 and 18 sigma away from the true state using the TP-PARAM and TP-ISO methods respectively. Additionally, the lower atmosphere temperature ($P < 0.1$ bar) is not retrieved in both cases.

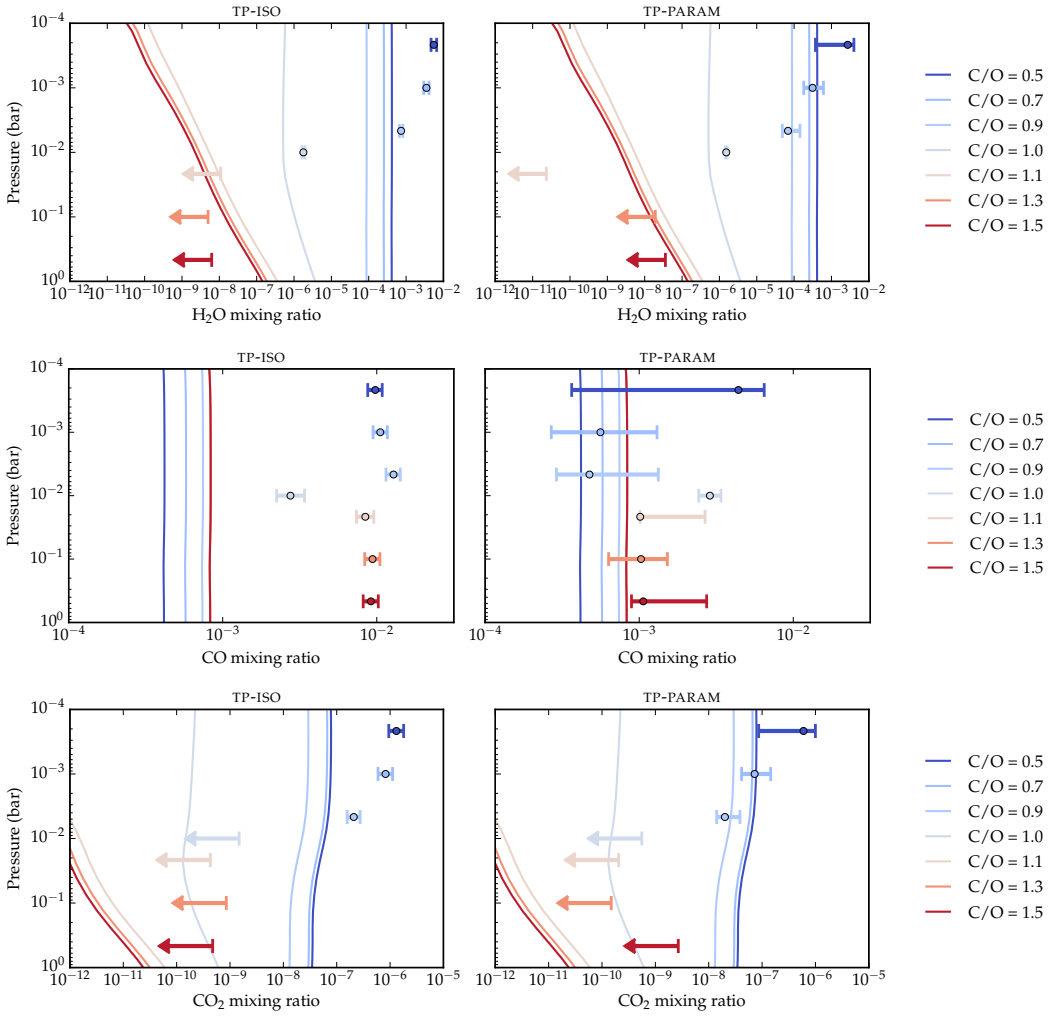

Figure 5.6: Retrieved H_2O (top), CO (middle) and CO_2 (bottom) abundance for C/O = 0.5 – 1.5 using the approach with isothermal profile (left) and parametrised TP profile (right). The solid lines show the input mixing ratio profiles for different C/O, with different colours corresponding to different C/O, as shown by the legend. The retrieved absolute mixing ratios for the different C/O are shown with error bars. Note that we retrieve constant-with-altitude mixing ratio profiles. Note also that the vertical position of the retrieved values are arbitrary.

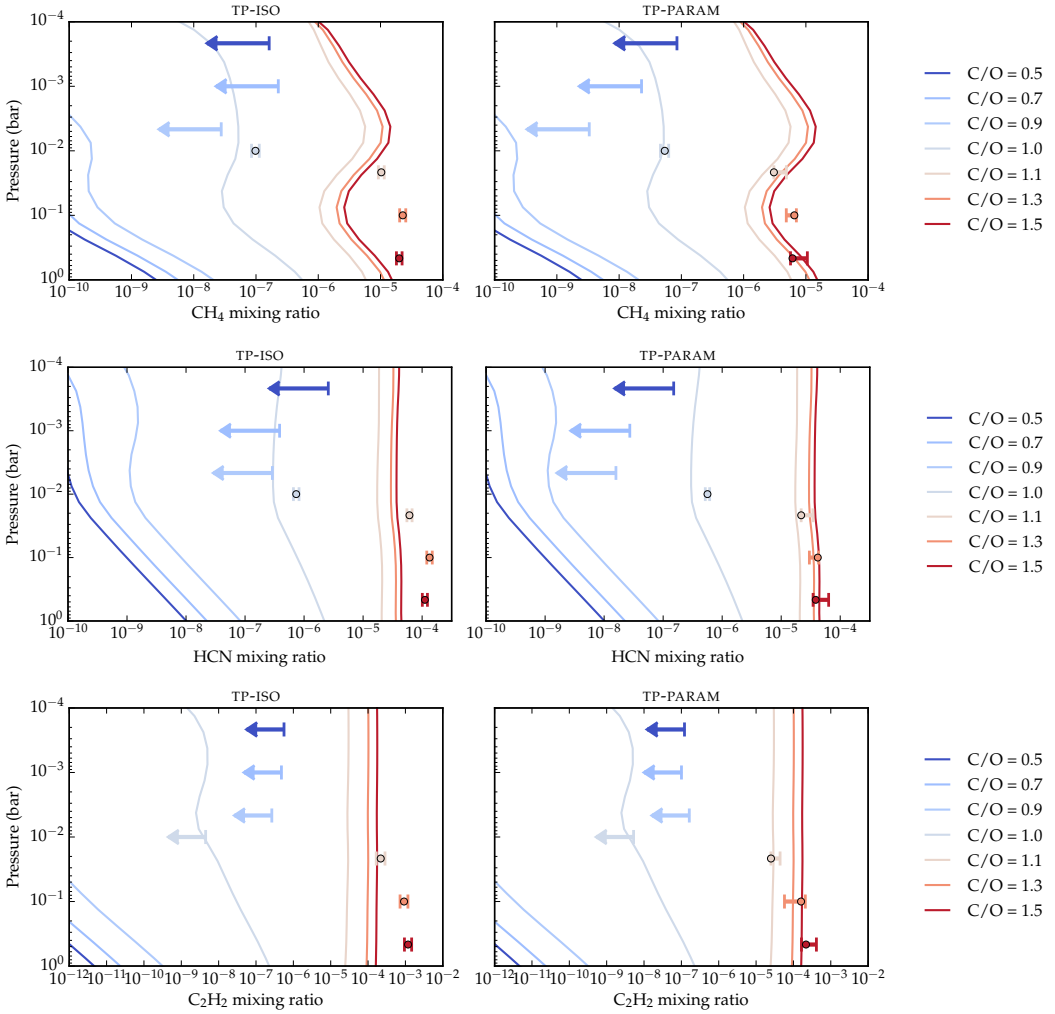

Figure 5.7: Retrieved CH_4 (top), HCN (middle) and C_2H_2 (bottom) abundance for C/O = 0.5 – 1.5. Caption as in Figure 5.6.

5.3.3 Retrieval of atmospheric abundances

The atmospheric retrieval results for the atmospheric abundances of H_2O, CO, CO_2, CH_4, HCN, C_2H_2 and NH_3 are shown in Table A.1 in Appendix A and in Figures 5.6 and 5.7. In these plots, the input mixing ratios for each molecule at each C/O are also shown with solid lines as a function of pressure. We note again that the retrieved abundances are constant-with-altitude, so that a single parameter is retrieved for each molecule using both the TP-ISO (left plots) and TP-PARAM (right plots) retrieval methods. Moreover, we found that NH_3 is never well retrieved, hence we do not show its retrieved values in these figures. This is not surprising given that NH_3 is never seen in the simulated transmission spectra (Figure 5.4).

In general, we found that the TP-ISO *method retrieves higher abundances by about one order of magnitude and significantly underestimates the error bars, causing strong biases,*

while the TP-PARAM *method gives considerably better results for all atmospheres with C/O greater and less than 1, but not for C/O = 1.*

Looking at the transmission spectra for C/O < 1 in Figure 5.4 it can be seen that H_2O has multiple features across the entire wavelength range and is therefore the dominant molecule. Indeed, we found that, for these C/O values, the retrieved abundance of H_2O has the smallest uncertainties, but only the approach using the parametrised TP profile gives unbiased results. Interestingly, the retrieval method using the isothermal approximation was found to bias the results significantly. For example, for C/O = 0.7 the true abundance for H_2O at 0.1 bar is 2.5×10^{-4} and is relatively constant with altitude. The retrieved abundance using the isothermal approximation was found to be $3 - 4 \times 10^{-3}$, and 16 sigma away from the true value. On the contrary, the retrieved abundance using the parametrised TP profile is $1.8 - 6.3 \times 10^{-4}$ and well within 1 sigma from the true value. For C/O = 0.5 and 0.9 we see similar results: using the TP-PARAM method the true state is within 1 to 2 sigma of the retrieved values, but if we use the TP-ISO method, the same retrieved values are 15 and 20 sigma away respectively from the true state.

The two other molecules that contribute to the spectrum, CO and CO_2, were found to be highly degenerate, but could be retrieved within 1 to 2 sigma using the TP-PARAM method. Using the TP-ISO method, abundances were however overestimated. For CO_2 we found that using the parametrised TP profile the true state is within 2 sigma of the retrieved state for C/O = 0.5, and within 1.5 and 1.1 sigma for C/O = 0.7 and 0.9 respectively. Using the isothermal profile, we obtain retrieved values that are significantly overestimated, and are 11.0, 10.2 and 9.6 sigma away from the true state for C/O = 0.5, 0.7 and 0.9 respectively. In these hot oxygen-rich atmospheres the retrieved abundances of CO and CO_2 must however be interpreted with caution, as both molecules have the only detectable feature in the same wavelength range ($\approx 4.0 - 5.5\,\mu$m). From Figure 5.8, showing the posterior distribution of CO and CO_2 for C/O = 0.7 using the parametrised TP-profile, it can be appreciated that the retrieved absolute abundances for these two molecules are highly degenerate. For C/O < 1 no other molecules could be retrieved, and only upper limits could be obtained.

The transmission spectra of these atmospheres with C/O > 1 show that the dominant molecules are CH_4, HCN, C_2H_2 and CO, while all other molecules remain hidden below these stronger absorbers (Figure 5.4). Only these dominant absorbers could be retrieved, while for all other molecules only upper limits could be placed. We found that also for

these carbon-rich atmospheres the TP-PARAM retrieval method gives considerably better results.

Figure 5.7 shows the retrieved CH_4, HCN and C_2H_2 abundances. For C/O > 1 we can see that the input abundance profiles change significantly as a function of pressure, especially for CH_4. In the case of CH_4 we found that the TP-ISO method significantly overestimates the abundances. For all C/O > 1 the retrieved abundances are higher than the true abundances at all pressures in the atmosphere. More reasonable results are obtained with the TP-PARAM method, where the retrieved abundances are always between the maximum and minimum true abundance.

For the same carbon-rich atmospheres, the retrieved abundances of HCN and C_2H_2 using the TP-PARAM approach are all within 1 to 2 sigma of the input abundance, while the values obtained with the TP-ISO are always overestimated by about one order of magnitude, and are 8 to 11 sigma away from the true state. Lastly, we note that the retrieved abundances of CO were within 1 sigma of the true state using the TP-PARAM method, while using the TP-ISO method the same values are an order of magnitude higher than the true state, and have underestimated error bars.

This case with C/O = 1 is the most peculiar as many molecules are visible in the spectrum, and their abundance varies significantly as a function of altitude. In the case of H_2O, CO_2 and CH_4 the true abundance profile changes by about one order of magnitude at the typical pressures probed by transmission spectra (10^{-1} – 10^{-4}) bar (see Figure 5.3). For H_2O, small differences are seen between the TP-PARAM and TP-ISO methods. The retrieved abundances are $1.4 - 1.6 \times 10^{-6}$ in the first case, and $1.6 - 1.9 \times 10^{-6}$ in the second case, while the input profile varies between 5×10^{-7} and 4×10^{-6} for pressures between 1 and 10^{-4} bar. For carbon monoxide, in both cases the retrieved abundances are overestimated by about one order of magnitude, with values 6 to 7.5 sigma away from the true state. This is somewhat surprising, considering that the input profile is constant with altitude. For CO_2 the retrieved abundance is within 1 sigma using the TP-PARAM method, and within 2 sigma using the TP-ISO approach. Finally, for CH_4 and HCN the retrieved abundances are very similar using both methods, and are found to be within the maximum and minimum abundances of the input profiles, which both vary significantly as a function of altitude.

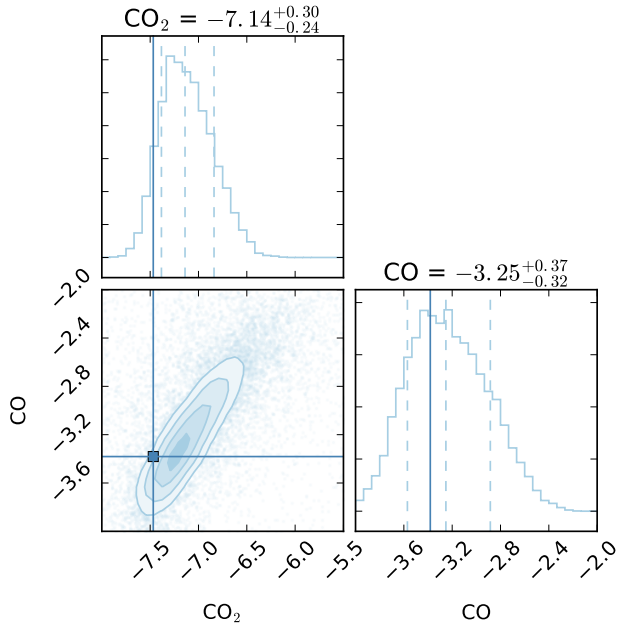

Figure 5.8: Posterior distributions of CO and CO_2 for C/O = 0.7 for the retrieval approach with a parametrised TP profile. Dashed lines in the histogram plots show the 1 sigma confidence intervals. The true state (absolute input abundance at 0.1 bar) is shown with a blue square box and straight blue lines. Note that the mixing ratios of these two molecules is approximately constant-with-altitude in this case, as seen in Figure 5.3.

5.4 Discussion

5.4.1 The impact of common approximations

The results presented in the previous section highlight how common assumptions used in current retrieval methods for exoplanets can potentially lead to wrong conclusions.

Strong biases are seen for all C/O ratios, where we see that the isothermal approximation causes in general an overestimation of the absolute abundances by one order of magnitude, and significantly underestimates error bars. The strongest biases are seen for H_2O, CO and CO_2 in the C/O < 1 atmospheres, and for HCN, CH_4, C_2H_2 for the C/O > 1 atmospheres. This is not surprising, given that these are the strongest absorbers for these C/O ranges, and therefore those with the smallest retrieved uncertainties.

For all these atmospheres, excluding C/O = 1, the retrieval method assuming a parametrised TP profile was found to describe the more complex temperature structure of the atmosphere, leading to retrieved values in general agreement with the true state within 1 sigma on average. This finding opens even new prospects for the use of this technique to characterise exoplanetary atmospheres, showing how *high signal-to-noise and broad wave-*

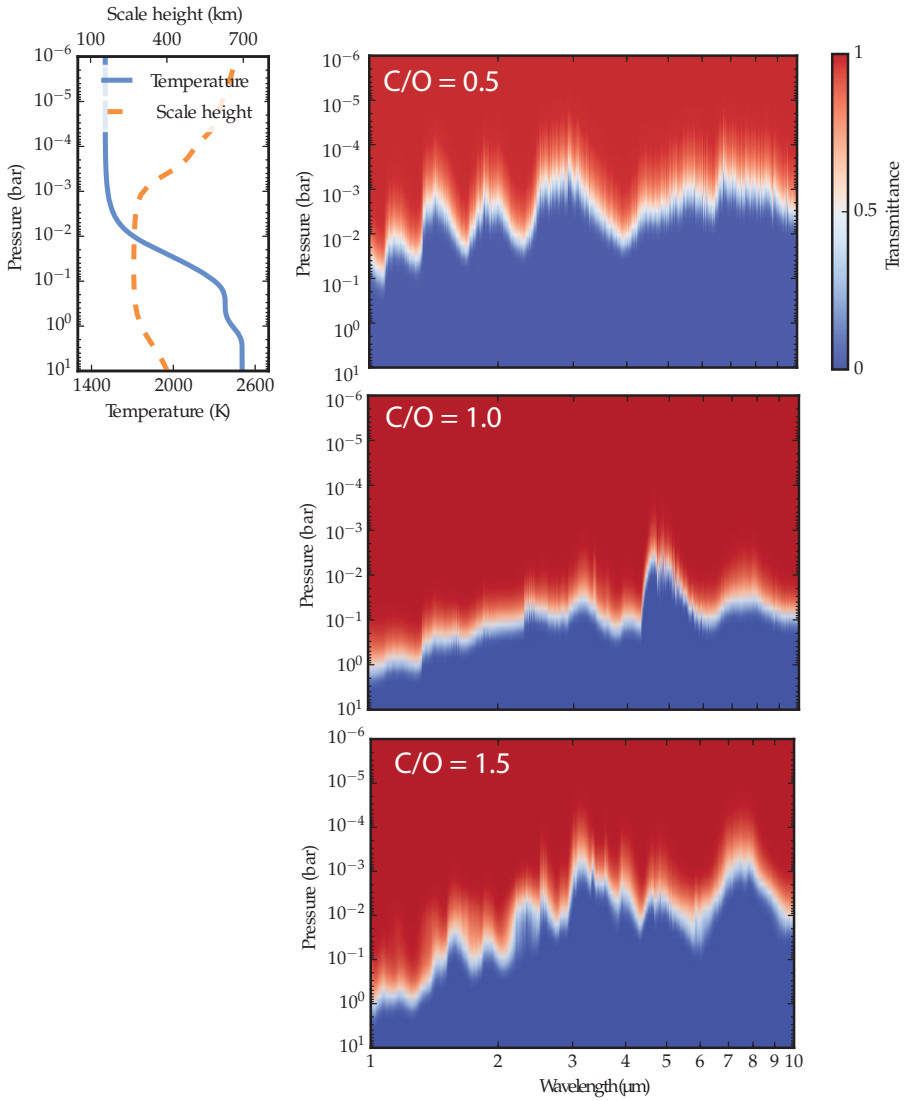

Figure 5.9: The first plot on the left shows the temperature profile (blue line) and scale height profile (dashed orange line) as a function of pressure. The other plots show the spectral transmittance as a function of pressure for the models with C/O of 0.5, 1.0 and 1.5. Note that the pressure axis is the same as the first plot. The transmittance is integrated over the path parallel to the line of sight. The transmittance plots allow us to see the pressures (and therefore the temperature and scale height) probed at different wavelengths for different C/O regimes.

length coverage transmission spectra can lead to significant constraints on the temperature profiles of the terminator region of hot Jupiter atmospheres.

In general, the retrieval of constant-with-altitude mixing ratio profiles seems sufficient to describe the more complex real profiles when the TP-PARAM approach is used, and is therefore a fair approximation in most cases. This is especially true for the C/O $<$ 1 atmospheres, where the true profiles of the most abundant molecules are constant, but it is also true for the C/O $>$ 1 atmospheres, where one of the most abundant molecules, CH_4, has a profile that varies significantly with altitude. The retrieved abundance of this molecule falls within the minimum and maximum true abundance, indicating that the features seen in the transmission spectra at 3.4 and 7.6 μm probe similar pressure regions in the atmosphere.

Retrieved parameters are more strongly affected for the C/O = 1 case, where the biases introduced by assuming a constant-with-altitude abundance profile dominate. Small differences in the retrieved values are seen using the TP-PARAM and TP-ISO methods, and the retrieved results are in both cases several sigma away from the true state. Interestingly, the TP profile retrieved using the TP-PARAM method is also several sigma away from the input profile. This indicates that the biases are driven by the assumption that the abundance profiles are constant-with-altitude, which is clearly wrong for most molecules. In this case, the different features of the same molecules seen at different wavelengths (e.g. H_2O and CO) probe different regions of the atmosphere, where the abundances can vary significantly. Trying to fit these features using the same abundances throughout the entire atmosphere clearly leads to strong biases. We did not explore here the possibility to fit a more complex abundance profile for the molecules, but future work in this direction will be required.

The retrieved abundances obtained with the TP-PARAM method will enable placing some limits on the C/O values of the observed atmospheres. Firstly, it will be clearly possible to differentiate between C/O greater or less than unity, and C/O = 1, as the spectra signatures change dramatically at this threshold. Tighter constraints on C/O can be obtained by linking the retrieved absolute abundances with atmospheric chemical models. However, our results indicate that it will be difficult. For C/O $<$ 1, the strongest tracer for C/O is water. Increasingly lower H_2O abundances are expected at increasing C/O, but the differences seen here are rather small, and comparable with the retrieved uncertainties (see Figure 5.6). Similarly, for C/O $>$ 1, the strongest tracers are HCN and C_2H_2 (and, to a lesser extent,

CH_4, which has however a non uniform abundance profile). However, even in this case the difference in absolute abundance is quite small, and comparable with the error bars of the retrieved values. This is not totally surprising, given that the simulated transmission spectra show very little variation between similar C/O in both the oxygen- and carbon-rich regimes. Higher signal-to-noise observations might further decrease these uncertainties, and therefore improve the inferred C/O, but we note that we are already very close to the systematic uncertainties. The use of additional techniques might prove effective to further constrain the C/O ratio, such as emission spectra through secondary eclipse measurements. Alternative approaches include the use of chemically consistent retrieval methods (see e.g. Greene et al., 2016). Such self consistent approaches, however, might lead to different sources of biases, as these models might not fully describe the atmospheric chemistry.

5.4.2 Understanding the biases

In order to understand why, and in which scenarios, a non-isothermal profile and constant-with-altitude abundance profiles might lead to strong biases, it is instructive to look at the spectral transmittance as a function of pressure for the atmospheres under study. Figure 5.9 shows the spectral transmittance integrated over the path parallel to the line of sight as a function of pressure, together with the temperature and scale height profiles. It can be seen that different spectral regions probe different pressure ranges, and therefore different temperatures and scale heights. Firstly, we note that the scale height does not increase exponentially with altitude between 10^{-3} and 1 bar, as one would expect in a purely isothermal atmosphere. On the contrary, the strong temperature gradient seen at these pressures causes the scale height to stay relatively constant at $\approx$ 200 km. For the atmosphere with C/O = 0.5 we see that most of the absorption occurs between 10^{-4} and 10^{-1} bar, while for the C/O = 1.1 case the transmission spectrum probes higher-pressure regions, from 1 bar to 10^{-3} bar. At these pressures the temperature varies from 1500 K to about 2500 K. We also note that the peak of the absorption features probe the higher-altitude and lower-temperature part of the atmospheres, while the troughs probe the regions of the atmosphere that are almost 1000 K hotter.

An isothermal approximation will clearly lead to several problems. Firstly, as we noted before, the scale height of an isothermal atmosphere will increase exponentially, while in this case it is roughly constant with pressure up to 1 mbar. Spectral features that probe different pressures, such as the strong water features seen for C/O < 1, will therefore vary

considerably if the scale height is constant with pressure or not. A second, equally important effect, is caused by the very different temperatures probed. Molecular opacity cross sections vary considerably between the temperature regions probed here (1500 K to 2500 K), and therefore assuming a single temperature will obviously lead to further biases.

Additionally, Figure 5.9 helps to explain why for the retrievals of the atmospheres with C/O < 1 we found that the fit of the low-altitude temperature improves for higher C/O, while that for the high-altitude part of the TP profile improves for lower C/O. As the C/O value increases from 0.5 to 0.9 we see that the water abundances decreases from about 4×10^{-4} to 1×10^{-4}. The effect in the transmission spectrum is a vertical shift towards lower absorption, which also translates into a vertical shift in the transmissivity plot. This means that as the water content drops, we probe increasingly higher pressure regions of the atmospheres, meaning that we increasingly lose information from the upper-altitude part of the atmosphere. This easily explains why the uncertainty on the retrieved upper-altitude temperature of these atmospheres progressively increases, while the constraint of the temperature in the bottom layers improves for higher C/O.

So far we have only considered cloud free, broad wavelength range observations. This is the case where common approximations are most likely to break down. Shorter wavelength ranges will for example tend to probe specific regions of the TP profiles. For instance, an atmosphere with C/O = 1.1 observed between 1 and 3 μm will only probe pressures between 1 and 0.1 bar, where the temperature is roughly constant at ≈ 2400 K. In this scenario, we expect the isothermal approximation to be sufficiently good. However, this is not always the case. If the atmosphere with C/O = 0.5 is observed between 2.5 and 4 μm, we will see a strong water feature with a peak absorption coming from a region with a temperature of about 1500 K, and with wings probing increasingly higher temperatures. Clearly, even in this case an isothermal approximation would give biased results, and our study indicates that the retrieved uncertainty of the abundance will be likely underestimated.

The presence of uniform clouds will increase the degeneracy of model parameters, somewhat hiding the underlying biases, as the effect of a cloud deck is that of making the atmosphere opaque. A cloud deck extending to 10 mbar would for example make the atmosphere opaque to incoming radiation for pressures higher than 10 mbar. This also means that it will be impossible to probe the temperature and mixing ratio profiles in this pressure regime. In the case under study, the TP profile for pressure lower than 10 mbar is relatively

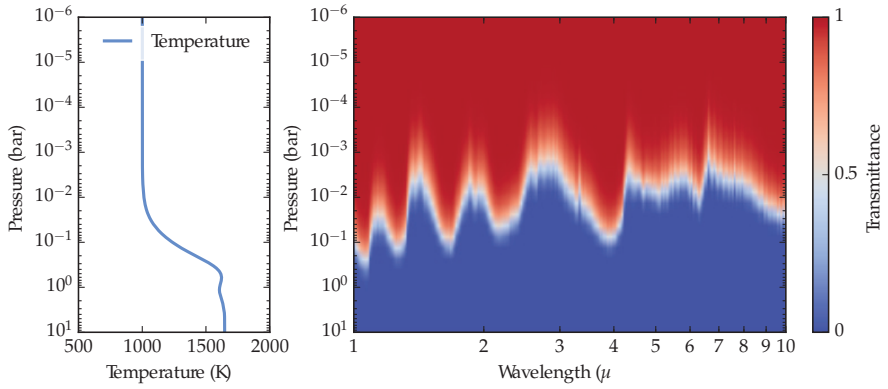

Figure 5.10: Temperature pressure profile and spectral transmittance for a planet with a cooler TP profile and C/O = 0.5. Caption as in Figure 5.9.

isothermal, and in the presence of clouds, an isothermal approximation would therefore be appropriate. Note however that cloud models commonly used in current retrievals were found to cause significant degeneracies. Line & Parmentier (2016) investigated the biases of retrieving a uniform cloud cover in the presence of patchy clouds and found significant degeneracies in the retrieved mean molecular weight.

We also note that similar biases are expected for cooler planets. Figure 5.10 shows the spectral transmittance as a function of pressure for an atmosphere with a cooler TP profile, with high altitude temperature of 1000 K. The spectrum was computed from a chemical model with C/O = 0.5, and assuming the same hot Jupiter used in this work. It can be seen that the spectrum probes the range of pressures (10^{-3} to 1 bar) where the TP profile changes more significantly. We therefore expect that the use of an isothermal profile to retrieve this spectrum will lead to similar biases to those found for the hotter planet case.

Lastly, we note that this study focused on two specific common assumptions in current retrieval methods, constant mixing ratio and temperature profiles, and the biases that these approximations can lead to. However, other strong assumptions are likely to bias our retrievals. One of the most important one is to neglect 3D dynamical effects. The simulated observations have in fact been generated using 1D chemical models, and assume a uniform chemistry and atmospheric temperature at the terminator region. Further studies that compare transmission spectra obtained with general circulation models and retrieved with the simpler 1D models are needed to address the biases of this assumption. A recent study in this direction is presented in Feng et al. (2016).

5.5 Conclusions

In this Chapter the biases caused by two common assumptions in the forward models used by current retrieval methods of transmission spectra of hot Jupiter atmospheres were investigated: the use of an isothermal profile and constant-with-altitude abundances. We investigated whether these assumptions will still be valid for high signal-to-noise, broad wavelength coverage spectra such as those expected by JWST. In order to do this, we simulated high quality observations using chemical schemes developed by Venot et al. (2012), which include detailed temperature and abundance profiles, and we retrieved them using two simpler forward models: the first one assumes an isothermal profile (TP-ISO), while the second one assumes a parametrised temperature profile (TP-PARAM). In both cases, constant-with-altitude abundances were retrieved. We found that:

- We can easily differentiate between C/O $<$ 1, C/O = 1 and C/O $>$ 1, as the chemistry changes more drastically at this threshold. However, we also found that tighter constraints are more difficult to obtain as the differences between the transmission spectra are relatively small. Higher signal-to-noise observations might lead to better constraints, but other biases, due to systematic uncertainties for example, might become more dominant. Emission spectra observations, possibly combined with transmission spectra, might give better constraints than transmission spectra alone.

- The non-uniform temperature profile could be well retrieved within about 1 sigma for all cases but C/O = 1 using the TP-PARAM method. This is an important result, opening the possibility to obtain detailed temperature structure information about the terminator region of a hot Jupiter.

- The retrieval approach that assumes an isothermal profile led to strong biases. We found that, on average, the retrieved abundances using this method are overestimated by about one order of magnitude and the error bars are underestimated. The TP-PARAM approach leads to much improved constraints, with retrieved abundances within 1–2 sigma of the input values in most cases.

- The retrieval assumption that abundance profiles are constant-with-altitude was found to be a good approximation for C/O $<$ 1 and C/O $>$ 1 atmospheres, but not for C/O = 1. In this latter case, most of the abundance profiles have strong variations, and a uniform abundance profile is a poor approximation that leads to significant

biases. Future work will therefore be needed to address the feasibility of fitting more complex abundance profiles.

These results show that when broad wavelength ranges and high signal-to-noise observations are used, the forward models used in our retrieval approaches need to allow for larger flexibility. One very simple solution is to adopt a parametrisation of the temperature-pressure profile, as the one used here, but other techniques, such as the two-stage approach used in Waldmann et al. (2015a) and described in Chapter 2, Section 6.1, could be considered in the future.

Chapter 6

Atmospheric retrieval of a hot Jupiter and a super-Earth

Spectral retrievals are becoming increasingly important and commonly used for the interpretation of observed transit and emission spectra of exoplanets. While such algorithms will show their full potential when high signal-to-noise, broad wavelength ranges observations will become available in the future with the advent of the James Webb Space Telescope or other space-based telescopes, as seen in the previous chapter, the best signal-to-noise spectra obtained today still require atmospheric retrievals for their correct interpretation.

Modelling low-resolution spectra in a short wavelength range requires in fact particular care in the handling of the priors, and requires a full mapping of the parameter space. This is because such models are highly degenerate, meaning that completely different models often explain the same dataset.

In this context, TauREx – thanks to its Bayesian framework and the use of the Nested Sampling algorithm - is the tool of choice for the interpretation of the transmission spectra available today. In this chapter, I will show how TauREx has been applied for the interpretation of spectra obtained with the Wide Field Camera 3 on-board the Hubble Space Telescope, the state-of-the-art instrument in use today to observe the atmospheres of transiting exoplanets.

The results presented in this chapter are published in Tsiaras et al. (2015) and Tsiaras et al. (2016) and are here reproduced with permission from the publisher.

6.1 Introduction

In Tsiaras et al. (2015) we presented a new pipeline for analysing data coming from the Wide Field Camera 3 (WFC3) on-board the Hubble Space Telescope (HST), and obtained

using the spatial scanning technique. During a spatial scanning exposure, the telescope slews slowly along the vertical axis of the detector instead of staring at the target, effectively smearing the spectrum over a larger number of pixels in the detector. As a result, the total number of collected photons is larger, therefore increasing the signal-to-noise ratio without incurring saturation. This technique has already been used in a number of exoplanetary spectra, and today it represents the de-facto standard method to observe exoplanetary spectra with WFC3/HST (e.g. Deming et al., 2013; Knutson et al., 2014a; Fraine et al., 2014).

This pipeline, compared to those already available, takes into account the geometric distortions of the instrument, whose impact may become important when combined to the scanning process. It can improve the photometric precision of existing data, while also pushing further the limits of this successful technique. It has been successfully benchmarked reanalysing the transit of HD 209458b (Tsiaras et al., 2015). The same pipeline was also used to analyse two new spectroscopic observations of the super-Earth 55 Cancri e obtained with WFC3/HST. These observations were obtained using the scanning mode, and adopting a very long scanning length and high scanning speed. The pipeline was successfully applied to produce a final transmission spectrum with relative uncertainties of only 22 ppm per visit, the smallest ever obtained for a WFC3 spectrum.

I performed the interpretation of the spectra of these two planets using TauREx: the analysis of the spectrum of HD 209458b suggested the presence of water vapour and clouds, while the retrieval of the spectrum of 55 Cancri e suggested that it is surrounded by an atmosphere dominated by hydrogen. This latter finding represents the first detection of an atmosphere around a super-Earth. In this chapter, I will present the atmospheric retrievals of the spectra of these two planets.

6.1.1 Acknowledgments

The results presented here have been published in Tsiaras et al. (2015) and Tsiaras et al. (2016), and have been obtained thanks to the additional contributions of other authors (in particular A. Tsiaras, I.P. Waldmann, and O.Venot). A. Tsiaras wrote the WFC3 pipeline and analysed the WFC3/HST observations, I.P. Waldmann assisted with the spectral retrievals, and O. Venot computed the atmospheric chemical models of 55 Cancri e.

Table 6.1: HD 209458b system parameters

Stellar parameters	
[Fe/H] [dex] (2)	0.00 ± 0.05
H [mag] (1)	6.591 ± 0.011
J [mag] (1)	6.366 ± 0.035
K [mag] (1)	6.308 ± 0.021
T_* [K] (2)	6065 ± 50
$M_* [M_\odot)$ (2)	1.119 ± 0.033
$R_* [R_\odot)$ (2)	1.155 ± 0.016
$\log(g_{*,\mathrm{surf}})$ [cgs] (2)	4.361 ± 0.008
Planetary parameters	
T_p [K] (2)	1449 ± 12
$M_p [M_J]$ (2)	0.685 ± 0.015
$R_p [R_J]$ (2)	1.359 ± 0.019
a [AU] (2)	0.04707 ± 0.00047
Transit parameters	
T_0 [HJD] (3)	$2452826.628521 \pm 0.000087$
P [days] (3)	$3.52474859 \pm 0.00000038$
R_p/R_* (2)	0.12086 ± 0.00010
a/R_* (2)	8.76 ± 0.04
i [deg] (2)	86.71 ± 0.05

(1) Cutri et al. (2003)
(2) Torres et al. (2008), ((2*) derived)
(3) Knutson et al. (2007)

6.2 HD 209458 b: water and possibly clouds in the atmosphere of a hot Jupiter

The hot Jupiter HD 209458b is the first transiting exoplanet detected (Charbonneau et al., 2000; Henry et al., 2000) and the first for which a transit and emission spectra were obtained (Charbonneau et al., 2002; Deming et al., 2005). The system parameters of this planet are shown schematically in Table 6.1. Previous observations in the optical have confirmed the presence of sodium in the atmosphere of this planet (Charbonneau et al., 2002; Snellen et al., 2008; Sing et al., 2008), while UV observations suggested that the planetary atmosphere is in hydrodynamic escape (Vidal-Madjar et al., 2003; Holmström et al., 2008; Ben-Jaffel & Sona Hosseini, 2010; Linsky et al., 2010). At longer wavelength, water vapour was identified with low-resolution spectra (Beaulieu et al., 2010; Deming et al., 2013), while carbon monoxide was detected with high-resolution spectroscopy (Snellen et al., 2010). Other species such as methane and carbon dioxide have also been suggested (Swain et al., 2009).

The spectrum, obtained analysing the spectroscopic images from WFC3/HST (Pro-

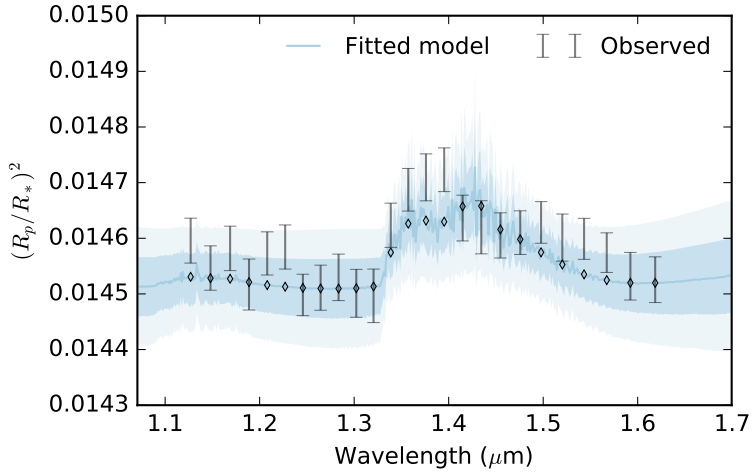

Figure 6.1: Infrared transmission spectrum of HD 209458b (black error bars), best fit obtained with the second retrieval containing H_2O and clouds (blue line). The shaded regions show the 1 and 2 sigma confidence intervals in the retrieved spectrum.

gram ID: 12181, PI: Drake Deming; Deming et al., 2013), is shown in Figure 6.1, together with the best fit model obtained with TauREx.

6.2.1 Atmospheric retrieval

The Nest Sampling algorithm implemented in TauREx was used to fully explore the parameter space and find the best fit to the WFC3 spectrum. Because of the limited number of data points in the observed spectrum, in order to reduce significantly the parameter space, the atmosphere is parametrised assuming an isothermal profile, with constant molecular abundances as a function of altitude. The fitted parameters are the temperature, the molecular abundances for the different species, the mean molecular weight, the radius at 10 bar, and the cloud top pressure – i.e. the pressure at which the cloud starts to be opaque. The cloud model used assumes an opaque and uniformly distributed cloud deck defined at a given pressure beyond which electromagnetic radiation is blocked at all wavelengths. Although this cloud model is extremely simple, it is considered sufficient to mimic the behaviour of real clouds in such a short wavelength range (1.1 – 1.7 μm) and for the signal-to-noise of the datasets under study. A broad range of absorbing molecules is considered, including H_2O, HCN, NH_3, CH_4, CO_2, CO, NO, SiO, TiO, VO, H_2S, C_2H_2.

We fit for the individual molecular abundances, assuming the bulk composition of the atmosphere to be made by a mixture of 85% hydrogen and 15% helium. We then couple the mean molecular weight to the atmospheric composition. We consider uniform priors

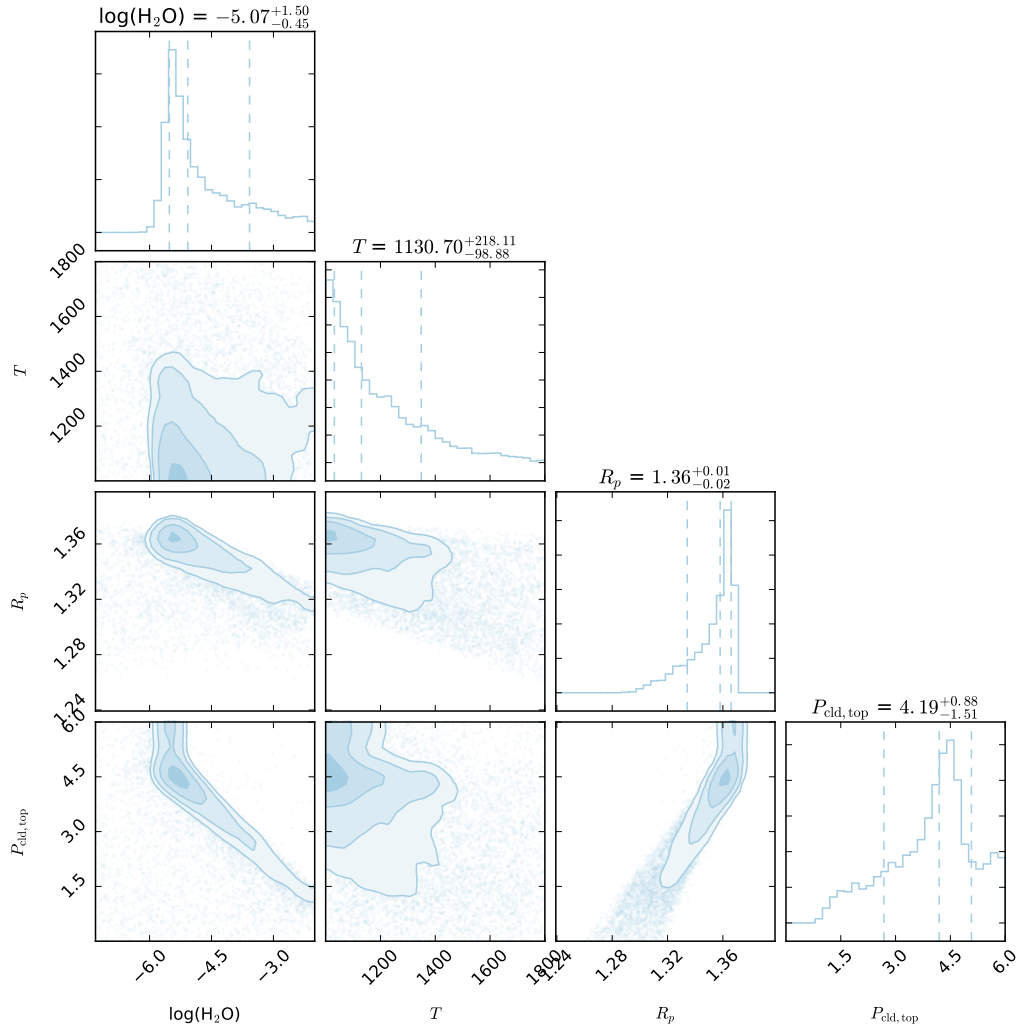

Figure 6.2: Posterior distributions of the second spectrum retrieval of HD 209568b including H_2O and clouds.

for the molecular volume mixing ratios ranging between 10^{-12} and 10^{-2}. This prior is justified by the fact that in hot Jupiters the absolute abundances of absorbing gases are significantly smaller compared to the H_2 and He content. We also assume uniform priors for the temperature ($T = 1000 - 1800$ K), 10 bar radius ($R_p = 1.3 - 1.4 R_J$) and cloud top pressure ($P_{\rm cloud} = 10^{-5} - 10^{-1}$ Pa). We run two retrievals, the first one including 12 molecules and aimed at identifying the most likely trace gases, and the second one including only the molecules identified in the first run, aimed at fully mapping the parameter space and at investigating the degeneracy of the model.

The first retrieval including all molecules shows that water is the strongest and most likely absorber, explaining the broad absorption feature at $\approx 1.35\,\mu$m. No other molecules

seem to contribute to the overall spectrum, while clouds may be present to explain the flat spectrum seen between 1.1 and 1.3 μm. We therefore run the second retrieval including only H_2O and clouds. Figure 6.1 shows the best fit to the data corresponding to the maximum a posteriori solution of this Bayesian retrieval, while Figure 6.2 shows the posterior distributions of this retrieval. We find that the retrieved absolute abundances of H_2O is $3 \times 10^{-6} - 3 \times 10^{-4}$. However, the posterior distributions (Figure 6.2) shows that this parameter is highly degenerate with the cloud top pressure and the 10 bar radius. It is therefore impossible with these data alone to constrain the absolute abundances of this absorber.

We found the 10 bar radius to be $1.36^{+0.01}_{-0.02}\,R_J$. The posterior distributions also show that the data can be best explained by a cloud deck at 0.15 bar, but we note that the distribution is very broad (and degenerate with the other fitted parameters), and a solution without clouds or with lower-pressure clouds is also acceptable.

The first retrieval including all molecules has a global evidence $\log \mathscr{E} = 209$, while the second retrieval including H_2O only has $\log \mathscr{E} = 210$. Despite the global evidence of the H_2O-only retrieval being marginally higher than that of the more complete model, this result shows that there is no statistical evidence that favour the presence of additional molecules in the spectrum (Trotta, 2008).

6.3 55 Cancri e: detection of the first atmosphere around a super-Earth

55 Cancri e is one of the five planets orbiting the star 55 Cancri, all discovered by radial velocity measurements (Butler et al., 1997; Marcy et al., 2002; McArthur et al., 2004; Fischer et al., 2008). It is an "exotic" super-Earth ($R_p = 8.09 \pm 0.25\,R_\oplus$; other system parameters shown in Table 6.2), orbiting so close to its host star that its surface temperature is in excess of 2000 K.

Super-Earths are an interesting class of planets that do not exist in our solar system, although they are found to be the most abundant planets in our Galaxy, especially around late-type stars (Mayor et al., 2011; Howard et al., 2012; Fressin et al., 2013; Dressing & Charbonneau, 2013). From early studies it seems that there might be a variety of cases, ranging from rocky to H_2/He composition (Sotin et al., 2007; Grasset et al., 2009; Valencia et al., 2007, 2013; Zeng & Sasselov, 2014). Before spectral observations of 55 Cancri e, there were only two super-Earths spectra observed with WFC3/HST, GJ1213b and

Table 6.2: 55 Cancri e system parameters

Stellar parameters	
$[Fe/H]$ [dex] (1)	0.31 ± 0.04
T_* [K] (2)	5196 ± 24
M_* $[M_\odot]$ (2)	0.905 ± 0.015
R_* $[R_\odot]$ (2)	0.943 ± 0.010
$\log(g_{*,\mathrm{surf}})$ [cgs] (2)	4.45 ± 0.001
Planetary parameters	
T_p [K] (3)	1950^{+260}_{-190}
M_p $[M_\oplus]$ (4)	8.09 ± 0.26
R_p $[R_\oplus]$ (5)	$1.990^{+0.084}_{-0.080}$
a [AU] (5)	$0.01545^{+0.00025}_{-0.00024}$
Transit parameters	
T_0 [BJD] (5)	$2455962.0697^{+0.0017}_{-0.0018}$
P [days] (5)	$0.7365417^{+0.0000025}_{-0.0000028}$
R_p/R_* (5)	$0.01936^{+0.00079}_{-0.00075}$
a/R_* (5)	$3.523^{+0.042}_{-0.040}$
i [deg] (5)	$85.4^{+2.8}_{-2.1}$

(1) Valenti & Fischer (2005), (2) von Braun et al. (2011)
(3) Crossfield (2012), (4) Nelson et al. (2014)
(5) Dragomir et al. (2014)

HD97658b (Kreidberg et al., 2014; Knutson et al., 2014b). In both cases, no significant spectral modulation was detected, suggesting the presence of thick clouds or a high mean molecular weight atmospheres.

Spitzer eclipse observations of 55 Cancri e suggest variability of the thermal emission from the dayside of the planet over time (Demory et al., 2016a), while a thermal brightness map, again obtained with Spitzer observations, suggests a large day-night temperature gradient (Demory et al., 2016b). While this large temperature variations might suggest the lack of an atmosphere, the modelling of the atmospheric and internal structure of this planet (Stevenson, 2013; Zeng & Sasselov, 2014; Hu & Seager, 2014; Forget & Leconte, 2014) shows that a gaseous envelope of H_2 and He might have been retained from the protoplanetary disc.

Primary transit observations of 55 Cancri e, obtained with WFC3/HST (ID: 13665, PI: Bjoern Benneke), were analysed with the pipeline presented in Tsiaras et al. (2015). The long scanning length and the high scanning speed intensify the geometrical distortions of the imaged spectra, that our pipeline can accurately model and correct for.

6.3.1 Atmospheric retrieval

In order to fit the WFC3 spectrum we use TauREx. As for HD 209458b, the atmosphere is parametrised assuming an isothermal profile with constant molecular abundances as a function of pressure. We include a wide range of molecules in the fit, including H_2O, HCN, NH_3, CH_4, CO_2, CO, NO, SiO, TiO, VO, H_2S, C_2H_2. The fitted parameters are the mixing ratios of these molecules, the atmospheric mean molecular weight, the surface pressure and radius.

We use uniform priors for the gas mixing ratios ranging from 1 to 10^{-8}. The mean molecular weight is coupled to the fitted composition, in order to account for both the fitted trace gases and possible unseen absorbers with signatures outside the wavelength range probed here. The uniform prior assumed for the mean molecular weight ranges from 2 to 10 amu. Lastly, we assume uniform priors for the surface radius, the surface pressure and the mean atmospheric temperature, ranging between $0.1 - 0.3\,R_J$, $10 - 10^7$ Pa, and $2100 - 2700$ K, respectively. We do not include a separate parameterisation for the cloud layer, as the pressure at the surface could be the pressure at the top of a cloud deck.

6.3.2 Ab-initio chemical simulations

In parallel to the spectral retrieval, we investigated the theoretical predictions for the chemical composition of an atmosphere enveloping 55 Cancri e. We assumed an atmosphere dominated by hydrogen and helium, with a mean molecular weight of 2.3 amu. We used a thermal profile with high-altitude atmospheric temperature of 1600 K. These parameters correspond to a scale height of 440 km, or 25 ppm, at the surface.

We used the same chemical scheme implemented in Venot et al. (2012) to produce vertical abundance profiles for 55 Cancri e, assuming a solar C/O ratio and a C/O ratio of 1.1. This chemical scheme can describe the kinetics of species with up to two carbon atoms. In Venot et al. (2015) it was shown that a more sophisticated chemical scheme including species with up to six carbon atoms produces comparable results, and that the simpler scheme can reliably model atmospheres with C/O ratios above unity. The scheme has been developed with combustion specialists and validated by experiments conducted in a wide range of temperatures ($300 - 2500$ K) and pressures ($0.01 - 100$ bar) (e.g. Battin-Leclerc et al., 2006; Bounaceur et al., 2007; Anderlohr et al., 2010; Wang et al., 2010).

The stellar flux was calculated in the following way. From 1 to 114 nm, we used the mean of the Sun spectra at maximum and minimum activity (Gueymard, 2004), scaled for

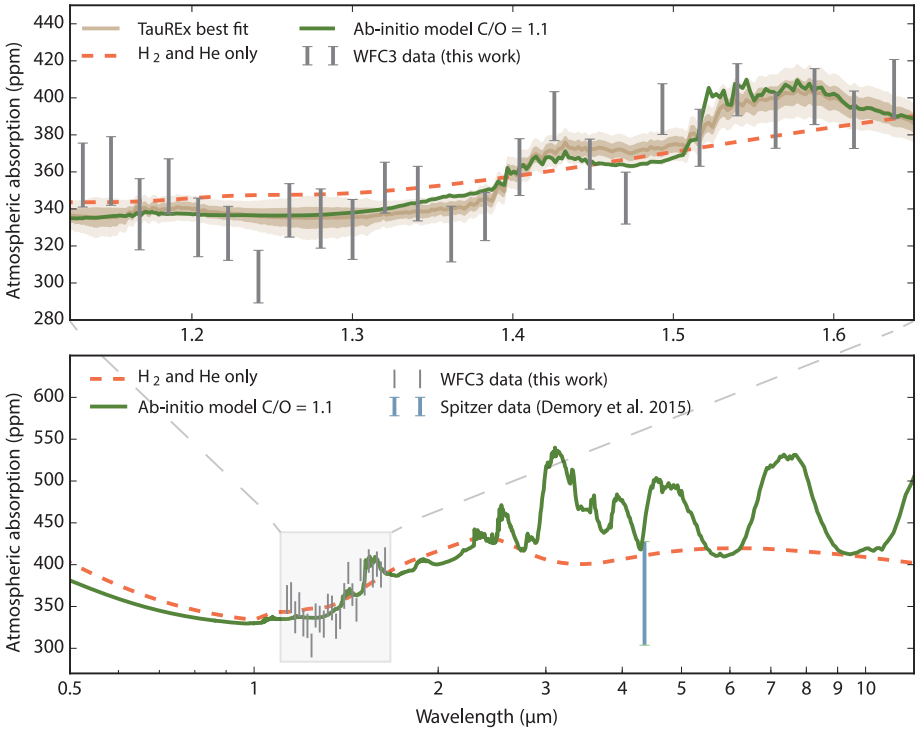

Figure 6.3: Top: Infrared transmission spectrum of the hot super-Earth 55 Cancri e (grey error bars), best fit obtained with TauREx (brown line), fitted model containing hydrogen and helium (dashed orange line) and an ab-initio model with C/O = 1.1 (green line) (Tsiaras et al., 2016). The shaded regions show the 1 and 2 sigma confidence intervals in the retrieved spectrum. Bottom: The same hydrogen/helium and ab-initio models plotted in a broader wavelength range. As we can see the two models can be better distinguished at longer wavelengths. The average transit depth of 55 Cancri e at 4.5 μm obtained with Spitzer Space Telescope (Demory et al., 2016a) is also shown in light blue.

the radius and effective temperature of 55 Cancri. From 115 to 900 nm, we used the stellar flux of ε Eridani (HD22049) from Segura et al. (2003) scaled also to the properties of 55 Cancri. ε Eridani is a K2V star ($T_* = 5084\,\mathrm{K}$ and $R_* = 0.735\,R_\odot$) quite close to 55 Cancri, which is a G8V star ($T_* = 5196\,\mathrm{K}$ and $R_* = 0.943\,R_\odot$), making ε Eridani a quite good proxy for 55 Cancri.

6.3.3 Results

The transmission spectrum of 55 Cancri e and the best fits to it, obtained with TauREx, are shown in Figure 6.3, while Figure 6.4 shows the posterior distributions of the fit to the spectrum, using a model that includes only HCN.

Regardless of the specific gas causing the absorption features on the right hand side of the WFC3 spectrum, the mean molecular weight (μ) of the atmosphere peaks at about 4 amu, as shown by the posterior distribution in Figure 6.4. Higher values for the mean

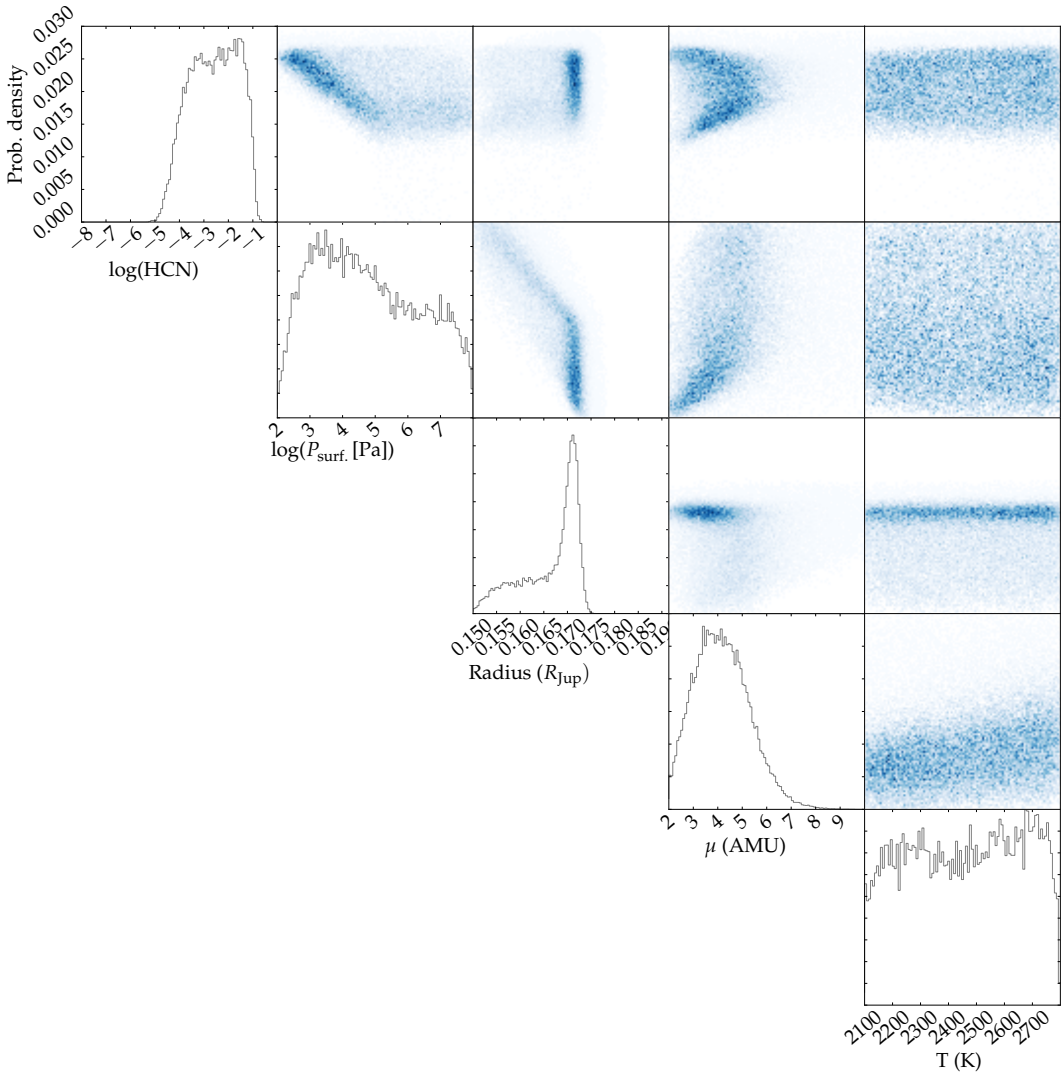

Figure 6.4: Posterior distributions of the retrieved atmospheric parameters and trace gases. Amongst all the molecules considered in the fit, here we only show the posterior of HCN, as all the other molecules show little or no contribution to the spectrum.

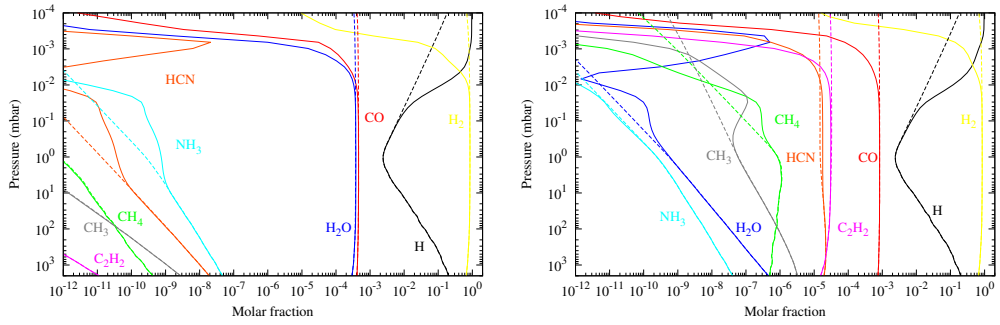

Figure 6.5: Vertical abundance profiles for the scenario with C/O solar (left) and C/O = 1.1 (right). The chemical compositions are calculated with the chemical scheme presented in (Venot et al., 2012), and includes the effects of photochemistry. The chemical equilibrium is also represented (dashed lines).

molecular weight would make the atmosphere more compact, and the features weakened. The relatively strong absorption seen between 1.4 and 1.6 μm indicates that μ is relatively low, and that the atmosphere is likely dominated by a mixture of hydrogen and helium. Therefore, the spectral absorbing features seen in the spectrum should be attributed to trace gases, rather than the main atmospheric component. The posterior of the surface pressure peaks at 0.1 bar, but we note that the parameter is only loosely constrained by the data.

Amongst the molecules considered in the fit, we find that the best absorber that can fit the data is HCN. This species can adequately fit the absorption features seen at approximately 1.42 and 1.54 μm. All the other absorbers show little or no contribution to the overall spectrum. Interestingly, we find no evidence of water vapour. Despite the result of the retrieval, we stress that with the current precision of the measurements and this restricted wavelength range, we cannot confirm the presence or absence of certain absorbers, in particular CO_2, C_2H_2 and CO.

We find that the posterior distribution of HCN tends to favour a scenario with a high absolute mixing ratio, but the acceptable values are broad, starting at 10^{-5}. We note the degeneracy between the HCN mixing ratio and the surface pressure: the lower the HCN abundance, the higher the surface pressure. Lastly, we find that the temperature is very poorly constrained by this data. For our best fitted model, which includes HCN, the scale height of the atmosphere is 242 km, or 14 ppm (assuming $T = 2100\,\mathrm{K}$, $\mu = 3.2\,\mathrm{amu}$, $R_p = 0.17\,R_J$). Given this value, the amplitude of the spectral modulation of 70 ppm, corresponds to 5 scale heights.

We also run other models to verify the validity of this result. We find that a straight line fit has a $\chi^2 = 89.4$, which, with 24 degrees of freedom, indicates that a straight line can be rejected with a 6 sigma confidence level. We also try to fit a model containing only a mixture of hydrogen and helium and no other trace gases. This model, which shows only the H_2-H_2 and H_2-He collision-induced absorption (dashed line in Figure 6.3) has a $\chi^2 = 39.6$.

As the Nested Sampling algorithm implemented in TauREx allows the precise computation of the global evidence of each model, we can also perform model comparison in a Bayesian framework. Table 6.3 summarises the global evidence and χ^2 values of the different models. We find that the straight line model has a log-evidence of 206.8, the model containing hydrogen and helium has $\log \mathscr{E} = 226.2$, and the model containing HCN has $\log \mathscr{E} = 228.8$. The Bayes factor $\mathcal{B}_m$, defined as the ratio between the evidences of two

Table 6.3: Log-evidence ($\log \mathscr{E}$) and χ^2 values for the different models shown in Figure 6.3.

Model	$\log \mathscr{E}$	χ^2
Straight line	206.8	89.4
Helium and hydrogen only	226.2	39.6
Best fit (including HCN)	228.8	23.7
Ab-initio model with C/O = 1.1	–	26.1

different models, shows that we can confidently reject the straight line model. We find that the Bayes factors for the models with HCN or hydrogen and helium are 20 times larger compared to the straight line model, suggesting very strong preference for the former two models according to the Jeffreys' scale (Jeffreys, H, 1961). Finally, the Bayes factor for the model with HCN is only 2.6 times larger compared to the model with hydrogen and helium, suggesting only a moderate preference for the model with HCN.

Interestingly, we find that the spectrum obtained using the results of the ab-initio model with a C/O greater than one is very close to the best fit found with TauREx (see Figure 6.3). In this case we obtain a χ^2 of 26.1. Figure 6.5 shows the vertical abundance profiles for the two cases with C/O ratio solar (left) and C/O = 1.1 (right). It can be clearly seen that the two scenarios are significantly different. For the solar C/O case, the dominant absorbing gases are CO and H_2O, with mixing ratios of $\approx 3 \times 10^{-4}$ and 4×10^{-4} respectively. On the contrary, for C/O = 1.1, while CO still remains the dominant species at 10^{-3}, H_2O decreases to $10^{-7} - 10^{-8}$, and HCN and C_2H_2 increase to about 10^{-5}.

The absolute abundance of HCN expected for a C/O = 1.1 scenario is roughly consistent with what we found in the retrieved spectrum, being at the end of the left tail of the posterior distribution of the fit (Figure 6.4). The transmission spectrum obtained using the abundances profiles for the C/O = 1.1 ratio is shown in Figure 6.3 (green line), and shows that is roughly consistent with the retrieved spectrum. The dominant absorbers in this wavelength range seen in this ab-initio model is HCN, while other relatively strong absorbers, such as CO, C_2H_2 and CH_4, are all hidden below the HCN absorption.

6.3.4 HCN as a tracer of high C/O ratio atmospheres

If the features seen at 1.42 and 1.54 μm are due to hydrogen cyanide (HCN), the implications for the chemistry of 55 Cancri e are considerable. Venot et al. (2015), using a new chemical scheme adapted to carbon-rich atmospheres, pointed out that the C/O ratio has a large influence on the C_2H_2 and HCN content in the exoplanet atmosphere, and that C_2H_2

and HCN can act as tracers of the C/O ratio. Indeed, in a large range of temperatures above 1000 K, at a transition threshold of about C/O = 0.9, the C_2H_2 and HCN abundance increases by several orders of magnitude, while the H_2O abundance decreases drastically, as sown in Figure 6.5.

We conclude that if the absorption feature is confirmed to be due to HCN, the implications are that this atmosphere has C/O ratio higher than solar. However, additional data in a broader spectral range are necessary to confirm this scenario. We also note that while there is a good line list available for the hot HCN/HNC system (Harris et al., 2006; Barber et al., 2014a), there is no comprehensive line list available for hot C_2H_2; provision of such a list is important for future studies of this interesting system.

6.3.5 Summary and conclusions

In this work, published in Tsiaras et al. (2016), I presented the first analysis of the two HST/WFC3 scanning-mode spectroscopic observations of the super-Earth 55 Cancri e. The observed spectrum was analysed with the TauREx retrieval code. These are the most important results:

- The planet has an atmosphere, as the detected spectral modulations are 6 sigma away from a straight line model.
- The atmosphere appears to be light-weighted, suggesting that a significant amount of hydrogen and helium is retained from the protoplanetary disk.
- There is no evidence of water vapour.
- The spectral features at 1.42 and 1.54 μm can best be explained by HCN, with a possible additional contribution of other molecules, such as CO, CO_2 and C_2H_2.
- This scenario is consistent with a carbon-rich atmosphere (e.g. C/O ratio = 1.1) dominated by carbon bearing species. The model for such an atmosphere was computed independently.

While these results have important implications to the study of 55 Cancri e and other super-Earths, further spectroscopic observations in a broader wavelength range in the infrared are needed to confirm our conclusions.

Chapter 7

Influence of stellar flares on exoplanets spectra

In chapters 4 and 5 I have investigated the effects that different approximations in the modelling of exoplanet spectra have on the final interpretation. We saw that the choice of accurate absorption cross sections is especially important, and that common assumptions in exoplanet retrievals need to be carefully considered when high signal-to-noise, broad wavelength range spectra will become available with the advent of JWST.

In this chapter, we shift the attention to potential external factors that could influence the observation of exoplanet spectra. I investigate how perturbations coming from the host star, specifically stellar flares, have on the final transmission spectra of cloud-free exoplanets. Stellar flares are known to be common in some stars: sudden increase of their irradiation can last several hours, and are likely to affect the chemistry of the orbiting exoplanets. This has also an effect on the observations of exoplanet spectra, as the timescale of this events is comparable to the timescale of the observations.

In here, I aim at answering the following questions: How can a stellar flare modify the composition of a planetary atmosphere and influence the resulting spectra? What are the potential biases of such a phenomenon in the interpretation of exoplanetary spectra?

In order to answer these questions, one-dimensional thermo-photochemical models of hypothetic planets located around the active star AD Leo were computed. The evolution of the chemical composition of the atmosphere was then evaluated using spectroscopic temporal data from the flare event, and synthetic transmission spectra were then simulated to evaluate the impact of the changing chemistry on the final observations. Significant effects were found in the chemical abundance of some species, that can lead to detectable variations

in the planetary spectra.

The methods and results presented in this chapter are published in Venot et al. (2016) and are here reproduced with permission from the publisher.

7.1 Introduction

M-dwarf stars are very abundant in the Galaxy, and, thanks to their relatively small radius, are amongst the best candidate targets for transit searches of small, possibly rocky, exoplanets (see e.g. Nutzman & Charbonneau, 2008). These stars are however known to be particularly active, exhibiting high stellar variability, such as star sports, granulation or stellar flares. Flares events are caused by magnetic processes, and result in violent and unpredictable outbursts of the photosphere. In some cases, during such events, the total emitted radiative energy can increase by several order of magnitudes across the entire wavelength range.

Segura et al. (2010) studied the effect of star activity on the habitability of orbiting exoplanets, and found that the enhanced UV radiation emitted during a flare would not affect the habitability of the planet. These authors, however, did not address the effect of the flare on the planetary spectra. A subsequent study (Tofflemire et al., 2012) found that a typical stellar flare would have no impact on the infrared detection and characterisation of exoplanets above a level of 5-10 mmag, but again did not consider the effects that the flare has on the atmospheric chemical composition, nor on the exoplanet spectra.

In this study the focus is on how a stellar flare, which has a timescale comparable to the timescale of exoplanet observations, has on both the chemical composition of typical H_2/He dominated super-Earths, and on their transmission spectra.

7.1.1 Acknowledgments

This work, published in Venot et al. (2016), has been possible thanks to the additional contributions of other authors (O. Venot, S. Carl, A.R. Hashim, L. Decin). In particular, O. Venot computed the atmospheric chemical models, while I computed the synthetic spectra and determined the impact that the varying chemistry has on the final synthetic observations of the planetary spectra.

7.2 Chemical models

To model the chemical composition of the atmosphere, we used the same 1D atmospheric chemical model presented in the previous chapters (Chapter 5, Section 2.1 and Chapter 6,

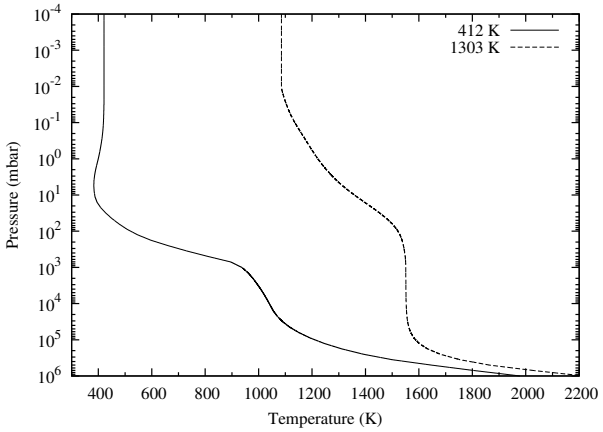

Figure 7.1: Thermal profiles of the hypothetic AD Leo b planets.

Section 3.2). These models were developed for hot atmospheres (Venot et al., 2012, and references therein), and have been extensively used to study the atmosphere of exoplanets (Venot et al., 2014; Agúndez et al., 2014; Venot et al., 2015; Venot & Agúndez, 2015; Tsiaras et al., 2016). The simpler, and computationally faster, model that includes species up to four carbon atoms was used instead of the more complicated scheme including species up to six carbon atoms, as this has little effect on the final synthetic spectra (Venot et al., 2015). This scheme includes 105 neutral species and 960 reactions (and their reverse reactions). We used a constant eddy diffusion coefficient, $K_{zz} = 10^8$ cm^{2}s^{-1}, to parametrise the vertical mixing (Fortney et al., 2013). This value is commonly used in the study of exoplanetary atmospheres (e.g Lewis et al., 2010; Moses et al., 2011; Line et al., 2011). The simulated planets have radius $R_p = 0.238\,R_J$ and mass $M_p = 0.02\,M_J$. Two thermal profiles were used: one with a high altitude temperature of 412 K, and one with 1303 K. These thermal profiles are taken from Fortney et al. (2013), which computed several temperature profiles for a GJ 1211b-like planet, corresponding to different irradiations of the planet. These TP profiles, shown in Figure 7.1, have been computed using a fully non-gray atmosphere code adapted to exoplanet atmospheres (Fortney et al., 2005, 2008; Morley et al., 2013), and consider a 50$\times$ solar metallicity, an internal temperature $T_{\rm int} = 60$ K, and a low Bond albedo of ~ 0.05 due to the absence of clouds. We consider these profiles to be appropriate to model the "AD Leo b" planets considered here, as both are super-Earths orbiting M stars with similar effective temperatures (AD Leo and GJ 1214 have $T_{\rm eff} = 3390 \pm 19$ and 3245 ± 31 K respectively; Rojas-Ayala et al., 2012). The distances from the host star of the two planets with these two TP profiles are 0.007 and 0.07 AU for the warmer and cooler planet respectively. The

existence of super-Earths so close to their parent stars could potentially raise some doubts, but we note the recent discovery of an exoplanet with similar properties (K2-22b; Sanchis-Ojeda et al., 2015).

Stellar fluxes for AD Leo were taken from Segura et al. (2010), who constructed a time-dependent sequence of spectra during a flare event that can be used in photochemical models. These spectra were obtained both in the visible and UV spectral region during a great flare event observed in 1985 (Hawley & Pettersen, 1991). Additional details on how the stellar fluxes were computed are discussed in Venot et al. (2016), Section 2.2. The time evolution of the stellar flare can be divided into three phases:

- The first impulsive phase, lasting 800 s. The stellar flux increases rapidly by one order of magnitude.
- The second impulsive phase, from 900 to 1500 s, during which the stellar flux decreases by less than one order of magnitude over the same timescale of the first phase.
- The gradual phase, during which the stellar flux continues to decrease, lasting from 1600 to 2600 s.

The different stellar fluxes during these three phases are shown in Figure 7.2, together with their quiescent fluxes.

7.3 Spectral models

Synthetic transmission spectra in the wavelength range 0.8–20 μm for the different atmospheric compositions were obtained using the forward model included in TauREx. The 1D radiative transfer model (described in Chapter 2) is based on a code that calculates the optical path through the planetary atmosphere, resulting in a transmission spectrum of transit depth as a function of wavelength. The wavelength and temperature-dependent absorption cross sections for the absorbing molecules were computed using the line lists from ExoMol (Barber et al., 2006; Harris et al., 2006; Yurchenko et al., 2011; Tennyson & Yurchenko, 2012; Yurchenko et al., 2013; Barber et al., 2014a), HITRAN (Rothman et al., 2013) and HITEMP (Rothman et al., 2010). We only considered the following molecules: H_2O, CO_2, OH, HCN, NH_3, CH_4 and CO. Although these molecules represent only a fraction of the 105 molecules considered in the thermo-photochemical model, they are the most abundant and therefore dominate the spectral features. We also included collisional-induced cross

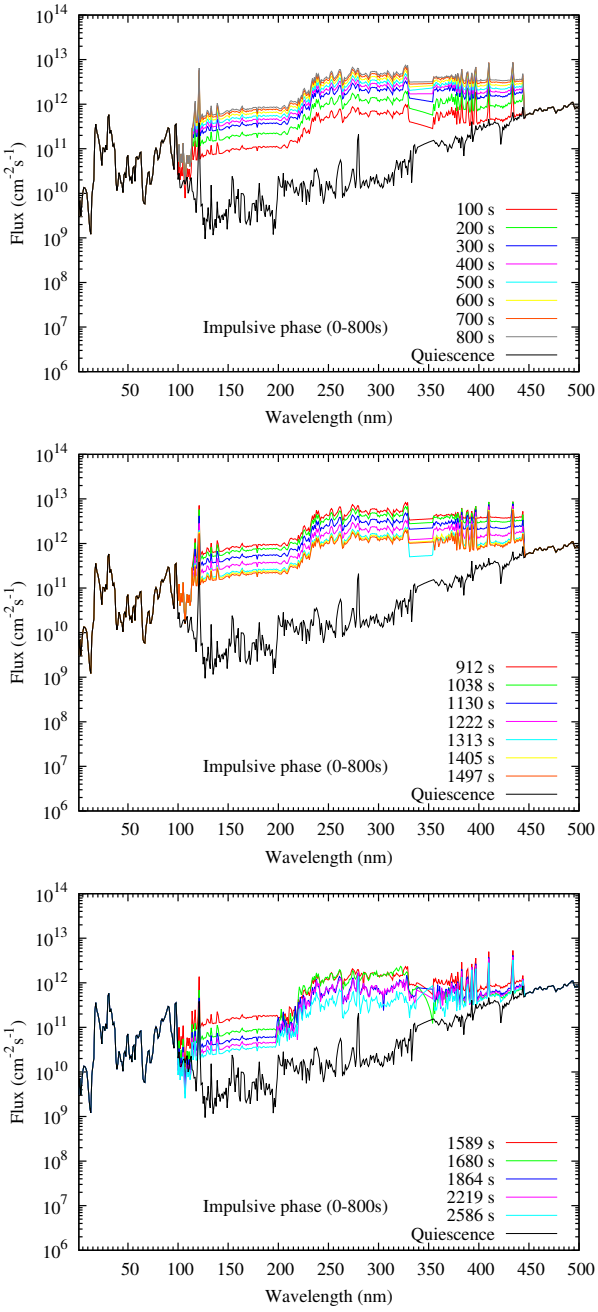

Figure 7.2: Stellar spectra of AD Leo at a distance of 1 AU from the star, during the three phases of the flare: the first impulsive (top), second impulsive (middle) and gradual (bottom). Spectra at different time steps during the flare are shown with different colours, as shown by the legend. These spectra are used as input for the photochemical models at the corresponding time steps. The quiescent spectra are shown in black.

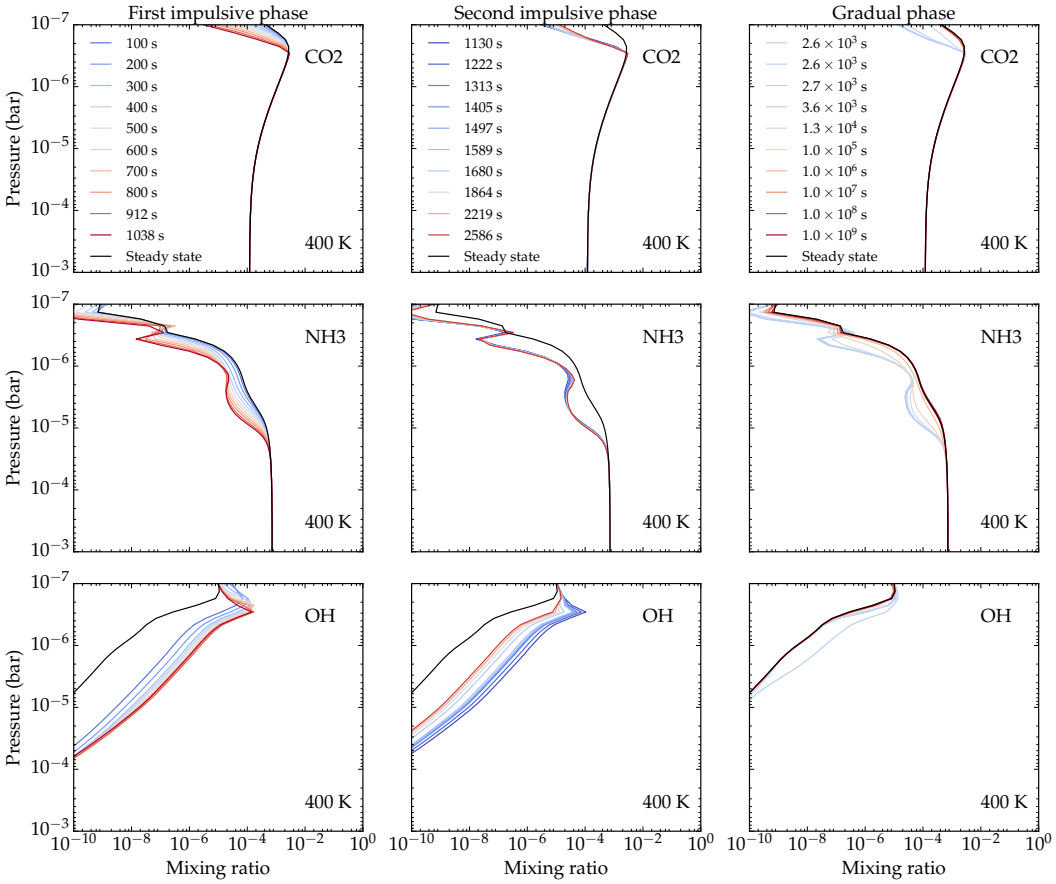

Figure 7.3: Evolution of CO_2, NH_3, and OH mixing ratios during the different phases of the flare event, for the cooler planet. The legend for all figures is in the upper panels.

sections for H_2–H2 and H_2–He (Borysow et al., 2001; Borysow, 2002). The atmosphere is assumed to be cloud-free.

7.4 Results

7.4.1 Chemical composition

The species that undergone the most significant changes are the most abundant ones: hydrogen (H), amidogen (NH_2), ammonia (NH_3), carbon dioxide (CO_2), nitric oxide (NO) and the hydroxyl radical (OH). The abundances of three of these molecules (NH_3, CO_2 and OH) during the different phases of the flare are shown in Figure 7.3 for the lower temperature case, and in Figure 7.4 for the higher temperature case. Additional cases are plotted in Venot et al. (2016).

In both cases, most species see their mixing ratios globally increase during the flare event, with a maximum abundance around 912 s (corresponding to the peak of the stellar

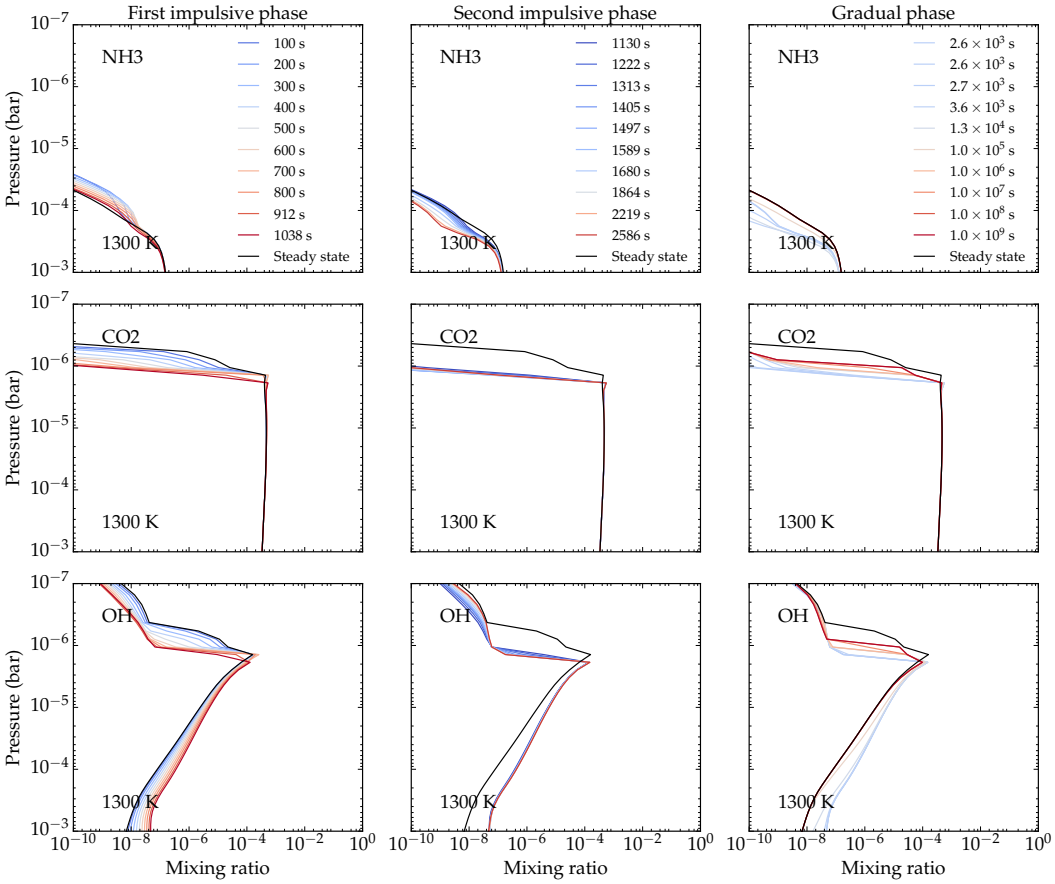

Figure 7.4: Evolution of CO_2, NH_3, and OH mixing ratios during the different phases of the flare event, for the hotter planet. The legend for all figures is in the upper panels.

flux). Exceptions are NH_3, CO_2, and NO, which can have a lower abundance during the stellar flare than at steady-state. The other species considered in this calculation, but not plotted (H_2O, HCN, and CH_4), experienced changes only in the very upper atmosphere ($P < 10^1$ mbar and $P < 4 \times 10^4$ mbar for the hotter and cooler TP profiles, respectively). Their abundances globally decrease by 1 to 5 orders of magnitude.

Figure 7.5 shows a comparison between the atmospheric composition of the initial and final steady-states. We can see that the final steady-state of the hot atmosphere (reached 10^{12} s after the end of the flare, i.e. ∼30,000 years) is significantly different from the initial steady-state. Smaller differences are found between the initial and final steady-states of the cooler atmosphere.

The interested reader is referred to Section 3.1 in Venot et al. (2016) for a more comprehensive discussion on the changing chemical composition of these atmospheres.

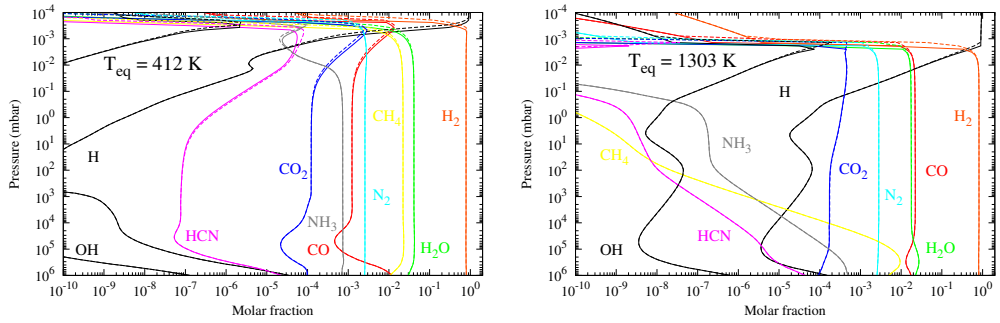

Figure 7.5: Atmospheric composition at the initial steady-state (dashed line) and at the final steady-state (solid line) with the thermal profile corresponding to cooler (top) and hotter (bottom) TP profiles. They correspond to the compositions obtained 1012 s after the flare event.

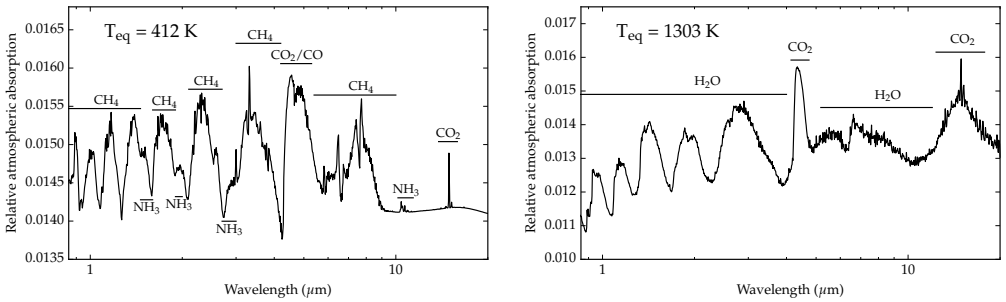

Figure 7.6: Synthetic transmission spectra for the initial steady-state for the cooler (left) and hotter (right) planets. The resolving power is 100 across the entire spectrum.

7.4.2 Synthetic spectra

We computed synthetic transmission spectra during the different phases of the flare and compared them to that of the initial steady-state. Figure 7.6 shows the instantaneous transmission spectra for the initial steady-state for the cooler and hotter planets respectively. The planetary radius is defined at 1 bar. The spectra are binned at a constant resolving power of $R = 300$. Figure 7.7 shows the differences between the initial steady-state and the transmission spectra obtained during the different phases of the flare. It can be seen that while small changes occur in the warm atmosphere, significant changes occur in the hotter planet atmosphere. The strongest changes are seen in the CO and CO_2 features at 4.6 and 14 μm in the hotter planet case, with total amplitude variations of, respectively, 220 and 200 ppm in the impulsive phase, and up to 500 ppm in the return to quiescence (or gradual) phase. Other changes with amplitude smaller than about 50 ppm are seen in other parts of the spectra, especially in the return to quiescence phase. Weaker variations of only 40 ppm are seen in the hotter planet, especially in the CO and CO_2 features at 4.6 μm and 14 μm. Other very

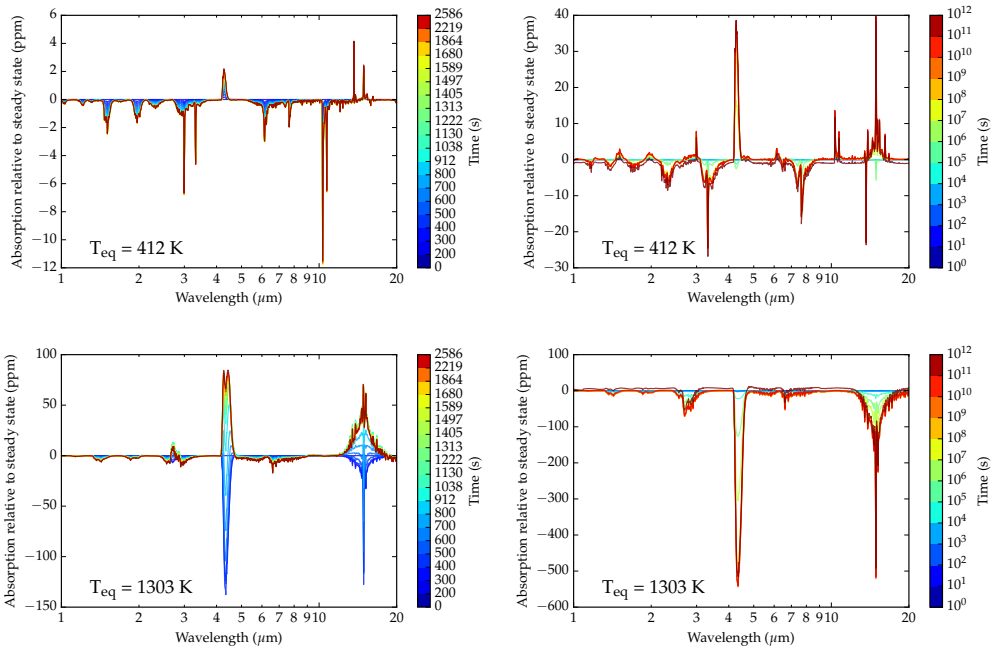

Figure 7.7: Difference in relative absorption between the initial steady-state and the instantaneous transmission spectra obtained during the different phases of the flare, for the cooler (top) and hotter (bottom) planets. The left plots refer to the impulsive and gradual phases, while the right plots to the return to quiescence phase. The colour legends of the right plots are represented with a logarithm scale. The resolving power is 100 across the entire spectrum.

small changes of about 20 ppm are seen in other parts of the spectra.

Two effects contribute to the observed changes. Firstly, changes in the molecular abundances vary the strength of the absorption features specific to the different molecules. Secondly, these changes modify the mean molecular weight for the different atmospheric layers, amplifying or reducing the amplitude of the absorption features throughout the entire spectrum. This is because the scale height, i.e. the altitude at which the atmospheric pressure decreases by $1/e$, is inversely proportional to the mean molecular weight ($H = k_B T/\mu g$; where k_B is the Boltzmann constant, T is the atmospheric temperature, g is the gravity acceleration and μ is the mean molecular weight). Relatively small variations in mean molecular weight are seen in the warm planet case as the flare evolves. For the hotter planet case, we found that μ decreases (and hence H increases) at high altitudes (i.e. low pressures) during the flare, indicating that the density of the upper atmosphere decreases. However, we found that the decrease in abundances in the atmosphere outweighs the effect of an increased scale height on the transmission spectra.

We note that while the mean molecular weight for each atmospheric layer is calculated

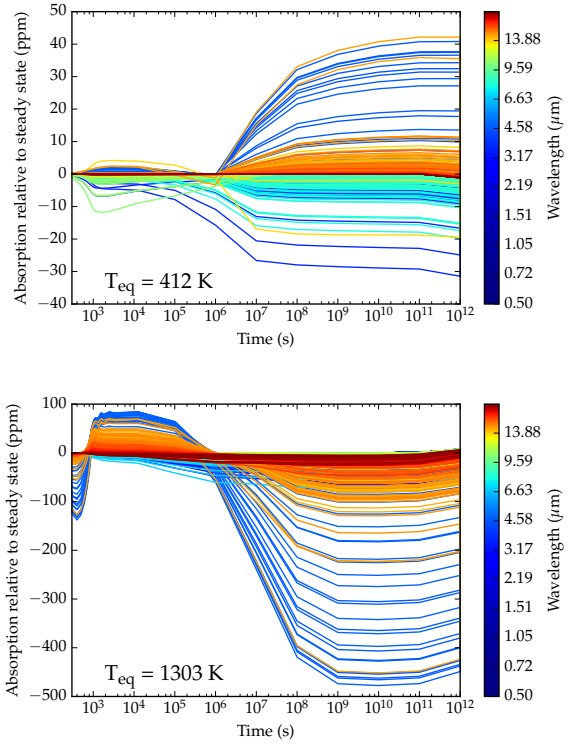

Figure 7.8: Systematic shifts as a function of time (x-axis) and wavelength (colourbar) with respect to the steady-state, for the cooler (top) hotter (bottom) planets. Each line represent a wavelength bin corresponding to a constant resolving power of 50. This is the shift, relative to the initial steady-state, as a function of wavelength and time that would be seen in the transit depth. It is equivalent to Figure 7.7, but with the x-axis and colourbar inverted.

taking into account the entire set of 105 molecules contained in the chemical model, we only compute the spectral opacities for a set of 7 molecules which are the most abundant. The changes observed in the transmission spectra are therefore limited to these molecules. Other absorbers that might contribute to additional observing features will however show much weaker – if not unobservable – features.

7.4.3 Effects on current and future observations

Transmission spectra of exoplanetary atmospheres are obtained by observing simultaneously the transit of the planet at different wavelengths. By tracing the transit depth as a function of wavelength it is possible to reconstruct the transmission spectrum of the planet, as different molecules absorb differently in different spectral regions, making the planet appear smaller or larger at different wavelengths. A varying atmospheric composition during the transit event will introduce a time and wavelength-dependent shift on the transit light curve.

Figure 7.8 shows the expected change in transit depth as a function of wavelength and time with respect to the transit depths measured during the initial steady-state. It must be noted that the timescale of the impulsive phase of the flare is of the same order of magnitude as the typical transit duration ($\sim 10^4$ s). Therefore, if a transit is observed during the flare, the transit depth would change during the transit, and shifts of up to 500 ppm are expected for the hotter planet. Clearly, these changes will be wavelength dependent, as shown by Figure 7.6. Changes of only ≈ 40 ppm would be seen in the cooler planet.

The bulk of recent observations of exoplanetary atmospheres have been obtained from space using the Hubble Space Telescope Wide Field Camera 3 (HST/WFC3), covering the $1.1 - 1.7$ μm range. Observations of hot Jupiters with WFC3 claim an uncertainty in transit depth per wavelength bin of 40 – 100 ppm (Fraine et al., 2014; Deming et al., 2013; Knutson et al., 2014a; Tsiaras et al., 2015), for resolving powers of about 50. For smaller super-Earths such as GJ1214b, an uncertainty of 60 ppm was achieved (Kreidberg et al., 2014), but only by combining observations of 12 transits, while for 55 Cancri e an uncertainty in transit depth of 22 ppm was achieved for a single transit observation (Tsiaras et al., 2016).

The predicted changes in the spectra during a flare are comparable to the sensitivity of the current and future instruments, especially for the hotter planet case, where relative changes of up to 500 ppm are seen. We note however that the current instrumentation cover spectral ranges that are either outside the spectral features with the strongest changes (HST/WFC3 and STIS), or have only photometric channels with relatively low sensitivity (Spitzer/IRAC). Only with future instruments such as those on-board JWST, which will cover a spectral range including the CO and CO_2 features that show the strongest variations, it will therefore be possible to observe changes of spectral features of planetary atmospheres due to stellar activity in the form of flares.

7.4.4 Recurrence of flares

A flare is not a unique event in the lifetime of a star. Studies on the flare activity of AD Leo (Pettersen et al., 1984), show that consecutive flares occur every 1 to 137 minutes. Such recurrent events have also been observed on other active stars (e.g. Lacy et al., 1976). In order to study the effect of recurrent stellar flares on the spectra of the exoplanet modelled here, we ran the same time-dependent chemical model described above, but imposed a period between consecutive flares of 5 h. We simulated a total of 47 flare cycles, representing about ten days. We found that the mixing ratios of the major species vary at each flare event,

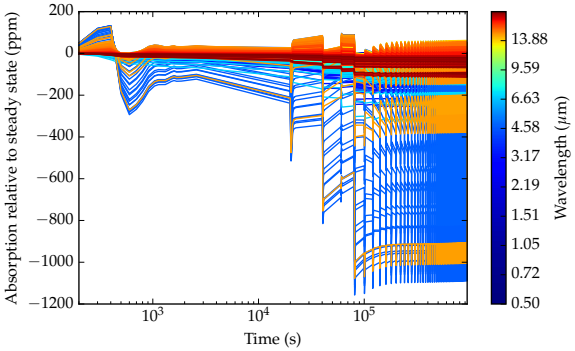

Figure 7.9: Systematic shifts as a function of time (x-axis) and wavelength (colourbar) with respect to the steady-state, for the case in which the flare comes back every 5 h. Caption as in Figure 7.8.

but a steady-state is reached after $\sim 10^5$ s for most molecules.

Figure 7.9 shows the shifts as a function of wavelength and time that would occur over the period of ten days assuming that the flare recurs every five hours. It can be seen that, compared to a single flare event, the changes in the transmission spectrum are even more dramatic. We see variations of up to 1200 ppm after one day, which would be detectable with future instruments onboard JWST, but also with current instruments. This shows even more clearly how strongly stellar activity influences the atmospheric chemistry and the resulting spectra.

7.5 Summary and conclusions

In this study, we investigated how the activity of a star can influence the chemical composition and resulting spectra of typical exoplanets. We focused on the effect of stellar flares, and found significant changes on the chemistry of the atmospheres of two typical planets around an active M star. These changes affect the transmission spectra of these planets, and the resulting differences are potentially observable with future observations, such as those expected from JWST.

A major simplification of this work is that we did not take into account that the thermal profile of the planets can vary during the flare event. It is expected that the temperature profile should be modified both by the increase of the irradiation and the change of the atmospheric composition. Such changes could potentially increase the differences seen here. However, if such changes in temperature are of the same order of magnitude as those found in Segura et al. (2010) (less than 8 K), we do not expect the results presented here to

be affected. Additional work should still be done in this direction.

Another aspect not discussed here is the effect of energetic particles. Flare events are accompanied by ejection of energetic particles (Segura et al., 2010). Such particles could potentially affect the chemical composition of the atmospheres studied here. Some studies suggest that the effects are only likely to affect the very low-pressure region of the atmospheres (Yelle, 2004; Koskinen et al., 2007), but other studies focusing on planets orbiting young stars with strong X-ray emission show that these emitted particles could potentially affect the atmosphere down to 1 mbar (Koskinen et al., 2007). It would therefore be interesting to quantify the effect of the increase of energetic particles expected during a stellar flare on the planetary spectra.

Concerning the spectra, we assumed a cloud-free atmosphere, in accordance to the thermal profiles and the atmospheric model. If if clouds were considered, their effect would be to flatten the spectra, especially at small wavelengths. However, a cloud deck would not hide the strongest features at long wavelengths (such as the ones caused by CO_2). Given that the spectral changes are seen at these strong features, the actual differences are expected to stay relatively constant.

Finally, although the investigation of this effect is beyond the scope of this study, we also note that the spectrum of the star varies as the flare evolves, introducing another systematic and time dependent change on the transit light curves. However, stronger variations will only be seen in the bluest bands (< 450 nm) compared to the red and in infrared, since emission in the blue is due to material heated by magnetic reconnection up to tens of thousands of degrees (Davenport et al., 2012).

Chapter 8

The chemistry and frequency of planetary debris around white dwarfs

In the previous chapters I have discussed at length how the spectra of exoplanet atmospheres can provide valuable information about their chemical composition and temperature structure. Atmospheric spectra of exoplanets, however, provide very little information about their internal composition, especially in the case of rocky, terrestrial planets. While the determination of mass and radius can provide some clues about their internal bulk composition, the detailed chemistry remains always unaccessible. In this context, metal polluted white dwarfs offer a compelling way to study the detailed composition of such planets.

White dwarf stars, due to their high surface gravity, present a pristine hydrogen or helium atmosphere, which can become contaminated by the accretion of small rocky planetesimals, which are in turn remnants of ancient planetary systems. Owing to the very short sinking timescales, the spectra of such metal polluted atmospheres, in principle, mirror the composition of the accreted material, allowing us to infer the detailed composition of the accreted rocky debris. This method is unique in its kind, as no other technique allows us to infer in detail the bulk composition of rocky extra-solar matter. In addition, the statistics of this phenomenon allows us to place tight constraints on the frequency of planetary debris at white dwarfs, and, in turn, around their progenitor main-sequence stars, providing an independent technique to further constrain planet formation and evolution scenarios.

In this chapter I will firstly show how metal lines seen in the atmosphere of white dwarfs have been interpreted over the years as a sign of external accretion of rocky material. Then, I will present the results of the first unbiased survey searching for circumstellar debris orbiting a homogeneous and well-defined sample of white dwarfs. This study provides the

first tight constraint on the frequency of circumstellar discs at white dwarfs, originating from the disruption of small asteroids and planetesimals. Lastly, I will also present the distribution of disc fractional luminosity as a function of cooling age for all known white dwarfs showing circumstellar dust, and suggest possible disc evolution scenarios.

The observations, methods and results presented in this chapter are published in (Rocchetto et al., 2015), and are here reproduced with permission from the publisher.

8.1 Introduction: the detailed composition of rocky planetary remnants

In the context of planetary system characterisation, white dwarfs offer a unique laboratory to study exoplanetary compositions. It is now clear that planetary systems around Sun-like and intermediate-mass stars survive, at least in part, the post-main sequence phases of their hosts (Zuckerman et al., 2010). Compelling evidence comes from metal polluted white dwarfs that commonly exhibit closely orbiting circumstellar dust discs originating from the disruption of large asteroids or planetesimals (Zuckerman & Becklin, 1987; Becklin et al., 2005; Gänsicke et al., 2006; von Hippel et al., 2007; Farihi et al., 2008b). Owing to high surface gravity and negligible radiative forces, heavy elements sink on relatively short timescales within the atmospheres of relatively cool ($T_{\rm eff} \lesssim 25\,000$ K) white dwarfs if compared to the evolution timescales (Fontaine & Michaud, 1979; Paquette et al., 1986). Therefore, the presence of metals in the atmospheres of cool white dwarfs must be a sign of recent external accretion (Zuckerman et al., 2003). The source of this accreting material was initially attributed to the interstellar medium (Dupuis et al., 1992, 1993a,b) or comets (Alcock et al., 1986), but both theories had trouble in explaining the high and ongoing accretion rates found at hydrogen dominated white dwarfs (Zuckerman et al., 2003). Today accretion from circumstellar material, resulting from the disruption of large asteroids or minor planets, is, by far, the most compelling explanation for atmospheric metals seen at a large fraction of cool white dwarfs (Jura, 2003; Veras et al., 2013).

Metal enriched white dwarfs have become a powerful tool to indirectly analyse the composition of exoterrestrial planetary matter, as their photospheres, in principle, mirror the composition of the accreted material. As an example of this technique, it was demonstrated that the relative abundances of 15 heavy elements in the atmosphere of GD 362 reflect the composition of a large asteroid that was similar in composition to the bulk Earth-Moon

system (Zuckerman et al., 2007). Notably, ultraviolet and optical spectroscopy have shown that metal-contaminated degenerates are, in general, refractory-rich and volatile-poor (Wolff et al., 2002; Dufour et al., 2007; Desharnais et al., 2008; Klein et al., 2010; Gänsicke et al., 2012), while infrared spectroscopy reveals that the circumstellar dust itself is silicate-rich and carbon-poor (Reach et al., 2005, 2009; Jura et al., 2009), and thus similar to materials found in the inner solar system (Lodders, 2003).

Such stars can also provide constraints about the presence of water in extrasolar planetary systems. For instance, the circumstellar disc identified around the white dwarf GD 61 (Farihi et al., 2013) resulted from the destruction of a rocky and water-rich extrasolar minor planet, demonstrating the existence of water in terrestrial zone planetesimals that could play an important role in delivering water to the surface of planets.

A fraction of metal polluted white dwarfs exhibits circumstellar dust discs, visible in the infrared as excess emission above the photospheric flux of the white dwarf atmosphere. The detailed modelling of such infrared excesses suggests that the circumstellar dust is arranged in the form of an optically thick but geometrically thin disc, with similar properties to the rings of Saturn (Rafikov & De Colle, 2006; Jura, 2003). These rings of warm dust are situated within the Roche limit of their host star, as also confirmed by the emission profiles of gaseous debris discovered at several dusty white dwarfs (Gänsicke et al., 2006; Brinkworth et al., 2009, 2012; Debes et al., 2012a). There is substantial theoretical evidence that these discs are created via the tidal disruption of post-main sequence planetary systems, perturbed by unseen distant planets (Debes & Sigurdsson, 2002; Veras et al., 2013, 2014). The transition from disruption to disc is still poorly understood but there are good models for the evolution of these metal dominated discs (Bonsor et al., 2011; Metzger et al., 2012; Debes et al., 2012b; Frewen & Hansen, 2014).

8.1.1 The statistics of metal-polluted white dwarfs

Despite the large number of metal polluted white dwarfs and circumstellar discs detected, the statistical frequency of the phenomenon still suffers from significant observational biases. While the fraction of metal polluted white dwarfs that are currently accreting has been constrained by several unbiased surveys to about 20–30% (Koester et al., 2014; Zuckerman et al., 2010, 2003), surveys aiming to detect infrared bright dust discs at white dwarfs suffer from considerable biases. The first searches for infrared excesses targeted relatively cool ($T_{\rm eff} \lesssim 25\,000$ K) stars known to be metal-polluted (Farihi et al., 2010a; Jura et al., 2007;

Debes et al., 2007), and this approach does not permit robust statistics of disc frequency over the entire white dwarf population. With little restriction of stellar effective temperature, wide field surveys such as SDSS, UKIDSS, and WISE have found disc frequencies between 0.4 and 1.9% (Girven et al., 2011; Steele et al., 2011; Debes et al., 2011b). On the other hand, more sensitive Spitzer observations of K-band bright white dwarfs, without regard to stellar temperature, resulted in a nominal disc frequency of 1.6% (Mullally et al., 2007). A higher frequency of 4.5% was obtained by Barber et al. (2014b), by targeting stars in a restricted temperature range where dust detections were expected based on prior surveys. Moreover, the observed sample was fragmented over several instruments and hence the result is difficult to compare with other surveys. It is therefore clear that the current picture looks somewhat incomplete, and additional work is needed to provide definite statistics of this phenomenon.

In the remainder of this chapter, I will present a study, originally published in Rocchetto et al. (2015), aimed at determining definite statistics of the frequency of infrared bright discs at white dwarfs. In Section 8.2 I describe the Spitzer observations of an unbiased sample of white dwarfs, in Section 8.3 I discuss how the spectral energy distributions of the sample stars were obtained and I describe those with detected discs. Sections 8.4 presents the derived disc frequency in the context of complementary HST ultraviolet observations, and Section 8.5 discusses the distribution of the fractional disc luminosity for all known dusty white dwarfs. Lastly, in Section 8.6 I give a short conclusion. Notes on the individual objects are presented in Appendix B.

Acknowledgments

This work has been possible thanks to the additional contributions of other authors (J. Farihi, B. T. Gaensicke and C. Bergfors). In particular, J. Farihi was the PI of the Spitzer observing program, and gave invaluable and meticulous help in the write up, B.T. Gaensicke was the PI of the Hubble survey, while C. Bergfors gave significant overall support.

8.2 Observations: An unbiased sample of white dwarfs

8.2.1 Sample selection

The sample of hydrogen dominated white dwarfs was selected from the catalogues compiled by Liebert, Bergeron, & Holberg (2005) and Koester et al. (2009), who performed model atmosphere analyses based on optical spectroscopy, providing effective temper-

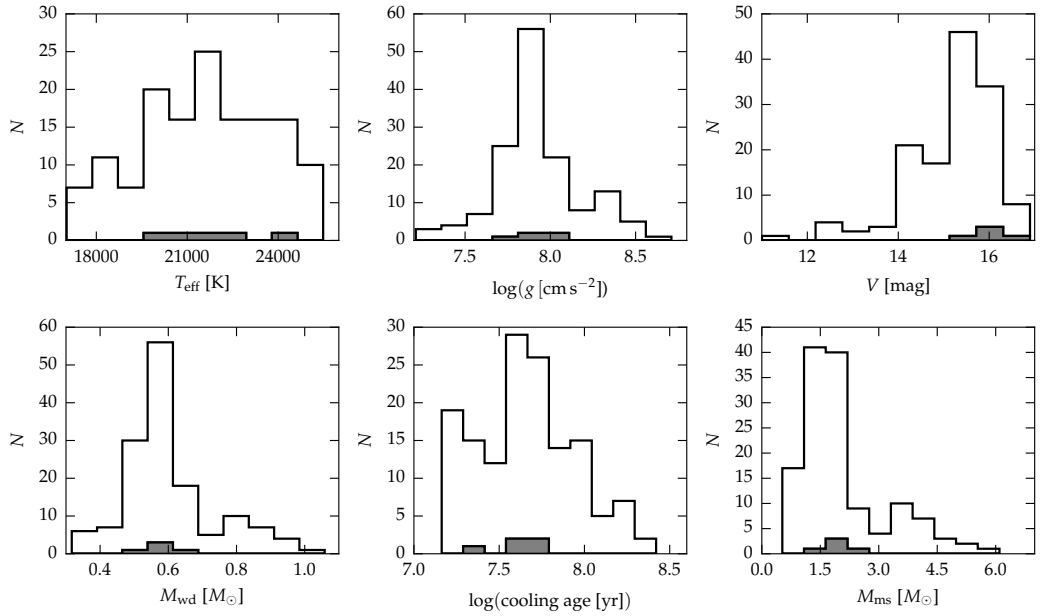

Figure 8.1: Fundamental properties of the 134 DA stars observed in the Spitzer survey. Effective temperature $T_{\rm eff}$ and surface gravity $\log g$ are taken from Liebert et al. (2005) or Koester et al. (2009), and V-band magnitudes are taken from the AAVSO Photometric All-Sky Survey (APASS). Final white dwarf mass $M_{\rm wd}$ and cooling age are derived using evolutionary cooling sequences (Fontaine, Brassard, & Bergeron, 2001); the initial main-sequence progenitor mass $M_{\rm ms}$ is derived from $M_{\rm wd}$ and the initial-to-final mass relation (Kalirai et al., 2008). The entire sample of white dwarfs is shown with the unfilled histograms, while the degenerates with infrared detected discs are shown in grey.

ature, surface gravity, mass and cooling age. The only criteria for selection were 1) $17\,000\,{\rm K} < T_{\rm eff} < 25\,000\,{\rm K}$ and corresponding young cooling ages of 15–270 Myr, 2) predicted fluxes $F_\lambda(1300\,{\rm \AA}) > 5 \times 10^{-14}\,{\rm erg\,cm^{-2}\,s^{-1}\,\AA^{-1}}$ for the corresponding HST/COS Snapshot survey. As such, the selection was performed only for temperature and brightness, resulting in an unbiased sample of 134 DA young single white dwarfs. The distributions of the fundamental stellar parameters are shown in Figure 8.1.

8.2.2 Spitzer observations

A total of 100 sample stars were observed between 2012 May and October in the 3.6 and 4.5 μm bandpasses using the Infrared Array Camera (IRAC; Fazio et al., 2004) on-board the Spitzer Space Telescope (Werner et al., 2004) as part of Program 80149[1]. An exposure time of 30 s was used for each individual frame, with 20 medium-size dithers in the cycling

[1]Five stars were also observed as part of the same Program but are not included in this study. The white dwarf 1929+012 was observed as an ancillary target (see Appendix B.1.2) (PI: J. Farihi) while four white dwarfs (0933+025, 1049+103, 1335+369, 1433+538) initially considered to be single stars and included in the original sample were found to host unresolved M dwarf companions and were therefore excluded. Infrared fluxes for these four binaries are reported in Appendix B.

pattern, resulting in 600 s total exposure time in each warm IRAC channel. The remaining 34 sample stars were previously observed during either the cold or warm mission, and their archival data were analysed.

We performed the analysis of all targets using the $0.6''$ pixel^{-1} mosaics (Post-Basic Calibrated Data) processed by the IRAC calibration pipeline version S 18.25.0 which produces a single, fully processed and calibrated image. Aperture photometry was performed using the point source extraction package APEX within MOPEX (Makovoz et al., 2006) and an aperture radius of 4 pixels with a 24–40 pixels sky annulus. Fluxes were corrected for aperture size, but not for colour. For blended sources, point response function (PRF) fitting was performed on the Basic Calibrated Data frames, using the package APEX MULTIFRAME within MOPEX. Fluxes obtained with PRF fitting were compared with the DAOPHOT PSF fitting routine within IRAF for a representative sample, and led to consistent results to within 4%, and always within the relative uncertainties. The measured flux uncertainty was computed by APEX and include the source photon noise and the variance in sky background. A 5% calibration uncertainty is conservatively added in quadrature to all IRAC fluxes. The flux determinations and uncertainties for the science targets, together with the physical parameters of the stars, are reported in Appendix B.

8.3 Data Analysis

We constructed the spectral energy distributions (SEDs) of the sample stars with additional short wavelength photometry from a variety of catalogues, including Sloan Digital Sky Survey (SDSS; Ahn et al., 2012), AAVSO Photometric All-Sky Survey (APASS; Henden et al., 2009), Two Micron All Sky Survey (2MASS; Skrutskie et al., 2006), UKIRT Infrared Deep Sky Survey (UKIDSS; Lawrence et al., 2007), and Deep Near Infrared Survey of the Southern Sky (DENIS; Epchtein et al., 1999). Additional near-infrared fluxes for a few targets were also obtained from the literature (Barber et al., 2012; Farihi, 2009).

The available optical and near-infrared fluxes were fitted with pure hydrogen white dwarf atmosphere models (Koe, 2010), kindly provided by the author. The fit was computed by matching the optical and near-infrared best-quality photometric data points with a model spectrum with $\log g = 8$ and effective temperature obtained from Liebert et al. (2005) or Koester et al. (2009), approximated to the closest available model. The Levenberg-Marquardt minimisation algorithm was used to find the best scaling factor.

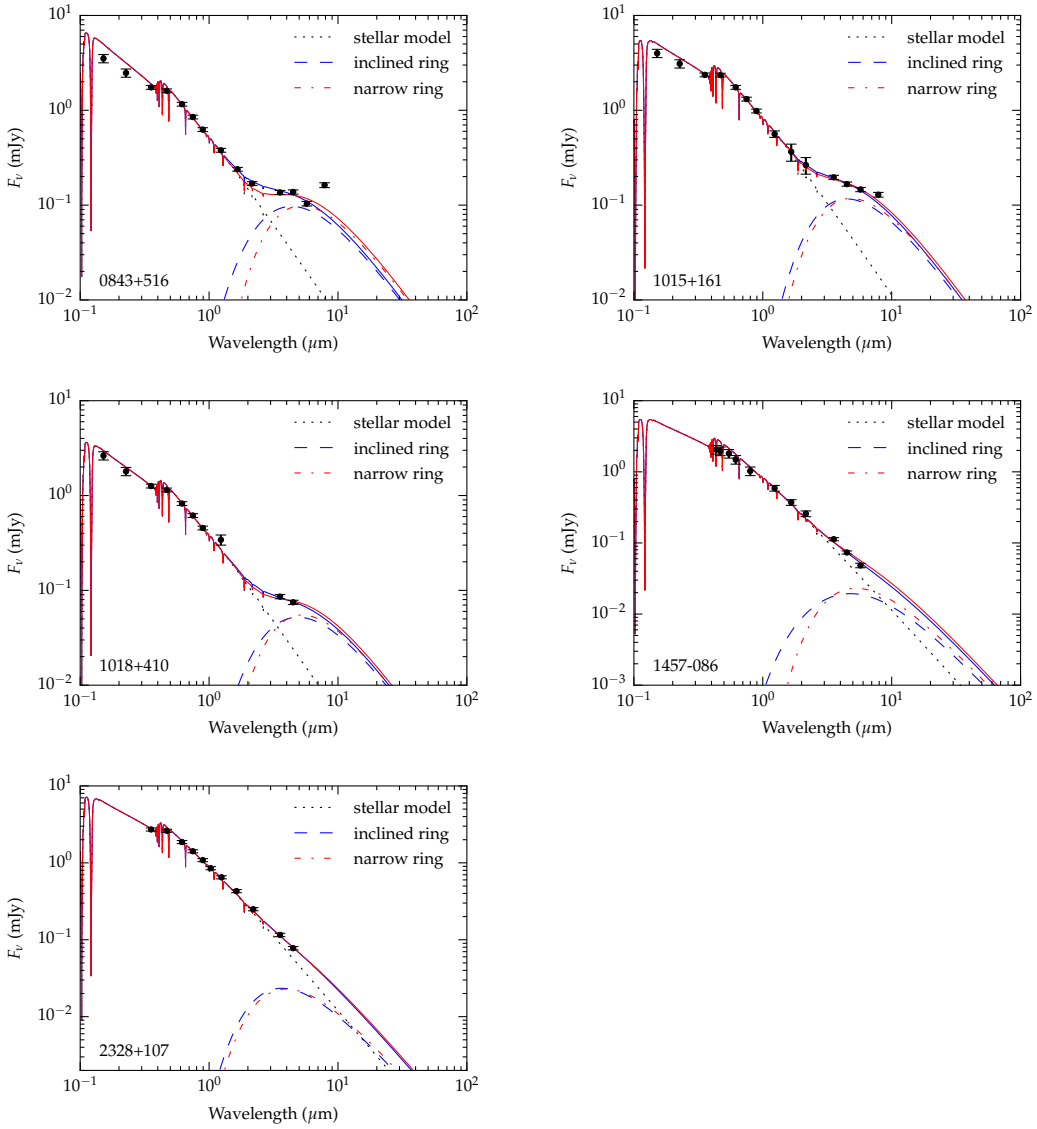

Figure 8.2: Infrared excesses fit by circumstellar disc models, with parameters listed in Table 8.1. Stellar atmosphere models are plotted as dot-dashed lines. Dashed lines represent emission from optically thick disc models, and solid lines represent the sum of the stellar and disc model fluxes. Two disc models are plotted for each star, one corresponding to highly inclined discs (blue), one to narrow rings (red). Circle symbols with error bars represent optical and infrared photometry, including IRAC fluxes.

8.3.1 Stars with infrared excesses

Amongst the 134 sample stars, a total of five white dwarfs show a significant ($> 4\sigma$) excess in the IRAC bands: 0843+516, 1015+161, 1457–086, 1018+410, and 2328+107. The first three stars have known infrared excesses (Xu & Jura, 2012; Farihi et al., 2009; Jura et al., 2007) and are also known to be metal polluted (Gänsicke et al., 2012; Koester et al., 2005), while the infrared excesses at 1018+410 and 2328+107 are reported for the first time. The

presence of a significant excess at the ancillary target 1929+012 (Melis et al., 2011; Debes et al., 2011a; Vennes et al., 2010) is also confirmed. This target, which is not part of the statistical sample, is discussed separately in Section 8.7, together with notes on individual excesses. The infrared excesses were individually modelled as thermal continua using an optically thick, geometrically thin disc (Jura, 2003). Stellar, substellar or planetary companions, and background contamination were confidently ruled out as the excess emission is either too strong (Farihi et al., 2009, 2008a) or the infrared colours are not compatible. The white dwarf radius was estimated from evolutionary models (Fontaine et al., 2001) while the distance was derived using synthetic absolute photometry (Holberg & Bergeron, 2006) compared with available optical and near-infrared magnitudes. Fluxes at 7.9 μm were excluded from the fits as these are often contaminated by a strong, 10 μm silicate emission feature, as seen in e.g. GD 362 (Jura et al., 2009) and G29-38 (Reach et al., 2005).

The free parameters in the disc model are the inner disc temperature $T_{\rm in}$, the outer disc temperature $T_{\rm out}$, and the disc inclination $i_{\rm disc}$. The radius of the disc at these temperatures can be easily estimated and is proportional to $T^{-3/4}$ (Chiang & Goldreich, 1997). In the absence of longer wavelength photometry extending to 24 μm, these three free parameters cannot be well constrained (Jura et al., 2007). A modest degree of degeneracy is found between the inclination and the radial extent of the disc, especially for subtle excesses (Bergfors et al., 2014; Girven et al., 2012). While the inner disc temperature and radius can be well constrained by the 3.6 μm emission, a large set of outer radii and disc inclinations can equally fit the longer-wavelength excess (see e.g. Jura et al., 2007). Two sets of representative models are therefore presented for each star with infrared excess, one corresponding to a relatively wide disc at high inclination and one to a more narrow ring at low inclination. An algorithm for bound constrained minimisation (Byrd et al., 1995) was used to estimate the best fit parameters in both cases.

Figure 8.2 shows the modelled and photometric SEDs of the dusty white dwarfs, and Table 8.1 gives the fitted disc parameters for the two sets of models. It can be seen that the outer radii of the model discs are all located within the Roche limit of the star (about $1.2\,R_{\odot} \approx 1.8\,R_{\oplus}$) where planetesimals larger than ≈ 1 km would be tidally destroyed. Moreover, the acceptable values for the inner disc temperatures are in agreement with the temperature at which solid dust grains rapidly sublimate in a metal-rich and hydrogen-poor disc (Rafikov, 2012). It is interesting to notice that all models with larger radial extent require

Table 8.1: White dwarfs with circumstellar discs in the Spitzer survey.

WD	$T_{\rm in}$ (K)	$T_{\rm out}$ (K)	$r_{\rm in}$ (R_*)	$r_{\rm out}$ (R_*)	$r_{\rm in}$ ($R_\odot$)	$r_{\rm out}$ ($R_\odot$)	$i_{\rm disc}$ (deg)
Model 1: High inclinations							
0843+516	1750	800	14	30	0.20	0.42	87
1015+161	1450	900	14	22	0.19	0.30	85
1018+410	1600	700	14	32	0.18	0.41	87
1457–086	2000	600	10	34	0.13	0.46	89
2328+107	1480	1300	14	16	0.21	0.24	88
Model 2: Narrow rings							
0843+516	1000	960	24	25	0.34	0.35	30
1015+161	1120	1000	18	20	0.25	0.28	70
1018+410	1000	940	22	24	0.28	0.31	40
1457–086	1000	940	20	21	0.27	0.28	75
2328+107	1300	1270	16	17	0.24	0.26	75

improbably high inclinations, all greater or equal to 85 deg (Table 8.1). Because it is highly unlikely for all the inclinations to be so confined, most if not all these discs must actually be relatively narrow. Section 8.5 discusses this evidence in greater detail.

8.4 The frequency of circumstellar debris

The detection of dust at five of 134 sample stars translates to a nominal excess frequency of $3.7^{+2.4}_{-1.0}$% for post-main sequence ages of 15–270 Myr. The upper and lower bounds are calculated using the binomial probability distribution and 1 sigma confidence level. These results are in broad agreement with results obtained in previous surveys, which were however affected by different sources of bias (for a full discussion see Rocchetto et al., 2015, Section 5). Notably, complementary HST/COS observations (discussed in the next section) demonstrate that at least 27% of white dwarfs with diffusion timescales of only a few weeks have photospheric metals that require ongoing accretion of a circumstellar reservoir. *This frequency is a firm lower limit on the fraction of 2–3 $M_\odot$ stars (the typical progenitors of the white dwarfs in our sample) that form planetary systems.*

These findings clearly indicate that the actual fraction of debris discs at white dwarfs is almost an order of magnitude higher, with nearly 90% of discs emitting insufficient flux to be detected by the current infrared facilities, such as Spitzer.

8.4.1 Complementary HST/COS observations

The entire sample of 134 stars observed with Spitzer was also approved as an HST/COS Snapshot program. Snapshot targets are observed during gaps between regular guest observer programs, and a total of 85 of the 134 white dwarfs were observed between 2010 September and 2013 February within programs 12169 and 12474 (PI: B.T. Gaensicke). The results of the HST survey itself are published elsewhere (Koester et al., 2014), and these are summarised briefly here. We found that amongst these 85 stars, 56% display atmospheric metals: 48 exhibit photospheric Si, 18 also show C, and 7 show further metals. Our analysis indicated that for 25 stars the metal abundances may be explained by radiative levitation alone, although accretion has likely occurred recently, leaving 23 white dwarfs (27%) that exhibit traces of heavy elements that can only be explained with ongoing accretion of circumstellar material, in agreement with previous estimates (Zuckerman et al., 2010, 2003). Spitzer observations of the 85 star subsample with HST data show that, amongst the 23 metal polluted white dwarfs that are currently accreting, there are two that exhibit detectable infrared excesses: 0843+516 and 1015+161. The co-observed subsample thus translates to approximately 10% of metal-enriched stars exhibiting detectable dust, indicating that about 90% of debris discs escape detection in the infrared. As expected, no infrared excesses are confidently found in the subsample of non-metal bearing degenerates observed with HST, strongly supporting the connection between infrared excesses and metal pollution.

Possible reasons for the apparent lack of infrared disc detections are still a matter of debate. The collective data for the known circumstellar discs at hydrogen dominated white dwarfs indicate that the DA degenerates with the highest accretion rates are significantly more likely to host an infrared detectable circumstellar disc (Xu & Jura, 2012; Farihi et al., 2009; Jura et al., 2007). The infrared excesses at 0843+516 and 1015+161 confirm this trend, as they both have the highest inferred Si accretion rates amongst the HST sample stars (3.6×10^7 g s^{-1} and 5×10^6 g s^{-1} respectively). This is consistent with a picture where white dwarfs accreting at the highest rates require the most massive and highest surface density discs, which are more likely to be detected in the infrared. One possible explanation for the dearth of infrared excesses is that mutual collisions may be enhanced in low surface density discs, and result in the partial or complete destruction of dust grains (Jura et al., 2007). A related possibility is increased, collisional grain destruction due to the impact of additional, relatively small planetesimals on a pre-exiting dust disc (Jura,

2008). This may explain the lower frequency of infrared excess detections at older and cooler white dwarfs (Bergfors et al., 2014, see also section 8.5), as the depletion of the reservoir of large asteroids for older white dwarfs would imply that smaller planetesimals are primarily accreted. In contrast, this work corroborates the interpretation that the lack of infrared excess detections at a large fraction of metal polluted white dwarfs is largely caused by the small total emitting surface area of dust grains, which implies a very low infrared fractional luminosity and hence undetectability with the current instrumentation. Sections 8.4.2 and 8.5 provide evidence supporting this hypothesis.

8.4.2 Hidden subtle excesses in the DAZ sample

Thanks to the large number of stars observed as part of this Spitzer survey, one can investigate the possibility that some DAZ white dwarfs have an infrared excess that is just below the current sensitivity limit. The distribution of excess and deficit infrared fluxes with respect to the model fluxes demonstrates a correlation between subtle infrared excesses and atmospheric metals indicating a population of tenuous circumstellar discs.

The observed excess and deficit fluxes with respect to the model flux at different wavelengths were expressed as a fraction of the photometric and model uncertainties, and were quantified by an excess significance, defined as:

$$\chi_{\rm excess} = \frac{F_{\rm obs} - F_{\rm model}}{\sqrt{\sigma_{\rm obs}^2 + \sigma_{\rm model}^2}} \tag{8.4.1}$$

where $F_{\rm obs}$ and $F_{\rm model}$ are the observed and photospheric model fluxes respectively, and similarly for the uncertainties, $\sigma_{\rm obs}$ and $\sigma_{\rm model}$.

The excess significance values were derived for each star in the subsample of 85 degenerates that have both Spitzer and HST observations. Two subgroups were created, one containing non-metal lined (DA) white dwarfs only and one containing degenerates exhibiting atmospheric metals (DAZ) only. All stars that show Si absorption in their spectra – including those whose metals can be explained by radiative levitation alone – were included in the DAZ sample, as these stars have likely accreted circumstellar material recently (Koester et al., 2014). The DA sample consists of 32 stars observed at 3.6 μm and 38 stars at 4.5 μm, while the DAZ sample consists of 34 stars observed at 3.6 μm and 45 stars at 4.5 μm. The differences in the subsample sizes are due to the lack of IRAC observations in channel 1 or 2 for some stars whose data were taken from the Spitzer archive. Infrared excesses that

have $\chi_{\rm excess} > 4$ and cases of strong contamination from nearby sources were excluded. The excess significance values were then plotted in separate histograms, for fluxes measured at 3.6 μm and 4.5 μm. The resulting four histograms are shown in Figure 8.3. It is important to stress that statistically confident excesses are defined for $\chi_{\rm excess} > 4$, corresponding to a significance level of 4σ. An excess with $1.5 < \chi_{\rm excess} < 4$ can be defined as a *candidate subtle excess*, which cannot yet be confirmed with confidence. However, one can test the observed distribution of all $\chi_{\rm excess} < 4$ for each subsample of DA and DAZ stars observed at 3.6 μm and 4.5 μm against expectations.

The number of candidate subtle excesses with $\chi_{\rm excess} > 1.5$ and their expected values were computed for each of these subsamples, and these are shown in Figure 8.3. The expected values were inferred assuming no correlation between atmospheric metals and infrared excesses and were computed in the following way. The observed excesses were randomly distributed in two subsamples equal in size to the observed DA and DAZ subsamples, at 3.6 and 4.5 μm. This computation was repeated 1000 times, and the average number of stars with $\chi_{\rm excess} > 1.5$ and its standard deviation were used as estimates.

Interestingly, departure from expectations was found. In the DA sample there is only one candidate excess at 4.5 μm, while the expected number is 1 ± 1 at 3.6 μm and 4 ± 1 at 4.5 μm. In contrast, the bulk of candidate excesses are found in the DAZ sample, where four and eight $\chi_{\rm excess} > 1.5$ values were found at 3.6 μm and 4.5 μm, while the expected numbers are 2 ± 1 and 5 ± 1 respectively.

This apparent correlation between atmospheric metals and subtle infrared excesses, especially at longer wavelengths, may reinforce the idea that most if not all circumstellar discs at metal polluted white dwarfs harbour dust that emits in the infrared, but its signature is too subtle to be detected with the current instrumentation.

8.5 The fractional luminosity of dust discs

Here it is shown that the distribution of the disc fractional luminosities ($\tau_{\rm disc} = L_{\rm IR}/L_*$) for the sample of known dusty white dwarfs reinforces the interpretation that a large fraction of circumstellar discs remain undetected in the infrared and may also provide insight into their evolution.

The value of fractional luminosity of all known infrared excesses at cool white dwarfs was estimated in a consistent manner. The photospheric flux was fitted as described in Section 8.3 and a blackbody distribution was then fitted to the infrared excesses seen in the

Table 8.2: Estimated fractional luminosities from thermal continuum for all known white dwarfs exhibiting infrared excess due to dust.

WD	$t_{\rm cool}$ (Myr)	$T_{\rm eff}$ (K)	$T_{\rm IR}$ (K)	$\tau_{\rm disc}$ (%)	Ref.
0106–328	160	16000	1470	0.08	1
0110–565	81	19200	1050	0.15	2
0146+187	418	11500	1120	1.47	3
0246+734[†]	995	8250	1000	0.31	4
0300–013	189	15200	1200	0.31	5
0307+078	531	10500	1200	0.18	1
0408–041	224	14400	1020	2.94	5
0435+410	116	17500	1250	0.28	2
0735+187	264	13600	1330	1.59	6
0842+231	92	18600	1350	0.60	6
0843+516	29	23900	1310	0.13	7
0956–017	283	13280	1220	2.73	8
1015+161	80	19300	1210	0.17	5
1018+410	45	22390	1210	0.11	9
1041+091	106	17910	1500	0.15	6
1116+026	358	12200	1000	0.48	5
1150–153	314	12800	940	2.02	10
1219+130	350	12300	1250	1.61	8
1225–079	531	10500	300	0.05	1
1226+110	45	22000	1070	0.40	11
1349–230	100	18200	1260	0.33	2
1455+298	1315	7400	400	0.19	12
1457–086	63	20400	1400	0.04	3
1541+651	409	11600	980	1.23	13
1551+175	178	15500	1690	0.19	4
1554+094	37	22800	1100	0.43	8
1615+164	274	13430	1370	1.28	6
1729+371	531	10500	820	2.02	14
1929+011	57	20890	1060	0.19	5
2115–560	653	9700	820	0.84	3
2132+096	301	13000	550	0.11	4
2207+121	122	17300	1100	0.74	7
2221–165	588	10100	950	0.67	1
2326+049	409	11600	1060	2.87	12
2328+107	37	21000	2000	0.05	9

[†] Needs to be confirmed.

References:

(1) Farihi et al. (2010b) (2) Girven et al. (2012) (3) Farihi et al. (2009) (4) Bergfors et al. (2014) (5) Jura et al. (2007) (6) Brinkworth et al. (2012) (7) Xu & Jura (2012) (8) Zuckerman et al. (2003); Farihi et al. (2012) (9) this work; (10) Jura et al. (2009) (11) Brinkworth et al. (2009) (12) Farihi et al. (2008b) (13) Barber et al. (2012) (14) Jura et al. (2007)

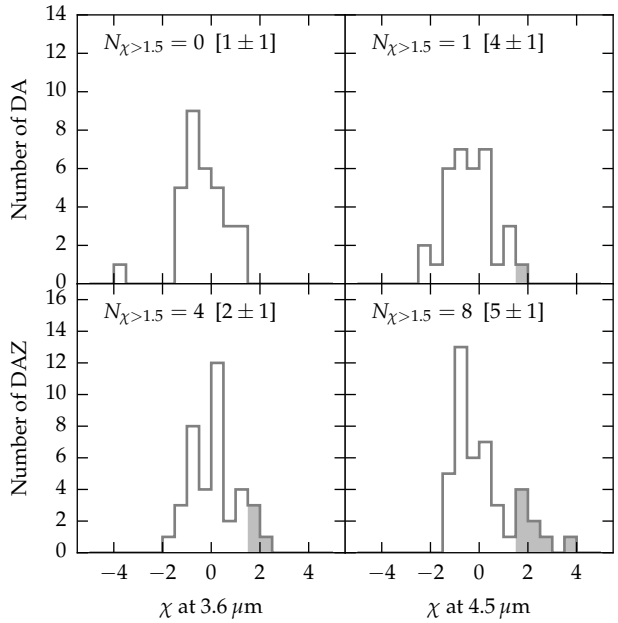

Figure 8.3: Histograms of excess significance $\chi_{\rm excess}$ for metal lined (DAZ) and non-metal bearing (DA) white dwarfs observed at 3.6 μm and 4.5 μm with IRAC. Stars with $\chi_{\rm excess} > 1.5$ are highlighted in grey, and their occurrence is given at the top left of each subplot, with expected values in square brackets. Stars with $\chi_{\rm excess} > 4$ and sources that are heavily contaminated by nearby objects have been excluded.

K-band and the three shortest wavelength IRAC bandpasses. The mid-infrared fluxes were all measured with Spitzer IRAC and were obtained from the literature and this work. Although a precise estimate of $\tau_{\rm disc}$ requires detailed modelling of the photospheric and disc flux, the values reported here are estimated to be good to 10%. Independently measured values for 18 discs known as of mid-2010 agree to within 10% (Farihi, 2011). Note also that $\tau_{\rm disc}$ is the *observed* fractional luminosity and it does not take into account the disc inclination, which cannot be determined confidently (Section 8.3.1). The fractional luminosity $\tau_{\rm disc}$ and the blackbody dust temperature $T_{\rm IR}$ for all 35 Spitzer detected white dwarfs with infrared excesses are listed in Table 8.2. The measured values of $\tau_{\rm disc}$ and their distribution are also plotted in Figure 8.4.

8.5.1 An undetected population of subtle excesses

The right hand histogram of the middle panel of Figure 8.4 shows the distribution of $\tau_{\rm disc}$, with 20 bins between 0% and 3%. It can be seen that the distribution rises sharply at the smallest fractional luminosities indicating that the majority of circumstellar discs are subtle. This suggests that many discs might have a fractional luminosity that is below the sensitivity

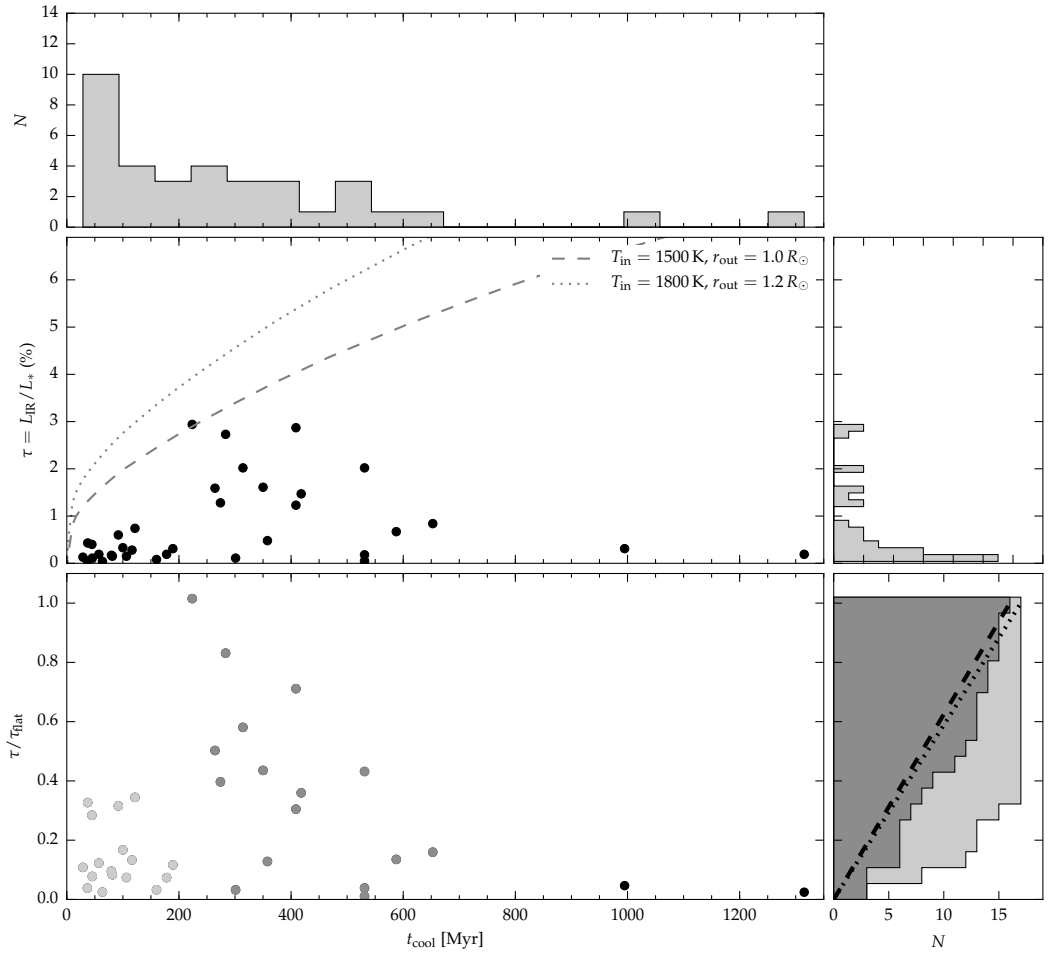

Figure 8.4: Fractional disc luminosities ($\tau_{disc} = L_{IR}/L_*$) for all known white dwarfs with detected dust discs. The central panel shows the value of τ_{disc} as a function of cooling age for $\log g = 8$. The two lines lines show the predicted fractional luminosity for a face on, opaque flat disc with radial extent given by assuming dust grains persist up to 1500 and 1800 K and extend up to $1.0\,R_\odot$ and $1.2\,R_\odot$. The dotted line is for $T_{in} = 1500\,K$ and $r_{out} = 1.0\,R_\odot$, while the dashed line is for $T_{in} = 1800\,K$ and $r_{out} = 1.2\,R_\odot$. The top panel shows the histogram of the number of infrared excesses as a function of cooling age, while the right panel shows the histogram of τ_{disc} with 20 bins between 0 and 3%. The bottom panel shows the ratio between the infrared fractional luminosity of all known dusty white dwarfs and the predicted maximum fractional luminosity for a flat, passive disc, assuming $T_{in} = 1500\,K$ and $r_{out} = 1.0\,R_\odot$. The light grey symbols have $t_{cool} < 200\,Myr$ (where t_{cool} is the cooling age of the white dwarf), while the dark grey symbols have $200\,Myr < t_{cool} < 700\,Myr$. The bottom-right panel shows the cumulative histograms of $\tau_{disc}/\tau_{disc,flat}$ for the light and dark grey target subsets, and the expected distributions for random inclinations are shown as dotted and dashed lines respectively.

limits of the current instrumentation and therefore escape detection.

Together with the likely presence of several additional subtle excesses seen in this Spitzer sample (see Section 8.4.2), the distribution of $\tau_{\rm disc}$ points to a large population of subtle excesses, suggesting that most if not all, currently accreting white dwarfs with metals have circumstellar dust. This significant population of dust discs are not detectable with current facilities, likely due to low surface areas and optical depth.

8.5.2 Narrow rings at young white dwarfs

The central panel of Figure 8.4 shows $\tau_{\rm disc}$ as a function of cooling age. The two lines represent the value of the maximum fractional disc luminosity $\tau_{\rm disc,flat}$ for a flat disc model (Jura, 2003) with a face-on configuration. This value was calculated assuming that the dust occupies all the space available between the distance from the star at which dust grains rapidly sublimate and the stellar Roche limit. The two lines correspond to two different assumptions about the maximum disc extent. The dotted line assumes that silicates rapidly sublimate at $T_{\rm in} = 1800\,{\rm K}$ and the stellar Roche limit is $r_{\rm out} = 1.2\,R_{\odot}$, while the dashed line assumes $T_{\rm in} = 1500\,{\rm K}$ and $r_{\rm out} = 1.0\,R_{\odot}$.

From the central panel it can be seen that infrared excesses have so far only been detected at white dwarfs older than $\approx 25\,{\rm Myr}$, and with corresponding $T_{\rm eff} \lesssim 25\,000\,{\rm K}$. This is somewhat expected, as at higher stellar temperatures, any disrupted asteroids debris will evaporate relatively quickly, so that their detection is less likely. However, the calculations done here and shown in the plot indicate that optically thick discs may exist at higher effective temperatures and cooling ages. Such discs would likely evolve more rapidly (Rafikov, 2011), and have a higher gas to dust ratios from sublimation prior to settling into optically thick rings.

As the white dwarf cools below 25 000 K the distance at which silicates rapidly sublimate decreases. Hence, the area available for dust grains increases, so that the disc luminosity should increase as well. *However, the expected increase assuming that discs are fully extended (shown by the two lines in the central panel of Figure 8.4), is not seen in the observed fractional luminosities*. Between 25 Myr and 200 Myr there are 17 dusty white dwarfs, all with fractional luminosities significantly less than the maximum value allowed for flat discs. This suggests that, at these young white dwarfs, large and extended discs do not form, or only persist for timescales significantly shorter than the disc lifetime.

This finding can be better explored and tested in the bottom-left panel of Figure 8.4,

showing the ratio $\tau_{\rm disc}/\tau_{\rm disc,flat}$ for $T_{\rm in} = 1500\,{\rm K}$ and $r_{\rm out} = 1.0\,R_{\odot}$ as a function of cooling age. Supposing that discs are fully extended, and assuming random inclinations, the distribution of $\tau_{\rm disc}/\tau_{\rm disc,flat}$ should be uniform between zero and one. However, this does not seem to be true, especially for white dwarfs younger than 200 Myr and older than 700 Myr. To better explore if discs have different radial extents at different ages it is useful to consider three cooling age ranges separately, for $t_{\rm cool} < 200\,{\rm Myr}$ (light grey symbols), for $200\,{\rm Myr} < t_{\rm cool} < 700\,{\rm Myr}$ (dark grey symbols), and for $t_{\rm cool} > 700\,{\rm Myr}$. It can be seen that:

1) At $t_{\rm cool} < 200\,{\rm Myr}$, the $\tau_{\rm disc}/\tau_{\rm disc,flat}$ values are all below 0.4, clearly demonstrating that inclination is not the only parameter influencing the distribution. A Kolmogorov–Smirnov (KS) test confirms this, resulting in a probability of 2×10^{-7} that the observed distribution is uniform in $\tau_{\rm disc}/\tau_{\rm disc,flat}$. As any observational bias would favour the detection of discs with lower inclinations and correspondingly higher fractional luminosities, this indicates that the observed discs must be relatively narrow at these young cooling ages.

2) Between $200\,{\rm Myr} < t_{\rm cool} < 700\,{\rm Myr}$ there is a notable increase in the observed fractional luminosities, with the brightest ($\tau_{\rm disc} \approx 3\%$) disc modelled as face on at 0408–041 (GD 56). In this range of cooling ages, a KS test reveals that the distribution of $\tau_{\rm disc}/\tau_{\rm disc,flat}$ is uniform with a probability of only 0.07, again indicating that a second parameter shapes the distribution, namely the disc radial extent. While not as striking as for the youngest cooling ages, this group likely contains several relatively narrow discs, as indicated by the cumulative distribution at smaller $\tau_{\rm disc}/\tau_{\rm disc,flat}$ values.

3) Lastly, at $t_{\rm cool} > 700\,{\rm Myr}$ there are only two detected discs with relatively low fractional luminosities ($\tau_{\rm disc} < 0.3\%$, $\tau_{\rm disc}/\tau_{\rm disc,flat} < 0.05$). The small $\tau_{\rm disc}$ values of these discs, together with the decreasing frequency of detections seen at older white dwarfs (top histogram of Figure 8.4), suggest that the number of large asteroid disruptions per time bin may be decreasing and hence also the fraction of detectable infrared excesses (Bergfors et al., 2014). However, it is interesting to note that the inferred metal accretion rates do not show a decreasing trend with cooling age (Koester et al., 2014), indicating that circumstellar material is present at older stars. This could be explained by the accretion of smaller asteroids, which can still provide continuous accretion, but are less likely to produce detectable amounts of dust (Wyatt et al., 2014).

8.5.3 Possible disc evolution scenarios

Interestingly, the existence of narrow rings is predicted by global models of white dwarf disc evolution. In general, the outer radii of flat and optically thick discs will rapidly decrease due to Poynting-Robertson (PR) drag. Specifically, for a range of initial, realistic surface density distributions, PR drag is significantly more efficient per unit mass on the outermost (and innermost) disc regions where grain density and optical depth is lowest. Solids are quickly gathered inward until they result in a region of moderate optical depth (Metzger et al., 2012; Bochkarev & Rafikov, 2011), giving rise to a sharp outer edge.

This edge forms rapidly, and marches appreciably inward within a few to several hundred PR drag timescales (Bochkarev & Rafikov, 2011), which can be as short as years for typical dusty white dwarfs and 1 μm dust at 1 $R_{\odot}$. The rate at which the outer edge migrates inward will be ultimately set by the dominant grain size, as the rate scales linearly with particle size and can be a factor of 1000 longer for centimetre vs. micron sizes. Stellar luminosity plays a smaller role, varying by less than a factor of 20 between typical 10 000 K and 20 000 K white dwarfs (Fontaine et al., 2001). Thus, if this mechanism is responsible for sculpting narrow rings, their radial extents as a function of white dwarf cooling age and luminosity may broadly constrain typical particle sizes.

The transition in disc brightness between the $t_{\rm cool} < 200$ Myr and 200 Myr $< t_{\rm cool} <$ 700 Myr cooling age ranges appears relatively abrupt, however, and it is not clear that a change in stellar luminosity (and PR drag timescale) alone can account for the difference between the two populations. It is perhaps tempting to interpret the brighter and presumably larger discs at 200 Myr $< t_{\rm cool} <$ 700 Myr as evidence of disc spreading, but the viscous timescale among particulates, for any reasonable particle size, is orders of magnitude insufficient and thus highly unlikely to have an effect on disc evolution (Metzger et al., 2012; Farihi et al., 2008b).

A distinct possibility for the narrowing of rings is disc truncation by impact. Theoretical models show that post-main sequence dynamical instabilities arise near $\sim$ 10 Myr, then peak around $\sim$ 100 Myr, decreasing towards later times (Veras et al., 2013; Mustill et al., 2014). This suggests that impacts on pre-existing discs by additional, perturbed and tidally shredded asteroids would be most prevalent among the $t_{\rm cool} < 200$ Myr disc sample. Disc impacts would have the favourable characteristic that they should occur preferentially at larger orbital radii, thus typically creating outwardly truncated discs. The

impactor fragments would destroy dust masses comparable to their initial, intact masses, creating gas that would quickly dissipate (Jura, 2008). Interestingly, the recently observed drop in infrared luminosity from the dusty white dwarf SDSS J095904.69−020047.6, and the disappearance of metallic gas emission lines from the similarly dusty and polluted star SDSS J161717.04+162022.4, may represent such impact events (Xu & Jura, 2014; Wilson et al., 2014).

Because viscous spreading is inefficient (see above), discs will not grow outward significantly during their lifetime, and the narrower discs might be the result of initial conditions; specifically, smaller disrupted parent bodies. However, assuming the same overall disc properties, flat configurations and random inclinations, the average disc mass should correlate with $\tau_{\rm disc}$ over each population, implying parent body masses all similar to within a factor of roughly 50. This would require relatively fine-tuned mass influxes per unit time in the post-main sequence, including a subtle change around 300 Myr to include slightly larger disruptions. Taken together, this seems unlikely to account for the disc brightness differences among the two cooling age ranges.

In summary, while none of the above mechanisms appears to be without possible drawbacks, the clear observational difference in disc populations demands a closer look with formation and evolutionary modelling. Insufficient data could be masking a more gradual change in the fractional disc luminosities that would favour the edge migration scenario, for example. The detection of a larger number of discs, especially those producing subtle excesses, would improve the statistics and reduce the possibilities.

8.6 Summary and conclusion

Spitzer observations of an unbiased sample of 134 DA white dwarfs yield a disc frequency near 4% in the effective temperature range $T_{\rm eff} = 17\,000\,{\rm K} - 25\,000\,{\rm K}$. However, complementary HST observations of 85 sample stars reveal that a much larger fraction of at least 27% host circumstellar material, indicating that about 90% of the discs remain undetected in the infrared. *This frequency is a firm lower limit on the fraction of 2–3 $M_\odot$ stars (the typical progenitors of the white dwarfs in our sample) that form planetary systems.*

Possible reasons for the lack of infrared detections were investigated. The distribution of observed fluxes compared to photospheric models at 3.6 and 4.5 μm for the subsamples with and without metals point to a population of excesses too subtle to currently be confirmed. Future observations, for example with *JWST* using spectroscopy, have the potential

to confirm the presence of faint dust discs.

The distribution of the fractional disc luminosities of all known dusty white dwarfs also points towards a hidden population of subtle discs. In addition, this distribution indicates that the disc population evolves over time. Only relatively narrow rings are found at $t_{\rm cool} \lesssim$ 200 Myr, while relatively extended discs, filling the majority of the space available between the distance at which silicates rapidly sublimate and the stellar Roche limit, occur only after a few hundred Myr. A marked decrease in the observed fractional disc luminosities, as well as a decrease in the frequency of detections towards cooler ages, is seen after this peak, suggesting that the number of large asteroids might be gradually depleted.

This study strongly reinforces the hypothesis that most if not all metal-enriched white dwarfs harbour circumstellar dust, but the majority remain unseen due to low surface area and the sensitivity limits of current instrumentation.

Appendix A

Additional material for Chapter 5

A.1 Retrieved absolute abundances

Table A.1: Retrieved absolute abundances with 1 sigma uncertainty for the seven molecules and seven C/O values considered in this study. For each retrieved parameter, we show in parenthesis how many sigma away the retrieved value is from the true state.

Mol.	C/O	Input value	Retrieved value	
		(at 0.1 bar)	TP-ISO	TP-PARAM
H_2O	0.5	4.08×10^{-4}	$4.66 \times 10^{-3} - 6.53 \times 10^{-3}$ (15.4)	$3.73 \times 10^{-4} - 1.98 \times 10^{-2}$ (1.0)
	0.7	2.52×10^{-4}	$2.98 \times 10^{-3} - 4.12 \times 10^{-3}$ (16.2)	$1.82 \times 10^{-4} - 5.37 \times 10^{-4}$ (0.4)
	0.9	8.63×10^{-5}	$6.61 \times 10^{-4} - 8.18 \times 10^{-4}$ (20.1)	$4.84 \times 10^{-5} - 9.75 \times 10^{-5}$ (0.3)
	1.0	1.20×10^{-6}	$1.63 \times 10^{-6} - 1.94 \times 10^{-6}$ (4.6)	$1.43 \times 10^{-6} - 1.62 \times 10^{-6}$ (4.0)
	1.1	4.56×10^{-8}	$< 1.04 \times 10^{-8}$ (1.4)	$< 5.44 \times 10^{-12}$ (6.2)
	1.3	2.44×10^{-8}	$< 5.13 \times 10^{-9}$ (1.5)	$< 2.58 \times 10^{-8}$ (1.1)
	1.5	1.85×10^{-8}	$< 5.74 \times 10^{-9}$ (1.3)	$< 1.49 \times 10^{-8}$ (0.8)
CO	0.5	4.13×10^{-4}	$8.74 \times 10^{-3} - 1.10 \times 10^{-2}$ (28.0)	$3.64 \times 10^{-4} - 5.33 \times 10^{-2}$ (0.9)
	0.7	5.70×10^{-4}	$9.46 \times 10^{-3} - 1.17 \times 10^{-2}$ (27.0)	$2.69 \times 10^{-4} - 1.17 \times 10^{-3}$ (0.0)
	0.9	7.34×10^{-4}	$1.15 \times 10^{-2} - 1.44 \times 10^{-2}$ (25.4)	$2.90 \times 10^{-4} - 7.79 \times 10^{-4}$ (0.4)
	1.0	8.22×10^{-4}	$2.24 \times 10^{-3} - 3.38 \times 10^{-3}$ (5.9)	$2.44 \times 10^{-3} - 3.40 \times 10^{-3}$ (7.5)
	1.1	8.24×10^{-4}	$7.39 \times 10^{-3} - 9.62 \times 10^{-3}$ (17.6)	$1.01 \times 10^{-3} - 1.01 \times 10^{-3}$ (799.1)
	1.3	8.24×10^{-4}	$8.35 \times 10^{-3} - 1.06 \times 10^{-2}$ (20.7)	$6.33 \times 10^{-4} - 1.67 \times 10^{-3}$ (0.5)
	1.5	8.24×10^{-4}	$8.17 \times 10^{-3} - 1.03 \times 10^{-2}$ (20.9)	$8.91 \times 10^{-4} - 1.26 \times 10^{-3}$ (1.4)
CO_2	0.5	3.60×10^{-8}	$9.46 \times 10^{-7} - 1.82 \times 10^{-6}$ (11.0)	$8.63 \times 10^{-8} - 4.13 \times 10^{-6}$ (1.5)
	0.7	3.07×10^{-8}	$5.95 \times 10^{-7} - 1.14 \times 10^{-6}$ (10.2)	$4.13 \times 10^{-8} - 1.27 \times 10^{-7}$ (1.5)
	0.9	1.35×10^{-8}	$1.57 \times 10^{-7} - 2.77 \times 10^{-7}$ (9.6)	$1.41 \times 10^{-8} - 2.87 \times 10^{-8}$ (1.1)
	1.0	2.10×10^{-10}	$< 2.78 \times 10^{-9}$ (1.1)	$< 3.24 \times 10^{-9}$ (0.3)
	1.1	8.01×10^{-12}	$< 3.32 \times 10^{-10}$ (0.6)	$< 7.64 \times 10^{-10}$ (2.5)
	1.3	4.28×10^{-12}	$< 7.23 \times 10^{-10}$ (0.9)	$< 8.56 \times 10^{-12}$ (0.4)
	1.5	3.25×10^{-12}	$< 4.05 \times 10^{-10}$ (1.0)	$< 4.51 \times 10^{-7}$ (1.3)
CH_4	0.5	6.43×10^{-11}	$< 1.85 \times 10^{-7}$ (0.6)	$< 1.30 \times 10^{-7}$ (0.7)

Mol.	C/O	Input value	Retrieved value	
		(at 0.1 bar)	TP-ISO	TP-PARAM
	0.7	1.44×10^{-10}	$< 2.86 \times 10^{-7}$ (0.5)	$< 1.97 \times 10^{-8}$ (0.3)
	0.9	5.39×10^{-10}	$< 2.03 \times 10^{-8}$ (0.1)	$< 7.50 \times 10^{-10}$ (0.5)
	1.0	4.36×10^{-8}	$8.48 \times 10^{-8} - 1.13 \times 10^{-7}$ (5.6)	$4.61 \times 10^{-8} - 6.35 \times 10^{-8}$ (1.4)
	1.1	1.16×10^{-6}	$9.23 \times 10^{-6} - 1.15 \times 10^{-5}$ (20.0)	$3.06 \times 10^{-6} - 3.06 \times 10^{-6}$ (43189.6)
	1.3	2.20×10^{-6}	$2.03 \times 10^{-5} - 2.55 \times 10^{-5}$ (20.4)	$4.85 \times 10^{-6} - 8.79 \times 10^{-6}$ (3.6)
	1.5	2.92×10^{-6}	$1.80 \times 10^{-5} - 2.22 \times 10^{-5}$ (18.6)	$5.66 \times 10^{-6} - 6.57 \times 10^{-6}$ (10.0)
HCN	0.5	1.32×10^{-9}	$< 2.15 \times 10^{-5}$ (0.4)	$< 1.15 \times 10^{-7}$ (0.0)
	0.7	2.94×10^{-9}	$< 2.94 \times 10^{-7}$ (0.1)	$< 2.68 \times 10^{-8}$ (0.4)
	0.9	1.10×10^{-8}	$< 5.54 \times 10^{-7}$ (0.3)	$< 3.03 \times 10^{-8}$ (0.9)
	1.0	8.84×10^{-7}	$6.60 \times 10^{-7} - 8.20 \times 10^{-7}$ (1.7)	$5.18 \times 10^{-7} - 6.14 \times 10^{-7}$ (5.3)
	1.1	2.09×10^{-5}	$5.53 \times 10^{-5} - 6.74 \times 10^{-5}$ (10.8)	$2.18 \times 10^{-5} - 2.18 \times 10^{-5}$ (1355.8)
	1.3	3.58×10^{-5}	$1.20 \times 10^{-4} - 1.49 \times 10^{-4}$ (12.1)	$3.00 \times 10^{-5} - 5.79 \times 10^{-5}$ (0.5)
	1.5	4.42×10^{-5}	$1.01 \times 10^{-4} - 1.23 \times 10^{-4}$ (9.2)	$3.47 \times 10^{-5} - 4.26 \times 10^{-5}$ (0.3)
C_2H_2	0.5	8.54×10^{-14}	$< 4.18 \times 10^{-7}$ (2.0)	$< 9.43 \times 10^{-8}$ (2.2)
	0.7	4.26×10^{-13}	$< 5.26 \times 10^{-7}$ (1.6)	$< 1.03 \times 10^{-7}$ (1.8)
	0.9	5.96×10^{-12}	$< 2.97 \times 10^{-7}$ (1.1)	$< 5.24 \times 10^{-7}$ (1.3)
	1.0	3.88×10^{-8}	$< 4.08 \times 10^{-9}$ (1.6)	$< 4.77 \times 10^{-9}$ (1.6)
	1.1	2.72×10^{-5}	$1.66 \times 10^{-4} - 2.88 \times 10^{-4}$ (7.5)	$2.51 \times 10^{-5} - 2.51 \times 10^{-5}$ (0.1)
	1.3	9.66×10^{-5}	$7.19 \times 10^{-4} - 1.17 \times 10^{-3}$ (9.2)	$5.76 \times 10^{-5} - 4.44 \times 10^{-4}$ (0.5)
	1.5	1.68×10^{-4}	$9.42 \times 10^{-4} - 1.47 \times 10^{-3}$ (8.7)	$1.64 \times 10^{-4} - 2.86 \times 10^{-4}$ (0.9)
NH_3	0.5	9.73×10^{-9}	$< 1.04 \times 10^{-7}$ (0.5)	$9.74 \times 10^{-12} - 3.33 \times 10^{-8}$ (0.7)
	0.7	9.73×10^{-9}	$< 4.43 \times 10^{-7}$ (0.3)	$< 1.11 \times 10^{-8}$ (0.9)
	0.9	9.65×10^{-9}	$< 3.53 \times 10^{-6}$ (10.9)	$< 9.63 \times 10^{-8}$ (0.5)
	1.0	9.62×10^{-9}	$< 3.76 \times 10^{-7}$ (1.4)	$< 9.12 \times 10^{-8}$ (3.2)
	1.1	8.70×10^{-9}	$< 1.80 \times 10^{-6}$ (36.4)	$< 2.89 \times 10^{-7}$ (4651.3)
	1.3	7.97×10^{-9}	$< 2.02 \times 10^{-6}$ (29.5)	$< 5.93 \times 10^{-7}$ (5.6)
	1.5	7.51×10^{-9}	$< 2.15 \times 10^{-6}$ (35.7)	$< 1.76 \times 10^{-7}$ (1452.3)

Appendix B

Additional material for Chapter 8

B.1 Notes on individual targets

B.1.1 Sample stars with infrared excesses

0843+516 and 1015+161. These are two metal polluted white dwarfs (Gänsicke et al., 2012) with known infrared excesses (Xu & Jura, 2012; Jura et al., 2007). The excesses at 7.9 μm lay significantly above the predicted fluxes of both disc models (Figure 8.2), likely due to the presence of a strong silicate emission feature.

1018+410. The infrared excess detected at this star is reported here for the first time. There is no reliable near-infrared photometry available for this star, and the 2MASS catalogue only reports $H = 16.7 \pm 0.1$ mag. However, *ugriz* photometry is available from SDSS and the photospheric level of the white dwarf can be well constrained. This degenerate was not observed in the complementary HST survey, and there are no high-resolution optical spectra available. Follow-up should confirm that this is a DAZ white dwarf.

1457–086. This is a known metal polluted white dwarf (Koester et al., 2005) that was found to host a narrow dust ring with high inclination and radial extent of only $0.01\,R_{\odot}$ (Farihi et al., 2009). The fit to the excess emission with both a narrow and highly inclined disc model show that the optical fluxes might be overestimated and 15% error bars were assumed to obtain a satisfactory fit. The star was not observed in the complementary HST survey and additional optical photometry may better constrain the photospheric flux.

2328+107. With a fractional luminosity of only 0.1% this infrared excess, reported here for the first time, joins an expanding list of attenuated discs (Bergfors et al., 2014; Farihi et al., 2010b). Both *JHK* and *ugriz* photometry are available from the 2MASS and SDSS online catalogues and allow to constrain the photospheric level confidently. This degenerate was not observed in the HST survey and high-resolution spectra are not yet

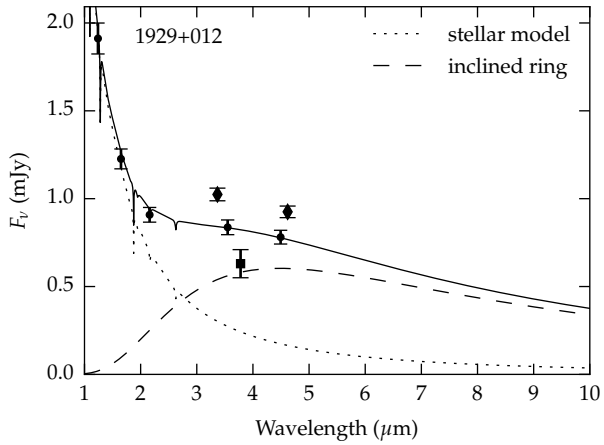

Figure B.1: SED and disc model for the ancillary target 1929+012. Fluxes from WISE All-Sky catalogue are plotted as diamond symbols and fluxes from ISAAC L' as square symbols. Circles symbols represent 2MASS and IRAC fluxes.

available but should reveal the presence of metals.

B.1.2 The infrared excess at 1929+012

This star, also known as GALEX 1931, is a metal polluted white dwarf (Vennes et al., 2010) with a known infrared excess (Gänsicke et al., 2012; Melis et al., 2011; Debes et al., 2011a), included in the Spitzer Program as an ancillary target. Photometry from WISE revealed infrared excess consistent with emission from a dust disc (Debes et al., 2011a), while ground-based photometry out to L-band indicated significant contamination from background sources in the WISE data (Melis et al., 2011). The Spitzer IRAC 3.6 and 4.5 μm photometry of GALEX 1931 confirms the excess emission but is heavily blended with two nearby sources, as shown by the J_sHK_sL' higher resolution images (Melis et al., 2011). Despite the success in extracting most of the flux, the relatively low spatial resolution of IRAC implies likely contamination in the derived fluxes. Therefore, 10% error bars are conservatively assumed in the resulting photometry. Figure B.1 shows the SED of 1929+012, including IRAC, ISAAC L' (Melis et al., 2011), and WISE All-Sky photometry. A fiducial disc model is also plotted. It can be seen that the IRAC fluxes are somewhat lower than the WISE values, but also higher than the ground-based L-band flux. These data indicate mild contamination is likely in the IRAC data and strong contamination is present in WISE.

B.2 Stellar parameters and flux determinations

Table B.1: Stellar parameters and flux determinations for the science targets

WD	$T_{\rm eff}$ (K)	$\log g$ [$\log(\mathrm{cm\,s^{-2}})$]	V (mag)	$M_{\rm wd}$ ($M_\odot$)	$M_{\rm ms}$ ($M_\odot$)	$\log(t_{\rm cool})$ [log (yr)]	$F_{3.6\,\mu\rm m}$ (μJy)	$F_{4.5\,\mu\rm m}$ (μJy)	$F_{5.7\,\mu\rm m}$ (μJy)	$F_{7.9\,\mu\rm m}$ (μJy)
0000+171	20 210	7.99	15.81	0.62	2.07	7.81	64 ± 3	44 ± 2	...	...
0013–241	18 530	7.90	15.38	0.56	1.56	7.90	103 ± 5	66 ± 3	...	...
0018–339	20 630	7.84	14.64	0.54	1.36	7.65	199 ± 10	119 ± 6	...	...
0028–474	17 390	7.65	15.15	0.44	...	7.87	166 ± 8	104 ± 5	...	...
0047–524	18 810	7.73	14.23	0.48	0.83	7.77	...	189 ± 9	...	65 ± 3
0048+202	20 360	7.89	15.38	0.57	1.57	7.72	101 ± 5	64 ± 3	...	...
0048–544	17 870	7.98	15.16	0.61	1.94	8.02	125 ± 6	81 ± 4	...	...
0059+257	21 370	8.04	15.90	0.65	2.35	7.75	68 ± 3	43 ± 2	...	...
0102+095	24 770	7.93	14.44	0.60	1.86	7.28	217 ± 10	146 ± 7	...	...
0110–139	24 690	7.99	15.75	0.63	2.16	7.36	60 ± 3	38 ± 2	...	...
0114–605	24 690	7.75	15.11	0.51	1.08	7.17	112 ± 5	75 ± 3	...	...
0124–257	23 040	7.79	16.18	0.52	1.18	7.35	36 ± 2	25 ± 1	...	...
0127+270	24 870	7.83	15.90	0.55	1.40	7.18	45 ± 2	28 ± 1	...	...
0129–205	19 950	7.88	15.30	0.56	1.54	7.75	103 ± 5	68 ± 3	...	...
0140–392	21 810	7.92	14.35	0.58	1.74	7.59	254 ± 12	165 ± 8	...	...
0155+069	22 010	7.67	15.47	0.47	0.66	7.39	108 ± 5	68 ± 3	...	...
0200+248	23 280	7.86	15.71	0.56	1.50	7.37	105 ± 11	71 ± 7	...	...
0201–052	24 630	7.64	...	0.46	0.62	7.19	41 ± 2	27 ± 1	...	...
0221–055	24 750	7.72	16.22	0.50	0.94	7.16	61 ± 3	42 ± 2	...	...

WD	$T_{\rm eff}$ (K)	$\log g$ [log(cm s^{-2})]	V (mag)	$M_{\rm wd}$ ($M_\odot$)	$M_{\rm ms}$ ($M_\odot$)	$\log(t_{\rm cool})$ [log (yr)]	$F_{3.6\,\mu{\rm m}}$ (μJy)	$F_{4.5\,\mu{\rm m}}$ (μJy)	$F_{5.7\,\mu{\rm m}}$ (μJy)	$F_{7.9\,\mu{\rm m}}$ (μJy)
0222–265	23 200	7.91	15.68	0.58	1.74	7.43	76 ± 4	51 ± 2	...	...
0227+050	19 340	7.76	12.80	0.50	0.97	7.73	...	775 ± 38	...	255 ± 12
0229+270	24 160	7.90	15.52	0.58	1.71	7.31	79 ± 4	51 ± 2	...	...
0231–054	17 310	8.45	14.31	0.90	4.60	8.42	...	236 ± 11	...	71 ± 3
0242–174	20 660	7.85	15.38	0.55	1.41	7.66	96 ± 4	62 ± 3	...	...
0300–232	22 370	8.39	15.68	0.86	4.32	8.04	72 ± 3	44 ± 2	...	...
0307+149	21 410	7.91	15.38	0.58	1.69	7.62	112 ± 5	71 ± 3	...	...
0308+188	18 450	7.72	14.19	0.48	0.79	7.80	319 ± 16	204 ± 10	...	...
0308–230	23 570	8.54	15.08	0.96	5.18	8.10	108 ± 5	69 ± 3	...	...
0331+226	21 450	7.78	15.28	0.52	1.12	7.52	110 ± 5	70 ± 3	...	...
0341+021	22 150	7.27	15.41	0.33	...	7.27	101 ± 5	60 ± 3	...	...
0349–256	20 970	7.91	15.67	0.58	1.67	7.67	71 ± 3	47 ± 2	...	...
0352+018	22 110	7.80	15.57	0.52	1.19	7.46	79 ± 4	50 ± 2	...	...
0358–514	23 380	7.93	15.72	0.59	1.82	7.43	69 ± 3	47 ± 2	...	...
0403–414	22 700	7.94	16.35	0.60	1.85	7.51	41 ± 2	26 ± 1	...	...
0410+117	21 070	7.84	13.91	0.54	1.37	7.60	...	246 ± 12	...	63 ± 3
0414–406	20 940	8.00	16.13	0.63	2.13	7.75	47 ± 2	29 ± 1	...	...
0416–105	24 850	7.92	15.37	0.59	1.80	7.26	87 ± 4	52 ± 2	...	...
0418–103	23 390	8.29	15.68	0.80	3.76	7.87	70 ± 3	44 ± 2	...	...
0421+162	19 620	8.03	14.29	0.64	2.23	7.89	275 ± 13	171 ± 8	115 ± 6	60 ± 5
0431+126	21 370	7.97	14.23	0.61	1.98	7.68	...	177 ± 8	...	34 ± 7

WD	$T_{\rm eff}$ (K)	$\log g$ [$\log(\mathrm{cm\,s^{-2}})$]	V (mag)	$M_{\rm wd}$ ($M_\odot$)	$M_{\rm ms}$ ($M_\odot$)	$\log(t_{\rm cool})$ [log (yr)]	$F_{3.6\,\mu\rm m}$ (μJy)	$F_{4.5\,\mu\rm m}$ (μJy)	$F_{5.7\,\mu\rm m}$ (μJy)	$F_{7.9\,\mu\rm m}$ (μJy)
0452–347	21 210	7.84	16.13	0.54	1.36	7.59	57 ± 3	36 ± 1	...	...
0455–532	24 430	7.55	...	0.43	...	7.22	30 ± 1	17 ± 1	...	...
0507+045.1	20 840	7.90	14.22	0.57	1.62	7.67	...	206 ± 10	...	136 ± 12
0843+516	23 870	7.90	16.04	0.58	1.70	7.35	136 ± 6	137 ± 7	103 ± 5	162 ± 9
0854+404	22 250	7.91	14.81	0.58	1.71	7.53	156 ± 7	97 ± 4	...	...
0859–039	23 730	7.79	13.19	0.53	1.21	7.28	703 ± 35	435 ± 21	...	...
0920+363	24 060	7.63	16.07	0.46	0.58	7.22	56 ± 2	35 ± 1	...	...
0933+025	22 360	8.04	15.93	0.65	2.37	7.66	3569 ± 178	2450 ± 122	...	...
0938+550	18 530	8.10	14.79	0.68	2.62	8.04	197 ± 9	123 ± 6	...	...
0944+192	17 440	7.88	14.51	0.55	1.47	8.00	246 ± 12	159 ± 8	...	...
0947+325	22 060	8.31	15.50	0.82	3.87	7.98	87 ± 4	57 ± 2	...	...
0954+697	21 420	7.91	15.96	0.58	1.69	7.62	63 ± 3	44 ± 2	...	...
1003–023	20 610	7.89	15.27	0.57	1.58	7.69	120 ± 6	73 ± 3	...	...
1005+642	19 660	7.93	13.69	0.58	1.75	7.82	460 ± 23	291 ± 14	...	...
1012–008	23 200	8.07	15.59	0.67	2.57	7.62	79 ± 4	49 ± 2	...	...
1013+256	21 990	8.00	16.32	0.63	2.15	7.65	41 ± 2	25 ± 1	...	...
1015+161	19 950	7.92	15.61	0.58	1.73	7.78	196 ± 9	166 ± 8	146 ± 7	128 ± 7
1017+125	21 390	7.88	15.67	0.56	1.53	7.60	73 ± 3	48 ± 2	27 ± 3	18 ± 6
1018+410	22 390	8.04	16.37	0.65	2.37	7.66	85 ± 4	74 ± 3	...	...
1020–207	19 920	7.93	15.04	0.58	1.74	7.79	130 ± 6	83 ± 4	...	...
1034+492	20 650	8.17	15.43	0.73	3.05	7.94	100 ± 5	58 ± 3	48 ± 3	28 ± 5

WD	$T_{\rm eff}$ (K)	$\log g$ [log(cm s^{-2})]	V (mag)	$M_{\rm wd}$ ($M_\odot$)	$M_{\rm ms}$ ($M_\odot$)	$\log(t_{\rm cool})$ [log (yr)]	$F_{3.6\,\mu{\rm m}}$ (μJy)	$F_{4.5\,\mu{\rm m}}$ (μJy)	$F_{5.7\,\mu{\rm m}}$ (μJy)	$F_{7.9\,\mu{\rm m}}$ (μJy)
1038+633	24 450	8.38	15.15	0.86	4.31	7.90	108 ± 5	67 ± 3	57 ± 3	17 ± 4
1049+103	20 550	7.91	15.81	0.58	1.67	7.71	3743 ± 187	2579 ± 129	...	...
1049–158	20 040	8.28	14.36	0.79	3.65	8.08	256 ± 12	155 ± 7	...	...
1052+273	23 100	8.37	14.12	0.86	4.23	7.98	304 ± 15	193 ± 9	129 ± 7	59 ± 7
1058–129	24 310	8.71	14.91	1.06	6.10	8.20	128 ± 6	81 ± 4	41 ± 6	29 ± 8
1102+748	19 710	8.36	15.05	0.84	4.13	8.18	136 ± 6	89 ± 4	...	...
1104+602	17 920	8.02	13.74	0.63	2.17	8.04	463 ± 23	293 ± 14	...	...
1115+166	22 090	8.12	15.05	0.70	2.80	7.77	133 ± 6	85 ± 4	69 ± 4	43 ± 6
1122–324	21 670	7.86	15.82	0.55	1.44	7.55	62 ± 3	40 ± 2	...	...
1129+155	17 740	8.03	14.09	0.64	2.22	8.06	376 ± 18	234 ± 11	159 ± 8	78 ± 7
1133+293	23 030	7.84	14.88	0.55	1.41	7.39	146 ± 7	90 ± 4	63 ± 6	43 ± 8
1134+300	21 280	8.55	12.45	0.96	5.23	8.26	1382 ± 69	889 ± 44	...	...
1136+139	23 920	7.83	...	0.54	1.37	7.28	22 ± 2	15 ± 1	...	...
1201–001	19 770	8.26	15.16	0.78	3.55	8.08	124 ± 6	80 ± 4	48 ± 3	17 ± 3
1204–322	21 260	8.00	15.62	0.62	2.12	7.72	80 ± 4	52 ± 2	...	...
1229–013	19 430	7.47	14.46	0.38	...	7.54	239 ± 12	152 ± 7	90 ± 5	64 ± 7
1230–308	22 760	8.28	15.73	0.80	3.71	7.90	69 ± 3	45 ± 2	...	...
1233–164	24 890	8.21	15.10	0.76	3.33	7.66	114 ± 5	72 ± 3	...	...
1243+015	21 640	7.82	16.46	0.53	1.27	7.52	33 ± 1	23 ± 1	...	...
1249+182	19 910	7.73	15.24	0.49	0.85	7.65	95 ± 4	60 ± 3	...	...
1257+048	21 760	7.95	14.94	0.60	1.89	7.62	150 ± 7	97 ± 4	...	...

WD	$T_{\rm eff}$ (K)	$\log g$ [$\log(\mathrm{cm\,s^{-2}})$]	V (mag)	$M_{\rm wd}$ ($M_\odot$)	$M_{\rm ms}$ ($M_\odot$)	$\log(t_{\rm cool})$ [log (yr)]	$F_{3.6\,\mu\rm m}$ (μJy)	$F_{4.5\,\mu\rm m}$ (μJy)	$F_{5.7\,\mu\rm m}$ (μJy)	$F_{7.9\,\mu\rm m}$ (μJy)
1310–305	20 350	7.82	14.48	0.53	1.25	7.66	223 ± 11	143 ± 7	...	...
1323–514	19 360	7.76	14.39	0.50	0.98	7.73	275 ± 13	175 ± 8	...	...
1325+279	21 270	8.04	15.80	0.65	2.35	7.76	63 ± 7	42 ± 4	...	...
1325–089	17 020	7.81	...	0.52	1.12	8.00	165 ± 8	107 ± 5	...	...
1330+473	22 460	7.95	15.23	0.60	1.91	7.55	106 ± 5	70 ± 3	...	...
1334–160	18 650	8.32	...	0.82	3.86	8.20	116 ± 6	98 ± 5	...	...
1335+369	20 510	7.78	...	0.51	1.08	7.62	2056 ± 102	1326 ± 66	...	...
1337+705	20 460	7.90	12.60	0.57	1.62	7.71	...	709 ± 35	...	254 ± 13
1338+081	24 440	7.65	16.44	0.47	0.66	7.20	30 ± 1	19 ± 1	...	...
1353+409	23 530	7.59	15.49	0.44	...	7.26	76 ± 3	50 ± 2	32 ± 4	26 ± 5
1408+323	18 150	7.95	13.97	0.59	1.81	7.97	...	251 ± 12	...	92 ± 7
1433+538	22 410	7.80	16.14	0.53	1.21	7.43	906 ± 45	632 ± 31	...	...
1449+168	22 350	7.79	15.39	0.52	1.15	7.42	91 ± 4	58 ± 2	...	...
1451+006	25 480	7.89	15.27	0.58	1.69	7.16	98 ± 5	63 ± 3	...	...
1452–042	23 530	8.19	16.29	0.74	3.21	7.74	51 ± 2	36 ± 2	...	...
1457–086	21 450	7.97	15.76	0.61	1.99	7.68	113 ± 5	73 ± 3	48 ± 3	...
1459+347	21 520	8.48	15.79	0.92	4.84	8.18	65 ± 3	41 ± 2	...	...
1507+220	19 870	7.75	14.95	0.49	0.91	7.66	150 ± 7	96 ± 4	...	...
1524–749	23 090	7.74	15.99	0.50	0.99	7.32	64 ± 3	39 ± 2	...	...
1525+257	22 290	8.28	15.73	0.80	3.70	7.93	69 ± 3	44 ± 2	...	...
1527+090	21 200	7.85	14.30	0.55	1.39	7.59	259 ± 13	164 ± 8	...	...

WD	$T_{\rm eff}$ (K)	$\log g$ [log(cm s^{-2})]	V (mag)	$M_{\rm wd}$ ($M_\odot$)	$M_{\rm ms}$ ($M_\odot$)	$\log(t_{\rm cool})$ [log (yr)]	$F_{3.6\,\mu{\rm m}}$ (μJy)	$F_{4.5\,\mu{\rm m}}$ (μJy)	$F_{5.7\,\mu{\rm m}}$ (μJy)	$F_{7.9\,\mu{\rm m}}$ (μJy)
1531–022	18 620	8.41	13.97	0.87	4.41	8.30	...	243 ± 12	...	103 ± 8
1533–057	20 000	8.50	15.38	0.93	4.94	8.29	94 ± 4	57 ± 2	...	...
1535+293	24 470	7.92	...	0.59	1.81	7.30	50 ± 2	30 ± 1	...	...
1547+057	24 360	8.36	15.94	0.85	4.16	7.88	56 ± 2	36 ± 1	...	...
1548+149	21 450	7.86	15.16	0.55	1.45	7.57	122 ± 6	77 ± 3	...	...
1614+136	22 020	7.21	15.23	0.32	...	7.23	113 ± 5	67 ± 3	54 ± 5	15 ± 5
1620–391	24 680	7.93	11.00	0.60	1.85	7.29	4988 ± 249	3113 ± 155	1936 ± 96	1073 ± 53
1633+676	23 660	7.97	16.25	0.62	2.04	7.45	34 ± 1	21 ± 1	...	...
1647+375	21 980	7.89	14.91	0.57	1.61	7.55	137 ± 6	86 ± 4	...	...
1713+332	22 120	7.43	14.39	0.38	...	7.32	...	184 ± 9	...	57 ± 6
1755+194	24 440	7.80	15.99	0.53	1.28	7.21	54 ± 3	33 ± 1	...	...
1914–598	19 760	7.84	...	0.54	1.31	7.74	288 ± 14	196 ± 10	...	...
1929+012	20 890	7.91	14.20	0.58	1.68	7.68	837 ± 41	781 ± 39	...	...
1943+163	19 760	7.79	13.95	0.51	1.11	7.71	...	230 ± 11	...	100 ± 7
1953–715	19 270	7.87	15.06	0.55	1.47	7.81	132 ± 6	84 ± 4	...	...
2021–128	20 750	7.82	15.20	0.53	1.26	7.62	112 ± 5	74 ± 3	...	...
2032+188	18 200	7.36	15.32	0.34	...	7.58	125 ± 6	80 ± 4	...	...
2039–202	19 740	7.79	12.33	0.51	1.08	7.70	...	1035 ± 51	...	371 ± 20
2046–220	23 410	7.83	15.37	0.54	1.36	7.33	88 ± 4	55 ± 2	...	...
2058+181	17 350	7.75	15.20	0.49	0.88	7.94	149 ± 7	97 ± 5	...	...
2134+218	18 000	7.86	14.48	0.55	1.39	7.93	...	158 ± 8	...	61 ± 6

WD	$T_{\rm eff}$ (K)	$\log g$ [$\log(\mathrm{cm\,s^{-2}})$]	V (mag)	$M_{\rm wd}$ ($M_\odot$)	$M_{\rm ms}$ ($M_\odot$)	$\log(t_{\rm cool})$ [log (yr)]	$F_{3.6\,\mu\rm m}$ (μJy)	$F_{4.5\,\mu\rm m}$ (μJy)	$F_{5.7\,\mu\rm m}$ (μJy)	$F_{7.9\,\mu\rm m}$ (μJy)
2149+021	17 930	7.86	12.76	0.54	1.38	7.94	...	776 ± 38	...	272 ± 15
2152–045	19 840	7.38	15.69	0.35	...	7.44	83 ± 4	52 ± 2	...	...
2200–136	24 730	7.61	15.33	0.45	0.53	7.19	106 ± 5	66 ± 3	...	...
2204+071	24 450	7.95	15.78	0.61	1.95	7.34	68 ± 3	41 ± 2	...	...
2210+233	23 230	8.24	15.84	0.77	3.48	7.82	59 ± 3	34 ± 1	...	...
2220+133	22 580	8.30	15.61	0.81	3.82	7.93	75 ± 3	45 ± 2	...	...
2220+217.1	18 740	8.24	15.22	0.77	3.43	8.13	79 ± 4	45 ± 2	...	...
2229+235	19 300	7.90	16.01	0.57	1.60	7.83	59 ± 3	38 ± 2	23 ± 2	...
2231–267	21 590	7.99	14.97	0.62	2.09	7.68	162 ± 8	108 ± 5	...	...
2238–045	17 540	8.18	16.90	0.73	3.04	8.17	22 ± 1	13 ± 0	...	...
2244+210	24 110	7.89	16.45	0.57	1.66	7.31	35 ± 1	21 ± 1	...	...
2257+162	24 580	7.49	15.97	0.41	...	7.24	417 ± 20	287 ± 14	195 ± 10	128 ± 8
2306+124	20 360	7.99	15.08	0.62	2.08	7.80	124 ± 6	76 ± 3	...	...
2322–181	21 680	7.90	15.28	0.57	1.65	7.59	99 ± 5	61 ± 3	...	...
2328+107	21 000	7.78	15.76	0.51	1.08	7.56	115 ± 5	77 ± 3	...	...
2359–324	22 480	7.74	16.26	0.50	0.97	7.38	43 ± 2	27 ± 1	...	...

Glossary

Symbol	Description	Equation Number
A_E	Einstein A coefficient, a measure of the probability of spontaneous emission for a given transition between states	3.1.0, 3.1.1
A	Equivalent atmospheric depth	2.3.7, 2.3.12, 2.3.13, 2.5.10
α	A scaling factor ranging from 0 to 1	2.6.12
a	Semi major axis	1.1.1, 1.1.3, 1.1.5, 1.1.8, 1.1.10–1.1.13, 6.2.0, 6.3.0
AU	Astronomical Unit	1.0.0, 1.1.15, 1.2.0, 7.2.0
α_{V}	Weighting factor between optical opacities	2.6.7, 2.6.8, 5.2.0
$\mathcal{B}$	Bayes factor	2.1.19–2.1.22, 6.3.0
B	Planck function	2.3.14, 2.3.19, 2.3.21, 2.3.22
b	Impact parameter in a transit light curve	1.1.6, 1.1.7, 1.1.12
$\mathscr{C}$	Closure operation	2.6.17
$\mathcal{C}$	Correlation length in terms of atmospheric scale heights	2.6.11
c	Speed of light	3.1.1, 3.1.4
χ^2	Chi squared	0.0.0, 4.3.0, 6.3.0
χ_{excess}	Excess significant	8.4.1
D	Total transit depth	2.3.13

Symbol	Description	Equation Number
ΔF	Change in star flux when the planet is in transit and out of transit	1.1.6–1.1.9, 1.1.13
$\mathscr{E}$	Bayesian evidence	0.0.0, 2.1.3, 2.1.12, 2.1.15, 2.1.19, 2.1.20, 6.2.0, 6.3.0, B.2.0
E	Exponential integral	2.6.7
exp	Expected value	2.1.5, 2.1.7
E_{low}	Lower-state energy	3.1.1, 3.1.2
ε	Error vector	2.1.1–2.1.3
erf	Error function	2.1.22
$F_{\text{no transit}}$	Star flux when the orbiting planet is transiting	1.1.6
F_{transit}	Star flux when the orbiting planet is not transiting	1.1.6
f_G	Doppler broadening profile	3.1.3, 3.1.7
F_{King}	King's factor in Rayleigh scattering, also known as depolarisation term, describing the effect of molecular anisotropy	2.3.12
f_L	Lorentzian broadening profile	3.1.5, 3.1.7
F	Flux	8.2.0, 8.4.1, B.2.0
F_p	Total emitted flux from the planet	2.3.22
F_*	Total emitted flux from the star	2.3.22
f_V	Voigt broadening profile	3.1.7, 3.2.1
G	Gravitational constant	1.1.1, 1.1.9–1.1.11, 2.6.4
g	Gravity	2.6.2, 2.6.4, 2.6.8, 7.4.0, B.2.0
g_{up}	Upper-state degeneracy	3.1.1, 3.1.2
γ_G	Doppler width	3.1.3, 3.1.4, 3.1.8, 3.2.3

Symbol	Description	Equation Number
γ_L	Lorentzian width	0.0.0, 3.1.5, 3.1.6, 3.1.8, 3.2.3, 4.2.2, 4.2.3
γ_V	Voigt width	3.2.3, 4.2.1
$g_{p,\mathrm{surf}}$	Planet surface gravity	1.1.12, 1.1.13
$g_{*,\mathrm{surf}}$	Star surface gravity	6.2.0, 6.3.0
$\mathcal{H}$	Scale height in units of transit depth (parts per million)	4.2.2
H	Atmospheric scale height	2.3.11, 2.6.2, 2.6.5, 4.2.2, 4.2.3, 7.4.0
h	Planck constant	3.1.1
$\mathcal{I}$	A generic integral	2.1.6, 2.1.9
I	Specific intensity	1.1.14, 1.1.15, 2.3.0–2.3.2, 2.3.4, 2.3.5, 2.3.14, 2.3.17, 2.3.19, 2.3.21, 2.3.22
i	Orbital inclination	1.1.3–1.1.5, 1.1.12, 1.1.13, 6.2.0, 6.3.0
i_{disc}	Disc inclination with respect to the normal to the line of sight	8.3.0
$\mathbf{J}$	Jacobian or weighting function matrix	2.1.2
J	Total angular momentum quantum number	0.0.0, 3.2.3
K_*	Velocity semi-amplitude	1.1.3–1.1.5, 1.1.13
k	Opacity absorption coefficient	2.5.0–2.5.2, 2.5.5, 2.5.7–2.5.10
κ_{IR}	Mean opacity in the infrared	2.6.7, 2.6.8, 5.2.0
κ_V	Mean opacity in the optical	2.6.7, 5.2.0
k_B	Boltzmann constant	2.4.0, 2.6.1, 2.6.2, 3.1.1, 3.1.2, 3.1.4, 7.4.0

Symbol	Description	Equation Number
K_{zz}	Diffusion coefficient	5.2.0, 7.2.0
l	Optical path along the line of sight	2.3.6–2.3.11, 2.5.10
λ	Wavelength	2.3.0–2.3.9, 2.3.12–2.3.22, 2.5.10
$L_{\rm IR}$	Infrared disc luminosity	8.5.0
L_*	Star luminosity	8.5.0
$\mathcal{M}$	Forward model	2.1.1–2.1.4, 2.1.6, 2.1.7, 2.1.12–2.1.14, 2.1.16–2.1.20, B.2.0
$m_{\rm mol}$	Molecule mass	2.6.3, 3.1.4
$M_\oplus$	Earth mass	1.1.15, 1.2.0, 6.3.0
M_J	Jupiter mass	1.1.5, 1.1.15, 1.2.0, 4.1.0, 4.2.1, 5.2.0, 6.2.0, 7.2.0
M_p	Planet mass	1.1.2–1.1.5, 1.1.10, 1.1.15, 2.4.0, 2.6.4, 4.1.0, 4.2.1, 5.2.0, 6.2.0, 6.3.0, 7.2.0
$M_\odot$	Solar mass	6.2.0
$M_{\rm mol}$	Molar mass	2.5.1
$M_{\rm ms}$	Initial main-sequence progenitor mass of a white dwarf	8.2.0, B.2.0
M_*	Star mass	1.1.0–1.1.5, 1.1.10, 1.1.11, 6.2.0, 6.3.0
μ	Equivalent to $\cos(\phi)$	1.1.14, 1.1.15, 2.3.14, 2.3.17–2.3.22
μ	Mean molecular weight	2.6.2, 2.6.3, 6.3.0, 7.4.0
$M_{\rm wd}$	White dwarf mass	8.2.0, B.2.0
n_T	Temperature dependence exponent	0.0.0, 3.1.6, 3.2.3, 4.2.3
N_A	Avogadro number	2.5.1

Symbol	Description	Equation Number
N_{air}	Loschmidt constant (2.6867805×10^{25} m^{-3})	2.3.12
N_{broad}	Number of broadening agents	3.1.6, 4.2.3
N_{gas}	Number of molecules	2.3.6, 2.3.8, 2.3.9, 2.3.15, 2.3.21, 2.3.22, 2.5.7–2.5.10, 2.6.3, 2.6.14, 2.6.16, 2.6.17
N_{layers}	Number of layers in the atmosphere	2.3.8, 2.3.9, 2.3.20–2.3.22, 2.4.0, 2.5.10, 2.6.0
N_{live}	Number of "live" random samples drawn from the full prior in Nested Sampling	2.1.17
N_{quad}	Number of Gaussian quadrature points	2.3.21, 2.3.22, 2.5.5, 2.5.9, 2.5.10, 4.4.0
n_R	Refractive index	2.3.12
N_{samp}	Number of samples in a Monte Carlo integration	2.1.5, 2.1.6, 2.1.8, 2.1.9
n_σ	Sigma significance	2.1.20, 2.1.22
ν	Wavenumber	2.5.0–2.5.2, 2.5.4–2.5.10, 3.1.1–3.1.8, 3.2.1, 3.2.2, 4.2.1, 4.2.3, 4.3.0, 4.4.0
o	Optical path length	2.3.2, 2.3.3
P	Orbital period	1.1.1, 1.1.3–1.1.5, 1.1.8, 1.1.10, 1.1.11, 1.1.13, 6.2.0, 6.3.0
P	Atmospheric pressure	2.3.11, 2.5.5, 2.6.0, 2.6.1, 2.6.5, 2.6.8, 2.6.11, 3.1.4, 3.1.6, 3.2.0–3.2.3, 4.2.3, 4.4.0, 5.2.0, 6.2.0, 7.4.0, B.2.0

Symbol	Description	Equation Number
$\mathcal{P}$	Probability	2.1.3, 2.1.4, 2.1.6, 2.1.7, 2.1.10–2.1.20, B.2.0
$p_{\rm val}$	p-value	2.1.20–2.1.22
$\mathcal{P}(\mathbf{x}\|\mathbf{y},\mathcal{M})$	Posterior distribution of $\mathbf{x}$, given $\mathbf{y}$ and $\mathcal{M}$	2.1.3, 2.1.6, 2.1.7
$\mathcal{P}(\mathbf{x},\mathcal{M})$	Prior distribution of $\mathbf{x}$	2.1.3, 2.1.12, 2.1.13, 2.1.17
$\mathcal{P}(\mathbf{y}\|\mathcal{M})$	Bayesian partition function, or Bayesian evidence (see also $\mathscr{E}$)	2.1.3, 2.1.12, 2.1.14, 2.1.16, 2.1.18
$\mathcal{P}(\mathbf{y}\|\mathbf{x},\mathcal{M})$	Likelihood distribution of $\mathbf{x}$	2.1.3, 2.1.4, 2.1.12, 2.1.13, 2.1.18
$P_{\rm cloud}$	Cloud top pressure	6.2.0
ϕ	Angle between the line of sight and the emergent intensity	2.3.14, B.2.0
Q	Partition function	3.1.1, 3.1.2
R	Spectral resolution	1.1.5, 1.3.0, 2.4.0, 2.5.0, 4.1.0, 4.2.1, 4.2.3, 4.3.0, 4.4.0, 4.5.0, 5.2.0
$R_\oplus$	Earth radius	6.3.0, 8.3.0
R_J	Jupiter radius	2.4.0, 4.1.0, 4.2.1, 5.2.0, 6.2.0, 6.3.0, 7.2.0
R_p	Planet radius	1.1.5, 1.1.6, 1.1.13, 2.3.7–2.3.10, 2.3.13, 2.3.22, 2.4.0, 2.5.10, 2.6.4, 2.6.5, 4.1.0, 4.2.1, 4.2.2, 5.2.0, 6.2.0, 6.3.0, 7.2.0
$R_\odot$	Solar radius	2.4.0, 4.1.0, 4.2.1, 6.2.0, 6.3.0, 8.3.0, 8.5.0, B.1.0
ρ_N	Number density	2.3.1–2.3.4, 2.3.6, 2.3.8, 2.3.9, 2.3.16, 2.3.21, 2.3.22, 2.6.0, 2.6.1

Symbol	Description	Equation Number
ρ	Mass density	2.5.2, 2.5.4–2.5.10
r_{in}	Inner radius of the circumstellar disc	8.3.0
r_{out}	Outer radius of the circumstellar disc	8.3.0, 8.5.0
ρ_*	Star density	1.1.8–1.1.10
R_*	Star radius	1.1.5–1.1.8, 1.1.10, 1.1.12, 2.3.13, 2.3.22, 2.4.0, 4.1.0, 4.2.1, 4.2.2, 6.2.0, 6.3.0, 8.3.0
S	Line intensity	3.1.1, 3.2.1
$\mathbb{S}$	The unit simplex space	2.6.15
σ	Opacity cross section	2.1.22, 2.3.1–2.3.4, 2.3.6, 2.3.9, 2.3.12, 2.3.16, 2.3.21, 2.3.22, 2.5.0, 2.5.1, 3.2.1, 3.2.2, 4.2.1
T	Temperature	2.3.14, 2.3.17, 2.3.19, 2.3.21, 2.3.22, 2.5.5, 2.6.0–2.6.2, 2.6.6, 2.6.8, 3.1.1, 3.1.2, 3.1.4, 3.1.6, 3.2.0–3.2.3, 4.1.0, 4.2.1, 4.2.3, 4.4.0, 5.2.0, 5.3.0, 6.2.0, 6.3.0, 7.2.0, 7.4.0, 8.1.0, 8.2.0, 8.3.0, 8.5.0, 8.6.0, B.2.0
$\mathcal{T}$	Transmittance	2.3.18–2.3.20, 2.5.2, 2.5.4–2.5.9
t	The shrinkage ratio in a Nested Sampling run, representing the shrinkage in prior volume between consecutive likelihood contours	2.1.17
T_0	Epoch of mid transit (in Julian days)	6.2.0, 6.3.0

Symbol	Description	Equation Number
$t_{\rm I}$	First contact point in a planetary transit light curve	1.1.7, B.2.0
$t_{\rm II}$	Second contact point in a planetary transit light curve	1.1.7, B.2.0
$t_{\rm III}$	Third contact point in a planetary transit light curve	1.1.7, B.2.0
$t_{\rm IV}$	Fourth contact point in a planetary transit light curve	1.1.7, B.2.0
t_F	Transit duration defined between the second ($t_{\rm II}$) and third ($t_{\rm III}$) contact points	1.1.7–1.1.9
t_T	Transit duration defined between the first ($t_{\rm I}$) and fourth ($t_{\rm IV}$) contact points	1.1.7–1.1.9
τ	Optical depth	2.3.2, 2.3.3, 2.3.5, 2.3.6, 2.3.8, 2.3.14–2.3.20, 2.3.22, 2.6.8, 4.2.1
$\tau_{\rm th}$	Thermal optical depth	2.6.6–2.6.8
$\tau_{\rm disc}$	Disc fractional luminosity	8.5.0
$t_{\rm cool}$	Cooling age of the white dwarf	8.5.0, 8.6.0, B.2.0
$T_{\rm eff}$	Effective temperature	7.2.0, 8.1.0, 8.2.0, 8.5.0, 8.6.0, B.2.0
$T_{\rm in}$	Inner temperature of the circumstellar disc	8.3.0, 8.5.0
$T_{\rm int}$	Internal heat flux	2.6.6, 2.6.8, 5.2.0, 7.2.0
$T_{\rm IR}$	Blackbody dust temperature of the circumstellar disc	8.5.0
$T_{\rm irr}$	Stellar irradiation at the top of the atmosphere	2.6.6, 2.6.8, 5.2.0
$T_{\rm out}$	Outer temperature of the circumstellar disc	8.3.0
T_p	Planet equilibrium temperature	6.2.0, 6.3.0
T_*	Star effective temperature	4.2.1, 6.2.0, 6.3.0
$T_{\rm surf}$	Surface temperature	2.3.19
u	Limb darkening coefficient	1.1.14, 1.1.15

Symbol	Description	Equation Number
V	Visual magnitude	B.2.0
V	Visual magnitude	1.1.15
var	Variance	2.1.9
v_p	Tangential velocity of planet	1.1.2, 1.1.3
v_*	Tangential velocity of star	1.1.2–1.1.4
w	Weight (usually a number between 0 and 1)	2.1.16–2.1.18, 2.3.21, 2.3.22, 2.5.5, 2.5.7–2.5.10
χ	Mixing ratio, or molar fraction, of a given species	2.3.6, 2.3.9, 2.3.16, 2.5.2, 2.5.4–2.5.8, 2.5.10, 2.6.0, 2.6.3, 2.6.13, 2.6.14, 2.6.16, 2.6.17, 3.1.6, 4.2.3
$\mathscr{X}$	Prior volume	2.1.13–2.1.18, B.2.0
X	A generic variable. Various definition given in the text.	2.1.9, 2.1.10, 2.1.15, 2.1.17, 2.6.15
$\mathbf{x}$	State vector	2.1.0–2.1.13, 2.1.17, 2.1.18, B.2.0
ξ	Defines the iso-likelihood contour in the prior volume $\mathscr{X}$	2.1.13
$\mathbf{y}$	Measurement or data vector	2.1.0–2.1.4, 2.1.6, 2.1.7, 2.1.12–2.1.14, 2.1.16, 2.1.18–2.1.20, B.2.0
z	Altitude	2.3.3–2.3.11, 2.3.14–2.3.16, 2.3.18, 2.3.21, 2.3.22, 2.5.1, 2.5.2, 2.5.4–2.5.10, 2.6.0, 2.6.4, 2.6.5

Bibliography

1966, A class of Markov processes associated with nonlinear parabolic equations

2010, Memorie della Societa Astronomica Italiana, 81, 921

Adams, E. R., Seager, S., & Elkins-Tanton, L. 2008, ApJ, 673, 1160

Adler-Golden, S., Berk, A., & Bernstein, L. S. 1998, Proc 7th Ann JPL

Agúndez, M., Venot, O., Iro, N., et al. 2012, Astronomy & Astrophysics, 548, A73

Agúndez, M., Venot, O., Selsis, F., & Iro, N. 2014, The Astrophysical Journal, 781, 68

Ahn, C. P., Alexandroff, R., Allende Prieto, C., et al. 2012, The Astrophysical Journal Supplement Series, 203, 21

Aitchison, J. 1982, Journal of the Royal Statistical Society Series B (...

—. 1986

—. 1994, Lecture Notes-Monograph Series

Alcock, C., Fristrom, C. C., & Siegelman, R. 1986, ApJ, 302, 462

Ali-Dib, M., Mousis, O., Petit, J.-M., & Lunine, J. I. 2014, The Astrophysical Journal, 785, 125

Alibert, Y., Mordasini, C., & Benz, W. 2004, Astronomy & Astrophysics, 417, L25

Alibert, Y., Mordasini, C., Benz, W., & Winisdoerffer, C. 2005, Astronomy & Astrophysics, 434, 343

Allard, N. F., Allard, F., & Kielkopf, J. F. 2005, Astronomy & Astrophysics, 440, 1195

Allard, N. F., Kielkopf, J. F., Spiegelman, F., Tinetti, G., & Beaulieu, J. P. 2012, Astronomy & Astrophysics, 543, A159

Alonso, R., Brown, T. M., Charbonneau, D., et al. 2007, Transiting Extrasolar Planets Workshop ASP Conference Series, 366, 13

Anderlohr, J. M., Pires da Cruz, A., Bounaceur, R., & Battin-Leclerc, F. 2010, Combustion Science and Technology, 182, 39

Anderson, H. L. 1986, Journal of Statistical Physics, 43, 731

Anglada-Escudé, G., Amado, P. J., Barnes, J., et al. 2016, Nature, 1

Auvergne, M., Bodin, P., Boisnard, L., et al. 2009, Astronomy & Astrophysics, 506, 411

Bakos, G., Lazar, J., Papp, I., Sari, P., & Green, E. M. 2002, The Publications of the Astronomical Society of the Pacific, 114, 974

Bakos, G., Noyes, R. W., Kovacs, G., et al. 2004, The Publications of the Astronomical Society of the Pacific, 116, 266

Bakos, G., Csubry, Z., Penev, K., et al. 2013, Publications of the Astronomical Society of Pacific, 125, 154

Baranne, A., Queloz, D., Mayor, M., et al. 1996, Astronomy and Astrophysics Supplement, 119, 373

Barber, R. J., Strange, J. K., Hill, C., et al. 2014a, Monthly Noticese of the Royal Astronomical Society, 437, 1828

Barber, R. J., Tennyson, J., Harris, G. J., & Tolchenov, R. N. 2006, Monthly Notices of the Royal Astronomical Society, 368, 1087

Barber, S. D., Kilic, M., Brown, W. R., & Gianninas, A. 2014b, The Astrophysical Journal, 786, 77

Barber, S. D., Patterson, A. J., Kilic, M., et al. 2012, ApJ, 760, 26

Barricelli, N. A. 1954, Numerical examples of processes of evolution (Methodos)

Barstow, J. K., Aigrain, S., Irwin, P. G. J., et al. 2013, Monthly Notices of the Royal Astronomical Society, 430, 1188

Barstow, J. K., Aigrain, S., Irwin, P. G. J., Kendrew, S., & Fletcher, L. N. 2015, Monthly Notices of the Royal Astronomical Society, 448, 2546

Barstow, J. K., & Irwin, P. G. J. 2016, arXiv.org, arXiv:1605.07352

Batalha, N., Kalirai, J., Lunine, J., Clampin, M., & Lindler, D. 2015, arXiv.org, arXiv:1507.02655

Batalha, N. M., Rowe, J. F., Bryson, S. T., et al. 2013, The Astrophysical Journal Supplement, 204, 24

Bates, D. R. 1984, Planetary and Space Science (ISSN 0032-0633), 32, 785

Battin-Leclerc, F., Bounaceur, R., Glaude, P.-A., & Belmekki, N. 2006, arXiv.org, arXiv:physics

Bean, J. L., Desert, J.-M., Seifahrt, A., et al. 2013, The Astrophysical Journal, 771, 108

Bean, J. L., Miller-Ricci Kempton, E., & Homeier, D. 2010, Nature, 468, 669

Beaulieu, J.-P., Albrow, M., Bennett, D., et al. 2007, The Messenger, 128

Beaulieu, J. P., Kipping, D., Batista, V., et al. 2010, Monthly Notices of the Royal Astronomical Society, 409, 963

Becklin, E. E., Farihi, J., Jura, M., et al. 2005, ApJ, 632, L119

Beichman, C., Benneke, B., Knutson, H. A., et al. 2014, Publications of the Astronomical Society of Pacific, 126, 1134

Ben-Jaffel, L., & Sona Hosseini, S. 2010, The Astrophysical Journal, 709, 1284

Bender, P. L., & Stebbins, R. T. 1996, Journal of Geophysical Research, 101, 9309

Benneke, B., & Seager, S. 2012, Nature, 753, 100

—. 2013, The Astrophysical Journal, 778, 153

Bergfors, C., Farihi, J., Dufour, P., & Rocchetto, M. 2014, Monthly Noticese of the Royal Astronomical Society, 444, 2147

Berta, Z. K., Charbonneau, D., Desert, J.-M., et al. 2012, Nature, 747, 35

Beuzit, J.-L., Feldt, M., Dohlen, K., et al. 2008, SPIE Proceedings, 7014, 701418

Bideau-Mehu, A., Guern, Y., Abjean, R., & Johannin-Gilles, A. 1973, Optics Communications, 9, 432

Birnbaum, G. 1979, Journal of Quantitative Spectroscopy & Radiative Transfer, 21, 597

Bochkarev, K. V., & Rafikov, R. R. 2011, ApJ, 741, 36

Bodenheimer, P., Lin, D. N. C., & Mardling, R. A. 2001, The Astrophysical Journal, 548, 466

Bodhaine, B. A., Wood, N. B., Dutton, E. G., & Slusser, J. R. 1999, Journal of Atmospheric and Oceanic Technology, 16, 1854

Boley, A. C., Hayfield, T., Mayer, L., & Durisen, R. H. 2010, Icarus, 207, 509

Bolstad, W. M. 2007, Introduction to Bayesian Statistics (Wiley-Interscience)

Bond, I. A., Abe, F., Dodd, R. J., et al. 2001, Monthly Notices of the Royal Astronomical Society, 327, 868

Bond, I. A., Udalski, A., Jaroszyński, M., et al. 2004, The Astrophysical Journal, 606, L155

Bonfils, X., Delfosse, X., Udry, S., et al. 2013, Astronomy & Astrophysics, 549, A109

Bonsor, A., Mustill, A. J., & Wyatt, M. C. 2011, Monthly Noticese of the Royal Astronomical Society, 414, 930

Borucki, W. J., Koch, D., Basri, G., et al. 2010, Science, 327, 977

Borucki, W. J., Koch, D. G., Basri, G., et al. 2011, The Astrophysical Journal, 736, 19

Borysow, A. 2002, Astronomy & Astrophysics, 390, 779

Borysow, A., Jorgensen, U. G., & Fu, Y. 2001, Journal of Quantitative Spectroscopy & Radiative Transfer, 68, 235

Boss, A. P. 1997, Science, 276, 1836

Bouanich, J. P., Aroui, H., Nouri, S., & Picard-Bersellini, A. 2001, Journal of Molecular Spectroscopy, 206, 104

Bouchy, F., Udry, S., Mayor, M., et al. 2005, Astronomy & Astrophysics, 444, L15

Bounaceur, R., Glaude, P. A., Fournet, R., et al. 2007, International Journal of Vehicle Design

Bridges, M., Feroz, F., Hobson, M. P., & Lasenby, A. N. 2009, Monthly Notices of the Royal Astronomical Society, 400, 1075

Brinkworth, C. S., Gänsicke, B. T., Girven, J. M., et al. 2012, ApJ, 750, 86

Brinkworth, C. S., Gänsicke, B. T., Marsh, T. R., Hoard, D. W., & Tappert, C. 2009, ApJ, 696, 1402

Brogi, M., de Kok, R. J., Birkby, J. L., Schwarz, H., & Snellen, I. A. G. 2014, Astronomy & Astrophysics, 565, A124

Brogi, M., Snellen, I. A. G., de Kok, R. J., et al. 2012, Nature, 486, 502

—. 2013, The Astrophysical Journal, 767, 27

Brooks, S., Gelman, A., Jones, G., & Meng, X.-L. 2011, Handbook of Markov Chain Monte Carlo (CRC Press)

Brown, T. M. 2001, Nature, 553, 1006

Burrows, A., Budaj, J., & Hubeny, I. 2008, The Astrophysical Journal, 678, 1436

Burrows, A., Hubeny, I., Budaj, J., Knutson, H. A., & Charbonneau, D. 2007, The Astrophysical Journal, 668, L171

Burrows, A., Hubeny, I., & Sudarsky, D. 2005, The Astrophysical Journal, 625, L135

Butler, R. P., Marcy, G. W., Williams, E., Hauser, H., & Shirts, P. 1997, The Astrophysical Journal, 474, L115

Butler, R. P., Vogt, S. S., Marcy, G. W., et al. 2004, The Astrophysical Journal, 617, 580

Byrd, R. H., Lu, P., Nocedal, J., & Zhu, C. 1995, SIAM Journal on Scientific Computing, 16, 1190

Campbell, B., Walker, G. A. H., & Yang, S. 1988, Astrophysical Journal, 331, 902

Cassan, A., Kubas, D., Beaulieu, J. P., et al. 2012, Nature, 481, 167

Cavalié, T., Moreno, R., Lellouch, E., et al. 2014, Astronomy & Astrophysics, 562, A33

Chahine, M. T. 1968, Journal of the Optical Society of America, 58, 1634

Chandrasekar, S. 1960, Radiative Transfer Dover Publications (New York)

Charbonneau, D., Brown, T. M., Latham, D. W., & Mayor, M. 2000, The Astrophysical . . . , 529, L45

Charbonneau, D., Brown, T. M., Noyes, R. W., & Gilliland, R. L. 2002, Nature, 568, 377

Chayes, F. 1971, Ratio Correlation, A Manual for Students of Petrology and Geochemistry (University of Chicago Press)

Chiang, E. I., & Goldreich, P. 1997, ApJ, 490, 368

Claret, A. 2000, Astronomy & Astrophysics, 363, 1081

Clark, R. N., & McCord, T. B. 1980a, Icarus, 41, 323

—. 1980b, Icarus, 43, 161

Cochran, W. D., & Hatzes, A. P. 1994, Astrophysics and Space Science, 212, 281

Combes, M., & Encrenaz, T. 1979, Icarus, 39, 1

Conrath, B. J. 1972, Journal of Atmospheric Sciences, 29, 1262

Conrath, B. J., Hanel, R. A., Kunde, V. G., & Prabhakara, C. 1970, Journal of Geophysical Research, 75, 5831

Cooper, C. S., & Showman, A. P. 2005, The Astrophysical Journal, 629, L45

Cowan, N. B., Greene, T., Angerhausen, D., et al. 2015, Publications of the Astronomical Society of Pacific, 127, 311

Cox, A. N. 2015, Allen's Astrophysical Quantities (Springer)

Crossfield, I. J. 2012, Astronomy & Astrophysics, 545, A97

Cutri, R. M., Skrutskie, M. F., Van Dyk, S., et al. 2003, VizieR On-line Data Catalog, 2246

Davenport, J. R. A., Becker, A. C., Kowalski, A. F., et al. 2012, The Astrophysical Journal, 748, 58

de Bergh, C., Lutz, B. L., Owen, T., & Maillard, J. P. 1990, Astrophysical Journal, 355, 661

de Kok, R. J., Brogi, M., Snellen, I. A. G., et al. 2013, Astronomy & Astrophysics, 554, A82

de Kok, R. J., & Stam, D. M. 2012, Icarus, 221, 517

de Kok, R. J., Stam, D. M., & Karalidi, T. 2011, Nature, 741, 59

De Rooij, W. A., & van der Stap, C. C. A. H. 1984, Astronomy and Astrophysics (ISSN 0004-6361), 131, 237

de Wit, J., & Seager, S. 2013, Science, 342, 1473

Debes, J. H., Hoard, D. W., Kilic, M., et al. 2011a, ApJ, 729, 4

Debes, J. H., Hoard, D. W., Wachter, S., Leisawitz, D. T., & Cohen, M. 2011b, The Astrophysical Journal Supplement Series, 197, 38

Debes, J. H., Kilic, M., Faedi, F., et al. 2012a, The Astrophysical Journal, 754, 59

Debes, J. H., & Sigurdsson, S. 2002, ApJ, 572, 556

Debes, J. H., Sigurdsson, S., & Hansen, B. 2007, The Astronomical Journal, 134, 1662

Debes, J. H., Walsh, K. J., & Stark, C. 2012b, ApJ, 747, 148

Deming, D., Seager, S., Richardson, L. J., & Harrington, J. 2005, Nature, 434, 740

Deming, D., Seager, S., Winn, J., et al. 2009, Publications of the Astronomical Society of Pacific, 121, 952

Deming, D., Knutson, H. A., Agol, E., et al. 2011, The Astrophysical Journal, 726, 95

Deming, D., Wilkins, A., McCullough, P., et al. 2013, The Astrophysical Journal, 774, 95

Demory, B.-O., Gillon, M., Madhusudhan, N., & Queloz, D. 2016a, Monthly Notices of the Royal Astronomical Society, 455, 2018

Demory, B.-O., Gillon, M., de Wit, J., et al. 2016b, Nature, 532, 207

Desert, J.-M., Lecavelier des Etangs, A., Hébrard, G., et al. 2009, The Astrophysical Journal, 699, 478

Desharnais, S., Wesemael, F., Chayer, P., Kruk, J. W., & Saffer, R. A. 2008, ApJ, 672, 540

Dhib, M., Bouanich, J. P., Aroui, H., & Broquier, M. 2000, Journal of Molecular Spectroscopy, 202, 83

Dominik, M., Jørgensen, U. G., Rattenbury, N. J., et al. 2010, Astronomische Nachrichten, 331, 671

Dong, S., & Zhu, Z. 2013, The Astrophysical Journal, 778, 53

Doyon, R., Hutchings, J. B., Beaulieu, M., et al. 2012, in SPIE Astronomical Telescopes + Instrumentation, ed. M. C. Clampin, G. G. Fazio, H. A. MacEwen, & J. M. Oschmann (SPIE), 84422R–13

Dragomir, D., Matthews, J. M., Winn, J. N., & Rowe, J. F. 2014, Formation, 293, 52

Dressing, C. D., & Charbonneau, D. 2013, The Astrophysical Journal, 767, 95

Dufour, P., Bergeron, P., Liebert, J., et al. 2007, ApJ, 663, 1291

Dupuis, J., Fontaine, G., Pelletier, C., & Wesemael, F. 1992, The Astrophysical Journal Supplement Series, 82, 505

—. 1993a, The Astrophysical Journal Supplement Series, 84, 73

Dupuis, J., Fontaine, G., & Wesemael, F. 1993b, The Astrophysical Journal Supplement Series, 87, 345

Dzigan, Y., & Zucker, S. 2012, The Astrophysical Journal Letters, 753, L1

Edwards, D. P., & Strow, L. L. 1991, Journal of Geophysical Research (ISSN 0148-0227), 96, 20

Ehrenreich, D., Bonfils, X., Lovis, C., et al. 2014, Astronomy & Astrophysics, 570, A89

Epchtein, N., Deul, E., Derriere, S., et al. 1999, Astronomy & Astrophysics, 349, 236

Farihi, J. 2009, Monthly Noticese of the Royal Astronomical Society, 398, 2091

—. 2011, White Dwarf Atmospheres and Circumstellar Environments, 117

Farihi, J., Barstow, M. A., Redfield, S., Dufour, P., & Hambly, N. C. 2010a, Monthly Noticese of the Royal Astronomical Society, 404, 2123

Farihi, J., Becklin, E. E., & Zuckerman, B. 2008a, ApJ, 681, 1470

Farihi, J., Gänsicke, B. T., & Koester, D. 2013, Science, 342, 218

Farihi, J., Gänsicke, B. T., Steele, P. R., et al. 2012, Monthly Noticese of the Royal Astronomical Society, 421, 1635

Farihi, J., Jura, M., & Zuckerman, B. 2009, ApJ, 694, 805

—. 2010b, ApJ, 714, 1386

Farihi, J., Zuckerman, B., & Becklin, E. E. 2008b, ApJ, 674, 431

Faure, A., Wiesenfeld, L., Drouin, B. J., & Tennyson, J. 2013, Journal of Quantitative Spectroscopy & Radiative Transfer, 116, 79

Fazio, G. G., Hora, J. L., Allen, L. E., et al. 2004, The Astrophysical Journal Supplement Series, 154, 10

Feller, W. 1971, An introduction to probability theory and its applications (John Wiley & Sons Inc)

Feng, Y. K., Line, M. R., Fortney, J., et al. 2016, arXiv.org, arXiv:1607.03230

Feroz, F., Balan, S. T., & Hobson, M. P. 2011a, Monthly Notices of the Royal Astronomical Society: Letters, 416, L104

—. 2011b, Monthly Notices of the Royal Astronomical Society, 415, 3462

Feroz, F., Gair, J. R., Graff, P., Hobson, M. P., & Lasenby, A. 2010a, Classical and Quantum Gravity, 27, 075010

Feroz, F., Gair, J. R., Hobson, M. P., & Porter, E. K. 2009a, Classical and Quantum Gravity, 26, 215003

Feroz, F., & Hobson, M. P. 2008, Monthly Noticese of the Royal Astronomical Society, 384, 449

Feroz, F., Hobson, M. P., & Bridges, M. 2009b, Monthly Notices of the Royal Astronomical Society, 398, 1601

Feroz, F., Hobson, M. P., & Trotta, R. 2010b, arXiv.org, arXiv:1001.0719

Ferraz-Mello, S., Tadeu Dos Santos, M., Beaugé, C., Michtchenko, T. A., & Rodríguez, A. 2011, Astronomy & Astrophysics, 531, A161

Ferruit, P., Birkmann, S., Böker, T., et al. 2014, in Proceedings of the SPIE, ed. J. M. Oschmann, M. Clampin, G. G. Fazio, & H. A. MacEwen, European Space Research and Technology Ctr. (Netherlands) (SPIE), 91430A

Fischer, D. A., Marcy, G. W., Butler, R. P., et al. 2008, The Astrophysical Journal, 675, 790

Fontaine, G., Brassard, P., & Bergeron, P. 2001, The Publications of the Astronomical Society of the Pacific, 113, 409

Fontaine, G., & Michaud, G. 1979, ApJ, 231, 826

Forget, F., & Leconte, J. 2014, Philosophical Transactions of the Royal Society A: Mathematical, 372, 20130084

Fortier, A., Alibert, Y., Carron, F., Benz, W., & Dittkrist, K. M. 2013, Astronomy & Astrophysics, 549, A44

Fortier, A., Benvenuto, O. G., & Brunini, A. 2007, Astronomy & Astrophysics, 473, 311

Fortney, J., Lodders, K., Marley, M. S., & Freedman, R. S. 2008, The Astrophysical Journal, 678, 1419

Fortney, J., Marley, M. S., Lodders, K., Saumon, D., & Freedman, R. 2005, The Astrophysical Journal, 627, L69

Fortney, J., Mordasini, C., Nettelmann, N., et al. 2013, The Astrophysical Journal, 775, 80

Fortney, J., Shabram, M., Showman, A. P., et al. 2010, The Astrophysical Journal, 709, 1396

Fox, K., Jennings, D. E., Stern, E. A., & Hubbard, R. 1988, Journal of Quantitative Spectroscopy and Radiative Transfer (ISSN 0022-4073), 39, 473

Fraine, J., Deming, D., Benneke, B., et al. 2014, Nature, 513, 526

Freedman, R. S., Marley, M. S., & Lodders, K. 2008, The Astrophysical Journal Supplement Series, 174, 504

Fressin, F., Torres, G., Charbonneau, D., et al. 2013, The Astrophysical Journal, 766, 81

Frewen, S. F. N., & Hansen, B. M. S. 2014, Monthly Noticese of the Royal Astronomical Society, 439, 2442

Friel, N., & Wyse, J. 2012, Statistica Neerlandica, 66, 288

Fukui, A., Narita, N., Kurosaki, K., et al. 2013, The Astrophysical Journal, 770, 95

Gabard, T., Grigoriev, I. M., Grigorovich, N. M., & Tonkov, M. V. 2004, Journal of Molecular Spectroscopy, 225, 123

Gamerman, D., & Lopes, H. F. 2006, Markov Chain Monte Carlo, Stochastic Simulation for Bayesian Inference, Second Edition (CRC Press)

Gänsicke, B. T., Koester, D., Farihi, J., Girven, J. M., & Parsons, S. G. 2012, Monthly Noticese of the Royal Astronomical Society, 424, 333

Gänsicke, B. T., Marsh, T. R., Southworth, J., & Rebassa-Mansergas, A. 2006, Science, 314, 1908

Gaudi, B. S. 2008, Extreme Solar Systems, 398, 479

Gelfand, A. E., & Smith, A. F. M. 1990, Journal of the American Statistical Association, 85, 398

Gelman, A., Carlin, J. B., Stern, H. S., et al. 2013, Bayesian Data Analysis, Third Edition (CRC Press)

Geman, S., & Geman, D. 1984, I,EEE Trans. Pattn. Anal. Mach. Intell., 6, 721

Gilliland, R. L., Chaplin, W. J., Dunham, E. W., et al. 2011, The Astrophysical Journal Supplement, 197, 6

Girven, J. M., Brinkworth, C. S., Farihi, J., et al. 2012, ApJ, 749, 154

Girven, J. M., Gänsicke, B. T., Steeghs, D., & Koester, D. 2011, Monthly Noticese of the Royal Astronomical Society, 417, 1210

Goody, R. M., & Yung, Y. L. 1995, Atmospheric Radiation, Theoretical Basis (Oxford University Press)

Goorvitch, D. 1994, The Astrophysical Journal Supplement Series, 95, 535

Gordon, S., & Mcbride, B. J. 1994, NASA Reference Publication

Gould, A. 2005, New Astronomy Reviews, 49, 424

Gould, A., & Loeb, A. 1992, Astrophysical Journal, 396, 104

Graff, P., Feroz, F., Hobson, M. P., & Lasenby, A. 2012, Monthly Notices of the Royal Astronomical Society, 421, 169

Grasset, O., Schneider, J., & Sotin, C. 2009, The Astrophysical Journal, 693, 722

Greene, T., Beichman, C., Eisenstein, D., et al. 2007, Techniques and Instrumentation for Detection of Exoplanets III. Edited by Coulter, 6693, 66930G

Greene, T. P., Line, M. R., Montero, C., et al. 2016, The Astrophysical Journal, 817, 17

Greenfield, S. M., & Kellogg, W. W. 1960, Journal of Meteorology, 17, 283

Grigoriev, I. M., Filippov, N. N., Tonkov, M. V., Gabard, T., & Le Doucen, R. 2001, Journal of Quantitative Spectroscopy & Radiative Transfer, 69, 189

Grimm, S. L., & Heng, K. 2015, The Astrophysical Journal, 808, 182

Gueymard, C. A. 2004, Solar Energy, 76, 423

Guilera, O. M., Fortier, A., Brunini, A., & Benvenuto, O. G. 2011, Astronomy & Astrophysics, 532, A142

Guillot, T. 2010, Astronomy & Astrophysics, 520, A27

Guillot, T., Santos, N. C., Pont, F., et al. 2006, Astronomy & Astrophysics, 453, L21

Haario, H., Laine, M., Mira, A., & Saksman, E. 2006, Statistics and Computing, 16, 339

Haario, H., Saksman, E., & Tamminen, J. 2001, Bernoulli, 7, 223

Hadded, S., Aroui, H., Orphal, J., Bouanich, J. P., & Hartmann, J. M. 2001, Journal of Molecular Spectroscopy, 210, 275

Han, C. 2007, The Astrophysical Journal, 661, 1202

Han, C., & Kim, Y.-G. 2001, The Astrophysical Journal, 546, 975

Hanel, R. A., & Conrath, B. J. 1970, Nature, 228, 143

Hanel, R. A., Conrath, B. J., Jennings, D. E., & Samuelson, R. E. 2003, Exploration of the Solar System by Infrared Remote Sensing, 534

Hansen, B. M. S. 2008, The Astrophysical Journal Supplement Series, 179, 484

Harris, G. J., Tennyson, J., Kaminsky, B. M., Pavlenko, Y. V., & Jones, H. R. A. 2006, Monthly Notices of the Royal Astronomical Society, 367, 400

Hastings, W. K. 1970, Biometrika, 57, 97

Haswell, C. A. 2010, Transiting Exoplanets by Carole A. Haswell. Cambridge University Press

Hatzes, A. P. 2014, Nature, 513, 353

Hawley, S. L., & Pettersen, B. R. 1991, Astrophysical Journal, 378, 725

Hedges, C., & Madhusudhan, N. 2016, Monthly Notices of the Royal Astronomical Society, 458, 1427

Hekker, S., & Aerts, C. 2010, Astronomy & Astrophysics, 515, A43

Helled, R., & Bodenheimer, P. 2011, Icarus, 211, 939

Henden, A. A., Welch, D. L., Terrell, D., & Levine, S. E. 2009, 214

Heng, K., Mendonça, J. M., & Lee, J.-M. 2014, The Astrophysical Journal Supplement, 215, 4

Heng, K., & Workman, J. 2014, The Astrophysical Journal Supplement, 213, 27

Henry, G. W., Marcy, G. W., Butler, R. P., & Vogt, S. S. 2000, The Astrophysical Journal, 529, L41

Hill, C., Yurchenko, S. N., & Tennyson, J. 2013, Icarus, 226, 1673

Holberg, J. B., & Bergeron, P. 2006, The Astronomical Journal, 132, 1221

Hollis, M. D. J., Tessenyi, M., & Tinetti, G. 2013, Computer Physics Communications, 184, 2351

Holmström, M., Ekenbäck, A., Selsis, F., et al. 2008, Nature, 451, 970

Hori, Y., & Ikoma, M. 2011, Monthly Notices of the Royal Astronomical Society, 416, 1419

Howard, A. W., Marcy, G. W., Bryson, S. T., et al. 2012, The Astrophysical Journal Supplement, 201, 15

Howell, S. B., Sobeck, C., Haas, M., et al. 2014, Publications of the Astronomical Society of Pacific, 126, 398

Hu, R., & Seager, S. 2014, The Astrophysical Journal, 784, 63

Hubeny, I., Burrows, A., & Sudarsky, D. 2003, The Astrophysical Journal, 594, 1011

Hubickyj, O., Bodenheimer, P., & Lissauer, J. J. 2005, Icarus, 179, 415

Humlícek, J. 1982, Journal of Quantitative Spectroscopy & Radiative Transfer, 27, 437

Husser, T. O., Wende-von Berg, S., Dreizler, S., et al. 2013, Astronomy & Astrophysics, 553, 6

Ida, S., & Lin, D. N. C. 2004, The Astrophysical Journal, 604, 388

Irwin, P. G. J., Teanby, N. A., De Kok, R., et al. 2008, Journal of Quantitative Spectroscopy & Radiative Transfer, 109, 1136

Iyer, A. R., Swain, M., Zellem, R. T., et al. 2016, The Astrophysical Journal, 823, 109

Jakosky, B. M., & Haberle, R. M. 1992, In: Mars (A93-27852 09-91), 969

Jeffreys, H. 1998, The Theory of Probability (OUP Oxford)

Jeffreys, H. 1961

Jorissen, A., Mayor, M., & Udry, S. 2001, Astronomy & Astrophysics, 379, 992

Jura, M. 2003, ApJ, 584, L91

—. 2008, The Astronomical Journal, 135, 1785

Jura, M., Farihi, J., & Zuckerman, B. 2007, ApJ, 663, 1285

—. 2009, The Astronomical Journal, 137, 3191

Kahn, H., & Harris, T. E. 1951, Natl. Bur. Stand. Appl. Math. Ser. 12, 27

Kalirai, J. S., Hansen, B. M. S., Kelson, D. D., et al. 2008, ApJ, 676, 594

Kane, S. R. 2007, Monthly Notices of the Royal Astronomical Society, 380, 1488

Kane, S. R., Mahadevan, S., von Braun, K., Laughlin, G., & Ciardi, D. R. 2009, Publications of the Astronomical Society of Pacific, 121, 1386

Kaplan, L. D. 1959, Journal of the Optical Society of America, 49, 1004

Karpenka, N. V., March, M. C., Feroz, F., & Hobson, M. P. 2013, Monthly Notices of the Royal Astronomical Society, 433, 2693

Kendrew, S., Scheithauer, S., Bouchet, P., et al. 2015, Publications of the Astronomical Society of Pacific, 127, 623

King, J. 1956, Scientific uses of earth satellites, University of Michigan Press

Kipping, D. M. 2008, Monthly Notices of the Royal Astronomical Society, 389, 1383

Kislyakova, K. G., Holmström, M., Lammer, H., Odert, P., & Khodachenko, M. L. 2014, Science, 346, 981

Klein, B., Jura, M., Koester, D., Zuckerman, B., & Melis, C. 2010, ApJ, 709, 950

Knutson, H. A., Benneke, B., Deming, D., & Homeier, D. 2014a, Nature, 505, 66

Knutson, H. A., Charbonneau, D., Noyes, R. W., Brown, T. M., & Gilliland, R. L. 2007, The Astrophysical Journal, 655, 564

Knutson, H. A., Madhusudhan, N., Cowan, N. B., et al. 2011, The Astrophysical Journal, 735, 27

Knutson, H. A., Dragomir, D., Kreidberg, L., et al. 2014b, The Astrophysical Journal, 794, 155

Koch, D. G., Borucki, W. J., Basri, G., et al. 2010, The Astrophysical Journal Letters, 713, L79

Koester, D., Gänsicke, B. T., & Farihi, J. 2014, Astronomy & Astrophysics, 566, 34

Koester, D., Rollenhagen, K., Napiwotzki, R., et al. 2005, Astronomy & Astrophysics, 432, 1025

Koester, D., Voss, B., Napiwotzki, R., et al. 2009, Astronomy & Astrophysics, 505, 441

Kopal, Z. 1950, Harvard College Observatory Circular

Koskinen, T. T., Aylward, A. D., Smith, C. G. A., & Miller, S. 2007, The Astrophysical Journal, 661, 515

Kreidberg, L., Bean, J. L., Desert, J.-M., et al. 2014, Nature, 505, 69

Kuntz, M. 1997, Journal of Quantitative Spectroscopy & Radiative Transfer, 57, 819

Lacis, A. A., & Oinas, V. 1991, Journal of Geophysical Research, 96, 9027

Lacy, C. H., Moffett, T. J., & Evans, D. S. 1976, Astrophysical Journal Supplement Series, 30, 85

Larson, H. P. 1980, In: Annual review of astronomy and astrophysics. Volume 18. (A81-20334 07-90) Palo Alto, 18, 43

Latham, D. W., Stefanik, R. P., Mazeh, T., Mayor, M., & Burki, G. 1989, Nature (ISSN 0028-0836), 339, 38

Lawrence, A., Warren, S. J., Almaini, O., et al. 2007, Monthly Noticese of the Royal Astronomical Society, 379, 1599

Lee, J. M., Fletcher, L. N., & Irwin, P. G. J. 2011, Monthly Noticese of the Royal Astronomical Society, 420, 170

Léger, A., Selsis, F., Sotin, C., et al. 2004, Icarus, 169, 499

Lellouch, E., Encrenaz, T., de Graauw, T., et al. 2000, Planetary and Space Science, 48, 1393

Lewis, N. K., Showman, A. P., Fortney, J., et al. 2010, The Astrophysical Journal, 720, 344

Liddle, A. R. 2004, Monthly Notices of the Royal Astronomical Society, 351, L49

Liebert, J., Bergeron, P., & Holberg, J. B. 2005, The Astrophysical Journal Supplement Series, 156, 47

Line, M. R., Fortney, J., Marley, M. S., & Sorahana, S. 2014, The Astrophysical Journal, 793, 33

Line, M. R., & Parmentier, V. 2016, The Astrophysical Journal, 820, 78

Line, M. R., Teske, J., Burningham, B., Fortney, J., & Marley, M. S. 2015, The Astrophysical Journal, 807, 183

Line, M. R., Vasisht, G., Chen, P., Angerhausen, D., & Yung, Y. L. 2011, The Astrophysical Journal, 738, 32

Line, M. R., Zhang, X., Vasisht, G., et al. 2012, The Astrophysical Journal, 749, 93

Line, M. R., Wolf, A. S., Zhang, X., et al. 2013, The Astrophysical Journal, 775, 137

Linsky, J. L., Yang, H., France, K., et al. 2010, The Astrophysical Journal, 717, 1291

Liou, K N. 2002

Lissauer, J. J., Dawson, R. I., & Tremaine, S. 2014, Nature, 513, 336

Lodders, K. 2003, ApJ, 591, 1220

—. 2004, The Astrophysical Journal, 611, 587

Macintosh, B. A., Graham, J. R., Palmer, D. W., et al. 2008, SPIE Proceedings, 7015, 701518

Macintosh, B. A., Graham, J. R., Barman, T. S., et al. 2015, Science, 350, 64

Madhusudhan, N. 2012, The Astrophysical Journal, 758, 36

Madhusudhan, N., Burrows, A., & Currie, T. 2011a, The Astrophysical Journal, 737, 34

Madhusudhan, N., & Seager, S. 2009, The Astrophysical Journal, 707, 24

Madhusudhan, N., Harrington, J., Stevenson, K. B., et al. 2011b, Nature, 469, 64

Makovoz, D., Roby, T., Khan, I., & Booth, H. 2006, Advanced Software and Control for Astronomy. Edited by Lewis, 6274, 10

Mandel, K., & Agol, E. 2002, Nature, 580, L171

Mansfield, C. R., & Peck, E. R. 1969, Journal of the Optical Society of America, 59, 199

Mantz, A. W., Malathy Devi, V., Chris Benner, D., et al. 2005, Journal of Molecular Structure, 742, 99

Marcy, G. W., & Butler, R. P. 1992, Astronomical Society of the Pacific, 104, 270

Marcy, G. W., Butler, R. P., Fischer, D. A., et al. 2002, The Astrophysical Journal, 581, 1375

Margolis, J. S. 1993, Journal of Quantitative Spectroscopy & Radiative Transfer, 50, 431

Markov, A. 1954, Tr. Mat. Inst. Steklov, 42, 1

Marley, M. S., & Sengupta, S. 2011, Monthly Notices of the Royal Astronomical Society, 417, 2874

Mayer, L., Wadsley, J., Quinn, T., & Stadel, J. 2005, Monthly Notices of the Royal Astronomical Society, 363, 641

Mayor, M., Lovis, C., & Santos, N. C. 2014, Nature, 513, 328

Mayor, M., & Queloz, D. 1995, Nature, 378, 355

Mayor, M., Pepe, F., Queloz, D., et al. 2003, The Messenger (ISSN0722-6691), 114, 20

Mayor, M., Marmier, M., Lovis, C., et al. 2011, arXiv.org, arXiv:1109.2497

Mazeh, T., Naef, D., Torres, G., et al. 2000, The Astrophysical Journal, 532, L55

McArthur, B. E., Benedict, G. F., Barnes, R., et al. 2010, The Astrophysical Journal, 715, 1203

McArthur, B. E., Endl, M., Cochran, W. D., et al. 2004, The Astrophysical Journal, 614, L81

McCartney, E. J. 1976, Wiley Series in Pure and Applied Optics

McCullough, P. R., Stys, J. E., Valenti, J. A., et al. 2005, The Publications of the Astronomical Society of the Pacific, 117, 783

McKean, H. P. 1967, Propagation of chaos for a class of non-linear parabolic equations, Vol. 7 (Lecture Series in Differential Equations)

McMillan, R. S., Moore, T. L., Perry, M. L., & Smith, P. H. 1994, Astrophysics and Space Science, 212, 271

Mégevand, D., Herreros, J.-M., Zerbi, F., et al. 2010, in Proceedings of the SPIE, ed. I. S. McLean, S. K. Ramsay, & H. Takami, Observatoire de l'Univ. de Genève, Switzerland (SPIE), 77354Y–77354Y–10

Melis, C., Farihi, J., Dufour, P., et al. 2011, ApJ, 732, 90

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. 1953, The Journal of Chemical Physics, 21, 1087

Metzger, B. D., Rafikov, R. R., & Bochkarev, K. V. 2012, Monthly Notices of the Royal Astronomical Society, 423, 505

Miller-Ricci Kempton, E., & Rauscher, E. 2012, Nature, 751, 117

Mira, A. 2001, Metron

Mordasini, C., Alibert, Y., Benz, W., Klahr, H., & Henning, T. 2012, Astronomy & Astrophysics, 541, A97

Mordasini, C., Alibert, Y., Benz, W., & Naef, D. 2009, Astronomy & Astrophysics, 501, 1161

Morley, C. V., Fortney, J., Kempton, E. M. R., et al. 2013, The Astrophysical Journal, 775, 33

Moses, J. I. 2013, arXiv.org, 5450

Moses, J. I., Madhusudhan, N., Visscher, C., & Freedman, R. S. 2013a, ApJ, 763, 25

Moses, J. I., Visscher, C., Fortney, J., et al. 2011, Nature, 737, 15

Moses, J. I., Line, M. R., Visscher, C., et al. 2013b, ApJ, 777, 34

Mousis, O., Lunine, J. I., Madhusudhan, N., & Johnson, T. V. 2012, The Astrophysical Journal Letters, 751, L7

Mousis, O., Fletcher, L. N., Lebreton, J. P., et al. 2014, Planetary and Space Science, 104, 29

Mukherjee, P., Parkinson, D., & Liddle, A. R. 2006, The Astrophysical Journal, 638, L51

Mullally, F., Kilic, M., Reach, W. T., et al. 2007, The Astrophysical Journal Supplement Series, 171, 206

Mustill, A. J., Veras, D., & Villaver, E. 2014, Monthly Noticese of the Royal Astronomical Society, 437, 1404

Nelson, B. E., Ford, E. B., Wright, J. T., et al. 2014, Monthly Notices of the Royal Astronomical Society, 441, 442

Nouri, S., Orphal, J., Aroui, H., & Hartmann, J. M. 2004, Journal of Molecular Spectroscopy, 227, 60

Nutzman, P., & Charbonneau, D. 2008, Publications of the Astronomical Society of Pacific, 120, 317

Öberg, K. I., Murray-Clay, R., & Bergin, E. A. 2011, The Astrophysical Journal Letters, 743, L16

O'Donovan, F. T., Charbonneau, D., Torres, G., et al. 2006, The Astrophysical Journal, 644, 1237

Olivero, J. J., & Longbothum, R. L. 1977, Journal of Quantitative Spectroscopy & Radiative Transfer, 17, 233

Paczynski, B. 1986, Astrophysical Journal, 301, 503

Paquette, C., Pelletier, C., Fontaine, G., & Michaud, G. 1986, The Astrophysical Journal Supplement Series, 61, 197

Parmentier, V., & Guillot, T. 2014, Astronomy & Astrophysics, 562, A133

Parmentier, V., Guillot, T., Fortney, J., & Marley, M. S. 2015, Astronomy & Astrophysics, 574, A35

Parmentier, V., Showman, A. P., & Lian, Y. 2013, Astronomy & Astrophysics, 558, A91

Partridge, H., & Schwenke, D. W. 1997, The Journal of Chemical Physics, 106, 4618

Pawlowsky-Glahn, V., & Egozcue, J. J. 2006, Geological Society, 264, 1

Peck, E. R., & Khanna, B. N. 1966, JOSA, 56, 1059

Peck, E. R., & Reeder, K. 1972, Journal of the Optical Society of America, 62, 958

Pepe, F. A., Cristiani, S., Rebolo Lopez, R., et al. 2010, in Proceedings of the SPIE, ed. I. S. McLean, S. K. Ramsay, & H. Takami, Observatoire Astronomique, Univ. de Genève, Switzerland (SPIE), 77350F–77350F–9

Perryman, M. 2014, The Exoplanet Handbook

Perryman, M., Hartman, J., Bakos, G. Á., & Lindegren, L. 2014, The Astrophysical Journal, 797, 14

Petigura, E. A., Howard, A. W., & Marcy, G. W. 2013, in Proceedings of the National Academy of Sciences of the United States of America, Astronomy Department, University of California, Berkeley, CA 94720. (National Acad Sciences), 19273–19278

Petrova, T. M., Solodov, A. M., Solodov, A. A., & Starikov, V. I. 2013, Journal of Quantitative Spectroscopy & Radiative Transfer, 129, 241

—. 2016, Journal of Molecular Spectroscopy, 321, 50

Pettersen, B. R., Coleman, L. A., & Evans, D. S. 1984, The Astrophysical Journal Supplement Series, 54, 375

Pierrehumbert, R. T. 2010, Principles of Planetary Climate

Pine, A. S. 1992, Journal of Chemical Physics (ISSN 0021-9606), 97, 773

Pine, A. S., Markov, V. N., Buffa, G., & Tarrini, O. 1993, Journal of Quantitative Spectroscopy & Radiative Transfer, 50, 337

Poddaný, S., Brát, L., & Pejcha, O. 2011, Detection and Dynamics of Transiting Exoplanets, 11, 06008

Pollacco, D. L., Skillen, I., Collier Cameron, A., et al. 2006, The Publications of the Astronomical Society of the Pacific, 118, 1407

Poor, C. L. 1930, Journal of the Optical Society of America, 20, 173

Queloz, D., Bouchy, F., Moutou, C., et al. 2009, Astronomy & Astrophysics, 506, 303

Rafikov, R. R. 2011, Monthly Noticese of the Royal Astronomical Society, 416, L55

—. 2012, ApJ, 760, 123

Rafikov, R. R., & De Colle, F. 2006, ApJ, 646, 275

Rauer, H., Catala, C., Aerts, C., et al. 2014, Experimental Astronomy, 38, 249

Reach, W. T., Kuchner, M. J., von Hippel, T., et al. 2005, ApJ, 635, L161

Reach, W. T., Lisse, C., von Hippel, T., & Mullally, F. 2009, ApJ, 693, 697

Reffert, S., & Quirrenbach, A. 2011, Astronomy & Astrophysics, 527, A140

Richard, C., Gordon, I. E., Rothman, L. S., et al. 2012, Journal of Quantitative Spectroscopy & Radiative Transfer, 113, 1276

Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2014, in Proceedings of the SPIE, ed. J. M. Oschmann, M. Clampin, G. G. Fazio, & H. A. MacEwen, Massachusetts Institute of Technology (United States) (SPIE), 914320–914320–15

Robinson, T. D., & Catling, D. C. 2012, The Astrophysical Journal, 757, 104

Rocchetto, M., Farihi, J., Gänsicke, B. T., & Bergfors, C. 2015, Monthly Notices of the Royal Astronomical Society, 449, 574

Rocchetto, M., Waldmann, I., Venot, O., Lagage, P. O., & Tinetti, G. 2016, The Astrophysical Journal, 833, 120

Rodgers, C. D. 1976, Reviews of Geophysics and Space Physics, 14, 609

Rodler, F., Lopez-Morales, M., & Ribas, I. 2012, The Astrophysical Journal Letters, 753, L25

Rojas-Ayala, B., Covey, K. R., Muirhead, P. S., & Lloyd, J. P. 2012, The Astrophysical Journal, 748, 93

Rothman, L. S., Gordon, I. E., Barber, R. J., et al. 2010, Journal of Quantitative Spectroscopy & Radiative Transfer, 111, 2139

Rothman, L. S., Gordon, I. E., Babikov, Y., et al. 2013, Journal of Quantitative Spectroscopy & Radiative Transfer, 130, 4

Rouan, D., Deeg, H. J., Demangeon, O., et al. 2011, The Astrophysical Journal Letters, 741, L30

Sabelfeld, K. K. 2004

Sackett, P. D. 2004, Planetary Systems in the Universe, 202, 44

Sahlmann, J., Lazorenko, P. F., Ségransan, D., et al. 2013, Astronomy & Astrophysics, 556, A133

Sahlmann, J., Lovis, C., Queloz, D., & Ségransan, D. 2011, Astronomy & Astrophysics, 528, L8

Sanchis-Ojeda, R., Rappaport, S., Pallè, E., et al. 2015, The Astrophysical Journal, 812, 112

Schreier, F. 1992, Journal of Quantitative Spectroscopy & Radiative Transfer, 48, 743

Seager, S. 2011, Contemporary Physics, 52, 602

Seager, S., & Mallén-Ornelas, G. 2003a, The Astrophysical Journal, 585, 1038

—. 2003b, The Astrophysical Journal, 585, 1038

Segura, A., Krelove, K., Kasting, J. F., et al. 2003, Astrobiology, 3, 689

Segura, A., Walkowicz, L. M., Meadows, V., Kasting, J., & Hawley, S. 2010, Astrobiology, 10, 751

Sellke, T., Bayarri, M. J., & Berger, J. O. 2001, The American Statistician, 55, 62

Sharp, C. M., & Burrows, A. 2007, The Astrophysical Journal Supplement Series, 168, 140

Shaw, J. R., Bridges, M., & Hobson, M. P. 2007, Monthly Notices of the Royal Astronomical Society, 378, 1365

Showman, A. P., Cooper, C. S., Fortney, J., & Marley, M. S. 2008, The Astrophysical Journal, 682, 559

Showman, A. P., Fortney, J., Lewis, N. K., & Shabram, M. 2012, Nature, 762, 24

Showman, A. P., Fortney, J., Lian, Y., et al. 2009, The Astrophysical Journal, 699, 564

Showman, A. P., & Guillot, T. 2002, Astronomy & Astrophysics, 385, 166

Sing, D. K., Vidal-Madjar, A., Désert, J. M., Lecavelier des Etangs, A., & Ballester, G. 2008, The Astrophysical Journal, 686, 658

Sing, D. K., Fortney, J., Nikolov, N., et al. 2016, Nature, 529, 59

Sivia, D. S. 2011, Elementary Scattering Theory, For X-ray and Neutron Users (Oxford University Press)

Skilling, J. 2004, in 24th AIP Conf. Proc., Killaha East, Kenmare, Kerry, Ireland (AIP), 395–405

Skilling, J. 2006, Bayesian Analysis, 1, 833

Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, The Astronomical Journal, 131, 1163

Smith, B. A., Soderblom, L. A., Beebe, R., et al. 1979, Science, 206, 927

Sneep, M., & Ubachs, W. 2005, Journal of Quantitative Spectroscopy & Radiative Transfer, 92, 293

Snell, H. E., Moncet, J.-L., Anderson, G. P., et al. 1995, in Proceedings of the SPIE, ed. J. C. Dainty, Atmospheric and Environmental Research, Inc., USA (SPIE), 88–95

Snellen, I. A. G., Albrecht, S., de Mooij, E. J. W., & Le Poole, R. S. 2008, Astronomy & Astrophysics, 487, 357

Snellen, I. A. G., Brandl, B. R., de Kok, R. J., et al. 2014, Nature, 509, 63

Snellen, I. A. G., de Kok, R. J., de Mooij, E. J. W., & Albrecht, S. 2010, Nature, 465, 1049

Solodov, A. M., & Starikov, V. I. 2008, Optics and Spectroscopy, 105, 14

Sotin, C., Grasset, O., & Mocquet, A. 2007, Icarus, 191, 337

Southworth, J. 2010, Monthly Noticese of the Royal Astronomical Society, 408, 1689

Sozzetti, A., Giacobbe, P., Lattanzi, M. G., et al. 2014, Monthly Notices of the Royal Astronomical Society, 437, 497

Spiegel, D. S., Haiman, Z., & Gaudi, B. S. 2007, The Astrophysical Journal, 669, 1324

Stam, D. M., De Rooij, W. A., Cornet, G., & Hovenier, J. W. 2006, Astronomy & Astrophysics, 452, 669

Steele, P. R., Burleigh, M. R., Dobbie, P. D., Jameson, R. F., & Barstow, M. A. 2011, Monthly Noticese of the Royal Astronomical Society, 416, 2768

Steinmetz, T., Wilken, T., Araujo-Hauck, C., et al. 2008, Science, 321, 1335

Stevenson, D. J. 2013, Planets, 195

Stevenson, K. B., Harrington, J., Nymeyer, S., et al. 2010, Nature, 464, 1161

Strege, C., Bertone, G., Feroz, F., et al. 2013, Journal of Cosmology and Astroparticle Physics, 04, 013

Strong, K., Taylor, F. W., Calcutt, S. B., Remedios, J. J., & Ballard, J. 1993, Journal of Quantitative Spectroscopy & Radiative Transfer, 50, 363

Struve, O. 1952, The Observatory, 72, 199

Swain, M., Vasisht, G., & Tinetti, G. 2008, Nature, 452, 329

Swain, M., Tinetti, G., Vasisht, G., et al. 2009, The Astrophysical Journal, 704, 1616

Tennyson, J., & Yurchenko, S. N. 2012, Monthly Notices of the Royal Astronomical Society, 425, 21

Tennyson, J., Yurchenko, S. N., Al-Refaie, A. F., et al. 2016, arXiv.org, arXiv:1603.05890

Terrile, R. J., Fink, W., Huntsberger, T., et al. 2005, 37, 31.19

Terrile, R. J., Tinetti, G., Lee, S., et al. 2006, American Geophysical Union, 21

Thiabaud, A., Marboeuf, U., Alibert, Y., Leya, I., & Mezger, K. 2015, Astronomy & Astrophysics, 580, A30

Thomas, G. E., & Stamnes, K. 2002, Radiative Transfer in the Atmosphere and Ocean, 546

Thompson, P., Cox, D. E., Hastings, J. B., & IUCr. 1987, Journal of Applied Crystallography, 20, 79

Thompson, W. J. 1993, Computers in Physics, 7, 627

Tinetti, G., Deroo, P., Swain, M., et al. 2010, The Astrophysical Journal Letters, 712, L139

Tinetti, G., Tennyson, J., Griffith, C. A., & Waldmann, I. 2012, Philosophical Transactions of the Royal Society A: Mathematical, 370, 2749

Tinetti, G., Vidal-Madjar, A., Liang, M.-C., et al. 2007, Nature, 448, 169

Tinetti, G., Drossart, P., Eccleston, P., et al. 2015, Experimental Astronomy, 40, 329

Todorov, K. O., Deming, D., Knutson, H. A., et al. 2013, The Astrophysical Journal, 770, 102

Tofflemire, B. M., Wisniewski, J. P., Kowalski, A. F., et al. 2012, The Astronomical Journal, 143, 12

Torres, G., Winn, J. N., & Holman, M. J. 2008, The Astrophysical Journal, 677, 1324

Trotta, R. 2007, Monthly Notices of the Royal Astronomical Society, 378, 819

—. 2008, Contemporary Physics, 49, 71

Tsiaras, A., Waldmann, I., Rocchetto, M., et al. 2015, arXiv.org, arXiv:1511.07796

Tsiaras, A., Rocchetto, M., Waldmann, I., et al. 2016, The Astrophysical Journal, 820, 99

Turing, A. M. 1950, Computing Machinery and Intelligence (Reprint) (Mind Design II: Philosophy)

Udalski, A., Paczynski, B., Zebrun, K., et al. 2002, Acta Astronomica, 52, 1

Udry, S., & Santos, N. C. 2007, Annual Review of Astronomy and Astrophysics, 45, 397

Valencia, D., Guillot, T., Parmentier, V., & Freedman, R. S. 2013, The Astrophysical Journal, 775, 10

Valencia, D., Sasselov, D. D., & O'Connell, R. J. 2007, ApJ, 665, 1413

Valenti, J. A., & Fischer, D. A. 2005, The Astrophysical Journal Supplement Series, 159, 141

Van de Hulst, H. C. 1957, Light Scattering by Small Particles (Courier Corporation)

Van de Hulst, H. C., & Twersky, V. 1957, Physics Today, 10, 28

Varanasi, P., & Chudamani, S. 1990, Journal of Quantitative Spectroscopy and Radiative Transfer (ISSN 0022-4073), 43, 1

Vazan, A., & Helled, R. 2012, The Astrophysical Journal, 756, 90

Vennes, S., Kawka, A., & Németh, P. 2010, Monthly Noticese of the Royal Astronomical Society, 404, L40

Venot, O., & Agúndez, M. 2015, Experimental Astronomy, 40, 469

Venot, O., Agúndez, M., Selsis, F., Tessenyi, M., & Iro, N. 2014, Astronomy & Astrophysics, 562, 51

Venot, O., Hébrard, E., Agúndez, M., Decin, L., & Bounaceur, R. 2015, Astronomy & Astrophysics, 577, A33

Venot, O., Hébrard, E., Agúndez, M., et al. 2012, Astronomy & Astrophysics, 546, A43

Venot, O., Rocchetto, M., Carl, S., Roshni Hashim, A., & Decin, L. 2016, The Astrophysical Journal, 830, 77

Venot, O., Fray, N., Bénilan, Y., et al. 2013, Astronomy & Astrophysics, 551, A131

Veras, D., Leinhardt, Z. M., Bonsor, A., & Gaensicke, B. T. 2014, arXiv.org, 2493

Veras, D., Mustill, A. J., Bonsor, A., & Wyatt, M. C. 2013, Monthly Noticese of the Royal Astronomical Society, 431, 1686

Vidal-Madjar, A., Lecavelier des Etangs, A., Désert, J. M., et al. 2003, Nature, 422, 143

Vogt, S. S., Allen, S. L., Bigelow, B. C., et al. 1994, Proc. SPIE Instrumentation in Astronomy VIII, 2198, 362

von Braun, K., Boyajian, T. S., ten Brummelaar, T. A., et al. 2011, The Astrophysical Journal, 740, 49

von Hippel, T., Kuchner, M. J., Kilic, M., Mullally, F., & Reach, W. T. 2007, ApJ, 662, 544

Voronin, B. A., Lavrentieva, N. N., Mishina, T. P., et al. 2010, Journal of Quantitative Spectroscopy & Radiative Transfer, 111, 2308

Wagner, F. W., Tosi, N., Sohl, F., Rauer, H., & Spohn, T. 2012, Astronomy & Astrophysics, 541, A103

Wagner, K., Apai, D., Kasper, M., et al. 2016, Science, 353, 673

Waldmann, I., Rocchetto, M., Tinetti, G., et al. 2015a, The Astrophysical Journal, 813, 13

Waldmann, I., Tinetti, G., Drossart, P., et al. 2012, Nature, 744, 35

Waldmann, I., Tinetti, G., Rocchetto, M., et al. 2015b, The Astrophysical Journal, 802, 107

Walker, G. A. H., Walker, A. R., Irwin, A. W., et al. 1995, Icarus, 116, 359

Wambsganss, J. 2004, Extrasolar Planets: Today and Tomorrow, 321, 47

Wang, H., Warner, S. J., Oehlschlaeger, M. A., et al. 2010, Combustion and Flame, 157, 1976

Wark, D. Q. 1961, Journal of Geophysical Research, 66, 77

—. 1970, Applied optics, 9, 1761

Wark, D. Q., & Hilleary, D. T. 1969, Science, 165, 1256

Wenger, C., & Champion, J. P. 1998, Journal of Quantitative Spectroscopy & Radiative Transfer, 59, 471

Werner, M. W., Roellig, T. L., Low, F. J., et al. 2004, The Astrophysical Journal Supplement Series, 154, 1

White, M. J., & Feroz, F. 2010, Journal of High Energy Physics, 07, 64

Wilson, D. J., Gänsicke, B. T., Koester, D., et al. 2014, Monthly Notices of the Royal Astronomical Society, 445, 1878

Winn, J. N. 2010, arXiv.org, arXiv:1001.2010

Wolff, B., Koester, D., & Liebert, J. 2002, Astronomy & Astrophysics, 385, 995

Wolszczan, A., & Frail, D. A. 1992, Nature (ISSN 0028-0836), 355, 145

Wyatt, M. C., Farihi, J., Pringle, J. E., & Bonsor, A. 2014, Monthly Notices of the Royal Astronomical Society, 439, 3371

Xu, S., & Jura, M. 2012, ApJ, 745, 88

—. 2014, The Astrophysical Journal Letters, 792, L39

Yamamoto, G. 1961, Journal of Meteorology, 18, 581

Yelle, R. V. 2004, Icarus, 170, 167

Yoo, J., DePoy, D. L., Gal-Yam, A., et al. 2004, The Astrophysical Journal, 603, 139

Yurchenko, S. N., Barber, R. J., & Tennyson, J. 2011, Monthly Noticese of the Royal Astronomical Society, 413, 1828

Yurchenko, S. N., Tennyson, J., Barber, R. J., & Thiel, W. 2013, Journal of Molecular Spectroscopy, 291, 69

Zellem, R. T., Griffith, C. A., Deroo, P., Swain, M., & Waldmann, I. 2014, The Astrophysical Journal, 796, 48

Zeng, L., & Sasselov, D. 2014, The Astrophysical Journal, 784, 96

Zhang, J., Lu, Z. H., & Wang, L. J. 2008, Applied optics, 47, 3143

Zuckerman, B., & Becklin, E. E. 1987, Nature, 330, 138

Zuckerman, B., Koester, D., Melis, C., Hansen, B. M., & Jura, M. 2007, The Astrophysical Journal, 671, 872

Zuckerman, B., Koester, D., Reid, I. N., & Hünsch, M. 2003, ApJ, 596, 477

Zuckerman, B., Melis, C., Klein, B., Koester, D., & Jura, M. 2010, ApJ, 722, 725