

The background of the cover is a white field filled with a pattern of blue concentric hexagonal lines. These lines form a series of nested hexagons that create a sense of depth and movement, particularly towards the center of the page. The lines are thin and evenly spaced, creating a subtle, geometric texture.

*Atmospheric Characterisation
of Directly Imaged Exoplanets*

Evert Nasedkin

Dissertation
submitted to the
Combined Faculty of Mathematics, Engineering and Natural Sciences
of Heidelberg University, Germany
for the degree of
Doctor of Natural Sciences

Put forward by
EVERT NASEDKIN
born in: Grande Prairie, Alberta, Canada
Oral examination: 24.07.2024

Atmospheric Characterisation of Directly Imaged Exoplanets

Referees:

Prof. Dr. Laura Kreidberg

Prof. Dr. Hans-Walter Rix

THEMA

Seit wir in den Nachthimmel blicken, haben wir uns über die Beschaffenheit von Planeten innerhalb und außerhalb unseres Sonnensystems Gedanken gemacht. Die Beobachtungen der letzten drei Jahrzehnte haben eine atemberaubende Vielfalt von Exoplaneten in Bezug auf ihre Größe, Temperatur und Zusammensetzung offenbart. Von den Tausenden von bekannten Exoplaneten kann nur eine kleine Handvoll direkt abgebildet werden. Diese seltenen Systeme bieten eine einzigartige Gelegenheit, ihre Atmosphären mit hoher Präzision und breiter Wellenlängenabdeckung zu charakterisieren. Um diese Planeten beobachten zu können, müssen sie genügend Wärme von ihrer Entstehung bewahren, um durch ihre thermischen Emission beobachtbar zu sein. Dazu müssen spezielle, kontrastreiche Bildgebungsverfahren eingesetzt werden, um das schwache Signal des Begleiters von dem nahen Stern zu trennen, den er umkreist. Nur junge, riesige Exoplaneten sind jung und hell genug, um direkt beobachtet werden zu können. Ihre relative Jugend ermöglicht es uns, ihre heutige Zusammensetzung mit dem Mechanismus in Verbindung zu bringen, durch den sie entstanden sind.

In dieser Arbeit stelle ich die Instrumente und Methoden vor, die zur Beobachtung und Charakterisierung solcher direkt abgebildeten Exoplaneten verwendet werden, und wende sie auf das Referenzsystem HR 8799 an. Die Entwicklung des Analysemoduls `pet i tRADTRANS` ermöglicht die schnelle Anpassung von Modellen an spektroskopische Daten mit Hilfe von bayes'schen Verfahren. Wir verwenden das Analysemodul um das *JWST* Early Release Science Target WASP-39 b, einen heißen Exoplaneten mit der Masse des Saturn, zu untersuchen. WASP-39 b hat eine Atmosphäre, die reich an Metallen ist, und die Entdeckungen von CO_2 und photochemisch erzeugtem SO_2 werden bestätigt. Durch den Vergleich von drei verschiedenen Algorithmen zur Nachbearbeitung von Hochkontrastbildern werden die Auswirkungen der zur Extraktion der Spektren eines Exoplaneten verwendeten Datenverarbeitungstechniken auf die Fähigkeit, auf atmosphärische Eigenschaften zu schließen, untersucht. Es hat sich gezeigt, dass die Berücksichtigung der Korrelation zwischen den Wellenlängen der spektroskopischen Messungen entscheidend ist, um unverfälschte Parameterschätzungen zu erhalten. Nach der Entwicklung der Werkzeuggefü für die Datenanalyse und Modellierung werden die Atmosphären der vier Planeten von HR 8799 systematisch charakterisiert. Mit Hilfe neuer VLT/GRAVITY-Beobachtungen und einer breiten Palette von Archivdaten erstellen wir die bisher vollständigsten Spektren dieser Objekte. Die Planeten werden mit dem `pet i tRADTRANS` Analysemoduls sowie mit selbstkonsistenten radiativ-konvektiven Gleichgewichtsrastern angepasst. Die Eigenschaften der

effektiven Temperatur, der Oberflächenschwere, des Radius, der Metallizität, des Kohlenstoff-Sauerstoff-Verhältnisses, der Masse und der bolometrischen Leuchtkraft werden für alle vier Planeten abgeleitet. Ihre Atmosphären sind sehr metallreich und weisen ein stellares bis superstellares C/O-Verhältnis auf. Zukünftige Studien werden notwendig sein, um diese atmosphärischen Eigenschaften mit ihrer Entstehungsgeschichte in Verbindung zu bringen, aber es ist klar, dass diese rätselhaften Planeten in den kommenden Jahren ein Ziel für weitere Beobachtungen bleiben werden.

SUMMARY

As long as we have gazed at the night sky we have wondered about the nature of planets both within our Solar System and beyond. The last three decades of observations have revealed a staggering diversity of exoplanets in terms of their size, temperature, and composition. Out of the thousands of known exoplanets, only a small handful can be directly imaged. These rare systems present a unique opportunity to characterise their atmospheres with high precision and broad wavelength coverage. In order to observe these planets, they must retain enough heat from their formation to be observable in thermal emission, using dedicated high-contrast imaging techniques to separate the faint signal of the companion from the nearby star which it orbits. Only young, giant exoplanets are bright enough to be directly observable. Their relative youth enables us to tie their present-day composition back to the mechanism via which they formed.

In this thesis, I present the tools and methods used to observe and characterise such directly imaged exoplanets and apply them to the benchmark system HR 8799. The development of the `petitRADTRANS` retrieval module allows for the rapid fitting of models to spectroscopic data in a Bayesian framework. This retrieval method is applied to the *JWST* Early Release Science target WASP-39 b, a hot, Saturn mass exoplanet. WASP-39 b is found to have an atmosphere enriched in metals, and the detections of CO₂ and photochemically produced SO₂ are confirmed. By comparing three different high-contrast imaging post-processing algorithms, the impacts of data processing techniques used to extract an exoplanet's spectrum on one's ability to infer atmospheric properties are explored. Accounting for the correlation between wavelength channels of spectroscopic measurements is found to be critical to producing unbiased parameter estimates. Having developed the data analysis and modelling framework, the atmospheres of the four HR 8799 planets are systematically characterised. Using new VLTI/GRAVITY observations, together with a broad range of archival data we compile the most complete spectra of these objects to date. The planets are fit using the `petitRADTRANS` retrieval framework, as well as using self-consistent radiative-convective equilibrium grids. Bulk properties of effective temperature, surface gravity, radius, metallicity, carbon-to-oxygen number ratio, mass, and bolometric luminosity are inferred for all four planets. Their atmospheres are found to be highly metal-rich, with stellar-to-superstellar C/O ratios. Future study will be necessary to link these atmospheric properties to their formation history, but it is clear that these enigmatic planets will remain a target of further observations for years to come.

ACKNOWLEDGEMENTS

It's hard to believe that I'm looking back at four years at MPIA already. There are so many people who have made getting to this point possible. First, my supervisor Paul Mollière has my endless gratitude, not in the least for reading and providing feedback on this thesis. Your patience, wisdom, and scientific insight have shaped me into the astronomer I am today, and will provide a template for how I can mentor others tomorrow. I truly could not have asked for a better advisor, and I look forward to many more years of collaboration and friendship.

To my Heidelberg family: Nico, Gabi, Eric, Max, Molly, Rhys, Verena, Rodrigo, and so many more, you have made Heidelberg home for the last four years. From ski trips to dinners, movie nights to bike tours, your support and friendship have meant the world to me. Finishing a PhD is often compared to climbing a mountain, and thanks to you guys I now have a pretty good idea of how they compare!

Science wouldn't be worthwhile if it wasn't for the people I work with. Laura, you have been an inspiration. The APEX department couldn't be in better hands. Niall and Chronis, we've had far too many late nights writing proposals, but at least by now we'll start to have late nights analysing JWST data. We'll get DISCO through eventually! Faustine, you were my first friend in Heidelberg and a guide into the strange land of exoplanets. To Sylvester, Mathias, Jason, Tomas, Will and the rest of the ExoGRAVITY team, thank you all for letting me continue to say that 'the HR 8799 paper is almost ready' for four years. This is an amazing collaboration to be a part of, and I'm excited to see what we can achieve in the future. Lastly, thank you to my examination committee, Laura Kreidberg, Hans-Walter Rix, Saskia Hekker and Andre Butz, particularly to Laura and Hans-Walter for reading and refereeing my thesis.

I could not have made it this far without the unending support of my partner, Carla. I can't express enough how much I appreciate your love and your time. Vos brillás más que cualquier estrella.

Last, but certainly not least, my family. You have supported me for the last ten years since I left home, and not a day goes by that I don't miss you all. I'm proud to be your son, and will always be grateful for how you encouraged me to follow my dreams.

| | | |
|-------------------------|--------------------------------|-------------|
| THEMA | | vi |
| SUMMARY | | viii |
| ACKNOWLEDGEMENTS | | x |
| I INTRODUCTION | | 1 |
| 1 | Introduction | 2 |
| 2 | A brief overview of exoplanets | 3 |
| 2.1 | Detection Methods | 3 |
| 2.1.1 | Radial velocity | 3 |
| 2.1.2 | Transits | 4 |
| 2.1.3 | Direct imaging | 8 |
| 2.2 | Exoplanet demographics | 10 |
| 3 | High-Contrast Imaging | 13 |
| 3.1 | Instrumentation | 13 |
| 3.1.1 | Adaptive optics | 13 |
| 3.1.2 | Coronagraphs | 16 |
| 3.1.3 | Integral Field Spectrographs | 17 |
| 3.2 | Data processing | 18 |
| 3.2.1 | Reference differential imaging | 19 |
| 3.2.2 | Angular differential imaging | 19 |
| 3.2.3 | Spectral differential imaging | 20 |
| 3.2.4 | Modern processing algorithms | 20 |
| 3.3 | Optical interferometry | 21 |

| | | |
|-------|----------------------------------|----|
| 3.3.1 | Theory | 22 |
| 3.3.2 | VLT/GRAVITY design | 25 |
| 4 | Planet formation | 28 |
| 4.1 | Circumstellar disks | 28 |
| 4.2 | Formation mechanisms | 32 |
| 4.2.1 | Gravitational Instability | 32 |
| 4.2.2 | Core Accretion | 34 |
| 4.3 | Evolution | 37 |
| 5 | Atmospheric physics | 39 |
| 5.1 | Solar System observations | 39 |
| 5.2 | Brown dwarf observations | 43 |
| 5.3 | Structure, interior and dynamics | 47 |
| 5.4 | Chemistry | 50 |
| 5.4.1 | Disequilibrium processes | 51 |
| 5.5 | Emission spectra | 53 |
| 5.5.1 | Radiative transfer | 54 |
| 5.5.2 | Absorption and emission lines | 56 |
| 5.5.3 | Continuum gas opacities | 59 |
| 5.5.4 | Clouds and condensation | 62 |
| 5.5.5 | Summary | 65 |
| 5.6 | Self-consistent models | 66 |
| 5.7 | Global circulation models | 67 |

| | | |
|-----------------------------------------------------|------------------------------|-----------|
| 5.8 | The inverse problem | 68 |
| 6 | Outline | 70 |
| II ATMOSPHERIC RETRIEVALS WITH PETITRADTRANS | | 71 |
| 7 | petitRADTRANS | 72 |
| 7.1 | Radiative transfer | 72 |
| 7.2 | Correlated-k | 76 |
| 7.3 | Opacity sources | 79 |
| 7.4 | Thermal structure | 79 |
| 7.5 | Chemistry | 83 |
| 7.6 | Clouds | 84 |
| 7.6.1 | Adaptive mesh refinement | 86 |
| 8 | Retrievals | 87 |
| 8.1 | Likelihood functions | 89 |
| 8.2 | Nested sampling | 90 |
| 8.3 | The pRT retrieval module | 91 |
| 9 | Case Study: WASP-39 b | 97 |
| 9.1 | Background and previous work | 97 |
| 9.2 | Data | 98 |
| 9.3 | Retrieval setup | 99 |
| 9.4 | Free chemistry | 101 |
| 9.5 | Chemical disequilibrium | 105 |
| 9.6 | Impacts data selection | 107 |

| | | |
|---------------------------------------------------------------------------------|------------------------------------------------------|------------|
| 9.7 | Degenerate parameterisations | 107 |
| 9.8 | Cloud modelling choices | 109 |
| 9.9 | Summary | 110 |
| III IMPACTS OF HIGH-CONTRAST IMAGE PROCESSING ON ATMOSPHERIC RETRIEVALS. | | 113 |
| 10 | Introduction | 115 |
| 11 | Observations | 117 |
| 11.1 | Data preprocessing | 119 |
| 11.2 | Stellar model for flux extraction | 119 |
| 12 | ADI data processing | 123 |
| 12.1 | Post-processing algorithms | 124 |
| 12.2 | Spectral covariance estimation | 128 |
| 12.3 | Bias correction | 131 |
| 12.4 | Impact of covariance on retrieved parameters | 132 |
| 13 | Injection testing | 140 |
| 13.1 | Choice of goodness-of-fit metric | 141 |
| 13.2 | Optimising spectral extractions | 142 |
| 14 | Retrieval tests | 146 |
| 14.1 | Outline of retrievals | 152 |
| 14.2 | Impact of algorithm selection on retrievals | 153 |
| 14.3 | Impacts of including the covariance in retrievals | 156 |
| 14.4 | Relation between measurement and posterior precision | 160 |

| | | |
|-----------------------------------------------------------------------------------------------|---------------------------------------|------------|
| 14.5 | Mitigation using nuisance parameters | 161 |
| 14.6 | Limitations | 163 |
| 15 | Summary & conclusions | 164 |
| 16 | Appendices | 166 |
| 16.A | SPHERE Data Reduction | 166 |
| IV FOUR-OF-A-KIND? SYSTEMATIC CHARACTERISATION OF THE HR8799 PLANETS WITH VLTI/GRAVITY | | 169 |
| 17 | Introduction | 171 |
| 18 | The planetary system of HR 8799 | 175 |
| 18.1 | HR 8799 A: The host star | 176 |
| 18.2 | Photometric studies | 177 |
| 18.2.1 | Variability | 179 |
| 18.2.2 | Orbital dynamics | 180 |
| 18.3 | Spectroscopic characterisation | 180 |
| 18.4 | Retrieval studies | 184 |
| 18.5 | Self-consistent atmospheric modelling | 184 |
| 18.6 | Formation | 186 |
| 19 | Observations | 190 |
| 19.1 | GRAVITY data | 190 |
| 19.2 | Archival data | 192 |
| 19.2.1 | SPHERE | 192 |
| 19.2.2 | GPI | 193 |

| | | |
|--------|-----------------------------------|-----|
| 19.2.3 | CHARIS | 193 |
| 19.2.4 | ALES | 194 |
| 19.2.5 | OSIRIS | 194 |
| 20 | Atmospheric modelling | 196 |
| 20.1 | Thermal structure | 196 |
| 20.1.1 | Spline profile | 197 |
| 20.1.2 | Guillot profile | 198 |
| 20.1.3 | Mollière profile | 199 |
| 20.1.4 | Zhang profile | 199 |
| 20.2 | Chemistry | 200 |
| 20.2.1 | Free chemistry | 200 |
| 20.2.2 | Interpolated (dis)equilibrium | 201 |
| 20.3 | Clouds | 202 |
| 20.4 | Retrieval setup | 203 |
| 20.5 | Retrieval ranking | 203 |
| 20.6 | Bayesian model averaging | 205 |
| 20.7 | Self-consistent forward modelling | 206 |
| 20.7.1 | ATMO | 207 |
| 20.7.2 | Exo-REM | 207 |
| 20.7.3 | Sonora | 207 |
| 20.7.4 | petitCODE | 208 |
| 20.7.5 | Grid Fits | 208 |

| | | |
|--------------------------------|---------------------------------------------|------------|
| 21 | Results | 212 |
| 21.1 | HR 8799 b | 219 |
| 21.2 | HR 8799 c | 221 |
| 21.3 | HR 8799 d | 224 |
| 21.4 | HR 8799 e | 225 |
| 21.5 | Impacts of modelling choices | 227 |
| 21.5.1 | Thermal structure | 227 |
| 21.5.2 | Chemistry | 228 |
| 21.5.3 | Clouds | 234 |
| 22 | Discussion | 239 |
| 22.1 | Highly enriched atmospheres | 239 |
| 22.2 | Impacts of data selection | 242 |
| 22.3 | Formation | 242 |
| 22.4 | Predictions for JWST | 244 |
| 23 | Conclusions | 248 |
| 24 | Appendices | 250 |
| 24.A | Data logs | 250 |
| 24.B | Reprocessing of the SPHERE and GPI datasets | 250 |
| 24.C | Retrieval validation | 253 |
| 24.D | Using the Hansen distribution with EDDYSED | 254 |
| 24.E | Complete retrieval results | 258 |
| V SUMMARY & OUTLOOK | | 263 |

| | | |
|--------|----------------------------------------------|-----|
| 25 | Summary | 264 |
| 25.1 | Retrieval development | 264 |
| 25.2 | Importance of data processing | 264 |
| 25.3 | New insight into the HR 8799 planets | 265 |
| 25.3.1 | Formation | 265 |
| 26 | Outlook | 268 |
| 26.1 | Future pRT development | 268 |
| 26.2 | Direct imaging with JWST | 269 |
| 26.2.1 | The Direct Imaging Spectral Community Survey | 269 |
| 26.3 | Closing | 271 |
| | List of Figures | 274 |
| | List of Tables | 277 |
| | Personal works | 278 |
| | Bibliography | 279 |



INTRODUCTION

“There is a single general space, a single vast immensity which we may freely call Void; in it are innumerable globes like this on which we live and grow. This space we declare to be infinite; in it are an infinity of worlds of the same kind as our own. For there is no reason nor defect of nature’s gifts, either of active or passive power to hinder the existence of other worlds through space, which is identical in natural character with our own space.”

Giordano Bruno, 1584



1. INTRODUCTION

The study of planetary systems goes back nearly as far as recorded history. Most of this time was spent studying a small subset of all planetary systems, also known as the Solar System planets. After the Copernican revolution of the 16th century made our small place in the vast cosmos apparent, we began to wonder about the worlds beyond our own. Is our Solar System unique? Is there other life in the universe?

Planetary science took off in earnest with the advent of the telescope, allowing for the observation of moons, rings, and atmospheric features. The complexity of our neighbours soon became apparent, from the endless storms of Jupiter to the scarred landscape of Mars. It was soon realised that the atmospheres of the Solar System were entirely unlike our own. The rocky planets span from the thin Martian atmosphere, barely clinging onto its dusty surface to the dense, acidic furnace of Venus. The diversity of these atmospheres, even for seemingly similar planets, is extraordinary, and so astronomers have spent the last two centuries measuring the composition, dynamics, and weather of our nearest neighbours.

For the last thirty years, we have been able to peer farther than ever before, and have begun the exploration of extrasolar worlds. As telescopes have grown in size and sensitivity, we can begin to make the same measurements of these distant planets as we made of Mars and Jupiter over the past century. From the first tentative signs of the presence of an atmosphere to the detailed characterisation performed today, our understanding of these atmospheres has developed rapidly. We are just now able to measure some exoplanet properties with similar precision to that of Solar System measurements. The more we unveil about these atmospheres, the more diversity we find, and it will be the task of the next years and decades to truly answer those early questions about our place among the stars.

In this thesis, I will continue the line of study that has captured our collective imaginations for millennia. This introduction will provide the reader with the necessary background and context for this work. In particular, I highlight the instruments and techniques used to obtain the data, the origins of the planets under investigation, and the links between observations and atmospheric models. The rest of this work details the development of novel techniques for atmospheric characterisation, the careful validation of assumptions often made about our data, and the application of both to a benchmark system of four directly imaged exoplanets: HR 8799.

2. A BRIEF OVERVIEW OF EXOPLANETS

The field of exoplanet research has seen an explosion of activity over the past thirty years. [Campbell et al. \(1988\)](#) claimed the first detection of an exoplanet using the radial velocity measurement technique, though the signal-to-noise ratio (S/N) of their observations was too low to confirm the detection. A few years later, [Wolszczan & Frail \(1992\)](#) found three super-earth exoplanets orbiting a pulsar by observing variations in the pulsar timing. Finally, it was in 1995 that [Mayor & Queloz \(1995\)](#) detected and confirmed 51 Peg b, the first exoplanet orbiting a main-sequence star, thus ushering in the era of exoplanet research.

Since then, over 5600 exoplanets have been detected through various means. The field has evolved in numerous directions, trying to answer questions about the formation history of these objects, their atmospheric composition and dynamics, and the population level trends. In the long term, the questions of habitability remain open, and we continue to wonder whether or not we are alone in the cosmos.

2.1. Detection Methods

Due to their small size and relatively cool temperatures, exoplanets are challenging to observe. Nevertheless, a suite of techniques for their detection and characterisation have been developed, the most widely-used of which I summarise here. Additional techniques, such as the use of gravitational microlensing to identify widely separated companions or high-resolution spectroscopy to characterise the atmospheres of hot Jupiters, have also provided significant insight into the exoplanet population but are not directly relevant to this work. These and other methods have been thoroughly reviewed in the literature, such as in reviews by [Wright & Gaudi \(2013\)](#), [Fischer et al. \(2014\)](#) or [Kaushik et al. \(2024\)](#).

2.1.1. Radial velocity

The radial velocity technique relies on the Doppler shift of the stellar spectrum induced by the gravitational pull of the planet on the host star. As the planet orbits the host star, the star itself is pulled by the gravitational attraction of the planet. This reflex motion causes a Doppler shift in the light emitted by the host star, red-shifting the light as the star moves away from the observer, and blue-shifting as it moves towards the observer. The period of the signal is thus the same as the period of the planet, with a signal strength proportional to the gravitational force exerted on the star by the planet. This is the same Doppler

technique used to characterise binary star systems. When applied to exoplanets, the key challenge is achieving the precision necessary to detect the small signals induced by low-mass exoplanets on the much more massive host star. Indeed, we see this bias reflected in the planets detected through this technique, where the easiest targets are high-mass planets on short period orbits around low-mass host stars.

We can quantify this picture. Consider a planet of mass M_p orbiting around a star of mass M_* with period T . We must also account for the orbital eccentricity e , which changes the separation of the planet and star, as well as the inclination angle between the orbital plane and the observer plane, i . The RV semi-amplitude K can be calculated as

$$K = \left(\frac{2\pi G}{T} \right)^{1/3} \frac{M_p \sin i}{(M_p + M_*)^{2/3}} \frac{1}{(1 - e^2)^{1/2}}, \quad (2.1)$$

where G is Newton's gravitational constant. Due to the inclination angle, it is only possible to measure the mass up to a factor of $\sin i$.

Early RV instruments achieved precisions of 10 m s^{-1} or so – enough to detect hot Jupiter type planets. Advances in instrumentation have improved the precision of RV instruments by orders of magnitude, with current generation instruments aiming for precisions approaching 1 cm s^{-1} at which point the dominant noise source is typically stellar activity. Nevertheless, such precision is required to identify earth-like exoplanets.

2.1.2. Transits

While the RV method enabled the first exoplanet discoveries, it has been the transit method, particularly with the *Kepler* and *TESS* missions, that has led to the detection of thousands of new planetary systems. Transmission spectroscopy, led by the use of the Hubble Space Telescope (*HST*) and James Webb Space Telescope (*JWST*), has become the standard method for atmospheric characterisation. An exoplanet transit occurs when the planet passes directly between the host star and the observer, thus blocking a fraction of the host star's light. As with the RV technique, larger planets on shorter period orbits are easier to detect with this method: they obscure more of the stellar light, transit more frequently, and are more likely to have a favourable geometry.

To first order, the transit depth $\Delta F/F$ is proportional to the area of the star covered by the area of the planet:

$$\frac{\Delta F}{F} \propto \left(\frac{R_p}{R_*} \right)^2, \quad (2.2)$$

where F is the flux emitted by the star and ΔF is the change in the flux due to the planet transiting the star. and R_* and R_{pl} are the stellar and planetary radii respectively. Thus to measure the planet radius, in principle all that needs to be done is to measure the brightness of the star in and out of transit. Needless to say, many other considerations impact this simple model. Stars are not uniform disks: limb darkening, spots, flares, and stellar variability can all impact the shape of the light curve during transit. Planets can cross the star at an oblique angle or only partially transit the host star. Transits are chromatic, and the depth depends on the wavelength of light due to absorption in the atmosphere. This chromaticity is critical for enabling transmission spectroscopy.

All of these additional complexities still do not outweigh the power of transit measurements. With sufficiently precise data and careful modelling, most of this complexity can be mitigated. This leaves obtaining precise measurements as the primary challenge. Transit observations generally fall into one of two categories: a survey observation or a characterisation observation. Surveys are designed to observe a given sample of stars and identify any transit events. This builds up the statistics necessary to understand exoplanet demographics and identify promising candidates for follow-up characterisation. These surveys can be ground-based, such as the HAT (Bakos et al., 2002), WASP (Pollacco et al., 2006) and KELT (Pepper et al., 2007) surveys, but the majority of transit discoveries have come from space-based survey missions, in particular from *Kepler* and *TESS*. *Kepler* was designed to observe a single field of around 10^5 stars for a period of several years, with sufficient precision to identify planets of $1 R_{\oplus}$ or smaller (Borucki et al., 2010). *TESS* took a different approach, observing nearly the full sky at a regular cadence to detect transiting planets around nearby, bright stars (Ricker et al., 2014; Stassun et al., 2018). Combined, these two missions have discovered over 3100 confirmed exoplanets.

Characterisation observations generally use larger telescopes and are focused on obtaining high-precision measurements. The goal of these observations is typically to determine the chemistry and thermal structure of the target's atmosphere using transmission spectroscopy. Even without spectroscopy, transit observations in combination with RV measurements provide critical insight into the composition and structure of exoplanets. With precise radii from the transit observations and masses derived from radial velocity measurements, the density of the planet can be inferred (e.g., Baines et al., 2007; Valencia et al., 2007; Agol et al., 2021). Rocky planets have densities similar to that of Earth at around 5 g cm^{-3} , while hydrogen-dominated gas planets have densities more comparable to that of Jupiter or Saturn, the latter of which has a density $< 1 \text{ g cm}^{-3}$. Intriguingly, many exoplanets lie comfortably between these two bounds, which can be explained by large water fractions or by large hydrogen

envelopes around a rocky core (Otegi et al., 2020; Mousavi-Sadr et al., 2023). At the other end of the spectrum, some exoplanets have densities approaching 10 g cm^{-3} , well above that of pure iron. While the measurement uncertainties are large, it is clear that the diversity of exoplanet compositions is larger still than what we find in our own Solar System.

Transmission spectroscopy

The same geometry that allows us to detect transiting exoplanets makes them favourable for spectroscopic characterisation. While before we assumed that the planet was a simple disk obscuring the host star, in reality, the atmosphere of the planet appears as a thin limb around the disk of the planet's interior. The atmosphere may be optically thin or thick at different wavelengths, depending on the composition, impact of aerosols, and pressure. It is typically only a small fraction of the planet radius, and thus the signal due to the atmosphere is even smaller than the transit signal itself, requiring relative photometric precision on the order of 100 ppm or less.

We can calculate the typical signal strength of the transmission spectra, a technique pioneered in Seager & Sasselov (2000). Consider an atmosphere in hydrostatic equilibrium and obeying the ideal gas law, with a scale height H ,

$$H = \frac{k_b T}{\mu g}. \quad (2.3)$$

This scale height depends on the thermal energy $k_b T$, where k_b and T are the Boltzmann constant and temperature, the surface gravity g , and the composition, which determines the mean molecular weight μ . We can consider the impact that the addition of n scale heights to the planet radius has on the transit depth.

$$\frac{\Delta I}{I} \approx \frac{(R_p + nH)^2}{R_*^2} - \frac{R_p^2}{R_*^2}, \quad (2.4)$$

$$\approx \frac{2nR_p H}{R_*^2}. \quad (2.5)$$

As we are concerned with measuring the absorption spectrum, we must determine how the scale height H varies as a function of wavelength λ . Fortney (2005) shows that the optical depth τ in the line-of-sight geometry of a transiting planet can be computed as a function of wavelength and altitude z :

$$\tau(\lambda, z) = \sigma(\lambda)n(z) \sqrt{2\pi R_{\text{pl}} H}, \quad (2.6)$$

for absorber cross section $\sigma(\lambda)$, scale height H , and $n(z) = n_0 \exp(-z/H)$ is the volume density at altitude z of the primary absorber. We can define a value of τ_{eq} where a solid disk of radius $R_{\text{pl}} + z_{\text{eq}}$ is the same as that of the absorption radius of the planet with its translucent atmosphere. The authors find this boundary occurs at $\tau_{\text{eq}} \approx 0.56$. We can then invert equation 2.6 to find the atmospheric height $z(\lambda)$, which combined with the planet radius yields the planetary radius as a function of wavelength:

$$z(\lambda) = H \log \left(\frac{X_{\text{abs}} \sigma_{\text{abs}}(\lambda) P_{\text{ref}}}{\tau_{\text{eq}}} \sqrt{\frac{2\pi R_{\text{pl}}}{k_b T \mu g}} \right). \quad (2.7)$$

Here X_{abs} is the volume mixing ratio of the absorber with cross section $\sigma_{\text{abs}}(\lambda)$, P_{ref} is the pressure where $z = 0$. As before k_b and T are the Boltzmann constant and temperature, g is the surface gravity, and μ is the mean molecular weight. This shows that although the cross section is in the logarithmic term, its orders-of-magnitude variation with wavelength leads to contributions to $z(\lambda)$ on the order of the scale height.

Transmission spectroscopy has opened the doors to understanding the composition and dynamics of these worlds. Charbonneau et al. (2002) obtained the first measurement of an exoplanet atmosphere, detecting the sodium doublet at 589.3 nm in the atmosphere of HD 209458 b. Since then HD 209458 b has served as a benchmark for hot Jupiter atmospheres; it was among the first exoplanets observed in thermal emission with *Spitzer* as it passed behind its host star (Deming et al., 2005), and to have been observed using broad wavelength infrared spectroscopy (Richardson et al., 2007). Sing et al. (2016) placed HD 209458 b in context with other hot Jupiters, compiling *HST* and *Spitzer* measurements from the UV to the infrared to form a spectral sequence. They identified systematic trends in the depth of the water absorption feature as a function of effective temperature, forming a continuum from hot, clear atmospheres to cooler cloudier atmospheres. This work was expanded by Mansfield et al. (2021), who produced a similar spectral sequence for hot Jupiters in thermal emission, finding evidence for compositional diversity as well as trends in atmospheric opacity sources as a function of temperature. A new era of transmission spectroscopy began with the observation of WASP-96 b with *JWST*, an Early Release Observation of the telescope (Radica et al., 2023). With unprecedented precision and wavelength coverage, *JWST* opens new windows into the atmospheres of hot Jupiters and enables the characterisation of worlds that were previously inaccessible.

While ‘easy’ to study due to their short periods and large atmospheric scale heights, hot Jupiters are not the only class of planets studied in transmission.

From ultra-hot Jupiters on extremely short orbits, such as WASP-121 b (Delrez et al., 2016; Mikal-Evans et al., 2023) to cool and cloudy sub-Neptunes (Kreidberg et al., 2014) to rocky planets (Lim et al., 2023), each class of exoplanet provides novel insight into the diversity of potential atmospheres.

2.1.3. Direct imaging

While transit and radial velocity measurements rely on the influence of a planet on its star to infer planet properties, direct imaging allows for the collection of photons from the planet itself. This is obviously a challenging prospect: planets are intrinsically small and dim and are located near very bright stars. Thus the challenge of direct imaging is to reduce the glare of the host star to be able to view light emitted from the planet. To date, it has only been possible to observe thermal emission from young planets that have retained some of the heat from their formation. This requirement necessarily implies that the directly imaged planet population is young, as after a few hundred million years most planets have cooled off too much to be detectable. Upcoming surveys with *JWST* may allow for the detection of older, colder targets (Carter et al., 2024).

The first direct detection was made in 2005, when Chauvin et al. (2005) detected the companion 2M 1207 b (also known as TWA 27 b). Orbiting a brown dwarf, 2M 1207 b is a $\sim 5 M_{\text{Jup}}$ planet on a 55 au orbit. This initial discovery already demonstrated the difficulty in drawing the line between brown dwarfs and giant exoplanets. They are spectrally similar, with similar masses and compositions. Nevertheless, for the purposes of this thesis, I will call any object under the $13 M_{\text{Jup}}$ deuterium burning threshold which orbits a central object that undergoes some fusion process a planet.

Following the discovery of 2M 1207 b, roughly 20 systems of directly imaged planets have been identified. These systems vary greatly in the number of planets per system, the planet-star separations, and the exoplanet compositions. Of the directly imaged planets, a few systems stand out as benchmarks. The HR 8799 system hosts four super Jupiters orbiting an AV/F0 host star. It was first discovered by Marois et al. (2008), and will be covered extensively in Chapter IV. The β Pictoris system has been an object of interest for several decades, as it was the first system to have a spatially resolved circumstellar disk (Smith & Terrile, 1984). With improvements to instrumentation, Lagrange et al. (2010) identified one giant planet in the system, and later identified a second inner companion by coupling radial velocity, astrometry, and direct interferometric measurements (Lagrange et al., 2020). These two systems span typical direct imaging properties. The coldest, outermost planet of the HR 8799 system, HR

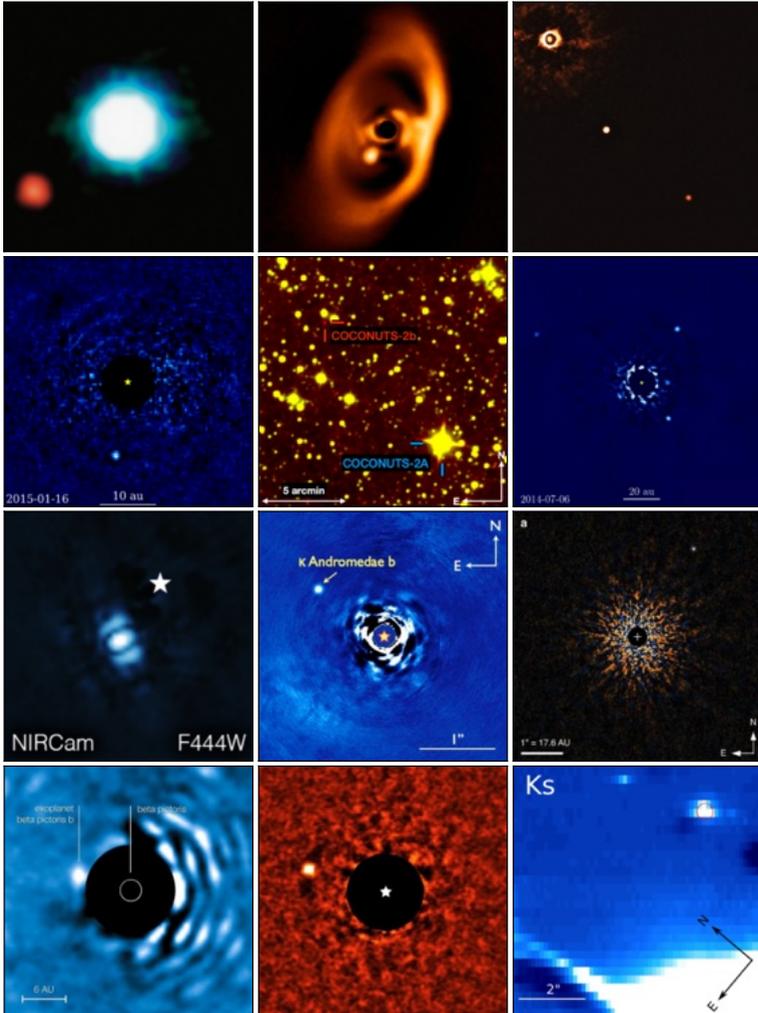


Figure 2.1: A family portrait of some of the directly imaged exoplanets. In order from left to right, top to bottom: 2M 1207 (Chauvin et al., 2005), PDS 70 b, c (Keppler et al., 2018), YSES 1 b, c (Bohn et al., 2020), 51 Eridani b (Macintosh et al., 2015), COCONUTS-2 b (Zhang et al., 2021), HR 8799 b, c, d, e (Marois et al., 2010), HD 95086 b, (Rameau et al., 2013), HIP 65426 b (Carter et al., 2023), κ Andromedae b (Bonnefoy et al., 2014a), GJ 504 b (Kuzuhara et al., 2013), β Pictoris b (Quanz et al., 2010), AF Leporis b (Mesa et al., 2023), and HD 106906 b (Daemgen et al., 2017).

8799 b, has a temperature of around 950 K, though it is somewhat different in its spectral shape compared to field brown dwarfs of the same temperature. At the other end of the spectrum, β Pictoris b has a mass of nearly $10 M_{\text{Jup}}$ and a temperature of about 1600 K, and is a close analogue of mid-L type brown dwarfs.

Even with only a small handful of targets compared to the abundance of transiting exoplanets, direct imaging has proven to be a rich opportunity to explore exoplanet atmospheres and planet demographics. For recent reviews of the subject, see [Currie et al. \(2023a\)](#), [Chauvin \(2024\)](#) and [Zurlo \(2024\)](#).

2.2. Exoplanet demographics

The most obvious and fundamental question in exoplanet demographics is ‘How many exoplanets are there?’ This deceptively simple question requires some careful thought. [Gaudi et al. \(2021\)](#) and [Zhu & Dong \(2021\)](#) review this question, and exoplanet demographics in general with greater depth, but I highlight a few key features of the exoplanet population here.

Many stars are host to more than one planet, and the number and character of the planets within a system often seem to be correlated. We can rephrase our question in a few ways. First, we can ask ‘What is the average number of planets hosted by a star?’ This is simply the total number of planets N_p , corrected for observational biases, divided by the number of stars N_* .

$$\bar{N}_p = \frac{N_p}{N_*}. \quad (2.8)$$

Equally informatively, we can phrase the question as ‘What fraction of stars host at least one planet?’ In this phrasing, we ignore planet multiplicity and divide the number of planetary systems N_s by the number of stars,

$$\bar{N}_s = \frac{N_s}{N_*}. \quad (2.9)$$

It is still observationally challenging to determine both N_p and N_s . Consider the observational biases present in different detection methods. RV searches are most sensitive to close in, massive planets with low inclinations relative to the star-planet-observer plane. Thus we are insensitive to systems that contain small planets on wide orbits. Similar biases are present in transit surveys, due to the sensitivity to larger radii planets on short period orbits. Direct imaging and gravitational microlensing have nearly inverse biases: both are more sensitive to larger planets on wide period orbits. Direct imaging in particular is so far

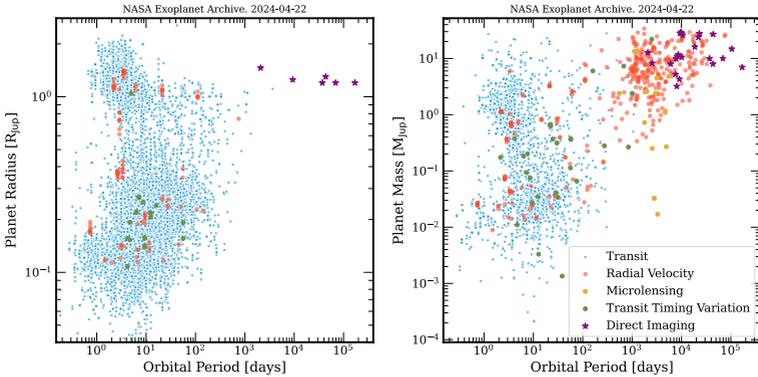


Figure 2.2: Left: Planet radius as a function of orbital period. Different coloured markers indicate the detection methods used for discovery. Right: Period-mass diagram. Data is taken from the NASA Exoplanet Archive, <https://exoplanetarchive.ipac.caltech.edu/>.

insensitive to planets smaller than $\sim 1 M_{\text{Jup}}$, and can only observe young planets, which may have different system architectures than older systems. Even interpreting N_* can be surprisingly difficult: many stars are likely unresolved binaries, which may not be apparent from photometric observations with small telescopes such as *Kepler* and *TESS*. Stellar binarity is also likely to influence the planet formation process, and will therefore produce different demographics than the single star population (Moe & Kratter, 2021). Finally, the choice of stars observed also introduces biases. Many surveys are designed to select low-activity FGK stars, which allow for easier detection of exoplanets, but may not be representative of the overall exoplanet population. Likewise, due to the relatively larger signal, transit observations are most sensitive for M-dwarf stars. While M-dwarfs are the most common stellar type, their low mass and high UV activity may result in different exoplanet populations than for higher mass stars.

With all of these caveats (and more) considered, it is still possible to estimate \bar{N}_p and \bar{N}_s for restricted regions of parameter space. In Figure 2.2, we present the current state of exoplanet discovery: over 5600 confirmed exoplanets as of writing. This figure shows the relationships between the planet radius and orbital period. Some of the observed features are physical, such as the radius gap between Neptune and Jupiter sized planets, as well as the hot-Neptune desert at short orbital periods. However, the dearth of planets at small radii and large orbital separations is largely due to the lack of sensitivity to that region of parameter space. Using similar data, Zhu & Dong (2021) estimate \bar{N}_p using a

Bayesian framework, finding that ‘ $\bar{N}_p = 1.23 \pm 0.06$ for planets with radii in the range of 1–20 R_{Jup} and orbital periods of up to 400 days.’ *Kepler* planets are often found in multiplanet systems, with planets in the same system behaving as ‘peas-in-a-pod’, sharing similar radii and clustered orbits (Weiss et al., 2018). The notable exception to this is hot Jupiters, which are more likely to reside in single planet systems.

Directly imaged planets occupy a relatively niche region of parameter space. With young ages and wide orbits, they are inherently distinct from the old, compact systems readily detected with transit surveys. Many efforts have been made over the years to observe both statistically representative samples, as well as targeted searches around stars more likely to host detectable planets. The deepest of these surveys were accomplished with extreme adaptive optics assisted high-contrast imagers situated on 8–10 m telescopes. In particular, the SPHERE infrared survey for exoplanets (SHINE) (Desidera et al., 2021; Vigan et al., 2021) and Gemini Planet Imager Exoplanet Survey (GPIES) (Nielsen et al., 2019) placed strong constraints on the occurrence rates of young, giant planets. When combined with RV surveys, they broadly estimate that ~20% of sun-like stars host giant exoplanets between 10–100 au, with higher mass stars more likely to host massive, widely separated companions. Currie et al. (2023a) provides a summary of direct imaging survey efforts, discussing how our understanding of this population has developed through improvements in instrumentation and survey design.

3. HIGH-CONTRAST IMAGING

The primary challenge of directly imaging exoplanets is not the faintness of the companion, but rather the overwhelmingly bright light of the host star. Thus high-contrast imaging (HCI) techniques are required to extract the companion signal from the glare of the host star. Even in thermal emission where the exoplanet spectral energy distribution (SED) is at a maximum, the exoplanet signal can be more than a factor of 10^4 fainter than that of the star. For an earth-like planet in reflected light, the contrast between the star and planet can be more than 10^{10} . This becomes increasingly challenging at small angular separations: the closer to the core of the stellar point-spread function (PSF), the more difficult it is to distinguish the two signals. Thus the field of HCI is largely dedicated to developing instrumentation and data processing techniques to minimise the host star signal while amplifying the companion signal.

3.1. Instrumentation

The second generation of planet finding instruments such as the Spectro-Polarimetric High-Contrast Exoplanet Research instrument (SPHERE) (Beuzit et al., 2008, 2019), the Gemini Planet Imager (GPI) (Larkin et al., 2014) and the Subaru Coronagraphic Extreme Adaptive Optics instrument (SCEAO) have enabled the regular detection and characterisation of young super Jupiter companions. These instruments were designed to reach contrasts of some $\sim 10^{-6}$ at 200 mas separation from the host star. To achieve this they are located at 8 m to 10 m observatories equipped with extreme adaptive optics systems and coronagraphs, and are operated to take advantage of differential imaging post-processing techniques. They often consist of both imaging cameras and integral field spectrographs (IFS), the latter of which enable the detailed characterisation of companions. Figure 3.1 shows the contrast curves for these and other similar instruments. These curves show for a given separation the maximum contrast between the host star and planet for which a 5σ detection could be made.

3.1.1. Adaptive optics

The first problem that must be solved when pushing to small angular separations is atmospheric turbulence. In seeing-limited conditions, the wavefront is highly distorted by the atmosphere. When time-averaged by a detector, the random motions of the PSF result in a smeared observation with a full width at half maximum (FWHM) on the order of $1''$. In the diffraction limit, a 10 m telescope has a resolution of about 30 mas. When trying to image a companion

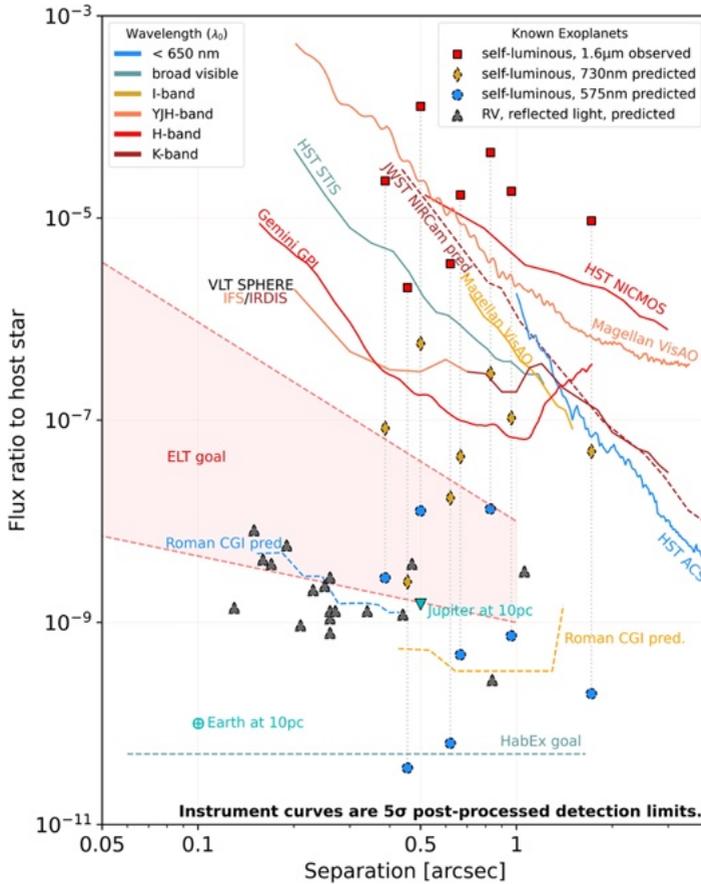


Figure 3.1: Contrast curves of current and future high contrast instruments. Each curve indicates the separation and contrast at which a 5σ detection can be obtained. Reproduced with permission from [Follette \(2023\)](#), see Figure 16 for details of how each contrast curve is computed.

at small angular separations, it is thus critical to apply adaptive optics corrections both to confine the light of the host star to a smaller angular region, but also that of the companion, ensuring that the faint signal is spread over a smaller detector region, which ensures it remains above the noise floor of the detector. Current extreme AO systems can routinely reach Strehl ratios (the ratio between the peak of the observed PSF and an ideal Airy pattern at the diffraction limit of the telescope) of $\sim 90\%$ in the H-band. To achieve this, several sub-systems are required: a wavefront sensor (WFS) which measures the aberrations due to

atmospheric turbulence, a real-time computer (RTC) which computes the appropriate corrections for the wavefront, and a deformable mirror (DM), which can change shape to correct for the distortions in the wavefront. Reviews of AO systems can be found in [Davies & Kasper \(2012\)](#) or [Milli et al. \(2016\)](#). In this section I will outline the fundamental operating principles and architecture of a modern extreme AO system.

The WFS measures the distortion of the wavefront by the atmosphere. This must be done at high spatial resolution across the FoV of the telescope, and at sufficient speed to correct for rapid changes in atmospheric turbulence. The speed and spatial resolution of WFS are limited by the amount of light available from a reference star (or laser guide star). As it is presently unfeasible to directly measure the phase of light in the near-IR or visible, phase variations must be computed from fluctuations in the intensity. Two technologies have emerged as the most common solutions this problem: the Shack-Hartmann WFS (SHWFS), and the Pyramid WFS. A more complete description of these and other wave front sensors can be found in [Rousset \(1999\)](#). The SFHWS uses a lenslet array in the pupil plane to sample the incident wavefront. For a planar wavefront, all of the projected images of a point source will be centered in the focal plane. If the wavefront at a given lenslet is shifted in phase, the position of the projected image will then be shifted in the focal plane. By measuring this displacement, a correction can be calculated by the RTC and sent to the DM. The ‘extreme’ label of extreme AO systems refers to the density of actuators available to manipulate the DM: the higher the actuator density, the higher the order of control available, although this comes at increasing the computational cost to determine the correction. The full control cycle occurs at a rate of around 1 kHz for corrections in the near-IR. For a pyramid WFS, a pyramidal prism is located at a focal plane in the optical path. Each face of the prism projects a unique pupil image through a series of lenses onto a detector. The flux measured in each pupil image can then be related back to the wavefront error. This method is more sensitive to low-order aberrations than the SHWFS, which is particularly important for coronagraphic imaging.

AO systems can allow for nearly diffraction limited imaging in the near IR, and as the technology develops the corrections will improve and can be applied to shorter wavelengths. However, they are only able to correct for distortions introduced prior to the WFS, and are not sensitive to non-common path aberrations throughout the remaining optical path. This leaves systematic residuals in the AO corrected images, commonly known as quasi-static speckles, or simply as speckles. These speckles vary slowly over time relative to the length of an observation, but are highly structured, thus complicating the detection of faint point sources in the image. To surpass this limitation, further post-

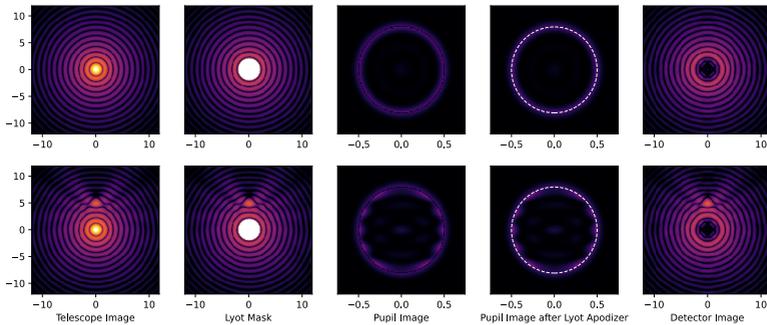


Figure 3.2: Example of the working principle of a Lyot coronagraph, made using `hcipy`. The top row shows a the optical path of a star passing through a Lyot coronagraphic setup, the bottom row also includes a companion with one tenth the brightness of the primary source. All images are on the same colour scale, with arbitrary units of intensity. From left to right we see the image as determined by the telescope PSF, in this case a simple circular aperture. A Lyot mask blocks out the central light of the PSF. The beam then passes through an apodizer at the pupil plane, which masks out residual starlight. Finally, on the right we see the image as projected onto a detector, with the core of the stellar PSF reduced in intensity, but allowing the companion light to pass through.

processing of the data is required.

3.1.2. Coronagraphs

Coronagraphs are optical elements used to block the on-axis light of the central star and transmit the off-axis light of a companion signal. Various designs have been introduced over the years, from the simple opaque focal plane mask of [Lyot \(1931\)](#) to modern apodizing phase plates optimised for a specific telescope pupil. [Galicher & Mazoyer \(2024\)](#) provides an overview of the theory and design of coronagraphic optics, which I will briefly explore here.

The most basic design for a coronagraph is that of Lyot ([Lyot, 1931](#)), an example of which I show in Figure 3.2. In this setup, the light from the telescope is focused onto an optical mask which blocks out the core of the PSF. Further on in the beam at the pupil plane an apodizing mask is inserted, further removing scattered light and increasing the contrast. The final image is then focused onto the detector plane, with the core of the PSF greatly reduced in intensity, but off axis light allowed to pass through unfiltered. This design requires very precise positioning of the stellar PSF onto the initial mask, and so it is highly

sensitive to the AO correction available. In addition, realistic telescope pupils do not produce clean circular Airy rings, and so the effectiveness of the design is further compromised. To mitigate these effects, there were several subsequent developments of focal plane coronagraphs (Zurlo, 2024).

Four quadrant phase-mask (4QPM). As the name implies, each quadrant of the focal plane receives a π phase shift, resulting in destructive interference of the host star, which is centred at the intersection of the four quadrants. Off axis light is simply phase shifted, but as it only passes through one quadrant is not destructively interfered. This type of coronagraph, together with a Lyot mask is available on *JWST*.

Vector vortex coronagraphs. Similar to the 4QPM, this coronagraph induces a phase shift at the focal plane. However, this design can induce a nearly continuous variation as a function of position angle, producing deeper contrast at small inner working angles (Mawet, 2007; Mawet et al., 2007).

Modern coronagraph designs are typically inserted in the pupil plane, modifying the phase of the wavefront to interfere the stellar signal but allow off axis light to pass through. These designs are further optimised for an individual telescope pupil geometry. By acting at the pupil plane, they are less sensitive to the precise positioning of the stellar PSF. These are *Apodizing Phase Plate* (APP) (Carlotti, 2013; Carlotti et al., 2014) or *Vector APP* (Doelman et al., 2021) designs, which are effective across a broad wavelength band and have high throughput, while being optimised for extreme contrast at small inner working angles. These are highly flexible designs that can produce nearly arbitrary dark regions around the central PSF, depending on the observing requirements. AO performance remains the limiting factor, but vector APP designs have been shown to reach raw contrasts of 10^{-4} (Doelman et al., 2021).

3.1.3. Integral Field Spectrographs

While imaging is the ideal tool for the detection of exoplanets in thermal emission, spectroscopic observations are necessary if we want to precisely characterise their atmospheres. While this can be achieved using traditional long slit spectroscopy (e.g., Snellen et al., 2014), most high contrast instruments are instead equipped with Integral Field Spectrographs (IFS). These instruments record spectra at different positions throughout the field of view, and can then be combined to produce a data cube with two spatial axis and one wavelength axis. While various optical designs can be used to achieve this, the most common design is the use of a lenslet array which samples the telescope image, and projects each image onto a slit for dispersal, thus providing a spectra at each lenslet location. The SPHERE instrument is one example of such an

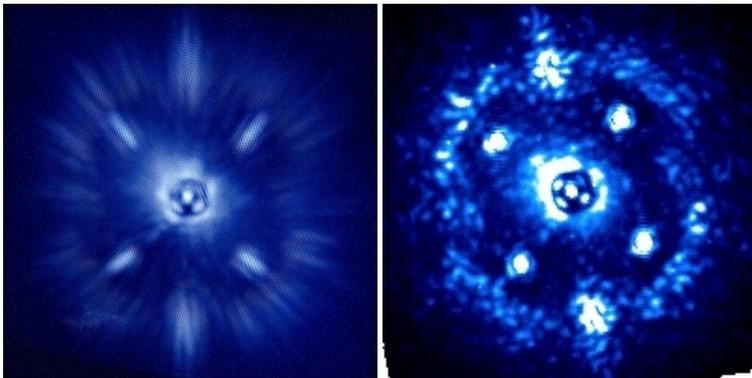


Figure 3.3: Left: detector image acquired by the SPHERE IFS. Each vertical line is a spatially separated, low resolution spectrum. The central star is obscured by a coronagraph, resulting in better sensitivity to companions at small separations. Right: a single frame of the resulting datacube, where each ‘spaxel’ (spatial pixel) would now have an additional axis of wavelength. The quasi-static speckles dominate the noise within the AO control ring. The four bright spots are known as satellite spots, and are used for centring the image.

instrument, and an example of the dispersed, spatially resolved spectra as well as a single frame of the resulting data cube can be seen in Figure 3.3. These instruments exist across a broad range of spectral resolutions and bandwidths, though most high-contrast instruments are designed to capture low-resolution ($R < 100$) spectra in the near-infrared. For additional contrast sensitivity, IFS instruments can be equipped with the same adaptive optics and coronagraphs as imaging instruments, resulting in excellent performance at small angular separations. The addition of spectral information can also aid in the detection of exoplanets through techniques such as molecular mapping (e.g. [Hoeijmakers et al., 2018](#); [Cugno et al., 2021](#); [Patapis et al., 2022](#)).

3.2. Data processing

Even with the use of AO and coronagraphic instruments, high-contrast data are typically still dominated by stellar speckles. These speckles vary slowly with time, and cannot simply be averaged out over the course of an observation. Therefore, more sophisticated methods are required for the removal of the remaining stellar PSF. These methods fall under the broad category of *differential imaging*, and make use of some varying data component to distinguish between

the stellar and planetary signals.

3.2.1. Reference differential imaging

Reference differential imaging (RDI) utilises the stability of a telescope PSF over time. In an ideal case, the stellar PSF can be measured by observing a reference star of similar brightness and subtracting the reference from the science target, leaving only the companion signal and residual noise. Unfortunately, the PSFs are rarely similar enough for this to function in such a straightforward fashion. Changes from imperfect AO correction, thermal effects, and variations in the optical path all lead to changes in the PSF both during the course of an observation and between observations. These can be partially mitigated by building a library of reference PSFs, building a model through the use of principal component analysis or similar techniques, and fitting this model to the science data for subtraction. Nevertheless, RDI typically remains in the domain of space telescopes, which have no atmosphere to distort the incoming wavefront and can be extremely thermally stable over time.

A recent, notable exception to this is SPHERE's so-called 'star-hopping' mode, pioneered by [Wahhaj et al. \(2021\)](#). When observations are taken in this mode, the telescope rapidly nods between the science target and a nearby reference target, leaving the adaptive optics control loop closed following an initial acquisition. This results in a series of science and reference images obtained nearly simultaneously, ensuring similar conditions in both the atmosphere and optical path. Compared to other image processing methods, these observations obtained a contrast sensitivity 2 magnitudes deeper between 100 mas and 300 mas.

3.2.2. Angular differential imaging

The most common HCI post-processing technique is angular differential imaging (ADI), first introduced by [Marois et al. \(2006\)](#). In this technique, the derotator used to keep an observed field stable on a detector as the earth rotates is disconnected. This allows the field to rotate as the telescope tracks the target over the course of an observation. By keeping the host star fixed at the centre of the field, the PSF remains stable while a companion rotates around the centre of the frame. Thus in every frame, the planet will appear in a different position. The frames can be median combined, producing an average PSF, with only a negligible contribution from the companion at any position. This average PSF is then subtracted from every frame, which can then be derotated and added, resulting in an averaging of the residual noise and the stacking of the companion signal.

3.2.3. Spectral differential imaging

Spectral differential imaging (SDI) (Racine et al., 1999) functions similarly to ADI. The PSF size depends on the wavelength of light. Thus as the wavelength increases, the speckle pattern expands radially outward, while the companion remains fixed in place. Equivalently, each frame can be rescaled such that the speckle pattern remains in a fixed position, while the companion shifts position throughout the frames. As with ADI, a median PSF image can then be determined and subtracted from each frame, leaving residual noise and the companion signal, which can be added to boost the signal. Recently Ruffio et al. (2023) extended this approach to jointly fit the stellar signal and companion signal directly to *JWST/NIRSpec* detector images rather than to extracted 1D spectra, finding that this approach dramatically improves the contrast sensitivity, particularly at small separations.

3.2.4. Modern processing algorithms

All of the previously mentioned data processing approaches have been dramatically improved over time. Using IFS data, ADI and SDI approaches are often combined, while ADI+RDI approaches have seen some use in processing data taken using *JWST* (Carter et al., 2023). Broadly speaking, most processing algorithms now fall under one of three categories.

Speckle subtraction These approaches inherit from early ADI work. They typically rely on principal component analysis (PCA) or a similar approach to learn the features of the stellar PSF, and use these features to produce a model which can be subtracted from the science images. This model can be built using the science frames themselves, or from a library of reference PSFs. The references can include both spatial variation as with ADI imaging, as well as spectral variation as with SDI. LOCI (Lafrenière et al., 2007; Maire et al., 2012; Marois et al., 2014), PCA (Amara & Quanz, 2012) and KL IP (Soummer et al., 2012; Pueyo, 2016) are all examples of this approach.

Inverse methods This approach uses a likelihood minimisation algorithm to fit a model of the background and companion PSF to the data. Ultimately, this produces a maximum likelihood estimate of the contrast and uncertainty at every location in the image. Some algorithms, such as ANDROMEDA (Cantalloube et al., 2015) first perform PSF subtraction before fitting the model, others such as PACO (Flasseur et al., 2018, 2020)

jointly fit a noise model and the planet signal.

Machine learning The last approach uses supervised machine learning to infer the presence of a companion. This requires a large training dataset, and typically perform inference on small patches of pixels throughout the field of view to build a detection map for the entire image. So far such approaches show promise for the detection of exoplanets, but have not yet been used to measure the companion flux, which is a necessary component for further characterisation.

[Cantalloube et al. \(2020a\)](#) performed a systematic comparison of these different algorithms to compare their ability to detect unseen companions. I will explore the use of such algorithms in Chapter III, focusing instead on how they impact the characterisation of exoplanets via their systematic effects on the extracted spectra.

3.3. Optical interferometry

While high contrast imaging allows us to observe exoplanets at 100 mas or so from their host star, the angular resolution available is still fundamentally limited by the telescope diameter. Optical interferometry provides an avenue to increase the sensitivity to exoplanets at much smaller separations, as the angular resolution is determined by the longest baseline available in the interferometer. [Michelson & Pease \(1921\)](#) was the first to use interferometry to measure the radius of a star, but it wasn't for another 50 years that the technology was suitably mature for [Labeyrie \(1975\)](#) to use two telescopes with sufficient stability to measure the angular size of Vega. Since these early observations, various facilities have been built to regularly perform optical and infrared interferometry. For example, the Very Large Telescope Interferometer (VLTI) has a maximum baseline of 134 m when using the 8 m Unit Telescopes (UTs), resulting in an angular resolution of around 5 mas in the K-band, as compared to 50 mas for any one of the 8 m telescopes in the diffraction limit. For this thesis, I make use of spectra obtained using VLTI/GRAVITY instrument, taking advantage of the high spatial resolution.. As it is impractical to build a monolithic telescope of equivalent diameter, this remains the best option for high angular resolution observations for the foreseeable future.

A more thorough treatment of fundamental interferometric theory can be found in [Haniff \(2007\)](#), upon whose work we will build our theoretical framework. [Roddier & Lena \(1984\)](#) also introduce considerations of the atmosphere

and the impact of practical optical systems on the production and measurement of interferometric fringes on a detector. In this section I introduce fundamental interferometric theory, as well as the details of the GRAVITY instrumental design.

3.3.1. Theory

An interferometer functions by passing the light emitted by some source through a pair of apertures, resulting in an fringe pattern caused by the constructive and destructive interference of the light waves. This pattern is sensitive to both the geometry of the interferometer and that of the emission source. In the most simple case, we can consider the double slit experiment of [Young \(1804\)](#), where a point source a emits monochromatic, plane-parallel radiation along the z axis, and passes through a screen with two pinhole apertures separated by a baseline B . The fringe pattern caused by summing the spherical waves emitted from the pinholes is then measured on a screen located at a distance z_1 from the aperture plane. This geometry is outlined in Figure 3.4. The amplitude of the electromagnetic wave \vec{E} at point x as a function of the diffraction angle α can be computed from the sum of the spherical waves emitted from the pinholes

$$\vec{E}(\alpha) = \frac{E_0}{r_1} e^{ikr_1} = \frac{E_0}{r_2} e^{ikr_2}, \quad (3.1)$$

$$= \frac{2E_0}{z_1} e^{ik(r_1+r_2)/2} \cos(k(r_1 - r_2)/2). \quad (3.2)$$

The distance from the i^{th} pinhole to point x is r_i , where $r_1 \approx r_2 \approx z_1$, and the *optical path difference* (OPD) is defined as $(r_1 - r_2)$, which for small angles is approximately αB . The intensity is the square of this amplitude,

$$I = |E(\alpha)|^2. \quad (3.3)$$

We can now define the *visibility* \mathcal{V} as the contrast between the peaks and troughs of the fringe pattern

$$\mathcal{V} = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}}. \quad (3.4)$$

The visibility is one of the key observables of interferometry, and is sensitive to the geometry of the emission source. The second key observable is the *phase* of the fringes, that is the position of the central fringe with respect to the zero-optical-path-difference location.

This formalism can be extended to a more realistic interferometric setup, accounting for the position of the emission source relative to the apertures, the

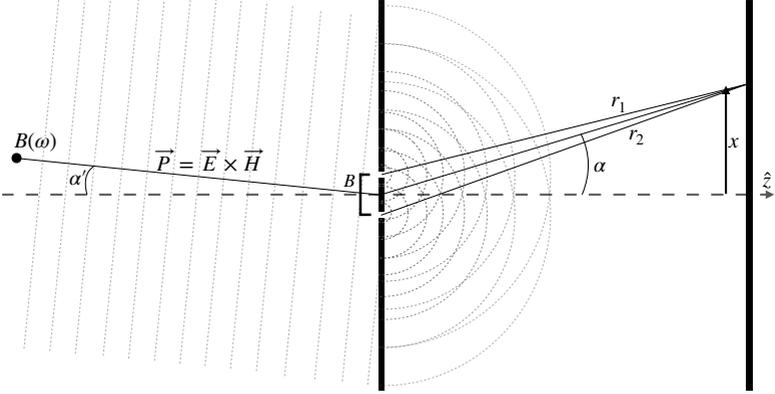


Figure 3.4: The geometric setup of the double slit experiment. The interference of the waves induced by the linear combination of the emission from the two slits results in a characteristic fringe pattern. The amplitude of the fringe at position x is determined by the optical path difference ($r_1 - r_2$), which can be determined from the baseline B and angle α .

geometry, coherence and bandwidth of the emission source, and the physical size and shape of the sub apertures. To address these, we adopt the framework of temporal and spatial coherence functions. These functions allow us to reframe our observables in terms of the correlations of the electric field from the source. The temporal coherence function $V_t(\tau)$ measures how correlated the electric fields along the emitted wave are in time at a fixed location. In particular, we are interested in the time-average of this quantity, as we are integrating over time when taking a measurement

$$\langle E^*(\vec{r}_1, t_1) \times E(\vec{r}_2, t_2) \rangle = V(\vec{r}_1 - \vec{r}_2, t_1 - t_2) = V(\vec{0}, \tau). \quad (3.5)$$

Importantly, by the Weiner-Khinchin theorem, this coherence function is equal to the Fourier transform of the spectral energy distribution of the emission source, $B(\omega)$, where ω is the frequency of the emission. Therefore, measurement of the correlation of the electric field at a point over time provides us with a means to obtain the emission spectrum of an astronomical source via an inverse Fourier transform of $V_t(\tau)$

$$\frac{V_t(\tau)}{V_t(0)} = \frac{\int B(\omega) e^{-i\omega\tau} d\omega}{\int B(\omega) d\omega}. \quad (3.6)$$

This also allows us to define a coherence time, which determines the maxi-

imum time delay that can separate different measurements. The coherence time depends directly on the bandwidth of the observed spectrum $\Delta\omega$: a broad bandwidth measurement will have a shorter coherence time than a narrow bandwidth observation. The coherence time is thus defined as

$$t_{\text{ch}} \approx 2\pi/\Delta\omega. \quad (3.7)$$

The spatial coherence function measures the correlations at a fixed time across varying location.

$$\langle E^*(\vec{r}_1, t_1) \times E(\vec{r}_2, t_2) \rangle = V(\vec{r}_1 - \vec{r}_2, t_1 - t_2) = V(\vec{\rho}, 0). \quad (3.8)$$

Analogous to the temporal coherence function, the spatial coherence function is related to the sky brightness by its Fourier transform by the Cittert-Zernike theorem

$$\frac{V_r(\vec{\rho})}{V_r(0)} = \frac{\int I(\vec{\alpha}) \exp\left(-2\pi i \left(\frac{\vec{\alpha} \cdot \vec{\rho}}{\lambda}\right)\right) d\alpha}{\int I(\vec{\alpha}) d\alpha}. \quad (3.9)$$

If we consider on-sky coordinates α and β , as well as the reciprocal coordinates u and v in Fourier space which are sampled by the interferometer baselines, we can rewrite equation 3.9 in a more practical way

$$V_r(u, v) = \frac{\int \int I(\alpha, \beta) \exp(-2\pi i (u\alpha + v\beta)) d\alpha d\beta}{\int \int I(\alpha, \beta) d\alpha d\beta}. \quad (3.10)$$

Thus by thoroughly sampling the uv plane by appropriate choices of baseline separations and orientations, an inverse fourier transform of the measurements will reconstruct the source brightness distribution.

To summarise: via measurements of fringe visibility at different times and different locations, we can calculate the spatial and temporal correlations of the emitted radiation. Applying inverse Fourier transforms, we can determine the source structure and spectrum, which are the key quantities of interest for most astrophysical signals. In practice, we will also need to account for the impact of the atmosphere on the impinging wavefront, as well as the transfer function of the optical system used in the interferometer. These effects can be mitigated through careful instrumental design, discussed further in the next section, as well as by careful calibration using both internal sources and well-characterised astrophysical signals.

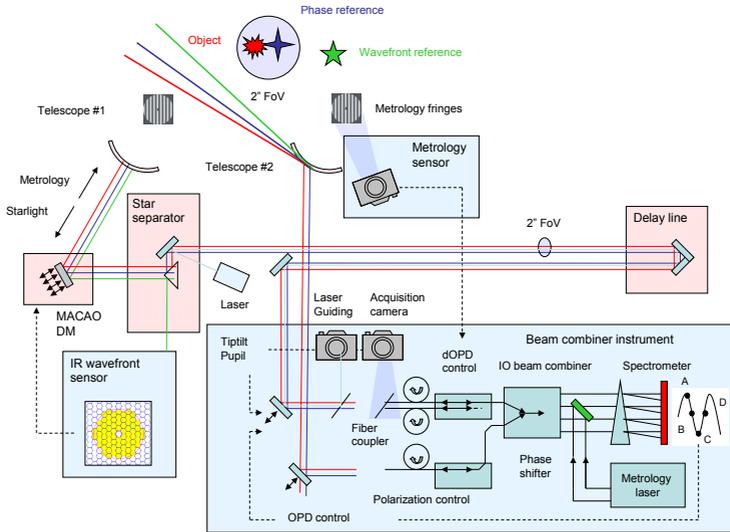


Figure 3.5: Schematic of a single baseline of VLTI/GRAVITY. THE VLTI components (shown in red boxes) are shared between the VLTI instruments, while the GRAVITY specific components are shown in blue boxes. Reproduced with permission from [Gravity Collaboration et al. \(2017\)](#).

3.3.2. VLTI/GRAVITY design

While the basic principles of interferometry are relatively straightforward, actually constructing an instrument sensitive enough to observe exoplanets is highly challenging. Modern interferometric systems have been enabled by developments in adaptive optics, allowing for diffraction-limited signals from large telescopes, real time control loops enabling the measurement and control of the fringe pattern, and low-noise infrared detectors with fast enough readout times such that the signal remains coherent. The optical and mechanical design of GRAVITY can be found in [Gravity Collaboration et al. \(2017\)](#), and we will summarise the key components of the instrument here. A schematic of the instrument is provided in Figure 3.5.

GRAVITY is a fibre-fed, K-band interferometric spectrograph that can coherently combine light from the four 8 m UTs or the four 1.8 m Auxiliary Telescopes (ATs) located at the VLT. It has the capability to operate in either single or dual field mode. In the latter mode, one of the fields is centred on a bright reference star in order to obtain a strong signal for wavefront sensing and

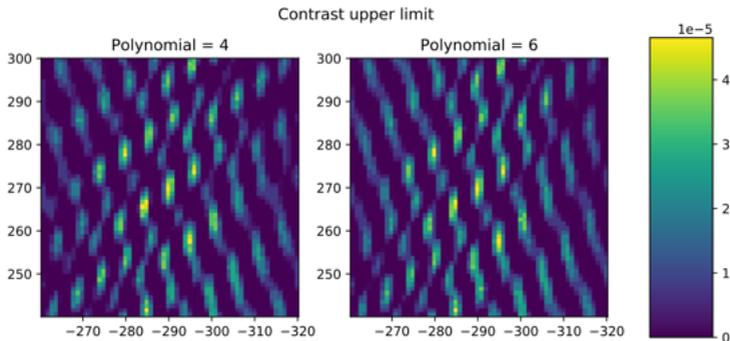


Figure 3.6: VLTI/GRAVITY science fringes for HR 8799 e, combining all 6 UT baselines.

fringe tracking, while the second field is centred on the faint science target for longer integrations. In single field mode the reference and science targets are the same.

VLTI

The main components of the VLTI are shared by the different interferometric instruments available at the facility. These consist primarily of the AO system and the delay lines. The AO system functions as in the case of high contrast imaging, splitting off shorter wavelength light from the reference star and passing it to a wavefront sensor. This sensor then measures the aberration of the wavefront on timescales of ms, and applies the appropriate corrections to a deformable mirror. With ideal correction, this allows for a near-diffraction limited image to be passed through the remainder of the instrument, and is particularly important for increasing the amount of light that is able to be passed into the single mode fibres. It also increases the coherence time, allowing for practical interferometric measurements.

Once the beam has been corrected by the AO system, it is passed through the delay lines which compensate for the OPD between the telescope mirrors. As the OPD must be identical to within the wavelength of interest, the delay lines can position the mirrors to 5 nm precision over a 180 m distance, and allows the OPD to remain stable over the course of an observation as it slowly varies due to the earth's rotation.

Beam combiner instrument

Following the delay lines the signal is passed to the beam combiner instrument. The goal of this instrument is to combine the light from the four telescopes, track and control the fringes of the reference star and thus integrate the science signal on the spectrograph. The fringe tracking correction is applied by tip-tilt-piston and lateral pupil control mirrors prior to coupling the beam to the optical fibers. The beam is then split, separating the reference phase and science star light, as well as separating the light between the fringe tracker and science spectrometer. Once split, the beams are coupled to the single mode fibres, which transport the light to the integrated optics system. This is an optical chip that introduces phase delays into each beam to sample the fringes, and combines each of the six baselines available from the four telescopes. The resulting outputs are then passed to the spectrometer where the light dispersed and spread over the detector. An example of the science fringes observed with VLTI/GRAVITY are included in Figure 3.6.

4. PLANET FORMATION

Like the stars they orbit, exoplanets are composed of material from the nebula in which they were born. Clearly planets both within our Solar System and without exhibit an incredible diversity of size, orbital configuration and composition, raising questions about the processes that led to such diversity. Is there a universal pathway to form planets? Do giant and terrestrial planets form in the same fashion? These - and many other questions - require careful study of both the circumstellar disks in which the planets form, as well as the processes of accretion and evolution that shape the final products of planet formation. In this section I will outline the environment in which planets are born, as well as the processes through which they form.

4.1. Circumstellar disks

Circumstellar disks, also known as protoplanetary disks, form during the collapse of the molecular gas cloud in the process of star formation. By conservation of angular momentum and internal friction, the cloud forms a rotating disk of gas and dust. The disks generally have masses of only a small fraction of their host star ($\sim 10\%$), and have radial extents from tens to hundreds of au. While they vary in structure and composition, they are composed primarily of molecular gas, with $\sim 1\%$ of the disk mass found in dust. Young Stellar Objects (YSOs) are classified based on their spectral energy distribution, as a proxy for their evolutionary state [Lada \(1987\)](#). Class 0 objects are prestellar cores, and appear as cold blackbodies. Over time the protostar collapses and heats up, becoming a Class 1 object with an optically thick envelope. As the envelope is expelled, the young star dominates the spectral energy distribution of the source, and the remaining gas collapses into the circumstellar disk. T Tauri stars with a circumstellar disks are the prototypes for this class of objects. These disks inherit material from the initial molecular cloud, but undergo significant processing during the disk lifetime, for instance from heating from the central star, adiabatic compression from the collapse of the gas cloud or from radiation from the local stellar environment.

Over the last decade, observations of planet-forming disks have shifted from inferring the presence of a disk through its infrared emission to the detailed characterisation of substructure on $\lesssim 10$ au scales. The Atacama Large Millimeter Array (ALMA) has spearheaded these observational efforts, providing some of the first measurements of disk substructure in the sub-millimetre around HL Tau ([ALMA Partnership et al., 2015](#)). The Disk Substructures at High Angular Resolution Project (DSHARP) extended this, with measurements

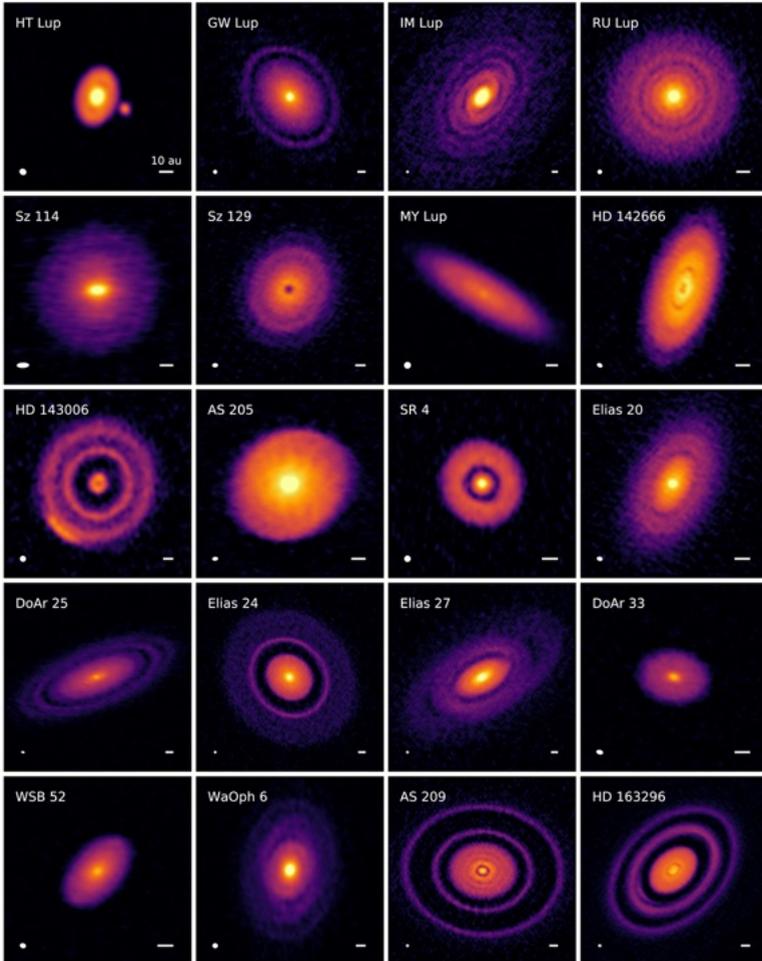


Figure 4.1: 1.25 mm (240 GHz) ALMA observations of bright circumstellar disks, observed as part of the DSHARP survey. Reproduced with permission from [Andrews et al. \(2018\)](#).

of disk substructure in bright, nearby protoplanetary disks, finding that such substructures are a universal feature of these disks ([Andrews et al., 2018](#)). Figure 4.1 shows the 1.25 mm (240 GHz) continuum emission of the DSHARP survey, highlighting the variety of morphologies present in the disks. These

observations highlight the necessity to include the impacts of substructure on planet formation: simple models with smooth density profiles are clearly not reflective of the disk population.

Substructures such as rings and gaps can both impact the planet formation process or be shaped by the presence of an unseen planet nearby. For example, due to viscous coupling the disk gas orbits at a sub-keplerian velocity. However, larger pebbles are not well-coupled to the gas and will follow keplerian orbits, leading to a speed differential between the pebbles and gas. Thus the pebbles experience aerodynamic drag from the gas, causing them to lose angular momentum and drift inward. Pressure bumps, such as in a ring, can stall this drift and cause a pileup of pebbles, leading to an overdensity that could trigger planet formation. Perhaps surprisingly however, most of the observed substructure occurs at separations far wider than the observed population of young, giant planets (Desidera et al., 2021; Nielsen et al., 2019). It thus remains to be seen how to link the large gas disks to the known exoplanet population.

In addition to the substructure observations in the disk continuum emission, ALMA has also pushed the boundaries in our understanding of disk chemistry. The Molecules with ALMA at Planet-forming Scales (MAPS) program examined ~ 50 molecular emission lines in 5 well-characterised disks (Öberg et al., 2021). Prior to the ALMA observations, the simple picture of disk chemistry used by the exoplanet community was the model of Öberg et al. (2011). This work explored the concept of snowlines in disks. As the disk midplane temperature decreases with separation, various species will condense out of the gas phase, changing the relative carbon-to-oxygen number ratios (C/O ratio) of both the gas and the solids in the disk. With high spatial resolution ALMA observations, it is clear that the distribution of molecules varies greatly both throughout a single disk, as well as between disks, as shown in Figure 4.2. Eistrup et al. (2018) use 1D models to show that not only do these molecular abundances vary spatially, but also in time as the disk evolves and ultimately disperses. As planets inherit their composition from the disks in which they form, it is critical to determine where the planet forming material originates, and how the chemistry of the gas and dust accreting onto the planet changes over time.

In addition to the disk emission observations from ALMA, scattered light observations in the visible to near-infrared from instruments such as VLT/SPHERE have revealed the structure of protoplanetary disks at small separations. These observations are sensitive to the scattering polarisation of dust at the surface of disks, typically at much smaller spatial scales than possible even with ALMA. Inner disk cavities appear to be a universal feature of Group I sources (Maaskant et al., 2013), following the classification of Meeus et al.

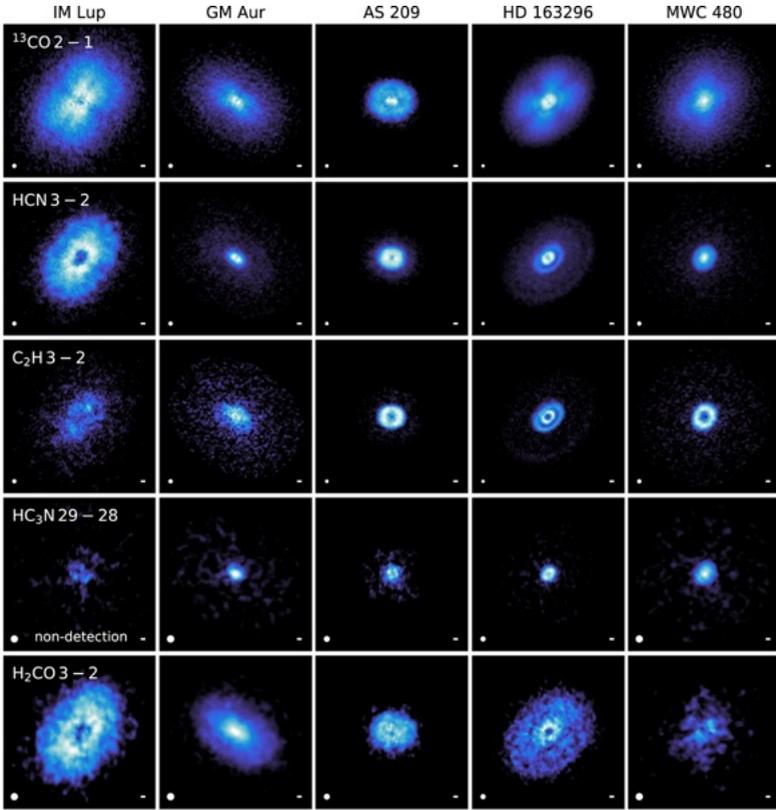


Figure 4.2: Zeroth moment maps for a selection of molecules for each of the disks observed as part of the MAPS survey. The small horizontal bar in the bottom right of each panel provides a 20 au scale, while the beam size is indicated in the bottom left. Reproduced with permission from [Öberg et al. \(2021\)](#).

(2001). In these disks the inner rim is located at a few tens of au from the host star, and is directly illuminated by it, leading to the far infrared excess from disk thermal emission. Conversely, Group II sources have shown less evidence for wide cavities and are fainter in scattered light ([Garufi et al., 2017a](#)). Nevertheless, in high spatial resolution imaging, morphological features in the dust appear to be universal, though not always directly correlated to gas phase features seen in ALMA images ([Garufi et al., 2017b](#)).

Protoplanetary disks are complex systems that evolve rapidly over their

~10 Myr lifetimes. Their composition and dynamics determine the outcomes of planet formation, which in turn shape the substructure and chemistry present in the disks. [Armitage \(2020\)](#) provides a classic introductory text on the topic. More recent and in depth reviews include that of [Zhao et al. \(2020\)](#), who provides detail on the early phases of disk formation. [Zurlo et al. \(2023\)](#) and [van der Marel \(2023\)](#) present the observational perspective on disks, emphasising the contributions of near infrared high-contrast imaging and ALMA sub-millimetre imaging respectively. A more theoretical approach focusing on dust formation and evolution is given by [Birnstiel \(2023\)](#), while [Öberg et al. \(2023\)](#) discusses general disk chemistry.

4.2. Formation mechanisms

Having set the stage by introducing the environment in which planets form, we can now turn our attention to the actual processes of planet formation. This is a complex topic, meeting at the intersection of hydrodynamics, thermodynamics and chemistry. Broadly speaking, planets are thought to be formed through either a ‘bottom-up’ or ‘top-down’ pathway. Each of these mechanisms produces different types of planets, and can be favoured under different conditions. One of the primary goals of exoplanet science is being able to determine the formation mechanism for a given planet or planetary system, drawing a line through the planet’s formation and evolution through to its present-day architecture and atmospheric state. For the purposes of this thesis we are primarily concerned with gas giant exoplanets, and so will focus our examination of planet formation on how to form such objects. As part of a larger series on the composition, mineralogy, and evolution of exoplanets, [Mordasini & Burn \(2024\)](#) provide a review of the observational constraints and state-of-the-art of formation modelling. In the same series, [Zhang \(2024\)](#) goes into more depth on the chemical environment of the protoplanetary disk in which the planets form.

4.2.1. Gravitational Instability

Gravitational instability (GI) is the ‘top-down’ formation pathway for exoplanets. In this scenario, giant planets form in a manner similar to stars, as a self-gravitating gas cloud fragments and collapses, ultimately forming a massive planet (e.g. [Perri & Cameron, 1974](#); [Cameron, 1978](#); [Mizuno, 1980](#)). This model was largely discarded for Solar System planets due to its failure to explain the similarity of their core masses, but was revived in the context of exoplanets by [Boss \(1997, 1998\)](#). This mechanism requires that the circumstellar disk

becomes unstable, following the Toomre criterion (Toomre, 1964)

$$\frac{c_s \Omega}{\pi G \Sigma} < Q_{\text{crit}} \approx 1. \quad (4.1)$$

Here, c_s is the speed of sound, Ω is the orbital frequency at a specified separation, and Σ is the local surface density. The critical value Q_{crit} has been determined through numerical simulations, and found to be approximately 1 for the formation of \sim Jupiter mass fragments (Raymond & Morbidelli, 2022, and references therein). The Toomre criterion is similar to other stability conditions, balancing the gravitational forces with the pressure of the system. It differs in its additional consideration of shear forces in a viscous, rotating disk, which provides additional support which the gravitational forces must overcome.

Satisfying this condition for instability is a necessary, but not sufficient condition for disk fragmentation. Gammie (2001) demonstrates the importance of cooling in the process of fragmentation. As the gas collapses, it will heat up, increasing the pressure and resisting further collapse. Only with sufficient cooling can the cloud complete its collapse and form a protoplanet. The cooling timescale is a function of both the temperature T and the optical depth of the gas, τ . Rafikov (2005) estimates the cooling timescale as

$$t_{\text{cool}} \approx \frac{\Sigma c_s^2}{\gamma - 1} \frac{f(\tau)}{2\sigma T}; f(\tau) = \tau + \frac{1}{\tau}. \quad (4.2)$$

Σ and c_s are the disk surface density and sound speed as above, and γ is the adiabatic index of the gas. The optical depth can be determined from the opacity of the disk κ , $\tau = \kappa \Sigma / 2$. In their analysis of the effect of the cooling timescale on disk fragmentation, Rafikov (2005) finds that GI is strongly disfavoured in inner disk regions at < 1 au, probable at around 10 au, and difficult at 100 au, due to the low surface densities at these separations, though the details depend on the specifics of the disk mass, opacity and temperature structure. They find that the mass of the fragment can be estimated as

$$M_f \approx \Sigma_{\text{min}} \lambda^2 \approx 0.15 M_{\odot} a_{\text{au}}^{-3/10} \left[\frac{f(\tau)}{\mu^4} \right] \left(\frac{M_*}{M_{\odot}} \right)^{1/10} \quad (4.3)$$

Here, $\lambda \approx 2\pi h$ is the length scale of the unstable mode for disk height h , a is the separation in au, μ is the mean molecular weight relative to atomic hydrogen, and M_* is the stellar mass in units of solar masses M_{\odot} . For typical values of these parameters, this results in relatively high mass fragments, on the order of $5M_{\text{Jup}}$ or greater.

Nero & Bjorkman (2009) explore the effects of local cooling timescales to determine whether the observed giant exoplanet population could be formed through GI. Their estimate of the cooling timescale is roughly an order of magnitude faster than that of Rafikov (2005), which means it is significantly easier for disk fragmentation to occur. They find that all of the giant planets would require large disk masses to be formed by GI. For the HR 8799 planets, they find that only the outermost planet could have formed in its present location via GI, though leaving open the possibility of migration from farther out. Indeed a few years later, Baruteau et al. (2011) find that massive, widely separated giant planets rapidly migrate inward on timescales of 10^4 years.

While this model provides a reasonable explanation for the formation of giant, widely separated exoplanets, it struggles to capture the details of the broader exoplanet population. Population synthesis models based on this formation pathway struggle to reproduce the overall exoplanet demographics: terrestrial and neptune mass planets are far more common than would be expected by GI alone. This model also requires very large disk masses, which are at the extreme end of the observed disk population. Finally, it primarily produces planets at wide separations. Even with rapid migration, it is challenging to produce the compact system architectures commonly observed in transiting and RV systems.

Schlaufman (2018) quantifies a boundary between GI and the core accretion mechanism. Data show a higher occurrence rate of giant planets around metal-rich stars if the planets form via core accretion (Fischer & Valenti, 2005). Schlaufman (2018) show that objects $> 10 M_{\text{Jup}}$ do not follow this trend, and thus must have formed via GI, while objects below that bound follow the host-star metallicity trend, and are therefore more likely to have formed via core accretion.

4.2.2. Core Accretion

Due to the challenges of GI in explaining both the Solar System planets and the broader exoplanet population, core accretion has become the preferred model of planet formation. This is the ‘bottom-up’ approach to planet formation. In this model, planets form by the gradual accumulation of pebbles and larger bodies known as planetesimals, until sufficient mass is reached that the protoplanet can start to accrete and maintain a hydrogen-helium envelope (e.g. Armitage, 2018; Raymond & Morbidelli, 2022; Drażkowska et al., 2023).

The canonical picture of core accretion was first outlined in Pollack et al. (1996). In their work, they simulate how a disk composed of gas and planetesimals will evolve over time, and in particular how they will accrete into a planet. They account for some degree of accretion physics: planetesimals

will enrich the planet envelope and core with metals, and will deposit energy into the planet, increasing its temperature. They identify three phases of the accretion process. First, planetesimals rapidly accrete. Their trajectories are governed largely by gravity rather than aerodynamic drag, and so the mutual attraction of planetesimals on similar orbits leads to gravitational focusing, effectively increasing their collisional cross section area. After the collision of two planetesimals into a larger body, its gravitational influence grows, further attracting nearby planetesimals. This phase is also known as runaway planetesimal accretion, and will continue until the local reservoir of planetesimals is depleted. As the velocity dispersion of the planetesimals increases to the order of the escape velocity, a new phase of growth begins, with the growth rate now proportional to $M^{-1/3}$. Thus the growth rate of larger objects decreases relative to smaller protoplanets, resulting in roughly equal mass bodies. This phase is known as oligarchic growth, with these larger bodies ultimately forming the cores of planets. Due to their gravitational region of influence, the oligarchs will be separated by ~ 5 -10 mutual Hill radii. During this phase, gas will begin to be accreted by the protoplanets as well. Once a critical threshold is reached where the gas mass is equal to the solid mass, the final phase of runaway accretion begins. This phase sees approximately exponential growth in the planet mass due to gas accretion, drawing from the disk reservoir. For numerical reasons, [Pollack et al. \(1996\)](#) did not explore the end state of this process. Subsequent works have found that this process is limited by the availability of material from the disk, as well as the thermodynamics of the accretion process, but it remains difficult to identify how this process turns off sufficiently early to form Saturn-mass planets ([Raymond & Morbidelli, 2022](#)). During this phase the planet will continue to accrete some remaining planetesimals which can be ablated by the planet envelope, and will contribute to the enhanced metallicity of the planet's atmosphere and core. While this simple framework is conceptually appealing, [Ida & Lin \(2004\)](#) noticed that the actual accretion rate of oligarchs outside of snowlines is very low, resulting in growth timescales longer than the disk lifetime. Given the observed population of exoplanets at wider separations, a mechanism for faster protoplanetary growth was required.

The pebble accretion mechanism postulates that rather than forming protoplanetary cores via the collision of large ($10^2 - 10^3$ km) planetesimals, they can be formed through the accumulation of smaller mm to cm sized grains and pebbles. This also avoids the problem of forming large planetesimals from dust and pebbles, which tend to fragment rather than coagulate at sizes of 1 cm or larger ([Zsom et al., 2010](#)). It was first recognised by [Ormel & Klahr \(2010\)](#) that the accretion radius for such small pebbles onto a protoplanet is much larger than the radius at which planetesimals can be accreted. This was more fully

explored in [Lambrechts & Johansen \(2012\)](#). They show that the entire region of the disk exterior to a protoplanet can act as a reservoir for pebbles: due to the aerodynamic drag between the gas on a sub-keplerian orbit and the keplerian pebbles, the pebbles lose angular momentum and rapidly drift inward. This provides a large flux of pebbles into the accretion radius of the protoplanet, which can then be efficiently accreted, growing the core to the critical mass within the disk lifetime. This growth continues until the ‘pebble isolation mass’ is reached, at which point the pressure bump induced by the gap opened by the protoplanet traps pebbles outside of its accretion radius, and only gas can continue to accrete onto the planet. [Lambrechts & Johansen \(2014\)](#) expand this to show that pebbles can grow efficiently near snowlines where volatiles sublime into the solid phase, which can also induce pressure bumps which further enhance the pebble surface density. This mechanism doesn’t entirely do away with planetesimals however: protoplanets smaller than ~ 500 km do not accrete pebbles efficiently, and therefore some large protoplanets must be present in the disk to begin the pebble accretion process ([Liu et al., 2019](#)).

Of course, planets do not remain stationary in the disk. [Bitsch et al. \(2015, 2019\)](#) discuss the impact of migration through the disk on the outcome of planet formation driven by pebble and planetesimal accretion. They show that planetesimal accretion is slower than the migration timescale through the disk and the disk lifetime itself, while pebble accretion can build planets fast enough to avoid catastrophic migration into the star. Likewise, the pebbles also change as they drift through the disk. As pebbles drift inward, they can evaporate as they cross snowlines, locally enriching the gas in volatiles ([Bitsch & Mah, 2023](#); [Danti et al., 2023](#)). The remaining small, refractory grains drift more slowly and pile up, leading to a similar enhancement in the dust-to-gas ratio of the disk. Both of these factors then contribute to the observed metal rich compositions of both giant exoplanets and the inner regions of circumstellar disks.

All of these formation scenarios are complicated by the fact that they occur in three dimensional circumstellar disks, with substructures determined between the interplay between (magneto)-hydrodynamics, radiative transfer, and N-body gravitation. In an effort to produce a more realistic picture of the accretion processes, [Morbidelli et al. \(2014\)](#) explored the hydrodynamics of accretion accounting for realistic heating and cooling, showing that gas accretion onto the planet is largely from the vertical direction, with most of the material spiralling outward from the planet in the circumplanetary disk (CDP). [Szulágyi et al. \(2014\)](#) uses this same model to show that in the limit of low viscosity, these 3D accretion flows produce lower growth rates than initially predicted in [Pollack et al. \(1996\)](#).

4.3. Evolution

Once a planet has formed, it begins the long process of radiating away the internal heat stored during the accretion process. As planets, by definition, are not massive enough to sustain nuclear fusion, in the absence of stellar irradiation they will monotonically cool over the course of their lifetime. Due to their similar internal structure, brown dwarf evolutionary models such as those of [Chabrier et al. \(2000\)](#), [Baraffe et al. \(2003\)](#) or [Saumon & Marley \(2008\)](#) have provided the basis for giant planet evolutionary models, and in turn inherit much of their physical basis from models of stellar evolution. These models compute T_{eff} as a function of time, for various masses and initial entropies. Based on assumptions of atmospheric physics and internal structure, an emission spectrum can be calculated, and the radiative power measured, leading to a gradual decrease in the effective temperature with each time step. These models generally account for absorption from common molecular species, some parameterisation of clouds or dust absorption, and determination of radiative and convective zones in the atmosphere by stability considerations and mixing length theory.

The goal of these models is to be able to determine any two of the following four properties, given the other two: mass, age, effective temperature, and radius. As colour-magnitude diagrams for brown dwarfs effectively trace their evolution in temperature over time, such evolutionary models should also capture the observed changes in colour and spectra with temperature. Most prominently is the transition between L-type and T-type brown dwarfs ([Kirkpatrick et al., 1999](#); [Kirkpatrick, 2005](#)). This transition occurs between 1200 K – 1400 K, where the colour of the brown dwarfs suddenly appears much more blue, breaking the trend of gradually reddening through the L-dwarf sequence. This transition is thought to be caused by two factors: as the temperature decreases, silicate clouds sink beneath the photosphere as the condensation point occurs at higher and higher pressures for the same temperature, and methane absorption becomes a significant component of the spectra as it becomes thermodynamically favoured over CO at cooler temperatures.

More recent evolutionary models such as [Phillips et al. \(2020\)](#) and [Morley et al. \(2024\)](#) extend the parameter space to cooler temperatures, use temperature-pressure profiles and chemistry derived from self-consistent radiative-convective equilibrium models, and use either state-of-the-art cloud models or account for reddening via a separate mechanism (discussed further in [Tremblin et al., 2015, 2016, 2017](#); [Phillips et al., 2020](#)).

Of course, these models rely heavily on understanding what the initial conditions are. These conditions are typically formulated in terms of the interior

entropy, or equivalently the amount of energy stored in the planet's interior following the formation process. Two extreme cases are usually considered, known as hot-start and cold-start. More realistic is a 'warm-start' case, existing somewhere between the two extremes. In the hot-start case, the planet must form via the rapid accretion of material, without time to radiate and cool. Conversely, a cold-start case, typically associated with a GI formation pathway, requires a slower accumulation of material, without the addition of significant energy to the interior of the planet. These assumptions have been explored by following the thermodynamics of accreting material in 3D simulation of a planet in a core accretion scenario, for example in [Szulágyi & Mordasini \(2017\)](#). In this work, the authors find that a significant fraction of the energy of the infalling gas is radiated away in an accretion shock on the planet envelope and the circumplanetary disk, resulting in cooler initial conditions than assumed for a true hot-start. This shock is also the source of H α emission observed in accreting protoplanets such as PDS 70 b and c ([Haffert et al., 2019](#)). Ultimately, a hotter object will cool more rapidly, and so after a few hundred million years, the two extreme scenarios become indistinguishable.

In order to constrain these scenarios, precise bolometric luminosities of very young objects with known masses and radii must be determined - unfortunately for directly imaged planets, it is difficult to precisely measure any of these parameters. Long term astrometric monitoring or radial velocity measurements are necessary for accurate dynamical mass estimates, and broad wavelength coverage must be obtained to measure the temperature. Unless the object is transiting, an independent radius estimate is nearly impossible. Finally, age estimates depend largely on inferring the age of the star, either by activity indicators or through membership in a young moving group. Such measurements have only been made for a small handful of targets, and so further measurements are necessary in order to understand the initial conditions for evolutionary models, which in turn inform our understanding of the planet formation mechanism.

5. ATMOSPHERIC PHYSICS

A planet's atmosphere is among its most characteristic features. They are a common, though not ubiquitous property of the Solar System planets, though even the absence of an atmosphere tells a story of a planet's formation and evolution through time. For terrestrial worlds, the atmosphere maintains the surface temperature, interacts with oceans and the interior, and is the location of ever-changing weather patterns. For gas giant planets, the atmosphere is nearly our only window into their interior. This thin skin at the outer reaches of the planet wholly shapes the observable thermal emission, and the patterns of the clouds and gasses determine the visible light reflected to us. Beyond a planet's mass and radius, the atmosphere is the only feature we have to characterise the planet's composition, temperature, dynamics and evolutionary history. It is therefore of critical importance to understand the physical processes that shape these dynamical systems, and in turn how the atmosphere shapes what we can observe.

In this section I will outline our current understanding of atmospheric physics focusing on gas giant planets. [Zhang \(2020\)](#) provides a comprehensive review of this subject, providing detail well beyond the scope of this work. I begin with how we learn about the atmospheres, those of the planets in our own Solar System as well as those of brown dwarfs, the cousins of directly imaged exoplanets. These observations then inform our model of the interior of gas giant planets, working from the interior out to the atmosphere, with particular attention paid to the atmospheric chemistry, and how this ultimately impacts calculations of thermal emission spectra. I will discuss modelling frameworks commonly used in the exoplanet community, building to the core problem of this thesis: how do we go backwards from measuring the spectrum of an exoplanet to inferring its atmospheric properties?

5.1. Solar System observations

The Solar System is divided between the rocky, inner planets, and gaseous worlds of the outer Solar System. Amongst the gas giants we see great diversity, from the enormous mass of Jupiter, to the low density of Saturn, to the frozen ice giants of Uranus and Neptune. [Sánchez-Lavega et al. \(2023\)](#) reviews the observations and models of the atmospheres of these worlds, while [Miguel & Vazan \(2023\)](#) provides an in depth exploration of their interior structure.

As with exoplanets, most of our understanding of these worlds has come from remote sensing observations, that is telescope observations of reflected sun light and thermal emission directly from the planets, such as the images of



Figure 5.1: Simultaneous observations of Jupiter observed in the thermal infrared, visible, and ultraviolet. International Gemini Observatory/NOIR-Lab/NSF/AURA/NASA/ESA, M.H. Wong and I. de Pater (UC Berkeley) et al.

Acknowledgments: M. Zamani

Jupiter shown in Figure 5.1. Unlike exoplanets, we have also sent orbiters and probes to make in-situ measurements of these planets. The earliest missions to pass by these planets were the *Voyager* probes on their Grand Tour of the Solar System, launched in 1977 (Kohlhase & Penzo, 1977). The images obtained during their flybys of the gas giants revolutionised our understanding of these objects and captured the public imagination. For Jupiter, we could view the clouds and storms with unprecedented detail. Magnetic fields around Uranus were detected for the first time, and were oriented nearly 90° from its rotation axis. Using its radiometer and infrared spectrometer, *Voyager 2* also found an anomalously low heat flux from Uranus compared to the other three planets. The wind speeds and storms were measured for all of the planets, showcasing just how dynamic and variable these objects are. More importantly than any single discovery, the *Voyager* missions highlighted the importance of comparison between the planets to understand their similarities and differences.

More recently, missions such as *Galileo* (Johnson et al., 1992), *Cassini* (Matson et al., 2002), and *Juno* (Bolton et al., 2017) have enabled the most precise measurements of the gas giant atmospheres to date. *Galileo* provided the first in-situ atmospheric composition measurements by dropping a probe into the atmosphere at 6.5° north latitude and 4.4° west longitude. The probe descended through the atmosphere for nearly an hour, reaching a pressure of over 20 bar, and finding unexpectedly low abundances of water. The *Cassini* Grand Finale provided similar direct measurements of Saturn's upper atmosphere using mass spectroscopy, finding that CH_4 , H_2O and NH_3 must be entering the upper atmosphere from its rings (Yelle et al., 2018; Serigano et al., 2020, 2022) While

not the first to make such measurements, *Juno* made precise measurements of the gravitational moments of Jupiter, effectively probing changes in its shape and density from a homogeneous sphere. These measurements showed that Jupiter’s core is likely dilute, contrary to contemporary models which generally divided the planet into clear, compositionally distinct layers. All of these novel measurement approaches highlight the importance of the Solar System planets for our understanding of the internal structure and composition of gas giant planets: such measurements are completely inaccessible for the exoplanet population.

Even with the advent of space missions, much of our understanding of the Solar System is still derived from telescope observations. The critical parameters necessary to describe an atmosphere are its temperature and its composition. The earliest recorded temperature measurement of the gas giants were by [Menzel et al. \(1926\)](#), who found the temperature of Jupiter to be between 120 K, reasonably close to the modern estimate of $T_{\text{eff}} = 88$ K, as derived from the bolometric luminosity

$$L_{\text{bol}} = 4\pi r_p^2 \sigma T_{\text{eff}}^4, \quad (5.1)$$

where r_p is the planet radius and σ is the Stefan-Boltzmann constant. As is apparent on Earth, the temperature varies significantly throughout the atmosphere, and so a more detailed model is required. Using data from the Voyager flyby, [Seiff et al. \(1998\)](#) measured the vertical thermal profile of the upper atmosphere of Jupiter, finding a strong temperature inversion above the tropopause at around 0.1 bar, where the temperature reaches a minimum at 120 K. Below the tropopause, the atmosphere closely follows an adiabat: that is to say a parcel of air that is moved upward or downward through the atmosphere is only heated or cooled by its adiabatic expansion or compression as the surrounding pressure varies.

More recent probes have enabled the characterisation of the thermal of Jupiter and Saturn’s atmospheres with an unprecedented level of precision. Using the mid-to-far infrared instrument CIRS on board *Cassini*, [Fletcher et al. \(2009\)](#) are able to measure the temperature as a function of pressure and latitude for both Jupiter and Saturn. They obtain direct temperature measurements from 1 bar up to 0.001 bar, and see clear temperature differences between the Northern and Southern hemispheres, with stronger seasonal variation apparent in Saturn. More recently [Fletcher et al. \(2023\)](#) used the MIRI/MRS on board *JWST* to determine the thermal structure, composition and seasonal variation in the Northern hemisphere of Saturn. They identify a strong banded structure visible in the mid infrared, together with an abundance of dynamical phenomena such as seasonal vortices, cyclones, and both eastward and westward

jets.

The distant, frozen worlds of Uranus and Neptune display sharply different characteristics than Saturn or Jupiter. Their lower mass and volatile rich composition suggest that they were unable to undergo runaway gas accretion, perhaps forming later than the inner two gas giants, after much of the protosolar gas disk had dispersed (Helled et al., 2020a). However, there remain many open questions about their formation and present day conditions. Uranus appears to be in thermodynamic equilibrium with the incident solar radiation, while Neptune exhibits a significant internal heat source. Stranger still is the obliquity of Uranus rotational axis, potentially caused by a giant impact (Rogoszinski & Hamilton, 2021). At 98° to its orbital planet, its poles receive more direct sunlight than the equator, driving strong seasonal variations. It was long thought that the atmosphere of Uranus was less dynamic than that of Neptune, which has strong visible vortexes and bands, while the Voyager images of Uranus appeared nearly featureless. More recent ground-based observations have found strong tropospheric circulation in Uranus' atmosphere (Molter et al., 2019). From the Voyager flybys, low order gravitational moments for both planets were measured. To fit these data, models find that both planets likely have interiors dominated by volatiles, with compositional gradients between the deep interior and the atmosphere, though the details of their internal structure remain controversial (Helled et al., 2011; Helling et al., 2019; Helled et al., 2020a). Their atmospheres show temperature gradients consistent with a dry adiabat, though again the lack of data allows for solutions ranging from sub- to super-adiabatic temperature gradients (Guillot, 1995).

The composition of the gas giant planets is reasonably well understood. Jupiter and Saturn are the most massive planets in the Solar System ($318 M_\oplus$ and $95 M_\oplus$ respectively), and both are nearly solar in composition, primarily composed of hydrogen and helium. Both show some enhancement of their metal abundances, (around $3\times$ what is found in the sun for Jupiter, and $10\times$ for Saturn). Chemical inventories have been taken for both atmospheres, detecting chemical species such as CH_4 , NH_3 , PH_3 , H_2S , isotopologues such as HD and CH_3D , and noble gasses including argon, krypton, and neon (Taylor et al., 2007). Due to their cold temperatures, water condenses out of the atmosphere and thus does not impact the reflection or emission spectra. Measurements of the planet metallicities are therefore driven by measurements of the methane abundance, which is the primary carbon bearing species in the atmosphere. Visually both planets are dominated by their clouds. Jupiter is covered in complex bands of clouds and storms, with the Great Red Spot its most prominent feature. The colouration of the bands is thought to vary, as pristine, white aerosol material is brought to the upper atmosphere by convective upwelling,

before mixing with photochemically produced material from the interaction of the atmosphere and solar UV radiation.

In contrast to the inner two gas giants, Uranus and Neptune are both composed of around 80% volatiles by mass. Due to strong compositional gradients, their atmospheres remain dominated by hydrogen and helium, with methane as the primary trace gas (Helled et al., 2020b,a). The atmospheres are too cold for water to occur in the gas phase. Instead, H₂S has been tentatively detected in the atmospheres of both planets, and although ammonia is expected to be present at these low temperatures, no detections have been made. These differences even within the Solar System highlight the diversity of structure and composition that can be expected in the exoplanet population.

Even without being able to spatially resolve them, such observations of our nearest neighbours inform our understanding of exoplanets, and the same techniques can be applied to distant systems. From these Solar System examples, we now have most of the observational pieces of the puzzle. We must turn towards brown dwarfs in order to fill in the rest of the gaps before putting everything together in a consistent theoretical framework.

5.2. Brown dwarf observations

Without a bright host star complicating observations, brown dwarfs provide excellent, easily observable analogues of the directly imaged exoplanet population. The study of these objects has developed alongside that of exoplanets, following the first incontrovertible brown dwarf discovery in the 1990s by Nakajima et al. (1995). They share similar masses and compositions, and share a similar formation mechanism to planets formed via GI. With a population in the thousands, it is also possible to measure statistical trends. Helling & Casewell (2014) provide a thorough overview of these objects, and Marley & Robinson (2015) outlines the modelling approaches used to characterise their atmospheres.

Like stars, brown dwarfs are classified by their spectral type (Kirkpatrick et al., 1999; Kirkpatrick, 2005). The L-dwarfs cover a temperature range from around 2300 K down to 1300 K, where the J-K colour begins to rapidly change. At this point the T-dwarf sequence begins, extending down to around 500 K. The coldest brown dwarfs are the Y-dwarfs, of which only a small number of nearby objects have been characterised to date. Figure 5.2 shows the SpeX spectral standards for the L and T dwarfs. These objects are used as benchmarks for defining the spectral types, though more recent examinations of these objects have shown that some of them may be poor choices as benchmarks due to their young age or for being potentially unresolved binaries or otherwise peculiar

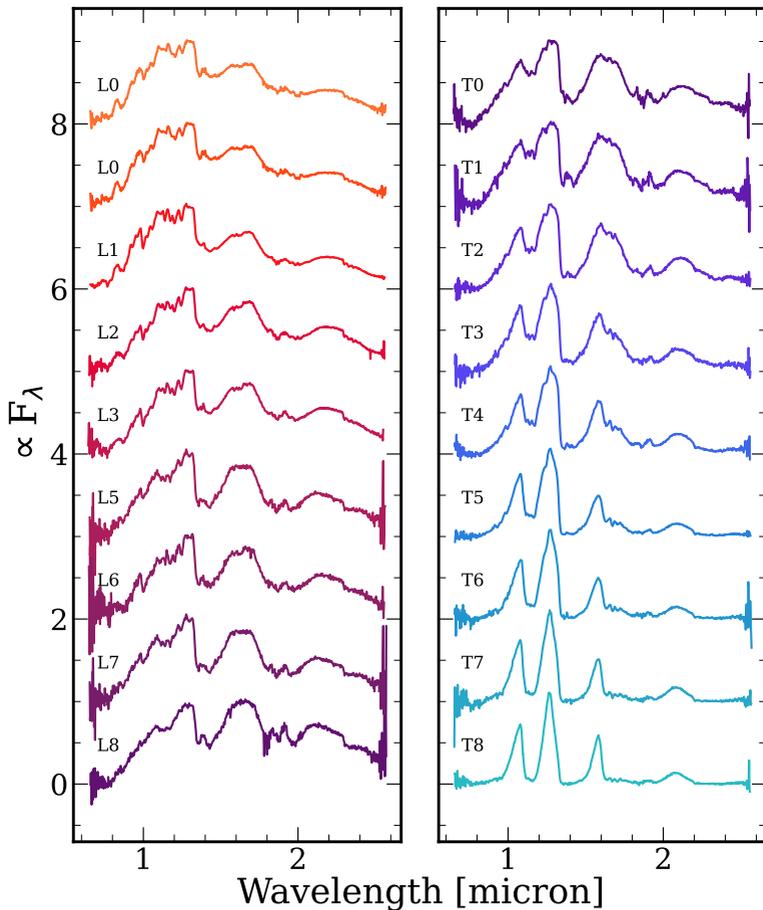


Figure 5.2: SpeX brown dwarf spectral standards from L0 to T8. The spectral type tracks the effective temperature, from ~ 2300 K for an L0 dwarf to ~ 500 K for a late T-dwarf. The spectral features are shaped by molecular and aerosol absorption. Water and carbon monoxide produce the large spectral features in warmer objects, while water and methane dominate the colder objects. Clouds reduce the amplitude of the molecular features and redden the spectra.

objects (e.g. [Burgasser et al., 2004](#); [Burgasser et al., 2006](#)).

The spectral sequence tracks the effective temperature of the L-dwarfs. As the temperature decreases, the objects become increasingly red in colour. This

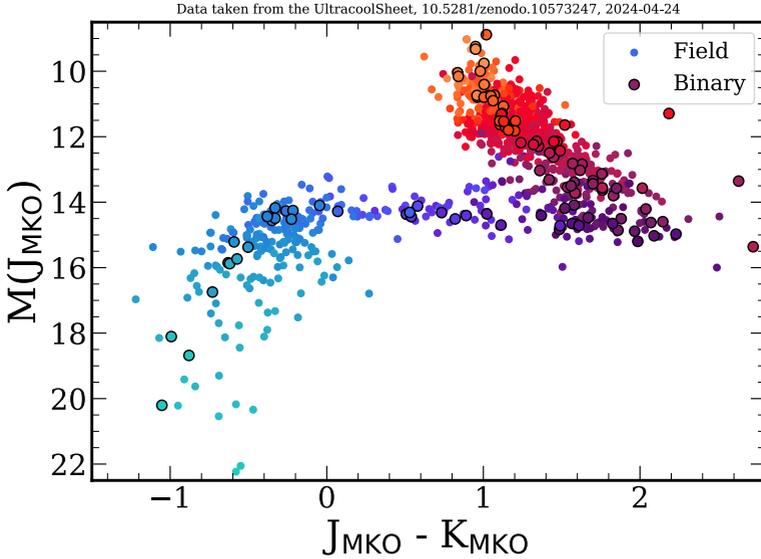


Figure 5.3: J-K colour magnitude diagram. The J-band magnitude roughly tracks effective temperature, while the J-K colour axis is a proxy for changes in the clouds and chemistry. The L/T transition occurs where the J-K colour suddenly becomes more blue due to changes in the clouds and chemistry of the atmospheres. All magnitudes are in the MKO system. The colour of the points indicates spectral type, from early L (orange) to late T (light blue). Data obtained from The UltracoolSheet at <http://bit.ly/UltracoolSheet>.

is partly due to the shift in the peak of the blackbody emission, but is also affected by the presence of clouds in the atmosphere. At temperatures below ~ 1600 K, silicates such as enstatite (MgSiO_3) and forsterite (Mg_2SiO_4) can condense out of the gas phase and form cloud layers (e.g. [Helling & Casewell, 2014](#)). These aerosol species have absorption features in the mid infrared, which were directly observed using the *Spitzer Space Telescope* ([Cushing et al., 2006](#); [Suárez & Metchev, 2022, 2023](#)). Other refractory species, such as metallic iron, also condense out at similar temperatures, leading to multiple cloud layers at different altitudes in the atmospheres. These clouds are thought to be one of the causes of the sudden change in colour that marks the transition to T-dwarfs, seen in Figure 5.3. As the silicates condense out at roughly a constant temperature, for cooler objects this means that they will condense at a higher pressure (lower altitude) than in a hotter object. If the object is cold enough,

this condensation pressure will be below the photosphere, and the clouds will no longer impact the emission spectrum.

Given that the brown dwarfs cannot sustain fusion, they will inevitably cool over time. Thus the L and T sequence doesn't only track temperature, but age. There are clear degeneracies here, as a heavier object will cool more slowly than a less massive object. Borrowing from the vernacular of the stellar community, it is common to refer to the hottest L dwarfs as 'early-type' objects, and T dwarfs as 'late-type' objects. This is complicated by the change in surface gravity with time. As the brown dwarf cools, it will contract. As mass is conserved, but the radius decreases, the surface gravity will also increase over time. Low-surface gravity is therefore thought to indicate young ages. Determining the surface gravity from the spectra alone is challenging. Objects in binary systems, where a dynamical mass can be estimated from the orbits, are important benchmarks for calibrating models.

The chemical composition of a brown dwarf depends on the interstellar environment in which it was born. Forming like stars, brown dwarfs generally lack pathways to the highly metal enriched atmospheres present in the exoplanet population [Chabrier et al. \(2014\)](#); [Maldonado & Villaver \(2017\)](#). While the environment determines the overall elemental abundances available in the atmosphere, the chemical species present depend strongly on the temperature of the object. L-dwarf spectral shapes are dominated in the near infrared by water and carbon monoxide ([Kirkpatrick et al., 1999](#); [Burrows et al., 2001](#)). Many other molecular and atomic species, have also been detected from their absorption signatures. Refractory species, such as TiO or FeH are more thermally stable, and so are visible in the coolest stars and hottest brown dwarfs. H₂O and CO are the dominant species in the mid L range, producing the spectral shape characteristic of this class of object. T-dwarfs are identifiable by their CH₄ absorption, a species which only becomes thermodynamically favoured at colder temperatures ($\lesssim 1000K$) ([Kirkpatrick et al., 1999](#)).

JWST has opened a new window into the atmospheres of ultracool dwarfs. With its unprecedented sensitivity, it has been able to provide detailed characterisation of the coldest Y dwarfs ([Luhman et al., 2024](#)), planetary mass objects have been characterised with both high resolution and broad wavelength coverage ([Miles et al., 2023](#); [Manjavacas et al., 2024](#)), and it is now not only possible to identify atomic and molecular features, but isotopologues as well ([Barrado et al., 2023](#)). The full extent to which *JWST* will transform our understanding of these objects is not yet clear, but it is apparent that there remain many lessons to be learned from high precision characterisation.

Between the cold Solar System giant planets and the isolated brown dwarfs, we are ready to put the pieces together and develop our physical model of these

objects and their atmospheres. Gas giant planets have a complicated interior structure that is the result of their formation history and differentiation over time (Miguel & Vazan, 2023). Models of their interiors must describe this structure while remaining consistent with their measured masses and radius, and will provide the lower boundary conditions for the atmosphere. Deep in the atmosphere of the giant planets the temperature structure follows an adiabatic profile, while the upper atmosphere is more strongly influence by heating from the sun. Their composition can vary greatly, affecting both the observed colours and spectra of the planets. Clouds are an ubiquitous feature of their atmosphere, where clouds of a given species will condense out of the gas phase at a consistent temperature. From brown dwarfs we understand the library of chemical species that we must consider, as well as the clouds that are present throughout different layers of the atmosphere. Our atmospheric model should account for the absorption of these molecular and aerosol species, as well as their distribution throughout the atmosphere, ultimately producing emission spectra that should reproduce the brown dwarf spectral templates, given appropriate choices of parameters.

5.3. Structure, interior and dynamics

Let us know explore the structure of gas giant planets from the inside out, following the review Miguel & Vazan (2023) as a guide for this topic. Giant planets are self-gravitating spheres composed primarily of hydrogen and helium. This can be modelled as a self-gravitating fluid under hydrostatic equilibrium:

$$\frac{\partial P}{\partial r} = -\rho g, \quad (5.2)$$

$$\frac{\partial T}{\partial r} = \frac{\partial P}{\partial r} \frac{T}{P} \nabla_T, \quad (5.3)$$

$$\frac{\partial m}{\partial r} = 4\pi r^2 \rho, \quad (5.4)$$

$$\frac{\partial L}{\partial r} = 4\pi r^2 \rho T \frac{\partial S}{\partial t}. \quad (5.5)$$

P is the pressure, the gravity $g = Gm/r^2$, T is the temperature, and ∇_T the temperature gradient, which depends on the process that dominates the energy transport. The local luminosity is L while S is the specific entropy. The time coordinate is t .

The temperature gradient is determined by the heat transport mechanism. In the simplest case, we can treat the temperature profile as adiabatic, apart from the outermost radiative layer. For an adiabatic profile the specific entropy

S is constant in across the layers, $\partial S / \partial r = 0$, perhaps driven by convection which can homogenise the composition over time. For this model, the composition must also remain constant throughout the interior. For an adiabat, the temperature gradient is

$$\nabla_{\text{ad}} = \left(\frac{\partial \log T}{\partial \log P} \right)_{\text{ad}} = \frac{(\gamma - 1)}{\gamma}, \quad (5.6)$$

where γ is the adiabatic index, which is the ratio of the heat capacity of a material at constant pressure to the heat capacity at constant volume. For an ideal diatomic gas, $\gamma = 7/5$. Non-adiabatic models can account for mass transport and mixing, multiple convective cells, and convective-conductive heat transport (e.g., [Vazan et al., 2015](#)).

In addition to the equations of hydrostatic equilibrium, and equation of state (EOS) is required. This is one of the largest sources of uncertainty in interior models due to the extreme pressures and temperatures involved, as well as the mix of materials that are present in the giant planet interiors. Considering an object composed of purely hydrogen and helium, the high pressures deep in the interior will lead to the phase transition of hydrogen into a metallic fluid. The helium may not be miscible with the hydrogen fluid, and so will form a stratified layer surrounding the hydrogen core ([Stevenson & Salpeter, 1977](#); [Brygoo et al., 2021](#)). The EOS for hydrogen-helium mixtures have been numerically computed, and the results compared to lab experiments. Nevertheless, even for simple compositions there is significant variation between models. The addition of ices and rocky material adds further complication. It remains unclear whether these materials could form a solid core, or would be fully mixed into the metallic hydrogen. There are likely compositional gradients throughout the interior of the giant planets ([Wahl et al., 2017](#); [Debras & Chabrier, 2019](#)). This has two main impacts on providing the boundary conditions for the atmosphere: first, that much of the heavy elements (including helium) will be sequestered deeper in the planet, so the atmospheric metallicity is not representative of the bulk metallicity. Second, these compositional gradients can inhibit convection, and therefore the heat transfer within the planet and ultimately the luminosity at the bottom of the atmosphere.

Although planetary atmospheres are highly dynamic, the assumption of hydrostatic equilibrium remains valid as the horizontal motion of the atmosphere dominates over vertical transport. The fundamental equation set for describing an atmosphere is composed of a continuity equation, a momentum equation, an energy equation, an equation of state, an equation of radiative transfer, and a series of transport equations for chemical species. For a full global circulation model, we might be interested in solving the primitive equations, which are the

simplification of the full Navier-Stokes fluid equations under hydrostatic balance. Zhang (2020) provides an exploration of these dynamical fluid equations and their application to atmospheric models. Marley & Robinson (2015) also reviews the atmospheres of brown dwarfs and giant planets from a theoretical perspective.

To a first approximation, we can consider the atmospheres of brown dwarfs and giant planets as 1-dimensional, thus greatly simplifying the computational complexity. The atmospheres vary much more greatly as a function of altitude than of latitude or longitude. This approximation is especially useful for self-luminous objects without significant irradiation, where the internal heat flux at the bottom of the atmosphere is thought to be uniform across the planet. The atmosphere can be composed of both convective and radiative regions. Convective regions follow a temperature profile determined by equation 5.6. The temperature gradient of radiative regions depends on the local net flux F_{rad} , the local temperature, and the opacity of the layer. For simplicity, considering the Rosseland mean opacity κ_R

$$\nabla_{\text{rad}} = \left(\frac{\partial \log T}{\partial \log P} \right)_{\text{rad}} = \frac{3\kappa_R P F_{\text{rad}}}{16g\sigma T^4}. \quad (5.7)$$

A region is convectively stable if the Schwarzschild criterion is satisfied, $\nabla_{\text{rad}} < \nabla_{\text{ad}}$, or equivalently

$$-\frac{\partial T}{\partial r} < \frac{g}{c_p}. \quad (5.8)$$

Here, for stability against convection the negative temperature gradient as a function of altitude must be less than gravity g divided by the local heat capacity at a constant pressure (c_p). The radiative convective boundary (RCB) occurs where $\nabla_{\text{rad}} = \nabla_{\text{ad}}$. In the approximation that the atmospheric opacity is grey (has no wavelength dependence), $P_{\text{RCB}} \approx g/\kappa_R$. There can be multiple convective layers in the atmosphere, due to the varying opacity as a function of altitude (Burrows et al., 1997; Mukherjee et al., 2022). For example, clouds may condense at a particular altitude, giving rise to a dramatic increase in the opacity, and thus changing whether the Schwarzschild criterion is satisfied or not.

If the entire atmosphere experiences strong convection such that the mixing timescale is shorter than the timescale for chemical reactions, its composition will be homogenised, mixing material upward from the bottom of the convective layer (Fegley & Lodders, 1996; Noll et al., 1997). Conversely, if the chemical timescale is faster than the mixing timescale, the atmosphere will be in a state of

local thermodynamic equilibrium (LTE), and the chemical abundances will be determined by the temperature, pressure, and available elements. If there are multiple convective regions separated by radiative regions, it should be expected to have a different composition, as the upper convective layer cannot draw material from deep within the atmosphere (Mukherjee et al., 2022).

5.4. Chemistry

The chemistry of the atmosphere depends on the available elements, the temperature structure, dynamic processes that mix material vertically and horizontally, and additional radiation sources that can induce photochemical reactions. In the most simple case, we can consider an atmosphere in chemical equilibrium. In such a case, the only variables necessary to consider are the elemental abundances, the temperature, and the pressure. As this is a local equilibrium, we can divide up the atmosphere into independent layers, each with a temperature and pressure. Equilibrium occurs when the Gibbs free energy G is minimised, where

$$G(P, T) = U + PV - TS = H - TS. \quad (5.9)$$

G is a function of pressure P and temperature T and volume V . U is the internal energy of the system and S is the entropy. H is the enthalpy, where $H = U + PV$.

Solving for equilibrium therefore is simply a matter of minimising G . A widely used implementation for this was introduced in the Chemical Equilibrium and Applications (CEA) program (Gordon & McBride, 1994), which directly descends from early analytical and computational work from Zeleznik & Gordon (1960). I outline the CEA method here. The Gibbs energy for a chemical species depends on its chemical potential μ and the quantity of the species in moles n . For N species, the Gibbs energy per unit mass g is

$$g = \sum_{i=1}^N \mu_i n_i. \quad (5.10)$$

We must apply additional constraints on the problem, such as mass conservation. In the general formulation using Lagrange multipliers λ to apply such constraints, we must minimise G

$$G = g + \sum_j^l \lambda_j (b_j - b_j^0), \quad (5.11)$$

where b_j^0 is the atomic mass mixing ratio per unit mass of total reactants, which must be conserved.

$$b_j = \sum_j^l a_{ij}n_j, \quad (5.12)$$

where a_{ij} is the number of moles of element j per species i , and so b_j is the atomic mass mixing ratio calculated from the abundance of the reaction products. Thus the mass conservation condition is

$$\sum_j^l a_{ij}n_j - b_i^0 = 0. \quad (5.13)$$

Taking the derivative of G , together with the mass conservation constraint and applying the minimisation condition

$$\partial G = \sum_i^N \left(\mu_i + \sum_j^l \lambda_i a_{ij} \right) \partial n_j + \sum_j^l (b_j - b_j^0) \partial \lambda_i = 0. \quad (5.14)$$

The required thermodynamic data for the chemical potentials at the relevant pressures and temperatures must be determined through lab measurements.

This procedure is complicated by the inclusion of condensible species. Species must be tested to determine if it is favourable for them to occur in the condensed phase or remain in the gas phase, by determining which phase minimises the Gibbs free energy

$$\left(\frac{\mu_j^0}{RT} \right)_c - \left(\frac{\mu_j^0}{RT} + \log \frac{n_j P}{n} \right)_g < 0. \quad (5.15)$$

Here the subscripts c and g refer to the condensed and gas phases respectively. R is the ideal gas constant, where n is the number of moles of reaction products, while n_j is the number of moles of reactant j . Solving this minimisation problem can be achieved iteratively. In the CEA code, a Newton-Raphson solver is used. In practice, attention must be paid to parts of parameter space that can introduce numerical issues: phase triple points, condensing species, ions, and species with very low abundance.

5.4.1. Disequilibrium processes

Chemical disequilibrium can be thought of as any process by which the chemical abundances are shifted away from their LTE values. This can occur

when the timescales of other processes, such as convection or photochemistry become shorter than the chemical reaction timescales that drive the mixture towards equilibrium. Such disequilibrium processes have been used to explain observed chemical abundances in both brown dwarfs and exoplanets (e.g., [Fegley & Lodders, 1996](#); [Marois et al., 2008](#); [Mukherjee et al., 2024](#)) This highlights the challenges of calculating chemical abundances out of equilibrium: rather than simply needing the chemical potentials and minimising the Gibbs free energy, we now need to consider the forwards and backwards reaction rates of every reaction in a complex chemical network. Measuring these reaction rates requires careful laboratory work, especially at conditions far from standard atmospheric conditions on earth. Numerically calculating all of the reactions is also computationally intensive, though not intractable. However, there are simple cases which can provide reasonable approximations of a full disequilibrium model, particularly when the timescales of the rate-limiting reaction involved are known such as in the $\text{H}_2\text{O}-\text{CO}-\text{CH}_4$ sub-network [Prinn & Barshay \(1977\)](#); [Zahnle & Marley \(2014\)](#). Consider an atmosphere undergoing strong convection. This process mixes the atmosphere, dredging up material from deep in the atmosphere and lofting it upwards. The typical assumption made is that this transport of material can be modelled as a diffusion process, parameterised by the eddy diffusion coefficient K_{zz} , which describes the vertical mixing strength (e.g. [Prinn & Barshay, 1977](#); [Allen et al., 1981](#); [Zahnle & Marley, 2014](#); [Mukherjee et al., 2022](#)). The timescale associated with this mixing is the time it takes for material to be mixed over one scale height of the atmosphere, H . In hydrostatic equilibrium,

$$H = \frac{k_b T}{\mu g}, \quad (5.16)$$

for temperature T , mean molecular mass μ and gravity g . k_b is the Boltzmann constant. The eddy diffusion coefficient has dimensions $[\text{L}]^2/[\text{T}]$, so we can define a mixing timescale t_m as

$$t_m = \frac{H^2}{K_{zz}}. \quad (5.17)$$

The chemical timescale t_c will depend on the reactions under consideration, and will be a strong function of pressure and temperature. Generally, t_c will be shorter at higher temperatures and higher pressures. We can equate the two timescales to determine where in the atmosphere vertical mixing begins to dominate over the chemical reaction timescale, homogenising the chemical abundances above this altitude.

This conceptual picture faces difficulties when we consider the different structures and mixing processes in an atmosphere. There may be multiple convective regions, separated by radiative regions, all of which may have different mixing strengths. Even within a given region, there is no a priori reason why K_{zz} should be constant with altitude. Mukherjee et al. (2022) explores this in depth. They self-consistently determine K_{zz} profiles as a function of pressure for a variety of atmospheric temperatures and compositions. They show that properly accounting for variations in mixing changes not only the chemical abundances, but the thermal structure of the atmosphere, and finding that their models better fit the observations of Miles et al. (2020) than a constant K_{zz} model. Other theoretical studies explore how vertical mixing should vary with other parameters such as surface gravity (Moses et al., 2013), or metallicity (Soni & Acharyya, 2023). Lee et al. (2023, 2024) have even coupled a simple chemical kinetics model with a full global circulation model in order to determine accurate chemical abundances when atmospheric dynamics are accounted for.

Photochemistry

Vertical mixing is not the only source of nonequilibrium atmospheric chemistry. Stars emit high-energy radiation that can photodissociate molecules, which can then recombine in novel ways. For example, H_2S is the dominant sulphur species in equilibrium chemistry at typical temperatures of exoplanets (~ 1000 K). Recent observations have instead detected absorption from SO_2 , which is significantly less abundant in equilibrium (Tsai et al., 2023; Powell et al., 2024). Instead, the H_2S is oxidized into SO_2 , as the photodissociation of water by the stellar UV radiation leads to the production of OH radicals. Similar theoretical studies have shown the impact of both carbon and sulphur photochemistry on the atmospheres of planets under a variety of irradiation conditions (Zahnle et al., 2016; Tsai et al., 2021).

5.5. Emission spectra

With a model for the thermal structure and chemistry, calculating the emission spectrum becomes an exercise in radiative transfer. To zeroth order, all objects with a finite temperature emit radiation as a blackbody, following Planck's law

$$B_\lambda(\lambda, T) = \frac{2hc^2}{\lambda^5} \frac{1}{e^{hc/\lambda k_b T} - 1}, \quad (5.18)$$

for wavelength λ and temperature T . h is Planck's constant, c is the speed of light and k_b is the Boltzmann constant. The thermal structure of the at-

mosphere provides sources of radiation at a range of temperatures, while the chemistry determines which species contribute as opacity sources, creating spectroscopic features that result in deviations from blackbody emission. While absorption due to molecular features is the most important contributor to the spectral shape, atomic line absorption and emission, collisionally induced absorption (CIA), aerosol opacity, Rayleigh scattering and other effects must all be accounted for. In this section we will outline the physics of absorption and emission that shape the observed spectra of directly imaged exoplanets and brown dwarfs, following the review of [Marley & Robinson \(2015\)](#).

5.5.1. Radiative transfer

At this point, the goal is to calculate the absorption, emission, and scattering of light as it passes through the atmosphere of an exoplanet, using the tools of radiative transfer. [Buglia \(1986\)](#) extensively covers the topic, and I will use their approach, together with that of [Marley & Robinson \(2015\)](#) to assemble the necessary concepts for understanding radiative processes in a self-luminous atmosphere.

The transfer of energy through the atmosphere has two primary effects. The first, and most obvious result, is the emission of light from the top of the atmosphere. The second, but equally important result is the transfer of energy within the atmosphere itself. To simplify the problem, we can consider the case of a plane-parallel atmosphere. For a thin layer of the atmosphere, radiation can flow from the immediate layers above or below, and the layer under consideration can absorb or scatter the incoming light, or emit light itself. The net flux, F_{net} is thus simply the combination of the upwards flux, F^+ and the downward flux F^-

$$F_{\text{net}} = F^+ - F^- . \quad (5.19)$$

The net flux will either heat or cool the layer, depending on the sign. The net heating rate q is found from the gradient of the flux with pressure

$$q = \frac{g}{c_p} \frac{dF_{\text{net}}}{dP} . \quad (5.20)$$

To determine the thermal structure, we are so far concerned only with the bolometric flux, which determines the heating and cooling of the atmosphere.

To determine the emission spectrum however, we must consider the spectrally resolved equation of radiative transfer,

$$\mu \frac{dI_\nu}{d\tau_\nu} = I_\nu(\tau_{\nu\mu}, \mu, \phi) - S_\nu(\tau_{\nu\mu}, \mu, \phi) . \quad (5.21)$$

The primary quantity of interest here is I_ν , the spectral radiance. The coordinate system we use here is in units of τ_ν , the frequency dependent optical depth in the vertical direction, can be found from the opacity k_ν and density of the absorber, $d\tau_\nu = -k_\nu \rho dz$. S_ν is the source function, representing the total energy (per frequency) added to the beam. μ is the zenith angle, and ϕ is the azimuthal angle; the atmosphere is plane-parallel but light can enter a layer at an angle and be scattered in a different direction.

The source function describes all of the sources of radiation in the atmosphere. This can include emission of thermal radiation from the atmosphere itself, incident stellar radiation, and scattering of light into the beam. These are the first, second and third terms on the right hand side of equation 5.22 respectively. We can write a general form of S_ν as

$$S_\nu(\tau_\nu, \mu, \phi) = (1 - \omega_\nu)B_\nu(T(\tau_\nu)) \quad (5.22)$$

$$+ \frac{\omega_\nu}{4\pi} \left(F_\nu^* e^{-\tau/\mu_*} \right) p_\nu(\tau_{nu}, \mu, \phi, -\mu_*, \phi_*) \quad (5.23)$$

$$+ \frac{\omega_\nu}{4\pi} \int_0^{2\pi} \int_{-1}^1 I_\nu(\tau_{nu}, \mu, \phi) p_\nu(\tau_{nu}, \mu, \phi, -\mu, \phi) d\mu' d\phi'. \quad (5.24)$$

ω_ν is frequency-dependant single scattering albedo. In the first term B_ν is the Planck function, where $T(\tau_\nu)$ is the atmospheric temperature profile. This term defines the radiation emitted by the atmosphere into the beam. In the second term, F_ν^* is the incident stellar flux, and p_ν is the scattering phase function, determining the flux of star light from the top of the atmosphere down. The last term is the scattering term from angles μ' , ϕ' into the beam at an angle μ , ϕ . With the source term defined, it is now possible to solve the integro-differential equation 5.21, though boundary conditions at the top and bottom of the atmosphere are still required. One example of a bottom boundary condition is supplied by [Marley & Robinson \(2015\)](#), originally from [Mihalas \(1970\)](#):

$$I_\nu(\tau_{nu,bot}, \mu, \phi) = B_\nu(T(\tau_{nu,bot})) + \mu \frac{dB_\nu}{d\tau_\nu} \Big|_{\tau_{nu,bot}}, \quad (5.25)$$

where the gradient term allows some flux from deeper in the interior.

At the top of the atmosphere we must consider both the upwelling flux F_ν^+ and the downwelling, F_ν^- . These are simply found by appropriately integrating over the angle dependence of I_ν :

$$F_\nu^+ = \int_0^{2\pi} \int_0^1 I_\nu(\tau_{nu}, \mu, \phi) \mu d\mu d\phi, \quad (5.26)$$

$$F_\nu^- = \int_0^{2\pi} \int_{-1}^0 I_\nu(\tau_{nu}, \mu, \phi) \mu d\mu d\phi. \quad (5.27)$$

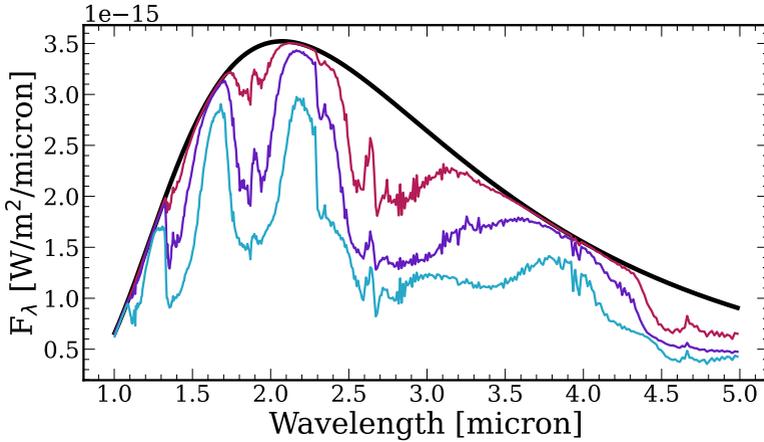


Figure 5.4: Impact of molecular absorption from H_2O and CO in an atmosphere with a non-inverted thermal profile, for varying quantities of each species.

We will explore how to numerically solve the RTE in chapter II.

One key insight to take from this introduction are that the emission spectrum depends heavily on the temperature gradient within the atmosphere. In optically thick regions deep in the atmosphere, the upwelling and downwelling fluxes are, to first order, equal. By the boundary conditions of a self-luminous atmosphere, the downwelling flux is 0 at the top of the atmosphere, and only the upwelling flux contributes, passing through cooler regions of gas which impart absorption features onto the emission from deeper regions. This implies that the observed absorption features provide information about the atmospheric temperature structure. Indeed, the shape of the features is sensitive to the temperature difference between layers. Wavelengths with strong opacity become optically thick at lower pressures, while wavelengths with weak opacities are more transparent, providing a window to deeper, hotter regions of the atmosphere (if there are no thermal inversions present). This is key to interpreting atmospheric features, providing information on the abundance of an absorber, its location in the atmosphere, and the temperature structure of the atmosphere itself.

5.5.2. Absorption and emission lines

Absorption and emission lines are the result of transitions in the energy states of molecules. At the temperatures involved in exoplanet atmospheres,

these are typically transitions in the coupled rotational-vibrational states of molecules and atoms. The energy of the transitions determines the frequency of the emitted or absorbed light. Transitions in only the vibrational state, with a constant rotational state are referred to as Q branch transitions, while the R and P branches occur for $\delta J = +1$ and $\delta J = -1$ respectively, where J is the rotational quantum number. The strength of the line depends on the number of molecules populating the initial and final energy states.

Consider a line with a central position ν_0 , strength $S = S(T)$ and line shape function $f(\nu - \nu_0)$. The opacity k_ν is then

$$k_\nu = S f(\nu - \nu_0), \quad (5.28)$$

normalised such that

$$\int_{-\infty}^{\infty} f(\nu - \nu_0) d\nu = 1. \quad (5.29)$$

The line strength depends on the number of molecules populating the initial state that transition to the final state E'' , which depends on the temperature

$$S(T) = S_0 \frac{Q(T_0)}{Q(T)} \exp \left[\frac{E''}{k_b} \left(\frac{1}{T_0} - \frac{1}{T} \right) \right] \frac{1 - \exp(-hc\nu_0/k_b T)}{1 - \exp(-hc\nu_0/k_b T_0)}. \quad (5.30)$$

Here, Q is the internal partition function for the rotational and vibrational states, which can be determined through laboratory measurements or via ab initio quantum chemical calculations.

It is important to remember that the actual structure of the observed absorption features is built from the superposition of many lines. Even for a single transition, there will be different contributions from different regions of the atmosphere. The result of this is that we observe absorption features if the temperature gradient is positive, i.e. the temperature increases with pressure, and emission features if the temperature gradient is inverted (increasing temperature with decreasing pressure). The energy of every photon absorbed by a molecule will ultimately be reemitted at a later time, though in a random direction. With an isothermal temperature structure of an atmosphere in LTE, this absorption and reemission would be equal throughout the atmosphere, due to Kirchoff's law, and there would be no spectroscopic features visible. With a temperature gradient however, spectroscopic features will emerge. If colder gas lies above warmer gas, light emitted at the wavelength of the absorption features, but only part of the light will be reemitted upward, reducing the flux in this region, and forming an absorption feature. Conversely, with hotter gas lying above colder gas, more light will be emitted due to the higher thermal energy, resulting in an emission feature.

Line broadening

Absorption lines are broadened by a variety of processes, all of which contribute to the overall shape of an observed spectrum. Due to Heisenberg's uncertainty principle $\Delta E \Delta t \gtrsim \hbar/2$, even well-defined lines have a finite width determined by the lifetime of the transition, known as the natural broadening of the line. If the lifetime of the transition from the higher state to the lower state is given by the decay rate γ , the line profile f_N follows a Lorentzian shape:

$$f_N(\nu - \nu_0) = \frac{\gamma/4\pi^2}{(\nu - \nu_0)^2 + (\gamma/4\pi^2)^2}. \quad (5.31)$$

This width is typically very narrow compared to other sources of line broadening.

As the molecules of the gas move back and forth relative to the observer due to thermal motion, this will introduce Doppler shifts. This results in the line shape f_D

$$f_D(\nu - \nu_0) = \frac{1}{\alpha_D} \sqrt{\frac{\ln 2}{\pi}} \exp\left(-\frac{\ln 2(\nu - \nu_0)^2}{\alpha_D^2}\right), \quad (5.32)$$

where α_D is the Half Width Half Maximum (HWHM) induced by the Doppler motion for temperature T and molecular mass m

$$\alpha_D = \frac{\nu_0}{c} \sqrt{2 \ln 2 \frac{k_b T}{m}}. \quad (5.33)$$

Pressure broadening is an important consideration, particularly due to the range of pressures present in an atmosphere. This is typically modelled as a pressure dependent Lorentzian profile, f_L

$$f_L(\nu - \nu_0) = \frac{1}{\pi} \frac{\alpha_L}{(\nu - \nu_0)^2 + \alpha_L^2}, \quad (5.34)$$

where the Lorentzian HWHM is

$$\alpha_L = \gamma P \left(\frac{T_0}{T}\right)^n. \quad (5.35)$$

for a width parameter γ , and temperature dependence parameter n , both associated with a molecular line list. Figure 5.4 demonstrates molecular absorption for a simple atmosphere model. In this toy model we include no additional sources of opacity and only four pressure levels, with the temperature decreasing with

decreasing pressure. Due to the low spectral resolution and high density of lines for H₂O and CO, the individual lines are generally not observed, but rather smear together to form broad features that define the characteristic shapes of near infrared spectra.

Realistic line shapes must account for broadening from various sources. The Voigt profile f_V is the result of the convolution of Doppler and pressure broadening

$$f_V(\nu - \nu_0) = \frac{\alpha_L}{\pi^{3/2}} \int_{-\infty}^{\infty} \frac{\exp(-y^2)}{(\nu - \nu_0 - \alpha_D y)^2 + \alpha_L^2} dy. \quad (5.36)$$

Strong lines, such as the sodium doublet, require specialised treatment to model their wing profiles. Such work has been discussed in [Allard et al. \(2016, 2019\)](#) for sodium and potassium.

Opacity databases

Given the complexity of both measuring and calculating molecular line lists for the nearly infinite array of chemical species, various databases have been developed to provide these line lists. These databases store line lists in various formats, compiling lab measurements from a variety of sources, as well as providing comparisons between lab measurements and quantum chemical calculations. The most widely used databases in the study of exoplanet atmospheres are Exomol [Tennyson & Yurchenko \(2012\)](#) and HITRAN/HITEMP [Gordon et al. \(2022\)](#), which has a history dating back over 50 years ([McClatchey, 1973](#)).

5.5.3. Continuum gas opacities

In addition to molecular line absorption, other processes can result in the absorption and scattering of light. These states typically have broad features compared to the narrow molecular transitions.

Collisionally induced absorption

CIA is induced by the inelastic scatter of two molecules in a gas, which can, at close separation, be regarded as a single molecule ([Hartmann et al., 2018](#)). These collisions occur on very short timescales, so by the Heisenberg uncertainty principle the energy spectrum is very broad. In the atmospheres of directly imaged planets, the two primary constituents are H₂ and He, so H₂-H₂ and H₂-He must be accounted for ([Richard et al., 2012](#); [Karman et al.,](#)

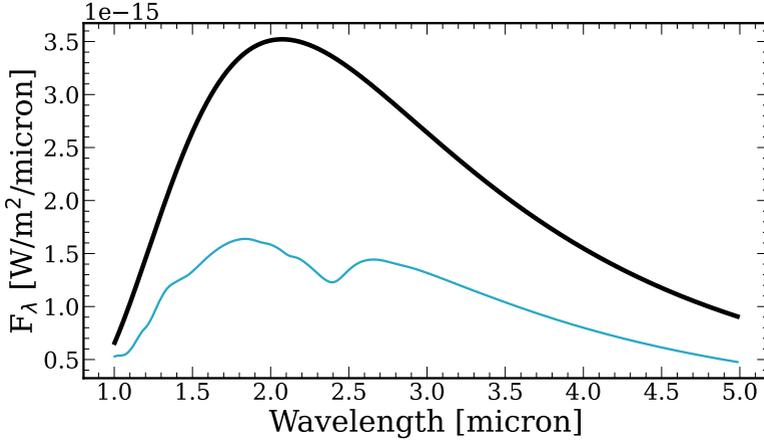


Figure 5.5: Impact of $\text{H}_2\text{-H}_2$ and $\text{H}_2\text{-He}$ CIA on a simple atmosphere with no molecular absorption.

2019). He–He collisions are already sufficiently rare to not significantly impact the spectral shape, and remaining CIA absorption is negligible.

Bound-Free

Bound-free opacity sources are caused by the ionisation of a bound electron by an incident photon. These transitions typically require either high temperatures or another source of high energy photons, transferring $h\nu - I$ of energy from the photon to the electron in the process, where I is the binding energy of the electron. The photon scattering cross section is 0 below the ionisation threshold, then rapidly reaches a maximum at the ionisation energy, before slowly decaying with increasing photon energy. Chapter 23 of [Shu \(1991\)](#) provides a derivation of the cross section for a for an electron with a primary quantum number n around a hydrogen-like atom with nuclear charge Z ,

$$\sigma_{bf} = n^{-5} \frac{8\pi}{3\sqrt{3}} \frac{Z^4 m_e e^{10}}{\hbar^3 (h\nu)^3} g_{bf}(\nu); h\nu > I, \quad (5.37)$$

where m_e and e are the electron mass and charge, h and \hbar are Planck's constant and the reduced Planck's constant, and g_{bf} is a slowly varying, order-unity Gaunt factor that corrects for the breakdown of the Born approximation near the ionisation threshold.

Due to the energies required for ionisation, this opacity source is more common either deep within the atmosphere of a self-luminous object, or in the upper atmosphere of a highly irradiated exoplanet. Nevertheless, it can contribute as a significant opacity source, particularly from H and H^- . As there are a variety of ionisation states, bound-free provides a continuum scattering source. Approximating this as entirely frequency dependent, the opacity can be estimated as

$$\bar{\kappa}_{bf} = 4.34 \times 10^{25} \frac{g_{bf}}{t} Z(1+X) \frac{\rho}{\text{g/cm}^3} \left(\frac{T}{\text{K}}\right)^{-7/2} \text{cm}^2 \text{g}^{-1}, \quad (5.38)$$

where g_{bf} is the Gaunt factor for bound-free transitions and t is an additional correction factor. In the classical limit $g_{bf} = 1$ and Z and X are the mass fractions of hydrogen and metals respectively.

Free-Free

Free-free transitions occur during the inelastic scattering of an electron or photon. In the case where a free electron scatters off a charged part a photon is emitted via bremsstrahlung. If a photon scatters off a charged particle, some of the photon energy can be absorbed by the particle. Photons can also scatter inelastically off atoms, exciting an electron into a higher energy state. The strength of the free-free absorption cross section is derived in Chapter 14 of [Shu \(1991\)](#) by determining the total rate of emission of a thermal bath of electrons and equating that to the luminosity through Kirchoff's law to find the specific emissivity j_ν for opacity κ_ν :

$$j_\nu = 4\pi\kappa_\nu B_\nu(T). \quad (5.39)$$

Equating to find κ_ν ,

$$\rho\kappa_\nu^{ff} = \sum_i n(Z_i)n_e \left(\frac{2m_e}{3\pi k_b T}\right)^{1/2} \left(\frac{4\pi Z_i^2 e^6}{3m_e^2 ch\nu}\right) g_{ff}(\nu) (1 - e^{h\nu/k_b T}), \quad (5.40)$$

where ρ is the density, $Z_i e$ is the ion charge, $n(Z_i)$ is the ion number density and n_e is the electron number density. g_{ff} is the Gaunt factor for free-free scattering. The remaining variables are as in equation 5.37.

Due to their abundance, the important species to consider in self-luminous atmospheres are H, H_2 , H_2^- and H^- . As with the bound-free opacity, the free-free opacity can be estimated using a grey approximation as

$$\bar{\kappa}_{ff} = 3.68 \times 10^{22} g_{ff}(1-Z)(1+X) \frac{\rho}{\text{g/cm}^3} \left(\frac{T}{\text{K}}\right)^{-7/2} \text{cm}^2 \text{g}^{-1}. \quad (5.41)$$

5.5.4. Clouds and condensation

The final important opacity source in an atmosphere is that of the clouds. Like gas continuum opacities, clouds tend to produce broad spectral features rather than sharp lines, though some cloud compositions do induce identifiable spectral features. When we refer to clouds, there are several different constituents that we may be referring to, that all fall under the broad umbrella of aerosols. The first is clouds produced similarly to water clouds on earth, that is by adiabatic cooling and condensation. In reality, such clouds are far more complicated: convection is not a requirement for condensation, and condensate species such as MgSiO_3 have no corresponding gas phase species, but form through surface reactions once conditions are thermodynamically favourable. If the particles are produced via photochemistry, they are usually referred to as hazes. On terrestrial planets, surface dust or salt spray from the oceans can loft additional aerosols into the atmosphere. Regardless of the mechanism, all of these aerosols add sources of continuum absorption and scattering opacity to the atmosphere, and for simplicity I will use the terms clouds and aerosols interchangeably. A full discussion of cloud physics in sub-stellar atmospheres is beyond the scope of this work, and we refer to the reviews of [Helling et al. \(2008b\)](#); [Helling \(2019\)](#).

The red colour of mid to late L-dwarfs has been attributed to clouds since their initial discovery. The early models developed by [Ackerman & Marley \(2001\)](#) and [Helling et al. \(2001\)](#); [Helling & Woitke \(2006\)](#) underpin the state-of-the-art today. These models follow two distinct paradigms for the cloud formation process. In the [Ackerman & Marley \(2001\)](#) picture (hereafter **AM01**), clouds form when the pressure and temperature of the atmosphere are sufficiently cool for a species to condense out of the gas form. Due to vertical mixing, parameterised by K_{zz} , the cloud particles are lofted upward into the atmosphere. A sedimentation fraction parameter f_{sed} is introduced, which determines the settling strength of the cloud particles, competing with the vertical mixing. The balance between the vertical mixing and sedimentation determine the mean particle size in each layer of the atmosphere above the condensation point, as larger particles settle faster than smaller particles. A log-normal distribution is typically used to describe the width of the particle size distribution about the mean. Thus the total mole fraction of species $q_t = q_v + q_c$ (the sum of the vapour and condensate mole fractions) can be found by solving the differential equation

$$K_{zz} \frac{\partial q_t}{\partial z} - f_{\text{sed}} w_* q_c = 0, \quad (5.42)$$

where w_* is the mass-weighted sedimentation velocity. The impact of this model

on a simple atmosphere is shown in Figure 5.6, highlighting the strong and broad opacity of the clouds.

In contrast to the AM01, the Helling & Woitke (2006) approach uses a parameterised microphysical approach, forming clouds at the top of the atmosphere before they settle downward. Small condensation nuclei are lofted high into the atmosphere, which super-saturated condensable species can nucleate onto. On rapid timescales these initial cloud seeds will grow, forming larger cloud particles with heterogeneous composition. As the particles grow, they will sink deeper into the atmosphere, where the change in temperature will process the grains. At sufficient temperatures, the grains will evaporate back into the gas phase and the cycle can repeat.

Both of these models rely on the condensation of species from the gas phase into solids or liquids, relying on known condensation curves (e.g. Visscher et al., 2010). Upon condensation, the species are removed from the gas phase, changing the bulk composition and spectral features. The formation of cloud particles will also be limited by the least available component. These compositional changes will lead to gas phase composition that are not indicative of the bulk composition, leading to changes of 15%–30% in the observed C/O ratio (Burrows & Sharp, 1999; Fonte et al., 2023; Calamari et al., 2024).

Barstow (2020) performed a systematic comparison of these parametric models for transiting exoplanet to determine how cloud modelling choices impact parameter inference. Fortunately, the resulting models were in good qualitative agreement, and atmospheric parameters such as the H₂O abundance were robust to the cloud model choice. On the other hand, this demonstrates the difficulty in directly measuring the cloud properties themselves and in testing different cloud models.

However, even these cloud models represent a simplification of the complex processes that underpin condensation and grain growth. Gao et al. (2020) and Powell et al. (2019); Powell & Zhang (2024) use the Community Aerosol and Radiation Model for Atmospheres (CARMA) microphysical model to explore what compositions can be expected in exoplanet atmospheres, and how they should be distributed throughout the atmosphere. While this work focuses on hot-Jupiter planets rather than self-luminous objects, it highlights the complexity of cloud formation. Due to differences in the surface energies, only some cloud species are likely to nucleate, while other have a strong energy barrier to nucleation. Thus they find only certain compositions to be likely at a given temperature. For example, at L-dwarf temperature ranges Mg₂SiO₄ is likely to be the dominant source of aerosol opacity. Not only does CARMA explore the nucleation and surface chemistry during the cloud formation process, but determines the particle size distribution throughout the clouds. This is one of

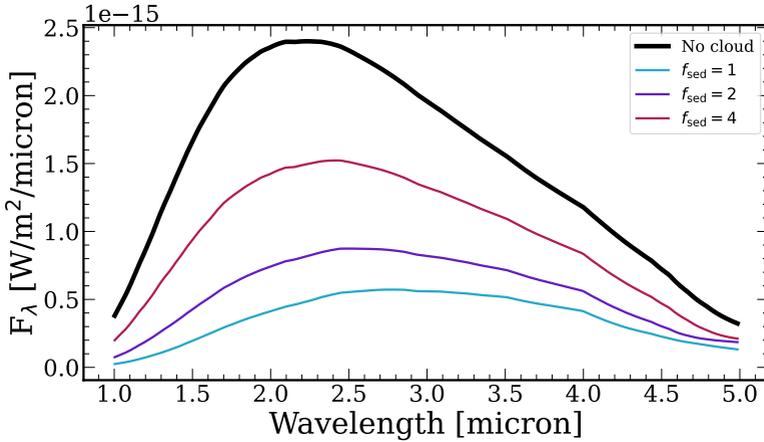


Figure 5.6: Impact of a silicate cloud on an emission spectrum for different f_{sed} , following the model of [AM01](#).

the major differences between the microphysical models, which predict complex, bimodal particle size distributions, and parametric models, which often use a broad log-normal distribution. As the cloud optical properties depend strongly on the particle size, accurately modelling this distribution is critical to correctly modelling cloud absorption and scattering.

In addition to the distribution of clouds throughout the atmosphere, their optical properties must also be considered in order to determine their impact on observed spectra. Cloud particles are of approximately the same size as the wavelengths of light observed in the near infrared, and thus scatter light in the Mie regime. Mie scattering assumes that the particles are spherical, which may or may not be a reasonable approximation. The distribution of hollow spheres (DHS) model ([Min et al., 2005](#)) extends this scattering approach to non-spherical particles. To determine the scattering amplitude and direction for Mie scattering, the Maxwell equations must be solved in the presence of a spherical particle with permittivity ϵ_1 and permeability μ_1 , surrounded by a medium with ϵ and permeability μ , subject to boundary conditions at the interface between the particle and the environment, boundedness at the coordinate origin, and that the wave asymptotes at infinity to a diverging spherical wave. This is a classic optics problem, discussed in texts such as [Bohren & Huffman \(1998\)](#). A full derivation of the solution is beyond the scope of this work, but there are a few key insights to keep in mind from the solutions. In the limit that the

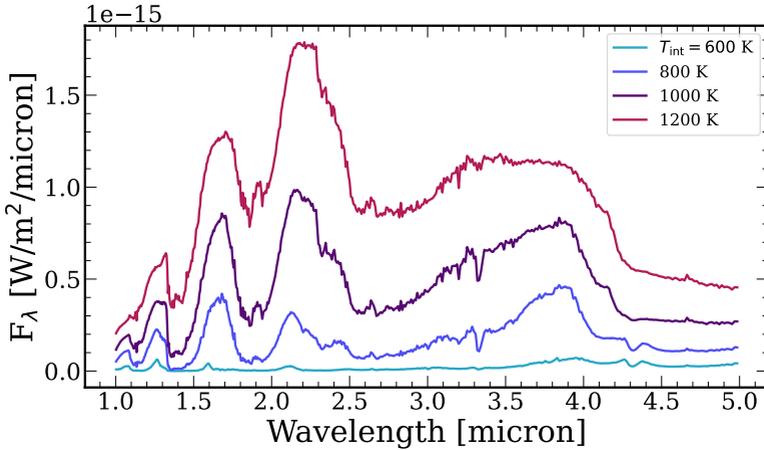


Figure 5.7: Complete emission spectral model of exoplanet atmospheres. This example incorporates the impact of molecular absorption, CIA opacity and cloud absorption and scattering for an objects with internal temperatures between 600 K and 1200 K, or roughly covering spectral types from the mid T-dwarfs to early L-dwarfs.

particle is much smaller than the wavelength of light, the solution reduces to that of Rayleigh scattering, which is proportional to λ^{-4} . If the particle size and wavelength are of similar length, then in the Mie scattering regime forward scattering is preferred, with larger particles forward scattering more strongly.

5.5.5. Summary

Having developed a model that incorporates all of the different sources of opacity and follows a physically motivated temperature profile, we can put all of the pieces together and calculate the emission spectrum of an exoplanet atmosphere using `petitRADTRANS` (pRT) (Mollière et al., 2019), with the results shown in Figure 5.7. The spectra computed for this example use the parameterised pressure-temperature profile of Guillot (2010), which for our parameter choice approximates an adiabat deep in the atmosphere with a radiative zone on top. The chemistry is determined using a table of pre-computed equilibrium abundances for each pressure and temperature point, with additional axes of metallicity ($[M/H]=0.5$) and C/O ($C/O=0.55$). We included silicate and iron clouds using the AM01 scheme. The point of this exercise is to emphasise how

all of the different atmospheric parameters interact with each other to create the complex observed spectra. This highlights the two different approaches to modelling these atmospheres. There are physics-based models, which incorporate as many processes as possible to the best extent of our knowledge in order to predict from a priori calculations what the observed spectra should be. This is known as the forward modelling approach, and includes both self-consistent 1D models and global circulation models, which we will discuss briefly in Sections 5.6 and 5.7 respectively. The second approach is to rely on the observations, and optimise model parameters to fit the data in order to develop a physical understanding of the mechanisms driving the spectral features. This approach is the topic of the remainder of this thesis.

5.6. Self-consistent models

Unfortunately, it is not possible to simulate in 3D the full complexity of all of the physical processes present in an atmosphere. With scales of micrometers in the condensation nuclei of clouds to global circulation patterns, and timescales ranging from rapid chemical reactions to slow variation in incident stellar radiation over time, a complete atmospheric model is currently infeasible. For all classes of models then, simplifications must be made.

A widely used class of models used to study brown dwarf and directly imaged exoplanet atmospheres is 1D, radiative-convective equilibrium (RCE) models. Such models which solve for the atmospheric structure in RCE and simultaneously apply some kind of chemical modelling are usually called self-consistent. From observations the full disk of the objects appear as point sources, integrating over any variation in their surface features. While dynamical processes are important, they are secondary to the bulk planet temperature and composition. In comparisons to observed spectra, models generated using the 1D approximation reproduce the data with a high degree of accuracy. This is a limited approximation: with sufficiently precise data, a variable surface (such as seen for Jupiter in Figure 5.1) produces a disk-integrated spectra that cannot be fully captured by a single 1D model (e.g. [Vos et al., 2023](#)).

The advantage of such models is that very few parameters are required as inputs, and these are often directly related to observables. The effective temperature T_{eff} , surface gravity $\log g$ and radius R are universally required. Different models approach clouds with different methods. Some, such as the recent Sonora Bobcat [Marley et al. \(2021\)](#) models neglect condensation entirely, focusing on reproducing objects where condensation is expected to play less of a role in shaping the spectrum. Conversely, the ATMO models [Phillips et al. \(2020\)](#) do not include clouds, but incorporate a parameter γ_{ad} that adjusts the

slope of the adiabatic temperature profile, which can act to produce a more isothermal temperature profile, reducing the amplitude of spectral features and generally reddening the spectrum. This reduced temperature gradient allows for good fits to L-dwarf near-infrared spectra without the need for incorporating condensation. The reddening of the near-infrared is generally interpreted as evidence for clouds (e.g., [Cushing et al., 2008](#); [Faherty et al., 2016](#); [Charnay et al., 2018](#)). Different cloud prescriptions, often based on either [Ackerman & Marley \(2001\)](#) or [Helling & Woitke \(2006\)](#) are incorporated into many models, with the strength of the cloudiness parameterised by f_{sed} or a similar parameter. There is a vast array of available models, each incorporating slightly different implementations of physical processes, using different line lists, and covering different regions of parameter space. These include ATMO ([Tremblin et al., 2016](#); [Phillips et al., 2020](#); [Petrus et al., 2023](#)), BT-Sett1 ([Allard et al., 2003](#); [Allard et al., 2012](#)), DRIFT-PHOENIX ([Hauschildt & Baron, 1999](#); [Helling et al., 2008a](#)), EGP ([Burrows et al., 1996](#); [Burrows & Volobuyev, 2003](#); [Marley et al., 2006](#); [Saumon & Marley, 2008](#)) and its successor, the Sonora series ([Marley et al., 2021](#); [Karalidi et al., 2021](#); [Morley et al., 2024](#); [Mukherjee et al., 2024](#)), Exo-REM ([Baudino et al., 2015](#); [Charnay et al., 2018](#)) which are computed using PICASO ([Mukherjee et al., 2023](#)), and pet itCODE ([Mollière et al., 2015, 2017](#)).

5.7. Global circulation models

Global circulation models (GCMs) represent a different approach to modelling atmospheres. Rather than assuming 1D atmosphere, GCMs model the complex hydrodynamics of a 3D atmosphere. Assumptions are required in other aspects to ensure computational feasibility. Typically only the primitive equations are solved, rather than the full system of Navier-Stokes equations; as phrased by [Showman & Kaspi \(2013\)](#), the GCM ‘assumes that dynamics introduces only small perturbations of the density, entropy, and pressure from a specified reference state, which we here take to be isentropic.’ Radiative transfer is often simplified to a grey or double-grey model, rather than performing the computations at high spectral resolution. Some form of a simplified chemical model is typically used, and clouds were often only post-processed into the results, rather than being included self-consistently. State-of-the-art GCMs such as [Tan & Showman \(2021a,b\)](#) and [Lee et al. \(2023, 2024\)](#) have made great strides in incorporating dynamically active clouds and chemistry, as well as using higher resolution radiative transfer. The advantage of GCMs is that they produce spatially resolved, time variable models of an atmosphere. This allows for the exploration of the complex dynamical phenomena present in a real atmosphere, from equatorial jets to zonal winds. Post-processing the GCM outputs

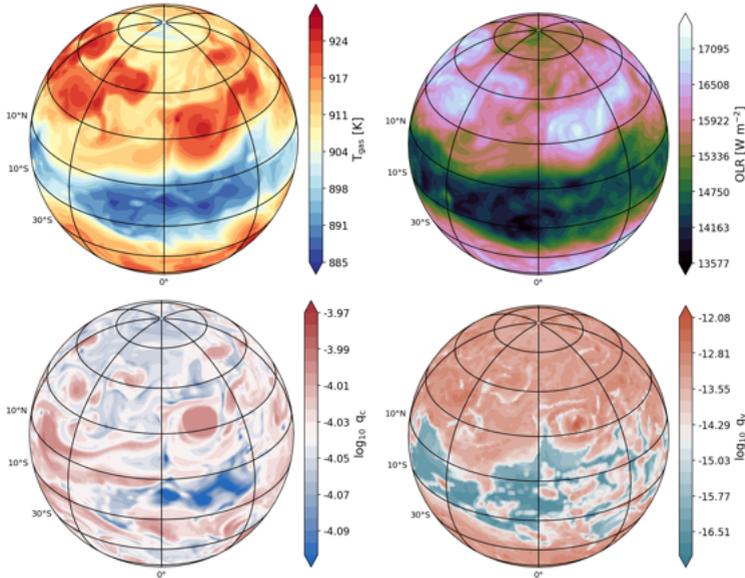


Figure 5.8: GCM results for a rapidly rotating sub-Jupiter mass object at 0.1 bar, detailing the spatial variation in temperature, outgoing longwave radiation, condensate mass fraction and condensate vapour fraction. Reproduced with permission from [Lee et al. \(2024\)](#).

to produce hemisphere-averaged emission spectra allows for the investigation of variability, with the ability to link the variable emission directly to the physical phenomena driving the changing flux. This is key to explaining recent studies of variable brown dwarfs ([Buenzli et al., 2012](#); [Vos et al., 2018](#); [Zhou et al., 2020](#), e.g.), and will be necessary in the near future for exploring the variability of directly imaged exoplanets. Figure 5.8 shows a recent GCM model of a brown dwarf from [Lee et al. \(2024\)](#), showcasing how the atmospheric dynamics and cloud properties vary spatially over the surface, ultimately leading to changes in the observed flux.

5.8. The inverse problem

With a thorough understanding of the thermal structure, composition, dynamics, and cloud distribution of an atmosphere, it is relatively straightforward to numerically perform the radiative transfer calculations required to compute

an emission spectrum. In astronomy we are faced with the inverse problem: we can observe the spectrum of an astrophysical object, and then need to determine the physical processes that produce it. This is a common problem across disciplines, with well-studied methods. While the fundamentals of atmospheric physics are reasonably well-understood, backing out the atmospheric state from an observed spectrum is non-trivial. Such inverse problems are often ill-posed, without a guarantee of the existence, uniqueness or stability of a solution. In the study of exoplanet atmospheres, techniques are borrowed from planetary science and earth observation, such as those presented in [Rodgers \(2000\)](#). If a Bayesian statistical framework is used to infer the atmospheric state from the observed spectrum, we refer to this procedure as an *atmospheric retrieval*. Developing an atmospheric retrieval framework and applying it to emission spectra of directly imaged planets is the primary focus of this work.

6. OUTLINE

At this point, all of the tools necessary to understand this thesis have been introduced. I will make use of high-contrast imaging data of young, giant exoplanets in order to characterise their atmospheric properties. Their atmospheres inherit their composition from the disk in which they form. While it is still unclear the exact mechanism by which they form it is hoped that improving our understanding of their atmospheres will provide insight into their formation and evolutionary history.

The goal of this thesis is to develop and apply such an atmospheric retrieval framework to determine the atmospheric state of exoplanets. In Chapter II I outline the retrieval package, and apply it in a case study for a well-studied transiting planet, WASP-39 b as part of the *JWST* Transiting Early Release Science program (Nasedkin et al., 2024). Chapter III examines the impact of data analysis choices on our ability to infer atmospheric parameters (Nasedkin et al., 2023). Finally, in Chapter IV I apply the retrieval and data analysis tools to the benchmark system of HR 8799 to systematically characterise the atmospheres of the four companions (Nasedkin et al., 2024).



ATMOSPHERIC RETRIEVALS WITH PETITRADTRANS

This chapter has been adapted from [Nasedkin et al. \(2024\)](#) and Welbanks et. al (2024, in prep.). I am the first author of [Nasedkin et al. \(2024\)](#) and the primary developer of the `petitRADTRANS` retrieval module described in this chapter. I have contributed a set of retrievals of WASP-39 b to the *JWST* Early Release Science program for a retrieval comparison study led by Luis Welbanks. The pRT retrievals are discussed in this chapter.

“Nature is written in that great book which ever is before our eyes, but we cannot understand it if we do not first learn the language in which it is written.”

Galileo Galilei, 1623



7. PETITRADTRANS

`petitRADTRANS` (pRT) is a fast radiative transfer code used for computing emission and transmission spectra of exoplanet atmospheres (Mollière et al., 2019), combining a Fortran back end with a Python-based user interface. It is widely used in the exoplanet community with 260 refereed citations in the literature at the time of writing and has been benchmarked against numerous similar tools, including many listed in MacDonald & Batalha (2023). The spectra calculated with pRT can be used as a forward model for fitting spectroscopic data using Monte Carlo techniques, commonly referred to as an atmospheric retrieval Madhusudhan & Seager (2009).

Atmospheric retrievals are a cornerstone of exoplanet atmospheric characterisation. pRT provides a powerful and user-friendly tool for researchers to fit exoplanet spectra with a range of built-in or custom atmospheric models. Various thermal structures, chemistry and cloud parameterisations, and opacity calculation methods can be combined and used to perform parameter estimation and model comparison for a given atmospheric spectrum. The Retrieval module combines the Radtrans forward modelling class with a nested sampler (Skilling, 2004) via a likelihood function to perform an atmospheric retrieval. The new retrieval module combines fast forward modelling with nested sampling codes, allowing for atmospheric retrievals on a large range of different types of exoplanet data. In this chapter I will outline the implementation details the spectroscopic calculations as originally published in Mollière et al. (2019, 2020)¹. The novel Retrieval module implementation is presented in Section 8. Finally in Section 9 I will present an application of the retrieval technique to the hot Jupiter WASP-39b.

7.1. Radiative transfer

In Section 5.5 the fundamental problem of radiative transfer was stated. Crucially, pRT is not a self-consistent model, and so the temperature structure is parameterised independently of the radiation transfer, and so it is not necessary to calculate the heating and cooling in the atmosphere due to radiation. This simplifies the problem to computing the top-of-atmosphere emission spectrum. While pRT can calculate spectra in both transmission and emission geometries, we will focus on the emission spectrum implementation here. There are two

¹For detailed descriptions of the actual use of pRT, links to the publicly available source code, and documentation the complete available functionality, see <https://petitradtrans.readthedocs.io>

methods implemented for solving the radiative transfer in an atmosphere divided into discrete pressure layers, with a temperature assigned to each layer. The first method uses a simplified statement of the radiative transfer problem, neglecting the scattering term from equation 5.21. This approximation significantly reduces the computation time, and is valid when redder-than-optical wavelengths are considered (where Rayleigh scattering is negligible) and when clouds are not a significant source of opacity in the atmosphere. The second implementation uses the method of [Feautrier \(1964\)](#), and iteratively solves for the scattering source function. It is this second method that will be described here, though the full implementation is detailed in [Mollière \(2017\)](#).

The fundamental coordinate that we will use for solving the radiative transfer equation is the frequency-dependent optical depth, τ_ν .

$$\tau_\nu = \int \kappa_\nu \rho dz. \quad (7.1)$$

Recalling that in Section 5.21 we had defined $\alpha_\nu = \kappa_\nu \rho$ for opacity κ_ν , we can rewrite τ_ν in the vertical direction as an integral of the density ρ and pressure P

$$\tau_\nu^{\text{vert}} = \int \frac{\kappa_\nu}{g} dP, \quad (7.2)$$

which we generalise to an arbitrary zenith angle θ , where $\mu = \cos \theta$,

$$\tau_\nu(\theta) = \frac{\tau_\nu^{\text{vert}}}{\mu}. \quad (7.3)$$

Consider now the differential equations for the inward I^+ and outward I^- pointing intensity, for an arbitrary source function S .

$$\frac{dI^+}{d\tau} = S - I^+, \quad (7.4)$$

$$\frac{dI^-}{d\tau} = -(S - I^-). \quad (7.5)$$

We can rewrite these to simplify the equations, defining I_J and I_H as a linear combination of I^+ and I^- such that

$$I_J = \frac{1}{2}(I^+ + I^-), \quad (7.6)$$

$$I_H = \frac{1}{2}(I^+ - I^-). \quad (7.7)$$

With these definitions, equation 7.4 becomes

$$\frac{dI_J}{d\tau} = -I_H, \quad (7.8)$$

$$\frac{dI_H}{d\tau} = S - I_J. \quad (7.9)$$

Taking the second derivative with respect to τ ,

$$\frac{d^2 I_J}{d\tau^2} = I_J - S. \quad (7.10)$$

This is a diffusion equation, which automatically provides a high-enough order to correctly solve for the intensity in the diffusive limit when discretised for numerical applications.

At this point we need to specify the boundary conditions on the problem. We will consider only the planetary radiation field, as the radiation of any incident light can be linearly combined with the planetary solution. Thus at the top of the atmosphere, there should be no radiation entering the planet atmosphere, and therefore

$$I^+(P = 0, \theta) = 0 \forall \theta. \quad (7.11)$$

At the bottom of the atmosphere the radiation field is assumed to be diffusive, and therefore

$$I^+(P \rightarrow \infty, \theta) = I^-(P \rightarrow \infty, \theta) \cdot \forall \theta. \quad (7.12)$$

Rephrasing these boundary conditions for I_J , we find

$$\frac{dI_J(0)}{d\tau} = I_J(0), \quad (7.13)$$

$$\frac{dI_J(P \rightarrow \infty)}{d\tau} = 0, \quad (7.14)$$

In order to solve these equations, we must discretise them. In pRT, we use a log-spaced pressure grid of N layers. Rewriting equation equation 7.10 in discrete form and dropping the subscript J for brevity,

$$-\frac{\left(\frac{I_{i+1}-I_i}{\tau_{i+1}-\tau_i}\right) - \left(\frac{I_i-I_{i-1}}{\tau_i-\tau_{i-1}}\right)}{\left(\frac{\tau_{i+1}+\tau_i}{2}\right) - \left(\frac{\tau_i+\tau_{i-1}}{2}\right)} + I_i = S_i. \quad (7.15)$$

In somewhat more compact notation,

$$M\mathbf{I}_J = \mathbf{S} \quad (7.16)$$

From equation 7.15 we see that the i^{th} term is coupled only to the $i + 1$ and $i - 1$ terms. Together with the decomposition of the radiation stream, this is among the key innovations of the Feautrier method. The matrix M can be expressed as a block tridiagonal matrix, which has favourable properties for solving equation 7.16, such as the tridiag solver of [Press et al. \(1992\)](#).

$$M = \begin{bmatrix} b_1 & c_1 & 0 & \dots & 0 \\ a_2 & b_2 & c_2 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & a_{N-1} & b_{N-1} & c_{N-1} \\ 0 & \dots & 0 & a_N & b_N \end{bmatrix} \quad (7.17)$$

Using such a solution for I_J , we can now calculate I_H from equation 7.9.

Having calculated I_J and I_H , we must now calculate the top-of-atmosphere emission spectrum in the z direction, which is defined as perpendicular to the plane-parallel atmosphere

$$F_v = \int_0^{2\pi} I_v \mathbf{n} \cdot \mathbf{e}_z d\Omega. \quad (7.18)$$

which, in terms of I_H is

$$F_v = \int_0^{2\pi} \int_{-\pi/2}^{\pi/2} I_v(P=0) \cos(\theta) \sin(\theta) d\phi d\theta, \quad (7.19)$$

$$= 2\pi \int_{-1}^1 I_v(P=0) \mu d\mu, \quad (7.20)$$

$$= -4\pi \int_0^1 I_H(P=0) \mu d\mu. \quad (7.21)$$

The mean intensity can be likewise calculated using I_J

$$J_v = \int_0^1 I_J(P) d\mu. \quad (7.22)$$

In practice equations 7.19 and 7.22 are only evaluated at a small number (3) of μ points, chosen through Gaussian quadrature. As most of the flux is concentrated in the vertical direction, this does not significantly reduce the accuracy when compared to 16 quadrature points. pRT includes additional numerical approaches to reduce the computation time of the radiative transfer calculations, notably accelerated lambda integration and Ng acceleration, which are discussed in [Mollière \(2017\)](#).

7.2. Correlated-k

The correlated-k (c-k) method is a computationally efficient approximation of a full line-by-line (lbl) radiative transfer calculation, originally introduced by Goody et al. (1989) and Lacis & Oinas (1991). For a given spectral resolution, the c-k method achieves a factor of ~ 100 – 1000 speed-up compared to an equivalently accurate lbl calculation. The errors introduced by the approximation are on the order of 1% when compared to high resolution lbl calculations binned to the same resolution.

Gas opacities are typically complicated functions of frequency, with the absorption strength varying by orders of magnitude between the line cores and wings. However, the opacities are also highly structured, with each line having a similar shape. The important insight of the c-k method is that when integrating opacities to compute an optical depth in a small frequency window, the order of integration is not important. Therefore the absorption coefficient strengths can be reordered into a cumulative frequency distribution, which is a much simpler function to integrate than the opacity function. We can write this cumulative distribution function $f(k_i)$ for a narrow frequency window by discretising the opacity k onto a grid. Within the frequency window, the distribution is then obtained by counting the number of points that fall into each k bin. Thus for the i^{th} bin and for M spectral sub-intervals

$$f(k_i) = \frac{1}{\nu_2 - \nu_1} \sum_j^M \left| \frac{\Delta \nu_j}{\Delta k_i} \right| W(k_i, k_i + \Delta k_i), \quad (7.23)$$

where $W(k_i, k_i + \Delta k_i)$ is a tophat function on the domain $[k_i, k_i + \Delta k_i]$. This function is constrained to the range $[0, 1]$, which means that the cumulative frequency distribution increments define the fraction of the frequency interval for which the absorption coefficient falls into the i^{th} bin. We call this increment $\Delta g_i = f(k_i)\Delta k_i$. Thus we can formally define g as a function of the absorption coefficient.

$$g(k_n) = \sum_i^n f(k_i)\Delta k_i. \quad (7.24)$$

This transformation is defined to be monotonic and bijective, with an inverse

$$k_n(g) = g^{-1}(k_n). \quad (7.25)$$

In comparison to the original opacity spectrum, this function is smooth and monotonic, and can be easily numerically integrated on a coarse grid in g space

- 16 points are used to define the g grid in pRT. [Goody et al. \(1989\)](#) present a formal proof of the validity of this transformation in the limits of strong and weak lines. The coefficients k for coordinates g can be tabulated as functions of pressure and temperature for individual gas species, thus allowing for the rapid computation of the radiative transfer.

Combining the c - k opacities of multiple species requires mixing the distributions in g space. [Mollière \(2017\)](#) introduced a method to mix the species by sampling the g distributions to determine their statistical properties in order to combine them. However, this approach meant that the radiative transfer calculation was non-deterministic, with a scatter of about 1% around the expected result. In order to overcome this, I implemented a method to fully combine two k -distributions similar to that of [Amundsen et al. \(2017\)](#), resulting in both faster computation times and a deterministic outcome. For computational efficiency, we only combine species that contribute more strongly than 1% the strength of the strongest absorber in each frequency bin.

The goal is now to combine the opacities of N_{spec} gas phase species, for N_ν frequency points and N_P pressure points. The opacity k_ν of each species is stored on a grid of N_g points in g space, We define a combined grid g_c as our coordinate, weighted by the intervals in g space Δg ,

$$g_c \left((i-1) \cdot N_g + j \right) = \Delta g_i \Delta g_j. \quad (7.26)$$

where i and j are the indices counting the points in the saved g grid. The indices of the left hand side are chosen such that the new coordinate grid is also monotonic, containing $i \times j$ points. This grid is then normalised by the sum of g_c , such that it is normalised to the interval $[0,1]$.

Consider now the calculation of the new k_ν for a single point in pressure-frequency space. The opacities are combined in sequence: the first two species are combined, then the opacity of the third species is combined with the opacity of the mixture of the first two species. Thus we will take the general approach of combining two arbitrary opacity distributions k_1 and k_2 , we can be applied recursively in order to mix multiple species. We also drop the ν subscript as this process can be iterated over every frequency interval. Using the same sorting approach to ensure the correct placement of the opacities on the g_c grid, defining the combined opacity k_c as the sum of the two input opacities

$$k_c \left((i-1) \cdot N_g + j \right) = k_1(i) + k_2(j). \quad (7.27)$$

The resulting array of opacities is sorted by the g coordinate using a `mergesort` algorithm ([Olagnon, 2011](#)). This produces a cumulative distribution function of the opacity k_c with $i \times j$ points in g space. The combined opacity distribution

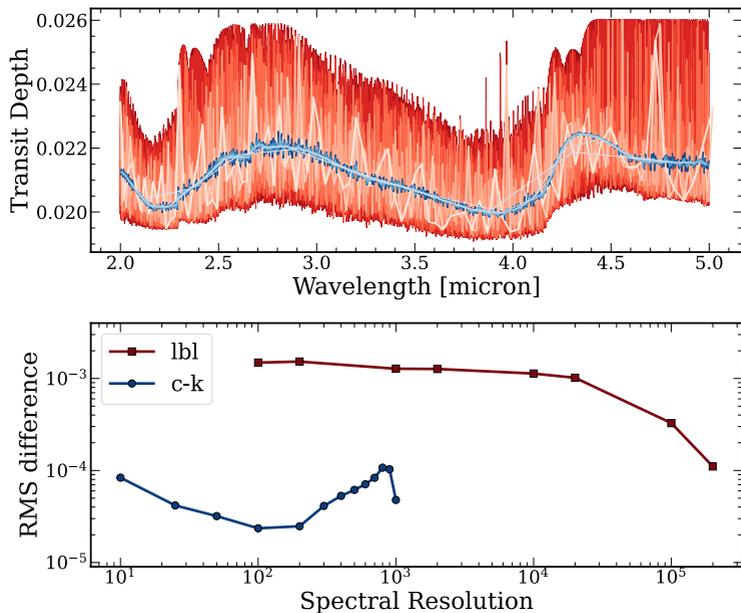


Figure 7.1: Comparison between a ground truth atmospheric model computed at a spectral resolution of 10^6 , and models calculated using binned c-k tables in blue, and models calculated using downsampled line-lists in red. The top panel shows the spectra at each resolution. The bottom panel shows the RMS difference between the binned spectra and the ground truth spectrum binned post-computation to the lower resolution bins. The c-k approach much more accurately reproduces the ground truth spectrum than the downsampling approach.

is linearly interpolated back to the original g grid with N_g points, and the process is repeated for every species, pressure, and frequency. Compared to the previous method, this approach resulted in a $\sim 5\times$ improvement in the computational speed and ensured that the c-k mixing is fully deterministic.

The final component of speeding up the calculation of spectra using the c-k approach is appropriately binning the spectrum. While it is easy to bin a high-resolution spectrum to lower resolution, for example by taking the mean flux in each low-resolution interval, this process still requires the computationally expensive step of calculating the high resolution spectrum. To avoid this, we use

the `exo-k` package of [Leconte \(2021\)](#) which calculates lower spectral resolution `c-k` tables from higher-resolution tables. These lower resolution inputs can then be used in the radiative transfer calculations, requiring fewer computations. The author provides a formal proof for the validity of this process, demonstrating that it does not result in a loss of accuracy. They also show that this method compares favourably to the widely used approach of downsampling a high-resolution line-by-line list to a lower resolution, which was validated using tests in `pRT`. Taking a spectrum computed using an `lbl` approach with a resolution of 10^6 , we calculated the difference between spectra computed using both the `c-k` method at various resolutions and using `lbl` spectra downsampled by various factors, shown in Figure 7.1.

7.3. Opacity sources

`pRT` can compute spectra either using line-by-line (`lbl`) calculations, or using correlated-`k` (`c-k`) tables for defining the opacities of molecular species. Opacities at high resolution are mostly calculated from the most complete and up-to-date line lists found in the ExoMol [Tennyson & Yurchenko \(2012\)](#); [McKemmish et al. \(2016\)](#); [Polyansky et al. \(2018a\)](#); [Chubb et al. \(2021\)](#) and HITEMP [Rothman et al. \(2010a\)](#); [Hargreaves et al. \(2020\)](#); [Gordon et al. \(2022\)](#) data bases, using the method described in [Mollière et al. \(2015\)](#). Correlated-`k` tables can be calculated from these high resolution opacities ([Mollière et al., 2019](#)) or downloaded in `pRT` format already from the Exomol website ([Chubb et al., 2021](#)). Contribution by `pRT` users is another source of the opacity database. A full list of the available species and sources is maintained in the `pRT` documentation.

7.4. Thermal structure

As `pRT` does not compute the thermal structure self consistently with the radiative transfer and chemistry, parametric models must be used. While it is possible to arbitrarily define the temperature for each pressure level (or indeed for users to define their own temperature pressure-profiles), `pRT` includes several profiles commonly used in the literature for ease of use. These are included in the `physics` sub-package, and can be combined with different chemical and cloud setups.

Isothermal

The simplest possible temperature structure is a constant temperature throughout the atmosphere. As noted previously, in the case of an emission

spectrum, this would produce only a blackbody curve. However, transmission spectra are often only sensitive to a narrow region of the atmosphere at low pressures, and an isothermal approximation is reasonable. Even with the advent of broad wavelength *JWST* transmission spectra, an isothermal atmosphere may still be preferred for its simplicity over a more complex model.

Madhusudhan and Seager 2009

Madhusudhan & Seager (2009) define a parametric temperature pressure profile, where the temperatures T_i at different pressures P_i are both free parameters. They define the profile as

$$P_0 < P < P_1 : P = P_0 e^{\alpha_1 (T - T_0)^{\beta_1}}, \quad (7.28)$$

$$P_1 < P < P_3 : P = P_2 e^{\alpha_2 (T - T_2)^{\beta_2}}, \quad (7.29)$$

$$P > P_3 : T = T_3, \quad (7.30)$$

$$(7.31)$$

where α_1 and α_2 are also a free parameters. In their work, $\beta_1 = \beta_2 = 0.5$. This can be solved to find the temperature for every pressure. Beginning with finding the boundary temperatures,

$$P > P_3 : T = T_3, \quad (7.32)$$

$$P_2 < P < P_3 : T_2 = T_3 - \frac{(\log P_3 - \log P_2)^{1/\beta_2}}{\alpha_2}, \quad (7.33)$$

$$P_1 < P < P_2 : T_1 = T_2 + \frac{(\log P_1 - \log P_2)^{1/\beta_2}}{\alpha_2}, \quad (7.34)$$

$$P_0 < P < P_1 : T_0 = T_1 - \frac{(\log P_1 - \log P_0)^{1/\beta_1}}{\alpha_1}. \quad (7.35)$$

We can then compute the full temperature profile

$$P > P_3 : T = T_3, \quad (7.36)$$

$$P_1 < P < P_3 : T = \frac{(\log P - \log P_2)^{1/\beta_2}}{\alpha_2} + T_2, \quad (7.37)$$

$$P_0 < P < P_1 : T = \frac{(\log P - \log P_0)^{1/\beta_1}}{\alpha_1} + T_0. \quad (7.38)$$

In practice, the version of this profile implemented in pRT is somewhat more flexible, and can compute the boundary temperatures as long as four temperatures at four pressures are defined.

The goal of this profile was to provide a data-driven framework for fitting the temperature profile, using as few parameters as possible while still being able to capture a diversity of shapes. The structure of the profile was motivated by physics: at high pressures the atmosphere becomes optically thick, and so a retrieval on the spectrum can only be sensitive down to the deepest layer of the atmosphere before the atmosphere is fully opaque. This layer sets T_3 , and the lack of a deep adiabatic structure does not significantly impact the observed spectrum. In the optically thin regions of the atmosphere, the profile must be flexible enough to capture the diversity of temperature profiles, from those following an adiabat to the inverted profiles of hot and ultra hot Jupiters. This requires at least two curves to describe such an inversion. By freely retrieving both the pressure and the temperature at that location, this profile parameterises both the strength and location of an inversion, if it is present.

Guillot 2010

The profile of [Guillot \(2010\)](#) is an analytic model, used to calculate the thermal structure of irradiated planets:

$$T_{\text{Guillot}}^4 = \frac{3T_{\text{int}}^4}{4} \left(\frac{2}{3} + \tau \right) + \frac{3T_{\text{irr}}^4}{4} \left(\frac{2}{3} + \frac{1}{\gamma\sqrt{3}} + \left(\frac{\gamma}{\sqrt{3}} - \frac{1}{\gamma\sqrt{3}} \right) e^{-\gamma\tau\sqrt{3}} \right), \quad (7.39)$$

where $T_{\text{irr}} = \sqrt{2}T_{\text{equ}}$ and $\tau = P \times \kappa_{\text{IR}}/g$. T_{equ} is the equilibrium temperature of an irradiated body, T_{int} is the intrinsic temperature of the planet and g is the surface gravity. κ_{IR} is the mean infrared opacity, and γ is the ratio between the optical and infrared opacities. With only four physically motivated parameters, the Guillot profile manages to capture much of the same structure as the [Madhusudhan & Seager](#) profile, ranging from a [Eddington \(1930\)](#) profile where the temperature always decreases with decreasing pressure to inverted atmospheres as observed in highly irradiated planets.

Line 2015 Spline

A spline temperature profile defines the temperature at N points throughout the atmosphere, and interpolates between these points to fully define the temperature structure ([Line et al., 2015](#)). The N temperature points are equidistantly spaced throughout the atmosphere in log-pressure space. We implement

both a linear interpolation method and a Piecewise Cubic Hermite Interpolating Polynomial using the `scipy.interpolate.PchipInterpolator` function. This cubic interpolation method was chosen to minimise the temperature overshoot or undershoot between nodes, effectively minimising the curvature of the entire profile.

Mollière 2020

[Mollière et al. \(2020\)](#) introduce a physically motivated temperature profile, split into three distinct altitude regions.

We begin with the photosphere region, which is the central of the three temperature profile components (the other two lying at lower or higher altitudes). Here, the temperature follows an Eddington profile, as in the first term of the Guillot profile in Equation 7.39. Rather than calculating τ as a function of an opacity parameter and the surface gravity, we parameterise τ as a function of pressure:

$$\tau = \delta P^\alpha, \quad (7.40)$$

where δ and α are independent parameters. Together with T_{int} , this part of the profile is defined by these three parameters.

The upper atmosphere is defined as the region above $\tau = 0.1$. At pressures lower than this level, four points equidistant in $\log P$ are defined. At $\tau = 0.1$ the temperature is fixed to that of the Eddington profile from the middle region. The remaining temperature points are free parameters of the model. For self-luminous atmospheres, these points are subject to the constraint that the temperature decreases with altitude ([Kitzmann et al., 2020](#)). The temperature profile is then interpolated from a cubic spline between the three points. Together with the central region, a total of 6 parameters are necessary to describe the temperature profile.

The lower region of the atmosphere is defined as a moist adiabat, up to the radiative-convective boundary, found using Equation 5.8. The moist adiabatic gradient is calculated as part of the equilibrium chemistry table, and thus this region of the profile is consistent with the chemistry model. At pressures deeper than the radiative-convective boundary, the temperature profile is forced onto this adiabatic gradient.

Zhang 2023

Finally, [Zhang et al. \(2023\)](#) recently introduced a novel P-T parameterisation. The atmosphere between 10^3 bar and 10^{-3} bar is divided up into six layers,

equidistant in log pressure. The temperature at the bottom of the atmosphere (T_{bot}) a free parameter. For the remaining layers, the temperature gradient $d \log T / d \log P|_i$ rather than the temperature itself are the model parameters. The temperature profile is then found by interpolating the gradient to the full pressure grid, and integrating to find the temperature at each pressure.

$$T_0 = T_{\text{Bot}}, \quad (7.41)$$

$$T_{i+1} = \exp \left(\log T_i + (\log P_{i+1} - \log P_i) \left(\frac{d \log T}{d \log P} \right)_i \right). \quad (7.42)$$

The atmosphere is isothermal above 10^{-3} bar. The key development of this profile was how it incorporated the results of radiative-convective equilibrium models into a retrieval framework, through careful prior selection. This is accomplished by fitting for the gradient of the temperature with respect to pressure, as opposed to directly retrieving the temperature as in the spline profile. By defining the priors such that the gradients must be similar to those expected from an RCE solution, this enforces physical results in a data-driven framework.

7.5. Chemistry

Two approaches to parameterising chemical abundances are included in pRT. The so-called ‘free chemistry’ approach assumes that the mass fraction of each species is vertically constant, and its value can be arbitrarily defined. In principle this method can be extended to arbitrary chemical profiles, simply requiring that the abundance of each species is specified at each discrete pressure level rather than as a single value. This approach is commonly used in remote sensing of solar system atmospheres, where the spectral resolution is high and measurements are very precise. With the abundance of moderate resolution data from *JWST* and the construction of ELT class telescopes, such approaches may soon be necessary for the interpretation of brown dwarf and exoplanets

In the second approach, equilibrium and disequilibrium chemistry can be interpolated from a pre-computed grid on-the-fly. These methods are included in the `chemistry.pre_calculated_chemistry` subpackage of pRT. The underlying equilibrium chemistry grid was computed using `easyChem` (Mollière et al., 2017), which is a Python-wrapped clone of the CEA code of Gordon & McBride (1994). The grid is composed of 100 equidistant temperature points between 60 K and 4000 K, 100 log-equidistant pressure points in $[10^{-8}, 1000]$ bar, 20 C/O points between 0.1 and 1.6, and 40 metallicity

points between -2 and 3 (in units of $[\text{Fe}/\text{H}]$ with respect to solar). The metallicity scales all elemental abundances by a factor of $10^{[\text{Fe}/\text{H}]}$, after which the oxygen abundance is scaled by the C/O parameter. The reactant species used to compute the equilibrium chemistry table were H, H_2 , He, O, C, N, Mg, Si, Fe, S, Al, Ca, Na, Ni, P, K, Ti, CO, OH, SH, N_2 , O, SiO, TiO, SiS, H_2O , C_2 , CH, CN, CS, SiC, NH, SiH, NO, SN, SiN, SO, S_2 , C_2H , HCN, C_2H_2 (acetylene), CH_4 , AlH, AlOH, Al_2O , CaOH, MgH, Mg, OH, PH_3 , CO_2 , TiO_2 , Si_2C , SiO_2 , FeO, NH_2 , NH_3 , CH_2 , CH_3 , H_2S , VO, VO_2 , NaCl, KCl, e-, H+, H-, Na+, K+, PH_2 , P_2 , PS, PO, P_4O_6 , PH, V, FeH, VO(c), VO(L), $\text{MgSiO}_3(\text{c})$, $\text{SiC}(\text{c})$, Fe(c), $\text{Na}_2\text{S}(\text{c})$, KCL(c), Fe(L), SiC(L), $\text{MgSiO}_3(\text{L})$, $\text{H}_2\text{O}(\text{L})$, $\text{H}_2\text{O}(\text{c})$, TiO(c), TiO(L), $\text{TiO}_2(\text{c})$, $\text{TiO}_2(\text{L})$, $\text{H}_3\text{PO}_4(\text{c})$, and $\text{H}_3\text{PO}_4(\text{L})$, where (c) indicates solid condensates and (L) indicates liquid condensate species. Using a Gibbs minimisation approach, the equilibrium abundances of each species was calculated, and the abundances of the most common species were stored in the lookup table, these include H_2 , He, CO, H_2O , HCN, C_2H_2 (acetylene), CH_4 , PH_3 , CO_2 , NH_3 , H_2S , VO, TiO, Na, K, SiO, e-, H-, H, and FeH. The mean molecular weight is also stored, as is the local moist adiabatic temperature gradient, ∇_{ad} . Chemical disequilibrium is implemented using a quench pressure approximation, modelling the impact of strong vertical mixing. At pressures lower than the specified quench pressure, the abundances of CO, H_2O and CH_4 remain constant.

These two approaches can be combined. The chemical equilibrium model can be used to calculate the abundances of one set of species, while a separate set can have their abundances defined by the user. This allows for a flexible approach where some species may not be expected in equilibrium, but may nevertheless play an important role in contributing to the opacity.

7.6. Clouds

Clouds are a key source of continuum opacity in exoplanet atmospheres, and several cloud parameterisations are included in pRT.

Grey cloud deck

The simplest implementation is that of a grey cloud deck. This model is parameterised only by the cloud top pressure, below which the atmosphere is opaque. Such a model is typically used in transmission geometries, which are sensitive to a relatively narrow dynamic range in pressure, though several recent studies have detected cloud optical features in mid infrared transmission spectra (e.g. [Grant et al., 2023](#); [Dyrek et al., 2024](#)).

Power law cloud

A somewhat more realistic model for the cloud opacity κ is to scale the opacity as a function of wavelength.

$$\kappa = \kappa_0 \left(\frac{\lambda}{\lambda_0} \right)^\gamma, \quad (7.43)$$

where $\lambda_0 = 0.35 \mu\text{m}$. This opacity is added to any Rayleigh scattering calculated in the atmosphere. For $\gamma = -4$, this follows the same wavelength dependence as Rayleigh scattering.

Fixed particle radius cloud

As an intermediate model between a fully parametric approach and a physics-motivated paradigm, a cloud can be defined by its mass fraction abundance and mean particle radius at every pressure layer within the atmosphere. Most commonly this would be used to define a log-normal particle size distribution centred on a mean particle radius that remains constant with altitude. The cloud optical properties can then be determined by defining a composition and particle geometry for the cloud.

Ackerman-Marley 2001

For a more physically motivated cloud model, pRT includes the clouds model of [Ackerman & Marley \(2001\)](#), as introduced in Section 5.5.4. A variety of common cloud species are included, where the cloud species determines both the optical properties of the clouds and the condensation pressure. This condensation pressure is calculated from the intersection of the (partial) pressure-temperature profile with the saturation vapor pressure curve of the condensate species. An arbitrary cloud base pressure can also be set, reflecting uncertainties in the modelling of the cloud formation process. Multiple cloud species can be included, with independent f_{sed} values for each. The cloud mass fraction at the base of the cloud can be determined either from equilibrium chemistry condensation, or as an independent parameter. The cloud mass fraction of the i^{th} species X_i is then usually parameterised such that the mass fraction decreases with decreasing pressure:

$$X_i(P) = X_{i,0} \left(\frac{P}{P_0} \right)^{f_{\text{sed}}}. \quad (7.44)$$

In the standard setup, a log-normal particle size distribution is used to determine the particle sizes about the mean size determined by balancing the vertical mixing

and sedimentation velocity. However, a [Hansen \(1971\)](#) size distribution has also been included, the implementation of which is presented in Section 24.D.

Arbitrary continuum opacities

Finally, an arbitrary source of continuum opacity can be defined, where the opacity κ must be calculated as a function of pressure and wavelength. This allows for a fully flexible implementation of a cloud or continuum opacity model.

Patchy clouds

Any of these cloud models can be used as part of a patchy cloud model. In pRT, this is typically implemented as a mix between a clear atmosphere column and a cloudy atmosphere column, though it is also possible to mix columns with different cloud opacities. In either case, the combination of the two columns is achieved from adding the weighted flux of each column F_a and F_b , weighted by a patchiness fraction f_p

$$F = f_p F_a + (1 - f_p) F_b. \quad (7.45)$$

Such patchy clouds have been found to be an important part of fitting sub-stellar atmospheres (e.g., [Vos et al., 2023](#)).

7.6.1. Adaptive mesh refinement

The adaptive mesh refinement (AMR) scheme originally introduced in [Mollière et al. \(2020\)](#) has been significantly updated to improve computational performance. Fundamentally, the procedure remains the same: in the regions where clouds are located, the resolution of the pressure grid is increased in order to better resolve the sharp change in opacity caused by condensation, without requiring the computational expense of computing the radiative transfer on a dense pressure grid throughout the atmosphere. The main change in the new approach is that for a given number of cloud layers in the atmosphere, defined by the number of cloud species included, the final pressure grid is guaranteed to have a fixed length. This means that pRT no longer needs to update the size of internal arrays for opacities, temperatures, and other parameters for each iteration in a retrieval.

8. RETRIEVALS

Unfortunately for atmospheric physicists, exoplanets are too far away to measure the atmospheric state directly. Instead, we obtain measurements of the emission or transmission spectrum, whose properties are inherently tied to the atmospheric state. The problem is to take this spectroscopic measurement, and invert it in order to obtain the atmospheric state that produced it. This procedure is generally referred to as an ‘atmospheric retrieval’, and inherits its techniques from earth and solar system science, such as those described in [Rodgers \(2000\)](#). In a trivial case, we can imagine that the spectrum \vec{F} is linearly determined by a forward model \mathbf{M} applied to an atmospheric state vector \vec{x}

$$\vec{F} = \mathbf{M}\vec{x}. \quad (8.1)$$

Clearly, all we need to do to solve the inverse problem is to invert the matrix \mathbf{M} such that

$$\mathbf{M}^{-1}\vec{F} = \vec{x}. \quad (8.2)$$

While this is an easy problem to state, actually solving it in practice can be very challenging. Consider the addition of an uncertainty term $\vec{\epsilon}$, so that

$$\vec{F} = \mathbf{M}\vec{x} + \vec{\epsilon}. \quad (8.3)$$

Already, the problem can no longer be trivially inverted, and indeed there is no guarantee that a solution exists. At this point, we must rely on techniques to find approximate solutions to the problem. Most of these techniques rely on minimising some likelihood function between the forward model and the observations. Many procedures exist for achieving this, from least-squares minimisation, through to optimal estimation ([Rodgers, 1976](#); [Twomey et al., 1977](#); [Rodgers, 2000](#)), Markov Chain Monte Carlo methods ([Metropolis et al., 1953](#); [Hastings, 1970](#)), and nested sampling [Skilling \(2004\)](#). These techniques have been widely applied to the study of exoplanet atmospheres. Non-exhaustively, retrieval methods with applications to exoplanet atmospheres have been described in [Madhusudhan & Seager \(2009\)](#); [Madhusudhan & Seager \(2011\)](#); [Benneke & Seager \(2012\)](#); [Lee et al. \(2013\)](#); [Line et al. \(2014\)](#); [Blecic et al. \(2016\)](#); [Mollière et al. \(2020\)](#); [Cubillos & Blecic \(2021\)](#); [MacDonald & Lewis \(2022\)](#); [Chubb & Min \(2022\)](#) and [Blain et al. \(2024\)](#). For a more complete list, [MacDonald & Batalha \(2023\)](#) catalogue retrieval codes that are available as of 2023.

The choice of method depends on the specific problem at hand. Optimal estimation methods perform well in the limit of abundant, high precision

data. However, they rely on the linearisation of the forward model, and assume that the posterior can be described using a multivariate gaussian distribution. MCMC methods are widely used (e.g. [Madhusudhan & Seager, 2011](#)), but suffer from the curse of dimensionality. It is also difficult to compare different models fit using a standard MCMC approach. Nested Sampling methods were designed to resolve these problems, providing estimates of the posterior probability distribution and the Bayesian evidence, which can be used to perform model comparison while also reducing the computational runtime when compared to standard MCMC methods. [Buchner \(2023\)](#) provides a review of the motivation and implementation of such nested sampling methods.

We can recast the inverse problem into a Bayesian formalism, which lends itself well to interpretation of the results. Consider Bayes theorem for some parameters $\vec{\theta}_i$, data \vec{D} and model M_i

$$P(\vec{\theta}_i|\vec{D}, M_i) = \frac{P(\vec{D}|\vec{\theta}_i, M_i)P(\vec{\theta}_i|M_i)}{P(\vec{D}|M_i)}. \quad (8.4)$$

We refer to $P(\vec{\theta}_i|\vec{D}, M_i)$ as the posterior probability distribution of the parameters $\vec{\theta}_i$ given the data. This directly tells us what the probability is that a model parameter falls within some range, based on the data. $P(\vec{\theta}_i|M_i)$ is the prior probability distribution, which can be used to constrain the problem. $P(\vec{D}|\vec{\theta}_i, M_i)$ is termed the likelihood. This term is generally what must be computed in order to calculate the posterior probability distributions, and can be interpreted as the probability of observing the data given the current set of parameters are true. By varying the parameters and keeping the data constant, we can explore the likelihood space and fill in the posterior distribution. Lastly $P(\vec{D}|M_i) \equiv \mathcal{Z}$ is the Bayesian evidence. It serves to normalise the posterior probability distribution, and does not rely on a particular parameter choice, but rather the choice of the model under investigation. Equivalently, it is the likelihood marginalised over the parameters. This provides a means to perform the likelihood ratio test between two models in order to determine which one is favoured over the other. This is the idea behind the widely-used ‘Bayes-factors’. The ratio $\mathcal{Z}_1/\mathcal{Z}_2$ between two models is the odds ratio of model 1 compared to model 2, where a ratio greater than 1 indicates that model 1 is more likely than model 2. The Bayes factor is the log of this ratio. [Kass & Raftery \(1995\)](#) and [Benneke & Seager \(2013\)](#) provide similar interpretations of the Bayes factor in terms of σ , which reflects the probability that the model preference occurs through random chance. As a rule-of-thumb, a $\log_{10} \Delta\mathcal{Z} = 2$ is decisive evidence, roughly equivalent to a 5σ detection.

In terms of solving the inverse atmosphere problem, our goal is now to a)

measure the posterior probability distributions of our model parameters, and b) calculate the Bayesian evidence so that we can perform model comparison. In order to do this, we must iterate through many variations of the parameters θ , and calculate the likelihood for each sample. The core of the problem is therefore to define the likelihood function, which when integrated over the parameters will produce the evidence \mathcal{Z} , and when minimised will provide the best-fit parameter values to describe the data.

8.1. Likelihood functions

The likelihood function provides a measurement of how well the data fits a model, given a fixed set of parameters. In practice it is more convenient to work with the log-likelihood function, which can be numerically easier to minimise. A natural example of such a log-likelihood is the χ^2 function. Given the assumption that the data D are independent and normally distributed with uncertainties σ , we can compare to a model S :

$$\chi^2 = \sum_i \frac{(D_i - S_i)^2}{\sigma_i^2}. \quad (8.5)$$

This can be interpreted as a distance between the data and the model, weighted by the uncertainties. A better fit to the data will result in a lower χ^2 value, tending towards the number of data points, and so minimising the χ^2 will result in the best fit to the data. We can generalise the χ^2 to account for correlated uncertainties, represented by the covariance matrix \mathbf{C} . For a k -dimensional multivariate normal distribution, the joint probability distribution of the i^{th} term is

$$P_{\vec{D}}(D_i) = \frac{\exp\left(\frac{1}{2} (D_i - S_i)^T \mathbf{C}^{-1} (D_i - S_i)\right)}{\sqrt{(2\pi)^N \det(\mathbf{C})}}. \quad (8.6)$$

We can find the log-likelihood by taking the log of the probability distribution. Up to a normalisation constant,

$$\mathcal{L} = (\vec{D} - \vec{S})^T \mathbf{C}^{-1} (\vec{D} - \vec{S}). \quad (8.7)$$

In pRT, the model S is an emission or transmission spectrum calculated on-the-fly based on a set of input parameters, while D and C are the input spectroscopic data and uncertainties being fit. However, we may need to estimate the uncertainties as part of the model fitting procedure, as the error bars provided from

instrumental measurements do not always reflect the statistical properties of the data. The second term which falls out naturally from considering the log of the joint multivariate probability distribution allows uncertainties to vary as a free parameter, acting to penalise overly large uncertainties.

$$-2 \log \mathcal{L} = (\vec{D} - \vec{S})^T \mathbf{C}^{-1} (\vec{D} - \vec{S}) + \log \left[(2\pi)^k \det(\mathbf{C}) \right], \quad (8.8)$$

where k is the dimensionality of the data vector \vec{D} .

8.2. Nested sampling

With a likelihood function defined, it is now possible to evaluate it in order to calculate posterior distributions and estimate the evidence (Buchner, 2023). The procedure of nested sampling (Skilling, 2004) is as follows.

Initialisation. During the initialisation phase, N points are drawn from the prior volume. These are collectively known as live points. The likelihood is calculated for each live point.

Likelihood-restricted prior sampling. Following the calculation of the likelihoods, the point with the lowest likelihood is dropped, becoming a dead point. A new sample is drawn with the condition that the new likelihood must be greater than the remaining minimum likelihood. This ensures that the sample will be drawn from a smaller region of the prior volume, decreasing roughly as $1/N$ for each dead point.

Iteration. This process repeats. For each iteration, the likelihood threshold increases, and thus the prior volume sampled by the current set of live points decreases exponentially, as $V_i = (1 - 1/N)^i$.

Termination. The termination condition is typically based on the convergence of the evidence estimate, \mathcal{Z} . The contributions to the evidence become smaller as the volume shrinks, as the contribution $\mathcal{Z}_i = \Delta V_i \mathcal{L}_i$. Once the evidence contributions become sufficiently small, the algorithm can terminate. The Bayesian evidence can be computed as the sum of the likelihoods multiplied by their respective volume contributions:

$$\mathcal{Z} = \sum_i \Delta V_i \mathcal{L}_i. \quad (8.9)$$

Thus we have obtained both the Bayesian evidence estimate which we can use to perform model comparison as well as the posterior probability distributions of the parameters obtained through the sampling process. Both `MultiNest` (Feroz & Hobson, 2008; Feroz et al., 2009, 2019; Buchner et al., 2014) and `UltraneSt` (Buchner et al., 2014; Buchner, 2019) samplers are available in `pRT`, with both offering MPI implementations that allow for easy parallelisation. While `UltraneSt` is recommended due to its improved accuracy in estimating the evidence, and is generally a more modern approach to nested sampling, testing has shown that `MultiNest` remains faster for the problem of exoplanet atmospheres.

While a general nested sampling algorithm is described above, most retrievals in `pRT` will rely on the `MultiNest` implementation, as wrapped by `pyMultiNest` (Buchner et al., 2014). The innovation of this method is in how it determines the restricted region for drawing subsequent live points. `MultiNest` improves upon ellipsoidal nested sampling (ENS), introduced by Mukherjee et al. (2006). In standard ENS, an ellipsoid is drawn around the remaining likelihood points, and the subsequent sample drawn from within the enclosed volume in prior space. In `MultiNest`, multiple ellipsoids can be drawn, which allows for the easy enclosure of multiple modes in posterior space. The details of this procedure, including how the ellipsoids are determined and how parameters are drawn from the enclosed volume (which may include overlapping ellipsoids) is described in Feroz et al. (2009). The authors later improve upon this method by defining importance nested sampling (INS), (Feroz et al., 2019). In INS, the unknown underlying probability distribution is analytically approximated and reweights the samples. (Feroz et al., 2019) expand on this for multi-ellipsoidal sampling, demonstrating that using all previous sample draws, including discard points, results in a better estimate of $\log \mathcal{Z}$. This method improves the computational efficiency of `MultiNest`, and provides an order of magnitude more accurate error estimate for the same computational cost.

8.3. The `pRT` retrieval module

The retrieval module combines the atmospheric model of `pRT` with a nested sampling algorithm in order to infer the posterior distributions of atmospheric parameters and to estimate the Bayesian evidence to perform model comparison¹. The general outline of the retrieval module is shown in Figure 8.1. A brief summary of the retrieval module is given in Nasedkin et al. (2024), which I expand upon here.

¹Examples and documentation are available at https://petitradtrans.readthedocs.io/en/latest/content/retrieval_examples.html

The module is subdivided into different submodules. Datasets, priors and other retrieval hyper parameters are set through the `RetrievalConfig` class, while the `models` module includes a range of complete atmospheric models that can be fit to the data. Users can also define their own model function, either by making use of temperature profiles from the `physics` module and chemistry parameterisations from the `chemistry` module or by implementing their own forward model. Multiple datasets can be included into a single retrieval, with each dataset receiving its own `Radtrans` object used for the radiative transfer calculation where some or all forward model parameters may be shared between the different data sets. This allows for highly flexible retrievals where multiple spectral resolutions, wavelength ranges and even atmospheric models can be combined in a single retrieval. All of these components are wrapped into the `retrieval` module, which initialises the retrieval, runs the sampling algorithm and generates output files. In this Section I will present the key features of each of these modules.

RetrievalConfig

The `RetrievalConfig` class is the primary class a user will interact with when setting up a retrieval. When initialising the class a user will supply meta-data for the retrieval, such as a name to identify it. Multiple `Data` objects can be associated with a single `RetrievalConfig` object, as often data is obtained from multiple different instruments before being jointly fit during the retrieval. Both emission and transmission spectra can be added, as well as broadband photometric data.

Following the addition of the data, parameters must be initialised. The parameters are stored in a dictionary of `Parameter` objects. Each `Parameter` has a name that will be used in the `model_generating_function`. They can take on either a fixed value, or be treated as a free parameter in the retrieval. For free parameters, a prior must be defined. In the context of `Multinest`, the prior is a transformation between a unit hypercube and physical parameter space. For each sample, `Multinest` draws random parameter values on the interval $[0,1]$, which are then passed as input to the prior transformation functions. To define a uniform prior over an interval $[a,b]$, the prior function $f(x)$ would simply be

$$f(x) = a + (b - a)x. \quad (8.10)$$

A Gaussian prior $\mathcal{N}(\mu, \sigma, x)$ with mean μ and standard deviation σ , which can be used to inform the retrieval based on previous measurements, would take the form of an inverse complementary error function,

$$\mathcal{N}(\mu, \sigma, x) = \mu + \sigma \sqrt{2} \operatorname{erfc}^{-1} [2(1 - x)], \quad (8.11)$$

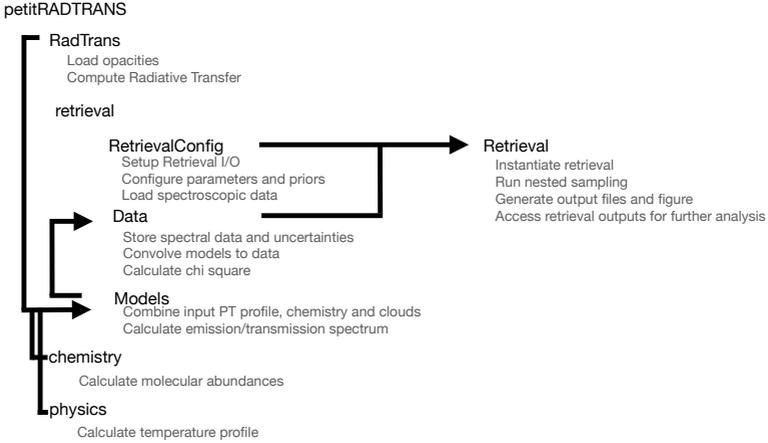


Figure 8.1: pRT program flow. The `Radtrans` class is the fundamental object in pRT, loading and storing opacities and calling the Fortran routines to calculate the radiative transfer. The `retrieval` module contains the classes used to perform an atmospheric retrieval. The retrieval inputs and parameters are set up in the `RetrievalConfig` object. Spectroscopic data is stored in a `Data` object. Associated with each `Data` object is a `model_generating_function`, which can be found in the `models` module or written independently. These functions return the wavelength and spectrum of the calculated model, which can be compared to the data. The `Retrieval` class reads in a `RetrievalConfig` object and its associated `Data` objects, initialises and runs a nested sampling retrieval, and generates output files and plots.

where

$$\operatorname{erfc}^{-1}[\operatorname{erfc}(x)] = x, \quad (8.12)$$

and

$$\operatorname{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-t^2} dt = 1 - \operatorname{erf}(x), \quad (8.13)$$

that is, inverse of the cumulative normal distribution function transforms from uniformly to normal distributed random variables. Such priors are included in pRT. Careful choice of priors is critical for reliable results, and is the subject of investigation in its own right. It is recommended to experiment with prior choices, ensuring that retrieval results are robust to such variation, or to determine how informative priors are impacting the inferred posterior distributions.

The `RetrievalConfig` class contains dedicated functions to setting up the parameters for chemical and cloud abundances. Instead of specifying parameters for each species individually, it is possible to define a list of species to include in the retrieval, and set whether the retrieval should be run using free chemistry or (dis)equilibrium chemistry. A combination of free and disequilibrium chemistry can also be used. Likewise for clouds, the user can specify the cloud species, whether the cloud should condense at its equilibrium location or have a freely retrieved base pressure, and whether the abundance should be determined from scaling the equilibrium abundance or be a freely retrieved parameter.

Finally, various plotting parameters can be defined in the `RetrievalConfig` object. These parameters can determine wavelength ranges and units to plot, which parameters should be included in corner plots, ranges of pressure and temperature to plot, and more. If the default plots are insufficient, the `Retrieval` class provides access to the `matplotlib` figure and axes objects, which can then be adjusted outside of `pRT`.

Data

The `Data` class is the primary object where spectroscopic or photometric data is stored. Input data can be either emission or transmission spectra, the only constraint is that the units of the input data and the model must match. Either standard 1σ Gaussian uncertainties can be used, or a full covariance matrix can be read in.

Many critical parameters must be defined for each `Data` object. This allows for a highly flexible retrieval framework: each dataset included in the retrieval could use an entirely unique atmospheric model, including different opacity choices (lbl or c-k), different binning, different scale factors or offsets and more. The `data_resolution` argument should be used to define the spectral resolution of the data based on the instrumental line spread function, which is most often not the same as the $\Delta\lambda/\lambda$ as stored in the input file. The calculated model will be convolved with a Gaussian kernel with this width, before being binned to the wavelength bins of the data. The model will be calculated at a spectral resolution specified by the `model_resolution` parameter, which will bin either the c-k or lbl opacities prior to computing the model to reduce the runtime. A scale factor can be applied to either the data points, the uncertainties, or both. Likewise, an offset can be applied to the data to calibrate between different datasets. A separate approach can also be used to retrieve uncertainties, by fitting for a free parameter b , following the approach of [Line et al. \(2015\)](#):

$$\sigma_T = \sqrt{\sigma^2 + 10^b}. \quad (8.14)$$

We calculate the total uncertainty for each wavelength σ_T bin from the nominal uncertainty σ and the inflation factor 10^b .

For photometric data, a transformation function must be supplied to compute the band-averaged photometry from the model spectrum. This is usually the transmission function of a filter. By default, pRT checks if the name of the `Data` object corresponds to a filter available from the SVO Filter Profile Service (Rodrigo et al., 2012; Rodrigo & Solano, 2020).

Often datasets will overlap in their wavelength ranges, and typically the same model will be used to fit both datasets. In such a case, an `external_radtrans_reference` can be supplied, where the `Radtrans` object associated with one dataset is used to provide the model for calculating the log likelihood for the other. This reduces the total number of model computations in a retrieval, and avoids redundantly calculating the same wavelength range multiple times for each set of parameters.

Models

The `models` module contains functions that take a `Radtrans` object and a set of parameters, and outputs the wavelength grid and calculated emission or transmission spectrum. The built-in models are typically distinguished by the thermal profile used, as each temperature profile model requires substantially different parameters. The chemistry is flexible, using the `chemistry` module to define abundances using free chemistry, (dis)equilibrium chemistry, or some combination of the two. Likewise the cloud prescriptions are highly general, and the model function will automatically determine the cloud model to use based on the supplied parameter set. Once the temperature profile, chemical abundances, and cloud structure have been defined, these are passed to the `Radtrans` object, which calculates the resulting spectrum. The model function must be calculated for each `Data` object associated with the `RetrievalConfig`.

Retrieval

All of the above classes are combined in the `Retrieval` class. In this class the `Multinest` or `Ultraneest` parameters such as the number of live points are supplied. The full log-likelihood function is defined in this class. It loops over each `Data` object associated with the `RetrievalConfig` object, computing the spectrum using the `model_generating_function` and calculating the log-likelihood with respect to each dataset. These log-likelihoods are subsequently summed together to determine the cumulative log-likelihood of the model compared to all of the datasets. Additional priors can be applied here, such as the likelihood penalty on the curvature in a spline temperature profile (Line

et al., 2014). Having defined the priors and the likelihood function, the sampler can then iterate through the parameter space, ultimately producing estimates of the Bayesian evidence and the posterior probability distribution, which can be marginalised over each parameter in order to determine the 1D posterior distributions.

Following the completion of the retrieval, the `Retrieval` object can be set to `evaluate` mode, and used to plot widely used diagnostic plots, examples of which are included in Section 9. An summary output file is also generated, which includes metadata about the parameters and datasets used in the retrieval, as well as the median and best-fit parameter results. Many functions are built into the `Retrieval` class to aid in the analysis of a retrieval. These include functions to calculate and store the best-fit spectrum. The χ^2 , as well as reduced χ^2 variations of any of these spectra can be calculated through comparisons to the data. The mass-fraction or volume-mixing-ratio abundances of a given sample, or the entire set of posterior samples, can be computed. Full API documentation of the features of the `Retrieval` class is available online.

9. CASE STUDY: WASP-39 B

The best way to understand the mechanics of an atmospheric retrieval analysis is by example. WASP-39 b is a hot Saturn-like planet that has been extensively studied as part of the *JWST* Transiting Early Release Science (ERS) program. As a contribution to that work, pRT retrievals will be included as part of a forthcoming retrieval comparison analysis in Welbanks et al. (in prep). Various retrieval codes were benchmarked against each other and against results from self-consistent forward models and GCMs in order to produce a robust, model independent portrait of this planet. In this Section I will present the results of the pRT retrievals of WASP-39 b, emphasising the key features of the retrieval module.

9.1. Background and previous work

WASP-39 b was discovered by [Faedi et al. \(2011\)](#). Using a combination of the WASP survey photometry, together with photometric followup from the LCOGT3 Faulkes Telescope North and spectroscopic radial velocity followup using the SOPHIE instrument at the OST, they determined that WASP-39 b is an inflated, Saturn mass planet with a radius $R_{\text{pl}} = 1.27 \pm 0.04 R_{\text{Jup}}$ and a mass of $M_{\text{pl}} = 0.28 \pm 0.03 M_{\text{Jup}}$, resulting in an extremely low density of $0.14 \pm 0.02 \rho_{\text{J}}$. The period of the planet is about 4 days, resulting in an equilibrium temperature of 1116^{+33}_{-32} K. The system is between 6-13 Gyr in age, with a stellar radius of $R_{\star} = 0.895 \pm 0.23 R_{\odot}$. The extremely low density means the atmosphere of WASP-39 b has a very large scale height, making it highly amenable to transmission spectroscopy observations. Subsequent broad band transit observations reinforced these planet parameters ([Maciejewski et al., 2016b,a](#)). Observations of the Rossiter-McLaughlin effect showed that the orbit of the planet is nearly perfectly aligned with the rotation of the host star ($i = 0^{\circ} \pm 11^{\circ}$) ([Mancini et al., 2018](#)). Shortly after its discovery, transmission spectra were obtained using *HST* ([Fischer et al., 2016](#)) and VLT/FORS2 ([Nikolov et al., 2016](#)); a cloud-free atmosphere is identified, and both sodium and potassium alkali lines are detected in the near infrared.

The combination of a large scale height and a clear atmosphere led to WASP-39b being selected as one of the *JWST* Transiting ERS targets. The goal of the ERS observations of WASP-39 b was to obtain high S/N across the full *JWST* wavelength range to enable the most robust atmospheric characterisation of a transiting exoplanet to date. To that end, it was observed with NIRSpec/PRISM ([Rustamkulov et al., 2023b](#)), NIRISS/SOSS ([Feinstein et al., 2023](#)), NIRCam ([Ahrer et al., 2023](#)), NIRSpec/G395H ([Alderson et al., 2023](#)) and the MIR-

I/LRS (Powell et al., 2024). Combined, these instruments cover wavelengths from 0.5–14 μm , with spectral resolutions ranging from $R = 100$ –3000. These observations were not without challenges. The NIRSpec/PRISM observations were saturated between 0.9–1.5 μm . While it was possible to recover a transmission spectrum in this region, it was of significantly lower quality than the remainder of the spectrum. Where instruments have overlapping wavelength ranges, there were often significant (~ 200 ppm) offsets between them, which required additional calibration to remove. However, even with the steep learning curve of acquiring and processing the spectra using new instruments and tools, the combined spectrum of WASP-39 b is one of the most comprehensive and precise spectra obtained to date, which has already altered our understanding of these atmospheres.

Perhaps the most surprising result of the ERS observations was the detection of SO_2 in both the near (Tsai et al., 2023) and mid infrared (Powell et al., 2024), though these detections were anticipated by photochemical calculations by Polman et al. (2023). This was the first detection of a sulphur bearing molecule in an exoplanet atmosphere. SO_2 is not expected from assumptions of equilibrium chemistry, but rather must be photochemically produced, in the net reaction $\text{H}_2\text{S} + \text{H}_2\text{O} + \gamma \rightarrow \text{SO}_2 + 3\text{H}_2$. Beyond the novel photochemistry, CO_2 was robustly detected for the first time in an exoplanet atmosphere (JWST Transiting Exoplanet Community Early Release Science Team et al., 2023), demonstrating the importance of broad wavelength coverage. A full range of modelling approaches has been used to analyse these objects, including self-consistent RCE models, GCMs, and retrievals (Constantinou et al., 2023; Constantinou & Madhusudhan, 2024). From the measurements of H_2O , CO and CO_2 abundances, these models generally agree on a sub-stellar C/O ratio around 0.3, and a metallicity of around $10\times$ solar, noting that WASP-39 itself is solar in composition.

9.2. Data

In this retrieval analysis, we use the full wavelength range available to study the atmosphere of WASP-39b using the pRT retrieval package. However, the data are highly inhomogenous in spectral resolution, and there are known offsets between different instruments, and even between different detectors in the same instruments. To mitigate this, a synthesised spectrum has been produced by the ERS team. This spectrum combines the observations from all of the *JWST* instruments, binning them to a constant spectral resolution. The regions of the PRISM spectrum that are saturated are masked out. A full description of the data synthesis project will be described in a forthcoming publication by Carter

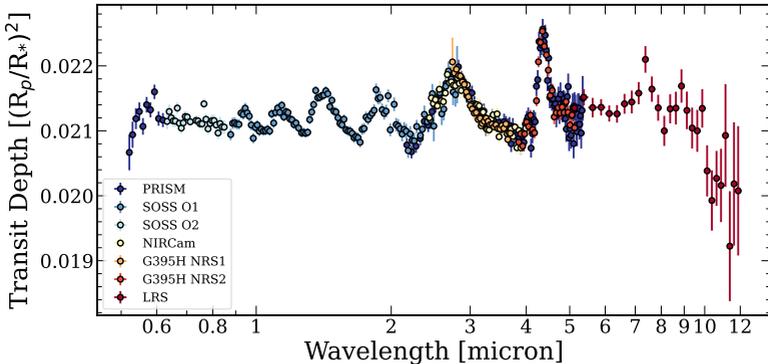


Figure 9.1: WASP-39 b transmission spectra as observed using the full set of *JWST* instruments. The raw spectra from each instrument have been binned to a constant spectral resolution, and offsets between the instruments have been applied. Using a series of test retrievals, the uncertainties have been fit for, finding moderately inflated uncertainties compared to the raw observations. Note that for transmission spectra, absorption features are upward, producing larger transit depths.

et al. (in prep). Offsets are fit for between each of the datasets, taking NIRISS SOSS O1 as the reference instrument. Using fits to grids of self-consistent models, the uncertainties are estimated as a free parameter, finding that the uncertainties should be inflated relative to the raw uncertainties produced by the *JWST* pipeline. The synthesised data are presented in Figure 9.1.

9.3. Retrieval setup

For retrievals of WASP-39 b, a fiducial model was agreed upon to provide a foundation for comparing different retrieval codes. This setup specified that in addition to H_2 and He, the atmosphere should be composed of H_2O , Na, K, CO, CO_2 , SO_2 , H_2S , CH_4 , NH_3 , HCN and C_2H_2 . The temperature structure should be non-isothermal. The clouds should consist of an opaque cloud deck, together with an optically thin haze layer. A cloud patchiness parameter should be retrieved to determine which fraction of the terminator region is covered by these two cloud components. One of the planet radius or the reference pressure should be specified, and the other allowed to vary. To implement this prescription in pRT, we used the temperature-pressure profile of [Guillot \(2010\)](#), as in equation 7.39. Our cloud model consisted of a grey cloud deck where

| Parameter | Prior |
|-------------------------------------------------------|----------------------------|
| $R_* [R_\odot]$ | 0.9324 |
| $P_{\text{ref}} [\text{bar}]$ | 0.01 |
| $\log g [\text{cm/s}^2]$ | $\mathcal{U}(2.0, 5.5)$ |
| $R_{\text{pl}} [R_{\text{Jup}}]$ | $\mathcal{U}(0.8, 1.6)$ |
| $T_{\text{int}} [\text{K}]$ | $\mathcal{U}(100, 1100)$ |
| $T_{\text{equ}} [\text{K}]$ | $\mathcal{U}(100, 1100)$ |
| γ_{Guillot} | $10^{-(x/2)^2/2}$ |
| $\log \kappa_{\text{IR}} [\text{cm}^2 \text{g}^{-1}]$ | $\mathcal{U}(-3, 1)$ |
| $\log P_{\text{cloud}} [\text{bar}]$ | $\mathcal{U}(-8, 3)$ |
| $\log \kappa_0 [\text{cm}^2 \text{g}^{-1}]$ | $\mathcal{U}(-4, 10)$ |
| $\log \gamma_{\text{Scat}}$ | $\mathcal{U}(-20, 2)$ |
| f_{cloud} | $\mathcal{U}(0, 1)$ |
| $\log X_i$ | $\mathcal{U}(-12.0, -0.3)$ |

Table 9.1: Priors for the fiducial free chemistry WASP-39 b retrieval setup. $\mathcal{U}(a, b)$ denotes uniform priors with bounds a and b , and $\mathcal{N}(\mu, \sigma)$ denotes a normal distribution centred at a mean μ with standard deviation σ . X_i is the mass fraction abundance of the i^{th} species.

the cloud top pressure was retrieved as a free parameter, and the haze model as described by equation 7.43, where both κ_0 and γ_{Scat} are free parameters. The reference pressure of the atmosphere was fixed to 0.01 bar, and we freely retrieved both the planet radius and $\log g$. The priors used for the free chemistry and disequilibrium retrievals are included in tables 9.1 and 9.2 respectively.

For the chemically consistent retrievals we used 4000 live points, and a sampling efficiency of 0.8. For the free chemistry retrievals, we used 1000 live points, as this was deemed sufficient from retrieval comparison testing. We used the c-k opacities at a model resolution of $R=300$. We performed validation retrievals, comparing the $R=300$ opacities to $R=1000$, finding that there were no significant changes to the posterior distributions. We convolved the model

| Parameter | Prior |
|-----------------------------------------------------------|----------------------------|
| R_* [R_\odot] | 0.9324 |
| P_{ref} [bar] | 0.01 |
| $\log g$ [cm/s^2] | $\mathcal{U}(2.0, 5.5)$ |
| R_{pl} [R_{Jup}] | $\mathcal{U}(0.8, 1.6)$ |
| T_{int} [K] | $\mathcal{U}(100, 1100)$ |
| T_{equ} [K] | $\mathcal{U}(100, 1100)$ |
| γ_{Guillot} | $10^{-(x/2)^2/2}$ |
| $\log \kappa_{\text{IR}}$ [$\text{cm}^2 \text{g}^{-1}$] | $\mathcal{U}(-3, 1)$ |
| [M/H] | $\mathcal{U}(-1.5, 3.0)$ |
| C/O | $\mathcal{U}(0.05, 1.55)$ |
| $\log P_{\text{quench}}$ [bar] | $\mathcal{U}(-8, 3)$ |
| $\log P_{\text{cloud}}$ [bar] | $\mathcal{U}(-8, 3)$ |
| $\log \kappa_0$ [$\text{cm}^2 \text{g}^{-1}$] | $\mathcal{U}(-4, 10)$ |
| $\log \gamma_{\text{Scat}}$ | $\mathcal{U}(-20, 2)$ |
| f_{cloud} | $\mathcal{U}(0, 1)$ |
| $\log X_{\text{SO}_2}$ | $\mathcal{U}(-12.0, -0.3)$ |

Table 9.2: Priors for the fiducial disequilibrium WASP-39 b retrieval setup. $\mathcal{U}(a, b)$ denotes uniform priors with bounds a and b , and $\mathcal{N}(\mu, \sigma)$ denotes a normal distribution centred at a mean μ with standard deviation σ .

with the line spread function for each instrument, before binning it to the wavelength bins provided by the data synthesis team.

9.4. Free chemistry

For the free chemistry retrievals, the mass fraction abundance of each trace species was retrieved as a free parameter. We included $\text{H}_2\text{-H}_2$ and $\text{H}_2\text{-He}$ col-

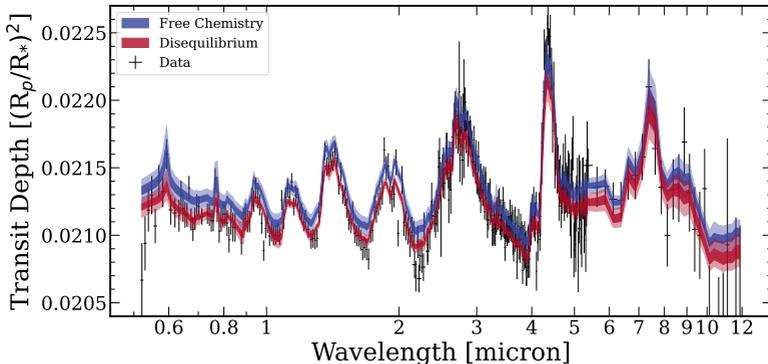


Figure 9.2: Retrieved spectra for WASP-39 b for a free chemistry model (blue), and a disequilibrium model (red). The light shading indicates the 3σ spread of the retrieved models, the dark shading is the 1σ interval.

litionally induced absorption and Rayleigh scattering from H_2 and He. Our sources of line opacities were H_2O (Polyansky et al., 2018b), CO_2 (Yurchenko et al., 2020), SO_2 (Underwood et al., 2016), CO (Rothman et al., 2010b), NH_3 (Coles et al., 2019), C_2H_2 (Chubb et al., 2020), H_2S (Azzam et al., 2016), HCN (Barber et al., 2014), CH_4 (Yurchenko et al., 2017), Na (Piskunov et al., 1995; Allard et al., 2019), and K (Piskunov et al., 1995; Allard et al., 2016). Figure 9.2 presents the retrieved spectrum of WASP-39 b as compared to the *JWST* data. In general the data are well reproduced: the average transit depth is consistent, and most of the absorption features (primarily due to H_2O , CO_2 and CO) are well fit. 1 and 3σ intervals are included in the figure, and we find that the models are generally consistent with the data; for the free retrieval, the reduced χ^2 is 1.05. While our fiducial model uses the temperature profile of Guillot (2010), in the region of the atmosphere to which the transmission spectrum is sensitive, we find that it closely follows an isothermal temperature profile. This is apparent in Figure 9.3, where the atmosphere in the free chemistry retrieval is isothermal above $\sim 10^{-2}$ bar. In comparisons to an identical retrieval run using an isothermal profile, the Guillot (2010) model is not favoured by the Bayes factor. However, we found that there was no preference for this profile over the use of an isothermal temperature profile. At 768 ± 28 K, the atmospheric temperature in the terminator region is found to be cooler than the equilibrium temperature of the planet.

We find a bulk metallicity of $[\text{M}/\text{H}] = 1.5 \pm 0.3$, or approximately $32\times$ solar. Using the water, CO and CO_2 abundances we find $\text{C}/\text{O} = 0.2 \pm 0.1$.

| Molecule | log VMR | $\Delta \log_{10} \mathcal{Z}$ |
|------------------|-------------------|--------------------------------|
| H ₂ O | -1.52 ± 0.237 | ... |
| CO | -2.1 ± 0.607 | ... |
| CO ₂ | -3.08 ± 0.256 | ... |
| CH ₄ | -9.09 ± 1.89 | ... |
| H ₂ S | -3.47 ± 0.678 | 1.4 |
| SO ₂ | -4.79 ± 0.238 | 9.3 |
| NH ₃ | -9.67 ± 1.69 | ... |
| HCN | -9.22 ± 2.03 | ... |
| Na | -2.83 ± 0.68 | ... |
| K | -5.14 ± 0.775 | ... |
| [M/H] | 1.5, 0.3 | ... |

Table 9.3: Volume mixing ratios from the full data, free chemistry retrieval.

The planet is cloudy, with $86 \pm 5\%$ cloud coverage and a log cloud top pressure of -2.5 ± 3.5 bar. While the spectral features of water and CO₂ are clear, we test for the significance of the detections of SO₂ and H₂S, finding that SO₂ is robustly detected with a \log_{10} Bayes factor of 9.3 and a log volume mixing ratio of -4.8 ± 0.2 , and H₂S is tentatively detected with a \log_{10} Bayes factor of 1.4 and an abundance of -3.5 ± 0.7 . The volume mixing ratios of each species are included in Table 9.3.

In addition, we also performed retrievals on each instrument individually, without accounting for any offsets or error inflation. In general, the individual retrievals were in excellent agreement, with the exception of the retrieval using only the NIRCam data, which found a sub-solar metallicity of 0.9 ± 1 . The water volume mixing ratio was also generally consistent with the full retrieval volume mixing ratio of -1.5 ± 0.2 . However, the MIRI/LRS retrieval did not show any evidence of water, instead finding a significant quantity of methane.

In our fiducial setup we retrieve both the planet radius and surface gravity, fixing the reference pressure to 0.01 bar. We also explored the impact of fixing these three parameters to their measured value or leaving them as free parameters.

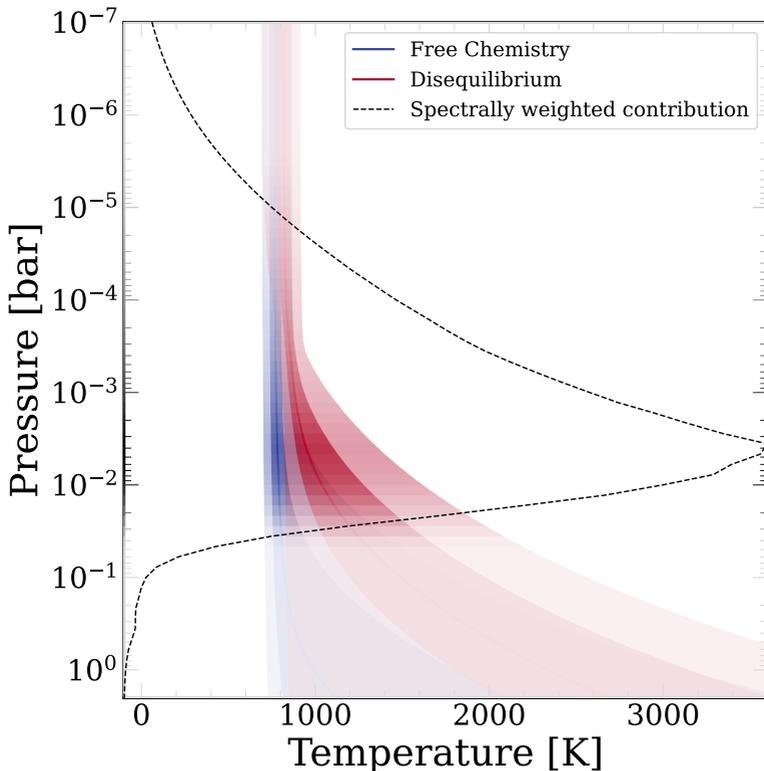


Figure 9.3: Thermal profile of the terminator region of WASP-39 b for the free chemistry retrieval (blue) and the disequilibrium chemistry retrieval (red). The dark shading indicates the 1σ confidence interval of the temperature profile, while the light shading indicates the 3σ confidence interval. The colour saturation is weighted by the averaged transmission contribution function (dashed line), indicating the region of the atmosphere to which the transmission spectrum is sensitive.

Regardless of the combination used, the retrieved parameter values were broadly consistent with the known values, though allowing the radius to be retrieved generally resulted in a larger radius than measured from the white light curve. The inferred metallicity and C/O ratios were found to be consistent to within the measured uncertainties.

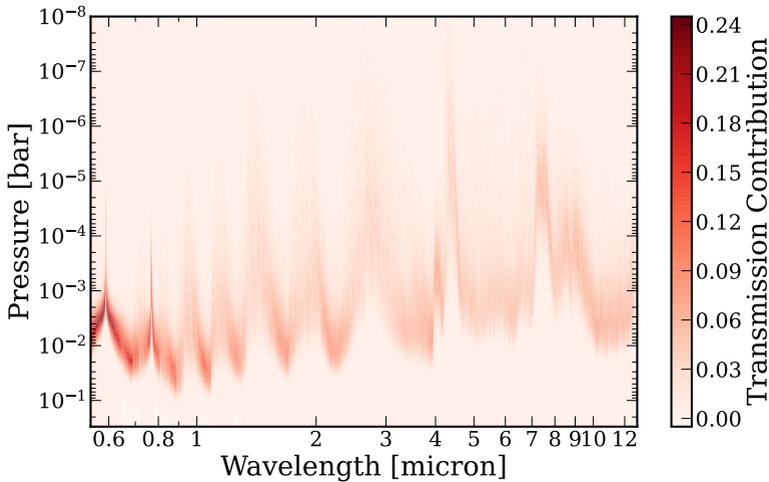


Figure 9.4: WASP-39 b transmission contribution function of the best-fit disequilibrium chemistry model. There is little sensitivity of the transmission spectrum to pressures higher than 10^{-2} bar, consistent with an opaque cloud deck located in that region.

9.5. Chemical disequilibrium

In addition to the free retrievals, we ran a series of disequilibrium chemistry retrievals using pRT. Our disequilibrium chemistry model uses an equilibrium chemistry table computed using `easyChem` (Mollière et al., 2017), interpolating along axes of metallicity, C/O, pressure and temperature. CO-H₂O-CH₄ disequilibrium is parameterised as a quench pressure, above which the mass fraction abundances become vertically constant, analogous to the homogenization of abundances caused when the vertical mixing timescale is smaller than the chemical reaction timescale (Zahnle & Marley, 2014). As SO₂ is photochemically produced, it is included as a freely retrieved species, which is required to fit the absorption feature at 4.2 μm . For these retrievals, we use 4000 live points and a sampling efficiency of 0.8. As with the free retrievals, we use the (Guillot, 2010) temperature profile, and the same set of opacity sources.

We performed disequilibrium retrievals from 1–12 μm , using precomputed offsets and error inflation. The overall goodness-of-fit was marginally better than the free retrievals, with $\chi^2/\nu = 1.04$. We find that by the Bayes factor, the disequilibrium model is strongly favoured over the free chemistry model, with

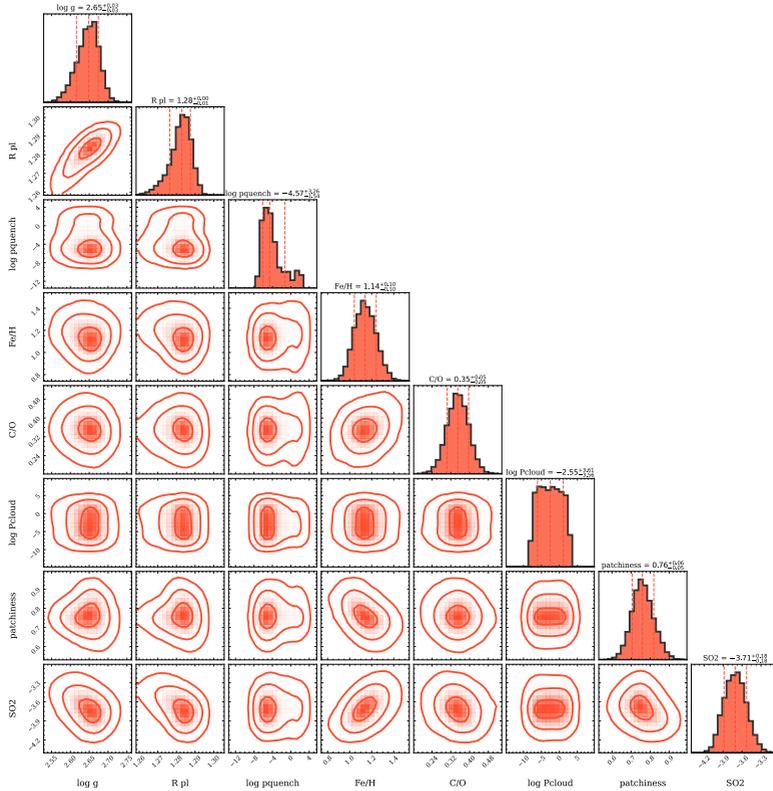


Figure 9.5: Posterior probability distributions for the WASP-39 b disequilibrium chemistry retrieval.

$\log_{10} = 5.5$. The disequilibrium model finds a slightly warmer temperature of 820 ± 38 K in the isothermal region of the atmosphere as compared to the free chemistry retrieval. The temperature profile also transitions to an adiabatic profile at lower pressures. By averaging the transmission contribution function shown in Figure 9.4 over wavelength, we find that there is some sensitivity of the transmission spectrum to the adiabatic temperature gradient at around 10^{-2} bar. The median retrieved $[M/H]$ is 1.14 ± 0.10 , which is compatible with estimates from self-consistent model grids (Alderson et al., 2023; Powell et al., 2024). The inferred C/O is 0.350 ± 0.046 . The planet is $76 \pm 5\%$ covered in clouds: including a grey cloud with a poorly constrained location ($\log P_{\text{top}} = -2.6 \pm 3.6$), as well as an enhanced Rayleigh scattering slope:

$\log \kappa_0 = -0.50 \pm 0.35$, $\gamma_{\text{scat}} = -1.24 \pm 0.22$. We detect photochemically produced SO_2 with a log volume mixing ratio of -5.0 ± 0.2 . The quench pressure is found to be high up in the atmosphere ($\log P_{\text{quench}} = -4.6 \pm 2.4$, indicating that the atmosphere is essentially in equilibrium. However, without a good parameterisation of photochemistry, we cannot properly interpret the chemical state of the upper atmosphere. The full set of retrieved parameters are presented in Figure 9.5.

9.6. Impacts data selection

Retrieving on each instrument individually highlights which parameters are constrained by specific regions of parameter space. We performed free chemistry retrievals on each instrument, and list the retrieved volume mixing ratios (VMRs) in Table 9.4. Retrieving on subsets of the data leads to relatively consistent results, with the metallicity typically ranging from 10-40 \times the solar value, depending on which combination of data is used. The metallicity inferred using only NIRCcam is significantly outlying from the rest of the individual instrument retrieval, at $[M/H] = -1 \pm 1$. Individual chemical abundances estimates vary greatly depending on whether a given instrument covers a particular molecular band or not. For example, the NIRCcam and NIRSpect/G395H instruments find only minimal traces of CO, while NIRSpect/PRISM and the MIRI/LRS identify it as among the most abundant species, despite CO opacity only contributing weakly between 5–6 μm in the LRS band. This highlights the importance of broad wavelength coverage when performing retrievals, and it is only with the full wavelength coverage that reliable abundance estimates can be made for a wide array of molecular species.

9.7. Degenerate parameterisations

It is crucial to understand how parameter choices impact retrieval results, particularly in the case where parameters are degenerate. Consider the surface gravity $\log g$ and the planet radius R_{pl} , which are defined at a reference pressure P_{ref} . R_{pl} and P_{ref} are fully degenerate. Setting the P_{ref} deeper in the atmosphere will result in a smaller radius. The surface gravity and planet radius are coupled through the definition of surface gravity g via

$$g = \frac{GM_{\text{pl}}}{R_{\text{pl}}^2}. \quad (9.1)$$

In principle, the planet radius is well defined by measurement of the white light curve measurement of the planetary transit, though this measurement does not

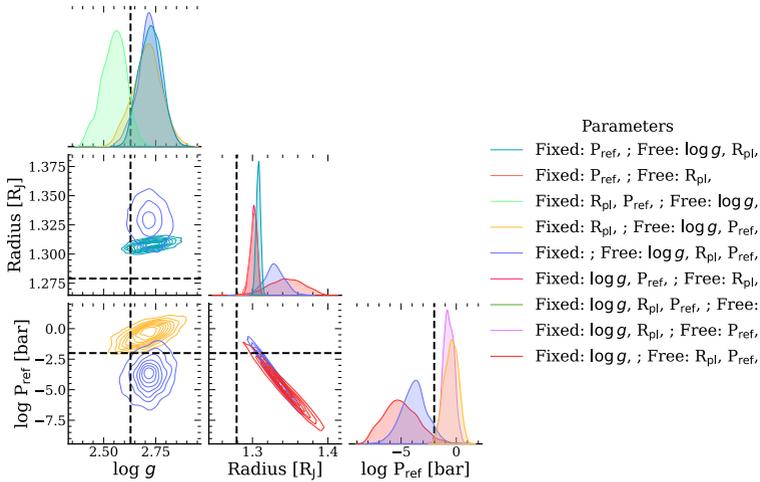


Figure 9.6: Corner plot comparing the impact of freely retrieving any of the three degenerate parameters of $\log g$, R_{pl} and P_{ref} . The black dashed lines indicate the fixed values used.

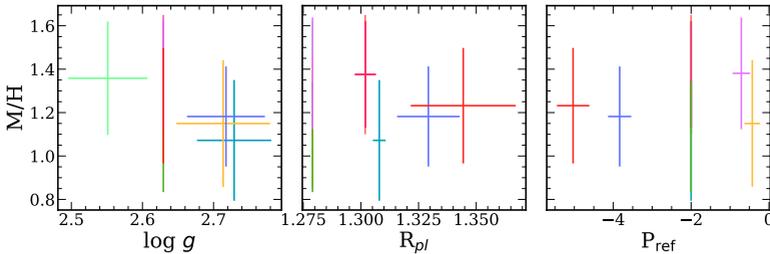


Figure 9.7: Scatter plot of the variation in retrieved $[M/H]$ due to the choice in keeping the three degenerate parameters of $\log g$, R_{pl} and P_{ref} fixed or free. The colour scheme is as in 9.6.

specify the pressure at which the radius is defined.

In order to test the impact of keeping these parameters, a series of retrievals was run, covering all permutations of keeping these parameters fixed or free. Figure 9.6 shows the results of these retrievals. Most retrievals find a surface gravity of $\log g = 2.72 \pm 0.05$, slightly higher than the measured value of 2.6294 from combining measurements of the planetary mass and radius. Conversely, if both the planet radius and pressure are fixed, the retrieved surface gravity is lower than the measured value, suggesting that a reference pressure of 0.01 may be too high in the atmosphere for the measured white light radius.

The retrieved radii for WASP-39 b are consistently larger than the measured white light value, falling into two groups at around $1.30 R_{\text{Jup}}$ and $1.34 R_{\text{Jup}}$. The radius measurements are more precise in cases where the reference pressure is fixed, and all of the cases with a fixed reference pressure favour the measurement of $1.30 R_{\text{Jup}}$. If the reference pressure is allowed to vary, the distribution is again bimodal. The more precise measurements are found by fixing the planet radius, and prefer values deeper in the atmosphere, around 0.3 bar. Overall, none of the retrieved posterior distributions are fully compatible with the measured surface gravity and radius. This may be due to additional degeneracies in the parameter space, such as the location of the cloud deck deep in the atmosphere.

While the retrieved values may be statistically discrepant from other measurements, the overall variation is small, and ultimately only has a minor impact on other retrieved parameters. Figure 9.7 shows how the inferred metallicity varies with keeping $\log g$, R_{pl} and P_{ref} fixed or free. In all cases, the metallicity measurements are compatible to within 1σ , though the inferred metallicity is anti-correlated with $\log g$. These parameter degeneracies were proposed as an explanation for why the metallicity found by free retrievals is somewhat higher than found by self-consistent grids. The most compatible results are achieved by freely retrieving $\log g$ and R_{pl} , $\log g$ and P_{ref} , or all three parameters. This experiment highlights the care that must be taken when defining retrieval setups, as the outcomes depend on the priors used.

9.8. Cloud modelling choices

Other studies have indicated the presence of silicate clouds in the transmission spectrum of hot Jupiter atmospheres. With temperatures of ~ 800 K in the terminator, silicate clouds are expected to be deep below region to which the WASP 39 b transmission spectrum is sensitive. However, other species such as KCl or Na_2S may condense at such temperatures. We perform retrievals on a combination of NIRSpec/PRISM and MIRI/LRS data to obtain the broad wavelength coverage required to assess cloud properties, using a treatment of

clouds based on [Ackerman & Marley \(2001\)](#). We compared these clouds, using physically measured optical constants, to the fiducial grey cloud model with an enhanced haze slope. The disequilibrium chemistry model with a Guillot PT profile was used.

For this set of retrievals we compared the impact of different cloud parameterisations, trying to determine if a) this choice can impact the inferred chemistry and thermal structure, and b) if we can identify the cloud composition as in [Grant et al. \(2023\)](#). We compared the patchy grey cloud with an enhanced scattering slope model as used in the free retrievals to patchy Ackermann-Marley clouds [Ackerman & Marley \(2001\)](#). This model is based on the settling of cloud particles as a function of their size, with a $\log K_{zz}$ parameter measuring the strength of vertical mixing, and an f_{SED} parameter that is used to set the sedimentation rate. We also retrieve a σ_{LN} , which sets the width of the log-normal particle size distribution. For these clouds, we use the optical constants associated with different compositions and particle geometries in order to determine if we can measure the cloud composition, specifically comparing crystalline MgSiO_3 , SiO_2 , Na_2S and KCl . The clouds condense out of the gas phase at the intersection of their condensation curves and the retrieved temperature profile. The Bayes factors for the different cloud compositions are listed in table 9.5. We found no evidence indicating that the clouds are composed of one of these species, and the grey cloud model with an enhanced scattering slope was strongly favoured by the Bayes factor ($> 5\sigma$). In addition to the retrievals performed on the combined PRISM and LRS datasets, we performed an additional comparison using the full, combined dataset with precomputed offsets and error inflation. Again, no evidence for silicate clouds was found.

| Cloud | $\Delta \log_{10} \mathcal{Z}$ |
|------------------|--------------------------------|
| Fiducial | 0 |
| KCl | -2 |
| SiO_2 | -3.2 |
| MgSiO_3 | -9.7 |

Table 9.5: Bayes factors for different cloud compositions, comparing the [Ackerman & Marley \(2001\)](#) model to the fiducial cloud model for the petitRADTRANS disequilibrium chemistry retrievals.

9.9. Summary

While pRT was already a widely used tool in the exoplanet community, the development of the pRT retrieval package has enabled fast, easy characterisation of a diverse array exoplanet atmospheres. The use of a highly flexible approach

combined with the computational efficiencies found in the use of *c-k* opacities and nested sampling ensure that *pRT* is prepared to deal with the current surfeit of data available from observatories such as *JWST*. Using these retrieval techniques, I was able to characterise the atmosphere of the benchmark hot Saturn, WASP-39 b. The results found through this method were consistent both with other retrieval codes and with self-consistent models. The metal rich atmosphere with a low C/O ratio was confirmed, as was the robust detection of CO₂ and SO₂. Novel cloud modelling comparisons were performed, finding that unlike other transiting planets such as WASP-17 b, a patchy grey cloud model with an enhanced haze scattering slope provides the best fit to the observed spectrum. Going forward, further developments of the retrieval package may include reduced computation time, 3D atmospheric structures, and fits to time variable data, bringing our understanding of exoplanet atmospheres ever closer to that of the dynamic worlds of our own solar system.

We gratefully acknowledge contributions to *petitRADTRANS* from Eleonara Alei, Karan Molaverdikhani, Sam de Regt, Francois Rozet, Aaron David Schneider, Tomas Stolker, Nick Wogan and Mantas Zilinskas. In addition, huge thanks to Luis Welbanks and the Transiting ERS team, and in particular the retrieval model synthesis team. It's been a pleasure to open the black box of retrievals working on this project, and hopefully it will provide an excellent template for future characterisation work.

| Species | log VMR | | | | | |
|------------------|---------------|--------------|--------------|---------------|---------------|---------------|
| | PRISM | NIRCAM | SOSS O1 | G395H | LRS | Full |
| H ₂ O | -1.55 ± 0.286 | -4.85 ± 1.48 | -1.6 ± 0.266 | -1.82 ± 0.289 | -9.78 ± 1.72 | -1.52 ± 0.237 |
| CO | -1.81 ± 0.378 | -8.02 ± 2.79 | -5.53 ± 3.31 | -7.41 ± 2.77 | -2.11 ± 1.25 | -2.1 ± 0.607 |
| CO ₂ | -2.8 ± 0.297 | -5.72 ± 1.16 | -3.65 ± 1.86 | -2.85 ± 0.331 | -9.6 ± 1.96 | -3.08 ± 0.256 |
| CH ₄ | -5.55 ± 0.316 | -10.3 ± 1.43 | -9.22 ± 1.87 | -8.87 ± 1.79 | -3.87 ± 0.625 | -9.09 ± 1.89 |
| H ₂ S | -2.99 ± 0.26 | -9.09 ± 2.2 | -8.27 ± 2.62 | -3.22 ± 0.267 | -8.9 ± 2.3 | -3.47 ± 0.678 |
| SO ₂ | -5.32 ± 0.809 | -9.21 ± 2.07 | -7.95 ± 2.81 | -5.03 ± 0.247 | -9.3 ± 2.17 | -4.79 ± 0.238 |
| NH ₃ | -9.15 ± 1.96 | -9.72 ± 1.73 | -9.51 ± 1.74 | -9.14 ± 1.83 | -9.97 ± 1.58 | -9.67 ± 1.69 |
| HCN | -7.51 ± 2.62 | -10 ± 1.63 | -7.1 ± 2.74 | -9.78 ± 1.76 | -9.75 ± 1.78 | -9.22 ± 2.03 |
| Na ^a | -2.51 ± 0.442 | -7.53 ± 2.95 | -7.41 ± 2.85 | -7.4 ± 2.78 | -7.72 ± 2.84 | -2.83 ± 0.68 |
| K | -7.5 ± 2.89 | -7.72 ± 2.87 | -8.83 ± 2.25 | -7.65 ± 2.81 | -8.75 ± 2.41 | -5.14 ± 0.775 |
| [M/H] | 1.6 ± 0.3 | -1 ± 1 | 1.2 ± 0.3 | 1.1 ± 0.3 | 1 ± 0.6 | 1.5 ± 0.3 |

Table 9.4: Volume mixing ratios for each dataset, free chemistry retrieval.



IMPACTS OF HIGH-CONTRAST IMAGE PROCESSING ON ATMOSPHERIC RETRIEVALS.

This chapter has been reproduced from [Nasedkin et al. \(2023\)](#). I am the first author of that work, supervised by Paul Mollière. The other coauthors provided suggestions on how to use the various data processing tools and provided comments on the manuscript.

“We succeeded in taking that picture [from deep space], and, if you look at it, you see a dot. That’s here. That’s home. That’s us.”

Carl Sagan, 1994



Abstract

Many post-processing algorithms have been developed in order to better separate the signal of a companion from the bright light of the host star, but the effect of such algorithms on the shape of exoplanet spectra extracted from integral field spectrograph data is poorly understood. The resulting spectra are affected by noise that is correlated in wavelength space due to both optical and data processing effects. Within the framework of Bayesian atmospheric retrievals, we aim to understand how these correlations and other systematic effects impact the inferred physical parameters. We consider three algorithms (KLIP, PynPoint, and ANDROMEDA), optimising the choice of algorithmic parameters using a series of injection tests on archival SPHERE and GPI data of the HR 8799 system. The wavelength-dependent covariance matrix was calculated to provide a measure of instrumental and algorithmic systematics. We perform atmospheric retrievals using petitRADTRANS on optimally extracted spectra to measure how these data processing systematics influence the retrieved parameter distributions. The choice of data processing algorithm and parameters significantly impact the accuracy of retrieval results, with the mean posterior parameter bias ranging from 1 to 3σ from the true input parameters. Including the full covariance matrix in the likelihood improves the accuracy of the inferred parameters, and cannot be accounted for using ad hoc scaling parameters in the retrieval framework. Using the Bayesian information criterion and other statistical measures as heuristic goodness-of-fit metrics, the retrievals including the full covariance matrix are favoured when compared to using only the diagonal elements.

10. INTRODUCTION

The field of high contrast imaging (HCI) has advanced dramatically over the last two decades. From the first detection of 2M 1207 b (Chauvin et al., 2005) to the ongoing large surveys such as GPIES (Nielsen et al., 2019) and SHINE (Desidera et al., 2021; Langlois et al., 2021; Vigan et al., 2021), we have seen improvements in instrumentation, adaptive optics and data processing that have led to the discovery of numerous new exoplanets. Such surveys have established the rarity of giant, widely separated companions, finding that $< 10\%$ of high-mass stars have planetary mass companions between 10-100 AU. However, much of this work has remained focussed on the detection of new companions at higher contrast ratios and smaller angular separations. The spectroscopic characterisation of known planets has seen less dedicated effort - to date there has not been a uniform survey of known objects to present a homogeneous sample of spectroscopic measurements. This can lead to systematic discrepancies between measurements made with different instruments, and challenges in fitting datasets with different spectral resolutions (Xuan et al., 2022). Such biases may impact the conclusions made from population studies, such as the exploration of the C/O ratios of the directly imaged planet population in Hoch et al. (2023). Individual characterisation efforts have nevertheless led to intriguing findings: measurements of water and carbon monoxide abundances in the HR 8799 planets (Konopacky et al., 2013; Lavie et al., 2017; Wang et al., 2020a), precise constraints on the C/O ratio and metallicity of β Pictoris b (Gravity Collaboration et al., 2020), measurements of isotope ratios Zhang et al. (2021) and the detection of a dusty envelope around PDS 70 b and c (Wang et al., 2021; Benisty et al., 2021).

This characterisation work remains challenging. Extensive post-processing is required to extract the faint signal of the target. Even in the most careful analysis systematic biases usually remain from both instrumental and processing effects. Integral field spectrograph (IFS) measurements introduce correlated noise as a function of wavelength due to pixel cross talk, interpolation effects and imperfect adaptive optics correction (speckles). Greco & Brandt (2016) provide a method for empirically estimating the correlation from IFS data. They demonstrated that accounting for such correlations is necessary when analysing exoplanet atmospheres, and failing to do so leads to biased and overconfident posterior distributions on measured parameters. Efforts such as the Exoplanet Imaging Data Challenge (Cantalloube et al., 2020a) explored the detection abilities of a suite of HCI algorithms but a systematic algorithmic comparison for spectral characterisation has not yet been performed.

Most post-processing techniques are based on Angular Differential Imaging

(ADI) (Marois et al., 2006, 2008), where the telescope is pupil stabilised and the field is allowed to rotate. This provides differential motion of the planet over the course of the observations and allows for the removal of stellar speckles by derotating and stacking the resulting images. Ongoing development of this method has been largely driven by the goals of increasing sensitivity at small angular separations. To this end, different algorithms have been developed to maximise the information available in imaging datasets, leveraging spatial and spectral information in order to separate the faint planet signal from the bright host star. Kiefer et al. (2021) explored how different approaches impact the signal-to-noise (S/N) of IFS observations, but did not examine the impact of the processing on the extracted spectral shape.

The use of atmospheric retrievals to study directly imaged planets is relatively new, with only a small but growing selection of targets being subject to such an analysis (e.g. Lee et al., 2013; Lavie et al., 2017; Mollière et al., 2020; Gravity Collaboration et al., 2020; Brown-Sevilla et al., 2023; Whiteford et al., 2023). While the effects of systematics are well understood for transmission spectroscopy using HST (Ih & Kempton, 2021), with significant efforts extending this to JWST (Barstow et al., 2015; Rocchetto et al., 2016; Lacy & Burrows, 2020), the impact of systematic uncertainties in ground-based high-contrast data on atmospheric retrievals has not been thoroughly explored. Even in the era of *JWST*, understanding systematics is critical to interpreting model fits to data. Ground-based observations will remain a key component of this understanding due to their higher spectral and angular resolution that cannot yet be achieved from space.

In this work we explore the systematic effects introduced through high-contrast data processing on the retrieval of atmospheric parameters. The details of our example datasets used are described in Section 11. Section 12 outlines our methods, exploring the different algorithm tested in Section 12.1, together with the measurement and interpretation of the covariance matrix in Sections 12.2 and 12.4. We determined the optimal parameters for spectral extraction through the injection and recovery of synthetic companions into the data in Section 13. The results of our retrieval comparisons are described in Section 14, while the implications and limitations of these results are discussed in Section 15.

11. OBSERVATIONS

While the first goal of our study is to demonstrate the effects that post-processing algorithms can have on inferred atmospheric parameters for general high-contrast spectroscopy, we still had to select demonstration datasets. We chose GPI and SPHERE observations of the well-known four-planet system in HR 8799 (Marois et al., 2008, 2010), where discrepancies between GPI and SPHERE datasets, covering the same wavelength range, had already been noted (Lavie et al., 2017; Mollière et al., 2020). HR 8799 has seen extensive photometric and spectroscopic observing campaigns, (e.g. Konopacky et al., 2013; Zurlo et al., 2016; Lavie et al., 2017; Greenbaum et al., 2018; Gravity Collaboration et al., 2019; Mollière et al., 2020; Wang et al., 2020a; Ruffio et al., 2021; Wang et al., 2023). The importance of this system, together with the abundance of high contrast data from multiple instruments make it an ideal object of study for our purposes. As a benchmark target, the companions have luminosity and spectra typical of this class of low surface gravity object and are representative of the current directly imaged exoplanet population.

SPHERE

The SPHERE data were taken during the commissioning run of the SPHERE instrument (Beuzit et al., 2008, 2019) in 2014, and were originally presented in Zurlo et al. (2016). It remains the best YJH band spectrum of HR 8799 to date in terms of signal-to-noise and spectral resolution. IFS frames in the YJH band were taken with a series of both 60 s and 100 s integrations, using pupil-stabilised observations to allow for ADI post-processing. Total field rotations of 15.37° and 29.65° were observed for the 60 s data cube and for the 100 s data cube, respectively. To compensate for the difference in exposure time, we multiply each 60 s exposure by a factor of $100/60$, in order to process the data as a whole. We rereduced the SPHERE data using the pipeline described in Vigan (2020): details of which are described in Appendix 16.A.

GPI

The GPI (Macintosh et al., 2014) observations of HR8799 were originally published in Greenbaum et al. (2018) and were taken on 17 November 2013, 18 November 2013, and 19 September 2016 for the K1, K2, and H bands respectively. As with the SPHERE data, the telescope was pupil-stabilised to take advantage of ADI post-processing. These were reduced using the standard GPI reduction pipeline (version 1.4.0). The median seeing of the observations

Table 11.1: Epochs of SPHERE ([1] Zurlo *et al.* (2016) 60.A-9249(C)) and GPI ([2] Greenbaum *et al.* (2018) GS-2015B-Q-500-1394) data used.

| Instrument | Date | Band | $\lambda/\Delta\lambda$ | Field Rotation [°] | Med. Seeing [as] | DIT [s] | NEXP | Ref. |
|-------------------|-------------|-------------|-------------------------------------------|---------------------------|-------------------------|----------------|-------------|-------------|
| SPHERE | 2014-08-12 | YJH | 29 | 29.65 15.37 | 0.87 | 100 60 | 32 48 | [1] |
| GPI | 2016-09-19 | H | 45 | 20.93 | 0.97 | 60 | 60 | [2] |

was $0''.97$; the observing conditions are more thoroughly described in [Ingraham et al. \(2014\)](#). While data were taken in the H, K1, and K2 bands of GPI, we only considered the H band observations due to the low S/N of the K-band observations. The observations from both GPI and SPHERE are summarised in Table 11.1.

11.1. Data preprocessing

In order to reduce the systematic variation between the datasets, we first rereduced the data with up-to-date pipelines. For both the SPHERE and GPI datasets, we then preprocess the IFS cubes using the Vortex Image Processing (VIP) library in order to select the optimal frames for further ADI processing. The `cube_detect_badfr_correlation` function computes the similarity between each frame and a reference frame in order to identify frames that are outliers when compared to the rest of the sequence. We choose the frame which maximises the mean similarity of all frames as the reference frame, and remove the most different 12% of frames from each the SPHERE and GPI datasets. Such variation in the data is typically due to changing observing conditions, introducing effects into the data such as the low-wind effect [Milli et al. \(2018\)](#) or the wind-driven halo [Cantalloube et al. \(2020b\)](#). This threshold is sufficient to remove frames which are significantly outlying and visually show differences when compared to the a typical frame. This leaves 69 ADI frames for the SPHERE dataset, and 51 for the GPI H-band dataset.

11.2. Stellar model for flux extraction

In order to obtain the absolute flux of the companions we use a model of the stellar spectrum to flux-calibrate the contrast measurements. HR8799 is an F0+VkA5mA5 C star ([Gray et al., 2003a](#)) located 41.3 ± 0.2 pc ([Gaia Collaboration et al., 2018](#)). Stellar photometry of HR 8799 from WISE and 2MASS is used to fit model stellar spectrum ([Cutri et al., 2021, 2003](#)). We exclude data points beyond $5 \mu\text{m}$ so that the fit is not impacted by the infrared excess from the debris disk ([Su et al., 2009; Faramaz et al., 2021](#)). Using the species package ([Stolker et al., 2020](#)), we fitted a BT-Nextgen model to the photometry within our wavelength range of interest. The best-fit model has parameters of $T_{\text{eff}} = 7200$ K, $\log g = 3.0$ and $[\text{Fe}/\text{H}] = 0.0$, slightly cooler than the models used in previous studies ([Zurlo et al., 2016; Greenbaum et al., 2018](#)). The full set of stellar parameters is listed in table 11.2. This spectrum is normalised to a 10 pc distance. The model is convolved to the instrumental spectral

Table 11.2: Stellar properties of HR 8799 A.

| HR 8799 | | |
|--------------------------|----------------------------------------|-------------|
| Parameter | Value | Note |
| α [J2000] | $23^h 07^m 28.7157^s \pm 0.0685^s$ | [1] |
| δ [J2000] | $+21^\circ 08' 03.3021'' \pm 0.0799''$ | [1] |
| μ_α [mas/yr] | 108.301 ± 0.168 | [1] |
| μ_δ [mas/yr] | -49.480 ± 0.152 | [1] |
| $\bar{\omega}$ [mas] | 24.2175 ± 0.0881 | [1] |
| d [pc] | 41.2925 ± 0.1502 | [1] |
| RV [km s ⁻¹] | -12.60 ± 1.4 | [2] |
| Spectral Type | F0+VkA5mA5 C | [3] |
| T_{eff} [K] | 7200 ± 50 | [4] |
| log g [cgs] | 3.0 ± 0.25 | [4] |
| [Fe/H] [dex] | 0.0 ± 0.2 | [4] |
| R_* [R_\odot] | 1.496 ± 0.0054 | [4] |
| L_* [L_\odot] | 5.230 ± 0.0498 | [4] |
| C/O | $0.54^{+0.12}_{-0.09}$ | [5] |

Notes. [1] [Gaia Collaboration et al. \(2018\)](#). [2] [Gontcharov \(2006\)](#). [3] [Gray et al. \(2003a\)](#). [4] BT-NextGen best-fit to photometry ([Hauschildt et al., 1999](#)). [5] [Wang et al. \(2020a\)](#).

HR 8799 is a λ Boötis star, for further discussion see [Mollière et al. \(2022\)](#).

resolutions and binned to the instrumental wavelength channels to allow for spectrophotometric calibration of contrast measurements.

In order place measurements of planet properties in context it is also necessary to understand the properties of the host star. [Wang et al. \(2020a\)](#) used HARPS observations to directly measure the C and O abundances of the star,

finding a C/O ratio of $0.54^{+0.12}_{-0.09}$. HR 8799 is a λ Boötis star, known to be depleted in iron (Gray & Corbally, 2002). Consistent with this, the authors fit Fe I and Fe II lines, finding a metallicity of $[\text{Fe}/\text{H}] = -0.52 \pm 0.08$. Both of the carbon and oxygen abundances were measured to be consistent with solar composition, suggesting that the iron metallicity is not representative of the bulk stellar composition, and that our BT-NextGen is still applicable. At the low spectral resolution considered in this study the metallicity does not significantly impact the SED of the star and variations in its measurement will not affect the calculation of the results, though will ultimately impact the context – and thus interpretation – of planetary metallicity measurements.

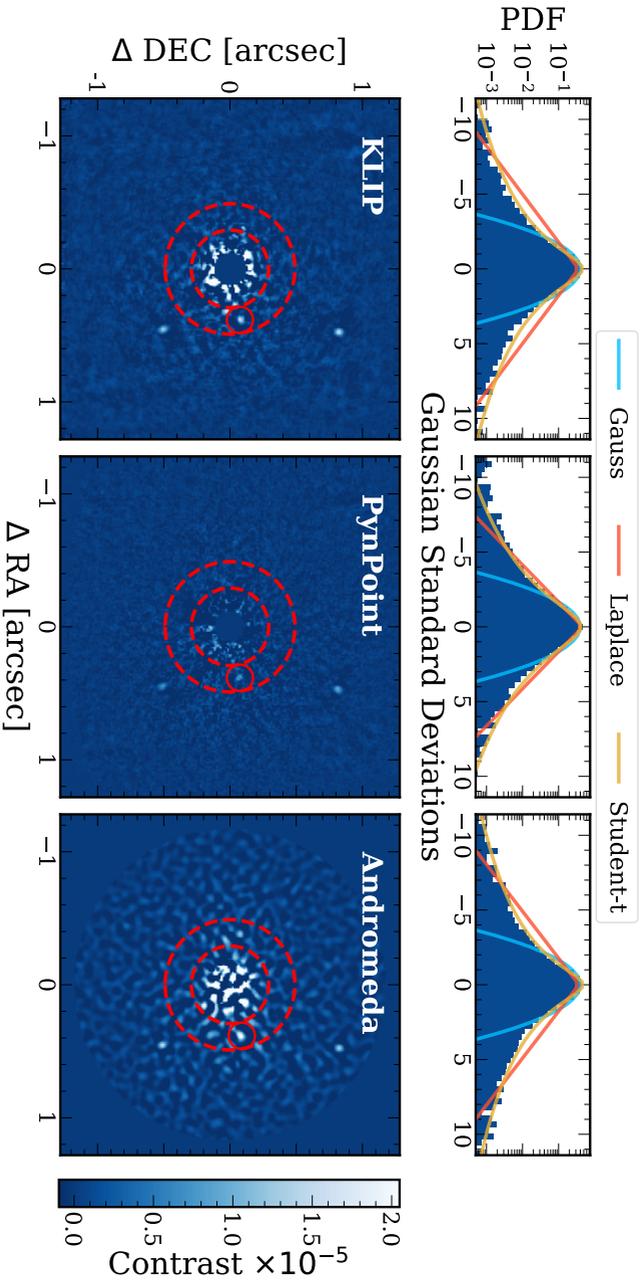


Figure 11.1: Single wavelength channel from the GPI H-band data of the HR 8799 system, post-processed with KLIP (left), PynPoint (center) and Andromeda (right). The solid red circle denotes the position of HR 8799 e as computed using the pyKLIP astrometry module. Marked in red dashed lines are the regions which are compared when finding the mean correlation between different wavelength channels. At the top of the figure, histograms of the residuals are plotted in units of σ . The light blue line is a Gaussian fit with the width defined as the Gaussian standard deviation of the residual frame below.

12. ADI DATA PROCESSING

The atmospheric properties of directly imaged exoplanets are presently accessible only through their thermal emission. For directly imaged planets, these spectra are usually obtained through low to moderate resolution IFS instruments equipped with coronagraphic optics. IFS data is complex, with a large array of systematic and random noise effects imprinted onto the data. Cross talk between neighbouring pixels due to optical effects (Antichi et al., 2009; Larkin et al., 2014) and scattered light can introduce correlations in wavelength space. Once the data has been reduced from raw detector frames to data cubes, quasi-static stellar speckles – light of the host star scattered by the telescope optics – is the dominant noise source (Marois et al., 2005, 2008). ADI processing is used to remove the stellar PSF and speckle noise, taking advantage of the stability of the PSF over time (Marois et al., 2006). ADI exploits the rotation of the planet through the frame, which produces a signal that is different from the stellar speckles, which remain fixed in position. By derotating and stacking the images, the residual speckles following post-processing are averaged out, while the planet signal is enhanced. The stability assumption is not without flaws, as the PSF varies due to thermal variation in the telescope, short and long-term atmospheric changes and more (Milli et al., 2016), but in practice it is robust enough to allow for planet detection. Obtaining an exoplanet spectrum is generally achieved by applying an ADI algorithm to each spectral channel on a 4D cube of IFS data. Modern ADI processing is more sophisticated than simply derotating and stacking the images, but the algorithms generally fall into three broad categories:

Speckle subtraction methods attempt to directly subtract the residual stellar speckles from each frame of the image cube. The planet signal is then measured either through aperture photometry or through fitting a model of the PSF to the signal and minimising the residuals. This is the most commonly used method, and includes algorithms such as (Template) Locally Optimised Combination of Images (LOCI and TLOCI, Lafrenière et al. 2007; Maire et al. 2012; Marois et al. 2014), low rank plus sparse decomposition (LLSG, Gomez Gonzalez et al. 2016) and various implementations of principal component analysis (PCA) based methods, including Karhunen-Loève Image Projection (KLIP, Soummer et al. 2012) and Standardised Trajectory Intensity Mean (STIM, Pairet et al. 2019). Such PCA-based methods construct an ordered library of principal components of the data: low orders describe the most important components of the stellar PSF, while higher orders describe high frequency noise. By building a library to describe the host star PSF, it can be more effectively subtracted from each frame before stacking the images, improving the S/N of the companion.

Inverse methods such as ANDROMEDA (Mugnier et al., 2009; Cantalloube et al., 2015), PACO (Flasseur et al., 2018) and TRAP (Samland et al., 2021) use likelihood minimisation to directly estimate the position and contrast of proposed signal at each point in the field. To do this, a parameterised forward model of the companion signature is fit to the data, and the parameters are optimised through a likelihood minimisation process. This yields a statistical interpretation of the residuals, and provides confidence region estimates that provide a metric for detection significance, under varying assumptions of the noise properties of the data.

Finally, *supervised machine learning* methods (Gomez Gonzalez et al., 2018; Hou Yip et al., 2019; Gebhard et al., 2022) are trained on large sets of data with injected targets and learn how to identify the presence of a companion in an image. These methods typically only produce a binary maps where a planet is either detected or not, and do not measure the strength of the planet signal. Thus these methods have not generally been used for exoplanet characterisation.

12.1. Post-processing algorithms

We chose to compare three widely used ADI techniques in order to determine the impact of such post-processing on the spectral shape and noise properties of the extracted exoplanet spectrum. In order to compare a diverse range of techniques we chose to use KLIP and PynPoint, which are different flavours of PCA-based speckle subtraction methods, and Andromeda, which is an inverse method. As our goal is to understand the impact of systematic effects, we chose these algorithms for their broad community use, typifying the effects likely present in existing work. A more complete examination of the diversity of algorithms, including spectral differential imaging (SDI) and ADI+SDI algorithms will be explored using a larger set of data in a forthcoming publication based on Phase 2 of the Exoplanet Imaging Data Challenge.

In order to assess our choice of algorithm, we compared extractions of known injected spectra at different positions and contrasts in order to optimise the parameter selection for extracting the true spectrum. In this section we present the specific steps we took to reduce SPHERE and GPI datasets of the HR 8799 system using each of these algorithms. While a wide range of parameters were explored, Table 12.1 summarises the parameter choices used in this analysis for each algorithm.

KLIP

KLIP is a PCA-based speckle subtraction algorithm, described in [Soummer et al. \(2012\)](#); [Wang et al. \(2015\)](#) and [Pueyo \(2016\)](#). A Karhunen-Loève transform of an optimised combination of reference images is used to define the basis of eigenimages, onto which the science frames are projected. Often this set of reference images is derived from the science observations, but in principle can be any representative measurements of the PSF. Mathematically, this is equivalent to building the basis of principal components. This projection is subtracted from the science frames in order to produce the final residual image. A forward model of the PSF is then injected in order to measure the position and contrast of a detected companion.

Our choice of KLIP parameters is guided by [Pueyo \(2016\)](#) and [Greenbaum et al. \(2018\)](#). For this study, we use KLIP in ADI mode. Comparison tests showed that the full ADI+SDI mode provided modest increases in S/N at low contrasts, but the overall shape of the spectrum remained similar. We set a region around the proposed location of the planet extending 13 pixels radially in each direction, and 18° on either side of the planet. The flux overlap parameter is used to set the aggressiveness of the subtraction, using a value of 0.1. Fixing these parameters may result in sub-optimal spectral extraction, particularly at very small separations where the rotational movement of the planet through the frame is small. However, we are primarily concerned with the overall trends in the spectral extractions and noise properties across different tools, and do not attempt to fine tune each algorithm for each individual injections.

We use the pyKLIP astrometric measurement tools to compute the location of each target within the field of view, which is used to provide our initial estimate for the planet position for each of the algorithms we consider. The extracted spectrum is highly sensitive to the inferred companion position, and so we use the KLIP astrometry as the location for all three algorithms. [Pueyo \(2016\)](#) outlines the procedure to extract the spectrum from KLIP processed data using the forward model extraction tool. For each target at each wavelength, a forward model is generated from the unsaturated PSF obtained during the observation. KLIP processing is then applied to subtract the stellar PSF and measure the contrast of the companion. This is converted into a flux measurement using the BT-NextGen model of the host star spectrum from Section 11.2.

PynPoint

PynPoint is a Python package designed for high contrast imaging data processing ([Amara & Quanz, 2012](#); [Stolker et al., 2019](#)). The standard PSF subtraction method used in the package is based on full-frame PCA. We pro-

cess each wavelength channel of the IFS data independently, filtering for bad pixels and running ADI-PCA on each stack of images. In contrast to KLIP, which builds a model of principal components in a local region near the planet, PynPoint builds its PC library from the full available field of view. The central $0.''12$ of each frame is masked out, due to the large residuals close to the host star.

Following the PSF subtraction, a PSF model with negative flux is injected at the position of the planet of interest, which is known from previously computed KLIP astrometry. The PSF model for the planet is simply the stellar PSF, which is either derived from satellite spots (for GPI data) or from unocculted observations of the host star (for SPHERE data). The position and magnitude of the negative planet are iteratively fit to the data using a simplex minimisation routine to minimise the χ^2 between the PSF model and the data. The minimisation is considered within an aperture with a radius of 4 pixels around the proposed location of the planet. The iteration continues until a tolerance of 0.01 is reached for both the planet position and contrast in magnitude units. We allow the planet position to vary by up to 3 pixels (offset) from the initial estimate from pyKLIP astrometry. This produces a best-fit value of the position and contrast-magnitude of the planet. While we allowed the number of principal components used to vary from 1 to 25, we found that the extraction quality degraded substantially after 15 components, which sets the upper bound we present in this work. This is then converted from magnitude to contrast, and multiplied by the BT-NextGen stellar model of Section 11.2 to find the absolute flux of the planet.

Andromeda

ANDROMEDA (ANgular DiffeRential OptiMal Exoplanet Detection Algorithm) is a maximum likelihood estimation algorithm for ADI data, and estimates the position and flux of point sources within the field of view (Mugnier et al., 2009; Cantalloube et al., 2015). We run the VIP implementation of the algorithm on each wavelength channel independently, and combine the extracted contrast and standard deviation to build the planet spectrum. ANDROMEDA begins by high-pass filtering the data to remove large spatial scale structure from each data frame. This step induces signal loss, and we chose a value of 0.3 for the filtering fraction parameter, leading to a $\sim 20\%$ energy loss as in Figure 1 of Cantalloube et al. (2015). We calculate the oversampling parameter for each wavelength channel to ensure the sampling is constant across wavelength, and additionally use this parameter to determine the outer working angle, which is

| Parameter | Value |
|----------------|-------------------------------------|
| pyKLIP | |
| nPC | 1 – 25 |
| flux_overlap | 0.1 |
| highpass | True |
| maxnumbasis | 150 |
| mode | ADI |
| PynPoint | |
| nPC | 1 – 15 |
| merit | Gaussian |
| aperture | 4 px |
| tolerance | 0.01 |
| cent_size | 0."12 |
| offset | 2 px |
| ANDROMEDA | |
| filtering_frac | 0.35, 0.30 |
| min_sep | 0.45 λ/d , 0.25 λ/d |
| width | 0.8 λ/d , 1.2 λ/d |
| iwa | 2.0 λ/d , 1.0 λ/d |
| owa | 60/ S , 45/ S |
| opt_method | lsq |

Table 12.1: Parameters used for each of the algorithms considered. Parameters that were not varied were set based on previously reported values in literature (Pueyo, 2016; Greenbaum et al., 2018; Cantalloube et al., 2015). For ANDROMEDA, the first column of parameters was used for the SPHERE data, and the second column for the GPI data. The oversampling parameter S is defined in eqn. 12.1. Further information about each of these parameters is available in the documentation of each package.

provided in λ/D . The oversampling parameter, S is defined to be

$$S = \frac{\lambda}{2\sigma_{\text{px}}D} \quad (12.1)$$

for pixel scale σ_{px} , telescope diameter D and wavelength λ .

Internally to ANDROMEDA pairs of images are chosen such that they are as close together in time as possible to preserve speckle self-similarity, while still ensuring movement of the proposed companion in order to avoid self-subtraction. This is done on an annular basis, as the motion of the planet depends on the separation from the host star. A scaling factor γ is fit using a least squares method to ensure that the mean of the intensity distribution of both images in the pair is equal. Using the assumption that the residual noise is white and Gaussian, ANDROMEDA then can perform a likelihood test to identify the presence of a companion, by minimising the difference between the residuals and a model of the companion signal.

Among the outputs of this algorithm are a contrast map, where each pixel represents the contrast of the planet if it was centred on that pixel, and a standard deviation map, specifying the uncertainty associated with each contrast estimate. This is different from the output of a speckle-subtraction algorithm, where the flux of an object must be estimated through aperture photometry or via fitting a PSF model to the residuals. To extract the spectrum, we sum the S/N map along the wavelength axis, and identify the maximum S/N pixel in a 10 pixel box around the known position of the planet. We then use this location to measure the contrast and standard deviation as a function of wavelength.

12.2. Spectral covariance estimation

Both high contrast imaging and IFS observations present challenges when deriving robust uncertainty estimates, as correlations are naturally present in the data. Due to aberrations in the telescope optics, imperfect correction for atmospheric turbulence from the adaptive optics systems, and imperfect stellar PSF subtraction, speckles from the stellar PSF are the dominant noise source for AO assisted, high-contrast datasets (Marois et al., 2006). These speckles move radially as a function of wavelength, scaling with the size of the stellar PSF. This induces a correlation between wavelength channels, as a speckle will take several channels of movement to pass over a pixel at a fixed separation. Crosstalk – light from a single lenslet in the lenslet array diffracting into neighbouring channels – will also couple these channels. Finally, as noted in Ruffio et al. (2021), additional correlation can be introduced through the interpolation of the 4D (λ, t, x, y) spectral cube during reconstruction from the detector

images. This interpolation to a fixed wavelength grid guarantees the correlation of the noise in the IFS cubes, as noise in neighbouring detector pixels will be interpolated to build the IFS spaxels.

Greco & Brandt (2016) demonstrate the necessity of accounting for these correlated errors when retrieving physical properties from IFS data. If these correlations are not accounted for, they find that the retrieved confidence intervals are both artificially small and unreliable, often excluding the true parameter values at $> 95\%$ confidence. This was reinforced by Ih & Kempton (2021), where they explored the impact of correlated noise on atmospheric retrievals for transiting planets, finding that the assumption of non-correlated noise leads to biased posteriors and overfitting of the data.

Measuring noise correlation

Greco & Brandt (2016) introduce a procedure for empirically measuring the correlation in IFS datasets, and demonstrated the importance of including the full covariance matrix when fitting IFS spectra. In this work we extend their method by measuring the spectrum of injected planets and the resulting covariance, as opposed to the parameterised noise instance used in their work. This allows us to explore how the noise properties vary across instruments and over different post-processing methods. For each PSF-subtracted dataset we compute the average correlation within a 6 pixel wide annulus centred at the separation of the companion of interest. As in their work, we find that the correlation matrix does not depend strongly on the width of this annulus. The companion itself is masked out, leaving only residual noise. Such an annulus is chosen in order to maintain consistent noise properties in the sample of pixels: in general the noise varies more strongly with radius than with position angle. Work such as Gebhard et al. (2022) explores choosing more a more representative sample of pixels to describe the noise at the location of the planet, but such methods are computationally expensive, and we see little azimuthal asymmetry in the residuals shown in Figure 11.1.

Within the annulus, we compute the elements of the correlation matrix, ψ_{ij} as

$$\psi_{ij} = \frac{\langle I_i I_j \rangle}{\sqrt{\langle I_i^2 \rangle \langle I_j^2 \rangle}} = \frac{C_{ij}}{\sqrt{C_{ii} C_{jj}}}, \quad (12.2)$$

where $\langle I_i \rangle$ is the mean pixel intensity in the i^{th} spectral channel, and C_{ij} is the covariance between the two channels.

Estimating uncertainties

In order to compute the covariance matrix from the correlation matrix, we must know the diagonal, or uncorrelated elements of the covariance matrix. Several methods of measuring the photometric uncertainty were considered. We estimate the uncorrelated error in each wavelength channel by combining the photometric uncertainty of the stellar PSF, σ_{star} , with the residual noise at the location of the planet, σ_{residual} . We include the stellar uncertainty because near the edges of the bands in which spectra are observed, the filter transmission drops and atmospheric absorption increases, resulting in an increase in the uncertainty on the host star photometry. To measure the uncertainty on the stellar photometry we measure the standard deviation of the background in an annulus far from the stellar PSF in each wavelength channel, and use this to calculate the signal to noise. This represents an optimistic estimate of the stellar uncertainty, as we are unable to monitor photometric variability due to atmospheric conditions over the course of the observation, which represents the dominant source of uncertainty for the stellar photometry. To measure the uncertainty on the planet photometry we take the standard deviation of the residuals in an annulus at the separation of the companion, masking out the planet itself.

The histograms of Figure 11.1 show that the assumption of Gaussian errors across the entire frame is inconsistent with the noise, and would underestimate the tails of the distribution. [Pairet et al. \(2019\)](#) demonstrate that a Laplacian provides a better fit to the tails of the residual distribution than a Gaussian, while [Mawet et al. \(2014\)](#) shows that the residuals tend to follow a Student-t distribution. We find that a Student-t distribution best matches the full frame residuals. However, as the likelihood function for a general Student-t distribution is not analytic, and a Gaussian distribution accurately captures the residuals to 2.15σ , we continued to follow the standard practice of defining uncertainties as the Gaussian standard deviation. Taking KLIP as an example, the best fit Student-t distribution has 1.75 DoF, a mean of 2.23×10^{-8} and a width $t = 5.29 \times 10^7$. At the point where the Gaussian distribution intersects the Student-t distribution, 89% of the residuals are enclosed, compared to 97% if the residuals were Gaussian distributed. We also note that, relative to the speckle subtraction algorithms, ANDROMEDA shows an excess of 10σ outliers in the residuals, leading to difficulties in distinguishing between true positives and false positives, consistent with the findings of [Cantalloube et al. \(2020a\)](#). The long tails of these distribution add additional noise to each frame and need to be accounted for the detection of planet candidates in order to avoid false positives. However, for a known companion where we are concerned with

inferring physical parameters to within $1-2\sigma$ confidence intervals, accounting for the 90% of the noise that is contained within the Gaussian fit to the residuals is sufficient for defining the uncertainties.

Thus the total uncorrelated uncertainty for the i^{th} wavelength channel is given as:

$$\mathbf{C}_{ii} = \sigma_{i,\text{star}}^2 \left(\frac{f_{i,\text{pl}}}{f_{i,\text{star}}} \right)^2 + \sigma_{i,\text{residual}}^2. \quad (12.3)$$

The method described here provides an empirical estimate of the covariance of the noise after high-contrast image processing. However, as it relies on measurements of mean pixel intensities in a residual image, it is only applicable for speckle-subtraction methods. As ANDROMEDA produces an estimate of planet contrast at each pixel location, rather than residual noise following PSF subtraction, this method cannot be directly applied to the processed ANDROMEDA frames. An example of such a frame is shown in Figure 11.1, where highly structured noise is visible in the frame. The noise pattern is highly correlated through wavelength space, and indeed would lead to very strong residual correlation. Rather than applying the procedure for measuring the covariance matrix for ANDROMEDA, we instead rely on the estimate of the standard deviation that is also provided by the algorithm, that is also measured during the likelihood minimisation.

Figure 12.1 shows the results of computing the correlation matrix for both the KLIP and PynPoint reductions. There is a strong, narrow correlation component along the diagonal with a width of around 2-3 pixels, with a weaker correlation extending out to 10 pixels in width. In the SPHERE data, the correlation decreases in the water absorption features at $1.15 \mu\text{m}$ and $1.4 \mu\text{m}$. The KLIP data typically displays stronger correlations than the PynPoint reductions. This difference may be because the pyKLIP implementation of the KLIP algorithm uses only the most correlated frames from the PSF library to build the PSF model, which introduces an additional source of correlation in the data.

12.3. Bias correction

PCA methods tend to see increased self-subtraction as the number of principal components increases, naturally leading to poorer extractions as the number of components increases [Lagrange et al. \(2010\)](#) Therefore we also considered an empirical estimate of the uncertainty by injecting and recovering a sample of planets in an annulus at the location of the planet. The standard

deviation of the recovered spectra provide a measure for the uncertainty due to planet position and variation in the effectiveness of the post-processing. This can also be used as a method to correct for bias introduced by self-subtraction caused by the post-processing algorithm: by comparing the recovered spectra to the known input, a scaling factor can be computed. This can then be used to mitigate the self-subtraction induced by the PCA processing [Marois et al. 2014](#); [Gerard & Marois 2016](#); [Ruffio et al. 2017](#). We found that when applying such a bias correction, the χ^2 between the injected and recovered spectra was often worse than that of the nominal spectral extraction. As the noise properties are not truly azimuthally symmetric, the average bias correction does not provide a good correction for any individual planet location, particularly at the relatively faint contrasts considered in this work. Any improvement in the spectral extraction is dominated by the both the changes to the shape of the spectrum introduced by the data processing and the random noise of the measurement. Therefore we choose not to include bias correction as a step in our post processing, and do not include uncertainty from injection and recovery tests in our error estimate. The finding that bias correction can reduce the accuracy of the spectral extraction is surprising, and warrants further investigation into where this widely used technique should be applied. We leave such a study to future work.

12.4. Impact of covariance on retrieved parameters

In the previous section we discuss how to measure a covariance matrix for IFS data. Here we explore how the covariance impacts Bayesian inference, that is, when estimating the parameters of the model used to explain the observations. [Greco & Brandt \(2016\)](#) demonstrated how failing to include the full covariance matrix when fitting atmospheric models to IFS data results in overly confident and biased parameter estimates. In 14.1 we expand their work to higher dimensional models using atmospheric retrievals, but first we want to pedagogically understand how the covariance is linked to the posterior probability distributions. We subsequently show how the precision of a posterior parameter estimate depends on the ratio between the length scale of the correlation in the data and the length scale over which the parameter introduces changes in the model spectra. If this ratio is larger than unity, the posterior width decreases relative to the case without correlation. If the ratio is about unity the posterior width will increase.

Consider a toy model, where the data \mathbf{y} is given by a simple sine function

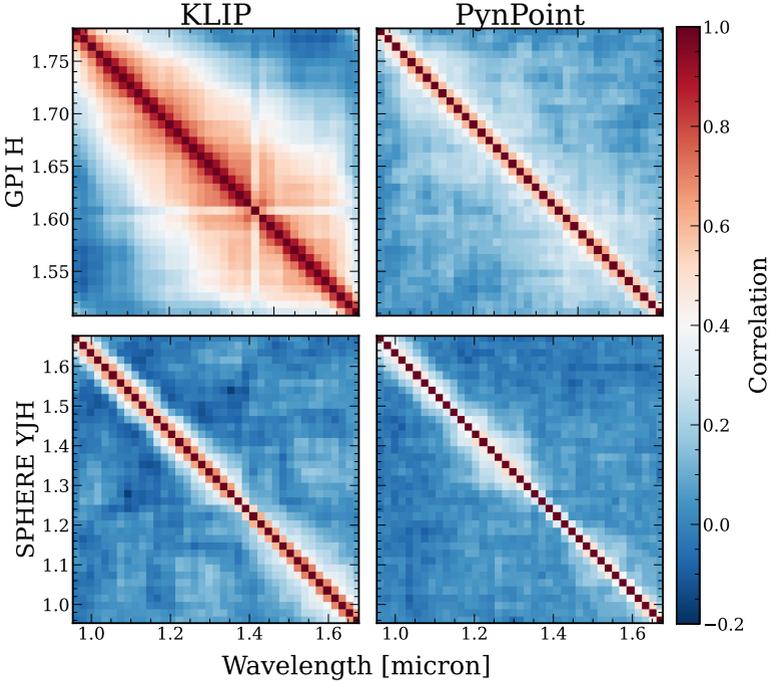


Figure 12.1: Correlation matrices for each dataset for HR 8799 e, with the GPI H-band data shown in the top row and the SPHERE YJH data on the bottom. Following the processing using KLIP (left) or PynPoint (right), we calculate the correlation and covariance matrices as described in Section 12.2. The correlation is computed as in Equation 12.2. The GPI data is more strongly correlated than the SPHERE data, particularly following the KLIP processing. The SPHERE data shows structure similar to the correlation matrix, with the correlation width following the shape of the water absorption spectrum.

with period T and offset D :

$$y_i = \sin(2\pi x_i/T) + D. \quad (12.4)$$

In the context of atmospheric parameters, we can think of this model as the first term of a Fourier series, which can be used to describe an atmospheric spectrum to arbitrary precision if extended to a high enough order. While the variation in the spectrum due to physical parameters is more complicated than this model, we can view an offset in the toy model as a change in the overall flux, while a

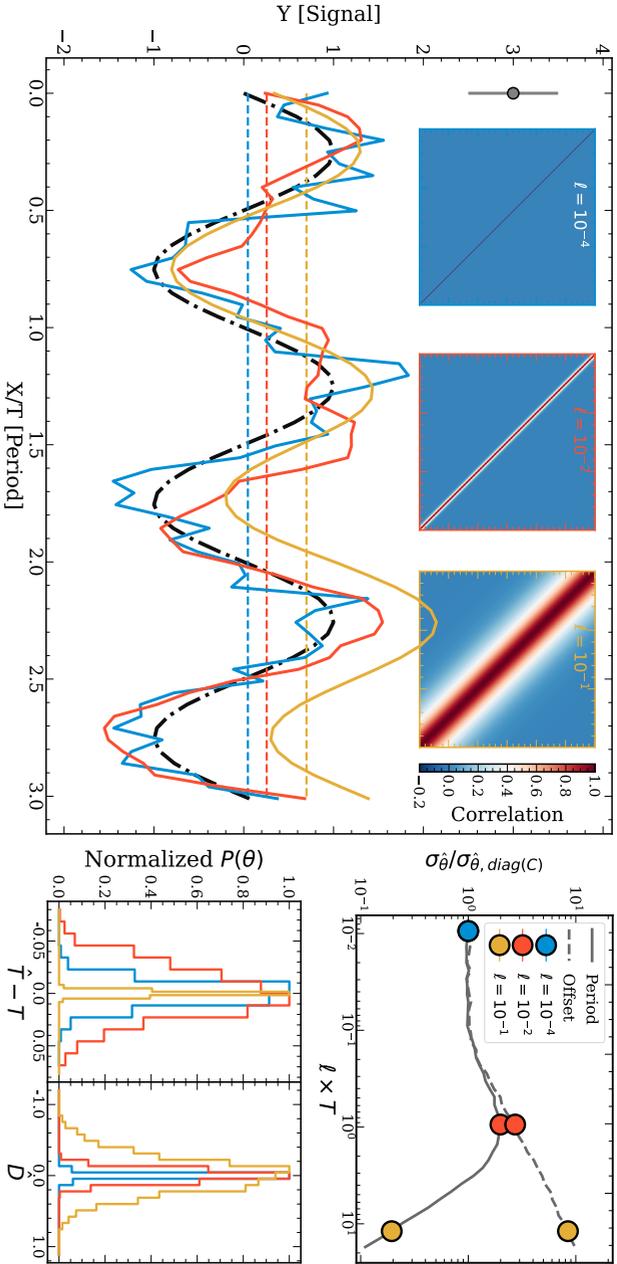


Figure 12.2: Toy model for demonstrating the effect of correlated uncertainties. **Left:** Data drawn from a toy sinusoidal model (Eqn. 12.4) when considering three cases of covariance in the data. The first (blue) is data drawn from a univariate Gaussian distribution with no correlations, as shown in the left inset. The second (orange) is drawn from a multivariate Gaussian distribution where the correlation length scale - as defined by the ℓ parameter of the Matérn kernel (Eqn. 12.5) - is less than the period of the model (centre inset). The third (yellow) is drawn from a multivariate Gaussian distribution with a correlation length scale greater than the period of the model (right inset). In the background the true input is plotted in gray. The gray datapoint indicates the 1σ error bar associated with each data point. **Right:** In the top panel, the posterior width of the period (solid gray line) and offset (dashed gray line) parameters scaled to the uncorrelated case as a function of the ratio between the correlation length scale and the period of the sine (so the model length scale). Marked in blue, orange and yellow are the draws plotted in the right panel. The histograms in the bottom panel are the posterior histograms for the period (left) and offset (right) for each of the highlighted cases.

change in the period would be reflected in the spectral shape, such as the near infrared water features.

The period and offset were arbitrarily chosen to be 30 and 0 respectively. Assuming that this toy model describes an observed experimental setup, we construct a synthetic dataset containing a total of 300 points, with x -coordinate values from 1 to 300, that is, ten periods. We applied different noise models to these toy data, and used nested sampling as implemented in `MultiNest` (Feroz & Hobson, 2008; Feroz et al., 2009; Feroz et al., 2019) to retrieve the value of the period parameters T and D . Such nested sampling algorithms improve on MCMC techniques to sample large parameter spaces efficiently, gradually restricting the sampling volume to regions of high likelihood. They are also robust to multimodal posterior probability distributions, and provide an estimate for the Bayesian evidence, \mathcal{Z} , as well as the posterior distributions and maximum likelihood fit.

We use the Matérn 3/2 kernel (Rasmussen & Williams, 2005) to describe the covariance of our dataset:

$$C_{3/2}(d)_{ij} = \sigma_i \sigma_j \left(1 + \frac{\sqrt{3}d_{ij}}{\ell} \right) \exp\left(-\frac{\sqrt{3}d_{ij}}{\ell} \right). \quad (12.5)$$

This is a model of for the correlation between two points separated by distance d_{ij} , where we can adjust characteristic correlation length scale through the correlation length parameter, ℓ . As ℓ decreases, the correlation matrix becomes more diagonal, while as ℓ increases the data becomes more strongly correlated across broad scales. The strength of the correlation is determined by the uncertainty on each point (σ_i), which we set to a constant value of 0.5 for each data point.

In Figure 12.2, we show three instances of the correlation matrix for $\ell = 10^{-4}$, 10^{-2} and 10^{-1} , ranging from uncorrelated to strongly correlated noise. Plotted in gray is the noise-less data, and the coloured lines show noise instances drawn from each of the three correlation matrices. For the diagonal case ($\ell = 10^{-4}$), we see that the data are randomly scattered around the true model. With no correlation, we see true, univariate Gaussian noise. As the correlation scale increases to $\ell = 10^{-2}$, we see that the data appear smoother, and the variations occur on larger spatial scales than in the case without any correlation. Finally, with $\ell = 10^{-1}$, we see that the data are offset from the ground truth, but do not have any small-scale scatter. This is the impact of covariance on the data: as the correlation length increases, every point is more strongly determined by the initial random draw any other point (effectively there are less points, as there are less independent measurements). In summary, we observe high frequency variation due to noise if the correlation length is small. In this case the mean of the data, parameterised by parameter D should be accurately and precisely

inferred. As the correlation length increases, the scatter of the mean D across multiple draws increases, but we see less small-scale variation, allowing a better estimate of the period, T .

We note here that the Matérn 3/2 kernel is only one model for the covariance, and is not perfectly suited for IFU data. A more robust model (such as described in [Greco & Brandt 2016](#)) would incorporate both a broad correlation term and a diagonal Gaussian term to the correlation matrix, which would introduce small-scale scatter in the data, even with a large correlation scale. Nevertheless, this is a suitable toy model to explore how changing the correlation length scale impacts parameter estimation.

To determine the impact on parameter inference, we vary the correlation matrix across a range of ℓ values from 10^{-4} to 10^0 , and use `pyMultiNest` ([Buchner et al., 2014](#)) with 400 live points to fit the true model, accounting for the covariance in the likelihood. We did not perturb the toy dataset by with an error model as defined by the covariance matrix, so we run noise-free retrievals. This is equivalent to running multiple inferences where the data are perturbed by draws from the covariance matrix and averaging over the posteriors of each inference. We set uniform priors on the period $P(\omega) = \mathcal{U}(0, 100)$ and offset $P(D) = \mathcal{U}(-10, 10)$. The results are not sensitive to the choice of number of live points ($n_{\text{live}} \gg n_{\text{param}}$) or priors. The upper right panel of Figure 12.2 shows the ratio between the width of the posterior distribution for this parameter and the width of the distribution in the case of univariate Gaussian noise (i.e. no correlation). We observe that as the correlation length scale approaches the length scale of the sine function (the period) the width of the posterior increases: the correlation introduces variations in the data on the scale of the period, making it difficult to estimate the parameter. This is the effect described in [Greco & Brandt \(2016\)](#), where accounting for the covariance matrix when fitting atmospheric models increases the posterior width. However, as the correlation scale continues to increase to scales larger than the period, we see that the posterior width decreases to values lower than in the case of uncorrelated noise. As is visible in the data in the left panel, without small scale variations to introduce uncertainty in the period, it becomes easier to estimate this parameter, at the cost of increased uncertainty in the estimate of the offset parameter D .

Effects of ignoring covariance

It is often the case that the full covariance matrix is not used when performing atmospheric retrievals, and we wanted to explore the impact of using only the diagonal terms when fitting a model to correlated data. Figure 12.3 shows

the best-fit reduced χ^2 as a function of the ratio between the period and the correlation length scale, as in the right panel of Figure 12.2. For each ℓ , we perform an ensemble of 25 retrievals using Multinest in order to reduce the scatter and to measure the uncertainty in the χ^2 due to the variation between individual noise instances. In this case, the data are perturbed by draws from the covariance matrix, in order to test the impact of using the incorrect covariance in the likelihood when the data are correlated. We define ν as the number of data points (300) minus the number of parameters (2). This procedure is repeated using both the full covariance matrix, \mathbf{C} , in the likelihood, as well as using only the diagonal elements of the matrix - that is, we assume that the data are uncorrelated. We find that the reduced χ^2 is a useful metric if the covariance is properly accounted for. If the data are correlated and only the uncorrelated uncertainties are used in the likelihood then the reduced χ^2 will be underestimated, and the scatter of the χ^2 increased. Often a $\chi^2/\nu < 1$ is interpreted either as overfitting of the data or overestimation of the uncertainties. However, we demonstrate here that for $n_{\text{param}} < n_{\text{data}}$ a $\chi^2/\nu < 1$ can be interpreted as an *underestimation* of the correlation of the data.

Applicability to atmospheric retrievals

We expect similar effects to be present in atmospheric retrievals with correlated data. Parameters that affect model spectra on wavelength scales smaller than the correlation scale may be retrieved to higher precision than expected if the uncertainties were uncorrelated, while parameters that are sensitive at approximately the correlation length scale will have larger posterior uncertainties. As seen in Figure 12.1, the correlation length can be a appreciable fraction of the total data, particularly in the case of the KLIP reduction of GPI data. Large scale correlations in the data can introduce offsets in the average flux measurement, which can lead to inconsistencies between datasets from different instruments or measured during different epochs. Correlations on moderate scales can alter the spectral shape, in turn impacting parameter estimates. For example, the surface gravity is particularly sensitive to the shape of the H-band, and changes to the shape of this band will lead to biased estimates. Thus for IFS data it is critical to account for the covariance matrix when fitting models to the data, in order to correctly capture the noise structure imprinted onto the signal.

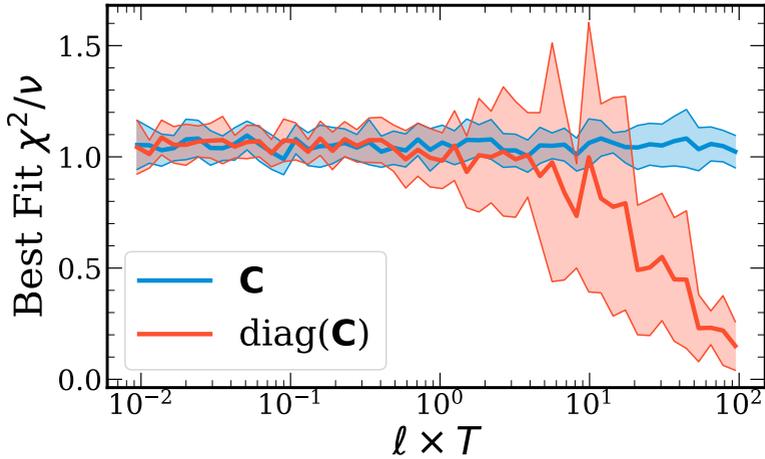


Figure 12.3: The best fit χ^2/ν as a function of the ratio between the correlation length scale (proportional to $1/\ell$) and the period, T . The χ^2 was computed for fits of equation 12.4 to data perturbed by draws from the covariance matrix, varying the correlation length scale. For each ℓ , 25 Multinest retrievals were run in order to compute the uncertainty on the χ^2 , shown as the shaded region around the mean. In blue, the covariance is properly accounted for in the likelihood, while in orange only the diagonal of the covariance is used in the likelihood. In order for the reduced χ^2 to be a useful metric, the covariance must be properly accounted for.

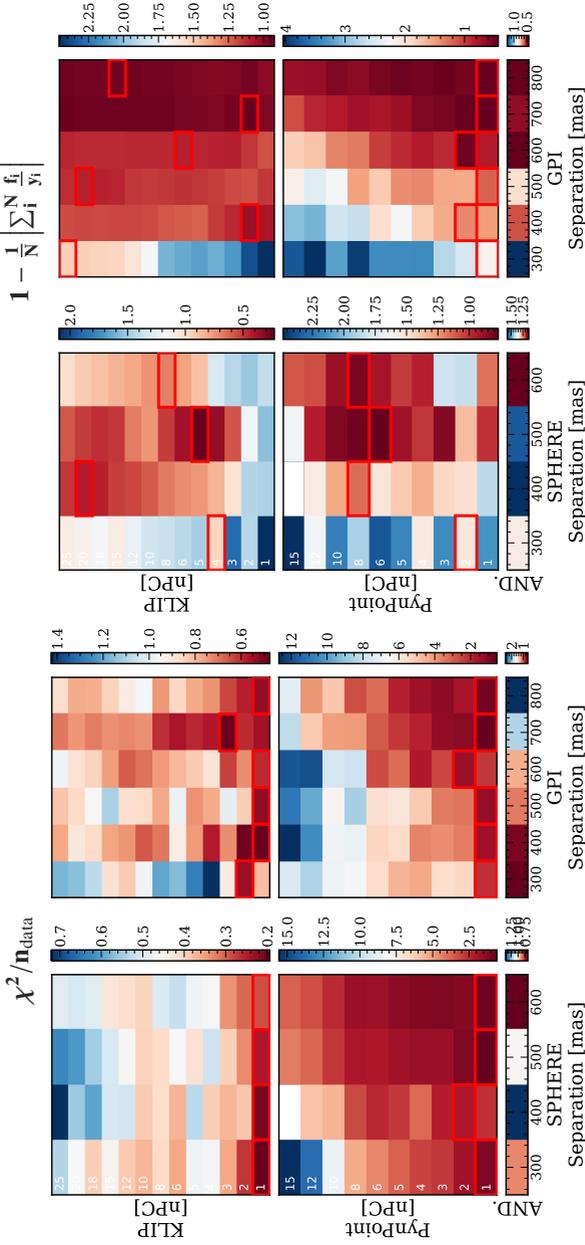


Figure 12.4: Goodness-of-fit metrics mapped across separation and number of principal components. The colour scale indicates median metric value from the injections at 4 separate position angles. The range of the colour scale for each sub plot is different in order to capture the variation within a single map. Highlighted in red are the optimal extractions for each separation. **Left:** χ^2/n_{data} map. **Right:** The same as the left panel, but calculated using the relative discrepancy (e) instead of the χ^2 .

13. INJECTION TESTING

In order to best extract a true signal, we want to optimise the data-processing parameters. However, without knowledge of the ground truth spectrum, it is unclear how these parameters should be tuned a priori. By injecting fake companions with a known spectrum, applying the post-processing, and comparing the extracted spectrum to the input we can then optimise the parameters, and use this setup to extract the true planet signal. In particular, we try to optimise the choice of the number of principal components used in PSF subtraction for KLIP and PynPoint as a function of the separation. This injection-extraction study also provides us with a metric for comparing the three algorithms described in Section 12.1.

Using the pyKLIP injection tool we injected companions into both the SPHERE and GPI HR 8799 datasets. The normalised stellar PSF was used both as a model for the planet PSF and to scale the bulk contrast of the injected companion. The spectrum was convolved with a gaussian kernel to the instrumental resolving power, and binned to the instrumental wavelength grid using the `rebin_give_width` function available in `petiRADTRANS`, which accounts for non-uniform bin sizes as the number of pixels per instrumental resolution element varies with wavelength. Only a single planet was injected at a time before the data processing, which was repeated for each planet position in order to avoid potential contamination from nearby signals. We injected the companions at varying positions into both the SPHERE and GPI datasets, with a spectrum generated using `petiRADTRANS` as described in Section 14. These were positions representative of the known separations of the inner three companions. Planets were injected at position angles from 120° to 240° from the location of HR 8799 e in 30° increments, and between 300 – 800 mas in 100 mas steps. This process was repeated for the SPHERE YJH and the GPI H-band datasets at mean contrasts from 10^{-7} to 10^{-4} .

Once the data were prepared, we ran each of the three data processing algorithms on each injected dataset, spanning a range of algorithmic parameters. While we could not exhaustively study the effect of each parameter, we chose to focus on the impact of the number of principal components used during PSF subtraction in order to optimise the spectral extraction. Other parameters, such as the `flux_overlap` parameter in KLIP, or the filtering fraction in ANDROMEDA were set based on suggested values from previous studies (Zurlo et al., 2016; Cantalloube et al., 2015) or from qualitative examination of the post-processed data. Several parameters, such as `thetolerance` and `merit` parameters of PynPoint were chosen to ensure accurate extractions within reasonable computation time. Various geometric parameters, such as the inner

and outer working angles together with the width parameter in ANDROMEDA, or the subsection and annuli parameters of KLIP were set based on recommendations from the documentation¹, and ensuring that the region under consideration would contain the entirety of the planet signal, extending to at least twice the FWHM of the signal. A full table of parameter choices for each algorithm is given in Table 12.1.

13.1. Choice of goodness-of-fit metric

We considered several goodness-of-fit metrics with which to determine the optimal extraction, including the signal-to-noise ratio (S/N), the relative discrepancy (e) and reduced χ^2 (χ^2/n_{data}). We take the median of each metric across the five different position angles where planets were injected. We exclude spectra that are over 20σ discrepant from the input, or that display strong outliers with contrast $> 2 \times 10^{-5}$, though the results are robust to including the outlying data. Each of these metrics identified different optimal spectra.

The mean S/N always identified the spectra processed using the largest number of components as optimal. However, the resulting spectra do not correctly retrieve the shape of the input spectrum, because they typically overestimate the flux, so we did not consider this metric further.

We define the mean relative discrepancy e between a measured flux \vec{s} and known input spectrum $\vec{\bar{s}}$ as

$$e = 1 - \frac{1}{N} \left| \sum_i^N \frac{s_i}{\bar{s}_i} \right|. \quad (13.1)$$

To identify the best fit spectrum we simply find the minimum value of this function. In contrast to the χ^2 or other distance metrics, the discrepancy is invariant of the magnitude of the measured quantity, and so provides a metric to compare spectra injected at different contrasts.

The χ^2 is a standard metric for measuring the similarity of distributions, but can also favour measurements with overestimated uncertainties. The χ^2 value between the extracted spectrum \mathbf{s} with covariance \mathbf{C} and the known injected spectrum $\bar{\mathbf{s}}$ was calculated for each post-processed dataset as

$$\chi^2 = (\mathbf{s} - \bar{\mathbf{s}})^T \mathbf{C}^{-1} (\mathbf{s} - \bar{\mathbf{s}}). \quad (13.2)$$

We present χ^2/n_{data} , dividing by the number of wavelength channels n_{data} to allow for a more straightforward comparison between instruments. We do not

¹https://pyklip.readthedocs.io/en/latest/fm_spect.html

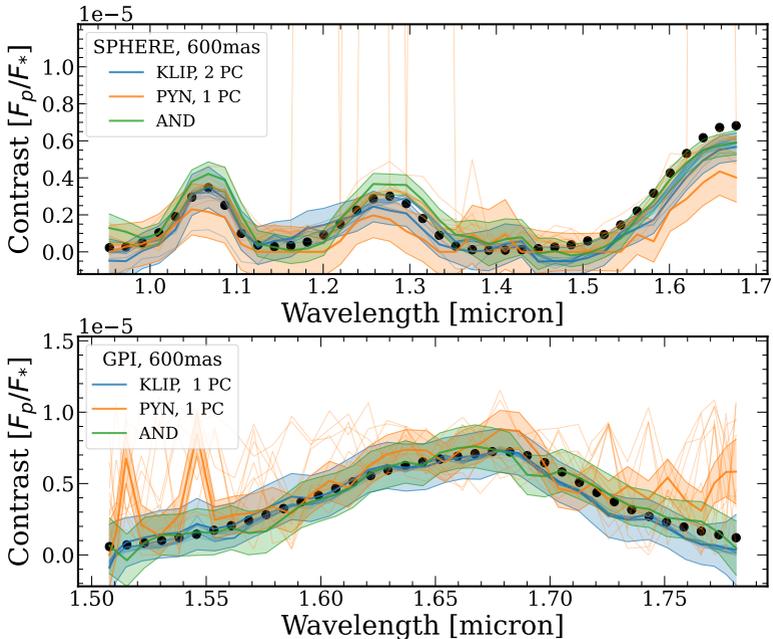


Figure 13.1: Typical spectral extractions for injected planets located at separations of 600 mas. These spectra are representative of the HR 8799 planets with ($F_p/F_* \sim 2 \times 10^{-6}$). The injections into the SPHERE data are shown on the top panel, the GPI on the bottom. Each injected planet was positioned 150° from HR 8799 e. Extractions for each algorithm are plotted, with the best fit spectrum (χ^2) and 1σ error bars from the diagonal of the covariance matrix highlighted by the shaded region. The faint lines show the variation in the extractions using different numbers of principal components.

subtract the degrees of freedom from the number of data points as is typical when computing the reduced χ^2 , as principal components are not free parameters in a statistical sense, thus making the definition of degrees of freedom challenging.

13.2. Optimising spectral extractions

We present in Figure 12.4 the map of χ^2/n_{data} (left) and mean relative discrepancy (right) as a function of both separation and number of principal

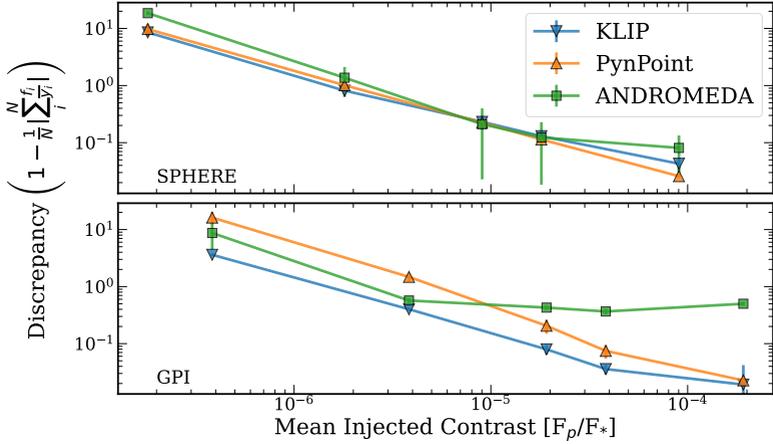


Figure 13.2: Best-fit discrepancy (eqn. 13.1) as a function of input contrast at 400 mas. The top panel shows the results for injections into the SPHERE data cube, while the bottom is for GPI. The injections were repeated at three position angles, and the uncertainty presented is the standard deviation of these measurements.

components used, taking the median across the injections at different parallactic angles. Extracted spectra for typical injected fake companions are shown in Figure 13.1. The precision and accuracy of our spectral measurements depends strongly on the separation, visible in the variation of the both metrics. There is strong position-dependent variation in the shape of the extracted spectra, including cases where the injected companion is not detected in any wavelength channel, as well as cases where the peak contrast is overestimated by a factor of 2. Examining Figure 11.1, we see that beyond about 400 mas from the host star the noise properties are relatively unstructured, while at 400 mas and closer the GPI data is dominated by the residual speckle noise. Such a trend is also present in the SPHERE data, though the speckle dominated regime extends out to only 300 mas. This transition in the underlying noise properties together with the greater angular displacement at wider separations results in the improved detections at wider separations. This can be disguised by the χ^2 metric, where large uncertainty estimates at small separations can result in a better χ^2/n_{data} , while the mean relative discrepancy provides a clearer trend as a function of separation. When using χ^2/n_{data} as the goodness-of-fit metric, we find that both KLIP and PynPoint favour low numbers of principal components.

Such results depend strongly on both the dataset, the reduction used, and the choice of metric. The SPHERE extractions are universally better than the GPI, largely due to the brightness of the injected spectra in the Y and J bands, the smaller inner working angle and pixel scale, and the longer integration time. PynPoint most strongly favours low numbers of principal components, across both metrics and at all separations. However, when we consider the relative discrepancy between the input and extracted spectra, and find that the number of PCs components favoured is much higher than when using the χ^2 , particularly for KLIP. These spectra may more closely match the shape of the input spectra, but may also underestimate the uncertainties, leading to them being disfavoured by the χ^2 . While the shape of the spectrum does depend on the number of components used, it is more strongly dependent on the particular location in the frame where it is injected.

Each algorithm displayed its own trends in the quality of its spectral extractions. KLIP produced smooth spectra, with systematics that were relatively consistent in shape over the full range of principal components used. However, it struggled to recover the brightest sources accurately, reproducing the over-subtraction effect described in Pueyo (2016). In contrast, PynPoint and ANDROMEDA performed worse at fainter contrasts, as demonstrated in Figure 13.2. The PynPoint spectra are more dominated by random scatter than by systematic variation, reflected in the typically diagonal correlation matrices. ANDROMEDA produced some of the best overall fits, but struggled to achieve the correct flux calibration, both over- and under-estimating the flux in different cases. For the bright injection case, PynPoint consistently performed the best, producing the lowest χ^2 values for each dataset and separation. KLIP struggled to extract the brightest spectra, over-subtracting the planet signal at the red end of the SPHERE data. This effect was more severe when larger numbers of principal components were used. However, KLIP also displayed a tendency to over-estimate the flux of the signals injected into the GPI data. ANDROMEDA was able to accurately extract the high S/N SPHERE injection, but systematically underestimated the flux in the GPI data.

The best fit of each algorithm performs relatively well at extracting the true spectrum, with typical best-fit reduced χ^2/n_{data} values approaching 1. Depending on the injected position angle and separation, the χ^2 for the same algorithm at the same separation can vary by a factor of ~ 10 , with typical standard deviations on the order of 10–100 depending on the dataset and algorithm. This variation in extraction suggests that injection recovery tests to measure and correct for algorithm throughput, such as detailed in Greenbaum et al. (2018), may introduce additional biases depending on the precise positioning of the injected companions. In our reproduction of this method, we find that it does

not provide better χ^2 values, and can introduce spurious wavelength-dependent signals. This variation is separation-dependent and impacts the extraction less strongly at wider separations, outside the speckle noise regime.

These results point to differences in the approach to data analysis required between detection and characterisation efforts. High numbers of principal components tend to whiten the noise and improve the detection significance, potentially allowing the discovery of fainter companions. However, this comes at the cost of reduced photometric accuracy, which is critical when attempting to recover the physical atmospheric parameters. For both speckle subtraction methods, low numbers of principal components produce the most accurate spectral extractions, though the precise number of components will depend on the brightness of the companion in question.

14. RETRIEVAL TESTS

Atmospheric retrievals provide a useful, data-driven tool for exploring the properties of exoplanet atmospheres. Retrieval results are dependant on the quality of the input data and the assumptions made about both the data and the model. For our investigation below we have two primary aims:

1. Exploring the impact of high-contrast image processing on the inferred atmospheric parameters through retrievals on synthetic data.
2. Characterising how correlated noise influences the fits of the synthetic data. With a known ground truth, we can explore how the use of the covariance matrix can help mitigate the impact of systematic effects introduced by the data processing.

To this end, we use a representative selection of optimised spectral extractions as described in Section 13. We choose to use the model injected at 600 mas, and positioned 150° rotated from HR 8799 e, combining both the SPHERE YJH and GPI H-band datasets. The best extracted spectrum as measured by the relative discrepancy were chosen as the baseline inputs to the retrievals. This represents a realistic, though challenging spectrum on which to perform atmospheric retrievals. For validation we also explored a set of retrievals on different locations and choice of extraction, finding that while the precision often varies with the S/N of the extracted spectrum, the overall trends of our results are reproducible. In contrast to Greco & Brandt (2016), these retrievals explore the full impact of IFS data processing on the spectra, as opposed to using data synthesised from a parametric estimate of the covariance. In contrast to their use of a 3-parameter BT-Sett1 model, we use an ~ 8 parameter forward model in order to understand the cumulative impacts of data post-processing on the inferred atmospheric parameters in the context of high-dimension on-the-fly retrievals. Such retrievals are highly flexible, and are more likely to try to fit spurious data features than more physically motivated fits from self-consistent grids.

Atmospheric model

The models we use in our atmospheric retrieval setup are computed using `petitRADTRANS` (Mollière et al., 2019), a fast, open-source radiative transfer code with which we can calculate the emission spectrum of an atmosphere¹. In this framework, the atmosphere of a planet is divided up into pressure bins.

¹<https://petitradtrans.readthedocs.io/>

Temperature and chemical structures are calculated and applied to each bin, and radiative transfer using the correlated-k method for the opacities (Goody et al., 1989; Lacis & Oinas, 1991) is performed to calculate the emission spectrum. The correlated-k opacities are binned from their native spectral resolving power of 1000 to a user-supplied model resolution using the `exo-k` package (Leconte, 2021), improving the computation time of the retrieval. A wavelength binning of at least twice the data resolution is used for the models, in order that the binned model spectrum is Nyquist sampled. This spectral model is convolved with a Gaussian kernel with the width of the instrumental spectra resolution and then binned to the wavelength grid of the input data for the retrieval using the `rebin_giv_width` function. At the spectral resolutions considered in this work, the effects of the convolution and binning on the spectrum can dominate the spectral shape over data processing effects. For this reason we ensured that we use the same convolution and binning procedure during the spectral injections as during the retrieval. However, future work should investigate incorporating better instrumental models and wavelength dependent kernels into retrieval frameworks.

Our baseline model uses a Guillot temperature profile (Guillot, 2010) and freely retrieved chemical abundances. This profile is a simple analytical model, constructed to estimate the thermal structure of irradiated planets:

$$T_{\text{Guillot}} = \frac{3T_{\text{int}}^4}{4} \left(\frac{2}{3} + \tau \right) + \frac{3T_{\text{irr}}^4}{4} \left(\frac{2}{3} + \frac{1}{\gamma \sqrt{3}} + \left(\frac{\gamma}{\sqrt{3}} - \frac{1}{\gamma \sqrt{3}} \right) e^{-\gamma \tau \sqrt{3}} \right), \quad (14.1)$$

where $T_{\text{irr}} = \sqrt{2}T_{\text{equ}}$ and $\tau = P \times \kappa_{\text{IR}}/g$. T_{equ} is the standard equilibrium temperature of an irradiated body, g is the surface gravity. P is the atmospheric pressure, divided up into a total of 80 log-spaced layers from 10^3 bar to 10^{-6} bar. The remainder of the parameters are as in Guillot (2010): T_{int} is the intrinsic internal temperature of the planet, κ_{IR} is the mean infrared opacity, and γ is the ratio between the optical and infrared opacities. All of these parameters are freely retrieved, rather than being derived from the opacities in each atmospheric layer.

This provides a simple but flexible model for the P-T profile, and is the model used to generate the injected spectrum. By setting the irradiation temperature to low values the Guillot model can reproduce the general shape of typical directly imaged planet temperature profiles. Setting the equilibrium temperature T_{equ} to zero provides the limiting case of the Eddington profile (Eddington, 1930). Together with the planet radius, log g , and the chemical abundances this model

| Parameter | Prior | Input |
|-------------------------------|-------------------------------------------------------|----------------------|
| Guillot, Free Chemistry | | |
| $\log g$ | $\mathcal{U}(2.0, 5.5)$ | 4.0 |
| R_{pl} | $\mathcal{N}(1.0 R_{\text{jup}}, 0.2 R_{\text{jup}})$ | $1.0 R_{\text{jup}}$ |
| T_{int} | $\mathcal{U}(300 \text{ K}, 2000 \text{ K})$ | 750 K |
| T_{equ} | $\mathcal{U}(0 \text{ K}, 300 \text{ K})$ | 100 K |
| γ | $\mathcal{N}(1, 0.2)$ | 0.5 |
| $\log \kappa_{\text{IR}}$ | $\mathcal{U}(-3.0, 1.0)$ | -1.0 |
| $\log X_{\text{H}_2\text{O}}$ | $\mathcal{U}(-7.0, 0.0)$ | -1.5 |
| $\log X_{\text{CO}}$ | $\mathcal{U}(-7.0, 0.0)$ | -2.0 |

Table 14.1: Priors for retrieval setup. $\mathcal{U}(a, b)$ denotes uniform priors with bounds a and b , and $\mathcal{N}(\mu, \sigma)$ denotes a normal distribution centred at a mean μ with standard deviation σ . The final column indicates the true values of the spectrum injected into the IFS cubes.

uses a total of 8 parameters. The only included sources of line opacities are H_2O from the ExoMol data base (Chubb et al., 2021; Tennyson & Yurchenko, 2012) and CO from HITEMP (Rothman et al., 2010a). Both Rayleigh scattering in an H_2 and He dominated atmosphere and collisionally induced absorption between H_2 - H_2 and H_2 -He are included as continuum opacity sources. The priors used for all parameters in the retrieval are presented in table 14.1.

Retrieval setup

pyMultiNest is used to generate samples and determine both the posterior parameter distributions and the Bayesian evidence of the retrieval (Buchner et al., 2014). This is a Python wrapper for the MultiNest sampler and likelihood integration method of Feroz & Hobson (2008). For all of the retrievals we use 4000 live points to thoroughly explore the parameter space, and a sampling efficiency of 0.8, as recommended in the pyMultiNest documentation for parameter estimation. We compute negative log likelihood, the value of which is minimised in order to find the best-fit set of parameters. Across many samples,

| Dataset | χ^2/n_{data} | $\log \mathcal{L}$ | BIC | d_m |
|---------------------------------------|--------------------------|--------------------|-------|-------|
| KLIP, \mathbf{C} | 0.69 | 2692 | -5289 | 2.35 |
| KLIP, $\text{diag}(\mathbf{C})$ | 0.41 | 2656 | -5218 | 3.29 |
| PynPoint, \mathbf{C} | 1.72 | 2646 | -5220 | 5.63 |
| PynPoint, $\text{diag}(\mathbf{C})$ | 1.43 | 2641 | -5208 | 9.60 |
| Andromeda, $\text{diag}(\mathbf{C})$ | 0.39 | 2650 | -5266 | 4.62 |
| Gaussian, $\text{diag}(\mathbf{C})$ | 1.06 | 2651 | -5267 | 2.21 |
| Noise Free, $\text{diag}(\mathbf{C})$ | 0.06 | 2686 | -5336 | 1.69 |

Table 14.2: Summary of retrievals run on synthetic data. We compare best fit reduced χ^2/n_{data} , the (negative) log likelihood which includes the covariance weighting term of equation 14.2, the Bayesian Information Criterion of equation 14.3 and the Mahalanobis d_M from equation 14.4. Retrievals were performed on data processed with each algorithm, both using the full and diagonal only terms of the covariance matrix. Toy models using univariate Gaussian scatter about the input and no scatter are also included, with the uncertainties defined as $\text{diag}(\mathbf{C})$ from the KLIP data.

this provides a measurement of the posterior probability distribution of model parameters given the data. Under the assumption of Gaussian distributed errors, the log likelihood function takes the form of a simple χ^2 likelihood distribution. Using the covariance matrix \mathbf{C} of the data from Section 12.2 with elements \mathbf{C}_{ij} , we compute the log likelihood function $\log \mathcal{L}$, which is the log-probability of measuring the observed spectrum \mathbf{S} given a forward model \mathbf{F} . A normalisation term is included which allows for a varying covariance matrix or uncertainty for each dataset and penalises samples with higher uncertainties. Thus our likelihood function is computed as:

$$-2 \log \mathcal{L} = (\mathbf{S} - \mathbf{F})^T \mathbf{C}^{-1} (\mathbf{S} - \mathbf{F}) + \log (2\pi \det(\mathbf{C})). \quad (14.2)$$

Goodness-of-fit metrics

In a retrieval on real exoplanet data without a ground truth value to compare to, we must turn to different metrics in order to determine which retrieval best describes the underlying spectrum. Table 14.2 lists χ^2 values of each best fit, as

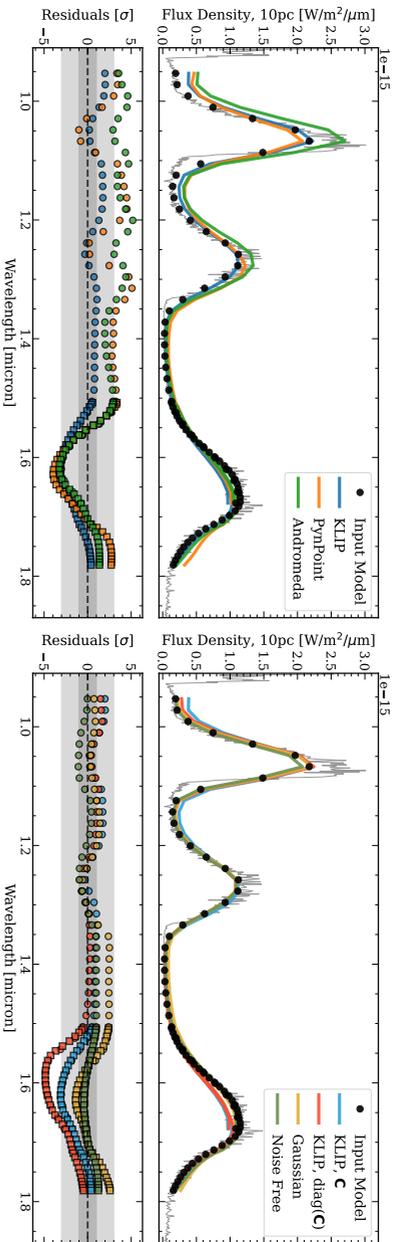


Figure 14.1: Best-fit models compared to the ground truth spectrum, in order to determine how well the retrieval can account for the systematics introduced through the data analysis. The residuals are calculated by sampling the posterior distributions to generate spectra, and taking the standard deviation at each wavelength. **Left:** The best fit from the retrievals for each KLIP (blue), PynPoint (orange) and ANDROMEDA (green). **Right:** The best fits of the KLIP retrievals with (blue) and without (yellow) including the covariance matrix, as well as the retrieval on the Gaussian noise (red) and Noise-free datasets (green). These are again compared to the true input spectrum.

well as the minimum negative log likelihood as computed in Equation 14.2 and the Bayesian Information Criterion (BIC, Wit et al. (2012)). Typically when performing model comparison in a Bayesian framework, we would turn to the Bayes factor in order to reject the null hypothesis. However, when comparing the impact of different data reductions on the retrieval outcomes, the Bayes factor as computed through the nested sampling evidence estimate is insufficient, as not all of the free parameters are included in the sampling process or in the prior volume, namely the those related to the post-processing algorithms. This would bias the evidence estimate, which depends on the choice of priors and thus the overall prior volume. A full treatment would require marginalising over these algorithmic parameters, and computing a forward model of the planet signal in the IFS data. At the present time, such a joint approach is computationally infeasible. Ruffio et al. (2019) and Wilcomb et al. (2020) demonstrate that this is possible given a linear model of the starlight, the planet signal and the residuals, which can be optimised and analytically marginalised over to determine posterior distributions. However, this approach loses information on the continuum shape of the spectrum, and relies on moderate-to-high spectral resolution to infer physical quantities. The atmospheric model is also not computed on the fly, and instead relies on a precomputed grid, limiting the parameter space available for exploration.

Therefore, instead of the Bayes factor, we rely on the BIC as a summary statistic:

$$\text{BIC} = k \log n - 2 \log \mathcal{L}_{\max} \quad (14.3)$$

for k free parameters and n data points. This formulation allows us to account for the free parameters of the atmospheric model, as well as the parameters of the data processing, where we add one parameter for each principal component used during PSF subtraction. Unlike the Bayes factor, the BIC is only a heuristic for model comparison, and differences in the BIC cannot be treated as a metric for statistical significance. Nevertheless, models with a lower BIC can be considered more strongly favoured. As the BIC depends on the likelihood, this also means we cannot directly compare retrievals which include or neglect the off-diagonal terms of the covariance matrix. Bayes factors and the BIC estimate whether a certain forward model is favoured when compared to another one, whereas turning covariance on or off corresponds to changing the functional form of the likelihood function. It is, therefore, not a question of forward model selection. Thus no single summary statistic can determine the overall goodness-of-fit of the retrieval.

The χ^2/n_{data} statistic is useful for understanding the impact of varying the covariance and quantifying the similarity of the model to the spectrum, while

the BIC is useful for heuristically evaluating the goodness-of-fit, accounting for possible over-fitting from the addition of extra parameters. In general however, we cannot directly compare the likelihood or the BIC when comparing the cases including the covariance to those using only the diagonal of the matrix with the usual motivation of model selection. Adding or neglecting the covariance does not correspond to a different forward model choice, instead it is equivalent to using a correct or incorrect functional form of the likelihood function. Therefore, assuming that the covariance is correctly measured, it is *always* better to include the full matrix in the likelihood in order to make statistically robust statements about the data. Thus even though the reduced χ^2 of the covariance case may be larger than that with only the diagonal, it still provides a more honest analysis of the data. It is also not surprising if the χ^2 increases if the covariance is added, since a Gaussian distribution defined by a covariance matrix with non-zero off-diagonal elements will always have a higher information content (e.g. [Rodgers, 2000](#)). As discussed in Section 12.4, including the covariance may either increase or decrease the width of the parameter posteriors.

For comparing retrieval results that include or neglect the covariance matrix we make use of the Mahalanobis distance d_M ([Mahalanobis, 1936](#)), to quantify the absolute distance between the posterior probability distributions $P(\theta | \vec{x})$ with means $\vec{\mu}$ and covariance S , and the true parameter values $\vec{\theta}$:

$$d_M(\vec{\theta}, P(\theta | \vec{x})) = \sqrt{(\vec{\theta} - \vec{\mu})^\top S^{-1} (\vec{\theta} - \vec{\mu})}. \quad (14.4)$$

This provides a metric for the overall accuracy of the retrieval when the true input parameters are known.

Finally, for this work we use the median parameter values and associated spectra as our point of comparison as opposed to the maximum likelihood fit. We find that although the spectrum generated by the median parameter values is a worse fit (by definition) than the best fit spectrum, the median parameters are a more accurate measurement of the input parameters.

14.1. Outline of retrievals

We performed three main tests to answer address the central theme of this paper:

1. Comparing retrievals on spectra extracted using different post processing algorithms.

2. Comparing retrievals that either include or ignore the covariance matrix in the likelihood function.
3. Testing if a lack of correlation information can be accounted for using additional ad-hoc data-processing parameters in the retrieval.

Our primary retrieval results are summarised in Table 14.2. In Section 14.2 we compare the cases of data that has been processed with KLIP, PynPoint and ANDROMEDA, both with and without the use of the covariance matrix from Section 12.2. As a benchmark, we also include a retrieval using the nominal input spectrum, perturbed with draws from a Gaussian distribution, where the covariance is given by the diagonal of the KLIP covariance matrix. This represents how the data would appear without systematics from HCI data processing and without the correlations introduced by the instrument optics. We also include a retrieval using the same uncertainties as in the Gaussian case, but without scatter about the input spectrum to validate our retrieval method and choice of goodness-of-fit metrics. For the sake of brevity, we refer to these as the ‘Gaussian’ and ‘noise-free’ cases respectively. We explore the impact of incorporating the covariance matrix in the retrieval framework in Section 14.3, using the KLIP, Gaussian and noise-free cases. Section 14.5 explores whether we can account for ignorance of the covariance in the data by introducing scaling factors and offsets in the retrieval.

Validation

To verify the validity of our results, we also ran a series of validation retrievals to test the sensitivity of our results to the choice of datasets, priors, and models. We ran retrievals on each dataset independently, as well as with broad and tight priors. Neither dataset was able to retrieve the parameters as precisely as the combined retrievals. The posterior distributions and fits were insensitive to our choice of priors. We ran additional retrievals using a spline temperature profile, as a proxy for our model not truly matching the underlying data. With 5 spline nodes, we were able to retrieve the $\log g$, R_{pl} , T_{int} and the water mass fraction to the same precision and accuracy as using the Guillot profile used to generate the data, thus concluding that the retrievals are flexible enough to account for some degree of imperfect model assumptions.

14.2. Impact of algorithm selection on retrievals

The largest variation in the extracted spectrum is due to the choice of post-processing algorithm, so our first aim is to explore how these differences in the

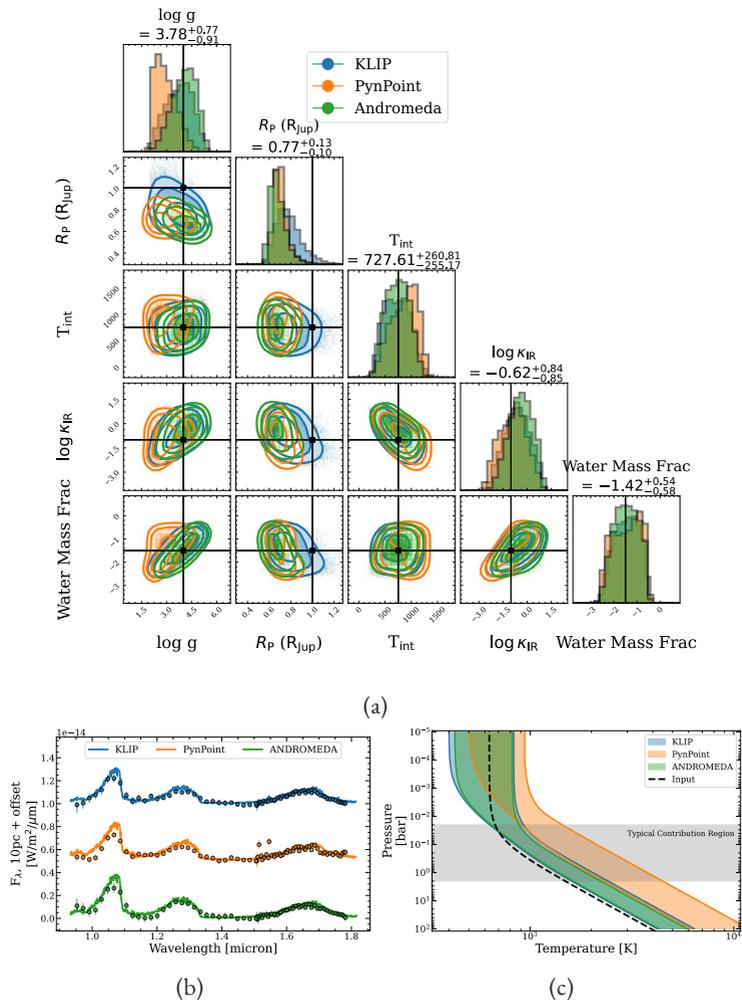


Figure 14.2: Results of retrievals comparing each of the three data processing algorithms. a: Posterior distributions for retrievals from each data processing algorithm. Contours are plotted for 2D Gaussian 1,2 and 3σ levels, corresponding to 36%, 86% and 99% confidence intervals, and the ground truth value is marked in black. The text labels correspond to the KLIP retrieval. b: Median-fit spectrum from retrievals on each algorithm, with the covariance (other than for ANDROMEDA). c: P-T Profiles for each retrieval. The shaded region indicates the 68% confidence region for the retrieved profile.

data lead to differences in the inferred parameters. In Figure 14.2, we compare the best fit results from each of the retrievals run on data processed using KLIP, PynPoint and ANDROMEDA to the ground truth spectrum injected into the IFS data. All three processing tools provide reasonable fits to the input spectrum, and share trends in the shape of their residuals, though KLIP provides the overall best reproduction of the input spectrum. The retrieved spectra tend to fit the input better at higher flux values, where the S/N is greater. Figure 14.2a shows the posterior distributions of most parameters; regardless of the retrievals setup the absolute uncertainties on all of the retrieved parameters are large. This highlights the importance having high S/N inputs to obtain precise constraints, as well as broad wavelength coverage to have sensitivity to a wide range of parameters. Not included in the plot are T_{equ} , γ and the CO mass fraction, none of which are constrained in any retrieval. For widely separated planets, T_{equ} is small and has little impact on the shape of the pressure temperature profile. As such, the Guillot profile is effectively reduced to the Eddington term, which does not depend on γ . Finally, there are no strong CO features present in the wavelength range considered in the injected planets. Thus we do not expect any of these parameters to impact the spectrum enough to be constrained by this retrieval. We neglect these unconstrained parameters when calculating the distance d_M .

KLIP performs the best of the three algorithms; accurately fitting the spectrum and retrieving the input parameters, measured from the χ^2/n_{data} and d_M respectively, as presented in table 14.2. PynPoint performs somewhat worse again, though the GPI data suffers from two outlying data points, and the measured uncertainties are generally smaller than for the KLIP or ANDROMEDA extractions. The physical interpretation of the PynPoint is significantly different than that of KLIP or ANDROMEDA: the inferred mass from the median $\log g$ and planet radius is more than a factor of 10 smaller when using the PynPoint parameters. This highlights the need for feedback between modelling and data analysis, as well as for comparison both between different data analyses and different models. This strongly impacts the measurement of $\log g$, which is sensitive to the shape of the H-band. Finally ANDROMEDA fails to reproduce the input spectrum well, but recovers the input parameters more accurately than PynPoint, though confidently excluding the true planet radius.

We find that for all of the retrieval setups, the median parameter values provide a better estimation of the true input parameters than the single maximum likelihood fit. For KLIP, we find that the median internal temperature estimate of 724 ± 260 K accurately, if imprecisely measure the true value of 750 K. However, the best fit value of 498 K is strongly biased from the true value, as are the remaining parameters. We therefore continue using only the

median parameter estimates, rather than the maximum likelihood fits.

The goodness-of-fit metrics provide a mechanism to select between the different retrievals. We find that all of the metrics favour the KLIP retrieval, though noting that the reduced χ^2 for ANDROMEDA is smaller, but does not account for the covariance. Based on the variation in the BIC and the Mahalabois distance, the effect that the algorithm choice has on the retrieved parameters is significant. Interpreting the Mahalabois distance as standard deviations from the truth, ANDROMEDA and PynPoint are 2.3 and 3.3 standard deviations less accurate than KLIP respectively. The trend of the BIC follows that of d_M , favouring KLIP, followed by Andromeda and then PynPoint. As the ground truth is not generally known, this reinforces the use of the BIC or a similar metric (such as the Bayes factor) as a robust metric for selecting between models, even when the data is also varied.

None of the retrievals retrieve the true input pressure temperature profile to within 1σ , as is evident from Figure 14.2c. However, in the region where the emission contribution is located, the retrieved PT profiles share a similar slope to the true input, at slightly higher temperatures. Such discrepancies highlight the importance of broad wavelength coverage in atmospheric retrievals, where the spectrum can probe different pressure, and thus temperature, layers of the atmosphere.

Effects of principal component optimisation

We also compare how the number of principal components used in the data processing impacts the retrieval results. This effect is more apparent at lower S/N so for this particular case we choose an injection at 400 mas, extracting the spectra with KLIP. Figure 14.3 highlights the impact that the choice of the number of PCs has on the precision of the posterior distributions. We compared the optimal extraction (6PCs for GPI and 8 for SPHERE) to an extraction using 25 PCs for each dataset. While both extractions retrieve the input parameters with similar accuracy, the optimised extraction is significantly more precise in its measurement of the planet radius.

14.3. Impacts of including the covariance in retrievals

While the choice of algorithm produces most of the difference in the spectral shape, we also consider how including the covariance into the log-likelihood calculation of the retrieval impacts the retrieved spectrum and inferred parameters. To understand this, we compare a KLIP retrieval with and without the use

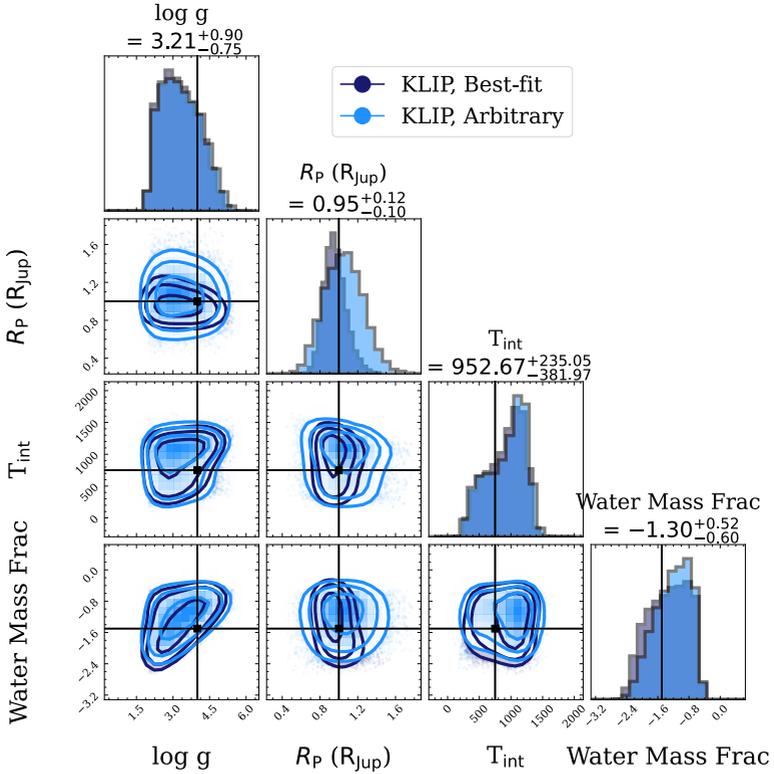


Figure 14.3: Corner plot comparing the retrieved parameter distributions for two different KLIP reductions. In light blue, the input spectrum was optimised using the relative discrepancy metric (Figure 12.4), while in dark blue an arbitrary extraction was chosen for each of the SPHERE and GPI datasets, reflecting a non-optimal parameter selection.

of the covariance matrices, and how these results compare to the Gaussian and noise-free cases.

The right panel of Figure 14.1 shows each of these datasets compared to the true input model, while Figure 14.4 shows the posterior distributions, best fits and PT profiles. We find that our retrievals reproduce the results of [Greco & Brandt \(2016\)](#): incorporating the covariance matrix improves the accuracy of the retrieval, at the cost of lower precision. Including the covariance matrix

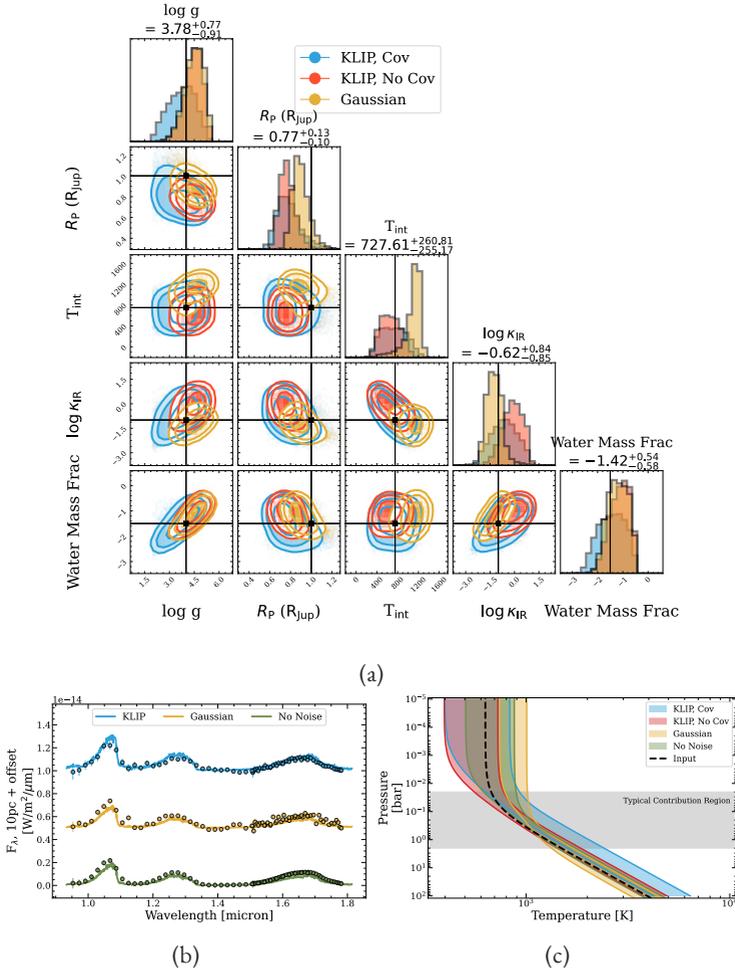


Figure 14.4: Results of retrievals using KLIP, comparing each the cases of computing the likelihood using the full covariance matrix (blue), the diagonal elements only (yellow) and using truly Gaussian scattered data (red). a) Posterior distributions for retrievals from KLIP. Contours are plotted for 2D Gaussian 1,2 and 3 σ levels, corresponding to 36%, 86% and 99% volume regions, and the ground truth value is marked in black. The text labels correspond to the retrieval including the covariance matrix. b) Median spectrum for the KLIP retrieval with covariance, as well as for the Gaussian (yellow) and noise free cases (green). c) P-T Profiles for KLIP, Gaussian, and noise-free retrievals. The shaded region indicates the 68% confidence region for the retrieved profile.

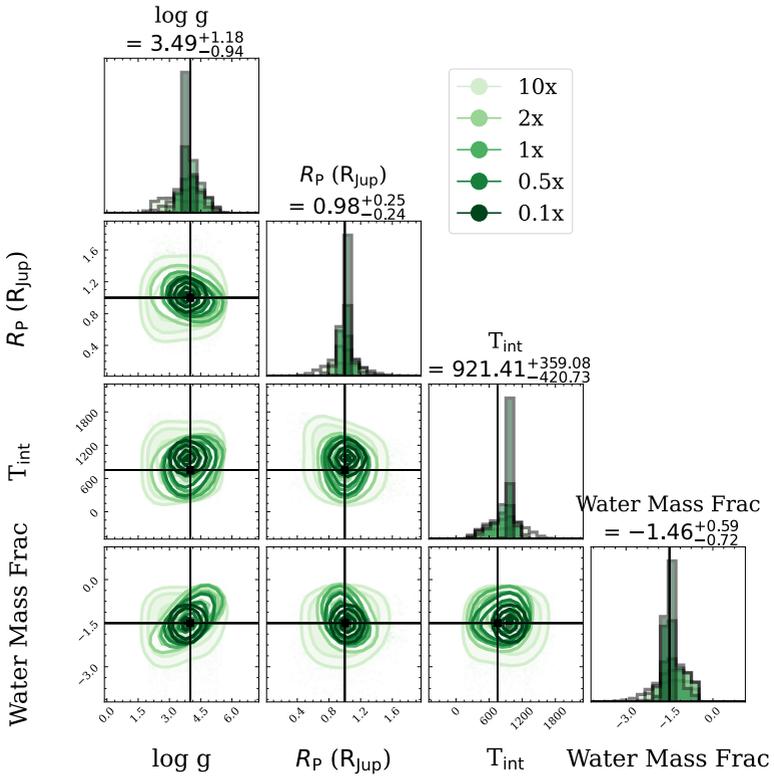


Figure 14.5: Posterior distributions for retrievals in the ‘Noise-free’ case. The nominal uncertainties are from $\text{diag}(\hat{C})$ of the KLIP extraction, and were scaled by factors from 0.1 to 10. The titles list the uncertainties for the 10x case.

reduced the posterior bias for all parameters. However, the $\text{diag}(\vec{C})$ case was able to more accurately retrieve the input PT profile, even though the parameter distributions were more discrepant from the true input parameters.

Quantitatively, including the covariance improved the d_M by 1 compared to the $\text{diag}(\vec{C})$, approaching the Gaussian measurement of 2.21. As in Section 12.4, the χ^2/n_{data} is underestimated. This is also reflected by the BIC, which favours the inclusion of the covariance matrix over both the $\text{diag}(\vec{C})$ case and the Gaussian case. In the KLIP extraction of the GPI data, there is rather broad covariance, which is easier to fit, as demonstrated by the toy model of 12.4.

Within the pressure range probed, all of the retrievals measure similar slopes, and correct temperatures to within 2σ , with the noise-free and $\text{diag}(\vec{C})$ cases most accurately retrieving the input profile. The correlation between atmospheric parameters of T_{int} and κ_{IR} in Figure 14.4a shows the difficulty in inferring atmospheric properties, and explains the inability of the retrievals to perfectly infer the temperature structure. For the KLIP data, the inclusion of the covariance matrix in the log-likelihood improves the accuracy of the constraint on both parameters.

We repeated this experiment using the PynPoint extractions. Consistent with the KLIP results, we find that including the covariance improves the accuracy of the retrieval, at a marginal cost to the precision of the retrieved parameters.

14.4. Relation between measurement and posterior precision

With the nominal uncertainties, the noise-free case can reproduce the input data to within 1σ across the entire wavelength range. However, at the level of precision of these measurements none of the retrievals are able to put strong constraints on any of the measured parameters, though this would be improved with more precise spectroscopic measurements. Using the noise-free case, we explored how the measurement precision affects the posterior precision by scaling the uncertainties by factors from 0.1 to 10, shown in Figure 14.5.

In all cases, the parameters are accurately retrieved, and the posterior precision increases as the uncertainties decrease. In the nominal case, the mean S/N per channel is 5 in the SPHERE wavelength range and 3 in the GPI-H band. With this precision and wavelength coverage, even in the optimistic case of no scatter in the data the internal temperature can only be constrained to within ± 223 K. This improves to ± 26 K if the S/N is improved by a factor of 10: while low S/N may be sufficient for accurate retrievals, high S/N is required

| Dataset | χ^2/n_{data} | $\log \mathcal{L}$ | $\log_{10} \mathcal{Z}$ | BIC | d_M |
|------------------------|--------------------------|--------------------|-------------------------|-------|-------|
| KLIP, C | 0.69 | 2692 | 1166 | -5289 | 2.35 |
| KLIP, diag(C) | 0.41 | 2656 | 1149 | -5218 | 3.29 |
| Offset GPI | 0.42 | 2657 | 1147 | -5214 | 3.29 |
| Offset SPH | 0.36 | 2659 | 1148 | -5218 | 3.40 |
| Scale GPI | 0.51 | 2666 | 1152 | -5233 | 4.48 |
| Scale SPH | 0.56 | 2659 | 1149 | -5218 | 3.93 |
| Scale GPI Err. | 1.07 | 2684 | 1158 | -5268 | 4.52 |
| Scale SPH Err. | 0.70 | 2660 | 1150 | -5220 | 4.73 |
| Scale Both. | 1.15 | 2684 | 1158 | -5263 | 4.54 |
| 10 ^b Both. | 0.43 | 2567 | 1148 | -5210 | 3.41 |
| GPI Only, C | 0.51 | 1320 | 571 | -2596 | 0.62 |
| SPH Only, C | 1.00 | 1373 | 593 | -2698 | 2.75 |

Table 14.3: Summary statistics for KLIP retrievals, including retrieved parameters to account for systematic biases. $\log_{10} \mathcal{Z}$ is the Bayesian evidence, the difference of which is the Bayes factor between two models. “Scale” indicates a multiplicative factor applied to the specified dataset, while “Offset” indicates an additive term.

to precisely measure the physical parameters. This finding complements Figure 14.3, highlighting that the main impact of the principal component optimisation is to improve the precision of the posterior distributions, as the choice of PCs impacts the precision of the spectroscopic measurement. Our choice of metrics are also validated, as the noise-free case is favoured by every metric (when comparing to diagonal only cases).

14.5. Mitigation using nuisance parameters

One challenge to the interpretation is the suggestion of overfitting by the reduced χ^2 values. A $\chi^2/n_{\text{data}} < 1$ suggests overfitting, although it can also be interpreted as overestimated uncertainties, or underestimated correlation (which effectively translates into overestimated uncertainties as well). The

retrievals both with and without covariance on the KLIP dataset both have $\chi^2/n_{\text{data}} < 1$. As the number of parameters is much lower than the number of data points and is identical to the ground truth model, this suggests that the uncertainty of the KLIP data is overestimated. In this section we explore the use of various parameterisations to account for systematics in the data, such as including offsets or scaling factors in the retrievals. As all of these comparisons use the same data, and additional parameters are properly included in the prior volume, we can now use the Bayes factor from Table 14.3 to quantitatively select the best model. We refer to Table 2 of [Benneke & Seager \(2013\)](#) for our interpretation of the Bayes factor: $\log \Delta \mathcal{Z}_{H_2, H_1} > 10$ is strong evidence in favour of model H_2 over H_1 .

Beginning with offsets, we fix one dataset and allow the other dataset to float, with a uniform prior of $\mathcal{U}(-10^{-14} \text{W/m}^2/\mu\text{m}, 10^{-14} \text{W/m}^2/\mu\text{m})$. We find that while allowing for offsets may marginally improve the fit to the data ($\chi_{\text{SPH, offset}}^2/n_{\text{data}} < \chi_{\text{KLIP, diag(C)}}^2/n_{\text{data}}$), it does not improve the accuracy or precision of the posteriors ($d_M > d_{M, \text{KLIP, diag(C)}}$), and is not favoured by the Bayes factor or BIC.

Next, we multiply one dataset and its corresponding uncertainties by a scaling factor ($\mathcal{U}(0.5, 2.0)$), fixing the remaining dataset. We find that a scaling factor of 0.73 ± 0.05 for the GPI dataset is somewhat favoured by the Bayes factor ($\Delta \log_{10} \mathcal{Z} = 3$), though the posterior precision and accuracy is somewhat reduced. Scaling the SPHERE data did not significantly improve the fit, and is not favoured by the Bayes factor.

This result is emphasised when we scale only the uncertainties for each of the datasets. We ran three retrievals: scaling the uncertainties of each dataset and fixing the other, or scaling both datasets with independent scaling factors simultaneously. To avoid hitting the prior boundaries, the scaling factor is given a uniform prior of $\mathcal{U}(0.05, 2.0)$. We find that a scaling factor of 0.28 ± 0.04 is favoured ($\Delta \log_{10} \mathcal{Z} = 9$) for the GPI dataset in both retrievals where the uncertainties are allowed to float, while allowing the SPHERE uncertainties to float does not change the fit. This implies that either the KLIP uncertainties for the GPI dataset are underestimated, or that the correlation is not correctly accounted for, as described in Section 12.4. Allowing the uncertainties to float improves the fit compared to the retrieval using only the diagonal components of the covariance matrix, but is still disfavoured compared to the retrieval using the full matrix.

[Line et al. \(2015\)](#) introduces a different parameterisation to scale the uncertainties and reduce overfitting. Using a parameter b , the uncertainty on the i^{th} wavelength bin is inflated as:

$$s_i^2 = \sigma_i^2 + 10^b, \quad (14.5)$$

where σ_i is the uncertainty on that bin. This will only allow for an increase in the size of the error bars, and allows us to account for model uncertainties and missing physics, rolling the additional uncertainty into the marginalised posterior parameter distributions. The prior range on b is set from -36 to -26: this encompasses the suggested range from [Line et al. \(2015\)](#) such that $0.01 \times \min(\sigma^2) < 10^b < 100 \times \max(\sigma^2)$. We retrieve b independently for each dataset used in the retrieval. Using this formalism, we find that the b parameterised retrieval is disfavoured compared to the KLIP retrievals, both with and without the use of the covariance matrix. Likewise, the d_M measured for this case (3.41) suggests a marginal decrease in accuracy relative to the baseline retrievals. This suggests that the extra parameters used to fit the b parameter are not justified to help resolve the problem of overfitting of the spectrum. As this formalism can only inflate the uncertainties, it is unsurprising that it cannot correct for the overestimated GPI uncertainties as shown by the scaling factor retrievals and the small reduced χ^2 values.

We conclude that scaling factors and offsets are inadequate for accounting for systematic offsets in the data due to the data processing and correlated noise. While allowing the uncertainties to float was marginally favoured by the Bayes factor compared to the $\text{diag}(\vec{C})$ KLIP retrieval, it was still strongly disfavoured compared to the retrieval using the full covariance matrix. Alternative methods, such as Gaussian process regression ([Wang et al., 2021](#); [Xuan et al., 2022](#)) may be able to overcome these limitations and allow for the characterisation of systematics in a Bayesian framework.

14.6. Limitations

This work reflects many of the best practices used in both data analysis and atmospheric retrievals, but remains an optimistic assessment of our ability to infer both accurate and precise physical parameters. Additional sources of bias are inevitably present in the data, such as the differences in spectra arising from different reduction pipelines as shown in appendix 16.A and from the process of building the 3D cubes from the 2D detector frames. Exoplanet data relies heavily on precise photometry, yet the host star which is used as a calibration source is obscured behind the coronagraph during the observations, making temporal monitoring of the PSF challenging. Finally, we use the same model for both injection and as the basis of the retrieval, ultimately ignoring many key physical processes present in real exoplanets. Even with these limitations, we remain optimistic about the prospects for retrievals to characterise directly imaged exoplanets, particularly in the era of high precision, broad wavelength spectroscopy as enabled by VLTI/GRAVITY, JWST, and the ELTs.

15. SUMMARY & CONCLUSIONS

Based on our comparison of high contrast imaging algorithms, it is clear that systematic variations are a more significant contribution to uncertainty than random errors for directly imaged exoplanet spectra. Such variations often lead to spectral correlation of the data, and knowledge of the length scale and strength of said correlation is crucial to accurate interpretation of the data. We used the methods of [Greco & Brandt \(2016\)](#) to compute covariance matrices for IFS data, and demonstrated that correlations in the data can both increase or decrease the posterior width of model parameters, depending on whether the parameter is sensitive to wavelength scales greater or less than the correlation scale. Using injection testing, we optimise our choice of algorithm parameters. We find that using only the S/N as a metric to determine the quality of spectral extraction does not produce optimal extractions. Instead, data processing parameters should be tuned using injection testing, with careful consideration of what goodness-of-fit metric should be used. Using the mean relative discrepancy, we optimised the number of principal components used in PSF subtraction in order to optimally extract the companion spectrum. Of course the number of principal components used in PSF subtraction is not the only source of systematic biases during spectral extraction: the precision of the astrometric solution, choices in processing both the science frames and the unsaturated PSF frames and the details of parameter choice all introduce biases on a similar level to the number of principal components, and must be independently optimised.

Each algorithm considered performed best under different conditions: the contrast, separation, observing conditions and data volume all impact which algorithm produces the optimal extraction. Care must be taken as the parameter choices that lead to the most sensitivity in order to detect companions are often different than the parameters required to robustly extract the spectrum in order to characterise the planets. During independent comparisons such choices led to statistically significant differences in both the shape and overall flux calibration of the extracted spectra. Without a priori knowledge of the spectrum, it is therefore necessary to compare multiple independent measurements in order to determine the underlying spectral shape.

By performing atmospheric retrievals on data processed using different algorithms we show that the variation between different data reductions is larger than the statistical posterior uncertainty. Model choice is highly dependant on data quality and quantity, and Bayesian comparisons should be performed to determine whether model complexity is suitable given the data. The ideal solution is to fully understand and correct for systematics during the data processing,

with broad data coverage and high spectral resolution. When comparing models, statistical tools such as the BIC or the Bayes factor should be used with care, and only when the free parameters are fully incorporated into the retrieval process to account for their impact on the prior volume and posterior distribution. In such a Bayesian framework, we find that the median parameter values are accurate measurements of the true input parameters, but that the maximum likelihood values are often strongly biased. Even using the median values, the difference in retrieved parameters from different data processing tools is significant, and can lead to dramatically different astrophysical interpretations. Retrievals should be performed on multiple data reductions to ensure that the retrieved parameters are robust to such variation.

We used the Mahalanobis distance, d_M to measure the distance between the true input parameters and the posterior distributions. Using this metric, we found that accounting for the covariance in the likelihood function of a retrieval framework can help mitigate correlations in the data, but not entirely resolve them. Compared to using only the diagonal terms using the full matrix will reduce the bias in all parameters, at the cost of slightly decreased precision, reproducing the results of [Greco & Brandt \(2016\)](#). When the data are correlated including the covariance is necessary to make meaningful statistical statements about models fitting the data. In all cases, including the full covariance matrix leads to improved accuracy of the inferred planet parameters. When testing the use of scaling factors and offsets to try to correct for these systematic biases we found worse results than when relying on the covariance matrix. Thus we recommend that the covariance matrix always be published along with IFS data of exoplanets.

The systematic biases of the spectral extractions fundamentally limit the accuracy with which we can understand exoplanet atmospheres, with effects that can be much more significant than the statistical posterior precision. One solution to the issues discussed in this work is to acquire higher quality data. Nevertheless, it will remain important to measure the covariance and to understand systematic effects imparted by data processing.

We would like to thank the anonymous referee for their insightful and detailed report, which substantially improved the quality of this manuscript. In addition we'd like to thank Alice Zurlo and Alex Greenbaum for supplying the data used in this work. Software used: `petiRADTRANS`, `pyKLIP`, `PynPoint`, `VIP-HCI`, `species`, `pyMultiNest`, `Python`, `numpy`, `astropy`, `phot_utils`, `matplotlib`, and `scicomap`.

16. APPENDICES

16.A. SPHERE Data Reduction

We begin the process of extracting planetary spectra by applying a range of pre-processing steps, hereafter referred to as the data reduction stage. Dark frames, detector flats and IFS flats are subtracted from the data. The spectral positions of each slice and the wavelengths are calibrated. Bad pixels and cross talk are corrected, and the science frames are background subtracted. ND filter transmission profiles are applied to stellar flux observations taken at the beginning and end of the ADI sequence observations. To centre the science frames, the satellite spots are used where available. If available, satellite spots are used to calibrate the wavelength centre of each channel. These frames have anamorphism corrected and are shifted to a common centre, and are output as a set of files with dimensions of (x, y, λ) .

For SPHERE data, we compare the results of using the standard SPHERE Data Center pipeline for the reduction to that of [Vigan \(2020\)](#). Recent updates to the pipeline have shown discrepancies in the wavelength solution for YJH data, however as this is only applicable to data taken in the satellite spot mode. Lacking satellite spots, we instead rely on the standard ESO wavelength solution, and remain unaffected by the changes.

As it is challenging to inject fake companions into the raw detector frames for an IFS, we instead use the extracted spectrum of HR 8799 e for our comparison. The raw frames were processed using both pipelines, and post-processed using KLIP in order to extract the planetary spectrum.

Figure 16.1 shows the results of this extraction. While qualitatively similar, the wavelength solutions are different between the pipelines, and both the location and amplitude of the $1.3 \mu\text{m}$ feature disagree. Finally, there is significant discrepancy between the shapes of the spectra in the H band. Without a clear metric for selecting a reduction pipeline, we choose to use the most up-to-date implementation of [Vigan \(2020\)](#) due to its ease of use and Python-based interface.

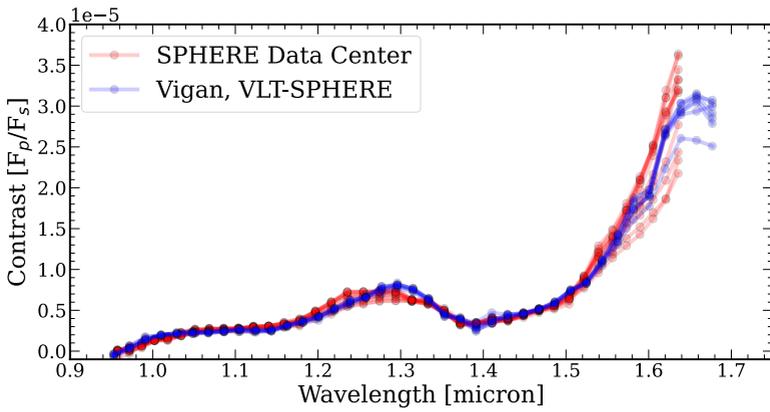


Figure 16.1: For the SPHERE data we compare the standard SPHERE Data Center data reduction (red) to the VLT-SPHERE pipeline described in [Vigan \(2020\)](#) (blue). As planets cannot be injected into the raw data, we compare the spectrum of HR 8799 e as extracted with KLIP, where each measurement in the figure represents a different number of principal components used in the spectra extraction.

FOUR-OF-A-KIND? SYSTEMATIC CHARACTERISATION OF THE HR8799 PLANETS WITH VLT/GRAVITY

This chapter has been reproduced from [Nasedkin et al. \(2024\)](#), which has been accepted for publication in *A&A*. I am the lead author of this work, supervised by Paul Mollière. The VLT/GRAVITY data were reduced by Sylvestre Lacour and Mathias Nowak. Tomas Stolker assisted in running several of the retrievals discussed in this study. Will Balmer, Jens Kammerer, Jinyi Shangquan and Jason Wang all assisted in obtaining the GRAVITY observations. All coauthors are members of the ExoGRAVITY or GRAVITY collaborations, and provided comments on the manuscript.

“Finally, let us have our imagination represent a wonderfully strange object such as a burning sun as if it were from up close. In one glance we broad lakes of fire lifting their flames up to the sky, raging storms whose fury redoubles the violence of the former, dense vapours that choke the fire and, raised by the force of the winds, constitute dark clouds which in turn crash down in fiery showers of rain.”

Immanuel Kant, 1755

Abstract

With four companions at separations from 16 to 71 au, HR 8799 is a unique target for direct imaging, presenting an opportunity for a comparative study of exoplanets with a shared formation history. Combining new VLTI/GRAVITY observations obtained within the ExoGRAVITY program with archival data, we performed a systematic atmospheric characterisation across all four planets. We explored different levels of model flexibility to understand the temperature structure, chemistry, and clouds of each planet using both `petitRADTRANS` atmospheric retrievals and fits to self-consistent radiative–convective equilibrium models. Using Bayesian model averaging to combine multiple retrievals (a total of 89 across all four planets), we find that the HR 8799 planets are highly enriched in metals, with $[M/H] \gtrsim 1$, and have stellar to superstellar atmospheric C/O ratios. The C/O ratio increases with increasing separation from $0.55^{+0.12}_{-0.10}$ for d to $0.78^{+0.03}_{-0.04}$ for b, with the exception of the innermost planet, which has a C/O ratio of 0.87 ± 0.03 . Such high metallicities are unexpected for these massive planets, and challenge planet-formation models. By retrieving a quench pressure and using a disequilibrium chemistry model, we derive vertical mixing strengths compatible with predictions for high-metallicity, self-luminous atmospheres. Bayesian evidence comparisons strongly favour the presence of HCN in HR 8799 c and e, as well as CH₄ in HR 8799 c, with detections at $> 5\sigma$ confidence. All of the planets are cloudy, with no evidence of patchiness. The clouds of c, d, and e are best fit by silicate clouds lying above a deep iron cloud layer, while the clouds of the cooler HR 8799 b are more likely composed of Na₂S. With well-defined atmospheric properties, future exploration of this system is well positioned to unveil further details of these planets, extending our understanding of the composition, structure, and formation history of these siblings.

17. INTRODUCTION

Directly imaged exoplanets provide an ideal laboratory for understanding the formation and evolution of planetary systems. These young systems provide unique insight into widely separated companions ($\gtrsim 10$ au): by directly measuring their emission spectra, we can peer into regions of their atmospheres inaccessible through other techniques. While spectroscopically similar to their brown dwarf cousins, these young, low-surface-gravity exoplanets display unique spectral shapes and colours (Faherty, 2018), indicative of differences in their chemistry, clouds, and formation history (Marley et al., 2010, 2012; Charnay et al., 2018).

The HR 8799 system is a benchmark target for directly imaged exoplanets, containing four planets (Marois et al., 2008, 2010), an inner debris disc (Boccaletti et al., 2023), and an outer Kuiper-belt-like disc (Su et al., 2009). This is among the best-studied systems of exoplanets, with a wide range of photometric and spectroscopic data. The spectroscopic data cover the near-infrared region (1–4 μm) at varying spectral resolution, while the photometric data extend out to 15 μm with the recent addition of JWST/MIRI observations (Boccaletti et al., 2023). Most of these studies, together with extensive modelling work, have tried to answer the following main questions:

1. How did the HR 8799 system form?
2. What are the dynamics of the system? Is it stable, and how do the planets and disc interact?
3. What are the atmospheres of each planet made of, and how have they evolved through time?

In the present work, we attempt to directly answer question (3), which has implications for question (1). Using Bayesian atmospheric retrievals (e.g. Madhusudhan, 2019) as well as fits to 1D self-consistent models, we infer the atmospheric properties of each of the four planets. To date, the only comprehensive retrieval study of all four of the HR 8799 planets was by Lavie et al. (2017). New, high-precision K-band spectra obtained with the VLTI/GRAVITY as part of the ExoGRAVITY program (Gravity Collaboration et al., 2019; Lacour et al., 2020), together with updated atmospheric models and opacity databases, provide motivation and the means to perform a systematic reanalysis of this system.

With effective temperatures in the range of 1000–1400 K, the HR 8799 planets sit in the middle of the L/T transition (Kirkpatrick et al., 1999). This

spectral transition is marked by changes in chemistry between L- and T-type objects, from CO-dominated carbon chemistry in the hotter objects to methane chemistry as the temperature falls below ~ 1300 K. While this transition is well established for brown dwarfs, detections of CH_4 in exoplanets remain elusive: there have been tentative detections in the atmosphere of HR 8799 b (Barman et al., 2015; Ruffio et al., 2021), but the only convincing detections have come from JWST observations of cool, transiting exoplanets (Bell et al., 2023; Madhusudhan et al., 2023) and the coldest directly imaged exoplanets, such as 51 Eridani b (Brown-Sevilla et al., 2023; Whiteford et al., 2023). This ‘missing methane’ is thought to be driven by convective upwelling in the atmospheres, dredging material from deeper, hotter regions of the atmosphere where CO is more favoured by equilibrium chemistry (Fegley & Lodders, 1996). Precise constraints on the abundance of both CO and CH_4 would allow better constraints on this vertical mixing, which is typically parameterised by the vertical diffusion coefficient K_{zz} .

The sharp change in colour in the L/T transition is thought to be caused by the sinking of silicate clouds through the atmosphere, as the cloud base shifts deeper as the effective temperature decreases (Burrows & Sharp, 1999). Once the cloud base sinks below the photosphere, the impact of the cloud opacity is increasingly removed from the spectrum, causing the blueward shift characteristic of T dwarfs as the effective temperature falls below 1300 K. Following the mid-infrared observations with *Spitzer*/IRS (Cushing et al., 2006, 2008), Suárez & Metchev (2022) identified a trend in the silicate absorption feature at $9\ \mu\text{m}$ as a function of temperature. The strength of this absorption feature was found to correlate positively with the near-infrared colour for L dwarfs, which is often used as a proxy for cloudiness. The HR 8799 planets lie comfortably below the temperature at which silicate clouds are expected to occur entirely below the photosphere, yet their red colour and near-infrared spectral shape are thought to be clear hallmarks of thick silicate cloud coverage (Mollière et al., 2020). However, Line et al. (2015) and Suárez & Metchev (2023) find that these clouds are not only sensitive to temperature, but also to surface gravity, which plays a role in determining the size and therefore settling speed of the aerosol particles. As young, giant exoplanets still retain significant heat from formation, their atmospheres remain inflated due to low surface gravity, which will in turn result in cloud properties that are unique to this class of object; observations of VHS 1256 b (Miles et al., 2023) remain the only spectroscopic observations of a silicate feature in a directly imaged planet. Burningham et al. (2021) and Vos et al. (2023) use atmospheric retrievals to identify the detailed structure and composition of the clouds, providing for the first time evidence to support the use of particular cloud compositions in these substellar atmospheres.

The mechanism through which four super-Jupiter planets can form in a single system is unclear. The presence of both an inner and an outer debris disc implies that the planets formed within a circumstellar disc; that is to say they did not form like stars. However, it is still unclear whether these objects formed through gravitational instability (GI) (Bodenheimer, 1974; Adams et al., 1989) or via core accretion (Pollack et al., 1996). Evolutionary models (Saumon & Marley, 2008) suggest that the current temperatures of the planets suggests hot-start boundary conditions for their evolution, which is more typically associated with GI (but also see Mordasini et al., 2017). GI models, such as that of Helled & Bodenheimer (2010), find that the amount of heavy elements accreted by the planets should be small, implying nearly stellar compositions for all four planets. Likewise, current composition estimates suggest that the planets share a C/O ratio with their host star (Hoch et al., 2023), but may be slightly enriched in metals, leading to tension with the predictions of the GI models.

In addition to understanding the formation mechanism, Mollière et al. (2022) present a framework through which we can infer the conditions of the formation environment from measured atmospheric parameters. However, these authors, and many others (e.g. Eistrup et al., 2018; Cridland et al., 2019, 2020; Turrini et al., 2021; Pacetti et al., 2022), demonstrate that this is not a straightforward task. The Öberg et al. (2011) model links the planet C/O ratio to the location of formation relative to snow lines in the disc. This model provides a simplified view through which we can understand the impact of disc conditions on the outcomes of planet formation, but the complex and time-evolving physics and chemistry of discs and forming planets make solving the inverse problem challenging. Nevertheless, the best hope for linking the atmospheric properties back to the protoplanetary disc is to infer robust atmospheric elemental abundances and link these to interior models to determine the bulk planetary composition (Guillot, 2005; Fortney et al., 2011), thereby determining what disc conditions could lead to the diversity of planet-formation outcomes.

While new data and modelling techniques are beneficial, interpreting such model–data comparisons for exoplanet spectra is far from trivial. Multiple techniques must be studied simultaneously to paint a consistent portrait of these worlds. Biases in inferred planet parameters are a common challenge in direct-imaging analyses: fits to emission spectra often find unphysically small radii that are inconsistent with evolutionary tracks (Marley et al., 2012). Retrievals using free molecular abundances tend to find higher C/O ratios than when disequilibrium chemistry models are considered (Lavie et al., 2017; Wang et al., 2020a), possibly due to additional oxygen sequestered in refractory clouds (Fonte et al., 2023). The inferred effective temperatures (T_{eff}) of each planet

can vary by hundreds of kelvin, within a region of parameter space where the chemical timescales can vary by orders of magnitude over tens of kelvin (Zahnle & Marley, 2014). Complicating matters further are the known discrepancies between spectral measurements (such as between SPHERE and GPI in the H-band; see Mollière et al., 2020), leading to uncertainties in both the shape and overall flux calibration of the spectra that are not reflected in the formal uncertainties. Attempting to address this problem, Nixon et al. (2023) demonstrate the use of Bayesian model averaging (BMA), which can be used to combine the posterior distributions of multiple models, allowing some degree of model uncertainty to be formally incorporated into the inferred parameter uncertainties. Finally, Greco & Brandt (2016) and Nasedkin et al. (2023) demonstrate the importance of properly accounting for the covariance in low-resolution IFS data — a thorough treatment of IFS data is necessary to ensure meaningful and unbiased posterior probability distributions.

Many of the questions of chemistry and formation will be addressed through the use of the various instruments aboard JWST. This telescope will open new observational windows, extending out to the mid-infrared, and allow new characterisation methods, such as molecular mapping of the system (Patapis et al., 2022). Simultaneous measurements of CO and CH₄ features between 3 and 5 μm will allow constraints to be placed on the vertical mixing in the atmosphere, and more precise estimates of the C/O and metallicity. Nevertheless, ground-based observations remain crucial: the innermost companion will remain challenging to measure spectroscopically without a coronagraph; across most of its spectra, HR 8799 e is below the 2×10^{-5} contrast threshold at 300 mas obtained in Ruffio et al. (2023).

The present study provides a comprehensive examination of the atmospheres of the HR 8799 companions, making use of new, high- S/N observations obtained with VLTI/GRAVITY, together with a combination of retrieval methods and self-consistent modelling. We present further context and background information on the HR 8799 system in Section 18. The data used in this work are described in Section 19, while the details of the pet i tRADTRANS (pRT) forward model are described in Section 20, with the self-consistent models introduced in Section 20.7. The results of the atmospheric retrievals and self-consistent grid fits are presented in Section 21. We discuss the limitations of this study, additional work to validate our results, and the implications of our findings in Section 22. The appendices contain details of the data and data analysis (24.A, 24.B), model validation (24.C), implementation details (24.D), and tables of the complete set of retrieval results (24.E).

18. THE PLANETARY SYSTEM OF HR 8799

While HR 8799 is one of the most well studied exoplanetary systems (as seen in Figure 18.1), there remains significant disagreement in the literature both with respect to the spectroscopic measurements and the inferred planet properties. In general, these super Jupiters all host very cloudy atmospheres, with significant impacts of disequilibrium chemistry. The spectra of these planets show characteristics of low surface gravity and have been classified at the L/T transition, though the variability typical of these objects has not yet been observed in the companions. The composition of the companions is generally found to be moderately enriched compared to the host star, and while H_2O and CO are the dominant absorbers, the C/O ratio estimates vary significantly between models. The measurements of the bulk planet properties discussed in

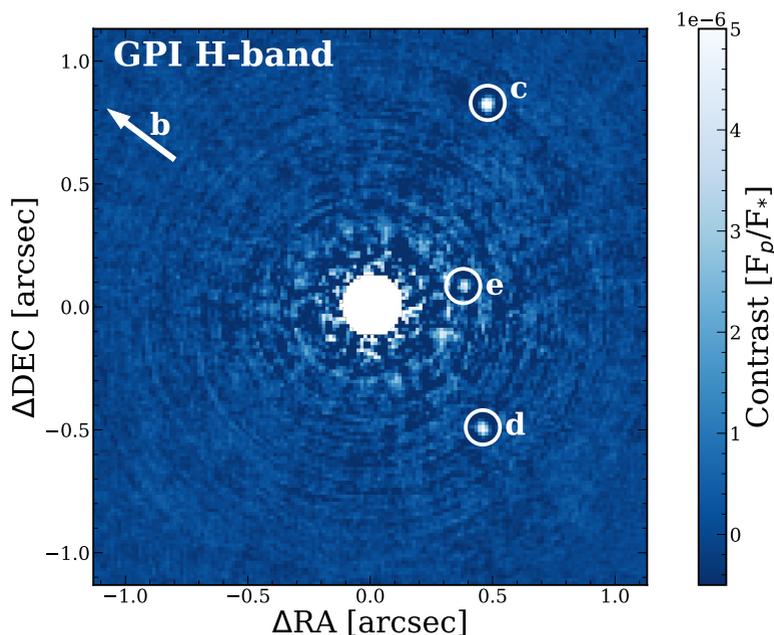


Figure 18.1: HR 8799 planets as imaged in the H-band with the Gemini/GPI IFU, originally published in [Greenbaum et al. \(2018\)](#). The IFU cube was processed using KLIP, and the image is mean combined along the spectral axis. HR 8799 b is outside the field of view of GPI.

this section, together with the results of this work are compiled in Tables 18.1, 18.2, 18.3, and 18.4 for planets b, c, d, and e respectively.

18.1. HR 8799 A: The host star

HR 8799 A is an A5V (Gehren, 1977; Cannon & Pickering, 1993) to F0+VkA5mA5 (Gray et al., 2003b) type star, host to four detected planets and an inner and outer debris disc (Su et al., 2009). It was one of the first identified γ Doradas pulsators (Kaye et al., 1999; Zerbi et al., 1999), and has been classified as a λ Boötis star (Sadakane, 2006; Moya et al., 2010) due to its depletion of heavy elements in the atmosphere. Due to this depletion, the [Fe/H] of HR 8799 A is measured to be subsolar, with measurements ranging from -0.47 ± 0.10 (Gray & Kaye, 1999; Sadakane, 2006) to between -0.32 ± 0.1 and -0.12 ± 0.1 , depending on the inclination angle (Moya et al., 2010). TESS photometry allowed the measurement of this inclination angle, finding a core rotation period of ≈ 0.7 days, which combined with $v \sin i$ and stellar radius measurements would result in a preliminary stellar inclination of $\approx 28^\circ$ (Sepulveda et al., 2023), and would favour the higher metallicity case presented by Moya et al. (2010). Using high resolution spectroscopy from the LBT/PEPSI and HARPS instruments, Wang et al. (2020a) found a very subsolar iron metallicity of -0.52 ± 0.08 , but found the relative carbon (C/H = 0.11 ± 0.12) and oxygen (O/H = 0.12 ± 0.14) abundances to be consistent with solar as is characteristic of λ Boötis stars. The derived C/O ratio from their measurements was $0.54^{+0.12}_{-0.09}$. For this work, we use a BT-Nextgen stellar model as fitted in Nasedkin et al. (2023) with best-fit parameters of $T_{\text{eff}} = 7200$ K, $\log g = 3.0$, and [Fe/H] = 0.0, slightly cooler than the models used in previous studies (Zurlo et al., 2016; Greenbaum et al., 2018). However, this temperature is in line with Sepulveda & Bowler (2022), though based on their dynamical mass estimate of the host star and the radius measurement of Baines et al. (2012) they find a higher surface gravity of 4.28.

Most indicators place the HR 8799 system between 25 and 60 Myr in age. Using the debris disc as evidence, Zuckerman & Song (2004) and Rhee et al. (2007) estimate an age of 30 Myr. Zuckerman et al. (2011) identified HR 8799 as ‘a likely member of the ~ 30 Myr old Columba Association’, thus providing an age and an association of stars with a shared formation history to which we can compare HR 8799 and its companions. However, using asteroseismology Moya et al. (2010) found that an age of ~ 1 Gyr or greater is also compatible with their measurements. While we continue to use the standard ~ 30 Myr age for the system, we acknowledge that there remains some uncertainty in the age and activity of the host star.

Table 18.1: Summary of literature and derived planet properties for HR 8799 b.

| Ref. | Clouds | M | log g [M_{Jup}] | T_{eff} [gK] | R [K] | [M/H] [R_{Jup}] | C/O | log L_{bol}/L_{\odot} |
|-------------------|-----------|---------------------|-------------------------------|----------------------------|------------------------|-------------------------------|---------------------------|--------------------------------|
| B11a | Slab | 0.1–3.3 | 3.5 ± 0.5 | 1100 ± 100 | 0.63 – 0.92 | ... | ... | -5.1 ± 0.1 |
| C11 | Thick | 5–15 | 4 – 4.5 | 800 – 1000 | ... | ... | ... | ... |
| G11 | Slab | 1.8 | 4 | 1100 | 0.69 | ... | ... | ... |
| M11 | Power law | 2–12 | 3.5 – 4.3 | 750 – 850 | ... | ... | ... | ... |
| M12 | AM01 | 26 | 4.75 | 1000 | 1.11 | 0 | ... | -4.95 ± 0.06 |
| L13 | Slab | 16^{+5}_{-4} | $5.0^{+0.1}_{-0.2}$ | 900^{+30}_{-60} | $0.66^{+0.07}_{-0.04}$ | ... | 0.96 | -5.1 ± 0.1 |
| B15 | Slab | ... | 3.5 | 1000 | ... | ... | 0.55 – 0.7 | ... |
| B16 | ER4 | ... | 3.4 – 3.8 | 1100 – 1200 | 0.6 – 0.7 | 0.5 | ... | ... |
| L17 | Mie | ... | 4 ± 0.1 | $^{a}320 \pm 20$ | 1.08 ± 0.02 | ... | 0.92 ± 0.01 | ... |
| W21 ^b | BT-Settl | ... | $4.8^{+0.4}_{-0.8}$ | $1423.3^{+212.6}_{-278.4}$ | ... | ... | ... | ... |
| R21 ^c | Slab | ... | $3.1^{+0.03}_{-0.03}$ | 1180^{+14}_{-14} | ... | 0 | $0.578^{+0.004}_{-0.005}$ | ... |
| Best $A \cap B^d$ | AM01 | $6.0^{+0.3}_{-0.3}$ | $4.10^{+0.03}_{-0.04}$ | 942^{+12}_{-13} | $1.11^{+0.03}_{-0.03}$ | $0.96^{+0.08}_{-0.08}$ | $0.78^{+0.03}_{-0.04}$ | $-5.08^{+0.04}_{-0.04}$ |
| BMA $A \cap B^e$ | AM01 | $6.0^{+0.4}_{-0.3}$ | $4.10^{+0.03}_{-0.04}$ | 942^{+12}_{-15} | $1.10^{+0.03}_{-0.03}$ | $0.96^{+0.08}_{-0.08}$ | $0.78^{+0.03}_{-0.04}$ | $-5.08^{+0.04}_{-0.04}$ |
| $A \cap B^f$ | AM01 | $6.0^{+0.3}_{-0.3}$ | $4.10^{+0.06}_{-0.06}$ | 936^{+22}_{-34} | $1.11^{+0.08}_{-0.08}$ | $1.1^{+0.1}_{-0.2}$ | $0.73^{+0.04}_{-0.04}$ | $-5.08^{+0.06}_{-0.06}$ |
| Grids | Various | 4.6 – 6.0 | 3.5 – 4.5 | 850 – 1100 | 0.73 – 1.2 | > 0.7 | 0.3 – 0.55 | -5.14 to -5.28 |

Notes on clouds: both the ‘Thick’ and ‘Slab’ clouds are based on [Burrows et al. \(2006\)](#), and are vertically extended throughout the atmosphere above a base pressure, with a decaying mass fraction, though the slab clouds account for a greater range of aerosol opacities. The ‘Power law’ clouds parameterise the vertical extent and position using a power law, and fix the base pressure to the location at 2300 K. ‘AM01’ clouds balance the cloud sedimentation and vertical mixing to determine the particle size, and use the sedimentation fraction to determine the vertical extent. ‘ER4’ is the Exo-Rem4 model from [Bonnefoy et al. \(2016\)](#). ‘Mie’ clouds do not use physical optical constants, but directly fit for mie scattering parameters. The ‘BT-Settl’ cloud model is based on radiation hydrodynamical simulations that solve for the diffusion and mixing of aerosol particles ([Allard et al., 2012](#)).

References: B11a: [Barman et al. \(2011\)](#); C11: [Currie et al. \(2011\)](#); G11: [Galicher et al. \(2011\)](#); M11: [Madhusudhan et al. \(2011\)](#); M12: [Marley et al. \(2012\)](#); L13: [Lee et al. \(2013\)](#); B15: [Barman et al. \(2015\)](#); B16: [Bonnefoy et al. \(2016\)](#); L17: [Lavie et al. \(2017\)](#); W21: [Wang et al. \(2021\)](#); R21: [Ruffio et al. \(2021\)](#).

^a Only T_{int} , a model parameter, is reported.

^b W21 used high resolution spectroscopy, and did not infer masses or radii, using masses of $7.2 \pm 0.7 M_{\text{Jup}}$ for the inner three planets and $5.8 \pm 0.5 M_{\text{Jup}}$ for HR 8799 b. A radius of $1.2 \pm 0.1 R_{\text{Jup}}$ was used for all planets.

^c R21 uncertainties were limited by the coarseness and boundaries of their model grid.

^d Single best retrieval parameters.

^e Bayesian model averaged parameters from group $A \cap B$.

^f Unweighted average parameters from group $A \cap B$.

18.2. Photometric studies

The HR 8799 system has been the subject of extensive photometric characterisation, from the red-optical out to the mid-infrared with JWST/MIRI. The outer three companions were originally detected in [Marois et al. \(2008\)](#),

with HR 8799 e following in [Marois et al. \(2010\)](#). Many photometric studies (e.g. [Lafrenière et al., 2009](#); [Fukagawa et al., 2009](#); [Metchev et al., 2009](#); [Currie et al., 2011](#); [Bergfors et al., 2011](#); [Galicher et al., 2011](#); [Soummer et al., 2011](#); [Skemer et al., 2012](#); [Currie et al., 2012](#); [Esposito et al., 2013](#); [Skemer et al., 2014](#); [Currie et al., 2014](#); [Maire et al., 2015](#); [Rajan et al., 2015](#); [Petit dit de la Roche et al., 2019](#); [Biller et al., 2021](#); [Boccaletti et al., 2023](#)) have identified the companions as L/T transition objects, with near-infrared colours compatible with more extended clouds than L-dwarfs of similar temperatures. This is generally explained as a result of the young age and low surface gravity, where the lower gravity allows the condensate particles to remain aloft above the photosphere at lower temperatures. Even in the earliest studies, disequilibrium chemistry was used as an explanation for the drop in the continuum flux due to CO absorption at $4\ \mu\text{m}$ ([Currie et al., 2011](#); [Janson et al., 2010](#)).

In addition to the four known companions there have been many searches for a fifth companion, interior to HR 8799 e. [Thompson et al. \(2023\)](#) used long time baseline astrometry and deep L' imaging with Keck/NIRC2 to search for this hypothesised companion, finding that an additional companion fits both the astrometry and photometry better than a four planet solution, but does not result in a significant detection. For now, we only examine the four confirmed companions further.

[Bonnefoy et al. \(2016\)](#) explores the implications of the near-infrared photometry for all four of the companions, comparing them to spectrally similar field objects from the SpeX PRISM library. Empirically, the HR 8799 planets are much more red in colour than field dwarfs of similar spectral type. They also show that using the dereddening coefficients for corundum (Al_2O_3), iron (Fe), enstatite (MgSiO_3), and forsterite (Mg_2SiO_4) from [Marocco et al. \(2014\)](#), the colours of the companions more closely match those of field dwarfs. However, they cannot quantitatively distinguish the chemical composition of the clouds, which requires mid-infrared spectroscopic observations of condensate absorption features ([Burningham et al., 2021](#); [Miles et al., 2023](#)).

[Marley et al. \(2012\)](#) and [Bonnefoy et al. \(2016\)](#) use estimates of the surface gravity and radius from spectroscopic fits to constrain the overall mass and luminosity of the planets, which they in turn compared to planetary evolution models, such as those of [Baraffe et al. \(2003\)](#) and [Saumon & Marley \(2008\)](#). With self-consistent, radiative equilibrium models, the planet radius is often difficult to fit, with the radius underestimated by over 30% compared to expectations from the evolutionary models (e.g. [Bonnefoy et al. 2016](#)).

Table 18.2: Summary of literature and derived planet properties for HR 8799 c

| Ref. | Clouds | M | log g [M_{Jup}] | T_{eff} [cgs] | R [R_{Jup}] | [M/H] [R_{Jup}] | C/O | log L_{bol}/L_{\odot} |
|------------------|-----------|---------------------|-------------------------------|---------------------------|---------------------------|-------------------------------|---------------------------|--------------------------------|
| C11 | Thick | 7–17.5 | 4 – 4.5 | 1000 – 1200 | ... | 0 | ... | -4.7 ± 0.1 |
| G11 | Slab | 1.1 | 3.5 | 1200 | 0.97 | ... | ... | -4.7 ± 0.1 |
| M11 | Power law | 3 – 11 | 4.0 – 4.3 | 950 – 1025 | ... | ... | ... | -4.7 ± 0.1 |
| M12 | AM01 | 8 – 11 | 4.1 ± 0.1 | 950 ± 60 | 1.32 – 1.39 | 0 | ... | -4.90 ± 0.1 |
| K13 | ... | 3 – 7 | 3.5 – 4.0 | 1100 ± 100 | 1 – 1.5 | ... | 0.65 ± 0.1 | ... |
| B16 | ER4 | ... | 3.8 – 3.9 | 1200 | 1.0 | 0.5 | ... | ... |
| L17 | Mie | ... | 4.5 ± 0.1 | ${}^a 960 \pm 20$ | 1.25 ± 0.02 | ... | 0.55 ± 0.01 | ... |
| G18 | Various | ... | 3.5 – 4.0 | 1100 – 1350 | 0.7 – 1.2 | ... | ... | -4.58 to -4.82 |
| W20 | Deck | ... | $3.97^{+0.03}_{-0.03}$ | 1054^{+7}_{-5} | 1.47 ± 0.02 | ... | $0.58^{+0.06}_{-0.06}$ | -4.59 ± 0.004 |
| W20 ^b | Deck | ... | $3.95^{+0.04}_{-0.12}$ | 1102 ± 2 | 1.20 | ... | $0.39^{+0.06}_{-0.06}$ | -4.69 ± 0.0002 |
| W21 ^c | BT-Settl | ... | $5.4^{+0.1}_{-0.2}$ | $1474.4^{+24.4}_{-36.3}$ | ... | ... | ... | ... |
| R21 ^d | Slab | ... | $3.63^{+0.03}_{-0.02}$ | 1200^{*+}_{-14} | ... | 0 | $0.562^{+0.004}_{-0.005}$ | ... |
| W23 | AM01 | ... | $4.17^{+0.41}_{-0.47}$ | ${}^e 1421^{+92}_{-72}$ | $1.01^{+0.09}_{-0.08}$ | ${}^f 0.51^{+0.40}_{-0.43}$ | $0.67^{+0.12}_{-0.15}$ | ... |
| Best $A \cap B$ | AM01 | $8.5^{+0.4}_{-0.4}$ | $4.26^{+0.02}_{-0.03}$ | 1158^{+11}_{-12} | $1.10^{+0.01}_{-0.01}$ | $1.27^{+0.06}_{-0.06}$ | $0.66^{+0.01}_{-0.01}$ | $-4.71^{+0.02}_{-0.02}$ |
| BMA $A \cap B$ | AM01 | $8.5^{+0.4}_{-0.5}$ | $4.26^{+0.02}_{-0.03}$ | 1159^{+11}_{-12} | $1.10^{+0.01}_{-0.01}$ | $1.27^{+0.05}_{-0.06}$ | $0.66^{+0.01}_{-0.01}$ | $-4.71^{+0.02}_{-0.02}$ |
| $A \cap B$ | AM01 | $8.6^{+0.4}_{-0.4}$ | $4.25^{+0.04}_{-0.14}$ | 1159^{+24}_{-76} | $1.10^{+0.23}_{-0.03}$ | $1.2^{+0.1}_{-0.1}$ | $0.63^{+0.05}_{-0.02}$ | $-4.70^{+0.03}_{-0.03}$ |
| Grids | Various | 1.24–10.3 | 3.5–4.5 | 1100–1200 | 0.8 – 1.31 | > 1.0 | 0.3 – 0.8 | -4.65 to -4.72 |

References: C11: Currie et al. (2011); G11: Galicher et al. (2011); M11: Madhusudhan et al. (2011); M12: Marley et al. (2012); L13: Lee et al. (2013); K13: Konopacky et al. (2013); B16: Bonnefoy et al. (2016); L17: Lavie et al. (2017); G18: Greenbaum et al. (2018); W20: Wang et al. (2020a); W21: Wang et al. (2021); R21: Ruffio et al. (2021); W23: Wang et al. (2023).

^a Only T_{int} , a model parameter, is reported.

^b W20 compared using strong and weak radius priors to enforce physicality.

^c W21 used high resolution spectroscopy, and did not infer masses or radii, using masses of $7.2 \pm 0.7 M_{\text{Jup}}$ for the inner three planets and $5.8 \pm 0.5 M_{\text{Jup}}$ for HR 8799 b. A radius of $1.2 \pm 0.1 R_{\text{Jup}}$ was used for all planets.

^d R21 uncertainties were limited by the coarseness and boundaries of their model grid.

^e W23 report the temperature at 3.3 bar rather than the effective temperature.

^f The metallicity of W23 was found by averaging the retrieved C/H and O/H ratios.

18.2.1. Variability

Young brown dwarfs are known to be highly variable (Radigan et al., 2014; Vos et al., 2019, 2023). L/T transition objects display stronger photometric variability – up to 30% –, though this amplitude is rare outside of the transition regime (Radigan, 2014). However, as we view the HR 8799 system nearly pole on, it is difficult to see the effects of rotational variation, in addition to the technical challenges of observing variability with high-contrast imaging instruments. Apai et al. (2016) and Biller et al. (2021) have placed upper limits on the photometric variability of the two outermost HR 8799 planets: 10% for b and

25% for c. Wang et al. (2022) used the Subaru/CHARIS instrument to attempt to monitor H-band variability in HR 8799 c and d, placing upper limits of 10% and 30% respectively. The atmospheric turbulence, stellar contamination, and significant post-processing required to measure the innermost companion has so far prevented measurements of variability for HR 8799 e.

18.2.2. Orbital dynamics

Within the context of directly imaged exoplanets, the HR 8799 companions orbit relatively near to their host star, from a projected separation of 16 au for e out to 71 au for b. Astrometric monitoring has allowed for the precise characterisation of the orbits of the companions, demonstrated in such studies as Wang et al. (2018b); Brandt et al. (2021) and Thompson et al. (2023). Using such orbital fitting techniques, Zurlo et al. (2022) inferred dynamical masses for each of the companions. While they explore a range of models, we use the fit assuming the planets are in a near-resonant 8:4:2:1 configuration and a host star mass of $1.47 M_{\odot}$. This model produced results typical of the range of models explored; the mass estimates for each companion are: $e = 7.6 \pm 0.9 M_{\text{Jup}}$, $d = 9.2 \pm 0.1 M_{\text{Jup}}$, $c = 7.7 \pm 0.7 M_{\text{Jup}}$, and $b = 5.8 \pm 0.4 M_{\text{Jup}}$. These dynamical mass estimates, as well as those of Brandt et al. (2021), are broadly consistent with mass estimates from evolutionary models, assuming hot start conditions (Marley et al., 2012). Further astrometric analysis of the GRAVITY data will be examined in a forthcoming paper from Chavez et al. (in prep).

18.3. Spectroscopic characterisation

In addition to the multitude of photometric observing campaigns, the spectroscopic characterisation of the HR 8799 planets has traced the development of dedicated exoplanet instrumentation, from long-slit spectrographs (Janson et al., 2010) to high-contrast integral field spectrographs (IFS) (Ingraham et al., 2014; Zurlo et al., 2016) to fibre-fed high resolution spectrometers (Wang et al., 2021). These observations cover a broad swath of wavelength ranges and spectral resolving powers, leading to often conflicting photometric calibration and inferred atmospheric parameters. In particular the H-band spectra as observed with SPHERE (Zurlo et al., 2016), GPI (Greenbaum et al., 2018) and CHARIS (Wang et al., 2022) display different flux peaks and different H-band shapes. As several atmospheric parameters such as $\log g$ and the water abundance are strongly impacted by the shape of this band, they have remained challenging to measure.

Many results for individual planets have been presented in the literature.

Table 18.3: Summary of literature and derived planet properties for HR 8799 d

| Ref. | Clouds | M | log g [M_{Jup}] | T_{eff} [cgs] | R [K] | [M/H] [R_{Jup}] | C/O | log L_{bol}/L_{\odot} |
|------------------|-----------|------------------------|-------------------------------|---------------------------|------------------------|-------------------------------|---------------------------|--------------------------------|
| C11 | Thick | 5–17.5 | 3.75 – 4.5 | 1000 – 1200 | ... | 0 | ... | -4.7 ± 0.1 |
| G11 | Slab | 6 | 4.0 | 1100 | 1.25 | ... | ... | ... |
| M11 | Power law | 3 – 11 | 3.5 – 4.2 | 850 – 1000 | ... | ... | ... | ... |
| M12 | AM01 | 8 – 11 | 4.1 ± 0.1 | 1000 ± 75 | 1.33 – 1.41 | 0 | ... | -4.80 ± 0.09 |
| B16 | ER4 | ... | 4.4 – 4.5 | 1200 – 1300 | 0.9 – 1.1 | 0.5 | ... | ... |
| L17 | Mie | ... | 4.2 ± 0.2 | $^{a}1420 \pm 10$ | 0.96 ± 0.05 | ... | 0 | ... |
| G18 | Various | ... | 3.5 – 4.0 | 1100 – 1600 | 0.65 – 1.4 | ... | ... | -4.58 to -4.82 |
| W21 ^b | BT-Settl | ... | $5.1^{+0.3}_{-0.4}$ | $1558.8^{+50.9}_{-91.4}$ | ... | ... | ... | ... |
| R21 ^c | Slab | $3.7^{+0.03}_{-0.03}$ | 1200^{+-}_{-14} | ... | ... | 0 | $0.551^{+0.004}_{-0.005}$ | ... |
| Best $A \cap B$ | AM01 | $9.19^{+0.08}_{-0.07}$ | $4.18^{+0.04}_{-0.03}$ | 1177^{+21}_{-21} | $1.26^{+0.05}_{-0.06}$ | $1.2^{+0.2}_{-0.1}$ | $0.61^{+0.03}_{-0.04}$ | $-4.63^{+0.04}_{-0.04}$ |
| BMA $A \cap B$ | AM01 | $9.19^{+0.08}_{-0.07}$ | $4.18^{+0.06}_{-0.04}$ | 1179^{+31}_{-28} | $1.26^{+0.06}_{-0.08}$ | $1.2^{+0.2}_{-0.2}$ | $0.60^{+0.04}_{-0.06}$ | $-4.62^{+0.05}_{-0.04}$ |
| $A \cap B$ | AM01 | $9.20^{+0.09}_{-0.08}$ | $4.19^{+0.07}_{-0.04}$ | 1179^{+38}_{-36} | $1.25^{+0.06}_{-0.09}$ | $1.2^{+0.4}_{-0.3}$ | $0.55^{+0.12}_{-0.10}$ | $-4.61^{+0.05}_{-0.05}$ |
| Grids | Various | 8.2–9.9 | 3.5–4.5 | 1200–1300 | 0.97–1.21 | > 0.0 | 0.2 – 0.55 | -4.59 to -4.65 |

References: C11: [Currie et al. \(2011\)](#); G11: [Galicher et al. \(2011\)](#); M11: [Madhusudhan et al. \(2011\)](#); M12: [Marley et al. \(2012\)](#); B16: [Bonnefoy et al. \(2016\)](#); L17: [Lavie et al. \(2017\)](#); G18: [Greenbaum et al. \(2018\)](#); W21: [Wang et al. \(2021\)](#); R21: [Ruffio et al. \(2021\)](#).

^a Only T_{int} , a model parameter, is reported.

^b W21 used high resolution spectroscopy, and did not infer masses or radii, using masses of $7.2 \pm 0.7 M_{\text{Jup}}$ for the inner three planets and $5.8 \pm 0.5 M_{\text{Jup}}$ for HR 8799 b. A radius of $1.2 \pm 0.1 R_{\text{Jup}}$ was used for all planets.

^c R21 uncertainties were limited by the coarseness and boundaries of their model grid.

HR 8799 b was first explored in [Bowler et al. \(2010\)](#) with Keck/OSIRIS, where they identify an L5-T2 spectral type, moderate levels of cloudiness and potential impacts of disequilibrium chemistry. [Barman et al. \(2011\)](#) added additional H-band OSIRIS observations, and inferred the low temperature, low surface gravity, and low CH₄ abundance of HR 8799 b through the triangular shape of the H-band feature. They also suggest that higher metallicity grids, up to 10× solar, may be able to better fit the data and provide more plausible radii than their solar metallicity models. This data was augmented with additional wavelength coverage in the K band in [Barman et al. \(2015\)](#), where they claim simultaneous detections of H₂O, CO, and tentatively CH₄. [Oppenheimer et al. \(2013\)](#) obtained low resolution spectra for all four planets in the Y, J and H bands using the Project 1640 instrumentation suite at the Palomar Hale Telescopes, and thus providing the only additional measurement for HR 8799 b. However these spectra are very low S/N , and are significantly discrepant from subsequent measurements.

[Janson et al. \(2010\)](#) were the first to spectroscopically explore HR 8799 c,

using the VLT/NACO L-band spectrometer. While they were limited in the available S/N , they still discussed the impact of disequilibrium chemistry on the overall shape of the spectrum, finding that there was strong CO absorption beyond 4 μm . These L-band observations were later succeeded by LBT/ALES observations (Doelman et al., 2022; Liu et al., 2023), where low resolution spectra at moderate S/N were obtained for the c, d, and e planets. Konopacky et al. (2013) presented the first conclusive evidence of CO and water absorption lines in a directly imaged exoplanet through K-band observations of HR 8799 c using Keck/OSIRIS, measuring the C/O ratio to be slightly above the stellar value at $0.65^{+0.10}_{-0.05}$. The Gemini/GPI instrument provided the first spectra obtained using a coronagraphic instrument in Ingraham et al. (2014), measuring both the c and d planets in the H and K bands. This was followed up with additional post-processing using KLIP (Soummer et al., 2012; Pueyo, 2016) in Greenbaum et al. (2018), where HR 8799 e was also detected. All three of the planets were found to best match mid-to-late L-type spectra, with HR 8799 c being most consistent with an L6 dwarf. Consistent with the photometric models, they found HR 8799 c to have a temperature between 1100 K and 1300K, with a $\log g$ around 4.0. As with the photometry, the self-consistent models they used to infer the planet properties struggled to obtain radius estimates consistent with predictions of evolutionary models. Wang et al. (2021) use high resolution spectroscopy to measure the rotation of c, d, and e, finding an upper limit of 14 km/s for c, and measurements of $10.1^{+2.8}_{-2.7}$ km/s for d and $15.0^{+2.3}_{-2.6}$ km/s for e. Wang et al. (2023) combine several of these datasets and perform a retrieval analysis to constrain the composition of HR 8799 c, finding $[\text{C}/\text{H}] = 0.55^{+0.36}_{-0.39}$, $[\text{O}/\text{H}] = 0.47^{+0.31}_{-0.32}$, and $\text{C}/\text{O} = 0.67^{+0.12}_{-0.15}$. These results depended strongly on the details of the forward model used in the retrieval, and the $[\text{C}/\text{H}]$ parameter could vary from 0.55 to 0.95, while the $[\text{O}/\text{H}]$ from 0.47 to 0.80, though they all represent significant enrichment compared to the host star abundances. These results are also significantly discrepant from those of Wang et al. (2020a), who found elemental abundance ratios for HR 8799 c of $[\text{C}/\text{H}] = 0.16^{+0.12}_{-0.13}$, $[\text{O}/\text{H}] = 0.13^{+0.08}_{-0.08}$, and $\text{C}/\text{O} = 0.58^{+0.06}_{-0.06}$, though they also found that enforcing strong mass priors led to both the metallicities and C/O ratio being subsolar. Ruffio et al. (2019); Ruffio et al. (2021) and Wang et al. (2021) explore HR 8799 c using moderate and high resolution spectroscopy respectively. Both works characterise the dynamics of the planets, with Ruffio et al. (2021) measuring the radial velocities for planets b, c, and d, finding them to be -9.1 ± 0.4 km/s, -11.1 ± 0.4 km/s, and -11.6 ± 0.8 km/s respectively, placing important constraints on the allowed orbits for the planets. They also confirm the presence of water and CO, but are unable to significantly detect CH_4 , which was consistent with the results of Wang et al. (2018a).

Table 18.4: Summary of literature and derived planet properties for HR 8799 e

| Ref. | Clouds | M | log g [M_{Jup}] | T_{eff} [cgs] | R [K] | [M/H] [R_{Jup}] | C/O | log L_{bol}/L_{\odot} |
|------------------|----------|----------------------------------------|----------------------------------------|------------------------------------------|----------------------------------------|----------------------------------------|----------------------------------------|-----------------------------------------|
| B16 | ER4 | ... | 3.7 – 4.1 | 1200 – 1300 | 0.9 – 1.0 | 0.5 | ... | ... |
| L17 | Mie | ... | 3.8 ± 0.3 | ^a 1230 ± 30 | 1.2 ± 0.1 | ... | 0 | ... |
| G18 | Various | ... | 3.5 – 4.0 | 1100 – 1650 | 0.6 – 1.4 | ... | ... | –4.58 – –4.75 |
| M20 | AM01 | 4.81 ^{+8.78} _{–3.33} | 4.00 ^{+0.46} _{–0.52} | 1154 ⁺⁴⁹ _{–48} | 1.12 ^{+0.09} _{–0.09} | 0.48 ^{+0.25} _{–0.29} | 0.60 ^{+0.07} _{–0.08} | ... |
| W21 ^b | BT-Settl | ... | 3.7 ^{+0.3} _{–0.1} | 1345.6 ^{+57.0} _{–53.3} | ... | ... | ... | ... |
| Best A∩B | AM01 | 7.5 ^{+0.6} _{–0.6} | 4.20 ^{+0.06} _{–0.06} | 1172 ⁺²⁹ _{–27} | 1.14 ^{+0.05} _{–0.05} | 1.9 ^{+0.1} _{–0.1} | 0.88 ^{+0.02} _{–0.02} | –4.71 ^{+0.05} _{–0.06} |
| BMA A∩B | AM01 | 7.5 ^{+0.7} _{–0.7} | 4.20 ^{+0.06} _{–0.06} | 1161 ⁺³³ _{–34} | 1.12 ^{+0.05} _{–0.05} | 1.9 ^{+0.1} _{–0.2} | 0.88 ^{+0.02} _{–0.02} | –4.72 ^{+0.06} _{–0.06} |
| A∩B | AM01 | 7.5 ^{+0.7} _{–0.7} | 4.3 ^{+0.1} _{–0.1} | 1198 ⁺⁴¹ _{–77} | 1.05 ^{+0.15} _{–0.08} | 1.8 ^{+0.3} _{–0.4} | 0.84 ^{+0.06} _{–0.07} | –4.71 ^{+0.07} _{–0.08} |
| Grids | Various | 1.07–8.8 | 3.5–4.5 | 1100–1400 | 0.75–1.24 | > 1.0 | > 0.55 | –4.70 – –4.78 |

References: B16: [Bonnefoy et al. \(2016\)](#); L17: [Lavie et al. \(2017\)](#); G18: [Greenbaum et al. \(2018\)](#); M20: [Mollière et al. \(2020\)](#); W21: [Wang et al. \(2021\)](#);

^a Only T_{int} , a model parameter, is reported.

^b W21 used high resolution spectroscopy, and did not infer masses or radii, using masses of $7.2 \pm 0.7 M_{\text{Jup}}$ for the inner three planets and $5.8 \pm 0.5 M_{\text{Jup}}$ for HR 8799 b. A radius of $1.2 \pm 0.1 R_{\text{Jup}}$ was used for all planets.

The first reliable spectroscopic measurements of HR 8799 d and e were published by [Zurlo et al. \(2016\)](#). These were obtained using the VLT/SPHERE instrument, and were the first YJH band observations of the inner two planets, and remain the highest quality observations in this band. Together with the modelling work in [Bonnefoy et al. \(2016\)](#), they classify both planets as L6-L8 dusty dwarfs, and confirm that only thick cloud models based on the Exo-REM self-consistent modelling code provide reasonable fits to the data, finding effective temperatures of 1200 K, log g in the range of 3.0-4.5, and metallicities of 0.5 for both planets. Compared to previous modelling work of [Madhusudhan et al. \(2011\)](#) and [Barman et al. \(2011\)](#), the Exo-REM models provided better fits to the data, due to improvements in the opacity databases, cloud treatments, and the inclusion of disequilibrium chemistry. Subsequent SPHERE observations, such as those in [Wahhaj et al. \(2021\)](#) have maintained consistent spectral shapes with these earlier observations. The [Gravity Collaboration et al. \(2019\)](#) performed the first interferometric observations of an exoplanet, measuring the K-band spectrum of HR 8799 e. HR 8799 e was detected as well, and they performed atmospheric analyses on all three planets using the full spectra at 1–2.5 μm . They found that the spectrum of HR 8799 d has a substantially different shape than the other two planets, but that all three shared supersolar metallicities and effective temperatures around 1100 K.

18.4. Retrieval studies

Atmospheric retrievals (e.g. Madhusudhan & Seager, 2009; Benneke & Seager, 2012; Waldmann et al., 2015; Burningham et al., 2017; Mollière et al., 2019) are widely used to solve the inverse atmosphere problem, inferring planet properties such as the thermal structure, chemical composition, and cloud properties from spectroscopic observations. The HR 8799 planets are among the first directly imaged planets to have retrieval methods applied to their spectra. Lee et al. (2013) performed the first retrieval study of HR 8799 b, using the spectrum published of Barman et al. (2011). This pilot study explored various levels of cloudiness, particle sizes, and compositions, finding that the planet is likely cloudy, with relatively large particle sizes ($1.5\mu\text{m}$) and a supersolar metallicity. They note the long-standing degeneracies between the cloud level and the planet radius, making it difficult to distinguish between different levels of cloudiness in the models. The first systematic characterisation of all four planets was performed in Lavie et al. (2017) using the HELIOS-Retrieval package, with the key goal of constraining the composition of all four planets using the data of Zurlo et al. (2016). After fitting for molecular abundances, they infer the elemental C/H and O/H ratios for each planet, finding oxygen enrichment for b, c, and e, and carbon enrichment for b and c. They find a strongly superstellar C/O ratio for b of 0.9, a stellar value for c, but were unable to constrain the ratio for the inner two planets. While previous works on HR 8799 e were limited due to a lack of high S/N K-band data, Mollière et al. (2020) made use of the GRAVITY spectrum obtained in Gravity Collaboration et al. (2019), together with the SPHERE data of Zurlo et al. (2016) and the GPI data of Greenbaum et al. (2018). Using the pRT retrieval framework and a novel temperature profile, they inferred a highly cloudy atmosphere, implementing clouds with multiple scattering. They infer modest enrichment of $[M/H]=0.48^{+0.25}_{-0.29}$ and a C/O ratio of $0.60^{+0.07}_{-0.08}$. Finally, Wang et al. (2020a) and Wang et al. (2023) both perform pRT retrieval studies of HR 8799 c. The latter study is unique in including high resolution data in the retrieval framework, allowing precise measurements of the elemental abundance ratios, finding modest enrichment of both carbon and oxygen.

18.5. Self-consistent atmospheric modelling

Motivated by the considerable volume of observations, extensive theoretical modelling work has been performed to better understand the physics of the atmospheres of the HR 8799 planets and similar substellar objects. Brown dwarf atmospheres saw extensive 1D modelling efforts (e.g. Chabrier et al., 2000; Al-

lard et al., 2001; Burrows et al., 2006; Saumon & Marley, 2008), driven largely by the need to trace the evolution of these continuously cooling objects over time. These studies demonstrated the necessity of accounting for silicate clouds in the atmospheres of L/T dwarfs, used to explain the red colour of these objects in the near infrared. Applied specifically to the young, low-gravity companions, Madhusudhan et al. (2011) developed one of the first models specifically for the HR 8799 companions to constrain their mass and age. They identify forsterite and iron as being the important contributors to the clouds, and infer planetary ages between 10 and 150 Myr, consistent with stellar measurements. Marley et al. (2012) provides a deep review of the state of modelling of the atmospheres of the HR 8799 planets, further developing the model of Saumon & Marley (2008). They find masses and ages for the planets consistent with the stellar properties, and that the companions share approximately consistent properties with L/T dwarfs of similar effective temperatures and surface gravities. Using the cloud model of Ackerman & Marley (2001, hereafter AM01), they infer an f_{sed} parameter of 2, implying that the clouds are moderately extended throughout the atmosphere. More recent self consistent models such as petitCODE (Mollière et al., 2015), ATMO (Tremblin et al., 2015; Phillips et al., 2020; Petrus et al., 2023), Exo-REM (Charnay et al., 2018), and Sonora (Marley et al. (2021); Karalidi et al. (2021); Morley et al. in prep) have been developed specifically to understand the thermal structure and clouds of directly imaged planets. There remain degeneracies between reddening and damping of spectral features via continuum opacity sources and through reductions in the temperature gradient, hypothesised to be due to diabatic convection (Tremblin et al., 2019).

Zahnle & Marley (2014) provide an in-depth exploration of the impacts of disequilibrium chemistry on cool, self-luminous atmospheres, providing predictions for the CO, CH₄, and NH₃ abundances as a function of vertical mixing and effective temperature, identifying the key transition between CH₄ and CO dominated chemistry at around 1100 K. Moses et al. (2016) uses a disequilibrium model including photochemistry to predict the chemical composition for a range of surface gravities and effective temperatures, and provides column abundance predictions for HR 8799 b, finding that the CO abundance should dominate over CH₄, assuming a solar composition. Soni & Acharyya (2023) extend this to superstellar metallicities and vertical mixing strengths, using the constraints on the CO and CH₄ abundances from Barman et al. (2015) to infer a vertical mixing strength of $\log K_{\text{zz}} \in [7, 10]$ for the 10 \times solar metallicity case. To better understand the planet structure, Thorngren et al. (2016) derive a mass-metallicity relationship. As the mass of the object increases, the metallicity tends to decrease, consistent with predictions of core accretion formation, as heavier objects accrete and retain more H₂ and He relative to a

lower mass object. From their relationship, they predict that a $6 M_{\text{Jup}}$ planet should have a Z_{pl}/Z_* ratio of between 3 to 5 (in a 68% confidence interval).

In addition to the 1D modelling efforts, global circulation models (GCMs) of self-luminous, substellar objects, such as those of [Showman & Kaspi \(2013\)](#); [Tan & Showman \(2021a,b\)](#) have been developed. These 3D models allow for the exploration of atmospheric dynamics, longitudinal variations, and time variability. Recent observations are beginning to validate these 3D models: [Suárez et al. \(2023\)](#) finds that brown dwarfs are cloudier when viewing the equator, which is consistent with the cloudiness predictions of rapidly rotating brown dwarfs in [Tan & Showman \(2021b\)](#). Likewise, the prediction of patchy clouds in the photosphere region leading to variability ([Showman & Kaspi, 2013](#)) seems to match the observations of high variability in low-gravity atmospheres with silicate clouds ([Vos et al., 2023](#)).

18.6. Formation

With four massive planets on wide orbits, HR 8799 provides a unique system with which to test formation scenarios. In general, these fall under the categories of either gravitational instability models (e.g. [Perri & Cameron, 1974](#); [Cameron, 1978](#); [Adams et al., 1989](#); [Laughlin & Rozyczka, 1996](#); [Boss, 1997](#)), where the planets form via the direct collapse of the gas into a substellar object, or core accretion ([Pollack et al., 1996](#); [Bodenheimer et al., 2000](#)), where a dense core of heavy material grows slowly until it is massive enough to experience runaway accretion and gather an extended hydrogen-helium envelope. GI models tend to produce larger planets on wider orbits with solar compositions, while core accretion scenarios form closer-in planets on more circular orbits, with the possibility of greater metal enrichment. [Dodson-Robinson et al. \(2009\)](#) tested both of these scenarios for HR 8799, finding that while core accretion may better explain the near-orbital resonances of the system, it struggled to form planets on such wide orbits (beyond 30 au), and could not rule out the possibility of direct gravitational collapse. Similarly, [Nero & Bjorkman \(2009\)](#) find that while HR 8799 b may have formed through gravitational instability, it is unlikely that disc fragmentation could have formed the inner three companions.

In addition to constraints from the mass and location of the companions, the present-day planet composition provides insight into the formation and evolution history. The template for this was developed in [Öberg et al. \(2011\)](#), demonstrating how the C/O ratio in the gas and dust varies as a function of position in the disc, which would in turn impact the outcome of the formation process. [Eistrup et al. \(2018\)](#) extended this model to include time evolution, and [Mollière et al. \(2022\)](#) presented a framework to link the measured planet prop-

erties to the disc environment in a Bayesian framework, which allows testing the effect of various formation assumptions. However, due to the uncertainty in the atmospheric measurements, combined with the many outstanding questions in formation modelling, this link remains tenuous.

The different formation scenarios can lead to dramatically different amounts of energy retained in the planet following the formation process. So-called ‘hot-start’ models result in young planets retaining the gravitational potential energy as internal heat, to be radiated and cooled over time (Marley et al., 2007; Mordasini et al., 2017). This scenario is typically associated with formation due to gravitational instability. In cold-start scenarios, often tied to core-accretion models, this energy is radiated away by accretion shocks as the gas flows from the circumstellar disc onto the forming planet, resulting in a lower internal energy (Mordasini et al., 2012; Szulágyi & Mordasini, 2017). This is a useful, though simplified picture of planet formation. Additional complication comes from the energetics of the accretion shock during core accretion, where different radiative efficiencies can lead to different initial entropies of the forming planet (Marleau et al., 2017). These shock-resolving models find typical internal energies that are an order of magnitude higher than in typical cold-start scenarios (Marleau et al., 2019), thus lying somewhere between the hot and cold start scenarios. Over time, all of these scenarios converge to the same cooling rate, though precise mass and luminosity estimates can distinguish between the two scenarios for the first ~ 100 Myr (Baraffe et al., 2003; Saumon & Marley, 2008). Current measurements of planet masses, temperatures, and radii generally favour hot or warm start models, but can only definitely exclude the coldest initial conditions, such as the cold-start models of Marley et al. (2007). The hot-start models of Baraffe et al. (2003) led to predictions of $7 M_{\text{Jup}}$ for the inner three planets, and $5 M_{\text{Jup}}$ for HR 8799 b, which are approximately consistent with the current dynamical mass estimates of Zurlo et al. (2022). Using the hot-start model of Saumon & Marley (2008), Marley et al. (2012) finds that the radii of all of the planets should be slightly larger than $1 R_{\text{Jup}}$, and that even assuming very cold initial conditions the planet radii should never fall below $1 R_{\text{Jup}}$, though this claim did not account for significantly nonsolar composition.

Finally, HR 8799 is home to both an inner and outer debris disc, imaged with *Spitzer* (Su et al., 2009), *Herschel* (Matthews et al., 2014), JWST (Boccaletti et al., 2023), and *ALMA* in the millimeter (Wilner et al., 2018). The inner debris disc has a temperature of around 150 K and is confined to within 10 au, while the cold outer debris disc is analogous to the Kuiper belt in our own Solar System (Geiler et al., 2018), but at a much wider separation (90–300 au). The structure of the outer disc appears to be sculpted by an additional gravitational component, though it is unclear whether this is due to HR 8799 b or an

additional unseen companion (Contro et al., 2015; Faramaz et al., 2021). The inner disc has been detected in thermal emission (Su et al., 2009) and resolved using MIRI coronagraphic imaging (Boccaletti et al., 2023). Modelling efforts have placed tentative limits of $\sim 1M_{\text{Jup}}$ on the allowed mass of companions interior to HR 8799 e (Goździewski & Migaszewski, 2018).

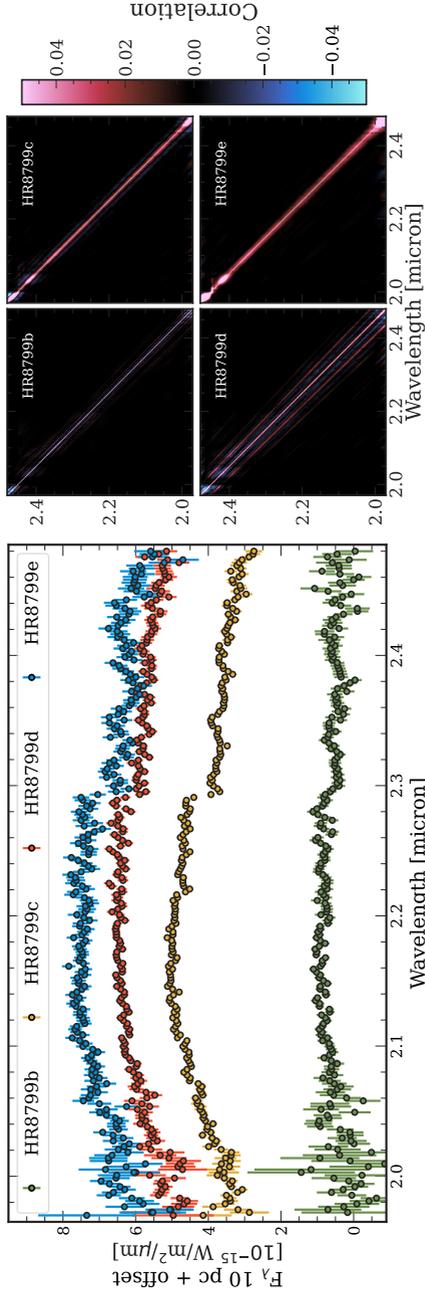


Figure 18.2: The flux-calibrated VLT/GRAVITY K-band spectra for each of the HR8799 planets, normalised to 10 pc. Each spectrum above that of HR 8799 b has an additional $1.5 \times 10^{-15} \text{ W/m}^2/\mu\text{m}$ offset. The panels on the right show the empirical correlation matrices for each of the four planets. The colour bar is scaled to highlight weak correlations in the GRAVITY data.

19. OBSERVATIONS

While the new GRAVITY spectra represent the best available K-band observations of the HR 8799 system, additional data are required to constrain planetary properties such as surface gravity and C/O ratio. We combine published datasets across a wide wavelength range from a variety of sources in order to present the most complete possible picture of this system. Archival photometric data of the companions are also included in our analysis, the details of which can be found in appendix 24.A. Also included in appendix 24.A is the stellar photometry used in fitting the BT-Nextgen model, with which the companion contrast measurements are flux calibrated. In this section we present a brief overview of the spectroscopic datasets included in the retrieval analysis, with the key observational parameters listed in Table 24.4. All of the observational data, together with the complete set of retrievals results is available on Zenodo¹.

19.1. GRAVITY data

In Figure 18.2 we present new VLTI/GRAVITY observations of HR 8799 e, together with the first interferometric observations of d, c, and b taken as part of the ExoGRAVITY project (Lacour et al., 2020), under ESO program ID 1104.C-0651. GRAVITY is a K-band spectroscopic interferometer that combines light from either the four 8 m Unit Telescopes (UTs) of the VLT, or the 1.8-m Auxiliary Telescopes (Gravity Collaboration et al., 2017). With baselines of up to 134 m, GRAVITY provides unprecedented spatial resolution, allowing for the detection of companions close to their host stars and the measurement of relative astrometry with a precision of few tens of μas . All observations of the HR 8799 system were obtained using the UTs, with the dual-field mode of GRAVITY. The medium resolution mode was used, which offers a resolution of $R \sim 500$ over a nominal wavelength range of 2.0 to 2.4 μm .

Two different strategies were used for the observations and data-reduction. Observations of HR 8799 e at all dates, except on 2 dates (11 November 2019 and 02 July 2023) were obtained using the on-axis strategy, in which a 50/50 beam-splitter is used to separate the field to between the science and fringe-tracking channels of GRAVITY. In this mode, observations with the science channel pointing at the location of the planet are interleaved with observations obtained with the fibre pointed at the central star. The on-star observations are then used to calibrate both the interferometric phase and amplitudes. This

¹<https://zenodo.org/records/10914429>

is similar to the observations reported in [Nowak et al. \(2020\)](#). The second strategy is the dual-field/off-axis strategy, in which the roof-mirror is used to split the field. The use of the roof-mirror is required to observe planets at larger separation, because the field of view of the beam-splitter does not reach these targets. In this case, the metrology zero point is calibrated using observations of the dedicated calibrator HD 25535, and the interferometric amplitude using an on-axis observation of the central star, typically done at the end of the observation sequence. This strategy is similar to the observation of Sgr A* by [GRAVITY Collaboration et al. \(2020\)](#).

The data-reduction was performed using the tools developed for the Exo-GRAVITY large program². The main steps of the reductions are as follows:

1. All data are first reduced with the GRAVITY pipeline ([Lapeyrère et al., 2014](#)), up to the ‘astroreduced’ data product, which keeps individual DITs separated.
2. For the on-axis observations, the phase reference is extracted from the on-star observations and subtracted from the on-planet observations. For the off-axis observations, this phase-reference is extracted from the observations of the binary-calibrator HD 25535. In both cases, the amplitude reference is taken using the on-star observations.
3. The stellar light (also called stellar speckle) is subtracted from the reduced data by fitting a fourth-order polynomial in wavelength multiplied by the amplitude reference. The astrometry of the planet is then extracted from the observations.
4. The contrast spectrum is then extracted using a model that also takes into account the residual starlight, and the planet astrometry previously extracted.

This procedure yields a contrast spectrum for each planet, at each observation date. The spectrum extraction, which consists entirely of linear operations on the complex coherent flux, also propagates the errors reported by the pipeline as covariance matrices. These covariance matrices allow for correlations over the wavelength channels and between the real and imaginary parts of the coherent flux. However, it should be noted that the GRAVITY pipeline does not report such covariances, and so the extraction code starts with fully diagonal covariance matrices.

For each planet, all the available spectra are then combined using a covariance-weighted combination. The final contrast spectrum $C = (c_{\lambda_1}, c_{\lambda_2}, \dots, c_{\lambda_n})^T$

²<https://gitlab.obspm.fr/mnowak/exogravity>

and its associated covariance matrix W are given by:

$$W = \left[\sum_t W_t^{-1} \right]^{-1} \quad (19.1)$$

$$C = W \cdot \left[\sum_t W_t^{-1} \cdot C_t \right], \quad (19.2)$$

where C_t and W_t represent the contrast spectrum and its associated covariance matrix on a given observing date t .

The contrast spectra are then converted to fluxes using a model of the stellar flux. For HR 8799, we used a BT-Nextgen model fit to the near infrared photometry, the details of which are more thoroughly discussed in Section 18.1 and are based on [Nasedkin et al. \(2023\)](#).

The faintest companion, HR 8799 b, was detected with a mean S/N of 3.4 per wavelength channel. HR 8799 c was observed with a mean S/N of 27.5 per channel, while HR 8799 d and HR 8799 e had a mean $S/N \approx 20$ and $S/N \approx 10$ respectively. These observations were taken over a 5 year period. With the 50 microarcsecond astrometric precision of GRAVITY, this will allow the detection of planet–planet orbital perturbances within a few years ([Covarrubias et al., 2022](#)), and we leave such analysis to future work.

19.2. Archival data

In addition to the new GRAVITY spectra, we also include archival data covering a broad wavelength range, presented in Fig. 19.1. [Mollière et al. \(2020\)](#) noted that the SPHERE ([Zurlo et al., 2016](#)) and GPI ([Greenbaum et al., 2018](#)) data are inconsistent with each other in the H-band. In order to reduce systematic variation and to account for correlations, we rereduce the data with up-to-date pipelines, and reprocesses the datasets optimally as described in [Nasedkin et al. \(2023\)](#) using KLIP ([Soummer et al., 2012](#); [Pueyo, 2016](#)). However, in order to best extract the planet signal we use KLIP in ADI+SDI mode, in comparison to ADI only mode as described in the previous study. Both the reprocessed SPHERE and GPI spectra can be found in Figs. 24.1 and 24.2. In total, our dataset includes nearly 400 data points for each planet: $N_b = 297$, $N_c = 391$, $N_d = 387$, $N_e = 388$.

19.2.1. SPHERE

Two sets of VLT/SPHERE ([Beuzit et al., 2008, 2019](#)) data are considered in this study: the first was taken during the commissioning run of the SPHERE

instrument on 12 August 2014, and was originally published in [Zurlo et al. \(2016\)](#). This is still the deepest SPHERE observation of HR 8799 covering the full YJH range, but due to the orientation of the field of view does not include HR 8799 c. This dataset was reprocessed as in [Nasedkin et al. \(2023\)](#) using KLIP in ADI+SDI mode, and we extract spectra and covariance matrices for both the e and the d companions. The second SPHERE dataset was published in [Flasseur et al. \(2020\)](#), who processed the dataset using the PACO-ASDI algorithm and were able to extract a spectrum for HR 8799 c in addition to d and e.

Additional SPHERE observations, such as presented in [Biller et al. \(2021\)](#) or [Wahhaj et al. \(2021\)](#) are available. However, in the case of [Biller et al. \(2021\)](#) the observations of the host star used for photometric calibration that were taken before and after the science observations are of insufficient S/N . While we attempted to calibrate the companion spectra using the satellite spots, this was unreliable. Finally, these observations only cover the Y and J bands, and lack the overlap with the GPI H-band spectrum, which is important for ensuring compatibility across instruments. Therefore we continue with only the datasets of [Zurlo et al. \(2016\)](#) and [Flasseur et al. \(2020\)](#).

19.2.2. GPI

Gemini/GPI ([Macintosh et al., 2014](#)) observations of HR8799, originally published in [Greenbaum et al. \(2018\)](#), were taken on 17 November 2013, 18 November 2013, and 19 September 2016 for the K1, K2 and H bands respectively. These were reduced using the standard GPI reduction pipeline (version 1.4.0), and reprocessed with KLIP using the same methods as the SPHERE data. As the new GRAVITY observations supersede the GPI data in the K-band, we only consider the GPI H-band data for this work.

19.2.3. CHARIS

Subaru/CHARIS ([Groff et al., 2015, 2017](#)) observations of HR 8799 c, d, and e were presented in [Wang et al. \(2020a\)](#) and [Wang et al. \(2022\)](#). These observations cover 1.2–2.4 μm range at low resolution. [Wang et al. \(2022\)](#) primarily examined these data for temporal variability, while here we combine the full two nights of observations in order to obtain the highest precision spectrum for each of the three planets. We take the mean spectrum for both nights, and add the errors in quadrature, dividing by the square root of the number of observations (i.e. by $\sqrt{2}$) to obtain a spectrum for each planet.

19.2.4. ALES

Doelman et al. (2022) presented L-band observations of HR 8799 c, d, and e obtained using the LBT/ALES instrument (Skemer et al., 2015). These supersede the VLT/NACO L-band observations of HR 8799 c of Janson et al. (2010), and are the first L-band spectra of HR 8799 d and e. These data also include covariance matrices, estimated using the analytic method of Greco & Brandt (2016).

19.2.5. OSIRIS

Archival Keck/OSIRIS (Larkin et al., 2006) data taken between 2009 and 2010 is included for HR8799b, as published in (Barman et al., 2011). HR 8799 b falls outside the field of view of most high-contrast-imaging IFUs, so OSIRIS is joined only by GRAVITY in measuring the near infrared spectrum of the planet. With an unbinned spectral resolution of $R \approx 4000$, and an integration time of 2700s in the H-band and 1800s in the K-band, when binned to a spectral resolution of $R \approx 60$ the OSIRIS data achieves a per-channel S/N comparable to or better than that of the GRAVITY observations. As the OSIRIS data were not taken using standard ADI observing modes, we did not attempt any rereduction or reprocessing of the archival data, apart from rescaling the flux and uncertainty by the current *Gaia* distance estimate of $41.2925 \text{ pm}0.15 \text{ pc}$ (Gaia Collaboration, 2020).

We also include the K-band spectra of Konopacky et al. (2013) (Figure 2 of that work). As published, this spectrum is not flux calibrated, and so we always fit for a flux-scaling term. For HR 8799 b and c, these K-band spectra allow us to explore the impact of different measurements on the retrieved atmospheric parameters, and to determine if our methods can reproduce the results of earlier work.

More recent observations of the HR 8799 planets using OSIRIS have been explored in Ruffio et al. (2021), but these spectra are continuum subtracted, requiring a somewhat different modelling framework than the rest of the data considered in this work. As such we do not fit these data, but we do use them as an additional check on the robustness of our fits when examining the best-fit models at higher resolution.

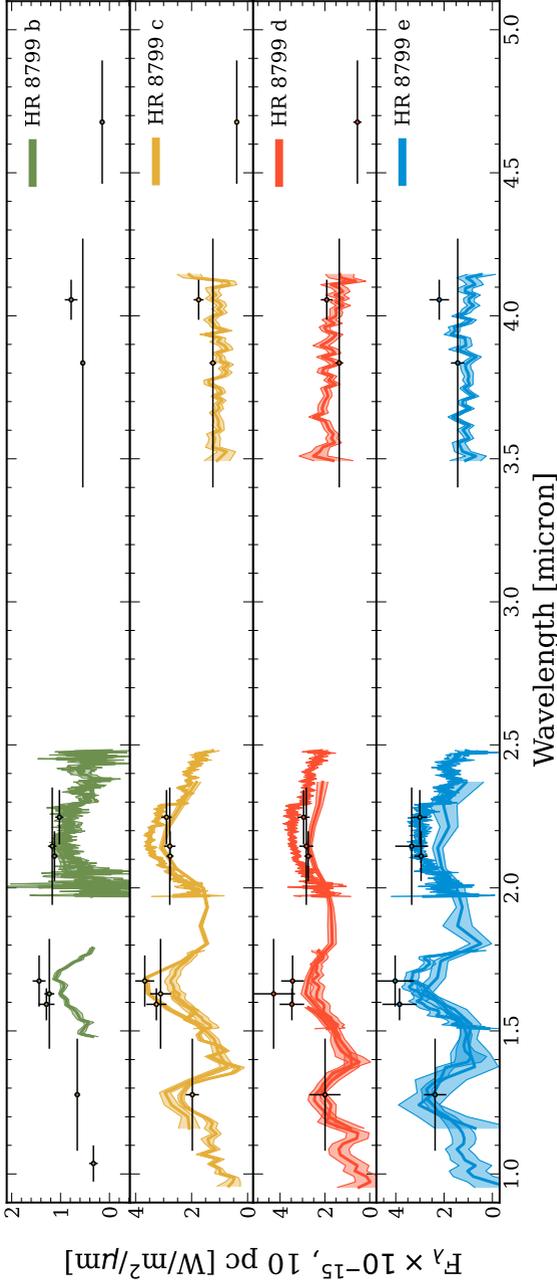


Figure 19.1: Data of the HR 8799 planets. The OSIRIS b and ALES d datasets have scaling factors of 1.0 and 1.2 applied respectively. Not shown are the MIRI photometry points from [Boccaletti et al. \(2023\)](#).

References: OSIRIS ([Barman et al., 2011](#)); SPHERE ([Zurlo et al., 2016](#)); GPI ([Greenbaum et al., 2018](#)); CHARIS ([Wang et al., 2020a](#); [Wang et al., 2022](#)); ALES ([Doelman et al., 2022](#)).

20. ATMOSPHERIC MODELLING

The forward models of our atmospheric retrieval setup were computed using pRT version 2.7 (Mollière et al., 2019), a fast, open-source radiative transfer code with which we calculate the emission spectrum of a planetary atmosphere¹. Our fiducial setup was based on that of Mollière et al. (2020), used to retrieve the atmospheric properties of HR 8799 e, though substantial improvements to the code have been made and are detailed further in Nasedkin et al. (2024). We explore a wide range of model parameterisations, summarising the parameters and prior distributions used in Table 20.1. As we consider several thermal profile parameterisations, we compare their prior distributions separately in Table 20.2.

To allow for both a data-driven and physically motivated approach, we retrieved either $\log g$ and R_{pl} with uniform priors or R_{pl} and M_{pl} , with Gaussian priors set by the dynamical mass estimates of Zurlo et al. (2022) and broad Gaussian priors centred at $1.1 R_{\text{Jup}}$, in line with estimates from evolutionary models (Marley et al., 2012).

As the computational cost of a retrieval varied greatly between the planets, it was unfeasible to run every model for every planet. As our primary point of comparison we explored the different temperature profile parameterisations for each planet, and ran both disequilibrium and free chemistry retrievals for each planet. Due to its low computational run time, we ran additional models for HR 8799 e, focusing on different cloud parameterisations.

20.1. Thermal structure

We compared a set of four temperature structures in our model comparison in order to distinguish the amount of model flexibility justified by the data and the impact of the temperature structure on other retrieved atmospheric parameters. While the thermal structure of these self-luminous objects is thought to be well-understood from 1D and 3D atmospheric models, this comparison will validate these predictions using an independent, data-driven methodology. At the same time, using the best physical understanding of the thermal structure may help constrain other parameters with greater accuracy and precision; it is necessary to compare both approaches to ensure consistent results. For all different temperature profiles we computed an effective temperature after the spectrum computation, by integrating F_{λ} over wavelength and applying the Stephan-Boltzmann law. To do this, we integrated a low resolution spectrum from 0.8 to 250 μm . The lower limit is set by the wavelength coverage of the

¹<https://petitradtrans.readthedocs.io/>

Table 20.1: Retrieval prior; temperature profile priors are included in Table 20.2. $\mathcal{N}(\mu, \sigma)$, $\mathcal{U}(\text{low}, \text{high})$.

| Parameter | Prior |
|--------------------------------------|------------------------------------------------------------|
| $\log g$ | $\mathcal{U}(2.5, 5.5)$ |
| Radius [R_{Jup}] | $\mathcal{U}(0.7, 2.0)$ |
| | $\mathcal{N}(1.1, 0.1)$ |
| Mass [M_{Jup}] | $\mathcal{N}(\mu_{\text{M, dyn}}, \sigma_{\text{M, dyn}})$ |
| $\log P_{\text{Quench}}$ [log bar] | $\mathcal{U}(-6.0, 3.0)$ |
| [M/H] | $\mathcal{U}(0.5, 2.5)$ |
| C/O | $\mathcal{U}(0.1, 1.6)$ |
| σ_{LN} | $\mathcal{U}(1.05, 3.0)$ |
| f_{sed} | $\mathcal{U}(0.0, 10.0)$ |
| $\log K_{\text{zz}}$ | $\mathcal{U}(5.0, 13.0)$ |
| $\log \text{Eq. Cloud Scaling}$ | $\mathcal{U}(-2.5, 2.5)$ |
| $\log \text{Cloud Mass Fracs.}$ | $\mathcal{U}(-6.5, 0.0)$ |
| $\log \text{Cloud } P_{\text{Base}}$ | $\mathcal{U}(-6.0, 3.0)$ |
| $\log \text{Mass Fracs.}$ | $\mathcal{U}(-7.0, 0.3)$ |

data; the optical band is unconstrained and leads to unrealistically large uncertainty on the effective temperature. The long wavelength limit is set by the wavelength coverage of the opacity databases.

20.1.1. Spline profile

To allow the data to fully determine the temperature profile of the atmosphere, we used a Piecewise Cubic Hermite Interpolating Polynomial as implemented in the `scipy.interpolate.PchipInterpolator` function. Following the prescription of [Line et al. \(2015\)](#), we penalised curvature in the

temperature profile by adding an additional term to the likelihood function,

$$\log p(\mathbf{T}) = \frac{1}{2\gamma} \sum_{i=1}^{N-1} (T_{i+1} - 2T_i + T_{i-1})^2 - \log(2\pi\gamma). \quad (20.1)$$

This is the additional penalty term, which we found by taking the sum of the discrete second derivative of the temperature profile. An additional hyperparameter, γ , was also included, with an inverse gamma distribution prior. If γ is large, (disfavoured by the prior), then the data truly demands strong curvature in the profile, while if γ is small, the data favours smoother profiles. Following [Line et al. \(2015\)](#), we set the parameters of the prior distribution on γ based on the work of [Lang & Brezger \(2004\)](#); [Rahman \(2005\)](#) and [Jullion & Lambert \(2007\)](#):

$$\Gamma^{-1}(\gamma) = \frac{\beta^\alpha}{\Gamma(\alpha)} \left(\frac{1}{\gamma}\right)^{\alpha+1} \exp\left(-\frac{\beta}{\gamma}\right), \quad (20.2)$$

for fixed α and β parameters given in Table 20.2. We repeated the retrievals and varied the number of nodes in the profile, which allowed us to use a Bayes factor comparison to determine the allowable level of complexity. This also allowed us to explore how the pressure-temperature profile can compensate for the presence of clouds by reducing the temperature gradient in the photospheric region.

20.1.2. Guillot profile

The [Guillot \(2010\)](#) (G10) profile is a simple analytical model, constructed to estimate the thermal structure of irradiated planets:

$$T_{\text{Guillot}}^4 = \frac{3T_{\text{int}}^4}{4} \left(\frac{2}{3} + \tau \right) + \frac{3T_{\text{irr}}^4}{4} \left(\frac{2}{3} + \frac{1}{\gamma\sqrt{3}} + \left(\frac{\gamma}{\sqrt{3}} - \frac{1}{\gamma\sqrt{3}} \right) e^{-\gamma\tau\sqrt{3}} \right), \quad (20.3)$$

where $T_{\text{irr}} = \sqrt{2}T_{\text{equ}}$ and $\tau = P \times \kappa_{\text{IR}}/g$. T_{equ} is the equilibrium temperature of an irradiated body, T_{int} is the intrinsic temperature of the planet, and g is the surface gravity. κ_{IR} is the mean infrared opacity, and γ is the ratio between the optical and infrared opacities. While these parameters can be physically interpreted, we treat them as nuisance parameters that control the shape of the profile, rather than self-consistently linking them to the chemical opacities. As

the HR 8799 planets are widely separated, T_{irr} is small, and thus the profile reduces to an [Eddington \(1930\)](#) profile, which corresponds to keeping only the first term on the righthand side of Equation (20.3).

20.1.3. Mollière profile

Introduced in [Mollière et al. \(2020\)](#) (M20), this is a physically motivated temperature profile, split into three distinct regions in altitude.

The middle level of the atmosphere contains the photosphere. In this region the temperature profile follows an Eddington profile, as in the first term of the Guillot profile in Equation 20.3. However, for this profile we parameterise the opacity τ as a function of pressure (P):

$$\tau = \delta P^\alpha \quad (20.4)$$

and retrieve parameters of $\log \delta$ and α , together with T_{int} , as in the [G10](#) profile.

The upper atmosphere is defined as the region above $\tau = 0.1$. Above this level, four pressure points are defined, equidistant in $\log P$. The deepest pressure point, at $\tau = 0.1$ is fixed to the temperature of the Eddington profile of the middle atmospheric region, while the remaining temperature points are freely retrieved parameters, subject to the constraint that the temperature decreases with altitude ([Kitzmann et al., 2020](#)), as inversions are not expected in self-luminous objects. The temperature profile is then interpolated from a cubic spline between the three points. Combined with the Eddington profile parameters, this results in a total of 6 parameters to describe the temperature profile.

The base of the atmosphere is defined as a moist adiabat, up to the radiative-convective boundary. This boundary occurs when the temperature gradient of the Eddington profile is Schwarzschild unstable:

$$\frac{dT}{dr} < \frac{T}{P} \frac{dP}{dr} \left(1 - \frac{1}{\gamma_{\text{ad}}} \right). \quad (20.5)$$

The moist adiabatic gradient is a function of the temperature, pressure, and chemical composition, and as such is interpolated from the disequilibrium chemistry table, discussed further in Section 20.2.2. Once the atmosphere is unstable to convection, the temperature profile is forced onto the moist adiabat.

20.1.4. Zhang profile

[Zhang et al. \(2023\)](#) (Z23) introduced a novel P-T parameterisation, incorporating the results of radiative-convective equilibrium models into the retrievals

via careful prior selection. This is accomplished by fitting for the gradient of the temperature with respect to pressure, as opposed to directly retrieving the temperature as in the spline profile. The prior locations and widths of the gradients were determined by empirically measuring the temperature gradients in self-consistent radiative-convective models, thus providing a means to enforce the physics of these models in a retrieval framework. The atmosphere between 10^3 bar and 10^{-3} bar was divided up into six layers, equidistant in log pressure. The temperature at the bottom of the atmosphere (T_{bot}) was freely retrieved. For the remaining layers, $d \log T / d \log P|_i$ were retrieved as free parameters. The temperature profile was then found by interpolating the gradient to the full pressure grid, and integrating to find the temperature at each pressure.

$$T_0 = T_{\text{Bot}} \quad (20.6)$$

$$T_{i+1} = \exp \left(\log T_j + (\log P_{i+1} - \log P_i) \left(\frac{d \log T}{d \log P} \right)_i \right) \quad (20.7)$$

The atmosphere was isothermal above 10^{-3} bar.

20.2. Chemistry

Understanding the atmospheric chemistry of the HR8799 planets is one of the key goals of this work. We compared a simplified disequilibrium chemistry model to a free chemistry retrieval with vertically constant abundances. We primarily used opacities from the ExoMol database (Tennyson & Yurchenko, 2012; Chubb et al., 2021), and include H_2O (Polyansky et al., 2018b), CO (Rothman et al., 2010a), CH_4 (Yurchenko et al., 2017), CO_2 (Yurchenko et al., 2020), NH_3 (Coles et al., 2019), HCN (Barber et al., 2014), H_2S (Azzam et al., 2016), PH_3 (Sousa-Silva et al., 2015), FeH (Wende et al., 2010), Na (Allard et al., 2019), K (Allard et al., 2016), SiO (Barton et al., 2013), TiO (McKemmish et al., 2019), and VO (McKemmish et al., 2016).

20.2.1. Free chemistry

In the free chemistry retrievals, we assumed a vertically constant mass fraction for each species, and retrieved the log mass fraction abundance ($\log X_i$ for each of H_2O , CO , CH_4 , CO_2 , HCN , H_2S , NH_3 , FeH , Na , and K), subject to the constraint that the sum of the mass fractions is less than one. Due to the lack of spectroscopic data in the Y and J bands, we did not retrieve FeH , Na or K for HR 8799 b in the free retrievals to reduce the number of free parameters. The hydrogen and helium mass fractions were calculated by using

the solar abundances (0.766 for H₂, 0.234 for He), and multiplying by one minus the sum of the retrieved mass fractions ($1 - \sum_i X_i$). The set of molecules included covers the most abundant trace species in the atmosphere, and in an atmosphere with strong vertical mixing in the photosphere the assumption of a vertically constant abundance is reasonable for H₂O, CO, and CH₄, though other species, such as FeH, have been found to have nonvertically constant abundances (Rowland et al., 2023). To measure the significance of a detection, we performed a ‘leave-one-out’ retrieval, and calculated the Bayes factor between the complete retrieval and the retrieval when excluding a single chemical species. This comparison was performed using the Zhang temperature profile and clouds condensing at their equilibrium saturation condition. While the detection significant may vary with different setups, this setup is representative of typical retrievals, and ensures consistent comparisons.

To determine the bulk properties of [M/H] and C/O for the free retrievals, we converted the retrieved mass fraction abundances to volume mixing ratios. [M/H] is defined as the ratio the planetary elemental abundances from the measurements of H₂O, CO, CH₄, CO₂, NH₃, and H₂S to the solar elemental abundances. The C/O ratio was likewise found from the number ratio of the carbon and oxygen atoms in the same set of molecular species.

20.2.2. Interpolated (dis)equilibrium

The disequilibrium model used a grid of equilibrium chemical abundances interpolated along dimensions of pressure and temperature, as well as [M/H] and C/O, which were freely retrieved parameters. The metallicity parameter scaled all of the elemental abundances, while the C/O scaled only the oxygen abundance. Initial test retrievals used a prior range of [-1.5, 1.5], but the high metallicity demanded by the data led to the choice of a prior range of [-0.5, 2.5] for the retrievals included in this work. The model of disequilibrium chemistry of the CO, CH₄, H₂O system was based on transport-induced quenching, resulting in a vertically constant abundance above a given pressure. This (log) quench pressure was one of the retrieved parameters. The equilibrium abundances used to build the grid were computed using easyChem (Mollière et al., 2017), which minimises the Gibbs free energy for the system at a given pressure, temperature, and atomic composition. We included all of the species listed in Section 20.2 as opacity sources, though 95 species are included in equilibrium chemical network used to determine the molecular abundances. For the alkali metals we used the wing profiles of Allard et al. (2016, 2019).

20.3. Clouds

We considered three cloud parameterisations in this work. The first was based of the model of [Ackerman & Marley \(2001\)](#), where cloud particles are lofted into the atmosphere through eddy diffusion (K_{zz}), and settle back down with a speed proportional to the parameter f_{sed} . We retrieved each of these parameters independently, together with σ_{LN} , which is the width of the log-normal particle size distribution.

While these parameters determine the structure of the clouds, we also determined the cloud opacity through the use of cloud optical constants for different cloud compositions, allowing us to differentiate between compositions, grain structure (amorphous or crystalline), and whether the particle shapes are spherical or based on the distribution of hollow spheres (DHS). In addition to the standard log-normal particle sized distribution used with [AM01](#) clouds, we incorporated the [Hansen \(1971\)](#) particle size distribution, which has been proposed to be a more accurate representation of the particle size than a log-normal distribution. Instead of σ_{In} , we retrieved the mean effective width b_{h} . The details of how these parameters shape the distribution, together with how they were incorporated into the [AM01](#) model are described in Appendix 24.D.

We tested a range of cloud compositions, including MgSiO_3 (both crystalline and amorphous particle shapes), Mg_2SiO_4 , Fe , Al_2O_3 , KCl , and Na_2S , as well as several combinations of these compositions. Each cloud composition had a mass fraction abundance at the cloud base that could either be scaled from an equilibrium value or freely retrieved, together with a unique f_{sed} , which allowed for different vertical extents for different cloud compositions.

Nominally, the cloud base occurs at the intersection of the cloud condensation curve and the temperature pressure profile. However, for our second parameterisation we included the cloud base pressure as a freely retrieved parameter, to determine if the clouds form where expected in the atmosphere. We followed the derivation of [AM01](#) and [Mollière et al. \(2017\)](#) to obtain the cloud abundance throughout the atmosphere. The abundance of the cloud species X_i was defined at this base pressure P_0 , and decreases with altitude to the power of f_{sed} :

$$X_i(P) = X_{i,0} \left(\frac{P}{P_0} \right)^{f_{\text{sed}}} . \quad (20.8)$$

At pressures higher than P_0 , the cloud is not condensed, and thus $X_i = 0$.

Finally, we also tested a simple grey cloud deck, where the cloud top pressure was freely retrieved, and the cloud acts as a source of opacity at the base of the photosphere.

For each of these models we could retrieve a cloud patchiness fraction, f_c . In this setup, we first calculated the usual cloudy spectrum, \vec{S}_{cd} . We then turned off the cloud opacity sources, and calculated another clear atmosphere spectrum, \vec{S}_{cl} . We then combined the two spectra, weighted by f_c :

$$\vec{S} = f_c \vec{S}_{cd} + (1 - f_c) \vec{S}_{cl}. \quad (20.9)$$

This approach divides the atmosphere into only clear and cloudy components. Other approaches, such as those of [Vos et al. \(2023\)](#) or [McCarthy et al. \(2024\)](#) have different patchiness fractions for different cloud layers, allowing for different degrees of cloudiness. Our approach reduces the number of parameters and is simple to implement in a retrieval framework, but future work should explore the patchiness of individual layers of clouds in the atmosphere.

20.4. Retrieval setup

We used the `pyMultiNest` ([Buchner et al., 2014](#)) wrapper of `MultiNest` ([Feroz & Hobson, 2008](#); [Feroz et al., 2019](#)) as the basis for our nested sampling routine, as testing showed that it runs significantly faster than the `UltraNest` sampler, which may provide more accurate estimates of the Bayesian evidence. For all retrievals we used 4000 live points to ensure dense posterior sampling and coverage of the parameter space. We set the sampling efficiency to 0.05, and used constant efficiency mode in order to reduce computation time. Comparisons to retrievals using 4000 live points and a sampling efficiency of 0.8 without constant efficiency showed that this choice does not bias the posterior estimates, and that the importance nested sampling evidence estimate is of sufficient precision for model comparison. The $(\log \mathcal{Z})$ evidence tolerance was set to 0.1, ensuring precise estimates of the evidence and ensuring convergence of the retrievals.

20.5. Retrieval ranking

Considering the number and range of models run, we must devise a system to systematically evaluate the quality of the retrieval. A true Bayesian approach would be to exclusively use the Bayes factor to evaluate the model fits. However, without well-defined prior odds for each model, we cannot quantitatively account for the prior probability of a given model. For example, based on the current understanding of these objects, the prior probability of a clear atmosphere model should be less than that of a cloudy model, but there is no clear way of assigning an objective probability. Instead we subjectively grouped some

models into a ‘low odds’ category. While these are useful for validating our assumptions about the planets and testing the inclusion of different datasets, they should not contribute significantly to a final combined parameter estimate. We further sorted the retrievals a posteriori, creating in total three tiers of retrieval results, illustrated in Fig. 20.1. We focused our overall analysis on a subset of retrievals that are both plausible models, use consistent datasets, and produce physically reasonable results.

Group A: This set of retrievals is defined as those with physically reasonable posterior values. Based on evolutionary models, it is expected that the HR 8799 planets have radii greater than $1 R_{\text{Jup}}$. Even high-mass, cold brown dwarfs over 1 Gyr in age are found to have minimum radii of $\sim 0.88 R_{\text{Jup}}$, with the minimum radius increasing with increasing metallicity (Burrows et al., 2011). Thus we exclude from group A any retrievals with a median retrieved radius less than $0.9 R_{\text{Jup}}$. We additionally enforce that the mass estimate should be broadly consistent with the dynamical mass estimates: the median retrieved mass must be greater than $1 M_{\text{Jup}}$ and less than $22 M_{\text{Jup}}$, approximately double the highest dynamical mass estimate of any of the planets (Zurlo et al., 2022). The exact positioning of these cuts does not significantly impact the results.

Group B: this set of retrievals includes those that we consider to have a high prior model probability $P(M)$; equivalently we are assigning a model prior probability $P(M) = 0$ to those models that we believe do not describe these atmospheres well. Specifically, we exclude models with a clear atmosphere, those with poorly parameterised temperature profiles used during validation studies (e.g. retrievals using only 2 nodes to define a spline temperature profile), and those using data inconsistent with our fiducial dataset. Thus while the retrievals using OSIRIS data for HR 8799 c are highly ranked by the Bayes factor, we exclude them from our analysis and from the Bayesian Model Average, as the Bayes factor is only a relevant metric when comparing like datasets. Likewise, a clear atmosphere would require a diabatic temperature profile to explain the reddening of the emission spectra, which we do not include in the retrievals and therefore the clear models are unlikely to be

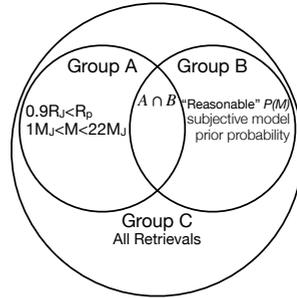


Figure 20.1: Illustration of how retrievals are grouped in this work. Group A retrievals are selected based on having physically plausible posterior distributions. Group B retrievals are models that are subjectively chosen to have a ‘reasonable’ prior probability. Group C includes the entire set of retrievals run in this work.

physically meaningful. As the Bayes factors are weighted heavily towards the best retrievals, a weighted posterior distribution effectively reduces to that of Group A.

Group C: the complete set of retrievals included in this work, regardless of prior or posterior likelihood. As the full set of retrievals includes highly unrealistic atmospheric models by design, we do not present combined posterior distributions, but only explore specific comparison retrievals used to validate different model assumptions. Ultimately we found that all of the retrievals fall into group A, universally finding reasonable estimates for the planet masses and radii.

The best set of retrievals is the intersection of groups A and B (indicated by $A \cap B$ when used to refer to a particular retrieval), which are retrievals that have physically plausible posterior values, and whose model we believe is a reasonable representation of the atmosphere. Tables 24.5 to 24.8 list the complete set of retrieval results, classifying the individual retrievals by group and sorting by the Bayes factor. We turn to [Kass & Raftery \(1995\)](#) for an interpretation of the Bayes factor in terms of frequentist statistical significance. Thus a $\Delta \log_{10} \mathcal{Z} > 1$ is considered substantial evidence, and $\Delta \log_{10} \mathcal{Z} > 2$ is considered strong evidence, equivalent to $> 5\sigma$ significance. Table 2 of [Benneke & Seager \(2013\)](#) present a similar, albeit slightly more conservative interpretation of the Bayes factor, with a similar threshold of $\log_{10} \mathcal{Z} = 2.1$ for ‘strong’ evidence in favour of one hypothesis over another, equivalent to 3.6σ significance.

20.6. Bayesian model averaging

We used the techniques of Bayesian Model Averaging (BMA) in order to combine estimates of a single parameter over a range of models, following the review of [Fragoso et al. \(2018\)](#). This has recently been applied to exoplanet spectroscopy in [Nixon et al. \(2023\)](#), demonstrating that these methods provide more realistic posterior uncertainties. They highlight that to naively use BMA, the use of multiple duplicate models must be avoided to avoid the repeated contribution of that model to the average. As we do not have any identical models in our retrieval suite, BMA remains a valid approach.

Consider Bayes theorem for the i^{th} model M_i for data \vec{D} , with parameters $\vec{\theta}_i$:

$$P(\vec{\theta}_i | \vec{D}, M_i) = \frac{P(\vec{D} | \vec{\theta}_i, M_i) P(\vec{\theta}_i | M_i)}{P(\vec{D} | M_i)}. \quad (20.10)$$

We are interested in obtain a joint posterior probability distribution $P(\vec{\theta} | \vec{D}, M)$

for the subset of parameters $\vec{\theta}$ that are shared between the set of models. From each model we require posterior probability distribution $P(\vec{\theta}_i|\vec{D}, M_i)$, the likelihood $P(\vec{D}|\vec{\theta}_i, M_i)$, and the evidence $P(\vec{D}|M_i) \equiv \mathcal{Z} = \int P(\vec{D}|\vec{\theta}_i, M_i)P(\vec{\theta}_i|M_i)d\vec{\theta}_i$. We then assume a prior probability distribution over the full set of models under consideration, and therefore each model has an associated prior probability $P(M_i)$. The choice of this prior probability should reflect the prior knowledge of the system under consideration. For example, the prior probability of a clear atmosphere model should be lower than that of a cloudy atmosphere model for the HR 8799 planets. However, quantifying this degree of certainty is highly subjective. We choose instead to use an uninformative prior distribution across N models for each planet:

$$P(M_i) = \frac{1}{N}. \quad (20.11)$$

This allows the data to determine which models should be favoured based on the evidence.

Considering all models in the range 1 to N , the posterior model probabilities given the data are

$$P(M_i|\vec{D}) = \frac{P(\vec{D}|M_i)P(M_i)}{\sum_{j=1}^N P(\vec{D}|M_j)P(M_j)}. \quad (20.12)$$

The marginal posterior distribution for a single parameter θ present in all of the models is thus

$$P(\theta|\vec{D}) = \sum_{j=1}^N P(\theta|\vec{D}, M_j)P(M_j|\vec{D}). \quad (20.13)$$

This combined posterior distribution folds in both the uncertainty from the data and prior distributions, but from the model uncertainty as well, providing a more robust estimate of the overall uncertainty on the inferred parameter.

20.7. Self-consistent forward modelling

In order to ensure that the retrieval results are robust and insensitive to the details of pRT, we fit each of the planet's spectra using several grids of 1D self-consistent models: ATMO (Phillips et al., 2020; Petrus et al., 2023), Sonora Bobcat, Cholla, and Diamondback (Marley et al., 2021; Karalidi et al., 2021), Morley et al. (in prep), Exo-REM (Charnay et al., 2018), and petitCODE (Mollière et al., 2015, 2017). These models represent the current state-of-the-art

in both cloudy and cloud-free self-consistent 1D models. The boundaries and intervals of each of these grids is presented in Table 20.3.

20.7.1. ATMO

We used an up-to-date grid of ATMO of models from [Petrus et al. \(2023\)](#), which in turn is based on prior versions from [Tremblin et al. \(2015\)](#) and [Phillips et al. \(2020\)](#). ATMO is a clear atmosphere model, based on the hypothesis that diabatic convection ([Tremblin et al., 2016, 2017](#)), not clouds, are responsible for the reddening of the near-infrared spectra of directly imaged exoplanets and brown dwarfs. This convection is instigated by disequilibrium chemical processes that reduce the temperature gradient, thus reddening the atmosphere. In ATMO, this is parameterised through an effective adiabatic index γ_{ad} , which modifies the temperature gradient. The inclusion of this parameter meant that ATMO is the only clear atmosphere grid that produced a reasonable fit to the spectra of the HR 8799 companions.

20.7.2. Exo-REM

The Exoplanet Radiative-convective Equilibrium Model (Exo-REM, [Baudino et al. \(2015\)](#); [Baudino et al. \(2017\)](#)) is a self-consistent model used to study directly imaged exoplanets and brown dwarfs ([Charnay et al., 2018](#)), but has also been extended to lower mass transiting planets ([Blain et al., 2021](#)). It implements a cloud microphysics model by combining AM01 with the timescale approach of [Rossow \(1978\)](#), which allows it to reproduce the L-T brown dwarf spectral sequence as a function of effective temperature. [Bonnefoy et al. \(2016\)](#) used this grid to explore the atmospheres of the HR 8799 companions, finding atmospheres mildly enriched in metals ($[M/H]=0.5$) and well constrained effective temperatures. However, they developed a set of custom grids that implement detailed cloud properties to model the atmospheres, which are likely more suited to the HR 8799 planets than the more general publicly available grid.

20.7.3. Sonora

The newly developed suite of Sonora models are designed to model the spectra and evolution of substellar atmospheres, covering the L-T-Y spectral sequence ([Marley et al., 2021](#)). Sonora comes in several flavours, implementing equilibrium chemistry in Sonora Bobcat ([Marley et al., 2021](#)), disequilibrium chemistry in Sonora Cholla ([Karlalidi et al., 2021](#)), and cloudy atmospheres in Sonora Diamondback ([Morley et al., 2024](#)). Like Exo-REM and ATMO, the

Sonora models are a 1D, radiative-convective equilibrium model that couples hydrostatic and thermochemical equilibrium temperature structure with a radiative transfer scheme to compute the atmospheric emission spectrum.

Sonora Bobcat and Cholla did not fit the HR 8799 spectra at all, validating the necessity of cloudy (or similar) atmospheric models. Thus we continued only with the cloudy Sonora Diamondback models in order to interpret the HR 8799 atmospheres. While similar to Exo-REM in implementing clouds, Diamondback currently fixes the C/O ratio to a solar value of 0.458 (Lodders et al., 2009), preventing the measurement of this parameter, and potentially leading to biases in the remaining parameters.

20.7.4. `petitCODE`

`petitCODE` is a radiative-convective and chemical equilibrium code used to compute the structures and spectra of exoplanet atmospheres (Mollière et al., 2015, 2017). We used the `cool-cloudy` and `hot-cloudy` grids computed for Stolker et al. (2020), spanning temperatures from 500-850 (`cool-cloudy`) and 1000-2000 K (`hot-cloudy`). The code setups are based on the work presented in Samland et al. (2017); Linder et al. (2019). Both grids implement the cloud model described in Ackerman & Marley (2001). While the cool grid only assumes Na_2S and KCl clouds, and the hot grid adds Mg_2SiO_4 and Fe clouds.

20.7.5. Grid Fits

We performed Bayesian fits using `species` to interpolate the grids (Stolker et al., 2020), and `MultiNest` to sample the parameter space. 400 live points were used for these fits, with uniform priors on all parameters covering the grid ranges as described in Table 20.3, and an additional Gaussian prior on the planet mass. We fit for covariance width and strength for all IFS datasets following the method of Wang et al. (2020b), as the empirical covariance matrices cannot be incorporated in `species`, other than for GRAVITY data. Fitting for the covariance parameters was universally favoured by the Bayes factor, and thus we only present the full fits. Consistent with the expectations of Greco & Brandt (2016) and Nasedkin et al. (2023), including these parameters also tended to broaden the posterior distributions, though posterior widths remain far narrower than the variation between the models.

In addition to the Bayesian fits, we performed a simple χ^2 minimisation over each grid to avoid potential issues with interpolating the spectra along the different parameter axes. Using this framework, we identify the single best-fit spectra, as presented in Table 21.2.

Table 20.2: Priors for temperature profiles.

| Spline | | Guillot (G10) | | Mollière (M20) | | Zhang (Z23) | |
|----------|-----------------------------------------------|---------------------------|--------------------------------------------|------------------|--------------------------------------------|-------------------------|----------------------------------------------|
| Name | Prior | Name | Prior | Name | Prior | Name | Prior |
| T_0 | $\mathcal{U}(0 \text{ K}, 300\text{K})$ | T_{int} | $\mathcal{U}(300 \text{ K}, 2500\text{K})$ | T_{int} | $\mathcal{U}(300 \text{ K}, 2000\text{K})$ | T_{bot} | $\mathcal{U}(2000 \text{ K}, 12000\text{K})$ |
| T_i | $\mathcal{U}(300 \text{ K}, 11900 \text{ K})$ | T_{equ} | $\mathcal{U}(10 \text{ K}, 100\text{K})$ | T_0 | $\mathcal{U}(0, T_{\text{Edd}})$ | $d \log T / d \log P_0$ | $\mathcal{N}(0.25, 0.025)$ |
| γ | $\Gamma^{-1}(1, 5 \times 10^{-5})$ | γ | $\mathcal{N}(0, 2)$ | T_1 | $\mathcal{U}(0, T_0)$ | $d \log T / d \log P_1$ | $\mathcal{N}(0.25, 0.045)$ |
| ... | ... | $\log \kappa_{\text{IR}}$ | $\mathcal{U}(-4, 1)$ | T_2 | $\mathcal{U}(0, T_1)$ | $d \log T / d \log P_2$ | $\mathcal{N}(0.26, 0.05)$ |
| ... | ... | ... | ... | α | $\mathcal{U}(1, 2)$ | $d \log T / d \log P_3$ | $\mathcal{N}(0.2, 0.05)$ |
| ... | ... | ... | ... | $\log \delta$ | $\mathcal{U}(0, 1)$ | $d \log T / d \log P_4$ | $\mathcal{N}(0.12, 0.045)$ |
| ... | ... | ... | ... | ... | ... | $d \log T / d \log P_5$ | $\mathcal{N}(0.07, 0.07)$ |

Notes
 $\mathcal{N}(\mu, \sigma)$: Gaussian prior.

 $\mathcal{U}(\text{low}, \text{high})$: Uniform prior.

 $\Gamma^{-1}(\alpha, \beta)$: Inverse Gamma function prior.

Table 20.3: Self-consistent grid boundaries and step sizes.

| Parameter | ATMO | Diamondback | Exo-REM | petitCODE, Cool | petitCODE, Hot |
|----------------------|--------------|---------------|-------------|-----------------|--------------------|
| T_{eff} [K] | [800, 3000] | [900, 2400] | [400, 2000] | [500, 850] | [1000, 1800] |
| | 50 | 100 | 50 | 50 | 100 |
| $\log g$ [dex] | [2.5, 5.5] | [3.5, 5.5] | [3.0, 5.0] | [3.0, 5.0] | [3.5, 5.5] |
| | 0.1 | 0.5 | 0.5 | 0.5 | 0.5 |
| [M/H] | [-0.6, 0.6] | [-0.5, 0.5] | [-0.5, 1.0] | [0.0, 1.4] | [-0.3, 0.3] |
| | 0.3 | 0.5 | 0.5 | 0.2 | 0.3 |
| C/O | [0.3, 0.7] | 0.458 | [0.1, 0.8] | 0.55 | [0.55, 0.75, 0.90] |
| | 0.24 | ... | 0.05 | ... | ... |
| f_{sed} | ... | [1.2, 3.4, 8] | ... | [0.5, 2.0] | [1.5, 4.5] |
| | ... | ... | ... | 0.5 | 1.5 |
| γ_{ad} | [1.01, 1.05] | ... | ... | ... | ... |
| | 0.02 | ... | ... | ... | ... |

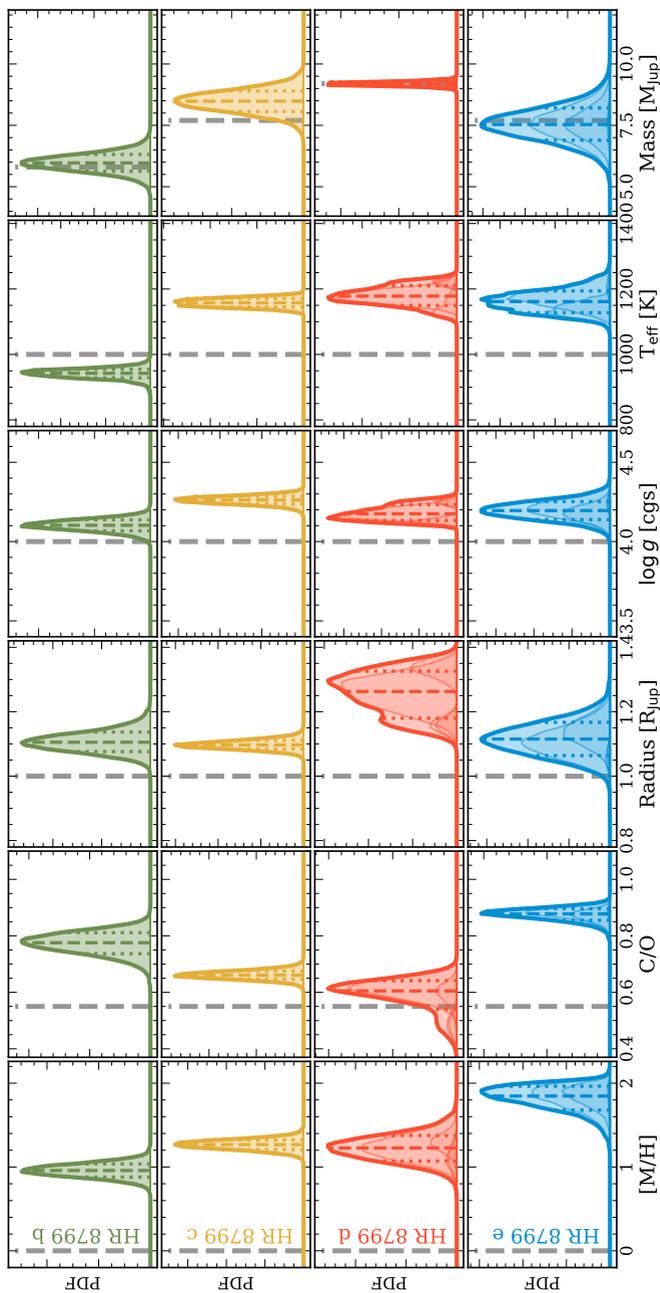


Figure 20.2: Bayesian-averaged posterior parameter distributions for each of the HR 8799 planets based on the group $A \cap B$ set of retrievals. In faint lines beneath the total posterior distribution are the individual contributions from different retrievals. The coloured dashed and dotted lines indicate the median and $\pm 34.1\%$ confidence regions respectively. The vertical grey lines indicate typical parameter values (e.g. solar metallicity and C/O) and serve as a visual reference for each parameter. For the planet mass, they indicate the dynamical mass estimate from [Zurlo et al. \(2022\)](#).

21. RESULTS

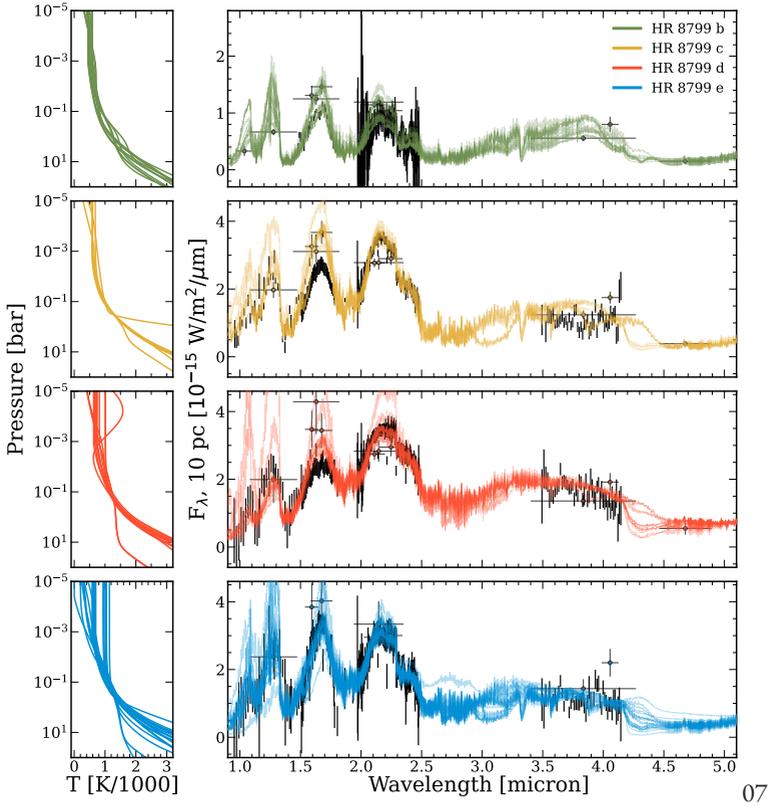


Figure 21.1: Best-fit temperature profiles and spectra for the group $A \cap B$ retrievals. From top to bottom are HR 8799 b, c, d, and e.

Based on both the atmospheric retrievals and the self consistent grid fits, the HR 8799 planets share enriched atmospheres, with stellar-to-superstellar C/O ratios. The properties of each atmosphere for subsets of the ensemble of retrievals are summarised in tables 18.1–18.4 for planets b–e, respectively. These tables also contain ranges of plausible parameter values for each planet based on the aggregate of the self-consistent models, while Fig. 20.2 shows the distributions for a subset of the key parameters. These estimates are synthesised from the results of the Bayesian fits and the χ^2 minimisation, rejecting solutions

Table 21.1: Grid-retrieval results

| Planet | Model | $\Delta \log_{10} Z$ | T_{eff} [K] | $\log g$ [cgs] | [M/H] | f_{set} | C/O | Radius [R_{Jup}] | $\log L/L_{\odot}$ | Mass [M_{Jup}] |
|--------|-------------------------|----------------------|-------------------------|------------------------------|------------------------|------------------------|---------------------------|--------------------------------|----------------------------|------------------------------|
| b | ATMO | -63 | 1020^{+9}_{-7} | $4.42^{+0.04}_{-0.04}$ | > 0.6 | ... | < 0.3 | $0.73^{+0.02}_{-0.02}$ | $-5.278^{+0.008}_{-0.009}$ | $5.6^{+0.4}_{-0.4}$ |
| | Diamondback | -98 | 959^{+10}_{-8} | $4.18^{+0.03}_{-0.03}$ | > 0.5 | < 1.0 | 0.488 | $0.95^{+0.02}_{-0.02}$ | $-5.155^{+0.007}_{-0.007}$ | $5.6^{+0.4}_{-0.4}$ |
| | Exo-Rem | 0 | 862^{+21}_{-6} | $4.0^{+0.02}_{-0.03}$ | $0.69^{+0.11}_{-0.03}$ | ... | $0.50^{+0.10}_{-0.02}$ | $1.19^{+0.03}_{-0.03}$ | $-5.14^{+0.02}_{-0.02}$ | $5.5^{+0.4}_{-0.4}$ |
| | petitCODIE ^a | -105 | > 850 | $3.48^{+0.01}_{-0.02}$ | > 1.4 | $2.53^{+0.01}_{-0.02}$ | 0.55 | $2.00^{+0.02}_{-0.04}$ | $-4.73^{+0.01}_{-0.02}$ | $4.8^{+0.8}_{-0.2}$ |
| c | ATMO | 0 | 1195^{+5}_{-6} | $4.34^{+0.03}_{-0.03}$ | > 0.6 | ... | $0.425^{+0.012}_{-0.014}$ | $1.0^{+0.01}_{-0.01}$ | $-4.723^{+0.003}_{-0.003}$ | $9.1^{+0.5}_{-0.5}$ |
| | Diamondback | -245 | 1237^{+2}_{-2} | < 3.5 | > 0.5 | $2.52^{+0.02}_{-0.02}$ | 0.488 | $0.99^{+0.01}_{-0.01}$ | $-4.683^{+0.002}_{-0.002}$ | $1.25^{+0.01}_{-0.01}$ |
| | Exo-Rem | -30 | 1065^{+1}_{-1} | $4.15^{+0.02}_{-0.02}$ | > 1.0 | ... | $0.50^{+0.003}_{-0.002}$ | $1.314^{+0.005}_{-0.005}$ | $-4.694^{+0.002}_{-0.002}$ | $9.6^{+0.5}_{-0.5}$ |
| | petitCODIE | -622 | $1400^{+0.2}_{-0.5}$ | $4.00^{+0.01}_{-0.01}$ | > 0.3 | > 4.5 | $0.55^{+0.00}_{-0.00}$ | $0.80^{+0.00}_{-0.00}$ | $-4.652^{+0.002}_{-0.002}$ | $2.57^{+0.02}_{-0.02}$ |
| d | ATMO | -38 | 1300^{+2}_{-2} | $4.39^{+0.01}_{-0.01}$ | $0.07^{+0.02}_{-0.02}$ | ... | < 0.3 | $0.926^{+0.004}_{-0.004}$ | $-4.652^{+0.003}_{-0.003}$ | $8.4^{+0.2}_{-0.2}$ |
| | Diamondback | -76 | 1234^{+8}_{-7} | $4.27^{+0.02}_{-0.02}$ | > 0.5 | $1.95^{+0.04}_{-0.06}$ | 0.488 | $1.05^{+0.01}_{-0.01}$ | $-4.632^{+0.002}_{-0.002}$ | $8.4^{+0.2}_{-0.2}$ |
| | Exo-Rem | 0 | 1155^{+5}_{-4} | $4.23^{+0.01}_{-0.01}$ | $0.78^{+0.03}_{-0.03}$ | ... | $0.200^{+0.002}_{-0.002}$ | $1.168^{+0.006}_{-0.005}$ | $-4.656^{+0.007}_{-0.005}$ | $9.3^{+0.6}_{-0.2}$ |
| | petitCODIE | -185 | 1400^{+1}_{-1} | $4.47^{+0.01}_{-0.01}$ | > 0.3 | $2.99^{+0.06}_{-0.07}$ | $0.55^{+0.00}_{-0.00}$ | $0.859^{+0.003}_{-0.003}$ | $-4.587^{+0.003}_{-0.002}$ | $8.9^{+0.2}_{-0.2}$ |
| e | ATMO | 0 | 1251^{+22}_{-20} | $4.38^{+0.06}_{-0.07}$ | > 0.6 | ... | $0.52^{+0.03}_{-0.03}$ | $0.87^{+0.03}_{-0.03}$ | $-4.778^{+0.006}_{-0.006}$ | $7.2^{+0.8}_{-0.8}$ |
| | Diamondback | -16 | 1244^{+6}_{-6} | < 3.5 | > 0.5 | $2.31^{+0.05}_{-0.05}$ | 0.488 | $0.92^{+0.01}_{-0.01}$ | $-4.738^{+0.005}_{-0.005}$ | $1.12^{+0.07}_{-0.05}$ |
| | Exo-Rem | -14 | 1064^{+3}_{-3} | $4.13^{+0.04}_{-0.04}$ | > 1.0 | ... | $0.50^{+0.01}_{-0.01}$ | $1.24^{+0.01}_{-0.01}$ | $-4.747^{+0.006}_{-0.005}$ | $8.0^{+0.8}_{-0.8}$ |
| | petitCODIE | -56 | 1399^{+2}_{-4} | $4.00^{+0.02}_{-0.03}$ | > 0.3 | $4.48^{+0.02}_{-0.03}$ | $0.55^{+0.00}_{-0.00}$ | $0.755^{+0.005}_{-0.004}$ | $-4.703^{+0.004}_{-0.004}$ | $2.3^{+0.1}_{-0.1}$ |

Notes

All values presented are the median values from the fits, with uncertainties given as the $\pm 34.1\%$ percentiles. Values without uncertainties were fixed during the fit.

a the `petitcodie-cool-cloudy` grid was used to fit HR 8799 b.

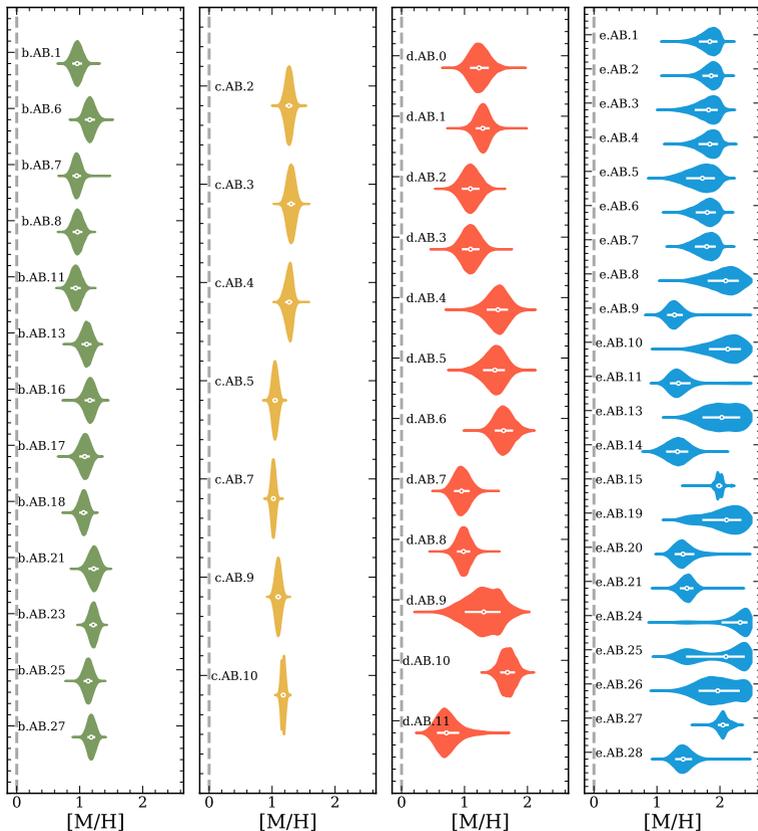


Figure 21.2: $[M/H]$ posterior distributions for all retrievals in group $A \cap B$. From left to right are the distributions for b, c, d, and e. The vertical line indicates solar metallicity. Model keys are as in tables 24.5–24.8, and are sorted by the Bayes factor from top to bottom.

with unphysical masses and radii. The full results of the grid fits are found in Table 21.1 for the Bayesian fits and Table 21.2 for the single-best χ^2 fits. In these tables, an index is assigned to each retrieval, with the format *planet.group.index*, which serves as the retrieval identifier throughout the text.

Nearly 100 retrievals were performed for this analysis across the four companions in order to derive robust constraints on primary planetary properties. The aggregate results of the Bayesian model average of group $A \cap B$ retrievals are

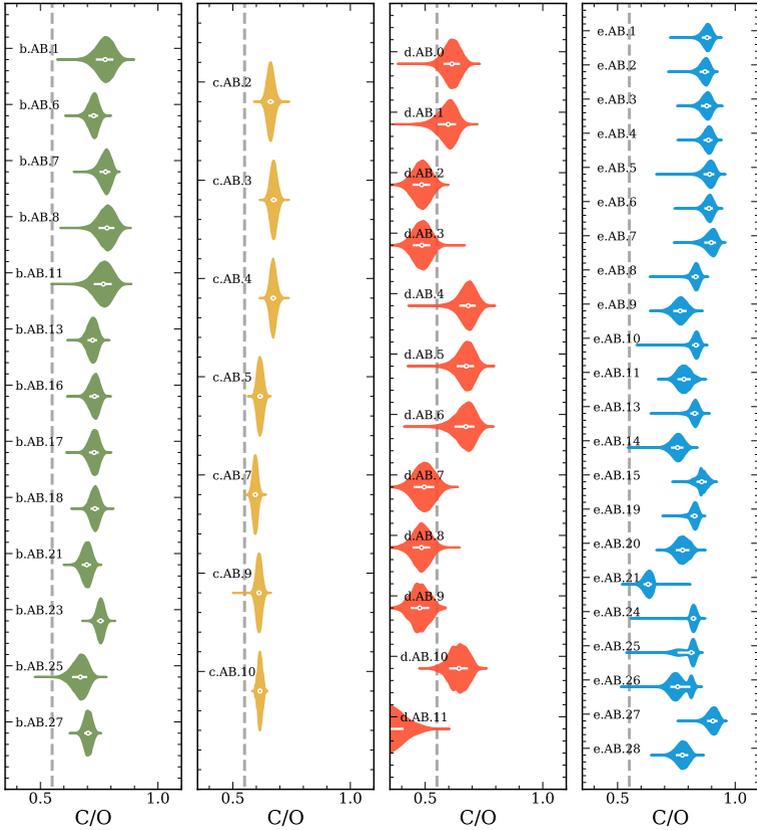


Figure 21.3: C/O posterior distributions for all retrievals in group $A \cap B$. From left to right are the distributions for b, c, d, and e. The vertical line indicates stellar C/O . Model keys are as in tables 24.5–24.8, and are sorted by the Bayes factor.

presented in Fig. 20.2, with the best fit models for the same sample of retrievals compared to the data in Fig. 21.1. Overall we find consistently good fits for all four planets, with best-fit $\chi^2/\nu < 2$ for each planet. The self-consistent grid-fits incorporate additional physical processes and require fewer free parameters than the retrievals, making them less flexible. The additional processes, such as radiative-convective equilibrium, chemistry, and cloud physics act as narrow priors on the bulk atmospheric parameters. While the lack of flexibility leads to

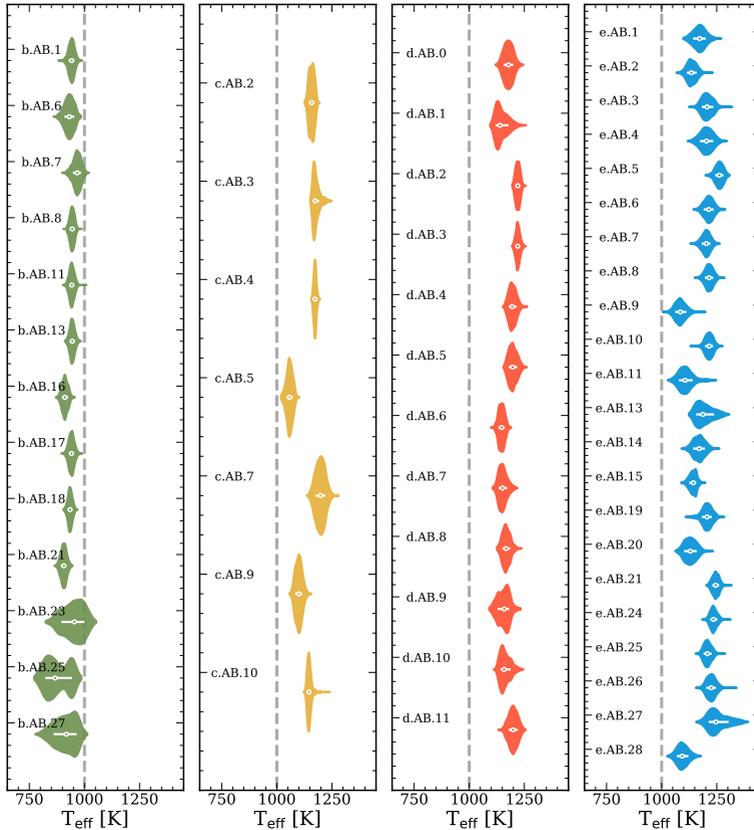


Figure 21.4: T_{eff} posterior distributions for all retrievals in groups $A \cap B$. From left to right are the distributions for b, c, d, and e. The vertical line indicates 1000 K. Model keys are as in tables 24.5–24.8, and are sorted by the Bayes factor.

worse fits when compared to retrievals, the fits of multiple self-consistent grids still result in mutually consistent results. Notably, `petitCODE` gives consistent parameter estimates even though it is strongly disfavoured by the goodness-of-fit metrics.

Both the pRT retrievals and the self-consistent grid fits show that all of the atmospheres are strongly enriched in metals, finding median $[M/H] \geq 1.0$ for each planet. This is driven by the carbon and oxygen abundances as measured through H_2O and CO , and particularly through the CO absorption feature at

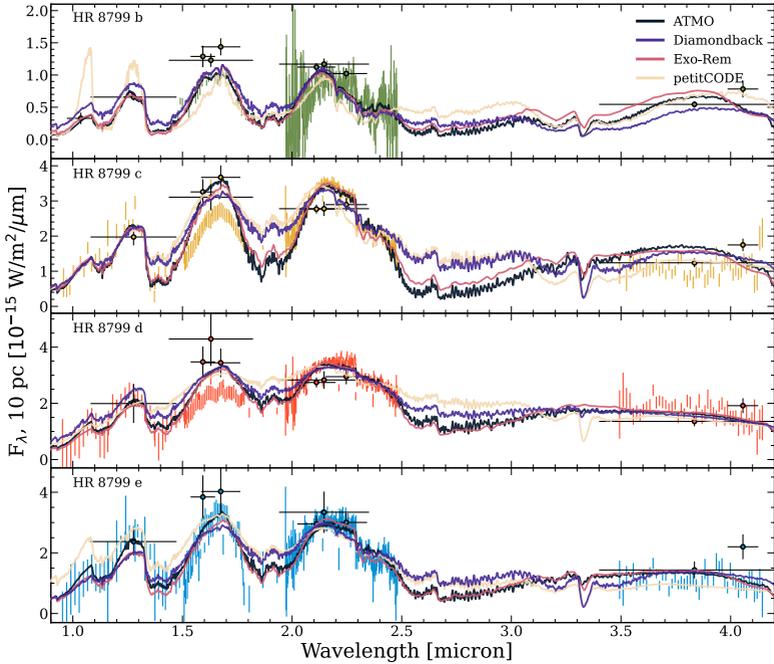


Figure 21.5: Grid fits from Exo-REM, Sonora Diamondback, ATMO and petitCODE. ALES data for HR 8799 d is scaled by the overall best-fit scaling parameter (Exo-REM, 1.18). Fits are the single best-fit χ^2 model from the grid.

2.3 μm . The C/O ratios decrease with decreasing separation from 0.78 ± 0.04 in HR 8799 b to 0.60 ± 0.05 for HR 8799 d, with HR 8799 e breaking the trend with carbon rich C/O ratio of 0.87 ± 0.02 . In addition to water and CO, HCN is confidently detected in HR 8799 c and e, while CH_4 is detected in HR 8799 c.

We find that cloudy atmospheres are universally favoured over clear atmospheres. Extended silicate clouds, together with an iron cloud deck are the preferred models for the three inner planets, with the cooler HR 8799 b showing evidence for Na_2S clouds.

The retrievals found masses, radii, and surface gravities consistent with evolutionary models. By construction, these form Group A of our retrievals. In general, retrieving the planet radius is challenging: the radius of directly imaged exoplanets is commonly underestimated by both retrievals and grid fits (e.g. Bonnefoy et al., 2016). These unphysical solutions are found in several of

Table 21.2: Grid-fit χ^2 results

| Planet | Model | χ^2 | T_{eff} [K] | $\log g$ [cgs] | [M/H] | f_{sed} | C/O | γ_{ad} | Radius [R_{Jup}] |
|----------------|------------------------|----------|-------------------------|-------------------|-------|------------------|-------|----------------------|--------------------------------|
| b | ATMO | 744 | 1000 | 4.5 | 0.6 | ... | 0.3 | 1.05 | 1.00 |
| | Diamondback | 879 | 1100 | 3.5 | 0.5 | 2.0 | 0.458 | ... | 1.07 |
| | Exo-REM | 510 | 850 | 3.5 | 0.5 | ... | 0.55 | ... | 1.05 |
| | petitCODE ^a | 968 | 850 | 3.5 | 1.4 | 2.0 | ... | ... | 0.79 |
| c | ATMO | 2162 | 1200 | 4.5 | 0.6 | ... | 0.3 | 1.01 | 0.97 |
| | Diamondback | 3900 | 1200 | 3.5 | 0.5 | 2.0 | 0.458 | ... | 1.08 |
| | Exo-REM | 1868 | 1100 | 3.5 | 1.0 | ... | 0.8 | ... | 1.22 |
| | petitCODE | 7374 | 1200 | 3.5 | 0.3 | 3.0 | 0.55 | ... | 1.09 |
| d ^b | ATMO | 943 | 1300 | 3.0 | 0.6 | ... | 0.55 | 1.01 | 0.94 |
| | Diamondback | 1336 | 1200 | 3.5 | 0.5 | 1.0 | 0.458 | ... | 1.12 |
| | Exo-REM | 936 | 1200 | 3.0 | 1.0 | ... | 0.55 | ... | 1.12 |
| | petitCODE | 2254 | 1200 | 4.0 | 0.0 | 1.5 | 0.55 | ... | 1.21 |
| e | ATMO | 669 | 1300 | 4.0 | 0.6 | ... | 0.55 | 1.03 | 0.81 |
| | Diamondback | 819 | 1200 | 3.5 | 0.5 | 2.0 | 0.458 | ... | 1.02 |
| | Exo-REM | 623 | 1100 | 3.5 | 1.0 | ... | 0.8 | ... | 1.15 |
| | petitCODE | 1136 | 1400 | 4.0 | 0.3 | 4.5 | 0.55 | ... | 0.74 |

Notes

All values presented are the single best-fit value according to the χ^2 . All of the models share a similar number of parameters, which are listed in Table 20.3. Additional parameters for data scaling and covariance fitting are shared between all models.

^a for HR 8799 b, the `petitcode-cool-cloudy` grid was used rather than the `petitcode-hot-cloudy` grid.

^b the fits for HR 8799 d included a scaling parameter for the ALES data.

the grid fits, while nearly every retrieval finds plausible values of the mass and radius. The small radius is often compensated for by adjusting the temperature, metallicity, or cloud properties. Nevertheless, in all cases the most favoured retrieval produced masses consistent with dynamical mass measurements and radii consistent with evolutionary models. In the group $A \cap B$ retrievals that used dynamical mass priors, the retrieved mass estimate is within 1σ of the

dynamical mass for b, d, and e, with the posterior width consistent with the prior width which. For HR 8799 c the retrieved mass is moderately higher than the dynamical mass estimate at $8.5 \pm 0.5 M_{\text{Jup}}$ compared to the dynamical mass of $7.7 \pm 0.7 M_{\text{Jup}}$. However, if the dynamical mass is not used as a constraint and $\log g$ is freely retrieved, the resulting mass estimate is found to be much larger than the dynamical mass estimate, highlighting the importance of including additional constraints in a retrieval framework. The estimate of the effective temperature of each planet is consistent with previous findings, with e, c, and d sharing temperatures around 1100 K, and b being cooler at around 950 K. While we do not perform a comparison to spectral templates due to the inferred high metallicity, low surface gravity, and potential complications from the viewing angle of the planets, we find that the spectral types found in [Bonnetfoy et al. \(2016\)](#) remain a good description of the four planets. HR 8799 d and e are similar to late-L type brown dwarfs, consistent with their inferred effective temperature. HR 8799 c is more likely fit by an early-T spectral type as it is a few kelvin cooler than e and d and is beginning to show spectral features from CH₄. At 950 K HR 8799 b is solidly in the T dwarf regime.

In this section, we present the measured properties for each of the four planets individually. Following that, in Section 21.5 we present a detailed discussion of the results and challenges of comparing different thermal structures, chemistry and cloud parameterisations, and data inclusion.

21.1. HR 8799 b

For HR 8799 b we included the GRAVITY and OSIRIS spectra, together with the full set of photometry, allowing the OSIRIS data to float as the published data are not flux calibrated. For the grid fits we included additional parameters to describe the covariance of the OSIRIS data.

HR 8799 b is the coldest and lowest mass planet in the HR 8799 system. The best estimate of these parameters via Bayesian averaging of group $A \cap B$ retrievals finds an effective temperature of 942_{-16}^{+12} K and a mass of $6.0_{-0.3}^{+0.4} M_{\text{J}}$, driven by the use of the dynamical mass estimate of [Zurlo et al. \(2022\)](#) as a prior. The radius is slightly inflated compared to Jupiter, with ($R_{\text{pl}} = 1.10_{-0.03}^{+0.03}$); combining the mass and radius estimates leads to $\log g = 4.10_{-0.04}^{+0.03}$. This is consistent with the Bayesian grid fits, though the single-best χ^2 fits found both lower ($\log g = 3.5$) and higher ($\log g=4.5$) solutions. The Bayesian averaged results (for group $A \cap B$) are driven by a single retrieval, with a $\Delta \log_{10} \mathcal{Z}$ of 4 relative to the next best retrieval. This single best retrieval uses the [Z23](#) temperature profile and free chemistry, finding a metallicity of $0.96_{-0.08}^{+0.08}$ and a C/O of $0.78_{-0.04}^{+0.03}$.

The grid fits find temperatures between 850 K–1100 K, with the single best fit, found using Exo-REM, finding $T_{\text{eff}} = 850$ K. Using the Bayesian fit, Exo-REM is again the most favoured model by the Bayes factor, and finds $T_{\text{eff}} = 862_{-6}^{+21}$ K and a radius of $1.19_{-0.03}^{+0.03} R_{\text{Jup}}$, consistent with expectations from evolutionary models (e.g. Marley et al., 2012). The ATMO and Sonora Diamondback models find a somewhat higher temperature and a smaller radius, but are disfavoured by the Bayes factor. petitCODE is divided into cool-cloudy and hot-cloudy grids, with the cool grid extending up to 850 K, and the hot grid beginning at 1000 K. As the effective temperature of HR 8799 b likely falls between these grids, it poorly fit the data, although the remaining parameters of $\log g$ and f_{sed} are compatible with the other self-consistent models.

Figure 21.2 shows the variation of the metallicity across the different models considered. We see that strong enrichment solution is almost always found, particularly by models preferred by the Bayes factor. The degree of enrichment does not systematically vary between disequilibrium and free chemistry retrievals. All of the self-consistent models favour high metallicity solutions, reaching the upper bounds of the grid in all cases.

Similarly to the metallicity, the C/O ratio is shown for the different retrievals in Fig. 21.3 and T_{eff} in 21.4. The C/O ratio is generally constrained to between 0.6 and 0.8. However, for free chemistry retrievals, the inferred C/O ratio only indicates the gas-phase composition. Additional oxygen is sequestered in the silicate clouds: accounting for this sequestration would result in a lower C/O ratio. Among the grid fits, shown in Fig. 21.5, the C/O ratio shows more variation, ranging from the lower bound of the ATMO grid at 0.3 to 0.55 from Exo-REM. The best fit models from Exo-REM are consistent with the stellar value of 0.54, and additional data covering the near infrared water features is likely necessary to improve these constraints. From Fig. 21.6 we find that both the free retrieval and disequilibrium retrievals display similar trends in the retrieved chemical abundances, finding that nearly 10% of the atmosphere is CO by mass, while water has a lower abundance of around 1% by mass. The best-fit free retrieval finds systematically lower abundances for both of these species compared to the best-fit disequilibrium retrieval. H₂S and CH₄ are found to be the next most abundant species in both the disequilibrium and free chemistry retrievals. However, even though their abundances are constrained by the posterior distribution, there is no evidence for their detection when comparing between the full free chemistry retrieval (b.AB.1) and retrievals without these species (b.A.0 and b.A.2). The b.AB.1 retrieval uses the Z23 temperature profile, free chemistry, and clouds condensing at their equilibrium position. This same setup is used for b.A.0 and b.A.2, apart from the exclusion of CH₄ and H₂S respectively. While the inferred CH₄ abundance of $\log X_{\text{CH}_4} = -5.0 \pm 0.4$ is

compatible with [Barman et al. \(2015\)](#), there is no statistical evidence to support the detection.

In order to compare the cloud composition for HR 8799 b, we considered a set of retrievals using the same temperature structure ([M20](#)) and cloud parameterisation, and vary the cloud optical properties and condensation curve. We found that Na_2S clouds are preferred over silicate, iron, and KCl clouds ($\Delta \log_{10} \mathcal{Z} \geq 2$). Patchy clouds were also explored, but no evidence was found for patchiness, regardless of cloud composition. At the temperature of HR 8799 b silicate clouds are expected to occur below the photosphere, with Na_2S or KCl clouds becoming the primary aerosol opacity source. This result is likely driven by the condensation temperature of Na_2S rather than the optical properties; unlike silicate clouds which have strong absorption features in the mid-infrared, crystalline Na_2S is featureless out to $15 \mu\text{m}$, apart from a characteristic scattering slope.

21.2. HR 8799 c

In our standard retrieval setup, we included data from SPHERE, GPI, CHARIS, GRAVITY, and ALES for HR 8799 c. We omitted the OSIRIS data, as it overlaps nearly completely with the GRAVITY data, is not flux calibrated, and requires either binning to lower resolution to be fit with the c-k opacity tables, or the use of the higher-resolution line-by-line opacities to fit the full resolution data, dramatically increasing computation time. Nevertheless, we performed several retrievals incorporating the OSIRIS data rather than the GRAVITY data to determine how this choice impacts the retrieved chemistry and clouds, and to determine if we can reproduce the findings of [Konopacky et al. \(2013\)](#). Overall, HR 8799 c proved challenging to fit: many retrieval setups either required far more model computations before convergence than the other three planets, or failed to converge entirely. The grid-fits for HR 8799 c also displayed the greatest variation between models.

Like HR 8799 b, the group $A \cap B$ retrievals of HR 8799 c are dominated by a single retrieval, with $\Delta \log_{10} \mathcal{Z} = 2$ compared to the next best retrieval. This retrieval used the [Z23](#) temperature profile and free chemistry, and requires high-altitude silicate clouds ($\log P_{\text{base}} = -3.4 \pm 1.8$ bar). The inferred effective temperature of 1159_{-12}^{+11} K is consistent with the range of temperatures found by the self-consistent grids, which spans from 1100 K to 1200 K, with the Bayesian fits averaging around ~ 1200 K. The retrieved mass ($8.5_{-0.5}^{+0.4} M_{\text{Jup}}$) is slightly higher than the dynamical mass estimate, though radius ($1.10 \pm 0.01 R_{\text{Jup}}$) is compatible with evolutionary models, and from these we derive a $\log g$ of $4.26_{-0.03}^{+0.02}$. The grid-fits also found plausible radii, favouring values slightly

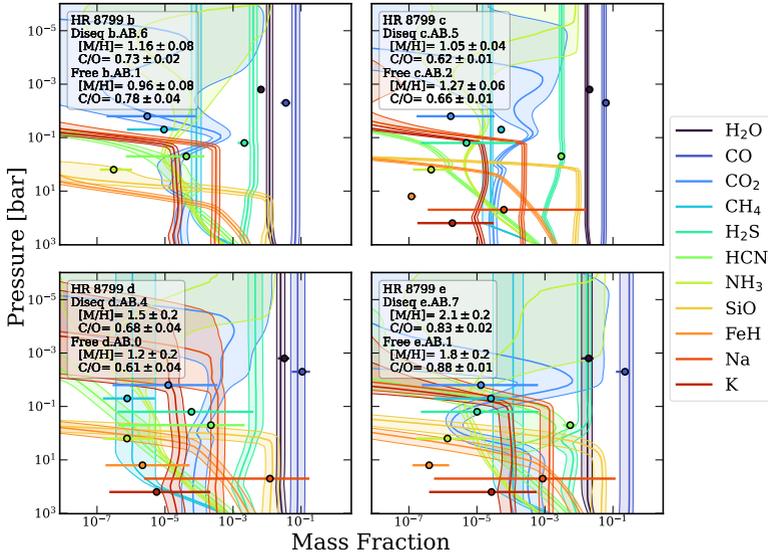


Figure 21.6: Mass fraction abundance profiles as a function of pressure in the atmospheres of HR 8799 b, c, d, and e. The solid lines indicate the median values of the most favoured disequilibrium retrieval for each planet, with the shaded region indicating the 90% confidence interval. The circular markers show the median values for each species from the most favoured free chemistry retrieval, with the error bars indicating the 90% confidence interval. The position along the pressure axis is arbitrary. The minimum mass fraction allowed in the free retrievals was 10^{-7} .

larger than $1 R_{\text{Jup}}$. All of the self-consistent models found temperatures of 1100–1200 K, and $\log g$ between 3.5 and 4.5. From the single-best χ^2 fits, the ATMO model found a $\log g$ of 4.5, while the remaining models find a lower solution of 3.5.

As with the other HR 8799 planets the retrievals favour highly enriched solutions, finding $[\text{M}/\text{H}] = 1.27^{+0.05}_{-0.05}$. The disequilibrium chemistry retrievals find slightly lower metallicities than the free chemistry retrievals, with the most-favoured disequilibrium retrieval finding a metallicity of 1.05 ± 0.04 . The data for HR 8799 c are highly discrepant in the H band (Figure 19.1), with the CHARIS data and photometry being around 50% brighter than the SPHERE and GPI data. As the metallicity is highly sensitive to the amplitude of the J, H, and K band peaks, such discrepancies need to be resolved in the data to ensure

reliable measurement of this parameter. All of the grid-fit solutions require high metallicity and are limited by the grid boundaries. The C/O ratios between the Bayesian fits and the χ^2 minimisation are consistent, typically favouring stellar to slightly substellar C/O. The retrievals present a more consistent picture, with most retrievals favouring a mildly super-stellar C/O ratio, with the average group $A \cap B$ C/O of $0.60^{+0.02}_{-0.01}$. The most favoured free chemistry retrieval for HR 8799 c find water and CO abundances consistent with most favoured disequilibrium retrieval. The free retrieval also finds an extremely high HCN mass fraction of $\log X_{\text{HCN}} = -2.54 \pm 0.05$, orders of magnitude higher than the predictions from equilibrium chemistry. This finding is strongly favoured by the Bayes factor, with $\log_{10} \mathcal{Z} = 30$ in favour of including HCN (c.AB.3 over c.A.8), with both retrievals using the Z23 temperature profile and clouds condensing at their equilibrium locations. The detection of HCN was largely driven by the ALES data; the wavelength dependence of the detection is discussed further in Section 20.2. If HCN is excluded, the overall metallicity is also increased, mostly due to a 3 dex increase in the H₂S abundance to $\log X_{\text{H}_2\text{S}} = -2.38 \pm 0.06$. While this solution is disfavoured, the higher H₂S abundance is more compatible with the equilibrium chemistry predictions. High resolution spectroscopy is likely required to precisely characterise the sulphur and nitrogen elemental abundances, and determine reliable abundances for these trace species. Although the CH₄ abundance is relatively low, with $\log X_{\text{CH}_4} = -4.3 \pm 0.06$, it is precisely constrained and detected with high confidence, $\Delta \log_{10} \mathcal{Z} = 11.5$ (c.AB.3 over c.A.6).

HR 8799 c is host to a highly cloudy atmosphere. The most favoured retrieval finds high altitude ($\log P_{\text{MgSiO}_3} = -3.4 \pm 1.8$ bar) MgSiO₃ cloud with a mass fraction of $\log X_{\text{MgSiO}_3} = -4.7 \pm 1.2$, together with a deep iron deck. The vertical mixing strength for the clouds is $\log K_{zz} = 8.0 \pm 0.9$, while the f_{sed} for both the silicate and iron clouds are compatible, with values between 5-6. Cloud composition could not be robustly determined for HR 8799 c, due to difficulties in retrieval convergence. However, we find that crystalline MgSiO₃ clouds (c.AB.5) provide a better fit by the χ^2 , and are favoured by the Bayes factor over patchy amorphous MgSiO₃ (c.AB.8), both using the M20 profile and disequilibrium chemistry. The use of patchy cloud layers may improve this fit, allowing individual layers to impact the spectrum independently, but this would come at the cost of substantially increasing the number of parameters to fit the continuum shape. The crystalline morphology provide a marginally better fit to the MIRI photometric data, but spectroscopic characterisation of the silicate feature is likely necessary to robustly distinguish these cases.

We ran a disequilibrium and free chemistry retrieval using the OSIRIS data in place of the GRAVITY data to check for consistency and to determine if we

could reproduce the findings of [Konopacky et al. \(2013\)](#). In order to use the correlated-k method of pRT, we binned the OSIRIS data by a factor of 4 to a spectral resolving power of $R \sim 1000$. We find that a high metallicity solution is still found using these data, with effective temperatures, and surface gravities consistent with the GRAVITY retrievals. For the disequilibrium retrieval, the C/O ratio is found to be significantly higher than any of the GRAVITY retrievals, as well as the free chemistry OSIRIS retrieval. While the free retrievals using the GRAVITY data find a slightly higher metallicity overall, the OSIRIS data finds a slightly higher CH₄ abundance of -3.81 ± 0.09 , as well as a much higher H₂S abundance. Overall, we find that the main findings of metal-rich atmospheres are reproducible regardless of whether we use the GRAVITY or OSIRIS data, and confirm the detection of water and CO in the atmosphere of HR 8799 c.

21.3. HR 8799 d

For HR 8799 d we included all available spectroscopic data as described in Section 19, but include a scaling factor for the ALES data set as otherwise it is incompatible with NACO photometric observations. This shifts the mean L-band flux of HR 8799 d to a similar magnitude as e and c.

Unlike HR 8799 b or c, d is well fit by a broad selection of models, and no single retrieval dominates the Bayesian average of figure 20.2. All of the preferred models ($\Delta \log_{10} \mathcal{Z} < 2$) used disequilibrium chemistry and the [M20](#) or [Z23](#) temperature profiles. Using the Bayesian average, the retrieved effective temperature is 1179_{-28}^{+31} K, compatible with the ranges found by the grid fits, which find T_{eff} from 1150 K – 1300 K. The planet mass ($9.2 \pm 0.1 M_{\text{Jup}}$) is tightly constrained by the dynamical mass prior, which in turn allowed for precise measurement of the planet radius ($1.26 \pm 0.07 R_{\text{Jup}}$) and $\log g$ (4.18 ± 0.05). This surface gravity is consistent with the estimates from the Bayesian grid fits, but is significantly higher than the 3.0-4.0 range found by the χ^2 fits. The self-consistent fits find marginally smaller radii than the retrievals, generally between 1.1 and $1.2 R_{\text{Jup}}$.

The metallicity of HR 8799 d is consistent with HR 8799 b and c, with $[M/H] = 1.2 \pm 0.2$. While most grid-fit solutions also favoured high-metallicity atmospheres, the Bayesian fit with the ATMO model find a solution consistent with solar metallicity, though the radius was inconsistent with evolutionary models ($0.926 \pm 0.004 R_{\text{Jup}}$). The best fit Exo-Rem model find a metallicity of 0.78 ± 0.03 , a radius of $1.168 \pm 0.006 R_{\text{Jup}}$ and an effective temperature of 1155 ± 5 K. The C/O ratio is always found to be consistent with the stellar value, with retrievals finding $C/O = 0.60_{-0.06}^{+0.04}$. The grid-fits are also typically

consistent with stellar, though the best fit Exo-Rem model found a substellar C/O ratio of 0.2.

Patchy clouds are marginally disfavoured by the Bayes factor, and the patchiness is poorly constrained, finding $f_{\text{cloud}} = 0.45 \pm 0.3$. The most favoured solutions require either amorphous or crystalline MgSiO_3 or crystalline MgSiO_3 , with a marginal preference for the amorphous structure. Each of these compositions displays slightly different near infrared slopes, shown in Figure 22.1. However, such a slope can be induced by various sources of continuum opacity that may not be fully accounted for in the retrieval. Thus mid-infrared observations of the silicate absorption features are necessary to robustly constrain the composition and particle geometry. There is no preference for a free cloud base pressure compared to the equilibrium position, suggesting that the AM01 model is sufficient to describe the clouds in this atmosphere.

While the disequilibrium retrievals are favoured over the free chemistry retrievals, there is excellent agreement between the freely retrieved abundances and the disequilibrium chemical profiles. Water and CO are both highly abundant, and the freely retrieved abundances agree with the disequilibrium profiles to within 1σ . No other species are both highly abundant and well constrained, so we do not perform leave-one-out retrievals to test for their presence. However, even at low abundances the freely retrieved CH_4 abundance is compatible with the disequilibrium profile.

21.4. HR 8799 e

The measurements of HR 8799 e largely reinforce existing literature values. The single most favoured retrieval, which also dominates the group $A \cap B$, used free chemistry and the Z23 temperature profile, together with a cloud base calculated using equilibrium condensation. This led to similar results as for HR 8799 c and d, and is consistent with the results of the grid fits. An effective temperature of 1138^{+30}_{-22} K is retrieved, compatible to within the uncertainties of the grid-fits, which found a temperature range of 1100 K to 1200 K. The mass posterior was determined by the dynamical mass prior, as was the planet radius, finding $M_{\text{pl}} = 7.5^{+0.6}_{-0.6}$ and $R_{\text{pl}} = 1.13^{+0.05}_{-0.05}$ respectively. This leads to a $\log g$ of $4.18^{+0.06}_{-0.05}$, slightly higher than the self-consistent estimates of 3.5–4.0. The overall best self-consistent model by the χ^2 was Exo-Rem, which finds an effective temperature of 1100K, a radius of $1.15 R_{\text{Jup}}$, and a somewhat low $\log g$ of 3.5. ATMO is the most favoured self-consistent model when using the Bayesian framework, though it found an unphysically small radius and higher temperature than other models. HR 8799 e is the only companion for which the MIRI photometry is not convincingly fit, as seen below in figure 22.5. Every

model underestimated the flux beyond 10 μm relative to the measurements, though this may be due to contamination from the host star or inner disc (Boccaletti et al., 2023).

Compared to the other three planets, HR 8799 e is found to have an even more metal rich atmosphere, with $[\text{M}/\text{H}] = 1.9^{+0.1}_{-0.2}$. Metallicities >1 were a universal feature of the retrievals for e. This was reinforced by the grid-fits, which uniformly find strong enrichment, running into the upper grid boundaries. Free chemistry retrievals are always preferred over the disequilibrium retrievals; from these we found the group $A \cap B$ C/O ratio is $0.87^{+0.02}_{-0.02}$. As the free chemistry retrieval C/O ratio only accounts for the gas-phase abundances, there is additional oxygen sequestered in the silicate clouds that could reduce the C/O ratio. However, the most favoured disequilibrium retrieval finds a similar value of 0.83 ± 0.02 , suggesting that HR 8799 e is somewhat of an outlier compared to the other three planets. Using a similar setup to Mollière et al. (2020) (e.AB.11), we find $\text{C}/\text{O} = 0.78^{+0.03}_{-0.03}$. The grid fits tend to find C/O ratios compatible with the stellar value, though the overall best fit Exo-Rem model also finds a higher value of 0.8.

In addition to the water and CO rich atmosphere, HCN is found to be highly abundant, with $\log X_{\text{HCN}} = -2.26 \pm 0.11$. This detection is strongly favoured by the Bayes factor, with $\log_{10} \mathcal{Z} = 7.5$. As with HR 8799 c and shown in Fig. 24.5, this detection was driven by using the HCN opacity to fit the ALES spectrum, though changes to the shape in the H and K-bands also provide a slightly better fit as well. This is slightly enriched compared to equilibrium chemistry predictions, but is expected for a metal rich planet with a relatively high C/O ratio (Giacobbe et al., 2021), and can be produced through photochemical reactions (Moses et al., 2013). In contrast, the presence of CH_4 is poorly constrained.

As with HR 8799 c and d, the most favoured retrieval for e favours silicate clouds with a deeper iron deck, both condensing at their equilibrium locations, with no preference for patchy clouds. A broad range of disequilibrium retrievals (e.AB.7–11) found consistent cloud properties, with condensation at the equilibrium location preferred, with a silicate abundance lower than predicted by equilibrium chemistry (between $10\times$ to $100\times$ less than equilibrium), and iron abundances consistent with equilibrium. Individual f_{sed} parameters for the silicate and iron clouds were not required, with variations in the Bayes factor driven more by the choice of patchiness and temperature profile.

21.5. Impacts of modelling choices

While we examined the primary atmospheric characteristics of each of the four HR 8799 planets, it is crucial to understand how the choice of model impacts these measurements. As described in Section 20, we performed retrievals using a broad selection of thermal structures, chemical parameterisations, and cloud models, each of which we is examined in detail below.

21.5.1. Thermal structure

For each planet we performed retrievals using four different temperature profile parameterisations. The [Z23](#) profiles is used in the most favoured retrieval for three of the four planets, while for HR 8799 d there is equal evidence for the [Z23](#) and [M20](#) profiles. The Guillot profile is found to be the second most preferred profile for HR 8799 c and e, while the spline profile using six nodes is strongly disfavoured by the Bayes factor. In Figure 21.7 we compare the retrieved and self-consistent temperature profiles for HR 8799 e. We find that there is excellent agreement between nearly all of the retrievals in the photosphere region, as well as with the best-fit Exo-Rem temperature profile. There is little variation in the photosphere region between the disequilibrium and free chemistry models. Only one model - the spline profile with free chemistry - found a profile more similar to that of the clear ATMO profile, though it is strongly disfavoured by the Bayes factor compared to the other free chemistry retrievals. The spline profile is the only profile that does not explicitly assume an adiabat deep in the atmosphere or rely on assumptions from self-consistent models, and so we cannot fully rule out the diabatic profiles of [Tremblin et al. \(2015\)](#). The bulk atmospheric properties are also reasonably consistent across the different temperature parameterisations. While there are statistically significant variations in the C/O ratio between the different parameterisations, they remain broadly consistent between 0.7 and 0.9.

Although the spline profile is disfavoured by the retrievals, it is a useful parameterisation to determine the amount of flexibility required by the model, and to explore the known degeneracies between the atmospheric thermal structure and clouds [Tremblin et al. \(2015, 2016\)](#). We performed a series of retrievals on HR 8799 b, varying the number of nodes in the spline profile and observing how the retrieved profile changes with the increased flexibility. We repeated this test for both a clear atmosphere model and an model with clouds condensing at the equilibrium base pressure. We find no significant differences in the temperature profiles between the clear and cloudy atmospheres. For HR 8799 b, the Bayes factor favours retrievals with three or four nodes in the spline profile.

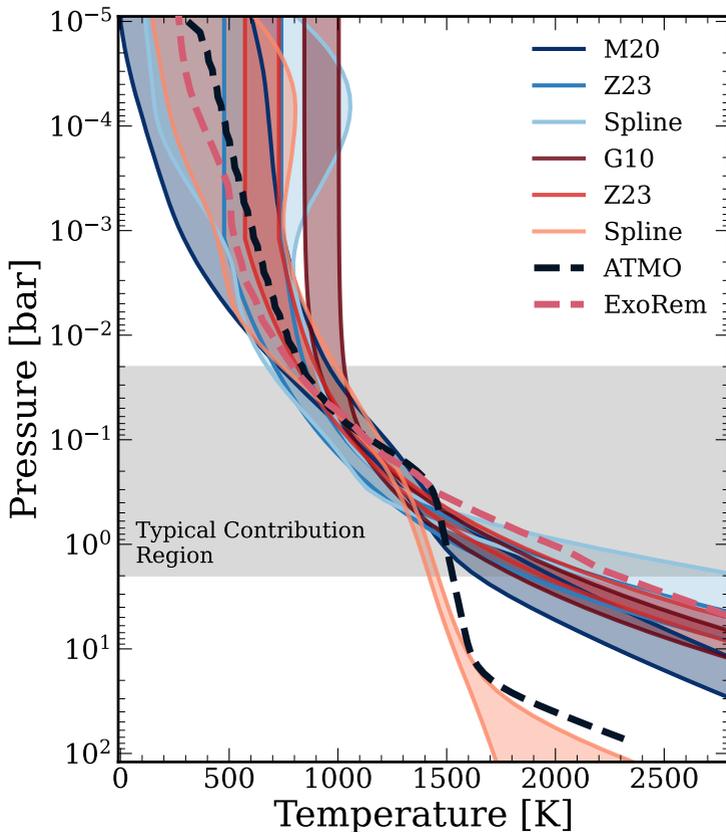


Figure 21.7: Temperature profiles for HR 8799 e. In blue are temperature profiles from disequilibrium retrievals, while in red are free chemistry retrievals. The shaded regions indicate 90% confidence intervals. Also included are the temperature profiles from the best fit self-consistent models.

Fewer nodes mean the profile cannot be accurately modelled, while more nodes add additional parameters without improving the fit to the spectra.

21.5.2. Chemistry

For all four planets we performed retrievals using a grid derived from an equilibrium chemistry solver with disequilibrium $\text{H}_2\text{O-CO-CH}_4$ quenching, as well as free chemistry retrievals where we directly retrieved the mass-fraction

abundance of various species. Fig. 21.6 shows the abundance profiles from the best fit disequilibrium and free chemistry retrievals for each planet. Both types of retrievals produce consistent metallicities and C/O ratios for each planet: overall there is excellent agreement in the water and CO abundances, which are the primary opacity sources in these atmospheres. Only HR 8799 b shows statistically significant discrepancy between the two methods for these species, with the free retrieval finding a slightly lower metallicity than the disequilibrium retrieval. Even for trace species, the free retrievals and disequilibrium retrievals are largely compatible, though only a few species have statistically significant detections in the free retrievals.

The strongest trace species detections are HCN in HR 8799 c and e, at abundances far higher than predicted by the equilibrium model. HR 8799 b also has a well constrained H₂S measurement, though it is not statistically significant. CH₄ is significantly detected in the atmosphere of HR 8799 c, demonstrating that with sufficient *S/N* and wavelength coverage, it is possible to constrain abundances at below 10⁻⁴ by mass. The free chemistry detection of CH₄ is at a moderately higher abundance than in the best fit disequilibrium model. While it is likely also present in the cooler atmosphere of b, additional wavelength coverage or higher *S/N* observations are required for a significant detection. For HR 8799 c, we include in Figures 24.4 and 24.5 in the appendix comparisons between the HR 8799 c data and models both with and without the contribution of CH₄ and HCN opacity, demonstrating the impact of these species on the spectral shape. While the HCN detection is driven primarily by the low flux of the ALES data, there is a significant change in the H-band shape, as well as a slight change in the peak amplitude of the K-band. As the ALES data are relatively low *S/N*, additional H and L band data should be obtained to confirm this detection. However, the CH₄ detection is driven by modest improvements in the fit throughout the K-band. Several abundant species predicted by the equilibrium network are not confidently detected by the free chemistry retrievals, such as CO₂, NH₃, and H₂S. Additional wavelength coverage or higher spectral resolution may allow for the characterisation of such species.

If we take the averaged free retrieval results at face value, we can derive elemental abundance ratios for each of the four planets, using a similar method to calculating the metallicity. Taking the volume mixing ratios of each molecular species, we can count the total number of C, N, O, and S atoms, and calculate the ratio relative to the planetary hydrogen abundance. Thus for example

$$\text{C/H} = \frac{X_{\text{CO}} + X_{\text{CO}_2} + X_{\text{CH}_4} + X_{\text{HCN}}}{X_{\text{H}_2} + 2X_{\text{H}_2\text{O}} + 2X_{\text{H}_2\text{S}} + X_{\text{HCN}} + 4X_{\text{CH}_4} + 3X_{\text{NH}_3}}, \quad (21.1)$$

where all abundances are measured in number fraction. These ratios are then normalised to the solar values from [Asplund et al. \(2009\)](#).

In Fig. 21.8, we present the elemental abundance ratios for each of the four planets. We find that most elements are enhanced relative to solar for all four planets. HR 8799 b appears depleted in nitrogen relative to the other planets, likely due to the nondetection of NH_3 , which will require observations of the $10\ \mu\text{m}$ feature to characterise. The HCN detections in HR 8799 c and e tightly constrain the nitrogen enhancement, though these planets still appear less enriched in nitrogen than in carbon or oxygen, though this is again likely due to additional nitrogen stored in N_2 and NH_3 , whose opacities are inaccessible at these wavelengths. The sulphur elemental ratio is poorly constrained for all of the planets apart from HR 8799 b, which has a precise - though not statistically significant - constraint on the H_2S abundance. HR 8799 b appears sulphur rich, while the remaining planets appear consistent with the solar value, or slightly depleted in sulphur, though this is largely due to a lack of measured chemical species.

The C/O ratio is a key consideration for planetary atmospheres. However, measuring the atmospheric C/O ratio and linking it to the bulk planet composition is far from trivial. [Lodders & Fegley \(2002\)](#) and [Lodders \(2003\)](#) explore the chemistry and condensation of substellar atmospheres, identifying the condensation sequence of refractory species throughout these atmospheres, finding that at typical L-dwarf temperatures there will be silicate clouds in the photosphere region. [Fonte et al. \(2023\)](#) demonstrate how oxygen is sequestered in silicate clouds and other refractory species. This was followed by the recent work from [Calamari et al. \(2024\)](#), who calculate the bulk planet C/O ratio from the atmospheric ratio, finding that the median sequestration of oxygen due to this condensation is $17.8^{+1.7}_{-2.3}\%$. They also identify a relation between the bulk and observed C/O ratio:

$$(\text{C/O})_{\text{obs}} \approx \frac{(\text{C/O})_{\text{bulk}}}{1 - 0.371 (\text{C/O})_{\text{bulk}}}. \quad (21.2)$$

Solving for $(\text{C/O})_{\text{bulk}}$, we find that for HR 8799 e, with an observed C/O ratio of $0.88^{+0.02}_{-0.02}$, should have a bulk C/O ratio of 0.66, much closer to the stellar value of 0.54. Likewise, HR 8799 d has the lowest observed C/O ratio of $0.61^{+0.03}_{-0.04}$, which translates to a modestly substellar bulk C/O ratio of 0.50. In general, this relation reduces the variation between the four planets, and brings the planetary C/O ratio more in line with the known stellar value.

While both chemistry models are compatible, they also share similar biases. The free chemistry model measures the gas phase abundance in the photosphere, and is primarily impacted by the atmosphere above the silicate clouds. Con-

versely, the underlying equilibrium model does remove oxygen from the gas phase due to condensation, though the additional flexibility in the cloud parameterisation means that it is only exactly correct for $f_{\text{sed}} = 1$. By parameterising disequilibrium via fixing the chemical abundances above a quench point, the model may lose this sensitivity, and therefore measures the abundances of CO, H₂O, and CH₄ in a similar fashion to the free chemistry model. Thus the C/O ratio as inferred by both models will be strongly impacted by the oxygen-depleted region above the silicate clouds, leading to over-estimates of the C/O ratio. Throughout this work we present these measurements, but we note that the adjustment introduced by Calamari et al. (2024) is likely a more accurate estimate of the bulk planet composition.

In addition to the elemental ratios, we also computed Z_{pl}/Z_* , which allows us to directly compare our metallicities to literature values, such as those of Thorngren et al. (2016). We converted the metallicity [M/H] of each atmosphere to Z_{pl} using the methods of Thorngren & Fortney (2019), adapting for our own notation:

$$10^{[\text{M}/\text{H}]} = \frac{1 + Y/X}{\left(Z_{\text{pl}}^{-1} - 1\right) \left(\frac{\mu_Z}{\mu_H}\right)}, \quad (21.3)$$

where X , Y , and Z are the solar hydrogen, helium and metal mass fractions and μ is the mean molecular weight of the metal content of the atmosphere. Rearranging and substituting in the measured atmospheric metallicity [M/H], we find:

$$Z_{\text{pl}} = \left(1 + \frac{1 + Y/X}{10^{[\text{M}/\text{H}]} \left(\frac{\mu_Z}{\mu_H}\right) (Z/H_{\odot})} \right)^{-1}. \quad (21.4)$$

We take the same assumptions as Thorngren & Fortney (2019), taking μ_Z to be 18, assuming most of the metal content is in water, μ_H to be 1 for atomic hydrogen, and Y/X to be 0.3383 as in Asplund et al. (2009). As the metallicity of HR 8799 A is near solar, Z_*/H is taken to be the solar value of $Z/H_{\odot} = 1.04 \times 10^{-3}$. To normalise to the stellar metallicity we follow Thorngren et al. (2016) and calculate the Z_* as

$$Z_* = 0.014 \times 10^{\text{Fe}/\text{H}}. \quad (21.5)$$

For HR 8799, we used solar metallicity to calculate Fe/H, but refer to the discussion in Section 18.1.

Disequilibrium chemistry has long been thought to play a key role in shaping the composition of the HR 8799 atmospheres (e.g. Marois et al., 2008).

With well-constrained chemical abundances, we can start to place limits on the strength of vertical mixing that drives this disequilibrium. The quench pressure we retrieve is defined as the level below which (in pressure) the abundances of H_2O , CO , and CH_4 become vertically constant. This parameterises dynamical mixing that homogenises the upper layers of the atmosphere. More rigorously, the quench point is defined as the point at which the chemical timescale t_{chem} and the mixing t_{mix} are equal. Following Zahnle & Marley (2014), the mixing timescale is defined through the ratio of the local atmospheric scale height H to the vertical eddy diffusion coefficient, K_{zz} :

$$t_{\text{mix}} = \frac{H^2}{K_{zz}}. \quad (21.6)$$

The chemical timescale depends on the reaction rates involved. Considering the $\text{CO}-\text{CH}_4$ reaction chain, Zahnle & Marley (2014) derive a timescale at the quench point (t_q) for CO . For strong mixing, pulling from material at depths below the point where the atmosphere is 1000 K, they find the timescale well described by an Arrhenius relation for quench pressure p in bar, metallicity m , where $m = 10^{[\text{M}/\text{H}]}$, and temperature T in kelvin:

$$t_{q1} = 1.5 \times 10^{-6} p^{-1} m^{-0.7} \exp(42000/T) \text{ s}. \quad (21.7)$$

For weak mixing, and therefore drawing from low temperatures with little CO , the timescale is found to be

$$t_{q2} = 40 p^{-2} \exp(25000/T) \text{ s}, \quad (21.8)$$

Combining the two, the total chemical timescale is defined as:

$$t_{\text{CO}} = \left(\frac{1}{t_{q1}} + \frac{1}{t_{q2}} \right)^{-1}, \quad (21.9)$$

which will favour the lower of the two values t_{q1} and t_{q2} . Equating the mixing and CO reaction timescales, we can infer the strength of vertical mixing in the atmospheres of the HR 8799 planets:

$$K_{zz} = \frac{H^2}{t_{\text{CO}}}, \quad (21.10)$$

where the scale height is defined as

$$H = \frac{k_{\text{B}}T}{\mu g}, \quad (21.11)$$

Table 21.3: Quench pressures, vertical mixing parameters, and sedimentation fractions for the HR 8799 planets.

| Planet | $\log P_q$ | K_{zzq} | $K_{zz,AM01}$ | f_{sed} | |
|--------|-------------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|-------------------------------|
| | [bar] | [cm ² /s] | [cm ² /s] | MgSiO ₃ | Fe |
| b | 1.7 ^{+0.2} _{-0.2} | 2.9 ^{+0.6} _{-0.7} | 8.6 ^{+0.8} _{-1.0} | 1.1 ^{+0.3} _{-0.3} | 4 ⁺² ₋₁ |
| c | 2.3 ^{+0.1} _{-0.1} | 6.3 ^{+0.3} _{-0.5} | 9.2 ^{+0.4} _{-0.4} | 3.3 ^{+0.3} _{-0.3} | 6 ⁺² ₋₂ |
| d | 1.0 ^{+0.6} _{-0.7} | 4.8 ^{+0.9} _{-1.3} | 9.2 ^{+0.7} _{-0.9} | 2 ⁺⁴ ₋₁ | 6 ⁺² ₋₂ |
| e | 2.1 ^{+0.4} _{-1.2} | 5.9 ^{+0.9} _{-0.6} | 8.6 ^{+0.7} _{-0.8} | 1.4 ^{+0.4} _{-0.3} | 6 ⁺² ₋₃ |

Notes

Measured from the group $A \cap B$ retrievals, providing median and $\pm 34\%$ confidence regions.

for temperature T , surface gravity g , mean molecular mass μ , and the Boltzmann constant k_B . In order to calculate these quantities for the HR 8799 planets, we take T_{eff} as a representative temperature to calculate the timescales and scale height, g as the measured surface gravity, and μ as the average mean molecular weight of the atmosphere. As K_{zz} is exponential in temperature, the choice of what temperature to use strongly influences the measured value. Using the temperature at the quench pressure, typically deep in the atmosphere, results in unphysically strong vertical mixing, with $\log K_{zz} \approx 20$. A more thorough analysis could try to measure the vertical mixing as a function of temperature throughout the atmosphere, but the current data quality is of insufficient resolution or S/N for such measurements. Thus we treat T_{eff} as a representative temperature with which to determine the vertical mixing strength. We include the results of these calculations, together with the retrieved K_{zz} used to parameterise the AM01 clouds in Table 21.3.

The quench pressure is well constrained in the Bayesian average of group $A \cap B$ retrievals for all four planets; the values of which are listed in Table 21.3. All of the planets quench below the photosphere, with d quenching at the highest altitude, around 10 bar. Including only the disequilibrium retrievals in the Bayesian average of group $A \cap B$, we derived K_{zz} from the quench pressure. We then turn to [Soni & Acharyya \(2023\)](#) for a comparison, who provide predictions for CH₄ and CO abundances for varying K_{zz} , T_{eff} , and $\log g$ across a range of metallicities. For our measured T_{eff} and CH₄ abundances, we should expect

$\log K_{zz}$ of around 6 for the warmer three companions, regardless of whether we use the measured CO or CH₄ abundance. For HR 8799 b a much lower value (less than ~ 2) is expected, assuming a $10\times$ solar metallicity. Our inferred K_{zz} values for c, d, and e are compatible with this prediction, finding $\log K_{zz}$ between 5 and 6. For HR 8799 b we also measure weak mixing, with $\log K_{zz} = 2.9^{+0.6}_{-0.7}$, which is again compatible with the predictions of [Soni & Acharyya \(2023\)](#). These measurements from the quench pressure are also inconsistent with the parameter used in the [AM01](#) clouds, which require stronger vertical mixing of $\log K_{zz} \sim 9$. This discrepancy is perhaps not surprising: 3D modelling predicts that K_{zz} should vary with altitude throughout the atmosphere, and the larger cloud particles likely respond to the atmospheric motion differently than the gas phase constituents. Ultimately, more precise constraints on the thermal structure and chemical abundances, as well as trace species detections are necessary to derive a more precise vertical mixing strength. Further modelling work is also necessary to provide a more physically motivated transport model than a vertically constant eddy diffusion coefficient.

21.5.3. Clouds

For the inner three planets, we find the most favoured solution is an optically thin silicate cloud lying above a deeper, optically thick iron cloud deck. In our standard setup, the clouds were parameterised as in [AM01](#), with the clouds condensing at the intersection of their condensation curve and the temperature profile, with their extent determined by f_{sed} . The cloud mass fraction was allowed to scale from equilibrium. In totally free retrievals, the cloud abundances, locations, and vertical extents were all free parameters of the model. This decouples the clouds from both the chemistry and the atmospheric thermal structure, allowing them to fit the spectral shape, but in potentially nonphysical configurations. In this framework the cloud extent was then parameterised as in [AM01](#), determined by f_{sed} and K_{zz} . Using this setup we find an optically thin silicate cloud lying above a compact iron cloud, while the [AM01](#) setup finds an iron cloud that extends high above the silicate cloud. Depending on the choice of other parameters, either the free cloud base or the equilibrium base can be preferred by the Bayes factor. HR 8799 e free chemistry retrievals strongly favour the equilibrium condensed clouds, while the disequilibrium retrievals favour the free cloud base setup. However, in general the clouds condensing at equilibrium are the most favoured setup for each planet. Without broad wavelength coverage and high spectral resolution to probe a high dynamic range in pressure, it is difficult to robustly distinguish between the different potential cloud structures.

There is a marginal preference for clouds parameterised using a [Hansen \(1971\)](#) particle size distribution over a log-normal distribution (e.AB.20 over e.AB.25, $\Delta \log_{10} \mathcal{Z} = 0.7$), though this was only compared for HR 8799 e using the Z23 profile, and assuming the clouds condense at their equilibrium saturation location. In this case a_h is calculated from f_{sed} and K_{ZZ} , and a lower f_{sed} for the MgSiO_3 cloud was retrieved than with the log normal distribution ($f_{\text{SED, Hansen}} = 1.18 \pm 0.20$). The effective distribution width parameter, b_h was found to be 0.016, which is narrower than the distributions found by [Burningham et al. \(2021\)](#).

In general our clouds are comparable to those of [Burningham et al. \(2021\)](#) and other similar studies (e.g. [Mollière et al., 2020](#); [Vos et al., 2023](#); [Balmer et al., 2023](#)). [Burningham et al. \(2021\)](#) find a combination of MgSiO_3 , SiO_2 , and Fe clouds provides the best-fit model to an ultracool field dwarf, 2MASSW J2224438–015852. The MgSiO_3 clouds in their model are located at 10^{-3} bar with a maximum optical depth of $\tau = 0.3$ at $1 \mu\text{m}$ and an effective particle radius of $\sim 0.04 \mu\text{m}$. The SiO_2 clouds are slightly deeper, at 10^{-2} bar, and are optically thick at $1 \mu\text{m}$. This is the same location where we find MgSiO_3 clouds condensing in the atmospheres of HR 8799 c, d, and e. The iron cloud is found to be deep in the atmosphere, though with an extended structure, with some contribution at the same altitude as the silicate clouds. While we did not fit for a three cloud model, the similar locations and optical depths of these clouds suggests similar structures between the objects, even though they differ in effective temperature by hundreds of kelvin.

Figure 21.10 highlights how the difference in particle radius contribute to the difference in the wavelength dependence of the cloud optical depth. The cloud particle radius in the [AM01](#) model is a function of many atmospheric factors, including the temperature, mixing strength, f_{sed} , and particle number density. We see that changes in the particle radius are correlated with changes in both the temperature and the particle density. [Luna & Morley \(2021\)](#) explore the impact of particle size and composition on the spectral signatures of clouds in young brown dwarfs in the mid-infrared, finding that small particle sizes will lead to visible features in the planetary spectrum. However, we see in Figure 21.10 that the particle sizes in regions of the atmospheres with significant cloud mass fraction tend to be larger ($> 1 \mu\text{m}$), and that there are no indications of deep cloud absorption features in the mid-infrared, even though silicate clouds are present in the atmosphere. Small silicate particles should produce deeper absorption features, which are not observed in the mid-infrared, suggesting that the impact of the small mean particle size in the upper atmosphere does not contribute strongly to the cloud opacity.

Due to the different slopes in the near-infrared opacity as a function of

wavelength, as shown in Figure 22.1, there is some sensitivity to different compositions and particle geometries. This can explain the mild preference for amorphous MgSiO_3 clouds in certain like-for-like retrieval comparisons. In reverse, the lack of features leads to a preference for Na_2S clouds in HR 8799 b. Performing additional tests on HR 8799 e, we find that Mg_2SiO_4 (e.AB.14) are mildly disfavoured by the retrievals, and Al_2O_3 (e.AB.24) clouds are strongly disfavoured, though this is again more likely from their condensation location rather than from the impact of aerosol spectral features. Observations of the silicate absorption features at $10\ \mu\text{m}$ would allow more precise measurement of this wavelength dependence, and in turn place better constraints on the cloud structure and composition.

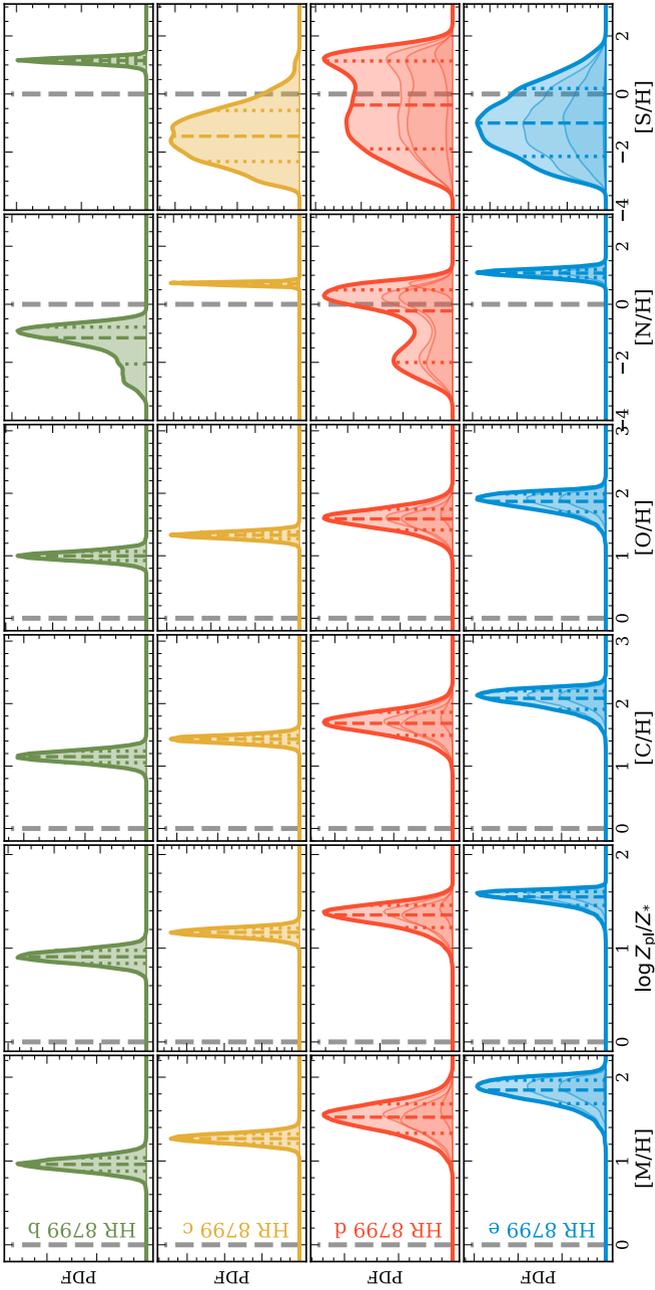


Figure 21.8: Metallicities and elemental number ratios for the HR 8799 planets derived from Bayesian averaged group $A \cap B$ free chemistry retrievals.

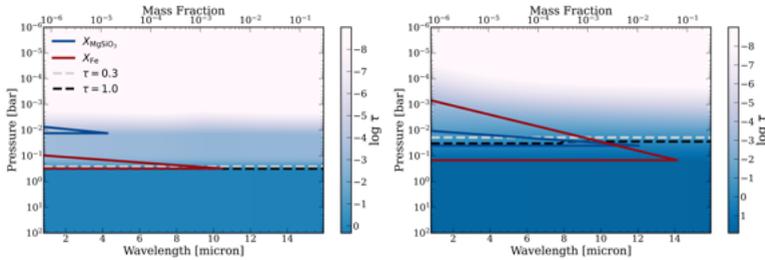


Figure 21.9: Cloud properties for HR 8799 e, showing the optical depth due to clouds as a function of pressure (colour map), with the $\tau = 0.3$ and $\tau = 1.0$ contours highlighted by the dashed lines. The solid lines indicate the mass fraction abundance of the MgSiO_3 and Fe clouds in blue and red respectively. **Left:** Cloud properties e.AB.13, using a freely retrieved cloud base pressure and abundances. **Right:** The same, but for Case e.AB.2, which uses the equilibrium condensation to determine the location of the cloud base. The abundances are determined using equilibrium chemistry, and retrieving a scaling factor, ($\log S_{\text{Fe}} = 0.0 \pm 1.1$, $\log S_{\text{MgSiO}_3} = -0.8 \pm 1.0$).

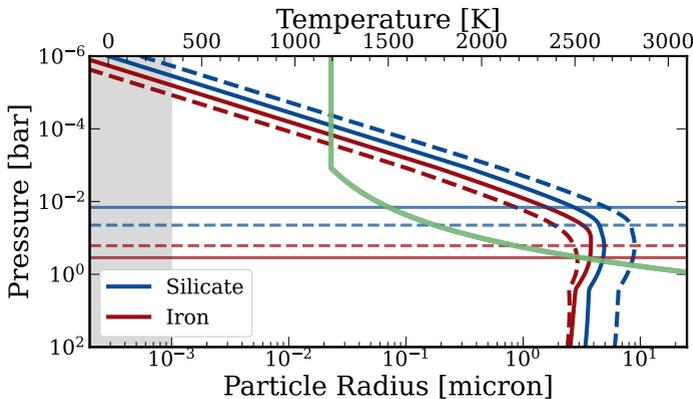


Figure 21.10: Effective particle radii as a function of altitude for silicate (blue) and iron (red) clouds. The solid lines indicate the radii for Case e.AB.13, which used a free cloud base pressure and abundance, while the dashed lines are for Case e.AB.2, which used equilibrium condensation and scaled equilibrium abundances. The horizontal lines indicate the cloud base pressure. The green line indicates the temperature profile. The shaded regions indicate unphysical particle sizes where the opacity contribution is set to 0.

22. DISCUSSION

22.1. Highly enriched atmospheres

While some enrichment is expected in giant planets formed through core accretion, metallicities of nearly $100\times$ the stellar value are far beyond the expectations for planets with masses larger than that of Jupiter. Having transformed the atmospheric metallicities from the retrievals to Z_{pl}/Z_* using equations 21.4 and 21.5, we can compare the HR 8799 planets to the broader population. Figure 22.2 shows how the inferred metallicities of the HR 8799 system compare to those of other directly imaged planets, and to a fit from [Thorngren et al. \(2016\)](#) derived from a sample of transiting exoplanets. The HR 8799 planets are clear outliers amongst the directly imaged planets, whose metallicities were taken from the literature. We again used equations 21.4 and 21.5 to convert from a measurement of $[M/H]$ to Z_{pl}/Z_* . Where stellar metallicities are not available, we assumed a metallicity $[\text{Fe}/H]=0$. Only the $2 M_{\text{Jup}}$ planet Af Lep b has a comparable degree of enrichment to the HR 8799 planets ([Zhang et al., 2023](#)). However, they are comparable to hot Jupiters observed in transmission, such as WASP 39 b ([Rustamkulov et al., 2023a](#)). [Looper et al. \(2008\)](#) and [Stephens et al. \(2009\)](#) demonstrate how high metallicity atmospheres facilitate condensation, in turn leading to the strong reddening seen in a subset of L dwarfs. The retrieval of highly metal rich and very cloudy atmospheres is consistent with this picture from brown dwarfs. Further supporting the high $[M/H]$ retrievals are their consistency with the self-consistent grids, which always favour their upper limits.

To validate these findings, we performed a series of test retrievals for HR 8799e, fixing the metallicity to solar composition, and compared cases where $[M/H]$ is fixed to values between 0.0 and 2.0 in steps of 0.5 dex. Figure 22.3 shows the best fit spectra from each of these retrievals, showing the clear differences in the J, H, and K bands between the different metallicity cases that are unable to be compensated for by varying other atmospheric parameters. We find that the cases of $[M/H] = 1.5$ and 2.0 are strongly favoured over the cases between 0.0 and 1.0, with solar composition disfavoured at $\Delta \log \mathcal{Z} > 10$. The remaining atmospheric parameters also significantly varied between the different retrievals: the C/O ratio increases with increasing $[M/H]$, maintaining a relatively constant abundance of H_2O in the atmosphere while allowing the CO abundance to increase. This combines with the decreasing f_{sed} , increasing the cloudiness of the planet to dampen the stronger molecular features at high metallicity. A more thorough treatment of the condensation and chemistry in the retrieval framework are likely necessary to accurately infer both of these parameters.

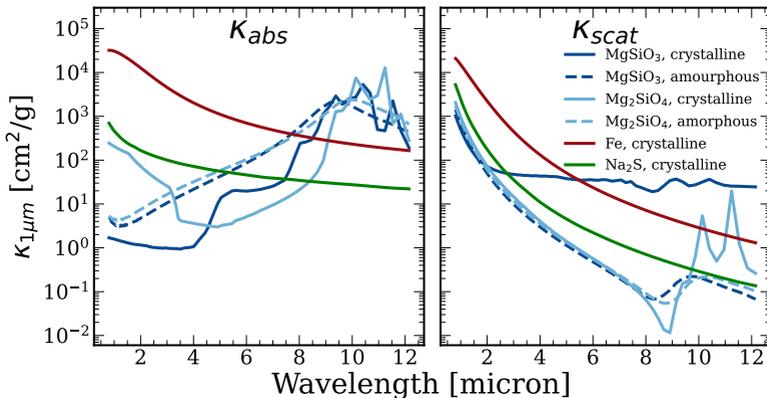


Figure 22.1: Cloud absorption (left) and scattering (right) opacities for different condensate compositions and structure, for 1- μm particles. Dark blue indicates MgSiO_3 , light blue is Mg_2SiO_4 . Solid (dashed) lines are for crystalline (amorphous) substances. The solid red (green) line is for crystalline iron (Na_2S).

The spectra shown in Figure 22.3 show that the retrieved metallicity is strongly dependant on both the height of the J, H, and K band peaks, as well as the shape of these features. However, the J bands is only covered by the low-resolution SPHERE data, with relatively poor S/N , and different instruments measure significantly different flux in the H band. Without compatible measurements in this spectral regime, robust conclusions about the metallicity are hard to draw. Additional constraints can be obtained from other wavelength regions; [Lodders & Fegley \(2002\)](#) find that the CO_2 abundance scales proportionally to $[\text{M}/\text{H}]^2$. With strong features between 3 and 4 μm , as well as in the mid-infrared, future observations should be able to place robust constraints on this parameter. Section 22.4 discusses the potential of JWST to make such observations.

Finally, we performed independent comparisons to the moderate resolution OSIRIS spectra presented in [Ruffio et al. \(2021\)](#). Following their methods, we used the parameters of the single best-fit disequilibrium models for HR 8799 b, c, and d to compute a high spectral resolution using line-by-line opacity lists. We convolved this model to the OSIRIS instrumental spectral resolution, and binned the model to the OSIRIS wavelength grid. The resulting spectra was multiplied by the atmospheric transmission function provided by [Ruffio et al.](#)

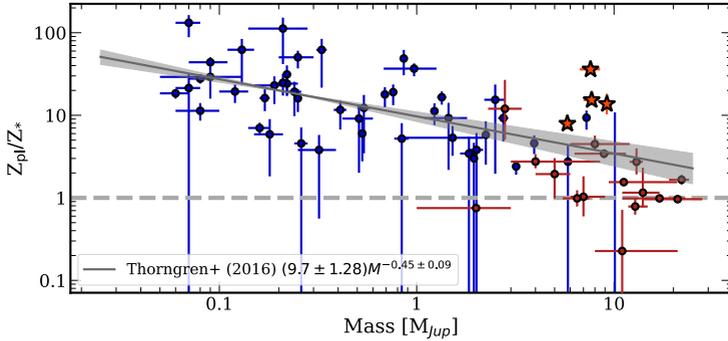


Figure 22.2: Metallicities of the exoplanet planet population. In blue are transiting planets, adapted from [Thorngren et al. \(2016\)](#) (dark blue). In red are directly imaged planets, with references listed below. Indicated with stars are the HR 8799 planets as measured in this work. The grey line indicates the fit by [Thorngren et al. \(2016\)](#).

Notes: β Pic b ([Gravity Collaboration et al., 2020](#)); PDS 70 b ([Wang et al., 2021](#)); 51 Eri b ([Whiteford et al., 2023](#)); VHS 1256 b ([Hoch et al., 2022](#)); HIP 65426 b ([Petrucci et al., 2021](#)); κ And b ([Bonnefoy et al., 2014b](#); [Wilcomb et al., 2020](#)); YSES 1 b ([Zhang et al., 2021](#)); AB Pic b ([Palma-Bifani et al., 2023](#)); AF Lep b ([Zhang et al., 2023](#)); GJ 504 b ([Bonnefoy et al., 2018](#)); HD 95086 b ([Desgrange et al., 2022](#)); Ross 458 c ([Burgasser et al., 2010](#)); DH Tau b ([Patience et al., 2012](#)); HN Peg b ([Leggett et al., 2008](#); [Suárez et al., 2021](#)); CT Cha b ([Schmidt et al., 2008](#)); GQ Lup b ([Demars et al., 2023](#)).

(2021). The continuum was measured by high-pass filtering the spectrum, and was subsequently subtracted from the model. We fit a scaling factor between the model and the OSIRIS data and computed the resulting χ^2 . This exercise was repeated, setting the metallicity of the model to solar. For HR 8799 c and b we find that the high metallicity model provides a better fit to the OSIRIS data than the solar model, with the caveat that fixing the metallicity during the retrieval may result in a better fit than setting it a posteriori, without changing other atmospheric parameters. For HR 8799 d the high metallicity model is only a marginally better fit than the solar metallicity model. In general, the fits from the GRAVITY and remaining archival data provide reasonable fits to the $R \approx 4000$ OSIRIS data, of similar quality to the fits displayed in [Ruffio et al. \(2021\)](#), and are included in the appendix in Figure 24.3.

22.2. Impacts of data selection

Given the inhomogeneity in the data in terms of spectral resolution, S/N , observing strategy and more, we performed a series of retrievals to examine the impact of different datasets on the retrieval results. We first performed retrievals using only the GRAVITY data to determine the constraining power of this new dataset. We found that using only the GRAVITY data we could rule out clear atmosphere solutions at $\Delta \log Z > 7$ (e.g. e.A.31 over e.A.32), using the Z23 profile and disequilibrium chemistry. Using only the GRAVITY observations for HR 8799 e, we could obtain estimates of the metallicity (2 ± 0.3), effective temperature (1143_{-32}^{+38} K), and C/O ratio ($0.71_{-0.2}^{+0.08}$), which are broadly consistent with the results from the combined dataset. We obtain similarly reliable estimates for b (b.A.32), but using only the GRAVITY data for d (d.A.12) and c (c.A.11), resulted in significantly lower metallicities. All of these retrievals use the same setup of the Z23 profile, disequilibrium chemistry, and silicate and iron clouds condensing at their equilibrium saturation point. Given the higher spectral resolution and S/N of the GRAVITY data, this leads to the conclusion that solutions to the full retrievals are largely driven by the fits to the GRAVITY spectra. The C/O ratios for d and b are also incompatible when using only the GRAVITY data, finding substellar values for both planets.

Conversely, we also performed retrievals that exclude the GRAVITY spectra. For HR 8799 e (e.A.33) and d (d.A.13), we find that the retrieved parameters again broadly agree with the retrieval including the GRAVITY data. Thus even if the retrievals are dominated by the GRAVITY data, the conclusions we draw are robust even when excluding the GRAVITY spectra.

22.3. Formation

The formation mechanism of the HR 8799 planets has seen much debate since their discovery: simply put, how can one form four such massive planets in the same system? The C/O ratio and metallicity are the best formation tracers observed to date in these planets. Nevertheless, neither gravitational instability nor core accretion scenarios have been ruled out. The high degree of enrichment in these objects would seem to suggest a core accretion formation scenario. However, Wang (2023) finds that even for $[M/H]=0.5$, approximately 100 earth masses of solids are required to enrich the atmosphere of HR 8799 e. With the even higher metallicities measured in this work, this number would increase, and it is unclear if it is possible to have nearly 1000 earth masses of metals available in a protoplanetary disc and accreted with high efficiency. Wang (2023) additionally find that late accretion of planetesimals would require less

material to result in a similar degree of enrichment, potentially only requiring a few hundred earth masses of solids to achieve the high metallicities of all four planets, though this material should quickly settle out of the atmosphere.

All four companions have stellar C/O ratios or higher, and the C/O ratio in this system varies with separation, decreasing from b to d, before a sharp increase in the C/O ratio for the innermost planet. Super-stellar C/O ratios have been tied to core-accretion formation together with pebble drift and evaporation (Schneider & Bitsch, 2021b,a; Mollière et al., 2022). A pathway to significant metal enrichment was found by Bitsch & Mah (2023), though it predicts that the CO rich pebbles should evaporate near the CO iceline, which is outside the radius of even HR 8799 b. Planetesimal accretion cannot be ruled out either: if large amounts of solids, with near-stellar composition, are accreted, more metal-rich planets will have C/O ratios that approach the stellar value. This is consistent with the trends in metallicity and C/O between the b, c, and d planets, though e remains an exception. Such a transition could be explained by the outer three planets, particularly d, trapping water ice and preventing these solids from reaching the innermost planet. Some combination of these mechanisms could explain both the atmospheric enrichment and the trends in the C/O ratio: for example, early enrichment from evaporating pebbles could lead to the high planetary metallicities, while late accretion of planetesimals could then drive the C/O ratio down towards the solar value. Alternatively, Chen et al. (2024) demonstrate that the opening of gaps in a protoplanetary disc can significantly alter the composition of the gas and ices available to accrete onto forming planets, and it seems likely that substructure induced by the four HR 8799 planets would strongly impact their eventual composition.

With effective temperatures well over 1000 K and radii significantly larger than that of Jupiter, the coldest initial condition scenarios of Marley et al. (2007) can be excluded. Beyond this constraint, the masses, luminosities, and radii seem largely consistent with a broad range of potential evolutionary tracks (e.g. Baraffe et al., 2003; Saumon & Marley, 2008; Mordasini et al., 2017). Further work to measure more formation tracers is clearly necessary to unravel this system. Mid-infrared spectroscopic observations could characterise the NH₃ abundance, at least in HR 8799 b, as well as the debris disc observed in Boccaletti et al. (2023). Higher spectral resolution could enable the measurement of carbon isotopes (e.g. as well as place better constraints on the metallicity of each of these atmospheres). Such measurements, combined with dedicated formation models of giant planets outside the water iceline, are necessary to determine whether these four planets share a formation pathway, or whether there were different mechanisms impacting different regions of the protoplanetary disc.

22.4. Predictions for JWST

The high spectral resolution modes of JWST may allow us to verify the measurements made in this work, particularly through observation of CO₂, CO, and CH₄ features in the near infrared and the silicate absorption features near 10 μm . To this end, we present the range of model predictions in these wavelength regions at the spectral resolution of the JWST instruments, for NIRSpec in figure 22.4 and for the MIRI/MRS in figure 22.5. The comparisons to the MIRI photometry shown in figure 22.5 demonstrate the typical degree of compatibility between the models and the data in these mid-infrared wavelengths. In the NIRSpec/G395H wavelength range we see significant discrepancy between models for the same planet in the amplitude of the CO₂ feature at 3.8 μm , as well as the CH₄ feature at 3.3 μm and the CO lines between 4.5 and 5 μm . These observations will also be able to confirm the presence of HCN in the atmospheres of HR 8799 c and e. Precise measurement of these features should provide robust constraints on the metallicity of these objects, verifying the degree of enrichment found via the ground-based observations. While silicate clouds are preferred in the retrieval comparison, none of the models show signs of deep silicate absorption features near 10 μm , but spectroscopic observations are required to validate these models. Mid-infrared observations will be particularly valuable for HR 8799 b, and will allow the clear detection of ammonia. If combined with a chemical model to determine the ratios of NH₃:HCN:N₂, this will allow for the measurement of the N/O ratio, which can also be used as a formation diagnostic (Turrini et al., 2021; Pacetti et al., 2022). Recently Ruffio et al. (2023) demonstrated the potential for high-contrast imaging with NIRSpec; even without the use of a coronagraph it should be possible to obtain flux calibrated, moderate resolution spectroscopy of HR 8799 b, c, and d using the NIRSpec IFU through a combination of forward modelling and reference differential imaging. In the case where the planet signal is unable to be separated from that of the host star, Patapis et al. (2022) demonstrated that it will be possible to at least identify trace species through molecular mapping in the mid-infrared, though this will be unable to characterise the broad wavelength features of the silicate clouds.

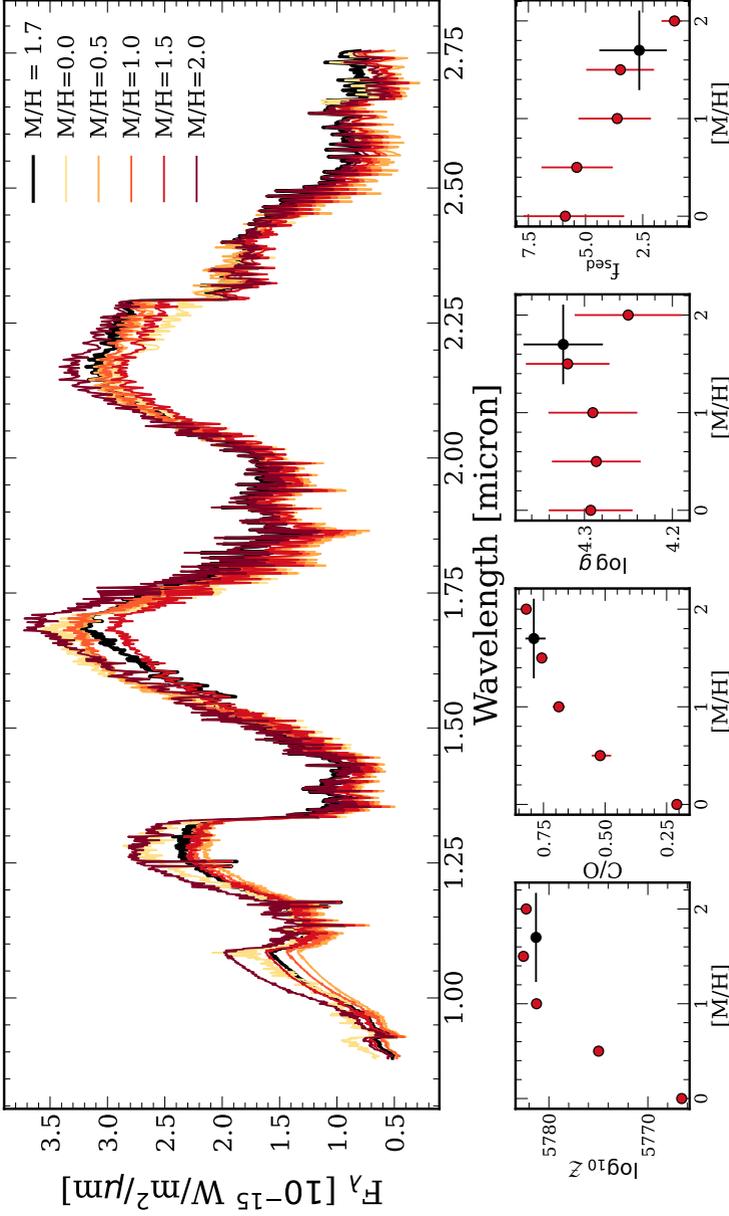


Figure 22.3: Comparison retrievals of HR 8799 e where the metallicity is fixed to a constant value. Black points denote the equivalent retrieval where the metallicity is a free parameter. The top panel shows how the best fit spectra differ across the set of retrievals, while the bottom row shows how different atmospheric parameters vary as a function of metallicity.

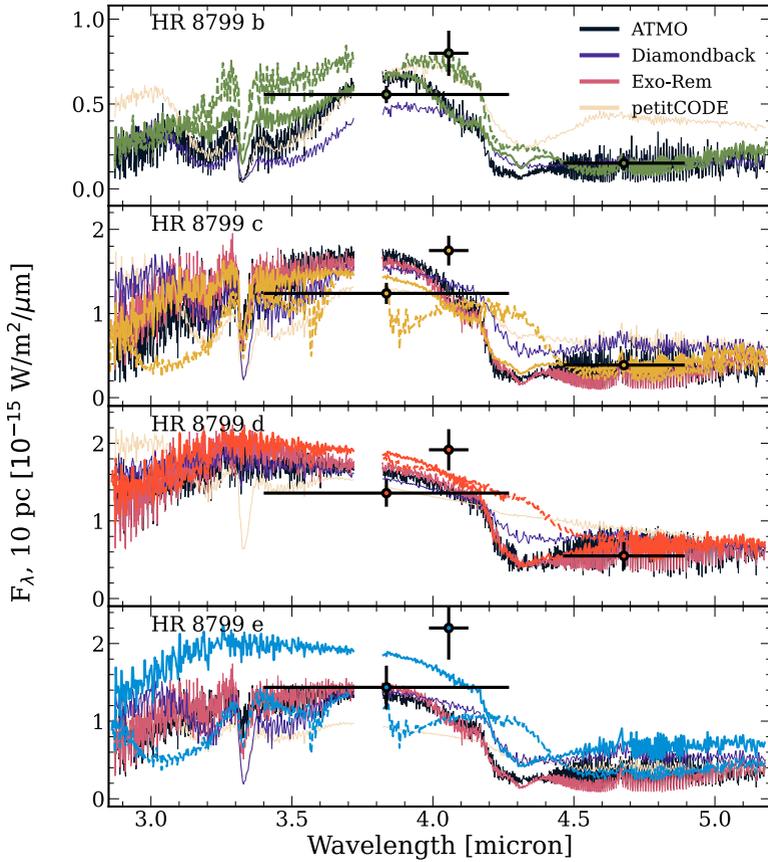


Figure 22.4: Predictions for NIRSpec/G395H based on most favoured disequilibrium (solid) and free chemistry (dashed) retrievals, together with the best-fit self-consistent models from each grid.

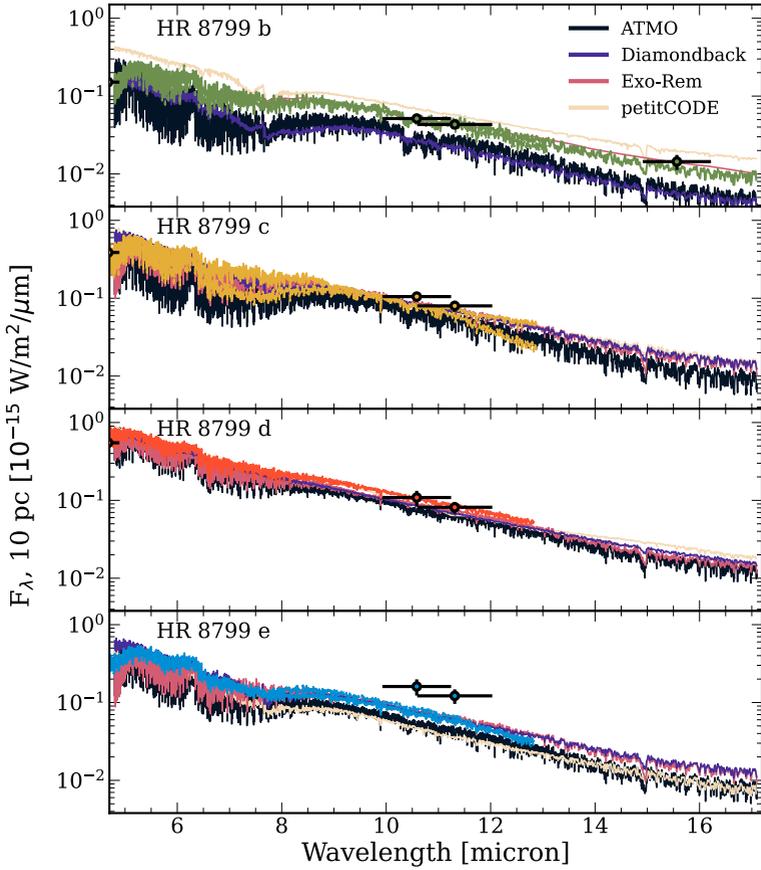


Figure 22.5: Predictions for MIRI/MRS based on the most favoured disequilibrium (solid) and free chemistry (dashed) retrievals, together with the best-fit self-consistent models from each grid.

23. CONCLUSIONS

After more than 15 years of study, the HR 8799 planets remain mysterious, though increasing data quality is allowing us to peer deeper into these atmospheres than ever before. We present new K-band spectra from the VLTI-/GRAVITY, which together with a large set of archival data form the basis of our atmospheric analysis. Using pet i tRADTRANS retrievals and fits to self-consistent grids, we inferred the atmospheric properties of all four companions, with reasonable agreement between the two methods. Our results are broadly consistent with the literature in terms of effective temperature, mass, surface gravity, and radius for all four planets. The use of the dynamical mass as a prior in the retrievals when determining $\log g$ allows us to reliably retrieve physically reasonable planet radii.

We find that all four planets are strongly enriched in metals, though there is still discrepancy between different models in constraining the precise value. This was validated by running retrievals using different temperature profiles and chemical models, and comparing to self-consistent grids. Further self-consistent modelling is necessary, particularly to extend model grids out to high metallicities. The C/O ratio is stellar to superstellar for all four planets. It decreases from the outermost planet to HR 8799 d, while HR 8799 e has a higher C/O ratio than the other companions. We confidently detect HCN in HR 8799 c and e, at abundances far higher than predicted by equilibrium chemistry; though this detection is largely driven by low- S/N data from LBT/ALES. CH_4 is also confidently detected in HR 8799 c for the first time. From the disequilibrium chemistry retrievals, H_2S appears to be a highly abundant species in all of the planets, but higher S/N and spectral resolution are required for a confident detection in a free retrieval framework. Using our retrieved quench pressure and chemical abundances, we are able to derive a vertical mixing strength, finding K_{zz} values compatible with high-metallicity predictions from [Soni & Acharyya \(2023\)](#). The mixing strength is stronger for the warmer planets, at $\log K_{zz} \approx 6$, and is lower for HR 8799 b with $\log K_{zz} \approx 2$.

All of the planets are highly cloudy. For the inner three planets, these clouds are composed of silicate clouds lying above the photosphere, and deep, dense iron clouds forming the base of the photosphere. Cooler than the other three planets, the most favoured model for HR 8799 b requires Na_2S clouds. All of the planets have effective temperatures consistent with literature values, with HR 8799 b still unique in its lower temperature and mass compared to its siblings.

We emphasise the use of robust model comparison in this work: while it may be difficult to present precise measurements of certain properties, the

use of multiple methods and models allows us to draw a robust portrait of each of these atmospheres. We also note that our conclusions rely on data with significant incompatibilities, particularly in the H-band flux. While we performed extensive analysis to mitigate the influence of any individual dataset, further observations are required to obtain reliable spectroscopic measurements in the near-infrared. While the HR 8799 planets share many similarities, much like our own Solar System there are differences in their atmospheric properties, which require further study.

We would like to thank Quinn Konopacky, Alice Zurlo, Alex Greenbaum, Beth Biller, Olivier Flasseur, David Doelman, and Pengyu Lui for providing the archival datasets used in this work. We are also grateful to Gilles Loupe and Malavika Vasist for providing helpful suggestions on how to implement the spline temperature profile. Thanks as well to Daniel Thorngren for providing the model for calculating the planet metallicity Z . Finally, we are grateful to our anonymous reviewer for their thorough and insightful report. SL acknowledges the support of the French Agence Nationale de la Recherche (ANR), under grant ANR-21-CE31-0017 (project ExoVLIT). J.J.W., A.C., and S.B. acknowledge the support of NASA XRP award 80NSSC23K0280. G.-D.M. acknowledges the support of the DFG priority program SPP 1992 “Exploring the Diversity of Extrasolar Planets” (MA 9185/1) and from the Swiss National Science Foundation under grant 200021_204847 “PlanetsInTime”. Parts of this work have been carried out within the framework of the NCCR PlanetS supported by the Swiss National Science Foundation. This work is based on observations collected at the European Southern Observatory under ESO programme 1104.C-0651. This publication makes use of VOSA, developed under the Spanish Virtual Observatory project supported by the Spanish MINECO through grant AyA2017-84089. VOSA has been partially updated by using funding from the European Union’s Horizon 2020 Research and Innovation Programme, under Grant Agreement n° 776403 (EXOPLANETS-A). Software used: `petitRADTRANS`, `pyKLIP`, `species`, `VIP-HCI`, `pyMultiNest`, `phot_utils`, `Python`, `numpy`, `matplotlib`, `astropy`, `sympy` and `Aspro`.

24. APPENDICES

24.A. Data logs

Table 24.1: Near infrared stellar Photometry of HR8799, using apparent flux normalised to 10 pc, retrieved from the Spanish Virtual Observatory (Bayo et al., 2008).

| Filter | λ [μm] | Flux [$\text{erg/s/cm}^2/\text{\AA}$] | Ref. |
|----------|--------------------------------|--------------------------------------------|------|
| 2MASS J | 1.235 | $2.198 \pm 0.055 \times 10^{-12}$ | S06 |
| 2MASS H | 1.662 | $8.754 \pm 0.145 \times 10^{-13}$ | S06 |
| 2MASS Ks | 2.159 | $3.433 \pm 0.057 \times 10^{-13}$ | S06 |
| WISE W1 | 3.353 | $6.840 \pm 1.367 \times 10^{-14}$ | W10 |
| WISE W2 | 4.603 | $2.316 \pm 0.166 \times 10^{-14}$ | W10 |

Notes

References: S06 [Skrutskie et al. \(2006\)](#); W10 [Wright et al. \(2010\)](#);

24.B. Reprocessing of the SPHERE and GPI datasets

To resolve the known discrepancies in the H-band flux between the archival SPHERE and GPI datasets, we reprocessed each using KLIP ([Soummer et al., 2012](#); [Pueyo, 2016](#)), ANDROMEDA ([Cantalloube et al., 2015](#)) and PynPoint ([Amara & Quanz, 2012](#); [Stolker et al., 2019](#)). We optimised the choice of algorithm parameters through a series of injection/extraction tests into each dataset, as in [Nasedkin et al. \(2023\)](#). Using two different goodness-of-fit metrics on injections representative of the true companion contrast and separation, we choose the number of principal components used in the PSF subtraction in order to extract the companion spectra with minimal bias. Figures 24.1 and 24.2 shows the extracted spectra for each of HR 8799 c, d, and e, for SPHERE and GPI respectively, compared to published literature spectra from [Zurlo et al. \(2016\)](#), [Greenbaum et al. \(2018\)](#), and [Flasseur et al. \(2018\)](#). As the goodness-

Table 24.2: Photometric data for HR 8799 b and c.

| Instrument/Filter | m | Ref. |
|------------------------------|------------------|------|
| HR 8799 b | | |
| Keck/NIRC2.H | 18.05 ± 0.09 | C12 |
| Keck/NIRC2.Ks | 17.03 ± 0.08 | M10 |
| Keck/NIRC2.Ms | 16.05 ± 0.3 | G11 |
| Paranal/NACO.Lp | 15.52 ± 0.1 | C14 |
| Paranal/NACO.NB405 | 14.82 ± 0.18 | C14 |
| Paranal/SPHERE/IRDIS.B_J | 19.78 ± 0.09 | Z16 |
| Paranal/SPHERE/IRDIS.D_H23_2 | 18.08 ± 0.14 | Z16 |
| Paranal/SPHERE/IRDIS.D_H23_3 | 17.78 ± 0.1 | Z16 |
| Paranal/SPHERE/IRDIS.D_K12_1 | 17.15 ± 0.06 | Z16 |
| Paranal/SPHERE/IRDIS.D_K12_2 | 16.97 ± 0.09 | Z16 |
| Subaru/CIAO.z | 21.22 ± 0.29 | C11 |
| JWST/MIRI.F1065C | 13.54 ± 0.04 | B23 |
| JWST/MIRI.F1140C | 13.64 ± 0.07 | B23 |
| JWST/MIRI.F1550C | 13.49 ± 0.25 | B23 |
| HR 8799 c | | |
| Keck/NIRC2.H | 17.06 ± 0.13 | C12 |
| Keck/NIRC2.Ks | 16.11 ± 0.08 | M10 |
| Keck/NIRC2.Ms | 15.03 ± 0.14 | G11 |
| Paranal/NACO.Lp | 14.65 ± 0.11 | C14 |
| Paranal/NACO.NB405 | 13.97 ± 0.11 | C14 |
| Paranal/SPHERE/IRDIS.B_J | 18.6 ± 0.13 | Z16 |
| Paranal/SPHERE/IRDIS.D_H23_2 | 17.09 ± 0.12 | Z16 |
| Paranal/SPHERE/IRDIS.D_H23_3 | 16.78 ± 0.1 | Z16 |
| Paranal/SPHERE/IRDIS.D_K12_1 | 16.19 ± 0.05 | Z16 |
| Paranal/SPHERE/IRDIS.D_K12_2 | 15.86 ± 0.07 | Z16 |
| JWST/MIRI.F1065C | 12.97 ± 0.18 | B23 |
| JWST/MIRI.F1140C | 13.59 ± 0.26 | B23 |
| JWST/MIRI.F1550C | 11.88 ± 0.23 | B23 |

Notes:

References: C12 [Currie et al. \(2012\)](#); M10 [Marois et al. \(2010\)](#); G11 [Galicher et al. \(2011\)](#); C14 [Currie et al. \(2014\)](#); Z16 [Zurlo et al. \(2016\)](#); C11 [Currie et al. \(2011\)](#); B23 [Boccaletti et al. \(2023\)](#).

Table 24.3: Photometric data for HR 8799 d and e.

| Instrument/Filter | m | Ref. |
|------------------------------|------------------|------|
| HR 8799 d | | |
| Keck/NIRC2.H | 16.71 ± 0.24 | C12 |
| Keck/NIRC2.Ks | 16.09 ± 0.12 | M10 |
| Keck/NIRC2.Ms | 14.65 ± 0.35 | G11 |
| Paranal/NACO.Lp | 14.55 ± 0.14 | C14 |
| Paranal/NACO.NB405 | 13.87 ± 0.15 | C14 |
| Paranal/SPHERE/IRDIS.B_J | 18.59 ± 0.37 | Z16 |
| Paranal/SPHERE/IRDIS.D_H23_2 | 17.02 ± 0.17 | Z16 |
| Paranal/SPHERE/IRDIS.D_H23_3 | 16.85 ± 0.16 | Z16 |
| Paranal/SPHERE/IRDIS.D_K12_1 | 16.2 ± 0.07 | Z16 |
| Paranal/SPHERE/IRDIS.D_K12_2 | 15.84 ± 0.1 | Z16 |
| JWST/MIRI.F1065C | 12.98 ± 0.14 | B23 |
| JWST/MIRI.F1140C | 12.98 ± 0.17 | B23 |
| JWST/MIRI.F1550C | 11.88 ± 0.23 | B23 |
| HR 8799 e | | |
| Keck/NIRC2.Ks | 15.91 ± 0.22 | C12 |
| Keck/NACO.Lp | 14.49 ± 0.21 | M10 |
| Keck/NACO.NB405 | 13.72 ± 0.2 | C14 |
| Paranal/SPHERE/IRDIS.B_J | 18.4 ± 0.21 | Z16 |
| Paranal/SPHERE/IRDIS.D_H23_2 | 16.91 ± 0.2 | Z16 |
| Paranal/SPHERE/IRDIS.D_H23_3 | 16.68 ± 0.21 | Z16 |
| Paranal/SPHERE/IRDIS.D_K12_1 | 16.12 ± 0.1 | Z16 |
| Paranal/SPHERE/IRDIS.D_K12_2 | 15.82 ± 0.11 | Z16 |
| JWST/MIRI.F1065C | 12.52 ± 0.26 | B23 |
| JWST/MIRI.F1140C | 12.52 ± 0.23 | B23 |
| JWST/MIRI.F1550C | 11.01 ± 0.42 | B23 |

Notes:

References: C12 [Currie et al. \(2012\)](#); M10 [Marois et al. \(2010\)](#); G11 [Galicher et al. \(2011\)](#); C14 [Currie et al. \(2014\)](#); Z16 [Zurlo et al. \(2016\)](#); C11 [Currie et al. \(2011\)](#); B23 [Boccaletti et al. \(2023\)](#).

of-fit metrics favoured the KLIP extractions, we used these as the basis of our retrieval analysis.

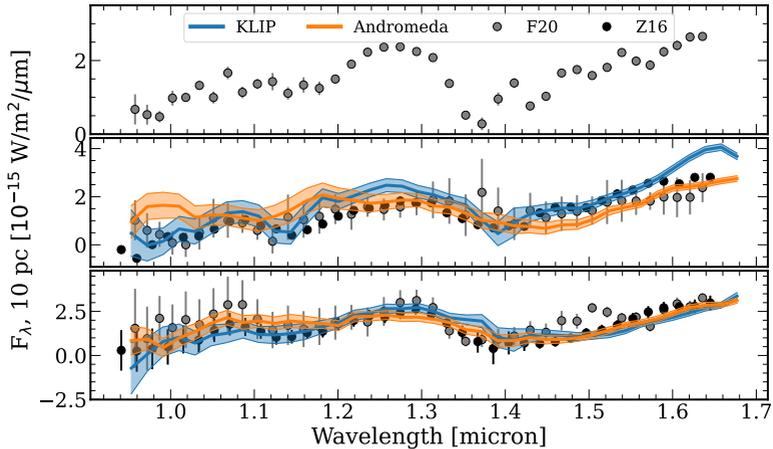


Figure 24.1: KLIP and ANDROMEDA extractions from SPHERE for HR 8799 c, d, and e compared to the spectra published in [Zurlo et al. \(2016\)](#) and [Flasseur et al. \(2018\)](#).

24.C. Retrieval validation

Extensive validation of the pRT retrieval module was performed as part of this work. Following updates described in [Nasedkin et al. \(2024\)](#), we verified that the results of [Mollière et al. \(2020\)](#) could be reproduced. We independently tested updates to the c-k mixing implementation, the adaptive mesh refinement implementation, updated opacity sources for H₂O and CO, bug fixes for convergence on multiple scattering in the clouds, the inclusion of photometric data, the inclusion of scaling factors on the SPHERE and GPI datasets, including or excluding the GPI K-band spectra, updates to each of the SPHERE, GPI, and GRAVITY datasets, different prior widths and the number of live points used in the retrieval. All of the posterior distributions were fully consistent to within 2σ with most falling well within 1σ of the published results, apart from the inclusion of new epochs of GRAVITY data, which led to a significantly higher retrieved metallicity ($[M/H]=1.1 \pm 0.32$) and an f_{sed} of 5 ± 2.6 .

We verified several model assumptions through retrievals that only include the GRAVITY datasets, or the GRAVITY data and photometry. We find

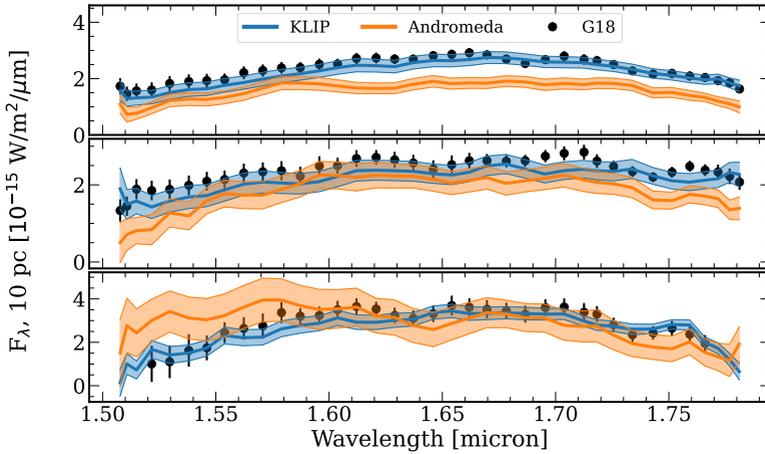


Figure 24.2: KLIP and ANDROMEDA extractions from GPI for HR 8799 c, d, and e compared to the spectra published in [Greenbaum et al. \(2018\)](#)

that the GRAVITY data alone could not distinguish between clear and cloudy models ($\Delta \log_{10} \mathcal{Z} < 1$), while cloudy models were strongly favoured once the broad wavelength coverage of the photometry was included ($\Delta \log_{10} \mathcal{Z} > 10$).

Models that use the dynamical mass estimates as priors for calculating the surface gravity were marginally favoured over those that freely retrieve $\log g$ and R_{pl} , but that the posteriors parameter distributions were generally consistent, with T_{int} and R_{pl} showing the greatest discrepancy. Using the dynamical mass as a prior and setting a Gaussian prior on the radius led to more reasonable estimates of the radius of HR 8799 e (0.97 ± 0.04) compared to the free retrieval (0.79 ± 0.05). However, the composition of the planet and the degree of cloudiness did not vary significantly between the two models.

Using the full dataset for HR 8799 b, we verified that retrievals including scattering clouds are strongly favoured over those without scattering ($\Delta \log_{10} \mathcal{Z} > 10$). Without scattering, both the temperature and composition ($[M/H]$ and C/O) are significantly discrepant from retrievals that include scattering clouds.

24.D. Using the Hansen distribution with EDDYSED

The EDDYSED cloud model from [Ackerman & Marley \(2001\)](#) is implemented in pRT, and is the most physically motivated model incorporated to date. Typically, it assumes a log-normal particle size distribution, where the

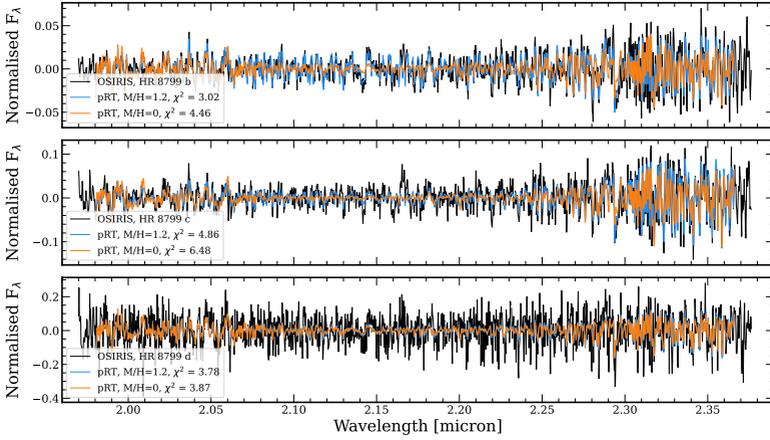


Figure 24.3: Comparison of best-fit disequilibrium models to OSIRIS data from [Ruffio et al. \(2021\)](#). From top to bottom is HR 8799 b, c, and d. In blue are the best-fit disequilibrium models, with the spectra generated using high-resolution line-by-line opacities, before being convolved, binned, and normalised for comparison. In orange is the same, but with the metallicity set to 0.

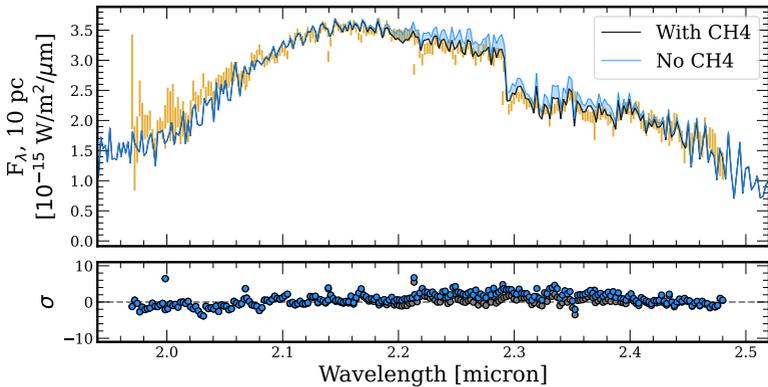


Figure 24.4: Comparison of best-fit disequilibrium models (black) of HR 8799 c to the data, with residuals shown in the bottom panel. In blue are the same spectra, but without opacity contributions from CH_4 .

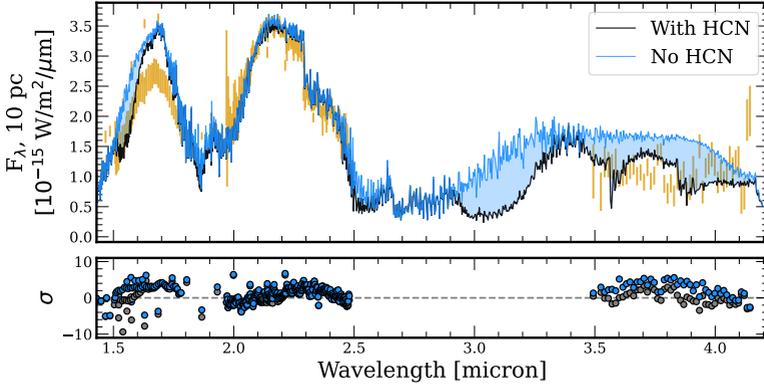


Figure 24.5: Comparison of best-fit disequilibrium models (black) of HR 8799 c to the data, with residuals shown in the bottom panel. In blue are the same spectra, but without opacity contributions from HCN.

geometric particle radius will vary throughout the atmosphere as a function of the vertical diffusion coefficient K_{zz} and the sedimentation fraction f_{sed} . Here, we substitute the log-normal particle size distribution with the Hansen distribution, originally introduced in Hansen (1971), and rederive the calculation for the particle radius as a function of K_{zz} and f_{sed} .

We begin with a review of the EDDYSED model: the distribution of the number of particles as a function of particle radius, $n(r)$ is approximated as a log-normal distribution with width σ_g and characteristic geometric radius r_g .

$$n(r) = \frac{N}{r \sqrt{2\pi} \log \sigma_g} \exp\left(-\frac{\log^2(r/r_g)}{2 \log^2 \sigma_g}\right), \quad (24.1)$$

N is the total number of cloud particles.

The goal of the EDDYSED model is to calculate r_g for each layer in the atmosphere, given K_{zz} and f_{sed} . It balances the upwards vertical mixing, parameterised by K_{zz} and the particle settling velocity, v_f

$$v_f = w_* \left(\frac{r}{r_w}\right)^\alpha. \quad (24.2)$$

Here w_* is the convective velocity scale. We note that $r_w \neq r_g$. r_w is the radius at which the particle settling velocity equals the convective velocity scale:

$$w_* = \frac{K_{zz}}{L}, \quad (24.3)$$

where L is the convective mixing length. Since w_* is known, and v_f can be found analytically as in [Ackerman & Marley \(2001\)](#); [Podolak \(2003\)](#), a linear fit can be used to find both α and r_w .

With both of these quantities known, we follow [AM01](#) and define f_{sed} as:

$$f_{\text{sed}} = \frac{\int_0^\infty r^{3+\alpha} n(r) dr}{r_w^\alpha \int_0^\infty r^3 n(r) dr} \quad (24.4)$$

For the log-normal distribution, one finds:

$$\int_0^\infty r^\beta n(r) dr = N r_g^\beta \exp\left(\frac{1}{2} \beta^2 \log^2 \sigma_g\right) \quad (24.5)$$

Which we can then use to solve for r_g :

$$r_g = r_w f_{\text{sed}}^{1/\alpha} \exp\left(-\frac{\alpha + 6}{2} \log^2 \sigma_g\right) \quad (24.6)$$

In order to use the Hansen distribution, we must recalculate the total number of particles N , and integrate the distribution for f_{sed} . We note here that the Hansen distribution is parameterised by the effective radius, \bar{r} , rather than the geometric mean radius. In this derivation we do not correct for this difference in definition, as both act as nuisance parameters in the context of an atmospheric retrieval.

We start by giving the Hansen distribution in full:

$$n_{\text{Hansen}}(r) = \frac{N (\bar{r} v_e)^{(2v_e-1)/v_e}}{\Gamma[(1-2v_e)/v_e]} r^{(1-3v_e)/v_e} \exp\left(-\frac{r}{\bar{r} v_e}\right) \quad (24.7)$$

In [Hansen \(1971\)](#), the authors use the parameters a and b to denote the mean effective radius and effective variance, which we write as \bar{r} and v_e respectively. These differ from the simple mean radius and variance by weighting them by the particle area, as the cloud particle scatters an amount of light proportional to its area. Thus:

$$\bar{r} = \frac{\int_0^\infty r \pi r^2 n(r) dr}{\int_0^\infty \pi r^2 n(r) dr} \quad (24.8)$$

and

$$v_e = \frac{\int_0^\infty (r - \bar{r})^2 r^2 n(r) dr}{\bar{r}^2 \int_0^\infty \pi r^2 n(r) dr} \quad (24.9)$$

As in EDDYSED, we fit for the settling velocity, which will provide us with α and r_w , which we can use to find f_{sed} , as in 24.4. However, we must now integrate the Hansen distribution. We find that:

$$\int_0^\infty r^\beta n_{\text{Hansen}}(r) dr = \frac{v_e^\beta (v_e \beta + 2v_e + 1) \left(\frac{1}{\bar{r}}\right)^{-\beta} \Gamma\left(\beta + 1 + \frac{1}{v_e}\right)}{\left(-v_e + v_e^{\beta+3} + 1\right) \Gamma\left(1 + \frac{1}{v_e}\right)} \quad (24.10)$$

We can then use Eqns. 24.4 and 24.10 to solve for \bar{r} :

$$\bar{r} = \left(\frac{f_{\text{sed}} r_w^\alpha v_e^{-\alpha} (v_e^{3+\alpha} - v_e + 1) \Gamma\left(1 + \frac{1}{v_e}\right)^\alpha}{(v_e \alpha + 2v_e + 1) \Gamma\left(\alpha + 1 + \frac{1}{v_e}\right)} \right)^{\frac{1}{\alpha}}. \quad (24.11)$$

Thus for a given K_{zz} , f_{sed} , and v_e , we can find the effective particle radius for every layer in the atmosphere.

However, in order to compute the cloud opacity, we still require the total particle count. For a volume mixing ratio of a given species, χ_i , we can integrate $n(r)$ to find N :

$$N = \frac{\chi_i}{(\bar{r}^3 v_e - 1)(2v_e - 1)} \quad (24.12)$$

24.E. Complete retrieval results

We include in the text abridged tables that present key parameters of interest. The complete set of inferred parameters for every retrieval is available online.

Legend: Chemistry/Profile/Clouds/Data/Info.

- Chemistry: (D)isequilibrium or (F)ree.
- Profile: (M)olliere, (Z)hang, (G)uillot or (S)pline(NNnodes).
- Clouds: Clear (CLR); (f)ree or (eq)uilibrium condensation location, (species)_(cd/am)_(P)atchy_(h)ansen. ‘*’ indicates f_{sed} was retrieved independently for each cloud species.
- Data: -(not included) or (only included). O indicates OSIRIS data was used in place of GRAVITY data.
- Info: ‘?’ indicates not included. ‘mr’ indicates mass and radius were used as parameters instead of $\log g$ and radius.

In the text, models will be referred to as *planet.group.index*.

Table 24.4: Spectroscopic Observation Log.

| Planet | Instrument | Date | λ | $\lambda/\Delta\lambda$ | APA | Seeing | Airmass | DIT | NEXP | Ref. |
|---------|------------|------------|------------|-------------------------|----------------|-----------------------|---------|----------------|--------|------|
| | | | | [μm] | [$^{\circ}$] | [as], Med. | Min. | [s] | | |
| b | GRAVITY | 2019-11-11 | 2.0–2.4 | 500 | ... | 0.98 | 1.54 | 8x100 | 3 | ... |
| | GRAVITY | 2021-08-26 | 2.0–2.4 | 500 | ... | 0.88 | 1.48 | 8x100 | 3 | ... |
| | GRAVITY | 2021-08-27 | 2.0–2.4 | 500 | ... | 0.86 | 1.45 | 8x100 | 2 | ... |
| | OSIRIS | 2009-07-22 | 2.0–2.4 | 60 | ... | ... | 1.0 | 900 | 30 | B11 |
| | OSIRIS | 2009-07-23 | 1.5–1.8 | 60 | ... | ... | 1.0 | 900 | 30 | B11 |
| | OSIRIS | 2009-07-30 | 1.5–1.8 | 60 | ... | ... | 1.0 | 900 | 30 | B11 |
| | OSIRIS | 2010-07-11 | 2.0–2.4 | 60 | ... | ... | 1.0 | 900 | 30 | B11 |
| c | OSIRIS | 2010-07-13 | 1.5–1.8 | 60 | ... | ... | 1.0 | 900 | 30 | B11 |
| | GRAVITY | 2019-11-11 | 2.0–2.4 | 500 | ... | 1.01 | 1.62 | 8x100 | 3 | ... |
| | GRAVITY | 2021-08-26 | 2.0–2.4 | 500 | ... | 1.04 | 1.50 | 8x100 | 3 | ... |
| | GRAVITY | 2021-08-27 | 2.0–2.4 | 500 | ... | 1.00 | 1.45 | 8x100 | 2 | ... |
| | GRAVITY | 2022-08-19 | 2.0–2.4 | 500 | ... | 0.64 | 1.54 | 4x100 | 10 | ... |
| | GRAVITY | 2023-07-02 | 2.0–2.4 | 500 | ... | 0.75 | 1.45 | 4x100 | 6 | ... |
| | SPHERE | 2015-07-04 | 0.9–1.6 | 30 | 16.4 | 1.43 | 1.44 | ... | 46 | F20 |
| | CHARIS | 2018-09-01 | 1.2–2.4 | 19 | 202.24 | 0.47 | 1.0 | 20 | 1201 | W22 |
| | CHARIS | 2018-09-02 | 1.2–2.4 | 19 | 206.55 | 0.42 | 1.0 | 20 | 1253 | W22 |
| | GPI | 2016-09-19 | 1.5–1.8 | 45 | 20.93 | 0.97 | 1.61 | 60 | 60 | G18 |
| ALES | 2019-09-18 | 2.8–4.2 | 35 | 85.64 | 0.8–1.1 | 1.02 | 3.934 | 1300 | D22 | |
| OSIRIS | 2010-2011 | 1.97–2.38 | 4000 | ... | ... | ... | 600 | 33 | K13 | |
| d | GRAVITY | 2019-11-09 | 2.0–2.4 | 500 | ... | 0.85 | 1.63 | 8x60 | 4 | ... |
| | GRAVITY | 2019-11-11 | 2.0–2.4 | 500 | ... | 1.14 | 1.70 | 8x100 | 3 | ... |
| | GRAVITY | 2021-08-26 | 2.0–2.4 | 500 | ... | 0.96 | 1.54 | 8x100 | 3 | ... |
| | GRAVITY | 2021-08-27 | 2.0–2.4 | 500 | ... | 1.18 | 1.48 | 8x100 | 2 | ... |
| | GRAVITY | 2022-09-15 | 2.0–2.4 | 500 | ... | 0.67 | 1.54 | 4x100 | 12 | ... |
| | SPHERE | 2014-08-12 | 0.9–1.6 | 30 | 29.65 | 0.87 | 1.43 | 100 | 32 | Z16 |
| | SPHERE | 2014-08-12 | 0.9–1.6 | 30 | 15.37 | 0.87 | 1.43 | 60 | 48 | Z16 |
| | CHARIS | 2018-09-01 | 1.2–2.4 | 19 | 202.24 | 0.47 | 1.0 | 20 | 1201 | W22 |
| | CHARIS | 2018-09-02 | 1.2–2.4 | 19 | 206.55 | 0.42 | 1.0 | 20 | 1253 | W22 |
| | GPI | 2016-09-19 | 1.5–1.8 | 45 | 20.93 | 0.97 | 1.61 | 60 | 60 | G18 |
| | ALES | 2019-09-18 | 2.8–4.2 | 35 | 85.64 | 0.8–1.1 | 1.02 | 3.934 | 1300 | D22 |
| | e | GRAVITY | 2018-08-28 | 2.0–2.4 | 500 | ... | 0.67 | 1.44 | 10x100 | 7 |
| GRAVITY | | 2019-11-09 | 2.0–2.4 | 500 | ... | 0.84 | 1.55 | 8x60 | 3 | M20 |
| GRAVITY | | 2019-11-11 | 2.0–2.4 | 500 | ... | 1.15 | 1.47 | 8x100 | 3 | M20 |
| GRAVITY | | 2021-08-26 | 2.0–2.4 | 500 | ... | 0.66 | 1.50 | 8x100 | 2 | ... |
| GRAVITY | | 2021-08-27 | 2.0–2.4 | 500 | ... | 0.67 | 1.47 | 8x100 | 2 | ... |
| GRAVITY | | 2021-09-27 | 2.0–2.4 | 500 | ... | 0.82 | 1.55 | 4x100 | 6 | ... |
| GRAVITY | | 2023-07-02 | 2.0–2.4 | 500 | ... | 0.75 | 1.46 | 4x100 | 6 | ... |
| SPHERE | | 2014-08-12 | 0.9–1.6 | 30 | 29.65 | 0.87 | 1.43 | 100 | 32 | Z16 |
| SPHERE | | 2014-08-12 | 0.9–1.6 | 45 | 15.37 | 0.87 | 1.43 | 60 | 48 | Z16 |
| CHARIS | | 2018-09-01 | 1.2–2.4 | 19 | 202.24 | 0.47 | 1.0 | 20 | 1201 | W22 |
| CHARIS | | 2018-09-02 | 1.2–2.4 | 19 | 206.55 | 0.42 | 1.0 | 20 | 1253 | W22 |
| GPI | | 2016-09-19 | 1.5–1.8 | 45 | 20.93 | 0.97 | 1.61 | 60 | 60 | G18 |
| ALES | | 2019-09-18 | 2.8–4.2 | 35 | 85.64 | 0.8–1.1 | 1.02 | 3.934 | 1300 | D22 |

Notes

References: B11: [Barman et al. \(2011\)](#); K13: [Konopacky et al. \(2013\)](#); Z16: [Zurlo et al. \(2016\)](#); G18: [Greenbaum et al. \(2018\)](#); G19: [Gravity Collaboration et al. \(2019\)](#); F20: [Flasseur et al. \(2020\)](#); M20: [Mollière et al. \(2020\)](#); W22: [Wang et al. \(2022\)](#)

IV Characterisation of the HR 8799 planets

Table 24.5: Abridged retrieval results HR 8799 b

| Index | Model | B.F. | χ^2/ν | T_{Eff} [K] | $\log g$ [cgs] | [M/H] | C/O | Radius [R_{Jup}] | $\log L/L_{\odot}$ | Mass [M_{Jup}] |
|---------|--------------------------|------|--------------|------------------------------------|----------------------------------------|----------------------------------------|------------------------------------------|----------------------------------------|-----------------------------------------|-------------------------------------|
| b.A.0 | F/Z/eq*FeMg_am/ALL/-CH4/ | 0 | 1.46 | 948 ⁺¹¹ ₋₁₄ | 4.11 ^{+0.03} _{-0.04} | 0.94 ^{+0.08} _{-0.08} | 0.78 ^{+0.03} _{-0.04} | 1.1 ^{+0.03} _{-0.03} | -5.06 ^{+0.04} _{-0.04} | 6 ^{+0.3} _{-0.4} |
| b.AB.1 | F/Z/eq*FeMg_am/ALL/ | 0 | 1.47 | 942 ⁺¹² ₋₁₃ | 4.1 ^{+0.03} _{-0.04} | 0.96 ^{+0.08} _{-0.08} | 0.78 ^{+0.03} _{-0.04} | 1.11 ^{+0.03} _{-0.03} | -5.08 ^{+0.04} _{-0.04} | 6 ^{+0.3} _{-0.3} |
| b.A.2 | F/Z/eq*FeMg_am/ALL/-H2S/ | 0 | 1.49 | 958 ⁺¹³ ₋₁₅ | 4.1 ^{+0.03} _{-0.04} | 1.1 ^{+0.2} _{-0.1} | 0.83 ^{+0.03} _{-0.04} | 1.1 ^{+0.03} _{-0.03} | -5.05 ^{+0.04} _{-0.04} | 5.9 ^{+0.4} _{-0.4} |
| b.A.3 | F/S1/eqMg_am/ALL/ | -2 | 1.48 | 977 ⁺¹¹ ₋₁₀ | 4.18 ^{+0.04} _{-0.04} | 1.25 ^{+0.1} _{-0.1} | 0.89 ^{+0.02} _{-0.03} | 1.01 ^{+0.02} _{-0.02} | -5.07 ^{+0.03} _{-0.03} | 6 ^{+0.4} _{-0.4} |
| b.A.4 | F/S2/eqMg_am/ALL/ | -4 | 1.50 | 970 ⁺¹⁴ ₋₁₃ | 4.17 ^{+0.04} _{-0.04} | 1.2 ^{+0.1} _{-0.1} | 0.87 ^{+0.03} _{-0.03} | 1.02 ^{+0.02} _{-0.02} | -5.08 ^{+0.04} _{-0.04} | 5.9 ^{+0.4} _{-0.4} |
| b.A.5 | F/S1/CLR/ALL/ | -4 | 1.53 | 985 ⁺¹¹ ₋₁₁ | 4.18 ^{+0.04} _{-0.04} | 1.43 ^{+0.08} _{-0.09} | 0.92 ^{+0.02} _{-0.02} | 1.01 ^{+0.02} _{-0.02} | -5.06 ^{+0.03} _{-0.04} | 6 ^{+0.5} _{-0.5} |
| b.AB.6 | D/M/eqNa_P/ALL/ | -4 | 1.52 | 931 ⁺²¹ ₋₂₃ | 4.12 ^{+0.05} _{-0.05} | 1.16 ^{+0.08} _{-0.08} | 0.73 ^{+0.02} _{-0.02} | 1.09 ^{+0.06} _{-0.04} | -5.07 ^{+0.06} _{-0.06} | 6 ^{+0.3} _{-0.3} |
| b.AB.7 | D/M/eqNa/ALL/ | -4 | 1.50 | 966 ⁺¹⁷ ₋₁₉ | 4.19 ^{+0.04} _{-0.04} | 0.95 ^{+0.07} _{-0.07} | 0.78 ^{+0.02} _{-0.03} | 1.0 ^{+0.05} _{-0.05} | -5.05 ^{+0.05} _{-0.05} | 6 ^{+0.3} _{-0.3} |
| b.AB.8 | F/Z/*FeMg_am/ALL/ | -4 | 1.48 | 944 ⁺¹⁰ ₋₁₀ | 4.15 ^{+0.03} _{-0.03} | 0.97 ^{+0.08} _{-0.08} | 0.78 ^{+0.03} _{-0.04} | 1.04 ^{+0.02} _{-0.02} | -5.11 ^{+0.03} _{-0.03} | 5.9 ^{+0.4} _{-0.3} |
| b.A.9 | F/S3/CLR/ALL/ | -4 | 1.49 | 962 ⁺¹³ ₋₁₂ | 4.17 ^{+0.04} _{-0.04} | 1.1 ^{+0.1} _{-0.1} | 0.85 ^{+0.03} _{-0.04} | 1.02 ^{+0.02} _{-0.02} | -5.09 ^{+0.04} _{-0.04} | 6 ^{+0.4} _{-0.5} |
| b.A.10 | F/S2/CLR/ALL/ | -4 | 1.52 | 962 ⁺¹⁵ ₋₁₄ | 4.17 ^{+0.04} _{-0.04} | 1.2 ^{+0.1} _{-0.1} | 0.87 ^{+0.03} _{-0.04} | 1.02 ^{+0.03} _{-0.03} | -5.09 ^{+0.04} _{-0.04} | 5.9 ^{+0.5} _{-0.5} |
| b.AB.11 | F/Z/*FeMg_cd/ALL/ | -4 | 1.48 | 942 ⁺¹¹ ₋₁₁ | 4.15 ^{+0.03} _{-0.03} | 0.94 ^{+0.08} _{-0.08} | 0.77 ^{+0.04} _{-0.04} | 1.05 ^{+0.02} _{-0.02} | -5.11 ^{+0.03} _{-0.03} | 5.9 ^{+0.3} _{-0.3} |
| b.A.12 | F/S3/eqMg_am/ALL/ | -4 | 1.51 | 965 ⁺¹² ₋₁₃ | 4.17 ^{+0.04} _{-0.04} | 1.1 ^{+0.1} _{-0.1} | 0.85 ^{+0.03} _{-0.04} | 1.02 ^{+0.02} _{-0.02} | -5.09 ^{+0.04} _{-0.03} | 6 ^{+0.4} _{-0.4} |
| b.AB.13 | D/M/eq*FeMg_am/ALL/ | -6 | 1.54 | 944 ⁺¹¹ ₋₁₀ | 4.06 ^{+0.03} _{-0.03} | 1.11 ^{+0.08} _{-0.08} | 0.72 ^{+0.02} _{-0.02} | 1.17 ^{+0.03} _{-0.03} | -5.03 ^{+0.04} _{-0.04} | 6 ^{+0.3} _{-0.3} |
| b.A.14 | F/S4/CLR/ALL/ | -6 | 1.49 | 966 ⁺¹³ ₋₁₂ | 4.17 ^{+0.04} _{-0.04} | 1.15 ^{+0.1} _{-0.1} | 0.86 ^{+0.03} _{-0.04} | 1.02 ^{+0.02} _{-0.02} | -5.09 ^{+0.04} _{-0.04} | 6 ^{+0.4} _{-0.5} |
| b.A.15 | F/S4/eqMg_am/ALL/ | -6 | 1.52 | 968 ⁺¹³ ₋₁₂ | 4.17 ^{+0.04} _{-0.04} | 1.15 ^{+0.1} _{-0.09} | 0.86 ^{+0.03} _{-0.03} | 1.02 ^{+0.02} _{-0.02} | -5.08 ^{+0.04} _{-0.04} | 6 ^{+0.4} _{-0.4} |
| b.AB.16 | D/M/eq*FeMg_cd/ALL/ | -6 | 1.55 | 911 ⁺¹⁴ ₋₁₃ | 4.05 ^{+0.04} _{-0.04} | 1.16 ^{+0.08} _{-0.08} | 0.73 ^{+0.02} _{-0.02} | 1.18 ^{+0.04} _{-0.04} | -5.09 ^{+0.04} _{-0.04} | 6 ^{+0.3} _{-0.3} |
| b.AB.17 | D/M/eq*FeMg_am_P/ALL/ | -7 | 1.55 | 941 ⁺¹² ₋₁₂ | 4.07 ^{+0.03} _{-0.03} | 1.08 ^{+0.08} _{-0.09} | 0.73 ^{+0.02} _{-0.02} | 1.16 ^{+0.03} _{-0.03} | -5.05 ^{+0.03} _{-0.03} | 6 ^{+0.3} _{-0.3} |
| b.AB.18 | D/Z/eq*FeMg_am/ALL/ | -7 | 1.57 | 934 ⁺¹⁰ ₋₁₀ | 4.06 ^{+0.03} _{-0.03} | 1.06 ^{+0.06} _{-0.07} | 0.73 ^{+0.02} _{-0.02} | 1.17 ^{+0.03} _{-0.03} | -5.05 ^{+0.03} _{-0.03} | 6 ^{+0.3} _{-0.3} |
| b.A.19 | F/S5/CLR/ALL/ | -7 | 1.50 | 955 ⁺¹³ ₋₁₄ | 4.16 ^{+0.04} _{-0.04} | 1.1 ^{+0.1} _{-0.1} | 0.84 ^{+0.04} _{-0.05} | 1.03 ^{+0.02} _{-0.02} | -5.09 ^{+0.04} _{-0.04} | 6 ^{+0.4} _{-0.5} |
| b.A.20 | F/S5/eqMg_am/ALL/ | -7 | 1.51 | 958 ⁺¹³ ₋₁₄ | 4.17 ^{+0.03} _{-0.04} | 1.09 ^{+0.1} _{-0.1} | 0.84 ^{+0.03} _{-0.04} | 1.03 ^{+0.02} _{-0.02} | -5.09 ^{+0.04} _{-0.03} | 6 ^{+0.4} _{-0.4} |
| b.AB.21 | D/M/eqKCl/ALL/ | -8 | 1.56 | 906 ⁺¹² ₋₁₃ | 4.08 ^{+0.03} _{-0.03} | 1.22 ^{+0.08} _{-0.08} | 0.7 ^{+0.02} _{-0.02} | 1.13 ^{+0.03} _{-0.03} | -5.13 ^{+0.04} _{-0.04} | 6 ^{+0.2} _{-0.3} |
| b.A.22 | F/S6/CLR/ALL/ | -9 | 1.50 | 960 ⁺¹³ ₋₁₃ | 4.17 ^{+0.04} _{-0.04} | 1 ⁺⁰ _{-0.1} | 0.83 ^{+0.04} _{-0.05} | 1.03 ^{+0.02} _{-0.02} | -5.09 ^{+0.04} _{-0.04} | 6 ^{+0.4} _{-0.4} |
| b.AB.23 | D/Z/*FeMg_am/ALL/ | -9 | 1.53 | 954 ⁺⁴⁵ ₋₅₉ | 4.17 ^{+0.03} _{-0.03} | 1.22 ^{+0.06} _{-0.06} | 0.76 ^{+0.01} _{-0.02} | 1.02 ^{+0.02} _{-0.02} | -5.1 ^{+0.08} _{-0.1} | 5.9 ^{+0.3} _{-0.3} |
| b.A.24 | F/S6/eqMg_am/ALL/ | -9 | 1.49 | 928 ⁺¹⁷ ₋₁₅ | 4.01 ^{+0.04} _{-0.04} | 0.82 ^{+0.07} _{-0.07} | 0.63 ^{+0.05} _{-0.05} | 1.21 ^{+0.04} _{-0.04} | -5.03 ^{+0.04} _{-0.04} | 5.8 ^{+0.4} _{-0.4} |
| b.AB.25 | D/S4/eq*FeMg_am/ALL/ | -10 | 1.53 | 867 ⁺⁷⁵ ₋₄₅ | 4.03 ^{+0.03} _{-0.03} | 1.13 ^{+0.07} _{-0.07} | 0.67 ^{+0.03} _{-0.04} | 1.21 ^{+0.03} _{-0.03} | -5.15 ^{+0.1} _{-0.1} | 6 ^{+0.3} _{-0.3} |
| b.A.26 | F/S7/CLR/ALL/ | -10 | 1.52 | 964 ⁺¹³ ₋₁₅ | 4.17 ^{+0.04} _{-0.04} | 1.1 ^{+0.1} _{-0.1} | 0.85 ^{+0.04} _{-0.04} | 1.02 ^{+0.02} _{-0.02} | -5.09 ^{+0.04} _{-0.04} | 6 ^{+0.4} _{-0.4} |
| b.AB.27 | D/M/tna/ALL/ | -11 | 1.56 | 918 ⁺⁴⁶ ₋₅₇ | 4.13 ^{+0.03} _{-0.03} | 1.18 ^{+0.06} _{-0.06} | 0.7 ^{+0.02} _{-0.02} | 1.07 ^{+0.03} _{-0.02} | -5.1 ^{+0.09} _{-0.1} | 6 ^{+0.2} _{-0.3} |
| b.A.28 | F/S8/CLR/ALL/ | -11 | 1.52 | 955 ⁺¹⁴ ₋₁₄ | 4.16 ^{+0.04} _{-0.04} | 0.98 ^{+0.1} _{-0.1} | 0.79 ^{+0.05} _{-0.1} | 1.03 ^{+0.03} _{-0.03} | -5.1 ^{+0.04} _{-0.04} | 6 ^{+0.4} _{-0.4} |
| b.A.29 | F/S9/CLR/ALL/ | -12 | 1.52 | 965 ⁺¹³ ₋₁₂ | 4.18 ^{+0.04} _{-0.04} | 0.95 ^{+0.1} _{-0.09} | 0.57 ^{+0.2} _{-0.1} | 1.02 ^{+0.02} _{-0.02} | -5.09 ^{+0.04} _{-0.04} | 6.1 ^{+0.4} _{-0.4} |
| b.A.30 | F/S0/CLR/ALL/ | -21 | 1.84 | 1020 ⁺¹² ₋₁₂ | 4.24 ^{+0.04} _{-0.04} | 1.92 ^{+0.04} _{-0.06} | 0.97 ^{+0.005} _{-0.005} | 0.93 ^{+0.02} _{-0.03} | -5.08 ^{+0.04} _{-0.04} | 5.8 ^{+0.5} _{-0.5} |
| b.A.31 | F/S0/eqMg_am/ALL/ | -21 | 1.83 | 1022 ⁺¹³ ₋₁₁ | 4.24 ^{+0.04} _{-0.04} | 1.93 ^{+0.04} _{-0.06} | 0.97 ^{+0.004} _{-0.005} | 0.93 ^{+0.02} _{-0.02} | -5.08 ^{+0.04} _{-0.04} | 5.8 ^{+0.5} _{-0.5} |
| b.A.32 | D/Z/eq*FeMg_am/Gr/ | -948 | 0.74 | 891 ⁺³⁰ ₋₂₇ | 4.18 ^{+0.07} _{-0.07} | 1.4 ^{+0.3} _{-0.4} | 0.42 ^{+0.1} _{-0.2} | 0.99 ^{+0.08} _{-0.08} | -5.3 ^{+0.09} _{-0.1} | 5.8 ^{+0.3} _{-0.3} |

Notes

All values presented are the median values from the fits, with uncertainties given as the $\pm 34.1\%$ percentiles.

Table 24.6: Abridged retrieval results HR 8799 c

| Index | Model | B.F. | χ^2/ν | T_{Eff} [K] | $\log g$ [cgs] | [M/H] | C/O | Radius [R_{Jup}] | $\log L/L_{\odot}$ | Mass [M_{Jup}] |
|---------|--------------------------|-------|--------------|------------------------------------|----------------------------------------|----------------------------------------|------------------------------------------|----------------------------------------|-----------------------------------------|-------------------------------------|
| c.A.0 | F/Z/eq*FeMg_am/O-Gr/mr | 0 | 1.72 | 1145 ⁺¹⁴ ₋₁₅ | 4.27 ^{+0.03} _{-0.04} | 0.95 ^{+0.1} _{-0.1} | 0.6 ^{+0.09} _{-0.1} | 1.06 ^{+0.03} _{-0.03} | -4.75 ^{+0.03} _{-0.03} | 8 ^{+0.5} _{-0.5} |
| c.A.1 | D/M/eq*FeMg_cd/O-Gr/mr | -1 | 1.83 | 1207 ⁺¹¹ ₋₁₀ | 4.34 ^{+0.03} _{-0.03} | 1.89 ^{+0.06} _{-0.06} | 0.87 ^{+0.005} _{-0.006} | 0.99 ^{+0.02} _{-0.02} | -4.73 ^{+0.03} _{-0.03} | 8.2 ^{+0.5} _{-0.5} |
| c.AB.2 | F/Z/eq*FeMg_am/ALL/mr | -430 | 2.88 | 1158 ⁺¹² ₋₁₂ | 4.26 ^{+0.02} _{-0.03} | 1.27 ^{+0.06} _{-0.06} | 0.66 ^{+0.01} _{-0.01} | 1.1 ^{+0.01} _{-0.01} | -4.71 ^{+0.02} _{-0.02} | 8.5 ^{+0.4} _{-0.4} |
| c.AB.3 | F/G/*FeMg_am/ALL/mr | -432 | 2.87 | 1173 ⁺¹⁸ ₋₈ | 4.26 ^{+0.02} _{-0.02} | 1.3 ^{+0.06} _{-0.06} | 0.67 ^{+0.01} _{-0.01} | 1.09 ^{+0.01} _{-0.01} | -4.71 ^{+0.03} _{-0.02} | 8.3 ^{+0.4} _{-0.4} |
| c.AB.4 | F/Z/*FeMg_am/ALL/mr | -435 | 2.91 | 1173 ⁺⁵ ₋₆ | 4.28 ^{+0.02} _{-0.02} | 1.27 ^{+0.05} _{-0.07} | 0.67 ^{+0.01} _{-0.01} | 1.07 ^{+0.01} _{-0.01} | -4.71 ^{+0.01} _{-0.01} | 8.5 ^{+0.4} _{-0.4} |
| c.AB.5 | D/M/eq*FeMg_cd/ALL/mr | -443 | 2.93 | 1057 ⁺¹⁴ ₋₁₃ | 4.05 ^{+0.02} _{-0.02} | 1.05 ^{+0.04} _{-0.04} | 0.62 ^{+0.01} _{-0.01} | 1.4 ^{+0.03} _{-0.03} | -4.69 ^{+0.03} _{-0.03} | 8.6 ^{+0.3} _{-0.4} |
| c.A.6 | F/G/*FeMg_am/ALL/-CH4/mr | -444 | 3.02 | 1191 ⁺⁴⁰ ₋₁₂ | 4.28 ^{+0.02} _{-0.02} | 1.2 ^{+0.06} _{-0.06} | 0.7 ^{+0.01} _{-0.01} | 1.07 ^{+0.01} _{-0.01} | -4.69 ^{+0.05} _{-0.02} | 8.4 ^{+0.3} _{-0.3} |
| c.AB.7 | D/S4/eq*FeMg_am/ALL/mr | -464 | 3.10 | 1198 ⁺²¹ ₋₂₃ | 4.29 ^{+0.02} _{-0.02} | 1.02 ^{+0.03} _{-0.03} | 0.6 ^{+0.01} _{-0.01} | 1.08 ^{+0.01} _{-0.01} | -4.67 ^{+0.03} _{-0.04} | 8.8 ^{+0.3} _{-0.3} |
| c.A.8 | F/G/*FeMg_am/ALL/-HCN/mr | -465 | 3.14 | 1197 ⁺⁴⁵ ₋₁₆ | 4.27 ^{+0.03} _{-0.04} | 1.36 ^{+0.05} _{-0.05} | 0.69 ^{+0.01} _{-0.01} | 1.1 ^{+0.05} _{-0.03} | -4.62 ^{+0.06} _{-0.05} | 8.7 ^{+0.2} _{-0.2} |
| c.AB.9 | D/M/eq*FeMg_am_P/ALL/mr | -468 | 3.25 | 1099 ⁺¹⁵ ₋₁₆ | 4.14 ^{+0.03} _{-0.03} | 1.1 ^{+0.05} _{-0.05} | 0.61 ^{+0.01} _{-0.01} | 1.28 ^{+0.04} _{-0.05} | -4.72 ^{+0.03} _{-0.03} | 8.9 ^{+0.4} _{-0.4} |
| c.AB.10 | D/Z/*FeMg_am/ALL/mr | -478 | 3.38 | 1145 ⁺⁹ ₋₇ | 4.23 ^{+0.02} _{-0.02} | 1.18 ^{+0.03} _{-0.03} | 0.62 ^{+0.01} _{-0.009} | 1.16 ^{+0.01} _{-0.01} | -4.69 ^{+0.02} _{-0.01} | 8.8 ^{+0.3} _{-0.3} |
| c.A.11 | D/Z/eq*FeMg_am/Gr/mr | -2678 | 0.84 | 1234 ⁺¹⁸ ₋₁₇ | 4.29 ^{+0.04} _{-0.03} | 0.79 ^{+0.1} _{-0.2} | 0.69 ^{+0.02} _{-0.02} | 1.01 ^{+0.03} _{-0.03} | -4.66 ^{+0.04} _{-0.03} | 7.6 ^{+0.4} _{-0.4} |

Notes

All values presented are the median values from the fits, with uncertainties given as the $\pm 34.1\%$ percentiles.

Table 24.7: Abridged retrieval results HR 8799 d

| Index | Model | B.F. | χ^2/ν | T_{Eff} [K] | $\log g$ [cgs] | [M/H] | C/O | Radius [R_{Jup}] | $\log L/L_{\odot}$ | Mass [M_{Jup}] |
|---------|-------------------------|-------|--------------|------------------------------------|----------------------------------------|--------------------------------------|----------------------------------------|----------------------------------------|-----------------------------------------|----------------------------------------|
| d.AB.0 | D/M/eq*FeMg_am/ALL/mr | 0 | 1.42 | 1177 ⁺²¹ ₋₂₁ | 4.18 ^{+0.04} _{-0.03} | 1.2 ^{+0.2} _{-0.1} | 0.61 ^{+0.03} _{-0.04} | 1.26 ^{+0.05} _{-0.06} | -4.63 ^{+0.04} _{-0.04} | 9.19 ^{+0.08} _{-0.07} |
| d.AB.1 | D/Z/*FeMg_am/ALL/mr | 0 | 1.39 | 1139 ⁺³⁸ ₋₁₉ | 4.13 ^{+0.02} _{-0.02} | 1.3 ^{+0.1} _{-0.1} | 0.6 ^{+0.03} _{-0.04} | 1.34 ^{+0.03} _{-0.03} | -4.6 ^{+0.06} _{-0.03} | 9.2 ^{+0.07} _{-0.07} |
| d.AB.2 | D/M/eq*FeMg_am_P/ALL/mr | -1 | 1.47 | 1220 ⁺¹⁰ ₋₁₀ | 4.25 ^{+0.02} _{-0.02} | 1.1 ^{+0.1} _{-0.1} | 0.49 ^{+0.04} _{-0.04} | 1.16 ^{+0.02} _{-0.02} | -4.58 ^{+0.03} _{-0.03} | 9.2 ^{+0.09} _{-0.09} |
| d.AB.3 | D/M/eq*FeMg_cd/ALL/mr | -1 | 1.47 | 1220 ⁺¹⁰ ₋₉ | 4.25 ^{+0.02} _{-0.02} | 1.1 ^{+0.1} _{-0.1} | 0.49 ^{+0.04} _{-0.04} | 1.16 ^{+0.02} _{-0.02} | -4.58 ^{+0.03} _{-0.03} | 9.21 ^{+0.09} _{-0.09} |
| d.AB.4 | F/G/*FeMg_cd/ALL/mr | -3 | 1.43 | 1194 ⁺¹⁸ ₋₁₅ | 4.18 ^{+0.02} _{-0.02} | 1.5 ^{+0.2} _{-0.2} | 0.68 ^{+0.03} _{-0.04} | 1.26 ^{+0.02} _{-0.03} | -4.58 ^{+0.03} _{-0.03} | 9.2 ^{+0.07} _{-0.07} |
| d.AB.5 | F/G/*FeMg_am/ALL/mr | -3 | 1.42 | 1196 ⁺²⁰ ₋₁₇ | 4.18 ^{+0.03} _{-0.02} | 1.5 ^{+0.2} _{-0.2} | 0.67 ^{+0.03} _{-0.04} | 1.26 ^{+0.03} _{-0.04} | -4.58 ^{+0.03} _{-0.03} | 9.2 ^{+0.07} _{-0.07} |
| d.AB.6 | F/Z/eq*FeMg_am/ALL/mr | -4 | 1.46 | 1146 ⁺¹³ ₋₁₃ | 4.17 ^{+0.02} _{-0.02} | 1.6 ^{+0.2} _{-0.1} | 0.67 ^{+0.04} _{-0.05} | 1.27 ^{+0.03} _{-0.03} | -4.62 ^{+0.03} _{-0.03} | 9.19 ^{+0.07} _{-0.07} |
| d.AB.7 | D/Z/eqMg_am/ALL/mr | -4 | 1.47 | 1150 ⁺²¹ ₋₁₆ | 4.17 ^{+0.04} _{-0.03} | 0.95 ^{+0.1} _{-0.1} | 0.5 ^{+0.04} _{-0.04} | 1.27 ^{+0.05} _{-0.06} | -4.67 ^{+0.04} _{-0.04} | 9.18 ^{+0.07} _{-0.07} |
| d.AB.8 | D/Z/eq*FeMg_am/ALL/mr | -5 | 1.49 | 1166 ⁺¹⁹ ₋₁₆ | 4.21 ^{+0.03} _{-0.03} | 0.98 ^{+0.1} _{-0.1} | 0.48 ^{+0.04} _{-0.04} | 1.22 ^{+0.05} _{-0.05} | -4.66 ^{+0.04} _{-0.04} | 9.19 ^{+0.07} _{-0.06} |
| d.AB.9 | D/Z/eq*FeMg_am/ALL/ | -6 | 1.50 | 1158 ⁺²² ₋₃₀ | 4.5 ^{+0.3} _{-0.3} | 1.3 ^{+0.3} _{-0.3} | 0.48 ^{+0.04} _{-0.04} | 1.23 ^{+0.08} _{-0.06} | -4.61 ^{+0.05} _{-0.06} | 20.70 ⁺¹⁹ ₋₁₀ |
| d.AB.10 | F/Z/*FeMg_am/ALL/mr | -6 | 1.44 | 1157 ⁺³⁰ ₋₁₆ | 4.16 ^{+0.01} _{-0.01} | 1.7 ^{+0.1} _{-0.1} | 0.64 ^{+0.04} _{-0.04} | 1.29 ^{+0.02} _{-0.02} | -4.57 ^{+0.04} _{-0.03} | 9.22 ^{+0.05} _{-0.05} |
| d.AB.11 | D/S4/eq*FeMg_am/ALL/mr | -12 | 1.51 | 1198 ⁺²¹ ₋₂₀ | 4.24 ^{+0.03} _{-0.04} | 0.71 ^{+0.2} _{-0.1} | 0.34 ^{+0.07} _{-0.06} | 1.17 ^{+0.05} _{-0.04} | -4.62 ^{+0.04} _{-0.04} | 9.19 ^{+0.07} _{-0.07} |
| d.A.12 | D/Z/eq*FeMg_am/Gr/mr | -2276 | 0.79 | 1172 ⁺¹⁹ ₋₁₈ | 4.18 ^{+0.03} _{-0.03} | 2 ^{+0.2} _{-0.2} | 0.78 ^{+0.02} _{-0.03} | 1.25 ^{+0.04} _{-0.04} | -4.61 ^{+0.04} _{-0.04} | 9.19 ^{+0.08} _{-0.08} |
| d.A.13 | D/Z/eq*FeMg_am/-Gr/mr | -3498 | 1.03 | 1162 ⁺²⁸ ₋₂₈ | 4.22 ^{+0.04} _{-0.04} | 0.46 ^{+0.7} _{-0.7} | 0.31 ^{+0.1} _{-0.1} | 1.19 ^{+0.06} _{-0.06} | -4.68 ^{+0.06} _{-0.06} | 9.2 ^{+0.08} _{-0.08} |

Notes

All values presented are the median values from the fits, with uncertainties given as the $\pm 34.1\%$ percentiles.

IV Characterisation of the HR 8799 planets

Table 24.8: Abridged retrieval results HR 8799 e

| Index | Model | B.F. | χ^2/ν | T_{Eff} [K] | $\log g$ [cgs] | [M/H] | C/O | Radius [R_{Jup}] | $\log L/L_{\odot}$ | Mass [M_{Jup}] |
|---------|-----------------------------|-------|--------------|------------------------------------|----------------------------------------|----------------------------------------|----------------------------------------|----------------------------------------|-----------------------------------------|---------------------------------------|
| e.A.0 | F/Z/eq*FeMg_am/ALL/-CH4/mr | 0 | 1.23 | 1139 ⁺²¹ ₋₂₀ | 4.18 ^{+0.05} _{-0.05} | 1.9 ^{+0.1} _{-0.1} | 0.88 ^{+0.01} _{-0.02} | 1.13 ^{+0.05} _{-0.05} | -4.73 ^{+0.05} _{-0.05} | 7.5 ^{+0.6} _{-0.6} |
| e.AB.1 | F/G/eq*FeMg_am/ALL/mr | 0 | 1.22 | 1172 ⁺²⁷ ₋₂₉ | 4.2 ^{+0.06} _{-0.06} | 1.8 ^{+0.1} _{-0.2} | 0.88 ^{+0.02} _{-0.02} | 1.1 ^{+0.05} _{-0.05} | -4.71 ^{+0.06} _{-0.05} | 7.5 ^{+0.7} _{-0.7} |
| e.AB.2 | F/Z/eq*FeMg_am/ALL/mr | 0 | 1.23 | 1134 ⁺²⁴ ₋₂₀ | 4.18 ^{+0.05} _{-0.05} | 1.9 ^{+0.1} _{-0.1} | 0.87 ^{+0.02} _{-0.02} | 1.14 ^{+0.05} _{-0.05} | -4.74 ^{+0.06} _{-0.05} | 7.5 ^{+0.6} _{-0.6} |
| e.AB.3 | F/G/fMg_am/ALL/mr | -1 | 1.22 | 1206 ⁺²⁷ ₋₂₃ | 4.24 ^{+0.05} _{-0.05} | 1.8 ^{+0.1} _{-0.2} | 0.88 ^{+0.02} _{-0.02} | 1.06 ^{+0.03} _{-0.03} | -4.67 ^{+0.05} _{-0.05} | 7.6 ^{+0.7} _{-0.7} |
| e.AB.4 | F/G/fFe/ALL/mr | -2 | 1.22 | 1203 ⁺³⁰ ₋₂₈ | 4.24 ^{+0.05} _{-0.05} | 1.8 ^{+0.1} _{-0.2} | 0.89 ^{+0.02} _{-0.02} | 1.07 ^{+0.03} _{-0.04} | -4.69 ^{+0.06} _{-0.05} | 7.6 ^{+0.7} _{-0.6} |
| e.AB.5 | F/G/grey_P/ALL/mr | -3 | 1.23 | 1261 ⁺¹⁸ ₋₁₈ | 4.35 ^{+0.05} _{-0.06} | 1.7 ^{+0.2} _{-0.3} | 0.89 ^{+0.02} _{-0.03} | 0.94 ^{+0.04} _{-0.03} | -4.64 ^{+0.04} _{-0.05} | 7.6 ^{+0.7} _{-0.7} |
| e.AB.6 | F/G/f*FeMg_am/ALL/mr | -3 | 1.23 | 1214 ⁺²² ₋₂₃ | 4.25 ^{+0.04} _{-0.04} | 1.8 ^{+0.1} _{-0.2} | 0.89 ^{+0.02} _{-0.02} | 1.05 ^{+0.03} _{-0.03} | -4.69 ^{+0.04} _{-0.04} | 7.5 ^{+0.6} _{-0.6} |
| e.AB.7 | F/S4/eq*FeMg_am/ALL/mr | -6 | 1.26 | 1202 ⁺¹⁹ ₋₂₀ | 4.28 ^{+0.05} _{-0.05} | 1.8 ^{+0.1} _{-0.2} | 0.9 ^{+0.02} _{-0.03} | 1.01 ^{+0.03} _{-0.03} | -4.7 ^{+0.05} _{-0.05} | 7.6 ^{+0.7} _{-0.7} |
| e.AB.8 | D/M/eq*FeMg_cd_P/ALL/mr | -6 | 1.32 | 1215 ⁺¹⁹ ₋₁₈ | 4.31 ^{+0.05} _{-0.05} | 2.1 ^{+0.2} _{-0.3} | 0.83 ^{+0.01} _{-0.02} | 1.0 ^{+0.03} _{-0.03} | -4.68 ^{+0.04} _{-0.05} | 7.9 ^{+0.7} _{-0.7} |
| e.AB.9 | D/Z/eq*FeMg_cd/ALL/mr | -7 | 1.27 | 1084 ⁺²⁷ ₋₂₃ | 4.09 ^{+0.06} _{-0.06} | 1.3 ^{+0.1} _{-0.1} | 0.77 ^{+0.03} _{-0.03} | 1.29 ^{+0.06} _{-0.07} | -4.82 ^{+0.06} _{-0.05} | 7.9 ^{+0.7} _{-0.7} |
| e.AB.10 | D/M/eq*FeMg_cd_P/ALL/mr | -7 | 1.32 | 1215 ⁺¹⁷ ₋₁₈ | 4.31 ^{+0.05} _{-0.05} | 2.1 ^{+0.2} _{-0.3} | 0.83 ^{+0.01} _{-0.02} | 1.0 ^{+0.03} _{-0.03} | -4.68 ^{+0.05} _{-0.05} | 7.9 ^{+0.6} _{-0.7} |
| e.AB.11 | D/M/eq*FeMg_cd/ALL/mr | -7 | 1.28 | 1106 ⁺³³ ₋₂₈ | 4.11 ^{+0.08} _{-0.06} | 1.3 ^{+0.2} _{-0.2} | 0.78 ^{+0.03} _{-0.03} | 1.26 ^{+0.07} _{-0.09} | -4.82 ^{+0.07} _{-0.06} | 7.9 ^{+0.7} _{-0.7} |
| e.A.12 | F/Z/eq*FeMg_am/ALL/-HCN/mr | -8 | 1.30 | 1148 ⁺²⁴ ₋₂₃ | 4.15 ^{+0.06} _{-0.05} | 1.91 ^{+0.07} _{-0.1} | 0.87 ^{+0.02} _{-0.02} | 1.17 ^{+0.06} _{-0.06} | -4.73 ^{+0.06} _{-0.06} | 7.5 ^{+0.7} _{-0.6} |
| e.AB.13 | D/Z/f*FeMg_am/ALL/mr | -9 | 1.33 | 1187 ⁺⁵⁰ ₋₃₀ | 4.26 ^{+0.04} _{-0.04} | 2 ^{+0.3} _{-0.3} | 0.83 ^{+0.02} _{-0.02} | 1.08 ^{+0.03} _{-0.03} | -4.69 ^{+0.07} _{-0.06} | 8.1 ^{+0.6} _{-0.6} |
| e.AB.14 | D/Z/eq*FeMg2/Gr/mr | -9 | 1.35 | 1170 ⁺²⁶ ₋₂₇ | 4.26 ^{+0.06} _{-0.06} | 1.3 ^{+0.2} _{-0.2} | 0.76 ^{+0.02} _{-0.03} | 1.06 ^{+0.05} _{-0.05} | -4.77 ^{+0.05} _{-0.06} | 7.8 ^{+0.7} _{-0.7} |
| e.AB.15 | F/Z/f*FeMg_am/ALL/mr | -10 | 1.29 | 1143 ⁺¹² ₋₁₈ | 4.19 ^{+0.05} _{-0.05} | 1.99 ^{+0.05} _{-0.05} | 0.86 ^{+0.02} _{-0.02} | 1.09 ^{+0.03} _{-0.03} | -4.73 ^{+0.04} _{-0.04} | 7 ^{+0.6} _{-0.7} |
| e.A.16 | D/Z/eq*FeMg_am/ALL/Fe-H15mr | -10 | 1.37 | 1202 ⁺²⁰ ₋₁₇ | 4.32 ^{+0.05} _{-0.05} | 1.5 ⁺⁰ ₋₀ | 0.76 ^{+0.02} _{-0.02} | 0.99 ^{+0.03} _{-0.03} | -4.73 ^{+0.05} _{-0.05} | 7.8 ^{+0.7} _{-0.7} |
| e.A.17 | D/Z/eq*FeMg_am/ALL/Fe-H20mr | -10 | 1.33 | 1179 ⁺²⁹ ₋₂₈ | 4.25 ^{+0.06} _{-0.06} | 2.0 ⁺⁰ ₋₀ | 0.82 ^{+0.01} _{-0.01} | 1.08 ^{+0.07} _{-0.07} | -4.77 ^{+0.06} _{-0.05} | 8 ^{+0.6} _{-0.6} |
| e.A.18 | D/M/grey_P/ALL/mr | -10 | 1.38 | 1245 ⁺¹⁶ ₋₁₆ | 4.36 ^{+0.05} _{-0.05} | 1.5 ^{+0.1} _{-0.1} | 0.63 ^{+0.02} _{-0.02} | 0.95 ^{+0.03} _{-0.03} | -4.68 ^{+0.04} _{-0.04} | 8 ^{+0.8} _{-0.8} |
| e.AB.19 | D/M/eq*FeMg_am/ALL/mr | -11 | 1.33 | 1206 ⁺²¹ ₋₂₂ | 4.3 ^{+0.05} _{-0.05} | 2.1 ^{+0.2} _{-0.2} | 0.83 ^{+0.01} _{-0.02} | 1.02 ^{+0.05} _{-0.04} | -4.69 ^{+0.05} _{-0.06} | 8.1 ^{+0.6} _{-0.6} |
| e.AB.20 | D/Z/eq*FeMg_am_h/ALL/mr | -11 | 1.33 | 1129 ⁺²⁹ ₋₂₇ | 4.15 ^{+0.06} _{-0.05} | 1.4 ^{+0.2} _{-0.1} | 0.78 ^{+0.03} _{-0.03} | 1.21 ^{+0.06} _{-0.06} | -4.77 ^{+0.07} _{-0.07} | 7.9 ^{+0.6} _{-0.6} |
| e.AB.21 | D/M/grey_P/ALL/mr | -11 | 1.37 | 1245 ⁺¹⁵ ₋₁₅ | 4.34 ^{+0.05} _{-0.07} | 1.5 ^{+0.1} _{-0.1} | 0.63 ^{+0.02} _{-0.02} | 0.97 ^{+0.06} _{-0.03} | -4.64 ^{+0.08} _{-0.06} | 8 ^{+0.8} _{-0.8} |
| e.A.22 | D/Z/eq*FeMg_am/ALL/Fe-H10mr | -11 | 1.41 | 1208 ⁺²¹ ₋₁₉ | 4.29 ^{+0.05} _{-0.05} | 1.0 ⁺⁰ ₋₀ | 0.69 ^{+0.02} _{-0.02} | 0.98 ^{+0.03} _{-0.03} | -4.72 ^{+0.05} _{-0.05} | 7.3 ^{+0.7} _{-0.6} |
| e.A.23 | D/Z/eq*FeMg_am/ALL/mr | -11 | 1.36 | 1209 ⁺²⁰ ₋₁₇ | 4.32 ^{+0.04} _{-0.05} | 1.7 ^{+0.7} _{-0.3} | 0.79 ^{+0.03} _{-0.05} | 0.99 ^{+0.03} _{-0.03} | -4.72 ^{+0.05} _{-0.05} | 7.9 ^{+0.6} _{-0.7} |
| e.AB.24 | D/Z/eqAl2O3/ALL/mr | -11 | 1.34 | 1235 ⁺¹⁷ ₋₁₄ | 4.36 ^{+0.04} _{-0.05} | 2.3 ^{+0.1} _{-0.3} | 0.82 ^{+0.01} _{-0.01} | 0.94 ^{+0.03} _{-0.03} | -4.73 ^{+0.04} _{-0.04} | 7.9 ^{+0.7} _{-0.7} |
| e.AB.25 | D/Z/eq*FeMg_am/ALL/mr | -11 | 1.36 | 1207 ⁺¹⁹ ₋₁₇ | 4.32 ^{+0.04} _{-0.04} | 2.1 ^{+0.3} _{-0.6} | 0.81 ^{+0.01} _{-0.06} | 0.99 ^{+0.03} _{-0.03} | -4.72 ^{+0.04} _{-0.05} | 8 ^{+0.6} _{-0.6} |
| e.AB.26 | D/Z/eq*FeMg_am/ALL/ | -12 | 1.36 | 1224 ⁺²¹ ₋₂₂ | 4.8 ^{+0.3} _{-0.2} | 2 ^{+0.3} _{-0.3} | 0.76 ^{+0.05} _{-0.03} | 0.96 ^{+0.03} _{-0.03} | -4.78 ^{+0.06} _{-0.06} | 19.91 ^{+1.3} _{-0.6} |
| e.AB.27 | F/G/grey/ALL/mr | -12 | 1.37 | 1245 ⁺⁵⁸ ₋₃₁ | 4.24 ^{+0.06} _{-0.06} | 2.05 ^{+0.09} _{-0.08} | 0.91 ^{+0.02} _{-0.02} | 1.04 ^{+0.04} _{-0.04} | -4.6 ^{+0.09} _{-0.07} | 7.4 ^{+0.8} _{-0.8} |
| e.AB.28 | D/S4/eq*FeMg_cd/ALL/mr | -12 | 1.27 | 1093 ⁺²⁷ ₋₂₄ | 4.08 ^{+0.05} _{-0.05} | 1.4 ^{+0.1} _{-0.1} | 0.78 ^{+0.02} _{-0.03} | 1.3 ^{+0.05} _{-0.05} | -4.73 ^{+0.06} _{-0.06} | 7.8 ^{+0.7} _{-0.7} |
| e.A.29 | D/Z/eq*FeMg_am/ALL/Fe-H05mr | -17 | 1.49 | 1217 ⁺²⁰ ₋₁₈ | 4.29 ^{+0.05} _{-0.05} | 0.5 ⁺⁰ ₋₀ | 0.52 ^{+0.04} _{-0.04} | 0.97 ^{+0.03} _{-0.03} | -4.71 ^{+0.04} _{-0.05} | 7.1 ^{+0.7} _{-0.6} |
| e.A.30 | D/Z/eq*FeMg_am/ALL/Fe-H00mr | -26 | 1.58 | 1227 ⁺¹⁷ ₋₁₆ | 4.29 ^{+0.05} _{-0.05} | 0.0 ⁺⁰ ₋₀ | 0.21 ^{+0.02} _{-0.02} | 0.95 ^{+0.03} _{-0.03} | -4.72 ^{+0.04} _{-0.05} | 6.9 ^{+0.6} _{-0.6} |
| e.A.31 | D/Z/eq*FeMg_am/Gr/mr | -2294 | 0.72 | 1173 ⁺³¹ ₋₂₈ | 4.24 ^{+0.06} _{-0.06} | 1.8 ^{+0.4} _{-0.5} | 0.85 ^{+0.02} _{-0.04} | 1.06 ^{+0.06} _{-0.05} | -4.71 ^{+0.07} _{-0.07} | 7.5 ^{+0.7} _{-0.7} |
| e.A.32 | D/Z/CLR/Gr/mr | -2301 | 0.79 | 1339 ⁺³⁶ ₋₃₈ | 4.3 ^{+0.07} _{-0.07} | 0.65 ^{+0.2} _{-0.2} | 0.71 ^{+0.05} _{-0.07} | 0.94 ^{+0.06} _{-0.05} | -4.53 ^{+0.08} _{-0.08} | 6.9 ^{+0.8} _{-0.6} |
| e.A.33 | D/Z/eq*FeMg_am/-Gr/mr | -3496 | 1.88 | 1143 ⁺³⁸ ₋₃₂ | 4.17 ^{+0.08} _{-0.08} | 2 ^{+0.3} _{-0.4} | 0.71 ^{+0.08} _{-0.2} | 1.15 ^{+0.09} _{-0.09} | -4.79 ^{+0.08} _{-0.09} | 7.5 ^{+0.7} _{-0.7} |

Notes

All values presented are the median values from the fits, with uncertainties given as the $\pm 34.1\%$ percentiles.



SUMMARY & OUTLOOK



25. SUMMARY

25.1. Retrieval development

pRT has grown to be among the most popular tools for modelling exoplanet spectra, with 260 refereed citations in the literature at time of publication. With the introduction of the retrieval package, pRT's capabilities have been expanded to meet the developing needs of the community. Through the use of nested sampling, correlated-k opacities, and a highly flexible retrieval framework, pRT is well positioned to serve as a characterisation tool for the high precision observations enabled by modern observatories.

This was demonstrated via application to WASP-39 b, an exemplary target for characterisation with *JWST*. The findings of significant metal enrichment ($\geq 10\times$ solar) and a subs-stellar C/O ratio were confirmed. The primary trace gasses of the atmosphere are H₂O and CO, with statistically significant detections of H₂S ($> 3\sigma$) and SO₂ ($\geq 5\sigma$). Patchy clouds in the terminator region were identified, with no significant evidence for clouds composed of KCl, SiO₂ or MgSiO₃. This work will form part of a retrieval comparison project, exploring how modelling choices impact inferences of atmospheric parameters.

25.2. Importance of data processing

Following the development of the retrieval package, a key consideration was understanding how data processing choices impact one's ability to infer atmospheric properties. We compared the KLIP, PynPoint and ANDROMEDA post-processing algorithms for high-contrast IFS data. The hyperparameters of each algorithm were optimised through injection and recovery testing, which highlighted the importance of tailoring the post-processing for unbiased spectral extraction, rather than simply maximising the S/N of the extracted spectrum. The noise properties imparted from the post-processing were unique to each algorithm. We therefore empirically measured the covariance matrix for each extracted spectrum and incorporated the covariance matrix into the likelihood function of the pRT retrieval.

Through a series of retrieval tests on synthetic data, we explored how the inferred parameters depend on the choice of data processing algorithm and the use of the covariance matrix. We found that the spectra extracted by each algorithm could result in significant variations in the posterior distributions. As such, injection testing should be used to verify that the spectral extraction is minimally biased, and retrievals should be run on multiple extractions in order to ensure consistent results. Replicating the results of [Greco & Brandt \(2016\)](#),

we found that if the covariance matrix is not included in the likelihood function, the inferred posterior distributions will be both biased and overconfident. It is therefore critical to either empirically measure the covariance matrix, or to fit for the covariance matrix as part of the retrieval itself.

25.3. New insight into the HR 8799 planets

Using the pRT retrieval package, we performed a systematic characterisation of the HR 8799 planets. Using new VLTI/GRAVITY observations, together with data reprocessed using the methods of Chapter III, we compiled the most comprehensive spectra of these objects to date. A series of retrievals were ran for each planet, exploring different parameterisation of the thermal structure, chemistry and clouds. Using Bayesian Model Averaging techniques, these retrievals were combined in a statistically robust fashion, producing robust estimates of key planet properties. We validated the retrievals through comparisons to self-consistent grid fits, finding that both approaches led to compatible findings.

All of the HR 8799 planets are enriched in metals, by factors of 10 to 60 times solar. The C/O ratio is stellar-to-superstellar for all four of the planets. As known from previous studies, the primary trace gasses in the atmospheres of the HR 8799 planets are water and CO, but we were also able to detect HCN at $> 5\sigma$ in HR 8799 c and e, as well as CH₄ at $> 5\sigma$ in HR 8799 c. Using a quench pressure approximation and mixing length theory, we were able to derive measurements of the vertical mixing strength that were compatible with literature predictions, finding relatively weak vertical mixing is required to drive the disequilibrium chemistry. While fitting the planet radius is often challenging when using emission spectra, the broad wavelength coverage and dynamical mass priors enabled the retrieval of radii consistent with predictions from evolutionary models. The atmospheres of all four planets are likely cloudy, showing damped absorption features and red colours. Silicate clouds provide the best fits for HR 8799 c, d, and e, though the resulting models do not display the silicate absorption feature near 10 μm as expected for this class of L/T transition object. HR 8799 b is better fit by an Na₂S cloud, which is also predicted by equilibrium chemistry for objects of similar effective temperatures (950 K).

25.3.1. Formation

Having precisely measured their atmospheric properties, we can turn again to the question of how the HR 8799 planets formed. It remains unclear whether this system formed by gravitational instability or core accretion. Indeed, with

four super-Jupiter planets in the same system, it is unclear how representative this system is of the directly imaged exoplanet population as a whole.

We can consider the evidence for each formation pathway. While in general planets formed through GI are expected to have a composition closer to that of stars, it remains possible that the particular properties of the HR 8799 system may enable this pathway. A high metallicity disk, as required to produce the metal rich planets, would lead to faster cooling timescales due to the increased opacity, as demonstrated by Rafikov (2005). From that work, for gravitational instability to occur at 100 au, “the disk has to possess at least $\tau \approx 2$ (marginally optically thick), $T_{\min} \approx 20$ K, $\Sigma_{\min} \approx 25$ g cm⁻², $L \gtrsim 10^{-2}L_{\odot}$ and $M_a \gtrsim 0.1M_{\odot}$ at 100 au”. While these are relatively extreme properties with respect to the average protoplanetary disk, HR 8799 is also an extreme system in terms of its mass. At the absolute minimum, there is 0.03 M_{\odot} of material stored in the four known companions. This means that even if the accretion efficiency of the HR 8799 planets was significantly higher than the $\lesssim 10\%$ efficiency expected in the solar system (e.g. Eriksson et al., 2022), the mass of the disk was likely sufficient to be gravitationally unstable. We can take this pathway one step farther, using equation 15 of Rafikov (2005) (equation 4.3 of this thesis). With a 10% accretion efficiency the disk mass between 15-70 au would have been at least 0.3 M_{\odot} . If the disk had an above-average metallicity, the opacity due to molecular and dust absorption is also likely to be high. Freedman et al. (2014) calculates the Rosseland mean opacity for varying metallicity and gas temperatures and pressures. At 100 K and a metallicity of 3 \times solar, κ_R might be $\sim 1.5\times$ greater than a solar composition gas, rising to more than a factor of 2 at 300 K. Combining this into equation 4.3, we can calculate the fragment mass for a 1.5 M_{\odot} host star, at a separation of 68 au. Even assuming only a small increase in $f(\tau)$ from 1 to 1.5 and assuming that the mean molar mass reflects a solar composition $\mu = 2.3$, the fragment mass could still be as small as 2.5 M_{Jup} . All of the HR 8799 planets fall above this threshold, suggesting that disk fragmentation is a possible pathway for their formation.

Conversely, Bitsch & Mah (2023) demonstrate a pathway to metal rich giant planets via pebble accretion. In their model, dust and small pebbles that are coupled to the gas flow will accrete onto a protoplanet during the runaway gas accretion phase without being trapped by the pressure bump induced by the planet. In this model, they find that giant planets can accrete far more metals than expected from the mass-metallicity relationship of Thorngren et al. (2016), thus potentially explaining the discrepancy between the HR 8799 system and the general directly imaged population.

Clearly more work is required to disentangle these potential formation pathways. Mollière et al. (2022) present a tool to infer the formation location

of exoplanets given their atmospheric properties and assuming a model of the disk chemistry and evolution. This should be applied to the HR 8799 planets, but equally important are the development of forward models adequate to perform such an inference for a system as complicated as the HR 8799 planets. The interactions of all four planets with the disk and with each other must be accounted for. Only with such a complete model can we make robust estimates of the properties of the disk that gave birth to this enigmatic system.

26. OUTLOOK

26.1. Future pRT development

While pRT is a state-of-the-art exoplanet modelling tool (particularly with the recent release of pRT3), there remain many avenues through which to improve the code further. The first goal is to improve the computational efficiency of the retrieval package. This will permit faster and easier retrievals at the current level of model complexity as a routine characterisation tool. Perhaps more importantly it will enable increases in retrieval complexity, both for fitting higher data with higher precision, spectral resolution, and broader wavelength coverage, as well as increases in model complexity through 3D treatments of atmospheres and incorporating time-dependent phenomena. 3D retrievals have already been demonstrated by [MacDonald & Lewis \(2022\)](#) and [Chubb & Min \(2022\)](#). As *JWST* begins to explore the variability of exoplanet and brown dwarf atmospheres, it will become necessary to build such time-dependence into the retrieval framework in order to interpret how atmospheric properties vary over time and spatial location. Faster retrievals can enable the inclusion of additional physics into the models. This could allow for the coupling of the chemistry and thermal structure, potentially enabling fully self-consistent models within a retrieval framework. In the other direction, it could also enable the rapid exploration of a high dimensional parameter space, allowing for a truly data-driven approach to model fitting. Both of these approaches would provide unique insight into the structure of exoplanet and brown dwarf atmospheres.

One path to follow in the pursuit of increased computational efficiency is the use of autodifferentiable code. This approach has been demonstrated by [Kawahara et al. \(2022\)](#), who make use of the JAX library to produce a GPU/TPU accelerated where the entire spectral model is automatically differentiable. This approach could allow for rapid retrievals, making use of gradient descent methods, as well as nested sampling methods such as JAXNS, which has been demonstrated to be far faster than `MultiNest` or similar existing methods ([Albert, 2020](#)).

Other machine learning approaches are also likely to result in dramatic improvements in computational speed. [Vasist et al. \(2023\)](#) incorporate pRT into a neural posterior estimation framework. Once a model is trained, the runtime of the actual retrieval where the model is fit to the data is negligible. While there is still computational expense in training the variety of models required, additional developments in this field may enable easier model generalisation.

26.2. Direct imaging with JWST

In the immediate future, *JWST* presents the best opportunity to characterise both transiting and directly imaged planets. While WASP-39 b is the exemplar transiting planet, VHS 1256 b has shown the power and challenges of broad wavelength, high spectral resolution observations (Miles et al., 2023; Petrus et al., 2024). These spectra are filled with molecular and isotopologue absorption lines, with detailed information about the thermal structure and direct observations of aerosol absorption features.

26.2.1. The Direct Imaging Spectral Community Survey

PIs: Niall Whiteford, Evert Nasedkin, Polychronis Patapis & Ioannis Argyriou.

While the detailed characterisation of individual planets or systems is key to interpreting their atmospheric properties, unravelling trends in planet formation, cloud formation and structure, and atmospheric chemistry will require a population level analysis. Such an analysis is also necessary to determine the cause of empirical trends, such as the colour differences between brown dwarfs and young, giant exoplanets. While similar, it has become clear that there are many spectral – and therefore physical – differences between these exoplanets and their brown dwarf cousins (Currie et al., 2023b). Differences in the surface gravity between young planets and brown dwarfs are thought to drive the divergence in near-infrared colour, through changes to their temperature structure, chemical state, and cloud properties (Tsuji & Nakajima, 2003). This requires measurement of H₂O, CO, and CH₄ molecular lines from 1–5 μm , together with silicate cloud absorption near 10 μm .

The unique wavelength coverage and spectral resolution of *JWST* enable the observation of these critical spectroscopic features which are inaccessible from the ground. We therefore proposed a *JWST* GO program, the Direct Imaging Spectral Community Survey (DISCO) to systematically study the atmospheric properties of both the directly-imaged exoplanets and brown dwarf populations as a function of effective temperature and surface gravity. Spectroscopic observations from 1 to 28 μm will allow us to unravel the enigmatic ‘L/T transition’ and understand the empirical differences between the two populations. By covering a temperature range from ~ 800 K to 2200 K, the observations will target fundamental atmospheric transitions: the onset and evolution of clouds, as well as stark changes in carbon chemistry (Zahnle & Marley, 2014). The sample covering this range in temperature and spectral type is shown in figure 26.1, highlighting the representative nature of our brown dwarf comparison objects and the anomalous colours of the young exoplanets.

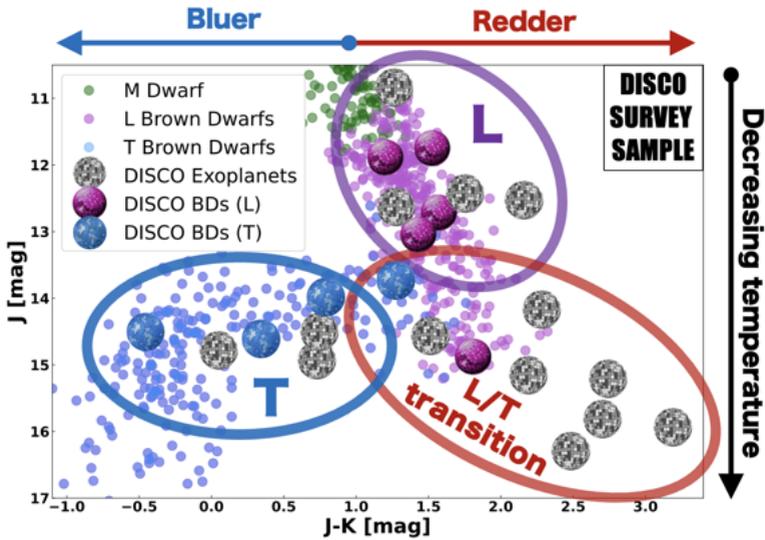


Figure 26.1: Colour–magnitude diagram of near-IR colours, outlining the L to T sequence of field brown dwarfs (Dupuy & Liu, 2012; Dupuy & Kraus, 2013; Liu et al., 2016; Best et al., 2018, 2020b,a) with the DISCO exoplanet and brown dwarf samples highlighted. Three distinct regions of the parameter colour space are covered by our sample: (1) L type, (2) the L/T transition, and (3) T type. These objects, across the 3 sub-samples, will allow us to study why exoplanets are systematically redder in colour (larger J-K mag, not extinction) than brown dwarfs at the same luminosity, and our observations will link these near-IR colours to mid-IR spectral features. The broad wavelength coverage of our survey will link these empirical trends to physical processes by measuring cloud properties and chemical states as they evolve across the sequence.

Such a program would address key open questions in atmospheric physics, but will also contextualise these objects and their place in the stellar formation story, distinguishing low mass stars from bona fide planets. Specifically, the DISCO survey would

1. Transform our understanding of the chemistry and clouds that shape the spectral signatures of sub-stellar atmospheres as a function of temperature.
2. Unravel the underlying processes that drive the divergent atmospheric evolution of directly-imaged exoplanets from brown dwarfs.
3. Facilitate improvements to the next generation of atmospheric and planet formation models with better understanding of fundamental properties.
4. Be the first broad wavelength spectral library of directly imaged exoplanets.

While this proposal has not (yet) been accepted, we firmly believe that a systematic survey such as this is necessary to answer fundamental questions about the atmospheric physics and formation history of directly imaged exoplanets.

26.3. Closing

We live in the most exciting of times to study the atmospheres of exoplanets. Having looked beyond the wanderers of our own solar system, we are now able to explore the vast array of worlds in our local neighbourhood and beyond, with each new discovery raising further questions. The current and next generation of observatories will expand our capabilities further, pushing to smaller and colder planets in the pursuit of finding somewhere else like home. We can only achieve this by overcoming monumental technical challenges, and it is through the continuous development of techniques such as those described in this thesis that we can achieve these goals.

ABBREVIATIONS

4QPM 4 Quadrant Phase Mask

ADI Angular differential imaging

ALES Arizona Lenslets for Exoplanet Spectroscopy

ALMA Atacama Large Millimeter Array

ANDROMEDA ANgular DiffeRential OptiMal Exoplanet Detection Algorithm

AO Adaptive optics

APP Apodizing Phase Plate

AT Auxiliary Telescope, VLT

BMA Bayesian model averaging

C/O Carbon-to-oxygen number ratio

c-k correlated-k

CHARIS Coronagraphic High Angular Resolution Imaging Spectrograph

CIA Collision induced absorption

CPD Circumplanetary disk

CSD Circumstellar disk

DM Deformable mirror

ELT Extremely Large Telescope

EOS Equation of state

ERS Early Release Science

FHWM Full width at half maximum

GCM Global circulation model

GI Gravitational Instability

GPI Gemini Planet Imager

HAT Hungarian Automated Telescope

HCI High-contrast imaging

HST Hubble Space Telescope

IFS Integral field spectrograph

JWST James Webb Space Telescope

KELT Kilodegree Exoplanet Little Telescope

KLIP Karhunen-Loève Image Processing

lbl line-by-line

-
- LBT** Large Binocular Telescope
LTE Local thermodynamic equilibrium
mas milliarcsecond
MIRI Mid-Infrared Instrument (JWST)
NS Nested sampling
OPD Optical path difference
PCA Principle component analysis
PSF Point-spread function
RCE Radiative-convective equilibrium
RDI Reference differential imaging
RTC Real time computer
RTE Radiative transfer equation
RV Radial velocity
S/N Signal to noise ratio
SED Spectral energy distribution
SDI Spectral differential imaging
SPHERE Spectro-Polarimetric High-Contrast Exoplanet Research
TESS Transiting Exoplanet Survey Satellite
UT Unit Telescope, VLT
VLT Very Large Telescope
WASP Wide Angle Search for Planets
WFS Wavefront sensor
YSO Young stellar object

LIST OF FIGURES

| | | |
|------|---------------------------------------------------------------------------------------|-----|
| 2.1 | Portrait of directly imaged exoplanets. | 9 |
| 2.2 | Exoplanet demographics. | 11 |
| 3.1 | Contrast curves of current and future high contrast instruments. | 14 |
| 3.2 | Coronagraph example. | 16 |
| 3.3 | Example of SPHERE IFS Data. | 18 |
| 3.4 | Double slit experiment geometry. | 23 |
| 3.5 | Schematic of a single baseline of VLTI/GRAVITY. | 25 |
| 3.6 | VLTI/GRAVITY fringes for HR8799e. | 26 |
| 4.1 | DSHARP disks observed with ALMA. | 29 |
| 4.2 | MAPS disks observed with ALMA. | 31 |
| 5.1 | Portraits of Jupiter. | 40 |
| 5.2 | Brown Dwarf Spectral Sequence. | 44 |
| 5.3 | Brown Dwarf Colour-Magnitude Diagram. | 45 |
| 5.4 | Molecular absorption in a simple atmosphere. | 56 |
| 5.5 | CIA in a simple atmosphere. | 60 |
| 5.6 | Cloud absorption in a simple atmosphere. | 64 |
| 5.7 | Modelling an emission spectrum. | 65 |
| 5.8 | GCM modelling of brown dwarf atmospheres. | 68 |
| 7.1 | Binning c-k and lbl spectra comparison. | 78 |
| 8.1 | pRT program flow. | 93 |
| 9.1 | WASP-39 b <i>JWST</i> Data. | 99 |
| 9.2 | pRT fits for WASP-39 b. | 102 |
| 9.3 | Retrieved WASP-39 b temperature profile. | 104 |
| 9.4 | WASP-39 b Transmission contribution function. | 105 |
| 9.5 | WASP-39 b posterior distributions. | 106 |
| 9.6 | Impact of $\log g$, R_{pl} and reference pressure on WASP 39-b retrievals. | 108 |
| 9.7 | Impact of $\log g$, R_{pl} and reference pressure on WASP 39-b metallicity. | 108 |
| 11.1 | Post-processed GPI data of HR 8799. | 122 |
| 12.1 | Correlation matrices for HR 8799e. | 133 |
| 12.2 | Toy model for demonstrating the effect of correlated uncertainties. | 134 |

| | | |
|-------|----------------------------------------------------------------------------------------|-----|
| 12.3 | χ^2/ν as a function of the ratio between the correlation length and the period. | 138 |
| 12.4 | Goodness-of-fit of spectral extractions. | 139 |
| 13.1 | Spectral extractions for injected planets located at separations of 600 mas. | 142 |
| 13.2 | Best-fit discrepancy as a function of input contrast at 400 mas. | 143 |
| 14.1 | Best-fit models compared to the ground truth spectrum. | 150 |
| 14.2 | Impacts of post-processing algorithm choice on retrieval results. | 154 |
| 14.3 | Impact of KL-mode choice on parameter inference. | 157 |
| 14.4 | Impacts of including the covariance matrix on retrieval results. | 158 |
| 14.5 | Posterior distributions for retrievals in the ‘Noise-free’ case. | 159 |
| 16.1 | SPHERE reduction comparison. | 167 |
| 18.1 | A GPI portrait of HR 8799. | 175 |
| 18.2 | VLT/GRAVITY Spectra of the HR 8799 planets. | 189 |
| 19.1 | Combined spectra of the HR 8799 planets. | 195 |
| 20.1 | HR 8799 retrieval classification. | 204 |
| 20.2 | Averaged posterior distributions for the HR 8799 planets. | 211 |
| 21.1 | Best fit temperature profiles and spectra for HR 8799 bcde. | 212 |
| 21.2 | Metallicity posterior distributions. | 214 |
| 21.3 | C/O posterior distributions. | 215 |
| 21.4 | T_{eff} posterior distributions. | 216 |
| 21.5 | Self-consistent grid fits of the HR 8799 planets. | 217 |
| 21.6 | Chemical abundance profiles of HR 8799 bcde. | 222 |
| 21.7 | Temperature profiles for HR 8799 e. | 228 |
| 21.8 | Elemental number ratios for the HR 8799 planets. | 237 |
| 21.9 | Cloud properties for HR 8799 e. | 238 |
| 21.10 | Effective particle radii as a function of altitude for silicate and iron clouds. | 238 |
| 22.1 | Absorption and scattering opacities for different condensates. | 240 |
| 22.2 | Metallicities of the exoplanet planet population. | 241 |
| 22.3 | Fixed metallicity retrievals. | 245 |
| 22.4 | Predictions for NIRSpec/G395H for HR 8799 bcde. | 246 |
| 22.5 | Predictions for MIRI/MRS for HR 8799 bcde. | 247 |
| 24.1 | Reprocessed SPHERE spectra for HR 8799 d and e. | 253 |
| 24.2 | Reprocessed GPI spectra for HR 8799 c, d, and e. | 254 |

| | | |
|------|--------------------------------------------------------|-----|
| 24.3 | Comparison of disequilibrium models to OSIRIS data. | 255 |
| 24.4 | Opacity contribution of CH ₄ for HR 8799 c. | 255 |
| 24.5 | Opacity contribution of HCN for HR 8799 c. | 256 |
| 26.1 | The DISCO Survey sample. | 270 |

LIST OF TABLES

| | | |
|------|--------------------------------------------------------------------|-----|
| 9.1 | Priors for WASP-39 free chemistry retrieval setup. | 100 |
| 9.2 | Priors for WASP-39 disequilibrium retrieval setup. | 101 |
| 9.3 | Volume mixing ratios from the full data, free chemistry retrieval. | 103 |
| 9.5 | No evidence for non-grey clouds in the atmosphere of WASP-39 b. | 110 |
| 9.4 | Volume mixing ratios for each dataset, free chemistry retrieval. | 112 |
| 11.1 | GPI and SPHERE observations of HR 8799. | 118 |
| 11.2 | Stellar properties of HR 8799 A. | 120 |
| 12.1 | Post-processing parameters. | 127 |
| 14.1 | Priors for retrieval setup. | 148 |
| 14.2 | Summary of retrievals run on synthetic data. | 149 |
| 14.3 | Summary statistics for KLIP retrievals. | 161 |
| 18.1 | Summary of literature and derived planet properties for HR 8799 b. | 177 |
| 18.2 | Summary of literature and derived planet properties for HR 8799 c | 179 |
| 18.3 | Summary of literature and derived planet properties for HR 8799 d | 181 |
| 18.4 | Summary of literature and derived planet properties for HR 8799 e | 183 |
| 20.1 | HR 8799 retrieval priors. | 197 |
| 20.2 | Priors for temperature profiles. | 209 |
| 20.3 | Self-consistent grid boundaries and step sizes. | 210 |
| 21.1 | Grid-retrieval results | 213 |
| 21.2 | Grid-fit χ^2 results | 218 |
| 21.3 | Vertical mixing parameters of the HR 8799 planets. | 233 |
| 24.1 | Near infrared stellar Photometry of HR8799. | 250 |
| 24.2 | Photometric data for HR 8799 b and c. | 251 |
| 24.3 | Photometric data for HR 8799 d and e. | 252 |
| 24.4 | Spectroscopic Observation Log. | 259 |
| 24.5 | Abridged retrieval results HR 8799 b | 260 |
| 24.6 | Abridged retrieval results HR 8799 c | 261 |
| 24.7 | Abridged retrieval results HR 8799 d | 261 |
| 24.8 | Abridged retrieval results HR 8799 e | 262 |

PERSONAL WORKS

Works used in this thesis

- Nasedkin, E., Mollière, P., Wang, J., et al. 2023, *A&A*, 678, A41
Nasedkin, E., Mollière, P., & Blain, D. 2024, *JOSS*, 9, 5875
Nasedkin, E., Mollière, P., Lacour, S., et al. 2024, *arXiv e-prints*, arXiv:2404.03776
Welbanks, L., Piaulet-G., C., Barstow, J. et al. (in prep)

Complete works of the author

- Liu, T., Li, P. S., Juvela, M., et al. 2018, *ApJ*, 859, 151
Mollière, P., Stolker, T., Lacour, S., et al. 2020, *A&A*, 640, A131
Nowak, M., Lacour, S., Lagrange, A. M., et al. 2020, *A&A*, 642, L2
Wang, J. J., Vigan, A., Lacour, S., et al. 2021, *AJ*, 161, 148
Kammerer, J., Lacour, S., Stolker, T., et al. 2021, *A&A*, 652, A57
Cugno, G., Patapis, P., Stolker, T., et al. 2021, *A&A*, 653, A12
Lacour, S., Wang, J. J., Rodet, L., et al. 2021, *A&A*, 654, L2
Patapis, P., Nasedkin, E., Cugno, G., et al. 2022, *A&A*, 658, A72
Mollière, P., Molyarova, T., Bitsch, B., et al. 2022, *ApJ*, 934, 74
Christiaens, V., Gonzalez, C., Farkas, R., et al. 2023, *JOSS*, 8, 4774
Hinkley, S., Lacour, S., Marleau, G. D., et al. 2023, *A&A*, 671, L5
Vasist, M., Rozet, F., Absil, O., et al. 2023, *A&A*, 672, A147
Tsai, S.-M., Lee, E. K. H., Powell, D., et al. 2023, *Nature*, 617, 483
Nasedkin, E., Mollière, P., Wang, J., et al. 2023, *A&A*, 678, A41
Balmer, W. O., Pueyo, L., Stolker, T., et al. 2023, *ApJ*, 956, 99
Blunt, S., Balmer, W. O., Wang, J. J., et al. 2023, *AJ*, 166, 257
Nowak, M., Lacour, S., Abuter, R., et al. 2024, *arXiv e-prints*, arXiv:2402.05019
Winterhalder, T. O., Lacour, S., Mérand, A., et al. 2024, *arXiv e-prints*, arXiv:2403.13055
Nasedkin, E., Mollière, P., & Blain, D. 2024, *JOSS*, 9, 5875
Nasedkin, E., Mollière, P., Lacour, S., et al. 2024, *arXiv e-prints*, arXiv:2404.03776

BIBLIOGRAPHY

- Ackerman, A. S., & Marley, M. S. 2001, [ApJ](#), **556**, 872
- Adams, F. C., Ruden, S. P., & Shu, F. H. 1989, [ApJ](#), **347**, 959
- Agol, E., Dorn, C., Grimm, S. L., et al. 2021, [PSJ](#), **2**, 1
- Ahrer, E.-M., Stevenson, K. B., Mansfield, M., et al. 2023, [Nature](#), **614**, 653
- Albert, J. G. 2020, [arXiv e-prints](#), [arXiv:2012.15286](#)
- Alderson, L., Wakeford, H. R., Alam, M. K., et al. 2023, [Nature](#), **614**, 664
- Allard, F., Guillot, T., Ludwig, H.-G., et al. 2003, in *Brown Dwarfs*, ed. E. Martín, Vol. 211, 325
- Allard, F., Hauschildt, P. H., Alexander, D. R., Tamanai, A., & Schweitzer, A. 2001, [ApJ](#), **556**, 357
- Allard, F., Homeier, D., & Freytag, B. 2012, [Philosophical Transactions of the Royal Society of London Series A](#), **370**, 2765
- Allard, N. F., Spiegelman, F., & Kielkopf, J. F. 2016, [A&A](#), **589**, A21
- Allard, N. F., Spiegelman, F., Leininger, T., & Mollière, P. 2019, [A&A](#), **628**, A120
- Allen, M., Yung, Y. L., & Waters, J. W. 1981, [J. Geophys. Res.](#), **86**, 3617
- ALMA Partnership, Brogan, C. L., Pérez, L. M., et al. 2015, [ApJ](#), **808**, L3
- Amara, A., & Quanz, S. P. 2012, [MNRAS](#), **427**, 948
- Amundsen, D. S., Tremblin, P., Manners, J., Baraffe, I., & Mayne, N. J. 2017, *Astronomy and Astrophysics*, **598**, A97
- Andrews, S. M., Huang, J., Pérez, L. M., et al. 2018, [ApJ](#), **869**, L41
- Antichi, J., Dohlen, K., Gratton, R. G., et al. 2009, [ApJ](#), **695**, 1042
- Apai, D., Kasper, M., Skemer, A., et al. 2016, [ApJ](#), **820**, 40
- Armitage, P. J. 2018, *A Brief Overview of Planet Formation* (Cham: Springer International Publishing), 2185
- Armitage, P. J. 2020, *Astrophysics of planet formation*, Second Edition (Cambridge University Press)
- Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, [ARA&A](#), **47**, 481, [_eprint: 0909.0948](#)
- Azzam, A. A. A., Yurchenko, S. N., Tennyson, J., & Naumenko, O. V. 2016, [MNRAS](#), **460**, 4063
- Baines, E. K., van Belle, G. T., ten Brummelaar, T. A., et al. 2007, [ApJ](#), **661**, L195
- Baines, E. K., White, R. J., Huber, D., et al. 2012, [ApJ](#), **761**, 57
- Bakos, G. Á., Lázár, J., Papp, I., Sári, P., & Green, E. M. 2002, [PASP](#), **114**, 974
- Balmer, W. O., Pueyo, L., Stolker, T., et al. 2023, [ApJ](#), **956**, 99
- Baraffe, I., Chabrier, G., Barman, T. S., Allard, F., & Hauschildt, P. H. 2003, [A&A](#), **402**, 701
- Barber, R. J., Strange, J. K., Hill, C., et al. 2014, [MNRAS](#), **437**, 1828
- Barman, T. S., Konopacky, Q. M., Macintosh, B., & Marois, C. 2015, [ApJ](#), **804**, 61, [_eprint: 1503.03539](#)
- Barman, T. S., Macintosh, B., Konopacky, Q. M., & Marois, C. 2011, [ApJ](#), **733**, 65
- Barrado, D., Mollière, P., Patapis, P., et al. 2023, [Nature](#), **624**, 263
- Barstow, J. K. 2020, [MNRAS](#), **497**, 4183
- Barstow, J. K., Aigrain, S., Irwin, P. G. J., Kendrew, S., & Fletcher, L. N. 2015, [MNRAS](#), **448**, 2546
- Barton, E. J., Yurchenko, S. N., & Tennyson, J. 2013, [MNRAS](#), **434**, 1469
- Baruteau, C., Meru, F., & Paardekooper, S.-J. 2011, [MNRAS](#), **416**, 1971
- Baudino, J. L., Bézard, B., Boccaletti, A., et al. 2015, [A&A](#), **582**, A83
- Baudino, J.-L., Mollière, P., Venot, O., et al. 2017, [ApJ](#), **850**, 150
- Bayo, A., Rodrigo, C., Barrado Y Navascués, D., et al. 2008, [A&A](#), **492**, 277
- Bell, T. J., Welbanks, L., Schlawin, E., et al. 2023, [Nature](#), **623**, 709

- Benisty, M., Bae, J., Facchini, S., et al. 2021, *ApJ*, **916**, L2
- Benneke, B., & Seager, S. 2012, *ApJ*, **753**, 100
- Benneke, B., & Seager, S. 2013, *ApJ*, **778**, 153
- Bergfors, C., Brandner, W., Janson, M., Köhler, R., & Henning, T. 2011, *A&A*, **528**, A134
- Best, W. M. J., Liu, M. C., Magnier, E. A., & Dupuy, T. J. 2020a, *AJ*, **159**, 257
- Best, W. M. J., Liu, M. C., Magnier, E. A., & Dupuy, T. J. 2020b, VizieR Online Data Catalog, [J/AJ/159/257](https://vizier.cesr-bcm.fr/vizieR/159/257)
- Best, W. M. J., Magnier, E. A., Liu, M. C., et al. 2018, *ApJS*, **234**, 1
- Beuzit, J.-L., Feldt, M., Dohlen, K., et al. 2008, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 7014, Proc. SPIE, 701418
- Beuzit, J. L., Vigan, A., Mouillet, D., et al. 2019, *A&A*, **631**, A155, _eprint: 1902.04080
- Billier, B. A., Apai, D., Bonnefoy, M., et al. 2021, *MNRAS*, **503**, 743
- Birnstiel, T. 2023, *arXiv e-prints*, [arXiv:2312.13287](https://arxiv.org/abs/2312.13287)
- Bitsch, B., Izidoro, A., Johansen, A., et al. 2019, *A&A*, **623**, A88
- Bitsch, B., Lambrechts, M., & Johansen, A. 2015, *A&A*, **582**, A112
- Bitsch, B., & Mah, J. 2023, *A&A*, **679**, A11
- Blain, D., Charnay, B., & Bézard, B. 2021, *A&A*, **646**, A15
- Blain, D., Sánchez-López, A., & Mollière, P. 2024, *AJ*, **167**, 179
- Blecic, J., Harrington, J., Cubillos, P., et al. 2016, in *American Astronomical Society Meeting Abstracts*, Vol. 227, American Astronomical Society Meeting Abstracts #227, 212.02
- Blunt, S., Balmer, W. O., Wang, J. J., et al. 2023, *AJ*, **166**, 257
- Boccaletti, A., M., P., B., et al. 2023, *arXiv e-prints*, [arXiv:2310.13414](https://arxiv.org/abs/2310.13414)
- Bodenheimer, P. 1974, *icarus*, **23**, 319
- Bodenheimer, P., Hubickyj, O., & Lissauer, J. J. 2000, *Icarus*, **143**, 2
- Bohn, A. J., Kenworthy, M. A., Ginski, C., et al. 2020, *ApJ*, **898**, L16
- Bohren, C. F., & Huffman, D. R. 1998, *Absorption and Scattering of Light by Small Particles* (John Wiley & Sons, Ltd), <https://onlinelibrary.wiley.com/doi/pdf/10.1002/9783527618156.fmatter>
- Bolton, S. J., Lunine, J., Stevenson, D., et al. 2017, *Space Sci. Rev.*, **213**, 5
- Bonnefoy, M., Currie, T., Marleau, G. D., et al. 2014a, *A&A*, **562**, A111
- Bonnefoy, M., Currie, T., Marleau, G. D., et al. 2014b, *A&A*, **562**, A111
- Bonnefoy, M., Zurlo, A., Baudino, J. L., et al. 2016, *A&A*, **587**, A58
- Bonnefoy, M., Perraut, K., Lagrange, A. M., et al. 2018, *A&A*, **618**, A63
- Borucki, W. J., Koch, D., Basri, G., et al. 2010, *Science*, **327**, 977
- Boss, A. P. 1997, *Science*, **276**, 1836
- Boss, A. P. 1998, *ApJ*, **503**, 923
- Bowler, B. P., Liu, M. C., Dupuy, T. J., & Cushing, M. C. 2010, *ApJ*, **723**, 850
- Brandt, G. M., Brandt, T. D., Dupuy, T. J., Michalik, D., & Marleau, G.-D. 2021, *ApJ*, **915**, L16
- Brown-Sevilla, S. B., Maire, A. L., Mollière, P., et al. 2023, *A&A*, **673**, A98
- Brygoo, S., Loubeyre, P., Millot, M., et al. 2021, *Nature*, **593**, 517
- Buchner, J. 2019, *PASP*, **131**, 108005
- Buchner, J. 2023, *Statistics Surveys*, **17**, 169
- Buchner, J., Georgakakis, A., Nandra, K., et al. 2014, *A&A*, **564**, A125, _eprint: 1402.0004
- Buenzli, E., Apai, D., Morley, C. V., et al. 2012, *ApJ*, **760**, L31
- Buglia, J. J. 1986, *Introduction to the theory of atmospheric radiative transfer*

- Burgasser, A. J., Geballe, T. R., Leggett, S. K., Kirkpatrick, J. D., & Golimowski, D. A. 2006, *ApJ*, **637**, 1067
- Burgasser, A. J., McElwain, M. W., Kirkpatrick, J. D., et al. 2004, *AJ*, **127**, 2856, [_eprint: astro-ph/0402325](#)
- Burgasser, A. J., Simcoe, R. A., Bochanski, J. J., et al. 2010, *ApJ*, **725**, 1405
- Burningham, B., Marley, M. S., Line, M. R., et al. 2017, *Monthly Notices of the Royal Astronomical Society*, **470**, 1177, [arXiv: 1701.01257](#)
- Burningham, B., Faherty, J. K., Gonzales, E. C., et al. 2021, *MNRAS*, **506**, 1944
- Burrows, A., Heng, K., & Nampaisarn, T. 2011, *ApJ*, **736**, 47
- Burrows, A., Hubbard, W. B., Lunine, J. I., et al. 1996, *Nuclear Physics B Proceedings Supplements*, **Vol. 51**, 51, 76
- Burrows, A., Hubbard, W. B., Lunine, J. I., & Liebert, J. 2001, *Reviews of Modern Physics*, **73**, 719
- Burrows, A., & Sharp, C. M. 1999, *ApJ*, **512**, 843, [_eprint: astro-ph/9807055](#)
- Burrows, A., Sudarsky, D., & Hubeny, I. 2006, *ApJ*, **640**, 1063
- Burrows, A., & Volobuyev, M. 2003, *ApJ*, **583**, 985
- Burrows, A., Marley, M., Hubbard, W. B., et al. 1997, *ApJ*, **491**, 856, [_eprint: astro-ph/9705201](#)
- Calamari, E., Faherty, J. K., Visscher, C., et al. 2024, *ApJ*, **963**, 67
- Cameron, A. G. W. 1978, *Moon and Planets*, **18**, 5
- Campbell, B., Walker, G. A. H., & Yang, S. 1988, *ApJ*, **331**, 902
- Cannon, A. J., & Pickering, E. C. 1993, *VizieR Online Data Catalog*, III/135A
- Cantalloube, F., Mouillet, D., Mugnier, L. M., et al. 2015, *A&A*, **582**, A89
- Cantalloube, F., Gomez-Gonzalez, C., Absil, O., et al. 2020a, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, **Vol. 11448**, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 114485A
- Cantalloube, F., Farley, O. J. D., Milli, J., et al. 2020b, *A&A*, **638**, A98
- Carlotti, A. 2013, *A&A*, **551**, A10
- Carlotti, A., Pueyo, L., & Mawet, D. 2014, *A&A*, **566**, A31
- Carter, A., Absil, O., Balmer, W., et al. 2024, *Into The Spotlight: Unveiling Wide-Separation Sub-Jupiters for Future JWST Characterization*, JWST Proposal. Cycle 3, ID. #5835
- Carter, A. L., Hinkley, S., Kammerer, J., et al. 2023, *ApJ*, **951**, L20
- Chabrier, G., Baraffe, I., Allard, F., & Hauschildt, P. 2000, *ApJ*, **542**, 464
- Chabrier, G., Johansen, A., Janson, M., & Rafikov, R. 2014, in *Protostars and Planets VI*, ed. **H. Beuther, R. S. Klessen, C. P. Dullemond, & T. Henning**, 619
- Charbonneau, D., Brown, T. M., Noyes, R. W., & Gilliland, R. L. 2002, *apj*, **568**, 377
- Charnay, B., Bézard, B., Baudino, J.-L., et al. 2018, *ApJ*, **854**, 172
- Chauvin, G. 2024, *Comptes Rendus Physique*, **24**, 139
- Chauvin, G., Lagrange, A. M., Dumas, C., et al. 2005, *A&A*, **438**, L25
- Chen, K., Kama, M., Pinilla, P., & Keyte, L. 2024, *MNRAS*, **527**, 2049
- Christiaens, V., Gonzalez, C., Farkas, R., et al. 2023, *JOSS*, **8**, 4774
- Chubb, K. L., & Min, M. 2022, *A&A*, **665**, A2
- Chubb, K. L., Tennyson, J., & Yurchenko, S. N. 2020, *MNRAS*, **493**, 1531
- Chubb, K. L., Rocchetto, M., Yurchenko, S. N., et al. 2021, *A&A*, **646**, A21
- Coles, P. A., Yurchenko, S. N., & Tennyson, J. 2019, *MNRAS*, **490**, 4638
- Constantinou, S., & Madhusudhan, N. 2024, [arXiv e-prints](#), [arXiv:2403.04825](#)
- Constantinou, S., Madhusudhan, N., & Gandhi, S. 2023, *ApJ*, **943**, L10

- Contro, B., Wittenmyer, R. A., Horner, J., & Marshall, J. P. 2015, *Origins of Life and Evolution of the Biosphere*, 45, 41
- Covarrubias, S., Blunt, S., & Wang, J. J. 2022, *Research Notes of the American Astronomical Society*, 6, 66
- Cridland, A. J., Eistrup, C., & van Dishoeck, E. F. 2019, *A&A*, 627, A127
- Cridland, A. J., van Dishoeck, E. F., Alessi, M., & Pudritz, R. E. 2020, *A&A*, 642, A229
- Cubillos, P. E., & Bleic, J. 2021, *MNRAS*, 505, 2675
- Cugno, G., Patapis, P., Stolker, T., et al. 2021, *A&A*, 653, A12
- Currie, T., Biller, B., Lagrange, A., et al. 2023a, in *Astronomical Society of the Pacific Conference Series*, Vol. 534, *Protostars and Planets VII*, ed. S. Inutsuka, Y. Aikawa, T. Muto, K. Tomida, & M. Tamura, 799
- Currie, T., Biller, B., Lagrange, A., et al. 2023b, in *Astronomical Society of the Pacific Conference Series*, Vol. 534, *Protostars and Planets VII*, ed. S. Inutsuka, Y. Aikawa, T. Muto, K. Tomida, & M. Tamura, 799
- Currie, T., Fukagawa, M., Thalmann, C., Matsumura, S., & Plavchan, P. 2012, *ApJ*, 755, L34
- Currie, T., Burrows, A., Itoh, Y., et al. 2011, *ApJ*, 729, 128, [_eprint: 1101.1973](#)
- Currie, T., Burrows, A., Girard, J. H., et al. 2014, *apj*, 795, 133
- Cushing, M. C., Roellig, T. L., Marley, M. S., et al. 2006, *ApJ*, 648, 614
- Cushing, M. C., Marley, M. S., Saumon, D., et al. 2008, *ApJ*, 678, 1372
- Cutri, R. M., Skrutskie, M. F., van Dyk, S., et al. 2003, *VizieR Online Data Catalog*, II/246
- Cutri, R. M., Wright, E. L., Conrow, T., et al. 2021, *VizieR Online Data Catalog*, II/328
- Daemgen, S., Todorov, K., Quanz, S. P., et al. 2017, *A&A*, 608, A71
- Danti, C., Bitsch, B., & Mah, J. 2023, *A&A*, 679, L7
- Davies, R., & Kasper, M. 2012, *ARA&A*, 50, 305
- Debras, F., & Chabrier, G. 2019, *ApJ*, 872, 100
- Delrez, L., Santerne, A., Almenara, J. M., et al. 2016, *MNRAS*, 458, 4025
- Demars, D., Bonnefoy, M., Dougados, C., et al. 2023, *A&A*, 676, A123
- Deming, D., Seager, S., Richardson, L. J., & Harrington, J. 2005, *Nature*, 434, 740
- Desgrange, C., Chauvin, G., Christiaens, V., et al. 2022, *A&A*, 664, A139
- Desidera, S., Chauvin, G., Bonavita, M., et al. 2021, *A&A*, 651, A70
- Desidera, S., Chauvin, G., Bonavita, M., et al. 2021, *A&A*, 651, A70
- Dodson-Robinson, S. E., Veras, D., Ford, E. B., & Beichman, C. A. 2009, *ApJ*, 707, 79
- Doelman, D. S., Snik, F., Por, E. H., et al. 2021, *Appl. Opt.*, 60, D52
- Doelman, D. S., Stone, J. M., Briesemeister, Z. W., et al. 2022, *AJ*, 163, 217
- Drażkowska, J., Bitsch, B., Lambrechts, M., et al. 2023, in *Astronomical Society of the Pacific Conference Series*, Vol. 534, *Protostars and Planets VII*, ed. S. Inutsuka, Y. Aikawa, T. Muto, K. Tomida, & M. Tamura, 717
- Dupuy, T. J., & Kraus, A. L. 2013, *Science*, 341, 1492
- Dupuy, T. J., & Liu, M. C. 2012, *ApJS*, 201, 19
- Dyrek, A., Min, M., Decin, L., et al. 2024, *Nature*, 625, 51
- Eddington, A. S. 1930, *MNRAS*, 90, 279
- Eistrup, C., Walsh, C., & van Dishoeck, E. F. 2018, *A&A*, 613, A14
- Eriksson, L. E. J., Ronnet, T., Johansen, A., et al. 2022, *A&A*, 661, A73
- Esposito, S., Mesa, D., Skemer, A., et al. 2013, *A&A*, 549, A52
- Faedi, F., Barros, S. C. C., Anderson, D. R., et al. 2011, *A&A*, 531, A40

- Faherty, J. K. 2018, in *Handbook of Exoplanets*, ed. H. J. Deeg & J. A. Belmonte (Springer, Cham), 188
- Faherty, J. K., Riedel, A. R., Cruz, K. L., et al. 2016, *ApJS*, **225**, 10
- Faramaz, V., Marino, S., Booth, M., et al. 2021, *AJ*, **161**, 271
- Feautrier, P. 1964, *Comptes Rendus Academie des Sciences (serie non specifiée)*, **258**, 3189
- Fegley, Bruce, J., & Lodders, K. 1996, *ApJ*, **472**, L37
- Feinstein, A. D., Radica, M., Welbanks, L., et al. 2023, *Nature*, **614**, 670
- Feroz, F., & Hobson, M. P. 2008, *MNRAS*, **384**, 449, _eprint: 0704.3704
- Feroz, F., Hobson, M. P., & Bridges, M. 2009, *MNRAS*, **398**, 1601, _eprint: 0809.3437
- Feroz, F., Hobson, M. P., Cameron, E., & Pettitt, A. N. 2019, *The Open Journal of Astrophysics*, **2**, 10
- Feroz, F., Hobson, M. P., Cameron, E., & Pettitt, A. N. 2019, *The Open Journal of Astrophysics*, **2**, 10
- Fischer, D. A., Howard, A. W., Laughlin, G. P., et al. 2014, in *Protostars and Planets VI*, ed. H. Beuther, R. S. Klessen, C. P. Dullemond, & T. Henning, 715
- Fischer, D. A., & Valenti, J. 2005, *ApJ*, **622**, 1102
- Fischer, P. D., Knutson, H. A., Sing, D. K., et al. 2016, *ApJ*, **827**, 19
- Flasseur, O., Denis, L., Thiébaud, É., & Langlois, M. 2018, *A&A*, **618**, A138
- Flasseur, O., Denis, L., Thiébaud, É., & Langlois, M. 2020, *A&A*, **637**, A9
- Fletcher, L. N., Orton, G. S., Teanby, N. A., & Irwin, P. G. J. 2009, *Icarus*, **202**, 543
- Fletcher, L. N., King, O. R. T., Harkett, J., et al. 2023, *Journal of Geophysical Research (Planets)*, **128**, e2023JE007924
- Follette, K. B. 2023, *PASP*, **135**, 093001
- Fonte, S., Turrini, D., Pacetti, E., et al. 2023, *MNRAS*, **520**, 4683
- Fortney, J. J. 2005, *MNRAS*, **364**, 649
- Fortney, J. J., Ikoma, M., Nettelmann, N., Guillot, T., & Marley, M. S. 2011, *ApJ*, **729**, 32
- Fragoso, T. M., Bertoli, W., & Louzada, F. 2018, *International Statistical Review*, **86**, 1
- Freedman, R. S., Lustig-Yaeger, J., Fortney, J. J., et al. 2014, *ApJS*, **214**, 25
- Fukagawa, M., Itoh, Y., Tamura, M., et al. 2009, *ApJ*, **696**, L1
- Gaia Collaboration. 2020, *VizieR Online Data Catalog*, I/350
- Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, *A&A*, **616**, A1, _eprint: 1804.09365
- Galicher, R., Marois, C., Macintosh, B., Barman, T., & Konopacky, Q. 2011, *ApJ*, **739**, L41
- Galicher, R., & Mazoyer, J. 2024, *Comptes Rendus Physique*, **24**, 133
- Gammie, C. F. 2001, *ApJ*, **553**, 174
- Gao, P., Thorngren, D. P., Lee, E. K. H., et al. 2020, *Nature Astronomy*, **4**, 951
- Garufi, A., Meeus, G., Benisty, M., et al. 2017a, *A&A*, **603**, A21
- Garufi, A., Benisty, M., Stolker, T., et al. 2017b, *The Messenger*, **169**, 32
- Gaudi, B. S., Meyer, M., & Christiansen, J. 2021, in *ExoFrontiers*, **2514-3433** (IOP Publishing), 2
- Gebhard, T. D., Bonse, M. J., Quanz, S. P., & Schölkopf, B. 2022, *A&A*, **666**, A9
- Gehren, T. 1977, *A&A*, **59**, 303
- Geiler, F., Krivov, A. V., Booth, M., & Löhne, T. 2018, *MNRAS*, **483**, 332
- Gerard, B. L., & Marois, C. 2016, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 9909, *Adaptive Optics Systems V*, ed. E. Marchetti, L. M. Close, & J.-P. Véran, 990958
- Giacobbe, P., Brogi, M., Gandhi, S., et al. 2021, *Nature*, **592**, 205
- Gomez Gonzalez, C. A., Absil, O., Absil, P. A., et al. 2016, *A&A*, **589**, A54

- Gomez Gonzalez, C. A., Absil, O., & Van Droogenbroeck, M. 2018, *A&A*, **613**, A71
- Gontcharov, G. A. 2006, *Astronomy Letters*, **32**, 759
- Goody, R., West, R., Chen, L., & Crisp, D. 1989, *J. Quant. Spec. Radiat. Transf.*, **42**, 539
- Gordon, I., Rothman, L., Hargreaves, R., et al. 2022, *Journal of Quantitative Spectroscopy and Radiative Transfer*, **277**, 107949
- Gordon, S., & McBride, B. J. 1994, Computer program for calculation of complex chemical equilibrium compositions and applications (NASA Reference Publication)
- Goździewski, K., & Migaszewski, C. 2018, *ApJS*, **238**, 6
- Grant, D., Lewis, N. K., Wakeford, H. R., et al. 2023, *ApJ*, **956**, L29
- Gravity Collaboration, Abuter, R., Accardo, M., et al. 2017, *A&A*, **602**, A94, [_eprint: 1705.02345](#)
- Gravity Collaboration, Lacour, S., Nowak, M., et al. 2019, *A&A*, **623**, L11, [_eprint: 1903.11903](#)
- GRAVITY Collaboration, Bauböck, M., Dexter, J., et al. 2020, *A&A*, **635**, A143
- Gravity Collaboration, Nowak, M., Lacour, S., et al. 2020, *A&A*, **633**, A110, [_eprint: 1912.04651](#)
- Gray, R. O., & Corbally, C. J. 2002, *AJ*, **124**, 989
- Gray, R. O., Corbally, C. J., Garrison, R. F., McFadden, M. T., & Robinson, P. E. 2003a, *AJ*, **126**, 2048
- Gray, R. O., Corbally, C. J., Garrison, R. F., McFadden, M. T., & Robinson, P. E. 2003b, *AJ*, **126**, 2048
- Gray, R. O., & Kaye, A. B. 1999, *AJ*, **118**, 2993
- Greco, J. P., & Brandt, T. D. 2016, *ApJ*, **833**, 134
- Greenbaum, A. Z., Pueyo, L., Ruffio, J.-B., et al. 2018, *AJ*, **155**, 226, [_eprint: 1804.07774](#)
- Groff, T., Chilcote, J., Brandt, T., et al. 2017, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 10400, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, ed. S. Shaklan, 1040016
- Groff, T. D., Kasdin, N. J., Limbach, M. A., et al. 2015, in *Techniques and Instrumentation for Detection of Exoplanets VII*, ed. S. Shaklan, Vol. 9605 (SPIE), 96051C, backup Publisher: International Society for Optics and Photonics
- Guillot, T. 1995, *Science*, **269**, 1697
- Guillot, T. 2005, *Annual Review of Earth and Planetary Sciences*, **33**, 493
- Guillot, T. 2010, *A&A*, **520**, A27
- Haffert, S. Y., Bohn, A. J., de Boer, J., et al. 2019, *Nature Astronomy*, **3**, 749
- Haniiff, C. 2007, *New A Rev.*, **51**, 565
- Hansen, J. E. 1971, *Journal of Atmospheric Sciences*, **28**, 1400
- Hargreaves, R. J., Gordon, I. E., Rey, M., et al. 2020, *ApJS*, **247**, 55
- Hartmann, J.-M., Tran, H., Armante, R., et al. 2018, *J. Quant. Spec. Radiat. Transf.*, **213**, 178
- Hastings, W. K. 1970, *Biometrika*, **57**, 97
- Hauschildt, P. H., Allard, F., & Baron, E. 1999, *ApJ*, **512**, 377
- Hauschildt, P. H., & Baron, E. 1999, *Journal of Computational and Applied Mathematics*, **109**, 41
- Helled, R., Anderson, J. D., Podolak, M., & Schubert, G. 2011, *ApJ*, **726**, 15
- Helled, R., & Bodenheimer, P. 2010, *Icarus*, **207**, 503
- Helled, R., Nettelmann, N., & Guillot, T. 2020a, *Space Sci. Rev.*, **216**, 38
- Helled, R., Nettelmann, N., & Guillot, T. 2020b, *Space Sci. Rev.*, **216**, 38
- Helling, C. 2019, *Annual Review of Earth and Planetary Sciences*, **47**, 583
- Helling, C., & Casewell, S. 2014, *A&A Rev.*, **22**, 80
- Helling, C., Oevermann, M., Lüttke, M. J. H., Klein, R., & Sedlmayr, E. 2001, *A&A*, **376**, 194
- Helling, C., & Woitke, P. 2006, *A&A*, **455**, 325

- Helling, C., Woitke, P., & Thi, W. F. 2008a, *A&A*, **485**, 547
- Helling, C., Ackerman, A., Allard, F., et al. 2008b, *MNRAS*, **391**, 1854
- Helling, C., Iro, N., Corrales, L., et al. 2019, *A&A*, **631**, A79, arXiv: 1906.08127
- Hinkley, S., Lacour, S., Marleau, G. D., et al. 2023, *A&A*, **671**, L5
- Hoch, K. K. W., Konopacky, Q. M., Barman, T. S., et al. 2022, *AJ*, **164**, 155
- Hoch, K. K. W., Konopacky, Q. M., Theissen, C. A., et al. 2023, *AJ*, **166**, 85
- Hoeijmakers, H. J., Schwarz, H., Snellen, I. A. G., et al. 2018, *A&A*, **617**, A144
- Hou Yip, K., Nikolaou, N., Coronica, P., et al. 2019, arXiv e-prints, arXiv:1904.06155
- Ida, S., & Lin, D. N. C. 2004, *ApJ*, **616**, 567
- Ih, J., & Kempton, E. M. R. 2021, *AJ*, **162**, 237
- Ingraham, P., Marley, M. S., Saumon, D., et al. 2014, *ApJ*, **794**, L15, eprint: 1409.5456
- Janson, M., Bergfors, C., Goto, M., Brandner, W., & Lafrenière, D. 2010, *ApJ*, **710**, L35
- Johnson, T. V., Yeates, C. M., & Young, R. 1992, *Space Sci. Rev.*, **60**, 3
- Jullion, A., & Lambert, P. 2007, *Computational Statistics & Data Analysis*, **51**, 2542
- JWST Transiting Exoplanet Community Early Release Science Team, Ahrer, E.-M., Alderson, L., et al. 2023, *Nature*, **614**, 649
- Kammerer, J., Lacour, S., Stolker, T., et al. 2021, *A&A*, **652**, A57
- Karalidi, T., Marley, M., Fortney, J. J., et al. 2021, *ApJ*, **923**, 269
- Karman, T., Gordon, I. E., van der Avoird, A., et al. 2019, *Icarus*, **328**, 160
- Kass, R. E., & Raftery, A. E. 1995, *Journal of the American Statistical Association*, **90**, 773
- Kaushik, M., Mattoo, A., & Rastogi, R. 2024, arXiv e-prints, arXiv:2404.09143
- Kawahara, H., Kawashima, Y., Masuda, K., et al. 2022, *ApJS*, **258**, 31
- Kaye, A. B., Handler, G., Krisciunas, K., Poretti, E., & Zerbini, F. M. 1999, *PASP*, **111**, 840
- Kepler, M., Benisty, M., Müller, A., et al. 2018, *A&A*, **617**, A44
- Kiefer, S., Bohn, A. J., Quanz, S. P., Kenworthy, M., & Stolker, T. 2021, *A&A*, **652**, A33
- Kirkpatrick, J. D. 2005, *ARA&A*, **43**, 195
- Kirkpatrick, J. D., Reid, I. N., Liebert, J., et al. 1999, *ApJ*, **519**, 802
- Kitzmann, D., Heng, K., Oreshenko, M., et al. 2020, *ApJ*, **890**, 174, eprint: 1910.01070
- Kohlhase, C. E., & Penzo, P. A. 1977, *Space Sci. Rev.*, **21**, 77
- Konopacky, Q. M., Barman, T. S., Macintosh, B. A., & Marois, C. 2013, *Science*, **339**, 1398
- Kreidberg, L., Bean, J. L., Désert, J.-M., et al. 2014, *Nature*, **505**, 69
- Kuzuhara, M., Tamura, M., Kudo, T., et al. 2013, *ApJ*, **774**, 11
- Labeyrie, A. 1975, *ApJ*, **196**, L71
- Lacis, A. A., & Oinas, V. 1991, *J. Geophys. Res.*, **96**, 9027
- Lacour, S., Wang, J. J., Nowak, M., et al. 2020, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 11446, *Optical and Infrared Interferometry and Imaging VII*, ed. P. G. Tuthill, A. Mérand, & S. Sallum, 11444600
- Lacour, S., Wang, J. J., Rodet, L., et al. 2021, *A&A*, **654**, L2
- Lacy, B. I., & Burrows, A. 2020, *ApJ*, **905**, 131
- Lada, C. J. 1987, in *Star Forming Regions*, ed. M. Peimbert & J. Jugaku, Vol. 115, 1
- Lafrenière, D., Marois, C., Doyon, R., & Barman, T. 2009, *ApJ*, **694**, L148
- Lafrenière, D., Marois, C., Doyon, R., Nadeau, D., & Artigau, É. 2007, *ApJ*, **660**, 770
- Lagrange, A. M., Bonnefoy, M., Chauvin, G., et al. 2010, *Science*, **329**, 57
- Lagrange, A. M., Rubini, P., Nowak, M., et al. 2020, *A&A*, **642**, A18
- Lambrechts, M., & Johansen, A. 2012, *A&A*, **544**, A32
- Lambrechts, M., & Johansen, A. 2014, *A&A*, **572**, A107

- Lang, S., & Brezger, A. 2004, *Journal of Computational and Graphical Statistics*, 13, 183
- Langlois, M., Gratton, R., Lagrange, A. M., et al. 2021, *A&A*, 651, A71
- Lapeyrère, V., Kervella, P., Lacour, S., et al. 2014, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 9146, *Optical and Infrared Interferometry IV*, ed. J. K. Rajagopal, M. J. Creech-Eakman, & F. Malbet, 91462D
- Larkin, J., Barczys, M., Krabbe, A., et al. 2006, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 6269, *Ground-based and Airborne Instrumentation for Astronomy*, ed. I. S. McLean & M. Iye, 62691A
- Larkin, J. E., Chilcote, J. K., Aliado, T., et al. 2014, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 9147, *Ground-based and Airborne Instrumentation for Astronomy V*, ed. S. K. Ramsay, I. S. McLean, & H. Takami, 91471K
- Laughlin, G., & Rozyczka, M. 1996, *ApJ*, 456, 279
- Lavie, B., Mendonça, J. M., Mordasini, C., et al. 2017, *AJ*, 154, 91, _eprint: 1610.03216
- Leconte, J. 2021, *A&A*, 645, A20
- Lee, E. K. H., Tan, X., & Tsai, S.-M. 2023, *MNRAS*, 523, 4477
- Lee, E. K. H., Tan, X., & Tsai, S.-M. 2024, *MNRAS*, arXiv:2311.16722 [astro-ph.EP]
- Lee, J.-M., Heng, K., & Irwin, P. G. J. 2013, *ApJ*, 778, 97, _eprint: 1307.1404
- Leggett, S. K., Saumon, D., Albert, L., et al. 2008, *ApJ*, 682, 1256
- Lim, O., Benneke, B., Doyon, R., et al. 2023, *ApJ*, 955, L22
- Linder, E. F., Mordasini, C., Mollière, P., et al. 2019, *A&A*, 623, A85
- Line, M. R., Fortney, J. J., Marley, M. S., & Sorahana, S. 2014, *ApJ*, 793, 33, _eprint: 1403.6412
- Line, M. R., Teske, J., Burningham, B., Fortney, J. J., & Marley, M. S. 2015, *ApJ*, 807, 183, _eprint: 1504.06670
- Liu, B., Ormel, C. W., & Johansen, A. 2019, *A&A*, 624, A114
- Liu, M. C., Dupuy, T. J., & Allers, K. N. 2016, *ApJ*, 833, 96
- Liu, P., Bohn, A. J., Doelman, D. S., et al. 2023, *A&A*, 674, A115
- Liu, T., Li, P. S., Juvela, M., et al. 2018, *ApJ*, 859, 151
- Lodders, K. 2003, *ApJ*, 591, 1220
- Lodders, K., & Fegley, B. 2002, *Icarus*, 155, 393
- Lodders, K., Palme, H., & Gail, H. P. 2009, *Landolt Börstein*, 4B, 712
- Looper, D. L., Kirkpatrick, J. D., Cutri, R. M., et al. 2008, *ApJ*, 686, 528
- Luhman, K. L., Tremblin, P., Alves de Oliveira, C., et al. 2024, *AJ*, 167, 5
- Luna, J. L., & Morley, C. V. 2021, *ApJ*, 920, 146
- Lyo, B. 1931, *L'Astronomie*, 45, 248
- Maaskant, K. M., Honda, M., Waters, L. B. F. M., et al. 2013, *A&A*, 555, A64
- MacDonald, R. J., & Batalha, N. E. 2023, *Research Notes of the American Astronomical Society*, 7, 54
- MacDonald, R. J., & Lewis, N. K. 2022, *ApJ*, 929, 20
- Maciejewski, G., Ginski, C., Gilbert, H., Mugrauer, M., & Neuhäuser, R. 2016a, *Information Bulletin on Variable Stars*, 6177, 1
- Maciejewski, G., Dimitrov, D., Mancini, L., et al. 2016b, *Acta Astron.*, 66, 55
- Macintosh, B., Graham, J. R., Ingraham, P., et al. 2014, *Proceedings of the National Academy of Science*, 111, 12661, _eprint: 1403.7520
- Macintosh, B., Graham, J. R., Barman, T., et al. 2015, *Science*, 350, 64
- Madhusudhan, N. 2019, *Annu. Rev. Astron. Astrophys.*, 57, 617
- Madhusudhan, N., Burrows, A., & Currie, T. 2011, *ApJ*, 737, 34
- Madhusudhan, N., Sarkar, S., Constantinou, S., et al. 2023, *ApJ*, 956, L13

- Madhusudhan, N., & Seager, S. 2009, *ApJ*, 707, 24, [_eprint: 0910.1347](#)
- Madhusudhan, N., & Seager, S. 2011, *ApJ*, 729, 41
- Mahalanobis, P. 1936, *Proceedings of the National Institute of Sciences of India*, 2, 49–55
- Maire, A. L., Skemer, A. J., Hinz, P. M., et al. 2015, *A&A*, 576, A133
- Maire, J., Gagné, J., Lafrenière, D., et al. 2012, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 8447, *Adaptive Optics Systems III*, ed. B. L. Ellerbroek, E. Marchetti, & J.-P. Véran, 844760
- Maldonado, J., & Villaver, E. 2017, *A&A*, 602, A38
- Mancini, L., Esposito, M., Covino, E., et al. 2018, *A&A*, 613, A41
- Manjavacas, E., Tremblin, P., Birkmann, S., et al. 2024, *AJ*, 167, 168
- Mansfield, M., Line, M. R., Bean, J. L., et al. 2021, *Nature Astronomy*, 5, 1224
- Marleau, G.-D., Klahr, H., Kuiper, R., & Mordasini, C. 2017, *ApJ*, 836, 221
- Marleau, G.-D., Mordasini, C., & Kuiper, R. 2019, *ApJ*, 881, 144
- Marley, M. S., Fortney, J., Seager, S., & Barman, T. 2006, [arXiv:astro-ph/0602468](#), [arXiv: astro-ph/0602468](#)
- Marley, M. S., Fortney, J. J., Hubickyj, O., Bodenheimer, P., & Lissauer, J. J. 2007, *ApJ*, 655, 541, [_eprint: astro-ph/0609739](#)
- Marley, M. S., & Robinson, T. D. 2015, *ARA&A*, 53, 279
- Marley, M. S., Saumon, D., Cushing, M., et al. 2012, *ApJ*, 754, 135, [_eprint: 1205.6488](#)
- Marley, M. S., Saumon, D., & Goldblatt, C. 2010, *ApJ*, 723, L117, [_eprint: 1009.6217](#)
- Marley, M. S., Saumon, D., Visscher, C., et al. 2021, *ApJ*, 920, 85
- Marocco, F., Day-Jones, A. C., Lucas, P. W., et al. 2014, *MNRAS*, 439, 372
- Marois, C., Correia, C., Galicher, R., et al. 2014, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 9148, *Adaptive Optics Systems IV*, ed. E. Marchetti, L. M. Close, & J.-P. Vran, 91480U
- Marois, C., Correia, C., Véran, J.-P., & Currie, T. 2014, in *Exploring the Formation and Evolution of Planetary Systems*, ed. M. Booth, B. C. Matthews, & J. R. Graham, Vol. 299, 48
- Marois, C., Doyon, R., Racine, R., et al. 2005, *JRASC*, 99, 130
- Marois, C., Lafrenière, D., Macintosh, B., & Doyon, R. 2008, *ApJ*, 673, 647
- Marois, C., Lafrenière, D., Doyon, R., Macintosh, B., & Nadeau, D. 2006, *ApJ*, 641, 556
- Marois, C., Macintosh, B., Barman, T., et al. 2008, *Science*, 322, 1348, [_eprint: 0811.2606](#)
- Marois, C., Zuckerman, B., Konopacky, Q. M., Macintosh, B., & Barman, T. 2010, *Nature*, 468, 1080, [_eprint: 1011.4918](#)
- Matson, D. L., Spilker, L. J., & Lebreton, J.-P. 2002, *Space Sci. Rev.*, 104, 1
- Matthews, B., Kennedy, G., Sibthorpe, B., et al. 2014, *ApJ*, 780, 97
- Mawet, D. 2007, in *In the Spirit of Bernard Lyot: The Direct Detection of Planets and Circumstellar Disks in the 21st Century*, ed. P. Kalas, 40
- Mawet, D., Riaud, P., Hanot, C., et al. 2007, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 6693, *Techniques and Instrumentation for Detection of Exoplanets III*, ed. D. R. Coulter, 66931M
- Mawet, D., Milli, J., Wahhaj, Z., et al. 2014, *ApJ*, 792, 97
- Mayor, M., & Queloz, D. 1995, *nat*, 378, 355
- McCarthy, A. M., Muirhead, P. S., Tamburo, P., et al. 2024, [arXiv e-prints](#), [arXiv:2402.15001](#)
- McClatchey, R. A. 1973, *AFCRL atmospheric absorption line parameters compilation*
- McKemmish, L. K., Masseron, T., Hoeijmakers, H. J., et al. 2019, *MNRAS*, 488, 2836
- McKemmish, L. K., Yurchenko, S. N., & Tennyson, J. 2016, *Mon. Not. R. Astron. Soc.*, 463, 771
- Meeus, G., Waters, L. B. F. M., Bouwman, J., et al. 2001, *A&A*, 365, 476

- Menzel, D. H., Coblentz, W. W., & Lampland, C. O. 1926, *ApJ*, **63**, 177
- Mesa, D., Gratton, R., Kervella, P., et al. 2023, *A&A*, **672**, A93
- Metchev, S., Marois, C., & Zuckerman, B. 2009, *ApJ*, **705**, L204
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. 1953, *J. Chem. Phys.*, **21**, 1087
- Michelson, A. A., & Pease, F. G. 1921, *ApJ*, **53**, 249
- Miguel, Y., & Vazan, A. 2023, *Remote Sensing*, **15**, 681
- Mihalas, D. 1970, *Stellar atmospheres* (WH Freeman)
- Mikal-Evans, T., Sing, D. K., Dong, J., et al. 2023, *ApJ*, **943**, L17
- Miles, B. E., Skemer, A. J. I., Morley, C. V., et al. 2020, *AJ*, **160**, 63, arXiv: 2004.10770
- Miles, B. E., Biller, B. A., Patapis, P., et al. 2023, *ApJ*, **946**, L6
- Milli, J., Mawet, D., Mouillet, D., Kasper, M., & Girard, J. H. 2016, in *Astrophysics and Space Science Library*, Vol. 439, *Astronomy at High Angular Resolution*, ed. H. M. J. Boffin, G. Hussain, J.-P. Berger, & L. Schmidtbreick, 17
- Milli, J., Banas, T., Mouillet, D., et al. 2016, in *Adaptive Optics Systems V*, ed. E. Marchetti, L. M. Close, & J.-P. Véran, Vol. 9909, International Society for Optics and Photonics (SPIE), 99094Z
- Milli, J., Kasper, M., Bourget, P., et al. 2018, in *Adaptive Optics Systems VI*, ed. L. M. Close, L. Schreiber, & D. Schmidt, Vol. 10703, International Society for Optics and Photonics (SPIE), 107032A
- Min, M., Hovenier, J. W., & de Koter, A. 2005, *A&A*, **432**, 909
- Mizuno, H. 1980, *Progress of Theoretical Physics*, **64**, 544
- Moe, M., & Kratter, K. M. 2021, *MNRAS*, **507**, 3593
- Mollière, P., Molyarova, T., Bitsch, B., et al. 2022, *ApJ*, **934**, 74
- Mollière, P. M. 2017, PhD thesis, Ruprecht-Karls University of Heidelberg, Germany
- Mollière, P., van Boekel, R., Bouwman, J., et al. 2017, *A&A*, **600**, A10, _eprint: 1611.08608
- Mollière, P., van Boekel, R., Dullemond, C., Henning, T., & Mordasini, C. 2015, *ApJ*, **813**, 47, _eprint: 1509.07523
- Mollière, P., Wardenier, J. P., van Boekel, R., et al. 2019, *A&A*, **627**, A67
- Mollière, P., Stolker, T., Lacour, S., et al. 2020, *A&A*, **640**, A131
- Mollière, P., Molyarova, T., Bitsch, B., et al. 2022, *ApJ*, **934**, 74
- Molter, E., de Pater, I., Sault, R., et al. 2019, in EPSC-DPS Joint Meeting 2019, Vol. 2019, EPSC
- Morbidelli, A., Szulágyi, J., Crida, A., et al. 2014, *Icarus*, **232**, 266
- Mordasini, C., Alibert, Y., Klahr, H., & Henning, T. 2012, *A&A*, **547**, A111
- Mordasini, C., & Burn, R. 2024, arXiv e-prints, arXiv:2404.15555
- Mordasini, C., Marleau, G.-D., & Mollière, P. 2017, *A&A*, **608**, A72
- Morley, C. V., Mukherjee, S., Marley, M. S., et al. 2024, arXiv e-prints, arXiv:2402.00758
- Moses, J. I., Madhusudhan, N., Visscher, C., & Freedman, R. S. 2013, *ApJ*, **763**, 25
- Moses, J. I., Marley, M. S., Zahnle, K., et al. 2016, *ApJ*, **829**, 66, arXiv: 1608.08643
- Mousavi-Sadr, M., Jassur, D. M., & Gozaliasl, G. 2023, *MNRAS*, **525**, 3469
- Moya, A., Amado, P. J., Barrado, D., et al. 2010, *MNRAS*, **405**, L81
- Moya, A., Amado, P. J., Barrado, D., et al. 2010, *MNRAS*, **406**, 566
- Mugnier, L. M., Cornia, A., Sauvage, J.-F., et al. 2009, *Journal of the Optical Society of America A*, **26**, 1326
- Mukherjee, P., Parkinson, D., & Liddle, A. R. 2006, *ApJ*, **638**, L51
- Mukherjee, S., Batalha, N. E., Fortney, J. J., & Marley, M. S. 2023, *ApJ*, **942**, 71
- Mukherjee, S., Fortney, J. J., Batalha, N. E., et al. 2022, *ApJ*, **938**, 107

- Mukherjee, S., Fortney, J. J., Morley, C. V., et al. 2024, *ApJ*, **963**, 73
- Nakajima, T., Oppenheimer, B. R., Kulkarni, S. R., et al. 1995, *Nature*, **378**, 463
- Nasedkin, E., Mollière, P., Wang, J., et al. 2023, *A&A*, **678**, A41
- Nasedkin, E., Mollière, P., & Blain, D. 2024, *JOSS*, **9**, 5875
- Nasedkin, E., Mollière, P., Lacour, S., et al. 2024, *arXiv e-prints*, [arXiv:2404.03776](https://arxiv.org/abs/2404.03776)
- Nero, D., & Bjorkman, J. E. 2009, *ApJ*, **702**, L163
- Nielsen, E. L., De Rosa, R. J., Macintosh, B., et al. 2019, *AJ*, **158**, 13
- Nikolov, N., Sing, D. K., Gibson, N. P., et al. 2016, *ApJ*, **832**, 191
- Nixon, M. C., Welbanks, L., McGill, P., & Kempton, E. M. R. 2023, *arXiv e-prints*, [arXiv:2310.03713](https://arxiv.org/abs/2310.03713)
- Noll, K. S., Geballe, T. R., & Marley, M. S. 1997, *ApJ*, **489**, L87
- Nowak, M., Lacour, S., Lagrange, A. M., et al. 2020, *A&A*, **642**, L2
- Nowak, M., Lacour, S., Abuter, R., et al. 2024, *arXiv e-prints*, [arXiv:2402.05019](https://arxiv.org/abs/2402.05019)
- Öberg, K. I., Facchini, S., & Anderson, D. E. 2023, *ARA&A*, **61**, 287
- Öberg, K. I., Murray-Clay, R., & Bergin, E. A. 2011, *ApJ*, **743**, L16
- Öberg, K. I., Guzmán, V. V., Walsh, C., et al. 2021, *ApJS*, **257**, 1
- Olagnon, M. 2011, ORDERPACK 2.0 – Unconditional, Unique, and Partial Ranking, Sorting, and Permutation Downloadable Fortran 90 source code
- Oppenheimer, B. R., Baranec, C., Beichman, C., et al. 2013, *ApJ*, **768**, 24, *eprint*: 1303.2627
- Ormel, C. W., & Klahr, H. H. 2010, *A&A*, **520**, A43
- Otegi, J. F., Bouchy, F., & Helled, R. 2020, *A&A*, **634**, A43
- Pacetti, E., Turrini, D., Schisano, E., et al. 2022, *ApJ*, **937**, 36
- Pairet, B., Cantalloube, F., Gomez Gonzalez, C. A., Absil, O., & Jacques, L. 2019, *MNRAS*, **487**, 2262
- Palma-Bifani, P., Chauvin, G., Bonnefoy, M., et al. 2023, *A&A*, **670**, A90
- Patapis, P., Nasedkin, E., Cugno, G., et al. 2022, *A&A*, **658**, A72
- Patience, J., King, R. R., De Rosa, R. J., et al. 2012, *A&A*, **540**, A85
- Pepper, J., Pogge, R. W., DePoy, D. L., et al. 2007, *PASP*, **119**, 923
- Perri, F., & Cameron, A. G. W. 1974, *Icarus*, **22**, 416
- Petit dit de la Roche, D. J. M., van den Ancker, M. E., Kissler-Patig, M., Ivanov, V. D., & Fedele, D. 2019, *Monthly Notices of the Royal Astronomical Society*, **491**, 1795, *eprint*: <https://academic.oup.com/mnras/article-pdf/491/2/1795/31163942/stz3117.pdf>
- Petrus, S., Bonnefoy, M., Chauvin, G., et al. 2021, *A&A*, **648**, A59
- Petrus, S., Chauvin, G., Bonnefoy, M., et al. 2023, *A&A*, **670**, L9
- Petrus, S., Whiteford, N., Patapis, P., et al. 2024, *ApJ*, **966**, L11
- Phillips, M. W., Tremblin, P., Baraffe, I., et al. 2020, *A&A*, **637**, A38, *eprint*: 2003.13717
- Piskunov, N. E., Kupka, F., Ryabchikova, T. A., Weiss, W. W., & Jeffery, C. S. 1995, *A&AS*, **112**, 525
- Podolak, M. 2003, *Icarus*, **165**, 428
- Pollacco, D. L., Skillen, I., Collier Cameron, A., et al. 2006, *PASP*, **118**, 1407
- Pollack, J. B., Hubickyj, O., Bodenheimer, P., et al. 1996, *icarus*, **124**, 62
- Polman, J., Waters, L. B. F. M., Min, M., Miguel, Y., & Khorshid, N. 2023, *A&A*, **670**, A161
- Polyansky, O. L., Kyuberis, A. A., Zobov, N. F., et al. 2018a, *MNRAS*, **480**, 2597
- Polyansky, O. L., Kyuberis, A. A., Zobov, N. F., et al. 2018b, *MNRAS*, **480**, 2597
- Powell, D., Louden, T., Kreidberg, L., et al. 2019, *ApJ*, **887**, 170
- Powell, D., & Zhang, X. 2024, *arXiv e-prints*, [arXiv:2404.08759](https://arxiv.org/abs/2404.08759)

- Powell, D., Feinstein, A. D., Lee, E. K. H., et al. 2024, *Nature*, 626, 979
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1992, Numerical recipes in FORTRAN. The art of scientific computing ("Cambridge University Press")
- Prinn, R. G., & Barshay, S. S. 1977, *Science*, 198, 1031
- Pueyo, L. 2016, *ApJ*, 824, 117
- Quanz, S. P., Meyer, M. R., Kenworthy, M. A., et al. 2010, *ApJ*, 722, L49
- Racine, R., Walker, G. A. H., Nadeau, D., Doyon, R., & Marois, C. 1999, *PASP*, 111, 587
- Radica, M., Welbanks, L., Espinoza, N., et al. 2023, *MNRAS*, 524, 835
- Radigan, J. 2014, *ApJ*, 797, 120
- Radigan, J., Lafrenière, D., Jayawardhana, R., & Artigau, E. 2014, *ApJ*, 793, 75
- Rafikov, R. R. 2005, *ApJ*, 621, L69
- Rahman, M. M. 2005, PhD thesis, University of British Columbia
- Rajan, A., Barman, T., Soummer, R., et al. 2015, *ApJ*, 809, L33
- Rameau, J., Chauvin, G., Lagrange, A. M., et al. 2013, *ApJ*, 779, L26
- Rasmussen, C. E., & Williams, C. K. I. 2005, Gaussian Processes for Machine Learning (The MIT Press)
- Raymond, S. N., & Morbidelli, A. 2022, in *Astrophysics and Space Science Library*, Vol. 466, *Demographics of Exoplanetary Systems, Lecture Notes of the 3rd Advanced School on Exoplanetary Science*, ed. K. Biazzo, V. Bozza, L. Mancini, & A. Sozzetti, 3
- Rhee, J. H., Song, I., Zuckerman, B., & McElwain, M. 2007, *ApJ*, 660, 1556
- Richard, C., Gordon, I. E., Rothman, L. S., et al. 2012, *J. Quant. Spec. Radiat. Transf.*, 113, 1276
- Richardson, L. J., Deming, D., Horning, K., Seager, S., & Harrington, J. 2007, *Nature*, 445, 892
- Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2014, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 9143, *Space Telescopes and Instrumentation 2014: Optical, Infrared, and Millimeter Wave*, ed. J. Oschmann, Jacobus M., M. Clampin, G. G. Fazio, & H. A. MacEwen, 914320
- Rocchetto, M., Waldmann, I. P., Venot, O., Lagage, P. O., & Tinetti, G. 2016, *ApJ*, 833, 120, [_eprint: 1610.02848](#)
- Roddier, F., & Lena, P. 1984, *Journal of Optics*, 15, 171
- Rodgers, C. D. 1976, *Reviews of Geophysics and Space Physics*, 14, 609
- Rodgers, C. D. 2000, *Inverse Methods for Atmospheric Sounding: Theory and Practice* (World Scientific Publishing Co. Pte. Ltd)
- Rodrigo, C., & Solano, E. 2020, in XIV.0 Scientific Meeting (virtual) of the Spanish Astronomical Society, 182
- Rodrigo, C., Solano, E., & Bayo, A. 2012, SVO Filter Profile Service Version 1.0, IVOA Working Draft 15 October 2012
- Rogozinski, Z., & Hamilton, D. P. 2021, *PSJ*, 2, 78
- Rossow, W. B. 1978, *Icarus*, 36, 1
- Rothman, L. S., Gordon, I. E., Barber, R. J., et al. 2010a, *J. Quant. Spec. Radiat. Transf.*, 111, 2139
- Rothman, L. S., Gordon, I. E., Barber, R. J., et al. 2010b, *J. Quant. Spec. Radiat. Transf.*, 111, 2139
- Rousset, G. 1999, in *Adaptive Optics in Astronomy*, ed. F. Roddier (Cambridge University Press), 91–130
- Rowland, M. J., Morley, C. V., & Line, M. R. 2023, *ApJ*, 947, 6
- Ruffio, J.-B., Macintosh, B., Wang, J. J., et al. 2017, *ApJ*, 842, 14
- Ruffio, J.-B., Macintosh, B., Konopacky, Q. M., et al. 2019, *AJ*, 158, 200
- Ruffio, J.-B., Konopacky, Q. M., Barman, T., et al. 2021, *AJ*, 162, 290
- Ruffio, J.-B., Perrin, M. D., Hoch, K. K. W., et al. 2023, *arXiv e-prints*, arXiv:2310.09902

- Rustamkulov, Z., Sing, D. K., Mukherjee, S., et al. 2023a, *Nature*, **614**, 659
- Rustamkulov, Z., Sing, D., Mukherjee, S., et al. 2023b, in American Astronomical Society Meeting Abstracts, Vol. 55, American Astronomical Society Meeting Abstracts, 124.01
- Sadakane, K. 2006, *PASJ*, **58**, 1023
- Samland, M., Bouwman, J., Hogg, D. W., et al. 2021, *A&A*, **646**, A24
- Samland, M., Mollière, P., Bonnefoy, M., et al. 2017, *A&A*, **603**, A57, [_eprint: 1704.02987](#)
- Sánchez-Lavega, A., Irwin, P., & García Muñoz, A. 2023, *A&A Rev.*, **31**, 5
- Saumon, D., & Marley, M. S. 2008, *ApJ*, **689**, 1327, [_eprint: 0808.2611](#)
- Schlaufman, K. C. 2018, *ApJ*, **853**, 37
- Schmidt, T. O. B., Neuhäuser, R., Seifahrt, A., et al. 2008, *A&A*, **491**, 311
- Schneider, A. D., & Bitsch, B. 2021a, *A&A*, **654**, A71
- Schneider, A. D., & Bitsch, B. 2021b, *A&A*, **654**, A72
- Seager, S., & Sasselov, D. D. 2000, *ApJ*, **537**, 916
- Seiff, A., Kirk, D. B., Knight, T. C. D., et al. 1998, *J. Geophys. Res.*, **103**, 22857
- Sepulveda, A. G., & Bowler, B. P. 2022, *AJ*, **163**, 52
- Sepulveda, A. G., Huber, D., Li, G., et al. 2023, *Research Notes of the American Astronomical Society*, **7**, 2
- Serigano, J., Hörst, S. M., He, C., et al. 2020, *Journal of Geophysical Research (Planets)*, **125**, e06427
- Serigano, J., Hörst, S. M., He, C., et al. 2022, *Journal of Geophysical Research (Planets)*, **127**, e07238
- Showman, A. P., & Kaspi, Y. 2013, *ApJ*, **776**, 85
- Shu, F. H. 1991, *The physics of astrophysics. Volume 1: Radiation.* ("University Science Books")
- Sing, D. K., Fortney, J. J., Nikolov, N., et al. 2016, *Nature*, **529**, 59
- Skemer, A. J., Hinz, P. M., Esposito, S., et al. 2012, *ApJ*, **753**, 14, [_eprint: 1203.2615](#)
- Skemer, A. J., Marley, M. S., Hinz, P. M., et al. 2014, *ApJ*, **792**, 17, [_eprint: 1311.2085](#)
- Skemer, A. J., Hinz, P., Montoya, M., et al. 2015, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 9605, *Techniques and Instrumentation for Detection of Exoplanets VII*, ed. S. Shaklan, 96051D
- Skilling, J. 2004, *Bayesian inference and maximum entropy methods in science and engineering*, **735**, 395
- Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, *AJ*, **131**, 1163
- Smith, B. A., & Terrile, R. J. 1984, *Science*, **226**, 1421
- Snellen, I. A. G., Brandl, B. R., de Kok, R. J., et al. 2014, *Nature*, **509**, 63
- Soni, V., & Acharyya, K. 2023, *ApJ*, **946**, 29
- Soummer, R., Hagan, J. B., Pueyo, L., et al. 2011, *ApJ*, **741**, 55
- Soummer, R., Pueyo, L., & Larkin, J. 2012, *ApJ*, **755**, L28
- Sousa-Silva, C., Al-Refaie, A. F., Tennyson, J., & Yurchenko, S. N. 2015, *MNRAS*, **446**, 2337
- Stassun, K. G., Oelkers, R. J., Pepper, J., et al. 2018, *AJ*, **156**, 102
- Stephens, D. C., Leggett, S. K., Cushing, M. C., et al. 2009, *ApJ*, **702**, 154
- Stevenson, D. J., & Salpeter, E. E. 1977, *ApJS*, **35**, 239
- Stolker, T., Bonse, M. J., Quanz, S. P., et al. 2019, *A&A*, **621**, A59
- Stolker, T., Quanz, S. P., Todorov, K. O., et al. 2020, *A&A*, **635**, A182
- Su, K. Y. L., Rieke, G. H., Stapelfeldt, K. R., et al. 2009, *ApJ*, **705**, 314
- Suárez, G., Vos, J. M., Metchev, S., Faherty, J. K., & Cruz, K. 2023, *ApJ*, **954**, L6
- Suárez, G., & Metchev, S. 2022, *MNRAS*, **513**, 5701

- Suárez, G., & Metchev, S. 2023, *MNRAS*, 523, 4739
- Suárez, G., Metchev, S., Leggett, S. K., Saumon, D., & Marley, M. S. 2021, *ApJ*, 920, 99
- Szulágyi, J., Morbidelli, A., Crida, A., & Masset, F. 2014, *ApJ*, 782, 65
- Szulágyi, J., & Mordasini, C. 2017, *MNRAS*, 465, L64
- Tan, X., & Showman, A. P. 2021a, *MNRAS*, 502, 678
- Tan, X., & Showman, A. P. 2021b, *MNRAS*, 502, 2198
- Taylor, F. W., Atreya, S. K., Encrenaz, T., et al. 2007, in *Jupiter The Planet, Satellites and Magnetosphere* ("Cambridge University Press"), 59
- Tennyson, J., & Yurchenko, S. N. 2012, *MNRAS*, 425, 21, _eprint: 1204.0124
- Thompson, W., Marois, C., Do Ó, C. R., et al. 2023, *AJ*, 165, 29
- Thorngren, D., & Fortney, J. J. 2019, *ApJ*, 874, L31
- Thorngren, D. P., Fortney, J. J., Murray-Clay, R. A., & Lopez, E. D. 2016, *ApJ*, 831, 64
- Toomre, A. 1964, *ApJ*, 139, 1217
- Tremblin, P., Amundsen, D. S., Chabrier, G., et al. 2016, *ApJ*, 817, L19, _eprint: 1601.03652
- Tremblin, P., Amundsen, D. S., Mourier, P., et al. 2015, *The Astrophysical Journal Letters*, 804, L17, _eprint: 1504.03334
- Tremblin, P., Chabrier, G., Baraffe, I., et al. 2017, *ApJ*, 850, 46, _eprint: 1710.02640
- Tremblin, P., Padiou, T., Phillips, M. W., et al. 2019, *ApJ*, 876, 144
- Tsai, S.-M., Malik, M., Kitzmann, D., et al. 2021, *ApJ*, 923, 264
- Tsai, S.-M., Lee, E. K. H., Powell, D., et al. 2023, *Nature*, 617, 483
- Tsuji, T., & Nakajima, T. 2003, *ApJ*, 585, L151
- Turrini, D., Schisano, E., Fonte, S., et al. 2021, *ApJ*, 909, 40
- Twomey, S., Herman, B., & Rabinoff, R. 1977, *Journal of the Atmospheric Sciences*, 34, 1085
- Underwood, D. S., Tennyson, J., Yurchenko, S. N., et al. 2016, *MNRAS*, 459, 3890
- Valencia, D., Sasselov, D. D., & O'Connell, R. J. 2007, *ApJ*, 665, 1413
- van der Marel, N. 2023, *European Physical Journal Plus*, 138, 225
- Vasist, M., Rozet, F., Absil, O., et al. 2023, *A&A*, 672, A147
- Vazan, A., Helled, R., Kovetz, A., & Podolak, M. 2015, *ApJ*, 803, 32
- Vigan, A. 2020, vlt-sphere: Automatic VLT/SPHERE data reduction and analysis
- Vigan, A., Fontanive, C., Meyer, M., et al. 2021, *A&A*, 651, A72
- Visscher, C., Lodders, K., & Fegley, Bruce, J. 2010, *ApJ*, 716, 1060
- Vos, J. M., Allers, K. N., Biller, B. A., et al. 2018, *MNRAS*, 474, 1041
- Vos, J. M., Biller, B. A., Bonavita, M., et al. 2019, *MNRAS*, 483, 480, _eprint: 1811.08370
- Vos, J. M., Burningham, B., Faherty, J. K., et al. 2023, *ApJ*, 944, 138
- Wahhaj, Z., Milli, J., Romero, C., et al. 2021, *A&A*, 648, A26
- Wahl, S. M., Hubbard, W. B., Militzer, B., et al. 2017, *Geophys. Res. Lett.*, 44, 4649
- Waldmann, I. P., Rocchetto, M., Tinetti, G., et al. 2015, *apj*, 813, 13
- Wang, J. 2023, *arXiv e-prints*, arXiv:2310.00088
- Wang, J., Mawet, D., Fortney, J. J., et al. 2018a, *AJ*, 156, 272
- Wang, J., Wang, J. J., Ma, B., et al. 2020a, *AJ*, 160, 150
- Wang, J., Wang, J. J., Ruffio, J.-B., et al. 2023, *AJ*, 165, 4
- Wang, J. J., Ruffio, J.-B., De Rosa, R. J., et al. 2015, pyKLIP: PSF Subtraction for Exoplanets and Disks
- Wang, J. J., Graham, J. R., Dawson, R., et al. 2018b, *AJ*, 156, 192, _eprint: 1809.04107
- Wang, J. J., Ginzburg, S., Ren, B., et al. 2020b, *AJ*, 159, 263
- Wang, J. J., Vigan, A., Lacour, S., et al. 2021, *AJ*, 161, 148

- Wang, J. J., Ruffio, J.-B., Morris, E., et al. 2021, *AJ*, 162, 148
- Wang, J. J., Gao, P., Chilcote, J., et al. 2022, *AJ*, 164, 143
- Weiss, L. M., Marcy, G. W., Petigura, E. A., et al. 2018, *AJ*, 155, 48
- Wende, S., Reiners, A., Seifahrt, A., & Bernath, P. F. 2010, *A&A*, 523, A58
- Whiteford, N., Glasse, A., Chubb, K. L., et al. 2023, *MNRAS*
- Wilcomb, K. K., Konopacky, Q. M., Barman, T. S., et al. 2020, *AJ*, 160, 207
- Wilner, D. J., MacGregor, M. A., Andrews, S. M., et al. 2018, *ApJ*, 855, 56
- Winterhalder, T. O., Lacour, S., Mérand, A., et al. 2024, *arXiv e-prints*, arXiv:2403.13055
- Wit, E., Heuvel, E. v. d., & Romeijn, J.-W. 2012, *Statistica Neerlandica*, 66, 217
- Wolszczan, A., & Frail, D. A. 1992, *nat*, 355, 145
- Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, *AJ*, 140, 1868
- Wright, J. T., & Gaudi, B. S. 2013, in *Planets, Stars and Stellar Systems: Volume 3: Solar and Stellar Planetary Systems*, ed. T. D. Oswalt, L. M. French, & P. Kalas (Dordrecht: Springer Netherlands), 489
- Xuan, J. W., Wang, J., Ruffio, J.-B., et al. 2022, *ApJ*, 937, 54
- Yelle, R. V., Serigano, J., Koskinen, T. T., et al. 2018, *Geophys. Res. Lett.*, 45, 10,951
- Young, T. 1804, *Philosophical Transactions of the Royal Society of London Series I*, 94, 1
- Yurchenko, S. N., Amundsen, D. S., Tennyson, J., & Waldmann, I. P. 2017, *A&A*, 605, A95
- Yurchenko, S. N., Mellor, T. M., Freedman, R. S., & Tennyson, J. 2020, *MNRAS*, 496, 5282
- Zahnle, K., Marley, M. S., Morley, C. V., & Moses, J. I. 2016, *ApJ*, 824, 137
- Zahnle, K. J., & Marley, M. S. 2014, *ApJ*, 797, 41
- Zeleznik, F., & Gordon, S. 1960, *An Analytical Investigation of Three General Methods of Calculating Chemical-Equilibrium Compositions*, Tech. rep., NASA
- Zerbi, F. M., Rodríguez, E., Garrido, R., et al. 1999, *MNRAS*, 303, 275
- Zhang, K. 2024, *arXiv e-prints*, arXiv:2404.15423
- Zhang, X. 2020, *Research in Astronomy and Astrophysics*, 20, 099
- Zhang, Y., Snellen, I. A. G., Bohn, A. J., et al. 2021, *Nature*, 595, 370
- Zhang, Z., Liu, M. C., Clayton, Z. R., et al. 2021, *ApJ*, 916, L11
- Zhang, Z., Mollière, P., Hawkins, K., et al. 2023, *AJ*, 166, 198
- Zhao, B., Tomida, K., Hennebelle, P., et al. 2020, *Space Sci. Rev.*, 216, 43
- Zhou, Y., Bowler, B. P., Morley, C. V., et al. 2020, *AJ*, 160, 77, _eprint: 2004.05168
- Zhu, W., & Dong, S. 2021, *ARA&A*, 59, 291
- Zsom, A., Ormel, C. W., Güttler, C., Blum, J., & Dullemond, C. P. 2010, *A&A*, 513, A57
- Zuckerman, B., Rhee, J. H., Song, I., & Bessell, M. S. 2011, *ApJ*, 732, 61
- Zuckerman, B., & Song, I. 2004, *ARA&A*, 42, 685
- Zurlo, A. 2024, *arXiv e-prints*, arXiv:2404.05797
- Zurlo, A., Gratton, R., Pérez, S., & Cieza, L. 2023, *European Physical Journal Plus*, 138, 411
- Zurlo, A., Vigan, A., Galicher, R., et al. 2016, *A&A*, 587, A57
- Zurlo, A., Goździewski, K., Lazzoni, C., et al. 2022, *A&A*, 666, A133