DISSERTATION | DOCTORAL THESIS

Titel | Title

Challenging atmospheric retrievals
Towards robust inference from exoplanet transmission spectra with next-generation telescopes

verfasst von | submitted by

Simon Schleich BSc MSc

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Doktor der Naturwissenschaften (Dr.rer.nat.)

Wien | Vienna, 2026

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 796 605 413
Dissertationsgebiet lt. Studienblatt | Field of study as it appears on the student record sheet:	Astronomie
Betreut von | Supervisor:	Ass.-Prof. Dr. Sudeshna Boro Saikia Univ.-Prof. Dipl.-Phys. Dr. Manuel Güdel Univ.-Prof. Dr. Aiko Voigt

**“The big candle lights the little candle,
and the little candle keeps burning.”**

Darach Ó Séaghdha

Motherfoclóir: Dispatches from a Not So Dead Language

Für meinen Großvater,
der mir gezeigt hat, dass Neugier kein Alter kennt

To my grandfather,
who showed me that curiosity does not age

Abstract

In the last three decades, the number of known planets outside the Solar System has grown into the thousands, spanning a vast parameter space. Characterising the atmospheres of these extrasolar planets offers a window into their formation histories and chemical compositions. Spectroscopic observations provide the primary means of probing these atmospheres, and the *James Webb* Space Telescope (JWST), alongside upcoming missions such as Ariel, has ushered in a new era of atmospheric characterisation, providing data with unprecedented signal-to-noise ratio (SNR), wavelength coverage, and spectral resolution.

To infer atmospheric properties from these observations, we rely on a technique named atmospheric retrieval, a Bayesian inference approach in which an atmospheric forward model is iteratively compared to an observed spectrum to estimate the posterior distributions of atmospheric parameters. The increased data quality from state-of-the-art observatories warrants a careful re-evaluation of this technique. In this thesis, we investigate the stability of atmospheric retrieval results in the context of transmission spectroscopy, with a focus on two key pillars: the forward model and the input data.

In the first part of this work, we investigate the impact of the pressure-temperature (p–T) prescription in the atmospheric forward model.Using synthetic spectra modelled after the hot Jupiter WASP-39 b, we construct a dataset spanning a range of vertical temperature structures, cloud decks, and instrument noise levels for JWST and the upcoming Ariel mission. We perform atmospheric semi-self-retrievals on these synthetic spectra, using heuristic multipoint p–T prescriptions of varying complexity. We evaluate the performance of these forward models through the Bayes factor and the accuracy of the estimated atmospheric parameters. The results consistently show that an isothermal p–T prescription is insufficient to retrieve known atmospheric parameters under state-of-the-art observing conditions. A multipoint profile is favoured in the majority of cases, with parameter recovery plateauing at a four-point profile. We conclude that a four-point p–T prescription represents the optimal balance between model complexity and retrieval performance, and should be preferred over isothermal prescriptions in transmission spectrum retrievals with JWST and Ariel.

In the second part of this work, we investigate the sensitivity of atmospheric retrievals to variations in the input transmission spectrum, using JWST NIRSpec PRISM observations of a single transit of WASP-39 b. Retrievals are performed on a baseline spectrum reduced as part of this work, using the Eureka! pipeline. We compare the resulting parameter posterior distributions to homogenised atmospheric retrieval performed on scattered instances of the spectrum produced in this work, and on independently reduced spectra derived from the same raw data. This analysis identifies three classes of parameter posterior distributions: stable Gaussian distributions for well-constrained parameters; stable uniform posteriors for weakly constrained parameters; and unstable, heavy-tailed posteriors. The presence of unstable posterior distributions makes robust interpretations of atmospheric characterisation results difficult, highlighting the importance of a careful assessment of credible interval sizes to properly reflect this uncertainty.

Taken together, the results of this thesis highlight the importance of rigorous forward model choices and careful treatment of input data in atmospheric retrievals. As the field moves into a new era of data quality and abundance, these findings underscore the need for systematic evaluation of retrieval methodologies to ensure the reliability of inferred atmospheric properties.

Declaration of published work

This thesis is result of the work I carried out during my PhD at the University of Vienna. In the three years of my project, I produced two first-author publications (the content of which I incorporated in this thesis). The list below summarises what parts of my thesis include content from these published works. Together with this, my thesis also includes parts of ongoing, unpublished work (currently in preparation to be submitted as a first-author publication).

- The work presented in Chapter 3 was carried out in collaboration with Sudeshna Boro Saikia, Quentin Changeat, Manuel Güdel, Aiko Voigt, and Ingo Waldmann. The initial analysis of simulated NIRSpec PRISM data has been published as "Knobs and dials of retrieving JWST transmission spectra - I. The importance of p–T profile complexity" in *Astronomy & Astropyhsics, 690, A336* (Schleich et al. 2024). The additional work performed on simulations of other JWST instruments and Ariel cases is currently in preparation under the title "Knobs and dials of retrieving JWST transmission spectra - III. Panchromatic preferences in p–T profile complexity".
- The work presented in Chapter 4 was carried out in collaboration with Sudeshna Boro Saikia, Quentin Changeat, Manuel Güdel, Aiko Voigt, and Ingo Waldmann. It was published as "Knobs and dials of retrieving JWST transmission spectra - II. Impact of pipeline-level differences on retrieval posteriors" in *Astronomy & Astropyhsics, 704, A223* (Schleich et al. 2025).

Simon Schleich
Vienna, March 2026

Contents

Chapter 1

Introduction

"This is the beginning of knowledge –
the discovery of something we did not know."

Frank Herbert

God Emperor of Dune

Over the past three decades, the study of planets and their atmospheres has grown far beyond the confines of our local neighbourhood. The eight planets of the Solar System and their satellites already exhibit a remarkable diversity in their atmospheric compositions. In the inner Solar System, we find small, dense planets with comparably tenuous atmospheres. Mercury, the closest planet to the Sun, has an extremely tenuous, collisionless exosphere directly bounded to its surface (Domingue et al. 2007). Venus and Earth, twins in their size and mass, have starkly contrasting atmospheres – Venus' atmosphere is dense and dominated by CO_2 (Marcq et al. 2017), while Earth's is comparably tenuous and filled with N_2 and O_2 (Salstein 1995). The atmosphere of Mars is even thinner, and mostly comprised of CO_2 (Vandaele et al. 2024). Beyond the inner Solar System and asteroid belt, we find the gas giants, Jupiter and Saturn, and colder ice giants, Uranus and Neptune. Their atmospheres are dominated by H_2 and He, enriched with some heavier volatiles, such as H_2O, CH_4 and NH_3 (Irwin 2003). Several of their moons have tenuous gaseous envelopes themselves. Titan, a moon of Saturn, has an atmosphere mainly composed of N_2 and CH_4 (Hörst 2017), and Io, in orbit around Jupiter, has a thin SO_2 atmosphere (Lellouch et al. 2007). However, in the context of the rapidly growing inventory of extrasolar planets, they represent only a limited and possibly atypical sample of the outcomes of planet formation and evolution. The discovery of thousands of extrasolar planets has transformed planetary science into a comparative and statistical discipline, enabling the exploration of exoplanetary parameters across a wide range of masses, system architectures, and stellar environments.

Unlike the planets of the Solar System, most exoplanets cannot be studied by direct, spatially resolved observations. Instead, we investigate their physical properties by examining subtle signatures they imprint on the light we observe from their host stars. These measurements inform us about atmospheric chemistry, temperatures, and the presence of aerosols, but only through the interaction of radiation with the atmosphere and the filtering properties of instrumentation and noise. As a result, the study of exoplanets is fundamentally an inference problem – observable quantities are mapped to physical parameters within a model that is a simplified representation of a complex system. Understanding the limitations of both the observational data we gather and the models we use to represent them is crucial to reliably draw conclusions about the nature of an exoplanet's atmosphere from remote observations.

As our sample sizes have grown, attention has increasingly shifted from individual objects to exoplanetary populations. Trends in radii, masses, and atmospheric composition have been used to probe processes such as planetary formation (e.g., Sun et al. 2024) and atmospheric photoevaporation (e.g., Fulton et al. 2017). Disentangling physical correlations from observational and methodological biases is therefore a key challenge in interpreting exoplanet atmospheric surveys. If we aim to characterise the atmospheres of extrasolar planets and place them in frameworks that investigate their characteristics on the level of planetary populations, we first need to find planets beyond the Solar System. The methods we use to detect exoplanets and the parameter space they cover are what we will look at first.

1.1 Observing extrasolar planets

Two definitions are commonly used to refer to the "first discovered exoplanet": Chronologically, the first confirmed planets outside the Solar System go back to Wolszczan and Frail (1992). Using the aptly named *pulsar timing* technique, they reported on two planet-sized bodies orbiting the pulsar PSR B1257+12. The first discovered exoplanet orbiting a main-sequence star was reported by Mayor and Queloz (1995), around the G-dwarf 51 Peg. Through Doppler-shifted spectral lines originating from the motion of the star-exoplanet pair around a common centre of gravity, they confirmed the presence of the Jupiter-mass companion 51 Peg b with the *radial velocity* method. Since these initial detections, more than 6 000 extrasolar planets have been discovered and confirmed (NASA EPA 2025). In addition to the pulsar timing and radial velocity techniques, a variety of observational methods have been employed to detect exoplanets, significantly expanding the collective database of known planets outside the Solar System.

Figure 1.1 illustrates the cumulative number of confirmed exoplanets by year. We see that the first three confirmed exoplanets were discovered using the pulsar timing method (Wolszczan and Frail 1992; Wolszczan 1994). The yield of this technique has been very modest afterwards, and today we count only eight exoplanets discovered using this method. In contrast, the radial velocity technique, starting with the discovery of 51 Peg b (Mayor and Queloz 1995), quickly became the dominant method for detecting exoplanets. It took almost another decade for the now-dominant *transit* technique to emerge with the discovery of OGLE-TR-53 b (Udalski et al. 2002). The radial velocity and transit methods, which have nearly exponentially expanded our inventory of known exoplanets, are what we will examine next.

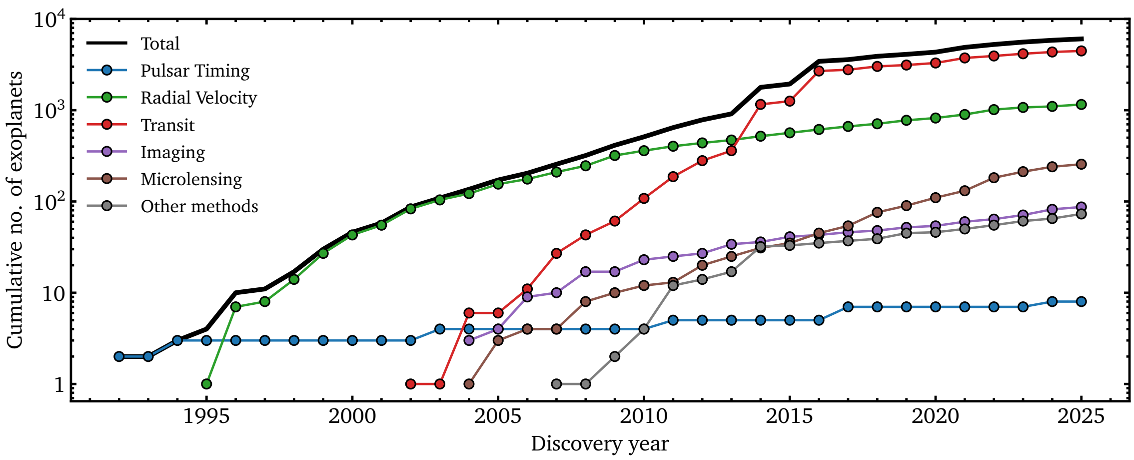

Figure 1.1: Cumulative number of discovered exoplanets by discovery year. Contributions to the total cumulative number (black line) are colour-coded by the associated discovery method. Data courtesy of NASA EPA (2025).

1.1.1 Radial velocity

Photons reaching an observer from a source in motion undergo a shift in wavelength corresponding to the relativistic Doppler effect. This shift is described by (Einstein 1905)

$$\lambda_{\text{obs}} = \lambda_0 \cdot \frac{1 + \frac{v}{c}\cos(\theta)}{\sqrt{1 - \left(\frac{v}{c}\right)^2}}, \tag{1.1}$$

where λ_{obs} and λ_0 are, respectively, the observed and rest-frame wavelengths, v is the magnitude of the velocity of the source's motion, θ is the angle between the source-observer line of sight and the direction of motion, and c is the speed of light. Equation 1.1 can be used to measure, for instance, the recession motion of stars by their radial velocities. When $v \ll c$, Equation 1.1 reduces to

$$\lambda_{\text{obs}} \approx \lambda_0 \cdot \left[1 + \frac{v}{c}\cos(\theta)\right]. \tag{1.2}$$

In a system of an exoplanet orbiting around its host star, the motion of the exoplanet will trigger a corresponding motion of the star around their common centre of mass, also called the *barycentre*. This motion happens on top of the previously mentioned recession velocity, inducing periodic signals in the measured radial velocity of a star. In the limiting case that the mass of the exoplanet is much smaller than the mass of the star ($M_{\text{p}} \ll M_*$), the amplitude of this signal follows the relation (e.g., Wright 2018)

$$K \approx \frac{M_{\text{p}}\sin(i)}{\sqrt{1-e^2}} \left(\frac{2\pi G}{P M_*}\right)^{1/3}, \tag{1.3}$$

where M_{p} and M_* are the mass of the exoplanet and of the host star, e and P are the eccentricity and period of the exoplanet's orbit, i is the inclination of the orbital plane from the line of sight of an observer ($i = 0$ represents a face-on orbit), and G is the gravitational constant. In monitoring the radial velocities of stars, it is therefore possible to detect exoplanets by the periodic red- and blueshift of the stellar spectrum. We see from Equation 1.3 that the strongest influence on K comes from the mass of the exoplanet, next to weaker correlations with the orbital period and eccentricity. This favours the detection of more massive exoplanets. As a reference point, Jupiter induces a radial velocity modulation amplitude of $K_{\text{J}} \approx 12\,\text{m}\,\text{s}^{-1}$ on the Sun (assuming $i = 0$). In contrast, the amplitude induced by Earth is only $K_{\oplus} \approx 0.1\,\text{m}\,\text{s}^{-1}$, which represents a critical precision threshold value in the search for Earth-analogue exoplanets orbiting a Sun-like star (Hara and Ford 2023).

Precisely constraining the mass of an exoplanet is a crucial factor when we investigate both the bulk properties of an exoplanet and its atmospheric properties (Batalha et al. 2019). However, in accurately deriving an exoplanet's mass from the radial velocity amplitude, we are hindered by the geometric factor $\sin(i)$. This means that we can only measure upper limits on the mass of an exoplanet if the inclination of the system to the observer is not known.

Constraining this using additional techniques, such as exoplanetary transits or astrometric measurements, can improve mass estimates (Wright 2018).

To date, the radial velocity technique is the second-most successful detection method for exoplanets, with a yield of more than 1 000 confirmed objects (as we see in Figure 1.1). The first generation of instruments used to detect exoplanets with this technique, such as OHP-ELODIE (Baranne et al. 1996) and Keck-HIRES (Vogt et al. 2000), operated with a precision comparable to K_J. However, the latest generation of spectrographs, such as Lowell-EXPRES (Jurgenson et al. 2016) and VLT-ESPRESSO (Pepe et al. 2021), has reached detection limits of only a few $K_\oplus$, promising to further widen the parameter space of known exoplanets. In terms of detection numbers, the success of the radial velocity technique is surpassed only by the transit method, which monitors stellar brightness.

1.1.2 Transits

If a star-exoplanet system has an exoplanet moving in an orbital plane aligned with the line of sight of an observer, then, at some point during its orbit, the exoplanet will pass between its host star and the observer. During this transit, the measured flux from the host star decreases as part of the stellar disk is blocked by the transiting exoplanet. This reduction in flux is directly related to the area blocked by the transiting body through

$$\delta_{\mathrm{p}} = \frac{4\pi R_{\mathrm{p}}^2}{4\pi R_*^2} = \left(\frac{R_{\mathrm{p}}}{R_*}\right)^2, \tag{1.4}$$

where δ_{p} represents the reduction in flux measured from the host star, and R_* and R_{p} represent the radius of the host star and its exoplanetary companion, respectively (as illustrated in Figure 1.2). In measuring δ_{p}, we therefore gain information about the size of an exoplanet (relative to the size of the star).

Observational constraints limit the range of exoplanets we can detect with the transit method. They therefore bias the parameter space of the exoplanets we do detect. Firstly, it requires a close alignment between the exoplanet's orbital plane and the observer's line of sight. For an exoplanet on a circular orbit, the sky-projected distance between the centre of the stellar disk and the mid-point of an exoplanet's transit, also called the *impact parameter*, is given by

$$b \equiv \frac{a\cos(i)}{R_*}. \tag{1.5}$$

It depends on the semi-major axis, a, and inclination, i, of the exoplanet's orbit, and the radius of the host star, R_*. In Figure 1.2, we see b in reference to the centre lines of the host star (dashed grey lines). The minimum inclination for which a non-grazing (i.e., fully occulting) transit can be observed is given by $b = 1 - R_{\mathrm{p}}/R_*$. In the limit of $R_{\mathrm{p}} \ll R_*$, this minimum inclination equivalently provides us with the probability, p, of the transit being detectable if

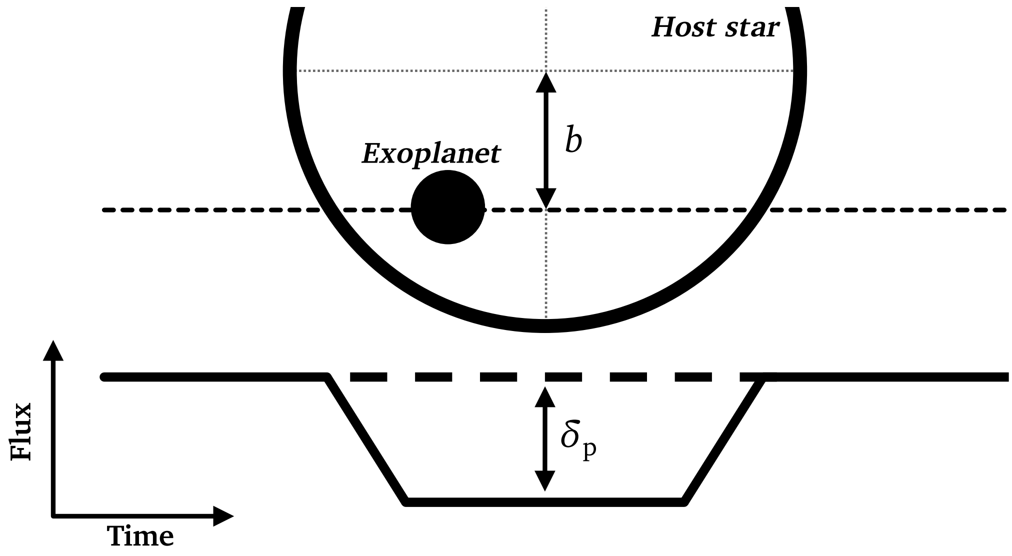

Figure 1.2: Sketch of an exoplanet transit. The centre of the transit is offset by the impact parameter, b, from the centre of the stellar disk. The exoplanet reduces the flux measured from the host star by a factor δ_p at mid-transit.

the orbit is randomly oriented (Winn 2010)

$$\cos(i_\mathrm{min}) = \frac{R_*}{a} = p. \tag{1.6}$$

As the ratio of R_* to a is commonly very small, i is restricted to values close to 90°. Secondly, this method relies on long-term, precise photometric monitoring of stars. Therefore, it favours the detection of exoplanets that frequently (i.e., on a short orbital period) produce a larger signal (i.e., have a large radius compared to their host star). As a reference, in the Solar System the transit signal induced by Jupiter is on the order of $\delta_\mathrm{J} \approx 10^{-2}$, while the corresponding value for Earth is $\delta_\oplus \approx 10^{-4}$. These signals are also induced only once every 11.83 years and once every year, respectively. In contrast, Equations 1.4 and 1.6 show that the parameter space favoured for transit detections is biased towards the detection of large, short-period exoplanets.

The first detection of an exoplanetary transit was not a discovery. Transits of HD 209458 b, a short-period Jupiter-mass exoplanet, were observed after its initial detection through the radial velocity technique (Charbonneau et al. 2000; Henry et al. 2000). As we have seen in Equation 1.6, the low probability of an appropriate viewing angle reduces the likelihood of detecting exoplanets through transits. However, this problem can be remedied through large-scale monitoring campaigns, conducted using both ground-based and space-based instruments. The first exoplanet discovered using the transit technique, OGLE-TR-53 b (Udalski et al. 2002), was the result of such a monitoring effort. It and many of the best-studied exoplanets today, such as WASP-39 b (Faedi et al. 2011), K2-18 b (Montet et al. 2015), or the planets of the TRAPPIST-1 system (Gillon et al. 2016), bear the names of these campaigns. Ground-based monitoring campaigns predate space-based programmes. The majority of exoplanets discovered via the transit method using ground-based instruments were and are being observed within two large-scale monitoring campaigns. The *Hungarian Automated Telescope Network* (HATNet, Bakos et al. 2004; Bakos et al. 2013), and the *Wide Angle Search for Planets* (WASP, Pollacco et al. 2006) have contributed on the order of 200 exoplanets through their discoveries. Results from these campaigns have been supplemented by multiple smaller campaigns, such as the *Optical Gravitational Lensing Experiment* (OGLE, Udalski et al. 1997) and the *Trans-atlantic Exoplanet Search* network (TrES, Alonso et al. 2004).

The remarkable success of the transit method began with the first space-based monitoring campaigns. The *Convention, Rotation, and planetary Transits* mission (CoRoT, Auvergne et al. 2009), operating between 2007 and 2013, has provided a modest yield of 34 exoplanets, but it paved the way for space missions dedicated to the detection of exoplanets. The *Kepler* space telescope (Borucki et al. 2010) and its extended mission *K2* (Howell et al. 2014), operating between 2009 and 2018, have already yielded on the order of 3 000 confirmed exoplanets, with approximately the same number of candidates still awaiting confirmation (Christiansen et al. 2025). Combined, Kepler and K2 present the most successful exoplanet-hunting programme to date. However, the *Transiting Exoplanet Survey Satellite* (TESS, Ricker et al. 2015) is on its way to surpass this. Launched in 2018, it has to date discovered about 700 exoplanets, but it has also (so far) yielded a staggering 7 800 candidates that await confirmation (Christiansen et al. 2025). In the future, the *PLAnetary Transits and Oscillations of stars* mission (PLATO, Rauer et al. 2025) promises to extend our inventory of known exoplanets even further. It aims to detect small exoplanets (comparable in size to Earth) with orbital periods on par with those of the rocky planets in the Solar System, diversifying the parameter space of known exoplanets further. The recently launched *Characterising ExOPlanet Satellite* mission (CHEOPS, Benz et al. 2021) will also vitally contribute to this parameter space. While its primary goal is not the discovery of exoplanets through the transit method, it will detect exoplanetary transits as follow-ups from already known systems, providing precise radius measurements and improving the bulk parameters of known exoplanets.

1.1.3 Other methods

As we see in Figure 1.1, the radial velocity and transit techniques are by far the most successful methods to detect exoplanets to date. However, multiple other methods have also added significant numbers to the inventory of known exoplanets.

Direct imaging

Detecting exoplanets through radial velocity shifts and transit events relies on measurements of the host star and how an exoplanet influences it. Contrary to this, the *direct imaging* of exoplanets resolves the exoplanet itself as a point-source of light, separate from the host star. This direct imaging of a stellar companion relies on it having a sufficiently large self-luminosity and spatial separation from the much brighter parent body to be detectable. The first substellar companion detected with this technique was the Brown Dwarf Gliese 229 B (Nakajima et al. 1995), an intermediate object between a giant exoplanet and a small star.

Since then, direct imaging has also been used to detect stellar companions categorised as giant exoplanets ($M_\mathrm{p} < 13\,\mathrm{M_J}$), but observational limitations naturally shape the typical targets for this technique. Preferably, they are recently-formed (and therefore hot and more self-luminous) Jupiter-mass exoplanets at larger orbital distances than probed with the transit or radial velocity technique. These selection criteria limit the pool of detectable exoplanets. However, since these young exoplanets are difficult to detect and characterise using indirect methods (due to stellar variability), direct imaging provides us with observable exoplanets that present ideal laboratories for investigating planetary formation processes (Pueyo 2018).

Since the first directly imaged exoplanet, 2MASS J1207334-393254 b (Chauvin et al. 2005), this technique has successfully added on the order of 100 exoplanets to our inventory (as indicated in Figure 1.1), including the well-studied system of HR 8799 containing four giant exoplanets (Marois et al. 2008; 2010) and WD-0806-661 b, a cold giant exoplanet orbiting a white dwarf (Luhman et al. 2011).

Gravitational microlensing

Gravitational lensing describes an effect in which a massive foreground object (the *lens*) distorts the photons arriving at an observer from a distant background object (the *source*). While this is commonly associated with images of distant galaxies, exoplanet detection occurs through lensing on a much smaller scale (hence *microlensing*). In a microlensing event, the lensed source (a background star) is not resolved. Instead, from the observer's point of view, a star passing in front of it acts as a lens, amplifying the source's brightness. The shape of this brightness amplification curve, measured over time, depends on the mass distribution within the lensing system. If exoplanetary companions accompany the lens, the amplification curve will exhibit distinct shapes that can be used to infer the exoplanet's mass and separation from the lensing host star (Batista 2018).

The first detections of exoplanets through microlensing events already revealed the unique parameter space it accesses: from Jupiter-mass (Bond et al. 2004) to cold Neptune-equivalent (Gould et al. 2006) and even Earth-sized (Beaulieu et al. 2006) exoplanets. To date, the microlensing technique has added on the order of a few hundred exoplanets to our parameter space. Its biggest drawback is the transient nature of microlensing events. The exoplanets discovered this way are unique in the parameter space of known exoplanets, but they are not open to follow-up analyses.

Remaining methods

In addition to the techniques discussed above, several other methods have been successfully used to detect exoplanets. Firstly, the gravitational impact of an exoplanet in a stellar system does not only manifest in measurable radial velocity changes of the host star. Exoplanets have also been detected by how they influence the astrometric (or angular) position of the host star in the sky (DENIS-P J082303.1-491201 b, Sahlmann et al. 2013), and by how they influence other companion bodies, such as in the variations they have on the timing of eclipsing binaries (DP Leo b, Qian et al. 2009) or the transits of other exoplanets (Kepler-19 c, Ballard et al. 2011). Secondly, exoplanets have also been discovered by their modulation of other measurable properties of their host stars. This includes the change in stellar pulsation timing (V 391 Peg b, Silvotti et al. 2007) and pulsation amplitude (KOI-55 b and c, Charpinet et al. 2011), and through their influence on the gas and dust kinematics of a circumstellar disk (HD 97048 b, Pinte et al. 2019).

With different observational constraints, the methods used to discover and confirm exoplanets have provided us with a parameter space vastly different from the few examples we have in our own Solar System. This 'zoo' of exoplanets is what we will have a look at next.

1.2 The zoo of exoplanets

Until the discovery of the first exoplanets three decades ago, the only references for planetary bodies orbiting stars were the planets in the Solar System. In its inner region, four rocky bodies with bulk compositions similar to Earth (i.e, terrestrial planets) orbit the Sun. In the outer Solar System, we find more massive but less dense gaseous bodies, collectively called gas giants. We can further divide these into conventional gas giants (Jupiter and Saturn) and ice giants (Neptune and Uranus). After three decades of exoplanet discoveries, one of the first things we have learned is that the structure of the Solar System does not agree with the vast majority of exoplanet systems we have discovered. We can see this in Figure 1.3, which shows confirmed exoplanets plotted in the orbital period and exoplanetary radius parameter space. Figure 1.3 also indicates the parameters of the planets in the Solar System for comparison. In this parameter space, they appear to form a boundary, with most discovered exoplanets sitting at shorter orbital periods. As we have seen in Section 1.1.2, this bias stems from the observational constraints imposed by the transit method, which has yielded the most exoplanets to date and is the primary method for measuring exoplanetary radii. While we can find a less concentrated spread of exoplanets in the parameter space of exoplanet mass against orbital period (NASA EPA 2025), the structure of the Solar System appears to still be an outlier compared to the exoplanetary systems we know today (Zhu and Dong 2021). We can divide these exoplanets populating this parameter space into different categories.

1.2.1 Hot Jupiters

At the top of Figure 1.3 sit exoplanets with radii and masses similar to the gas giants in our Solar System. At long periods, we find giant exoplanets most similar to the Solar System's gas giants. However, the bulk of discovered exoplanets in this region orbit their host stars on much shorter periods. Consequently, they receive a significantly increased amount of irradiation. We call these exoplanets *hot Jupiters*. The first exoplanet discovered to orbit a main-sequence star was of this type (51 Peg b, Mayor and Queloz 1995). While the bulk properties of these exoplanets are similar to those of Jupiter and Saturn, with atmospheres dominated by hydrogen

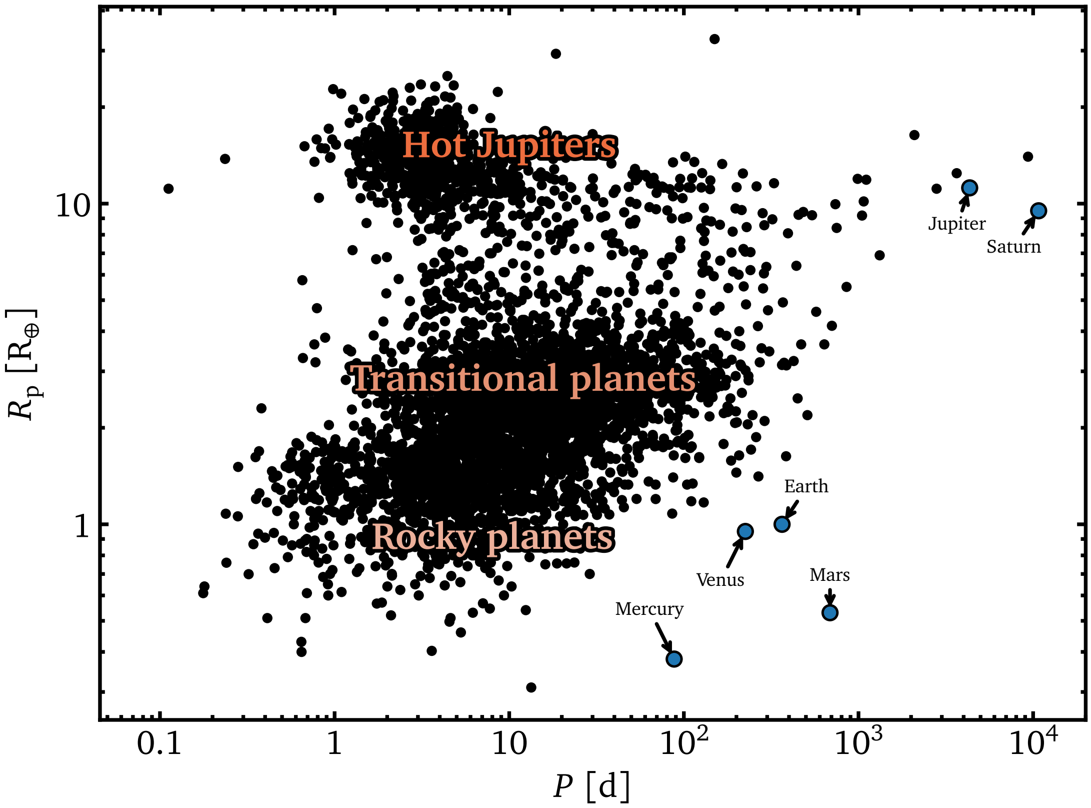

Figure 1.3: Period-radius plot of confirmed exoplanets, showing orbital period (in d, x-axis) against exoplanet radius (in $R_\oplus$, y-axis). Solar System planet parameters (excluding Uranus and Neptune) are shown in blue for reference. For better visualisation, the plot is constrained to $P \leqslant 2 \times 10^4$ d. Data courtesy of NASA EPA (2025).

and helium, their proximity to the parent star places them in an environment where we do not find gas giants in the Solar System.

It is still unclear where and how hot Jupiters form. Gas giants in general are believed to form through either core accretion or gravitational instabilities within the protoplanetary disk. In the former, rocky cores are proposed to accrete many times their mass in gas from the protoplanetary disk (Pollack et al. 1996), while in the latter, parts of the protoplanetary disk fragment into gravitationally bound clumps (Boss 1997). These theories must be linked to the close orbital distances of hot Jupiters, raising questions about their formation locations. The main hypotheses for the origin of hot Jupiters are split between *in situ* formation and migration. In the first case, conditions of the protoplanetary disk might prevent the fragmentation of the gas, making *in situ* formation through gravitational instability unlikely (Rafikov 2005). Core accretion could operate this close to the host star, but forming a sufficiently large core could be difficult (Schlichting 2014; Batygin et al. 2016). In contrast, hot Jupiters have also been proposed to form at much larger orbital distances, where conditions for either formation mechanism are more achievable, and then migrate to their observed orbital positions. This migration can be triggered by torques applied from the gas disk (Lin and Papaloizou 1986) or through the circularisation of perturbed, highly elliptical orbits (Weidenschilling and Marzari 1996). However, these processes are sensitive to the conditions of the protoplanetary disk and could lead to tidal disruptions or migration too far inwards (Trilling et al. 1998; Duffell et al. 2014).

One way to address questions about the formation of hot Jupiters is through investigations into their atmospheres. The location of an exoplanet's formation will imprint markers in an exoplanet's atmosphere in the form of atomic ratios (Öberg et al. 2011; Mordasini et al. 2016). Hot Jupiters are ideal targets for atmospheric characterisation. Their comparatively large sizes and low mean molecular weight (MMW) atmospheres (resulting in a large atmospheric scale-height) yield a high atmospheric signal-to-noise ratio (SNR). In this work, I will focus on the characterisation of hot Jupiters, but they represent only a small fraction of exoplanets that are known to us. We can see an illustration of the distribution of confirmed exoplanets by their radius in Figure 1.4. This distribution has a small peak at $R_\mathrm{p} \approx 12\,\mathrm{R}_\oplus$ (approximately the size of Jupiter), but the vast majority of known exoplanets are clustered at sizes of $R_\mathrm{p} < 4\,\mathrm{R}_\oplus$.

Before we consider the bulk of known exoplanets, there is one noteworthy absence of objects in the radius-period parameter space. If we look at Figure 1.3, we can find a general dearth of objects in the regime $5\,\mathrm{R}_\oplus < R_\mathrm{p} < 8\,\mathrm{R}_\oplus$, especially pronounced for orbital periods of $P < 3\,\mathrm{d}$. This region, often referred to as the *sub-Jovian desert*, is produced by two mechanisms: tidal disruption and photoevaporation. As discussed by Owen and Lai (2018), the first creates a mass barrier for migrating gas giants, forming the upper boundary of this region. The second mechanism strips smaller, close-in exoplanets of significant parts of their atmospheres, causing them to shrink and form the lower boundary of the sub-Jovian desert (Owen and Lai 2018), where the next category of exoplanets begins.

1.2.2 Transitional planets

Exoplanets with radii of $R_\mathrm{p} \in [1.0, 4.0]\,\mathrm{R}_\oplus$ fall within a size regime not present in the Solar System, sitting between the largest rocky bodies and the smallest gas giants. Exoplanets in this size category have been labelled *super-Earths*, *sub-Neptunes*, or *mini-Neptunes*. Here, we will use the umbrella term *transitional planets*. Disentangling their nature and structure is a question currently at the forefront of exoplanetary research, as we lack directly comparable bodies in the Solar System.

Explaining the bulk properties of these exoplanets using a combination of their interior and atmospheric composition is a degenerate problem (Rogers and Seager 2010), leading to several proposed explanations for their structure. This is supplemented by the *radius valley*, a defining feature of the population of transitional planets contributing to their unique nature. If we look at the distribution of exoplanets by their radius (Figure 1.4), it appears that exoplanets smaller than $4\,\mathrm{R}_\oplus$ make up a continuous, dominating population of all known exoplanets. However, studies of this smaller size regime have found a dearth of exoplanets centred at approximately $1.8\,\mathrm{R}_\oplus$ (Fulton et al. 2017; Van Eylen et al. 2018). We see this in the inset panel of Figure 1.4, which shows the distribution of exoplanets with radii smaller than $5\,\mathrm{R}_\oplus$. Although there are observational biases attached to this distribution, especially with the occurrence rate of exoplanets smaller than approximately $1.2\,\mathrm{R}_\oplus$ (Fulton et al. 2017), it illustrates two population peaks separated by a gap.

Two prevalent theories try to explain this gap, tied to the fundamental nature of its underlying populations. The first considers both peaks to consist of exoplanets of the same type: rocky cores covered by H_2-rich envelopes, referred to as *gas dwarfs* (Rogers et al. 2011; Rigby et al. 2024). Exoplanets with larger radii (and masses) can hold on to these envelopes, forming the peak at $R_\mathrm{p} \approx 2.2\,\mathrm{R}_\oplus$. The second peak at $R_\mathrm{p} \approx 1.6\,\mathrm{R}_\oplus$ is made up of smaller exoplanets that were stripped of their envelopes by photoevaporation or core-powered mass loss (Owen and Wu 2017; Ginzburg et al. 2018). The gas-dwarf scenario has several open questions associated with it: Rocky cores might not be able to accumulate significant H_2-envelopes without accreting significant fractions of other volatiles (Fortney et al. 2013), and it is unclear if their surfaces should be molten, forming magma oceans that imprint molecular signatures into their atmospheres (Kite et al. 2020; Misener et al. 2023). In contrast to a homogeneous population, the second theory explaining the radius valley posits that the two populations are inherently different. Smaller exoplanets are rocky with tenuous envelopes, while larger transitional planets fall in the broadly labelled category of *water worlds*. These water worlds have been proposed to have terrestrial atmospheres (dominated by N_2 and CO_2) sitting atop planet-covering oceans, or atmospheres containing significant amounts of H_2O (Léger et al. 2004; Adams et al. 2008; Zeng et al. 2019). A third possibility places tenuous H_2-envelopes above water oceans, dubbed *hycean worlds* (hydrogen-ocean, Madhusudhan et al. 2021).

One of the current frontiers of observations with the *James Webb* Space Telescope (JWST, Gardner et al. 2023) is to determine the bulk atmospheric properties of exoplanets in this size regime, which can inform us about their interiors and help break the degeneracy in their nature (Bean et al. 2021). However, the smaller size and expected higher MMW of these exoplanets make it more challenging to determine their atmospheric compositions. This challenge increases even more when we move to the last type of exoplanets we will discuss here – dense and rocky bodies.

1.2.3 Rocky planets

At the bottom of the period-radius parameter space are exoplanets with radii close to the size of Earth. Explaining the nature of exoplanets falling into the transitional regime is a degenerate problem. However, as their size approaches Earth's radius, their bulk density tells us that they are composed primarily of solid materials (elements such as Fe, Mg, or Si), possibly beneath comparatively small and tenuous atmospheres (Rogers 2015). I label these exoplanets here as *rocky planets*. Due to observational biases, small rocky exoplanets are preferentially detected around low-mass stars, and most of them are very close to their host stars (with orbital periods on the order of days). This favourably impacts the SNR of atmospheric observations (Barstow et al. 2016), but atmospheric characterisation is complicated due to stellar variability (Rackham et al. 2018; Thompson et al. 2024). It is also questionable if these exoplanets possess atmospheres

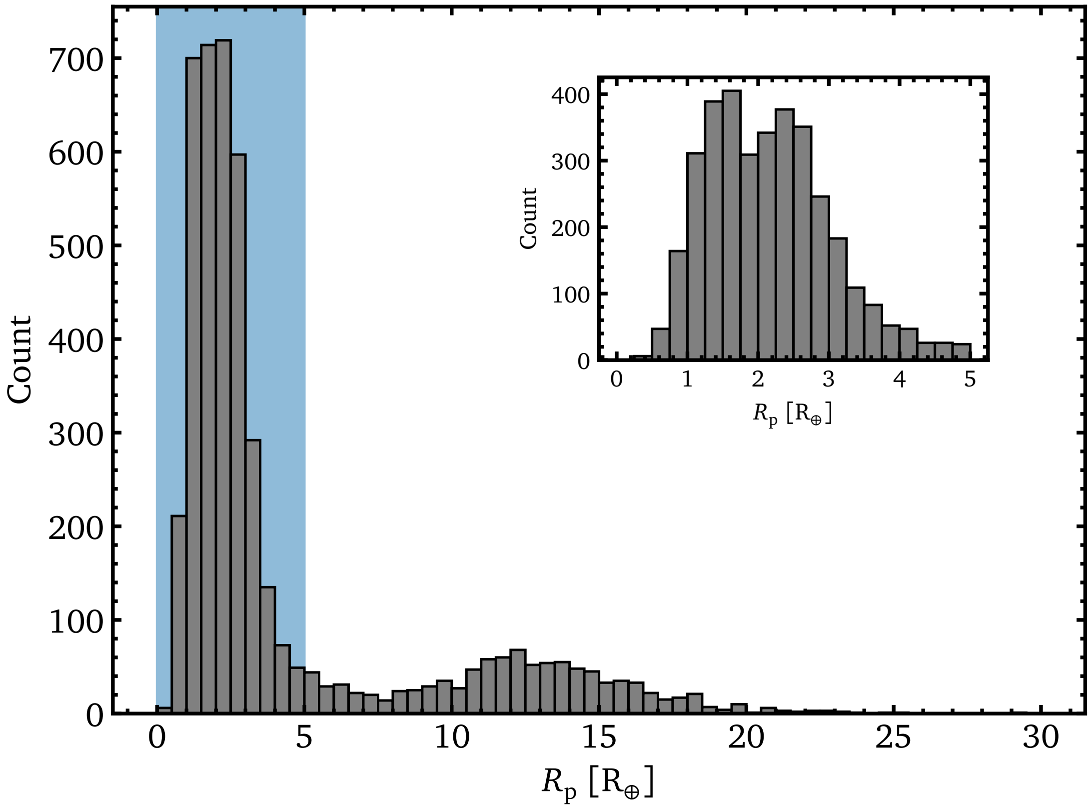

Figure 1.4: Distribution of confirmed exoplanets by their radius (in $R_\oplus$). The main panel shows all confirmed exoplanets with known radii, binned into radius elements of size $\Delta R_p = 0.5\,R_\oplus$. The blue-shaded size range is represented in the inset panel, showing exoplanets with $R_p \leqslant 5\,R_\oplus$. This sub-selection is binned in radius elements of size $\Delta R_p = 0.25\,R_\oplus$ Data courtesy of NASA EPA (2025).

to begin with – at orbital distances this small, increased instellation levels, especially in the X-ray and ultraviolet (XUV) wavelength regime, could strip away the atmospheres of rocky exoplanets within astronomically short time scales. This would leave them as barren rocks, although models do not agree on these scenarios (Nakayama et al. 2022; Van Looveren et al. 2025). This disparity can be resolved by observing atmospheres around these exoplanets, without necessarily characterising them. The detection of a rocky exoplanet atmosphere represents another fundamental frontier in observations with state-of-the-art instruments such as JWST.

Rocky planets with especially close orbits belong to an even more exotic subcategory in this parameter space, called *lava planets*. Although they are not necessarily restricted to radii comparable to Earth, their orbital proximity differentiates them from their size counterparts in the transitional planet regime. These exoplanets have surface temperatures high enough to melt rock, thereby forming a magma ocean on the exoplanet's surface. Subsequently, their atmospheres could contain silicates in vapour form (Schaefer and Fegley 2009; Lichtenberg and Miguel 2025). Few examples of such exoplanets are known and well characterised, but spectroscopic observations of their atmospheres suggest scenarios ranging from absent atmospheres to vaporised silicates or even thick volatile atmospheres (Crossfield et al. 2022; Zieba et al. 2022; Hu et al. 2024).

1.3 Investigating exoplanetary atmospheres

The discovery of the first extrasolar planets more than thirty years ago has produced a rapidly accelerating field of research. The parameter space of the exoplanets we have discovered so far is populated by objects that are significantly different from the planets we know in the Solar System. However, detecting the presence of planets in stellar systems outside our own is only the first step in understanding fundamental questions about their formation and nature. Placing the planets of our Solar System in the broader context of planetary populations requires us to understand more than the bulk properties of exoplanets. To address pressing questions about their nature and formation pathways, we need ways to more closely investigate their atmospheres.

1.3.1 Atmospheric retrieval

To derive the properties of exoplanetary atmospheres, we use a technique called *atmospheric retrieval*. This method uses an inverse modelling approach to estimate atmospheric characteristics, such as the vertical temperature distribution or the chemical composition, from observational data. Atmospheric retrieval has its origins in the study of the atmospheres we find within the Solar System (Rodgers 2000). Inverse modelling methods are used to characterise Earth's atmosphere from remote observations (e.g., Stubenrauch et al. 2008; Li et al. 2024). They are similarly used to characterise the atmospheric thermal, chemical, and aerosol properties of rocky bodies in the Solar System, including Venus (Marcq et al. 2006; Helbert et al. 2023), Mars (Guerlet et al. 2022; Kleinböhl et al. 2024), and Titan (Fan et al. 2019).

The information content of exoplanet atmospheric observations is not comparable with the information contained in Solar System atmospheric spectra. Early characterisation efforts on the atmospheres of hot Jupiters used simple atmospheric models combined with radiative transfer calculations to infer the presence of atomic or molecular species in an exoplanet's atmosphere from low-resolution spectroscopic data (Charbonneau et al. 2002; Tinetti et al. 2007). Madhusudhan and Seager (2009) proposed a way of using a parametric atmospheric forward model combined with millions of model evaluations to characterise exoplanet atmospheres driven by observational data. Since then, atmospheric retrievals have advanced significantly, and numerous frameworks have been developed for these purposes. For a comprehensive list of state-of-the-art atmospheric retrieval frameworks, I refer to MacDonald and Batalha (2023).

As inverse modelling techniques, atmospheric retrieval frameworks follow the core principle of statistical inference driven by observational data, and require three components: (1) a reliable dataset, (2) an atmospheric forward model, and (3) a parameter estimation method. I illustrate the general workflow of an atmospheric retrieval framework in Figure 1.5. It rests on two main pillars – the inference-guiding observational data and the forward model. The

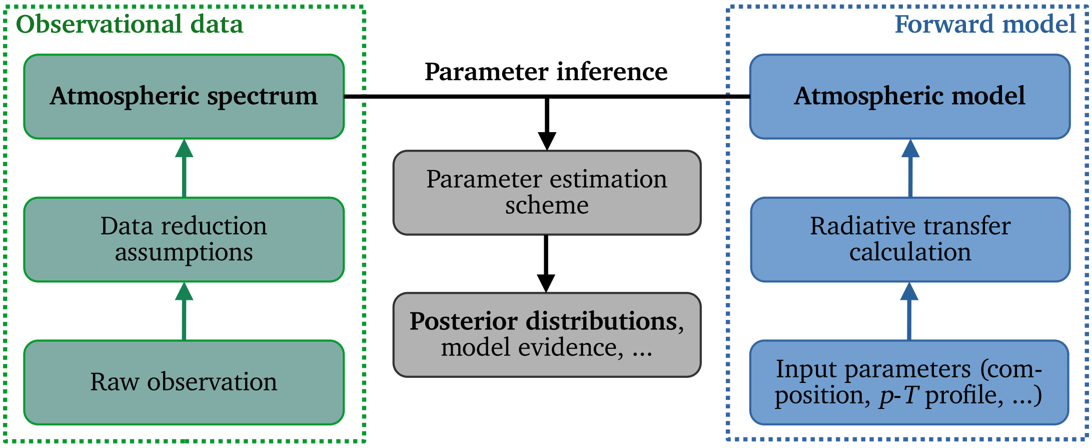

Figure 1.5: Structure of atmospheric retrieval frameworks. An atmospheric forward model is used to generate synthetic observations via radiative transfer calculations. These are compared to observed data. A parameter-estimation scheme is used to explore the parameter space of the forward model, resulting in posterior distributions for its parameters.

type of atmospheric observational data influences how the forward model is constructed, given the expected information it encodes. In general terms, it includes fixed parameters (such as planetary and stellar bulk properties) and variable parameters that encode prescriptions for the atmospheric chemical composition, the pressure-temperature structure, and the presence of aerosols. Synthetic observations calculated from this model are then compared with the observational data in an iterative scheme, yielding an ensemble of forward-model parameters at the end of the parameter-estimation process.

Chemistry

Forward models need to account for the chemical structure of the exoplanet's atmosphere. In the heuristic approach of 'free chemistry', molecular profiles are fitted individually for each species considered. These profiles are defined in the domain of the forward model, characterised by their vertical (or pressure-dependent) shape. Each profile is determined by n free parameters, representing the vertical nodes on which it is anchored. The vertical positions of these nodes can be free parameters themselves, raising this number to $2n$ per species. Because this introduces a large number of free parameters into the forward model, molecular profiles in this approach are often considered vertically homogeneous, thereby coarsely determining the presence and abundance of atmospheric absorbers (Changeat et al. 2019; Al-Refaie et al. 2022). Self-consistent chemical models can substitute for this flexible but coarse implementation of chemistry. They reduce the number of free parameters by trading them off against more rigid model assumptions, relying on underlying chemical networks and often derived under the assumption of thermochemical equilibrium. Atmospheric molecular profiles are inferred based on the minimisation of Gibbs-free energy (Stock et al. 2018; Woitke et al. 2018; Agúndez et al. 2020). Recently, chemical forward models have been developed that seek to circumvent these constraints by introducing additional free parameters to mimic disequilibrium processes

(Constantinou and Madhusudhan 2024) or by incorporating disequilibrium chemistry using reduced chemical networks (Al-Refaie et al. 2024).

Pressure-temperature structure

The vertical temperature structure in retrievals is captured in the atmospheric pressure-temperature (p–T) profile. Similar to chemical profiles, 'free' n-point p–T profiles are constructed by interpolating between n temperature values retrieved within the atmospheric pressure domain. The simplest form of this is an isothermal profile, characterised by a constant temperature throughout the atmosphere. Heuristic p–T profiles offer significant flexibility in determining the atmospheric temperature structure, but at the cost of a large number of free parameters. The degree of complexity of such a profile needs to be carefully investigated to ensure that the information content of the data is preserved and that the estimation of other forward-model parameters is not inadvertently biased (e.g. Rocchetto et al. 2016; Lueber et al. 2024). In contrast, parametric p–T prescriptions provide a trade-off between a smaller number of free parameters and more rigid assumptions. One common parametric implementation is the 'Guillot' profile (Guillot 2010; Line et al. 2012). It assumes thermochemical equilibrium in the atmosphere and approximates atmospheric radiation in three channels, split between two optical and one infrared channel, and depends on five free parameters. The Guillot profile is physically motivated through the assumption of radiative equilibrium, but it produces isothermal atmospheres at low pressures, which might not be an accurate reflection of an exoplanet's atmosphere (Barstow and Heng 2020; Fortney et al. 2021). The 'Madhusudhan' profile (Madhusudhan and Seager 2009) is another common parametric p–T profile. It uses three layers defined by generalised exponential profiles, designed to reproduce temperature profiles observed in the Solar System planets and from self-consistent atmospheric models of exoplanets. The Madhusudhan profile provides more flexibility in resolving temperature inversions at the cost of one additional free parameter compared to the Guillot profile, and has problems reproducing the deep atmospheric temperature structure from self-consistent models (Blecic et al. 2017).

Clouds

The inclusion of clouds inevitably complicates atmospheric retrievals, as it adds additional free parameters into the forward model. A simple prescription to account for the presence of clouds in an exoplanet's atmosphere is to assume a wavelength-independent opacity step function at a specified pressure, often referred to as *grey clouds*. Flat-opacity clouds add only a single free parameter to the forward model, but they can break down when the underlying data covers larger wavelengths (Barstow and Heng 2020). With increasing quality in observational data and the expected diversity in signatures from aerosols (Gao et al. 2021), more complex parametric and self-consistent prescriptions to account for the distribution and nature of aerosols have been implemented (Ormel and Min 2019; Ma et al. 2023; Changeat et al. 2025).

This diversity in atmospheric forward-model parametrisations might seem daunting. However, it pays to consider the type of underlying data that is fed into the parameter estimation process. This determines the kind of atmospheric information we are mainly probing and informs us on how to set up our retrieval schemes. The different types of observational data, and what we have learned from them, are what we will examine next.

1.3.2 Transmission spectroscopy

Transmission spectroscopy refers to spectroscopic observations of an exoplanet's transit. As we have seen in Equation 1.4, the drop in flux measured during transit directly encodes information about the radius of the exoplanet. While exoplanet detections using the transit method rely on photometric observations to determine an exoplanet's bulk size, panchromatic observations of the transit show that this radius depends on the observed wavelength. We can express this as

$$\delta_\mathrm{p}(\lambda) = \left(\frac{R_\mathrm{p}(\lambda)}{R_*}\right)^2, \tag{1.7}$$

where $\delta_\mathrm{p}(\lambda)$ and $R_\mathrm{p}(\lambda)$ are the wavelength-dependent transit depth and radius of the exoplanet, respectively, and R_* is the radius of the host star. We see an example of this in Figure 1.6, which shows the measured radius of a hot Jupiter from a panchromatic transit recorded with JWST. The panchromatic variations of the measured radius originate from the fact that the flux measured during a transit passes through the exoplanet's atmosphere, where it is attenuated primarily by molecular absorption. Figure 1.6 also shows where the absorption by H_2O and CO_2 is expected to be very pronounced, corresponding to significant increases in the measured exoplanetary radius (Zhang et al. 2024a).

Transmission spectroscopy has been primarily used to investigate the atomic and molecular inventory of exoplanetary atmospheres. The first exoplanet to have a detected atmospheric signature was the hot Jupiter HD 209458 b. Using observations with the Space Telescope Imaging Spectrograph (STIS), Charbonneau et al. (2002) reported the detection of absorption by Na and Vidal-Madjar et al. (2003) recovered the signature of escaping H through Lyman-α absorption. These atomic absorption features are constrained to a very narrow wavelength range. The first broadband absorption features in exoplanetary atmospheres were discovered shortly after. Barman (2007) reported on the absorption by H_2O in the atmosphere of HD 209458 b, by covering a larger wavelength range than the previous atomic absorption feature analyses. Tinetti et al. (2007) used data from the *Spitzer* space telescope Infrared Array Camera (IRAC) to uncover absorption of H_2O in the atmosphere of the hot Jupiter HD 189733 b. Observations taken with the Wide Field Camera 3 (WFC3) onboard the *Hubble* Space Telescope (HST) have been a major driving force in the study of exoplanet atmospheres in transmission. WFC3 data have been used to study exoplanetary populations, focusing on the detection of H_2O absorption signatures and the influence of atmospheric clouds (Sing et al. 2016; Tsiaras et al. 2018).

With the advent of JWST, the wavelength coverage, spectral resolution, and SNR of exoplanet transmission spectra have been significantly improved. Observational capabilities of JWST extend between approximately 1 µm and 12 µm, making it more amenable for the absorption

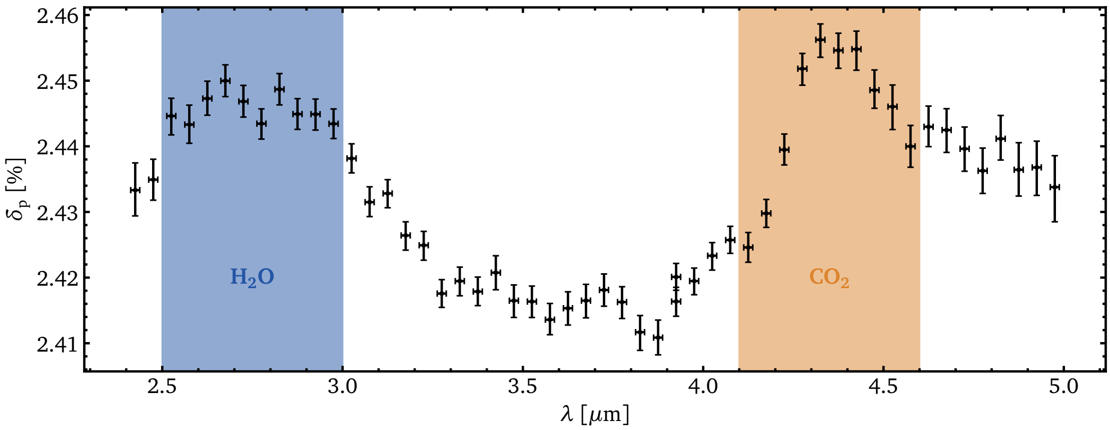

Figure 1.6: Example of an exoplanet transmission spectrum, showing wavelength (in µm, x-axis) against transit depth (in %, y-axis). The data shown here are from an observation of the hot Jupiter HD 189733 b (JWST NIRCam, Zhang et al. 2024a).

signatures of molecular ro-vibrational lines. One of the best-studied hot Jupiters observed with JWST in transmission is WASP-39 b. The absorption signature of CO_2 was first detected in its atmosphere using JWST (JTERS23). Its chemical composition has since been characterised by absorption signatures from Na, K, H_2O, CO and SO_2 (Ahrer et al. 2023; Alderson et al. 2023; Feinstein et al. 2023; Rustamkulov et al. 2023; Powell et al. 2024). The abundances of SO_2 inferred from these observations indicate a production by photochemical processes (Tsai et al. 2023), making this the first detection of a disequilibrium process in an exoplanet's atmosphere. This detection has also been reported in the atmospheres of other hot gas giants since then (Dyrek et al. 2024; Barat et al. 2025; Gressier et al. 2025), opening the door to detailed investigations into the chemical processes that shape giant planet compositions (Crossfield et al. 2025).

As we move down the parameter space of exoplanets shown in Figure 1.3, atmospheric characterisation becomes more difficult, even with state-of-the-art instruments. Atmospheric signals for these exoplanets become smaller, and many of them show indications of high-altitude cloud cover that mute spectral features (Brande et al. 2024). A famous example of this is GJ 1214 b. Extensive studies with both HST and JWST have repeatedly revealed a mostly flat transmission spectrum, making it difficult to determine its atmospheric properties (Berta et al. 2012; Kreidberg et al. 2014a; Ohno et al. 2025). Transitional planets such as K2-18 b and TOI-270 d have proven more amenable to characterisation efforts. Observations with HST have provided conflicting results between H_2O and CH_4 absorption features in the atmosphere of K2-18 b (Tsiaras et al. 2019; Bézard et al. 2022). JWST has been able to break this degeneracy, enabling the characterisation of bulk atmospheric constituents (such as CH_4 and CO_2) in the atmospheres of transitional planets (Madhusudhan et al. 2023; Holmberg and Madhusudhan 2024).

Characterising the atmospheres of rocky planets using transmission spectroscopy presents a hurdle currently not surmountable with state-of-the-art data. A prominent example of this is the TRAPPIST-1 system, which contains seven approximately Earth-sized exoplanets. Observations with HST and JWST have revealed flat transmission spectra for the exoplanets in this system (de Wit et al. 2018; Gressier et al. 2022; Espinoza et al. 2025; Piaulet-Ghorayeb et al. 2025), consistent with the absence of light, hydrogen-dominated envelopes around these exoplanets. While they cannot rule out tenuous secondary atmospheres, theory suggests that the exoplanets in this system would be bare rocks, stripped of their atmospheres (Van Looveren et al. 2024). Their host stars further complicate the atmospheric characterisation of rocky planets via transmission spectra. The activity of low-mass stars can mimic or mask the transmission signal of these targets (e.g. Rackham et al. 2018), leading to a high likelihood of contaminated transmission spectra (Lim et al. 2023; Radica et al. 2025).

It is worth noting that the high-quality transmission spectra obtained with JWST potentially contain more information than is extracted using one-dimensional terminator models. Firstly, efforts have been made to extract asymmetric information from the morning and evening limbs during exoplanetary transits (Espinoza et al. 2024; Murphy et al. 2024). Secondly, assumptions about the data of transmission spectra neglect the thermal emission from the exoplanet's night side, which points towards the observer during a transit (Kipping and Tinetti 2010). Lustig-Yaeger et al. (2025) recently presented a first foray into extracting thermal and chemical information about the night-side of a hot Jupiter using the infrared excess during its transit.

1.3.3 Emission spectroscopy

180° after a primary transit, an exoplanet will disappear behind its host star in an event we call a *secondary eclipse*. If we compare the flux we measure inside and outside this occultation, we can gain information about the flux coming from the exoplanet itself through the dip in the light curve (similar to what we have seen in Figure 1.2). The depth of this dip, Δ_e, encodes information about the flux contribution coming from the exoplanet's dayside, (Winn 2010)

$$\Delta_e(\lambda) = \frac{F_\mathrm{p}}{F_*}(\lambda) = \frac{I_\mathrm{p} R_\mathrm{p}^2}{I_* R_*^2}(\lambda). \tag{1.8}$$

Here, F_p and F_* represent the specific flux of the exoplanet and host star, I_p and I_* their specific intensities, R_p and R_* their radii, and λ the measured wavelength. The eclipse depth shown in Equation 1.8 encodes information about the flux from the exoplanet's dayside relative to the flux of the host star at the same wavelength. While this combines thermal emission and reflected light, emission spectroscopy at infrared wavelengths primarily measures the thermal emission of an exoplanet's atmosphere. An exoplanet emitting as a blackbody would produce a smooth curve, representative of the exoplanet's effective temperature. However, the presence of an atmosphere alters the observed flux from an exoplanet's dayside. Consequently, emission spectroscopy is sensitive to the molecular composition of the exoplanet's atmosphere and its thermal structure. We see this in Figure 1.7, which shows the eclipse spectrum of a hot Jupiter recorded with JWST. The reduction in eclipse depth centred at 4.3 µm is caused by the absorption of CO_2, meaning that the observed flux comes from higher and cooler region of the

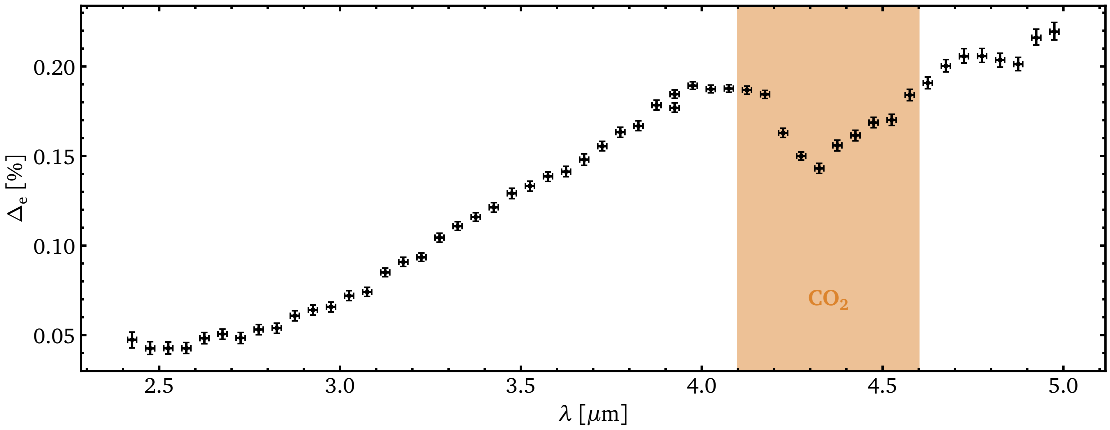

Figure 1.7: Example of an exoplanet emission spectrum, showing wavelength (in µm, x-axis) against exoplanet dayside flux ratio (in %, y-axis). The data shown here are from an observation of the hot Jupiter HD 189733 b (JWST NIRCam, Zhang et al. 2024a).

atmosphere at this wavelength (Zhang et al. 2024a). In the presence of a thermal inversion, molecular absorption features become emission features, making emission spectroscopy an ideal technique to characterise the thermal structure of an exoplanet's atmosphere.

As with transmission spectroscopy, the detection and characterisation of exoplanets based on secondary eclipses are most amenable to hot Jupiters. Charbonneau et al. (2005) reported the first photometric measurement of infrared emission from an exoplanet by measuring the secondary eclipse of the TrES-1 b. The first multiband measurements, translating to a very low resolution spectrum, were reported by Deming et al. (2005) for the hot Jupiter HD 189733 b. Emission spectroscopy performed with HST has enabled the study of the molecular inventory and thermal structure of hot Jupiters, extracting information about their atmospheric H_2O content and the presence of thermal inversions (e.g., Kreidberg et al. 2014b; Haynes et al. 2015; Changeat et al. 2022). With the advent of JWST, the broader accessible wavelength range has enabled more detailed characterisations of several hot Jupiters, determining their atmospheric metallicity (August et al. 2023; Schlawin et al. 2024; Gressier et al. 2025) and identifying the presence of CH_4 in cooler gas giants below temperatures of 1000 K (Bell et al. 2023; Wiser et al. 2025).

Characterising smaller exoplanets through emission spectroscopy becomes increasingly complex, as the flux contribution from the exoplanet's atmosphere to the total measured flux becomes vanishingly small (Mukherjee et al. 2025). Particularly hot exoplanets are among the more amenable targets. One outstanding example is the observation of the lava planet 55 Cnc e. Analyses of emission spectra taken with JWST have hinted at a possibly volatile-rich atmosphere of this exoplanet (Hu et al. 2024), but variability in the eclipse spectra makes conclusive interpretations difficult (Patel et al. 2024). The search for an atmosphere enveloping an Earth-sized exoplanet has also been a focus of emission measurements performed with JWST. Although not amenable to characterisation, observations of these exoplanets can point towards the presence or absence of atmospheres. So far, these emission photometry and spectroscopy observations have been mostly consistent with airless bodies or very tenuous atmospheres (Greene et al. 2023; Zieba et al. 2023; Weiner Mansfield et al. 2024; August et al. 2025; Wachiraphan et al. 2025). Future observational campaigns, such as the 500-hour *Rock Worlds* JWST programme (Redfield et al. 2024), will further investigate the prevalence of atmospheres around rocky exoplanets.

As with analysis of transmission spectra, more detailed analyses of the light curves during a secondary eclipse can provide spatial information about the exoplanetary atmosphere. During the ingress and egress of a secondary eclipse, the dayside of the exoplanet is partially eclipsed by its host star, allowing a mapping of its brightness distribution. This technique is commonly referred to as *eclipse mapping* (Williams et al. 2006; Rauscher et al. 2007) and is the only method of deriving both longitudinal and latitudinal information about an exoplanet's atmosphere. It has been successfully applied to derive information about the dayside temperature distribution and hot-spot position of two hot Jupiters: HD 189733 b (Wit et al. 2012; Majeau et al. 2012; Rauscher et al. 2018) and WASP-18 b (Coulombe et al. 2023; Challener et al. 2025).

Lastly, combined-light flux measurements are not the only method of observing exoplanetary atmospheres. Spectroscopic observations of directly imaged exoplanets directly probe atmospheric emission. Because they have no other observables, models used to infer the properties of these exoplanets have to account for additional unknown parameters, such as the planetary mass or radius, in the inference. Spectroscopic observations of these objects were historically restricted to ground-based instruments (Barman et al. 2011; Konopacky et al. 2013; Greenbaum et al. 2018; Zhang et al. 2024b). However, the increased sensitivity of JWST has enabled the characterisation of cooler and smaller planetary-mass companions (Miles et al. 2023; Matthews et al. 2024; Balmer et al. 2025; Voyer et al. 2025).

1.3.4 Phase curve spectroscopy

Transmission and emission spectra represent observations of exoplanetary atmospheres over fractions of their orbits, focusing on the atmospheric terminator or the dayside region, respectively. Observing the totality of an exoplanet's orbit (or *phase curve*) adds significant information to measurements, as it tracks spatial information about the atmosphere, such as day-to-night heat distribution and hot-spot offsets driven by atmospheric dynamics (Cowan et al. 2007). Figure 1.8 illustrates this, showing the broad-band phase curve of a hot Jupiter recorded with JWST, centred on the exoplanet's primary transit. The long-term sinusoidal trend of this observation encodes the contribution of the exoplanet's dayside to the measured flux. The peak of this curve is shifted compared to the time of the exoplanet's eclipse, indicating that the hottest part of its atmosphere is longitudinally offset from the sub-stellar point (Bell et al. 2024).

Observing phase curves and inferring atmospheric properties from them is non-trivial. Compared to transmission or emission spectroscopy, observations of a complete orbital phase take significant time investment and rely on long-term instrument stability (Knutson et al. 2012).

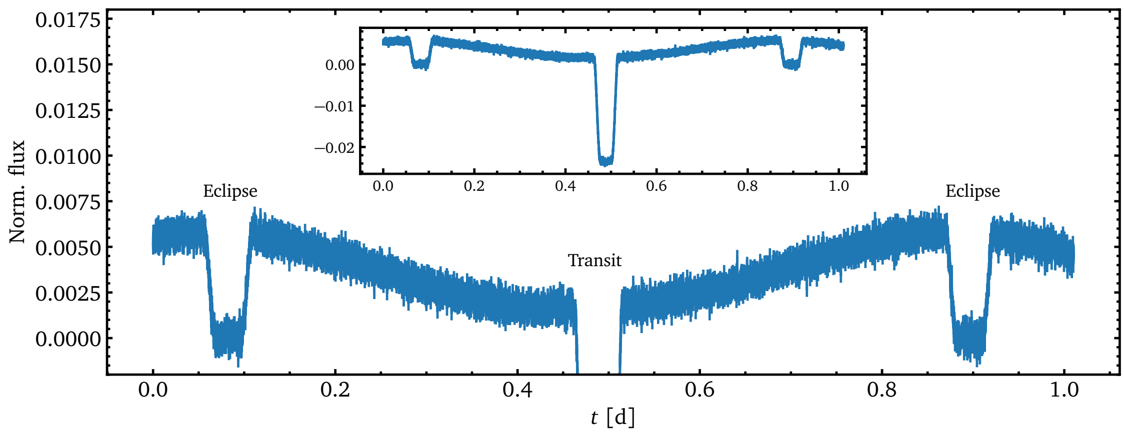

Figure 1.8: Example of an exoplanet phase curve showing time (in d, x-axis) against normalised flux (relative to the stellar flux, y-axis). The data shown here are from an observation of the hot Jupiter WASP-43 b (MIRI LRS, Bell et al. 2024).

With HST, phase curves of a few amenable hot Jupiters have been observed (Stevenson et al. 2014; Kreidberg et al. 2018; Arcangeli et al. 2019; Mikal-Evans et al. 2022), resolving diurnal variations in temperature and atmospheric H_2O abundances. Advances with JWST have enabled the phase-curve characterisation of smaller exoplanets, such as the hazy and high-MMW atmosphere of the transitional planet GJ 1214 b (Kempton et al. 2023) or the atmosphere-less Earth-sized planet GJ 367 b (Zhang et al. 2024c).

Traditional analyses of phase-curve data employ one-dimensional atmospheric retrieval approaches to recover atmospheric characteristics at individualised points in the orbital phase (e.g, Stevenson et al. 2014; Kreidberg et al. 2018). More recently, novel approaches have been developed to perform unified retrievals that properly couple the atmospheric model across multiple dimensions, including the spatial (e.g., Irwin et al. 2020; Chubb and Min 2022) or wavelength domain (Changeat et al. 2024).

1.4 Research questions and thesis outline

With the recent launch of JWST and the upcoming launch of the Ariel space mission, the science yield from observations of extrasolar planets is poised to increase by orders of magnitude. This revolution in data quality inescapably urges us to re-evaluate the methodologies we use to derive atmospheric properties in the context of state-of-the-art observational data. In this work, we will examine how this leap in data quality affects two fundamental pillars of atmospheric retrieval – the parameter-defining atmospheric forward model and the inference-guiding observational data. Chapter 2 provides an overview of the methods used in this thesis. This encompasses the instruments of JWST and how we perform data reduction, the atmospheric radiative transfer model and its components, and the statistical framework we use to infer atmospheric properties. Using these methods, we will evaluate the state-of-the-art methodology to characterise exoplanet atmospheres from transmission spectra by addressing the following questions:

How strong can model assumptions bias the results of atmospheric retrievals?
The parametrisation of the forward model is a central pillar of atmospheric retrievals. In the SNR regime of exoplanet atmospheric observations, the information contained in the spectra can be limited. Overly complex models might not be supported by the data, while underparameterised forward models can bias atmospheric retrieval results. Rocchetto et al. (2016) showed that an isothermal p–T profile can bias atmopsheric retrieval results of transmission spectra. They considered idealised observational scenarios, simulating multi-instrument JWST transmission spectra of a cloud-free hot Jupiter. Since then, observational data from JWST have become available, demonstrating that combining datasets from different instruments is nontrivial. In Chapter 3, we will investigate whether the biases reported by Rocchetto et al. (2016) translate into single-instrument, single-transit observations with realistic noise levels, and whether they are uniform across all instruments. To investigate this, we will examine atmospheric retrievals of simulated hot-Jupiter transmission spectra with realistic observational noise and assess the impact of using p–T profiles of varying complexity on the reported retrieval parameter estimates.

How stable are atmospheric retrieval results against perturbations of the underlying data?
The underlying observational data guide the parameter estimation process in atmospheric retrievals. By definition, Bayesian inference frameworks assume these data are 'correct'. However, analyses of JWST data have shown that using different data reduction pipelines can affect the extracted atmospheric spectrum (e.g., Ahrer et al. 2023; Coulombe et al. 2023). Differences in these pipeline-coupled spectra propagate into the posterior distributions of forward model parameters (e.g., Powell et al. 2024; Davenport et al. 2025; Schmidt et al. 2025). In Chapter 4, we will investigate how these differences manifest on the level of an individual data reduction pipeline. We will identify differences in a transmission spectrum derived from the same raw data using the same data reduction pipeline but with different assumptions in individual reduction steps. We will perform end-to-end data reduction of a transit observation performed with JWST, deriving the transmission spectrum of a hot Jupiter. By scattering this spectrum, we will investigate how stable atmospheric retrieval results are under uncorrelated data perturbations. We will contrast this with a stability analysis of results we derive from already existing transmission spectra of the same exoplanet. These were derived from the same raw data using the same pipeline, allowing us to quantify the impact of pipeline-level differences in the data on atmospheric retrieval results.

Chapter 5 summarises the findings of this thesis. We will provide answers to the research questions posed above and consider the broader implications that they have for atmospheric characterisation efforts in the era of JWST and Ariel.

CHAPTER 2

METHODOLOGIES

"*For a moment, nothing happened.*
Then, after a second or so, nothing continued to happen."

Douglas Adams

The Hitchhiker's Guide to the Galaxy

This Chapter provides an overview of the methods used throughout this work. Section 2.1 describes the instrument used in this work, the observational data that comes from it, and how raw observational data is reduced into finished data products. Section 2.2 covers the methodology used to characterise exoplanet atmospheres from these observations, including a description of the radiative transfer calculations used to derive transmission spectra from an atmospheric model and the atmospheric retrieval framework used in this work. Section 2.3 discusses how to merge the observational data and spectral models using the statistical framework of Bayesian inference.

2.1 The data: Transmission spectra

The work presented here focuses on the atmospheric characterisation of hot Jupiters using their transmission spectra. The source of these spectra is the *James Webb* Space Telescope (JWST, Gardner et al. 2023), a space-based observatory launched on 25 December 2021 in orbit around the Sun-Earth L2 point. Before the launch of JWST, exoplanet atmospheric spectroscopy from space was driven by two other observatories: HST and the *Spitzer* Space Telescope (Spitzer). With the launch of JWST, these observations have entered a new era in terms of wavelength coverage, spectral resolution, and sensitivity.

2.1.1 The *James Webb* Space Telescope

JWST is a 6.5 m infrared (IR) space telescope jointly developed between the National Aeronautics and Space Administration (NASA), the European Space Agency (ESA), and the Canadian Space Agency (CSA). Initially conceived as the *Next Generation Space Telescope* (HST & Beyond Committee et al. 1996), it represents the culmination of approximately 30 years of development and serves as a successor to HST. Compared with the previous state of the art, JWST has radically improved the quality of exoplanetary atmospheric observations. We see this in Figure 2.1, which compares transmission spectra of a hot Jupiter taken with HST and Spitzer to observations of the same exoplanet with JWST. In the top panel, we see transmission spectra from HST's Space Telescope Imaging Spectrograph (STIS) and Wide Field Camera 3 (WFC3). These transmission spectra represent a combination of three and two transits, respectively (Sing et al. 2016), and each cover only narrow wavelength ranges of $\Delta\lambda \approx 0.7\,\mu\mathrm{m}$. In comparison to that, the JWST NIRISS transmission spectrum is the result of a single transit (Feinstein et al. 2023), covering a wavelength regime about twice as wide ($\Delta\lambda \approx 2\,\mu\mathrm{m}$). The improved data quality of JWST is even more apparent when we look at wavelengths longward of $3\,\mu\mathrm{m}$. In the bottom panel of Figure 2.1, We see observations taken with the $3.6\,\mu\mathrm{m}$ and $4.5\,\mu\mathrm{m}$ channels of Spitzer's Infrared Array Camera (IRAC) instrument (Sing et al. 2016). In this wavelength regime, they represent less an atmospheric spectrum, and more individual photometric radii. Compared to that, the JWST NIRSpec observation (Alderson et al. 2023) spanning $\Delta\lambda \approx 2.5\,\mu\mathrm{m}$ already allows a visible identification of atmospheric absorption features.

These examples show the capabilities of two of JWST's instruments in observing exoplanet

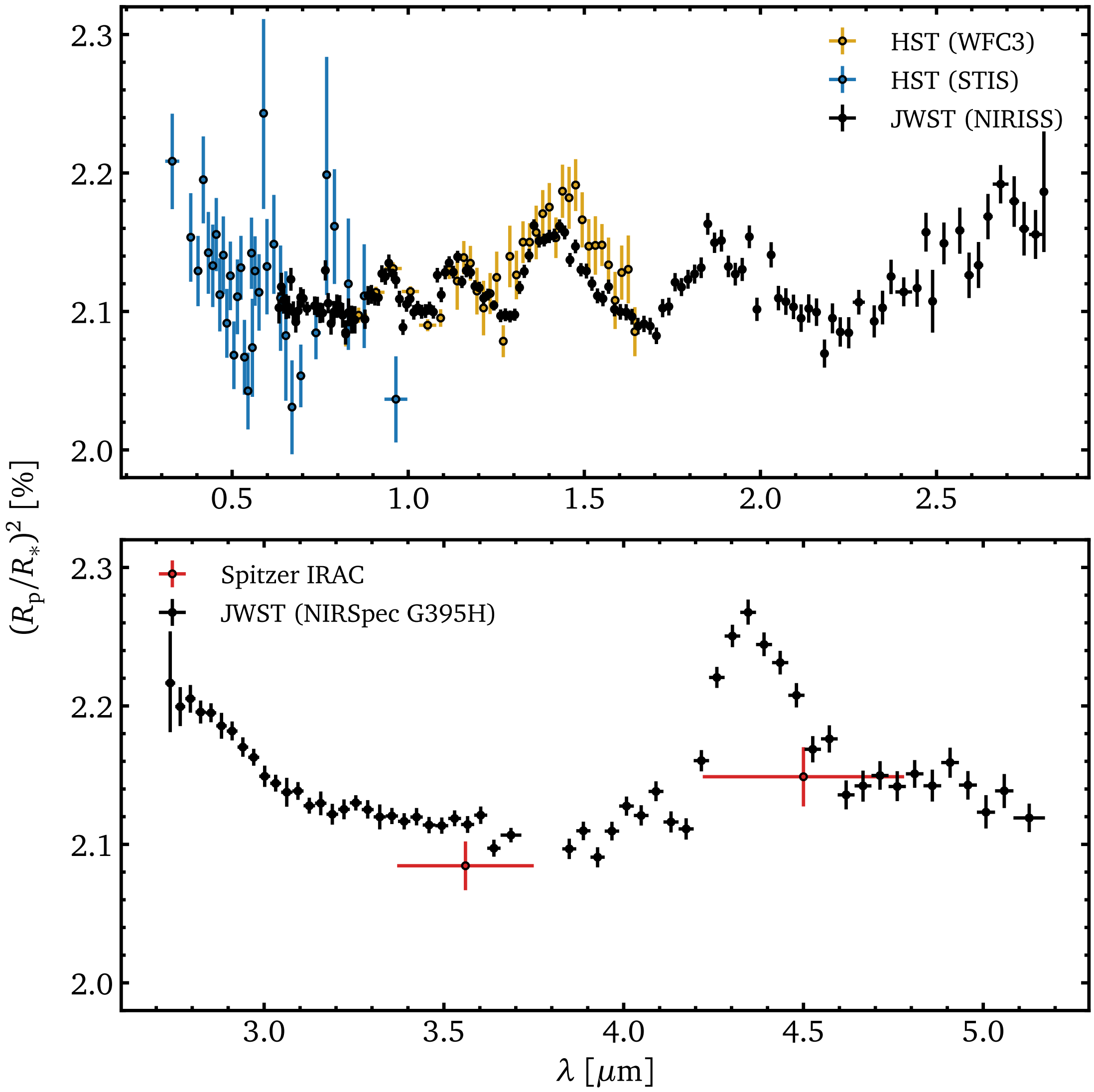

Figure 2.1: Illustration of transmission spectra of the hot Jupiter WASP-39 b. Both panels show wavelength (in µm, x-axis) against transit depth (in %, y-axis). (Upper panel) A comparison between transmission spectra taken by HST and JWST. Results from HST STIS (Sing et al. 2016) are shown in blue, from HST WFC3 (Sing et al. 2016) in yellow. Results from JWST NIRISS presented in Carter & May et al. (2024) are shown in black. (Lower panel) A comparison between transmission spectra taken by Spitzer and JWST. Spitzer IRAC results (Sing et al. 2016) are shown in red. Results from JWST NIRSpec (F290LP/G395H) presented in Carter & May et al. (2024) are shown in black. The data were retrieved from the Atmospheric Spectroscopy Table (`https://doi.org/10.26133/NEA36`).

atmospheres through primary transits. In total, JWST carries four IR instruments, each providing imaging and spectroscopic capabilities. Together, they have 17 operating modes encompassing both static and time-series observation (TSO) capabilities (Gardner et al. 2023). Three of the instruments (NIRCam, NIRISS, and NIRSpec) operate in the near-infrared (NIR), and one of them (MIRI) operates in the mid-infrared (MIR). This results in a total operational wavelength range from 0.6 µm to 28 µm. JWST is a multi-purpose observatory serving the astronomical scientific community from cosmology and early-universe observations to Solar

System physics. The focus of the work I discuss here is on observations obtained with NIRSpec, but all instruments of JWST have been successfully used to study the atmospheres of exoplanets.

As part of the early release science (ERS) programme to investigate JWST's capabilities of studying transiting exoplanets, it observed the hot Jupiter WASP-39 b (Faedi et al. 2011) with all NIR instruments. Results from these analyses have directly emphasised the leap in data quality from previous state-of-the-art observatories. One of the most important findings from these early investigations was the identification of SO_2 in the atmosphere of WASP-39 b (Rustamkulov et al. 2022; Alderson et al. 2023). This was the first observational indicator of photochemical processes altering the composition of an exoplanet's atmosphere away from thermochemical equilibrium (Tsai et al. 2023). A follow-up transit observation in the MIR with JWST confirmed this finding (Powell et al. 2024), making WASP-39 b one of the most well-studied exoplanets with JWST. As each of its instruments covers different wavelength regimes at varying spectral resolutions, its capabilities enable the study of exoplanetary atmospheres in unprecedented detail.

NIRCam

The *Near-InfraRed Camera* (NIRCam) instrument performs observations with several broad-, medium-, and narrow-band filters, split into a short-wavelength (SW) channel from 0.6 µm to 2.3 µm, and a long-wavelength (LW) channel from 2.4 µm to 5.0 µm (Horner and Rieke 2004; Rieke et al. 2023). NIRCam has five observational modes and was primarily designed to perform imaging (Gardner et al. 2023). It is one of two instruments onboard JWST used to directly image exoplanets with its coronographic capabilities (the other is MIRI). To spectroscopically observe exoplanet transits, it uses a *grism time series* (GTS) mode in its LW channel. This results in spectra between 2.4 µm and 5.0 µm (depending on the implemented filter) at a spectral resolution of $R \approx 1\,500$ (Greene et al. 2017). LW-channel spectroscopy is paired with SW-channel observations. These are either photometric or in the form of low-resolution spectroscopy at $R \approx 300$ (Schlawin et al. 2016; Rieke et al. 2023). The ERS transmission spectrum of WASP-39 b observed with NIRCam primarily covers a broad absorption feature by H_2O centred at approximately 2.8 µm (Ahrer et al. 2023).

NIRCam has a K-band sensitivity threshold of 4.0 mag to 4.5 mag (Rieke et al. 2023), enabling transit observations of the brightest host stars among the instruments of JWST. This has allowed observations of the lava planet 55 Cnc e (Hu et al. 2024), the young, Neptune-sized AU Mic b (Implementation Cycle 4; PID: 5311; PI: A. Feinstein) and the archetypical hot Jupiter HD 189733 b (Zhang et al. 2024b). The host stars in these systems have K-band magnitudes of 4.02 mag, 4.53 mag and 5.54 mag, respectively (Cutri et al. 2003), making them brightest host stars observed with JWST to date.

NIRISS

The *Near-InfraRed Imager and Slitless Spectrograph* (NIRISS) instrument operates in four observational modes, primarily designed for spectroscopy. It also has redundant imaging capabilities that overlap with those of NIRCam (Doyon et al. 2012; 2023). The *single-object slitless spectroscopy* (SOSS) mode is specifically designed to perform spectroscopic observations of exoplanetary transits, with a J-band sensitivity threshold of 6.3 mag and a spectral resolution of $R \approx 700$ (Doyon et al. 2023). The grism used for SOSS produces three orders of cross-dispersed spectra covering 0.6 µm to 2.83 µm. The first (0.84 µm to 2.83 µm) and second (0.6 µm to 1.4 µm) spectral orders are marginally overlapping due to technical constraints, making spectral extraction non-trivial, while the third order is not meant for science yield, with most of its transmission blueward of 0.6 µm (Albert et al. 2023).

The broad wavelength coverage of NIRISS encompasses major absorption features from H_2O and CH_4 in its first order, as well as from K in its second order. During commissioning, NIRISS observed the hot Jupiter HAT-P-14 b, tentatively identifying absorption by H_2O in its atmosphere, together with a high cloud deck (Albert et al. 2023; Liu et al. 2025). As part of the ERS, NIRISS has produced a transmission spectrum of WASP-39 b. This spectrum is also illustrated in the upper panel of Figure 2.1 (upper panel). Feinstein et al. (2023) reported clear identifications of absorption features by H_2O and K in the atmosphere of WASP-39 b from this observation. NIRISS has also been used to disentangle the degeneracy between absorption features from H_2O and CH_4 in the sub-Neptune K2-18 b (Benneke et al. 2019; Blain et al. 2021; Bézard et al. 2022). Together with complementary observations from NIRSpec, absorption by CH_4 has clearly been identified in its atmosphere (Madhusudhan et al. 2023; Schmidt et al. 2025).

NIRSpec

The *Near-InfraRed Spectrograph* (NIRSpec) is designed to perform both single-object and multi-object spectroscopy in 4 observational modes, covering wavelengths from 0.6 µm to 5.5 µm (Jakobsen et al. 2022; Böker et al. 2023). It observes exoplanet transits in the *bright object time series* (BOTS) mode using a fixed slit together with a variety of dispersion and filter element combinations, as described in Birkmann et al. (2022). Using a prism as the dispersion element in combination with a clear filter, it can perform low-resolution ($R \approx 100$) spectroscopy from 0.6 µm to 5.5 µm with a high J-band sensitivity threshold of approximately 10.5 mag, making the observation of bright host stars difficult. NIRSpec BOTS also offers combinations of medium-resolution ($R \approx 1000$) or high-resolution ($R \approx 3000$) gratings with filters to cover parts of this total wavelength range. This decreases the J-band sensitivity threshold to below 6 mag for some combinations (Gardner et al. 2023). NIRSpec disperses its spectra onto two detectors (labelled NRS1 and NRS2). When using the prism or medium-resolution gratings, these spectra are dispersed solely onto NRS 1, while observations taken with the high-resolution gratings are dispersed over both detectors, creating a gap of approximately 0.1 µm in the spectrum (Birkmann et al. 2022).

NIRSpec has been used extensively to observe transiting exoplanets. Around 40% of observational programmes during the first four cycles of JWST make use of one of its observational modes (TrExoLiSTS 2026). The NIRSpec prism mode provides the widest continuous wavelength coverage of any NIR instrument modes available for JWST. As part of the ERS, WASP-39 b was observed in this mode, deliberately achieving saturation below approximately 2.2 µm to improve the SNR of the observation. This data has led to the first detection of CO_2 in the atmosphere of an exoplanet (JTERS23). Rustamkulov et al. (2023) presented an investigation into retrieving information from the saturated part of the detector, providing a continuous NIR spectrum constraining several atmospheric trace species, including Na, H_2O, CO_2, and SO_2.

The combination of the high-resolution G395H grating and F290LP filter is another observational mode extensively used to observe exoplanet transits with JWST. It covers the long-wavelength wing of the broad H_2O absorption feature centred at 2.8 µm, as well as significant absorption features of methane at 3.4 µm and of CO_2 at 4.2 µm. Alderson et al. (2023) presented an analysis of a WASP-39 b transmission spectrum taken in this observational mode, identifying absorption by H_2O, CO_2, and SO_2. The G395H/F290LP mode also complements observations with NIRISS. HAT-P-14 b (Liu et al. 2025) and K2-18 b (Madhusudhan et al. 2023; Schmidt et al. 2025) are examples of exoplanetary atmospheres observed with both modes. This provides continuous wavelength coverage comparable to that of NIRSpec prism, but with a lower sensitivity threshold (allowing for observations of brighter host stars) and higher spectral resolution. A drawback of this is the uncertainty in detector-level (for NRS1 and NRS2) and instrument-level (between NIRISS and NIRSpec) offsets between individually observed transmission spectra (e.g., Madhusudhan et al. 2023; Carter & May et al. (2024)).

MIRI

The *Mid-InfraRed Instrument* (MIRI) is the only instrument onboard JWST that performs observations in the MIR, in a wavelength range from 5 µm to 28.5 µm (Rieke et al. 2015; Wright et al. 2023). It has four operational modes, encompassing imaging (including coronography) and spectroscopy, and is the only instrument for which all observational modes have been used to observe exoplanet atmospheres. MIRI's coronography mode allows it to directly image exoplanets, complementing the capabilities of NIRCam in the NIR.

A primary driver of exoplanet observations with MIRI is the *low-resolution spectroscopy* (LRS) mode, which operates with a fixed slit or in a slitless mode (Kendrew et al. 2015). LRS provides spectra with a resolution of $R \approx 100$ in a wavelength range of 5 µm to 12 µm, and has been used to observe transmission and emission spectra, as well as exoplanet phase curves. After the extensive ERS observational campaign of WASP-39 b found evidence of the absorption of SO_2 in its atmosphere, a MIRI LRS observation was used to further confirm the presence of SO_2 through absorption features at 7.7 µm and 8.5 µm (Powell et al. 2024). LRS has been used to study a wide range of exoplanets in transmission, emission and through phase curves, including hot Jupiters (Dyrek et al. 2024; Inglis et al. 2024), sub-Neptunes (Kempton et al. 2023; Madhusudhan et al. 2025), and rocky planets (Alam et al. 2024; Tusay et al. 2025). The *medium-resolution spectroscopy* (MRS) capabilities of MIRI have been used to observe eclipses of HD 189733 b (PID: 1633; PI: D. Demming) and 55 Cnc e (PID: 7875; PI: M. Zhang). However, there are no published results of these observational programmes as of the writing of this thesis.

The photometric imaging capabilities of MIRI have also been used to study exoplanet atmospheres, specifically in the search for atmospheres around rocky planets, using the F1280W and F1500W filters centred at 12.8 µm and 15 µm, respectively. Imaging in the F1500W filter covers a strong absorption band of CO_2, which is expected to be one of the dominant atmospheric constituents for rocky planets (Lichtenberg and Miguel 2025). A comparison with measurements in the F1280W filter (outside the CO_2 features) allows a determination of the atmospheric pressure or CO_2 abundance (Deming et al. 2009; Ih et al. 2023). This has been achieved for the hot rocky planet TRAPPIST-1 b, where MIRI observations point towards the absence of an atmosphere (Ducrot et al. 2024). The *Rock Worlds* Director's Discretionary Time (DDT) programme has committed 500 additional observational hours to continue the work of observing eclipses of rocky exoplanets with the F1500W filter to find indications of CO_2 absorption (Redfield et al. 2024).

2.1.2 Data reduction

The raw observational data from JWST are stored in the Mikulski Archive for Space Telescopes (MAST) for access by the scientific community. In addition to that, the `jwst` pipeline (Bushouse et al. 2024) is used to provide automatically reduced and calibrated data products. To further fine-tune and process these data products, many frameworks are being developed to handle exoplanet TSO data from JWST, with varying approaches and levels of accessibility. The `JExoREs` pipeline (Holmberg and Madhusudhan 2023) builds on the automatically reduced data products retrievable from the MAST as `*.calints` files. Data reduction frameworks such as `tshirt` (Glidic et al. 2022), `FIREFLy` (Rustamkulov et al. 2022), `tiberius` (Kirk et al. 2017; Kirk et al. 2021), and the work presented in Voyer et al. (2025) incorporate and modify the low-level data reduction stages performed by the `jwst` pipeline. The `SPARTA` pipeline (Kempton et al. 2023; Xue et al. 2025) performs data reduction fully independent of the `jwst` pipeline. Irrespective of the underlying approach, the data reduction of JWST observations relies on reference files (such as wavelength maps and bias frames) managed through the calibration reference data system (CRDS)[1], which maps sets of reference files into a unique *context* (`*.pmap`).

In this work, we use `Eureka!`[2] (Bell et al. 2022), an end-to-end data reduction pipeline wrapping around `jwst`, as well as performing independent light-curve extraction and fitting. The stages of `Eureka!`, including the correction, reduction, and extraction steps of exoplanet atmospheric observations, are what we will examine next.

[1] `https://jwst-crds.stsci.edu/`
[2] Version `0.10.dev0+g3c10926.d20230426`

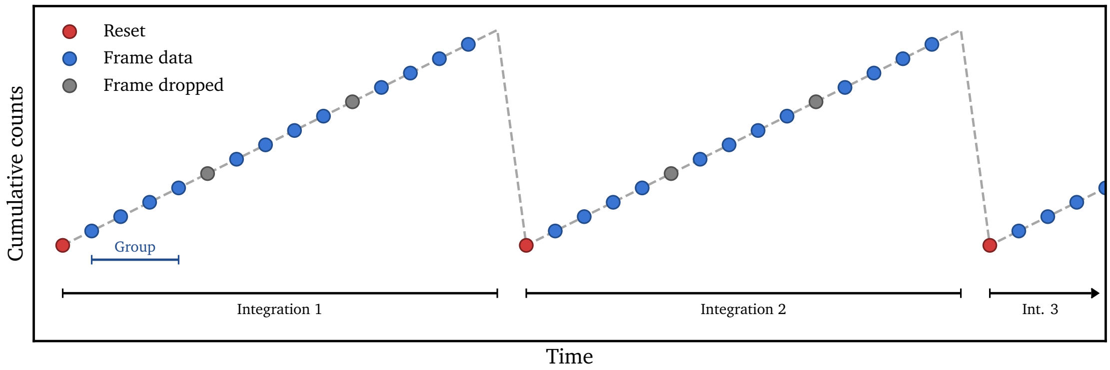

Figure 2.2: Illustration of JWST up-the-ramp readout pattern. A series of integrations makes up one exposure with JWST.

Data structure of JWST observations

To understand the reduction process of raw observational data taken with JWST, it is useful to review the structure of these data products (JDox 2026). At the highest level, data recorded by JWST is separated into *exposures*. While this is a generic term used in astronomy, in the context of JWST it specifically refers to the pixel readout during a single 'expose' command. The data structure within a single exposure is divided into multiple smaller units, illustrated in Figure 2.2. A single exposure consists of a series of *integrations*, separated by detector resets. These integrations are read out in an *up-the-ramp* scheme, subdivided into individual *groups* of *frames*. The number of frames per group, and if frames get dropped between groups, depends on the readout pattern of the instrument selected for a given exposure. The lowest level of data interacted with in the data reduction process is group-level data. For TSOs, the data reduction and calibration process transforms the group-level ramps into signal rates for each integration. In the context of this work, these signal rates represent the high-cadence brightness monitoring required to observe exoplanetary transits in detail.

Stages of `Eureka!`

The execution of `Eureka!` is split into several stages, corresponding to pipeline stages of `jwst` and light-curve data analysis. We see a flowchart of these stages in Figure 2.3, which illustrates the data products generated by each step, along with the associated dimensions where relevant. The summary of the individual stages below is based on the official documentation of `jwst` (JWST pipeline documentation 2025) and `Eureka! v0.10` (Eureka! documentation 2023). It is worth pointing out that, while broadly following the same steps and ideas, there are variations in the execution of these steps for the individual instruments of JWST (the biggest being between the reduction of MIRI data and data from the NIR instruments). The descriptions below default to the reduction of data from NIRSpec.

Stage 1: Detector processing

The first stage of `Eureka!` acts primarily as a wrapper for the `jwst` pipeline. Baseline

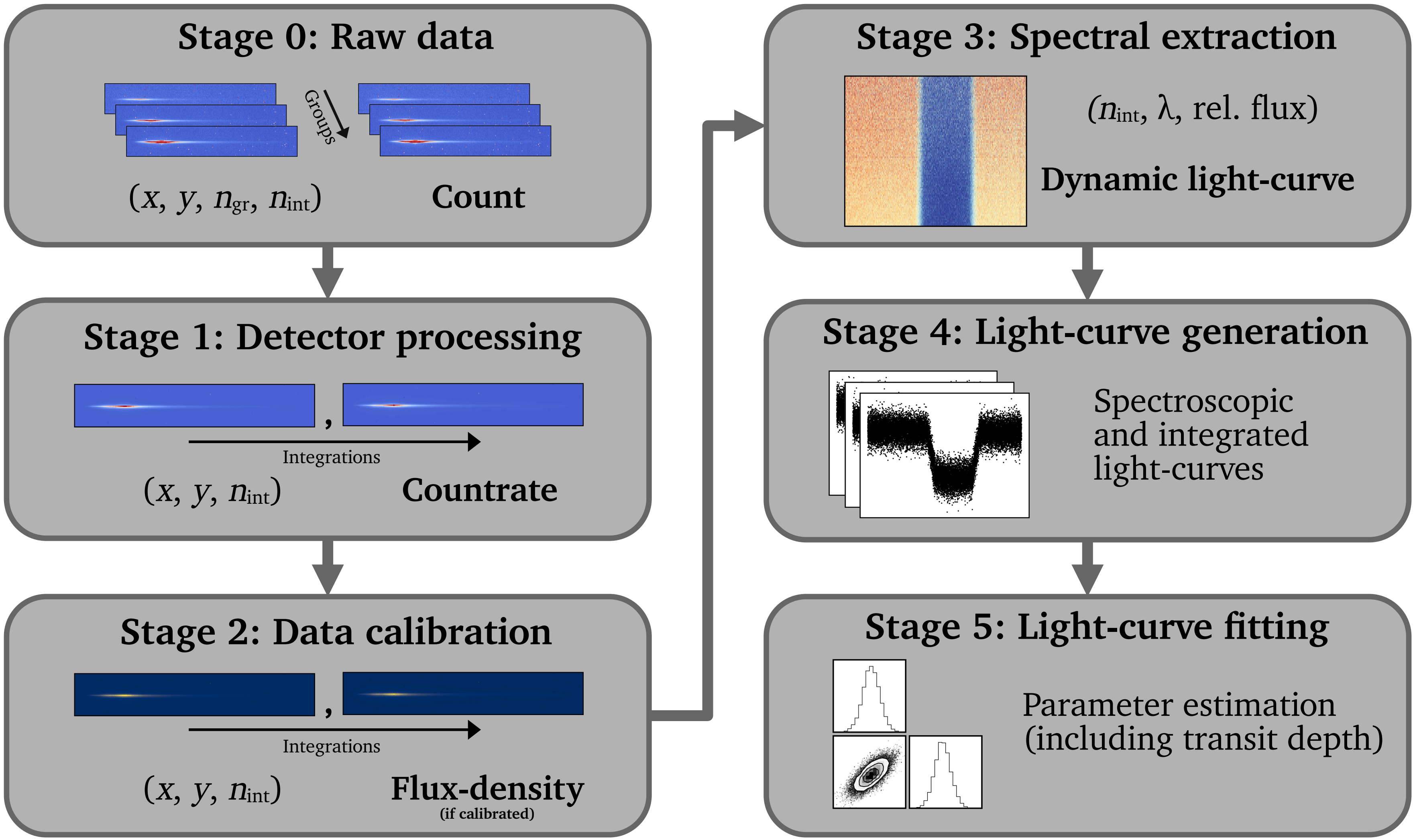

Figure 2.3: Flow chart of JWST data reduction. Each panel represents a stage in the data reduction process. Data dimensions and associated units after the completion of the stage are provided where relevant. x and y refer to the detector image pixel directions, n_{gr} and n_{int} to the number of groups and integrations, respectively, and λ to the wavelength.

initialisation steps are applied first, including the data quality mask. Pixels are checked for saturation based on the full-well capacity of the 16-bit detectors ($2^{16} = 65\,536$ counts). A superbias frame is subtracted from every group in this step (based on a specific CRDS context), removing fixed detector biases from the science data.

The detector-level processing of this stage applies a *reference pixel correction* step. This is based on the fact that the readout electronics impose a signal at every readout on each pixel. Because the readout electronics consists of multiple amplifiers, this signal can vary with each readout amplifier (Rauscher et al. 2022). The reference pixel correction uses physical reference pixels at the edges of each instrument's detector. The mean signal of the reference pixels associated with each amplifier is subtracted from the pixels it is representative of. The data are then corrected for the non-linearity of the detector response. As the signal on the detector increases, the detector's response to incoming photons (converting them into electrons) becomes nonlinear. Pre-calculated polynomials for each pixel of the detector correct for this, transforming the signal back into a linear ramp (Robberto 2011). After this, the dark current is subtracted from the data using CRDS-context-specific dark frames.

A crucial step in the first stage of data reduction is the detection of jumps in the ramp-level raw data. These are usually caused by cosmic-ray impacts. Jumps in the up-the-ramp data are flagged based on the two-point difference method described in Anderson and Gordon (2011), following

$$\frac{|d_i - \mathrm{med}(d)|}{\sigma_i} > r. \tag{2.1}$$

Here, d_i represents the distance between the group-level values y_i and y_{i+1} for a pixel, $\mathrm{med}(d)$ is the median of all d_i for each pixel, σ_i is the expected noise for d_i, and r is a rejection threshold value. If a jump is detected, the corresponding pixel is removed, and the step is repeated until no more jumps are identified. This step can only be performed on data with three or more groups, as it requires at least two distances. If this is not the case, the jump rejection step is skipped during data reduction.

Observations by JWST also suffer from additional readout amplifier-induced noise that temporally follows a 1/f power spectrum. This manifests (visually) in stripping patterns imprinted in the fast-readout direction of the detector. As this behaviour changes on a group-by-group level, the 1/f noise must be removed before ramp fitting collapses the group data (Rauscher 2015; JDox 2026). In `Eureka!`, this is addressed through group-level background subtraction (GLBS). This step performs a background-subtraction routine at the group level, masking the source and determining the background level through the median level or a polynomial fit.

At the end of the first stage, the group-level data are transformed into a count rate (counts per second) for each pixel. Each detector ramp is linearly fitted using a least-squares algorithm with 'optimal weighting', as described in Fixsen et al. (2000). A last step rescales the science data if a non-standard gain setting was used in the observational setup. With this, the raw four-dimensional data $(x, y, n_{\text{gr}}, n_{\text{int}})$ have been collapsed into a three-dimensional data cube of (x, y) detector images for each integration.

Stage 2: Data calibration

Stage 2 begins with a step transforming the positions from detector frame coordinates to a world coordinate system, and maybe more importantly, assigns an indicator to the direction of the dispersion axis - either horizontally or vertically aligned with the image frame. Subsequently, the two-dimensional data are correlated with a wavelength map, and a source type (defaulting to a point source for TSOs) is assigned to the observation for further processing. The data are also flat-field corrected at this stage. Flat-fielding involves dividing the science frame by a flat-field image to correct for the non-uniform detector response to a uniform illumination source. It is worth noting that the order of the first few steps (after transforming the coordinate systems) depends on the observing mode.

Stage 2 also includes a photometric calibration step, which allows a transformation of units. For point sources, the count-rate data resulting from stage 1 can be converted into flux-density values. It is noteworthy that this step is unnecessary for relative measurements such as transmission or emission spectroscopy. In these cases, we care about relative changes in the signal, which does not require us to calibrate the data away from counts as a unit. Lastly, stage 2 includes a spectral-extraction step that can be applied to retrieve source spectra for each integration. We will do this in the next stage and do not need to perform it here. By the end of stage 2, we have produced a wavelength-calibrated three-dimensional dataset containing two-dimensional detector images of pixels (λ, y) for each integration.

Stage 3: Spectral extraction

With stage 3, `Eureka!` leaves the `jwst` pipeline framework. This stage is primarily concerned with extracting the spectral signal from the calibrated integration frames. For spectral extraction, we need to specify a signal and background region in the cross-dispersion direction, y, of the detector. The signal region is centred on the source point spread function (PSF), which is determined through a Gaussian fit. Stage 3 performs outlier rejection in the temporal direction using a double-iteration σ-threshold routine. Column-by-column background subtraction can be performed in this stage by specifying a polynomial of degree n to fit to the background region determined in this stage.

This leaves a background-free, calibrated and corrected detector image containing pixels with coordinates (λ, y). Each of these pixels has a signal, $f_{\lambda,y}$, and variance, $\sigma_{\lambda,y}$, associated with it. A standard spectral extraction routine might produce an object's wavelength-dependent spectral signal, s_λ, as

$$s_\lambda = \sum_y f_{\lambda,y}. \tag{2.2}$$

This represents the sum of the pixel-based signal over all pixels in the extraction aperture at a specific wavelength. The variance of this extraction is the sum of the individual, pixel-level uncertainties, or

$$\sigma_\lambda^2 = \sum_y \sigma_{\lambda,y}^2. \tag{2.3}$$

The spectral extraction routine implemented in `Eureka!` is based on the 'optimal spectra extraction' routine described in Horne (1986). It modifies Equation 2.2 by incorporating the wavelength-dependent spatial profile of the source, P_λ. In `Eureka!`, the spatial profile can be constructed using a variety of methods, including a polynomial or Gaussian fit, or through a median over all integrations. Through a positivity ($P_{\lambda,y} \geqslant 0 \, \forall \, y$) and normalisation ($\sum_y P_{\lambda,y} = 1$) requirement, the spatial profile provides a way of assigning a contribution factor to each pixel in the extraction aperture to the total signal. This spectral signal is then extracted following

$$s_{\lambda,\mathrm{opt}} = \frac{\sum_y \frac{f_{\lambda,y}}{\sigma_{\lambda,y}^2} \cdot P_{\lambda,y}}{\sum_y w_{\lambda,y}}, \tag{2.4}$$

where the weight, $w_{\lambda,y}$, is dependent on the pixel-level variance and spectral profile through

$$w_{\lambda,y} = \frac{P_{\lambda,y}^2}{\sigma_{\lambda,y}^2}. \tag{2.5}$$

The variance of the optimal spectral extraction is subsequently given by the inverse of the weight as

$$\sigma_{\mathrm{opt},\lambda}^2 = \sum_y w_{\lambda,y}^{-1}. \tag{2.6}$$

The construction of the spatial profile and optimal spectral extraction rely on the pixels included in the extraction aperture. These two processes are applied iteratively, checking for outliers through a rejection threshold, r_t, using the inequality

$$\frac{|f_{\lambda,y} - s_\lambda P_{\lambda,y}|}{\sigma_{\lambda,y}} > r_t, \tag{2.7}$$

until no new pixels are rejected between two iterations. The optimal spectral extraction routine achieves a better SNR over the standard routine by assigning pixels far from the peak of the spatial profile less weight (Horne 1986). At its end, stage 3 of `Eureka!` has produced a time-series of wavelength-dependent signals for our transit observation.

Stage 4: Light-curve generation

Stage 4 of `Eureka!` extracts the time-series data crucial to transmission spectroscopy. Light-curves are extracted in two ways. Firstly, spectroscopic light-curves are extracted by binning the data along the dispersion direction into N spectroscopic channels. At the highest-resolution level, N can be equal to the number of pixels in the dispersion direction of the detector. In addition, a broadband (also called white) light-curve is created by integrating over the wavelength domain, producing a photometric transit observation. The creation of light-curves is associated with another iteration of outlier rejection along the time axis. This can be performed on both the unbinned data and, if chosen, on the λ-binned light-curves. The outlier rejection is performed using a rolling median box filter with a user-specified width and rejection threshold. At this stage, `Eureka!` can record spectroscopic drift and pointing jitter, and optionally correct for a wavelength drift through spectral cross-correlation.

The white light-curve can provide information about the non-spectroscopic parameters of the exoplanet-star pair. It summarises a photometric transit depth and other relevant information. The spectroscopic light-curves are the most essential data products of this stage. In the following step, `Eureka!` performs model fits to these light-curves to ultimately determine the depth of the light-curve at the mid-transit time. This directly translates to the spectroscopic radius of the exoplanet. Arranging these radii with respect to the observed wavelength gives us the characteristic transmission spectrum of an exoplanet. At the end of stage 4, we have therefore arrived at a wavelength-dependent time-series dataset ready to be characterised by a transit model.

Stage 5: Light-curve fitting

The fifth stage of `Eureka!` is now disconnected from any data reduction steps. It is concerned with parameter estimation through fitting the extracted light-curve data. The general fitting routine combines astrophysical and systematic effects to form a light-curve model. Bayesian inference is used to move through the parameter space of the model using a sampling algorithm. By design, `Eureka!` is highly modular in this stage. It allows defining a wide variety of expressions to fit the light-curve data.

For this work, the astrophysical transit model is based on the `batman` Python package (Kreidberg 2015). We can summarise the necessary parameters of `batman` in the vector $\boldsymbol{p}_{\text{transit}} = (P, t_0, i, a, e, \omega, \delta)$. Here, P represents the orbital period of the exoplanet, t_0 the time of the transit mid-point, i the orbital inclination, a the semi-major axis, e the orbital eccentricity, ω the argument of periapsis, and δ the transit depth. These parameters are used to calculate the observed flux during transit (Mandel and Agol 2002; Kreidberg 2015),

$$F(t; \boldsymbol{p}_{\text{transit}}) = 1 - \iint\limits_{S(t;\boldsymbol{p}_{\text{transit}})} I(\mu)\, dS. \tag{2.8}$$

Here, $F(t; \boldsymbol{p}_{\text{transit}})$ represents the time-dependent reduction in measured flux during the transit, relative to an out-of-transit baseline, $S(t; \boldsymbol{p}_{\text{transit}})$ the area of the stellar disk obscured by the transiting exoplanet, and $I(\mu)$ the sky-projected intensity of the star. This intensity plays an important role in the transit model. It is usually assumed to be radially symmetric, but not uniform. The radial variation of the intensity over the stellar disk is characterised by *limb* *darkening*, describing its dependence on the angle between the line of sight and the emergent intensity, θ, often summarised in $\mu = \cos\theta$. Eureka! can incorporate the characterisation of this intensity variation, $I(\mu)$, in a variety of functional forms. Notable among those are the commonly used quadratic limb-darkening law described in Kipping (2013), and the multi-parameter laws introduced in Claret (2000).

Observations of exoplanetary transits are not free of systematics. The astrophysical model defined through batman is insufficient to capture the overall behaviour of the time-series signal. We can account for systematics stemming from the instrument itself by modifying the astrophysical model. Commonly, this includes ramping effects materialising as long-term temporal trends. Eureka! can account for these systematic trends by incorporating polynomials and exponential functions in time. It also allows the inclusion of Gaussian processes and fit parameters to account for changes in the PSF of the source (such as mirror-tilt events, Alderson et al. 2023). Lastly, Eureka! also allows for the inclusion of a photometric scatter parameter, which represents additional scatter necessary to bring the joint model fit to a reduced-χ^2 value of unity.

The astrophysical and systematic models are combined into a total model representing the signal behaviour with respect to time. This total model is fit to the light-curve data produced in stage 4. A sampling algorithm, such as Markov Chain Monte Carlo (MCMC, Metropolis et al. 1953; Hastings 1970) or Nested Sampling (NS, Skilling 2006), moves through the parameter space of the joint model, providing parameter estimates at the end of the fitting process. As this process closely overlaps with the inference framework of atmospheric retrieval, I refer to Section 2.3 for a more detailed overview.

Stage 6: Exoplanet spectral extraction

The final stage of Eureka! acts primarily as a clean-up and transformation stage. It takes the parameter estimates from the prior stage and transforms them into more user-friendly outputs. No further modifications are made to the fitted data here. At the end of the Eureka! pipeline, we have produced an exoplanet's transmission spectrum in the form of its wavelength-dependent radius (an example of which we have seen in Figure 2.1).

2.2 The forward model: 1D radiative transfer

With the observational data in place, we now want to consider how we can build a model that represents this data. Our goal is to create a model that can accurately simulate a transmission spectrum observed by JWST, accounting for the interaction of light with the exoplanet's atmosphere. Repeated evaluations of this model will allow us to estimate its parameter values that best represent our observed data. To make this iterative process fast and flexible, we perform our calculations in one dimension.

2.2.1 Basic concepts of radiative transfer in atmospheres

Our model needs to characterise the interaction of radiation with the atmosphere of an exoplanet. In the following section, we define the basic concepts of these interactions, labelled under the umbrella term of *radiative transfer*. This follows the explanations in Goody and Yung (1995) and Liou (2002).

Radiative quantities

First and foremost, it will be beneficial to define the radiative quantities relevant to us. We declare *spectral intensity* as

$$I_\lambda = \frac{dE_\lambda}{\cos(\theta)\, dA\, d\Omega\, dt\, d\lambda}, \tag{2.9}$$

in units of $\mathrm{J\,m^{-2}\,sr^{-1}\,s^{-1}\,m^{-1}}$. Equation 2.9 defines I_λ as the differential amount of radiative energy, dE_λ, in a specified wavelength interval, λ to $\lambda + d\lambda$, which crosses an area element, dA, in a time interval, dt, and in directions constrained to a solid angle, $d\Omega$, oriented to dA by an angle, θ. The spectral intensity is a directional quantity, usually confined to a 'pencil' (or beam) of radiation, illustrated in Figure 2.4. Marginalising (or integrating) the spectral intensity over the entire (hemispheric) solid angle removes the directionality of this quantity. This integral results in the *spectral flux density*,

$$F_\lambda = \int_\Omega I_\nu \cos(\theta) d\Omega = \frac{dE_\lambda}{dA\, dt\, d\lambda}, \tag{2.10}$$

in units of $\mathrm{J\,m^{-2}\,s^{-1}\,m^{-1}}$. Furthermore, if we consider the *spectral flux density* across all wavelengths, we arrive at the *(total) flux density*,

$$F = \int_0^\infty F_\lambda\, d\lambda = \frac{dE_\lambda}{dA\, dt}, \tag{2.11}$$

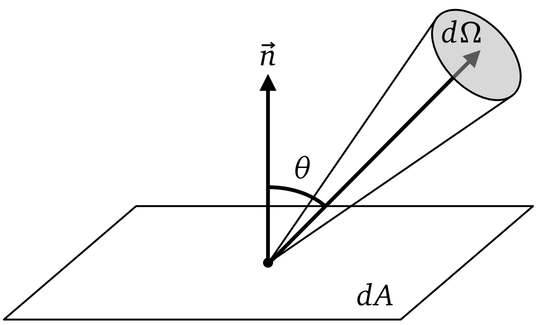

Figure 2.4: Illustration of a pencil of radiation, confined to a solid angle, $d\Omega$, moving through an area element, dA, at an angle, θ, oriented to the area normal vector, $\vec{n}$.

in units of $\mathrm{J\,m^{-2}\,s^{-1}}$, and if we consider the flux density from a total area, we derive the *(total) flux*,

$$f = \int_A F\,dA = \frac{dE_\lambda}{dt}. \tag{2.12}$$

in units of $\mathrm{J\,s^{-1}}$. These are the most essential quantities that we consider in the transfer of radiation, or radiative energy, through a medium. They are often used interchangeably, but aspects like the directionality of I_λ and wavelength-marginalisation of F are key distinctions between them.

Another relevant term in the context of radiative transfer is the concept of *optical depth*. It characterises how opaque a medium is to radiation. In a medium with mass density ρ, a given mass extinction coefficient k_λ, and a depth s, we can define the optical depth as

$$\tau(\lambda, s) = \int_0^s \rho\, k_\lambda\, d\hat{s}. \tag{2.13}$$

The equation of radiative transfer

Radiative transfer describes how the specific intensity of a beam of radiation, I_λ, is changed to $I_\lambda + dI_\lambda$ after traversing a medium. In this process, energy might be removed or added to the beam. If it is removed, we call the process *extinction*, and if it is added, we call the process *emission*. Therefore, I denote the total change in spectral intensity as

$$dI_\lambda = dI_{\lambda,\mathrm{ext.}} + dI_{\lambda,\mathrm{emi.}}, \tag{2.14}$$

We first want to look at the intensity changes by extinction processes. Two processes are primarily used to characterise extinction: (1) *absorption*, where the radiative energy is absorbed by the material and converted into another form; and (2) *scattering*, where the radiative energy is specifically diverted out of the pencil of radiation by particles after it was absorbed. We define $dI_{\lambda,\mathrm{ext.}}$ through

$$dI_{\lambda,\mathrm{ext.}} = -k_\lambda \rho I_\lambda\, ds, \tag{2.15}$$

where ρ denotes the mass density of the material, k_λ the mass extinction cross section (combining extinction from absorption and scattering, in units of $\mathrm{m^2\,kg^{-1}}$), and ds represents the travelled path. The negative sign arises because the intensity is attenuated (energy is removed), whereas all other properties of the equation are positive.

Opposite to this, the intensity might be intensified by (1) *emission* from the material it passes through, and (2) *scattering* from other directions into the pencil of radiation. Following the same approach as for attenuation, the increase in intensity is given by

$$dI_{\lambda,\mathrm{emi.}} = j_\lambda \rho\, ds, \tag{2.16}$$

where we define a source function coefficient, j_λ, with the same meaning as the mass extinction cross section, k_λ. It has the units of $\mathrm{J\,kg^{-1}\,s^{-1}\,sr^{-1}\,m^{-1}}$. We can then express the total alteration of the intensity of the radiation pencil as a combination of Equations 2.15 and 2.16,

$$dI_\lambda = (j_\lambda - k_\lambda I_\lambda)\rho\, ds. \tag{2.17}$$

It is sometimes convenient to define a source function as $J_\lambda := j_\lambda / k_\lambda$. This then allows us to rearrange Equation 2.17 into the *general equation of radiative transfer*,

$$\frac{1}{k_\lambda \rho}\frac{dI_\lambda}{ds} = J_\lambda - I_\lambda. \tag{2.18}$$

It is also worth pointing out that in a non-scattering medium in local thermodynamic equilibrium (LTE), the source function, J_λ is given by Planck's function, $B_\lambda(T)$, where T denotes the temperature of the medium. If we substitute this into Equation 2.18, we arrive at *Schwarzschild's equation*,

$$\frac{1}{k_\lambda}\frac{dI_\lambda}{ds} = B_\lambda(T) - I_\lambda \tag{2.19}$$

Beer-Bougher-Lambert law

Now we will consider a simplifying case – one where we do not care about contributions from emission or diffuse scattering. In this case, the equation of radiative transfer simplifies to

$$\frac{dI_\lambda}{k_\lambda \rho\, ds} = -I_\lambda. \tag{2.20}$$

This is a simple differential equation. Suppose we are now interested in the attenuation of the intensity along a path of length s. In that case, the solution to Equation 2.20, under the assumption that the intensity at the beginning of the path is $I_\lambda(0)$, is

$$I_\lambda(s) = I_\lambda(0)\exp\left(-\int_0^s k_\lambda \rho\, ds\right). \tag{2.21}$$

If we assume that the medium along the path is homogeneous, the mass extinction cross section, k_λ, is constant. We can recognise that the integral in Equation 2.21 is the optical depth we have defined in Equation 2.13. With this, we arrive at the *Beer-Bouguer-Lambert law*,

$$I_\lambda(s) = I_\lambda(0)\exp\left(-\tau_\lambda\right). \tag{2.22}$$

This relation is the central equation in the technique of transmission spectroscopy, where we are concerned with the change in spectral intensity, I_λ, of stellar light as it passes through the atmosphere of an exoplanet on its way to an observer. It states that this change follows a simple exponential decrease, dependent on the nature and length of the medium it traverses. As it does not depend on any directional quantities, Equation 2.22 also applies to the flux-density and total flux.

Lastly, we can rearange Equation 2.22 to arrive at the definition of *monochromatic transmittance*, $\mathcal{T}$,

$$\mathcal{T}_\lambda = \frac{I_\lambda(s)}{I_\lambda(0)} = \exp\left(-\tau_\lambda\right). \tag{2.23}$$

This is a conveniently normalised number that tells us how much incident radiant intensity is transmitted through a medium.

2.2.2 TauREx

To perform radiative transfer calculations through an atmospheric model and compare the results to observations, I use the fully-Bayesian retrieval framework `TauREx 3` (Al-Refaie et al. 2021; 2022, referred to as `TauREx` henceforth)[3]. It allows us to set up an atmospheric forward model, defined by the physical properties of the star-exoplanet system and the exoplanet's atmosphere, and then iteratively perform radiative transfer calculations through this model to compare it to observational data.

Atmospheric forward model setup

Atmospheres are three-dimensional objects, but computational requirements in the inference process generally limit us to simplified models. In this work, we will use one-dimensional atmospheric columns as a representative stand-in for the total atmosphere. The base of this model is a reference exoplanetary body. It is defined by its mass, M_p, enclosed in its radius, R_0. We can assume this radius corresponds to the solid surface of a rocky planet or a fully opaque, high-pressure layer of a gas giant. On top of this reference radius sits the atmospheric model domain. It is bounded by p_BOA and p_TOA, pressure values that represent the top of the atmosphere (TOA) and bottom of the atmosphere (BOA), respectively. The model domain between these is divided into a fixed number of layers, each with associated local properties. These local properties define the altitude profile of the atmospheric model domain through

$$\begin{aligned} z_0 &= 0 \\ z_l &= z_{l-1} + \Delta z_l \\ \Delta z_l &= -H_{l-1} \log\left(\frac{p_l}{p_{l-1}}\right) \end{aligned} \tag{2.24}$$

In these relations, l represents the index of the atmospheric model layer and z_l the associated altitude. This altitude is calculated with respect to R_0 (hence $z_0 = 0$). The change in altitude from one layer to another, Δz_l, then depends on the scale height of the previous layer, H_{l-1}, and the pressure of layer l, p_l, and of the previous layer, p_{l-1}. The local scale height is defined

[3] Version 3.2.4

through

$$H_l = \frac{k_B\, T_l}{\mu_l\, g_l}, \tag{2.25}$$

where k_B, T_l, μ_l, and g_l are the Boltzmann constant, as well as the temperature, MMW, and acceleration due to gravity of layer l, respectively. These parameters form the skeleton of our atmospheric model.

Throughout this work, we will use consistent ways to characterise key vertical structure components of the forward model. One of these is the temperature structure of the atmosphere, encoded in a p–T profile. For this work, we will define this profile through a multipoint, or n-point, approach. A multipoint profile consists of n p–T pairs, (p_n, T_n), forming a vertical temperature profile using a smoothing window (with a default size of 10 layers) to avoid discontinuities. If $n = 1$, this is equivalent to a vertically homogeneous (os isothermal) temperature profile. Another important part of the forward model is the vertical chemical structure. In this work, we will use constant profiles defined through a mixing ratio (MR). For an atmospheric gas-phase constitutent m (such as H_2O), the mixing ratio is defined as

$$X_m = \frac{n_m}{n_{\text{tot}}}, \tag{2.26}$$

where n_m and n_{tot} represent the number density of the constituent and the total number density, respectively. TauREx uses the ideal gas law to calculate the total number density,

$$n_{l,\text{tot}} = \frac{p_l}{k_B\, T_l}, \tag{2.27}$$

for each layer of the forward model domain. The MRs are required to sum to unity. Outside of the individually defined atmospheric constituents, TauREx therefore also relies on *fill-gases*, defined by ratios between each other. In gas giant exoplanets, these are commonly assumed to be H_2 and He. Lastly, we want to consider the presence of atmospheric aerosols. Clouds and hazes can impact radiative transfer through both absorption and scattering processes. In this work, we will consider a simplified characterisation of aerosols through a flat-opacity cloud deck. This is implemented as an absorption step function in the atmospheric forward model.

These parameters define our forward model, providing an environment for radiative transfer calculations within this domain. In the next section, we want to take a look at how we produce observables from these calculations to relate them to our measurement data.

Effective planetary radius

When we observe an exoplanet's atmosphere through spectroscopically measuring its primary transit, the key quantity we are interested in is its wavelength-dependent radius, $R_\text{p}(\lambda)$. To fit observables in radiative transfer models, we make use of a concept called the *effective planetary radius*. The derivation I present here follows explanations in de Wit and Seager (2013).

We start from an observational assumption – when we observe the primary transit of exoplanets, the apparent radius of the planet we measure may change with respect to the observed wavelength. We have seen in Equation 1.7 that the measured radius of an exoplanet varies with wavelength due to the interaction of the observed stellar radiation with the exoplanet's atmosphere. We now impose that the wavelength-dependent radius of the exoplanet, $R_\mathrm{p}(\lambda)$, can be expressed with two components,

$$R_\mathrm{p}(\lambda) = R_0 + r(\lambda). \tag{2.28}$$

Here, R_0 represents a wavelength-independent reference radius at which the exoplanet is opaque in all wavelengths, and $r(\lambda)$ represents a wavelength-dependent attenuation term. The reference radius might be, for example, the solid part of a rocky planet, or the high-pressure interior of a gas giant. Using this, we express the wavelength-dependent transit depth measured during a primary transit, $\delta(\lambda)$, as

$$\delta(\lambda) = \frac{[R_0 + r(\lambda)]^2}{R_*^2} = \frac{R_0^2 + \overbrace{2\,r(\lambda)R_0 + r(\lambda)^2}^{=a(\lambda)}}{R_*^2} = \frac{R_0^2 + a(\lambda)}{R_*^2}, \tag{2.29}$$

where R_* represents the radius of the host star, and we have collected all wavelength-dependent factors into an attenuation term, $a(\lambda)$, that modifies the reference size, R_0, of the exoplanet. Physically, we know that the attenuation of stellar flux during an exoplanet's transit traces back to the radiative transfer through the exoplanet's atmosphere. Since Equation 2.23 describes the intensity transmitted through a medium, we know that the reduction of the intensity follows $1 - \mathcal{T} = 1 - \exp(-\tau(\lambda, r))$. Here, we consider that the ray of light moves through a planetary atmosphere at a distance, r, from the centre of the body, being attenuated by the optical depth, $\tau(\lambda, r)$, of the atmosphere. In this geometry, illustrated in Figure 2.5 (panel A), the optical depth (following from Equation 2.13) is calculated as

$$\tau(\lambda, r) = 2\int_0^\infty \rho(\hat{r})\,k_\lambda(\hat{r})\,dx \quad \text{with} \quad \hat{r} = \sqrt{r^2 + x^2}. \tag{2.30}$$

The factor 2 in this Equation accounts for the assumption of a symmetric path in the approach and recession to the point $r = \hat{r}$.

Now we make a crucial equivalence: the reduced intensity recorded by an observer can be described by two scenarios. Firstly, the physically accurate description of radiation from the host star passing through an infinitely expanded atmosphere, which gradually loses opacity with distance from the centre of the body. Secondly, a solidly opaque disk with a fixed radius. From the point of view of an observer that cannot spatially resolve the atmosphere of the exoplanet passing in front of the host star, these two assumptions produce the same amount of

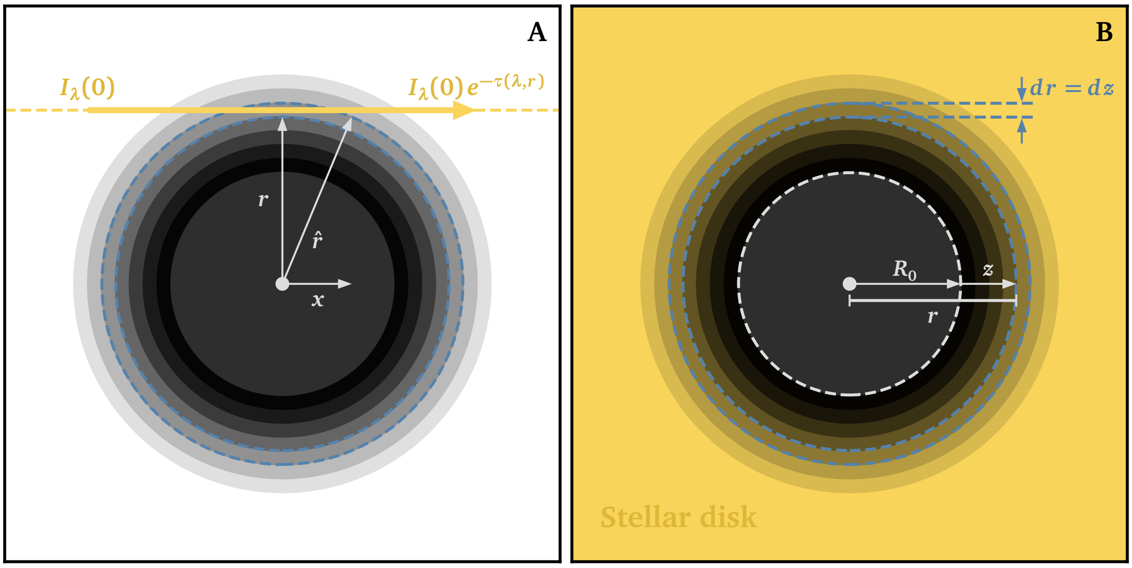

Figure 2.5: Illustration of radiative transfer through an exoplanet's atmosphere. (Left) Side view, with the host star on the left and the observer on the right of the panel. (Right) Front view, with the host star behind the exoplanet.

intensity reduction. We can therefore set up the equation

$$\underbrace{\pi R_p^2(\lambda)}_{\text{solid disk}} = \underbrace{\int_0^\infty 2\pi r [1 - \exp(-\tau(\lambda, r))]\, dr}_{\text{diffuse atmosphere}}. \tag{2.31}$$

We solve this integral by splitting it into two parts. The distance from the centre of the body, r, can be expressed through the sum of the reference radius, R_0, and the altitude above it, z. This is also shown in Figure 2.5 (panel B). We therefore make the following adjustments to Equation 2.31:

$$\begin{aligned} r(z) = R_0 + z \quad &\rightarrow \quad z(r) = r - R_0, \\ \frac{dr}{dz} = 1 \quad &\rightarrow \quad dr = dz. \end{aligned}$$

We can substitute the variable in Equation 2.31 and adjust the integration boundaries from $(0, \infty)$ to $(-R_0, \infty)$, resulting in

$$\begin{aligned} \pi R_p^2(\lambda) &= 2\pi \int_{-R_0}^{\infty} (R_0 + z) [1 - \exp(-\tau(\lambda, z))]\, dz \\ &= 2\pi \left[\underbrace{\int_{-R_0}^{0} (R_0 + z) [1 - \exp(-\tau(\lambda, z))]\, dz}_{A} + \underbrace{\int_0^\infty (R_0 + z) [1 - \exp(-\tau(\lambda, z))]\, dz}_{B} \right]. \end{aligned} \tag{2.32}$$

We solve part A of this split integral first. The integration is influenced by the assumption we made about the optical depth,

$$\tau(\lambda, z) = \begin{cases} \infty \text{ for } z \leqslant 0 \\ \tau(\lambda, z) \text{ otherwise,} \end{cases} \tag{2.33}$$

or, in other words, the exoplanetary body is fully opaque at any radius, r, below the reference radius, R_0. This allows us to solve the integral A straightforwardly:

$$\begin{aligned} A &= \int_{-R_0}^{0} (R_0 + z)[1 - \overbrace{\exp(-\tau(\lambda, z))}^{=0}]\, dz = \int_{-R_0}^{0} (R_0 + z)\, dz \\ &= \left[R_0 z + \frac{z^2}{2} \right]_{-R_0}^{0} = R_0^2 - \frac{R_0^2}{2} \\ A &= \frac{R_0^2}{2} \end{aligned} \tag{2.34}$$

We plug this back into Equation 2.32, giving us

$$\begin{aligned} \pi R_p^2(\lambda) &= 2\pi \left[\frac{R_0^2}{2} + \int_0^{\infty} (R_0 + z)\, [1 - \exp(-\tau(\lambda, z))]\, dz \right], \\ R_p^2(\lambda) &= R_0^2 + 2 \int_0^{\infty} (R_0 + z)\, [1 - \exp(-\tau(\lambda, z))]\, dz. \end{aligned} \tag{2.35}$$

As a final step, we compare Equation 2.35 with Equation 2.29 and find an expression for the attenuation term, $a(\lambda)$,

$$a(\lambda) = 2 \int_0^{\infty} (R_0 + z)[1 - \exp(-\tau(\lambda, z))]\, dz. \tag{2.36}$$

This results in the central relation of chromatic planetary radius behaviour used in atmospheric retrievals. In the next section, we will look at how our assumptions in creating the atmospheric forward models inform the calculation of this attenuation term.

Optical depth calculation

Combining Equation 2.29 and Equation 2.36,

$$\delta(\lambda) = \frac{R_0^2 + 2\int_0^{\infty} (R_0 + z)[1 - \exp(-\tau(\lambda, z))]\, dz}{R_*^2}, \tag{2.37}$$

provides us with several parameters that are key to calculating the transit depth of our model atmosphere. The previously discussed reference radius of the exoplanet, R_0, as well as the radius of the host star, R_*, define a wavelength-independent baseline transit depth. The atmospheric altitude profile, encoded in z, plays a significant role because it defines the domain in which our radiative transfer calculations are performed. From Equation 2.24 and Equation 2.25, we see that it is strongly affected by the temperature, T, and MMW, μ. One of the defining parts determining the chromatic transit depth is the calculation of the atmospheric optical depth. In a heterogeneous atmosphere, we can generalise its calculation by a combination of multiple contributions,

$$\tau(\lambda, z) = \sum_i \tau_i(\lambda, z). \tag{2.38}$$

Here, i represents the different types of processes that can contribute to the total optical depth of the atmosphere.

Molecular absorption

In the IR spectroscopy of exoplanet atmospheres, absorption by rotational and vibrational (or *ro-vibrational*) transitions of molecules represents one of the key processes shaping an atmospheric spectrum (Tennyson 2019). The optical depth contribution from absorption by an atmospheric gas-phase constituent, m, follows our already established definition in Equation 2.13. Restating it here, it reads

$$\tau_m(\lambda, z) = \int \sigma_m(\lambda, z)\, X_m(z)\, n_{\text{tot}}(z)\, dz, \tag{2.39}$$

where $X_m(z)$ and $n_{\text{tot}}(z)$ are the altitude-dependent MR and total number density profiles, respectively, and $\sigma_m(\lambda, z)$ represents the absorption cross section of the atmospheric constituent (in $\mathrm{m^2\,molecule^{-1}}$), which depends on the local pressure and temperature, encoded in the altitude profile. In Figure 2.6, we see the characteristic, wavelength-dependent absorption cross sections of several molecules often considered in the characterisation of exoplanet atmospheres. The distinct features at specific wavelengths (such as at 4.2 µm for CO_2) indicate that molecular absorption plays a key role in shaping the transmission spectra of exoplanetary atmospheres. Computing the temperature- and pressure-dependent cross sections for molecules is a computationally intensive and complex task, requiring the calculation of individual transition lines and absorption profiles. TauREx uses pre-computed cross sections at a resolution of $R = 15\,000$ (Chubb et al. 2021). These are derived from theoretically calculated and experimentally determined line-lists available databases such as the ExoMol project (Tennyson et al. 2020), and the HITRAN (Gordon et al. 2022) and HITEMP (Rothman et al. 2010) archives.

Rayleigh scattering

Rayleigh scattering describes the scattering of electromagnetic radiation by particles much smaller than the wavelength of the radiation (Liou 2002). In TauREx, the Rayleigh scattering absorption cross section for an atmospheric constituent, m, is calculated through (Keady and Kilcrease 2002; Liou 2002)

$$\sigma_m^{\text{Ray}}(\lambda) = \frac{128\,\pi^5}{\lambda^4} \cdot F_m\, \alpha_m(\eta)^2. \tag{2.40}$$

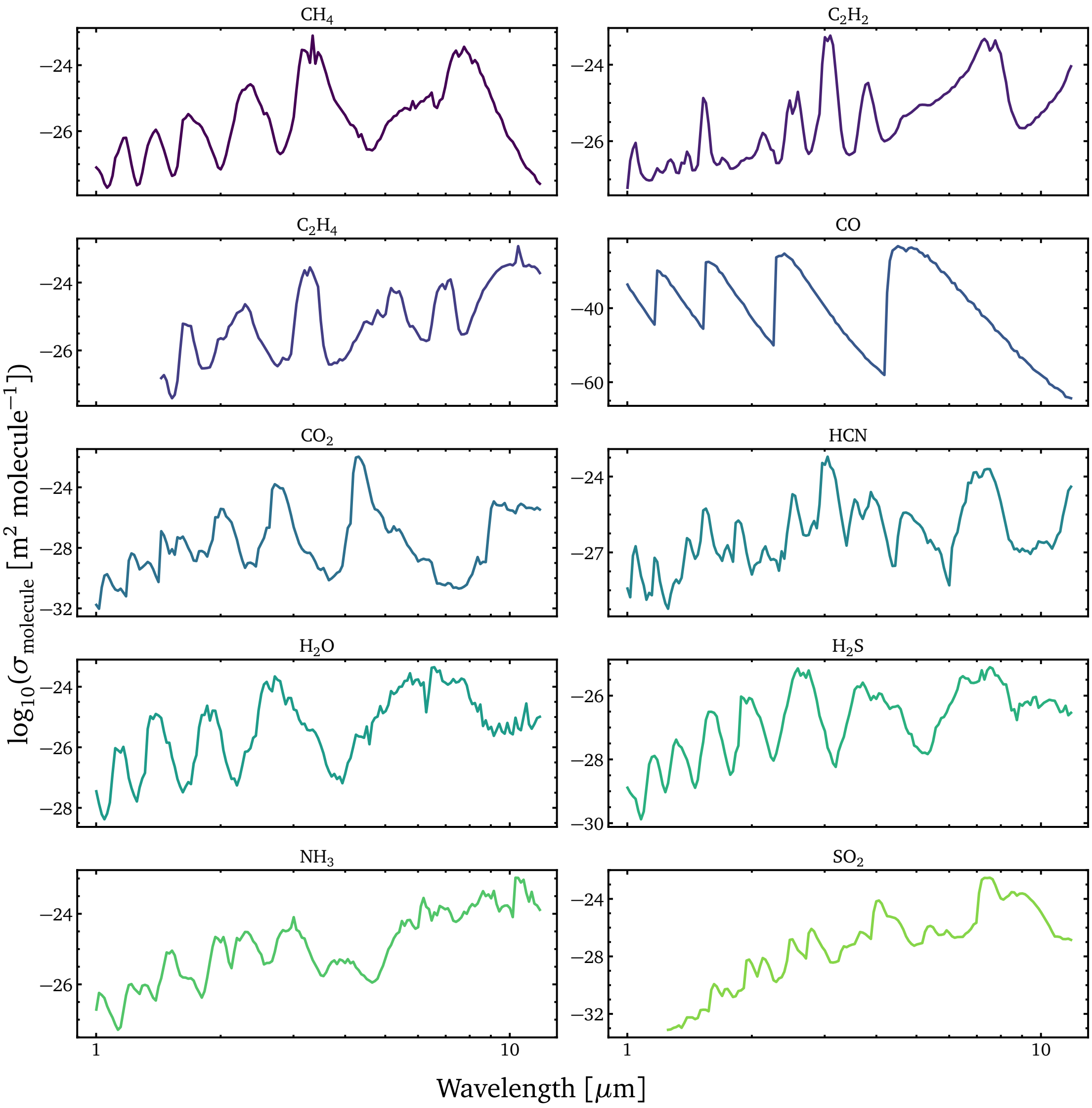

Figure 2.6: Illustration of molecular absorption cross-sections. Each panel shows wavelength (in µm, x-axis) against cross-section (in cm^2 molecule^{-1}, y-axis). The sources for each cross sections are: CH_4 (Yurchenko et al. 2017); C_2H_2 (Chubb et al. 2020); C_2H_4 (Mant et al. 2018); CO (Li et al. 2015); CO_2 (Yurchenko et al. 2020); HCN (Barber et al. 2014); H_2O (Polyansky et al. 2018); H_2S (Azzam et al. 2016); NH_3 (Coles et al. 2019); SO_2 (Underwood et al. 2016). Each cross-section is retrieved from the ExoMol database in the provided `TauREx` format (at $R = 15\,000$) for $p = 10^{-3}$ bar and $T = 1000$ K. These are binned onto 200 log-uniformly distributed wavelength points in the range $[0.6, 12.0]$ µm using the `FluxBinner`-class in `TauREx`.

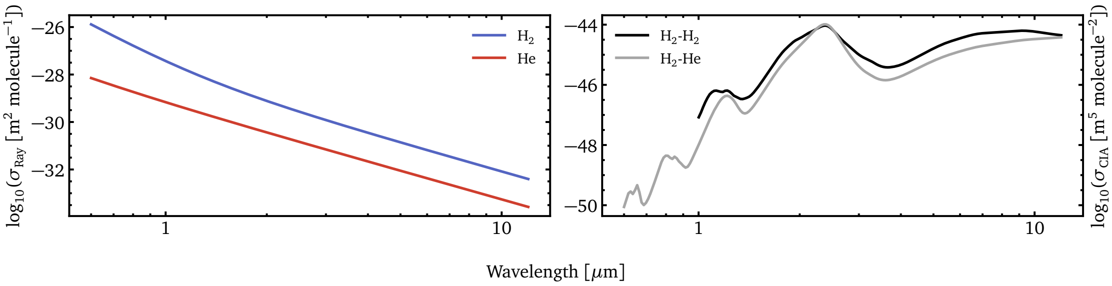

Figure 2.7: Illustration of Rayleigh scattering and CIA cross-sections from 0.6 µm to 12 µm. Each panel shows wavelength (in µm) on the x-axis. (Left) Rayleigh scattering cross sections (in m^{2}molecule^{-1}) for H_2 (blue) and He (red) as defined in the `scattering.py` routines for `TauREx`. (Right) CIA cross-section (in m^{5}molecule^{-2}) of H_2–H_2 (black, Abel et al. 2011; Fletcher et al. 2018) and H_2–He (grey, Abel et al. 2012) pairs at a temperature of $T = 1000\,\mathrm{K}$. Collision-induced absorption (CIA) from H_2–H_2 pairs is only available for $\lambda \geqslant 1\,\mu\mathrm{m}$.

where $\alpha_m(\eta)$ represents the molecular polarisability of m, which depends on its refractive index, η, and F_m represents a depolarisation correction factor. Due to the dependence of $\sigma_m^{\rm Ray}(\lambda)$ on η, these cross sections are only calculated for a select set of species. This includes the typical atmospheric background gases H_2, He, N_2, and O_2, as well as CH_4, CO_2, CO, H_2O, and NH_3 (Hollis et al. 2013). Since the Rayleigh scattering cross section only depends on wavelength and the properties of the scattering particle, its optical depth contribution takes the form

$$\tau_m^{\rm Ray}(\lambda, z) = \sigma_m^{\rm Ray}(\lambda) \int \chi_m(z)\, n_{\rm tot}(z)\, dz. \tag{2.41}$$

Due to the strong wavelength-dependence of Rayleigh scattering, its contribution to opacity is most prominent towards shorter wavelengths, where it can dominate the continuum of transmission spectra. In Figure 2.7 (left panel), We see the $\sigma_m^{\rm Ray}(\lambda)$ for H_2 and He as they are defined in `TauREx`, calculated for wavelengths from 0.6 µm to 12 µm. The typical proportionality of λ^{-4} we see in Equation 2.40 is slightly modified by particle-dependent properties.

Collsion-induced absorption

Collision-induced absorption (CIA) is a process where pairs of molecules acquire a temporary dipole during collisions, enabling them to absorb electromagnetic radiation in spectral regions in which they might be inactive in isolation (Frommhold 1994). In `TauREx`, the a contribution to the optical depth from CIA between two atmospheric constituents, m and n, is defined as

$$\tau_{mn}^{\rm CIA}(\lambda, z) = \int \sigma_{mn}^{\rm CIA}(\lambda, z)\, \chi_m(z)\chi_n(z)\, n_{\rm tot}(z)^2\, dz, \tag{2.42}$$

with constituent-specific MRs, $\chi_m(z)$ and $\chi_n(z)$, and the CIA cross section, $\sigma_{mn}^{\rm CIA}(\lambda, z)$. `TauREx` uses CIA cross sections available at the HITRAN archive, compiled from theoretical calculations and experimental data (Terragni et al. 2025). We see the CIA absorption cross-sections for H_2–H_2 and H_2–He pairs in Figure 2.7 (right panel). Its contribution to the total optical depth dominates at high pressures due to its proportionality to $n_{\rm tot}(z)^2$.

Clouds

The presence of aerosols can heavily alter a transmission spectrum compared to an equivalent, aerosol-free case (Marley et al. 2013). As reviewed in Gao et al. (2021), aerosol models in atmospheric retrievals span a wide range of complexity. Single-variable parametrisations can capture first-order effects of aerosols on an atmospheric spectrum (such as reducing the amplitude of absorption features), while more complex models offer a more physical interpretation at the cost of a larger number of parameters. In this work, we will consider a simple characterisation of clouds through a flat-opacity cloud deck. Its contribution to the total optical depth is defined through a step function,

$$\tau_{\text{cloud}}(\lambda, z) = \begin{cases} \infty & \text{if} \quad p(z) \geqslant P_{\text{cloud}}, \\ 0 & \text{otherwise.} \end{cases} \tag{2.43}$$

Here, P_{cloud} defines a pressure level considered the top of the cloud deck. The model atmosphere becomes fully opaque at pressures larger than P_{cloud}, equally for all wavelengths.

2.3 The inference: Bayesian statistics

Both the extraction of the transit depth from exoplanet TSO data and the inference of exoplanet atmospheric parameters from panchromatic transit depth values rely on a statistical framework categorised under the umbrella term *Bayesian statistics*. Statistics in the Bayesian sense treats probability as a *degree of belief*, and is centred around an equation known as *Bayes theorem* (Trotta 2008). This equation is used to both infer the parameters of individual models and to evaluate the performance of competing models. As it is central to the whole statistical framework, we will start with its derivation. This follows Trotta (2008) and Sharma (2017).

2.3.1 Bayes theorem

Let us consider two events, which I will denote with A and B. These could represent the result of a coin toss, the roll of a specific number on a six-sided die, or a measurement of data. Subsequently, we denote the probability of these two events occurring as $P(A)$ and $P(B)$, respectively. Taking this, we can define the *joint probability* of these two events occurring as $P(A, B)$. This describes a combination of event A happening, or $P(A)$, and then event B happening given that event A has happened, or $P(B|A)$. The second term is called the *conditional probability*, $P(B|A)$. The joint probability for both events is then the product of both probabilities, or

$$P(A, B) = P(A) \cdot P(B|A). \tag{2.44}$$

The joint probability $P(B,A)$ is equivalent to $P(A,B)$. Using this, we derive an expression for the conditional probability $P(B|A)$,

$$\begin{aligned} P(A,B) &= P(B,A), \\ P(A) \cdot P(B|A) &= P(B) \cdot P(A|B), \\ P(B|A) &= \frac{P(A|B) \cdot P(B)}{P(A)}. \end{aligned} \tag{2.45}$$

Equation 2.45 is commonly called *Bayes theorem*. It is central to inverse modelling processes.

We can explicitly define A and B in the context of atmospheric parameter inference from spectroscopic observations. Event A describes the observation of an exoplanet's atmosphere, labelled as a set of measurements $\boldsymbol{d} = \{d_i\}_{i \in I}$. Event B denotes the model we use to represent an exoplanet's atmosphere, $\mathcal{M}$. This model consists of assumptions we hold to be true and parameters we can freely vary. Therefore, we map event B to a set of parameters representative of the atmospheric model, or $\boldsymbol{\theta} = \{\theta_i\}_{i \in I}$. This maps the elements in Equation 2.45 to

$$\begin{aligned} P(B) &\equiv \pi(\boldsymbol{\theta}), \\ P(A|B) &\equiv \mathcal{L}(\boldsymbol{\theta}), \\ P(B|A) &\equiv P(\boldsymbol{\theta}|\boldsymbol{d}), \\ P(A) &\equiv E. \end{aligned} \tag{2.46}$$

Our prior degree of belief in the model parameters is encoded in $P(B) = \pi(\boldsymbol{\theta})$, called the *prior probability* or parameter prior. Commonly, priors for individual parameters are defined by a uniform distribution,

$$\pi(\theta_i) = \mathcal{U}(a_i, b_i) = \begin{cases} \frac{1}{b_i - a_i}, & \theta_i \in [a_i, b_i] = \Theta_i, \\ 0, & \text{otherwise} \end{cases} \tag{2.47}$$

where Θ_i is the domain of θ_i, or by a normal distribution,

$$\pi(\theta_i) = \mathcal{N}(\mu_i, \sigma_i^2) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \cdot \exp\left(-\frac{(\theta_i - \mu_i)^2}{2\sigma_i^2}\right), \tag{2.48}$$

where μ_i and σ_i^2 represent the mean and variance. Priors that define the distribution of $\ln(\theta_i)$, equivalently called *log-uniform* and *log-normal*, are also common.

We want to update our prior degree of belief in the light of the observations, forming a *posterior probability*, $P(B|A) = P(\boldsymbol{\theta}|\boldsymbol{d})$. We do this using the conditional probability, $P(A|B) = \mathcal{L}(\boldsymbol{\theta})$, which is called the *likelihood* of the data. Typically, this likelihood is defined as a normal distribution on the data,

$$\mathcal{L}(\boldsymbol{\theta}) = \prod_{i \in I} \frac{1}{\sqrt{2\pi\sigma_i^2}} \cdot \exp\left[-\frac{(d_i - m_i(\boldsymbol{\theta})^2}{2\sigma_i^2}\right], \tag{2.49}$$

where d_i is the i-th data points, σ_i^2 is its associated variance, and $m_i(\boldsymbol{\theta})$ is the associated model predicition of d_i

The last term, $P(A) = P(\boldsymbol{d}) \equiv E$ is called *evidence*, or sometimes *marginal likelihood* or *prior predictive*. It represents the probability of observing the data, accounting for all possible values of $\boldsymbol{\theta}$ under the prior. We can assume that a vector of model parameter values is represented by $\boldsymbol{\theta}^j = \{\theta_i^j\}_{i \in I}$. We know that the model posterior, as a probability, is a normalised quantity,

$$\sum_j P(\boldsymbol{\theta}^j|\boldsymbol{d}) = \sum_j \frac{\mathcal{L}(\boldsymbol{\theta}^j)\, \pi(\boldsymbol{\theta}^j)}{p(\boldsymbol{d})} = 1, \tag{2.50}$$

which provides us with an expression for the evidence,

$$p(\boldsymbol{d}) \equiv E = \sum_j \mathcal{L}(\boldsymbol{\theta}^j)\pi(\boldsymbol{\theta}^j). \tag{2.51}$$

In the case of a continuous parameter space, this is equivalent to the expression

$$E = \int_{\Theta_{\mathcal{M}}} \mathcal{L}(\boldsymbol{\theta})\, \pi(\boldsymbol{\theta})\, d\boldsymbol{\theta}, \tag{2.52}$$

where $\Theta_{\mathcal{M}} = \{\boldsymbol{\theta} : \pi(\boldsymbol{\theta}) > 0\}$ represents the support of the prior (or the integration domain). With these terms, we arrive at the commonly used equation to relate parameter posterior probabilities to the likelihood of the data,

$$P(\boldsymbol{\theta}|\boldsymbol{d}) = \frac{\mathcal{L}(\boldsymbol{\theta}) \cdot \pi(\boldsymbol{\theta})}{E}. \tag{2.53}$$

2.3.2 Moving through the parameter space

Our goal in using Bayesian inference to characterise exoplanet atmospheres is to explore $P(\boldsymbol{\theta}|\boldsymbol{d})$. In this context, we can find two major algorithmic strategies to move through this parameter space: Markov Chain Monte Carlo (MCMC, Metropolis et al. 1953; Hastings 1970) and Nested Sampling (NS, Skilling 2006). As we have seen in Section 2.1.2, MCMC can be used to fit models to light-curve data in the data reduction of JWST observations with `Eureka!`. However, this plays only a minor role in my work. We will revisit the details of using the MCMC algorithm `emcee` (Foreman-Mackey et al. 2013) in this context in Chapter 4. For comprehensive reviews of MCMC algorithms and their applications, I refer to Craiu and Rosenthal (2014) and Sharma (2017). The focus of my work is on the posterior analysis produced by `TauREx`, which is based on an NS algorithm. Below, I will therefore focus on describing the operations of NS and how it is implemented in `TauREx`. This overview is based on the work of Skilling (2006), Feroz et al. (2009), and Ashton et al. (2022).

The goal of NS is the calculation of the Bayesian evidence (Equation 2.52). The accumulation of parameter posterior samples, which is a crucial result for our planned model analysis, is a by-product of this process. NS uses a relation between the likelihood and prior of the model parameter space to rewrite this evidence integral. We can define a *prior volume*, V, as

$$V(\Lambda) = \int_{\mathcal{L}>\Lambda} \pi(\boldsymbol{\theta})\, d\boldsymbol{\theta}. \tag{2.54}$$

Here, Λ represents a likelihood threshold value and $\pi(\boldsymbol{\theta})$ the prior on the parameter space. Equation 2.54 defines the prior probability mass enclosing a region of the parameter space where $\mathcal{L} > \Lambda$. $V(\Lambda)$ is a monotonically decreasing function of Λ in the domain $V(\Lambda) \in [0,1]$. We can use the inverse of Equation 2.54 to express the likelihood as a function of the prior volume, $\mathcal{L}(V)$. This enables us to rewrite the multi-dimensional expression for the Bayesian evidence (Equation 2.52) as a one-dimensional integral,

$$E = \int_{\Theta_{\mathcal{M}}} \mathcal{L}(\boldsymbol{\theta})\, \pi(\boldsymbol{\theta})\, d\boldsymbol{\theta} = \int_0^1 \mathcal{L}(V)\, dV. \tag{2.55}$$

We see an illustration of this in Figure 2.8. The iso-likelihood contours in a two-dimensional parameter space (left panel) translate into likelihood values associated with prior volumes (right panel). The area under the curve in the right panel of Figure 2.8 represents the Bayesian evidence we want to evaluate. If we evaluate a set of likelihoods $\mathcal{L}_i = \mathcal{L}(V_i)$ for $i \in [1, M]$, then we can approximate the Bayesian evidence through numerical integration (or quadrature) as a weighted sum,

$$E \geqslant \sum_{i=1}^{M} \mathcal{L}_i\, w_i, \tag{2.56}$$

where w_i represents the weight associated with $\mathcal{L}(V_i)$ The core algorithm of NS revolves around the iterative evaluation of Equation 2.56, walking up the curve in Figure 2.8 (right panel) and computing the model evidence.

In `TauREx`, NS is implemented through the `MultiNest` algorithm (Feroz et al. 2009)[4]. It uses a trapezium approach to determine the weights in Equation 2.56,

$$w_i = \frac{V_{i-1} - V_{+1}}{2}, \tag{2.57}$$

and performs the sampling routine as follows. A number of points, $n_{\rm live}$ (referred to as *live points*), are drawn from the full prior, $\pi(\boldsymbol{\theta})$, with an initial prior volume of $V_0 = 1$. These samples are ordered by their likelihood values, and the sample with the smallest likelihood ($\mathcal{L}_1$) is discarded from the set of live points. Subsequently, a replacement point, $n_{\rm new}$, is drawn with the requirement $\mathcal{L}_{n_{\rm new}} > \mathcal{L}_1$. This reduces the prior volume to $V_1 = \alpha_1 \cdot V_0$, where $\alpha_1 \in [0,1]$

[4]specifically the `PyMultiNest` Python package (Buchner et al. 2014)

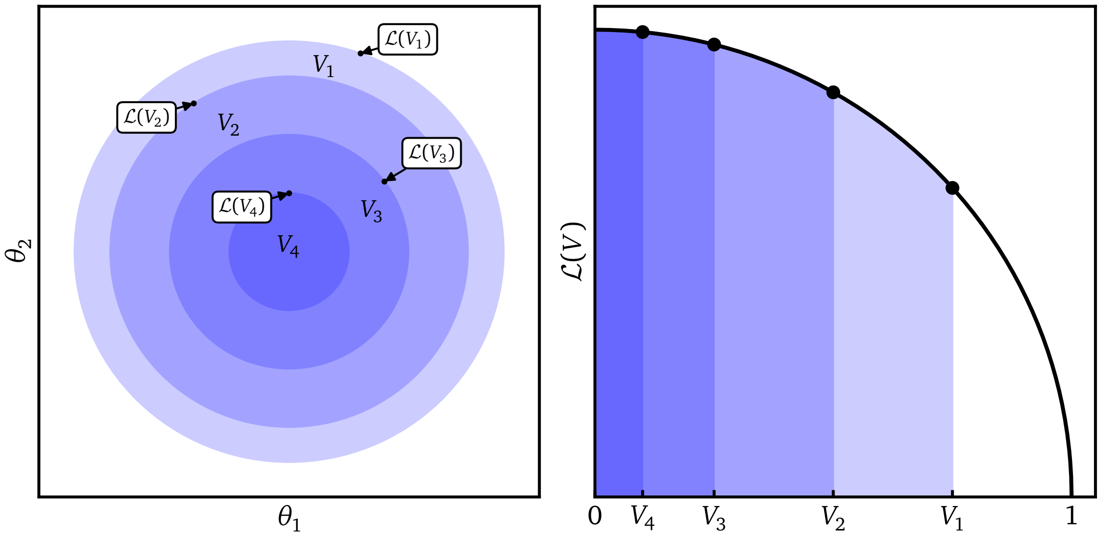

Figure 2.8: Schematic illustration of the prior volume and Bayesian evidence. (Left) Iso-likelihood contours, $\mathcal{L}(V_i)$ in a two-dimensional parameter space, (θ_1, θ_2), enclosing prior volumes, $V_i = V(\mathcal{L}_i)$. (Right) The transformed $\mathcal{L}(V)$-function, associating the prior volumes, V_i, with the enclosed likelihood values, $\mathcal{L}(V_i)$.

is a compression factor. In this iterative process, NS therefore allows us to step through the parameter space, compressing the sampled area to the prior volume with the highest likelihood.

We repeat the process of updating the set of live points until the evidence is determined to some specified precision, ε. At the i-th iteration, the largest contribution to the evidence that can be made by the remaining prior volume is

$$\Delta E_{i,\max} = \mathcal{L}_{\max} \cdot V_i, \tag{2.58}$$

where $\mathcal{L}_{\max}$ is the maximum likelihood of the current set of live points. The associated prior volume, V_i, can be approximated by

$$V_i = \exp(-i/n_{\text{live}}). \tag{2.59}$$

If $\Delta E_{i,\max} \leqslant \varepsilon \cdot E$, the algorithm stops. At this point, the sampling process has provided us with a total of $\mathcal{N}$ samples. These include a total of M inactive points (which have been replaced in the sampling process) and a final set of n_{live} live points, or $\mathcal{N} = M + n_{\text{live}}$. We can then refine the evidence with a final increment from the current live points through

$$\Delta E_{\text{final}} = \sum_{j=1}^{n_{\text{live}}} \mathcal{L}_j \, w_{M+j}, \tag{2.60}$$

where $w_{M+j} = X_M/n_{\text{live}}$ for all j. Once we have found E, we can analyse the posterior sample distribution, $\mathcal{N}$. Each point in this set (also called the *trace*) is associated with a posterior weight, calculated through

$$p_i = \frac{\mathcal{L}_i w_i}{E}. \tag{2.61}$$

The posterior samples and their associated weights now allow us to calculate statistical measures on the sampled model parameters.

2.3.3 Results of Bayesian inference

In the context of Bayesian inference, we are primarily concerned with two outcomes of the inference process: parameter estimates and model performance. The former evaluates the results of an individual model, informing us about the region of the model's parameter space that can best reproduce a dataset. The latter is helpful for comparing the performance of multiple models against one another.

Estimated model parameters

During the inference process for a specific atmospheric forward model, the sampling algorithm produces multidimensional vectors of forward-model parameters, $\boldsymbol{\theta}^j = \{\theta_i^j\}_{i \in I}$, at each iteration. After the sampling process is finished, these vectors make up the *posterior samples* (or simply posteriors) of the model. To report the results of an atmospheric retrieval, we summarise the posteriors in the form of *credible intervals*. If we say that θ_i is one of the parameters of our model, defined on the domain Θ_i, then a credible interval, $C \subset \Theta_i$, is defined by

$$\int_C P(\theta_i|\boldsymbol{d})\, d\theta_i = \alpha. \tag{2.62}$$

Here, $P(\theta_i|\boldsymbol{d})$ is the posterior probability distribution of θ_i, and $\alpha \in [0, 1]$ defines the fraction of the posterior probability mass contained in C. Given the observed data and the model, there is a posterior probability of α that θ_i is in C.

Credible intervals are not uniquely determined. We can place a credible interval of size α in any subset $C \in \Theta_i$ that fulfils Equation 2.62. Throughout this work, we will examine parameter estimates for each model parameter using a centred credible interval (CCI). The CCI is centred on the median (50%-quantile) of the marginalised parameter posterior, and we calculate its boundaries as

$$\mathrm{CCI}_\alpha = \left[q\left(0.5 - \frac{\alpha}{2}\right), q\left(0.5 + \frac{\alpha}{2}\right)\right], \tag{2.63}$$

where α denotes the size of the CCI, and q represents the quantiles associated with the empirical cumulative distribution function (CDF) of the parameter. Since NS produces weighted posteriors, the empirical CDF of each parameter is defined through the cumulative sum of these weights, ordered in the prior space of the parameter.

The size, α, of any credible interval is a subjective choice we need to make. Contrary to the frequentist concept of the *confidence interval*, we do not assume an underlying analytical distribution to the credible interval. It merely represents the statement that "we assign a degree of belief (probability) of α to the assumption that the value of the parameter sits within the credible interval". A common choice of credible interval is CCI_{68}, which draws parallels to the standard deviation of a normal distribution. However, as posterior distributions do not necessarily follow a standard normal distribution, the boundaries of a CCI_{68} do not simply scale the same as an $n\sigma$ confidence interval. We therefore need to be careful in documenting the size of the credible interval we derive from our results.

Model performance

Model comparison allows us to assess the performance of a set of competing models, $\boldsymbol{\mathcal{M}} = \{\mathcal{M}_i\}_{i\in I}$, with the goal of finding a balance between model parsimony and fit quality. We can employ several metrics to evaluate model performance, which supplement each other with different kinds of information.

The Bayes factor

We can use Bayes theorem not only to update our degree of belief in the parameters of a specific model. It also allows us to compare our degrees of belief across different models used to represent the same data. Starting again from Bayes theorem (Equation 2.45), we substitute event A as our observation, but now event B is not the set of parameters making up a specific model, but rather a whole set of models, $\boldsymbol{\mathcal{M}}$. The posterior probability of any model $\mathcal{M}_i \in \boldsymbol{\mathcal{M}}$ follows the relation

$$P(\mathcal{M}_i|\boldsymbol{d}) = \frac{P(\boldsymbol{d}|\mathcal{M}_i)P(\mathcal{M}_i)}{\sum_{j=1}^{I} P(\boldsymbol{d}|\mathcal{M}_j)\,P(\mathcal{M}_j)}. \tag{2.64}$$

We can interpret this the same way we did Equation 2.46: $P(\mathcal{M}_i)$ represents the prior probability we assign to $\mathcal{M}_i$, $P(\boldsymbol{d}|\mathcal{M}_i)$ represents the likelihood of the data over the entire parameter space of the model (the model evidence, E_i, we have defined in Equation 2.52), and $P(\mathcal{M}_i|\boldsymbol{d})$ is the posterior probability of $\mathcal{M}_i$. The denominator of Equation 2.64 is now the likelihood of the data over the entire ensemble of models. If we want to compare our degrees of belief in two models after we have considered the data, we can use Equation 2.64 to construct the expression

$$\frac{P(\mathcal{M}_i|\boldsymbol{d})}{P(\mathcal{M}_j|\boldsymbol{d})} = \frac{P(\boldsymbol{d}|\mathcal{M}_i)}{P(\boldsymbol{d}|\mathcal{M}_j)} \cdot \frac{P(\mathcal{M}_i)}{P(\mathcal{M}_j)}. \tag{2.65}$$

Equation 2.65 represents the ratio of posterior probabilities between two models, $\mathcal{M}_i$ and $\mathcal{M}_j$, also called the *posterior odds*. It is a measure of how well one model has performed in representing the data compared to another model, and is therefore inherently a relative metric.

Table 2.1: Categories for evaluating the Bayes factor between two models, $\mathcal{M}_i$ and $\mathcal{M}_j$, following Kass and Raftery (1995). Posterior odds, B_{ij}, are rounded to the nearest integer. The two-model posterior probability is computed assuming $\boldsymbol{\mathcal{M}} = \{\mathcal{M}_i, \mathcal{M}_j\}$. For negative values of $\ln(B_{ij})$, the interpretation of the Bayes factor switches in favour of model $\mathcal{M}_j$.

$\ln(B_{ij})$	Posterior odds	Two-model posterior probability	Description
0	1:1	0.500	
1	3:1	0.731	Positive evidence for $\mathcal{M}_i$
2	7:1	0.881	
3	20:1	0.953	Strong evidence for $\mathcal{M}_i$
4	55:1	0.982	
5	148:1	0.993	Very strong evidence for $\mathcal{M}_i$
6	403:1	0.998	

The ratio of model evidence values defines the *Bayes factor*,

$$B_{ij} = \frac{P(\mathbf{d}|\mathcal{M}_i)}{P(\mathbf{d}|\mathcal{M}_j)} = \frac{E_i}{E_j} = B_{ji}^{-1}. \tag{2.66}$$

It is a conversion factor that transforms the prior odds of two models into posterior odds. If we assume that all models in the ensemble have equal prior odds, then the posterior odds of two models are simply equal to the Bayes factor,

$$\frac{P(\mathcal{M}_i|\mathbf{d})}{P(\mathcal{M}_j|\mathbf{d})} = B_{ij} \tag{2.67}$$

Judging the value of B_{ij} is subjective. In and of itself, it provides us with a way of calculating the posterior odds ratio between two models trying to represent the same data. The interpretation of B_{ij}, therefore, is subject to the desired standards of scientific evidence. In this work, I will use the categories suggested by Kass and Raftery (1995), which are based on the natural logarithm of the Bayes factor, $\ln(B_{ij})$. I list these categories in Table 2.1. Following them, an odds ratio of less than 3:1 does not consitute an evidence worth mentioning. On the other end of this scale, an odds ratio of more than 148:1 constitutes very strong evidence. I note that we will not analyse the Bayes factor in the context of commonly used *detection significances*, or σ-values. A common method to make Bayes factors more 'interpretable' in the context of atmospheric retrievals is to translate them into σ-values. This is done by inverting an expression of the Bayes factor in terms of the frequentist p-value, presented in Sellke et al. (2001, Equation 2). However, this conversion is applied erroneously, as recently pointed out by Thorngren et al. (2025). There is also no need to specifically convert the Bayesian interpretation of posterior model probability into a frequentist significance value. Therefore, we will stick to interpreting Bayes factors as what they represent: the posterior odds ratio of two competing models.

The application of the Bayes factor has some caveats we need to be aware of. Firstly, the definition of B_{ij} in Equation 2.67 assumes that equal prior odds for the models we are comparing. While assigning proper prior odds to models can be difficult, this represents an additional source of uncertainty that influences the posterior odds analysis. Secondly, as B_{ij} is computed from the model evidence, Equation 2.52 shows us that it is strongly sensitive to $\pi(\boldsymbol{\theta})$. A weaker (or broader) prior distribution extending into regions of the parameter space with low likelihood will lower the evidence, possibly unduly penalising models (Thorngren et al. 2025). Lastly, B_{ij} is a relative metric. It can find the most favoured model out of $\boldsymbol{\mathcal{M}}$, but it does not guarantee that this model is a good fit to the data. Therefore, we will supplement the use of the Bayes factor with other metrics. This will provide us with a more comprehensive picture to judge the performance of individual models in an ensemble.

Reduced χ-square

One of the most widely used quality-of-fit metrics is the reduced χ-square statistic,

$$\chi^2_\nu = \frac{1}{\nu} \cdot \sum_{i=1}^{D} \frac{(d_i - m_i)^2}{\sigma_i^2}, \tag{2.68}$$

where, for a total of D measurements, the vector (d_i, σ_i, m_i) contains the observed data, standard deviation, and model prediction of the i-th data point. The factor $\nu = D - n$ represents the degrees of freedom of the model, defined as the difference between the total number of data points, D, and the number of free parameters of the model, n. Equation 2.68 informs us about the average squared discrepancy between the data and prediction per degree of freedom, normalised to the error of the data.

χ^2_ν is a point estimate. It uses only a single value, m_i, to compare against the data. Bayesian inference provides us with a range of values, the credible interval. To determine χ^2_ν in this work, we will use the median model, constructed from the medians of all parameter posterior distributions. The χ^2_ν value provides us with a way of checking how well the posterior prediction for a model fits the data, complementing the relative comparison provided by the Bayes factor.

Akaike information criterion

The Akaike information criterion (AIC, Akaike 1974) was derived as an estimator of the information loss incurred when using an approximate model to represent a true data-generating process. It is calculated as

$$\mathrm{AIC} = -2\ln\left(\hat{\mathcal{L}}\right) + 2\,n, \tag{2.69}$$

where $\hat{\mathcal{L}}$ represents the maximum likelihood of the model and n the number of model parameters. It provides us with a way of comparing the performance of models relative to each other without depending on the model priors. In this work, we will use the corrected Akaike information criterion (Burnham and Anderson 2004, referred to as Ψ from here onwards),

$$\Psi = \underbrace{-2\ln\left(\hat{\mathcal{L}}\right) + 2n}_{\text{AIC}} + \frac{2n(n+1)}{D-n-1}, \tag{2.70}$$

where D represents the number of data points. Ψ modifies the AIC when $D/n < 40$, and converges to the AIC if D becomes large.

Like χ^2_ν, Ψ is a point estimate, using the maximised likelihood. However, it represents a relative metric, similar to the Bayes factor. As an estimate of information loss, smaller values of Ψ are better. In an ensemble of models, $\boldsymbol{\mathcal{M}} = \{\mathcal{M}_i\}_{i\in I}$, each model is therefore associated with

$$\Delta_i = \Psi_i - \Psi_{\text{min}}, \tag{2.71}$$

where Ψ_{min} is the defined as $\Psi_{\text{min}} = \min\{\Psi(\mathcal{M}_\text{i}) : \mathcal{M}_\text{i} \in \boldsymbol{\mathcal{M}}\}$. Models with $\Delta_i \leqslant 2$ are considered to have substantial support from the data, comparable to the model with Ψ_{min}, while models with $\Delta_i \geqslant 4$ and $\Delta_i \geqslant 10$ have considerably less and essentially no support, respectively (Burnham and Anderson 2004).

Parameter recovery accuracy

When characterising exoplanet atmospheres using atmospheric retrieval, we do not have the luxury of knowing the true values of the parameters we estimate. However, when we perform retrievals on synthetically generated transmission spectra, we have full control over the input parameters used to generate them. In these cases, we can perform parameter self-retrievals. Let us assume that the forward model used to generate a synthetic transmission spectrum has I parameters, $\boldsymbol{\phi} = \{\phi_i\}_{i\in I}$. We can use a retrieval model, characterised by J parameters, $\boldsymbol{\theta} = \{\theta_j\}_{j\in J}$, for an atmospheric retrieval of this synthetic dataset. The overlap of parameters between the forward model and the retrieval model, or $\boldsymbol{\phi} \cap \boldsymbol{\theta}$, contains K parameters, $\boldsymbol{\xi} = \{\xi_k\}_{k\in K}$. For each $\xi_k \in \boldsymbol{\xi}$, we can determine if the atmospheric retrieval has successfully found the underlying, true parameter if $\phi_{k,\text{GT}} \in \text{C}_\alpha(\xi_k)$. Here, $\phi_{k,\text{GT}}$ is the value of ϕ_k used to calculate the forward model. The rate of these successful self-retrievals depends on the credible interval we choose. For the parameter accuracy evaluation we will perform in this work, we choose $\text{CCI}_{99.7}$ as the credible interval.

CHAPTER 3

INTRINSIC LIMITS OF RETRIEVALS

p–T profile biases in transmission spectroscopy

"*Since all models are wrong, [we] must be alert to what is importantly wrong.*"

George E. P. Box

Science and Statistics (1976), JASA, 71, 356

DECLARATION OF PUBLISHED WORK

This Chapter includes text already published in Schleich et al. (2024). As a result, there may be partial overlap with this publication.

In this Chapter, we will explore the role of the forward model in atmospheric retrievals, which determines the parameters we estimate in the inference process. In this context, it is crucial to understand both (1) how much model complexity is supported by the underlying data, and (2) in which ways the setup of the forward model can bias the retrieval results and influence our conclusions. To address these aspects of the forward model setup, we focus on how to characterise the pressure-temperature (p–T) profile in the retrieval of the transmission spectrum of a hot Jupiter. Part of this work was previously published in Schleich et al. (2024). It also contains previously unpublished work in preparation for submission as Schleich et al. (in prep.)

3.1 Context and motivation

As shown in Chapter 2, the technique of atmospheric retrieval rests on several main pillars: the observational data, the atmospheric forward model, and the statistical sampling method. If we want to test the reliability of atmospheric retrieval results, each of these pillars provides an avenue to assess it. While the statistical sampling of atmospheric retrievals is fully guided by the underlying observational data, the parameter estimates are necessarily constrained by the forward model.

Section 2.2 has provided us with an overview of various ways in which the physical characteristics of an exoplanet's atmosphere are incorporated into these forward models. One of these is the vertical p–T profile. Within the Solar System, where atmospheres are accessible to in-situ measurements, we have already found a variety of complex p–T structures (e.g., Seiff et al. 1998; Fulchignoni et al. 2005; Koskinen et al. 2015; Limaye et al. 2017). In the context of exoplanets, where we rely solely on indirect measurements, we necessarily have to use simplified models to represent data with lower information content. In Section 1.3.1, we have already seen an ensemble of commonly used p–T prescriptions. Heuristic, or 'free', p–T profiles retrieve individual temperature values at different pressure levels within the atmosphere (which can themselves be free parameters). They offer significant flexibility in retrieving the atmospheric temperature structure, but at the cost of an increasing number of free parameters. Parametric or analytic profiles offer a trade-off between a smaller number of free parameters and more rigid assumptions (e.g. Madhusudhan and Seager 2009; Guillot 2010).

Extracting information about the p–T structure of an exoplanet's atmosphere is commonly associated with the analysis of emission spectra. They probe large pressure ranges on an exoplanet's dayside by observing its thermal emission. Line et al. (2012) and Line et al. (2013) investigated the impact of using parametric and heuristic p–T profiles in the retrieval of these datasets. By performing retrievals on synthetic spectra, they find that a preference for the parametric profile transitions to a preference for the heuristic profile with increasing SNR and spectral resolution. Blecic et al. (2017) draws a similar conclusion when analysing the one-dimensional retrieval of synthetic hot Jupiter emission spectra created from 3D radiative-hydrodynamic simulations. They compared the performance of two parametric profiles and found that, in high spectral resolution and SNR cases, a more flexible parametrisation produces a better representation of the mean underlying p–T structure.

In contrast, transmission spectra probe the atmospheric temperature profile only indirectly through the atmospheric scale height (as we have seen in Equation 2.24). The instruments of HST, the previous state-of-the-art for space-based transmission spectroscopy, only cover a narrow wavelength range ($\Delta\lambda < 1\,\mu$m). Consequently, an isothermal p–T prescription has found ample application in analyses of these data (e.g., Tsiaras et al. 2018; Roudier et al. 2021), which probe only a narrow vertical region of an exoplanet's atmosphere. Improving the data quality changes this assumption. Rocchetto et al. (2016) investigated the impact of using different p–T prescriptions on observations of a quality expected from JWST. Their results, derived for an idealised scenario represented by a cloud-free hot Jupiter observed with the full wavelength coverage of JWST, indicate a strong bias in the retrieved MRs of atmospheric constituents when an isothermal p–T profile is used.

The recent launch of JWST and the upcoming launch of the Ariel space mission (Tinetti et al. 2021) motivate us to investigate this closer. Their instruments can observe transmission spectra over a much broader wavelength range and at significantly improved SNR. In this Chapter, we systematically investigate the role of the p–T profile in the retrieval of exoplanet transmission spectra. We do this using simulated data, which provides us with complete information about the true parameters used to generate it. As an exoplanetary model, we use the hot Jupiter WASP-39 b, the most well-studied exoplanet by JWST to date. We simulate observations for all four instruments of JWST (NIRCam, NIRSpec, NIRISS, and MIRI), and for the upcoming Ariel space mission. The resulting synthetic spectra cover a variety of input p–T profiles, as well as attenuation of the spectrum by atmospheric aerosols in the form of a cloud deck. We perform homogeneous atmospheric retrievals on these spectra, using a large range of heuristic p–T profiles. This allows us to quantify the necessary and supported p–T profile complexity to correctly recover an exoplanet's atmospheric parameters from its transmission spectrum.

3.2 Methods

To investigate the behaviour of parameter estimates in atmospheric retrievals, we need control over the true values of these parameters. In real observational data, the ground-truth of these parameters is unknown. We therefore synthetically generate datasets for atmospheric retrievals, which consist of two parts. The atmospheric transmission spectrum is created by performing radiative transfer through an atmospheric forward model. This spectrum is then combined with error bars, representing our uncertainty in the individual transit depth values.

Our synthetic observations are modelled after the star-planet system of WASP-39. It consists of a late-G type host star orbited by a hot Jupiter, WASP-39 b, at a period of approximately 4 d (Faedi et al. 2011; Mancini et al. 2018). WASP-39 b has been extensively observed with JWST within the ERS programme, providing us with transmission spectra from all four of its

Table 3.1: Parameters of the WASP-39 system. Stellar and planetary parameters and uncertainties are taken from Mancini et al. (2018).

Parameter	Value	Unit
Star (WASP-39)		
Stellar mass (M_*)	0.918 ± 0.047	$M_\odot$
Stellar radius (R_*)	1.013 ± 0.022	$R_\odot$
Effective temperature (T_{eff})	$5\,485 \pm 50$	K
Metallicity ([Fe/H])	0.01 ± 0.09	dex
Surface gravity ($\log g$)	4.41 ± 0.15	dex
Planet (WASP-39 b)		
Planetary mass (M_p)	0.281 ± 0.032	M_J
Planetary radius (R_p)	1.279 ± 0.040	R_J
Equilibrium temperature (T_{eq})	$1\,116^{+33}_{-32}$	K
Orbital period (P)	$4.0552941 \pm 3.4 \times 10^{-6}$	d
Semi-major axis (a)	$0.0486 \pm 5 \times 10^{-4}$	AU
Inclination (i)	$87.83^{+0.25}_{-0.22}$	$^\circ$

instruments. The system parameters that form the foundation for our synthetic forward models are given in Table 3.1.

3.2.1 Synthetic input data

We then create a grid of synthetic transmission spectra by performing radiative transfer calculations through one-dimensional atmospheric models. We set the atmospheric pressure domain of these models as $\log_{10}(p\,[\mathrm{bar}]) \in [1; -9]$, with 110 atmospheric layers. As a standardised setup for all cases, we construct a primary atmosphere dominated by H_2 and He. Additionally, we include CH_4, CO, CO_2, H_2O, and H_2S as atmospheric absorbers with constant MRs throughout the atmosphere. We also include contributions from CIA of H_2-H_2 and H_2-He, and Rayleigh scattering as implemented in `TauREx`. We can see the values defining the gas phase of our atmospheric model and the sources associated with their optical depth contribution in Table 3.2. Figure 3.1 shows how these contributions form a synthetic transmission spectrum, extending from 0.5 µm to 12 µm and using the system parameters of WASP-39 b, illustrated for a constant atmospheric temperature of $T_{\mathrm{eq}} = 1116\,\mathrm{K}$. The spectrum is dominated by the absorption features of H_2O, with minor regions at longer wavelengths where CO_2 and H_2S become the dominant absorbers. The continuum of the spectrum is dominated by Rayleigh scattering at short wavelengths ($\lambda < 1\,\mu\mathrm{m}$), and then by CIA. CH_4 and CO are spectrally less active, only showing minor regions where their contribution surpasses the underlying continuum. Using this fixed chemical setup, we create a small grid of forward models to systematically investigate the influence of different p–T prescriptions in the retrieval of state-of-the-art transmission spectra.

Table 3.2: Atmospheric species for synthetic forward models. Values for volume-mixing ratios, X_M, are inspired by the results presented in JTERS23

Parameter	Value	Source[a]
He / H_2	0.13	H_2–H_2 (Abel et al. 2011) H_2–He (Abel et al. 2012; Fletcher et al. 2018)
X_{CH_4}	1×10^{-7}	Yurchenko et al. (2017)
X_{CO}	4×10^{-3}	Li et al. (2015)
X_{CO_2}	5×10^{-5}	Yurchenko et al. (2020)
X_{H_2O}	5×10^{-3}	Polyansky et al. (2018)
X_{H_2S}	5×10^{-4}	Azzam et al. (2016)

[a] Sources for He / H_2 refer to CIA cross sections, sources for X_M refer to molecular absorption cross sections.

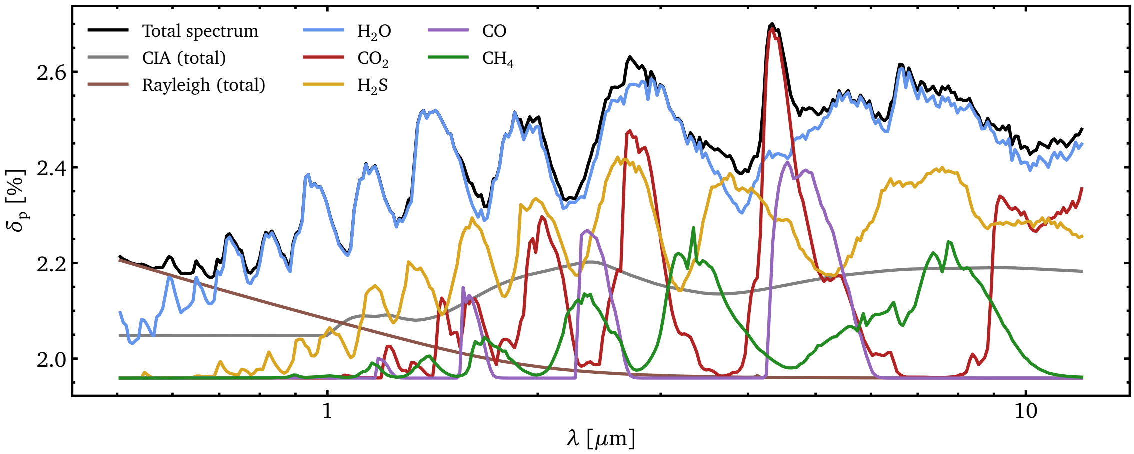

Figure 3.1: Absorption contributions from atmospheric constituents, showing wavelength (x-axis, in µm) against transit depth (y-axis, in %). Individual contributions are colour-coded for molecular constituents of the atmosphere, as well as for the total contributions from CIA and Rayleigh scattering. The total transmission spectrum is shown as a black solid line (calculated using an isothermal temperature profile with $T = 1116\,\mathrm{K}$).

Pressure-temperature profiles

Our focus in this Chapter is on the role of the p–T profile in atmospheric retrievals of transmission spectra. Consequently, we move beyond the simplified isothermal assumption and create synthetic transmission spectra using two different p–T profiles, defined by eleven (p, T) pairs distributed log-uniformly within the atmosphere. This defines the first parameter in our grid of atmospheric forward models. The left panel of Figure 3.2 illustrates the resulting profiles. We differentiate between them by their behaviour with decreasing pressure. The first profile (shown in blue) is less complex, with a purely negative temperature gradient. We refer to it as 'monotonic' from here on. The other one (shown in red) has a more complex vertical structure, characterised by a temperature inversion between a negative temperature gradient in the lower atmosphere and a positive temperature gradient in the upper atmosphere. We refer to this profile as 'inverse' from here on. The shapes of the two profiles are inspired by the cases of the hot Jupiters WASP-76 b and WASP-77 A b presented in Changeat et al. (2022) for

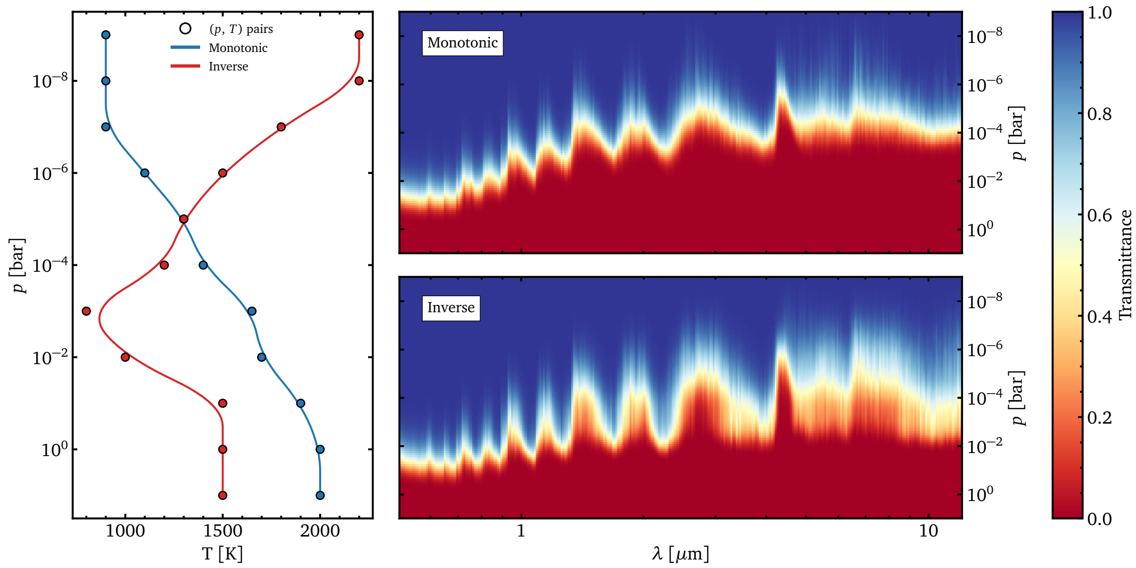

Figure 3.2: Summary of the temperature structure used to generate synthetic transmission spectra. (Left) p–T structure within the atmospheric forward model domain, showing temperature (x-axis, in K) against pressure (y-axis, in bar). The 'monotonic' profile is shown in blue, while the 'inverse' profile is shown in red. The circles show the eleven p-T nodes used as input. The solid lines show the interpolated vertical temperature structure. (Right) Transmittance maps for the monotonic (upper panel) and inverse (lower panel) case, showing wavelength (x-axis, in µm) against transit depth (y-axis, in %).

the 'monotonic' and 'inverse' profiles, respectively.

We see the impact of using these two p–T profiles on our synthetic transmission spectra in the right panels of Figure 3.2. It shows the wavelength- and pressure-dependent transmittance for the monotonic and inverse case. As the monotonic case is hotter at higher pressures, the model's scale height is larger. This inflates the atmosphere, pushing the region probed in transmission to higher altitudes. In contrast, the inverse p–T profile is cooler at lower altitudes, meaning that lower regions in the atmosphere are probed. We can see this prominently at approximately 2.5 µm. The overall shapes of the transmittance map and the resulting transmission spectra are closely overlapping.

Atmospheric cloud deck

In constructing our synthetic spectra, we also want to be mindful of the impact of clouds. Rocchetto et al. (2016) investigated the role of the p–T profile in an idealised scenario, not considering any cloud cover. Observations have shown us that aerosol formation is ubiquitous in the atmospheres of exoplanets, spanning a wide range from spectrally clear to heavily clouded, where increasing cloudiness leads to muted spectral features (Sing et al. 2016; Gao et al. 2021). We want to account for this in our synthetic forward models, quantifying how atmospheric cloudiness influences the p–T profile analysis. We choose a simple model prescription for clouds in the form of a flat-opacity cloud deck, defined as an optical-depth step function at a fixed cloud-top pressure (Equation 2.43). This is the second parameter that defines our grid of forward models. We include three different cases of the cloud-top pressure to cover

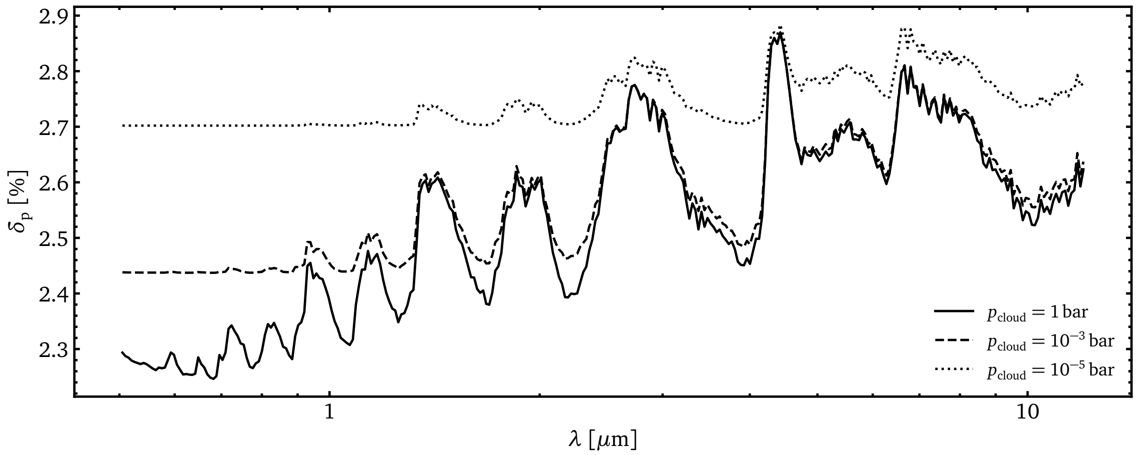

Figure 3.3: Impact of different cloud-top pressure levels on a transmission spectrum, showing wavelength (x-axis, in µm) against transit depth (y-axis, in %). The transmission spectra were generated using the inverse p–T profile, with a cloud-top pressure of 1 bar (solid line), 10^{-3} bar (dashed line), and 10^{-5} bar (dotted line).

a wide range of possible modulations to the transmission spectrum, placing the cloud top at $\log_{10}(p_{\rm cloud}\,[\rm bar]) \in \{0, -3, -5\}$. We label these as high-pressure, medium-pressure, and low-pressure clouds, respectively. Figure 3.3 illustrates the influence of these cases on a synthetic transmission spectrum. In the high-pressure cloud case (solid line), the synthetic spectrum shows significant modulations across the whole wavelength range. In the medium-pressure case (dashed line), the spectrum becomes more muted, especially at wavelengths shorter than 1 µm. Figure 3.2 (right panels) show us that this is where the deepest regions of the atmosphere ($p < 10^{-3}$ bar) are probed, which become obscured by the medium-pressure cloud deck. In the low-pressure cloud case (dotted line), the spectrum is mostly flat, only showing modulations at wavelengths larger than 2 µm, significantly reducing the information it contains.

Transit-depth precision

To transform our synthetic transmission spectra into simulated observations, we need to combine them with wavelength-dependent uncertainties in the transit depth. The third parameter defining our grid of forward models is therefore the precision of the transmission spectrum, $\sigma_{\rm td}$, corresponding to the instrument cases we consider. These are summarised in Table 3.3, and Figure 3.4 shows their associated transit-depth uncertainty maps.

First, we consider an observation taken with NIRSpec in its prism mode. For this case, we generate the transit depth uncertainty in two different ways. On the one hand, we use uncertainty values associated with a `Eureka!`-based data reduction of the NIRSpec PRISM observation of WASP-39 b. This 'realistic' noise is based on the data reduction performed in Chapter 4. As a small difference from the finalised data reduction process, the light-curve fitting for this noise case was performed with an additional point-scatter parameter (e.g., JTERS23, Rustamkulov et al. 2023), slightly increasing the transit-depth uncertainty compared to the transmission spectrum presented in Section 4.2.1. We smooth this error curve by fitting a

Table 3.3: Instrument cases used to generate synthetic transmission spectra.

Instrument	λ_{min} [μm]	λ_{max} [μm]	R^a	σ_{td} [ppm]^a
JWST				
NIRSpec PRISM	2.25^b	5.29	410	104 / 238^c
NIRSpec G395H	2.73	5.17	1999	425
NIRISS SOSS	0.63	2.81	834	260
NIRCam F322W2	2.42	4.02	1101	295
MIRI LRS	5.13	11.86	34	239
JWST (full)^d	0.63	11.86	1391	339
Ariel				
WASP-39, 2 orbits	0.55	7.64	100	369
HD 209458	0.55	7.64	100	100

^{a} The resolution, R, and transit depth uncertainty, σ_{td}, are represented by median values.
^{b} The NIRSpec PRISM observation is saturated at wavelengths shorter than λ_{min}.
^{c} The first and second value represent the `Pandexo` and DRS case, respectively.
^{d} This case uses the noise levels from NIRISS SOSS, NIRSpec G395H and MIRI LRS.

fourth-order polynomial to it, correcting for outliers and deriving a more general wavelength-dependent noise curve this way. We label this case as *data-reduction scaled* (DRS). We also use the JWST noise-simulation tool `Pandexo` (Batalha et al. 2017) to generate a second set of transit-depth error bars for the NIRSpec PRISM setup. We run `Pandexo` using the observational parameters corresponding to the NIRSpec PRISM observation of WASP-39 b performed in observational cycle 1 (PID: 1366), and the default assumptions from ExoCTK[1] (excluding a noise floor) with the system parameters listed in Table 3.1. For future reference, we label this case *Pandexo*. Excluding a noise floor from it represents a more optimistic assumption about the transit-depth uncertainty than in the DRS case. Both noise curves are shown in Figure 3.4 with grey and golden markers representing the DRS and Pandexo case, respectively. Including both allows us to investigate the influence of the transit-depth SNR on the retrieval accuracy from a single instrument. The error bars in the Pandexo case are smaller than in the DRS case by approximately a factor of 2.3, representative of a SNR increase associated with a combination of approximately five transits.

Secondly, we consider the other instrument modes of JWST that have been used to measure a transmission spectrum of WASP-39 b. This includes the first and second spectral orders of NIRISS, NIRSpec with the F290LP/G395H filter and grism combination, the NIRCam long-wavelength channel using the F322W2 filter, and MIRI LRS. For the NIR observations, we take the error bars presented in Carter & May et al. (2024). The authors performed a homogeneous data reduction analysis of all NIR observations of WASP-39 b. We use the values of σ_{td} based on free limb-darkening coefficients (LDCs)[2] at the native instrument resolution. To transform these observation-specific transit depth uncertainties into more general error curves, we smooth

[1] Exoplanet Characterisation Toolkit, `https://exoctk.stsci.edu/pandexo/`
[2] Retrieved from `https://doi.org/10.5281/zenodo.10161743`

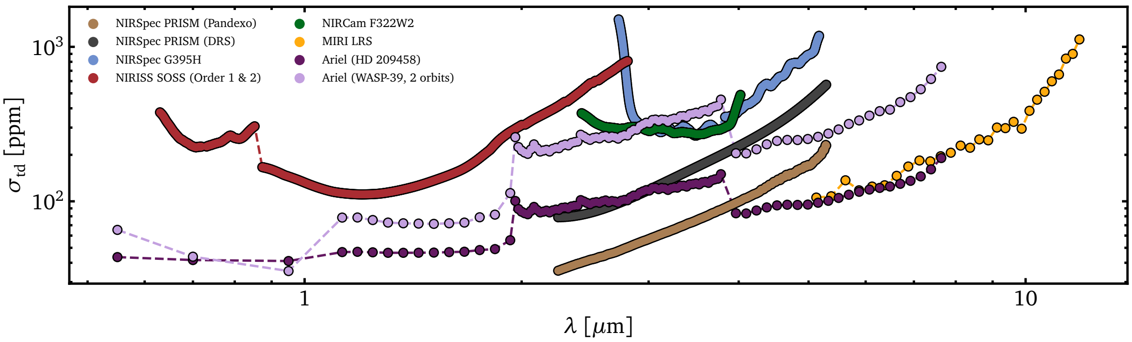

Figure 3.4: Instrument noise cases used to create the synthetic transmission spectra, showing wavelength (x-axis, in µm) against transit depth uncertainty (y-axis, in pmm). The noise curves are labelled by each instrument case, as described in the text.

them to remove outliers. For a more detailed description of this, I refer to Appendix A. For the case of MIRI LRS, we take the values of σ_{td} presented in Powell et al. (2024)[3] based on a data reduction with `Eureka!`. These data are not available at the native instrument resolution. Rather, it is binned into 28 spectral channels with a width of 0.25 µm.

Lastly, we consider noise levels for observations of the upcoming Ariel mission. We simulate this noise using the most recent version[4] of ArielRad (Mugnai et al. 2020). When using WASP-39 as the host star, we mimic an observation resulting from two stacked transits by reducing the simulated noise level by a factor of $\sqrt{2}$. This represents a Tier 2 target for Ariel, designed to investigate the atmospheric structure and composition of individual exoplanets (Tinetti et al. 2021; Edwards and Tinetti 2022) We label this case as 'Ariel-WASP'. As a comparison, we also perform ArielRad noise simulations using the canonical hot Jupiter host star HD 209458. A single transit using HD 209458 as the host star places our simulated observation in Ariel's Tier 3 target category, representing the highest expected data quality, apart from dedicated phase curve observations (Tinetti et al. 2021; Edwards and Tinetti 2022). We label this case as 'Ariel-HD' for future reference.

Final grid of synthetic spectra

To summarise the extent of our grid of forward models: we combine two input p–T cases and three cloud-pressure cases with nine noise scenarios. This provides us with a total of 54 synthetic transmission spectra. Using these, we can investigate the impact of using varying p–T prescriptions in atmospheric retrievals across a wide range of wavelength coverage, SNR, and expected underlying atmospheric structure. The smallest wavelength coverage we consider comes from NIRCam F322W2 ($\Delta\lambda = 1.60$ µm), while the widest coverage is associated with the full JWST case ($\Delta\lambda = 11.24$ µm). For spectra simulating JWST observations, the spectral resolution is generally high. The MIRI LRS case is an exception to this, with a resolution approximately two orders of magnitude lower than the NIR JWST cases. For simulated Ariel observations, the median spectral resolution masks a discontinuous combination of photometric

[3]Retrieved from https://doi.org/10.5281/zenodo.10055845
[4]ArielRad v2.4.26, ExoRad v2.1.111, and the Payload config files v0.0.17.

and low-resolution instruments

As we are interested in investigating the potential biases of using different p–T prescriptions in the retrieval of atmospheric properties, we do not scatter these synthetic spectra. Adding randomisation through a self-scattering step would introduce additional, non-controllable biases into the retrieval results. We use the non-scattered instances of each spectrum as an idealised case, representing the averaging of results from a larger set of randomised instances (e.g., Feng et al. 2018; Changeat et al. 2019).

3.2.2 Atmospheric retrieval

To investigate how the p–T prescription impacts retrieval results, we now perform semi-self-retrievals using TauREx. In all of them, we use 750 live points and a tolerance of 0.5 for the natural logarithm of the evidence. Table 3.4 provides an overview of the retrieval parameters and their associated prior distributions.

Seven of the model parameters are self-retrieved – the reference radius, R_{p}, the five MRs of atmospheric trace constituents, X_{M}, and the cloud-top pressure of the flat-opacity cloud deck, p_{cloud}. These are homogeneous in all retrievals. To handle the p–T structure of the atmosphere, we consider heuristic, or 'free', p–T profiles of varying complexity. In the simplest case, this encompasses an isothermal profile, adding one free parameter to the model. Increasing the complexity, we construct multipoint profiles with two, four, six, and eight pressure nodes. For all these profiles, the first and last pressure nodes are anchored at the bottom of the atmosphere (BOA) and top of the atmosphere (TOA), respectively. For the remaining pressure nodes, we make the informed choice to slightly cluster them around the expected probed region of the atmosphere. As we see in Figure 3.2 (right panels), this region extends approximately between 10^{-2} bar and 10^{-6} bar. Consequently, we place the nodes at values of $\log_{10}(p_i\,[\mathrm{bar}])$ given by

- $\{1, -9\}$ for the two-point profile,
- $\{1, -3, -7 - 9\}$ for the four-point profile,
- $\{1, -1, -3, -5, -7, -9\}$ for the six-point profile,
- $\{1, -1, -2, -3, -4, -5, -7, -9\}$ for the eight-point profile.

Lastly, we also perform a full self-retrieval using an eleven-point p–T profile. The temperature nodes of this case are placed at the same pressure values as the nodes of the input profile (Figure 3.2). Using these six heuristic p–T prescriptions will give us a detailed overview of the supported and necessary vertical temperature structure when retrieving the transmission spectrum of a hot Jupiter. The number of free parameters in these cases increases from eight (the isothermal case) to 18 (eleven-point full self-retrieval). In total we perform 324 atmospheric retrievals to cover our grid of synthetic spectra.

Retrieval performance metric

To assess the performance of all atmospheric retrievals, we use the performance metrics

Table 3.4: Parameters for the retrieval of synthetic transmission spectra. In multipoint retrievals, T_i refers to the temperature at each individual pressure node.

Parameter	Prior type	Range	Associated unit
Self-retrieval			
R_p	uniform	$[1.0, 1.5]$	$\mathrm{R_J}$
X_M	log-uniform	$[10^{-12}, 10^{-0.1}]$	–
p_cloud	log-uniform	$[10^{-9}; 10]$	bar
Multipoint *p*–*T* profile			
T_i	uniform	$[300, 3000]$	K

described in Section 2.3.3, specifically the Bayes factor, $\ln(B_{ij})$, and the parameter recovery accuracy. We will highlight the Bayes factor calculated using two reference cases, the eleven-point self-retrieval, $\ln(B_\mathrm{m,11pt})$, and the isothermal retrieval, $\ln(B_\mathrm{m,iso})$. In the first case, we are evaluating model preference under an optimistic assumption about the information content of each synthetic spectrum – that it contains sufficient information about the atmosphere to recover the full complexity of the underlying p–T structure. In the second case, we are evaluating model preference based on parsimony – if retrievals with complex p–T profiles have posterior odds below a given threshold, the simpler model is preferred.

Next to the Bayes factor, we also look at the accuracy of the retrieval, defined by successful parameter recoveries. For each self-retrieved parameter, θ, we count it as successfully recovered if the $\mathrm{CCI}_{99.7}$ of the estimated parameter encompasses the known input value, θ_GT. For easier visualisation, we calculate the distance between the true parameter value and the median of the parameter posterior distribution,

$$\Delta\theta = \theta_\mathrm{GT} - \mathrm{med}(\theta). \tag{3.1}$$

We then scale $\Delta\theta$ by the width of the credible interval. Because the $\mathrm{CCI}_{99.7}$ is not necessarily symmetric around the median in the parameter space of θ, we calculate this scale as

$$\Omega(\theta) = \begin{cases} \frac{\Delta\theta}{\mathrm{med}(\theta) - q_\theta(0.0015)}, & \theta_\mathrm{GT} < \mathrm{med}(\theta), \\ \frac{\Delta\theta}{q_\theta(0.9985) - \mathrm{med}(\theta)}, & \theta_\mathrm{GT} > \mathrm{med}(\theta). \end{cases} \tag{3.2}$$

Here, $q_\theta(0.0015)$ and $q_\theta(0.9985)$ represent the edges of the $\mathrm{CCI}_{99.7}$ A parameter then counts as successfully retrieved if $\Omega(\theta)$ is less than or equal to unity. Parameters with $\Omega(\theta) > 1$ are overestimated, as the true parameter value is smaller than the range of the CCI. For $\Omega(\theta) < -1$, the opposite is true. We summarise the retrieval performance by counting the successfully retrieved parameters, n_success, against the total number of parameters identical between the forward model and retrievals, n_tot. For each instrument case we have simulated, this gives us $n_\mathrm{tot} = 42$ from seven self-retrieved parameters (five molecular MRs, X_M, the planetary reference radius, R_p, and the cloud-top pressure, p_cloud), two input p–T cases (monotonic and inverse) and three input cloud cases. In addition to this, we assess the accuracy of the retrieved p–T profiles. To do this, we empirically compare the credible intervals of the retrieved p–T profiles to the true temperature structure underlying the synthetic spectra.

3.3 Results

3.3.1 NIRSpec PRISM

The selection of p–T prescriptions we have chosen provides us with a fine-grained overview of the balance between model complexity and data information content across all synthetic spectral cases we consider. We start with analysing the retrieval results for the NIRSpec PRISM instrument cases. As we see in Table 3.3, the difference between the DRS and Pandexo noise case allows us to investigate the impact of an increased observational SNR, while holding the underlying instrument constant.

Model preference
We first investigate how our models perform when assessed using the Bayes factor. In Figure 3.5, we see the evaluation of $\ln(B_{m,11\mathrm{pt}})$ in the DRS and Pandexo noise case, summarised for all cloud-top pressure values and underlying p–T profiles. Two distinct behaviours manifest in these results. Firstly, the Bayes factor peaks at different models based on the underlying p–T structure of the synthetic spectra, as we see in the left and right panels of Figure 3.5, respectively. In the monotonic case, the two- and four-point retrievals are strongly to moderately preferred, while the posterior odds of more complex models decline towards unity. In the inverse case, the preference for the two-point profile vanishes. We still see a positive preference for the four-point profile in the high-pressure cloud cases. Here, the spectra contain enough information to favour the four-point profile over the other multipoint cases. When the spectral features of the synthetic spectra become more muted (medium-pressure clouds), this preference disappears, leaving the multipoint models with inconclusive posterior odds.

A notable exception to this behaviour is the performance of the isothermal profile. In the majority of cases, it is strongly disfavoured against all other multipoint profiles, based on $\ln(B_{\mathrm{m,11pt}})$. Since the Bayes factor is a relative measure, this underlines the preference peaks we observe in the monotonic and inverse cases. The isothermal profile is clearly insufficiently flexible to capture the impact of either of the underlying p–T structures on the synthetic transmission spectra. Only when they become strongly muted (low-pressure clouds) is the Bayes factor insufficient to reasonably distinguish between any of the models.

Comparing the retrievals performed with the DRS noise case to those with the Pandexo noise case, we see that the two distinct behaviours we identified before are exacerbated. The peaks of the preference curves are more pronounced, encompassing the medium-pressure cloud cases as well. The isothermal profile is now strongly disfavoured even in the low-pressure cloud cases. This aligns with the fact that the increased SNR in the Pandexo case indicates that the spectra contain more information. These Bayes factor evaluations already tell us

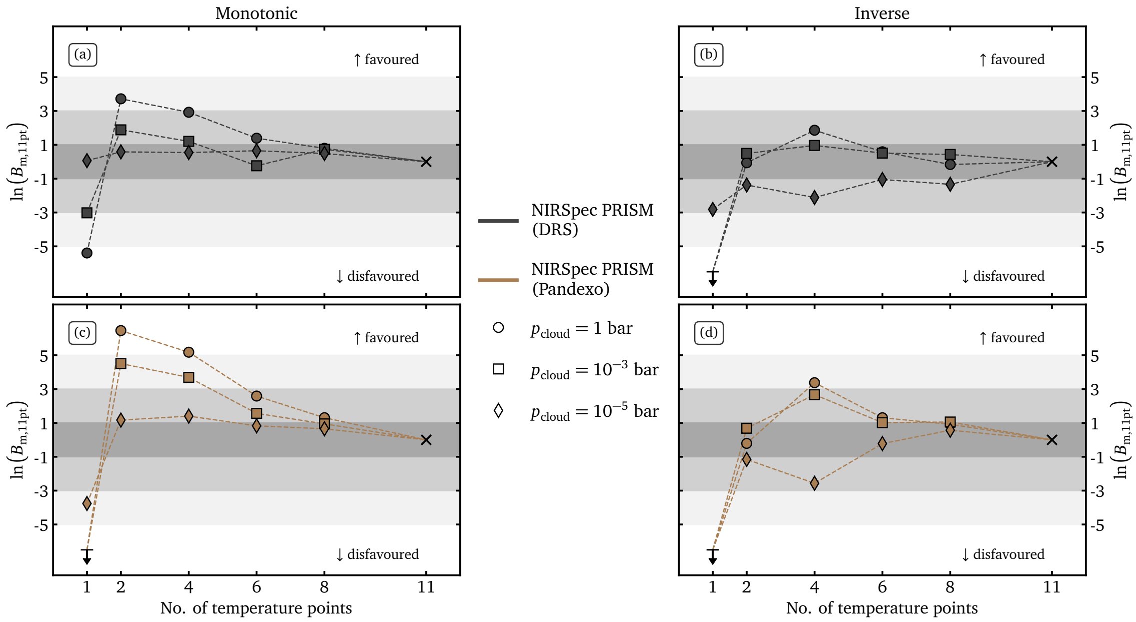

Figure 3.5: Bayes factor comparison for retrievals of the DRS and Pandexo synthetic datasets. All panels show the number of temperature points in the atmospheric retrieval (x-axis) against the Bayes factor in reference to the full self-retrieval (y-axis). The data points are separated into retrievals performed on synthetic spectra generated with high ($p_{cloud} = 1$ bar, circles), medium ($p_{cloud} = 10^{-3}$ bar, squares), and low ($p_{cloud} = 10^{-5}$ bar, diamonds) cloud-top pressure values. The grey shaded areas denote the Bayes factor thresholds given in Table 2.1, and cases fulfilling $|\ln(B_{m,11pt})| > 6.5$ are dented by black arrows. (Top) Results for retrievals of the DRS spectra generated with a monotonic (left, panel a) and inverse (right, panel b) p–T profile. (Bottom) Results for retrievals of the Pandexo spectra generated with a monotonic (left, panel c) and inverse (right, panel d) p–T profile.

that we should substitute isothermal p–T prescriptions for more flexible profiles. This aligns with the results presented in Rocchetto et al. (2016), which showed that isothermal retrievals performed on simulated, cloud-free JWST spectra of hot Jupiters can significantly bias the retrieved atmospheric MRs.

Parameter accuracy

Evaluating the accuracy of the retrieved parameters provides us with a more detailed view of the performance of individual p–T prescriptions. In Figure 3.6, we see the accuracy of the estimated parameter associated with retrievals performed on the DRS spectra. The retrieval results of the seven self-retrieved parameters in each model are split by the underlying input p–T profile and cloud-top pressure. As we have determined, the increase in SNR from the DRS case to the Pandexo cases exacerbates the model preference behaviour. Therefore, the behaviour we observe in the retrieval accuracy of the DRS cases provides us with a lower bound for the Pandexo case.

We focus on the MRs of the atmospheric trace species (H_2O, CO_2, H_2S, CO, and CH_4) first. When the information content in the synthetic spectra is high (high- and medium-pressure clouds, marked by circles and squares), the isothermal profile clearly fails to retrieve the correct MRs. This manifests in the accuracy curves shown in Figure 3.6 flaring outside of the grey

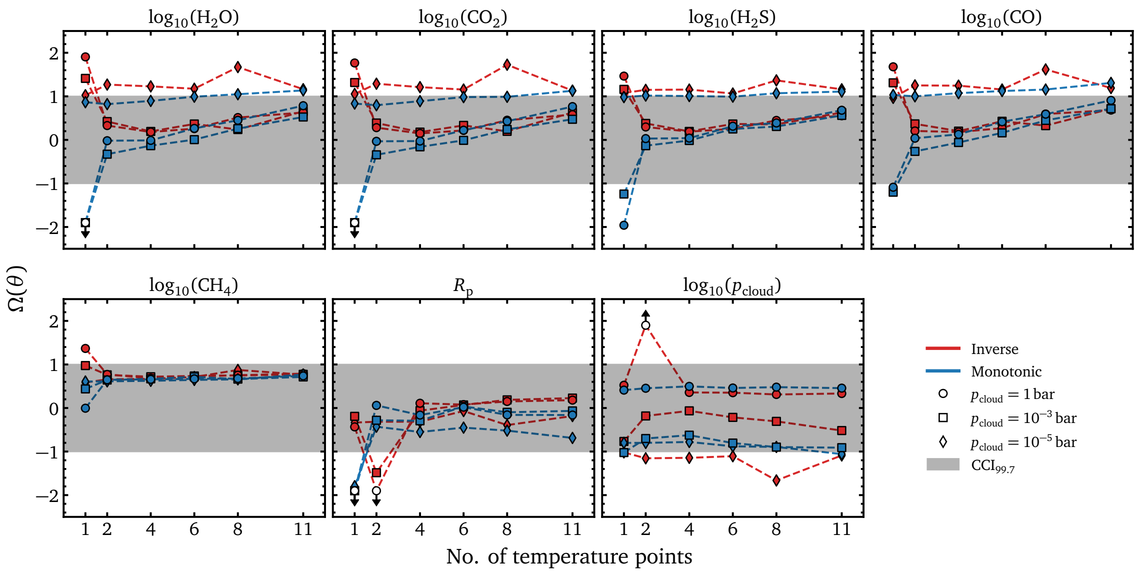

Figure 3.6: Parameter retrieval accuracy in the DRS case. The individual panels represent the self-retrieved parameters in each retrieval: MRs of $H_2O, CO_2, H_2S, CO,$ and CH_4, planetary reference radius, and cloud-top pressure. All panels show the number of temperature points in the atmospheric retrieval (x-axis) against the scaled distance of the input parameter value from the parameter posterior median ($\Omega(\theta)$, y-axis). The accuracy curves are colour-coded by the underlying p–T case, where blue represents the monotonic profile and red the inverse profile. Individual markers denote the high-pressure (circles), medium-pressure (squares), and low-pressure (diamonds) cloud-top input cases. Values of $|\Omega(\theta)| > 2$ are denoted by empty markers with arrows.

area, marking $|\Omega(\theta)| = 1$. We see this most prominently for H_2O and CO_2. These species create the most distinct features in the synthetic spectra (Figure 3.1). For H_2S and CO, this bias is less prominent, but the isothermal retrievals still fail to recover the correct MRs. CH_4 is a slight outlier in this. Due to the low underlying abundance and lack of distinct absorption features in the synthetic spectra, the posterior distributions associated with CH_4 are significantly broader. Following from this, the ranges of the $CCI_{99.7}$ are wider, more easily encompassing the underlying true value. Notably, we find that the isothermal bias in retrieving atmospheric MRs is anti-symmetric between the monotonic and inverse case. In the monotonic case (blue curves), $\Omega(\theta) < -1$, indicating that the MRs are underestimated. In the inverse case (red case), the opposite is true. In contrast to the isothermal retrievals, the more complex p–T prescriptions consistently retrieve the atmospheric MRs in the presence of high- and medium-pressure clouds. This picture shifts when we consider the low-pressure cloud cases (shown with diamonds), where all p–T prescriptions struggle to accurately retrieve the input parameters.

For the reference radius, R_p, and cloud-top pressure, $\log_{10}(p_{cloud})$, we observe a less distinct behaviour. The isothermal profile fails to retrieve the correct radius in the monotonic case, while the two-point profile does the same in the inverse case. It is also the only profile that incorrectly retrieves the cloud-top pressure in the inverse case with high-pressure clouds. Otherwise, the cloud-top pressure is consistently correctly retrieved by all p–T prescriptions. An exception to this is the inverse case with low-pressure clouds, where we again see that all retrievals fail to produce the correct cloud-top pressure.

Retrieved *p*–*T* profiles

Based on the Bayes factor on parameter retrieval accuracy, we can already tell that, in the context of the multipoint *p*–*T* profiles, the two- and four-point cases should be our focus. None of the more complex *p*–*T* profiles provides an advantage in the accuracy of the retrieved parameter. Because of their increased number of free parameters, they are disfavoured based on their posterior odds.

In Figure 3.7, we see the retrieved *p*–*T* profiles for the isothermal, two-point, and four-point assumptions in the context of the DRS case, compared to the true underlying vertical temperature structure. In the monotonic *p*–*T* case (top row), we can clearly see that two temperature

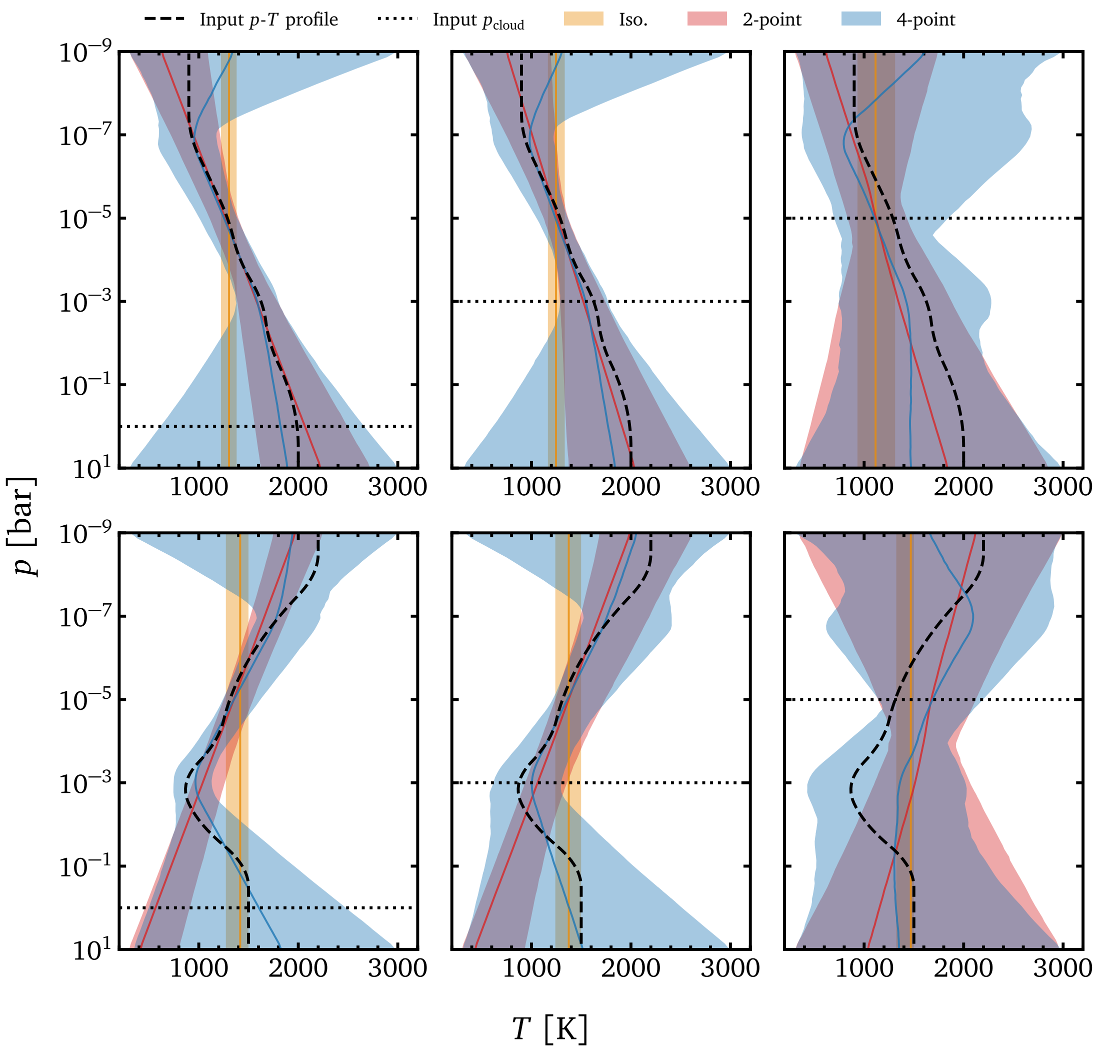

Figure 3.7: Retrieved *p*–*T* profiles in the DRS case. Each panel shows temperature, (in K, x-axis) against pressure (in bar, y-axis). Retrieved profiles are shown for the isothermal (orange), two-point (red) and four-point (blue) cases. For retrieved *p*–*T* profiles, the solid line indicates the median and the shaded area the corresponding $\mathrm{CCI}_{99.7}$. In all panels, the input *p*–*T* profile is shown as a dashed black line. (Top) Retrieval results for the monotonic case. From left to right, the three panels are ordered by increasing input cloud-top pressure, indicated by the horizontal dotted line. (Bottom) Same, but for the inverse case.

points are sufficient to capture the underlying p–T structure, which (at first order) is a simple negative temperature gradient. The retrieved four-point profile also encompasses the input p–T profile, meaning that it successfully retrieved the underlying temperature structure. Comparing this to the inverse case (bottom row), the two-point profile is insufficient to reproduce the temperature inversion point at 1 mbar, significantly underestimating the temperature in the high-pressure regime of the atmospheric forward model as a result. In both cases, the isothermal profile, by its construction, is unable to capture any vertical temperature gradient.

Both the behaviour of the Bayes factor preference curve (Figure 3.5) and the asymmetries in the parameter accuracy (Figure 3.6) are directly correlated to this inability to capture the true vertical temperature structure. We can compare the expected pressure regime probed in transmission with the retrieved temperature profiles. As we have seen in Figure 3.2, the probed pressure regime in the transmission spectra extends approximately between 10^{-2} bar and 10^{-6} bar in a cloud-free case. In the monotonic case, we see that the isothermal retrieval solution underestimates the true atmospheric temperature in the probed pressure regime. In the inverse case, the opposite is true. This is anti-correlated with the bias in atmospheric MRs from isothermal retrievals we have observed before. The source of this anti-correlation is the influence of the local atmospheric scale height, H_l, in the transit geometry. Equation 2.25 tells us that H_l is directly proportional to the local temperature. Increasing the temperature, and therefore H_l, increases the path length of the local atmospheric layer (Equation 2.24). This directly influences the calculation of the optical depth (Equation 2.39), compensating for a longer path length by decreasing the molecular MRs. If the temperature is underestimated, this bias reverses.

This probed region of the atmosphere also influences the width of the retrieved temperature profile CCIs. All multipoint profiles are anchored with two temperature nodes at the BOA and TOA. For the two-point profile, the posterior distributions of these two nodes are comparatively narrow, as the retrieval converges to the temperature gradient registered in the probed atmospheric region. In the case of the four-point profile, this information constrains the posterior distributions of the two intermediate temperature nodes, which sit closer to the probed atmospheric region. Consequently, the posterior distributions of the BOA and TOA temperature nodes are unconstrained, as they span nearly the entire temperature prior range. This behaviour holds until the case in which the low-pressure cloud deck masks much of the probed atmospheric region, leading to the multipoint profiles becoming highly unconstrained.

Increased SNR

As we have observed in the evaluation of the Bayes factor, increasing the observational SNR from the DRS case to the Pandexo case exacerbates the shape of the preference curve. This is also true for the parameter accuracy, which is summarised in Figure B.1. In the case of high-pressure and medium-pressure cloud decks, the values of $\Omega(\theta)$ are larger in the Pandexo noise case, because the CCIs derived from the parameter posteriors are narrower. This is also true for the retrieved p–T profiles. For the multipoint profiles, the credible intervals of the temperature nodes closest to the probed atmospheric region are narrower. For the four-point profile, the credible intervals of the BOA and TOA temperature nodes remain unconstrained, comparable to the DRS case. When a low-pressure cloud deck masks most of the information contained in the synthetic spectra, the behaviour of the parameter accuracy curves and the shape of the retrieved p–T profiles more closely overlap with what we have seen for the DRS case.

3.3.2 Panchromatic preferences

Analysing the simulated observations for NIRSpec PRISM has provided us with a comprehensive overview of the preferred p–T parametrisation to use in a specific instrument scenario. We extend this to cover a large range of instruments from JWST and Ariel. This provides a comprehensive overview of the importance of p–T complexity over a wide range of spectral resolutions, wavelength coverage, and SNR.

Model preference

We first extend the analysis of model preference based on the Bayes factor. Figure 3.8 summarises the previously established Bayes factor analysis for our whole grid of synthetic spectra, encompassing 324 individual retrievals (including the two prior NIRSpec PRISM cases). The correlation between the shape of the Bayes factor preference curve and the underlying p–T profile stays consistent in our extended grid. For the monotonic input case, it peaks at the two-point profile, which is sufficient to capture the simpler, purely negative temperature gradient. In the inverse input case, the peak of the preference curve is shifted to the four-point profile. Here, more than two temperature nodes are necessary to capture the more complex temperature inversion in the input profile.

The magnitude of the model preference, meaning the value of the Bayes factor, is correlated with the instrument case underlying the synthetic spectrum. Datasets that span a wide wavelength range, combined with high spectral resolution, exhibit the strongest model preference peaks. This includes the full JWST range, NIRISS SOSS, NIRSpec G395H, and Ariel-HD. The behaviour of this preference curve is tied to the underlying, true p–T structure. As we see in Figure 3.8, in the monotonic case we find posterior odds of up to 150:1 and 90:1 for two-point and four-point retrievals, respectively. This indicates that they are significantly preferred over more complex profiles. The peak in model preference is muted in the inverse case, where the performance of individual multipoint profiles based on $\ln(B_{m,11\mathrm{pt}})$ is less conclusive. However, we do see a plateau in the model preference curve. Notably, for simulated spectra both spanning a wide wavelength range and high spectral resolution (NIRISS SOSS, JWST full, Ariel-HD, and Ariel-WASP), this plateau begins at the four-point profile, already excluding the two-point retrievals based on their posterior odds. In contrast, for cases with high spectral resolution but smaller wavelength coverage (NIRSpec G395H, NIRCam F322W2), this plateau starts at the two-point profile. This is directly linked to the extent of the atmospheric vertical region probed by the data. In the former case, the broader wavelength coverage (combined with

Figure 3.8: Bayes factor values for the total grid of synthetic retrievals. p–T profiles used in the retrievals are given on the x-axis, and different simulated instrument cases on the y-axis. The cells are colour-coded by their Bayes factor in reference to the eleven-point retrieval (numbers in bold). Below these, the smaller numbers represent the Bayes factor in reference to the isothermal retrieval. The left and right columns represent the monotonic and inverse cases, respectively. In each column, the rows represent different input cloud-top cases.

a low-lying cloud deck) probes a large region of the atmosphere, indicating that neither the isothermal nor the two-point profile can adequately capture the complexity of the underlying p–T profile. In the latter case, the limited wavelength coverage results in a narrower probed region of the atmosphere, and simpler models suffice. Here, the isothermal profile provides insufficient flexibility to capture the underlying p–T profile, but the information in the spectra is sufficiently represented by a simple gradient (i.e., two points).

The magnitude of the model preference is also strongly correlated with the cloud deck underlying the synthetic spectrum. We see this trend by examining the individual rows in Figure 3.8. The gradient in model preference values is most pronounced in the high-pressure cloud top cases, where each spectrum contains the most information. Considering cloud decks at increasingly lower pressures, this model preference behaviour becomes more muted. In the medium-pressure cloud case, models with more complex p–T prescriptions are still strongly preferred over the isothermal retrievals. However, the model preference now plateaus at the two-point retrievals in both the monotonic and inverse cases. With a low-pressure cloud deck at 10^{-5} bar, the preference curves fully flatten out in most cases, showing no more distinct preference over the multipoint retrievals compared to the isothermal case. Only the full-wavelength JWST case still has noticeably lower posterior odds for the isothermal retrievals.

The MIRI LRS simulation is an outlier, showing no discernible preference among the different p–T cases, even in the high-pressure cloud-top cases. It is distinguished from the other simulated spectra by its large wavelength coverage but comparatively low spectral resolution. As we see in Table 3.3, the simulated MIRI LRS spectra cover approximately $\Delta\lambda = 6.7\,\mu\text{m}$, more than four times the wavelength coverage of the narrowest simulated spectra, the NIRCam F322W2 cases. However, the simulated NIRCam spectra have a spectral resolution approximately 30 times that of the MIRI LRS spectra. Both the MIRI LRS and NIRCam F322W2 cases have comparable average transit-depth uncertainties. This indicates that, to recover a p–T profile more complex than an isothermal approximation, spectral resolution plays a more significant role compared to total wavelength coverage.

Retrieval accuracy

Next, we analyse the retrieval accuracy for all self-retrieved parameters. In Figure 3.9, we see a summary of this for each instrument case. For a full overview of each individual self-retrieved parameter, I refer to Appendix B. Generally, the parameter recovery rates follow trends very similar to those of the model preference curves. Isothermal retrievals exhibit the worst performance among all retrieval cases. For most instruments, we see that the recovery rate is on the order of 50% or lower. Only the MIRI LRS and Ariel-WASP cases are outliers in this, showing a better performance for the isothermal profile. Both of them have a comparatively low average spectral resolution, indicating that they contain low amounts of information to infer a vertical temperature structure. Even in these outlier cases, the isothermal profile still performs worse than all other p–T prescriptions. We also observe that the recovery rate generally peaks at the four-point retrievals, which most consistently achieves a near-perfect recovery score. The

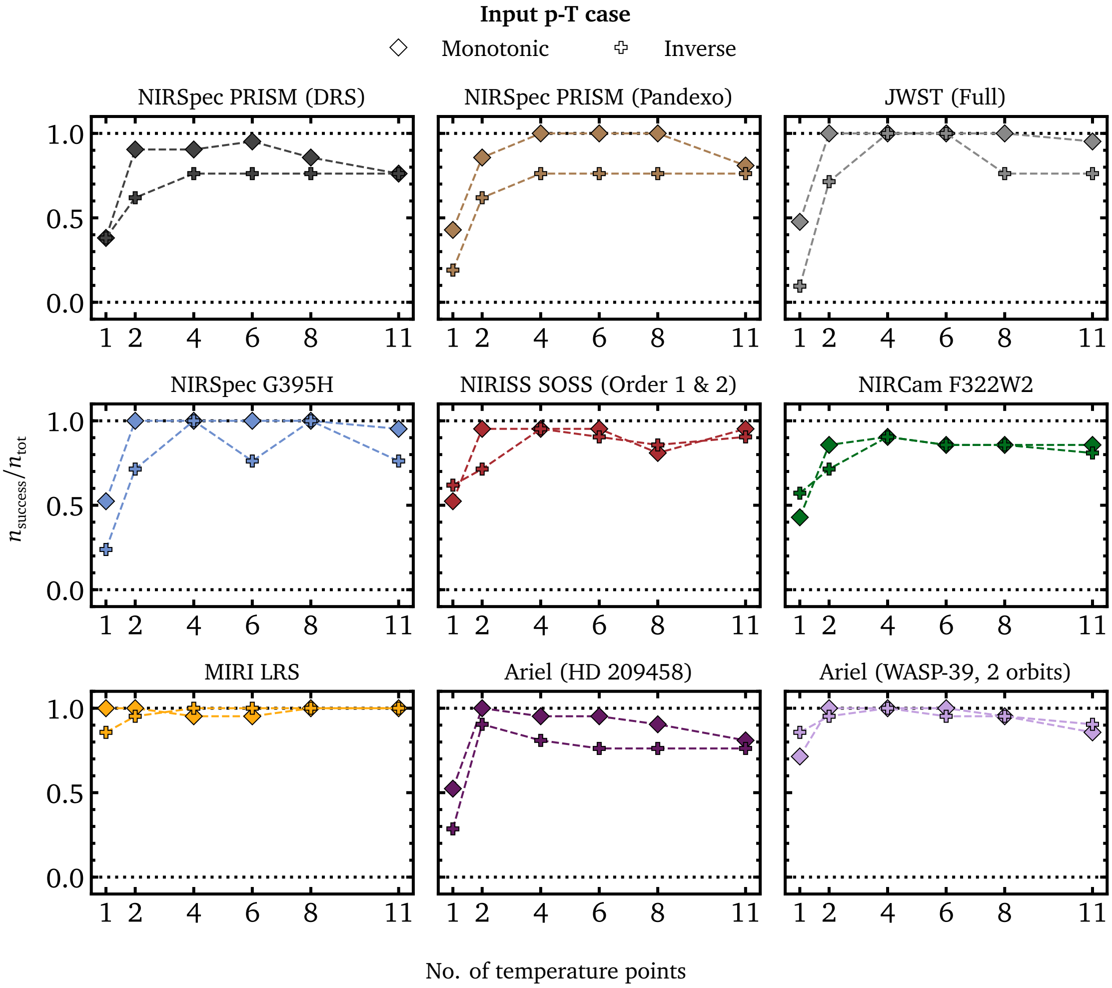

Figure 3.9: Fraction of successfully retrieved forward model parameters for different simulated instrument cases. Each panel shows the number of temperature points used in the retrieval (x-axis) against the normalised number of successfully retrieved parameters (y-axis). Results are split between the monotonic case (diamonds) and the inverse case (pluses).

performance of the two-point profile, as well as of the more complex p–T prescriptions, does not reach this level.

We have already seen that the underlying p–T profile plays a significant role in determining a bias in these self-retrieved parameters. In Figure 3.9, the parameter recovery rates are therefore separated for the monotonic and inverse cases, each with a respective target of $n_{\mathrm{tot}} = 21$ parameters. The general trends in parameter accuracy persist when we examine them separately by input p–T case. However, the performance of the two-point profile drops for the inverse case. This is especially noticeable in the high information-content cases of the full JWST range, NIRISS SOSS, and NIRSpec G395H. For the results in the monotonic case, the two-point profile performs on par with the four-point profile, scoring a recovery rate of at or close to 100%. In the inverse case, this drops to only 70%, while the four-point profile stays at its higher score. As illustrated in the initial NIRSpec PRISM results, the two-point profile most often fails to retrieve the planetary reference radius, R_{p}.

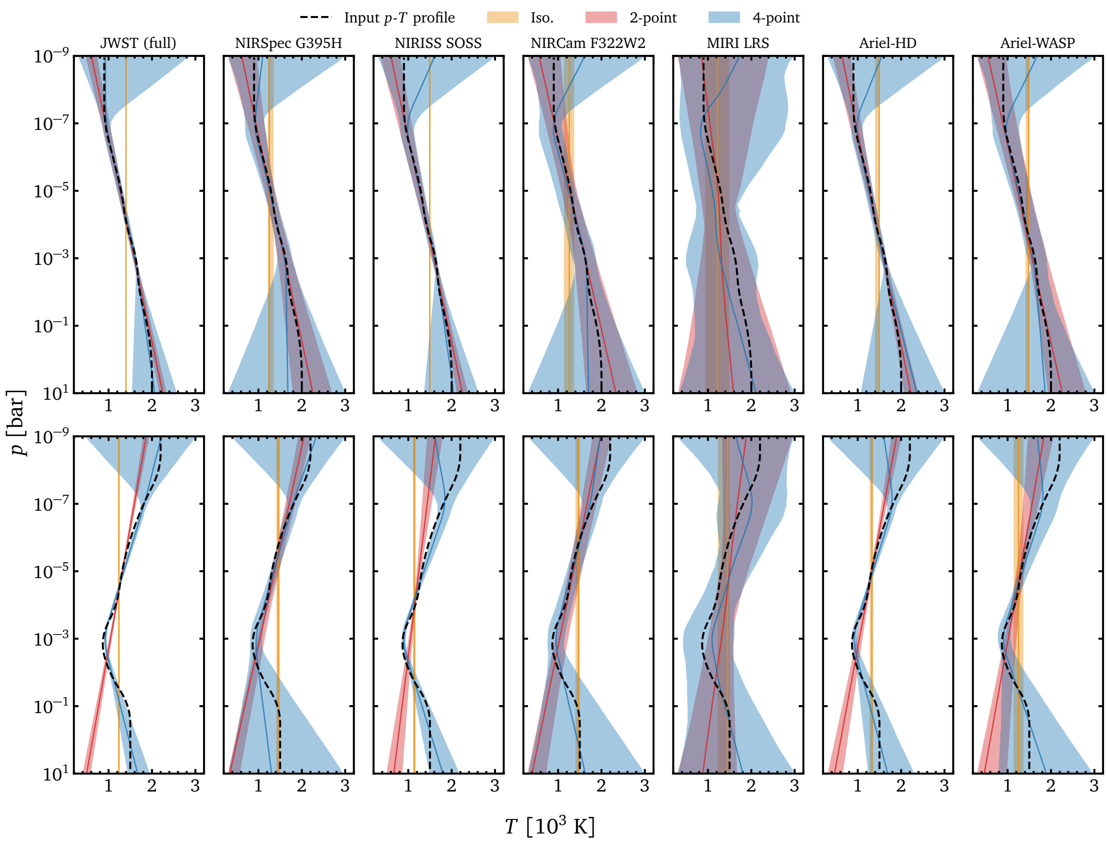

Figure 3.10: Retrieved p–T profiles for additional simulated instrument cases. All underlying spectra were generated with a high-pressure cloud deck. Retrieval results are shown for the isothermal (orange), two-point (red) and four-point (blue) cases. For each retrieved p–T profile, the solid line indicates the median and the shaded area the corresponding $\mathrm{CCI}_{99.7}$. Each panel shows temperature (in K, x-axis) against pressure (in bar, y-axis). Columns are separated by the simulated instrument case indicated in the label, where the top panel shows retrievals on spectra with an underlying monotonic p–T case, and the bottom panel shows the same, but with an underlying inverse p–T case.

We can map the exact nature of these biases by extending our analysis of the retrieved p–T profiles to the underlying input p–T structures across all simulated instrument cases. As before, we focus on the results from the isothermal, two-point, and four-point retrievals. None of the more complex p–T prescriptions shows significant support in terms of their posterior odds, or any advantage in their parameter recovery accuracy. In Figure 3.10, we see the retrieved p–T profiles in the monotonic and inverse case for all instruments. As we have seen in Figure 3.7, increasing the cloud-top pressure consistently broadens the CCIs for the retrieved multipoint p–T profiles. We therefore highlight the retrieved p–T profiles in the high-pressure cloud case, where each spectrum contains the most amount of information.

The general trends we have observed in the model preference behaviour and parameter accuracy curves continue here. Spectra with wide wavelength coverage and high SNRs (the full JWST range, NIRISS SOSS, and Ariel-HD) produce retrieved temperature profiles with very narrow CCIs. In the monotonic case, this shows that both the two-point and the four-point profiles very closely overlap with the input p–T structure. In the inverse case, this narrow retrieved temperature profile results in the two-point profile missing the underlying vertical p–T structure by a significant amount. This contributes to the reduced parameter recovery score for the two-point profile in the inverse case we have observed in Figure 3.9. This mainly manifests in a wrongly retrieved reference radius, as the two-point profile underestimates the temperature at the BOA. As follows from Equation 2.37, it compensates for the less vertically extended atmosphere by increasing the reference radius. MIRI LRS proves again to be an outlier in this, where the retrieved p–T profiles have significantly wider credible intervals. Especially in the monotonic case, we see that the multipoint retrievals span a significant fraction of the prior space ($T \in [300, 3000]$ K), placing next to no constraints on the retrieved p–T structure. This is reflected in the accuracy curve for the MIRI LRS retrievals, which indicates an almost uniform performance of all p–T cases.

3.4 Limitations

To properly contextualise our results, it is important to be aware of the limitations imposed on the investigation we have carried out here. Firstly, we need to consider the setup of our radiative transfer forward model. The underlying p–T profiles we selected were of arbitrary complexity, covering a negative temperature gradient and a more complex structure with a temperature inversion point. Additionally, we chose to model a H_2 / He – dominated atmosphere, which, due to its low MMW, provides a prime target for detailed atmospheric characterisation. The chemical structure of our mock atmosphere was inspired by previously described results for the hot Jupiter WASP-39 b (JTERS23) and did not incorporate physiochemical-informed choices of the atmospheric chemical structure based on the chosen p–T profile.

The way we constructed the chemical profiles of our mock atmosphere also deserves specific attention. We constructed our grid of synthetic transmission spectra assuming MRs that are constant within the model domain. This is important to guarantee that, when performing atmospheric retrievals using the same approximation, we do not introduce any additional biases by oversimplifying the chemical structure. However, the anti-correlated bias between the temperature and chemistry implies that an oversimplified chemical structure could induce the same bias in retrieved temperature. Changeat et al. (2019) showed that using constant MR profiles when the underlying true chemical profile is more complex introduces this degenerate bias in retrieving an isothermal p–T profile.

It is also worth highlighting that, for the multipoint p–T prescriptions, the location of the pressure nodes influences the performance of the associated retrieval, especially in reproducing ground-truth p–T profiles, as is the case here. By fixing the pressure nodes, we inherently risk missing a significant 'feature' of the underlying p–T structure, such as the inversion point of the inverse profile. To counteract this, we could include the pressure nodes of the profile as free parameters as well. However, this effectively doubles the number of free parameters associated with any multipoint profile, significantly increasing the computational requirements of running such retrievals (Changeat et al. 2021). An intermediate solution to this could be to determine the pressure region with the maximum optical depth contribution in situ during the retrieval and cluster any pressure nodes in this region. Waldmann et al. (2015) presented an approach like this for the retrieval of exoplanetary emission spectra, performing a multi-stage approach to dynamically adjust the complexity of the retrieved p–T profile.

Lastly, we have only investigated non-scattered, synthetic data. Datasets derived from real observations will have additional noise characteristics attached to them. Lueber et al. (2024) presented an in-depth analysis of the information content of JWST transit observations of the hot Jupiter WASP-39 b, exploring the question of p–T complexity in the case of real data. In several cases, they found a preference for an isothermal prescription based on a Bayes factor analysis. This is in contrast to the results of this Chapter, which favour implementing a p–T profile at least as complex as a two-point profile. If atmospheric retrievals of real datasets produce inconsistent results based on the complexity of the p–T prescription, it would indicate an unresolved underlying bias. The anti-correlation between the p–T profile and molecular MRs we have found here is one possible source for this. However, other sources of model biases, such as oversimplified chemical profiles (Changeat et al. 2019), three-dimensional effects (Espinoza and Jones 2021; Pluriel et al. 2022), or poorly constrained exoplanetary masses (Changeat et al. 2020; Di Maio et al. 2023), could equally impact the transmission results.

In this Chapter, we have explored how impactful the setup of the forward model can be on the results of atmospheric retrievals. It is a crucial pillar in any inverse Bayesian problem, determining the parameters that represent the data, and which of these are part of the parameter estimation process. However, Bayesian inversion processes must always assume that the data are 'true', since they fully guide the inference process. If there are uncertainties about how the data are prepared, it becomes non-trivial to estimate their impact. How data preparation (or, in other words, reduction) differences propagate into the results of atmospheric retrievals is what I will discuss in the next Chapter.

Chapter 4

Observational limits of retrievals

Variability from data perturbations

"*You don't have a right thing.*
You've got a whole plateful of maybe a little less wrong."

James S. A. Corey

Leviathan Wakes

Declaration of published work

This Chapter includes text already published in Schleich et al. (2025). As a result, there may be partial overlap with this publication.

In this Chapter, we will investigate the role of the input data in the atmospheric retrieval process. In Bayesian inference, the parameter estimation process is fully guided by the underlying dataset and the associated likelihood ascribed to each data point. In this process, we assume that the dataset guiding the inference is 'true', meaning free of systematics and other errors. However, transforming observations of exoplanetary transits into spectra used for inference processes takes a significant number of steps. It is therefore imperative to know how reliable characterisation results are, considering that one single instance of an exoplanet atmospheric spectrum might not be the true instance. Parts of this work were previously published in Schleich et al. (2025).

4.1 Context and motivation

The input dataset guides the statistical sampling of the atmospheric forward model parameters, directly influencing the parameter posterior distributions through the likelihood of the data. With the advent of JWST, the quality of datasets that we can use to infer the atmospheric characteristics of exoplanets has dramatically increased. This has brought with it new challenges that need to be addressed to appropriately adjust the methodology we use to infer atmospheric properties.

The assumptions we make, and the techniques we apply at different stages of the data reduction process, propagate into differences in the resulting atmospheric spectrum and may influence atmospheric characterisation efforts. At the level of different instruments, combining transmission spectra introduces the problem of disagreement between mean transit depths. Accounting for possible offsets between spectra can lead to different conclusions about atmospheric characteristics (e.g., Madhusudhan et al. 2023; Edwards et al. 2024). We can remove this problem by considering only the results of individual instruments. However, this introduces the question of temporal and chromatic binning of the data (e.g., Morello et al. 2022; Davey et al. 2025), and the physical nature of the host star in the form of its limb-darkening (Morello et al. 2017; Keers et al. 2024), which can act as additional sources of biases. At even lower data levels, the utilisation of different data reduction pipelines and assumptions can introduce discrepancies in atmospheric spectra, which propagate into differences in forward model parameter estimates (e.g., Mugnai et al. 2024; Davenport et al. 2025). Maximising the potential of state-of-the-art observatories such as JWST requires us to be aware of, investigate, and, ideally, account for all these different sources of bias.

Our goal in this Chapter will be to investigate how random and systematic differences in a transmission spectrum propagate in atmospheric characterisation results. For this, we examine a transit observation of WASP-39 b (Faedi et al. 2011), a Saturn-mass hot Jupiter orbiting a G-type host star. It is one of the most deeply studied exoplanets by JWST to date. As part of the ERS programme for transiting exoplanets (PID: 1366, PI: N. Batalha, Co-PIs: J. Bean and K. Stevenson), it has been targeted for panchromatic transit observations with all three NIR instruments of JWST. These observations have resulted in the detection of multiple atomic and molecular species in the atmosphere of WASP-39 b, notably the first detection of CO_2 in the atmosphere of an exoplanet (JTERS23), as well as absorption from H_2O, Na, and K (e.g., Ahrer et al. 2023; Alderson et al. 2023; Feinstein et al. 2023; Rustamkulov et al. 2023). However, the tentative detection of sulphur dioxide (SO_2) in the atmosphere of WASP-39 b (Alderson et al. 2023) sparked significant additional interest in this exoplanet. It proved to be the first detection of a product of photochemistry in an exoplanet's atmosphere (Tsai et al. 2023). A DDT programme using MIRI to observe a transit of WASP-39 b in the MIR (PID: 2783, PI: D. Powell) successfully confirmed the identification of absorption signatures by SO_2 (Powell et al. 2024). In sum, WASP-39 b is one of the few exoplanets to date that has a combined panchromatic transmission spectrum covering the full wavelength range of JWST instruments.

However, placing precise constraints on the identified atmospheric constituents of WASP-39 b has proven difficult. Characterisation results depend on the forward model setup and the assumption taken in the data reduction process (e.g., Constantinou et al. 2023; Lueber et al. 2024). To investigate this sensitivity in detail, we will first produce a transmission spectrum of WASP-39 b from scratch by reducing the raw NIRSpec PRISM observation of its transit. We then analyse the impact of perturbations in the dataset used for atmospheric retrievals on the parameter posterior distributions.

4.2 Observational data

The dataset we analyse here is a NIRSpec bright object time series (BOTS), using the prism as the dispersion element. It covers one primary transit of WASP-39 b, and was taken on 10 July 2022 (14:04 – 23:38 UT). This observation spans approximately 8 hours, encompassing 21 500 integrations, each with five groups. Our data reduction process starts with the raw observational data (i.e., the non-calibrated Stage 1b, or `.uncal`, files) obtained from the MAST. As a primary input dataset, we produce our own transmission spectrum from this raw observational data. To test how sensitive the results of atmospheric retrievals are to random perturbations of the input data, we create self-scattered instances of this transmission spectrum. We then compare these results to already available spectra derived from the same raw data, using the same pipeline.

4.2.1 Data reduction

To perform the data reduction, we use the pipeline `Eureka!` (Bell et al. 2022), which itself is based on the `jwst` pipeline (Bushouse et al. 2024). As both pipelines were, and still are, under heavy development, it is imperative that we version-control our results. For documentation purposes, I therefore note that we use `Eureka!` v0.10[1] and `jwst==1.8.2`, with the CRDS context `jwst_1202.pmap`. Apart from the steps highlighted below, we follow the standard JWST data reduction routine for TSOs described in Section 2.1.2, using the default settings of `Eureka!`.

[1] specifically `eureka==0.10.dev0+g3c10926.d20230426`

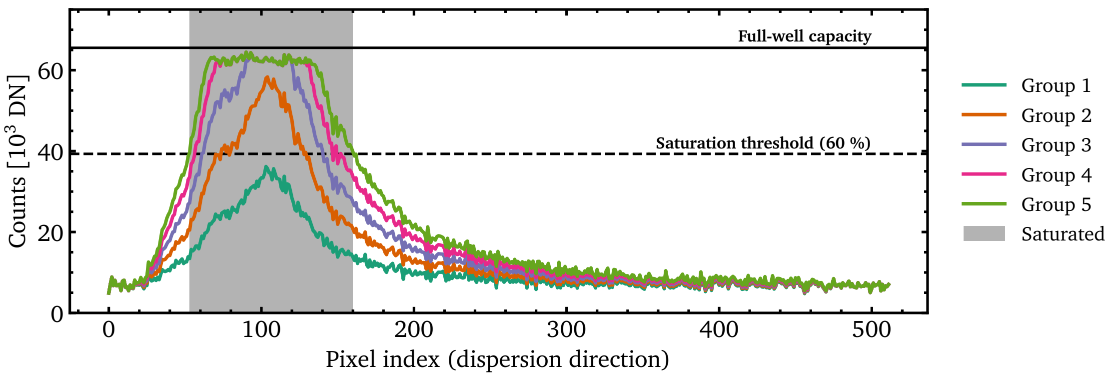

Figure 4.1: Integration-averaged detector counts for the centre of the spectral trace, showing detector x-position (x-axis) against count value (y-axis). The values are colour-coded by the five groups in this observation. The solid and dashed black lines denote the full-well capacity and 60% of it, respectively. The grey-shaded area denotes the region where the signal exceeds the 60% threshold for group 5.

Data reduction and calibration

We start at stage 1 of `Eureka!` with the raw observational data of the JWST observation mentioned above. It is noteworthy that in this observation, saturation was intentionally achieved at shorter wavelengths to increase the SNR at longer wavelengths (JTERS23). In Figure 4.1, we see the integration-averaged detector counts for pixels in detector row 16 (approximately where the signal peak is located) for each group. At group 5, full-well saturation is reached for $x \in [80, 120]$. We will use this to restrict the detector region we consider for spectral extraction. We skip the reference pixel correction, as there are no reference pixels in this subarray (Birkmann et al. 2022). For the jump rejection step, we use a threshold of 10σ to reject outliers, rather than the default 4σ. For observations with as few groups as is the case here, a threshold as small as 4σ leads to excessive fractions of detector pixels being flagged as outliers (JTERS23; Rustamkulov et al. 2023). A higher threshold circumvents this problem. We also perform the group-level background subtraction (GLBS) step introduced in `Eureka!` v0.7, which accounts for $1/f$-noise introduced during detector readout (JTERS23). We define the background region for this step as being outside of $y \in [5; 22]$. After inspecting the detector images of this observation, we also manually mask several pixels that, at the time of data reduction, are not flagged as bad pixels[2]. To synthesise the background, we fit a second-order polynomial to each row of the detector, using an outlier rejection threshold of 5σ. We skip the gain-scale step, which is not relevant for our relative measurements, and then collapse the group-level data to integration-level data by fitting the observational ramps, transforming the data into count-rate detector images.

We then proceed to stage 2, where we perform additional calibration and unit-conversion steps. At this stage, we skip the flat field step – at the time of data reduction, the reference files for this were incomplete, meaning that this step did not work as intended (e.g. Alderson et al. 2023; Sarkar et al. 2024). We also leave out two more steps that are not necessary for

[2]This pixel map can be found at `https://doi.org/10.5281/zenodo.15697940`

our purposes – the photometric calibration (from count rate to flux density) and 1D spectral extraction at this stage. At the end of stage 2, we have applied the wavelength map to our two-dimensional countrate images.

After these steps, we proceed to stage 3 of `Eureka!` to extract the one-dimensional spectra from the data. At this point, we constrain the data to non-saturated regions of the detector. We conservatively set 60% of the full-well capacity as the saturation threshold to avoid strong nonlinearity effects. This restricts the data to pixels falling into $x \in [160; 512]$ (in the dispersion direction), as we see in Figure 4.1. On this restricted data cube, we then perform a double-iteration outlier rejection step along the time axis, with a threshold of 5σ. We then perform the 'optimal spectral extraction' routine implemented in `Eureka!`. For this, we use an aperture half-width of six pixels, centred on pixel row 16, and a background half-width of nine pixels. This is the optimal combination to keep the spectral aperture half-width as large as possible (Horne 1986), while maximising the overall spectral precision. In this step, we construct the spatial profile using the median frame and perform outlier rejection with a threshold of 10σ for the spatial profile and 60σ for the optimal spectral extraction routine.

Extraction of spectroscopic light curves

In stage 4, we extract the light curve data from our observation, both in spectroscopic form and as an integrated white-light curve. We restrict the data to wavelengths below 5.3 µm as the throughput of PRISM becomes negligible beyond this range (Jakobsen et al. 2022). Combined with the restriction we imposed on the lower end of the subarray in the previous stage, this effectively constrains our wavelength range to $\lambda \in [2.24, 5.30]$ µm. We perform a final round of outlier rejections in the time direction here, using a rolling median from a box with a width of 100 data points, using a rejection threshold of 5σ. After this, we extract the light curves from each pixel column of the detector image, and normalise each light-curve by its median.

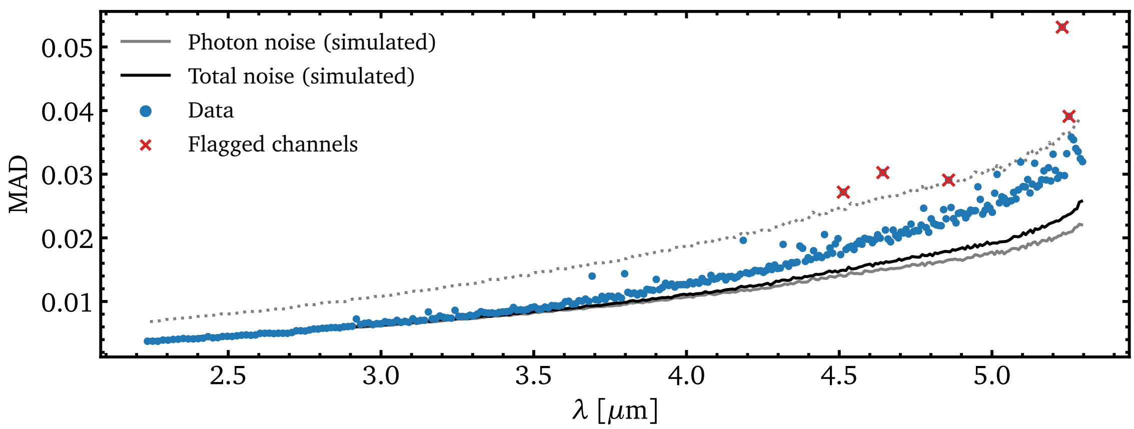

Figure 4.2: Spectroscopic light-curve precision, showing wavelength (in µm, x-axis) against median absolute deviation (MAD, y-axis). Blue points represent the measured light-curve precision. The solid grey and black lines represent the simulated photon noise and total noise, respectively. The grey dotted line indicates 1.75 times the photon noise, and channels exceeding this limit are marked with red crosses.

These light curves might still be influenced by errors in the data, such as problematic pixels we did not account for or spurious events. As a last correction mechanism, we therefore compare the precision of our extracted light curves with the expected precision from an instrument simulation. We measure this precision as the median of absolute differences (MAD) for each time series,

$$\mathrm{MAD} = \mathrm{median}(|\mathrm{x}_{\mathrm{t}+1} - \mathrm{x}_{\mathrm{t}}|), \tag{4.1}$$

where t represents the time axis, and x_t and x_{t+1} are two successive data points. Figure 4.2 shows these channel-level precision values, compared to the expected noise levels from an instrument simulation. We perform this simulation using the JWST Exoplanet Simulator (JexoSim, Sarkar and Madhusudhan 2021), considering all noise sources included in JexoSim and ten realisations. The simulation setup mimics the real observational setup with five groups per integration and a fixed bin size of one pixel. The system parameters of WASP-39 b are queried from within JexoSim. In Figure 4.2, the solid black line shows the precision predicted from JexoSim using all noise sources. In comparison, the solid grey line indicates the contribution of photon noise. It is readily apparent that photon noise is the dominant noise component over almost the entire wavelength regime. Only towards the longer end of our wavelength regime do we see that the total noise decouples from the photon noise contribution. This is where we expect to see an increased importance of read-out noise (Sarkar and Madhusudhan 2021). For our purposes, we consider detector columns to be excessively noisy if the light curve precision falls outside of a factor of 1.75 times the expected photon noise (marked as a dotted grey line in Figure 4.2).

Extracting transit-depth information from our data will require us to fit a transit model to the light curves. For non-chromatic parameters (such as the transit mid-point time, t_0, or inclination, i), we make use of a channel-integrated white-light curve with much higher precision. To construct this light curve, we exclude the wavelength channels flagged during the precision analysis (red crosses in Figure 4.2) and perform an additional time-series outlier rejection, using a rolling median with a threshold of 3σ.

Light curve fitting

To fit both the white-light curve and the spectroscopic light curves we have extracted, we use a combined model that accounts for the astrophysical signal and systematics. The parameters of this model are summarised in Table 4.1.

As described in Section 2.1.2, the astrophysical model depends on the radius-ratio of the exoplanet and host star, R_p/R_*, the transit mid-point time, t_0, the exoplanet's orbital inclination, i, eccentricity, e, and period, P, the argument of the orbit's periapsis, ω, and the scaled semi-major axis, a/R_*, summarised in the parameter vector $\boldsymbol{p}_\mathrm{transit}$. These parameters form the baseline of the light-curve model, which is further modified by the non-uniform intensity of the stellar disk. We account for this stellar limb-darkening by using the four-parameter law

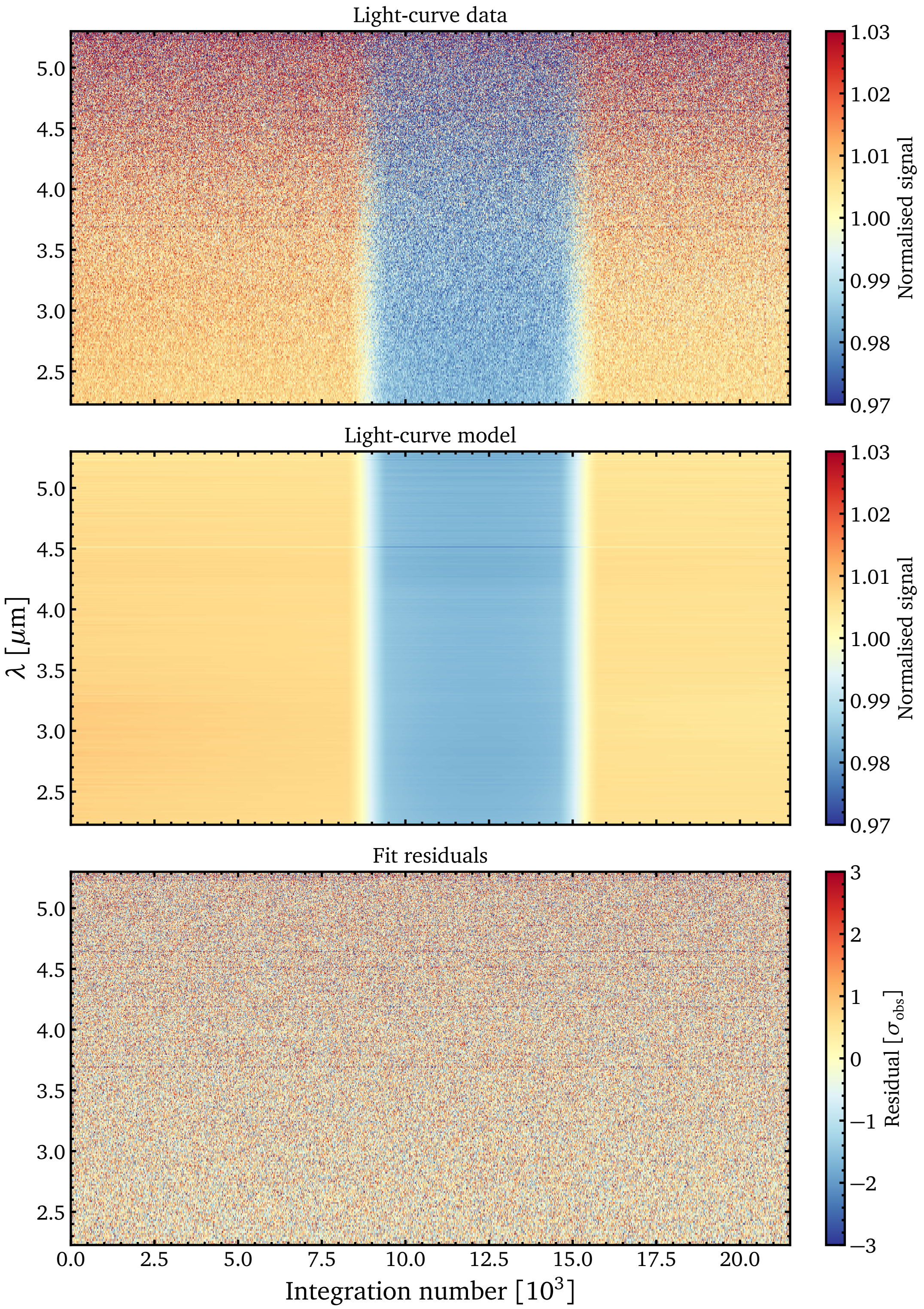

Figure 4.3: Dynamic light-curve fit results for the NIRSpec PRISM transit of WASP-39 b. All panels show integration number (x-axis) against wavelength (y-axis, in µm). (Top) Observational data, where the colour map indicates the median-normalised signal. (Middle) Panchromatic transit light-curve model, where the colour map represents the median-normalised signal. (c) Residual map between the observation and model. The colour map shows residuals normalised against the measurement uncertainty, σ_{obs}.

Table 4.1: Light-curve fitting parameters. For the application of the combined model parameters, 'all' refers to the parameter being free in both the fitting of the spectroscopic and the white-light curves, 'white' indicates that the parameter is free in the white-light curve fit and fixed for the spectroscopic light-curve fits, and 'fixed' denotes parameters that are fixed in all cases. The values for normally distributed priors are taken from Mancini et al. (2018).

Parameter	Unit	Prior / Value	Application
Astrophysical model			
R_p	R_*	$\mathcal{N}\left(0.148, 0.015^2\right)$	all
t_0	d	$\mathcal{U}\left(59770.81, 59770.86\right)$	white
i	deg	$\mathcal{N}\left(87.83, 0.25^2\right)$	white
a	$\mathrm{R}_\odot$	$\mathcal{N}\left(11.4, 1.0^2\right)$	white
P	d	4.0552765	fixed
e	–	0	fixed
ω	deg	90	fixed
Systematics model			
c_0	–	$\mathcal{N}(1, 0.05^2)$	all
c_1	–	$\mathcal{N}(0, 0.01^2)$	all
c_2	–	$\mathcal{N}(0, 0.01^2)$	all

described in Claret (2000), where the intensity of the stellar disk is defined as

$$\frac{I(\mu)}{I(1)} = 1 - \sum_{k=1}^{4} a_k \left(1 - \mu^{k/2}\right). \tag{4.2}$$

Here, a_k represents the LDCs, and $\mu = \cos(\theta)$, where θ represents the angle between the line of sight and the emerging intensity. We choose the four-parameter law over the commonly employed quadratic law, as the latter has been shown to introduce biases in the retrieved transit depth (Morello et al. 2017; Keers et al. 2024). The LDCs of this four-parameter prescription are fixed, and pre-calculated from a grid of 3D stellar models (`Stagger`, Magic et al. 2015) using the Python package `ExoTiC-LD` (Grant and Wakeford 2024). The last part of our light-curve model requires us to account for instrument systematics. We fit a quadratic polynomial, S, with coefficients c_0, c_1, and c_2 to each of the spectroscopic and to the white-light curve, accounting for temporal systematic trends. Our final light-curve model has the form

$$F(t) = F_\mathrm{transit}(t; \boldsymbol{p}_\mathrm{transit}, a_1, a_2, a_3, a_4) \times S(t; c_0, c_1, c_2) \tag{4.3}$$

We then fit each light-curve model using the MCMC algorithm `emcee` (Foreman-Mackey et al. 2013). For the white-light curve, we perform a fit using seven free parameters $(R_\mathrm{p}, t_0, i, a, c_0, c_1, c_2)$. In contrast, we fit the spectroscopic light curves using four free parameters $(R_\mathrm{p}, c_0, c_1, c_2)$. We fix the values of t_0, i, and a to the results of the white-light curve fit, as we can assume that they will not change with respect to different wavelengths. The fit parameters are summarised in Table 4.1. Globally, we use 20 walkers with 5 000 steps and a burn-in phase of 500 steps. We analyse the convergence of the fit for each channel using the mean autocorrelation time (mACT), calculated across all walkers and parameters. The individual spectroscopic channels have an mACT between approximately 40 and 50 steps, and the white-light curve has an mACT of 60 steps, which we can deem sufficient to ensure convergence (Foreman-Mackey et al. 2013). In Figure 4.3, we see a summary of the light-curve fitting results in the form of a two-dimensional light curve. It shows the signal (in the form of the median-normalised flux) with respect to both time and wavelength. The top panel represents the observational data. In the centre of the panel, we clearly see the transit of WASP-39 b as a panchromatic reduction in the observed signal. We also see that the measurements become noisier at longer wavelengths, where the SNR decreases. The central panel shows the fitted light-curve model for each channel, and the bottom panel shows the resulting fit residuals, scaled to the associated uncertainty value of each data point, σ_{obs}.

Extraction of the transmission spectrum

The last step in our data reduction routine is the extraction of the transmission spectrum. We do this at the pixel-resolution level, where we have performed the light-curve fitting. We then use these values to produce the transmission spectrum using a weighted arithmetic mean (WAM) with a fixed bin size of three pixels. This accounts for the typically assumed resolution element size of 2.2 pixels for NIRSpec (Jakobsen et al. 2022). The final transmission spectrum is calculated as

$$\bar{\delta}_{\mathrm{p}} = \frac{\sum_{i=1}^{3} w_i\, \delta_{\mathrm{p},i}}{\sum_{i=1}^{3} w_i} \quad \text{with} \quad \sigma_{\bar{\delta}} = \frac{1}{\sum_{i=1}^{3} w_i}, \tag{4.4}$$

where $\bar{\delta}_{\mathrm{p}}$ and $\delta_{\mathrm{p},i}$ represent the binned and pixel-level transit-depth values, respectively, and $\sigma_{\bar{\delta}}$ describes the uncertainty of $\bar{\delta}_{\mathrm{p}}$. We define the associated weights, w_i, as the inverse of the variance values, σ_i^2, associated with each δ_i.

It is noteworthy that the MCMC sampling we have performed to fit the light-curve models has yielded a posterior distribution for each model parameter, including R_{p}. `Eureka!` reports the associated parameter estimates as CCI_{68}. These intervals are not necessarily symmetric around the median. However, the likelihood we use to perform atmospheric retrievals (Equation 2.49) is a normal distribution. Consequently, we make the approximation of determining σ_i for each pixel-level transit depth as the arithmetic mean between the lower and upper wings of the CCI. Recently, Davey et al. (2025) investigated the impact of asymmetric error bars on the atmospheric retrieval of JWST-quality data. They reported that an average asymmetry of approximately 80% can critically influence atmospheric retrievals. The asymmetry in the error bars from our light-curve fitting results reaches a maximum of 7.5%, which is well below this threshold. We therefore deem that our approximation should not significantly impact the retrieval results.

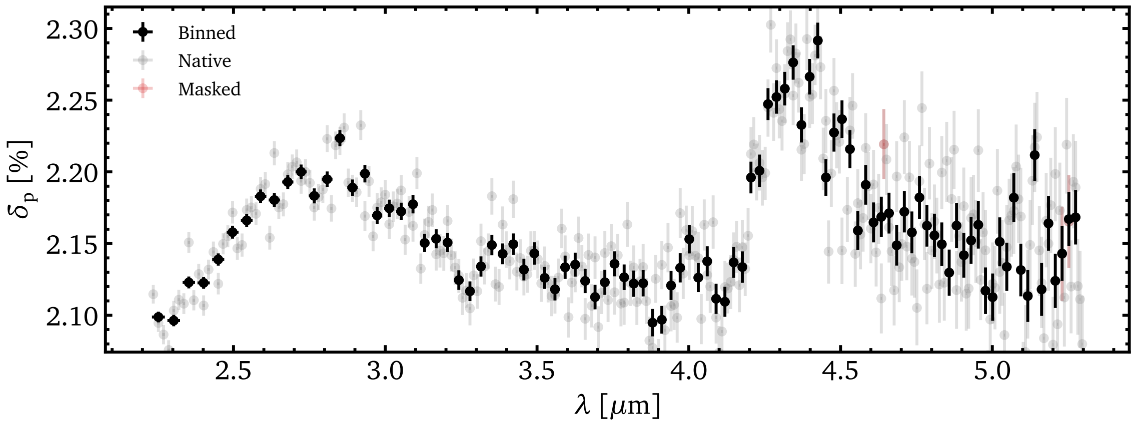

Figure 4.4: NIRSpec PRISM transmission spectrum of WASP-39 b, showing wavelength (in µm, x-axis) against transit depth (in %, y-axis). Grey data points represent the transit-depth values extracted at the detector's pixel resolution. Black data points show the transmission spectrum binned with a fixed bin size of three pixels. Data points marked in red are excluded from the binning process.

We see the final transmission spectrum resulting from our data reduction process in Figure 4.4, together with the underlying pixel-resolution spectrum. The data points marked in red are the channels we have found to be excessively noisy (see also Figure 4.2), which we exclude from the binning process. Henceforth, we label this spectrum as 'SP-TW' (for *Sp*ectrum - *T*his *W*ork).

4.2.2 Panchromatic perturbation

While we assume that our spectral instance represents the 'true' state of the observation, unaccounted biases can influence the results of our parameter estimation process. This is compounded by the fact that the individual data points of a transmission spectrum influence the posterior distributions in a non-uniform manner. To judge the reliability of our retrieval results, we therefore assess how strongly our parameter estimates depend on these data points.

One method used to judge the weight of individual data points in the parameter estimation process is the leave-one-out cross-validation (LOO-CV) technique. It works by fitting a forward model to a dataset with one datum removed (Gelman et al. 2014). With the typical number of data points from pre-JWST data, this would require tens of retrievals. However, with the current state-of-the-art data, this jumps to several hundreds to thousands of retrievals at native JWST instrument resolution. Current implementations of LOO-CV make use of numerical approximations to circumvent this computational boundary (Vehtari et al. 2017; Welbanks et al. 2023; Murphy et al. 2025). An additional complication for tests like this is that atmospheric absorbers produce correlated signals within an atmospheric transmission spectrum through their panchromatic absorption cross-sections. We have seen an example of this in Figure 2.6. A comprehensive investigation into the dependence of the posterior distributions on the underlying data would therefore also require us to perform retrievals excluding all possible combinations of data points. For a spectrum with n data points, this would necessitate 2^n retrievals, making this computationally infeasible for native-resolution transmission spectra

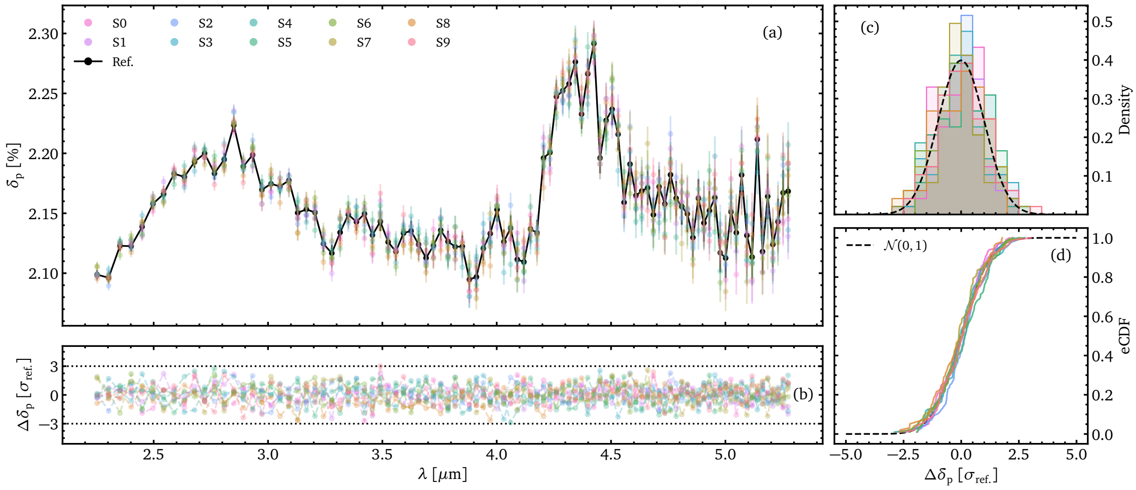

Figure 4.5: Summary of the scattered instances of SP-TW. Results from scattered instances are colour-coded consistently in all panels. (a) Scattered transmission spectra, showing wavelength (in µm, x-axis) against transit depth (in %, y-axis). The original instance of SP-TW (labelled 'Ref.') is shown in black. (b) Differences between the scattered instances and the reference spectrum, showing wavelength (in µm, x-axis) against residuals (in units of $\sigma_{\mathrm{Ref.}}$, y-axis). The dotted black lines represent levels of $\pm 3\sigma_{\mathrm{Ref.}}$. (c) Distributions of normalised residuals of the scattered instances, binned in steps of $0.5\sigma_{\mathrm{ref.}}$. (d) Empirical cumulative distribution functions (eCDF) for the normalised residuals. In panels (c) and (d), the black dashed line represents the PDF and CDF of a standard normal distribution, respectively.

from JWST.

Therefore, we opt to perform a simplified test to initially assess the stability of our retrieval results to perturbations in the underlying dataset. We generate a set of ten new spectra by successively creating randomised instances of SP-TW. We create each of these instances by individually drawing new transit-depth values from a normal distribution, $\mathcal{N}(\mu_{\mathrm{td},\lambda}, \sigma^2_{\mathrm{td},\lambda})$, where ($\mu_{\mathrm{td},\lambda}$, $\sigma_{\mathrm{td},\lambda}$) represent the pairs of transit-depth mean and standard deviation making up SP-TW. While this test is not sensitive to correlated noise, it assesses the stability of posterior distributions under perturbations in the data following a normal distribution. If we want to consider this in the context of perturbations stemming from the data reduction process, we can interpret these scattered instances as deviations following random noise. We see the original instance of SP-TW and all ten scattered instances in Figure 4.5, together with the residuals between the new mean transit-depth values and the baseline they were scattered from (normalised by $\sigma_{\mathrm{td},\lambda}$). Applying a one-sample Kolmogorov-Smirnov (K-S) test to these residuals gives us p-values between 0.14 and 0.97, providing us with the sanity check that all scattered instances of SP-TW really are self-scattered from a normal distribution.

4.2.3 Published transmission spectra

In Section 4.2.1, we have derived a transmission spectrum of WASP-39 b from a single-transit NIRSpec PRISM observation. However, previous analyses of the same dataset already exist. We know that the underlying data guide Bayesian inference processes. If the same data is prepared independently using ostensibly the same method (i.e., independent reductions of observational

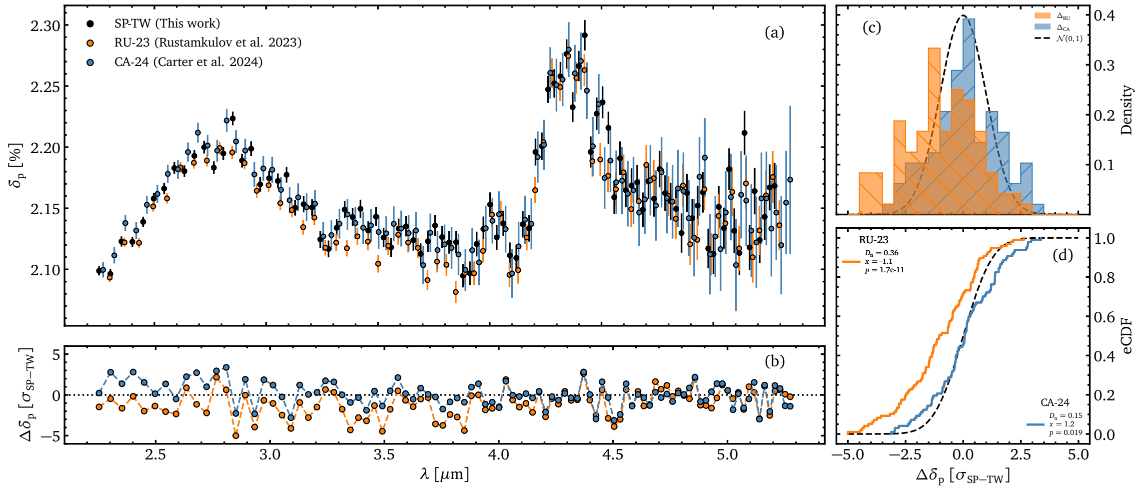

Figure 4.6: Comparison of existing transmission spectra of WASP-39 b. In each panel, results associated with SP-TW, RU-23, and CA-24 are colour-coded in black, orange, and blue, respectively. (a) Transmission spectra, showing wavelength (in µm, x-axis) against transit depth (in %, y-axis). (b) Differences between RU-23 and CA-24 against SP-TW, showing wavelength (in µm, x-axis) against residuals (in units of $\sigma_{\mathrm{SP-TW}}$, y-axis). The dotted line indicates $\Delta\delta_{\mathrm{p}} = 0$. (c) Distributions of normalised residuals of RU-23 and CA-24, binned in steps of $0.5\sigma_{\mathrm{SP-TW}}$. (d) Empirical cumulative distribution functions (eCDF) for the residuals of RU-23 and CA-24. The results of a one-sample K-S test against a standard normal distribution are provided in the legend of the figure. In panels (c) and (d), the black line represents the PDF and CDF of a standard normal distribution, respectively.

data using the same pipeline), then we have to assume that homogeneous inferences applied to all these data will give us the same answer. However, this is not always the case. Mugnai et al. (2024) showed that spectra derived with different pipelines can systematically bias atmospheric retrieval results. Similarly, Edwards et al. (2024) highlighted the impact of changes in data reduction assumptions on uniform analyses of multi-instrument datasets. Here, we go one step further down this ladder and compare the results of data reductions performed with the same reduction pipeline. We therefore want to perform comparative atmospheric retrievals. Using SP-TW as a baseline, we perform homogenised atmospheric retrievals on existing transmission spectra derived from the same observational data to investigate the reliability of our retrieval results. These spectra are summarised in Figure 4.6.

Rustamkulov et al. (2023)

The first NIRSpec PRISM transmission spectrum of WASP-39 b was presented in JTERS23. However, this spectrum was derived only for wavelengths longer than 3 µm due to conservative estimations on detector saturation. The first spectrum covering the full PRISM wavelength range was presented in Rustamkulov et al. (2023), with a specific focus on recovering the saturated region of the detector. Of the four pipeline results presented in their work, we use the `Eureka!`-based reduction to achieve the closest possible overlap with our data reduction[3]. Henceforth, we will refer to this spectrum as 'RU-23'.

[3]The data for this were retrieved from https://doi.org/10.5281/zenodo.7388031

Carter & May et al. (2024)

Carter & May et al. (2024) presented a reanalysis of all NIR transit observations of WASP-39 b from the ERS programme. Out of the results presented in their work, we use the transmission spectrum derived at native instrument resolution (`bin_scale1`) with fixed LDCs[4]. It is worth noting that the `Eureka!`-based reduction presented in their work is based on the 'baseline' spectroscopic time series from Rustamkulov et al. (2023), which was derived with the `FIREFLy` data reduction pipeline. The closest overlap we can achieve between the results from Carter & May et al. (2024) and SP-TW therefore begins with the light-curve fitting. From here onwards, we will refer to this spectrum as 'CA-24'.

Comparison of the spectra

Table 4.2 lists the differences between the data reduction steps applied in this work to derive SP-TW, the `Eureka!-ExoTEP` data reduction presented in Rustamkulov et al. (2023) to derive RU-23, and the `FIREFLy-Eureka!` data reduction presented in Carter & May et al. (2024) to derive CA-24. I note that, in the case of CA-24, the steps prior to the light-curve fitting refer to the data reduction of `FIREFLy`-based dataset discussed in Rustamkulov et al. (2023). Steps not listed are performed under the same assumptions in all three cases.

In their original forms, both RU-23 and CA-24 cover the full wavelength range of NIRSpec PRISM, including the saturated part of the detector. For our analysis, we constrain both spectra to the wavelength range covered by SP-TW (approximately 2.2 µm to 5.3 µm) to enable a fair comparison with respect to the contained information. The global shape of all three spectra closely overlaps. Two main absorption peaks, a broad one centred at 2.7 µm and a narrower one centred at 4.4 µm, are clearly identifiable in all three cases. A minor increase in transit depth, centred at 4.1 µm and covering only a few data points, is also visible in all three spectra, but more ambigiously than the two major features.

We do note that the wavelength maps of all three spectra do not fully coincide. We are not comparing the transmission spectra at the pixel level, but rather at the resolution of the instrument. Consequently, differences in binning will propagate into slightly offset wavelength bin centres. To evaluate the wavelength-dependent differences between SP-TW and the other two spectra, we calculate the point-wise residuals by linearly interpolating RU-23 and CA-24 onto the wavelength map of SP-TW. Applying a one-sample K-S test to the normalised residual distributions shows us that neither of them is consistent with a standard normal distribution. The p-value of the K-S performed on the RU-23 residuals can reject this underlying hypothesis with a significance of more than 3σ. We also clearly see this in Figure 4.6. The distribution of the RU-23 residuals is clearly offset from 0 (panel c), and the resulting empirical CDF is shifted to the left, indicating an overall negative offset. In the transmission spectrum, this manifests most prominently in the region from approximately 2.8 µm to 3.9 µm, where the transit depth of RU-23 is consistently smaller than both CA-24 and SP-TW.

Applying the K-S test to CA-24 shows that the hypothesis of its residuals following a normal

[4]The data for this were retrieved from https://doi.org/10.5281/zenodo.10161742

Table 4.2: Main data reduction differences between SP-TW, RU-23, and CA-24. CRDS context `jwst_1202.pmap` was used to derive SP-TW.

Step	SP-TW	RU-23	CA-24
`jwst` pipeline version	1.8.2	1.6.0	1.6.2
Stage 1 and 2			
Bias subtraction	`jwst_nirspec_superbias_0299`	Custom	Custom
Dark current subtraction	Yes	Yes	No
Reference pixel correction	No	Top / bottom 6 px	Top / bottom 6 px
Jump rejection	$10\,\sigma$	No	No
GLBS	2nd order polynomial	Median (top / bottom 6 px)	Yes, but unspecified
Stage 3			
Spectral half-width	6 px	4 px	Variable
Time series outliers	100 px box-car ($5\,\sigma$)	20 px box-car ($3\,\sigma$)	unspecified
Light-curve fitting			
Systematic trend	Quadratic	Linear	Linear
Limb-darkening	4-parameter (fixed)	Quadratic (fitted)	Quadratic (fitted)
Error bar inflation	No	Yes ($\bar{\chi}^2_\nu$ to 1)	unspecified
Pre-fit binning	No (pixel-level)	No (pixel-level)	Instrument resolution

distribution is rejected at a level of 2σ. The differences between it and SP-TW are more subtle than for RU-23. However, in the empirical CDF we see that the tails of the CA-24 residual distribution deviate from the expected normal distribution CDF. The fact that neither of the two residual distributions is consistent with a normal distribution implies that both of them show systematic differences compared to SP-TW. As we see in Figure 4.6 (panel b), these differences are most pronounced at short wavelengths. This is a result of the much smaller transit-depth uncertainties in this wavelength regime. Towards longer wavelengths, these differences vanish as the error-bar sizes increase for all spectra.

4.3 Atmospheric retrieval of WASP-39 b

Before we perform comparative atmospheric retrievals on both the randomly scattered instances of SP-TW and the already existing transmission spectra of WASP-39 b, we need to define a uniform model setup for all cases. This will allow us to robustly compare the posterior distributions resulting from these retrievals.

4.3.1 Method

We perform all atmospheric retrievals using TauREx, with a global setup closely overlapping the retrieval methodology we have used in Chapter 3.

Retrieval setup

To retrieve the atmospheric characteristics of WASP-39 b, we make use of the WASP-39 system parameters we have already seen in Table 3.1. We define the atmospheric forward model in the pressure regime $\log_{10}(p\,[\mathrm{bar}] \in \{1, -9\}$ with 110 atmospheric layers and using H_2 and He as background gases, with $He/H_2 = 0.13$. Based on the results in Chapter 2, we choose a heuristic multipoint profile with four fixed pressure nodes to represent the atmospheric p–T structure. We place these nodes at $\log_{10}(p_i\,[\mathrm{bar}]) \in \{1, -3, -7, -9\}$. We define the atmospheric profiles of additional trace species with vertically constant MRs. In the radiative transfer calculations, we consider molecular absorption using cross-sectional data from the Exomol project (Tennyson et al. 2020; Chubb et al. 2021), HITRAN (Gordon et al. 2022), and HITEMP (Rothman et al. 2010), as well as CIA from H_2–H_2 (Abel et al. 2011) and H_2–He (Abel et al. 2012; Fletcher et al. 2018) pairs, and Rayleigh scattering as incorporated in TauREx. We also consider the opacity contribution from a flat-opacity cloud deck defined through the cloud-top pressure, p_{cloud}. The free parameters of our forward models are the reference radius, R_{p}, the temperature values at the four pressure nodes, T_i, the individual molecular MRs, X_{M}, and the cloud-top pressure, p_{cloud}. These parameters, and the priors we use for them, coincide with the parameters given in Table 3.4. For all retrievals, we use 700 live points and an error tolerance of 0.5 for the natural logarithm of the evidence.

Model tuning

The reference radius, p–T profile, and cloud prescription are fixed parts of our model setup. However, we tune the atmospheric forward model by investigating the preferred inventory of trace gases. To do this, we start with a baseline model containing CO (Li et al. 2015), CO_2 (Yurchenko et al. 2020), H_2O (Polyansky et al. 2018), and CH_4 (Yurchenko et al. 2024). These species are major spectrally active C- and O-bearing gases in H_2–He dominated atmospheres in the temperature regime we expect for WASP-39 b (e.g., Mollière et al. 2022), and have previously been reported in characterisations of the atmosphere of WASP-39 b (e.g., Tsiaras et al. 2018; Wakeford et al. 2018; JTERS23). We then investigate the inclusion of additional molecular opacity sources by comparing the baseline model's performance with that of models with an extended molecular inventory.

We judge this performance based on three metrics. Firstly, the Bayes' factor, $\ln(B_{\mathrm{m},0})$, between an altered model, m, and the baseline, index with '0'. We consider an altered model to have considerable support against the baseline if $\ln(B_{\mathrm{m},0}) > 3$, which corresponds to posterior odds of more than 20 : 1 in favour of the altered model. This initial search builds a list of extended models preferred over the baseline. We then perform a second round of model tuning by considering all unique combinations of the preferred molecules, and evaluate $\ln(B_{\mathrm{m},0})$ for each of these combinations. Next to the Bayes factor, we also use the corrected AIC, Ψ, and the reduced Chi-square square value, χ^2_ν. These two values provide comparative metrics based on point estimates of the retrieval results, namely the maximised likelihood and median model for Ψ and χ^2_ν, respectively.

Parameter estimation results

To summarise the weighted posterior samples resulting from the NS parameter inference, we again use CCIs derived from the marginalised parameter traces. For the retrievals we perform here, we calculate CCI_{95} for each retrieved parameter.

It is noteworthy that a common credible interval size used in the evaluation of atmospheric retrieval results is a centred interval containing 68% of the posterior probability mass, or CCI_{68}, which is then reported as a '1σ' interval. The implied equivalence between the Bayesian and frequentist result statistic only holds true when the associated posterior distribution is close to being a normal distribution. If it is heavily asymmetric, then the edges of the σ-equivalent CCIs do not simply scale with the z-score. Consequently, the '1σ' credible interval can induce false confidence in the parameter estimation result. This compounds with limitations on parameter accuracy originating from current stellar and planetary models (e.g., Rackham et al. 2018; MacDonald et al. 2020), current opacity models (Niraula et al. 2022; 2023), and from inherent model uncertainties (Barstow et al. 2020; Nixon et al. 2024).

4.3.2 Retrievals using SP-TW

To determine the appropriate molecular inventory of our baseline atmospheric forward model containing $CO, CO_2, H_2O,$ and CH_4, we test the inclusion of the hydrocarbons C_2H_2 (Chubb

Table 4.3: Model performance metrics for the model tuning performed on the SP-TW dataset. The columns show the ID and corresponding molecular inventory of the forward models used for model tuning. The Bayesian evidence, $\ln(E)$, is used to calculate the Bayes factor, $\ln(B_{\mathrm{m},0})$, in reference to the baseline model, M_0. The cAIC, Ψ, is used to calculate the relative metric $\Delta_m = \Psi_{\min} - \Psi_m$ (where $\Psi_{\min}$ corresponds to M_{20}). The last two columns indicate the number of free parameters in the model, as well as the reduced χ-square metric, χ^2_ν, of the median model.

ID	Model description	$\ln E$	$\ln B_{\mathrm{m},0}$	Ψ	Δ_m	# FP	$\bar{\chi}^2_\nu$
			Baseline				
0	CO, CO_2, CH_4, H_2O	642.02	–	-1306.02	55.00	10	3.26
1	Removed CO	634.52	-7.50	-1304.88	56.14	9	3.26
2	Removed H_2O	515.48	-126.54	-1063.76	297.26	9	6.00
3	Removed CH_4	642.73	0.71	-1309.45	51.57	9	3.21
4	Removed CO_2	193.96	-448.06	-423.07	937.95	9	13.28
			Initial opacity search				
5	Added SO_2	647.65	**5.63**	-1323.51	37.51	11	3.06
6	Added H_2S	661.40	**19.38**	-1350.08	10.95	11	2.75
7	Added HCN	647.20	**5.18**	-1322.10	38.92	11	3.08
8	Added C_2H_2	648.95	**6.92**	-1321.43	39.60	11	3.09
9	Added C_2H_4	640.31	-1.72	-1303.85	57.17	11	3.29
10	Added CH_3	641.84	-0.18	-1303.80	57.22	11	3.29
11	Added NH_3	641.02	-1.00	-1303.70	57.32	11	3.29
			Combinations of $\{C_2H_2, HCN, H_2S, SO_2\}$				
12	Added $\{H_2S, HCN\}$	663.67	21.65	-1348.69	12.33	12	2.77
13	Added $\{H_2S, SO_2\}$	662.52	20.49	-1353.36	7.66	12	2.72
14	Added $\{HCN, SO_2\}$	651.16	9.14	-1329.87	31.15	12	2.99
15	Added $\{C_2H_2, SO_2\}$	656.27	14.24	-1345.07	15.95	12	2.81
16	Added $\{C_2H_2, HCN\}$	649.54	7.52	-1327.72	33.30	12	3.02
17	Added $\{C_2H_2, H_2S\}$	663.65	21.63	-1354.81	6.21	12	2.70
18	Added $\{H_2S, SO_2, HCN\}$	664.08	22.06	-1351.26	9.76	13	2.74
19	Added $\{C_2H_2, HCN, SO_2\}$	657.63	15.60	-1344.19	16.83	13	2.83
20	Added $\{C_2H_2, H_2S, SO_2\}$	666.06	24.03	**-1361.02**	–	13	2.63
21	Added $\{C_2H_2, H_2S, HCN\}$	664.01	21.99	-1352.96	8.06	13	2.72
22	Added $\{C_2H_2, H_2S, HCN, SO_2\}$	665.88	23.85	**-1358.32**	2.70	14	2.66

et al. 2020) and C_2H_4 (Mant et al. 2018), the methyl radical CH_3 (Adam et al. 2019), the nitrogen-bearing compounds HCN (Barber et al. 2014) and NH_3 (Coles et al. 2019), and the sulfur-bearing compounds H_2S (Azzam et al. 2016) and SO_2 (Underwood et al. 2016). The results of this model tuning process are summarised in Table 4.3 in the form of $\ln(B_{\mathrm{m},0})$, Ψ, and χ^2_ν.

Firstly, we perform a validation test by running atmospheric retrievals individually, excluding the baseline molecules. Removing CO_2 and H_2O produces the most significant differences in $\ln(B_{\mathrm{m},0})$, with values of -446.06 and -126.54, respectively. This is an expected result – in the transmission spectrum of WASP-39 b, we can visually identify the absorption features of CO_2

(centred at 4.4 µm and H_2O (centred at 2.7 µm). Consequently, removing these two molecules from the forward model results in significantly worse model performance. Removing CO is also signifcantly disfavoured, with $\ln(B_{m,0}) = -7.5$. The removal of CH_4 leads to an inconclusive Bayes factor, not favouring the adjustment of the baseline to a significant degree. As a result, we leave the baseline model intact when searching for additional molecular contributions.

Next, we perform our initial model tuning step. We find a strong model preference (translating to posterior odds of more than 150:1) for models where the baseline molecular inventory has been individually extended with H_2S, SO_2, HCN, and C_2H_2. Out of these, including H_2S has the strongest effect on the Bayes factor, resulting in a value $\ln(B_{m,0}) = 19.38$. A combined model containing all four of these molecules is also very strongly favoured, with $\ln(B_{m,0}) = 23.58$ At this point, it is worth noting that the bottom-up model-tuning approach we use here contrasts with the top-down approach described in Benneke and Seager (2013). When we add contributions to a small baseline, the value of the Bayes factor can be exaggerated, leading to the spurious identification of molecules (Welbanks et al. 2026). We see this behaviour in the initial extension of the baseline model. However, we also perform retrievals with all possible combinations of the molecules selected in the first step. This allows the molecules to compete with each other, and confirms that the fully extended model is still preferred to a sufficiently large degree. All models containing H_2S produce the biggest increase in posterior probability compared to the baseline. This is also reflected in the χ^2_ν of the respective models. When considering all three model performance metrics, we see that model M_{22}, extended with $\{H_2S, SO_2, HCN, C_2H_2\}$, and model M_{20}, extended with $\{H_2S, SO_2, C_2H_2\}$, perform on an equivalent level. They result in, respectively, a Bayes factor of 23.58 and 24.03, and a χ^2_ν of 2.66 and 2.63. M_{20} is the model with the smallest Ψ, making it the most preferred out of the whole ensemble based on the corrected AIC. However, the Δ_m between M_{20} and M_{22} is 2.70. This makes M_{22} the only model with any significant support compared to M_{20} based on Ψ, as the next-smallest value is 6.21 for M_{17}, already classified as having considerably less support.

We adopt the fully extended model M_{22}, containing molecular absorption contributions from $\{CH_4, C_2H_2, CO, CO_2, HCN, H_2O, SO_2\}$, as the fiducial model. We are not making any claims about the detection or detection significance of any atmospheric constituents based on the model-tuning process, beyond the posterior probabilities associated with the Bayes factor. Using the Bayes factor to derive σ-values and report the detection of atmospheric constituents runs the risk of misrepresenting its relative nature (Schmidt et al. 2025; Welbanks et al. 2026). Here, we want to analyse the impact of perturbations in the transmission spectrum on parameter posterior distributions derived from atmospheric retrievals. With an equal performance between M_{22} and M_{20}, we therefore choose the fully extended model including HCN to provide us with an additional point of comparison.

Retrieval results

With a fiducial model in place, we can now analyse the associated retrieval posteriors and parameter estimates. Figure 4.7 provides us with a comprehensive overview of the retrieved

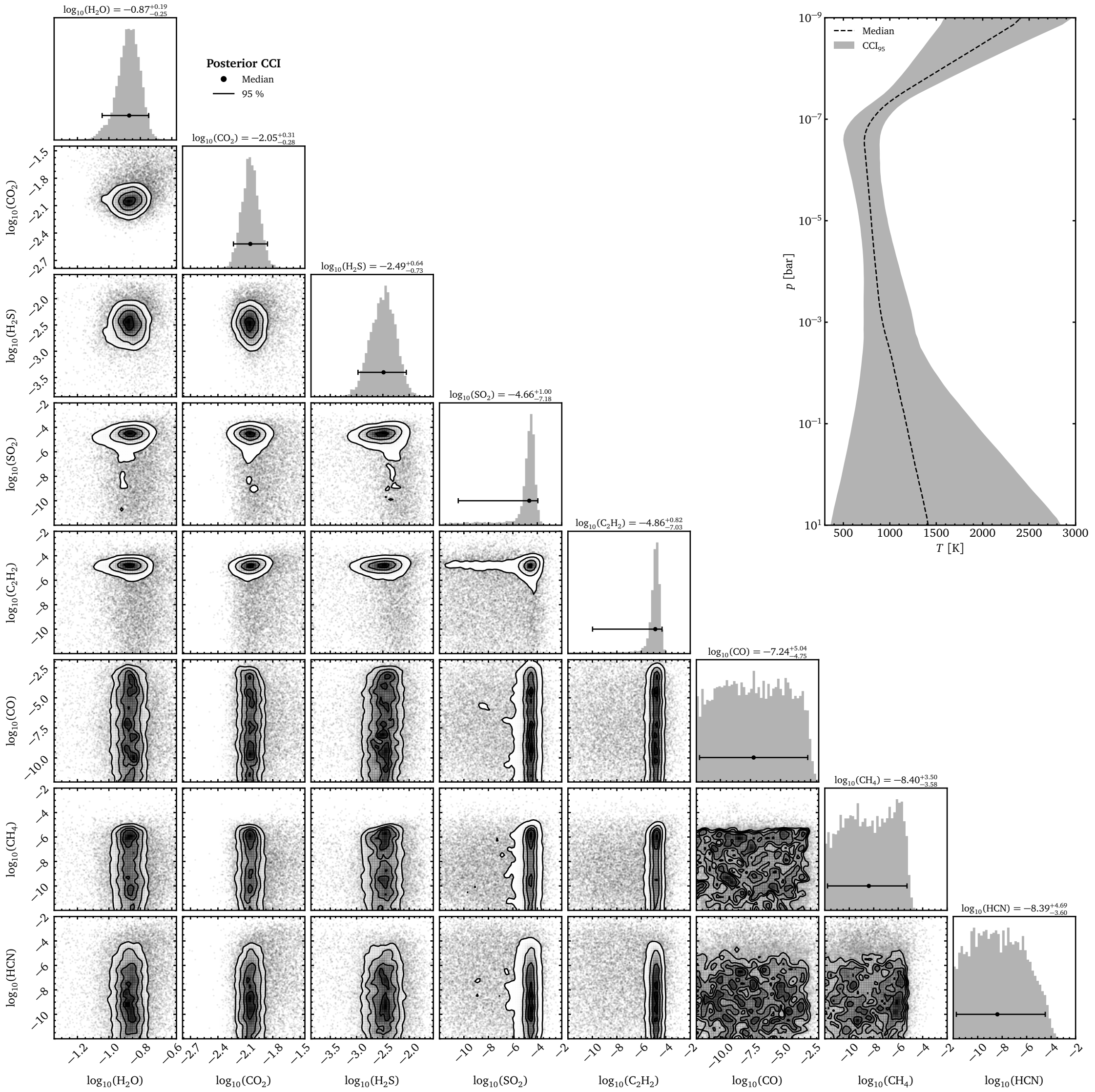

Figure 4.7: Atmospheric retrieval results of the fiducial model applied to SP-TW, showing the posterior distributions of the molecular MRs. The main diagonal shows the associated marginalised posteriors with resulting parameter estimate medians (points) and CCI_{95} (error bars). The inset plot in the top right shows the median retrieved p–T profile (dashed line) and CCI_{95} (shaded region).

molecular MRs and the associated p–T profile. In Figure 4.8, we see the CCI_{95} of the total model (top panel), as well as the individual molecular contributions to the median model (bottom panel). The parameter estimates of all model parameters are listed in Table 4.4.

The main molecules that shape the transmission spectrum are $H_2O, CO_2,$ and H_2S. We have already been able to visually identify the broad absorption of H_2O, which extends from 2.3 µm to 3.5 µm. The resulting inferred MR is $\log_{10}(X_{H_2O}) \in [-1.04, -0.75]$. The absorption by H_2O also dominates beyond 4.7 µm, where another major cross-sectional peak then falls outside the wavelength range of our transmission spectrum. Similarly, CO_2 produces the most prominent feature in the transmission spectrum, centred at 4.4 µm. A secondary absorption peak of it is visible at approximately 2.8 µm. From the atmospheric retrieval, we infer a MR of

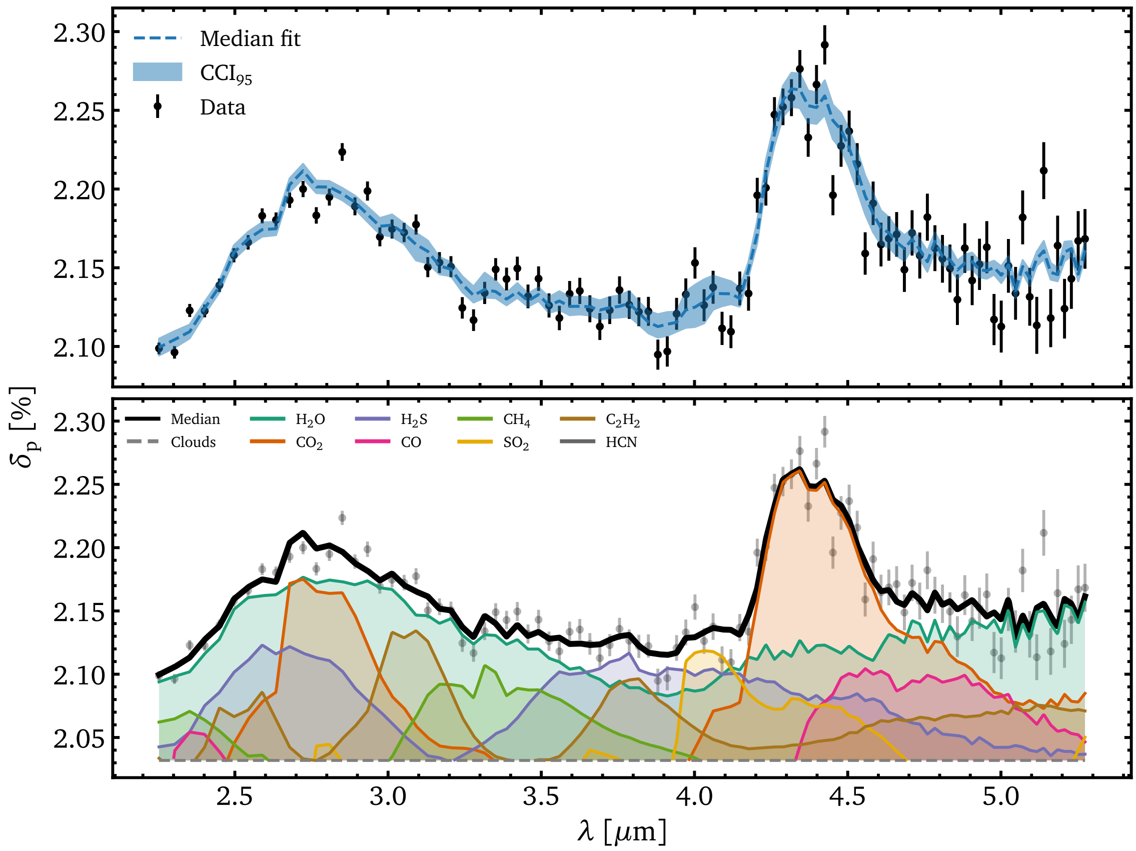

Figure 4.8: Median model fit of the fiducial model to SP-TW. Both panels show wavelength (in µm, x-axis) against transit depth (in %, y-axis). (Top) Median model (dashed blue line) and CCI_{95} (shaded area) of the fiducial model. SP-TW is shown as black data points. (Bottom) Contributions to the median model of individual molecular opacity sources (colour-coded by molecule). The flat-opacity cloud deck is indicated with a dashed grey line. The total median model is shown in black, and SP-TW as grey data points.

$\log_{10}(X_{CO_2}) \in [-2.23, -1.86]$. The third dominant constituent, both by its inferred MR value and its contribution to the optical depth, is H_2S. It shows a broad absorption feature from 3.5 µm to 4.5 µm (where it has the dominant contribution to the transmission spectrum), and a narrower feature centred at 2.6 µm. Its inferred MR is $\log_{10}(X_{H_2S}) \in [-2.97 - 2.05]$. All three of these species have posterior distributions that are close to normal distributions, and the CCI_{95} constrains them to within 1 dex. For H_2O and CO_2, this is further narrowed to 0.5 dex This indicates that the underlying transmission spectrum contains sufficient information to confidently constrain their molecular MRs.

In contrast to these well-constrained MRs, we only find upper limits for CO and CH_4. In their marginalised posteriors, this manifests as close-to-uniform distributions with sharp edges at the respective upper limit. In Figure 4.8, we see that CO has no significant absorption features in the wavelength range of our transmission spectrum. Its main contribution comes from the region longwards of 4.5 µm. Consequently, the credible interval of the CO MR is very wide, encompassing most of the prior range. We can place its upper limit at $\log_{10}(X_{CO}) \leqslant -2.74$. Compared to CO, CH_4 has a well-known absorption feature centred at 3.4 µm, which we can see in the contributions to the median model. Because the transmission spectrum does not prominently display this feature, the posterior distribution of CH_4 is more constraining in its upper limit, which we can place at $\log_{10}(X_{CH_4}) \leqslant -5.28$.

Lastly, our fiducial model also contains SO_2, C_2H_2, and HCN. For all three of these molecules, the parameter estimates encompass a wide range of the prior space, with corresponding CCI_{95} sizes of approximately 5.5 dex to 7.0 dex. This indicates that the MRs of these molecules are not very well constrained by the underlying data. However, we also note a significant difference in the shape of the marginalised posterior distributions of these molecules. For HCN, the case is similar to CO and CH_4. Its uniform posterior distribution points to an upper limit in the parameter estimate, with $\log_{10}(X_{HCN}) \leqslant -4.48$. The contribution of HCN to the median model is not visible in Figure 4.8 because its low median MR places it below the cloud deck. As a reference, Figure 2.6 illustrates the cross-section profile of HCN. In contrast, the marginalised posteriors of SO_2 and C_2H_2 are close to normal distributions at first glance. SO_2 produces a distinct absorption feature at 4.0 µm, and C_2H_2 has two minor absorption peaks at approximately 3.1 µm and 3.8 µm, the latter of whch sits closely on top of H_2S. The skewed nature of their posterior distributions is underlined by the fact that the median values of their CCIs are $\log_{10}(X_{SO_2}) = -4.66$ and $\log_{10}(X_{C_2H_2}) = -4.86$, respectively, which sit very close to the upper boundary of their CCI_{95}. In these cases, calculating a CCI with a width of 68% and declaring it a '1σ' uncertainty would provide a misleading sense of confidence in the parameter estimate. If we do this for SO_2, the $CCI_{1\sigma}$ (written as a commonly used point estimate and uncertainty value) is $\log_{10}(X_{SO_2}) = -4.66^{+0.37}_{-0.85}$. This scales the true lower boundary of the CCI_{95} to 7σ, rather than the actual equivalent of 2σ. When reporting parameter estimates from atmospheric retrievals, this highlights the importance of properly calculating and indicating the corresponding credible interval, rather than scaling an apparent σ-value.

In Figure 4.7, we also see the retrieved four-point p–T profile of WASP-39 b, indicated by its median (dashed line) and CCI_{95} (shaded area). The temperature at the two intermediate pressure nodes (placed at 10^{-3} and 10^{-7} bar, respectively) is well constrained, with a CCI-width of approximately 500 K. The temperature profile in this region behaves very closely to an isotherm. In contrast, the temperature values at the two boundary nodes, placed at the TOA and BOA, are much less constrained. At the TOA ($p = 10^{-9}$ bar) the credible interval of the temperature has a width of 1500 K, encompassing half the prior space. In this pressure region, the atmosphere is fully transparent. Consequently, there is little information to constrain the temperature value. The absorption of XUV radiation is expected to heat the thermospheres of hot Jupiters by significant amounts (Fortney et al. 2021), making this behaviour of the p–T profile a plausible possibility. However, the setup of our atmospheric forward model indicates that this could also be a model degeneracy. As we have seen in Chapter 3, the biasing behaviour of simplified p–T profiles is symmetric, indicating that the constant molecular MR profiles we use in our forward model could cause this temperature behaviour. The temperature at the BOA ($p = 10$ bar) is even less constrained. Its CCI_{95} has a width of 2400 K, covering almost the entire prior range. We can trace this to the constraint of the flat-opacity cloud deck, which is constrained to $\log_{10}(p_{\text{cloud}}\,[\text{bar}]) \leqslant -3.29$. Consequently, pressure regions below this threshold are not accessible in the transmission spectrum, resulting in an unconstrained parameter posterior.

Contextualisation with previous retrievals

We contextualise the parameter estimates from our fiducial model by comparing them to previously reported analyses of the atmospheric characteristics of WASP-39 b. For consistency, we compare our results to previous works that (1) have characterised a transmission spectrum of WASP-39 b, and (2) have used a 'free chemistry' approach with constant MRs.

Prior to the launch of JWST, one of the major species accessible in transmission spectroscopy performed with HST was H_2O. Tsiaras et al. (2018) analysed two transits of WASP-39 b observed with HST WFC3, inferring a H_2O mixing ratio of $\log_{10}(X_{\text{H}_2\text{O}}) = -5.94 \pm 0.61$. This is significantly lower than the MR resulting from our retrieval. They also infer an isothermal temperature of $T = 1258.71 \pm 289.53$. This is more compatible with the constrained temperature region from our retrieval. In contrast to their results, Wakeford et al. (2018) analysed a combined dataset from HST STIS and the WFC3, as well as Spitzer IRAC and VLT FORS2 data to infer $\log_{10}(X_{\text{H}_2\text{O}}) = -1.37^{+0.05}_{-0.13}$. They also derive an isothermal temperature of $T = 920^{+70}_{-60}$ K. These results are in closer agreement with our parameter estimates, and the inferred MR of H_2O is significantly different from the results of Tsiaras et al. (2018). Analysing the same dataset, Welbanks et al. (2019) report an ever larger value of $\log_{10}(X_{\text{H}_2\text{O}}) = -0.65^{+0.14}_{-1.83}$, shrinking to $\log_{10}(X_{\text{H}_2\text{O}}) = -2.43^{+0.27}_{-0.24}$ with a constrained prior distribution. They use a six-parameter p–T prescription (Madhusudhan and Seager 2009), but do not directly comment on the respective retrieved p–T profiles. While our reported results agree with some of these previous analyses, the range of reported H_2O abundances covers more than five orders of magnitude. This indicates poor overall agreement in the retrieval results for HST-era observations.

The advent of JWST has enabled access to a much larger range of potential atmospheric constituents through their absorption features in the NIR and MIR. Constantinou et al. (2023) reported results from atmospheric retrievals performed on the approximately 3–5 µm, `Eureka!`-reduced NIRSpec PRISM transmission spectrum of WASP-39 b first presented in JTERS23. They report MRs of $\log_{10}(X_{\text{H}_2\text{O}}) = -3.29^{+0.59}_{-0.56}$, $\log_{10}(X_{\text{CO}_2}) = -5.11^{+0.63}_{-0.54}$, and $\log_{10}(X_{\text{H}_2\text{S}}) = -4.11^{+0.49}_{-0.46}$ as the dominant atmospheric trace species with prominent absorption features. These abundances are smaller by two to three orders of magnitude compared to our retrieval results. Constantinou et al. (2023) also report $\log_{10}(X_{\text{SO}_2}) = -6.40^{+0.39}_{-0.35}$, and notably $\log_{10}(X_{\text{CO}}) = -4.11 \pm 0.61$. These more stringent abundance constraints lie within the broader posterior distribution reported in our results. The parametric p–T profile (Madhusudhan and Seager 2009) retrieved in their work is consistent with an isotherm with $T = 738^{+55}_{-54}$ K at the top of their atmospheric forward model domain ($p = 10^{-6}$ bar), which is in agreement with the p–T profile we retrieve from SP-TW. It is worth pointing out that the results reported by Constantinou et al. (2023) vary by approximately 1 dex when considering different pipelines or when extending the input transmission spectrum with data from HST WFC3.

Lueber et al. (2024) studied the information content of all NIR transmission spectra of WASP-39 b taken as part of the ERS. When using the full-range NIRSpec PRISM data presented in Rustamkulov et al. (2023), they report value of $\log_{10}(X_{H_2O}) = -3.10^{+0.20}_{-0.19}$. They also find a tighter constraint on the MR of CO and SO_2, with values of $-2.85^{+0.17}_{-0.28}$ and $-5.68^{+0.31}_{-0.62}$, respectively, as well as MRs of CO_2 and H_2S smaller by approximately 2 orders of magnitude compared to our results. When changing the cloud model in their retrievals, Lueber et al. (2024) find changes of approximately 0.5 dex in their reported MRs, as well as a loss of preference for the inclusion SO_2 and CO in their forward model. The latter stands in stark contrast to the otherwise reported tight posterior constraints. In both cloud-model cases, Lueber et al. (2024) report an isothermal temperature of approximately 1000 K within a pressure domain extending to 10^{-6} bar, which is in close agreement with the p–T profile we retrieve from SP-TW in this region.

It is also worth noting that all the results we have compared our posterior distributions to were reported as point estimates with uncertainties equivalent to a $\mathrm{CCI}_{1\sigma}$. We cannot feasibly reproduce the corresponding posterior distributions of each previous result to derive a CCI_{95}. However, if we recall the example of the CCI_{68} for the SO_2 abundance in our fiducial model, calculating proper credible intervals rather than scaling the reported uncertainties could significantly change the associated parameter estimates. This is also especially true if the previously reported results are heavily asymmetric (such as the H_2O MR in Welbanks et al. 2019). Consequently, the disagreement between these results and our retrievals might be lessened with properly homogenised credible intervals.

4.3.3 Sensitivity to random scatter

We now want to test the robustness of the parameter posterior distributions from our fiducial model. As we have seen, the posterior distributions for SO_2 and C_2H_2 are heavily asymmetric, with long tails towards low abundance values. Because the inference process is driven by the underlying data, this could indicate that these two parameters are estimated from only a few data points. Indeed, if we examine the contributions of SO_2 and C_2H_2 to the median fiducial model, we see that the dominant parts of their absorption features extend over only approximately five data points, respectively. To test the parameter posterior behaviour under random noise, we therefore apply our fiducial model to ten self-scattered instances of SP-TW.

We see a summary of all resulting marginalised posterior distributions in Figure 4.9. These distributions, and consequently their derived parameter estimates, can be categorised into three types. Firstly, we find stable and well-constrained posteriors, which do not change significantly under random perturbations of the input spectrum. In the context of the molecular abundances of our fiducial model, this refers to H_2O, CO_2, and H_2S. In all scattered instances, their respective marginalised posteriors are close to normal distributions, and their derived CCIs agree with the parameter estimates from the initial instance of SP-TW. H_2S produces an individual outlier in this (scattered instance 7), where the posterior distribution appears normal, but the credible interval skews towards lower abundances. However, even in this case, the width

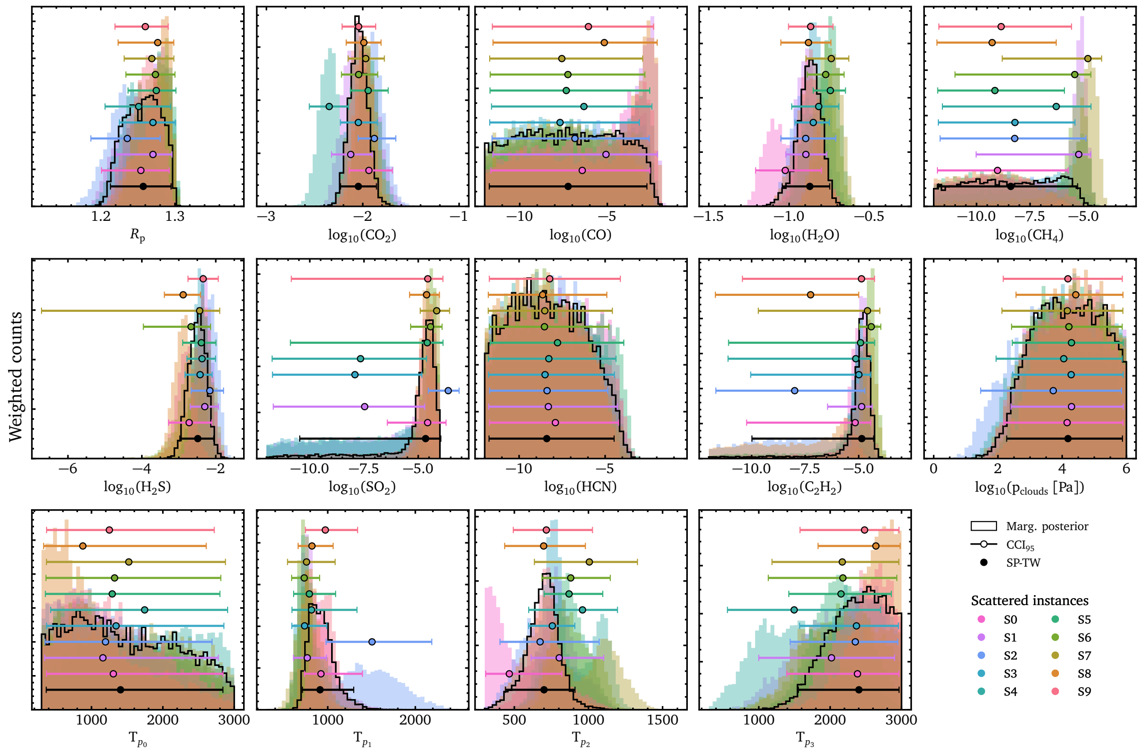

Figure 4.9: Marginalised posterior distributions from atmospheric retrievals on scattered instances of SP-TW. Each panel shows the sampled parameter values (x-axis) against the weighted counts (y-axis). Results from the reference instance (SP-TW) are indicated in black, while results from the scattered instances are colour-coded. Data points and error bars indicate the median and CCI_{95} of each marginalised posterior distribution.

of the CCI is approximately 4 dex, which is still narrower than the other skewed distributions we have identified (by about 2 dex). The abundances of all stable, well-constrained molecules (H_2O, CO_2, and H_2S) are constrained by contributions across wide wavelength ranges. The reference radius, R_p, also falls in the category of stable and well-constrained.

The second category of posteriors we identify are stable, unconstrained posteriors. These distributions consistently extend to one of the edges of the prior space, and follow a uniform shape until a sharp cut-off. Consequently, parameter estimates derived from these cases are in agreement with each other. The MRs of CO and HCN fall into this category. As they are not mapped to any significant absorption features over the wavelength range of our spectrum, perturbing it does not change the parameter posterior estimate. The posterior distributions of CO produce visually identifiable peaks at approximately $\log_{10}(X_{CO}) = -2$ in some instances, but the distributions are not skewed enough to draw the lower edge of the CCI_{95} away from the boundary of the prior space. The cloud-top pressure, p_{cloud}, also falls in this category.

Lastly, we find marginalised posterior distributions that are unstable and skewed, significantly changing in shape when the underlying spectrum is scattered. Most notably, this is the case for the parameter estimates of the SO_2 and C_2H_2 MRs. They only produce absorption features in small parts of the transmission spectrum. Consequently, scattered spectral instances

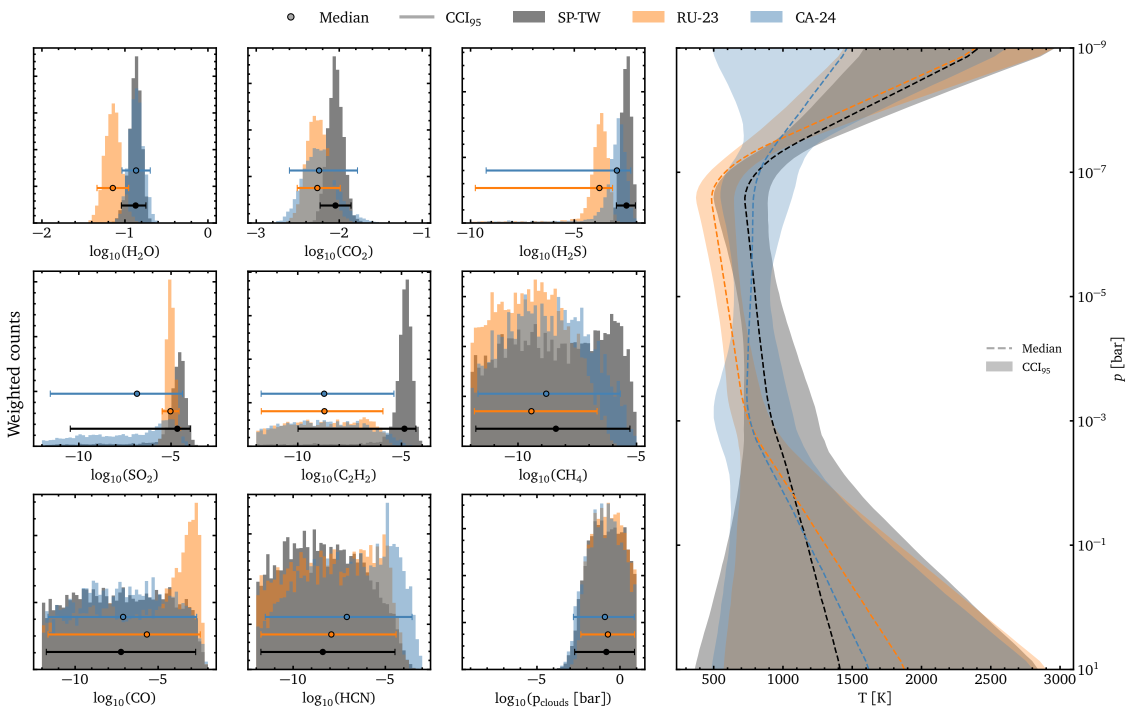

Figure 4.10: Results of atmospheric retrievals performed on SP-TW (black), RU-23 (orange), and CA-24 (blue). (Left) The grid of smaller panels shows the marginalised posterior distributions of the molecular MRs and cloud-top pressure. Data points and error bars represent the distribution median and CCI_{95}, respectively. (Right) Retrieved four-point p–T profiles, represented by the median (dashed line) and CCI_{95} (shaded area).

will induce strong variations in the derived credible intervals if they influence these regions. Next to the heavily-tailed posterior distributions we have identified in the retrieval of the initial instance of SP-TW, we see several instances of narrow parameter estimates for SO_2 and C_2H_2 in Figure 4.9. This mirrors the well-constrained MRs of H_2O, CO_2, and H_2S. In contrast, several other instances in both cases turn into upper limits, with broad CCI_{95} ranges and a centred median. To a lesser extent, this is also true for CH_4. In the retrieval of the initial instance of SP-TW, CH_4 was inferred as an upper limit. Several of the scattered instances produce a parameter estimated that more closely resembles the tailed cases of SO_2 and C_2H_2. Even in these unstable cases, we find that all CCI_{95} values are in agreement with the results retrieved from the original instance of SP-TW. We can conclude that random perturbations in the input data do not produce disagreement in parameter estimates. However, the derived parameter constraints could be overconfidently small. Assigning specific MR values to atmospheric trace species based on individual retrieval results should be treated with caution, especially if the corresponding credible intervals are heavily skewed.

4.3.4 Homogenised atmospheric retrievals on existing spectra

After investigating the impact random noise has on the posterior distributions of our fiducial model, we now investigate the impact of systematic differences. We do this by applying our fiducial model to RU-23 and CA-24. All three of these spectra were derived from the same underlying raw data, and they are constrained to the same wavelength regime. Consequently, we assume that they contain the same information about the characteristics of the atmosphere of WASP-39 b. Based on this assumption, we do not retune the fiducial model, but rather apply it homogeneously on SP-TW, RU-23, and CA-24. In Figure 4.10, we see a comparison of the marginalised posterior distributions for the molecular MRs and cloud-top pressure for these retrievals, together with the inferred p–T profiles. The credible intervals for each retrieved parameter are summarised in Table 4.4 for all cases.

Globally, the atmospheric retrievals performed on SP-TW, RU-23, and CA-24 result in marginalised posterior distributions and derived parameter estimates that are in agreement within the CCI_{95}. In the behaviour of the posterior distributions, we find the same three types we have identified for the self-scattered spectra mimicking random noise, although the parameters they manifest for are slightly different. H_2O and CO_2 are, again, well-constrained and stable parameters. In all three cases, the posterior distributions are close to normal distributions, and the MRs are constrained to ranges within 0.5 to 1.0 dex. The abundances of CO and HCN are constrained to upper limits in every case, representing stable and unconstrained posteriors. The same is true for the cloud-top pressure, p_{cloud}, and also for the MR of CH_4. It is noteworthy that the retrievals performed on RU-23 produce an upper limit of CH_4 smaller by an order of magnitude than in the other two cases. As is visible in Figure 4.6, the transit-depth values of RU-23 are systematically smaller than for SP-TW and CA-24 in the region of the methane absorption feature around 3.4 µm. This could explain the reduced upper limit on the parameter estimate.

The biggest difference we find in the shape of the marginalised posterior distributions is for the MRs of SO_2 and C_2H_2. In the case of SO_2, the heavily tailed posterior distribution from the retrieval of SP-TW is contrasted by a tightly constrained abundance interval from RU-23, as well as from an almost flat posterior distribution resulting from the retrieval of CA-24. For C_2H_2, the skewed posterior of SP-TW is replaced by two unconstrained distributions from retrievals of RU-23 and CA-24. Their parameter estimates are still in agreement with each other, but they reflect the cases we have previously identified as skewed and unstable. Notably, H_2S now also falls into this category. The tight posterior constraint from SP-TW is replaced by heavily tailed parameter estimates from RU-23 and CA-24, both of which span more than 5 dex. At this parameter, we also find the only disagreement in the credible intervals. The parameter estimated on the MR of H_2S based on RU-23 misses the credible interval derived from SP-TW, although this difference is smaller than 0.2 dex.

As a final point of comparison, the retrieved p–T profiles agree between all three cases. Each of them indicates a close-to-isothermal behaviour in the centre of the atmospheric domain, while struggling to constrain the temperature value at the BOA temperature node. As the parameter estimate for p_{cloud} is consistent across the three retrievals, this region of the atmosphere is inaccessible, leading to unconstrained posterior distributions. A marginal difference is the credible interval for the temperature at the TOA pressure node. Retrievals performed on SP-TW

Table 4.4: Parameter credible intervals for atmospheric retrievals performed on SP-TW, RU-23, and CA-24. The reported values for individual parameters are given as 95% centred credible intervals, followed by the median. The pressure nodes associated with the four temperature values are located at $\log_{10}(p_i\,[\mathrm{bar}]) = \{1, -3, -7, -9\}$.

Parameter	Unit	Input spectrum					
		SP-TW		**RU-23**		**CA-24**	
		95% CCI	Median	95% CCI	Median	95% CCI	Median
R_p	$\mathrm{R_J}$	$[1.21, 1.29]$	1.26	$[1.19, 1.28]$	1.24	$[1.22, 1.30]$	1.26
$\log_{10}(CO_2)$	-	$[-2.23, -1.86]$	-2.05	$[-2.51, -1.99]$	-2.26	$[-2.60, -1.78]$	-2.24
$\log_{10}(CO)$	-	$[-11.74, -2.74]$	-7.24	$[-11.63, -2.46]$	-5.68	$[-11.63, -2.68]$	-7.11
$\log_{10}(H_2O)$	-	$[-1.04, -0.75]$	-0.87	$[-1.34, -0.95]$	-1.15	$[-1.04, -0.69]$	-0.87
$\log_{10}(CH_4)$	-	$[-11.78, -5.28]$	-8.40	$[-11.84, -6.66]$	-9.43	$[-11.69, -5.69]$	-8.81
$\log_{10}(H_2S)$	-	$[-2.97, -2.05]$	-2.49	$[-9.77, -3.15]$	-3.79	$[-9.25, -2.28]$	-2.95
$\log_{10}(SO_2)$	-	$[-10.46, -3.96]$	-4.66	$[-5.48, -4.57]$	-5.03	$[-11.55, -4.38]$	-6.85
$\log_{10}(HCN)$	-	$[-11.75, -4.48]$	-8.39	$[-11.74, -4.42]$	-7.93	$[-11.51, -3.55]$	-7.09
$\log_{10}(C_2H_2)$	-	$[-9.99, -4.30]$	-4.86	$[-11.76, -5.89]$	-8.71	$[-11.76, -5.36]$	-8.73
$\log_{10}(p_\mathrm{cloud})$	bar	$[-3.27, 0.88]$	-1.18	$[-3.65, 0.88]$	-1.28	$[-3.20, 0.86]$	-1.09
T_{p_0}	K	$[366.33, 2841.23]$	1407.14	$[571.13, 2901.82]$	1874.77	$[481.25, 2805.83]$	1615.97
T_{p_1}	K	$[711.31, 1296.84]$	912.26	$[607.68, 895.15]$	729.77	$[446.74, 1088.07]$	732.90
T_{p_2}	K	$[440.38, 909.23]$	700.23	$[321.21, 650.24]$	450.71	$[497.63, 1128.41]$	789.30
T_{p_3}	K	$[1554.85, 2965.68]$	2402.63	$[1427.31, 2962.16]$	2382.54	$[486.12, 2617.69]$	1462.25

and RU-23 indicate a thermal inversion in this region. In contrast, the result from CA-24 shows a much broader posterior, consistent with an isothermal profile and encompassing the p–T structure of the other two cases.

4.4 Limitations

To properly contextualise the results we have achieved here, we need to be aware of several limitations that apply to them. Firstly, the reported MRs of atmospheric constituents for WASP-39 b vary widely. In a comparison to previous analyses of the NIRSpec PRISM data we have used, the results we have achieved with our fiducial model are notably higher ($\sim$ 14% H_2O and $\sim$ 1% CO_2). When we account for retrievals of pre-JWST observational data, and of data from other instruments of JWST, we find discrepancies of up to five orders of magnitude for the retrieved model parameters (e.g., Tsiaras et al. 2018; Wakeford et al. 2018; Lueber et al. 2024). Exoplanet atmospheres are inherently more complex than atmospheric retrieval models can account for, making it difficult to draw conclusions about atmospheric characteristics from simplified forward models. Here, we circumvent this problem by applying a homogeneous forward model in our retrievals. This addressess disagreement stemming from the underlying data reduction process, without necessarily claiming to retrieve the 'correct' atmospheric properties of our target. On this point, it is also important to keep in mind that we do not address the impact of individual data-reduction steps on atmospheric retrieval results. The impact of using different pipelines on atmospheric characterisation results has been reported previously (e.g., Mugnai et al. 2024; Powell et al. 2024; Davenport et al. 2025). We focus on differences arising from performing retrievals on different spectra independently reduced using the same pipeline.

The last point concerns the application of the fiducial model. As we homogeneously use it on SP-TW, RU-23, and CA-24, we overlook the possibility of performing the model-tuning steps on each spectrum individually. The assumption that the model-tuning process is independent of the underlying transmission spectrum (even if it was derived from the same raw data) is not necessarily valid. The systematic differences we have identified between the three spectra we have analysed might propagate into the selection process of constituent molecules. In

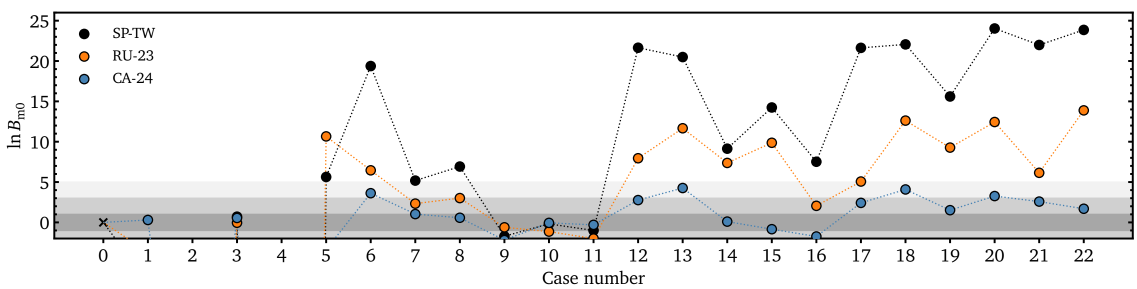

Figure 4.11: Dynamic model-tuning process for SP-TW, RU-23, and CA-24, showing the ID of different models (x-axis), against the Bayes factor compared to the initial model (y-axis). Results are colour-coded for SP-TW (black), RU-23 (orange), and CA-24 (blue).

Figure 4.11, we see how the model tuning process behaves when we treat it dynamically for all three spectra. The relative changes in the Bayes factor are fairly consistent between the three spectra. However, the absolute values of $\ln(B_{\mathrm{m},0})$ are significantly smaller for all retrievals performed on CA-24. In this process, we can identify an overlapping favoured inclusion of SO_2 (M_5) and H_2S (M_6). Only the model tuning based on SP-TW and RU-23 favourably includes C_2H_2 (M_8).

Chapter 5

Conclusions

"Thunder rolled ... it rolled a 6."

Terry Prachett

Guards! Guards!

State-of-the-art observations from the ground and from space provide us with data of unprecedented quality, providing us with insights into exoplanetary atmospheres not seen before. To infer the properties of these atmospheres, we make use of a statistical inference technique called atmospheric retrieval. In this method, we iteratively compare evaluations of radiative transfer forward models with high-precision spectroscopic data from modern observatories such as JWST and Ariel. The detailed information we can derive from these observations can inform us about the formation history of exoplanetary systems, trends in whole populations of exoplanets, and how our own planetary neighbourhood, the Solar System, fits into this wider picture.

In this thesis, we examined two main aspects that influence the results of atmospheric retrievals: the setup of the radiative transfer forward model and the inference-guiding dataset. Firstly, we investigated the relation between the complexity of the atmospheric forward model and the results of atmospheric retrievals. We created a grid of synthetic transmission spectra for a hot Jupiter, and investigated the impact of using pressure-temperature (p–T) prescriptions of varying complexity on atmospheric retrieval posteriors. This has allowed us to quantify the supported model complexity for state-of-the-art data and to identify biasing effects arising from underparameterised p–T prescriptions. We also explored the impact of perturbations in the underlying data on atmospheric retrieval results. We produced a transmission spectrum of the hot Jupiter WASP-39 b from a JWST NIRSpec PRISM observation and investigated how atmospheric retrieval results change when this dataset is perturbed.

5.1 Biases from the p–T profile prescription

In Chapter 3, we investigated the recoverable complexity of p–T profiles from simulated state-of-the-art transmission spectra. We simulated spectra for a hot Jupiter modelled after WASP-39 b, convolving high-resolution radiative transfer forward models with simulated and real observational noise levels. These encompass the instruments of JWST and simulations for the upcoming Ariel space mission, covering nine different scenarios in total. We created these synthetic spectra with two p–T structure assumptions and three cloud-top pressure levels. We tested the performance of atmospheric retrievals using heuristic multipoint p–T prescriptions with varying complexity, ranging from an isothermal profile to an eleven-point full self-retrieval. In total, we performed 324 atmospheric retrievals covering our grid of synthetic transmission spectra. We assessed the performance of these models using the Bayes factor and the accuracy of parameter recovery relative to the input values.

We conclude that an isothermal profile should not be used to retrieve state-of-the-art transmission spectra from JWST or Ariel. In approximately 75% of the tested cases, it leads to severe biases in the retrieved model parameters, particularly in the atmospheric molecular mixing ratios. This bias arises from a degeneracy between the temperature and MRs in the transit-observation geometry. The temperature controls the atmospheric scale height, while the molecular MRs influence the optical depth integral. Following from this, their degeneracy manifests as an anticorrelation in their parameter estimates. When the isothermal profile underestimates the local temperature of the probed atmospheric region, the atmospheric scale height shrinks, thereby shortening the path length in the optical-depth integral. Consequently, the molecular MRs are increased (and overestimated) to compensate for the otherwise missing optical depth. If the isothermal profile overestimates the local temperature, the inverse is true.

When assessing the supported model complexity for the p–T prescription, we find that two to four p–T points are adequate to retrieve the information content of JWST and Ariel data. Using more than four (and up to eleven) p–T points does not provide additional improvement to the model performance, while introducing additional free model parameters. The amplitude of the model preference is correlated with the signal-to-noise ratio (SNR) and wavelength coverage of the transmission spectrum, and the cloud deck of the synthetic input data. When the transmission spectrum is almost cloud-free, the Bayes factor is most clearly able to provide a preference for either the two- or four-point profile, based on the p–T structure used to create the synthetic spectrum. With a decreasing cloud-top pressure, many of the spectral features are masked or muted. Consequently, the model posterior odds distinguishing between preferred model setups become less pronounced. When comparing the individual instrument scenarios, the spectral resolution has the most significant impact on the model preference evaluation, followed by the wavelength coverage. We can see this when comparing the results for a simulated NIRCam F322W2 and MIRI LRS spectrum. The MIRI spectrum has approximately four times the wavelength coverage of the NIRCam spectrum, but its spectral resolution is approximately 30$\times$ lower. Both spectra have a comparable mean transit-depth uncertainty. Retrievals performed on the NIRCam spectra can produce model preference curves and parameter accuracy estimates comparable to those from cases containing more information (such as the NIRSpec G395H data). In contrast, the retrievals performed on the MIRI spectra result in more or less indistinguishable preference and performance between all tested p–T prescriptions, even though the wavelength coverage of MIRI is among the widest between all our tested instrument cases.

Based on the results of this investigation, we can recommend using a p–T prescription in the range of a two- to four-point profile for high-resolution spectroscopic data expected from JWST and Ariel. We recommend not using an isothermal prescription as it biases the retrieved atmospheric MRs. Future work could refine the selection of an optimal multipoint p–T prescription. Physically self-consistent models can help us understand the expected p–T structure of exoplanetary atmospheres, and informed priors based on this anticipated temperature structure could supplement the heuristic approach we have investigated here.

5.2 Stability of results under data perturbations

In Chapter 4 of this thesis, we investigated a second important part of any atmospheric retrieval framework: the role of the underlying dataset. We produced a transmission spectrum of the hot Jupiter WASP-39 b from a JWST NIRSpec PRISM observation, using the end-to-end data reduction pipeline `Eureka!`. We derived an atmospheric forward model by tuning the model parameters using this dataset. Our fiducial model contained a four-point p–T profile to characterise the vertical temperature of the exoplanet's atmosphere, molecular absorption contributions from $H_2O, CO_2, CO, CH_4, H_2S, SO_2, C_2H_2$, and HCN. We then investigated how the posterior distributions of the model parameters change under perturbations to this dataset.

The dataset perturbations are characterised in two ways. Firstly, we produce self-scattered instances of the transmission spectrum we have reduced. For each instance, we draw new data points from normal distributions based on the transit-depth mean and uncertainty values of the original spectral instance. These represent uncorrelated noise arising from random variations introduced during the data reduction process. Secondly, we consider three transmission spectra of WASp-39 b reduced from the same raw observation. The first of these is the spectrum we produced in this work, while the other two are published transmission spectra produced using the same data reduction pipeline (Rustamkulov et al. 2022, Carter & May et al. (2024)). These three spectra show systematic differences. Performing homogenised retrievals under all these scenarios provides an overview of the impact of random and systematic data perturbations.

We find that the behaviour of the parameter posterior distributions can be broadly categorised into three types, both for systematic and for random perturbations of the dataset. First, we identify several parameters characterised by well-constrained posterior distributions that are close in shape to a normal distribution. In our fiducial model, this encompasses the reference radius, R_{p}, as well as the MRs of H_2O, CO_2, and H_2S. The parameter estimates derived from these posterior distributions are stable under the data perturbations we investigated. Second, we find posterior distributions that translate into upper or lower limits for the estimated parameters, characterised by a shape close to a uniform distribution. In our fiducial model, the cloud-top pressure, p_{cloud}, and the MRs of CO and HCN fall into this category. These are also stable under perturbations of the dataset. Lastly, we identify skewed posterior distributions with broad tails. The MRs of SO_2, C_2H_2, and CH_4 fall into this category. These posteriors yield parameter estimates that are unstable to perturbations of the dataset, at times mimicking well-constrained or unconstrained posteriors. We also find that the p–T profile is consistent between the results obtained from different transmission spectra. It is well constrained within the probed atmospheric region, while the temperature values at the top and bottom of the atmosphere remain largely unconstrained.

Despite these variations in the behaviour of the parameter posterior distributions, it is important to point out that the resulting parameter estimates agree across all the cases of data perturbation we considered. However, this is based on a subjective choice of the credible interval size that defines the parameter estimate. In our case, we reported CCI_{95}, a credible interval that encompasses 95% of the posterior probability mass, centred on the distribution median. This analysis will change when the parameter credible intervals are constructed differently. The impact of this will be most prominent in the case of the unstable posterior distributions. As these are heavily skewed, deriving credible intervals with narrow widths, such as the commonly employed '1σ' centred credible interval (CCI), can provide a misleading sense of accuracy in the inferred parameter values. Directly computing credible interval sizes, rather than scaling the boundaries of smaller ones, can reveal the asymmetries of these distributions, thereby helping identify unstable parameters.

5.3 Outlook

In this work, we investigated several aspects of an atmospheric retrieval framework that affect the reported posterior distributions of parameters. Adjusting the methodologies we use to characterise the nature of exoplanetary atmospheres from observational data will be crucial to provide reliable characterisation results with the advent of data with unprecedented precision and wavelength coverage. However, there are more steps beyond the work we conducted here to further ensure this reliability.

The results presented in Chapter 3 have clearly shown us that state-of-the-art transmission spectra support and require the inference of p–T structures more complex than an isothermal profile. However, the way a multipoint profile is constructed offers avenues for improvement. In our case, we made an informed decision to slightly cluster the pressure nodes of each p–T profile within the pressure region we expect to be probed by the transmission spectrum *a priori*. For non-synthetic data, we do not know the true p–T structure of the exoplanet's atmosphere. A heuristic p–T profile with few pressure nodes will therefore always risk missing a significant feature of the vertical temperature structure, such as a temperature inversion. One way to address this would be to treat the pressure values at each p–T node as free parameters. However, this effectively doubles the number of model parameters associated with the p–T profile, significantly increasing the computational requirements on retrievals (e.g., Changeat et al. 2021). Another approach would be to dynamically select the location of the pressure nodes in the p–T profile. Instead of fixing them *a priori*, redistributing them *in situ* during the retrieval process would ensure that they stay clustered in the probed region of the atmosphere. This region can be determined by identifying the maximum of an optical-depth contribution function (e.g., Mollière et al. 2019). Waldmann et al. (2015) presented retrievals of atmospheric emission spectra based on this concept, performing multi-stage retrievals to dynamically adjust the complexity of the retrieved p–T profile. How applicable this is to the retrieval of transmission spectra needs to be investigated in detail, as they inherently contain less information about the thermal structure of an exoplanet's atmosphere than emission spectra.

In the context of the retrieved p–T structure, it is also worth considering how the vertical chemical profiles of the atmospheric forward model are characterised. We assumed the chemical profiles in our retrievals to be constant, which is a coarse, data-driven approach to identify important molecular absorbers in exoplanetary atmospheres. Accounting for a more complex vertical chemical structure is a difficult task, requiring a trade-off between physiochemical self-consistency and computational efficiency. The strong assumptions of equilibrium chemistry approaches have become less applicable with the identification of SO_2 as a product of photochemical processes in the atmospheres of several hot gas giants. Recently, Constantinou and Madhusudhan (2024) presented a modified equilibrium approach, where additional free parameters mimic photochemical disequilibrium processes. Al-Refaie et al. (2024) presented the direct implementation of a chemical kinetics model into atmospheric retrievals coupled to the utilisation of a reduced chemical network. Both of these approaches produce more physically plausible vertical profiles by imposing more stringent model assumptions. In contrast, extending the heuristic multipoint approach we have used for the atmospheric p–T profile to also cover the vertical chemical structure would provide more flexibility. However, this comes at the cost of introducing significantly more free parameters into the model than might be supported even by state-of-the-art data (Changeat et al. 2019). A step to supplement these approaches would be to incorporate additional information from complex, self-consistent physical models to challenge or validate results from simplified retrievals. Ensembles of modelling results for an exoplanets atmospheric structure could also be used to construct informed priors on retrieval parameters, without the necessity of incoporating computationally expensive and complex physical models directly into the atmospheric forward model.

In Chapter 4, we have seen that even using datasets reduced from the same raw data, and with the same pipeline, does not produce fully consistent parameter posterior distributions. However, it was outside the scope of this investigation to assess the impact of individual data-reduction steps and techniques on atmospheric retrieval results. JWST was launched recently, and pipelines to reduce its raw observational data are still under development, adding techniques to refine the final data products (e.g. NSClean, Rauscher 2015). Many of the individual steps in each stage also require user-specified input parameters, such as the rejection threshold in the detection of detector jumps or the type of background model. Clarifying exactly how these individual steps influence the extracted atmospheric spectrum and how this propagates into the parameter posterior distributions of atmospheric retrievals will further help determine the reliability of the reported atmospheric characteristics.

Lastly, it is worth noting that accounting for the uncertainty in data reduction steps or in the selection of the atmospheric forward model could provide a valuable alternative to searching for the 'correct' choice in these instances. In the context of the atmospheric forward model, rather than choosing a single model setup (as we have seen with the model tuning process in Section 4.3), it may be beneficial to account for the uncertainty in model selection by considering the results of multiple atmospheric retrievals of the same dataset using different model setups. Wakeford et al. (2017) incorporated this in form of Bayesian model averaging (BMA) into their analysis of the transmission spectrum of HAT-P-26 b. BMA provides a way to combine model results by assigning weights to each model based on its evidence. The analysis presented in Wakeford et al. (2017) used BMA to derive a weighted mean for each model parameter, but this method can also be used to determine full combined posteriors, allowing the calculation of proper credible intervals rather than weighted uncertainties. BMA may not be the best method to employ when the set of candidate models does not cover the true data-generating process (Yao et al. 2018), as is the case for atmospheric retrievals, where one-dimensional forward models are used to represent an inherently three-dimensional problem. As it is based on the model evidence, BMA is also strongly dependent on the parameter priors. Nixon et al. (2024) discussed several methods to account for model uncertainty in atmospheric retrievals. In addition to BMA, they provide alternative methods for model combination based on PSIS-LOO (Vehtari et al. 2017; Welbanks et al. 2023) and the stacking of predictive distributions (Yao et al. 2018). Future work investigating appropriate ways to combine atmospheric retrieval results from different forward models will enable us to more robustly quantify uncertainty in the estimated atmospheric parameters.

In the realm of data reduction, accounting for uncertainty is less trivial. Uncertainty in the light-curve model, such as the characterisation of stellar limb-darkening or the heterogeneities, can be treated with the same approach as in the atmospheric forward model. Carter & May et al. (2024) reviewed the differences in their derived JWST transmission spectra when using fixed or fitted LDCs. Their recommendation to use results derived with fixed LDCs can be substituted by accounting for the uncertainty in the light-curve model. Accounting for differences in the data arising from non-inference-based reduction stages is more challenging. Here, the most robust way forward might be to catalogue the impact of individual data reduction steps and to adopt a consistent approach to handle large datasets, thereby guaranteeing homogeneous population-level characterisation efforts. Understanding the differences in data products arising from small changes in individual data reduction steps and incorporating this uncertainty into the results derived from atmospheric retrievals will require substantial future work.

We are entering an era in which high-quality data from state-of-the-art observatories are providing far more insight into exoplanetary atmospheres than ever before. Only by making our inference methods robust can we guarantee that our understanding of exoplanetary systems, how they form, and how our own Solar System fits into this picture, is not skewed by biased characterisation techniques.

Bibliography

Abel, M., Frommhold, L., Li, X., and Hunt, K. L. C. (2011). "Collision-Induced Absorption by H2 Pairs: From Hundreds to Thousands of Kelvin". *The Journal of Physical Chemistry A* 115.25, pp. 6805–6812. DOI: 10.1021/jp109441f.

Abel, M., Frommhold, L., Li, X., and Hunt, K. L. C. (2012). "Infrared absorption by collisional H2–He complexes at temperatures up to 9000 K and frequencies from 0 to 20 000 cm-1". *The Journal of Chemical Physics* 136.4, p. 044319. DOI: 10.1063/1.3676405.

Adam, A. Y., Yachmenev, A., Yurchenko, S. N., and Jensen, P. (2019). "Variationally Computed IR Line List for the Methyl Radical CH3". *The Journal of Physical Chemistry A* 123.22, pp. 4755–4763. DOI: 10.1021/acs.jpca.9b02919.

Adams, E. R., Seager, S., and Elkins-Tanton, L. (2008). "Ocean planet or thick atmosphere: on the mass-radius relationship for solid exoplanets with massive atmospheres". *The Astrophysical Journal* 673.2, p. 1160. DOI: 10.1086/524925.

Agúndez, M., Martínez, J. I., Andres, P. L. de, Cernicharo, J., and Martín-Gago, J. A. (2020). "Chemical equilibrium in AGB atmospheres: successes, failures, and prospects for small molecules, clusters, and condensates". *Astronomy & Astrophysics* 637, A59. DOI: 10.1051/0004-6361/202037496.

Ahrer, E.-M., Stevenson, K. B., Mansfield, M., Moran, S. E., Brande, J., et al. (2023). "Early release science of the exoplanet WASP-39b with JWST NIRCam". *Nature* 614.7949, pp. 653–658. DOI: 10.1038/s41586-022-05590-4.

Akaike, H. (1974). "A new look at the statistical model identification". *IEEE Transactions on Automatic Control* 19.6, pp. 716–723. DOI: 10.1109/TAC.1974.1100705.

Al-Refaie, A. F., Changeat, Q., Venot, O., Waldmann, I. P., and Tinetti, G. (2022). "A comparison of chemical models of exoplanet atmospheres enabled by TauREx 3.1". *The Astrophysical Journal* 932.2, pp. 123–147. DOI: 10.3847/1538-4357/ac6dcd.

Al-Refaie, A. F., Changeat, Q., Waldmann, I. P., and Tinetti, G. (2021). "TauREx 3: a fast, dynamic, and extendable framework for retrievals". *The Astrophysical Journal* 917.1, pp. 37–60. DOI: 10.3847/1538-4357/ac0252.

Al-Refaie, A. F., Venot, O., Changeat, Q., and Edwards, B. (2024). “FRECKLL: full and reduced exoplanet chemical kinetics DistiLLed”. *The Astrophysical Journal* 967.2, pp. 132–147. DOI: 10.3847/1538-4357/ad3dee.

Alam, M. K., Gao, P., Adams Redai, J., Wallack, N. L., Wogan, N. F., et al. (2024). “JWST COMPASS: the first near- to mid-infrared transmission spectrum of the hot super-Earth L 168-9 b”. *Astronomical Journal* 169.1, p. 15. DOI: 10.3847/1538-3881/ad8eb5.

Albert, L., Lafrenière, D., Doyon, R., Artigau, É., Volk, K., et al. (2023). “The near infrared imager and slitless spectrograph for the James Webb Space Telescope. III. Single object slitless spectroscopy”. *Publications of the Astronomical Society of the Pacific* 135.1049, p. 075001. DOI: 10.1088/1538-3873/acd7a3.

Alderson, L., Wakeford, H. R., Alam, M. K., Batalha, N. E., Lothringer, J. D., et al. (2023). “Early release science of the exoplanet WASP-39b with JWST NIRSpec G395H”. *Nature* 614.7949, pp. 664–669. DOI: 10.1038/s41586-022-05591-3.

Alonso, R., Brown, T. M., Torres, G., Latham, D. W., Sozzetti, A., et al. (2004). “TrES-1: the transiting planet of a bright K0 V star”. *The Astrophysical Journal* 613.2, p. L153. DOI: 10.1086/425256.

Anderson, R. E. and Gordon, K. D. (2011). “Optimal cosmic-ray detection for nondestructive read ramps”. *Publications of the Astronomical Society of the Pacific* 123.908, p. 1237. DOI: 10.1086/662593.

Arcangeli, J., Désert, J.-M., Parmentier, V., Stevenson, K. B., Bean, J. L., et al. (2019). “Climate of an ultra hot Jupiter - spectroscopic phase curve of WASP-18b with HST/WFC3”. *Astronomy & Astrophysics* 625, A136. DOI: 10.1051/0004-6361/201834891.

Ashton, G., Bernstein, N., Buchner, J., Chen, X., Csányi, G., et al. (2022). “Nested sampling for physical scientists”. *Nature Reviews Methods Primers* 2.1, pp. 1–22. DOI: 10.1038/s43586-022-00121-x.

August, P. C., Buchhave, L. A., Diamond-Lowe, H., Mendonça, J. M., Gressier, A., et al. (2025). “Hot rocks survey I: a possible shallow eclipse for LHS 1478 b”. *Astronomy & Astrophysics* 695, A171. DOI: 10.1051/0004-6361/202452611.

August, P. C., Bean, J. L., Zhang, M., Lunine, J., Xue, Q., et al. (2023). “Confirmation of subsolar metallicity for WASP-77Ab from JWST thermal emission spectroscopy”. *Astrophysical Journal Letters* 953.2, p. L24. DOI: 10.3847/2041-8213/ace828.

Auvergne, M., Bodin, P., Boisnard, L., Buey, J.-T., Chaintreuil, S., et al. (2009). “The CoRoT satellite in flight: description and performance”. *Astronomy & Astrophysics* 506.1, pp. 411–424. DOI: 10.1051/0004-6361/200810860.

Azzam, A. A. A., Tennyson, J., Yurchenko, S. N., and Naumenko, O. V. (2016). "ExoMol molecular line lists – XVI. The rotation–vibration spectrum of hot H2S". *Monthly Notices of the Royal Astronomical Society* 460.4, pp. 4063–4074. DOI: 10.1093/mnras/stw1133.

Bakos, G., Noyes, R. W., Kovács, G., Stanek, K. Z., Sasselov, D. D., et al. (2004). "Wide-field millimagnitude photometry with the HAT: a tool for extrasolar planet detection". *Publications of the Astronomical Society of the Pacific* 116.817, pp. 266–278. DOI: 10.1086/382735.

Bakos, G. Á, Csubry, Z., Penev, K., Bayliss, D., Jordán, A., et al. (2013). "HATSouth: a global network of fully automated identical wide-field Telescopes1". *Publications of the Astronomical Society of the Pacific* 125.924, pp. 154–183. DOI: 10.1086/669529.

Ballard, S., Fabrycky, D., Fressin, F., Charbonneau, D., Desert, J.-M., et al. (2011). "The Kepler-19 system: a transiting 2.2 R$\oplus$ planet and a second planet detected via transit timing variations". *The Astrophysical Journal* 743.2, pp. 200–220. DOI: 10.1088/0004-637X/743/2/200.

Balmer, W. O., Kammerer, J., Pueyo, L., Perrin, M. D., Girard, J. H., et al. (2025). "JWST-TST high contrast: living on the wedge, or, NIRCam bar coronagraphy reveals CO2 in the HR 8799 and 51 Eri exoplanets' atmospheres". *The Astronomical Journal* 169.4, p. 209. DOI: 10.3847/1538-3881/adb1c6.

Baranne, A., Queloz, D., Mayor, M., Adrianzyk, G., Knispel, G., et al. (1996). "ELODIE: a spectrograph for accurate radial velocity measurements". *Astronomy & Astrophysics Supplement Series* 119, pp. 373–390. DOI: 10.1051/aas:1996251.

Barat, S., Désert, J.-M., Mukherjee, S., Goyal, J. M., Xue, Q., et al. (2025). "A metal-poor atmosphere with a hot interior for a young sub-Neptune progenitor: JWST/NIRSpec transmission spectrum of V1298 tau b". *Astronomical Journal* 170.3, p. 165. DOI: 10.3847/1538-3881/adec89.

Barber, R. J., Strange, J. K., Hill, C., Polyansky, O. L., Mellau, G. Ch., et al. (2014). "ExoMol line lists – III. An improved hot rotation-vibration line list for HCN and HNC". *Monthly Notices of the Royal Astronomical Society* 437.2, pp. 1828–1835. DOI: 10.1093/mnras/stt2011.

Barman, T. (2007). "Identification of Absorption Features in an Extrasolar Planet Atmosphere". *The Astrophysical Journal* 661.2, p. L191. DOI: 10.1086/518736.

Barman, T. S., Macintosh, B., Konopacky, Q. M., and Marois, C. (2011). "Clouds and chemistry in the atmosphere of the extrasolar planet HR 8799 b". *The Astrophysical Journal* 733.1, pp. 65–83. DOI: 10.1088/0004-637X/733/1/65.

Barstow, J. K., Aigrain, S., Irwin, P. G. J., Kendrew, S., and Fletcher, L. N. (2016). "Telling twins apart: exo-earths and venuses with transit spectroscopy". *Monthly Notices of the Royal Astronomical Society* 458.3, pp. 2657–2666. DOI: 10.1093/mnras/stw489.

Barstow, J. K., Changeat, Q., Garland, R., Line, M. R., Rocchetto, M., et al. (2020). "A comparison of exoplanet spectroscopic retrieval tools". *Monthly Notices of the Royal Astronomical Society* 493.4, pp. 4884–4909. DOI: 10.1093/mnras/staa548.

Barstow, J. K. and Heng, K. (2020). "Outstanding challenges of exoplanet atmospheric retrievals". *Space Science Reviews* 216.5, pp. 82–107. DOI: 10.1007/s11214-020-00666-x.

Batalha, N. E., Lewis, T., Fortney, J. J., Batalha, N. M., Kempton, E., et al. (2019). "The precision of mass measurements required for robust atmospheric characterization of transiting exoplanets". *The Astrophysical Journal Letters* 885.1, p. L25. DOI: 10.3847/2041-8213/ab4909.

Batalha, N. E., Mandell, A., Pontoppidan, K., Stevenson, K. B., Lewis, N. K., et al. (2017). "PandExo: A Community Tool for Transiting Exoplanet Science with JWST & HST". *Publications of the Astronomical Society of the Pacific* 129.976, p. 064501. DOI: 10.1088/1538-3873/aa65b0.

Batista, V. (2018). "Finding planets via gravitational microlensing". In: *Handbook of exoplanets*. Ed. by H. J. Deeg and J. A. Belmonte. Springer, pp. 659–687. ISBN: 978-3-319-55333-7. DOI: 10.1007/978-3-319-55333-7_120.

Batygin, K., Bodenheimer, P. H., and Laughlin, G. P. (2016). "In situ formation and dynamical evolution of hot Jupiter systems". *The Astrophysical Journal* 829.2, pp. 114–131. DOI: 10.3847/0004-637X/829/2/114.

Bean, J. L., Raymond, S. N., and Owen, J. E. (2021). "The nature and origins of sub-Neptune size planets". *Journal of Geophysical Research: Planets* 126.1, e2020JE006639. DOI: 10.1029/2020JE006639.

Beaulieu, J.-P., Bennett, D. P., Fouqué, P., Williams, A., Dominik, M., et al. (2006). "Discovery of a cool planet of 5.5 Earth-masses through gravitational microlensing". *Nature* 439.7075, pp. 437–440. DOI: 10.1038/nature04441.

Bell, T. J., Ahrer, E.-M., Brande, J., Carter, A. L., Feinstein, A. D., et al. (2022). "Eureka!: an end-to-end pipeline for JWST time-series observations". *Journal of Open Source Software* 7.79, p. 4503. DOI: 10.21105/joss.04503.

Bell, T. J., Crouzet, N., Cubillos, P. E., Kreidberg, L., Piette, A. A. A., et al. (2024). "Nightside clouds and disequilibrium chemistry on the hot Jupiter WASP-43b". *Nature Astronomy* 8.7, pp. 879–898. DOI: 10.1038/s41550-024-02230-x.

Bell, T. J., Welbanks, L., Schlawin, E., Line, M. R., Fortney, J. J., et al. (2023). "Methane throughout the atmosphere of the warm exoplanet WASP-80b". *Nature* 623.7988, pp. 709–712. DOI: 10.1038/s41586-023-06687-0.

Benneke, B. and Seager, S. (2013). "HOW TO DISTINGUISH BETWEEN CLOUDY MINI-NEPTUNES AND WATER/VOLATILE-DOMINATED SUPER-EARTHS". *The Astrophysical Journal* 778.2, p. 153. DOI: 10.1088/0004-637X/778/2/153.

Benneke, B., Wong, I., Piaulet, C., Knutson, H. A., Lothringer, J., et al. (2019). "Water vapor and clouds on the habitable-zone sub-Neptune exoplanet K2-18 b". *The Astrophysical Journal Letters* 887.1, p. L14. DOI: 10.3847/2041-8213/ab59dc.

Benz, W., Broeg, C., Fortier, A., Rando, N., Beck, T., et al. (2021). "The CHEOPS mission". *Experimental Astronomy* 51.1, pp. 109–151. DOI: 10.1007/s10686-020-09679-4.

Berta, Z. K., Charbonneau, D., Désert, J.-M., Miller-Ricci Kempton, E., McCullough, P. R., et al. (2012). "The flat transmission spectrum of the super-Earth GJ 1214 b from Wide Field Camera 3 on the Hubble Space Telescope". *Astrophysical Journal* 747.1, pp. 35–52. DOI: 10.1088/0004-637X/747/1/35.

Bézard, B., Charnay, B., and Blain, D. (2022). "Methane as a dominant absorber in the habitable-zone sub-Neptune K2-18 b". *Nature Astronomy* 6.5, pp. 537–540. DOI: 10.1038/s41550-022-01678-z.

Birkmann, S. M., Ferruit, P., Giardino, G., Nielsen, L. D., Muñoz, A. G., et al. (2022). "The near-infrared spectrograph (NIRSpec) on the James Webb Space Telescope - IV. Capabilities and predicted performance for exoplanet characterization". *Astronomy & Astrophysics* 661, A83. DOI: 10.1051/0004-6361/202142592.

Blain, D., Charnay, B., and Bézard, B. (2021). "1D atmospheric study of the temperate sub-neptune K2-18b". *Astronomy & Astrophysics* 646, A15. DOI: 10.1051/0004-6361/202039072.

Blecic, J., Dobbs-Dixon, I., and Greene, T. (2017). "The implications of 3D thermal structure on 1D atmospheric retrieval". *The Astrophysical Journal* 848.2, pp. 127–151. DOI: 10.3847/1538-4357/aa8171.

Böker, T., Beck, T. L., Birkmann, S. M., Giardino, G., Keyes, C., et al. (2023). "In-orbit Performance of the Near-infrared Spectrograph NIRSpec on the James Webb Space Telescope". *Publications of the Astronomical Society of the Pacific* 135.1045, p. 038001. DOI: 10.1088/1538-3873/acb846.

Bond, I. A., Udalski, A., Jaroszyński, M., Rattenbury, N. J., Paczyński, B., et al. (2004). "OGLE 2003-BLG-235/MOA 2003-BLG-53: a planetary microlensing event". *The Astrophysical Journal* 606.2, p. L155. DOI: 10.1086/420928.

Borucki, W. J., Koch, D., Basri, G., Batalha, N., Brown, T., et al. (2010). "Kepler planet-detection mission: introduction and first results". *Science* 327.5968, pp. 977–980. DOI: 10.1126/science.1185402.

Boss, A. P. (1997). "Giant planet formation by gravitational instability". *Science* 276.5320, pp. 1836–1839. DOI: 10.1126/science.276.5320.1836.

Brande, J., Crossfield, I. J. M., Kreidberg, L., Morley, C. V., Barman, T., et al. (2024). "Clouds and clarity: revisiting atmospheric feature trends in Neptune-size exoplanets". *Astrophysical Journal Letters* 961.1, p. L23. DOI: 10.3847/2041-8213/ad1b5c.

Buchner, J., Georgakakis, A., Nandra, K., Hsu, L., Rangel, C., et al. (2014). "X-ray spectral modelling of the AGN obscuring region in the CDFS: Bayesian model selection and catalogue". *Astronomy & Astrophysics* 564, A125. DOI: 10.1051/0004-6361/201322971.

Burnham, K. P. and Anderson, D. R. (2004). "Multimodel Inference: Understanding AIC and BIC in Model Selection". *Sociological Methods & Research* 33.2, pp. 261–304. DOI: 10.1177/0049124104268644.

Bushouse, H., Eisenhamer, J., Dencheva, N., Davies, J., Greenfield, P., et al. (2024). *JWST calibration pipeline*. Zenodo. DOI: 10.5281/zenodo.10569856.

Carter, A. L., May, E. M., Espinoza, N., Welbanks, L., Ahrer, E., et al. (2024). "A benchmark JWST near-infrared spectrum for the exoplanet WASP-39 b". *Nature Astronomy* 8.8, pp. 1008–1019. DOI: 10.1038/s41550-024-02292-x.

Challener, R. C., Weiner Mansfield, M., Cubillos, P. E., Piette, A. A. A., Coulombe, L.-P., et al. (2025). "Horizontal and vertical exoplanet thermal structure from a JWST spectroscopic eclipse map". *Nature Astronomy* 9.12, pp. 1821–1832. DOI: 10.1038/s41550-025-02666-9.

Changeat, Q., Al-Refaie, A. F., Edwards, B., Waldmann, I. P., and Tinetti, G. (2021). "An Exploration of Model Degeneracies with a Unified Phase Curve Retrieval Analysis: The Light and Dark Sides of WASP-43 b". *The Astrophysical Journal* 913.1, p. 73. DOI: 10.3847/1538-4357/abf2bb.

Changeat, Q., Bardet, D., Chubb, K., Dyrek, A., Edwards, B., et al. (2025). "Cloud and haze parameterization in atmospheric retrievals: insights from Titan's Cassini data and JWST observations of hot Jupiters". *Astronomy & Astrophysics* 699, A219. DOI: 10.1051/0004-6361/202453186.

Changeat, Q., Edwards, B., Al-Refaie, A. F., Tsiaras, A., Skinner, J. W., et al. (2022). "Five key exoplanet questions answered via the aynalysis of 25 hot-Jupiter atmospheres in eclipse". *The Astrophysical Journal Supplement Series* 260.1, p. 3. DOI: 10.3847/1538-4365/ac5cc2.

Changeat, Q., Edwards, B., Waldmann, I. P., and Tinetti, G. (2019). "Toward a more complex description of chemical profiles in exoplanet retrievals: a two-layer parameterization". *The Astrophysical Journal* 886.1, pp. 39–56. DOI: 10.3847/1538-4357/ab4a14.

Changeat, Q., Ito, Y., Al-Refaie, A. F., Yip, K. H., and Lueftinger, T. (2024). "Toward atmospheric retrievals of panchromatic light curves: ExPLOR-ing generalized inversion techniques for transiting exoplanets with JWST and ariel". *The Astronomical Journal* 167.5, pp. 195–211. DOI: 10.3847/1538-3881/ad3032.

Changeat, Q., Keyte, L., Waldmann, I. P., and Tinetti, G. (2020). "Impact of Planetary Mass Uncertainties on Exoplanet Atmospheric Retrievals". *The Astrophysical Journal* 896.2, p. 107. DOI: 10.3847/1538-4357/ab8f8b.

Charbonneau, D., Allen, L. E., Megeath, S. T., Torres, G., Alonso, R., et al. (2005). "Detection of thermal emission from an extrasolar planet". *The Astrophysical Journal* 626.1, p. 523. DOI: 10.1086/429991.

Charbonneau, D., Brown, T. M., Latham, D. W., and Mayor, M. (2000). "Detection of planetary transits across a Sun-like star". *The Astrophysical Journal* 529.1, p. L45. DOI: 10.1086/312457.

Charbonneau, D., Brown, T. M., Noyes, R. W., and Gilliland, R. L. (2002). "Detection of an extrasolar planet atmosphere*". *Astrophysical Journal* 568.1, pp. 377–385. DOI: 10.1086/338770.

Charpinet, S., Fontaine, G., Brassard, P., Green, E. M., Van Grootel, V., et al. (2011). "A compact system of small planets around a former red-giant star". *Nature* 480.7378, pp. 496–499. DOI: 10.1038/nature10631.

Chauvin, G., Lagrange, A.-M., Dumas, C., Zuckerman, B., Mouillet, D., et al. (2005). "Giant planet companion to 2MASSW J1207334-393254". *Astronomy & Astrophysics* 438.2, pp. L25–L28. DOI: 10.1051/0004-6361:200500116.

Christiansen, J. L., McElroy, D. L., Harbut, M., Ciardi, D. R., Crane, M., et al. (2025). "The NASA exoplanet archive and exoplanet follow-up observing program: data, tools, and usage". *The Planetary Science Journal* 6.8, pp. 186–200. DOI: 10.3847/PSJ/ade3c2.

Chubb, K. L. and Min, M. (2022). "Exoplanet atmosphere retrievals in 3D using phase curve data with ARCiS: application to WASP-43b". *Astronomy & Astrophysics* 665, A2. DOI: 10.1051/0004-6361/202142800.

Chubb, K. L., Rocchetto, M., Yurchenko, S. N., Min, M., Waldmann, I., et al. (2021). "The ExoMolOP database: Cross sections and k-tables for molecules of interest in high-temperature exoplanet atmospheres". *Astronomy & Astrophysics* 646, A21. DOI: 10.1051/0004-6361/202038350.

Chubb, K. L., Tennyson, J., and Yurchenko, S. N. (2020). "ExoMol molecular line lists – XXXVII. Spectra of acetylene". *Monthly Notices of the Royal Astronomical Society* 493.2, pp. 1531–1545. DOI: 10.1093/mnras/staa229.

Claret, A. (2000). "A new non-linear limb-darkening law for LTE stellar atmosphere models. Calculations for -5.0 <= log[M/H] <= +1, 2000 K <= Teff <= 50000 K at several surface gravities". *Astronomy & Astrophysics* 363, pp. 1081–1190. URL: https://ui.adsabs.harvard.edu/abs/2000A&A...363.1081C.

Coles, P. A., Yurchenko, S. N., and Tennyson, J. (2019). "ExoMol molecular line lists – XXXV. A rotation-vibration line list for hot ammonia". *Monthly Notices of the Royal Astronomical Society* 490.4, pp. 4638–4647. DOI: 10.1093/mnras/stz2778.

Collette, A. (2013). *Python and HDF5*. O'Reilly Media, Inc. ISBN: 978-1-4919-4498-1. URL: https://www.oreilly.com/library/view/python-and-hdf5/9781491944981/.

Constantinou, S. and Madhusudhan, N. (2024). "VIRA: an exoplanet atmospheric retrieval framework for JWST transmission spectroscopy". *Monthly Notices of the Royal Astronomical Society* 530.3, pp. 3252–3277. DOI: 10.1093/mnras/stae633.

Constantinou, S., Madhusudhan, N., and Gandhi, S. (2023). "Early Insights for Atmospheric Retrievals of Exoplanets Using JWST Transit Spectroscopy". *The Astrophysical Journal* 943, p. L10. DOI: 10.3847/2041-8213/acaead.

Coulombe, L.-P., Benneke, B., Challener, R., Piette, A. A. A., Wiser, L. S., et al. (2023). "A broadband thermal emission spectrum of the ultra-hot Jupiter WASP-18b". *Nature* 620.7973, pp. 292–298. DOI: 10.1038/s41586-023-06230-1.

Cowan, N. B., Agol, E., and Charbonneau, D. (2007). "Hot nights on extrasolar planets: mid-infrared phase variations of hot Jupiters". *Monthly Notices of the Royal Astronomical Society* 379.2, pp. 641–646. DOI: 10.1111/j.1365-2966.2007.11897.x.

Craiu, R. V. and Rosenthal, J. S. (2014). "Bayesian computation via Markov Chain Monte Carlo". *Annual Review of Statistics and Its Application* 1.Volume 1, 2014, pp. 179–201. DOI: 10.1146/annurev-statistics-022513-115540.

Crossfield, I. J. M., Ahrer, E.-M., Brande, J., Kreidberg, L., Lothringer, J., et al. (2025). "Mapping the SO2 Shoreline in Gas Giant Exoplanets". *The Astrophysical Journal* 994.2, p. 184. DOI: 10.3847/1538-4357/ae17cb.

Crossfield, I. J. M., Malik, M., Hill, M. L., Kane, S. R., Foley, B., et al. (2022). "GJ 1252b: a hot terrestrial super-Earth with no atmosphere". *Astrophysical Journal Letters* 937.1, p. L17. DOI: 10.3847/2041-8213/ac886b.

Cutri, R. M., Skrutskie, M. F., van Dyk, S., Beichman, C. A., Carpenter, J. M., et al. (2003). *VizieR online data catalog: 2MASS all-sky catalog of point sources*. URL: https://ui.adsabs.harvard.edu/abs/2003yCat.2246....0C.

Davenport, B., Kempton, E. M.-R., Nixon, M. C., Ih, J., Deming, D., et al. (2025). "TOI-421 b: A Hot Sub-Neptune with a Haze-free, Low Mean Molecular Weight Atmosphere". *The Astrophysical Journal Letters* 984.2, p. L44. DOI: 10.3847/2041-8213/adcd76.

Davey, J. J., Yip, K. H., Changeat, Q., and Waldmann, I. P. (2025). "Investigating the influence of asymmetric errors on retrievals of exoplanet transmission spectra". *RAS Techniques and Instruments* 4, rzaf052. DOI: 10.1093/rasti/rzaf052.

de Wit, J. and Seager, S. (2013). “Constraining exoplanet mass from transmission spectroscopy”. *Science* 342.6165, pp. 1473–1477. doi: 10.1126/science.1245450.

de Wit, J., Wakeford, H. R., Lewis, N. K., Delrez, L., Gillon, M., et al. (2018). “Atmospheric reconnaissance of the habitable-zone Earth-sized planets orbiting TRAPPIST-1”. *Nature Astronomy* 2.3, pp. 214–219. doi: 10.1038/s41550-017-0374-z.

Deming, D., Seager, S., Winn, J., Miller-Ricci, E., Clampin, M., et al. (2009). “Discovery and characterization of transiting super earths using an all-sky transit survey and follow-up by the James Webb Space Telescope”. *Publications of the Astronomical Society of the Pacific* 121.883, pp. 952–968. doi: 10.1086/605913.

Deming, D., Seager, S., Richardson, L. J., and Harrington, J. (2005). “Infrared radiation from an extrasolar planet”. *Nature* 434.7034, pp. 740–743. doi: 10.1038/nature03507.

Di Maio, C., Changeat, Q., Benatti, S., and Micela, G. (2023). “Analysis of the planetary mass uncertainties on the accuracy of atmospherical retrieval”. *Astronomy & Astrophysics* 669, A150. doi: 10.1051/0004-6361/202244881.

Domingue, D. L., Koehn, P. L., Killen, R. M., Sprague, A. L., Sarantos, M., et al. (2007). “Mercury’s Atmosphere: A Surface-Bounded Exosphere”. *Space Science Reviews* 131.1, pp. 161–186. doi: 10.1007/s11214-007-9260-9.

Doyon, R., Hutchings, J. B., Beaulieu, M., Albert, L., Lafrenière, D., et al. (2012). “The JWST fine guidance sensor (FGS) and near-infrared imager and slitless spectrograph (NIRISS)”. In: *Space Telescopes and Instrumentation 2012: Optical, Infrared, and Millimeter Wave*. Vol. 8442, 84422R. doi: 10.1117/12.926578.

Doyon, R., Willott, C. J., Hutchings, J. B., Sivaramakrishnan, A., Albert, L., et al. (2023). “The near infrared imager and slitless spectrograph for the James Webb Space Telescope. I. Instrument overview and In-flight performance”. *Publications of the Astronomical Society of the Pacific* 135.1051, p. 098001. doi: 10.1088/1538-3873/acd41b.

Ducrot, E., Lagage, P.-O., Min, M., Gillon, M., Bell, T. J., et al. (2024). “Combined analysis of the 12.8 and 15 μm JWST/MIRI eclipse observations of TRAPPIST-1 b”. *Nature Astronomy*. doi: 10.1038/s41550-024-02428-z.

Duffell, P. C., Haiman, Z., MacFadyen, A. I., D’Orazio, D. J., and Farris, B. D. (2014). “The migration of gap-opening planets is not locked to viscous disk evolution”. *The Astrophysical Journal Letters* 792.1, p. L10. doi: 10.1088/2041-8205/792/1/L10.

Dyrek, A., Min, M., Decin, L., Bouwman, J., Crouzet, N., et al. (2024). “SO2, silicate clouds, but no CH4 detected in a warm Neptune”. *Nature* 625.7993, pp. 51–54. doi: 10.1038/s41586-023-06849-0.

Edwards, B. and Tinetti, G. (2022). “The Ariel target list: the impact of TESS and the potential for characterizing multiple planets within a system”. *Astronomical Journal* 164.1, p. 15. DOI: 10.3847/1538-3881/ac6bf9.

Edwards, B., Tsiaras, A., Changeat, Q., and Yip, K. H. (2024). “On the difficulties of obtaining absolute transit depths with HST WFC3: KELT-11 b, an example”. *RAS Techniques and Instruments* 3.1, pp. 415–436. DOI: 10.1093/rasti/rzae023.

Einstein, A. (1905). “Zur Elektrodynamik bewegter Körper”. *Annalen der Physik* 322.10, pp. 891–921. DOI: 10.1002/andp.19053221004.

Espinoza, N., Allen, N. H., Glidden, A., Lewis, N. K., Seager, S., et al. (2025). “JWST-TST DREAMS: NIRSpec/PRISM transmission spectroscopy of the habitable zone planet TRAPPIST-1 e”. *Astrophysical Journal* 990, p. L52. DOI: 10.3847/2041-8213/adf42e.

Espinoza, N. and Jones, K. (2021). “Constraining Mornings and Evenings on Distant Worlds: A new Semianalytical Approach and Prospects with Transmission Spectroscopy”. *The Astronomical Journal* 162.4, p. 165. DOI: 10.3847/1538-3881/ac134d.

Espinoza, N., Steinrueck, M. E., Kirk, J., MacDonald, R. J., Savel, A. B., et al. (2024). “Inhomogeneous terminators on the exoplanet WASP-39 b”. *Nature* 632.8027, pp. 1017–1020. DOI: 10.1038/s41586-024-07768-4.

Eureka! documentation (2023). *Eureka! pipeline documentation*. URL: https://eurekadocs.readthedocs.io/en/v0.10/index.html (visited on 01/09/2026).

Faedi, F., Barros, S. C. C., Anderson, D. R., Brown, D. J. A., Cameron, A. C., et al. (2011). “WASP-39b: a highly inflated Saturn-mass planet orbiting a late G-type star”. *Astronomy & Astrophysics* 531, A40. DOI: 10.1051/0004-6361/201116671.

Fan, S., Shemansky, D. E., Li, C., Gao, P., Wan, L., et al. (2019). “Retrieval of chemical abundances in titan’s upper atmosphere from cassini UVIS observations with pointing motion”. *Earth and Space Science* 6.7, pp. 1057–1066. DOI: 10.1029/2018EA000477.

Feinstein, A. D., Radica, M., Welbanks, L., Murray, C. A., Ohno, K., et al. (2023). “Early release science of the exoplanet WASP-39b with JWST NIRISS”. *Nature* 614.7949, pp. 670–675. DOI: 10.1038/s41586-022-05674-1.

Feng, Y. K., Robinson, T. D., Fortney, J. J., Lupu, R. E., Marley, M. S., et al. (2018). “Characterizing Earth Analogs in Reflected Light: Atmospheric Retrieval Studies for Future Space Telescopes”. *The Astronomical Journal* 155.5, p. 200. DOI: 10.3847/1538-3881/aab95c.

Feroz, F., Hobson, M. P., and Bridges, M. (2009). “MultiNest: an efficient and robust Bayesian inference tool for cosmology and particle physics”. *Monthly Notices of the Royal Astronomical Society* 398.4, pp. 1601–1614. DOI: 10.1111/j.1365-2966.2009.14548.x.

Fixsen, D. J., Offenberg, J. D., Hanisch, R. J., Mather, J. C., Nieto-Santisteban, M. A., et al. (2000). "Cosmic-ray rejection and readout efficiency for large-area arrays". *Publications of the Astronomical Society of the Pacific* 112.776, p. 1350. DOI: 10.1086/316626.

Fletcher, L. N., Gustafsson, M., and Orton, G. S. (2018). "Hydrogen Dimers in Giant-planet Infrared Spectra". *The Astrophysical Journal Supplement Series* 235.1, p. 24. DOI: 10.3847/1538-4365/aaa07a.

Foreman-Mackey, D. (2016). "corner.py: Scatterplot matrices in Python". *Journal of Open Source Software* 1.2, p. 24. DOI: 10.21105/joss.00024.

Foreman-Mackey, D., Hogg, D. W., Lang, D., and Goodman, J. (2013). "emcee: The MCMC Hammer". *Publications of the Astronomical Society of the Pacific* 125.925, p. 306. DOI: 10.1086/670067.

Fortney, J. J., Dawson, R. I., and Komacek, T. D. (2021). "Hot Jupiters: origins, structure, atmospheres". *Journal of Geophysical Research: Planets* 126.3, e2020JE006629. DOI: 10.1029/2020JE006629.

Fortney, J. J., Mordasini, C., Nettelmann, N., Kempton, E. M.-R., Greene, T. P., et al. (2013). "A framework for characterizing the atmospheres of low-mass low-density transiting planets". *The Astrophysical Journal* 775.1, pp. 80–93. DOI: 10.1088/0004-637X/775/1/80.

Frommhold, L. (1994). *Collision-induced absorption in gases*. Cambridge Monographs on Atomic, Molecular and Chemical Physics. Cambridge: Cambridge University Press. ISBN: 978-0-521-01967-5. DOI: 10.1017/CBO9780511524523.

Fulchignoni, M., Ferri, F., Angrilli, F., Ball, A. J., Bar-Nun, A., et al. (2005). "In situ measurements of the physical characteristics of Titan's environment". *Nature* 438.7069, pp. 785–791. DOI: 10.1038/nature04314.

Fulton, B. J., Petigura, E. A., Howard, A. W., Isaacson, H., Marcy, G. W., et al. (2017). "The California-Kepler survey. III. A gap in the radius distribution of small planets*". *The Astronomical Journal* 154.3, pp. 109–128. DOI: 10.3847/1538-3881/aa80eb.

Gao, P., Wakeford, H. R., Moran, S. E., and Parmentier, V. (2021). "Aerosols in exoplanet atmospheres". *Journal of Geophysical Research: Planets* 126.4, e2020JE006655. DOI: 10.1029/2020JE006655.

Gardner, J. P., Mather, J. C., Abbott, R., Abell, J. S., Abernathy, M., et al. (2023). "The James Webb Space Telescope Mission". *Publications of the Astronomical Society of the Pacific* 135.1048, p. 068001. DOI: 10.1088/1538-3873/acd1b5.

Gelman, A., Hwang, J., and Vehtari, A. (2014). "Understanding predictive information criteria for Bayesian models". *Statistics and Computing* 24.6, pp. 997–1016. DOI: 10.1007/s11222-013-9416-2.

Gillon, M., Jehin, E., Lederer, S. M., Delrez, L., de Wit, J., et al. (2016). "Temperate Earth-sized planets transiting a nearby ultracool dwarf star". *Nature* 533.7602, pp. 221–224. DOI: 10.1038/nature17448.

Ginzburg, S., Schlichting, H. E., and Sari, R. (2018). "Core-powered mass-loss and the radius distribution of small exoplanets". *Monthly Notices of the Royal Astronomical Society* 476.1, pp. 759–765. DOI: 10.1093/mnras/sty290.

Glidic, K., Schlawin, E., Wiser, L., Zhou, Y., Deming, D., et al. (2022). "Atmospheric characterization of hot Jupiter CoRoT-1 b using the Wide Field Camera 3 on the Hubble Space Telescope". *The Astronomical Journal* 164.1, p. 19. DOI: 10.3847/1538-3881/ac6cdb.

Goody, R. M. and Yung, Y. L. (1995). *Atmospheric radiation: theoretical basis*. 2nd ed. New York: Oxford University Press. ISBN: 978-0-19-510291-8. URL: https://research.ebsco.com/linkprocessor/plink?id=ed2ac048-f64e-3a02-853d-8fed7ab9935b.

Gordon, I. E., Rothman, L. S., Hargreaves, R. J., Hashemi, R., Karlovets, E. V., et al. (2022). "The HITRAN2020 molecular spectroscopic database". *Journal of Quantitative Spectroscopy and Radiative Transfer* 277, p. 107949. DOI: 10.1016/j.jqsrt.2021.107949.

Gould, A., Udalski, A., An, D., Bennett, D. P., Zhou, A.-Y., et al. (2006). "Microlens OGLE-2005-BLG-169 implies that cool Neptune-like planets are common". *The Astrophysical Journal* 644.1, p. L37. DOI: 10.1086/505421.

Grant, D. and Wakeford, H. R. (2024). "ExoTiC-LD: thirty seconds to stellar limb-darkening coefficients". *Journal of Open Source Software* 9.100, p. 6816. DOI: 10.21105/joss.06816.

Greenbaum, A. Z., Pueyo, L., Ruffio, J.-B., Wang, J. J., Rosa, R. J. D., et al. (2018). "GPI spectra of HR 8799 c, d, and e from 1.5 to 2.4 μm with KLIP forward modeling". *The Astronomical Journal* 155.6, p. 226. DOI: 10.3847/1538-3881/aabcb8.

Greene, T. P., Bell, T. J., Ducrot, E., Dyrek, A., Lagage, P.-O., et al. (2023). "Thermal emission from the Earth-sized exoplanet TRAPPIST-1 b using JWST". *Nature* 618.7963, pp. 39–42. DOI: 10.1038/s41586-023-05951-7.

Greene, T. P., Kelly, D. M., Stansberry, J., Leisenring, J., Egami, E., et al. (2017). "$\lambda = 2.4 – 5$ μm spectroscopy with the JWST NIRCam instrument". *Journal of Astronomical Telescopes, Instruments, and Systems* 3.3, p. 035001. DOI: 10.1117/1.JATIS.3.3.035001.

Gressier, A., Mori, M., Changeat, Q., Edwards, B., Beaulieu, J. P., et al. (2022). "Near-infrared transmission spectrum of TRAPPIST-1 h using hubble WFC3 G141 observations". *Astronomy & Astrophysics* 658, A133. DOI: 10.1051/0004-6361/202142140.

Gressier, A., Batalha, N. E., Wogan, N., Alderson, L., Doud, D., et al. (2025). "JWST-TST DREAMS: sulfur dioxide in the atmosphere of the Neptune-mass planet HAT-P-26 b from NIRSpec G395H transmission spectroscopy". *Astronomical Journal* 170.5, p. 292. DOI: 10.3847/1538-3881/ae0929.

Guerlet, S., Ignatiev, N., Forget, F., Fouchet, T., Vlasov, P., et al. (2022). "Thermal structure and aerosols in Mars' atmosphere from TIRVIM/ACS onboard the ExoMars trace gas orbiter: validation of the retrieval algorithm". *Journal of Geophysical Research: Planets* 127.2, e2021JE007062. DOI: 10.1029/2021JE007062.

Guillot, T. (2010). "On the radiative equilibrium of irradiated planetary atmospheres". *Astronomy & Astrophysics* 520, A27. DOI: 10.1051/0004-6361/200913396.

Hara, N. C. and Ford, E. B. (2023). "Statistical methods for exoplanet detection with radial velocities". *Annual Review of Statistics and Its Application* 10.Volume 10, 2023, pp. 623–649. DOI: 10.1146/annurev-statistics-033021-012225.

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., et al. (2020). "Array programming with NumPy". *Nature* 585.7825, pp. 357–362. DOI: 10.1038/s41586-020-2649-2.

Hastings, W. K. (1970). "Monte Carlo sampling methods using Markov Chains and their applications". *Biometrika* 57.1, pp. 97–109. DOI: 10.2307/2334940.

Haynes, K., Mandell, A. M., Madhusudhan, N., Deming, D., and Knutson, H. (2015). "Spectrscopic evidence for a temperature inversion in the dayside atmosphere of hot Jupiter WASP-33b". *The Astrophysical Journal* 806.2, pp. 146–158. DOI: 10.1088/0004-637X/806/2/146.

Helbert, J., Haus, R., Arnold, G., D'Amore, M., Maturilli, A., et al. (2023). "The second venus flyby of BepiColombo mission reveals stable atmosphere over decades". *Nature Communications* 14.1, p. 8225. DOI: 10.1038/s41467-023-43888-7.

Henry, G. W., Marcy, G. W., Butler, R. P., and Vogt, S. S. (2000). "A transiting "51 peg-like" planet*". *The Astrophysical Journal* 529.1, p. L41. DOI: 10.1086/312458.

Hollis, M. D. J., Tessenyi, M., and Tinetti, G. (2013). "tau: A 1D radiative transfer code for transmission spectroscopy of extrasolar planet atmospheres". *Computer Physics Communications* 184.10, pp. 2351–2361. DOI: 10.1016/j.cpc.2013.05.011.

Holmberg, M. and Madhusudhan, N. (2023). "Exoplanet spectroscopy with JWST NIRISS: diagnostics and case studies". *Monthly Notices of the Royal Astronomical Society* 524.1, pp. 377–402. DOI: 10.1093/mnras/stad1580.

Holmberg, M. and Madhusudhan, N. (2024). "Possible hycean conditions in the sub-Neptune TOI-270 d". *Astronomy & Astrophysics* 683, p. L2. DOI: 10.1051/0004-6361/202348238.

Horne, K. (1986). "An optimal extraction algorithm for CCD spectroscopy." *Publications of the Astronomical Society of the Pacific* 98, pp. 609–617. DOI: 10.1086/131801.

Horner, S. D. and Rieke, M. J. (2004). “The near-infrared camera (NIRCam) for the James Webb Space Telescope (JWST)”. In: *Optical, Infrared, and Millimeter Space Telescopes*. Vol. 5487, pp. 628–634. DOI: 10.1117/12.552281.

Hörst, S. M. (2017). “Titan’s atmosphere and climate”. *Journal of Geophysical Research: Planets* 122.3, pp. 432–482. DOI: 10.1002/2016JE005240.

Howell, S. B., Sobeck, C., Haas, M., Still, M., Barclay, T., et al. (2014). “The K2 mission: characterization and early results”. *Publications of the Astronomical Society of the Pacific* 126.938, pp. 398–409. DOI: 10.1086/676406.

HST & Beyond Committee, Dressler, A., Brown, R. A., Davidsen, A. F., Ellis, R. S., et al. (1996). *Exploration an the search for origins: a vision for ultraviolet-optical-infrared space astronomy*. Tech. rep. NAS 1.15:112149. URL: https://ntrs.nasa.gov/citations/19970007007.

Hu, R., Bello-Arufe, A., Zhang, M., Paragas, K., Zilinskas, M., et al. (2024). “A secondary atmosphere on the rocky exoplanet 55 Cancri e”. *Nature* 630.8017, pp. 609–612. DOI: 10.1038/s41586-024-07432-x.

Hunter, J. D. (2007). “Matplotlib: A 2D Graphics Environment”. *Computing in Science & Engineering* 9.3, pp. 90–95. DOI: 10.1109/MCSE.2007.55.

Ih, J., Kempton, E. M.-R., Whittaker, E. A., and Lessard, M. (2023). “Constraining the Thickness of TRAPPIST-1 b’s Atmosphere from Its JWST Secondary Eclipse Observation at 15 μm”. *The Astrophysical Journal Letters* 952.1, p. L4. DOI: 10.3847/2041-8213/ace03b.

Inglis, J., Batalha, N. E., Lewis, N. K., Kataria, T., Knutson, H. A., et al. (2024). “Quartz clouds in the dayside atmosphere of the quintessential hot Jupiter HD 189733 b”. *The Astrophysical Journal Letters* 973.2, p. L41. DOI: 10.3847/2041-8213/ad725e.

Irwin, P. G. J. (2003). “Vertical structure of temperature, composition, and clouds”. In: *Giant planets of our Solar System: an introduction*. Ed. by P. G. J. Irwin. Berlin, Heidelberg: Springer, pp. 67–132. ISBN: 978-3-540-37713-9. DOI: 10.1007/3-540-37713-1_4.

Irwin, P. G. J., Parmentier, V., Taylor, J., Barstow, J., Aigrain, S., et al. (2020). “2.5D retrieval of atmospheric properties from exoplanet phase curves: application to WASP-43b observations”. *Monthly Notices of the Royal Astronomical Society* 493.1, pp. 106–125. DOI: 10.1093/mnras/staa238.

Jakobsen, P., Ferruit, P., Oliveira, C. A. de, Arribas, S., Bagnasco, G., et al. (2022). “The Near-Infrared Spectrograph (NIRSpec) on the James Webb Space Telescope - I. Overview of the instrument and its capabilities”. *Astronomy & Astrophysics* 661, A80. DOI: 10.1051/0004-6361/202142663.

JDox (2026). *JWST user documentation*. URL: https://jwst-docs.stsci.edu/#gsc.tab=0 (visited on 01/12/2026).

Jurgenson, C., Fischer, D., McCracken, T., Sawyer, D., Szymkowiak, A., et al. (2016). "EXPRES: a next generation RV spectrograph in the search for Earth-like worlds". In: *Ground-based and Airborne Instrumentation for Astronomy VI*. Vol. 9908. eprint: arXiv:1606.04413, 99086T. DOI: 10.1117/12.2233002.

JWST pipeline documentation (2025). *Package documentation — jwst 1.17.0*. URL: https://jwst-pipeline.readthedocs.io/en/1.17.0/ (visited on 01/09/2026).

JWST Transiting Exoplanet Community Early Release Science Team, Ahrer, E.-M., Alderson, L., Batalha, N. M., Batalha, N. E., et al. (2023). "Identification of carbon dioxide in an exoplanet atmosphere". *Nature* 614.7949, pp. 649–652. DOI: 10.1038/s41586-022-05269-w.

Kass, R. E. and Raftery, A. E. (1995). "Bayes Factors". *Journal of the American Statistical Association* 90.430, pp. 773–795. DOI: 10.1080/01621459.1995.10476572.

Keady, J. J. and Kilcrease, D. P. (2002). "Radiation". In: *Allen's astrophysical quantities*. Ed. by A. N. Cox. 4th ed. New York: AIP Press, pp. 95–121. ISBN: 978-0-387-98746-0.

Keers, R. E., Shapiro, A. I., Kostogryz, N. M., Glidden, A., Niraula, P., et al. (2024). "Reliable Transmission Spectrum Extraction with a Three-parameter Limb-darkening Law". *The Astrophysical Journal Letters* 977.1, p. L7. DOI: 10.3847/2041-8213/ad8b51.

Kempton, E. M.-R., Zhang, M., Bean, J. L., Steinrueck, M. E., Piette, A. A. A., et al. (2023). "A reflective, metal-rich atmosphere for GJ 1214b from its JWST phase curve". *Nature* 620.7972, pp. 67–71. DOI: 10.1038/s41586-023-06159-5.

Kendrew, S., Scheithauer, S., Bouchet, P., Amiaux, J., Azzollini, R., et al. (2015). "The mid-infrared instrument for the James Webb Space Telescope, IV: the low-resolution spectrometer". *Publications of the Astronomical Society of the Pacific* 127.953, pp. 623–633. DOI: 10.1086/682255.

Kipping, D. M. (2013). "Efficient, uninformative sampling of limb darkening coefficients for two-parameter laws". *Monthly Notices of the Royal Astronomical Society* 435.3, pp. 2152–2160. DOI: 10.1093/mnras/stt1435.

Kipping, D. M. and Tinetti, G. (2010). "Nightside pollution of exoplanet transit depths". *Monthly Notices of the Royal Astronomical Society* 407.4, pp. 2589–2598. DOI: 10.1111/j.1365-2966.2010.17094.x.

Kirk, J., Wheatley, P. J., Louden, T., Doyle, A. P., Skillen, I., et al. (2017). "Rayleigh scattering in the transmission spectrum of HAT-P-18b". *Monthly Notices of the Royal Astronomical Society* 468.4, pp. 3907–3916. DOI: 10.1093/mnras/stx752.

Kirk, J., Rackham, B. V., MacDonald, R. J., López-Morales, M., Espinoza, N., et al. (2021). "ACCESS and LRG-BEASTS: a precise new optical transmission spectrum of the ultrahot Jupiter WASP-103 b". *Astronomical Journal* 162.1, p. 34. DOI: 10.3847/1538-3881/abfcd2.

Kite, E. S., Fegley Jr., B., Schaefer, L., and Ford, E. B. (2020). “Atmosphere origins for exoplanet sub-Neptunes”. *The Astrophysical Journal* 891.2, pp. 111–127. doi: 10.3847/1538-4357/ab6ffb.

Kleinböhl, A., Kass, D. M., Schreier, M., Piqueux, S., Suzuki, S., et al. (2024). “Far infrared radiative properties of Mars atmospheric aerosols and their application to Mars climate sounder retrievals of aerosol profiles, aerosol columns and surface temperatures”. *Icarus* 419, p. 116000. doi: 10.1016/j.icarus.2024.116000.

Knutson, H. A., Lewis, N., Fortney, J. J., Burrows, A., Showman, A. P., et al. (2012). “3.6 and 4.5 μm phase curves and evidence for non-equilibrium chemistry in the atmosphere of extrasolar planet HD 189733b”. *The Astrophysical Journal* 754.1, p. 22. doi: 10.1088/0004-637X/754/1/22.

Konopacky, Q. M., Barman, T. S., Macintosh, B. A., and Marois, C. (2013). “Detection of Carbon Monoxide and Water Absorption Lines in an Exoplanet Atmosphere”. *Science* 339.6126, pp. 1398–1401. doi: 10.1126/science.1232003.

Koskinen, T. T., Sandel, B. R., Yelle, R. V., Strobel, D. F., Müller-Wodarg, I. C. F., et al. (2015). “Saturn’s variable thermosphere from Cassini/UVIS occultations”. *Icarus* 260, pp. 174–189. doi: 10.1016/j.icarus.2015.07.008.

Kreidberg, L. (2015). “batman: BAsic Transit Model cAlculatioN in Python”. *Publications of the Astronomical Society of the Pacific* 127.957, p. 1161. doi: 10.1086/683602.

Kreidberg, L., Bean, J. L., Désert, J.-M., Benneke, B., Deming, D., et al. (2014a). “Clouds in the atmosphere of the super-Earth exoplanet GJ 1214b”. *Nature* 505.7481, pp. 69–72. doi: 10.1038/nature12888.

Kreidberg, L., Bean, J. L., Désert, J.-M., Line, M. R., Fortney, J. J., et al. (2014b). “A precise water-abundance measurement for the hot Jupiter WASP-43b”. *The Astrophysical Journal Letters* 793.2, p. L27. doi: 10.1088/2041-8205/793/2/L27.

Kreidberg, L., Line, M. R., Parmentier, V., Stevenson, K. B., Louden, T., et al. (2018). “Global climate and atmospheric composition of the ultra-hot Jupiter WASP-103b from HST and Spitzer phase curve observations”. *Astronomical Journal* 156.1, p. 17. doi: 10.3847/1538-3881/aac3df.

Léger, A., Selsis, F., Sotin, C., Guillot, T., Despois, D., et al. (2004). “A new family of planets? “Ocean-planets””. *Icarus* 169.2, pp. 499–504. doi: 10.1016/j.icarus.2004.01.001.

Lellouch, E., McGrath, M. A., and Jessup, K. L. (2007). “Io’s atmosphere”. In: *Io after galileo: a new view of Jupiter’s volcanic moon*. Ed. by R. M. C. Lopes and J. R. Spencer. Berlin, Heidelberg: Springer, pp. 231–264. isbn: 978-3-540-48841-5. doi: 10.1007/978-3-540-48841-5_10.

Li, G., Gordon, I. E., Rothman, L. S., Tan, Y., Hu, S.-M., et al. (2015). "Rovibrational line lists for nine isotopologues of the CO molecule in the x1σ+ ground electronic state". *The Astrophysical Journal Supplement Series* 216.1, pp. 15–33. DOI: 10.1088/0067-0049/216/1/15.

Li, L., Chen, Y., Fan, L., Sun, D., He, H., et al. (2024). "A high-precision retrieval method for methane vertical profiles based on dual-band spectral data from the GOSAT satellite". *Atmospheric Environment* 317, p. 120183. DOI: 10.1016/j.atmosenv.2023.120183.

Lichtenberg, T. and Miguel, Y. (2025). "Super-Earths and Earth-like exoplanets". In: *Treatise on geochemistry*. Ed. by A. Anbar and D. Weis. Elsevier, pp. 51–112. ISBN: 978-0-323-99763-8. DOI: 10.1016/B978-0-323-99762-1.00122-4.

Lim, O., Benneke, B., Doyon, R., MacDonald, R. J., Piaulet, C., et al. (2023). "Atmospheric reconnaissance of TRAPPIST-1 b with JWST/NIRISS: evidence for strong stellar contamination in the transmission spectra". *The Astrophysical Journal* 955, p. L22. DOI: 10.3847/2041-8213/acf7c4.

Limaye, S. S., Lebonnois, S., Mahieux, A., Pätzold, M., Bougher, S., et al. (2017). "The thermal structure of the Venus atmosphere: Intercomparison of Venus Express and ground based observations of vertical temperature and density profiles". *Icarus* 294, pp. 124–155. DOI: 10.1016/j.icarus.2017.04.020.

Lin, D. N. C. and Papaloizou, J. (1986). "On the Tidal interaction between protoplanets and the protoplanetary disk. III. Orbital migration of protoplanets". *The Astrophysical Journal* 309, pp. 846–858. DOI: 10.1086/164653.

Line, M. R., Wolf, A. S., Zhang, X., Knutson, H., Kammer, J. A., et al. (2013). "A systematic retrieval analysis of secondary eclipse spectra. I. A comparison of atmospheric retrieval techniques". *The Astrophysical Journal* 775.2, pp. 137–159. DOI: 10.1088/0004-637X/775/2/137.

Line, M. R., Zhang, X., Vasisht, G., Natraj, V., Chen, P., et al. (2012). "Information content of exoplanetary transit spectra: an initial look". *The Astrophysical Journal* 749, pp. 93–103. DOI: 10.1088/0004-637X/749/1/93.

Liou, K. N. (2002). *An introduction to atmospheric radiation*. 2nd ed. International Geophysics Series v. 84. Amsterdam: Academic Press. ISBN: 978-0-12-451451-5. URL: https://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=195011&site=ehost-live.

Liu, R., Wang, L.-C., Rustamkulov, Z., and Sing, D. K. (2025). "Unveiling the Atmosphere of the super-Jupiter HAT-P-14 b with JWST NIRISS and NIRSpec". *The Astronomical Journal* 169.6, p. 335. DOI: 10.3847/1538-3881/adcba7.

Lueber, A., Novais, A., Fisher, C., and Heng, K. (2024). "Information content of JWST spectra of WASP-39b". *Astronomy & Astrophysics* 687, A110. DOI: 10.1051/0004-6361/202348802.

Luhman, K. L., Burgasser, A. J., and Bochanski, J. J. (2011). "Discovery of a candidate for the coolest known brown dwarf*". *The Astrophysical Journal Letters* 730.1, p. L9. DOI: 10.1088/2041-8205/730/1/L9.

Lustig-Yaeger, J., Sotzen, K. S., Stevenson, K. B., Tsai, S.-M., Challener, R. C., et al. (2025). "JWST-TST DREAMS: the nightside emission and chemistry of WASP-17b". *The Astrophysical Journal Letters* 994.1, p. L4. DOI: 10.3847/2041-8213/ae17ae.

Ma, S., Ito, Y., Al-Refaie, A. F., Changeat, Q., Edwards, B., et al. (2023). "YunMa: enabling spectral retrievals of exoplanetary clouds". *The Astrophysical Journal* 957.2, pp. 104–123. DOI: 10.3847/1538-4357/acf8ca.

MacDonald, R. J. and Batalha, N. E. (2023). "A catalog of exoplanet atmospheric retrieval codes". *Research Notes of the American Astronomical Society* 7.3, pp. 54–59. DOI: 10.3847/2515-5172/acc46a.

MacDonald, R. J., Goyal, J. M., and Lewis, N. K. (2020). "Why Is it So Cold in Here? Explaining the Cold Temperatures Retrieved from Transmission Spectra of Exoplanet Atmospheres". *The Astrophysical Journal* 893, p. L43. DOI: 10.3847/2041-8213/ab8238.

Madhusudhan, N. and Seager, S. (2009). "A temperature and abundance retrieval method for exoplanet atmospheres". *The Astrophysical Journal* 707.1, pp. 24–39. DOI: 10.1088/0004-637X/707/1/24.

Madhusudhan, N., Constantinou, S., Holmberg, M., Sarkar, S., Piette, A. A. A., et al. (2025). "New constraints on DMS and DMDS in the Atmosphere of K2-18 b from JWST MIRI". *The Astrophysical Journal Letters* 983.2, p. L40. DOI: 10.3847/2041-8213/adc1c8.

Madhusudhan, N., Piette, A. A. A., and Constantinou, S. (2021). "Habitability and biosignatures of hycean worlds". *The Astrophysical Journal* 918.1, pp. 1–26. DOI: 10.3847/1538-4357/abfd9c.

Madhusudhan, N., Sarkar, S., Constantinou, S., Holmberg, M., Piette, A. A. A., et al. (2023). "Carbon-bearing molecules in a possible hycean atmosphere". *The Astrophysical Journal Letters* 956, p. L13. DOI: 10.3847/2041-8213/acf577.

Magic, Z., Chiavassa, A., Collet, R., and Asplund, M. (2015). "The Stagger-grid: A grid of 3D stellar atmosphere models - IV. Limb darkening coefficients". *Astronomy & Astrophysics* 573, A90. DOI: 10.1051/0004-6361/201423804.

Majeau, C., Agol, E., and Cowan, N. B. (2012). "A two-dimensional infrared map of the extrasolar planet HD 189733b". *The Astrophysical Journal Letters* 747.2, p. L20. DOI: 10.1088/2041-8205/747/2/L20.

Mancini, L., Esposito, M., Covino, E., Southworth, J., Biazzo, K., et al. (2018). "The GAPS programme with HARPS-N at TNG - XVI. Measurement of the Rossiter–McLaughlin effect of transiting planetary systems HAT-P-3, HAT-P-12, HAT-P-22, WASP-39, and WASP-60". *Astronomy & Astrophysics* 613, A41. doi: 10.1051/0004-6361/201732234.

Mandel, K. and Agol, E. (2002). "Analytic light curves for planetary transit searches". *Astrophysical Journal* 580.2, p. L171. doi: 10.1086/345520.

Mant, B. P., Yachmenev, A., Tennyson, J., and Yurchenko, S. N. (2018). "ExoMol molecular line lists – XXVII. Spectra of C2H4". *Monthly Notices of the Royal Astronomical Society* 478.3, pp. 3220–3232. doi: 10.1093/mnras/sty1239.

Marcq, E., Encrenaz, T., Bézard, B., and Birlan, M. (2006). "Remote sensing of Venus' lower atmosphere from ground-based IR spectroscopy: latitudinal and vertical distribution of minor species". *Planetary and Space Science* 54.13, pp. 1360–1370. doi: 10.1016/j.pss.2006.04.024.

Marcq, E., Mills, F. P., Parkinson, C. D., and Vandaele, A. C. (2017). "Composition and chemistry of the neutral atmosphere of Venus". *Space Science Reviews* 214.1, p. 10. doi: 10.1007/s11214-017-0438-5.

Marley, M. S., Ackerman, A. S., Cuzzi, J. N., and Kitzmann, D. (2013). "Clouds and hazes in exoplanet atmopsheres". In: *Comparative climatology of terrestrial planets*. Ed. by S. J. Mackwell, A. A. Simon-Miller, J. W. Harder, and M. A. Bullock. University of Arizona Press, pp. 367–391. isbn: 978-0-8165-9975-2.

Marois, C., Macintosh, B., Barman, T., Zuckerman, B., Song, I., et al. (2008). "Direct imaging of multiple planets orbiting the star HR 8799". *Science* 322.5906, pp. 1348–1352. doi: 10.1126/science.1166585.

Marois, C., Zuckerman, B., Konopacky, Q. M., Macintosh, B., and Barman, T. (2010). "Images of a fourth planet orbiting HR 8799". *Nature* 468.7327, pp. 1080–1083. doi: 10.1038/nature09684.

Matthews, E. C., Carter, A. L., Pathak, P., Morley, C. V., Phillips, M. W., et al. (2024). "A temperate super-Jupiter imaged with JWST in the mid-infrared". *Nature* 633.8031, pp. 789–792. doi: 10.1038/s41586-024-07837-8.

Mayor, M. and Queloz, D. (1995). "A Jupiter-mass companion to a solar-type star". *Nature* 378.6555, pp. 355–359. doi: 10.1038/378355a0.

McKinney, W. (2010). "Data Structures for Statistical Computing in Python". *Proceedings of the 9th Python in Science Conference*, pp. 56–61. doi: 10.25080/Majora-92bf1922-00a.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). "Equation of state calculations by fast computing machines". *Journal of Chemical Physics* 21.6, pp. 1087–1092. doi: 10.1063/1.1699114.

Mikal-Evans, T., Sing, D. K., Barstow, J. K., Kataria, T., Goyal, J., et al. (2022). “Diurnal variations in the stratosphere of the ultrahot giant exoplanet WASP-121b”. *Nature Astronomy* 6.4, pp. 471–479. doi: 10.1038/s41550-021-01592-w.

Miles, B. E., Biller, B. A., Patapis, P., Worthen, K., Rickman, E., et al. (2023). “The JWST early-release science program for direct observations of exoplanetary systems II: a 1 to 20 μm spectrum of the planetary-mass companion VHS 1256–1257 b”. *The Astrophysical Journal Letters* 946.1, p. L6. doi: 10.3847/2041-8213/acb04a.

Misener, W., Schlichting, H. E., and Young, E. D. (2023). “Atmospheres as windows into sub-Neptune interiors: coupled chemistry and structure of hydrogen–silane–water envelopes”. *Monthly Notices of the Royal Astronomical Society* 524.1, pp. 981–992. doi: 10.1093/mnras/stad1910.

Mollière, P., Wardenier, J. P., Boekel, R. van, Henning, T., Molaverdikhani, K., et al. (2019). “petitRADTRANS - A Python radiative transfer package for exoplanet characterization and retrieval”. *Astronomy & Astrophysics* 627, A67. doi: 10.1051/0004-6361/201935470.

Mollière, P., Molyarova, T., Bitsch, B., Henning, T., Schneider, A., et al. (2022). “Interpreting the Atmospheric Composition of Exoplanets: Sensitivity to Planet Formation Assumptions”. *The Astrophysical Journal* 934, p. 74. doi: 10.3847/1538-4357/ac6a56.

Montet, B. T., Morton, T. D., Foreman-Mackey, D., Johnson, J. A., Hogg, D. W., et al. (2015). “Stellar and planetary properties of K2 campaign 1 candidates and validation of 17 planets, including a planet receiving earth-like insolation”. *The Astrophysical Journal* 809, pp. 25–40. doi: 10.1088/0004-637X/809/1/25.

Mordasini, C., van Boekel, R., Mollière, P., Henning, Th., and Benneke, B. (2016). “The imprint of exoplanet formation history on observable present-day spectra of hot jupiters”. *The Astrophysical Journal* 832, pp. 41–73. doi: 10.3847/0004-637X/832/1/41.

Morello, G., Tsiaras, A., Howarth, I. D., and Homeier, D. (2017). “High-precision Stellar Limb-darkening in Exoplanetary Transits”. *The Astronomical Journal* 154.3, p. 111. doi: 10.3847/1538-3881/aa8405.

Morello, G., Dyrek, A., and Changeat, Q. (2022). “Is binning always sinning? The impact of time-averaging for exoplanet phase curves”. *Monthly Notices of the Royal Astronomical Society* 517.2, pp. 2151–2164. doi: 10.1093/mnras/stac2828.

Mugnai, L. V., Pascale, E., Edwards, B., Papageorgiou, A., and Sarkar, S. (2020). “ArielRad: the Ariel radiometric model”. *Experimental Astronomy* 50.2, pp. 303–328. doi: 10.1007/s10686-020-09676-7.

Mugnai, L. V., Swain, M. R., Estrela, R., and Roudier, G. M. (2024). “Comparing transit spectroscopy pipelines at the catalogue level: evidence for systematic differences”. *Monthly Notices of the Royal Astronomical Society* 531.1, pp. 35–51. doi: 10.1093/mnras/stae1073.

Mukherjee, S., Schlawin, E., Bell, T. J., Fortney, J. J., Beatty, T. G., et al. (2025). "A JWST panchromatic thermal emission spectrum of the warm neptune archetype GJ 436b". *The Astrophysical Journal Letters* 982.2, p. L39. doi: 10.3847/2041-8213/adba46.

Murphy, M. M., Beatty, T. G., Schlawin, E., Bell, T. J., Line, M. R., et al. (2024). "Evidence for morning-to-evening limb asymmetry on the cool low-density exoplanet WASP-107 b". *Nature Astronomy* 8.12, pp. 1562–1574. doi: 10.1038/s41550-024-02367-9.

Murphy, M. M., Beatty, T. G., Welbanks, L., and Fu, G. (2025). "HST Transmission Spectra of the Hot Neptune HD 219666 b: Detection of Water and the Challenge of Constraining Both Water and Methane". *The Astronomical Journal* 169.6, p. 286. doi: 10.3847/1538-3881/adc684.

Nakajima, T., Oppenheimer, B. R., Kulkarni, S. R., Golimowski, D. A., Matthews, K., et al. (1995). "Discovery of a cool brown dwarf". *Nature* 378.6556, pp. 463–465. doi: 10.1038/378463a0.

Nakayama, A., Ikoma, M., and Terada, N. (2022). "Survival of terrestrial N2–O2 atmospheres in violent XUV environments through efficient atomic line radiative cooling". *The Astrophysical Journal* 937.2, pp. 72–90. doi: 10.3847/1538-4357/ac86ca.

NASA EPA (2025). *NASA exoplanet archive*. url: https://exoplanetarchive.ipac.caltech.edu/ (visited on 11/11/2025).

Niraula, P., de Wit, J., Gordon, I. E., Hargreaves, R. J., Sousa-Silva, C., et al. (2022). "The impending opacity challenge in exoplanet atmospheric characterization". *Nature Astronomy* 6.11, pp. 1287–1295. doi: 10.1038/s41550-022-01773-1.

Niraula, P., Wit, J. de, Gordon, I. E., Hargreaves, R. J., and Sousa-Silva, C. (2023). "Origin and Extent of the Opacity Challenge for Atmospheric Retrievals of WASP-39 b". *The Astrophysical Journal Letters* 950.2, p. L17. doi: 10.3847/2041-8213/acd6f8.

Nixon, M. C., Welbanks, L., McGill, P., and Kempton, E. M.-R. (2024). "Methods for Incorporating Model Uncertainty into Exoplanet Atmospheric Analysis". *The Astrophysical Journal* 966.2, p. 156. doi: 10.3847/1538-4357/ad354e.

Öberg, K. I., Murray-Clay, R., and Bergin, E. A. (2011). "The effects of snowlines on C/O in planetary atmospheres". *The Astrophysical Journal Letters* 743.1, p. L16. doi: 10.1088/2041-8205/743/1/L16.

Ohno, K., Schlawin, E., Bell, T. J., Murphy, M. M., Beatty, T. G., et al. (2025). "A possible metal-dominated atmosphere below the thick aerosols of GJ 1214 b suggested by its JWST panchromatic transmission spectrum". *Astrophysical Journal Letters* 979.1, p. L7. doi: 10.3847/2041-8213/ada02c.

Ormel, C. W. and Min, M. (2019). "ARCiS framework for exoplanet atmospheres - The cloud transport model". *Astronomy & Astrophysics* 622, A121. DOI: 10.1051/0004-6361/201833678.

Owen, J. E. and Lai, D. (2018). "Photoevaporation and high-eccentricity migration created the sub-jovian desert". *Monthly Notices of the Royal Astronomical Society* 479.4, pp. 5012–5021. DOI: 10.1093/mnras/sty1760.

Owen, J. E. and Wu, Y. (2017). "The evaporation valley in the Kepler planets". *The Astrophysical Journal* 847.1, pp. 29–43. DOI: 10.3847/1538-4357/aa890a.

Patel, J. A., Brandeker, A., Kitzmann, D., Roche, D. J. M. P. dit de la, Bello-Arufe, A., et al. (2024). "JWST reveals the rapid and strong day-side variability of 55 cancri e". *Astronomy & Astrophysics* 690, A159. DOI: 10.1051/0004-6361/202450748.

Pepe, F., Cristiani, S., Rebolo, R., Santos, N. C., Dekker, H., et al. (2021). "ESPRESSO at VLT - on-sky performance and first results". *Astronomy & Astrophysics* 645, A96. DOI: 10.1051/0004-6361/202038306.

Piaulet-Ghorayeb, C., Benneke, B., Turbet, M., Moore, K., Roy, P.-A., et al. (2025). "Strict limits on potential secondary atmospheres on the temperate rocky exo-Earth TRAPPIST-1 d". *The Astrophysical Journal* 989.2, p. 181. DOI: 10.3847/1538-4357/adf207.

Pinte, C., van der Plas, G., Ménard, F., Price, D. J., Christiaens, V., et al. (2019). "Kinematic detection of a planet carving a gap in a protoplanetary disk". *Nature Astronomy* 3.12, pp. 1109–1114. DOI: 10.1038/s41550-019-0852-6.

Pluriel, W., Leconte, J., Parmentier, V., Zingales, T., Falco, A., et al. (2022). "Toward a multidimensional analysis of transmission spectroscopy - II. Day-night-induced biases in retrievals from hot to ultrahot Jupiters". *Astronomy & Astrophysics* 658, A42. DOI: 10.1051/0004-6361/202141943.

Pollacco, D. L., Skillen, I., Cameron, A. C., Christian, D. J., Hellier, C., et al. (2006). "The WASP project and the SuperWASP cameras". *Publications of the Astronomical Society of the Pacific* 118.848, p. 1407. DOI: 10.1086/508556.

Pollack, J. B., Hubickyj, O., Bodenheimer, P., Lissauer, J. J., Podolak, M., et al. (1996). "Formation of the giant planets by concurrent accretion of solids and gas". *Icarus* 124.1, pp. 62–85. DOI: 10.1006/icar.1996.0190.

Polyansky, O. L., Kyuberis, A. A., Zobov, N. F., Tennyson, J., Yurchenko, S. N., et al. (2018). "ExoMol molecular line lists XXX: a complete high-accuracy line list for water". *Monthly Notices of the Royal Astronomical Society* 480.2, pp. 2597–2608. DOI: 10.1093/mnras/sty1877.

Powell, D., Feinstein, A. D., Lee, E. K. H., Zhang, M., Tsai, S.-M., et al. (2024). "Sulfur dioxide in the mid-infrared transmission spectrum of WASP-39b". *Nature* 626.8001, pp. 979–983. DOI: 10.1038/s41586-024-07040-9.

Pueyo, L. (2018). "Direct imaging as a detection technique for exoplanets". In: *Handbook of exoplanets*. Ed. by H. J. Deeg and J. A. Belmonte. Springer, Cham, pp. 705–765. ISBN: 978-3-319-55333-7. DOI: 10.1007/978-3-319-55333-7_10.

Qian, S.-B., Liao, W.-P., Zhu, L.-Y., and Dai, Z.-B. (2009). "Detection of a giant extrasolar planet orbiting the eclipsing polar dp leo". *The Astrophysical Journal Letters* 708.1, p. L66. DOI: 10.1088/2041-8205/708/1/L66.

Rackham, B. V., Apai, D., and Giampapa, M. S. (2018). "The transit light source effect: false spectral features and incorrect densities for M-dwarf transiting planets". *The Astrophysical Journal* 853.2, pp. 122–140. DOI: 10.3847/1538-4357/aaa08c.

Radica, M., Piaulet-Ghorayeb, C., Taylor, J., Coulombe, L.-P., Benneke, B., et al. (2025). "Promise and peril: stellar contamination and strict limits on the atmosphere composition of TRAPPIST-1 c from JWST NIRISS transmission spectra". *The Astrophysical Journal* 979, p. L5. DOI: 10.3847/2041-8213/ada381.

Rafikov, R. R. (2005). "Can giant planets form by direct gravitational instability?" *The Astrophysical Journal* 621.1, p. L69. DOI: 10.1086/428899.

Rauer, H., Aerts, C., Cabrera, J., Deleuil, M., Erikson, A., et al. (2025). "The PLATO mission". *Experimental Astronomy* 59.3, pp. 26–137. DOI: 10.1007/s10686-025-09985-9.

Rauscher, B. J. (2015). "Teledyne H1RG, H2RG, and H4RG noise generator". *Publications of the Astronomical Society of the Pacific* 127.957, p. 1144. DOI: 10.1086/684082.

Rauscher, B. J., Fixsen, D. J., and Mosby, G. (2022). "Simple improved reference subtraction for H4RG, H2RG, and H1RG near-infrared array detectors". *Journal of Astronomical Telescopes, Instruments, and Systems* 8, p. 028002. DOI: 10.1117/1.JATIS.8.2.028002.

Rauscher, E., Menou, K., Seager, S., Deming, D., Cho, J. Y.-K., et al. (2007). "Toward eclipse mapping of hot jupiters". *The Astrophysical Journal* 664.2, p. 1199. DOI: 10.1086/519213.

Rauscher, E., Suri, V., and Cowan, N. B. (2018). "A More Informative Map: Inverting Thermal Orbital Phase and Eclipse Light Curves of Exoplanets". *Astronomical Journal* 156.5, p. 235. DOI: 10.3847/1538-3881/aae57f.

Redfield, S., Batalha, N., Benneke, B., Biller, B., Espinoza, N., et al. (2024). *Report of the working group on strategic exoplanet initiatives with HST and JWST*. DOI: 10.48550/arXiv.2404.02932.

Ricker, G. R., Winn, J. N., Vanderspek, R., Latham, D. W., Bakos, G. Á., et al. (2015). "Transiting exoplanet survey satellite (TESS)". *Journal of Astronomical Telescopes, Instruments, and Systems* 1, p. 014003. DOI: 10.1117/1.JATIS.1.1.014003.

Rieke, G. H., Wright, G. S., Böker, T., Bouwman, J., Colina, L., et al. (2015). "The mid-infrared instrument for the James Webb Space Telescope, I: Introduction". *Publications of the Astronomical Society of the Pacific* 127.953, p. 584. DOI: 10.1086/682252.

Rieke, M. J., Kelly, D. M., Misselt, K., Stansberry, J., Boyer, M., et al. (2023). "Performance of NIRCam on JWST in flight". *Publications of the Astronomical Society of the Pacific* 135.1044, p. 028001. DOI: 10.1088/1538-3873/acac53.

Rigby, F. E., Pica-Ciamarra, L., Holmberg, M., Madhusudhan, N., Constantinou, S., et al. (2024). "Toward a self-consistent evaluation of gas dwarf scenarios for temperate sub-Neptunes". *The Astrophysical Journal* 975.1, pp. 101–124. DOI: 10.3847/1538-4357/ad6c38.

Robberto, M. (2011). *Implementation of a new algorithm for correction of non-linearity in JWST detectors*. Tech. rep. STScI, Technical Report JWST–STScI–002344. URL: https://ui.adsabs.harvard.edu/abs/2011jwst.rept.2344R.

Rocchetto, M., Waldmann, I. P., Venot, O., Lagage, P.-O., and Tinetti, G. (2016). "Exploring biases of atmospheric retrievals in simulated JWST transmission spectra of hot Jupiters". *The Astrophysical Journal* 833.1, pp. 120–133. DOI: 10.3847/1538-4357/833/1/120.

Rodgers, C. D. (2000). *Inverse methods for atmospheric sounding: theory and practice*. Vol. 00002. Series on Atmospheric, Oceanic and Planetary Physics. Singapore: World Scientific. ISBN: 978-981-02-2740-1. URL: https://research.ebsco.com/linkprocessor/plink?id=6516d526-1c5f-3109-ba33-9c505141b4eb.

Rogers, L. A. and Seager, S. (2010). "A framework for quantifying the degeneracies of exoplanet interior compositions". *The Astrophysical Journal* 712.2, pp. 974–992. DOI: 10.1088/0004-637X/712/2/974.

Rogers, L. A. (2015). "Most 1.6 earth-radius planets are not rocky". *The Astrophysical Journal* 801.1, pp. 41–54. DOI: 10.1088/0004-637X/801/1/41.

Rogers, L. A., Bodenheimer, P., Lissauer, J. J., and Seager, S. (2011). "Formation and structure of low-density exo-neptunes". *The Astrophysical Journal* 738.1, pp. 59–75. DOI: 10.1088/0004-637X/738/1/59.

Rothman, L. S., Gordon, I. E., Barber, R. J., Dothe, H., Gamache, R. R., et al. (2010). "HITEMP, the high-temperature molecular spectroscopic database". *Journal of Quantitative Spectroscopy and Radiative Transfer*. XVIth Symposium on High Resolution Molecular Spectroscopy (HighRus-2009) 111.15, pp. 2139–2150. DOI: 10.1016/j.jqsrt.2010.05.001.

Roudier, G. M., Swain, M. R., Gudipati, M. S., West, R. A., Estrela, R., et al. (2021). "Disequilibrium chemistry in exoplanet atmospheres observed with the Hubble Space Telescope". *Astronomical Journal* 162.2, pp. 37–50. DOI: 10.3847/1538-3881/abfdad.

Rustamkulov, Z., Sing, D. K., Mukherjee, S., May, E. M., Kirk, J., et al. (2023). "Early release science of the exoplanet WASP-39b with JWST NIRSpec PRISM". *Nature* 614.7949, pp. 659–663. DOI: 10.1038/s41586-022-05677-y.

Rustamkulov, Z., Sing, D. K., Liu, R., and Wang, A. (2022). "Analysis of a JWST NIRSpec lab time series: characterizing systematics, recovering exoplanet transit spectroscopy, and constraining a noise floor". *The Astrophysical Journal Letters* 928.1, p. L7. DOI: 10.3847/2041-8213/ac5b6f.

Sahlmann, J., Lazorenko, P. F., Ségransan, D., Martín, E. L., Queloz, D., et al. (2013). "Astrometric orbit of a low-mass companion to an ultracool dwarf". *Astronomy & Astrophysics* 556, A133. DOI: 10.1051/0004-6361/201321871.

Salstein, D. A. (1995). "Mean properties of the atmosphere". In: *Composition, chemistry, and climate of the atmosphere*. New York: Van Nostrand Reinhold. ISBN: 0-442-01264-0.

Sarkar, S. and Madhusudhan, N. (2021). "JexoSim 2.0: end-to-end JWST simulator for exoplanet spectroscopy – implementation and case studies". *Monthly Notices of the Royal Astronomical Society* 508.1, pp. 433–452. DOI: 10.1093/mnras/stab2472.

Sarkar, S., Madhusudhan, N., Constantinou, S., and Holmberg, M. (2024). "Exoplanet transit spectroscopy with JWST NIRSpec: diagnostics and homogeneous case study of WASP-39 b". *Monthly Notices of the Royal Astronomical Society* 531.2, pp. 2731–2756. DOI: 10.1093/mnras/stae1230.

Schaefer, L. and Fegley, B. (2009). "Chemistry of silicate atmospheres of evaporating super-earths". *The Astrophysical Journal* 703.2, p. L113. DOI: 10.1088/0004-637X/703/2/L113.

Schlawin, E., Rieke, M., Leisenring, J., Walker, L. M., Fraine, J., et al. (2016). "Two NIRCam channels are better than one: how JWST can do more science with NIRCam's short-wavelength dispersed hartmann sensor". *Publications of the Astronomical Society of the Pacific* 129.971, p. 015001. DOI: 10.1088/1538-3873/129/971/015001.

Schlawin, E., Mukherjee, S., Ohno, K., Bell, T. J., Beatty, T. G., et al. (2024). "Multiple clues for dayside aerosols and temperature gradients in WASP-69 b from a panchromatic JWST emission spectrum". *The Astronomical Journal* 168.3, p. 104. DOI: 10.3847/1538-3881/ad58e0.

Schleich, S., Boro Saikia, S., Changeat, Q., Güdel, M., Voigt, A., et al. (2024). "Knobs and dials of retrieving JWST transmission spectra - I. The importance of p–T profile complexity". *Astronomy & Astrophysics* 690, A336. DOI: 10.1051/0004-6361/202451845.

Schleich, S., Boro Saikia, S., Changeat, Q., Güdel, M., Voigt, A., et al. (2025). "Knobs and dials of retrieving JWST transmission spectra - II. Impacts of pipeline-level differences on retrieval posteriors". *Astronomy & Astrophysics* 704, A223. DOI: 10.1051/0004-6361/202556553.

Schleich, S., Changeat, Q., Boro Saikia, S., Güdel, M., and Voigt, A. (in prep.). "Knobs and dials of retrieving JWST transmission spectra - III. Panchromatic preferences in p–T profile complexity".

Schlichting, H. E. (2014). "Formation of close in super-Earths and mini-Neptunes: required disk masses and their implications". *The Astrophysical Journal Letters* 795.1, p. L15. DOI: 10.1088/2041-8205/795/1/L15.

Schmidt, S. P., MacDonald, R. J., Tsai, S.-M., Radica, M., Wang, L.-C., et al. (2025). "A comprehensive reanalysis of K2-18 b's JWST NIRISS+NIRSpec transmission spectrum". *The Astronomical Journal* 170.6, p. 298. DOI: 10.3847/1538-3881/ae019a.

Seiff, A., Kirk, D. B., Knight, T. C. D., Young, R. E., Mihalov, J. D., et al. (1998). "Thermal structure of Jupiter's atmosphere near the edge of a 5-μm hot spot in the north equatorial belt". *Journal of Geophysical Research: Planets* 103.E10, pp. 22857–22889. DOI: 10.1029/98JE01766.

Sellke, T., Bayarri, M. J, and Berger, J. O. (2001). "Calibration of ρ Values for Testing Precise Null Hypotheses". *The American Statistician* 55.1, pp. 62–71. DOI: 10.1198/000313001300339950.

Sharma, S. (2017). "Markov chain monte carlo methods for bayesian data analysis in astronomy". *Annual Review of Astronomy and Astrophysics* 55.Volume 55, 2017, pp. 213–259. DOI: 10.1146/annurev-astro-082214-122339.

Silvotti, R., Schuh, S., Janulis, R., Solheim, J.-E., Bernabei, S., et al. (2007). "A giant planet orbiting the 'extreme horizontal branch' star V 391 pegasi". *Nature* 449.7159, pp. 189–191. DOI: 10.1038/nature06143.

Sing, D. K., Fortney, J. J., Nikolov, N., Wakeford, H. R., Kataria, T., et al. (2016). "A continuum from clear to cloudy hot-jupiter exoplanets without primordial water depletion". *Nature* 529.7584, pp. 59–62. DOI: 10.1038/nature16068.

Skilling, J. (2006). "Nested sampling for general bayesian computation". *Bayesian Analysis* 1.4, pp. 833–859. DOI: 10.1214/06-BA127.

Stevenson, K. B., Désert, J.-M., Line, M. R., Bean, J. L., Fortney, J. J., et al. (2014). "Thermal structure of an exoplanet atmosphere from phase-resolved emission spectroscopy". *Science* 346.6211, pp. 838–841. DOI: 10.1126/science.1256758.

Stock, J. W., Kitzmann, D., Patzer, A. B. C., and Sedlmayr, E. (2018). "FastChem: a computer program for efficient complex chemical equilibrium calculations in the neutral/ionized gas phase with applications to stellar and planetary atmospheres". *Monthly Notices of the Royal Astronomical Society* 479.1, pp. 865–874. DOI: 10.1093/mnras/sty1531.

Stubenrauch, C. J., Cros, S., Lamquin, N., Armante, R., Chédin, A., et al. (2008). "Cloud properties from atmospheric infrared sounder and evaluation with cloud-aerosol lidar and infrared pathfinder satellite observations". *Journal of Geophysical Research: Atmospheres* 113.D8, pp. 0–17. DOI: 10.1029/2008JD009928.

Sun, Q., Wang, S. X., Welbanks, L., Teske, J., and Buchner, J. (2024). "A Revisit of the Mass–Metallicity Trends in Transiting Exoplanets". *Astronomical Journal* 167.4, p. 167. DOI: 10.3847/1538-3881/ad298d.

Tennyson, J. (2019). *Astronomical spectroscopy: an introduction to the atomic and molecular physics of astronomical spectroscopy*. 3rd ed. Advanced textbooks in physics. London: World Scientific Publishing Europe Ltd. ISBN: 978-1-78634-694-0.

Tennyson, J., Yurchenko, S. N., Al-Refaie, A. F., Clark, V. H. J., Chubb, K. L., et al. (2020). "The 2020 release of the ExoMol database: Molecular line lists for exoplanet and other hot atmospheres". *Journal of Quantitative Spectroscopy and Radiative Transfer* 255, p. 107228. DOI: 10.1016/j.jqsrt.2020.107228.

Terragni, J., Gordon, I. E., Adkins, E. M., Boulet, C., Campargue, A., et al. (2025). "Collision induced absorption in HITRAN2024: enhanced and improved data for atmospheric and planetary studies". *Journal of Quantitative Spectroscopy & Radiative Transfer* 347, p. 109631. DOI: 10.1016/j.jqsrt.2025.109631.

Thompson, A., Biagini, A., Cracchiolo, G., Petralia, A., Changeat, Q., et al. (2024). "Correcting exoplanet transmission spectra for stellar activity with an optimized retrieval framework". *The Astrophysical Journal* 960.2, p. 107. DOI: 10.3847/1538-4357/ad0369.

Thorngren, D. P., Sing, D. K., and Mukherjee, S. (2025). *Bayesian Model Comparison and Significance: Widespread Errors and how to Correct Them*. DOI: 10.48550/arXiv.2510.00169.

Tinetti, G., Eccleston, P., Haswell, C., Lagage, P.-O., Leconte, J., et al. (2021). *Ariel: Enabling planetary science across light-years*. DOI: 10.48550/arXiv.2104.04824.

Tinetti, G., Vidal-Madjar, A., Liang, M.-C., Beaulieu, J.-P., Yung, Y., et al. (2007). "Water vapour in the atmosphere of a transiting extrasolar planet". *Nature* 448.7150, pp. 169–171. DOI: 10.1038/nature06002.

TrExoLiSTS (2026). *Transiting Exoplanets List of Space Telescope Spectroscopy*. URL: https://www.stsci.edu/~nnikolov/TrExoLiSTS/JWST/trexolists.html (visited on 01/05/2026).

Trilling, D. E., Benz, W., Guillot, T., Lunine, J. I., Hubbard, W. B., et al. (1998). "Orbital evolution and migration of giant planets: modeling extrasolar planets". *The Astrophysical Journal* 500.1, pp. 428–440. DOI: 10.1086/305711.

Trotta, R. (2008). “Bayes in the sky: bayesian inference and model selection in cosmology”. *Contemporary Physics* 49.2, pp. 71–104. doi: 10.1080/00107510802066753.

Tsai, S.-M., Lee, E. K. H., Powell, D., Gao, P., Zhang, X., et al. (2023). “Photochemically produced SO2 in the atmosphere of WASP-39b”. *Nature* 617.7961, pp. 483–487. doi: 10.1038/s41586-023-05902-2.

Tsiaras, A., Waldmann, I. P., Zingales, T., Rocchetto, M., Morello, G., et al. (2018). “A population study of gaseous exoplanets”. *The Astronomical Journal* 155.4, pp. 156–171. doi: 10.3847/1538-3881/aaaf75.

Tsiaras, A., Waldmann, I. P., Tinetti, G., Tennyson, J., and Yurchenko, S. N. (2019). “Water vapour in the atmosphere of the habitable-zone eight-earth-mass planet K2-18 b”. *Nature Astronomy* 3.12, pp. 1086–1091. doi: 10.1038/s41550-019-0878-9.

Tusay, N., Wright, J. T., Beatty, T. G., Desch, S., Colón, K., et al. (2025). “A disintegrating rocky world shrouded in dust and gas: mid-infrared observations of K2-22 b using JWST”. *The Astrophysical Journal Letters* 987.1, p. L6. doi: 10.3847/2041-8213/addfd0.

Udalski, A., Kubiak, M., and Szymanski, M. (1997). “Optical gravitational lensing experiment. OGLE-2 – the second phase of the OGLE project”. *Acta Astronomica* 47, pp. 319–344. doi: 10.48550/arXiv.astro-ph/9710091.

Udalski, A., Zebrun, K., Szymanski, M., Kubiak, M., Soszynski, I., et al. (2002). “The optical gravitational lensing experiment. Search for planetary and low- luminosity object transits in the galactic disk. Results of 2001 campaign – supplement”. *Acta Astronomica* 52, pp. 115–128. doi: 10.48550/arXiv.astro-ph/0207133.

Underwood, D. S., Tennyson, J., Yurchenko, S. N., Huang, X., Schwenke, D. W., et al. (2016). “ExoMol molecular line lists – XIV. The rotation–vibration spectrum of hot SO2”. *Monthly Notices of the Royal Astronomical Society* 459.4, pp. 3890–3899. doi: 10.1093/mnras/stw849.

Van Eylen, V, Agentoft, C., Lundkvist, M. S., Kjeldsen, H, Owen, J. E., et al. (2018). “An asteroseismic view of the radius valley: stripped cores, not born rocky”. *Monthly Notices of the Royal Astronomical Society* 479.4, pp. 4786–4795. doi: 10.1093/mnras/sty1783.

Van Looveren, G., Boro Saikia, S., Herbort, O., Schleich, S., Güdel, M., et al. (2025). “Habitable zone and atmosphere retention distance (HaZARD) - stellar-evolution-dependent loss models of secondary atmospheres”. *Astronomy & Astrophysics* 694, A310. doi: 10.1051/0004-6361/202452998.

Van Looveren, G., Güdel, M., Boro Saikia, S., and Kislyakova, K. (2024). “Airy worlds or barren rocks? On the survivability of secondary atmospheres around the TRAPPIST-1 planets”. *Astronomy & Astrophysics* 683, A153. doi: 10.1051/0004-6361/202348079.

Vandaele, A. C., Aoki, S., Bauduin, S., Daerden, F., Fedorova, A., et al. (2024). "Composition and chemistry of the martian atmosphere as observed by Mars express and ExoMars trace gas orbiter". *Space Science Reviews* 220.7, p. 75. DOI: 10.1007/s11214-024-01109-7.

Vehtari, A., Gelman, A., and Gabry, J. (2017). "Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC". *Statistics and Computing* 27.5, pp. 1413–1432. DOI: 10.1007/s11222-016-9696-4.

Vidal-Madjar, A., Lecavelier des Etangs, A., Désert, J.-M., Ballester, G. E., Ferlet, R., et al. (2003). "An extended upper atmosphere around the extrasolar planet HD 209458 b". *Nature* 422.6928, pp. 143–146. DOI: 10.1038/nature01448.

Vink, R., Gooijer, S. de, Beedie, A., Burghoorn, G., nameexhaustion, et al. (2025). *pola-rs/polars: Python Polars 1.35.2*. Zenodo. DOI: 10.5281/zenodo.17564463.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., et al. (2020). "SciPy 1.0: fundamental algorithms for scientific computing in Python". *Nature Methods* 17.3, pp. 261–272. DOI: 10.1038/s41592-019-0686-2.

Vogt, S. S., Marcy, G. W., Butler, R. P., and Apps, K. (2000). "Six new planets from the keck precision velocity survey*". *The Astrophysical Journal* 536.2, pp. 902–915. DOI: 10.1086/308981.

Voyer, M., Changeat, Q., Lagage, P.-O., Tremblin, P., Waters, R., et al. (2025). "MIRI-LRS spectrum of a cold exoplanet around a white dwarf: water, ammonia, and methane measurements". *The Astrophysical Journal Letters* 982.2, p. L38. DOI: 10.3847/2041-8213/adbd46.

Wachiraphan, P., Berta-Thompson, Z. K., Diamond-Lowe, H., Winters, J. G., Murray, C., et al. (2025). "The thermal emission spectrum of the nearby rocky exoplanet LTT 1445A b from JWST MIRI/LRS". *The Astronomical Journal* 169.6, p. 311. DOI: 10.3847/1538-3881/adc990.

Wakeford, H. R., Sing, D. K., Deming, D., Lewis, N. K., Goyal, J., et al. (2018). "The Complete Transmission Spectrum of WASP-39b with a Precise Water Constraint". *The Astronomical Journal* 155.1, p. 29. DOI: 10.3847/1538-3881/aa9e4e.

Wakeford, H. R., Sing, D. K., Kataria, T., Deming, D., Nikolov, N., et al. (2017). "HAT-P-26b: A Neptune-mass exoplanet with a well-constrained heavy element abundance". *Science* 356.6338, pp. 628–631. DOI: 10.1126/science.aah4668.

Waldmann, I. P., Rocchetto, M., Tinetti, G., Barton, E. J., Yurchenko, S. N., et al. (2015). "Tau-REx II: Retrieval of Emission Spectra". *The Astrophysical Journal* 813.1, p. 13. DOI: 10.1088/0004-637X/813/1/13.

Weidenschilling, S. J. and Marzari, F. (1996). "Gravitational scattering as a possible origin for giant planets at small stellar distances". *Nature* 384.6610, pp. 619–621. DOI: 10.1038/384619a0.

Weiner Mansfield, M., Xue, Q., Zhang, M., Mahajan, A. S., Ih, J., et al. (2024). "No thick atmosphere on the terrestrial exoplanet Gl 486b". *The Astrophysical Journal Letters* 975.1, p. L22. DOI: 10.3847/2041-8213/ad8161.

Welbanks, L., Madhusudhan, N., Allard, N. F., Hubeny, I., Spiegelman, F., et al. (2019). "Mass–Metallicity Trends in Transiting Exoplanets from Atmospheric Abundances of H2O, Na, and K". *The Astrophysical Journal Letters* 887.1, p. L20. DOI: 10.3847/2041-8213/ab5a89.

Welbanks, L., McGill, P., Line, M., and Madhusudhan, N. (2023). "On the Application of Bayesian Leave-one-out Cross-validation to Exoplanet Atmospheric Analysis". *The Astronomical Journal* 165.3, p. 112. DOI: 10.3847/1538-3881/acab67.

Welbanks, L., Nixon, M. C., McGill, P., Tilke, L. J., Wiser, L. S., et al. (2026). "Challenges in the detection of gases in exoplanet atmospheres". *Nature Astronomy* 10.2, pp. 234–247. DOI: 10.1038/s41550-025-02730-4.

Williams, P. K. G., Charbonneau, D., Cooper, C. S., Showman, A. P., and Fortney, J. J. (2006). "Resolving the surfaces of extrasolar planets with secondary eclipse light curves". *Astrophysical Journal* 649.2, p. 1020. DOI: 10.1086/506468.

Winn, J. N. (2010). "Exoplanet transits and occultations". In: *Exoplanets*. Ed. by S. Seager, pp. 55–77. DOI: 10.48550/arXiv.1001.2010.

Wiser, L. S., Bell, T. J., Line, M. R., Schlawin, E., Beatty, T. G., et al. (2025). "A precise metallicity and carbon-to-oxygen ratio for a warm giant exoplanet from its panchromatic JWST emission spectrum". *Proceedings of the National Academy of Sciences* 122.39, e2416193122. DOI: 10.1073/pnas.2416193122.

Wit, J. de, Gillon, M., Demory, B.-O., and Seager, S. (2012). "Towards consistent mapping of distant worlds: secondary-eclipse scanning of the exoplanet HD 189733b". *Astronomy & Astrophysics* 548, A128. DOI: 10.1051/0004-6361/201219060.

Woitke, P., Helling, C., Hunter, G. H., Millard, J. D., Turner, G. E., et al. (2018). "Equilibrium chemistry down to 100 K - impact of silicates and phyllosilicates on the carbon to oxygen ratio," *Astronomy & Astrophysics* 614, A1. DOI: 10.1051/0004-6361/201732193.

Wolszczan, A. and Frail, D. A. (1992). "A planetary system around the millisecond pulsar PSR1257 + 12". *Nature* 355.6356, pp. 145–147. DOI: 10.1038/355145a0.

Wolszczan, A. (1994). "Confirmation of Earth-mass planets orbiting the millisecond pulsar PSR B1257 + 12". *Science* 264.5158, pp. 538–542. DOI: 10.1126/science.264.5158.538.

Wright, G. S., Rieke, G. H., Glasse, A., Ressler, M., García Marín, M., et al. (2023). "The mid-infrared Instrument for JWST and its in-flight performance". *Publications of the Astronomical Society of the Pacific* 135.1046, p. 048003. DOI: 10.1088/1538-3873/acbe66.

Wright, J. T. (2018). “Radial velocities as an exoplanet discovery method”. In: *Handbook of exoplanets*. Ed. by H. J. Deeg and J. A. Belmonte. Springer, Cham, pp. 619–631. ISBN: 978-3-319-55333-7. DOI: 10.1007/978-3-319-55333-7_4.

Xue, Q., Zhang, M., Coy, B. P., Brady, M., Ji, X., et al. (2025). “The JWST rocky worlds DDT program reveals GJ 3929 b to likely be a bare rock”. *The Astrophysical Journal Letters* 995.2, p. L52. DOI: 10.3847/2041-8213/ae2098.

Yao, Y., Vehtari, A., Simpson, D., and Gelman, A. (2018). “Using stacking to average bayesian predictive distributions (with discussion)”. *Bayesian Analysis* 13.3, pp. 917–1007. DOI: 10.1214/17-BA1091.

Yurchenko, S. N., Mellor, T. M., Freedman, R. S., and Tennyson, J (2020). “ExoMol line lists – XXXIX. Ro-vibrational molecular line list for CO2”. *Monthly Notices of the Royal Astronomical Society* 496.4, pp. 5282–5291. DOI: 10.1093/mnras/staa1874.

Yurchenko, S. N., Amundsen, D. S., Tennyson, J., and Waldmann, I. P. (2017). “A hybrid line list for CH4 and hot methane continuum”. *Astronomy & Astrophysics* 605, A95. DOI: 10.1051/0004-6361/201731026.

Yurchenko, S. N., Owens, A., Kefala, K., and Tennyson, J. (2024). “ExoMol line lists – LVII: High accuracy ro-vibrational line list for methane (CH4)”. *Monthly Notices of the Royal Astronomical Society* 528.2, pp. 3719–3729. DOI: 10.1093/mnras/stae148.

Zeng, L., Jacobsen, S. B., Sasselov, D. D., Petaev, M. I., Vanderburg, A., et al. (2019). “Growth model interpretation of planet size distribution”. *Proceedings of the National Academy of Sciences* 116.20, pp. 9723–9728. DOI: 10.1073/pnas.1812905116.

Zhang, M., Hu, R., Inglis, J., Dai, F., Bean, J. L., et al. (2024a). “GJ 367b is a dark, hot, airless sub-Earth”. *The Astrophysical Journal Letters* 961.2, p. L44. DOI: 10.3847/2041-8213/ad1a07.

Zhang, M., Paragas, K., Bean, J. L., Yeung, J., Chachan, Y., et al. (2024b). “Retrievals on NIRCam transmission and emission spectra of HD 189733b with PLATON 6, a GPU code for the JWST era”. *The Astronomical Journal* 169.1, pp. 38–57. DOI: 10.3847/1538-3881/ad8cd2.

Zhang, Z., Mukherjee, S., Liu, M. C., Fortney, J. J., Mader, E., et al. (2024c). “Disequilibrium chemistry, diabatic thermal structure, and clouds in the atmosphere of COCONUTS-2b”. *The Astronomical Journal* 169.1, p. 9. DOI: 10.3847/1538-3881/ad8b2d.

Zhu, W. and Dong, S. (2021). “Exoplanet statistics and theoretical implications”. *Annual Review of Astronomy and Astrophysics* 59.Volume 59, 2021, pp. 291–336. DOI: 10.1146/annurev-astro-112420-020055.

Zieba, S., Zilinskas, M., Kreidberg, L., Nguyen, T. G., Miguel, Y., et al. (2022). “K2 and Spitzer phase curves of the rocky ultra-short-period planet K2-141 b hint at a tenuous rock vapor atmosphere”. *Astronomy & Astrophysics* 664, A79. DOI: 10.1051/0004-6361/202142912.

Zieba, S., Kreidberg, L., Ducrot, E., Gillon, M., Morley, C., et al. (2023). "No thick carbon dioxide atmosphere on the rocky exoplanet TRAPPIST-1 c". *Nature*, pp. 1–4. DOI: 10.1038/s41586-023-06232-z.

Appendices

Appendix A

JWST NIR transit depth uncertainty

To produce synthetic transmission spectra for the NIR instruments of JWST (excluding the NIRSpec PRISM case), we make use of the homogeneous data reduction presented in Carter et al. (2024). These cover observations of WASP-39 b transits performed with NIRSpec G395H, NIRISS SOSS, and NIRCam F322W2. More specifically, we choose the results[1] associated with the highest spectral resolution (`bins_scale1`) and derived by fitting the LDCs. The latter results in slightly larger overall error bars, providing more conservative transit-depth uncertainties.

The error bars reported in Carter & May et al. (2024) contain wavelength-dependent outliers, as we can expect for real measurements. To mitigate these observation-specific outliers, we smooth over the error bar curves twice, using a rolling median with a box width of 5% of the total wavelength range. In Figure A.1, we see the resulting smoothed error bar curves compared to the results for free and fixed LDCs. In all panels of the plot, it is clearly visible that the error bars from data reductions using free-parameter LDCs are higher.

[1]The data for these were retrieved from `https://zenodo.org/records/10161742`

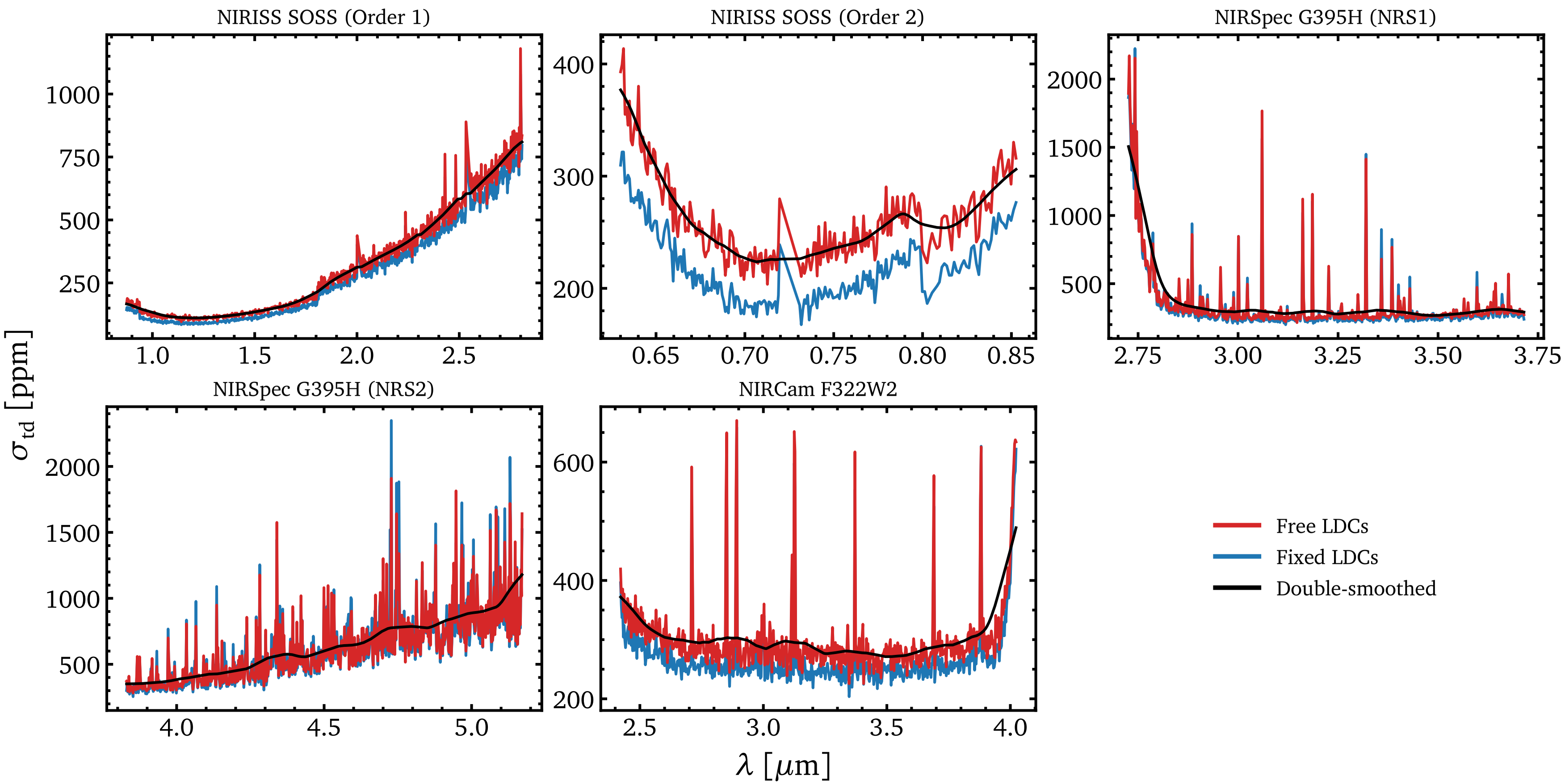

Figure A.1: Transit depth errors from Carter & May et al. (2024) for transit observations of WASP-39 b with different NIR instruments of JWST, showing uncertainty (in ppm) against wavelength (in μm). In each panel, the error bars for fits with the LDCs as free parameters and with the LDCs fixed are plotted in red and blue, respectively. Smoothed curves of the free-LDC uncertainties are shown in black.

Appendix B

Additional synthetic retrieval plots

Figures B.1 through B.8 summarise the parameter estimation accuracy and p–T structure recovery for all simulated instrument cases we have seen in Table 3.3. They follow the same structure as Figure 3.6 and Figure 3.7, showing the results from two underlying p–T cases (monotonic and inverse) and three input cloud decks (high pressure, medium pressure, and low pressure).

NIRSpec PRISM - Pandexo

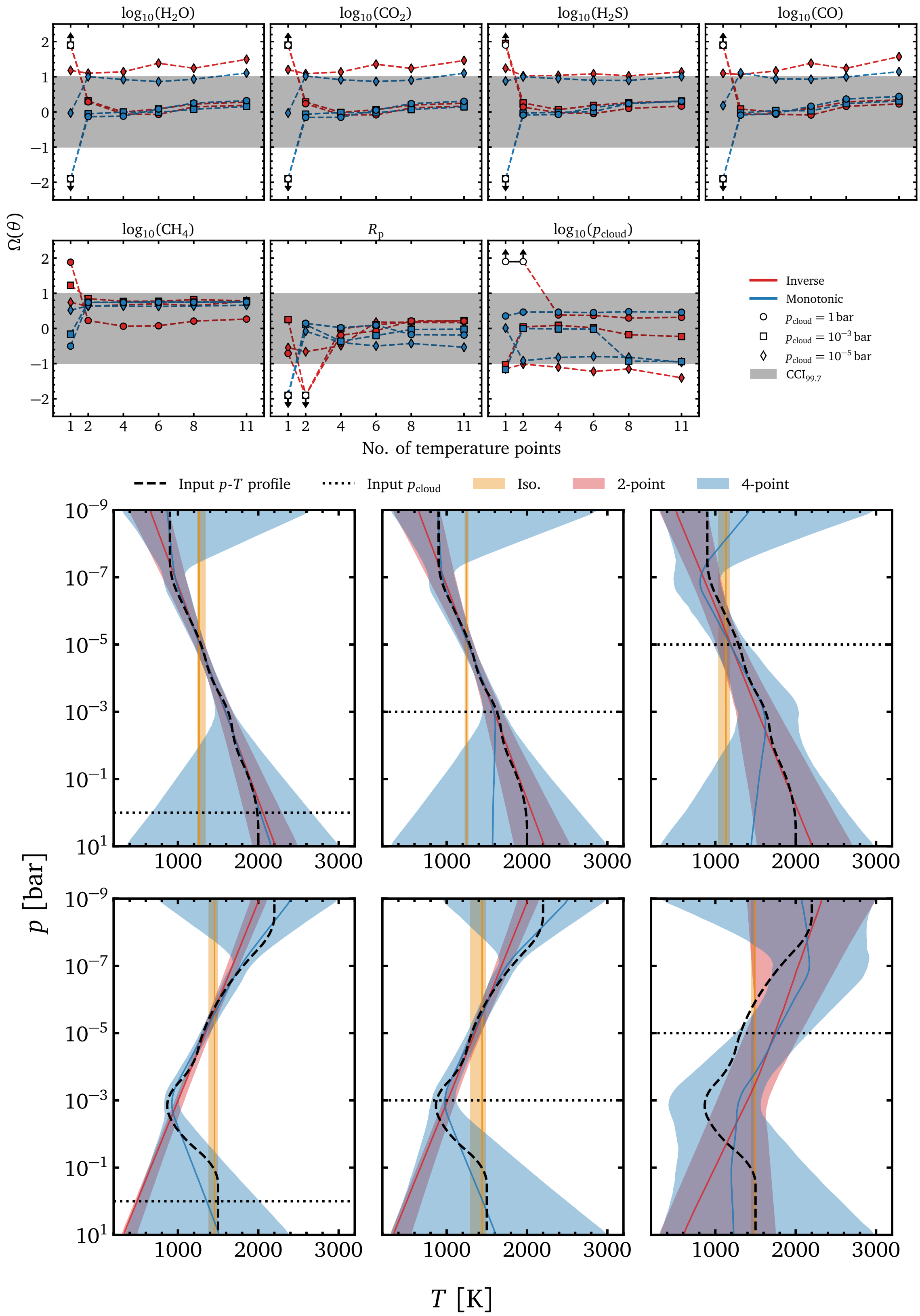

Figure B.1: Parameter accuracy and retrieved p–T profiles for the NIRSpec PRISM Pandexo noise case.

Full JWST range

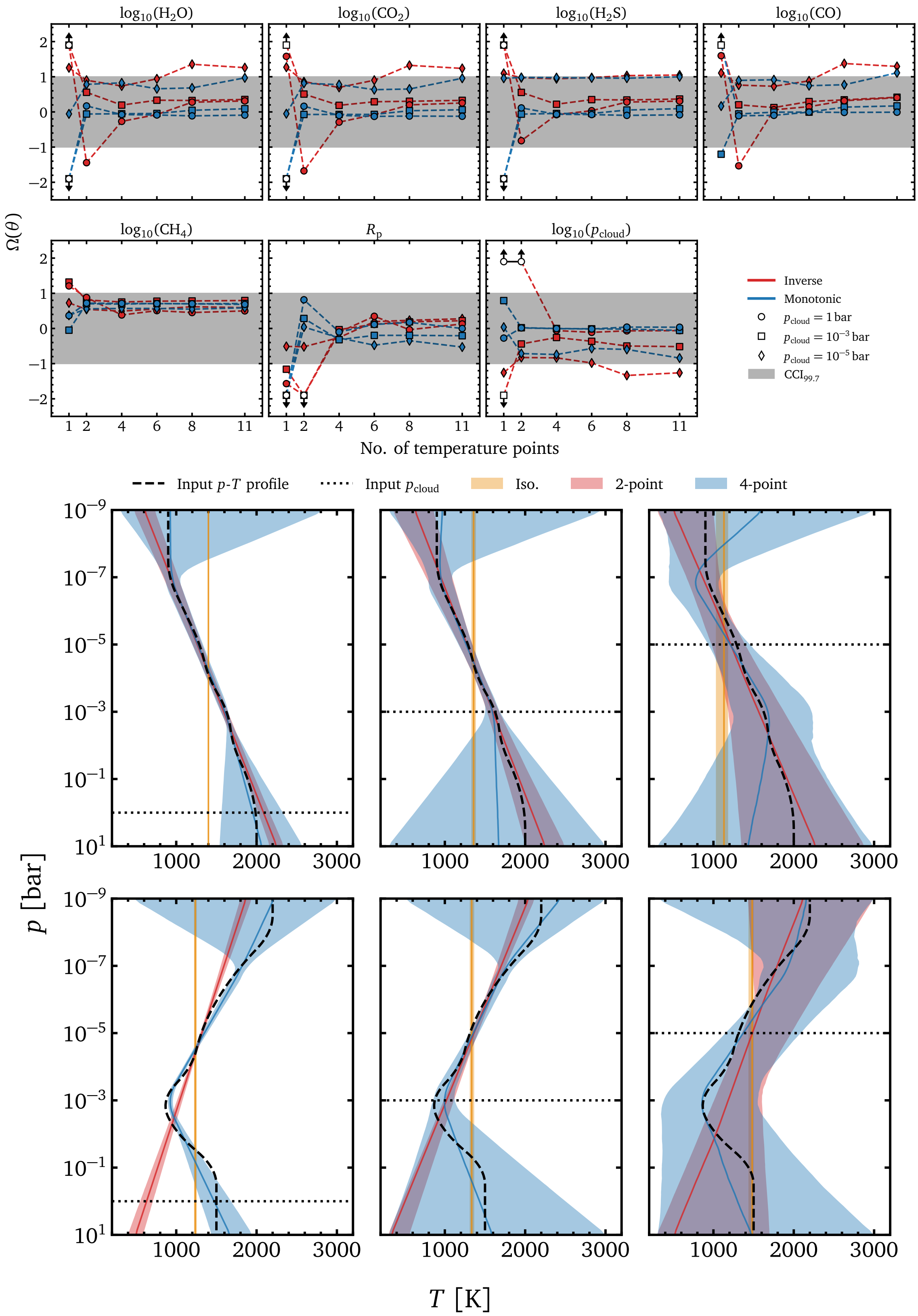

Figure B.2: Parameter accuracy and retrieved p–T profiles for the JWST (full) noise case.

NIRISS SOSS (order 1 and 2)

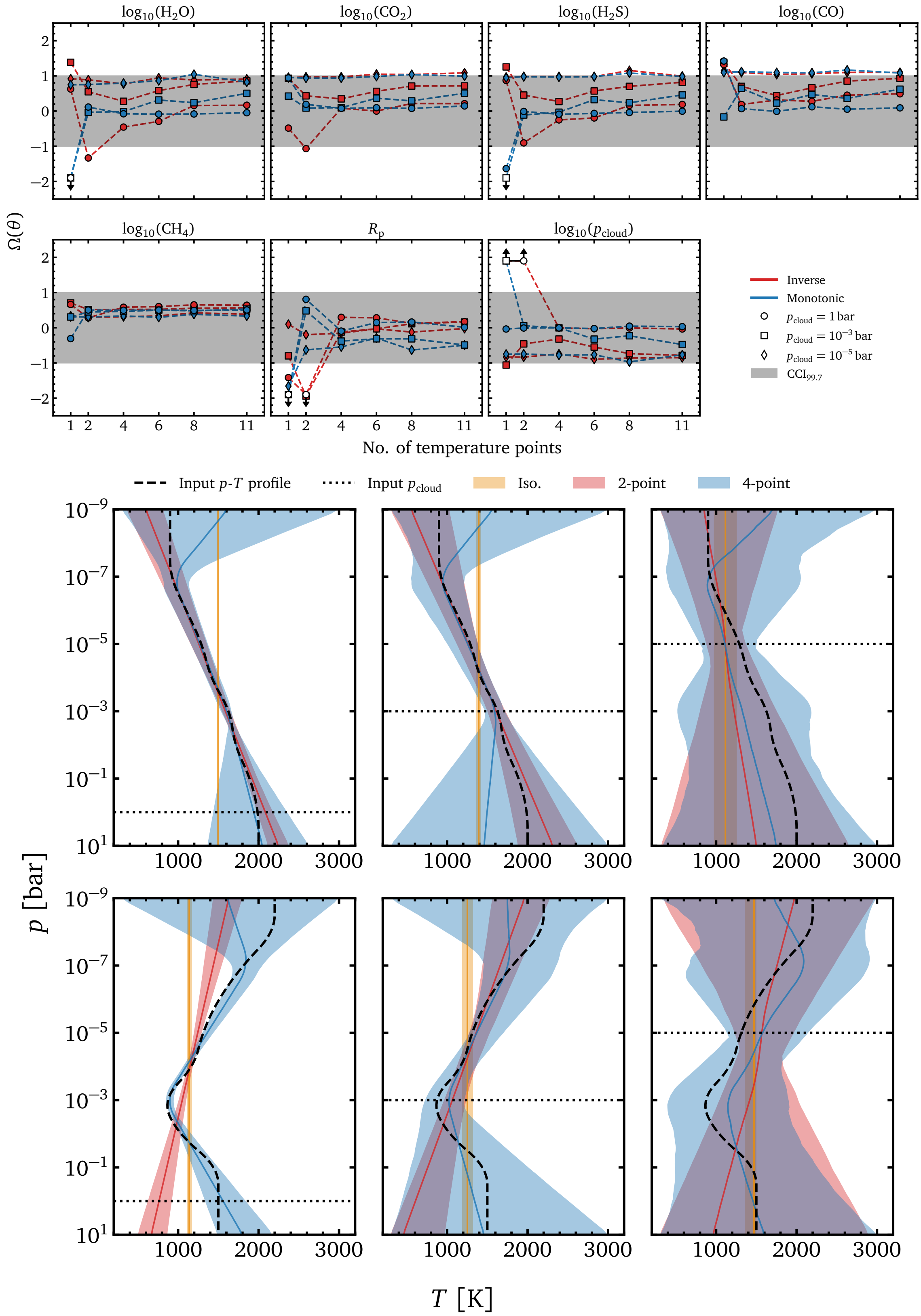

Figure B.3: Parameter accuracy and retrieved p–T profiles for the NIRISS SOSS noise case.

NIRSpec F290LP/G395H

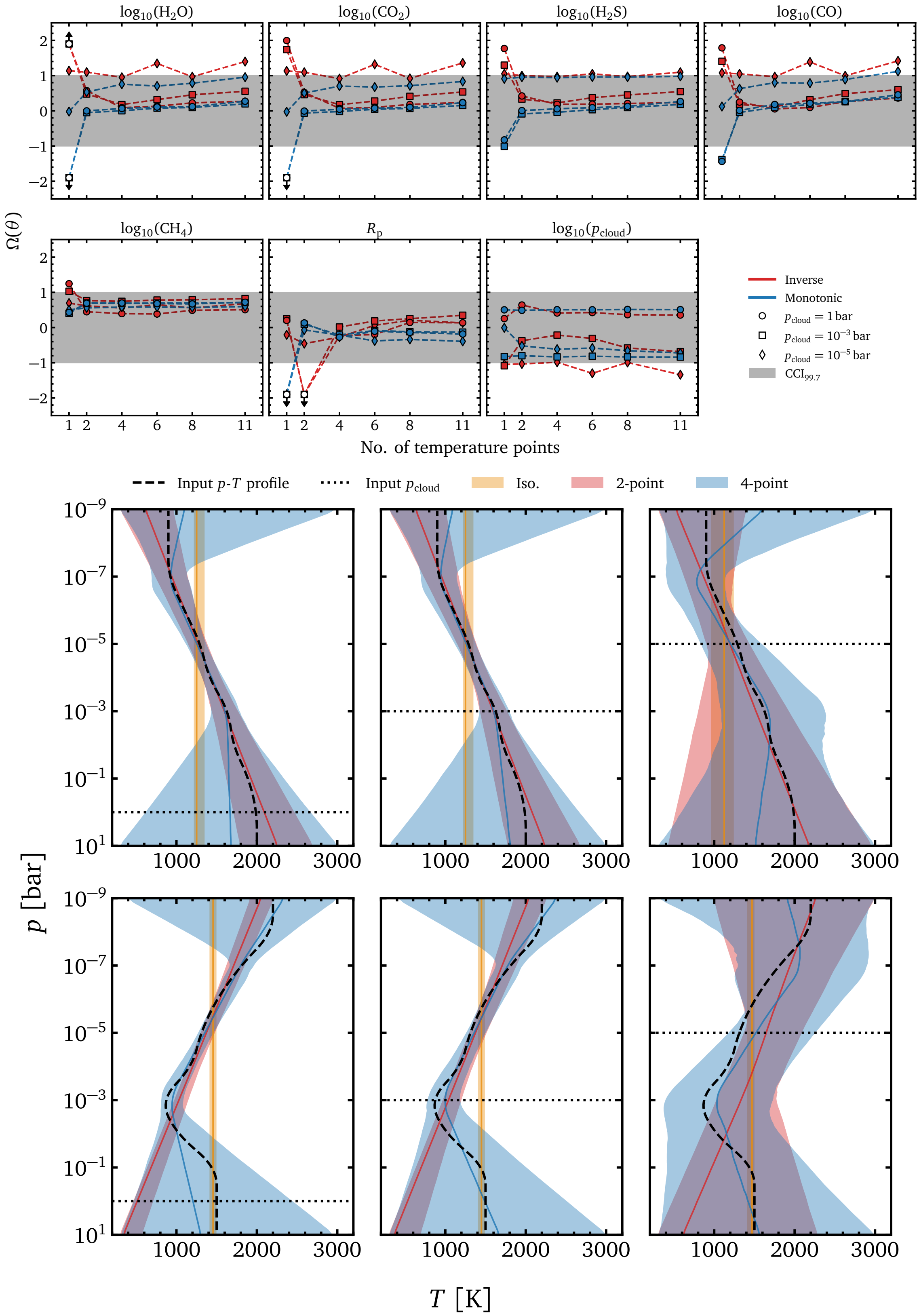

Figure B.4: Parameter accuracy and retrieved p–T profiles for the NIRSpec F290LP/G395H noise case.

NIRCam F322W2

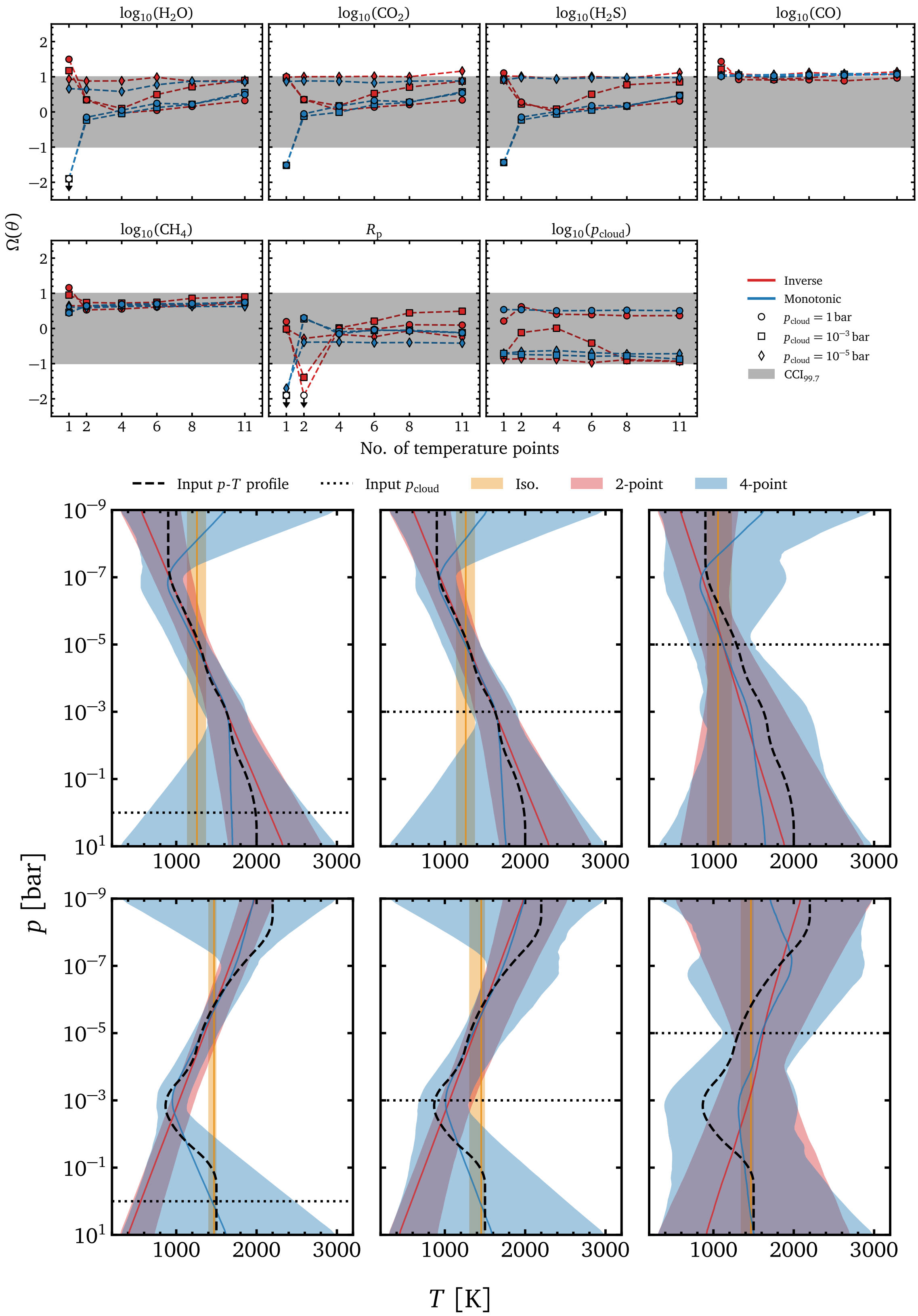

Figure B.5: Parameter accuracy and retrieved p–T profiles for the NIRCam F322W2 noise case.

MIRI LRS

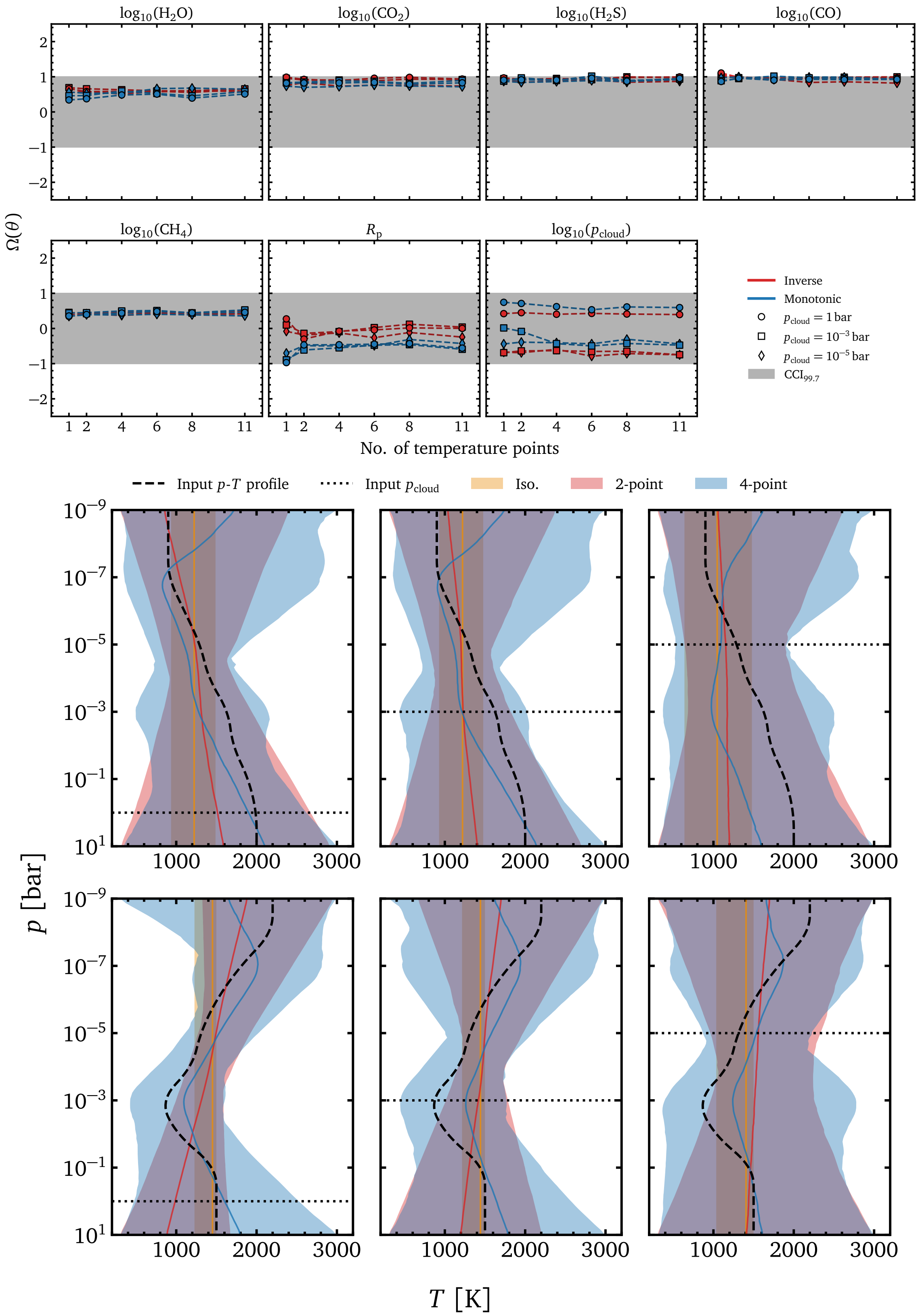

Figure B.6: Parameter accuracy and retrieved p–T profiles for the MIRI LRS noise case.

Ariel - WASP-39 (2 orbits)

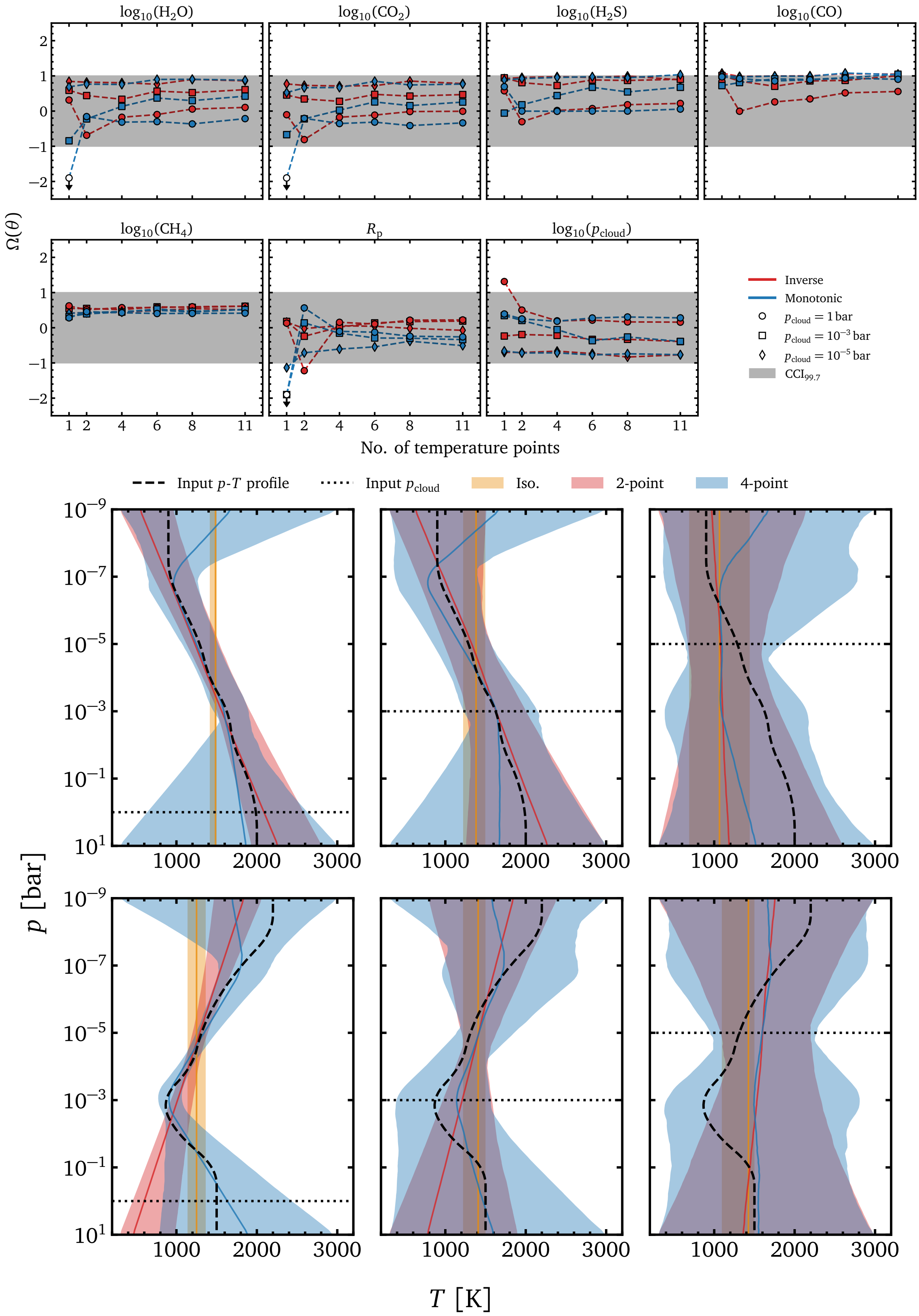

Figure B.7: Parameter accuracy and retrieved p–T profiles for the Ariel-WASP noise case.

Airel - HD 209458

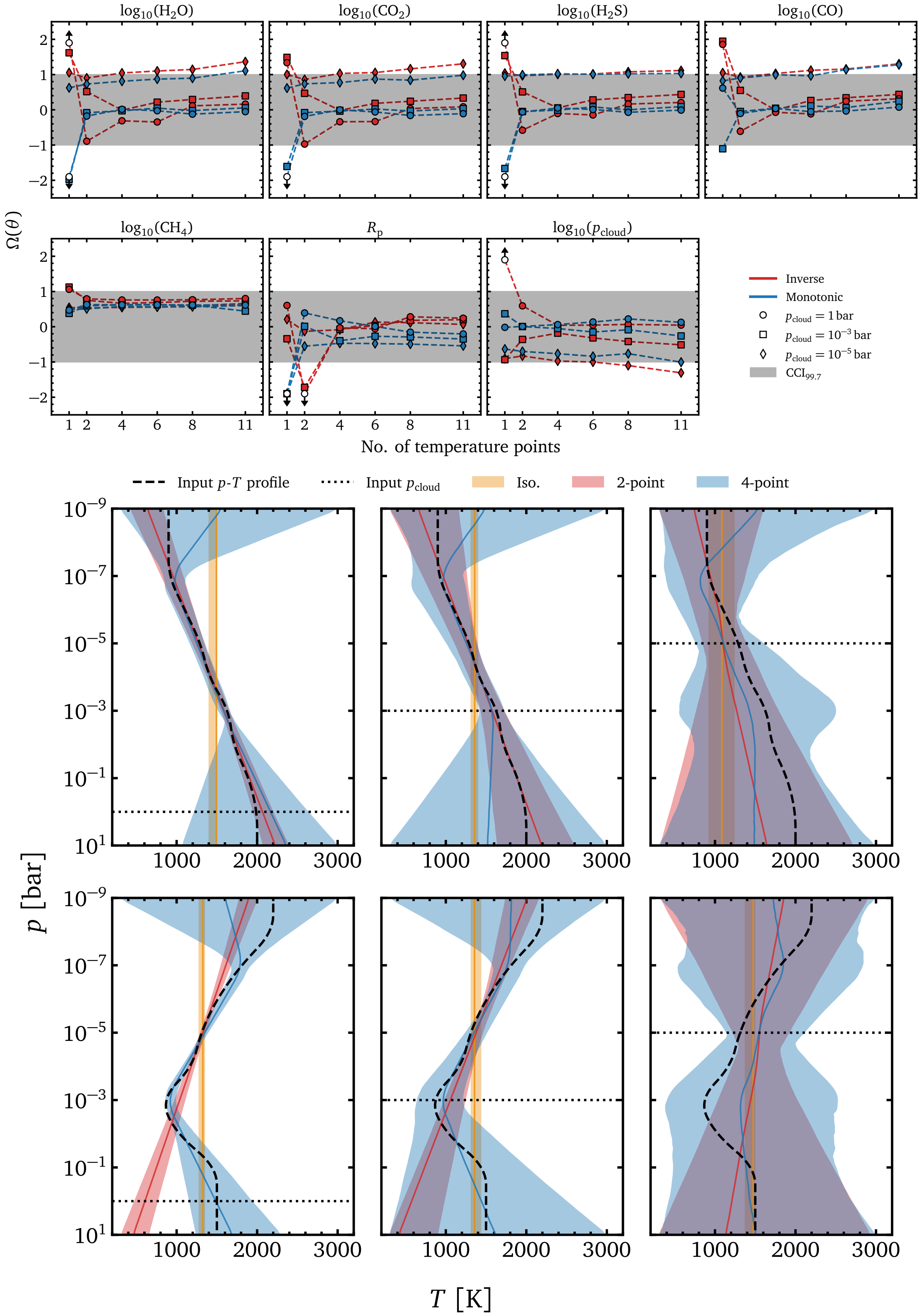

Figure B.8: Parameter accuracy and retrieved p–T profiles for the Ariel-HD noise case.

Data and software statement

The work conducted in my thesis would not have been possible without a significant portion of publicly available data and software. While I also credit them in the individual chapters of my thesis, I provide a summary overview of the sources of these here.

Own work

Data products associated with my own work presented in this thesis can be found in the following places (in chronological order): `https://doi.org/10.5281/zenodo.10497341` (Schleich et al. 2024); `https://doi.org/10.5281/zenodo.15697940` (Schleich et al. 2025). Parts of this thesis are based on unpublished results, in preparation for submission. The associated data products are available at `https://doi.org/10.5281/zenodo.15978362`.

Data

I used publicly available data associated with the following publications (in chronological order):

- JTERS23 – `https://doi.org/10.5281/zenodo.6959426`,
- Rustamkulov et al. (2023) – `https://doi.org/10.5281/zenodo.7388031`,
- Carter & May et al. (2024) – `https://doi.org/10.5281/zenodo.10161742`,
- Powell et al. (2024) – `https://doi.org/10.5281/zenodo.10055844`.

I also used JWST observational data from the programme #1366 (PI: N. Batalha, Co-PIs: J. Bean and K. Stevenson).

Software

The work I conducted during my PhD is built on two pieces of open source software – the atmospheric retrieval code `TauREx3` (Al-Refaie et al. 2021; 2022), and the JWST data reduction pipeline `Eureka!` (Bell et al. 2022), based on the pipeline `jwst` (Bushouse et al. 2024). I am immensely grateful to all the people who have contributed to the development of these two codes. I have also made use of the codes `Pandexo` (Batalha et al. 2017) and `JexoSim` (Sarkar and Madhusudhan 2021) to perform noise simulations for JWST observations, ArielRad

(Mugnai et al. 2020) to perform noise simulations for the Ariel mission, and `ExoTiC-LD` (Grant and Wakeford 2024) to calculate limb-darkening profiles for stars.

Last, but certainly not least, I also gratefully acknowledge the open source packages for the Python programming language I have used in my work (in alphabetical order): `corner` (Foreman-Mackey 2016); `h5py` (Collette 2013); `matplotlib` (Hunter 2007); `numpy` (Harris et al. 2020); `pandas` (McKinney 2010); `polars` (Vink et al. 2025); `scipy` (Virtanen et al. 2020). I also acknowledge the use of Grammarly (version 8.934.0) in reviewing the grammar and syntax of my manuscript.

List of Figures

List of Tables

List of Abbreviations

p*–*T	pressure-temperature
BMA	Bayesian model averaging
BOA	bottom of the atmosphere
CCI	centred credible interval
CDF	cumulative distribution function
CIA	collision-induced absorption
CRDS	calibration reference data system
CSA	Canadian Space Agency
DDT	Director's Discretionary Time
ERS	early release science
ESA	European Space Agency
HST	*Hubble* Space Telescope
IR	infrared
IRAC	Infrared Array Camera
JWST	*James Webb* Space Telescope
K-S	Kolmogorov-Smirnov
LDC	limb-darkening coefficient
MAST	Mikulski Archive for Space Telescopes
MCMC	Markov Chain Monte Carlo
MIR	mid-infrared
MMW	mean molecular weight
MR	mixing ratio
NASA	National Aeronautics and Space Administration
NIR	near-infrared
NS	Nested Sampling
PSF	point spread function
SNR	signal-to-noise ratio
Spitzer	*Spitzer* Space Telescope
STIS	Space Telescope Imaging Spectrograph
TOA	top of the atmosphere
TSO	time-series observation
WFC3	Wide Field Camera 3

Zusammenfassung

In den letzten drei Jahrzehnten ist die Zahl der uns bekannten Planeten außerhalb des Sonnensystems auf Tausende angestiegen und deckt einen weiten Parameterbereich ab. Die Charakterisierung der Atmosphären dieser Exoplaneten bietet einen Einblick in ihre Entstehungsgeschichte und ihre chemische Zusammensetzung. Spektroskopische Beobachtungen sind das wichtigste Mittel zur Untersuchung dieser Atmosphären, und das *James-Webb*-Weltraumteleskop (JWST) hat zusammen mit kommenden Missionen wie *Ariel* eine neue Ära der Charakterisierung von Atmosphären eingeläutet, indem es Daten mit beispiellosem Signal-Rausch-Verhältnis (SNR), Wellenlängenabdeckung und spektralen Auflösung liefert.

Um aus diesen Beobachtungen Rückschlüsse auf die Eigenschaften von exoplanetaren Atmosphären zu ziehen, stützen wir uns auf die Methode eines atmospärischen Inversionsprozesses (*atmospheric retrieval*), einen bayesianischen Inferenzansatz, bei dem ein atmosphärisches Vorwärtsmodell iterativ mit einem beobachteten Spektrum verglichen wird, um die A-posteriori-Verteilungen der Modellparameter zu gewinnen. Die verbesserte Datenqualität modernster Observatorien rechtfertigt eine sorgfältige Reevaluation dieser Methode. In dieser Arbeit untersuchen wir die Stabilität von atmosphärischen Parameterinferenzergebnissen im Kontext der Transmissionsspektroskopie, wobei wir uns auf zwei wichtige Säulen der darunter liegenden Methodik konzentrieren: das atmosphärische Vorwärtsmodell und die Beobachtungsdaten.

Im ersten Teil dieser Arbeit untersuchen wir den Einfluss des modellierten Druck-Temperatur-Profils (p–T-Profil) auf die Parameterinferenz von Transmissionsspektren. Anhand von synthetischen Spektren, die nach dem "heißen Jupiter" WASP-39 b modelliert wurden, erstellen wir einen Datensatz, der einen weiten Bereich vertikaler Temperaturstrukturen, Wolkendecken und Instrumentenszenarios für JWST und die bevorstehende *Ariel*-Mission abdeckt. Wir führen atmosphärische Semi-Selbst-Parameterinferenz dieser synthetischen Spektren durch, wobei wir heuristische Multipunkt-p–T-Profile unterschiedlicher Komplexität verwenden. Wir bewerten die Leistungsfähigkeit dieser atmosphärischen Modelle anhand des Bayes-Faktors und der Genauigkeit der geschätzten atmosphärischen Parameter. Die Ergebnisse zeigen durchweg, dass ein isothermes p–T-Profil nicht ausreicht, um die bekannten atmosphärischen Parameter unter modernsten Beobachtungsbedingungen korrekt zurückzugewinnen. In den meisten Fällen wird ein Multipunktprofil bevorzugt, wobei die Richtigkeit der Parameterinferenz bei einem Profil mit vier Punkten ein Plateau erreicht. Wir kommen zu dem Schluss, dass ein Multipunktprofil mit vier Punkten das optimale Gleichgewicht zwischen Modellkomplexität

und Inferenzpräzision darstellt und bei der Parameterinferenz von Transmissionsspektren mit JWST und *Ariel* gegenüber einem isothermen Profil vorzuziehen ist.

Im zweiten Teil dieser Arbeit untersuchen wir die Sensitivität von atmosphärischen Parameterinferenzergebnissen gegenüber Schwankungen im der Inferenz zu Grunde liegenden Transmissionsspektrums unter Verwendung von JWST NIRSpec PRISM-Beobachtungen eines individuellen Transits von WASP-39 b. Die Parameterinferenz wird zuerst anhand eines im Rahmen dieser Arbeit reduzierten Transmissionsspektrums durchgeführt, das mit der `Eureka!`-Datenreduktionspipeline erstellt wurde. Wir vergleichen die resultierenden a-posteriori-Parameterverteilungen mit homogenisierten atmosphärischen Parameterinferenzergebnissen, die über gestreute Instanzen des in dieser Arbeit erzeugten Spektrums sowie über unabhängig von denselben Beobachtungsdaten reduzierte Spektren erlangt wurden. Diese Analyse identifiziert drei Klassen von a-posteriori-Parameterverteilungen: stabile Normalverteilungen für gut eingeschränkte Parameter, stabile Gleichverteilungen für schwach eingeschränkte Parameter und instabile, stark asymmetrische a-posteriori-Verteilungen. Das Vorhandensein instabiler a-posteriori-Verteilungen erschwert eine robuste Interpretation der Ergebnisse der atmosphärischen Charakterisierung und unterstreicht die Notwendigkeit einer sorgfältigen Bewertung der Breite von Kredibilitätsintervallen, um diese Unsicherheit angemessen widerzuspiegeln.

Zusammenfassend unterstreichen die Ergebnisse dieser Arbeit die Bedeutung einer rigorosen Vorwärtsmodellauswahl und einer sorgfältigen Behandlung der Beobachtungsdaten bei atmosphärischen Inversionsanalysen. Da das Forschungsfeld in eine neue Ära erwarteter Datenqualität und -fülle eintritt, betonen diese Erkenntnisse die Notwendigkeit einer systematischen Evaluierung von Inversionsmethoden, um die Zuverlässigkeit der abgeleiteten atmosphärischen Eigenschaften zu gewährleisten.

Acknowledgements

This project was funded by the FGGA Emerging Field Grant 2021. The author acknowledges financial support by the University of Vienna and the Österreichische Forschungsgemeinschaft (ÖFG). This work is based in part on observations made with the NASA/ESA/CSA *James Webb* Space Telescope. The data were obtained from the Mikulski Archive for Space Telescopes at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-03127 for JWST. These observations are associated with programme #1366. The author acknowledges the Transiting Exoplanet Community Early Release Science Program team for developing their observing program with a zero-exclusive-access period. The computational results have been achieved in part using the Austrian Scientific Computing (ASC) infrastructure.

And thus we have reached the end of this project, three and a half years in the making. All of this would not have been possible without the guidance, support, and help of a lot of people.

First and foremost, I need to thank my supervisors, without whom I would not have had a research project to work on. Thanks to Sudeshna, for letting me explore all the questions I had about atmospheric retrieval, the observational data, and everything in between. I greatly enjoyed learning from and with you. Thanks to Manuel, for insightful discussion and feedback on my work, and also for being a great lecturer during my whole university journey, including my bachelor's and master's studies. And thanks to Aiko, for providing me with a perspective not pigeon-holed in Astrophysics. Hot Jupiters might not be the closest analogues to Earth, but the *Climate Dynamics and Modelling* group has provided me with a welcoming environment to discuss my research and present my work.

I also want to thank my close collaborators, Quentin and Ingo, for all their expertise, input, and patience. Special thanks go to Quentin, for hosting me in Baltimore and Groningen, and for putting up with my incessant questions. I am glad I did not scare you away immediately during our first meeting at ESTEC.

I have also benefited greatly from the scientific environment I was lucky enough to find my way into. At the University of Vienna, I enjoyed the wonderful company of the members of the *Star-Planet Alliance* (I still vote for that name!), especially the people I have had the pleasure of sharing an office with over the years. They had to listen to my ramblings about the design of plots and the pitfalls of statistics and data reduction too many times not to be mentioned here. Thanks to Gwenaëlle and Yanina (our honorary office member), for the silly things that entertained us over the years; to Nils and Louis, for our exploits in MTG; and to Ivan and Shiv, for inciting a passion in running and bouldering that has kept me healthy. A big thank-you also to the people who remotely struggled to make sense of atmospheric retrievals together with me: Maël, Jack, and Gordon.

I cannot forget to express my deepest gratitude to all the people that helped make this thesis possible on not a scientific, but a human level. The biggest thanks goes to my family – they believed in the fact that I will see the end of this journey, even when I sometimes did not. Thank you for all your support, and sorry for not visiting as often as I should have.

Thanks to Michael, Frodo, Markus, Steffi, Clemens, Tanja, and Lukas, for many relaxing evenings playing boardgames and other silly things. I will be your DM any day of the year (when there is not a dissertation breathing down my neck). Another big thank-you goes to Kevin and Lukas, for all the fun we have had remotely.

Lastly, and maybe most importantly, to Greta: I might have gone insane finishing this without you. From the bottom of my heart - thank you. *Sussurriamo sempre parole nelle bottiglie.*