

**Université de Montréal**

**Caractérisation de l'atmosphères des exoplanètes par  
spectroscopie de transit à haute dispersion avec SPIRou.**

par

**Anne Boucher**

Département de physique  
Faculté des arts et des sciences

Thèse présentée en vue de l'obtention du grade de  
Philosophiæ Doctor (Ph.D.)  
en Physique

30 avril 2022

© Anne Boucher, 2022



# Université de Montréal

Faculté des arts et des sciences

---

Cette thèse intitulée

## **Caractérisation de l'atmosphères des exoplanètes par spectroscopie de transit à haute dispersion avec SPIRou.**

présentée par

**Anne Boucher**

a été évaluée par un jury composé des personnes suivantes :

*Patrick Dufour*

---

(président-rapporteur)

*David Lafrenière*

---

(directeur de recherche)

*René Doyon*

---

(membre du jury)

*Alain Lecavelier des Étangs*

---

(examineur externe)

*Olivier Sonnentag*

---

(représentant du doyen de la FESP)



## Résumé

---

L'objectif principal de cette thèse est de caractériser l'atmosphère de Jupiters chaudes par spectroscopie de transmission à haute résolution, dans l'infrarouge proche, avec l'instrument SPIRou. L'historique de formation, d'évolution et de migration des planètes est empreint dans leur composition chimique, et de retrouver cette composition permet d'en élucider le mystère. La spectroscopie de transit et d'émission a prouvé être fortement efficace à cette tâche, autant pour la détermination de la composition que pour la détermination d'autres caractéristiques atmosphériques comme le profil de température et la dynamique, accessibles à haute résolution. Les Jupiter Chaudes, planètes géantes gazeuses qui orbitent très près de leur étoile, offrent des conditions d'observations très favorables pour ce type d'étude. Encore beaucoup d'éléments nous échappent quant aux processus physiques, chimiques et dynamiques qui gouvernent l'atmosphère de ces objets astronomiques. Des études détaillées de ceux-ci, telles que celles présentées dans cette thèse, sont nécessaires pour mieux comprendre ces mécanismes.

Dans un premier temps, nous avons fait l'étude de deux transits de HD 189733 b, une des exoplanètes les plus étudiées. Cela nous a permis de valider nos méthodes d'analyse avec des données provenant du spectropolarimètre infrarouge à haute résolution SPIRou, installé au télescope Canada-France-Hawaii, représentant d'ailleurs la première caractérisation d'une atmosphère en spectroscopie de transit pour le *SPIRou Legacy Survey*. Nous avons confirmé la présence d'un signal d'eau à un niveau de signification de plus de  $5\sigma$ , basé sur les résultats de *test-t*. Nous présentons la première analyse de spectroscopie de transit haute résolution faite par méthode de récupération, basée sur l'inférence bayésienne et appliquée à une grille de modèles SCARLET interpolée. Celle-ci a permis d'inférer une abondance de  $\log_{10}\text{H}_2\text{O} \simeq -4.4$ . Les résultats obtenus sont cohérents avec la littérature et indiquent que l'atmosphère de HD 189733 b est relativement claire (sans nuages) et possède un C/O super-solaire (correspondant à une formation au-delà de la ligne de glace de l'eau). Un fort décalage vers le bleu de l'absorption par l'eau a été observé, indiquant la présence de forts vents allant du côté jour vers le côté nuit ou un signal dominé par le côté soir du terminateur (limbe arrière), ou une combinaison des deux.

Dans un second temps, nous présentons la première analyse à haute résolution dans l'infrarouge proche de trois transits de la très peu dense sous-Saturne WASP-127 b. Une étude récente

à basse résolution a montré un spectre de transmission riche et une abondance super-solaire de CO<sub>2</sub> dans son atmosphère. La contribution de CO et de CO<sub>2</sub> n'a cependant pu être démêlée étant donné la couverture spectrale et la résolution limitée des données HST et de Spitzer utilisées, menant à des scénarios équiprobables de faible et fort C/O. La couverture de la bande de CO à 2.3 μm par SPIRou a permis de faire cette différenciation, et nos résultats ont exclu la présence de CO en abondance supérieure à 10<sup>-4.3</sup>, impliquant que le signal détecté à 4.5 μm dans les données Spitzer provient majoritairement du CO<sub>2</sub>. De plus, un calcul de test-*t* sur les données SPIRou a confirmé la détection de H<sub>2</sub>O à un niveau ajusté de 4.9σ, mais également une détection potentielle de OH, à 2.4σ. Cette présence extrêmement inattendue de OH pourrait potentiellement être expliquée si la température du côté jour est assez grande, aidée par l'irradiation amplifiée de l'étoile qui quitte la séquence principale, ou encore par du mélange vertical. Nous présentons également la première méthode de récupération complète appliquée à la spectroscopie de transit à haute résolution, en utilisant la suite de modèles petitRADTRANS, et sur trois ensembles de données différents : les données SPIRou, les données HST et Spitzer de l'étude initiale, et les deux ensembles de données combinés. Une comparaison des différents résultats obtenus confirme que l'analyse conjointe permet d'avoir de meilleures contraintes sur les paramètres atmosphériques. Alors que l'étude initiale favorisait un fort C/O, nos résultats pointent vers un C/O très sous-solaire, produit par un C/H sous-solaire et un O/H plutôt stellaire. Les scénarios de formation qui supportent une telle composition sont ceux au-delà des lignes de glace de H<sub>2</sub>O et de CO<sub>2</sub> (~ 10ua), avec une accrétion supplémentaire de matériel riche en O via la migration et le croisement des lignes de glaces. L'accrétion du matériel est soit dominée par le gaz et tardive (après 5 à 7 millions d'années — Ma), ou encore, mixte (de gaz et de glaces) et plus précoce, avec un mélange cœur-enveloppe substantiel.

Bien qu'il en reste beaucoup à faire, ces travaux de recherche ont démontré que la spectroscopie de transit à haute résolution dans l'infrarouge proche est utile pour explorer les conditions atmosphériques des Jupiter et sous-Saturne chaudes, et plus spécifiquement, avec l'instrument SPIRou. La combinaison de données à faible et à haute résolution s'avère un outil très puissant pour l'étude des atmosphères, et le sera encore plus avec les capacités révolutionnaires de JWST.

**Mots clés :** Atmosphères d'exoplanètes — Composition atmosphérique d'exoplanètes — Astronomie d'exoplanètes — Circulation atmosphérique — Jupiters chaudes — Spectroscopie moléculaire — Analyse de données astronomiques — Nuages atmosphériques — Composition atmosphérique — Spectroscopie de transmission — Spectroscopie haute résolution

# Abstract

---

The main objective of this thesis is to characterize the atmosphere of hot Jupiters with high resolution transmission spectroscopy in the near-infrared with the SPIRou instrument. The formation, evolution and migration history of exoplanets is imprinted in their chemical composition, and finding this composition makes it possible to trace back this history. Transit and emission spectroscopy have proven to be highly effective for this task, in the determination of the composition, but also of other atmospheric characteristics such as the temperature profile and the dynamics, accessible at high resolution. Hot Jupiters – gas giant planets orbiting very close to their star – offer highly favourable observation conditions for this type of study. Many pieces of the puzzle are still missing regarding the physical, chemical and dynamical processes that govern the atmospheres of these astronomical objects, and detailed studies, such as the ones presented in this thesis, are necessary to better understand those mechanisms.

First, we present the results of our analysis of two HD 189733 b transits, one of the most studied exoplanets to date. This study allowed to validate our analysis method with SPIRou data, a high-resolution near-infrared spectro-polarimeter installed at the Canada-France-Hawaii Telescope. It also represents the first characterization of an atmosphere with transit spectroscopy as part of the SPIRou Legacy Survey. Our results confirmed the H<sub>2</sub>O detection in the planet’s atmosphere at a  $5\sigma$  level, based on a  $t$ -test. We present the first analysis of a Bayesian retrieval framework applied to high-resolution transmission spectroscopy, using a grid of SCARLET models. We constrained the H<sub>2</sub>O abundance to  $\log_{10}\text{H}_2\text{O} \simeq -4.4$ . The results are consistent with the literature and agree on the atmosphere of HD 189733 b being relatively clear (without clouds) and having a super-solar C/O (corresponding to a formation beyond the H<sub>2</sub>O ice line). A strong blue shift of the water absorption signal was observed, indicative of strong day-to-night winds or a signal dominated by the terminator’s evening side (trailing limb), or a combination of both.

Second, we present the first high-resolution analysis in the near-infrared of three transits of the super low-density sub-Saturn WASP-127 b. A recent low-resolution study showed a rich transmission spectrum and super-solar abundance of CO<sub>2</sub> in its atmosphere. However, the contribution of CO and CO<sub>2</sub> could not be disentangled given the limited spectral coverage and

resolution of the HST and Spitzer data, leading to equiprobable low and high C/O scenarios. The coverage of the CO band at  $2.3 \mu\text{m}$  by SPIRou made it possible to differentiate between the two cases, and our results excluded CO abundances greater than  $10^{-4.3}$ , implying that the signal at  $4.5 \mu\text{m}$  seen in the Spitzer data mostly comes from  $\text{CO}_2$ . Moreover, a test- $t$  analysis on the SPIRou data confirmed the detection of  $\text{H}_2\text{O}$  at an adjusted level of  $4.9\sigma$ , but also a tentative detection of OH, at  $2.4\sigma$ . The presence of OH, although extremely unexpected, could potentially be explained from a high enough dayside temperature, increased by the amplified irradiation of the star leaving the main sequence, or from vertical mixing. We also present the first full-retrieval framework applied to transmission spectroscopy at high resolution, using the petitRADTRANS model suite on three different datasets: on the SPIRou data, on the HST and Spitzer data from the original study, and on both datasets combined. A comparison of the different results obtained confirms that the joint analysis provides better constraints on the atmospheric parameters. While the initial study favored a high C/O, our results point toward a highly subsolar C/O, produced by a sub-stellar C/H and a roughly stellar O/H. Formation scenarios that support such a composition are those beyond the  $\text{H}_2\text{O}$  and  $\text{CO}_2$  ice lines ( $\sim 10 \text{ ua}$ ), with further accretion of O-rich material via migration and ice lines crossing. The primordial/bulk accretion was either gas-dominated and late (after 5-7 Myr), or earlier and mixed (with gas and ice), with substantial core-envelope mixing.

Although much remains to be done, this research work has demonstrated that high-resolution near-infrared transit spectroscopy is useful for exploring the atmospheric conditions of hot Jupiters and sub-Saturns, and more specifically, with the SPIRou instrument. The combination of low and high resolution data has shown to be a very powerful tool for such studies, and will be even more so with the revolutionary capabilities of JWST.

**Keywords** : Exoplanet atmospheres — Exoplanet atmospheric composition — Exoplanet astronomy — Atmospheric circulation — Hot Jupiters — Molecular spectroscopy — Astronomy data analysis — Atmospheric clouds — Atmospheric composition — Transmission spectroscopy — High resolution spectroscopy

# Table des matières

---

<b>Résumé</b> .....	5
<b>Abstract</b> .....	7
<b>Liste des tableaux</b> .....	13
<b>Table des figures</b> .....	15
<b>Liste des sigles et des abréviations</b> .....	19
<b>Remerciements</b> .....	23
<b>Introduction</b> .....	25
<b>Chapitre 1. Éléments théoriques et notions de base</b> .....	29
1.1. Détection et caractérisation d'exoplanètes .....	29
1.1.1. Exoplanètes.....	29
1.1.1.1. Jupiter chaude .....	31
1.1.1.2. Sous-Saturne .....	31
1.1.2. Méthodes de détection .....	32
1.1.2.1. Méthode de vitesses radiales.....	33
1.1.2.2. Méthode du transit .....	36
1.1.3. Méthodes de caractérisation .....	38
1.1.3.1. Spectre en émission .....	39
1.1.3.2. Courbe de phase.....	42
1.1.3.3. Spectre de transmission .....	43
1.2. Formation des planètes géantes .....	46
1.2.1. Instabilités gravitationnelles dans le disque (IG) .....	47
1.2.2. Accrétion de coeur (AC) .....	47
1.2.3. Accrétion de cailloux .....	49
1.2.4. Migration.....	50

1.3.	Physique des atmosphères.....	51
1.3.1.	Transfert radiatif .....	51
1.3.2.	Balance énergétique .....	54
1.3.3.	Profil vertical .....	55
1.3.4.	Composition.....	60
1.3.4.1.	Hypothèses de base.....	60
1.3.4.2.	Considérations additionnelles.....	62
1.3.4.3.	Nuages et aérosols.....	68
1.3.5.	Dynamique .....	70
1.4.	Approches observationnelles .....	73
1.4.1.	Spectroscopie à faible dispersion.....	73
1.4.2.	Spectroscopie à haute dispersion .....	79
1.4.3.	Analyses Conjointes.....	92
1.5.	Détermination des propriétés .....	94
1.5.1.	Méthode de récupération .....	99
1.6.	Contexte de ces travaux .....	100
1.7.	Contributions .....	102
<b>Chapitre 2. Characterizing exoplanetary atmospheres at high resolution with</b>		
<b>SPIRou : Detection of water on HD 189733 b .....</b>		
		105
	RÉSUMÉ.....	106
	ABSTRACT.....	106
2.1.	Introduction .....	107
2.2.	Observations and data reduction .....	110
2.3.	Analysis.....	113
2.3.1.	Telluric absorption correction.....	113
2.3.2.	Transmission spectrum construction.....	116
2.3.3.	Atmospheric Model .....	118
2.4.	Atmospheric Signal Extraction :	
	Methods and results .....	120
2.4.1.	Model processing .....	120
2.4.2.	Cross-Correlation .....	121

2.4.2.1.	Algorithm .....	121
2.4.2.2.	Results .....	123
2.4.3.	Log-Likelihood Mapping Method .....	129
2.4.3.1.	Results .....	130
2.4.4.	Markov Chain Monte Carlo .....	130
2.5.	Discussion .....	133
2.5.1.	Effects of telluric residuals .....	135
2.5.2.	Retrieved Atmospheric Parameters .....	137
2.5.3.	Radial velocity offset .....	138
2.6.	Concluding remarks .....	140
<b>Chapitre 3.</b>	<b>CO or no CO ? Narrowing the CO abundance constraint and recovering the H<sub>2</sub>O detection in the atmosphere of WASP-127 b using SPIRou .....</b>	<b>143</b>
	RÉSUMÉ.....	144
	ABSTRACT.....	144
3.1.	Introduction .....	145
3.2.	The unusual WASP-127 system.....	147
3.3.	Observations.....	150
3.4.	Data Reduction and Analysis .....	153
3.4.1.	Telluric absorption correction.....	154
3.4.2.	Transmission spectrum construction.....	155
3.5.	Atmospheric Signal Extraction :	
	Models and Methods .....	157
3.5.1.	Atmospheric Model .....	157
3.5.1.1.	Model processing .....	159
3.5.2.	Cross-Correlation and Log-Likelihood Mapping.....	160
3.5.2.1.	Algorithm .....	160
3.5.2.2.	Detection significance .....	161
3.6.	Results.....	162
3.6.1.	Full Retrievals .....	166
3.6.1.1.	Joint Retrieval .....	168

3.7. Discussion.....	174
3.7.1. Molecular (non-)detections.....	174
3.7.1.1. CO and CO <sub>2</sub> .....	174
3.7.1.2. OH.....	175
3.7.1.3. FeH.....	176
3.7.1.4. N-bearing molecules.....	177
3.7.2. Comparison between the types of retrievals.....	177
3.7.3. Radial Velocity Offset.....	179
3.7.4. Implications for the Formation of WASP-127 b.....	180
3.8. Concluding remarks.....	182
3.9. Appendix : Injection-Recovery tests.....	186
<b>Chapitre 4. Conclusion.....</b>	<b>189</b>
<b>Bibliographie.....</b>	<b>193</b>

## Liste des tableaux

---

2.1	SPIRou observations of HD 189733 .....	110
2.2	Adopted parameters for the system HD 189733 .....	115
2.3	MCMC Retrieval Parameter Priors and Results for HD 189733 b .....	133
2.4	Overview of the results from the model grid for the different methods, using their respective best-fit model. ....	135
3.1	Overview of previous detections on WASP-127 b .....	147
3.2	WASP-127 System Parameters .....	149
3.3	SPIRou observations of WASP-127 .....	152
3.4	MCMC Retrieval Parameter Priors and Results for WASP-127 b .....	171



## Table des figures

---

1.1	Schéma d'une orbite planétaire. ....	32
1.2	Schéma de la méthode de détection d'exoplanètes par vitesses radiales.....	34
1.3	Exemples de courbes de vitesses radiales. ....	35
1.4	Schéma de la méthode de détection d'exoplanètes par transit.....	36
1.5	Flux relatifs émis et réfléchis de différents objets astronomiques. ....	38
1.6	Visualisation du concept de spectroscopie d'émission. ....	40
1.7	Spectres IR de la Terre, et température de luminosité.....	40
1.8	Exemple de spectre d'émission de WASP-33 b, et son inversion de température.....	41
1.9	Visualisation du concept de spectroscopie de transmission.....	44
1.10	Exemple de spectre de transmission pour WASP-19 b. ....	44
1.11	Géométrie d'un système en transit. ....	46
1.12	Visualisation des concepts d'absorption et d'émission de raies spectrales.....	52
1.13	Spectre de sections efficaces moléculaires.....	53
1.14	Exemple de différents types de profils TP. ....	57
1.15	Ensemble de profils TP extraits d'un modèle GCM 3D de HD 209458 b.....	58
1.16	Carte de températures équatoriales d'un modèle GCM de WASP-76 b.....	59
1.17	Relation Masse-Métallicité d'un ensemble d'exoplanètes, comparées aux planètes du système solaire.....	61
1.18	Variation des ratios d'abondance atomiques solides et gazeux en fonction la distance de l'étoile. ....	62
1.19	Schéma du principe de la dérive de cailloux.....	64
1.20	Variation des ratios d'abondances atomiques, solides et gazeux, en fonction de la distance de l'étoile pour plusieurs pas de temps.....	66

1.21	Profil de température de plusieurs Jupiter chaudes et courbes de condensation de plusieurs éléments chimiques. ....	67
1.22	Visualisation des différentes couches de nuages pour différents objets. ....	68
1.23	Influence des nuages gris et des aérosols/brumes sur le spectre de transmission des planètes. ....	69
1.24	Visualisation de la dynamique sur HD 189733 b, théorique et observée. ....	71
1.25	Visualisation de la dégénérescence entre le rayon planétaire et l'abondance des éléments. ....	75
1.26	Schéma de l'effet de source lumineuse du transit. ....	77
1.27	Séquence de spectres de transmission de JCs observés avec HST et Spitzer, de clairs à nuageux. ....	78
1.28	Variation de la position du signal planétaire en émission en fonction de la phase orbitale. ....	81
1.29	Visualisation de l'effet Rossiter-McLaughlin, par la variation de la forme des raies, et l'influence sur la courbe de vitesse radiale. ....	82
1.30	Étapes pour générer une séquence modèle mieux représentative du signal observé. ...	85
1.31	Exemple de signal de CC pour la raie d'absorption du CO dans HD 209458 b, en fonction de la phase orbitale. ....	86
1.32	Exemple de carte de CCF en fonction de $K_P$ et $v_{\text{sys}}$ , pour la détection de H <sub>2</sub> O sur HD 189733 b. ....	87
1.33	Image colorisée d'un spectre échelle obtenu avec SPIRou. ....	88
1.34	Influence des vents et de la rotation de la planète sur la forme et la position des CCF. ...	90
1.35	Exemple de méthode de récupération des paramètres à partir d'une analyse combinée sur des données simulées de CRIRES et WFC3. ....	93
2.1	Courbes de masse d'air des deux transits de HD 189733 b. ....	111
2.2	Courbes de S/N moyenné de la bande <i>H</i> des deux transits de HD 189733 b. ....	112
2.3	Étapes de réduction appliquées aux séquences de transit de HD 189733 b. ....	114
2.4	Carte des maxima de S/N pour les CCF d'une grille de modèles de HD 189733 b. ...	123
2.5	Carte de CCF en fonction de $K_P$ et de $v_{\text{rad}}$ pour le meilleur modèle de HD 189733 b. ...	124
2.6	Carte de CCF du test d'injection négative du meilleur modèle de HD 189733 b. ....	125

2.7	Carte de CCF 2D du meilleur modèle de HD 189733 b, en fonction de la VR et de la phase orbitale. ....	127
2.8	Carte de test- $t$ du meilleur modèle de HD 189733 b en fonction de $K_p$ et $v_{\text{rad}}$ , et histogramme des valeurs de test- $t$ du signal à l'intérieur et à l'extérieur de la trace de la planète. ....	128
2.9	Carte des maxima de test- $t$ pour une grille de modèles de HD 189733 b. ....	129
2.10	Carte de $\Delta\text{BIC}$ pour une grille de modèles de HD 189733 b. ....	131
2.11	Distributions postérieures des paramètres atmosphériques de HD 189733 b récupérés par la méthode MCMC avec une grille de modèles. ....	132
2.12	Profil T-P de HD 189733 b récupéré. ....	134
3.1	Courbes de masse d'air des trois transits de WASP-127 b. ....	150
3.2	Courbes de S/N moyenné de la bande $H$ des trois transits de WASP-127 b. ....	151
3.3	Étapes de réduction appliquées aux séquences de transit de WASP-127 b. ....	153
3.4	Carte de CCF 2D du meilleur modèle de H <sub>2</sub> O de WASP-127 b, en fonction de la VR et de la phase orbitale. ....	164
3.5	Carte de test- $t$ du meilleur modèle de H <sub>2</sub> O de WASP-127 b en fonction de $K_p$ et $v_{\text{rad}}$ , et histogramme des valeurs de test- $t$ du signal à l'intérieur et à l'extérieur de la trace de la planète. ....	165
3.6	Carte de test- $t$ de la détection potentielle de OH dans l'atmosphère de WASP-127 b. ....	166
3.7	Distributions postérieures des paramètres atmosphériques de WASP-127 b récupérés par la méthode de MCMC pour les différents ensembles de données, ainsi que leur profil de température associé. ....	169
3.8	Meilleurs modèles récupérés par les différents ensembles de données. ....	170
3.9	Distribution postérieure des valeurs de C/O récupérées par l'analyse conjointe. ....	170
3.10	Distributions postérieures des valeurs de C/H et O/H récupérées par l'analyse conjointe. ....	172
3.11	Limites de détection des abondances de différentes molécules d'atmosphères modélisées avec ou sans aérosols. ....	187



## Liste des sigles et des abréviations

---

AC	Accrétion de Coeurs
AIC	Absorption Induite par Collision
BERV	Vitesse radiale barycentrique de la Terre, de l'anglais <i>Barycentric Earth Radial Velocity</i>
CC	Correlation-Croisée
CCF	Fonction de Correlation-Croisée, de l'anglais <i>Cross-Correlation Function</i>
CFHT	Télescope Canada-France-Hawaii, de l'anglais <i>Canada-France-Hawaii Telescope</i>
CLV	Variation Centre-Bord, de l'anglais <i>Center-to-Limb Variation</i>
CRIRES	<i>Cryogenic Infrared Echelle Spectrograph</i>
DST	Déviation Standard
ELT	<i>Extremely Large Telescope</i>
ETL	Équilibre Thermodynamique Local
GCM	Modèle de circulation globale, de l'anglais <i>Global Circulation Model</i>
HR	Haute résolution
HST	Télescope Spatial Hubble, de l'anglais <i>Hubble Space Telescope</i>
IG	Instabilités Gravitationnelles
IR	Infrarouge
IRP	Infrarouge Proche

LBD	Limite du brûlage du deutérium
LG	Ligne de Glace
JC	Jupiter Chaude
JWST	Télescope Spatial James Webb, de l'anglais <i>James Webb Space Telescope</i>
Ma	Million d'années
MCMC	Monte Carlo par chaînes de Markov, de l'anglais <i>Markov Chain Monte Carlo</i>
PCA	Analyse aux composantes principales, de l'anglais <i>Principal Component Analysis</i>
RME	Effet Rossiter-McLaughlin, de l'anglais <i>Rossiter-McLaughlin Effect</i>
S/N	Signal-sur-bruit, de l'anglais <i>Signal-to-Noise</i>
SFD	Spectroscopie à Faible Dispersion
SHD	Spectroscopie à Haute Dispersion
SLS	<i>SPIRou Legacy Survey</i>
SPIRou	Spectro-Polarimètre InfraRouge
SS	Système Solaire
STIS	<i>Space Telescope Imaging Spectrograph</i>
TMT	<i>Thirty Meter Telescope</i>
TP	Température-Pression
VR	Vitesse radiale
UV	Ultra-Violet
VLT	<i>Very Large Telescope</i>
WFC3	<i>Wide Field Camera 3</i>

*Pour deux femmes fortes, Gisèle et Marie-Paul*



## Remerciements

---

Je voudrais d'abord remercier mon directeur de recherche David Lafrenière, qui a toujours fait preuve de patience et d'écoute, et qui a toujours été présent pour me supporter dans mes questionnements et inquiétudes. Merci à son intuition et ses connaissances scientifiques incroyables qui m'ont grandement aidé à mener à terme ce projet d'une envergure inattendue.

Un gros merci à tous mes collègues, qui ont été d'un grand aide pour la complétion de ce projet, que ce soit par les discussions sur les techniques d'analyse, ou encore pour me ramener à l'ordre avec mes conclusions optimistes un peu trop hâtives.

Plus particulièrement, à mon très cher ami Antoine qui a une compréhension infinie des mathématiques et qui approche chaque problème avec calme et positivisme. Il a une capacité incroyable à rapidement contextualiser les situations, et à avoir une bonne vue d'ensemble des concepts. Il m'a énormément aidé avec le développement des codes, et aidé à penser à différentes façons de pousser les analyses. Il était toujours prêt à discuter de mes résultats, aussi minimes qu'ils soient, et à m'encourager lorsque les choses me semblaient impossibles. Il aura été un excellent compagnon de bureau et d'étude de prédoc. Merci pour tout.

Stefan était toujours prêt à répondre à mes questions, à discuter, et m'envoyer des modèles à toutes heures de la journée. Sa curiosité et sa rigueur de travail sont tout à son honneur, et je suis choyée d'avoir pu apprendre de lui. Merci pour ton énorme générosité, tes dons de fruits, et surtout tes dons de pommes, qui ont su combler les petits creux d'avant ou d'après-midi. Merci à Étienne Artigau et Neil J. Cook pour tout leur soutien technique, et particulièrement à Neil d'avoir été plus que tolérant avec les requêtes de nouvelles sur la réduction des données. Merci à René Doyon qui m'a donné mon premier stage, et qui m'a intégrée à la belle équipe d'astrophysique à l'UdeM. Son enthousiasme entraînant et inépuisable, ainsi que ses visions de grandeurs ont fait avancer beaucoup de projets et c'est d'autant plus inspirant d'en faire partie. Un grand merci à tous mes autres collègues et collaborateurs.

Toutes mes années d'études n'auraient pas pu être possible sans le support constant de toute ma famille, et ma belle famille, mais surtout de mes parents Rhéa et Jacques, qui ont toujours cru en moi et qui m'ont toujours encouragé d'entreprendre les projets que je voulais. À Frédéric, pour ton support, ton aide, ta patience, ta compréhension, merci d'être là, et d'avoir partagé ces

dernières années avec moi. Mention spéciale à Carbone, pour avoir été une source de réconfort et de calme tout au long de mes études.

Merci à Luc de m'avoir introduite à l'astronomie quand j'étais petite. Le souvenir d'avoir vu Jupiter et Saturne de mes yeux dans un télescope a profondément marqué mon cheminement. Et à Patrick pour m'avoir fais comprendre au secondaire que la science pouvait être « cool ».

Merci à mes ami.e.s d'astrophysique avec qui j'ai partagé de très beaux moments durant toutes ces années : ma *girl* Roxane, mon *partner* de voyage, de plongée, et de travail dans les cafés François-René, Corinne, et Patrice, le duo solaire de feu, Amélie et sa sagesse infinie, Jonathan la référence absolue de toutes questions d'astro ou de café, Benoit mon *partner* de bureau et de *Hackathon*, Marie Eve avec qui j'ai redécouvert mon côté artistique et seconde *partner* de voyage, Marcel avec qui on ne s'ennuie jamais, Annabelle, Marie-Lou, et Frédérique, qui ont su rendre les dîners au café d'astro toujours aussi plaisants. À tous mes ami.e.s de physique pour avoir fait du baccalauréat en physique une expérience fort enrichissante, et même plaisante, ce qui m'a encouragé à poursuivre et de me rendre jusque là. À tout.e.s mes autres ami.e.s, qui m'ont permis de décrocher quand c'était le temps.

Merci à la gang de Genève qui m'a si bien accueillie dans votre institution et qui a fait de mon stage en Suisse une expérience inoubliable. Je remercie d'ailleurs les Fonds québécois de recherche sur la nature et les technologies d'avoir financé ce stage.

Finalement, merci au Conseil de recherches en sciences naturelles et en génie du Canada et à l'Institut de Recherche sur les Exoplanètes de m'avoir soutenu financièrement tout au long de mon parcours doctoral.

# Introduction

---

Une des grandes questions de l'humanité qui m'a toujours interpellée est la suivante : « Sommes-nous seuls dans l'Univers ? ». Cette question est posée depuis des millénaires et reste encore à ce jour sans réponse. Heureusement, nous sommes maintenant dans une ère technologique assez avancée qui devrait nous permettre d'y répondre dans les prochaines décennies. La façon la plus directe pour y arriver est l'étude des « *exoplanètes* », c'est-à-dire les planètes qui orbitent les étoiles autres que le Soleil.

Les premières candidates de planètes extra-solaires ont été détectées et proposées par Campbell et al. (1988) ( $\gamma$  Cep, 61 Vir et  $\epsilon$  Eri) et Latham et al. (1989) (HD 114762). Elles n'ont cependant été confirmées que quelques années plus tard. La première découverte non ambiguë d'une exoplanète a été faite par Wolszczan & Frail en 1992. Non pas un, mais deux objets de masse planétaire ont été observés orbitant le pulsar PSR B1257+12. Enfin, la découverte de la première exoplanète orbitant une étoile de type solaire, 51 Pegasi, a été faite en 1995 par Mayor & Queloz, une découverte importante qui leur a valu le prix Nobel de physique en 2019. Cette découverte a été capitale dans le domaine de la recherche des exoplanètes.

Cela fait donc seulement une trentaine d'années qu'on sait qu'une autre planète comme la nôtre, quelque part dans le cosmos, *pourrait* exister et que l'idée qu'il serait possible d'y trouver de la vie est devenue réellement tangible. Depuis ces découvertes, l'effort scientifique mis dans la recherche sur les exoplanètes n'a cessé de croître et maintenant plus de cinq milles exoplanètes ont été confirmées (5017 en date du 30 avril 2022, selon le The Extrasolar Planets Encyclopaedia). En revanche, la question demeure : Sommes-nous seuls ?

Comme une panoplie d'exoplanètes de tous genres ont été découvertes, le focus actuel a dévié quelque peu de cette question. Bien qu'encore beaucoup d'efforts soient mis dans la découverte de nouvelles exoplanètes, on cherche maintenant aussi à caractériser celles qu'on connaît déjà, entre autres par l'étude de leurs atmosphères, qui renferment beaucoup d'informations et qui nous sont accessibles depuis très peu. De déterminer les processus physiques, chimiques et dynamiques qui entrent en jeu dans l'atmosphère d'exoplanètes peut nous aider à faire la lumière sur l'historique de formation et d'évolution des planètes.

La caractérisation d'une couche de gaz sur un objet se trouvant à des dizaines de parsecs n'est pas une tâche simple, mais elle est possible grâce à nos méthodes de télédétection (Burrows & Sharp, 1999). Seager & Sasselov (2000) sont les premiers à avoir proposé de le faire via la spectroscopie en transmission lors des transits des planètes géantes. Cette méthode, et les autres qui en découlèrent ont prouvé être payantes et c'est ce dont il sera question dans cet ouvrage.

Dans cet ouvrage il ne sera cependant pas question de répondre à *la* question de la vie ailleurs dans l'univers, car pour y parvenir il faudra attendre l'avènement, au début de la prochaine décennie, de la nouvelle génération de télescopes géants terrestres, comme le *European-Extremely Large Telescope* (E-ELT) ou encore le *Thirty Meter Telescope* (TMT), et plus tard de nouveaux télescopes spatiaux tels que le *Large UV/Optical/Infrared Surveyor* (LUVOIR) et le *Habitable Exoplanet Observatory* (HABEx). Ces télescopes géants seront aptes à aller chercher les faibles signaux dans l'atmosphère des planètes de type terrestre, qui pourraient potentiellement montrer des biosignatures (des indices directs ou indirects qui nécessitent la présence d'activités biologiques, c.-à-d. *la vie* !). Mais avant d'y arriver, il est primordial de bien comprendre ce qu'on observe, d'où l'importance de toutes les branches de l'exoplanétologie.

D'ailleurs, beaucoup de grandes questions sur l'étude des exoplanètes sont restées longtemps ou restent encore sans réponses claires. Par exemple : Comment peut-on former une planète massive avant que le disque protoplanétaire ne disparaisse ? Comment les Jupiter chaudes font-elles pour en arriver là où elles sont, si près de leur étoile ? Est-ce que d'autres systèmes comme notre système solaire existent et comment se sont-ils formés ? Comment les abondances élémentaires des planètes se comparent-elles à celles du disque protoplanétaire ? Est-ce que toutes les planètes géantes ont un cœur solide ? La liste pourrait continuer longtemps, mais ici, on se concentrera plutôt sur les questions suivantes : De quoi est formée l'atmosphère des Jupiters chaudes et sous-Saturnes chaudes ? Quelle est la dynamique atmosphérique globale sur ces planètes ? Comment les paramètres atmosphériques observés sont-ils reliés à la formation et l'évolution de la planète ? Comment l'étoile affecte-t-elle la planète ?

Cet ouvrage présente les résultats de recherches en spectroscopie de transit faites à haute résolution dans l'infrarouge avec des données de l'instrument SPIRou, pour les Jupiter chaudes HD 189733 b, et WASP-127 b. Le chapitre 1 portera sur les notions générales et la théorie de base derrière l'étude des atmosphères d'exoplanètes. À la section 1.1, les méthodes de détection et caractérisation des exoplanètes seront présentées, suivies par un survol des processus de formation de planètes à la section 1.2. La section 1.3 portera sur la physique des atmosphères. Les principaux résultats observationnels qui leur sont associés et la détermination des paramètres atmosphériques seront présentés dans les sections 1.4 et 1.5, respectivement. Les chapitres 2 et 3 constituent deux articles publiés ou à paraître dans *The American Astronomical Society Journals*. On y retrouve les résultats de l'analyse sur la Jupiter chaude très connue HD 189733 b (Numéro

de référence : AAS31484) et sur la sous-Saturne gonflée WASP-127 b, respectivement. Un survol des éléments importants de ce travail ainsi qu'une ouverture sur la suite du projet conclura ce manuscrit au chapitre 4.



# Chapitre 1

---

## Éléments théoriques et notions de base

Les notions présentées dans ce chapitre sont basées, en plus des références données dans le texte, sur les livres de références de Perryman 2011, Seager & Deming 2010, Carroll & Ostlie 2007, et Deeg & Belmonte 2018.

### 1.1. Détection et caractérisation d'exoplanètes

#### 1.1.1. Exoplanètes

On appelle *exoplanètes* toutes planètes qui orbitent une étoile autre que le Soleil. Selon l'Union Astronomique Internationale (UAI), une *planète* est définie comme un objet céleste qui (1) orbite le Soleil, (2) est assez massif pour être en équilibre hydrostatique (sphérique) et (3) a nettoyé son orbite. Cependant, cette définition, adoptée en 2006, ne s'applique qu'aux planètes du Système Solaire (SS). La définition d'une exoplanète est cependant beaucoup plus générale, c.-à-d. : tout objet astronomique ayant une vraie masse en deçà de la masse limite du brûlage du deutérium (LBD; estimée à  $\sim 13 M_J$ ), et en orbite autour d'une étoile, peu importe son mécanisme de formation. Par souci de complétude, on définit les *Naines brunes* comme les objets intermédiaires aux planètes et étoiles : ce sont des objets sous-stellaires qui ont une masse supérieure à la LBD, mais inférieure à la masse limite du brûlage de l'hydrogène (entre 75 et 95  $M_J$ , dépendant de la métallicité). On retrouve aussi les objets dits *flottants* qui n'orbitent aucune étoile, mais qui ont des masses inférieures à la LBD. Une définition précise et formelle reste encore à ce jour à être formulée pour ce genre d'objets (voir Basri & Brown (2006); Soter (2006)). Dans ce document, le mot « planète » sera utilisé pour n'importe quel objet ayant une masse  $M_p \lesssim 13 M_J$  (aucune fusion nucléaire), gravitationnellement lié et orbitant un objet stellaire ou une naine brune (Lecavelier des Etangs & Lissauer, 2022).

Depuis les premières détections, une grande variété de systèmes ont été découverts montrant des propriétés très diversifiées. Que ce soit autour d'étoiles de faibles masses (et naines brunes),

d'étoiles plus évoluées (ex. : les naines blanches et les pulsars) en passant par les étoiles de type solaire, on retrouve des exoplanètes autour de nombreux types d'étoiles. Leurs paramètres orbitaux et leur multiplicité varient aussi grandement. Entre autres, légèrement plus du quart ont une excentricité  $e \gtrsim 0.3$  et la majorité a un demi-grand axe  $a < 10 \text{ua}$ <sup>1</sup>. En date du 30 avril 2022, 822 systèmes planétaires multiples ont été confirmés. Environ deux tiers d'entre eux possèdent seulement deux planètes détectées. Le système TRAPPIST-1 (Gillon et al., 2017), comptant 7 planètes, est celui avec le plus grand nombre de planètes confirmées, après le SS et Kepler-90 (Cabrera et al., 2014; Shallue & Vanderburg, 2018) qui en comptent 8.

Cette diversité de paramètres, produit de la stochasticité de la formation et l'évolution des planètes, s'observe aussi dans la structure et la composition interne des planètes. On peut ainsi regrouper les planètes en trois groupes distincts : les planètes « terrestres » (rocheuses ;  $\lesssim 1-2 R_{\oplus}$  et  $\lesssim 5 M_{\oplus}$ ), les planètes « de glaces » ( $2-6 R_{\oplus}$  et  $\sim 14-17 M_{\oplus}$ ) et « géantes » ( $> 6 R_{\oplus}$  et  $\sim 30-320 M_{\oplus}$ ; gazeuses)<sup>2</sup>. Les planètes terrestres, comme Mercure, Vénus, Terre et Mars, sont principalement formées de roches et de métaux, et possèdent généralement une atmosphère mince. Leurs atmosphères minces sont plus susceptibles de changer suite à l'échappement des gaz légers et des réactions (photo-)chimiques qui peuvent y avoir lieu. Les planètes gazeuses n'ont pas de surface rocheuse et ont des masses très supérieures à celles des planètes terrestres, soit  $0.03 M_J \lesssim M_p \lesssim 13 M_J$  ( $\sim 10 M_{\oplus} < M_p \lesssim 4000 M_{\oplus}$ ). Les atmosphères des géantes gazeuses (comme Jupiter et Saturne) sont composées presque entièrement d'hydrogène et d'hélium, tandis que les planètes de glaces (comme Uranus et Neptune) sont formées principalement de glaces d'eau, de méthane et d'ammoniaque, mais peuvent aussi posséder une enveloppe d'hydrogène et d'hélium. Ces atmosphères sont plus représentatives de la composition primitive, lors de leur formation, étant moins facilement perturbées par l'environnement. Parmi les diverses planètes découvertes, il existe les super-Terres ( $1-10 M_{\oplus}$ ), cette classe d'objets particulièrement abondante, mais qu'on ne retrouve pas dans notre SS. Environ la moitié des étoiles de type solaire en posséderait une ou plus avec des périodes  $\lesssim 100 \text{j}$  (Mayor et al., 2011; Petigura et al., 2013). Il existe également les planètes *Super-Puff*, qui ont des masses de seulement quelques  $M_{\oplus}$ , mais des rayons plus grands que celui de Neptune, leur donnant une densité extrêmement faible. Elles sont cependant plus froides et moins massives que les Jupiter chaudes de faible densité (voir ci-bas), et ont toujours du mal à être expliquées par les modèles de formation, étant donné leur faible masse, mais leur atmosphère riche en H et He (ex. : Libby-Roberts et al., 2020). Enfin, on peut penser aux planètes de lave, ces planètes rocheuses ( $R_p \lesssim 2 R_{\oplus}$  et  $M_p \lesssim 10 M_{\oplus}$ ; faisant donc

---

1. Selon les bases de données de Exoplanets.org

2. On peut classifier les planètes en plus petits groupes, selon leur masses ou leur rayons, en comparant avec les planètes du SS. Par exemple, on peut parler des planètes de type super-Terre, de type sous-Neptune, de type Jupiter, etc., selon leurs paramètres observés.

partie de l'ensemble des super-Terres) qui sont tellement proche de leur étoile, que la température de surface atteint entre 2500 à 3000 K. Ces énormes températures mèneraient à la fusion et la vaporisation des roches, créant ainsi des océans d'éléments réfractaires sur l'ensemble du côté jour de la planète.

1.1.1.1. Jupiter chaude. Une Jupiter chaude (JC) est un type de planètes qu'on ne retrouve pas dans le SS. Leurs principales caractéristiques (taille, masse) ressemblent à celles de Jupiter, mais elles ont de très courtes périodes orbitales ( $P \lesssim 10$  jours ou  $a \lesssim 0.1$  ua). Comme elles sont très proches de leur étoile, elles ont des températures de surface très grandes (jusqu'à 1000–2500 K), ce qui leur a valu le nom de *Jupiter Chaudes*. De plus, la majorité de leurs orbites ont été circularisées par les forces de marées intenses près de l'étoile et ces planètes sont également sujettes à avoir une rotation synchrone (la période de rotation de la planète est la même que sa période de révolution autour de l'étoile, ce qui mène la planète à toujours montrer la même face à l'étoile). On en retrouve davantage autour des étoiles de type solaires qu'autour des étoiles de faible masse et naines brunes. Le consensus est que ces planètes se sont formées assez loin dans le disque protoplanétaire et auraient migré (via interaction avec le disque, ou avec d'autres planètes) jusqu'à leur position actuelle de manière non destructive. Étant donné leur température de surface extrême, leur verrouillage gravitationnel et leur courte période orbitale, on s'attend à ce que les JCs aient des atmosphères plutôt « exotiques » (ex. : gros gradient de température jour-nuit, inversion de température, convection en profondeur, etc.), mais qui sont plus faciles à étudier. En effet, elles ont plus de chance de transiter, le font sur des échelles de temps plus courtes et ont des hauteurs d'échelle relativement grandes (voir l'éq. 1.1.11 et explications ; Section 1.1.3.3). La première à avoir été découverte est 51 Peg b (Mayor & Queloz, 1995).

1.1.1.2. Sous-Saturne. Les sous-Saturnes (ou super-Neptunes) sont des planètes ayant des masses en deçà de celle de Saturne ( $\sim 95M_{\oplus}$  ; ou au dessus de celle de Neptune,  $\sim 17M_{\oplus}$ ). Relativement peu de ces planètes ont été observées : l'énorme écart de masse entre Neptune et Saturne/Jupiter, serait causé par l'accrétion accélérée de gaz qui s'enclenche une fois que la masse de la planète/protoplanète atteint  $10\text{--}20M_{\oplus}$ . De plus, une absence accrue de ce type de planète est observée à courtes périodes orbitales ( $P < 2\text{--}4$  j), dénommée le désert des Neptunes chaudes, de l'anglais *sub-Jovian/Neptunian desert*. À petites distances de leur étoile, ces planètes seraient fortement irradiées et perdraient leur atmosphère, formant ainsi des mini-Neptunes (comme GJ 3470 b ; Bonfils et al., 2012) ou encore des super-Terres, en laissant seulement le cœur rocheux (aussi appelé planètes *chthoniennes* – Lecavelier des Etangs et al. 2004, comme CoRoT-7 b, Léger et al. 2009). Il se pourrait également que ce soit des processus encore inconnus qui feraient que les planètes comme Neptune aient plus de difficultés à migrer près de leur étoile. Certaines planètes se retrouvent tout de même dans le désert des Neptunes chaudes,

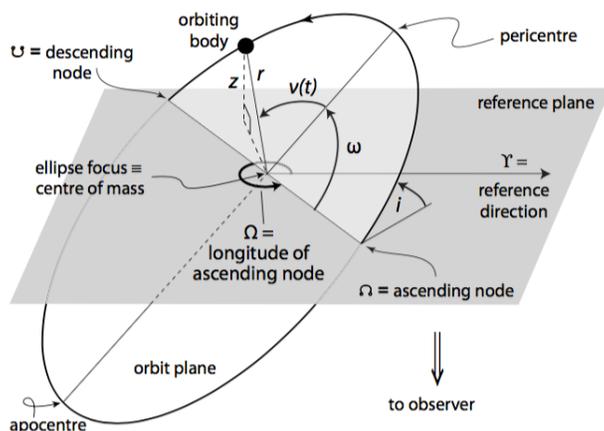


FIGURE 1.1 Schéma 3D d'une orbite planétaire elliptique. Figure tirée de Perryman (2011).  $i$  est l'inclinaison,  $\Omega$  est la longitude du nœud ascendant,  $\omega$  est l'argument du périastre et  $\nu(t)$  est l'anomalie vraie. Le plan de référence (gris) est tangent au plan de la sphère céleste, c.-à-d. ce que l'observateur voit.

comme NGTS-4 b, une planète de  $20 M_{\oplus}$  et un rayon de  $0.8 R_{\text{Nept}}$ , possédant toujours une atmosphère, probablement grâce à son cœur massif (West et al., 2019).

### 1.1.2. Méthodes de détection

Il existe une multitude de méthodes pour détecter les exoplanètes, qu'on peut largement répartir en 3 grandes catégories : dynamique, lentillage gravitationnel et photométrique. Parmi les méthodes dynamiques, on retrouve la méthode de vitesse radiale (mesure des variations de la vitesse radiale de l'étoile) et l'astrométrie (mesure des changements de position de l'étoile), toutes deux causées par l'influence gravitationnelle de la planète. On retrouve aussi le chronométrage d'orbite (mesure des variations de temps de transit d'une planète causées par l'influence gravitationnelle d'autres planètes du même système). Ensuite, le lentillage gravitationnel utilise le fait que lorsqu'un système étoile-planète déforme l'espace-temps et dévie, telle une lentille, la lumière d'un objet brillant de fond : une brève microamplification se produit lorsque la planète passe, en parcourant son orbite, précisément sur la ligne de visée de son étoile et l'objet de fond. Enfin, pour les méthodes photométriques, il y a d'abord l'imagerie haut contraste (ou imagerie directe ; une image est directement prise de la planète), et ensuite plusieurs autres méthodes indirectes, la principale étant la méthode du transit (détection d'une diminution périodique de la luminosité de l'étoile causée par le passage de la planète devant l'étoile). Cependant, nous irons plus en profondeur seulement pour les deux méthodes qui sont particulièrement utiles pour l'étude de l'atmosphère des exoplanètes, soit la méthode de vitesse radiale et la méthode du transit.

1.1.2.1. Méthode de vitesses radiales. La méthode des vitesses radiales est une méthode de détection indirecte qui se base sur les effets dynamiques générés par l'interaction gravitationnelle entre deux objets. En effet, une planète n'est pas réellement en orbite autour de son étoile. Plutôt, la planète et l'étoile sont en orbite autour du barycentre du système (le centre de masse). L'étoile possède donc aussi un mouvement orbital qui sera plus ou moins accentué dépendant des paramètres du système<sup>3</sup>. Si l'orbite n'est pas vue de face ( $i \neq 0$ , voir Figure 1.1), ceci va générer une perturbation périodique de la *vitesse radiale* (VR ;  $v_r = \dot{z}$ ) de l'étoile hôte, la composante de vitesse dans la direction de la ligne de visée de l'observateur. Grâce à l'*effet Doppler*, on peut détecter cette variation de VR. Les raies en absorption (ou en émission) dans le spectre d'une étoile vont alors se déplacer par rapport à leur position au repos (correspondant à une VR nulle). En général, c'est en mesurant le décalage Doppler qu'on estime la VR des étoiles et la présence d'une planète est déduite par l'observation d'une perturbation périodique (voir Figure 1.2). La longueur d'onde observée (reçue)  $\lambda_{\text{obs}}$  est donnée par l'équation relativiste suivante :

$$\lambda_{\text{obs}} = \lambda_0 \left( \frac{1 - v_s/c}{1 - v_{\text{obs}}/c} \right), \quad (1.1.1)$$

où  $\lambda_0$  est la longueur d'onde émise,  $v_{\text{obs}}$  et  $v_s$  sont les vitesses de l'observateur et de la source, respectivement, et  $c$  est la vitesse de la lumière.

L'importance du signal en VR de l'étoile est donné par la *demi-amplitude en vitesse radiale*,  $K$ , qui est définie comme suit :

$$K = \left( \frac{2\pi G}{P} \right)^{1/3} \frac{M_p \sin i}{(M_* + M_p)^{2/3}} \frac{1}{(1 - e^2)^{1/2}}, \quad (1.1.2)$$

où  $G$  est la constante gravitationnelle,  $P$  est la période,  $i$  est l'inclinaison,  $e$  est l'excentricité de l'orbite, et  $M_p$  et  $M_*$  sont les masses de la planète et de l'étoile, respectivement<sup>4</sup>. Pour donner une référence d'échelle, la demi-amplitude du Soleil induite par Jupiter est de  $K_J \simeq 12.5$  m/s, tandis que pour la Terre elle est de  $K_{\oplus} \simeq 0.09$  m/s. Cette méthode est donc plus sensible aux planètes massives sur des orbites à courtes périodes, orbitant des étoiles de faibles masses. Les planètes de type Jupiter chaude et les étoiles de types naines rouges sont donc les cibles les plus avantageuses à étudier.

Un des grands avantages de cette méthode est qu'on peut déterminer la masse de la planète, à condition de connaître la masse de l'étoile. La masse de la planète sera incertaine d'un facteur

3. Comme les étoiles sont souvent beaucoup plus massives que leurs planètes, le centre de masse du système est souvent très près du centre de l'étoile. Cependant, plus la planète est massive, plus le centre de masse s'éloigne du centre de l'étoile. Par exemple, le centre de masse entre le Soleil et Jupiter est tout juste au-delà de la surface du Soleil (Vallisneri et al., 2020).

4. L'indice « \* » sera utilisé pour dénoter les paramètres de l'étoile hôte, tandis que « p » dénotera les paramètres de la planète (sauf si explicité autrement). Les variables  $R$ ,  $M$  et  $T$  représenteront le rayon, la masse et la température de l'objet considéré, respectivement. ( $R$ , sans indice, pourra aussi représenter le pouvoir de résolution d'un spectrographe, selon le contexte.)

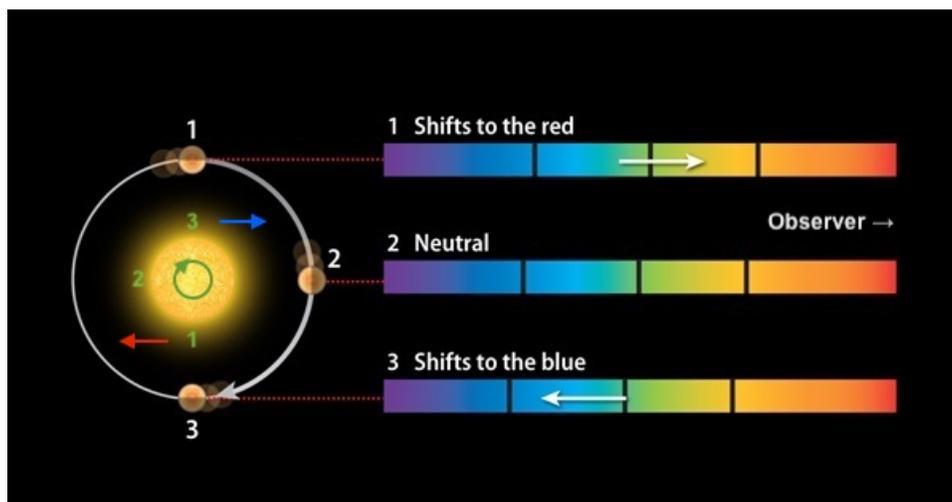
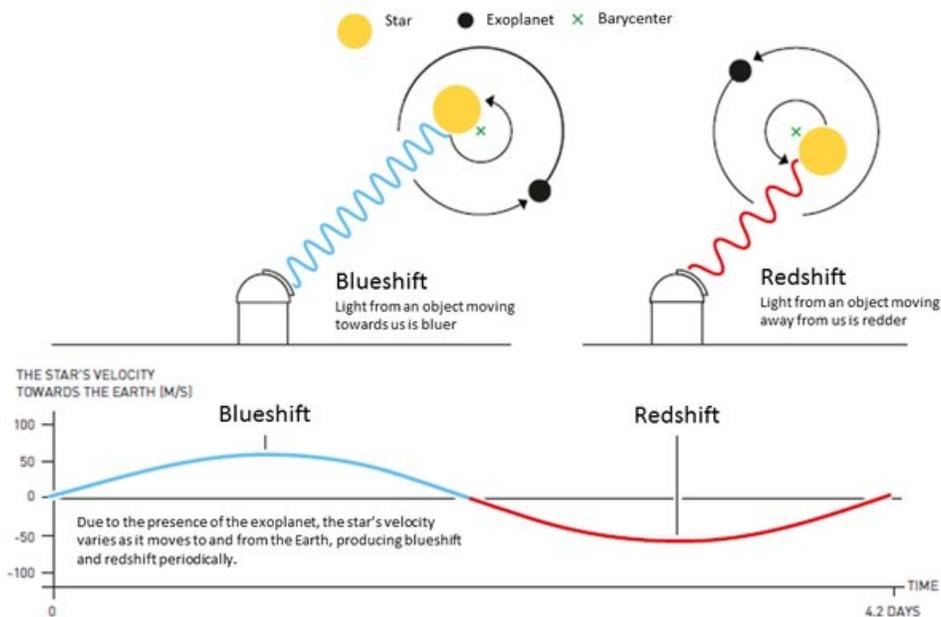


FIGURE 1.2 Méthode de détection d'exoplanète par vitesse radiale. Lorsque la VR est nulle (2), le spectre n'est pas décalé. Par contre, lorsque l'étoile s'éloigne (1) ou s'approche (3) de l'observateur, à cause du mouvement de recul causé par l'orbite de la planète, le spectre est décalé vers le rouge ou le bleu, respectivement. Crédits : Space Book, Las Cumbres Observatory.

$\sin i$ , à moins que l'inclinaison soit également connue. Un exemple de courbe de VR est montré à la Figure 1.3.

Il ne faut cependant pas oublier d'inclure les autres contributions à la VR, par exemple la vitesse systémique du système (notée  $\gamma$ ) par rapport au barycentre du SS, ainsi que  $d$ , les décalages de VR causés par la dérive possible des instruments lors de la prise des données. Il

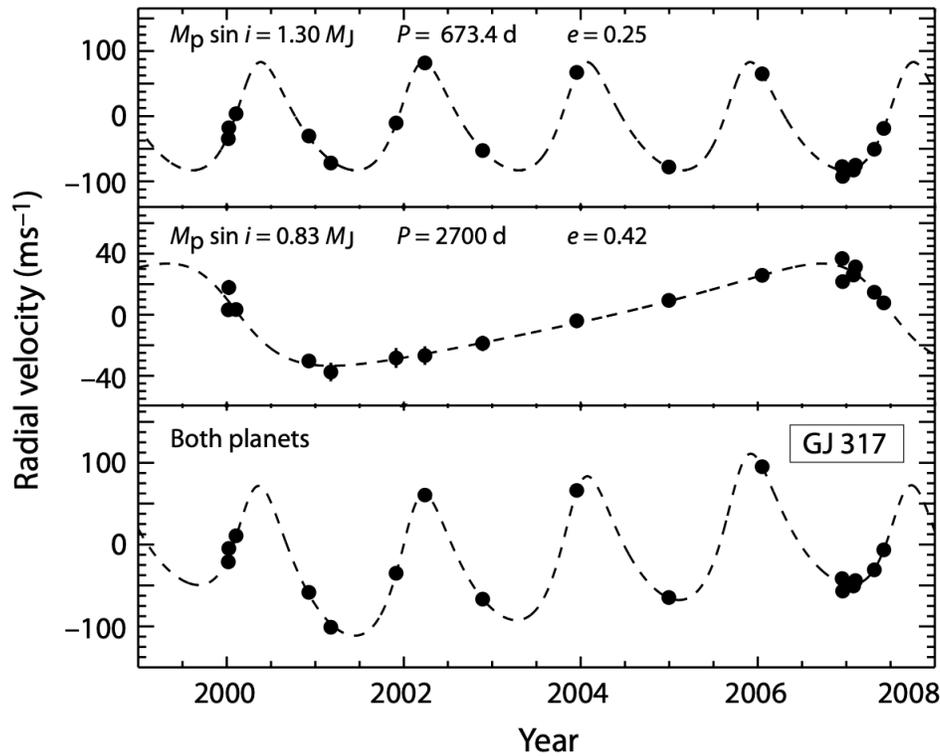


FIGURE 1.3 Exemple de courbe de vitesse radiale pour la naine M GJ 317 et un modèle à deux planètes. Les valeurs observées (points noirs) ont un bon accord avec les modèles planétaires individuels (lignes en tiret; deux panneaux du haut) et combinés (panneau du bas). Tiré de Johnson et al. (2007).

faut aussi prendre en compte le mouvement de rotation et orbital de la Terre autour du Soleil (qui donne une contribution additionnelle jusqu'à 0.5 et 30 km/s, respectivement), souvent appelé BERV (de l'anglais *Barycentric Earth Radial Velocity*). Enfin, les sources de bruit, qu'on appelle « *jitter* » sont aussi à considérer. L'activité stellaire, les inhomogénéités de l'atmosphère (taches et facules, où le rayonnement est plus faible et plus intense, respectivement) et les oscillations stellaires, la granulation de surface, le rougissement gravitationnel, des compagnons planétaires inconnus et autres erreurs systématiques contribuent à ce *jitter*.

La mesure de la vitesse radiale de l'étoile est obtenue par la mesure de la position des raies du spectre. On voudra donc obtenir des spectres à relativement haut pouvoir de résolution ( $R = \lambda/\Delta\lambda$ ; qui est un indice sur la quantité de détails présents dans un spectre, ou encore, la capacité de distinguer des raies spectrales adjacentes dans un spectre) pour maximiser la précision sur la position des raies spectrales.

La première exoplanète orbitant une étoile de type solaire (51 Peg b; Mayor & Queloz 1995) a été détectée grâce à la méthode de vitesse radiale, avec l'instrument ELODIE, un spectromètre

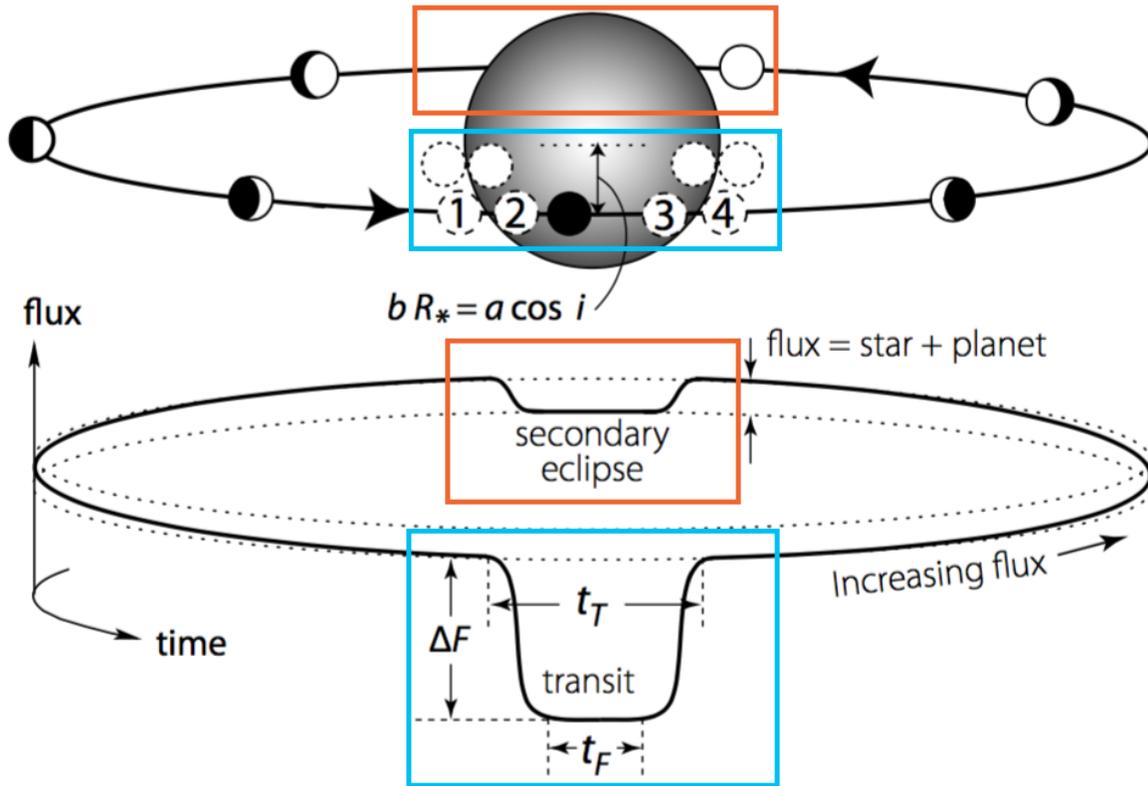


FIGURE 1.4 Schéma d'un transit, en bleu, et d'une éclipse secondaire, en orange (haut) et de leur courbe de lumière associée (bas). Figure tirée de Perryman (2011).

échelle avec  $R = 42000$ , installé à l'observatoire de Haute-Provence (1.93-m). Cette méthode est la deuxième plus prolifique après la méthode du transit.

1.1.2.2. Méthode du transit. Cette méthode de détection photométrique et indirecte se base sur le principe que lorsqu'un objet passe devant une source de lumière, l'intensité lumineuse reçue par un observateur diminue. On peut déduire la présence de planètes avec ce même principe : lorsqu'une planète transite, elle bloque une partie de la lumière de l'étoile. Ceci se traduit en une diminution temporaire et périodique de l'intensité lumineuse stellaire reçue sur Terre (d'au plus quelques pour cent ; voir Figure 1.4), qui est autrement constante<sup>5</sup>. Pour observer un tel transit, il est nécessaire d'avoir un alignement étoile-planète-observateur presque parfait : l'inclinaison du plan orbital doit avoisiner  $90^\circ$ .

La profondeur du transit  $\Delta F/F$  est donnée par

$$\frac{\Delta F}{F} \simeq \left( \frac{R_p}{R_*} \right)^2, \quad (1.1.3)$$

5. Il existe d'autres sources de variabilité stellaire, mais celles-ci ne présentent pas les mêmes caractéristiques.

qu'on peut relier à la fraction de l'aire du disque stellaire qui est bloquée par la planète ( $F$  est le flux de l'étoile et  $\Delta F$  la diminution de flux causé par la planète lors du transit). En considérant l'assombrissement centre-bord (la baisse de luminosité qu'on observe entre le centre d'une étoile et son bord, causée par des effets de profondeur optique étant donné que la densité et la température diminuent avec l'altitude dans l'atmosphère de l'étoile), l'équation ci-haut devient exacte et permet alors, si le rayon de l'étoile ( $R_*$ ) est connu, de déterminer le rayon de la planète ( $R_p$ ). En combinant avec la masse obtenue par VR, on peut calculer sa densité et estimer la composition globale de son intérieur. Le signal sera plus grand pour les petites étoiles et pour les grosses planètes.

Lorsque la planète passe devant son étoile, on parlera de *transit* (ou d'*éclipse primaire*), mais lorsqu'elle passe derrière son étoile, on parlera d'*éclipse secondaire* (Fig. 1.4). Similairement, lorsque la planète passe derrière l'étoile, il y a une baisse de flux associée à la perte de la lumière émise et réfléchiée par la planète, quoique la dernière soit très faible dans la plupart des cas. Le flux restant lors de l'éclipse secondaire ne provient que de l'étoile.

Il y a cependant certaines mises en garde pour cette méthode. D'abord, le signal (la baisse de luminosité) doit être périodique ; il est donc important d'avoir plus d'un transit pour confirmer la nature planétaire de celui-ci. Pour les planètes ayant de grands demi-grands axes, il est beaucoup moins probable d'avoir un alignement menant à un transit et encore moins d'en observer plusieurs étant donné les plus longues périodes orbitales. Cette méthode est donc plus sensible aux petites orbites. Ensuite, il faut s'assurer que la baisse de luminosité provienne bien d'une planète. Particulièrement, les systèmes d'étoiles doubles ayant un plan orbital vu par la tranche ( $i \simeq 90^\circ$  ; aussi appelés *binaires éclipsantes*) peuvent générer de fausses détections d'exoplanètes<sup>6</sup>. Enfin, cette méthode peut également subir la contamination de structures présentes à la surface de l'étoile, telles que les taches stellaires, granulation, oscillations non radiales et autres phénomènes pouvant faire varier la luminosité apparente de l'étoile.

La première détection faite par la méthode du transit a été celle de HD 209458 b, par Charbonneau et al. (2000) et Henry et al. (2000). *Kepler* (téléscope spatial lancé en 2009 pour faire de la détection d'exoplanètes par photométrie), est l'instrument le plus prolifique côté détection d'exoplanètes par méthode du transit. Il aura regardé à un seul endroit du ciel pour pouvoir détecter des planètes ayant de grandes périodes, et des systèmes plus éloignés du voisinage solaire. Il a, à ce jour, permis de détecter plus de 2300 exoplanètes confirmées<sup>7</sup>, dont 30 dans la zone

---

6. Les systèmes d'étoiles doubles vont également causer une baisse de luminosité périodique. Cette baisse est généralement beaucoup plus importante que les transits planétaires, dépendant de la nature des deux objets. Cependant, si l'alignement n'est pas parfait et qu'on a un transit *rasant* (ou *grazing transit*), la baisse de lumière associée pourrait être confondue à celle d'un transit planétaire. Une autre possibilité est qu'un tel système éclipsant situé en arrière plan d'une étoile cible soit confondu dans l'ouverture photométrique de cette dernière, introduisant un signal de faible amplitude s'apparentant à un transit d'exoplanète.

7. Site Web de Kepler

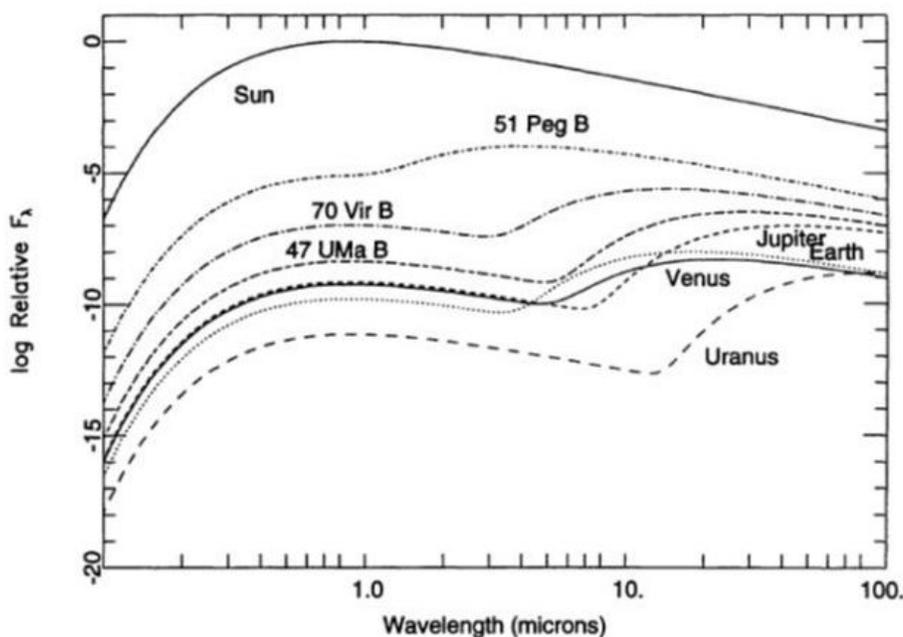


FIGURE 1.5 Flux relatif du Soleil, Vénus, Terre, Jupiter, Uranus, ainsi que les planètes orbitant 51 Pegasi, 70 Virginis et 47 Ursa Majoris. Crédits : Présentation de Arron Anderson

habitable de leur étoile (là où l'eau peut se retrouver sous forme liquide). Son successeur, TESS (*Transiting Exoplanet Survey Satellite*), a été lancé en 2018. Contrairement à Kepler, TESS regarde plutôt partout dans la sphère céleste, mais pour moins longtemps. Cela veut donc dire qu'on trouve plutôt des planètes ayant de petites périodes, mais autour d'étoiles plus brillantes.

### 1.1.3. Méthodes de caractérisation

Il existe une multitude de façons de « caractériser » les exoplanètes, par exemple en étudiant leur orbite, leurs interactions avec leur étoile et/ou les autres planètes du système, leur composition interne, leur champ magnétique, etc. Ici, nous nous concentrerons sur les méthodes de caractérisation de leur atmosphère.

On pourrait penser que la méthode la plus naturelle d'observer l'atmosphère d'une planète serait d'en prendre une image (et/ou un spectre) directement. Cependant, la méthode d'imagerie directe (imagerie haut-contraste) est limitée aux planètes géantes (et/ou naines brunes), qui sont encore jeunes et brillantes, et qui sont relativement loin de leur étoile (ou isolées), étant donné l'énorme contraste étoile-planète. Certaines planètes ont tout de même pu être caractérisées par l'entremise de cette méthode, par exemple Beta Pictoris b (ex. : Currie et al., 2013), HR 8799 b (ex. : Barman et al., 2015) et 51 Eri b (ex. : Macintosh et al., 2015).

En effet, le flux émis par la planète est très faible comparativement à celui émis par l'étoile ; la lumière de la planète est « noyée » dans celle de l'étoile, tel que présenté sur la Figure 1.5.

Sur cette figure, on peut voir qu'il y a deux pics dans le spectre des planètes. La partie visible du spectre (pic de gauche) correspond à la lumière de l'étoile qui a été réfléchi(e) (et diffusée) par la planète et la partie dans l'infrarouge (IR ; pic de droite) provient de l'émission thermique de la planète. La grande différence de flux entre l'étoile et la planète (le grand contraste) est ce qui rend difficile la détection de la radiation de la planète. Par exemple, Jupiter a un contraste de  $F_p/F_\star \sim 10^{-9}-10^{-10}$  dans le visible. Heureusement, on peut voir que le contraste s'améliore lorsqu'on considère l'émission thermique, à plus grandes longueurs d'onde. Le contraste de Jupiter, à son pic d'émission autour de  $\sim 20-30 \mu\text{m}$ , diminue et le ratio de flux augmente à  $\sim 10^{-6}$ . Le contraste est encore meilleur pour les JCs, comme 51 Peg b ( $F_{\text{JC}}/F_\star \sim 10^{-3}$ ), d'où l'intérêt d'étudier ces objets et de le faire dans l'IR.

Le rapport de flux planète-étoile est donné par :

$$\frac{F_{\oplus,p}}{F_{\oplus,\star}} = \frac{F_p R_p^2}{F_\star R_\star^2}, \quad (1.1.4)$$

où  $F_{\oplus,p(\star)} = F_{p(\star)}(R_{p(\star)}/D_\oplus)^2$  est le flux de la planète (ou de l'étoile) observé de la Terre,  $F_{p(\star)}$  est le flux de surface de la planète (de l'étoile), et  $D_\oplus$  est la distance du système à la Terre.

Avec la capacité des instruments et méthodes d'aujourd'hui, il est malheureusement très difficile d'imager (et donc de caractériser) des planètes similaires à la Terre ou même à Jupiter (avec des contrastes  $\lesssim 10^{-6}$ ). Toutefois, des efforts sont déployés dans la recherche et la technologie de l'imagerie directe pour rendre possible l'étude des systèmes planétaires plus semblables au nôtre. Il faudra toutefois attendre l'avènement des télescopes de 30 mètres avant de pouvoir espérer le faire pour une planète de type terrestre. En attendant, on utilise la méthode des transits pour analyser l'atmosphère des exoplanètes, soit en transmission durant le transit ou en émission avant/après l'éclipse secondaire. On peut également suivre l'émission des planètes qui ne transitent pas sur toute leur orbite et avoir leur courbe de phase.

1.1.3.1. Spectre en émission. Comme mentionné à la Section 1.1.2.2, lorsqu'un système transite ( $i \simeq 90^\circ$ ), la planète passe périodiquement devant et derrière son étoile. Lorsque la planète est cachée par son étoile (éclipse secondaire), on observe seulement la lumière de l'étoile. Cependant, juste avant (ou après) l'éclipse, la lumière reçue par l'observateur est une combinaison de celle de l'étoile et du côté jour de la planète (encadrés orange de la Fig. 1.4). En prenant une mesure différentielle des deux signaux, on peut révéler l'émission de la planète seule (voir Figure 1.6).

La profondeur de l'éclipse est une mesure du ratio des flux émis par la planète et de l'étoile. Si on suppose que l'étoile et la planète émettent comme des corps noirs ( $B(T,\lambda)$ ) à des températures effectives et d'équilibre  $T_\star$  et  $T_p$ , respectivement, et qu'on voit un seul hémisphère,

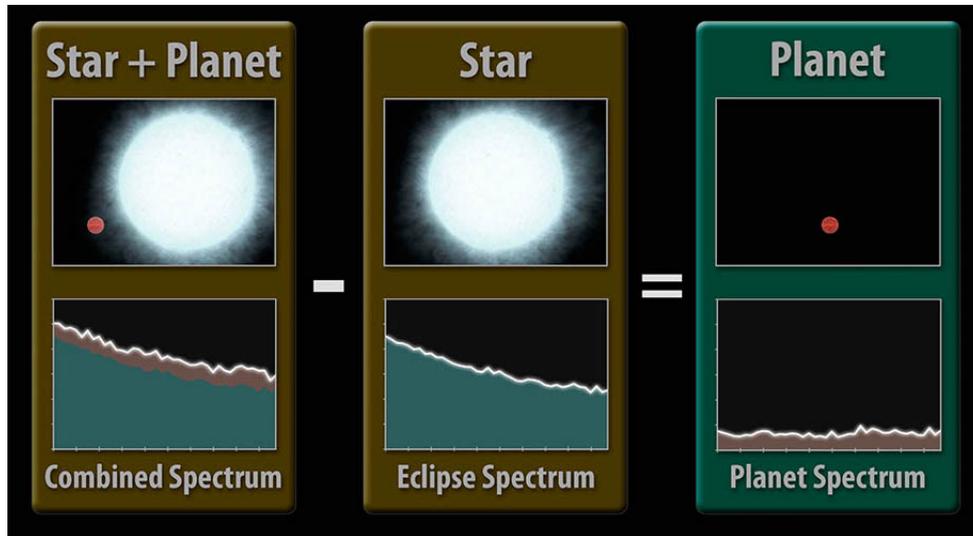


FIGURE 1.6 Visualisation du concept d'isoler le spectre d'une planète par spectroscopie d'émission. On prend d'abord un spectre qui combine les contributions de la planète et de l'étoile (gauche). Durant l'éclipse secondaire, on perd la contribution de la planète, et ne reste que celle de l'étoile (milieu). En faisant une analyse différentielle, on isole le signal de la planète (droite). Crédit : NASA/JPL-Caltech/R. Hurt (SSC/Caltech).

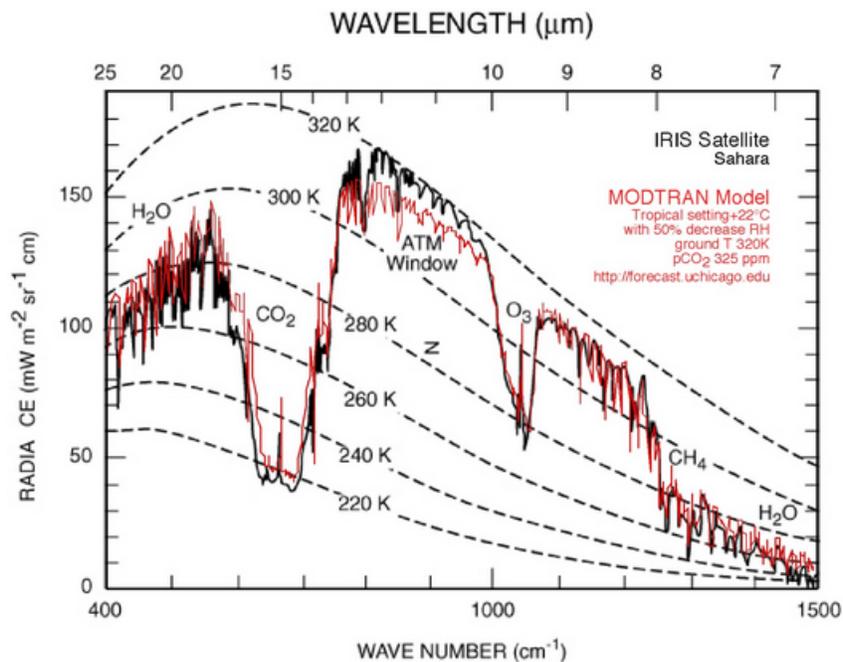


FIGURE 1.7 Spectre IR de la Terre, comparé à un modèle MODTRAN. Les courbes en tirets montrent des spectres d'émission de corps noirs à différentes températures. On peut voir qu'on peut associer une température propre aux différentes parties du spectre, représentant leur température de luminosité. Celles-ci nous indiquent à quelle température les bandes d'absorption/d'émission sont formées. Figure de MODTRAN, University of Chicago.

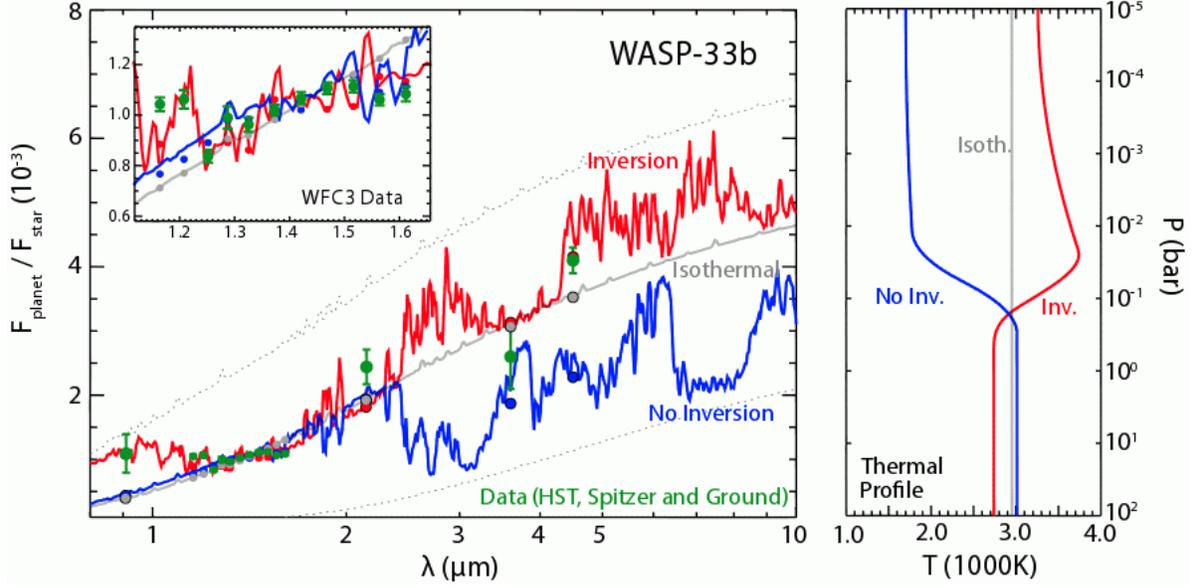


FIGURE 1.8 Exemple de spectre d'émission de la planète WASP-33b. En mesurant le ratio de flux  $F_p/F_*$  (points verts), on peut comparer à des modèles théoriques (courbes rouge, grise et bleue) et extraire beaucoup d'informations, tel que le profil vertical (panneau de droite; voir Section 1.3.3), et la composition (Section 1.3.4). Crédit : Figure tirée de Haynes et al. (2015).

$F_*(\lambda) = \pi B(T_*, \lambda)$  et  $F_p(\lambda) = \pi B(T_p, \lambda)$ . On obtient alors :

$$\frac{F_{\oplus,p}(\lambda)}{F_{\oplus,*}(\lambda)} = \frac{B_p(T_p, \lambda) R_p^2}{B_*(T_*, \lambda) R_*^2} \stackrel{\text{R-J}}{=} \left[ \frac{T_p}{T_*} \right] \left( \frac{R_p}{R_*} \right)^2. \quad (1.1.5)$$

Pour la première égalité, une version dépendante de la longueur d'onde de l'équation 1.1.4 a été utilisée. La seconde égalité est valide seulement dans l'intervalle de fréquence où l'approximation Rayleigh-Jeans est valide ( $hc/\lambda kT \ll 1$ ) :  $B(T, \lambda) \approx 2\pi ck_B T / \lambda^4$ . Pour les planètes, cette limite est atteinte avec des observations dans l'IR moyen à lointain ( $\gtrsim 10 \mu\text{m}$ ). On peut alors estimer la température du côté jour de la planète. Par contre, les spectres peuvent ne pas ressembler du tout à des corps noirs, et c'est pourquoi on voudra plutôt utiliser les températures de luminosité  $T_{\text{lum}}(\lambda)$  de la planète (voir Figure 1.7) et de l'étoile au lieu de  $T_p$  et  $T_*$ . En mesurant la profondeur de l'éclipse pour chaque courbe de lumière spectroscopique (pour chaque longueur d'onde), on peut reconstruire le spectre en émission de l'atmosphère de la planète (voir Figure 1.8). Ceci se traduit alors comme :

$$\frac{F_{\oplus,p}(\lambda)}{F_{\oplus,*}(\lambda)} = \left[ \frac{T_{\text{lum},p}(\lambda)}{T_{\text{lum},*}(\lambda)} \right] \left( \frac{R_p}{R_*} \right)^2. \quad (1.1.6)$$

Comme la majorité du flux émis par la planète se trouve dans l'IR, les études de « spectroscopie d'émission » ont été faites principalement dans l'IR.

Il est possible de faire la même chose pour des planètes qui ne transitent pas : le même principe de combinaison des spectres de la planète et de l'étoile s'applique<sup>8</sup>. Il faudra cependant mettre plus d'effort pour soustraire le signal de l'étoile puisque sur toutes les données, une fraction du signal de la planète sera présent et pourrait venir contaminer le spectre stellaire à soustraire. Comme le signal planétaire est très petit et précieux, on voudra minimiser sa contamination et sa soustraction subséquente, qui pourrait l'éliminer complètement. C'est pourquoi on utilisera surtout des données haute résolution, qui permettent de démêler les signaux grâce à leurs différents décalages Doppler (cette méthode sera expliquée en détail à la Section 1.4.2).

1.1.3.2. Courbe de phase. Cette méthode consiste à obtenir des données photométriques ou spectroscopiques continuellement sur une orbite complète (ou partielle). Ceci correspond donc aux régions à l'extérieur des encadrés oranges et bleus sur la Figure 1.4. En effet, la forme de la courbe de lumière permet de déterminer le contraste de température entre le côté jour et le côté nuit, d'estimer l'albédo de la planète, et de déterminer si le point le plus chaud est décalé du point sous-stellaire (et donc la présence d'une dynamique globale). Tel que mentionné brièvement à la section 1.1.1.1, l'orbite de la plupart des Jupiters chaudes et des autres planètes qui sont très près de leur étoile sera, au fil du temps, "circularisée" par les forces de marées, menant à ce qu'on appelle le verrouillage gravitationnel. Alors, l'orbite de la planète devient circulaire (excentricité nulle) et la période de révolution de la planète devient égale à sa période de rotation, ce qui implique que la planète montre toujours la même face à son étoile (comme la Lune qui montre toujours la même face à la Terre). Le côté jour permanent de la planète est alors continuellement irradié par l'étoile, possédera alors ce qu'on appelle un "point chaud". Sans vents, ce point chaud se trouve au point sous-stellaire (Figure 1.11), mais avec la présence de forts vents (comme les courants-jets), le point chaud peut être décalé du point sous-stellaire (dans la direction des vents, voir Section 1.3.5). En analysant l'émission thermique du côté jour et du côté nuit, on peut avoir une idée de l'efficacité de la circulation de l'atmosphère : un grand contraste de température indique une faible redistribution d'énergie, tandis qu'un petit contraste indique une bonne redistribution.

En analysant la dépendance de la courbe de phase avec la longueur d'onde, on peut retrouver (via la spectroscopie en émission) le spectre de la planète en fonction de la phase orbitale et ainsi construire une carte de température de luminosité de la surface de la planète. C'est ce que Stevenson et al. (2014) ont fait pour WASP-43 b : ils ont donc déterminé la variation du spectre en émission thermique pour une rotation complète<sup>9</sup>.

Deux points qui sont ressortis des études de courbe de phase sont les suivants. 1) Plus la température d'irradiation (ou  $T_{eq}$ ) est grande, moins efficace sera la redistribution d'énergie (et

---

8. On favorisera les moments où la planète montre son côté jour, mais le principe s'applique aussi pour le côté nuit, mais qui est encore plus difficile à détecter.

9. Une vidéo montrant cette variation est disponible : [Movie S1](#)

donc plus grand sera le contraste jour-nuit ; Zhang et al. 2018). Ceci impactera aussi le décalage du point chaud (il sera moins décalé pour des températures plus grandes). 2) La majorité des courbes de phase des JCs montrent justement des points chauds décalés vers l'est du point sous-stellaire. Ce décalage est prédit par les modèles de circulation globale (ou GCMs, de l'anglais *Global Circulation Models*). Par exemple, Knutson et al. (2007) ont créé une carte de température de la JC HD189733b. Ils ont pu déterminer que le point chaud est décalé vers l'est, avec un faible contraste en température jour-nuit ( $\sim 300$  K d'écart), ce qui indique que la redistribution d'énergie se fait efficacement.

1.1.3.3. Spectre de transmission. Lorsque la planète transite devant son étoile, une partie de la lumière provenant de l'étoile qui devait nous rejoindre est alors bloquée par la planète (encadrés bleus de la Fig. 1.4). En fait, une fraction de cette lumière réussit à traverser la partie optiquement mince de l'atmosphère de la planète, tout en prenant quelques caractéristiques spectrales au passage (Figure 1.11 ; expliqué plus en détail à la Section 1.3.1). De façon similaire à la spectroscopie d'émission, on peut déterminer la composition de l'atmosphère et éventuellement voir certains effets dynamiques en isolant la contribution de la planète.

En fait, l'opacité de l'atmosphère, qui dépend de sa composition, varie avec la longueur d'onde  $\lambda$  et influencera la profondeur du transit (voir Figure 1.9). Ceci peut se traduire par un changement du rayon effectif de la planète,  $R_p(\lambda)$ . La profondeur du transit, reliée à la fraction du disque stellaire bloquée par la planète et qui est ce qu'on mesure, peut alors s'écrire comme<sup>10</sup> :

$$\frac{\Delta f}{f} = \frac{\pi R_p^2(\lambda)}{\pi R_\star^2}. \quad (1.1.7)$$

Un exemple de spectre de transmission est montré à la Figure 1.10.

On peut cependant décomposer le numérateur en 2 composantes, soit une partie complètement opaque (transmittance nulle ;  $\text{Tr}(r,\lambda) = e^{-\tau(r,\lambda)} = 0$ ) à toutes les longueurs d'onde (pour une distance au centre de la planète  $r$  allant de 0 à  $R_{p,0}$ , le rayon de référence) et une portion semi-transparente qui dépend de  $\lambda$ , contribuant une hauteur d'atmosphère effective supplémentaire  $h_{\text{eff}}(\lambda)$  (pour  $r$  allant de  $R_{p,0}$  à  $h_{\text{eff}}(\lambda)$ ) :

$$\pi R_p^2(\lambda) = \pi [R_{p,0} + h_{\text{eff}}(\lambda)]^2 \quad (1.1.8)$$

$$= \int_0^\infty 2\pi r (1 - \text{Tr}(r,\lambda)) dr \quad (1.1.9)$$

En combinant les équations 1.1.7 et 1.1.9, on obtient :

$$\frac{\Delta f}{f} = \frac{\pi R_p^2}{\pi R_\star^2} + \frac{\int_{R_p}^\infty 2\pi r (1 - e^{-\tau(r,\lambda)}) dr}{\pi R_\star^2}. \quad (1.1.10)$$

---

10. Pour simplifier les équations présentées et focaliser sur le principe de base, l'effet de l'assombrissement centre-bord de l'étoile sera négligé ici.

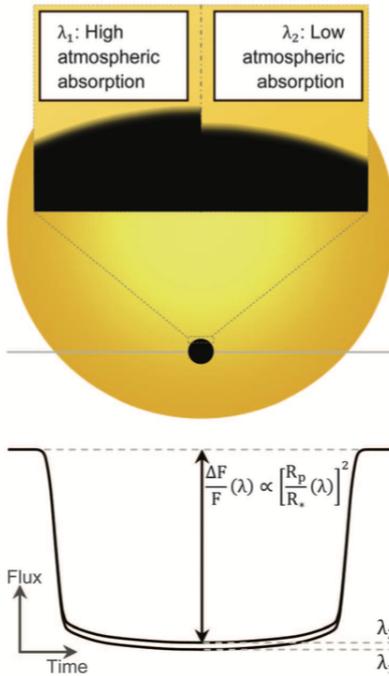


FIGURE 1.9 Variation de la profondeur du transit ( $\Delta f/f$ )( $\lambda$ ), induit par la dépendance en longueur d'onde de l'opacité de l'atmosphère de la planète. Figure tirée de de Wit & Seager (2013).

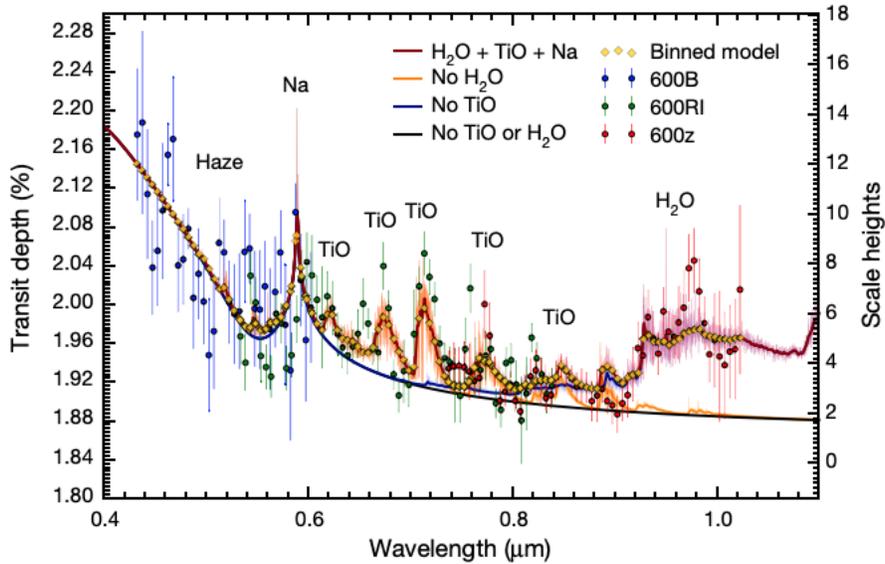


FIGURE 1.10 Exemple de spectre de transmission pour la planète WASP-19b. On voit que la profondeur du transit observée (et donc le rayon effectif de la planète; points bleus, verts et rouges) varie avec la longueur d'onde. En comparant avec des modèles d'atmosphère, on peut ainsi dévoiler la présence de certaines molécules (comme H<sub>2</sub>O, TiO, Na, et des aérosols). Figure tirée de Sedaghati et al. (2017).

Seulement le dernier terme au membre de droite va générer des caractéristiques spectrales. C'est donc l'importance de ce terme qui va surtout nous intéresser. En développant l'éq. 1.1.8, on a  $\pi R_P^2(\lambda) = \pi[R_{p,0}^2 + 2R_{p,0}h_{\text{eff}}(\lambda) + h_{\text{eff}}^2(\lambda)]$ , où le premier terme au membre de droite représente la partie complètement opaque de la planète et les deux autres représentent la partie semi-transparente de l'atmosphère. Si on veut seulement avoir un ordre de grandeur du signal qu'on peut espérer observer, on peut poser que  $R_P \gg h_{\text{eff}}$  (ce qui est généralement le cas). On a alors  $\text{Aire}_{\text{atm}} \approx 2\pi R_P h_{\text{eff}}$ , puisque le terme en  $h_{\text{eff}}^2$  est négligeable. De plus, il est commun de poser une épaisseur  $h_{\text{eff}} \simeq 5H$  (Seager et al., 2009), où  $H = kT/\mu g$  est la hauteur d'échelle (développé dans la Section 1.3.3). En combinant le tout avec la définition de  $g$  et de la densité  $\rho$ , on trouve :

$$\frac{\Delta f_{\text{atm}}}{f} = \frac{30kT}{4\pi\mu G\rho R_*^2}, \quad (1.1.11)$$

où  $T$ ,  $\mu$  et  $\rho$  sont la température, le poids moléculaire moyen de l'atmosphère et la densité de la planète, respectivement. Par exemple, pour une JC orbitant une étoile de type solaire, le signal est de l'ordre de  $\sim 300$  ppm ( $T \sim 1000$  K,  $\mu \sim 2$  u,  $\rho \sim 1$  g/cm<sup>3</sup>), tandis que pour une planète comme la Terre dans la zone habitable, le signal n'est que de  $\sim 1$  ppm ( $T \sim 250$  K,  $\mu \sim 29$  u,  $\rho \sim 5.5$  g/cm<sup>3</sup>). En comparaison, Hubble est capable d'atteindre un bruit de  $\sim 30$  ppm (Kreidberg et al., 2014). Notre meilleure chance de détecter du signal est donc pour les planètes avec des atmosphères chaudes, ayant de petits poids moléculaires moyens, et de faibles densités (autrement dit : de grandes valeurs de  $H$ ) et qui orbitent des petites étoiles.

Le rapport de flux en- et hors-transit est alors :

$$\frac{F_{\oplus,\text{trans}}(\lambda)}{F_{\oplus,*}(\lambda)} = \frac{f - \Delta f}{f} = 1 - \frac{\Delta f}{f}. \quad (1.1.12)$$

Étant donnée la géométrie de cette mesure, la spectroscopie de transit va surtout aller chercher du signal en haute atmosphère (Figure 1.11). Plus les rayons passent près du centre de la planète, plus ils auront de couches d'atmosphère à traverser (qui peuvent être à différentes températures) et plus l'intensité sera atténuée. Aussi, l'atmosphère qui est sondée est celle du terminateur, c.-à-d. l'anneau d'atmosphère à la limite entre le côté jour et le côté nuit. Comme on ne résout ni le disque stellaire ni le disque planétaire, il est possible que la mesure du spectre, qui est une moyenne sur tout l'anneau, ne soit pas représentative des vraies conditions (jour vs nuit). Par exemple, comme vu précédemment, les JCs ont de forts courants-jets qui décalent le point le plus chaud du point sous-stellaire et le point le plus froid du point anti-stellaire. Cela a pour effet de créer une différence de température entre le côté matin du terminateur et le côté soir (celui du soir étant plus chaud que celui du matin). L'atmosphère du côté soir/chaud sera plus étendue et aura donc un signal plus grand (de par l'équation de la hauteur caractéristique  $H$ ) et inversement pour le côté matin/froid. La seule façon de résoudre cette différence est d'analyser

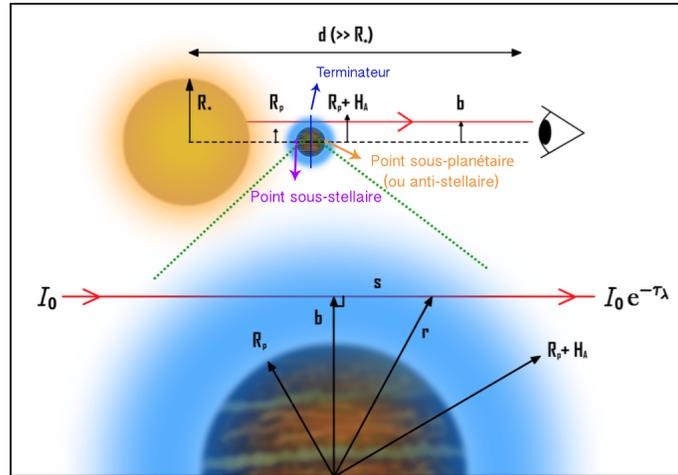


FIGURE 1.11 Géométrie d'un système en transit. Les rayons partant de l'étoile (d'intensité  $I_0$ ) traversent l'atmosphère de la planète et sont atténués (diminués d'un facteur  $e^{-\tau_\lambda}$ ). Dans ce schéma, le côté jour est la moitié gauche de la planète (face à l'étoile), et le côté nuit à droite (face à l'observateur). Figure tirée de MacDonald & Madhusudhan (2017).

l'entrée en transit (*ingress*) et la sortie (*egress*) séparément. Comme ces moments ne durent généralement pas très longtemps, cela demande que le signal soit très fort et/ou d'avoir beaucoup d'observations pour maximiser le signal.

De plus, cette géométrie du système, où la lumière traverse l'atmosphère avec un grand angle par rapport à sa normale (*slant viewing*) pose un désavantage. Selon Fortney (2005), cet angle de vue incliné peut rendre certaines raies d'absorption plus faibles, voir même invisibles. Ceci nous oblige à prendre en considération certains condensats et aérosols mineurs qui, observés d'un angle normal, ont de petites profondeurs optiques, mais qui ne peuvent plus être négligées, lorsqu'observées avec un angle incliné. En comparaison, la spectroscopie en émission peut sonder des couches d'atmosphère plus profondes, puisque la ligne de visée est parallèle à la normale de la surface.

## 1.2. Formation des planètes géantes

La formation des planètes implique un grand nombre de processus complexes et aléatoires qui seraient bien longs à développer en détail ici. C'est pourquoi, dans cette section, nous survolerons très brièvement les mécanismes de formation principaux des planètes géantes.

D'abord, le même principe de base s'applique à tous les systèmes planétaires : un nuage de gaz (principalement composé d'hydrogène moléculaire) et de poussière s'effondre pour former des étoiles et/ou naines brunes. Un disque circumstellaire est inévitablement formé étant donné la conservation du moment cinétique et l'accrétion axiale de matière sur le disque. C'est l'accrétion inefficace de la matière sur l'étoile et la viscosité dans le disque qui permettent la formation

de planète dans ce dernier, d'où son nom « disque protoplanétaire ». Suivant cette séquence d'évènements, on peut donc aussi énoncer que la composition du disque protoplanétaire est la même que celle du nuage qui s'est effondré pour former l'étoile.

Il y a deux scénarios principaux pour la formation de planètes géantes : soit par instabilités gravitationnelles dans le disque ou par accréation de cœur.

### **1.2.1. Instabilités gravitationnelles dans le disque (IG)**

Ce scénario suit une approche « descendante », où une planète se forme par l'effondrement d'un disque instable, lorsque celui-ci est massif et/ou froid (Bodenheimer et al., 1980). Dans ce scénario, les instabilités évoluent en bras spiraux denses (ondes de densités), qui vont ensuite se fragmenter et se contracter pour former directement des planètes géantes (Boss, 1997). Cette méthode implique surtout la formation de planètes géantes gazeuses ou géantes de glaces (possible par la photo-évaporation des enveloppes gazeuses ; Boss et al. 2002). Ce scénario est compatible avec la présence de cœurs solides si on considère la coagulation et la sédimentation subséquente des grains de poussière et de glace. De plus, cette méthode de formation se ferait relativement tôt dans la vie du disque (lorsqu'il est encore massif), et le processus serait relativement rapide (en quelques centaines de milliers d'années seulement ; Boss 2000), comparativement au scénario d'accréation de cœur (section suivante).

Cependant, ce processus est surtout optimal à des distances de l'ordre de 50 à 100 ua de l'étoile (Armitage, 2007), ce qui peut s'expliquer pour plusieurs raisons. De un, à plus grandes distances de l'étoile, la densité du gaz décroît et il devient de plus en plus difficile pour le gaz de s'effondrer sur lui-même (Kratte et al., 2010). De deux, à plus petites distances de l'étoile, le gaz ne refroidit pas assez rapidement pour qu'il y ait formation de bras spiraux ou d'amas liés gravitationnellement (Matzner & Levin, 2005).

### **1.2.2. Accréation de coeur (AC)**

Dans ce scénario « ascendant », une planète géante se forme d'abord par la construction d'un cœur rocheux qui accrete ensuite une énorme quantité de gaz. C'est d'ailleurs le scénario avec lequel la majorité des planètes découvertes par VR semble être le plus cohérente (Matsuo et al., 2007).

Plus précisément, l'AC se fait en plusieurs étapes. D'abord, on aura la formation d'un cœur rocheux qui suivra les mêmes étapes que la formation des planètes rocheuses<sup>11</sup>, mais qui devra atteindre une masse minimale entre 5 à  $20 M_{\oplus}$  pour pouvoir garder son enveloppe gazeuse (Ikoma et al., 2000). Ensuite, il y aura une accréation lente de gaz, avec encore un peu d'accréations de planétésimaux. Après quelques millions d'années, le cœur va atteindre une masse critique qui rendra le taux d'accréation de gaz supérieur à celui des planétésimaux et qui continuera d'augmenter de façon exponentielle (Pollack et al., 1996). C'est alors qu'on entre dans la phase d'accréation de gaz accélérée. Une fois que la masse de l'enveloppe atteint entre quelques pour cent (Perryman, 2011) à  $3 M_{\oplus}$  (Cridland et al., 2020), il devient difficile de garder les planétésimaux qui sont accrés intacts. Ils se font détruire par la friction du gaz et la pression dynamique. Ils pourront ainsi contribuer à la composition chimique de l'enveloppe. Enfin, l'accréation de gaz sera arrêtée lorsqu'un trou dans le disque aura été creusé, ou alors une fois le disque dissipé. Ces planètes géantes se formeront généralement au-delà de la ligne de glace (« snow line »), l'endroit dans le disque où la température est assez basse pour qu'il puisse y avoir la formation de glaces (d'eau ou d'autres composés volatils ; Öberg et al. 2011). Des simulations montrent qu'il n'y a pas assez de matériel solide dans le disque interne pour créer des cœurs suffisamment gros et suffisamment rapidement (Bodenheimer et al., 2000). Cette augmentation de matériel solide, au-delà de la ligne de glace, en plus de la diminution de l'influence gravitationnelle de l'étoile, permet de former plus facilement des planètes géantes à cet endroit. Cette méthode se fait sur un plus long laps de temps, de l'ordre de quelques millions d'années.

Toutefois, cette hypothèse apporte aussi ses incertitudes, surtout par rapport à l'accroissement des planétésimaux. Il existe certaines incohérences lorsqu'on regarde l'échelle de temps nécessaire pour croître un grain de poussière sous-micrométrique jusqu'à des objets de quelques km de large, et qu'on la compare aux temps de formation planétaire observés (qui sont beaucoup plus courts), surtout pour les objets à larges séparations, où la densité de solide diminue (Rafikov, 2011). C'est pourquoi l'AC permet surtout d'expliquer la formation de planètes à l'intérieur de  $\sim 50$  ua (Armitage, 2007). De plus, lorsque les grains atteignent des tailles de l'ordre du cm, il est difficile d'aller plus gros puisque ces grains vont rebondir les uns sur les autres si la vitesse entre eux est  $\gtrsim 0,1$  m/s et se fragmenter si elle est  $\gtrsim 1$  m/s (ex. : Windmark et al., 2012).

---

11. La formation de planètes terrestres peut se décrire très rapidement en 3 phases : (1) La poussière va s'accumuler dans le plan-milieu du disque protoplanétaire ; (2) Des roches ayant des tailles de plus en plus grosses se formeront via des processus de collisions et des instabilités dans le disque, jusqu'à former des objets de quelques km de large, ou plus gros, qu'on appelle les « planétésimaux » ; (3) Ces objets fusionneront (ou en accumuleront davantage) pour former des planètes terrestres, ou encore pour former le cœur de planètes géantes. L'atmosphère des planètes terrestre proviendra soit de l'accréation de gaz primordial ou encore par dégazage de seconde génération. En comparaison et dans tous les cas, l'atmosphère des planètes géantes apparaît par capture directe de gaz dans le disque protoplanétaire (par accréation de gaz et/ou de glaces).

### 1.2.3. Accrétion de cailloux

Pour surpasser les problèmes associés à ces deux méthodes, certaines études ont développé le scénario de formation par accrétion de cailloux (de l'anglais *Pebble accretion*; Lambrechts & Johansen 2012). Ce principe est basé sur le fait que la formation de cailloux par coagulation jusqu'à des tailles de l'ordre du mm au cm est très rapide : elle peut se faire en quelques milliers ou dizaines de milliers d'années, à 1 ua ou plus loin dans le disque, respectivement (ex. : Weidenschilling, 1980; Nakagawa et al., 1981; Rodmann et al., 2006; Öberg & Bergin, 2016, et autres références dans celui-ci). Ces cailloux sont couplés au gaz (par friction; comparative-ment aux plus grosses roches qui sont découplées du gaz), et leur capture par les protoplanètes est faite dans tout le rayon de Hill (la région jusqu'où un objet a une influence gravitationnelle; ex. : Ormel et al. 2010). Comme ces cailloux sont faciles et rapides à produire, ils sont présents en grande quantité dans le plan milieu du disque, mais leur grand taux d'accrétion découle de la friction du gaz sur ceux-ci (Lambrechts & Johansen, 2012). En effet, c'est un élément clé dans la croissance des coeurs puisque la friction change le processus de diffusion des cailloux lorsqu'ils passent près d'un coeur, et favorise l'accrétion (Johansen & Lambrechts, 2017). L'accrétion de cailloux peut même aider les processus de formation par instabilités gravitationnelles. La turbulence du gaz dans le disque, et la formation de zones de densité aideront à concentrer les cailloux en filaments denses, qui s'effondreront sous l'effet d'autogravitation (Rafikov, 2004; Youdin & Johansen, 2007; Johansen & Lambrechts, 2017). En somme, il est beaucoup plus facile pour les planétésimaux existants d'accréter de petits objets (cailloux) que d'autres planétésimaux. Le temps pour former un coeur de planète géante (ou encore une planète terrestre) est diminué d'un facteur 30, à 5 ua, et d'un facteur 100, à 50 ua, comparativement au processus d'accrétion sur coeur avec des planétésimaux. Ce processus supporte fortement la formation de planètes géantes gazeuses via accrétion sur coeur avec cailloux, puisque leur coeur peut être formé avant que le disque protoplanétaire se dissipe après 1–10 Ma. De plus, la présence de ces cailloux est supportée par les observations de jeunes disques protoplanétaires (ex. : Rodmann et al., 2006).

Un problème réapparaît si l'on considère la « dérive de cailloux » (de l'anglais *Pebble drift*; ex. : Andrews & Birnstiel, 2018). Le gaz va subir une force vers l'extérieur à cause du gradient de pression radiale présent dans les disques « lisses » (la densité de gaz décroît de façon monotone avec la distance dans le disque). Ceci aura pour effet de « ralentir » le gaz, et de le faire orbiter à des vitesses sous-képlériennes. Pour la classe d'objets ayant la bonne taille pour être couplé au gaz (les cailloux), cela créera une force de friction qui va « voler » du moment cinétique aux particules et les envoyer spiraler vers l'intérieur du disque. À un certain point, la vitesse à laquelle les cailloux sont envoyés vers l'intérieur devient plus grande que la vitesse de collision des cailloux, ce qui prévient leur coagulation pour former des objets plus gros. On peut cependant contourner ce problème en introduisant des structures dans le disque, c.-à-d. des

anneaux (Pinilla et al., 2012a,b). Ils ont été observés dans plusieurs systèmes grâce au *Atacama Large Millimeter/submillimeter Array* (ALMA), tel que HL Tau et TW Hydrae (ALMA Partnership et al., 2015; Andrews et al., 2016, respectivement). Les trous dans le disque (creusés par les objets les plus gros) produiront des excès de pression à l’extérieur de leur orbite qui pousseront les cailloux à s’accumuler et s’agglomérer dans le maximum de pression et produiront des anneaux. La dérive de cailloux est ainsi freinée, ce qui empêche d’épuiser la réserve (Andrews et al., 2016), et permet la formation de planètes.

#### 1.2.4. Migration

La migration de planètes permet de déplacer les planètes potentiellement loin de l’emplacement où elles ont été formées. C’est ce qui expliquerait, entre autres, la position des Jupiter chaudes et autres planètes se trouvant à  $\lesssim 0.2ua$  de leur étoile. Une formation *in-situ* serait très improbable dû au manque de matériel solide très près de l’étoile pour former un cœur assez massif. Il existe trois grandes catégories de migration, qui impliquent toutes des interactions entre la planète et le disque protoplanétaire, un/des compagnons stellaires (ex. : systèmes binaires) ou d’autres planètes (d’un système multiple).

Il y a trois types de migration dans le disque, dépendamment de la masse de la planète et si celle-ci creuse un trou dans le disque (type II; pour les planètes massives), creuse partiellement un trou (type III; pour les sous-géantes) ou n’en creuse pas du tout (type I; pour les planètes de plus faibles masses). La présence de gaz et de force de marées permettrait aussi de circulariser les orbites des planètes. La migration des Jupiter chaudes vers l’étoile doit cependant être arrêtée avant que celle-ci ne soit détruite dans l’étoile. Ce freinage, encore mal compris, serait fait soit à l’aide de vents stellaires, par la troncature du disque par photoévaporation (Trilling et al., 2002), ou par d’autres effets d’instabilités et/ou influences magnétiques. La migration avec disque pourra davantage influencer la composition de la planète puisqu’il y aura possibilité d’accréter plus de matières à différents endroits dans le disque, que ce soit sous forme de gaz, de glace ou de roche.

De leur côté, les interactions entre planètes peuvent mener à des changements de configuration orbitale. Par exemple, l’interaction gravitationnelle entre deux (ou plusieurs) planètes ayant des orbites rapprochées pourrait rapprocher l’orbite d’une planète au détriment de l’autre qui se fera éjecter<sup>12</sup>. Les orbites peuvent également entrer en résonance, ou encore être affectées par des oscillations de Kozai-Lidov (Kozai, 1962; Lidov, 1962)<sup>13</sup>.

---

12. L’influence d’une étoile compagnon, ou le passage d’une étoile dans une jeune région dense –“*stellar fly-by*”– peuvent également venir perturber les systèmes.

13. Un système à deux corps peut se faire perturber par un troisième objet distant, causant une libration de l’argument du périastre  $\omega$  (voir Figure 1.1), et conduisant à un échange périodique entre les valeurs d’excentricité ( $e$ ) et d’inclinaison ( $i$ ) de l’orbite. Cet effet peut mener des orbites quasi circulaires à devenir très excentriques et même à renverser des orbites modérément inclinées à osciller entre prograde et rétrograde.

## 1.3. Physique des atmosphères

Comme pour la diversité des systèmes planétaires, on s'attend à voir des atmosphères qui sont toutes autant variées. Un des buts principaux de la caractérisation des atmosphères d'exoplanètes est de mieux comprendre et identifier les éléments clés des mécanismes de leur formation primordiale.

Nous sommes à l'ère où nous pouvons déterminer une panoplie de paramètres dans l'atmosphère des exoplanètes. Avec ces informations, on voudra, d'un côté, valider les modèles de formation planétaires en se basant sur nos observations, et d'un autre, établir quels sont les meilleurs scénarios de formation et d'évolution pour les planètes observées grâce à ces mêmes modèles.

Dans cette section, les éléments théoriques de base sur l'atmosphère des exoplanètes seront présentés (avec un accent sur les planètes géantes gazeuses). D'abord, les notions sur le transfert radiatif et sur la balance énergétique atmosphérique seront expliquées. Ensuite, les « observables » des atmosphères seront présentés, dans l'ordre : le profil vertical, la composition chimique et la dynamique.

### 1.3.1. Transfert radiatif

Toute l'information que l'on peut espérer avoir d'une exoplanète proviendra de la lumière qu'elle émet, ou encore de celle qui provient de l'étoile et qui aura été réfléchiée ou qui aura traversé son atmosphère. Dans tous les cas, c'est l'interaction entre la lumière et la matière qui nous informera des conditions dans l'atmosphère de la planète. Il est donc important de bien comprendre les processus qui influenceront ce qu'on peut observer.

Le transfert radiatif permet d'établir comment la radiation est affectée lorsqu'elle passe dans un médium, le médium en question ici est l'atmosphère de la planète. Plusieurs processus d'absorption et d'émission de photons par les particules dans le gaz influenceront la radiation se propageant dans l'atmosphère, ce qui se traduit par la formation de raies (en absorption ou en émission) dans le spectre émergent. Beaucoup d'informations peuvent être tirées de ce spectre, comme la température, la pression et la composition. Le changement d'intensité  $dI_\lambda$  d'un faisceau de lumière  $I_\lambda$ , qui est fonction de  $\lambda$ , traversant un volume de gaz est donné par l'équation de transfert radiatif :

$$dI_\lambda = -\kappa_\lambda \rho I_\lambda ds + j_\lambda \rho ds, \quad (1.3.1)$$

où  $\kappa_\lambda$  est le coefficient d'absorption (ou l'opacité),  $j_\lambda$  est le coefficient d'émission,  $\rho$  est la densité du gaz, et  $ds$  est la distance traversée. Le premier terme au membre de droite représente ce qui est perdu et le deuxième représente ce qui est ajouté au faisceau lumineux. On définit la profondeur optique comme  $d\tau_\lambda = -\kappa_\lambda \rho ds$  et  $\tau_\lambda = \int_S \kappa_\lambda \rho ds$ , où  $S$  est l'étendue du trajet optique

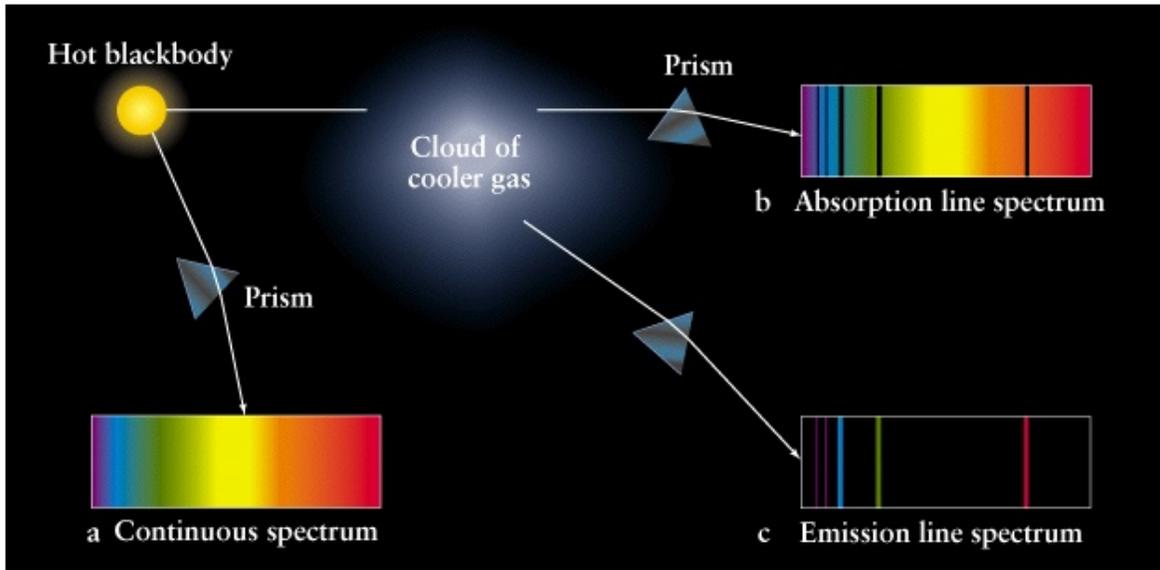


FIGURE 1.12 (A) Une source à large spectre isotherme observée avec un élément de diffraction donnera un spectre continu. (B) un gaz froid placé devant la source donnera un spectre d'absorption. (C) Un gaz chaud donnera un spectre en émission. Une combinaison de ces situations est possible, et donnera des spectres « hybrides ». Figure tirée du cours de Dmitri Pogosian, à l'Université d'Alberta.

au travers du médium. Les quantités  $\kappa_\lambda$  et  $\rho$  dépendent des conditions locales du gaz, soit la composition, la température  $T$  et la pression  $P$ . En utilisant la définition de  $d\tau_\lambda$ , on peut réécrire cette équation sous sa forme plus générale :

$$\frac{dI_\lambda}{d\tau_\lambda} = I_\lambda - S_\lambda, \quad (1.3.2)$$

où  $S_\lambda = j_\lambda/\kappa_\lambda$  est la fonction source.

Dans une atmosphère complètement isotherme, aucune raie n'est formée :  $I_\lambda = S_\lambda$ , alors l'intensité ne change pas ( $dI_\lambda/d\tau_\lambda = 0$ ; voir Figure 1.12A). Des raies en absorption seront formées lorsque la lumière d'une source à large spectre traverse un médium plus froid qui absorbe ladite lumière et la ré-émet dans des directions aléatoires ( $I_\lambda < S_\lambda$ ; Figure 1.12B). On peut l'observer par spectroscopie de l'émission planétaire lorsque la température de l'atmosphère diminue avec l'altitude. En comparaison, des raies en émission seront formées lorsque la lumière de la source à large spectre traverse un médium plus chaud : l'intensité sera plus grande à cause de l'émission du médium ( $I_\lambda > S_\lambda$ ; Figure 1.12C). On peut alors l'observer, toujours avec la spectroscopie d'émission, lorsque la température de l'atmosphère augmente avec l'altitude (inversion de température). En spectroscopie de transit, le terme d'émission (thermique de la planète) est complètement dominé par la radiation d'étoile. On peut ignorer le second terme de l'éq. 1.3.2 et résoudre pour l'intensité de l'étoile qui sera absorbée par l'atmosphère planétaire. On obtient :  $I_\lambda = I_{\lambda,0}e^{-\tau_\lambda}$ , qui sont des signaux en absorption.

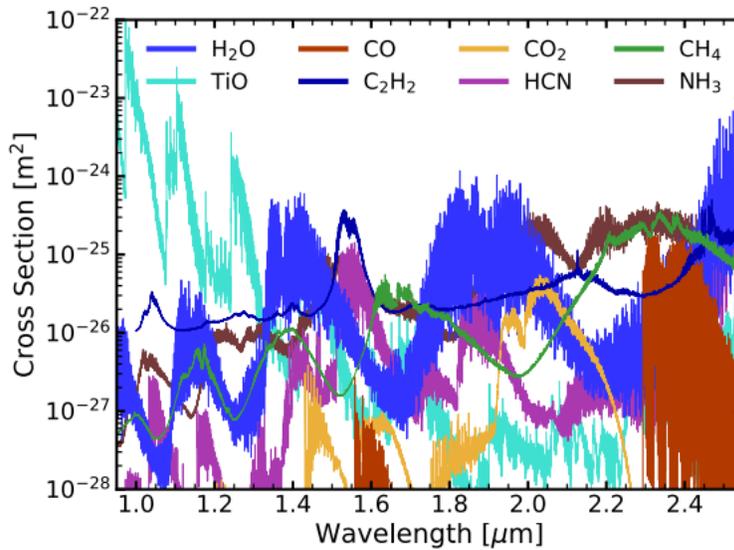


FIGURE 1.13 Spectres des sections efficaces de différentes molécules sur la plage de longueur d'onde de SPIRou. Figure tirée de Pelletier et al. (2021).

Chaque élément présent dans le gaz laissera son empreinte avec des raies spectrales uniques. C'est justement par leur nature unique qu'on peut déterminer avec plus ou moins de certitude la composition du gaz. Plus précisément, ces raies spectrales proviennent soit de transitions électroniques des électrons des atomes, ou encore de la rotation et/ou de la vibration des molécules. Cependant, certaines raies ou bandes d'absorption<sup>14</sup> peuvent se superposer et rendre l'identification plus ardue. C'est pourquoi avoir des spectres à haut signal-sur-bruit (S/N), avec le plus de détails et sur une plage de longueur d'onde la plus large possible est préférable, puisqu'il y aura moins de chance de confusion. Il est à noter que ce ne sont pas nécessairement les gaz les plus abondants qui vont générer le plus de caractéristiques spectrales dans une atmosphère. Ce sont plutôt ceux qui ont une grande opacité, c.-à-d. une grande section efficace d'absorption ( $\sigma_\lambda$ ; relié à l'opacité via  $\kappa_{\lambda\rho} = n\sigma_\lambda$ , où  $n$  est la densité en nombre; voir Figure 1.13), qui vont laisser une plus grande empreinte. Par exemple, pour les planètes géantes gazeuses ayant des atmosphères dominées par l'hydrogène et l'hélium, on verra l'empreinte des molécules ayant une grande opacité malgré leur plus faible abondance (comme le H<sub>2</sub>O, CO, CO<sub>2</sub>, etc.), puisque l'hydrogène et l'hélium sont plutôt inertes (ne montrent pas beaucoup de raies d'absorption (ou émission) à ces températures, excepté pour l'absorption induite par collision; AIC).

14. On appelle une bande d'absorption une forêt de raies d'absorption qui sont assez proches pour creuser une bande dans un spectre, si la résolution du spectre n'est pas assez grande pour résoudre les détails de toutes ces raies.

### 1.3.2. Balance énergétique

L'énergie apportée dans l'atmosphère de la planète provient de l'étoile hôte et de l'intérieur même de la planète. Pour compenser cette énergie injectée<sup>15</sup>, l'atmosphère va s'ajuster de façon à émettre une quantité équivalente : on parlera alors de la « balance énergétique ».

On peut écrire le bilan énergétique sous la forme suivante :  $E_{em}(t) = (1 - A_B)E_{inc}(t) + E_{int}(t)$ , où  $E_{em}(t)$  est l'énergie émise par la planète,  $E_{inc}(t)$  est l'énergie incidente,  $E_{int}(t)$  est l'énergie interne et  $A_B$  est l'albédo de Bond. L'*albédo* est une mesure de la réflectivité de la surface d'une planète (et/ou de son atmosphère ; de ses constituants)<sup>16</sup>. En explicitant les termes, on obtient :

$$\frac{4\pi}{f} R_P^2 F_P = (1 - A_B) F_* \left( \frac{R_*}{d} \right)^2 \pi R_P^2 + L_{p,int}, \quad (1.3.3)$$

où  $F_P$  et  $F_*$  sont les flux émis par la planète et l'étoile, respectivement,  $L_{p,int}$  est la luminosité interne de la planète,  $d$  est la distance moyenne de la planète sur son orbite (le rayon, si circulaire) et  $f$  est un facteur qui tient en compte la fraction de la surface de la planète (l'angle solide) sur laquelle le flux est réémis par rapport à  $4\pi$ .

Grâce à cet équilibre, on peut faire une estimation approximative de la température de la planète, qu'on nommera la température d'équilibre ( $T_{eq}$ ). La température d'équilibre est la température effective d'une planète si on ne considère que l'énergie incidente provenant de son étoile (on ignore l'énergie interne). En partant de l'équation 1.3.3, en ignorant le dernier terme, et en utilisant la loi de Stefan-Boltzmann ( $F = \sigma_B T_{eff}^4$ ; où  $\sigma_B$  est la constante de Stefan-Boltzmann, et  $T_{eff}$  est la température efficace d'un objet) pour décrire les flux  $F_*$  et  $F_P$  de l'étoile et la planète, respectivement, on obtient :

$$T_{eq,p} = T_{eff,*} \left( \frac{R_*}{2d} \right)^{1/2} [f(1 - A_B)]^{1/4}. \quad (1.3.4)$$

Par exemple, pour la Terre ( $d = 1 \text{ ua}$ )  $T_{eq,\oplus} = 254 \text{ K}$ , tandis que pour Jupiter ( $d \sim 5.2 \text{ ua}$ )  $T_{eq,J} = 110 \text{ K}$ , les deux ayant  $A_B \approx 0.3$  et en assumant une redistribution complète ( $f = 1$ ).

Cependant, l'atmosphère n'est pas nécessairement à cette température, et celle-ci n'est utile qu'en tant que première approximation. Plutôt, selon l'advection et les vents, l'efficacité de recirculation de la chaleur, l'endroit où cette énergie est déposée, les nuages, etc., les atmosphères d'exoplanète auront une structure en température plus complexe.

15. Ces quantités d'énergie sont « stables » sur des temps très longs, mais sont susceptibles de varier, surtout en début et fin de vie de l'étoile, ou via d'autres perturbations du système.

16. Plus précisément, l'albédo de Bond,  $A_B$ , représente la fraction de l'énergie incidente (provenant principalement de l'étoile) qui est réfléchiée dans l'espace, et donc perdue, par la planète. Ceci inclut toute la radiation réfléchiée dans toutes les directions et pour toutes les longueurs d'onde.

### 1.3.3. Profil vertical

Le profil vertical fait référence au profil de température en fonction de la pression (ou de l'altitude ; profil TP). Il est intimement lié à la conservation d'énergie et au mécanisme de transport d'énergie. Ceux-ci nous permettent donc de mieux connaître les processus radiatifs présents dans les atmosphères ainsi que leurs liens avec les processus chimiques et dynamiques. Le profil TP est relié au spectre de la planète via les équations suivantes : l'équilibre hydrostatique, l'équilibre radiatif et l'équation de transfert radiatif (éq. 1.3.1). L'équation pour l'équilibre hydrostatique est donnée par :

$$\frac{dP}{dz} = -\rho g, \quad (1.3.5)$$

où  $\rho$  est la densité du gaz et  $g$  l'accélération gravitationnelle (la variation de  $g$  avec  $z$  est négligée). Cette équation nous donne la variation de la pression avec l'altitude, qui est un équilibre entre la gravité qui tente de compresser l'atmosphère et le gradient de pression du gaz qui l'en empêche. Pour un gaz parfait (à température constante  $T$ ), on obtient  $P = P_0 e^{-z/H}$ , où on définit  $H = kT/\mu g$  comme étant la *hauteur d'échelle*,  $\mu$  étant le poids moléculaire moyen et  $k$  la constante de Boltzmann. Ceci nous dit que la pression d'une atmosphère décroît exponentiellement avec l'altitude. On peut aussi la voir comme la distance sur laquelle la pression diminue d'un facteur  $1/e$ .

L'équilibre radiatif est une formulation de la conservation d'énergie dans l'atmosphère et exprime que le flux radiatif total dans chaque couche doit être conservé puisqu'il n'y a ni création ni destruction d'énergie dans l'atmosphère ( $dF(\tau)/d\tau = 0$ ). Résoudre ces trois équations simultanément nous permet de trouver nos trois inconnus, soit le champ de radiation  $I_\lambda(\tau)$ , la température  $T(\tau)$  et la pression  $P(\tau)$ .

La structure thermique va dépendre du mécanisme d'énergie dominant, soit la radiation ou la convection (la conduction étant souvent négligeable). Pour savoir s'il y aura convection, on compare les gradients de température radiatif (de l'atmosphère) et adiabatique. La condition de stabilité pas rapport à la convection est donnée par :

$$\left(\frac{dT}{dz}\right)_{\text{ad}} > \left(\frac{dT}{dz}\right)_{\text{rad}}. \quad (1.3.6)$$

Si le gradient radiatif est plus grand que l'adiabat, la convection se déclenche et le profil de température est donné par le gradient thermique adiabatique  $\Gamma$  :

$$\Gamma = -\frac{dT}{dz} = \frac{g}{c_p}, \quad (1.3.7)$$

où  $c_p$  est la capacité thermique à pression constante. Dans le cas contraire, le transport d'énergie va se faire par radiation, les opacités vont gouverner le profil de température et il sera donné par

l'équilibre radiatif. La solution dépendra alors des conditions frontières d'irradiation (chauffée par le bas, ou par le haut et le bas, lorsqu'irradiée par l'étoile hôte, et où l'énergie est absorbée) et de la quantité d'absorbeurs présents.

Qualitativement, un profil TP peut être décrit comme suit : très haut dans l'atmosphère ( $P \lesssim 10^{-5}$ ), la profondeur optique est assez faible à toutes les longueurs d'onde pour laisser passer librement la radiation entrante et sortante, ce qui ne génère pas de structures spectrales. C'est un peu plus bas, où les couches d'atmosphère sont encore optiquement minces ( $\tau < 1$ ), que les raies spectrales sont principalement formées. Dépendant du niveau d'irradiation de l'étoile hôte, il peut y avoir une inversion de température : la présence de gaz et d'aérosols fortement absorbeurs va réchauffer le gaz ambiant et on aura alors des raies en émission. Comme elles sont fortement irradiées par leur étoile, on s'attend à ce que quelques JCs ou ultra-chaudes (JUCs;  $> 1600\text{--}2200\text{ K}$ ; Baxter et al. 2020, Parmentier et al. 2018) aient des gradients de température proche de l'isotherme et/ou aient des inversions de température. Profondément dans l'atmosphère ( $P \gtrsim 1\text{--}10\text{ bar}$ ), on devient optiquement épais ( $\tau \gtrsim 1$ ). En assumant l'équilibre thermodynamique local (ETL; qui nous dit que la radiation agit comme un corps noir, où  $I_\lambda = S_\lambda = B_\lambda(T)$ ), on a alors un profil isotherme (aucune raie ne sera formée). Si on va encore plus profondément, le transport d'énergie par radiation n'est plus efficace, la convection démarre et le profil de température est dicté par l'équation 1.3.7. Différents types de profils sont montrés sur la Figure 1.14.

Au final, le profil vertical nous importe puisqu'il nous permet de déterminer, via des méthodes numériques, les abondances des éléments chimiques dans l'atmosphère. La structure en température va influencer la densité de chaque composante (atomique et moléculaire; dans chaque niveau d'excitation et état d'ionisation dans le gaz et/ou de rotation et vibration – via les équations de Boltzmann et de Saha, respectivement). On peut ainsi calculer leur opacité dans chaque couche de l'atmosphère et résoudre l'équation de transfert radiatif. C'est donc le profil TP et les abondances qui vont déterminer la structure du spectre observé. Ultimement, cela nous permettra de déterminer la température de surface d'une planète rocheuse, nous permettant d'établir des critères d'habitabilité.

Les mesures de profil TP sont principalement obtenues via des spectres d'émission thermique du côté jour (ex. : Kreidberg et al., 2014; Stevenson et al., 2014; Line et al., 2014; Mikal-Evans et al., 2020; Mansfield et al., 2021). Le spectre en émission sonde directement la température de luminosité de l'atmosphère en fonction de la longueur d'onde qui peut se traduire par différentes profondeurs (selon où se trouve la photosphère pour chaque longueur d'onde; Madhusudhan 2019; tel que vu à la Figure 1.7, et Section 1.1.3.1). Ce sont ensuite le gradient de température du profil et la composition chimique qui détermineront la forme et l'amplitude des structures observées dans le spectre.

La majorité des observations ont montré des atmosphères qui n'ont pas d'inversion de températures. À ce jour, on connaît encore peu de planètes montrant une inversion de température

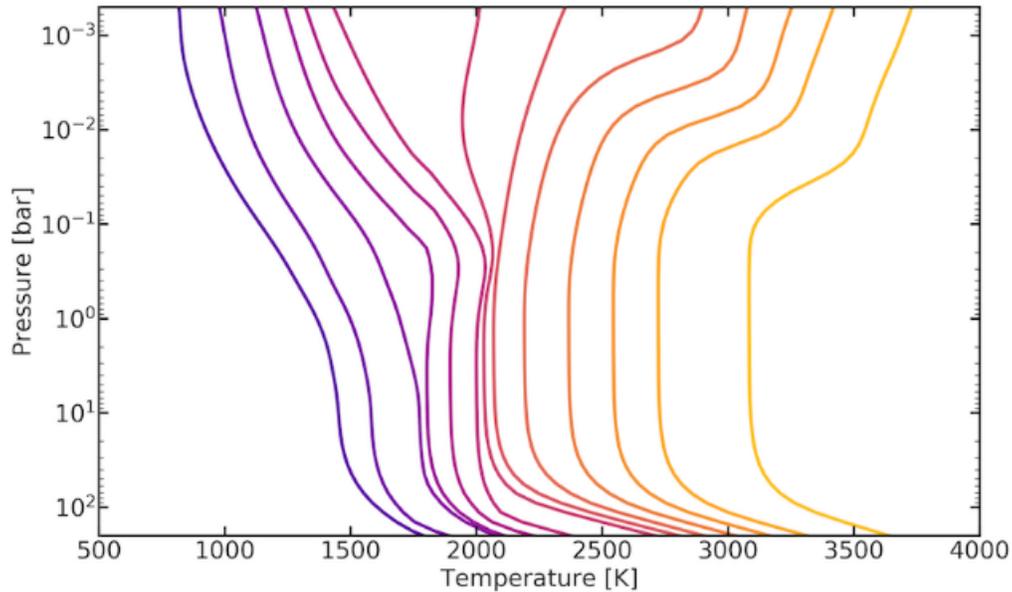


FIGURE 1.14 Profils TP pour des modèles d’atmosphère 1D, avec  $T_{\text{eq}}$  qui augmente (de 500 à 3600 K, avec des sauts de 50 ou 200 K). On voit que les modèles les plus froids (indigo, à gauche) diminuent continuellement, tandis que les plus chauds montrent des zones isothermes et/ou des inversions de températures (vers le jaune, à droite). De plus, en profondeur, près de  $10^2$  bar, on peut voir le profil adiabatique où la convection est enclenchée. Tiré de Mansfield et al. (2021).

dans leur atmosphère du côté jour. Parmi celles-ci, on a les JUCs WASP-18 b (Sheppard et al., 2017; Arcangeli et al., 2018), WASP-121 b (Evans et al., 2017), WASP-33 b (Haynes et al., 2015, voir Figure 1.8), WASP-76 b (Wardenier et al., 2021), et WASP-12 b (Madhusudhan et al., 2011), ainsi que la JC HAT-P-32 b (Zhao et al., 2014). Les autres JUCs connues ne montrent pas de signe d’inversion (Beatty et al., 2017; Parmentier et al., 2018). Comme ces planètes sont fortement irradiées, on explique cela par un manque d’absorbeurs dans l’UV et le visible, qui causeraient ces inversions. Plus récemment, Mansfield et al. (2021) ont présenté une étude comparative entre les spectres d’émission des JCs et JUCs, et ont montré une tendance entre la force de la bande d’absorption de l’eau à  $1.4 \mu\text{m}$  et la température. Cela aide à expliquer l’absence de signal en émission des plus chaudes JUCs, qui auront plutôt des spectres de quasi-corps noir, à cause de la dissociation moléculaire et l’apparition de l’opacité du continu du  $\text{H}^-$ .

Les contraintes sur la température provenant des spectres en transmission sont également possibles et seront complémentaires à celles faites en émission. On peut penser à la détermination de la température par la pente de Rayleigh (ex. : Lecavelier Des Etangs et al., 2008a), la forme des raies de Na (ex. : Sing et al., 2008a), ou du spectre de transmission de la planète, mais il est tout de même important de tenir compte de certains effets. D’abord, de par sa géométrie, le signal en spectroscopie de transmission provient des couches plus hautes de l’atmosphère, là où

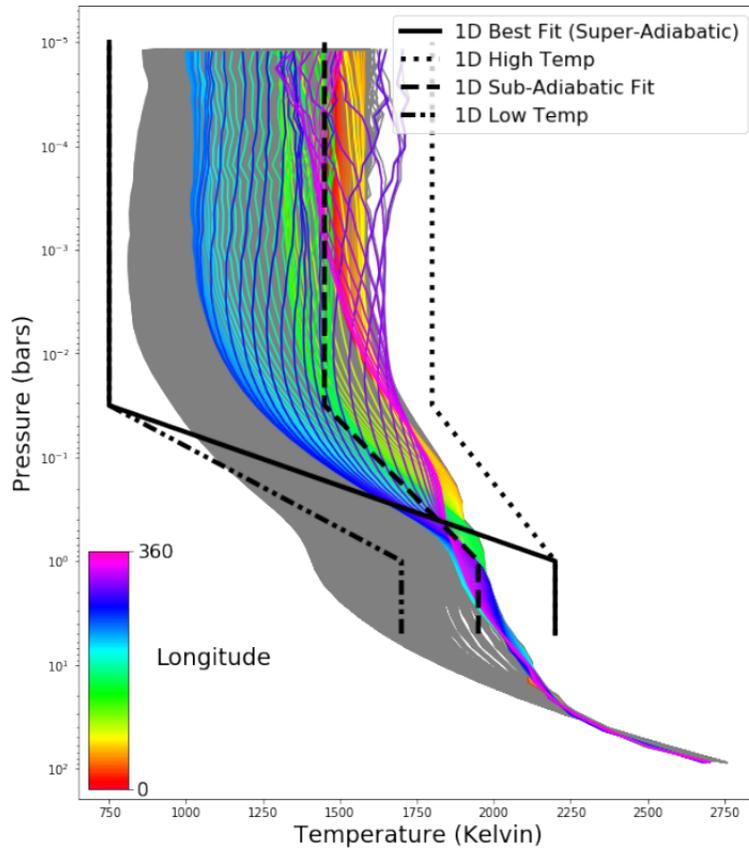


FIGURE 1.15 Profils de température d'un modèle 3D de HD209458b en rotation synchrone. Les profils arc-en-ciel sont les profils à l'équateur, colorés en fonction de leur longitude, tandis que les gris sont pour toute la planète. Les modèles en noir sont les modèles 1D simples qui représentent le mieux l'entièreté de la planète. Le meilleur modèle 1D (trait plein) n'est même pas physiquement possible (gradient de température super-adiabatique). Tiré de Beltz et al. (2021), Figure 5.

les gradients de température sont souvent faibles (voir nuls). Ensuite, tel que mentionné brièvement à la Section 1.1.3.3, le signal en transmission est moyenné sur l'ensemble du terminateur. C'est d'ailleurs là où on retrouve les plus grands contrastes de températures et d'abondances (par la nature même du terminateur, qui délimite le côté jour du côté nuit). Donc un rayon typique passera une moitié de son trajet dans le côté jour et la seconde moitié dans le côté nuit. À cela s'ajoutent les grandes différences entre les deux « hémisphères » du terminateur, soit le côté matin et le côté soir. À ce propos, certaines études utilisant des données de faible dispersion ont montré que la température mesurée par spectroscopie de transit était systématiquement plus petite que  $T_{eq}$  (MacDonald et al., 2020). Ils ont démontré que ceci provenait de l'utilisation de modèles d'atmosphères 1D (en une dimension) pour représenter des terminateurs ayant de gros gradients de composition entre le côté matin et le côté soir, plutôt que des modèles 3D. Ceci peut aussi affecter les abondances retrouvées, qui seront surestimées par potentiellement plus

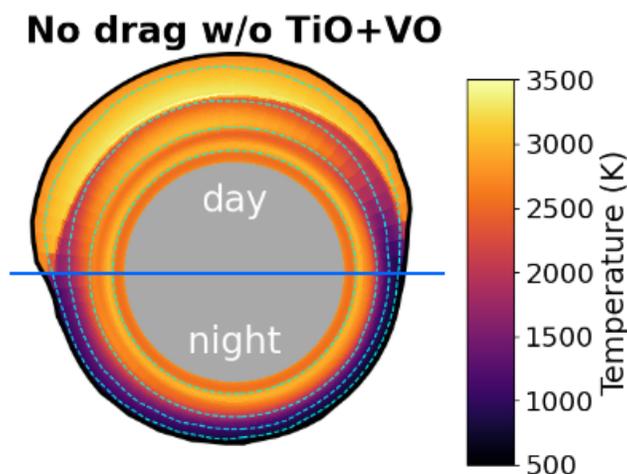


FIGURE 1.16 Carte de températures équatoriales (vue du pôle) d'un GCM atmosphérique de WASP-76b. Les lignes en tiret sont les isobars de  $10^1$ ,  $10^{-1}$ ,  $10^{-3}$ , et  $10^{-5}$  bar. Sur cette figure, on voit bien la différence entre le côté jour (haut) et côté nuit (bas). On peut même voir le déplacement du point chaud vers l'est, qui amplifie les asymétries (de températures, et/ou de compositions) présentes au terminateur (ligne bleue horizontale). Tirée de Wardenier et al. (2021).

d'un ordre de grandeur. Tous ces facteurs rendent une détermination précise du profil TP plus complexe.

C'est d'ailleurs pourquoi, en plus de simplifier les modèles déjà complexes, certaines études de spectroscopie de transit de JCs utilisaient des profils TP isothermes. Ceci était justifié par le fait qu'en spectroscopie de transit, une grosse partie du signal provient des régions  $1 \gtrsim P \gtrsim 10^{-5}$  bar, là où le profil TP peut avoir une forme isotherme, sans gros gradient de température. Cependant, Rocchetto et al. (2016) ont montré que cela pouvait mener à une forte surestimation des abondances retrouvées. Il est donc recommandé d'utiliser au minimum des profils TP non isothermes. De plus en plus, les études montrent qu'il est encore plus avantageux d'avoir des modèles de circulation globale (GCM, de l'anglais *Global Circulation Model*) qui incluent l'information et l'évolution d'un système atmosphérique en 3D (Caldas et al., 2019; Pluriel et al., 2020; Falco et al., 2022; Pluriel et al., 2022). Ceci implique que pour chaque point de la surface de la planète, on peut générer un profil TP, et les combiner pour obtenir un spectre émergent plus représentatif de la réalité. Un exemple d'ensemble de profils TP pour un modèle 3D est présenté à la Figure 1.15. L'efficacité de l'utilisation de modèles 3D a été prouvée par plusieurs (ex. : Zhang et al., 2017; Flowers et al., 2019; Beltz et al., 2021; Wardenier et al., 2021). Ces modèles peuvent adéquatement traiter des différences d'abondance, de température, de vitesse de vents, etc., présents au terminateur, et partout ailleurs sur la planète (voir Figure 1.16). Cependant, le principal désavantage de ces modèles 3D est qu'ils sont beaucoup plus coûteux en temps de calcul à produire. Pluriel et al. (2022) ont démontré que sous certaines conditions,

l'utilisation de modèles 1D (ou 2D) pouvait être satisfaisante, surtout pour des  $T_{\text{eq}}$  plus basses que 1400 et 2000 K, selon s'il y a présence de source d'opacité dans le visible ou non, respectivement. Néanmoins, de plus en plus d'efforts seront mis à produire des modèles 3D le plus rapidement possible, et avec le moins d'approximation possible (ex. : Nixon & Madhusudhan, 2022).

### 1.3.4. Composition

Il suffit de regarder des images des planètes du SS pour deviner qu'elles possèdent des atmosphères très différentes les unes des autres, juste par leurs couleurs ! Bien qu'elles proviennent du même nuage moléculaire, elles ont chacune leur composition bien propre, et on peut déjà dire qu'elles se sont formées et/ou ont évoluées différemment. Il en est de même pour les exoplanètes, où une grande variété de compositions peut être observée. Nous explorerons donc les différentes manières de façonner la composition d'une planète.

1.3.4.1. Hypothèses de base. La composition de l'atmosphère, de même que l'intérieur des planètes, est d'abord et avant tout déterminée par la composition primordiale de la nébuleuse stellaire. Les principaux matériaux disponibles pour former une planète et leur atmosphère sont du gaz et des solides. Le gaz est principalement composé d'hydrogène – H – et d'hélium – He (ex. : Fortney et al. 2007; Seager et al. 2007) et des quantités traces d'éléments plus lourds (principalement dominés par les éléments contenant du carbone – C, de l'azote – N, et/ou de l'oxygène – O; Burrows & Sharp 1999). Pour les solides, ce sera plutôt des éléments volatils sous forme de glaces (lorsque la température ambiante est sous la température de condensation) et de poussières (éléments réfractaires, cailloux, planétésimaux, etc.). La métallicité d'une planète sera donc déterminée par le rapport entre la quantité de gaz et de solides qui aura été capturée.

Ensuite, pour déterminer la composition atmosphérique, on se base sur l'équilibre chimique (Burrows & Sharp, 1999; Lodders & Fegley, 2002), cette condition étant facilement atteinte lorsque les réactions peuvent se faire rapidement (conditions de hautes températures et pression : en profondeur). Il est aussi habituel et simple d'assumer une composition solaire (Asplund et al., 2009). À l'équilibre, cela implique que O est l'élément le plus abondant, après H et He, suivis par C, ce qui donne un C/O de  $\sim 0.5$ . De plus, H est principalement sous forme de dihydrogène ( $\text{H}_2$ ); C est soit sous forme de CO (à grande température :  $T \geq 1300$  K; Madhusudhan et al. 2014) ou  $\text{CH}_4$  (autrement) et en quantité moins importante sous forme de  $\text{CO}_2$  (dépendamment de la température); N est soit sous forme de  $\text{N}_2$  ou  $\text{NH}_3$ ; l'eau ( $\text{H}_2\text{O}$ ) est le porteur dominant d'O à toutes les températures. On retrouve également du  $\text{C}_2\text{H}_2$ , du HCN et autres hydrocarbures en quantité considérable dépendant de la température (Lodders & Fegley, 2002). Un autre type de molécule qui est présent dans certaines JCs (surtout chez les ultras chaudes) sont les hybrides

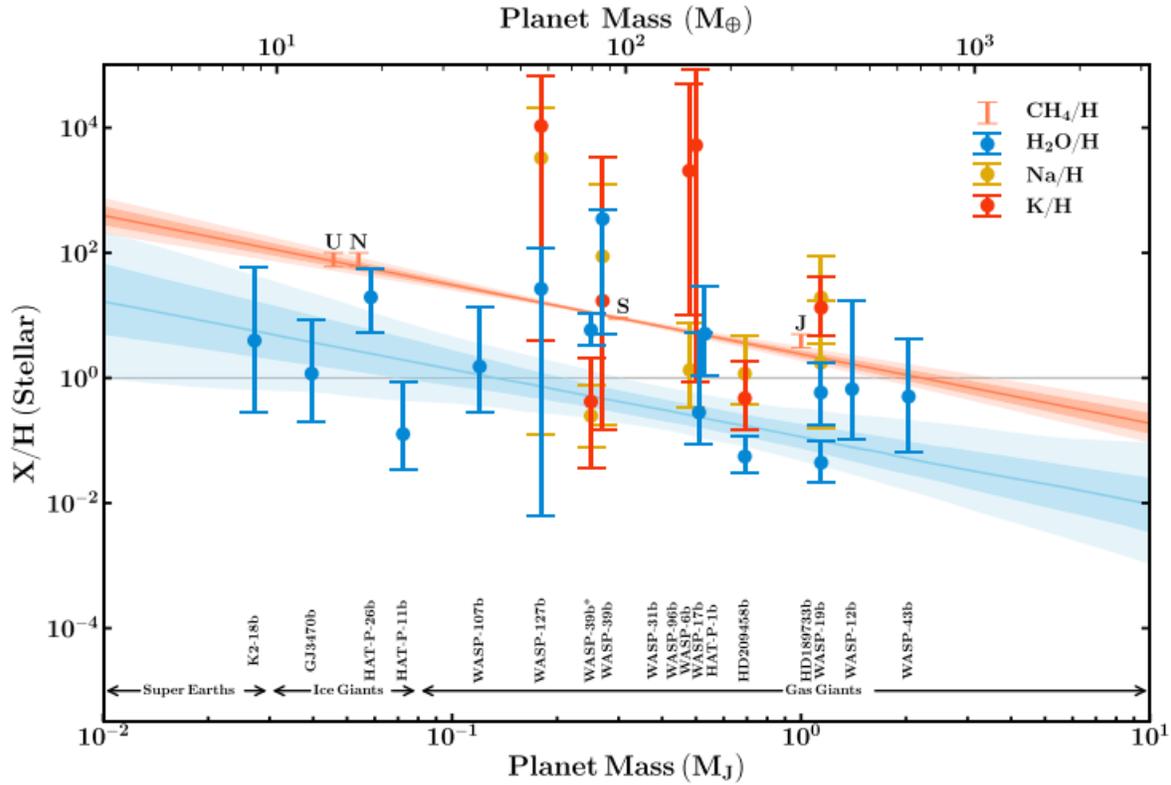


FIGURE 1.17 Relation Masse-Métallicité pour différentes planètes géantes à partir de leur abondance (relative à leur étoile hôte) de H<sub>2</sub>O (bleu), de Na (jaune) et/ou de K (orange). La relation pour les planètes du SS est basée sur leur abondance de CH<sub>4</sub> (corail). On voit que les abondances de H<sub>2</sub>O sont généralement sous la courbe du SS, tandis qu'elles sont généralement supérieures pour le Na et le K. Tiré de Welbanks et al. (2019), Figure 11.

métalliques, tels que le FeH et le TiH, ou encore les oxydes métalliques, tels que le TiO, le VO et le AlO, les absorbeurs de haute altitude qu'on croit à l'origine des inversions de température (ex. : MacDonald & Madhusudhan, 2019).

En fait, le modèle d'AC nous dit que moins la planète est massive (moins de gaz et donc plus de roches en proportion), plus sa métallicité (son abondance en éléments plus lourds, autres que H et He) sera grande, et vice versa. Alors, le rapport de métallicité entre la planète et l'étoile peut s'écrire ainsi :  $Z_P/Z_* \propto 1/M_P$  (Atreya et al., 2018). Cette relation de masse-métallicité est observée pour les planètes géantes du SS (tel que montré sur la Figure 1.17), mais aussi pour beaucoup d'autres planètes et sur un large éventail de masses ( $20M_{\oplus} < M < 20M_J$ ; Thorngren et al. 2016). Cette tendance est également observée dans les études de population de modèles de formation, qui montre que les planètes de plus faible masse sont généralement plus polluées par l'accrétion de planétésimaux (Fortney et al., 2013; Mordasini et al., 2016). Cependant, Welbanks et al. (2019) ont montré que la relation entre l'abondance de H<sub>2</sub>O (qu'on peut relier à la métallicité par le O/H) et  $M_P$  était différente de la relation masse-métallicité pour le C dans les

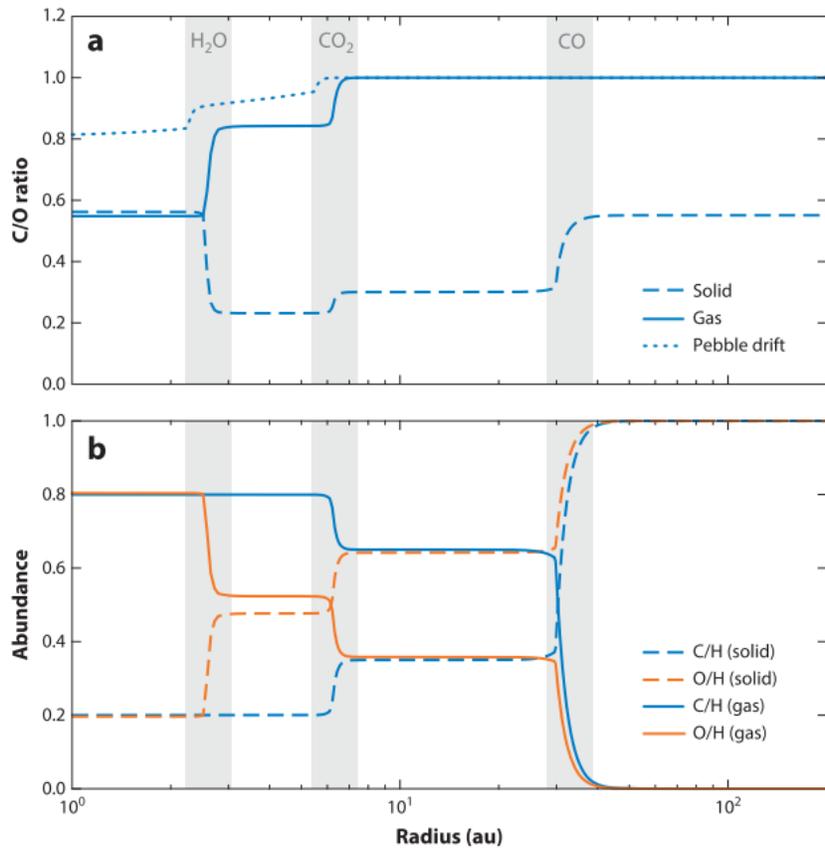


FIGURE 1.18 (a) Variation du C/O des phases gazeuse et solide en fonction de la distance de l'étoile, et influencée par les lignes de glaces (zones grises), où le gaz se condense. (b) Variation de l'abondance de C et de O des phases gazeuse et solide, relativement aux abondances solaires. Tiré de Madhusudhan (2019), Figure 8.

planètes géantes du SS (Figure 1.17). Les abondances en  $H_2O$  pour beaucoup de JC semblent plus faibles que celles observées dans le SS, comparativement aux abondances d'alcalins tels que le Na et le K, qui semblent suivre ou être supérieures à la relation du SS.

De plus en plus d'études montrent des compositions qui sont radicalement différentes de solaire, et beaucoup d'entre elles montrent des  $C/O \gtrsim 1$  (ex. : Madhusudhan et al., 2011; Madhusudhan, 2012; Moses et al., 2013a; Brogi & Line, 2019; Gandhi et al., 2019; Pelletier et al., 2021). Cela nous montre bien que la supposition d'une abondance solaire à l'équilibre chimique est trop simple et ne peut pas tout expliquer. Il faut donc se tourner vers des modèles plus complexes.

1.3.4.2. Considérations additionnelles. Après l'effondrement initial du disque, le disque va se refroidir et ses propriétés vont évoluer en conséquence. C'est d'ailleurs principalement la température dans le disque qui va influencer sa composition. À un certain moment et une certaine position dans le disque, il sera assez froid pour que certaines molécules se condensent. Cette

position est appelée la *ligne de glace* (LG) et diffère pour chacun des éléments (Öberg et al., 2011). À ces endroits, il y aura donc un changement de la composition du gaz et de la poussière qui peuvent être accrétés par les planètes en formation, tel que présenté sur la Figure 1.18. À l'intérieur d'une LG (où il fait plus chaud), les éléments seront sous forme de gaz, tandis qu'à l'extérieur (où il fait plus froid), ils seront sous forme de glaces (solides). Les principales LGs proviennent des éléments volatils les plus abondants tels que H<sub>2</sub>O, CO<sub>2</sub>, CH<sub>4</sub>, CO, etc., et ils condenseront dans cet ordre, la LG<sub>H<sub>2</sub>O</sub> (ligne de glace de l'eau) étant la plus près de l'étoile. De cette façon, la quantité d'O dans la phase gazeuse diminue plus on se déplace vers l'extérieur du disque et qu'on traverse les différentes LGs, tandis qu'elle augmente dans les solides (et de façon similaire pour le C ; Figure 1.18b).

On constate alors comment les différents scénarios de formation peuvent influencer la composition chimique, et plusieurs cas de figure en ressortent.

1) Une composition solaire pour le C/O et la métallicité (via les ratios C/H et O/H). Ceci peut être obtenu dans le cas où il y a eu formation par IGs (Section 1.2.1), ou encore via le modèle d'AC (Section 1.2.2). Dans le second cas, on peut l'observer via une accréation dominée par le gaz à l'intérieur de la LG<sub>H<sub>2</sub>O</sub> et de la ligne d'évaporation des poussières contenant du C, là où la majorité des volatils sont sous forme gazeuse.

2) Un C/O stellaire ou sous-stellaire, mais une métallicité super-stellaire. Dans ce cas, cela implique une formation par AC dominée par l'accréation de solides (de glaces et/ou de planétésimaux) à l'extérieur des LGs. Les glaces, qui sont généralement riches en O, qui sont accrétées par une planète possédant une atmosphère assez massive, seront sublimées et contribueront à l'ensemble de la composition atmosphérique (Mordasini et al., 2015; Cridland et al., 2019), ce qui diminuera le C/O et augmentera la métallicité. Toutefois, la métallicité finale ([X/H]) sera inversement proportionnelle à la masse finale de la planète (Madhusudhan et al., 2014). Un tel cas a été observé dans l'atmosphère de HAT-P-26 b (MacDonald & Madhusudhan, 2019), ayant un C/O < 0.33, qui correspond à la valeur la plus faible estimée jusqu'à maintenant, et une métallicité super-solaire (via le O/H).

3) Un C/O super-stellaire, mais une métallicité sous-stellaire. Comparativement au cas précédent, cela implique plutôt une formation par AC dominée par l'accréation de gaz au-delà des LGs. Plus on s'éloigne de l'étoile, plus le O et le C disponibles se font capturer dans les glaces et moins le gaz en est riche. Alors, une accréation dominée par le gaz, qui est pauvre en O, entraînera une augmentation du C/O et une diminution de la métallicité. Une migration vers l'intérieur après que le disque se soit dissipé (et sans accréation subséquente de planétésimaux/glaces) permettra de garder ces ratios d'abondance intacts (Öberg et al., 2011; Madhusudhan et al., 2014).

Cependant, ces scénarios sont basés sur l'hypothèse d'un disque statique, alors qu'en vérité, les disques sont tout sauf statiques.

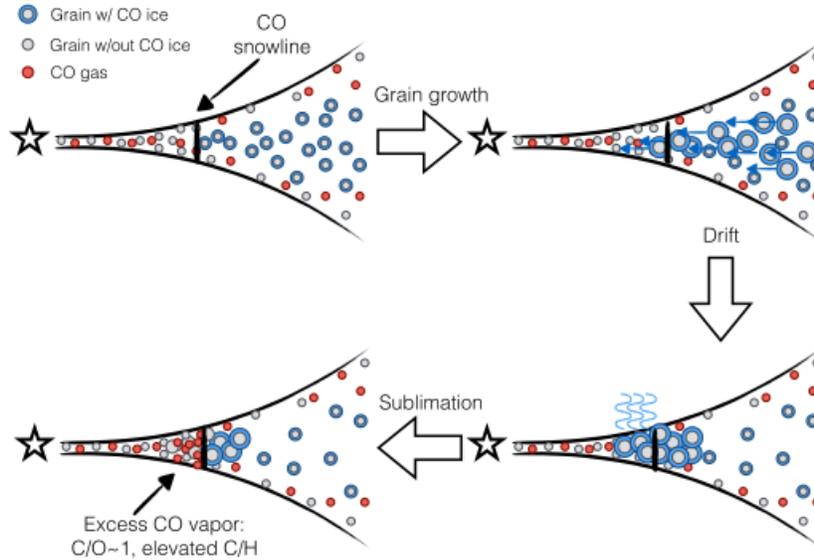


FIGURE 1.19 Schéma simplifié de la dérive de cailloux, qui focalise sur la LG du CO. À la LG, la quantité de CO sous forme gazeuse est d'abord déterminée par l'équilibre entre la condensation solide sur les grains de poussière et la sublimation. À mesure que les grains s'agglomèrent et grossissent pour former des cailloux, ils commencent à dériver vers l'intérieur. La glace de CO qui recouvre les cailloux qui traverseront la LG sera alors sublimée. Ceci augmentera alors l'abondance locale de CO dans le gaz, et mènera à un ratio  $C/O \sim 1$  et un  $C/H$  plus élevé que ce qu'on obtiendrait avec un disque statique. Tirée de Öberg & Bergin (2016).

D'une part, si on considère l'accrétion de cailloux comme la principale façon de bâtir le cœur de planètes géantes, on doit aussi considérer la dérive des cailloux (décrite à la Section 1.2.3 ; Öberg & Bergin 2016; Madhusudhan et al. 2017; Booth et al. 2017). Cette dérive permet une redistribution des composés volatils et réfractaires de l'extérieur du disque vers l'intérieur. Grâce à ce mouvement radial vers l'intérieur du disque, les glaces se feront vaporiser/sublimier lorsqu'elles traverseront leur LG. On observera alors une augmentation de la quantité de ces éléments sous forme gazeuse, qui pourra ensuite être capturée (plus facilement) par la planète. Par exemple, une formation entre les LGs de  $H_2O$  et du CO, avec un excès de CO (dû à la dérive) et avec peu de mélange entre le cœur et l'enveloppe peut mener à une atmosphère avec un  $C/O$  ET un  $C/H$  super-stellaire (voir Figure 1.19). C'est d'ailleurs le cas de  $\tau$  Bootis b, où un  $C/O \sim 1$  et une métallicité super-solaire ont été observés (Pelletier et al., 2021). Un autre exemple serait la tendance à voir des abondances d'eau plus faibles sur la Figure 1.17 de Welbanks et al. (2019) qui peut être expliquée par le cas (3) : une accrétion de gaz au-delà de la  $LG_{H_2O}$  mènera à une abondance plus faible en  $H_2O$  (et donc en O) et à un  $C/O$  super-stellaire. Par contre, une accrétion subséquente de matériel riche en métaux alcalins, potentiellement causé par la dérive de cailloux (Booth et al., 2017), pourrait ainsi expliquer les abondances super-solaires observées de Na et K.

D'une autre part, l'état des disques (leur structure radiale en température et en composition) évoluera beaucoup dans le temps (Eistrup et al., 2018; Cevallos Soto et al., 2022). 1) La quantité de matériel disponible (gaz, glace, et poussière) diminuera avec le temps à mesure que le matériel sera accrété non seulement par les planètes, mais surtout par l'étoile (ou encore éjecté par les vents stellaires). Donc plus le temps avance, moins il y aura de matériel disponible pour créer des planètes plus massives. 2) La position des lignes de glace dans le disque variera avec le temps puisque celles-ci dépendent principalement de la température dans le disque et avec son refroidissement, elles auront tendance à se déplacer vers l'intérieur, c.-à-d. se rapprocher l'étoile. Ceci aura donc comme effet de varier les ratios d'abondance des éléments clés (en l'occurrence O, C et N) en fonction du temps et de la position dans le plan du disque. 3) La viscosité du disque, le taux d'accrétion, et le profil de taille des cailloux dans le disque vont tous influencer les propriétés du disque, comme la densité et la température, le ratio gaz-poussière, etc., ce qui va ensuite influencer les propriétés des planètes qui vont y être formées (Cevallos Soto et al., 2022). 4) Le niveau de radiation dans le disque influencera la composition chimique dans le disque. En effet, le niveau d'ionisation du plan médian du disque peut fortement faire évoluer la chimie dans le disque après quelques  $10^5$  a, et continuera jusqu'à 7 Ma, surtout entre les LG de  $H_2O$  et de  $O_2$  (Eistrup et al., 2018). Le ratio C/O ne peut alors pas être considéré comme fixe, et peut même devenir plus faible que celui des glaces. De plus, à basse température, le CO qui entre en collision avec les grains de poussière peut se transformer en  $CO_2$ , ou autres glaces plus complexes, ce qui représente une façon d'épuiser la quantité de CO dans le gaz. La présence de OH est cependant nécessaire pour que cette réaction ait lieu, celui-ci créé par photo-dissociation de l'eau. Ces variations sont montrées sur la Figure 1.20.

Outre les variations chimiques produites dans les disques, plusieurs processus physiques et chimiques vont générer des déséquilibres chimiques à l'intérieur même de l'atmosphère planétaire. Ces déséquilibres s'installent lorsque les processus hors équilibre sont plus rapides que les réactions d'équilibre (surtout en haute atmosphère). Ces processus comprennent principalement le mélange vertical et la convection (mélanges turbulents ; souvent observé via le déséquilibre entre le CO et le  $CH_4$  ; Madhusudhan et al. 2014), et les réactions photochimiques (ex. : Zahnle et al., 2009; Moses et al., 2013a), mais aussi l'échappement atmosphérique (discuté dans la Section 1.3.5), et l'activité biologique (Schwieterman et al., 2018). Les études ont montré que la chimie hors équilibre est surtout présente à faible température ( $\lesssim 1300$  K), tandis que les planètes plus chaudes ( $\gtrsim 2000$  K) tendent à être en équilibre chimique.

Concernant la photochimie, c'est le processus observé lorsque la radiation de l'étoile (surtout UV) vient dissocier certaines molécules et/ou les éjecter. On attribuera d'ailleurs la création d'aérosols à la photochimie (voir 1.3.4.3). Un effet de la photochimie a été montré dans Fleury et al. (2019) : ils ont observé une conversion du CO en  $CO_2$  dans les atmosphères de JCs dominées par le  $H_2$  (de façon thermochimique pour le  $^{13}CO$ , mais photochimique pour le  $^{12}CO$ ).

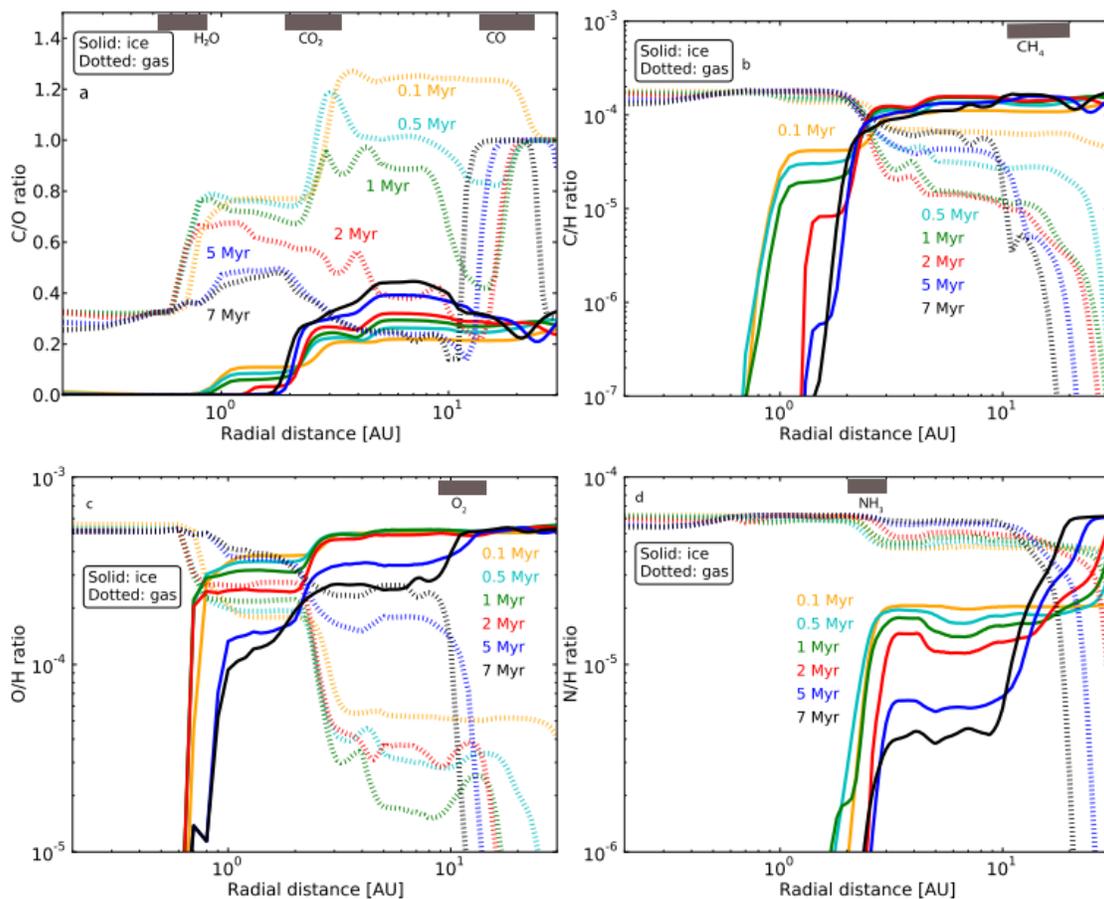


FIGURE 1.20 Ratios des abondances de C, O et N pour les phases gazeuse (lignes en tirets) et de glace (lignes pleines) en fonction de la distance de l'étoile, pour six pas de temps : 0.1, 0.5, 1, 2, 5 et 7 Ma. Les panneaux a, b, c, et d représentent les ratios C/O, C/H, O/H et N/H, respectivement. Les régions grises indiquent la zone radiale sur laquelle se déplacent les LGs de différents éléments volatils. On constate que les ratios d'abondances peuvent varier de façon importante en fonction du temps. Tirée de Eistrup et al. (2018), Figure 9.

Des températures en deçà de 1000 K sont nécessaires, sinon le CO<sub>2</sub> est dissocié et transformé en H<sub>2</sub>O. Même s'il sont présents, ces effets de déséquilibres se feront généralement moins sentir sur les planètes massives : elles sont capables d'atteindre des taux de réactions très rapides en profondeur, ce qui peut recréer les molécules qui ont été détruites ou éjectées et ainsi rétablir l'équilibre, et leur plus grand potentiel gravitationnel permet de plus facilement limiter leur « échappement atmosphérique ». Comparativement, comme l'atmosphère des planètes rocheuses (moins massives) est *généralement* plus mince et plus froide, elles sont donc aussi plus susceptibles de changer avec le temps et plus propices aux déséquilibres chimiques. Il est donc plus

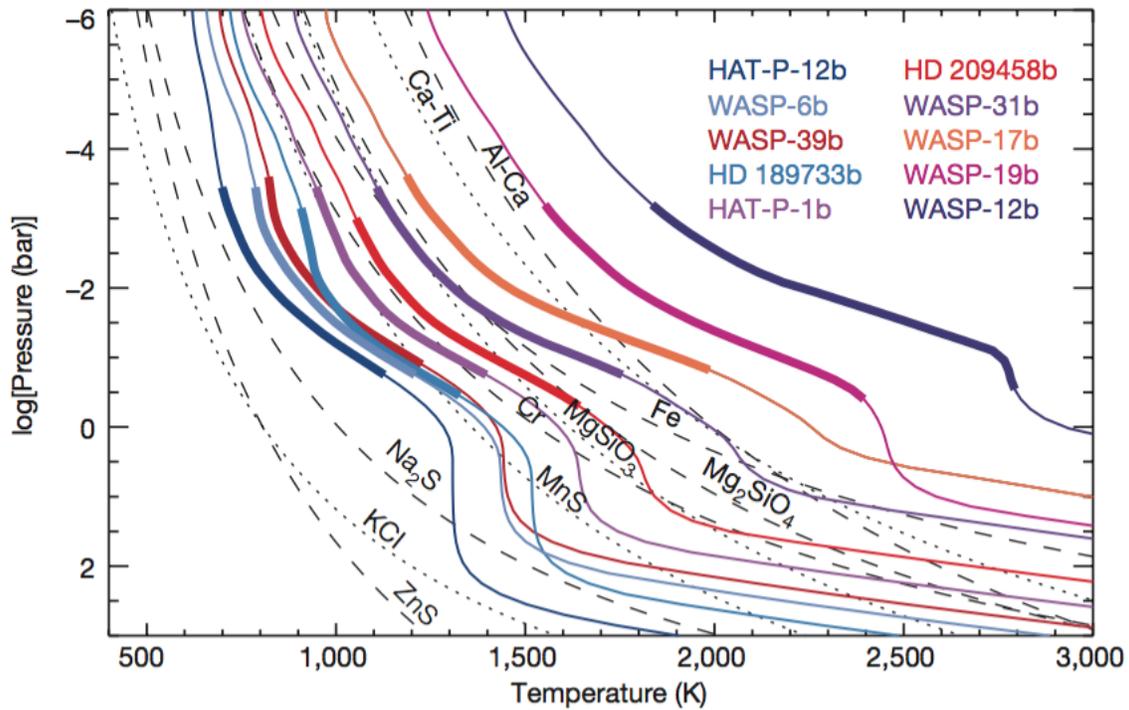


FIGURE 1.21 Profils de température-pression théorique de 10 JCs. Les régions plus épaisses représentent la zone sondée en spectroscopie de transit. Les lignes en tiret et pointillées représentent les courbes de condensation de différents éléments chimiques. Figure tirée de Sing et al. (2016).

difficile de prédire la composition de celles-ci. Il y a tout de même des moyens de déterminer l'élément dominant de ces atmosphères (par exemple, distinguer si elles sont dominées par l'hydrogène ou par l'eau ; Miller-Ricci et al. 2009).

Au final, la composition chimique des planètes dépend de *où* et *quand* celles-ci se forment dans le disque ainsi que de *quoi*, c.-à-d. la quantité relative de gaz et de poussière qui aura été accrétée (Madhusudhan et al., 2014). La composition finale sera également influencée par l'historique de migration de la planète, au cours de laquelle elle pourra accréter davantage de matériel, gazeux ou solide. Enfin, l'atmosphère poursuivra sa transformation même une fois la planète formée, les processus thermiques et photochimiques affecteront la composition moléculaire et donc ce qu'on peut observer aujourd'hui. C'est pourquoi il est impossible de dériver tout l'historique de la planète basé seulement sur le produit fini (l'atmosphère observée). Malgré tout, la composition atomique devrait rester représentative de où le matériel a été accréte dans le disque et de sa composition. Il est donc tout de même raisonnable d'explorer cette avenue puisque certaines tendances ressortent et nous permettent de mettre certaines contraintes sur ces hypothèses.

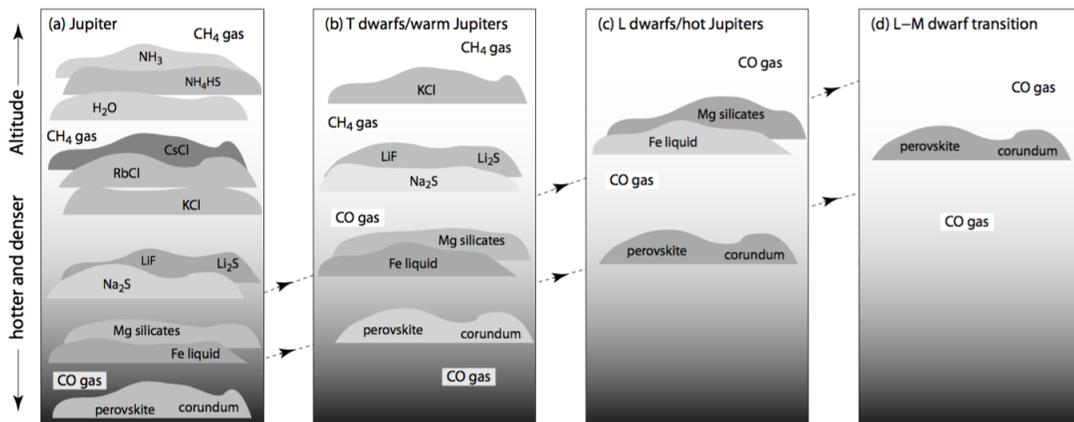


FIGURE 1.22 Schéma des couches de nuages en fonction de l'altitude dans les exoplanètes et naines brunes à différentes températures : (a) Jupiter; (b) naines T et Jupiter tièdes; (c) naines L et Jupiter Chaudes; (d) naines L-M. La structure reste similaire partout, mais le niveau des nuages des éléments réfractaires monte en altitude dans l'atmosphère des objets plus chauds. Les éléments volatils peuvent ne jamais se condenser, surtout si la température augmente. Figure tirée de Lodders (2004).

1.3.4.3. Nuages et aérosols. Toutes les planètes (ou presque) possédant une atmosphère posséderont des nuages. Ce mot englobe autant les nuages que les aérosols (ou brumes ; de l'anglais "haze"). Les nuages sont définis comme une masse de particules solide ou liquide en suspension dans l'atmosphère. Quant à elles, les brumes sont définies comme étant de très petites particules solides qui diminuent la visibilité et rendent opaque l'atmosphère. Les nuages se forment par la condensation du gaz en solide ou liquide. Pour ce faire, il faut que les conditions y soient propices : lorsque du gaz chaud monte, il prend de l'expansion et se refroidit, et si la température le permet, il y aura condensation et formation de nuage. La base des nuages se trouve à l'intersection entre la courbe de pression de vapeur de saturation (la courbe d'équilibre entre la forme gazeuse et liquide/solide d'un composé donné) et le profil TP de la planète. On peut ainsi déduire quels nuages sont présents et à quelle pression/altitude (voir les Figures 1.21 et 1.22).

Le désavantage majeur des nuages est qu'ils vont diffuser la lumière et cacher les couches d'atmosphère en dessous d'eux (ex. : spectre de transmission complètement plat de la super-Terre GJ 1214 b; Kreidberg et al., 2014). Ils agissent comme une source additionnelle d'opacité et leur présence modifie le spectre émergent, surtout les spectres de transmission. En fait, ils vont venir masquer les raies d'absorption dans les spectres (voir le bas de la Figure 1.23, p.69). Les spectres d'émission sont moins affectés par les nuages étant donné la température plus chaude du côté jour, et les lignes de visée sont proches de la normale aux couches atmosphériques, permettant ainsi de sonder les régions plus chaudes et plus profondes, mais peuvent tout de même l'être (Kitzmann et al., 2011; Parmentier et al., 2021). Les particules plus larges seront souvent modélisées comme une source d'opacité grise (diffusion indépendante de la longueur d'onde, ou

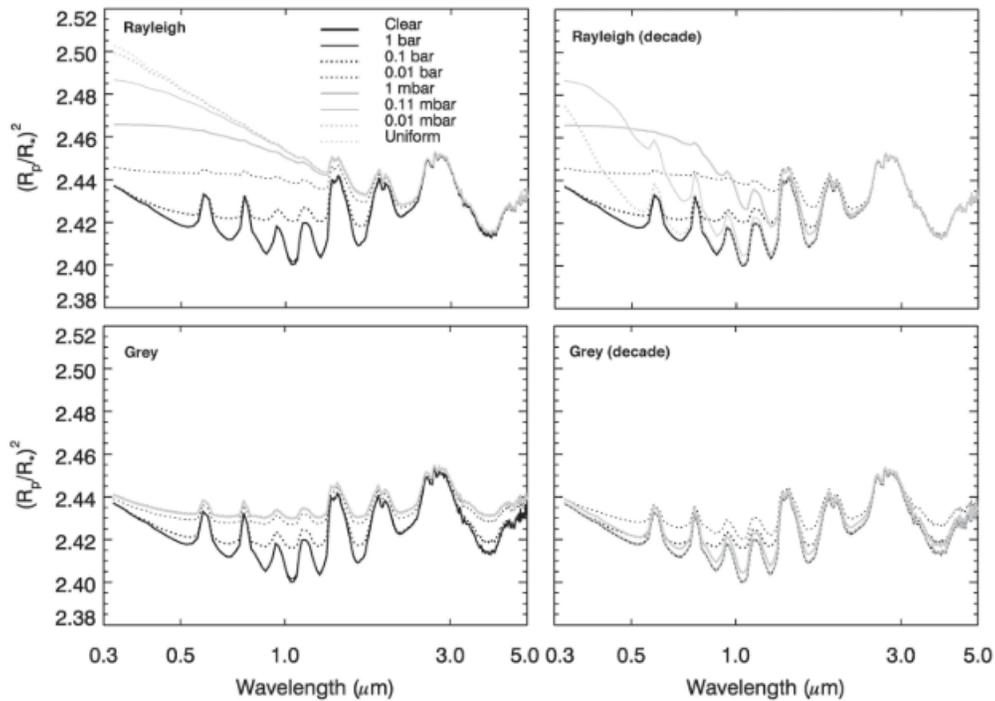


FIGURE 1.23 Effet des différentes propriétés des nuages (chromaticité [haut/bas], altitude du sommet des nuages [différentes courbes] et étendue de la couche de nuage [gauche/droite]) sur le spectre de transmission des JCs. Dans chaque panneau, les différentes courbes représentent différents niveaux du sommet des aérosols (en haut) et des nuages gris (en bas), pour une couche de nuages continue (à gauche) et une couche de nuage « mince », qui s'étale seulement sur un ordre de grandeur en pression plus bas (à droite). On voit que dans tous les cas, les structures spectrales sont affaiblies : les aérosols créent une pente où l'opacité diminue avec la longueur d'onde, tandis que les nuages gris remontent le bas des spectres et les rendent plats. Figure tirée de Barstow et al. (2017).

achromatique). Pour les brumes/aérosols, elles sont souvent associées à la diffusion de Rayleigh, qui agit surtout aux petites longueurs d'onde ( $\sigma_{\text{diff}} \propto \lambda^{-4}$  pour des tailles de particules petites et  $\ll \lambda$ ). Comme leur opacité dépend de la longueur d'onde (chromatique), elles vont donner une forme caractéristique en pente aux spectres, et vont aussi cacher certains traits spectraux (haut de la Figure 1.23). Il peut cependant y avoir de la contamination par l'activité stellaire, qui va imiter cette signature en pente dans les spectres. Le casse-tête vient se complexifier si on considère une couverture partielle (fractionnaire) des nuages (Line & Parmentier, 2016; Barstow, 2020).

La présence de nuages pose problème puisqu'il devient plus difficile de distinguer, pour une profondeur de raie donnée, si elle a été créée par une atmosphère sans nuages avec une plus faible abondance d'absorbeurs ou avec des nuages, mais aussi une abondance plus grande (Lecavelier Des Etangs et al., 2008a; Benneke & Seager, 2012,; voir aussi Section 1.4.1). C'est pourquoi il est important d'inclure les nuages dans les modèles si on veut la meilleure interprétation possible de nos données. Cependant, ils sont difficiles à modéliser et le transfert radiatif qui

inclut des nuages en devient d'autant plus complexe à résoudre. C'est un problème aux équations couplées : la présence de nuage va influencer le profil TP, ce qui va modifier les conditions pour la condensation d'autres espèces chimiques, ce qui peut générer ou enlever d'autres nuages. Par conséquent, on étudie de plus en plus l'influence des nuages et brumes sur les spectres d'atmosphère d'exoplanètes, afin de lever les dégénérescences entre la présence de nuage et l'abondance des éléments chimiques (ex. : Line & Parmentier, 2016; Madhusudhan et al., 2016; MacDonald & Madhusudhan, 2017; Pinhas & Madhusudhan, 2017; Kempton & Bean, 2017).

La pente de l'absorption des aérosols peut tout de même permettre d'estimer le poids moléculaire moyen de l'atmosphère de façon indépendante (c'est aussi possible par la forme des structures d'absorption individuelles et/ou de la profondeur de transit relative entre les différentes structures en absorption d'une même molécule; Benneke & Seager 2012). La mesure de la pente de diffusion des aérosols permet aussi une mesure de l'abondance des gaz inertes (comme l' $H_2$  et le  $N_2$ ), qui permet ensuite la détermination des abondances absolues des éléments présents. Sans cela, c'est plutôt des abondances relatives qu'on pourra extraire.

### 1.3.5. Dynamique

On observe de la circulation atmosphérique dans l'atmosphère de toutes les planètes du SS. Aucune atmosphère ne peut réellement être « sans-vent », puisque c'est un mécanisme de redistribution d'énergie. Les gradients de température (et donc de pression) vont générer des vents (mouvements globaux de masse de gaz), qui tendent à vouloir les éliminer. La circulation sera principalement influencée par l'irradiation de l'étoile et la période de rotation de la planète.

Pour les planètes près de leur étoile, comme les JCs, et qui sont verrouillées gravitationnellement, on s'attend à ce qu'il y ait de forts gradients de température entre le côté jour et le côté nuit (créant des écoulements zonaux; direction est-ouest), ainsi qu'entre l'équateur et les pôles (créant des écoulements méridionaux; direction nord-sud). La photométrie, les spectres et les courbes de lumière IR de plusieurs JCs ont permis de mettre des contraintes 3D sur leur structure en température, et indiquent de la circulation atmosphérique intense (Showman & Polvani, 2011). Les courbes de phases ont permis de montrer une tendance des planètes les plus irradiées (températures d'équilibre plus grandes) à être moins efficaces à transporter l'énergie (des contrastes de température jour-nuit plus grands). Elles ont aussi permis de montrer que beaucoup de JCs ont un point chaud décalé du point sous-stellaire, causé par les écoulements zonaux (courants-jets vers l'est, aussi appelé *super-rotation*). L'exemple de HD 189733 b (Knutson et al., 2007) est montré sur la Figure 1.24. Avec les données haute résolution, on peut directement mesurer la vitesse des vents dans l'atmosphère des JCs. Les écoulements zonaux vont aussi créer une différence de décalage en VR du signal à l'entrée et la sortie de transit qui peut être détectée si le signal est assez fort (ex. : des vents de  $2.3_{-1.5}^{+1.3}$  km s<sup>-1</sup> et  $5.3_{-1.4}^{+1.0}$  km s<sup>-1</sup> pour le côté

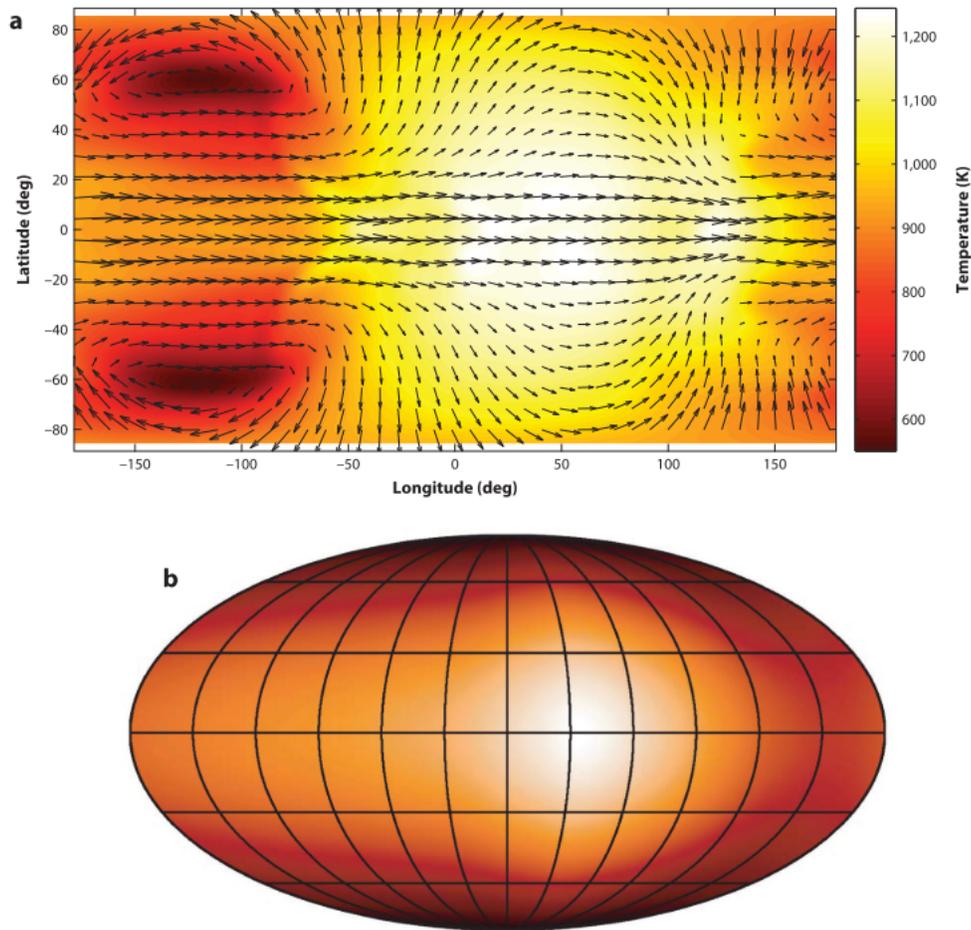


FIGURE 1.24 Exemple de la dynamique dans HD189733b. (a) Prédictions théoriques d'un GCM montrant le courant-jet qui pousse le point chaud à l'est du point sous-stellaire (Showman & Polvani, 2011). Les vecteurs de vents sont représentés par les flèches, et l'échelle de couleurs est associée à la température, en K. (b) Reconstruction de la carte de température de luminosité à partir des observations de courbe de phase obtenues avec Spitzer à  $8\ \mu\text{m}$  (Knutson et al., 2007). Les modèles et les observations sont cohérents entre eux. Figure tirée de Madhusudhan (2019).

matin et le côté soir, respectivement, ont été observée sur HD189733 b; Louden & Wheatley, 2015). De plus, un décalage vers le bleu en VR des raies de CO ( $\sim 2\text{km/s}$ ) par rapport à la vitesse systémique de la planète HD 209458 b et moyenné sur tout le terminateur, a été observé, suggérant que de forts vents soufflent du côté jour au côté nuit (Snellen et al., 2010).

Ces résultats (la présence de vents et jets équatoriaux) sont prédits par les GCMs de JCs (Figure 1.24; Showman et al., 2009; Miller-Ricci Kempton & Rauscher, 2012; Rauscher & Kempton, 2014; Zhang et al., 2017). Ces modèles 3D permettent une représentation de la dynamique qui est plus fidèle, mais comme pour les profils TP, ceux-ci sont beaucoup plus coûteux en termes de calcul numérique et c'est pourquoi ils ne sont pas toujours utilisés. Beltz et al.

(2021) ont d'ailleurs montré que les informations 3D sur la dynamique sont capturées par les spectres à haute résolution. Par exemple, Flowers et al. (2019) ont pu observer les effets de la rotation de la planète et des vents/de la super-rotation (voir Section 1.4.2).

Un autre effet dynamique à considérer est les pertes atmosphériques. Comme mentionné brièvement à la Section 1.3.4.2, plusieurs processus d'« évaporation » atmosphérique ont été observés et feront varier la composition de l'atmosphère des exoplanètes, influenceront son évolution, et vont, en fait, déterminer si elles en possèdent une ou non.

D'abord, les pertes thermiques (ou « échappement thermique » ; « *Jeans escape* ») sont causées par l'échappement des atomes, ions et molécules ayant des vitesses plus grandes que la vitesse d'échappement et qui sont assez hauts dans l'atmosphère pour qu'il n'y ait plus de collisions pour les « stopper ». Les vitesses nécessaires seront atteintes par les particules qui sont à la fin de la distribution en vitesse de Maxwell-Boltzmann. Ensuite, l'« échappement hydrodynamique » est observé lorsque l'atmosphère est tellement chauffée par la radiation qu'un vent thermique d'éléments légers entraîne les éléments plus lourds avec eux (cas asymptotique de l'échappement thermique). Enfin, on peut également penser à d'autres processus dits « non thermiques ». Par exemple, les champs magnétiques vont accélérer les ions et les pousser à s'échapper ou à entrer en collision avec d'autres particules qui pourront s'échapper ; les produits de la photodissociation, induite par la radiation UV, peuvent acquérir de grandes vitesses et s'échapper de l'atmosphère ; les vents stellaires peuvent interagir avec les particules et les entraîner loin de la planète.

L'échappement d'atomes d'hydrogène de l'exosphère de HD 209458 b a été observé (Vidal-Madjar et al., 2003), via la détection de la raie d'absorption Ly- $\alpha$  dans l'UV, et dans plusieurs autres exoplanètes depuis, dont HD 189733 b (Lecavelier Des Etangs et al., 2010). Ces études de la raie de Ly- $\alpha$  permettent non seulement de mieux comprendre la composition de l'exosphère des planètes, mais aussi de contraindre l'échappement de gaz de celle-ci, son taux de perte de masse et même la morphologie du gaz qui s'échappe (qui viendra façonner la forme de la courbe du transit du Ly- $\alpha$ , et qui peut être très asymétrique ; Bourrier et al., 2016). Si l'atmosphère des planètes s'échappe assez longtemps, on peut bien s'imaginer qu'elles peuvent éventuellement complètement perdre leur atmosphère, ce qui peut représenter la perte d'une fraction plus ou moins importante de leur masse initiale. C'est pourquoi les meilleures cibles pour étudier ces échappements seraient les Neptunes chaudes qui orbitent les naines M. D'une part, pour être capable d'observer ce phénomène d'échappement, il faut qu'une atmosphère soit encore présente et en train d'être perdue. En général, les JCs et les planètes géantes subiront des échappements atmosphériques, mais leur taux de perte de masse sera trop faible pour affecter leur évolution de façon importante, ce qui nous permet d'observer le phénomène en action, même pour les planètes âgées de plusieurs milliards d'années (Owen, 2019). D'une autre part, leur exosphère étendue offre un meilleur contraste face à la taille relativement petite de leur étoile hôte de faible

masse (ex. : Lecavelier Des Etangs et al., 2010; Lecavelier des Etangs et al., 2012; Bourrier et al., 2013; Ehrenreich et al., 2015; Bourrier et al., 2018).

Plus récemment, le He, précisément la raie d'absorption de l'hélium métastable à 1083 nm, a été observée avec HST dans l'exosphère de WASP-121b (Spake et al., 2018), et dans plusieurs autres cibles depuis (ex. : Allart et al., 2018; Alonso-Floriano et al., 2019; Guilluy et al., 2020; Zhang et al., 2022). Cette raie, également observable avec les instruments aux sols, a ouvert une nouvelle avenue et a facilité les observations pour faire le pont entre l'atmosphère et l'exosphère de certaines planètes (Allart et al., 2018). Cette raie semble cependant être présente seulement sous certaines conditions d'irradiation bien précises (Oklopčić & Hirata, 2018; Oklopčić, 2019).

## 1.4. Approches observationnelles

L'acquisition des spectres en émission et en transmission peut se faire soit avec la spectroscopie à faible dispersion ou avec la spectroscopie à haute dispersion. Dans tous les cas, en prenant une série de spectres avant et pendant un transit (ou éclipse secondaire, ou courbe de phase), on réussit à avoir la variation de la profondeur du transit (ou de l'éclipse secondaire, ou le signal en émission de la planète relatif à celui de l'étoile) simultanément à plusieurs longueurs d'onde. On peut ainsi retrouver le spectre de la planète, en se basant sur les sections 1.1.3.1 à 1.1.3.3. Pour chacune des méthodes d'observation (faible et haute dispersion), les avantages, les désavantages et les différences seront présentés, et un survol des résultats principaux obtenus grâce à ces méthodes suivra.

### 1.4.1. Spectroscopie à faible dispersion

Avec la *spectroscopie à faible dispersion* (SFD) on obtient des spectres à faible résolution,  $R \lesssim 5000$ . La quantité  $R = \lambda/\Delta\lambda$  est le « pouvoir de résolution » d'un spectrographe,  $\Delta\lambda$  étant la « résolution spectrale », et indique à quel point un spectrographe peut résoudre les détails d'un spectre (on peut aussi l'écrire sous la forme  $R = c/\Delta v$ ).

Avec les observations à faible résolution, on perd les détails fins du spectre, qui comportent beaucoup d'information ; pour voir les raies individuellement, on doit se tourner vers la spectroscopie à haute résolution (HR ; voir Section 1.4.2). Tout de même, on réussit à voir les bandes d'absorption : les ensembles de raies d'absorption individuelles se combinent pour former des bandes d'absorption, ce qui donne la forme générale de l'enveloppe des spectres. Aussi, c'est la méthode qui offre le plus de photons par élément de résolution. Ceci nous permet d'aller chercher des signaux plus faibles.

Les instruments les plus utilisés sont ceux à bord du *Hubble Space Telescope* (HST), télescope spatial de 2.4 m. D'abord, on compte le *Space Telescope Imaging Spectrograph* (STIS), un spectro-imageur couvrant des UV jusqu'à  $\sim 1 \mu\text{m}$  et ayant une résolution faible à moyenne

dans le plus rouge,  $R \sim 500\text{--}10\,000$ , mais pouvant atteindre jusqu'à  $\sim 120\,000$  dans les UV et le visible, et ensuite le *Wide Field Camera 3*; (WFC3;  $R \sim 130$ ), une caméra à grand champ (permettant de faire de la spectroscopie sans fente) qui possède deux détecteurs, un dans l'UV et le visible et l'autre dans l'infrarouge proche ( $0.85\text{--}1.7\ \mu\text{m}$ ; IRP), ce qui est optimal pour l'observation d'exoplanètes. La première caractérisation de l'atmosphère d'une exoplanète a été la détection du doublet du sodium à  $589.3\ \text{nm}$  par Charbonneau et al. (2002) dans l'atmosphère de HD 209458 b, avec STIS, mais depuis, c'est plutôt WFC3 qui est le plus couramment utilisé (ex. : Iyer et al., 2016). Outre l'utilisation des spectres, il est aussi possible d'utiliser des points photométriques pour aller chercher de l'information dans d'autres plages du spectre électromagnétique. Les données du *Infrared Array Camera* (IRAC) du télescope *Spitzer* sont souvent utilisées (bandes à  $3.6, 4.5, 5.8$  et  $8,0\ \mu\text{m}$ ) puisqu'elles sont complémentaires aux données de HST (ex. : Tinetti et al., 2007; Knutson et al., 2011; Désert et al., 2011; Line et al., 2014; Saba et al., 2021; Wallack et al., 2021). On se rappellera que la majorité du signal des exoplanètes, qui sont en général des objets plus froids, est émis dans l'IR, d'où l'intérêt des données Spitzer, et l'intérêt qu'auront les données du JWST dans l'étude des exoplanètes.

La SFD se fait surtout (mais pas exclusivement) dans l'espace puisqu'on peut atteindre une bien meilleure précision. Le principal avantage de l'espace est l'absence de l'atmosphère terrestre, qui autrement introduit une transmission et une qualité d'image variables dans le temps et très difficiles à calibrer, résultant en un important bruit systématique résiduel dans le signal mesuré. Il y a aussi généralement moins de bruit de fond thermique, ce qui est optimal lorsqu'on observe dans l'IR. Par contre, le temps d'observation spatial est souvent plus contingenté étant donné les temps de vie « limités » des missions. De plus, c'est plutôt dispendieux et périlleux d'envoyer de gros télescopes dans l'espace. La taille de ces télescopes est donc beaucoup plus petite que ceux au sol, ce qui limite la sensibilité des observations. C'est d'ailleurs ce qui rend le JWST aussi impressionnant : il révolutionnera les observations dans l'espace à cause de son aire collectrice record de  $25\ \text{m}^2$ . Malheureusement, sa distance de la Terre limitera beaucoup plus son temps d'opération puisque, contrairement à HST, il ne pourra pas être réparé. Un autre avantage d'aller dans l'espace est que c'est le seul endroit où il est possible d'obtenir des données dans les plages de longueurs d'onde qui sont absorbées par l'atmosphère de la Terre, comme l'UV et les bandes d'eau saturées de l'atmosphère terrestre. Il est donc nécessaire d'utiliser des télescopes spatiaux pour aller chercher les signaux tels que celui de  $\text{Ly-}\alpha$  pour une meilleure caractérisation de l'exosphère (ex. : Lecavelier des Etangs et al., 2012), et la pente de diffusion des aérosols (ex. : Lecavelier Des Etangs et al., 2008a,b; Lothringer et al., 2022).

Un léger désavantage des instruments spatiaux d'aujourd'hui est leur bande passante relativement limitée. Bien que STIS et WFC3 couvrent des UV jusqu'au visible et à l'IRP, et permettent de détecter les éléments chimiques visibles dans cette gamme de longueurs d'onde, il reste qu'on est plus limité dans l'IRP pour faire la détection des principales molécules porteuses

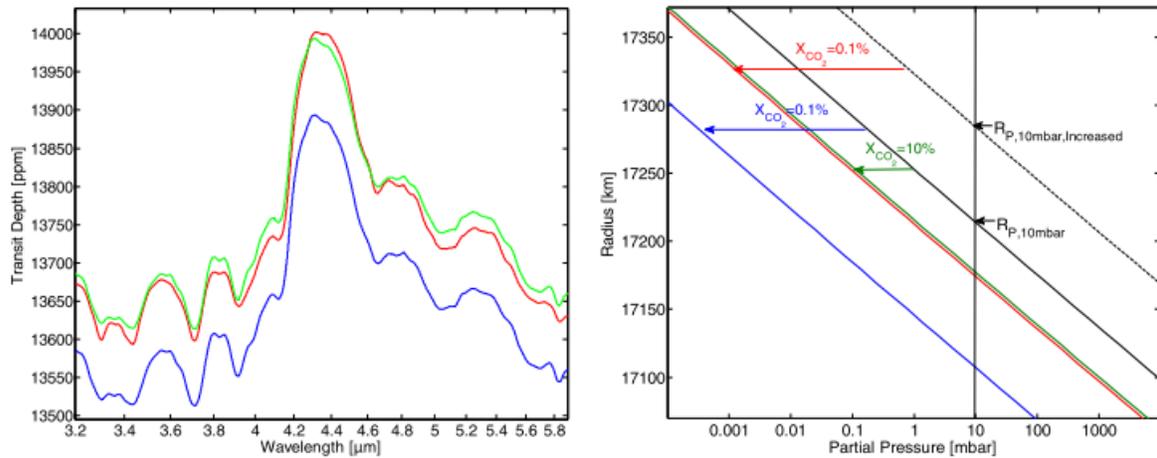


FIGURE 1.25 Dégénérescence entre l'abondance (ou rapport de mélange),  $X_{\text{CO}_2}$  et le rayon de la planète  $R_{\text{P},10}$ . (Gauche) Spectre de transmission du scénario A (rouge; 10%  $\text{CO}_2$  et 90%  $\text{N}_2$ ) et B (bleu; 0.1%  $\text{CO}_2$  et 99.9%  $\text{N}_2$ ), et B avec un plus grand rayon (vert). À moins d'avoir accès à une autre section de spectre qui puisse ancrer la solution, on ne peut pas déterminer  $X_{\text{CO}_2}$  à mieux que plusieurs ordres de grandeur. (Droite) Pression totale de la planète pour deux  $R_{\text{P}}$  différents (noir) et la pression partielle du  $\text{CO}_2$  en fonction de la distance au centre de la planète (mêmes couleurs que gauche). Deux planètes avec différentes abondances peuvent avoir la même pression partielle (rouge et vert) en fonction de l'altitude, et donc des structures spectrales similaires. Figure tirée de Benneke & Seager (2012).

de carbone et d'oxygène. Par exemple, le CO et le  $\text{CO}_2$  n'ont pas de caractéristiques spectrales importantes dans l'intervalle de longueur d'onde de HST, et celles-ci se chevauchent dans les bandes de Spitzer (Diamond-Lowe et al., 2014; Brogi et al., 2017; Spake et al., 2021). Aussi, un tel intervalle limité peut même empêcher la détection de  $\text{H}_2\text{O}$  (qui est généralement l'élément le plus facile à détecter) à cause de la présence potentielle de nuages ou d'aérosols, mais qui serait autrement visible/détectable plus loin dans l'IR (voir Figure 1.27). Encore une fois, JWST viendra révolutionner la SFD grâce à son intervalle de longueur d'onde qui va s'étendre de 0.6 à 12  $\mu\text{m}$ .

Par rapport à la présence de nuages, il est bien connu qu'il existe une dégénérescence parfaite entre l'abondance absolue des éléments chimiques ( $X_i$ , l'abondance de l'élément  $i$ ) et la présence de nuages ( $P_{\text{cloud}}$ ; tel que présenté dans la Section 1.3.4.3; Lecavelier Des Etangs et al. 2008a; Benneke & Seager 2012). Par exemple, il sera difficile de différencier si un spectre complètement plat (sans caractéristiques spectrales) provient d'une atmosphère n'ayant aucune vapeur d'eau et aucun nuage, ou d'une atmosphère ayant beaucoup d'eau, mais aussi beaucoup de nuages qui viennent cacher les structures du spectre. Non seulement cela, mais c'est aussi dégénéré avec le rayon de la planète ( $R_{\text{P}}$ ), et la pression de référence,  $P_0$  (Griffith, 2014; Heng & Kitzmann, 2017; Fisher & Heng, 2019). Pour illustrer ceci, considérons deux planètes A et B ayant le même rayon, mais B ayant des abondances moléculaires plus petites que A (molécules « actives »; voir Figure 1.25). Alors, B aura une profondeur de transit globalement plus petite

que celle de A. Par contre, si on augmente le rayon de B, on retrouvera une profondeur de transit similaire à A, et on ne peut plus différencier le cas A du cas B. De plus, dépendamment de quelle pression de référence est utilisée, on obtiendra différentes valeurs de  $R_p$ . En gros, tout ce qui fait varier la profondeur de transit et l'amplitude du signal (qui dépend de  $H$ , et qui elle dépend des éléments déjà énumérés plus haut, mais aussi de la température étant donné son effet sur la structure en densité et par conséquent l'opacité) sera propice à varier les résultats s'ils ne sont pas tous pris en compte. C'est pourquoi il est aussi plus facile de calculer l'abondance relative des éléments que de calculer leur abondance absolue à HR (en comparant la profondeur relative des bandes moléculaires – ou des raies). Toutefois, Welbanks & Madhusudhan (2019) ont montré qu'en ayant des modèles qui incluent peu d'hypothèses simples et avec un spectre qui couvre un intervalle de longueur d'onde assez large (visible et IR), on peut atteindre une détermination des abondances chimiques et des propriétés des nuages/aérosols assez précises. Aussi, cette dégénérescence peut être partiellement levée, ou complètement, si la pente de diffusion dans le domaine visible peut être attribuée au  $H_2$  (Lecavelier Des Etangs et al., 2008a; Fisher & Heng, 2018), mais ne peut être brisé si la diffusion est causée par la présence de nuages ou d'aérosols (Heng, 2016). Autrement, une analyse combinée en émission et en transmission permet de mieux contraindre la structure en température et  $R_p$ , et ainsi la composition. Puisque les spectres en émission permettent d'aller chercher le profil de température plus profondément dans l'atmosphère et sont moins sensibles aux nuages, il est plus facile de directement déterminer les abondances absolues, de par la forme et l'amplitude des structures spectrales qui dépendent du profil TP.

Une dernière complication doit être considérée en transit, soit l'effet de source lumineuse du transit (de l'anglais *Transit Light Source Effect*, voir Figure 1.26). Il s'agit du fait que la mesure de la profondeur de transit est une mesure différentielle avec le signal de l'étoile, intégré sur sa surface. Cependant, ce signal peut subir des variations/être contaminé par la présence des taches stellaires et de facules/plages (zones plus froides et plus chaudes à la surface de l'étoile, respectivement). On mesure généralement le spectre de l'étoile comme étant le spectre de l'étoile hors transit, intégré sur tout le disque. Si l'étoile ne possède aucune tache ou facule, la zone occultée par la planète sera bien représentée par le spectre moyenné sur le disque. Par contre, si l'étoile en possède, on aura alors 2 cas de figure : a) Si la planète traverse ces zones, on pourra généralement en déterminer l'importance et corriger pour cet effet ; b) Si la planète ne passe *pas* devant ces zones, alors le spectre moyenné de l'étoile ne représentera pas bien la zone occultée par la planète, ce qui introduira des erreurs dans le spectre de transmission retrouvé de la planète (Rackham et al., 2018). Cet effet est plus important pour les étoiles plus actives, comme les naines M, où il est possible de voir apparaître dans le spectre stellaire les mêmes signaux que sur la planète (comme de  $H_2O$ ) provenant des taches assez froides. Ce problème s'avère pire si les raies stellaires et planétaires se chevauchent, ce qui est possible lors d'un

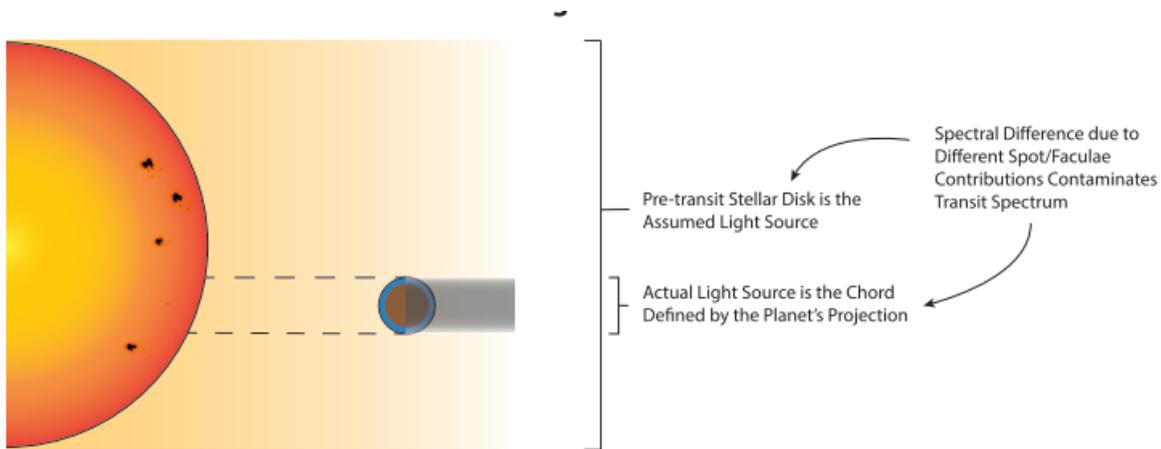


FIGURE 1.26 Schéma de l'effet de source lumineuse du transit. Durant un transit, la planète masque la portion de la photosphère stellaire directement derrière elle (du point de vue de l'observateur). Pour mesurer une profondeur de transit, on devrait donc utiliser le spectre de cette portion de photosphère seulement (la vraie source). Par contre, on utilise généralement le spectre moyen de toute la surface de l'étoile (la source supposée). Alors, toutes les différences entre la vraie source et la source supposée généreront des variations dans la profondeur de transit mesurée et affecteront le spectre de transmission de la planète. Le passage des taches et des faculae produira également une modulation dans le flux (non montré), et son amplitude dépendra de la taille et de la densité de ces structures sur la surface stellaire. Figure tirée de Rackham et al. (2018).

transit (Deming & Sheppard, 2017). Rackham et al. (2019) ont montré que cet effet était moins importants pour les étoiles de types solaires (de type FGK). Genest et al. (2022) ont montré qu'avec la SHD (voir les détails à la prochaine section), on pouvait démêler le signal provenant de l'étoile de celui provenant de la planète et résoudre en partie ce problème, si la vitesse orbitale de la planète varie suffisamment.

En plus de rendre possible la cartographie de la surface des planètes via leur température (Knutson et al., 2007; Stevenson et al., 2014), la SFD a permis la détection d'un bon nombre d'éléments chimiques. La première détection de l'atmosphère d'une exoplanète a été la détection du doublet du Na à basse résolution avec HST/STIS dans HD 209458 b (Charbonneau et al., 2002). Les détections qui ont suivi comprennent : la détection de H, via la raie de Ly- $\alpha$  dans l'UV (ex. : Vidal-Madjar et al., 2003; Lecavelier Des Etangs et al., 2010; Bourrier et al., 2013; Ehrenreich et al., 2015,; qui permet l'étude de l'échappement atmosphérique), et de métaux ionisés (comme le O I, et le C II; Vidal-Madjar et al., 2004); les éléments inertes H<sub>2</sub> et N<sub>2</sub>, via la pente de diffusion de Rayleigh, dans l'UV proche et le visible (ex. : Lecavelier Des Etangs et al., 2008a; Benneke & Seager, 2012); la détection de métaux alcalins (comme le Na et le K), ainsi que certaines molécules comme le TiO et le VO, dans le domaine visible (ex. : Redfield et al., 2008; Sing et al., 2011; Chen et al., 2017); le He métastable dans l'IRP à 1080.3 nm (ex. : Spake et al., 2018, également en lien avec l'échappement atmosphérique, et qui permet

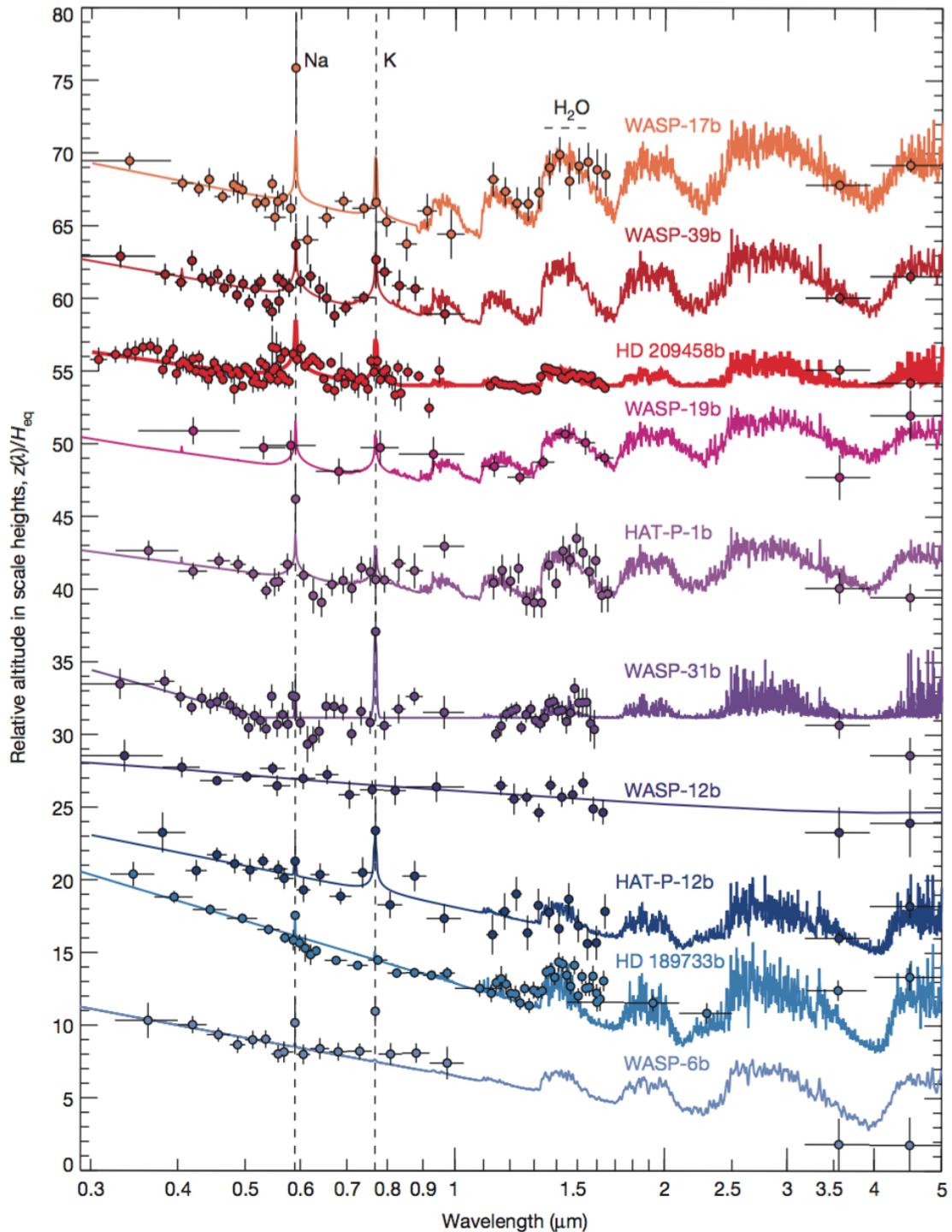


FIGURE 1.27 Séquence de spectres de transmission de JCs observés avec HST et Spitzer. Les lignes pleines montrent le meilleur ajustement des modèles atmosphériques. Les spectres ont été décalés pour améliorer la vision et ont été mis en ordre de la différence d'altitude entre le côté bleu du visible et l'IR moyen. Les planètes avec des atmosphères plutôt « claires » (sans nuages; en haut) montrent des raies de métaux alcalins et de H<sub>2</sub>O bien définies. Les planètes ayant des atmosphères plutôt nuageuses (vers le bas) ont de fortes pentes de diffusion, des raies alcalines minces et des signaux de H<sub>2</sub>O partiellement ou complètement obscurcis. Figure tirée de Sing et al. (2016).

de relier l'exosphère aux couches plus basses de l'atmosphère); beaucoup de détections d' $\text{H}_2\text{O}$  dans l'IRP (ex. : Kreidberg et al., 2014; McCullough et al., 2014; Line et al., 2017; Tsiaras et al., 2018; Mikal-Evans et al., 2020; Mansfield et al., 2021). Une liste plus exhaustive de ce type de détection a été comptabilisée dans Madhusudhan (2019), mais celle-ci a continué de grandir depuis. La détection de molécules contenant du carbone (comme le CO) est plus ardue, mais a été tentée par plusieurs en utilisant des données Spitzer, qui couvrent plusieurs structures spectrales de plusieurs absorbeurs (Désert et al., 2009; Madhusudhan et al., 2011; Stevenson et al., 2017). Tout de même, la majorité des détections concluantes de CO ont été faites à partir des instruments au sol (voir la prochaine Section).

Cette multitude de détections sur une panoplie de cibles, surtout celle de  $\text{H}_2\text{O}$  qui est devenue routinière, a permis de faire des études comparatives et d'extraire certaines tendances de manière consistante, dont certaines ont été brièvement mentionnées dans les sections précédentes. Par exemple, la majorité des JCs montrent des signes de la présence d'aérosols et de diffusion Rayleigh et/ou encore de nuages gris (Barstow et al., 2017). Aussi, la faiblesse des signatures d'eau observée dans plusieurs JCs peut être expliquée soit par la présence de nuages (Sing et al., 2016,; voir Figure 1.27 qui montre un continuum d'atmosphères allant de sans à avec nuages) ou soit par une sous-abondance due à la position de la formation dans le disque (Pinhas et al., 2019; Welbanks et al., 2019, voir Figure 1.17), qui peut aussi expliquer une tendance de certaines atmosphères à avoir de forts C/O (Brewer et al., 2017). De plus, l'analyse comparative des spectres en émission de JCs faite par Mansfield et al. (2021) a également permis de montrer qu'il existait une relation entre l'amplitude de la bande de  $\text{H}_2\text{O}$  à  $1.4 \mu\text{m}$  et la température de l'atmosphère.

### 1.4.2. Spectroscopie à haute dispersion

La *spectroscopie à haute dispersion* (SHD) permet d'obtenir des spectres à HR ( $R > 20000$ ) et donc d'aller chercher le détail des raies d'absorption individuelles. Contrairement à la SFD, il est alors possible de différencier le signal de différents absorbeurs qui ont des bandes d'absorption qui se chevauchent (tant qu'une majorité de raies ne se chevauchent pas). La SHD correspond en fait à un idéal; la configuration ultime serait d'être capable d'en faire dans l'espace. Cependant, ce n'est pas encore envisageable. D'abord, les systèmes de SHD sont nécessairement plus volumineux que leurs analogues basse résolution, mais en plus, on doit considérer que plus on disperse la lumière, plus le nombre de photons reçus par élément de résolution sera petit et plus on aura besoin de gros miroir pour avoir un signal suffisant. Cela rend ces systèmes plus difficiles à envoyer dans l'espace et, à ce jour, il n'y en a pas de disponible. On doit donc se contenter de faire la SHD au sol, pour l'instant.

Ensuite, comme on observe au sol, cela implique qu'on doit travailler avec l'absorption tellurique et la turbulence atmosphérique, qui représente un des plus grands challenges/ennuis de la méthode. L'absorption tellurique (absorption causée par l'atmosphère terrestre, dominée par la vapeur d'eau, et plus présente dans l'IR) vient creuser des portions, parfois non négligeables, du spectre qu'on observe. Comme on connaît bien ce phénomène et le spectre de transmission terrestre, on peut essayer de le modéliser, d'enlever sa contribution et ainsi de reconstruire le spectre original. Par contre, pour certaines raies trop profondes ou complètement saturées, il est simplement impossible de retrouver le signal original. Tout de même, il existe de multiples méthodes pour corriger les telluriques. Les telluriques sont généralement corrigés en utilisant des modèles synthétiques de transmission atmosphérique terrestre (ex. : avec *Molecfit* de Smette et al. 2015 ; Allart et al. 2017), en enlevant le signal qui possède une dépendance sur la masse d'air (ex. : Brogi et al., 2018), en utilisant des méthodes d'analyse aux composantes principales (de l'anglais *Principal Component Analysis*, PCA), par exemple avec les PCAs directement, la *Singular Value Decomposition*, et *SYSREM* (Damiano et al., 2019; de Kok et al., 2013a; Alonso-Floriano et al., 2019, respectivement), ou encore une combinaison de ces méthodes, ex. : *APER0*, (ex. : tel qu'utilisé dans Boucher et al., 2021). Quant à la turbulence, c'est principalement un problème en imagerie directe, mais on peut résoudre quelque peu ce problème avec des systèmes d'optique adaptative et des observatoires en très haute altitude. Toutefois, dans certaines conditions extrêmes, la turbulence pourrait introduire des erreurs systématiques dans l'extraction de la VR du signal. Il faut donc bien s'assurer de corriger adéquatement les telluriques pour limiter la contamination, surtout si on cherche un signal d'H<sub>2</sub>O, qui pourrait facilement être considéré à tort comme provenant de la planète.

Par contre, un avantage d'être sur la Terre est qu'on a le pouvoir de faire de plus gros télescopes (ayant une plus grosse surface collectrice ; ex. : le *Extremely Large Telescope* – ELT – de 39.3 m de diamètre, et le *Thirty Meter Telescope* – TMT – de 30 m de diamètre), qui peuvent être plus « facilement » réparés et améliorés au fil du temps. Ces énormes capteurs de photons permettent de contrer partiellement le problème de signaux faibles typique de la HR et augmentent le nombre de photons reçus par élément de résolution.

Comme à haute résolution le nombre de photons reçus par élément de résolution sera plus petit, le S/N de la planète par élément de résolution sera donc également plus petit. Par contre, à HR, on a accès à un plus grand nombre de raies individuelles. En combinant le signal de toutes les raies disponibles, on réussit à augmenter le S/N global de la planète, proportionnellement à  $\sqrt{N}$ , où  $N$  est le nombre total de raies. Il est donc avantageux d'avoir une plus grande plage spectrale, afin de maximiser  $N$ . D'élargir la plage spectrale permet aussi d'aller chercher un plus grand nombre d'espèces chimiques, et leur signal associé. Les méthodes pour extraire ce signal sont expliquées à la Section 1.5.

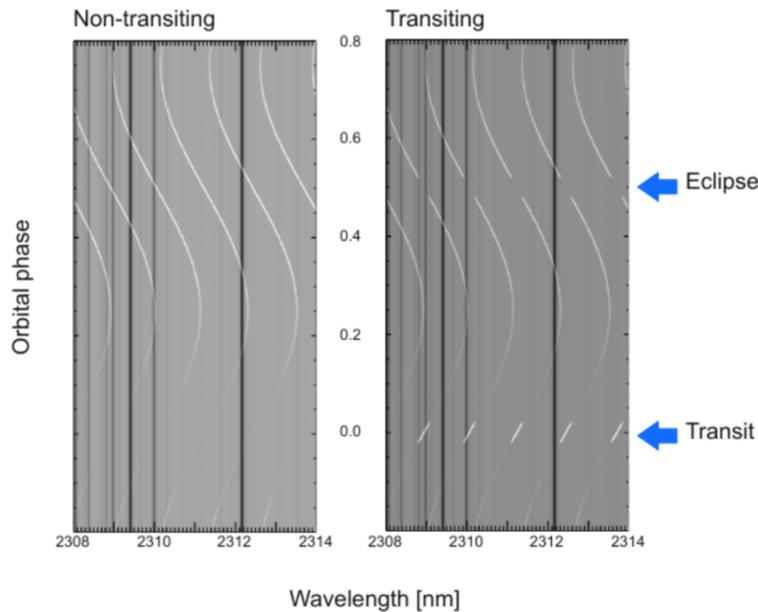


FIGURE 1.28 Modèle de spectres HR d'une JC, qui transite (droite) ou non (gauche), en fonction de la phase orbitale. Les raies sombres sont des raies d'absorption telluriques tandis que les raies claires sont des raies d'émission ou d'absorption (pendant le transit) du CO par l'atmosphère de la planète. On voit que le décalage Doppler des raies de la planète change à mesure que cette dernière orbite son étoile. Le signal est maximal lorsqu'une majorité du côté jour de la planète fait face à l'observateur, et en transmission lors du transit. Figure tirée de Schwarz et al. (2013).

C'est avec la SHD que la vitesse radiale devient importante pour la spectroscopie de transit ou d'émission. Lorsque la planète orbite autour de son étoile, la VR de l'étoile varie, et c'est ce qui est utilisé pour détecter les exoplanètes (Section 1.1.2.1). Par contre, ce qui va nous intéresser ici est la variation en VR de la planète, dont l'ordre de grandeur est donné par la demi-amplitude  $K_p$ . De la même façon que  $K$  est proportionnelle à  $M_p$  pour l'étoile (éq. 1.1.2),  $K_p$  est proportionnelle à  $M_*$  pour la planète. Ceci donne des variations de VR bien plus importantes (typiquement de l'ordre des dizaines de km/s) : on voit les planètes bouger évidemment plus facilement que leur étoile. Ces plus grandes variations en VR vont induire, de façon équivalente, un plus grand décalage Doppler sur les raies générées par l'atmosphère de la planète. En transit, celles-ci seront d'abord décalées vers le bleu à mesure que la planète s'approche, et seront ensuite décalées vers le rouge à mesure que la planète s'éloigne. Puisqu'on peut résoudre les raies individuellement, on peut différencier les raies de la planète, des raies telluriques et de l'étoile, qui, en comparaison, seront essentiellement immobiles sur une échelle de quelques heures, et qu'on pourra plus facilement soustraire. Ce principe est illustré à la Figure 1.28.

D'ailleurs, en combinant la valeur de la vitesse radiale de la planète mesurée ( $K_p$ ) avec les paramètres déjà connus du système ( $K_*$ ,  $i$  et  $P$ ), il est possible de déterminer les masses de

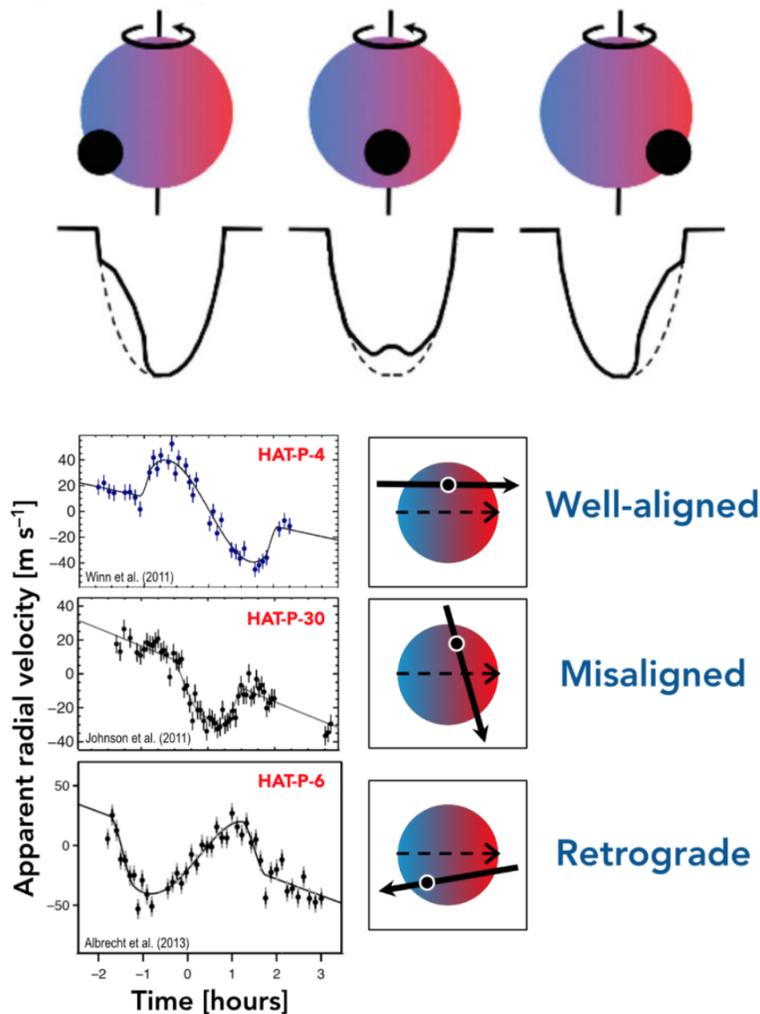


FIGURE 1.29 Visualisation de l'effet Rossiter-McLaughlin (RME) (Haut) Variation de la forme des raies stellaires causée par l'occultation par la planète de zones ayant différentes VR. (Bas) Ces variations mènent à un changement de la VR totale (intégrée sur la surface) perçue, et donneront une forme particulière à la courbe de VR. La géométrie de l'orbite influencera l'allure de cette courbe. La figure du haut est tirée de Collier Cameron et al. (2010) et la figure du bas a été extraite de la page web de Princeton University.

l'étoile et de la planète (équation 1.1.2). La masse de plusieurs planètes a pu être calculée et/ou confirmée de cette façon, ex. :  $\tau$  Boötis b (Brogi et al., 2012) et HD 189733 b, (de Kok et al., 2013a).

En fait, la SHD est très efficace pour déterminer les paramètres orbitaux des planètes grâce à l'effet Rossiter-McLaughlin (RME; Rossiter 1924; McLaughlin 1924). Cet effet apparaît lorsque la planète en transit vient occulter différentes parties de l'étoile qui ont des vitesses

radiales différentes par rapport à l'observateur : puisque l'étoile tourne, les zones masquées seront d'abord décalées vers le bleu, et ensuite vers le rouge (en assumant une orbite prograde et alignée), et qui viendra influencer la VR mesurée de l'étoile (voir Figure 1.29). La forme résultante de la courbe de VR permet alors d'estimer l'angle de spin-orbite (l'angle entre l'axe de rotation de l'étoile et celui de l'orbite de la planète) (ex. : Triaud et al., 2010; Triaud, 2018; Moutou et al., 2020; Allart et al., 2020).

Il est aussi possible d'utiliser la SHD pour déterminer le profil TP, que ce soit en émission (ex. : Brogi & Line, 2019; Gandhi et al., 2019; Pelletier et al., 2021) ou en transmission (particulièrement via le doublet de sodium qui peut être sondé même à plus faible résolution; Sing et al., 2008a,b; Seidel et al., 2020a), puisque la forme des raies est influencée par la structure en température. Comme on peut directement aller sonder la forme et le cœur des raies (à plus faibles pressions) avec la SHD, on peut en faire une détermination plus précise. De la même façon, il est possible d'inférer la présence d'une inversion de température (Vidal-Madjar et al., 2011; Huitson et al., 2012; de Kok et al., 2013b; Schwarz et al., 2015). Il faut cependant être alerte parce qu'un signal d'inversion « impure » (diminution de  $T_P$  avec l'altitude, suivi d'une augmentation de  $T_P$  avec l'altitude) va venir créer un signal ayant des ailes en absorption, mais un cœur en émission. Ce scénario observé sans une précision assez bonne et un signal assez grand pourrait être mépris pour une plus faible abondance, puisque la profondeur de la raie aura l'air plus faible.

L'influence des nuages et des aérosols sur les spectres exoplanétaires a aussi été explorée avec la SHD (Gandhi & Jermyn, 2020; Sánchez-López et al., 2020; Harada et al., 2021). Ils confirment que la SHD permet de voir au-dessus des nuages puisqu'on a accès aux raies directement, qui vont sonder plus haut dans l'atmosphère.

Malheureusement, le phénomène de superposition des signaux stellaire et planétaire reste un problème même à HR. En effet, tel que vu plus haut, le RME va venir déformer les raies d'absorption de l'étoile (Figure 1.29). C'est bien lorsqu'on veut faire des analyses de paramètres orbitaux, mais ce l'est moins lorsqu'on veut mesurer le spectre de l'étoile précisément (surtout si les méthodes courantes impliquent de prendre la moyenne/médiane du signal). Brogi et al. (2016) ont montré l'importance de prendre en compte le RME et de corriger pour les raies stellaires et leur distorsion durant les transits. Dans leur étude, une soustraction du flux hors transit au flux en transit non corrigé mène à un signal parasite ressemblant à une raie de CO (dans le cas de HD 189733 b). Ceci brouille la détection du CO dans le spectre de la planète. Comme pour la SFD, cet effet est surtout important lorsque les éléments chimiques sont à la fois présents dans l'étoile et la planète. Genest et al. (2022) ont aussi montré l'importance de corriger pour le RME, qui vient introduire des signaux parasites dans la carte de corrélation-croisée  $K_P$  versus  $v_{\text{sys}}$  (tel que décrit ci-bas).

Du sol, le signal de l'étoile est affecté par la variation de la transmission atmosphérique terrestre et la variation de la qualité d'image. On ne peut donc pas calibrer de façon absolue le flux de l'étoile pour le soustraire correctement à chaque pose. Plutôt, on doit le renormaliser, mais l'information du niveau du continu spectral est alors perdue. Les variations spectrales du signal causées par la planète seront toutes mesurées relativement au flux reçu de l'étoile à chaque instant, causant ainsi la perte du niveau du continu pour les spectres de transmission. On ne peut donc pas mesurer une « profondeur de transit » à proprement parler : on dit que les données sont « autocalibrées ». On peut se baser sur le contraste des raies, leur amplitude et leur « comportement » général, mais pas sur leur niveau de référence (Brogi & Line, 2019). Ceci mène à des dégénérescences, similaires à celles décrites dans la section précédente. La détermination du rayon planétaire est alors impossible, à moins d'avoir des données de transit SFD additionnelles qui pourraient ancrer les solutions. De plus, Rasmussen et al. (2021) ont montré qu'après normalisation du continu, appliquer différents filtres optimisés pour enlever les effets systématiques à plus basses fréquences permettait d'augmenter davantage l'intensité des détections moléculaires dans l'atmosphère des exoplanètes.

Ces mêmes étapes de réduction/extraction (qui supprime le niveau du continu des spectres HR) viendront principalement affecter les signaux stellaires et telluriques, mais également le signal planétaire. C'est pourquoi il est important de générer une séquence de données synthétique (avec un modèle donné) et d'appliquer les mêmes étapes sur cette séquence que sur les données (ex. : Brogi & Line, 2019; Pelletier et al., 2021; Boucher et al., 2021). Cela permet une meilleure comparaison entre les données observées et le modèle, qui ont subi le même traitement. Un exemple de traitement sur une séquence modèle, ainsi que son effet sur le signal original sont montrés à la Figure 1.30. En ignorant cette étape, un modèle fondamentalement non représentatif de l'atmosphère (parfait et non affecté par l'analyse) pourrait alors être utilisé à tort pour représenter les données (affectées par l'analyse). On penserait que l'autocalibration des données supprime la sensibilité aux abondances absolues, mais Brogi & Line (2019) ont montré que la SHD est surprenamment sensible à celles-ci, même si elle l'est beaucoup plus aux ratios d'abondances. Ils soutiennent que les méthodes de récupération avec la SHD permettraient d'extraire d'excellentes contraintes sur les ratios d'abondances, tels que le C/O et/ou la métallicité.

Une détection d'un spectre de transmission planétaire avec la SHD est très différente de la SFD. On ne peut pas voir le spectre directement comme à basse résolution. Il faut plutôt utiliser des techniques qui vont tenir compte de toutes les raies pour maximiser le signal, comme la corrélation-croisée (CC; les détails de la méthode sont présentés à la Section 1.5). Cette détection est souvent représentée comme une fonction de corrélation-croisée (CCF) en fonction de la phase orbitale (Figure 1.31), mais si le signal n'est pas suffisamment fort pour être vu ainsi, il faut alors combiner le signal de chaque exposition. Il est alors coutume de présenter cette

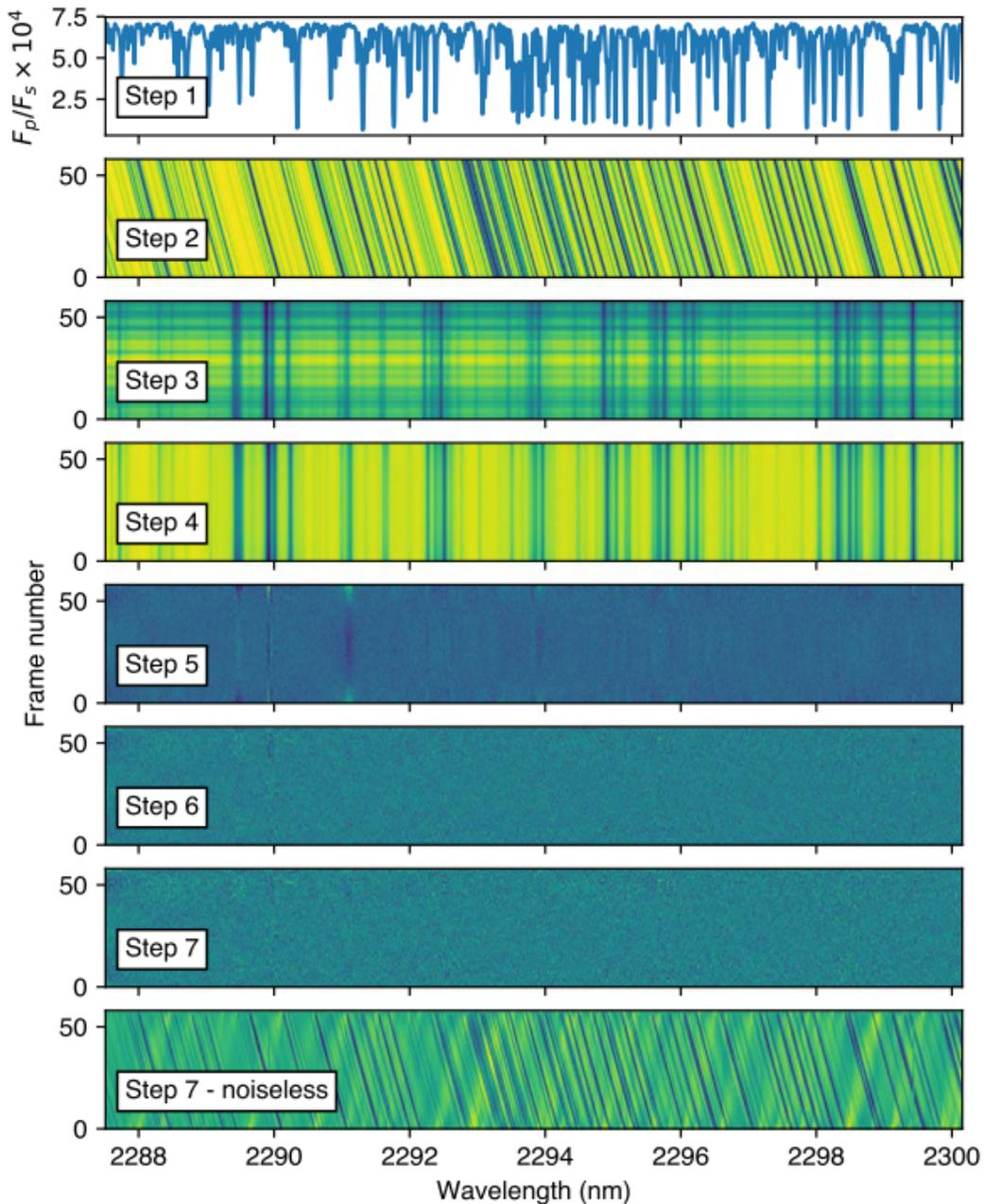


FIGURE 1.30 Visualisation des étapes du processus pour générer une séquence modèle. 1) Spectre du modèle utilisé. 2) Le signal est ensuite décalé à la vitesse de la planète en fonction du temps pour toute la séquence des observations (l'échelle de couleur représente l'intensité du spectre de transmission, l'axe du temps va vers le haut). 3) Le signal précédent est injecté dans les données observées « brutes ». 4) à 7) Application des mêmes étapes de réductions que sur les vraies données. Généralement, le niveau du continu spectral est soustrait (4 et 5), le signal statique dans le temps est enlevé (raies telluriques et stellaires ; 6), et les résidus additionnels sont corrigés (7). 7) représente le produit final, incluant les sources majeures de bruits (astrophysiques et instrumentaux). 8) Produit final, sans bruit. La différence entre 2) et 8) montre l'effet de tout le traitement de données sur le signal de la planète. Figure tirée de Brogi & Line (2019).

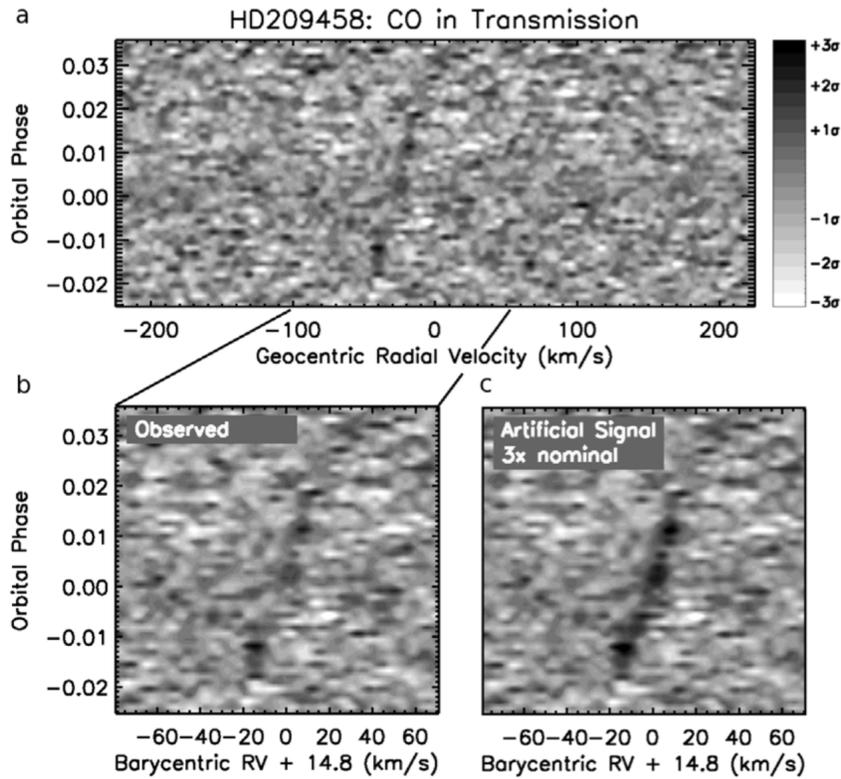


FIGURE 1.31 Signal de corrélation-croisée du CO du spectre de transmission de HD 209458 b en fonction du décalage en vitesse radiale et de la phase orbitale. On voit le signal se « déplacer » à mesure que la planète bouge durant le transit. Le signal observé (a et b) se trouve à la valeur de VR attendue et ressemble au signal simulé (c). Figure tirée de Snellen et al. (2010).

combinaison du signal<sup>17</sup> en fonction des paramètres orbitaux  $K_P$  et de la vitesse systémique du système ( $v_{\text{sys}}$ , ou  $v_{\text{rad}}$  lorsqu'on travaille dans le référentiel de la planète), voir Figure 1.32. Un signal planétaire devrait alors se retrouver à la position  $K_P$  et  $v_{\text{sys}}$  (ou à  $v_{\text{rad}}=0$ ).

La caractérisation atmosphérique par SHD est beaucoup plus récente qu'avec la SFD. La première caractérisation par SHD dans l'IRP a été faite par Snellen et al. (2010). Il s'agit de la détection de CO (à  $\sim 2.3 \mu\text{m}$ ) dans l'atmosphère de HD 209458 b en transmission, avec CRIRES (*Cryogenic Infrared Echelle Spectrograph*; voir Figure 1.31), qui est, comme son nom l'indique, un spectrographe échelle IR ( $\sim 0.92$  à  $5.2 \mu\text{m}$ , mais limité à  $\sim 15 \text{ nm}$  en simultané), HR ( $R \sim 100\,000$ ), et qui est installé au *Very Large Telescope* (VLT; 8.2 m).

Depuis cet important résultat, les détections faites par SHD ont explosé, et ce, grâce à une panoplie de spectrographes HR, ayant tous de très larges plages spectrales en simultané. Parmi ceux dans l'IRP, on retrouve NIRSPEC ( $0.95\text{--}5.4 \mu\text{m}$ , où la couverture en simultanée a à peu près la largeur d'une bande atmosphérique,  $R \simeq 25\,000$ ; McLean et al. 1998), un spectrographe

17. Combiner le signal à différents  $K_P$  se traduit par changer l'angle de combinaison du signal (sur la Figure 1.31), tandis que pour  $v_{\text{sys}}$ , cela se traduit par un décalage à gauche ou à droite.

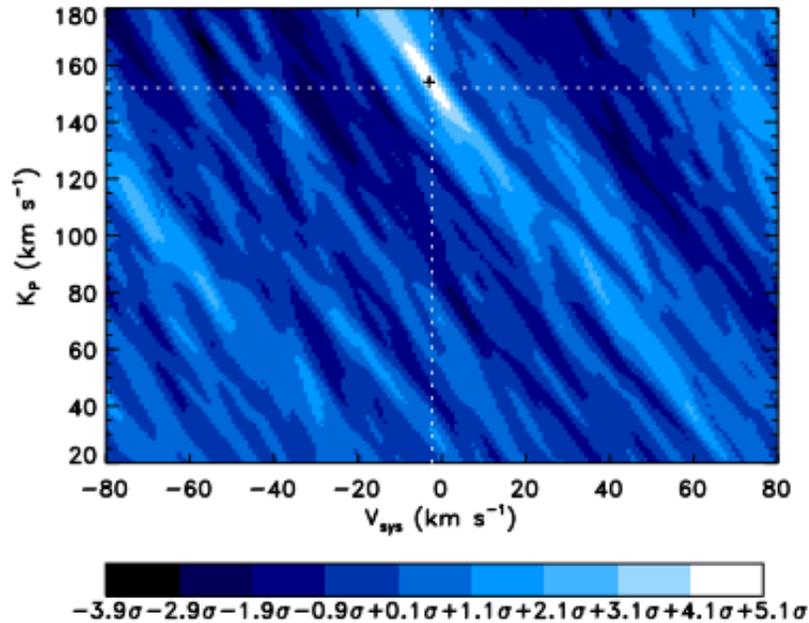


FIGURE 1.32 Signal total de CC de l'émission de  $\text{H}_2\text{O}$  au côté jour de HD 189733 b. L'échelle de couleur montre l'intensité du signal de  $\text{H}_2\text{O}$  en absorption, en fonction de  $v_{\text{sys}}$  sur l'axe des x et  $K_p$  sur l'axe des y. On voit que le signal est trouvé à la position attendue (au croisement des lignes blanches pointillées). Figure tirée de Birkby et al. (2013).

installé au télescope Keck II de 8-m; CARMENES, ayant un canal dans l'IRP ( $0.96\text{--}1.71\ \mu\text{m}$ ,  $R \simeq 80000$ ; Quirrenbach et al. 2014) installé à l'observatoire de Calar Alto, sur le télescope de 3.5-m; et GIANO ( $0.95\text{--}2.45\ \mu\text{m}$ ,  $R = 50000$ ; Oliva et al. 2006), un spectrographe monté sur le Telescopio Nazionale Galileo, de 3.6-m.

Le dernier, mais non le moindre : SPIRou, le Spectro-Polarimètre InfraRouge ( $\sim 0.95\text{--}2.50\ \mu\text{m}$ ,  $R \sim 70000$ ; Donati et al. 2020). Comme son nom l'indique, il s'agit d'un spectro-polarimètre échelle alimenté par fibre installé au Télescope Canada-France-Hawaii (CFHT, 3.6-m). Les spectrographes échelle (comme ceux présentés ci-haut) permettent d'aller chercher des ordres de diffraction plus grands. Avec la dispersion croisée, les hauts ordres de diffraction qui habituellement se chevaucheraient peuvent ainsi être isolés. Ces différents ordres de diffractions sont montrés sur la Figure 1.33. Finalement, on peut rapidement ajouter à cette liste quelques contributions des instruments IGRINS ( $1.45\text{--}2.5\ \mu\text{m}$ ,  $R \lesssim 45000$ ; Flagg et al., 2019; Line et al., 2021), et IRD ( $0.97\text{--}1.75\ \mu\text{m}$ ,  $R \lesssim 70000$  Nugroho et al., 2021) au télescope Subaru (8-m).

Un des grands résultats apportés par la SHD est la possibilité de mesurer la dynamique des atmosphères (Showman et al., 2012). Être capable de résoudre les raies nous donne accès non seulement à leur position, mais aussi à leur forme et à leur élargissement. Les variations de

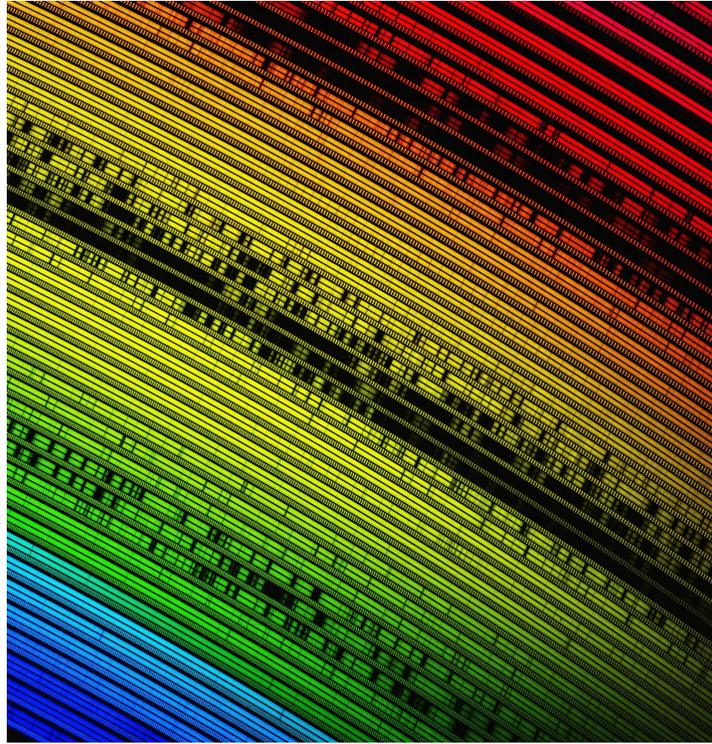


FIGURE 1.33 Image infrarouge monochrome obtenue par le spectrographe SPIRou, colorisée du bleu au rouge selon les différentes longueurs d'onde, allant de 1 à 2.5  $\mu\text{m}$ . Pour chaque ordre spectral, il y a deux spectres (provenant des deux fibres « science ») et un spectre Fabry-Perot (de la fibre de calibration). Les grandes zones sombres sont les bandes de raies telluriques. Un peu partout dans le spectre on peut apercevoir d'autres raies telluriques, mais aussi des raies stellaires, tandis que le signal de la planète, si présent, est invisible. Crédit : Étienne Artigau, présenté au concours de l'ACFAS « La preuve par l'image ».

forme et d'élargissement seront créées par la structure en température (ex. : Beltz et al., 2021) et la dynamique de l'atmosphère (les mouvements globaux du gaz).

Tel qu'établi à la Section 1.3.5, la dynamique des atmosphères se manifeste sous plusieurs formes. D'abord, les flots globaux de gaz allant du côté jour au côté nuit (vents à l'échelle de la planète) généreront un décalage vers le bleu du signal. On peut déterminer la vitesse des vents en mesurant les déviations par rapport à la VR orbitale théorique de la planète ( $v_{\text{rad}} \neq 0$ ). De tels décalages ont été observés dans le signal atmosphérique de plusieurs exoplanètes (ex. : Snellen et al., 2010; Seidel et al., 2021), mais sont compatibles avec les prédictions et les résultats des modèles GCMs (Miller-Ricci Kempton & Rauscher, 2012). La présence de vents verticaux en haute atmosphère a même été décelée par Seidel et al. (2021), par l'extraction de la forme de la raie du Na de WASP-76 b. Ensuite, la rotation planétaire et de la super-rotation équatoriale (courant-jet vers l'est) de HD 189733 b ont pu être résolues spatialement en allant chercher des mesures différentielles des parties jour et nuit du terminateur séparément (Louden & Wheatley, 2015, avec HARPS). En plus de confirmer la présence d'un vent jour-à-nuit dans

l’atmosphère de HD 189733 b, d’autres études ont réussi à mettre des contraintes sur la vitesse de rotation de la planète (ex. : Brogi et al., 2016; Flowers et al., 2019, les deux montrant que la rotation de HD 189733 b est cohérente avec une rotation synchrone). Flowers et al. (2019) ont utilisé des modèles GCM en 3D, qui viennent capturer les subtilités de la dynamique sur l’ensemble du terminator (au lieu d’en assumer les conditions constantes partout). En fait, leur utilisation permet d’enlever le décalage vers le bleu qui apparaît dans le signal atmosphérique lorsque des modèles 1D qui ne considèrent pas la dynamique sont utilisés (tel qu’illustré par la Figure 1.34). La conclusion tirée par Beltz et al. (2021) est quelque peu différente. Leur analyse de données CRIRES en émission pour HD 209458 b suggèrent que la largeur des raies ne peut être directement liée à la rotation de la planète seulement (étant donné la présence de vents). Ils ont toutefois montré que les spectres HR sont sensibles aux structures et à la dynamique 3D de l’atmosphère, que l’utilisation de modèles 3D améliore les détections, mais qu’ils étaient principalement sensibles à la structure en température, et secondairement à la chimie et à la dynamique. La présence de nuages vient causer davantage de problèmes, même avec des modèles 3D (Harada et al., 2021). Wardenier et al. (2021) ont montré que des asymétries dans les signaux (par exemple, provenant d’un signal dominé par le côté soir du terminator) pouvaient également mener à des déplacements du pic de CCF, et parfois de façon importante, dans les cartes  $K_P$  versus  $v_{\text{sys}}$ . Les déplacements du pic (surtout les décalages vers le bleu), autrement que de provenir de vents jour-à-nuit, ont aussi été attribués à des erreurs de  $T_0$ , de  $K_P$ , ou autres paramètres orbitaux (Brogi & Line, 2019; Line et al., 2021).

La SHD a aussi été prolifique dans la détection de divers éléments chimiques. On compte évidemment la détection de  $\text{H}_2\text{O}$  dans plusieurs cibles, et utilisant les instruments mentionnés ci-haut. Entre autres, du  $\text{H}_2\text{O}$  a été détecté dans 51 Peg b, HD 179949 b, HD 189733 b, HD 209458 b (Birkby et al., 2017; Brogi et al., 2014; Webb et al., 2020; Birkby et al., 2013; Brogi et al., 2018; Hawker et al., 2018; Alonso-Floriano et al., 2019), et dans d’autres planètes qui ne transitent pas, comme HD 88133 b, et HD 102195 b (Guilluy et al., 2019; Piskorz et al., 2018). Le CO a surtout été détecté en émission du côté jour, par exemple, sur  $\tau$  Boötis b (Brogi et al., 2012; Pelletier et al., 2021), sur la très jeune CL Tau b (Flagg et al., 2019), la JUC WASP-33 (Yan et al., 2022a), mais aussi sur HD 189733 b (de Kok et al., 2013a; Rodler et al., 2013), seul, ou avec de l’ $\text{H}_2\text{O}$  (Brogi et al., 2016; Flowers et al., 2019), ou encore la combinaison  $\text{H}_2\text{O} + \text{CO} + \text{HCN}$  (Cabot et al., 2019). Le HCN et la chimie de l’azote ont aussi été détectés ou confirmés quelques fois, comme sur HD 209458 b (ex. : MacDonald & Madhusudhan, 2017; Hawker et al., 2018; Gandhi et al., 2020). Un inventaire assez exhaustif de la composition de HD 209458 b a été compilé par Giacobbe et al. (2021), qui ont montré la présence de 5 molécules carbonées et azotées :  $\text{H}_2\text{O}$ , CO, HCN,  $\text{CH}_4$ ,  $\text{NH}_3$ , et  $\text{C}_2\text{H}_2$ , ce qui représente une première en SHD pour ces trois dernières. Une tentative de détection de  $\text{CH}_4$  avait précédemment été faite par Guilluy

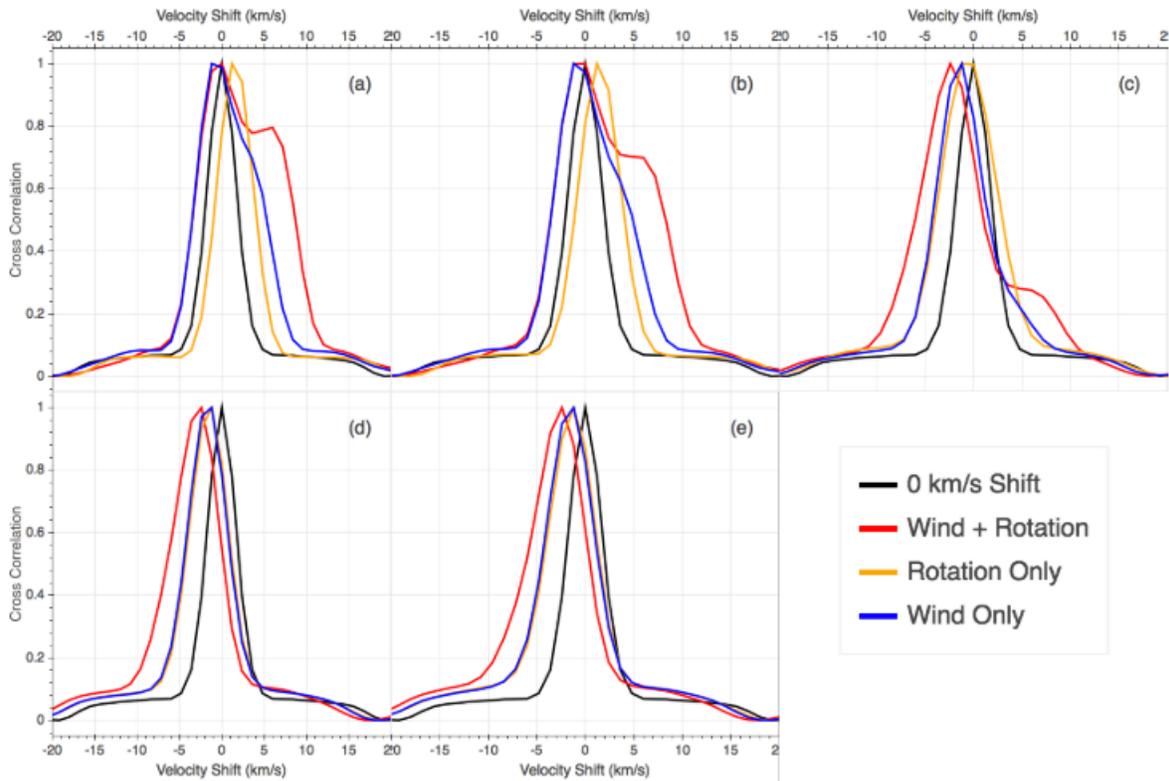


FIGURE 1.34 CCF entre un modèle d’atmosphère stationnaire (en rotation synchrone) avec un modèle stationnaire (noir), un modèle qui inclus la rotation de la planète et un vent jour-à-nuit (rouge), seulement de la rotation (jaune), et seulement un vent jour-à-nuit (bleu). De gauche à droite, on a le cas a) au début de l’entrée en transit, b) au milieu de l’entrée en transit, c) au milieu du transit, d) milieu de la sortie et e) à la fin de la sortie du transit. On voit clairement la dépendance de la forme de la CCF avec la position durant le transit (et donc des régions plus ou moins illuminées par l’étoile à différents moments). On voit aussi à quel point la CCF d’un modèle peut être affectée par l’utilisation d’un modèle stationnaire au lieu d’un modèle plus complexe (3D qui inclut la dynamique). Figure tirée de Flowers et al. (2019).

et al. (2019), mais elle a été réfutée par Gandhi et al. (2020), qui ont jugé coupable la liste de raies d’absorption utilisée pour générer leur modèle.

En effet, pour détecter des signaux moléculaires dans des données HR, il faut des modèles qui ont une résolution encore plus grande (question d’avoir assez de précision pour contrôler différents aspects comme le décalage Doppler et le profil instrumental). La position des raies et leur intensité doivent être extrêmement précises si on veut être capable de retrouver le signal voulu. Autrement dit, si le modèle n’est pas suffisamment précis, voir erroné, on ne trouvera jamais le signal même si la molécule est présente. Malheureusement, les listes de raies d’absorption ne sont pas toutes créées également et ne sont pas parfaites. Certaines sont basées sur des données expérimentales, d’autres sont calculées numériquement, et il est difficile de déterminer quelles listes de raies sont optimales, d’autant plus que leur capacité à reproduire les observations semble changer d’un jeu de données à un autre (Brogi & Line, 2019; Gandhi et al., 2020; Webb et al.,

2020). Il faut donc être vigilant lorsqu'on utilise différentes listes de raies : les résultats et les conclusions qu'on tire de ces analyses seront toujours, au mieux, aussi précises que le modèle utilisé.

Outre ces molécules, il est devenu habituel de faire la détection de He métastable à 1080.3 nm avec la SHD, surtout avec CARMENES (Allart et al., 2018; Salz et al., 2018; Allart et al., 2019; Alonso-Floriano et al., 2019; Zhang et al., 2020), mais également avec NIRSPEC (Kirk et al., 2020), GIARPS (la combinaison de GIANO et HARPS-N; Guilluy et al., 2020), et SPIRou (Allart et al. 2022, en préparation).

À cette liste de détections déjà bien grande, on ajoute celle des métaux alcalins et des métaux plus lourds et leurs ions (comme le Fe, Ti, Cr, Ca, etc.), qui sont observables surtout dans le visible. Les instruments souvent utilisés sont : HARPS et HARPS-N (les deux ayant  $\sim 0.38\text{--}0.69\ \mu\text{m}$ ,  $R \simeq 115\,000$ ), ESPRESSO ( $\sim 0.38\text{--}0.69\ \mu\text{m}$ ,  $R \simeq 200\,000$ ), le canal visible de CARMENES ( $0.52\text{--}0.96\ \mu\text{m}$ ,  $R \simeq 100\,000$ ), UVES au VLT ( $0.3\text{--}1.1\ \mu\text{m}$ ,  $R \simeq 80\,000$  à  $110\,000$ ), et Subaru HDS ( $0.3\text{--}1.0\ \mu\text{m}$ ,  $R \lesssim 160\,000$ ; Nugroho et al., 2017). Ce sont surtout des recherches de Na qui sont faites (ex. : Wyttenbach et al., 2015; Chen et al., 2020; Seidel et al., 2020a; Allart et al., 2020; Khalafinejad et al., 2021; Seidel et al., 2021), mais certaines ciblent également le K (ex. : Keles et al., 2020). La première détection de Na dans HD 209458 b de Charbonneau et al. (2002) a aussi potentiellement été infirmée par le fait que le RME et la variation centre-bord (CLV; de l'anglais *Center-to-Limb Variation*, et qui représente la baisse luminosité de la photosphère d'un disque stellaire sur le limbe, comparativement au centre du disque, c.-à-d. la conséquence de l'assombrissement centre-bord, mentionné à la Section 1.1.2.2) n'avaient pas été pris en compte dans la soustraction de signal (Casasayas-Barris et al., 2020). Cela montre bien l'importance de prendre en compte ces effets.

Finalement, les métaux lourds (particulièrement le Fe I et le Fe II; Gibson et al. 2020; Cabot et al. 2021; Gibson et al. 2022) ont surtout été observés sur les JUCs. C'est Hoeijmakers et al. (2018) qui ont parti le bal avec la détection de Fe I, Fe II et Ti II dans l'atmosphère de KELT-9 b, l'exoplanète la plus chaude connue à ce jour ( $T_p \sim 4050\text{ K}$ ), et qui est en train de perdre son atmosphère (échappement observé par la raie de  $H\alpha$ , Yan & Henning, 2018). Des études subséquentes ont montré la présence de Mg I, Fe I, Fe II and Ti II, et des tentatives de détections de Ca I, Cr I, Co I, et de Sr II (Hoeijmakers et al., 2019). De façon similaire, KELT-20 b (aussi connue sous le nom de MASCARA-2 b; Talens et al. 2018) possède, au minimum, les éléments suivants : Fe I, Fe II, Cr I, Cr II, Na I, Mg I, Ca II (Stangret et al., 2020; Nugroho et al., 2020; Hoeijmakers et al., 2020; Borsa et al., 2022; Yan et al., 2022b), et H ( $H\alpha$ ,  $H\beta$ , et  $H\gamma$ ; Casasayas-Barris et al. 2019).

WASP-33 b, une autre JUC bien connue, possède, elle aussi, du Na (Cont et al., 2022), du Fe I (Nugroho et al., 2020; Herman et al., 2022), et du H (Yan et al., 2021), et est une des rares planètes pour laquelle une détection robuste de TiO a été faite (Nugroho et al., 2017; Cont

et al., 2021). Elle partage aussi un élément chimique « rare » avec KELT-20 b, le FeH. Kesseli et al. (2020) ont effectué une recherche du FeH avec CARMENES sur 12 exoplanètes, mais n’ont trouvé de signal faible que pour ces deux cibles. Ils ont établi qu’il faut avoir une forte gravité de surface et être assez chaud ( $>1800$  K), mais pas trop ( $<3000$  K) pour être capable d’en détecter.

Un autre élément difficile à détecter est le radical hydroxyle, le OH. Celui-ci a été détecté pour la première fois dans l’atmosphère très riche de WASP-33 b (Nugroho et al., 2021). La seule autre détection a été faite sur WASP-76 b (Landman et al., 2021). Le OH serait le produit de la dissociation de l’eau, mais cela prend de hautes températures pour que le produit se fasse thermochimiquement ( $T_p \gtrsim 2500$  K; Parmentier et al. 2018). La photochimie, dans un environnement très irradié, pourrait cependant en produire en grande quantité (Eistrup et al., 2018).

Finalement, dans l’atmosphère de WASP-76 b, on a détecté du Na (Žák et al., 2019; Seidel et al., 2019; Deibert et al., 2021b), de l’He et du Ca II (Casasayas-Barris et al., 2021), en plus du Fe I (Ehrenreich et al., 2020). Cette détection de Fe I est assez particulière, puisque le signal de CCF (en fonction de la phase orbitale) est très asymétrique. Cette asymétrie proviendrait du fait que le gaz de Fe I est chaud et décalé vers l’est du côté jour, mais condensé du côté nuit, ce qui donnerait une forme particulière au signal observé. Ce résultat a été contesté par certains qui prône plutôt la présence de nuages opaques d’autres éléments, plutôt que provenant de la condensation du Fe (Savel et al., 2022), ou partiellement appuyé par d’autres (Wardenier et al., 2021, qui montrent que la condensation peut expliquer les résultats, mais n’est pas essentielle, puisque les résultats peuvent aussi être expliqués par un fort gradient de température entre les deux côtés du terminateur). Dans tous les cas, l’asymétrie a été confirmée (Kesseli & Snellen, 2021; Seidel et al., 2021). Taberero et al. (2021) ont également fait la première détection robuste de Li I, en plus du Na I, Mg I, Ca II, Mn I, K I, et Fe I, avec ESPRESSO.

### 1.4.3. Analyses Conjointes

La contribution principale des données SFDs est la détection du niveau du continu et des caractéristiques spectrales les plus larges des divers éléments chimiques (bandes d’absorption moléculaires, doublets alcalins, diffusion des aérosols, etc.). Par contre, si ces bandes d’absorption se chevauchent ou si des nuages sont présents, des ambiguïtés surgissent lorsqu’il est temps d’identifier les éléments présents et de calculer leur abondance. Comparativement, les données SHD permettent d’accéder directement aux structures fines des spectres, soit les raies d’absorption (des ailes jusqu’au cœur; qui sondent des pressions plus basses aux pressions plus hautes), via leur position et leur contraste/amplitude. Ceci se fait au prix de perdre le niveau absolu du continu (via les processus de réduction/analyse des données), nécessaire pour calculer les abondances absolues. C’est tout de même la meilleure méthode pour identifier les éléments présents

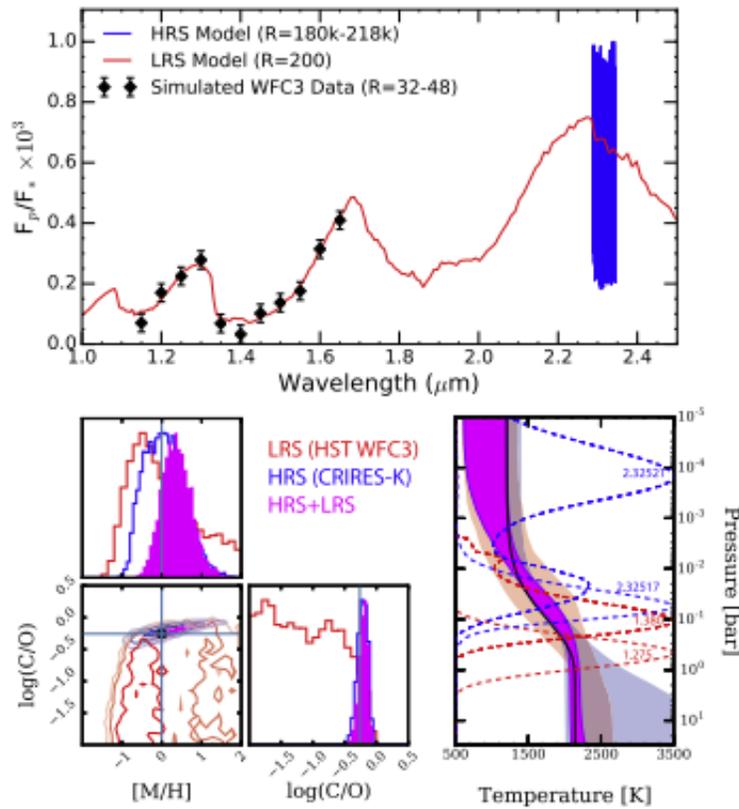


FIGURE 1.35 (Haut) Observations simulées de HST/WFC3 (noir) et VLT/CRIRES (bleu). (Bas) Distributions postérieures de la métallicité ( $[M/H]$ ), du ratio C/O (gauche) et les intervalles de confiance (68%) du profil TP (droite); le tout obtenu par la méthode de récupération en utilisant seulement les données HST (rouge), VLT (bleu) et combiné (magenta). Les lignes en tiret sont les courbes de contributions pour les longueurs d'onde indiquées : les données HST sondent plus bas dans l'atmosphère et sur une région plus étroite, tandis que les données VLT ont une plage dynamique beaucoup plus grande, ce qui permet de sonder un plus large intervalle de pression. L'analyse combinée améliore grandement la précision des mesures. Figure tirée de Brogi & Line (2019).

de façon claire et sans ambiguïté. Ces limitations sont toutefois complémentaires et combiner les méthodes peut maximiser l'utilisation de l'information comprise dans chacune, et ainsi aider à lever certaines dégénérescences.

Brogi et al. (2017) ont présenté la première étude conjointe de SFD et SHD, en combinant des données en émission de HST (WFC3), Spitzer, et CRIRES de HD 209458 b. Ils ont montré que cette combinaison améliore les contraintes sur la structure en température et les abondances. Les ambiguïtés causées par le chevauchement de certaines bandes (comme le CO et le CO<sub>2</sub> dans IRAC2) peuvent être levées par les données SHD qui vont chercher le détail de certaines raies appropriées (assumant un signal suffisant). Ils notent toutefois que leur méthode est lente et qu'une diminution du temps de calcul est nécessaire pour pouvoir appliquer des algorithmes de récupération complets (voir Section 1.5).

Beaucoup de ces études conjointes ont été faites dans le domaine visible, focalisant sur la détection du Na. Pino et al. (2018) ont développé une méthode d'analyse conjointe pour étudier HD 189733 b avec des données de HST et HARPS. Ils réussissent à trouver un modèle qui réconcilie les deux ensembles de données simultanément, lesquels semblaient pourtant pointer vers deux scénarios différents (un signal sans structures en SFD versus un signal avec des structures prononcées en SHD). Ils ont également montré que la SHD est sensible à la présence des aérosols.

Khalafinejad et al. (2021) ont étudié l'atmosphère de WASP-69 b avec des données CARMENES combinées aux données d'archives de HARPS-N, OSIRIS/GTC et WFC3/HST. Ils ont ainsi pu utiliser l'information du continu de la SFD comme données à priori pour déterminer l'abondance de Na détecté, la pression de référence et la température, et conclure qu'un scénario d'équilibre thermodynamique permet d'expliquer leurs observations, malgré la présence d'aérosols qui diminuent l'intensité de certaines structures.

Enfin, Brogi & Line (2019) ont montré la possibilité d'appliquer des méthodes de récupération de paramètres (de l'anglais *retrieval framework*; par exemple, en utilisant l'inférence Bayésienne, et une méthode d'échantillonnage comme la méthode Monte-Carlo par chaînes de Markov) à des ensembles de données combinés de SHD et SFD. Ils ont développé une méthode de calcul qui est statistiquement plus appropriée pour ce genre d'étude, et qui considère intrinsèquement le poids des différents ensembles de données (un problème qui avait été relevé dans Brogi et al. (2017)). Leurs simulations montrent que la détermination des ratios d'abondances est à peine meilleure que celle basée uniquement sur la SHD (parce qu'elle y est déjà très sensible), mais que la combinaison de SHD avec SFD améliore la détermination du profil TP (la SHD permettant une meilleure précision à basse pression, et la SFD à plus haute pression, voir Figure 1.35). De plus, leur étude a montré que l'absence de certains gaz clés pouvait biaiser les résultats, et/ou élargir artificiellement les contraintes.

On peut donc conclure que de combiner les données à faible et haute résolution permet grandement d'améliorer les contraintes sur les paramètres atmosphériques. Cela permet réellement d'avoir une meilleure vue d'ensemble sur les conditions présentes.

## 1.5. Détermination des propriétés

Une fois le spectre planétaire obtenu, soit à faible ou haute résolution, l'étape suivante est d'en extraire les informations. Il y a plusieurs façons de s'y prendre. Dans la plupart des cas, des modèles atmosphériques seront nécessaires. Ceux-ci sont construits avec les principes de base de la physique et de la chimie atmosphérique (voir Section 1.3), de manière exacte ou paramétrée et peuvent être autocohérents ou non. Les modèles autocohérents (*self-consistent*) vont vérifier (souvent par méthode itérative) que la physique et la chimie sont respectées pour des

valeurs de paramètres donnés (abondances, profil TP, etc.), et ce, dans toute l’atmosphère. Ce sont ces modèles qui vont habituellement inclure le plus de détails. En comparaison, les modèles qui ne sont pas autocohérents ne font pas cette vérification ; il n’y a pas de rétroaction pour vérifier si les conditions (pour un ensemble de paramètres donnés) sont physiquement possibles. Par exemple, on peut paramétrer les profils TP, ou encore demander que certaines équations soient satisfaites seulement à certains endroits dans l’atmosphère. Certains processus peuvent être traités de manière autocohérente tandis que d’autres non. Pour simplifier le problème, les atmosphères seront souvent modélisées en 1D, en couches parallèles superposées (modèle *plan parallèle*). Les interactions entre les couches et les conditions frontières vont permettre de déterminer la solution. Avec ces modèles, l’état de l’atmosphère est choisi et spécifié par une combinaison de paramètres : cela suppose que cet état est connu complètement et sans incertitudes. Par contre, beaucoup d’approximations sont souvent faites (ex. : équilibre chimique, équilibre radiatif, abondances solaires, etc.), faute de connaître les conditions exactes d’avance – c’est ce qu’on veut trouver – et par limitations en puissance de calcul. Il y a tout de même des méthodes pour aller chercher ces incertitudes (voir ci-bas).

Il existe deux grandes approches pour mesurer la qualité d’un modèle. Une sera pour visualiser et quantifier une détection, tandis que l’autre sera plutôt pour faire de la sélection de modèles et mettre des contraintes sur les paramètres.

Pour la quantification des détections en SHD (et concernant surtout les détections moléculaires), les méthodes utilisées sont la CC, le test-*t* et le  $\Delta$ BIC, et une brève description de chacune sera présentée ici. Tel que mentionné brièvement à la section 1.4.2, la CC est une mesure de la similarité de deux signaux en fonction de leur déplacement l’un par rapport à l’autre (ici le déplacement est déterminé par le décalage Doppler). La forme générale de la CCF est la suivante :

$$\text{CCF}(\boldsymbol{\theta}) = \sum_{i=1}^N f_i \cdot m_i(\boldsymbol{\theta}), \quad (1.5.1)$$

où  $f$  et  $m(\boldsymbol{\theta})$  représentent les données observées et le modèle, respectivement, et pour des paramètres de modèles  $\boldsymbol{\theta}$  donnés (qui peuvent inclure les paramètres atmosphériques et dynamiques, dont le décalage Doppler). L’indice  $i$  désigne les longueurs d’onde et la somme est faite sur les  $N$  points des données. La CCF de plusieurs spectres peut également être sommée pour avoir la CCF totale, par exemple pour l’ensemble des spectres acquis lors d’un transit. Il est possible d’utiliser des versions dérivées/pondérées de cette équation, où un poids est assigné aux données ou au modèle, ou aux deux.

Cette technique permet de combiner le signal sur l’ensemble du spectre, et donc de toutes les raies en même temps. Dans notre étude, les deux signaux utilisés sont les spectres de transmission planétaires observés et synthétiques (modèles). Plus les deux signaux se ressembleront,

plus la valeur de CC sera grande. Donc si une molécule est présente dans les données et que l'on compare avec un modèle qui l'inclut, un pic positif apparaîtra à la vitesse de la planète dans la CCF. On aura un pic négatif (creux) si le signal est similaire, mais de signe inverse, comme ce serait le cas avec des raies en émissions, mais comparées à un modèle en transmission (ou vice-versa). Des exemples de CCF sont présentés sous différents formats aux figures 1.31 (2D, séquence temporelle en fonction de la VR), 1.32 (2D, signaux combinés en fonction des paramètres orbitaux), et 1.34 (1D, en fonction de la VR). Dans un cas parfait, un pic apparaît seulement à la position de la planète et la CC est zéro partout ailleurs (comme la courbe noire de la Figure 1.34), mais en réalité, il y a du bruit, qui vient cacher/brouiller le signal et qui peut introduire d'autres pics et creux à différentes vitesses (Figures 1.31 et 1.32). Pour déterminer à quel point le signal qui nous importe (celui de la planète) se démarque des autres pics de bruit, on peut calculer le S/N de la CCF totale (en 1D ou en 2D ; Snellen et al. 2010; Brogi et al. 2012; Rodler et al. 2013; Birkby et al. 2013; Brogi et al. 2018; Cabot et al. 2019; Alonso-Floriano et al. 2019; Sánchez-López et al. 2019; Gibson et al. 2020; Webb et al. 2020; Pelletier et al. 2021; Boucher et al. 2021). Il est habituel de comparer le pic de la CCF totale avec la déviation standard (DST) du bruit, c.-à-d. des valeurs de CC loin du pic (pour ne pas biaiser la DST vers de hautes valeurs artificiellement).

La méthode du test- $t$  (aussi appelé test de Student ; Student 1908) est un test statistique qui vérifie l'hypothèse nulle ( $H_0$ ) que deux distributions gaussiennes aient la même valeur moyenne. Il existe aussi le test- $t$  de Welch (Welch, 1947, une version généralisée du test de Student), plus appropriée lorsque les deux distributions ont des variances et/ou un échantillonnage différents. Dans notre étude, les deux distributions considérées proviennent de notre carte de CCF. La première est la distribution des valeurs de CC qui suivent la trace de la planète, à l'intérieur d'un petit intervalle de vitesse donné, et la seconde est celle à l'extérieur de cet intervalle, loin de la trace de la planète. Le test- $t$  évalue ensuite la probabilité (en échelle de  $\sigma$ ) que ces deux échantillons proviennent de la même distribution « mère ». Un fort signal planétaire se traduit par une grande valeur de test- $t$ , qui nous indique à quel niveau on est loin de  $H_0$ , ce qui représente par définition le niveau de notre détection. Ce test a été employé maintes fois (ex. : Birkby et al., 2013; Birkby et al., 2017; Brogi et al., 2018; Cabot et al., 2019; Alonso-Floriano et al., 2019; Sánchez-López et al., 2019; Webb et al., 2020; Boucher et al., 2021; Giacobbe et al., 2021).

Cependant, ces deux métriques comportent leur lot d'ambiguïtés, tel que pointé par Cabot et al. (2019). En effet, les valeurs absolues du S/N et du test- $t$  ( $\sigma$ ) sont plutôt arbitraires, et peuvent facilement être ajustées. L'étendue et l'échantillonnage (via le sur- ou le sous-échantillonnage) des vitesses testées (autant pour  $v_{\text{sys}}$  que pour  $K_P$ ), l'étendue des régions exclues autour du pic ou encore des régions considérées à l'intérieur ou à l'extérieur de la trace de la planète, un calcul sur une carte 1D ou 2D, sont tous des facteurs qui feront varier les valeurs de S/N et  $\sigma$ . De se fier simplement à ces valeurs peut mener à un sur-ajustement et introduire des

biais dans les résultats. À défaut d’avoir une métrique moins arbitraire, il est préférable d’utiliser ces dernières comme indicateur d’une détection significative ou non, sans trop s’attarder sur la valeur précise de celle-ci.

La troisième métrique est le critère d’information bayésien (BIC ; de l’anglais *Bayesian Information Criterion*), qui est une mesure de « l’information extraite » par un modèle donné (utilisée dans Boucher et al., 2021, au Chapitre 2). On le calcule comme suit,  $BIC = k \ln(n) - 2 \ln \mathcal{L}$ , où  $k$  est le nombre de paramètres,  $n$  est le nombre de points de données et  $\ln \mathcal{L}$  est la log-vraisemblance de notre modèle, et le meilleur modèle est associé à la plus petite valeur de BIC. Pour faire de la sélection ou de la comparaison de modèles, on calcule le  $\Delta BIC$ . Pour des valeurs de  $k$  et  $n$  fixes, cela se traduit alors comme  $\Delta BIC = 2\Delta(\ln \mathcal{L}) = 2(\ln \mathcal{L} - \ln \mathcal{L}_0)$ , où  $\ln \mathcal{L}_0$  est la log-vraisemblance du meilleur modèle. On peut alors « très fortement » rejeter les modèles ayant des  $\Delta BIC > 10$  (Kass & Raftery, 1995), fortement les rejeter avec  $\Delta BIC \sim 6$ , tandis qu’un  $\Delta BIC \lesssim 2$  n’est pas digne de mention.

Mais pour calculer le  $\Delta BIC$ , on aura de besoin de calculer  $\ln \mathcal{L}$ . La log-vraisemblance (de l’anglais *log-likelihood*), dérivée de la vraisemblance, est souvent utilisée dans l’inférence statistique pour faire l’estimation de paramètres. La fonction de vraisemblance décrit la fonction de densité de probabilité d’observer un résultat (les données), étant donné un ensemble de paramètres donné (mesure de la vraisemblance qu’un ensemble de paramètres puisse générer les observables). Fondamentalement, et tout comme la CCF, c’est une mesure de l’ajustement des paramètres aux données.

Zucker (2003) a dérivé un formalisme pour transposer une valeur de CCF à une valeur de  $\ln \mathcal{L}$ , en partant de la définition du  $\chi^2$ , ayant souvent une forme similaire à la suivante :

$$\chi^2 = \sum \frac{(f_i - m_i)^2}{\sigma_i^2}, \quad (1.5.2)$$

où la somme sur  $i$  et la dépendance de  $m_i$  sur  $\theta$  sont implicites, et  $\sigma$  est l’incertitude des données  $f$ . Leur approche servait à combiner le signal de plusieurs ensembles de données en VR de façon cohérente.

Cette idée a été reprise par Brogi & Line (2019), qui ont développé une nouvelle forme de transposition CCF-à- $\ln \mathcal{L}$ . Ils sont partis d’une fonction de vraisemblance gaussienne de la forme :

$$\begin{aligned}
\mathcal{L} &= \prod_i \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{[f_i - am_i]^2}{2\sigma_i^2}\right) \\
&= \left(\frac{1}{\sqrt{2\pi\sigma_i^2}}\right)^N \exp\left(-\sum_i \frac{[f_i - am_i]^2}{2\sigma_i^2}\right),
\end{aligned} \tag{1.5.3}$$

où  $\sigma_i$  est la déviation standard (ou plus généralement l'incertitude) de  $f_i$  et  $a$  est un facteur d'échelle pour le modèle (qu'on fixe généralement à 1, mais pas toujours, ex. : Gibson et al. 2020; Nugroho et al. 2021). La log-vraisemblance peut alors être calculée de cette dernière équation :

$$\ln \mathcal{L} = -N \ln \sigma - \frac{1}{2\sigma^2} \sum [f_i - m_i]^2. \tag{1.5.4}$$

À partir de là, ils calculent une seconde version de la vraisemblance, un estimateur du maximum de vraisemblance, en annulant la dérivée partielle de l'équation 1.5.4 par rapport à  $\sigma$ . Ceci enlève la dépendance de la vraisemblance sur la partie « non-connue » du bruit des données. Ils arrivent finalement à la formule suivante :

$$\ln \mathcal{L} = -\frac{N}{2} \ln \frac{1}{N} \left[ \sum f_i^2 - 2 \sum f_i m_i + \sum m_i^2 \right]. \tag{1.5.5}$$

Si les flux sont bien normalisés, alors  $\sum f_i = 0$  et  $\sum m_i = 0$ . Sous cette condition, alors les premier et dernier termes de l'équation 1.5.5 représentent la variance des données et du modèle, respectivement. On retrouve la CC au terme du milieu (cf. éq. 1.5.1). On voit donc que d'avoir une variance trop grande pour les données (donc beaucoup de bruit) et pour le modèle va pénaliser le  $\ln \mathcal{L}$ . On s'assure alors qu'un modèle ayant une variance très grande et qui viendrait augmenter arbitrairement le terme de CCF, en sera également pénalisé.

L'équipe de Gibson et al. (2020) ont aussi suivi cette approche, mais en n'enlevant pas toute la dépendance sur  $\sigma$ . Ils ont plutôt établi que  $\sigma$  pouvait être écrit comme  $\beta\sigma$ , où  $\beta$  représente la partie inconnue du bruit, et  $\sigma$  celle qu'on connaît. En posant plutôt la dérivée partielle de  $\beta$  à zéro (donc en enlevant la dépendance de la partie non-connue du bruit), on arrive à l'équation de log-vraisemblance suivante :

$$\ln \mathcal{L} = -\frac{N}{2} \ln \left[ \frac{1}{N} \left( \sum \frac{f_i^2}{\sigma_i^2} + \sum \frac{m_i^2}{\sigma_i^2} - 2 \sum \frac{f_i m_i}{\sigma_i^2} \right) \right] \tag{1.5.6}$$

qui peut d'ailleurs être écrite sous la forme plus compacte :

$$\ln \mathcal{L} = -\frac{N}{2} \ln \frac{\chi^2}{N}, \tag{1.5.7}$$

via l'équation 1.5.2. L'équation 1.5.7 a ainsi été utilisée comme mesure de signal directement ou via le  $\Delta\text{BIC}$  (Gibson et al., 2020; Boucher et al., 2021).

Pour la SFD, il est plus commun d'utiliser le  $\chi^2$  pour mesurer la « qualité » du modèle utilisé (ex. : Kreidberg et al., 2015; Spake et al., 2021; Mansfield et al., 2022), mais cette métrique a aussi été utilisé en SHD (ex. : Rodler et al., 2013; Brogi et al., 2016; Beltz et al., 2021). Mais outre leurs valeurs absolues, le  $\chi^2$  et ces  $\ln \mathcal{L}$  sont surtout utilisés comme fonction d'optimisation pour les méthodes de récupérations.

### 1.5.1. Méthode de récupération

Une fois qu'on a une suite de modèles qui transpose les quantités d'intérêt sur les données/observables (principalement les abondances chimiques, le profil TP, et la dynamique), il ne nous reste qu'à les extraire.

Pour se faire, on peut utiliser les métriques mentionnées ci-dessus (ex. : la minimisation du  $\chi^2$ , la maximisation de la vraisemblance, ou de la CC, etc.) avec des grilles de modèles. Ces grilles de modèles sont précalculées, et peuvent faire varier d'un à plusieurs paramètres, ce qui permet d'explorer l'espace des paramètres. Par contre, le grand désavantage est que plus on ajoute de paramètres, plus la taille de la grille augmentera de façon exponentielle. Le même problème survient si on veut améliorer l'échantillonnage d'un paramètre donné (que ce soit pour le rendre plus étendu ou avoir des pas de grille plus fins). De plus, pour explorer les régions entre 2 points d'une grille, il est nécessaire d'interpoler les modèles, ce qui n'est pas optimal si les variations causées par un des paramètres ne sont pas triviales.

Par contre, si on veut une méthode plus robuste à la comparaison de modèles et à l'exploration de l'espace des paramètres, on utilisera les méthodes de « récupération » (ou de recouvrement, de l'anglais "*retrieval method*"). Contrairement aux grilles de modèles, cette méthode permet d'explorer « dynamiquement » l'espace des paramètres, ce qui mènera à focaliser sur les régions d'intérêt. Cela nécessite un estimateur bayésien qui viendra optimiser les paramètres du modèle (ou leur intervalle) selon une fonction de log-vraisemblance (Brogi & Line, 2019). Donc à chaque itération (ou pas), un modèle sera généré suivant un certain ensemble de paramètres et sera comparé aux observations, via la log-vraisemblance ; un choix statistique sera fait pour soit garder l'itération en mémoire ou l'exclure ; un nouvel ensemble de paramètres sera ensuite échantillonné, basé sur la dernière itération. Le choix de garder ou non l'itération et la sélection du prochain pas est déterminé par la méthode d'échantillonnage (ex. : la méthode de Monte-Carlo par chaînes de Markov – MCMC, ou encore l'échantillonnage imbriqué, de l'anglais *nested sampling*). Une fois l'espace des paramètres bien échantillonné, on obtient une distribution de probabilité postérieure pour chacun des paramètres. En marginalisant sur tous les

autres paramètres, on peut estimer la valeur la plus probable d'un paramètre (la moyenne) ainsi que ses incertitudes.

Bien que cette méthode soit le standard pour l'analyse SFD, elle ne l'est pas encore pour la SHD (quoiqu'elle est en plein essor). La raison est que c'est beaucoup plus long en temps de calcul. Il est toujours possible de faire certaines hypothèses pour accélérer la génération des modèles (ex. : fixer certains paramètres, restreindre leur intervalle possible, ou paramétrer le plus possible), ou encore d'optimiser les codes et les méthodes (Gibson et al., 2022).

Madhusudhan & Seager (2009) sont les premiers à avoir utilisé une méthode de ce type pour les atmosphères d'exoplanètes, avec des données SFD. Au lieu de calculer tous les processus physiques dans l'atmosphère (modèles autocohérents), ils ont paramétré leurs profils TP et ont seulement exigé que la balance d'énergie soit satisfaite au haut de l'atmosphère. Ils ont ainsi moins d'informations directes sur la physique qui est en jeu, mais peuvent tout de même retrouver l'état de l'atmosphère (déterminé par les paramètres de leur modèle qui sont contraints de manière statistique), et ce, plus rapidement. Depuis, plusieurs autres analyses globales utilisant cette méthode ont vu le jour et se spécialisent dans les atmosphères des exoplanètes (ex. : Benneke & Seager, 2012; MacDonald & Madhusudhan, 2017; Wakeford et al., 2017; Barstow et al., 2017; Lavie et al., 2017) et des naines brunes (Line et al., 2014, 2015, 2017). Benneke (2015) a construit une analyse basée sur une méthode de récupération qui utilise des modèles autocohérents. Leur méthode peut ainsi combiner la cohérence des modèles atmosphériques complexes et le traitement statistique des méthodes de récupération, qui offrent un meilleur traitement des incertitudes des observations, mais cela devient évidemment plus coûteux en temps de calcul. Plus récemment, d'autres études ont réussi à intégrer l'utilisation des méthodes de recouvrement similaires (avec des modèles autocohérents, ou non), mais avec la SHD (Brogi et al., 2017; Brogi & Line, 2019; Gandhi et al., 2020; Gibson et al., 2020; Pelletier et al., 2021; Gibson et al., 2022).

Par contre, ces études ont toutes été faites avec des données en émission. La première étude de spectroscopie de transit à HR faite avec une méthode de recouvrement hybride (une méthode MCMC avec une grille de modèles interpolée) est présentée au Chapitre 2 (Boucher et al., 2021), tandis que (au moment d'écrire ces lignes) la première étude complète (une méthode MCMC qui génère les modèles à chaque itération) est présentée au Chapitre 3.

## 1.6. Contexte de ces travaux

L'exoplanétologie est un domaine extrêmement populaire en général, et l'étude des atmosphères, plus particulièrement, a pris beaucoup d'ampleur dans les dernières décennies. Notamment, la branche de la SHD a explosé avec l'arrivée des spectrographes échelle dans l'IRP et continue d'apporter d'importants résultats. Les avancées potentielles de ce champ d'études sont

nombreuses. C'est en partie ce qui a motivé cette thèse de doctorat, qui porte sur la caractérisation de l'atmosphère d'exoplanètes. Plus précisément, nous nous sommes concentrés sur la spectroscopie de transit à haute résolution d'une Jupiter chaude et d'une sous-Saturne chaude, avec l'instrument SPIRou.

Globalement, on s'intéresse à l'étude de l'atmosphère des exoplanètes pour plusieurs raisons. On vise à faire avancer la compréhension de la formation, l'évolution et l'historique de migration des exoplanètes, mais plus précisément des planètes géantes ; à mieux comprendre les processus physiques et chimiques présents, plus précisément ceux des objets très irradiés et/ou des systèmes vieux ; à étudier la dynamique présente dans ces mêmes atmosphères, où la haute résolution est primordiale ; et enfin, à agrandir l'échantillon de spectres de transmission dans la littérature, qui par le fait même agrandit la plage spectrale disponible pour les analyses combinées et études de populations futures. De plus, la nouveauté de ce champ de recherche en astrophysique à l'Université de Montréal et la nouveauté de l'instrument SPIRou nécessitaient des efforts supplémentaires. Ces travaux de recherche ont donc permis de développer et construire les algorithmes de traitement de données et d'extraction des paramètres pour les données SPIRou, en plus de valider le tout avec de vraies données.

C'est réellement une ère excitante pour l'étude des atmosphères, avec l'arrivée du JWST, et surtout qu'on n'a jamais été aussi près d'avoir une réponse à notre fameuse question avec l'arrivée prochaine des télescopes géants.

Le choix d'utiliser SPIRou s'est fait naturellement. Cet instrument a été conçu en grande partie par l'équipe de l'Université de Montréal, en collaboration avec 6 autres pays. Notre équipe a donc une très bonne compréhension du fonctionnement de l'instrument. C'est aussi parce qu'il est un des seuls de sa catégorie avec à la fois une haute résolution, un large intervalle de longueur d'onde en continu, en plus d'un grand débit lumineux, ce qui lui donne une grande richesse d'informations spectrales.

Pour la première analyse (présentée au Chapitre 2), nous avons choisi HD 189733 b (Bouchy et al., 2005) : La raison principale est que la grande brillance de son étoile et sa grande profondeur de transit offrent des conditions optimales pour faire des analyses détaillées. C'est entre autres pour ces raisons que, tout comme HD 209458 b (Charbonneau et al., 2000), HD 189733 b est une des JCs les plus étudiées à ce jour : en transit, en émission, à faible résolution, à haute résolution, en vitesse radiale, en polarimétrie, etc. Les détections répétées lui ont donné son statut de référence et d'étalon. Cela a motivé sa sélection en tant que cible de validation pour les données SPIRou, ainsi qu'en tant que première analyse de spectroscopie de transit pour le SPIRou Legacy Survey (SLS ; PI : Jean-François Donati). Les paramètres du système étant très bien connus, on peut ainsi mieux ancrer la recherche.

Comme deuxième analyse (présentée au Chapitre 3), nous avons choisi WASP-127 b pour deux raisons principales. La première est que dès sa découverte, le système montrait des points

prometteurs pour l'analyse en spectroscopie de transit : son grand rayon et sa grande température d'équilibre pointaient vers une très grande hauteur d'échelle, équivalant à un très grand signal atmosphérique. La deuxième est qu'une équipe l'ayant étudié par SFD a obtenu des résultats qui penchaient vers deux conclusions différentes et contradictoires. Le manque de couverture dans l'IRP a limité leur capacité à différencier entre les deux scénarios de façon indépendante. SPIROU fait partie de la courte liste d'instruments pouvant lever cette incertitude, grâce à sa couverture de la bande *K*. Nous avons donc sauté sur l'occasion d'étudier un système très inhabituel.

## 1.7. Contributions

J'ai rédigé les deux articles, ainsi que l'introduction et la conclusion présentés dans cette thèse. J'ai fait la revue de littérature présentée dans l'introduction. J'ai généré toutes les figures présentées dans les articles, tandis que les figures de l'introduction proviennent de la littérature (leur référence étant indiquée dans la légende). L'idée de faire de la spectroscopie de transmission à haute résolution avec SPIROU provient de David Lafrenière.

J'ai écrit et construit la majorité du code informatique Python utilisé pour cette thèse (ceci exclut les bibliothèques de codes publiques utilisées, telles que `numpy`, `matplotlib`, `scipy`, `astropy`, `emcee`, etc., ainsi que les suites de modèles SCARLET et petitRADTRANS). À moins de mention contraire et à l'exception de la réduction des données brutes et la correction des telluriques par APERO, j'ai bâti et effectué tout le reste de la réduction, de l'analyse des données, ainsi que des méthodes de récupération des paramètres présentés dans ces articles. Ceci comprend l'extraction des spectres de transmission des données, le calcul et le traitement des séquences de modèles, l'analyse de corrélation-croisée, de test-*t* et de log-vraisemblance, ainsi que les tests d'injection.

Pour l'article deux (A2), présenté au Chapitre 3, j'ai rédigé 3 demandes de temps pour l'observatoire du CFHT, avec l'aide de David Lafrenière et Antoine Darveau-Bernier, et j'ai planifié les observations obtenues. Toutes les données brutes reçues (pour les deux articles, incluant les données du SLS) ont été réduites et corrigées des telluriques avec APERO, par Étienne Artigau et Neil J. Cook, les auteurs et programmeurs de ce logiciel de réductions des données SPIROU et de correction des telluriques. Étienne Artigau a aussi développé un outil pour calculer la vitesse radiale des spectres, et aidé à rédiger la section sur les étapes de correction des telluriques. Neil J. Cook a été très proactif dans la distribution des données nouvellement réduites, et supervisait la qualité des données réduites.

L'ensemble des co-auteurs.trices a contribué au développement de certaines parties de l'analyse et/ou à la révision du texte des articles. Ceci a permis de clarifier ou rectifier certains points et d'élever la qualité des deux articles autant au niveau du contenu qu'au niveau de la structure.

Plus particulièrement : Antoine Darveau-Bernier a aidé au développement, à la programmation et à l'amélioration des codes d'analyse des données, et fait plusieurs suggestions pour pousser l'analyse. Il a également développé un outil pour rapidement obtenir les paramètres des exoplanètes dans la littérature, l'ExoFile. Stefan Pelletier a généré toutes les grilles de modèles avec SCARLET, la suite de modèles développée par Björn Benneke, utilisée dans l'article un (A1) du chapitre 2. Il a écrit la base du code de correction des PCAs pour les spectres de transmission. Nos discussions ont mené à l'optimisation de mes analyses, et ont permis de rectifier beaucoup de points qui font que A2 est beaucoup plus robuste. Il a révisé les formulations en anglais et consolidé la discussion des deux articles. David Lafrenière a grandement aidé au développement de l'étude en général, à trouver des pistes de solutions pour les problèmes rencontrés, à suggérer des tests pour valider les résultats. Il a supervisé l'écriture des articles et de l'introduction, ce qui a grandement consolidé et amélioré le texte. Romain Allart a aidé à trouver de nouvelles façons d'analyser les données lorsqu'aucun signal n'était détecté pour A1. Il a également comparé mes résultats avec ses résultats précédents pour A2. Michael Radica a aidé à développer la discussion sur l'interprétation des résultats dans les deux articles. René Doyon a suivi l'évolution de ce projet de recherche et donné des suggestions de tests tout au long du projet. Il a poussé, avec l'aide de Björn Benneke, à ce que WASP-127 b soit observée dans le SLS, ce qui m'a permis d'avoir un transit additionnel de WASP-127 b, qui représente des données hautement nécessaires et précieuses. Jean-François Donati et le reste de l'équipe SPIRou m'ont laissé mener les analyses de spectroscopie de transit pour le SPIRou Legacy Survey. Jonathan Gagné nous a enligné sur la discussion par rapport au décalage vers le rouge gravitationnel et le décalage vers le bleu convectif. Claire Moutou a également effectué les nuits d'observation pour les transits HD 189733 b. Florian Debras a collaboré à partager le temps d'observation de WASP-127 b.



## Chapitre 2

---

# Characterizing exoplanetary atmospheres at high resolution with SPIRou : Detection of water on HD 189733 b

Ce manuscrit a été publié dans la revue arbitrée *The Astronomical Journal* le 9 novembre 2021 (# AAS31484).

ANNE BOUCHER<sup>1</sup>, ANTOINE DARVEAU-BERNIER<sup>1</sup>, STEFAN PELLETIER<sup>1</sup>, DAVID LAFRENIÈRE<sup>1</sup>, ÉTIENNE ARTIGAU<sup>1</sup>, NEIL J. COOK<sup>1</sup>, ROMAIN ALLART<sup>1</sup>, MICHAEL RADICA<sup>1</sup>, RENÉ DOYON<sup>1</sup>, BJÖRN BENNEKE<sup>1</sup>, LUC ARNOLD<sup>2</sup>, XAVIER BONFILS<sup>3</sup>, VINCENT BOURRIER<sup>4</sup>, RYAN CLOUTIER<sup>5</sup>, JOÃO GOMES DA SILVA<sup>6</sup>, EMILY DEIBERT<sup>7</sup>, XAVIER DELFOSSE<sup>3</sup>, JEAN-FRANÇOIS DONATI<sup>8</sup>, DAVID EHRENREICH<sup>4</sup>, PEDRO FIGUEIRA<sup>6,9</sup>, THIERRY FORVEILLE<sup>3</sup>, PASCAL FOUQUÉ<sup>2,8</sup>, JONATHAN GAGNÉ<sup>1,10</sup>, ERIC GAIDOS<sup>11</sup>, GUILLAUME HÉBRARD<sup>12</sup>, RAY JAYAWARDHANA<sup>13</sup>, BAPTISTE KLEIN<sup>14</sup>, CHRISTOPHE LOVIS<sup>4</sup>, JORGE H. C. MARTINS<sup>6</sup>, EDER MARTIOLI<sup>12,15</sup>, CLAIRE MOUTOU<sup>8</sup>, AND NUNO C. SANTOS<sup>6,16</sup>

<sup>1</sup> Institut de Recherche sur les Exoplanètes (iREx), Université de Montréal, Département de Physique, C.P. 6128 Succ. Centre-ville, Montréal, QC H3C 3J7, Canada

<sup>2</sup> CFHT Corporation ; 65-1238 Mamalahoa Highway, Kamuela, HI 96743, USA

<sup>3</sup> CNRS, IPAG, Université Grenoble Alpes, F-38000 Grenoble, France

<sup>4</sup> Observatoire Astronomique de l'Université de Genève, Chemin Pegasi 51b, CH-1290 Versoix, Switzerland

<sup>5</sup> Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA

<sup>6</sup> Instituto de Astrofísica e Ciências do Espaço, Universidade do Porto, CAUP, Rua das Estrelas, 4150-762 Porto, Portugal

<sup>7</sup> David A. Dunlap Department of Astronomy & Astrophysics, University of Toronto, 50 St. George Street, ON M5S 3H4, Canada

<sup>8</sup> Université de Toulouse, CNRS, IRAP, 14 av. Belin, F-31400 Toulouse, France

<sup>9</sup> European Southern Observatory, Alonso de Cordova 3107, Vitacura, Santiago, Chile

<sup>10</sup> Planétarium Rio Tinto Alcan, Espace pour la Vie, 4801 av. Pierre de Coubertin, Montréal, QC, Canada

<sup>11</sup> Department of Earth Sciences, University of Hawai'i at Manoa, Honolulu, HI 96822 USA

<sup>12</sup> Sorbonne Université, CNRS, UMR 7095, Institut d'Astrophysique de Paris, 98 bis bd Arago, F-75014 Paris, France

<sup>13</sup> Department of Astronomy, Cornell University, Ithaca, NY 14853, USA

<sup>14</sup> Sub-department of Astrophysics, Department of Physics, University of Oxford, Oxford OX1 3RH, UK

<sup>15</sup> Laboratório Nacional de Astrofísica, Rua Estados Unidos 154, 37504-364, Itajubá - MG, Brazil

<sup>16</sup> Departamento de Física e Astronomia, Faculdade de Ciências, Universidade do Porto, Rua do Campo Alegre, 4169-007 Porto, Portugal

**RÉSUMÉ.** Nous présentons la première détection d'atmosphère d'exoplanète réalisée dans le cadre du SPIRou Legacy Survey, un grand programme d'observation de 300 nuits exploitant les capacités de SPIRou, le nouveau spectropolarimètre installé sur le Télescope Canada-France-Hawaï (TCFH; 3,6 m). Nous avons observé deux transits de HD 189733 b, une Jupiter chaude largement étudiée qui est connue pour montrer des signaux importants en absorption de vapeur d'eau dans son spectre de transmission. En combinant les deux transits, nous détectons avec succès l'absorption de vapeur d'eau de la planète à  $5.9\sigma$  en utilisant un test- $t$  de corrélation-croisée, ou avec un  $\Delta\text{BIC} > 10$  en utilisant un calcul de log-vraisemblance. En utilisant une méthode de récupération bayésienne en supposant un modèle d'atmosphère avec un profil T-P paramétré, nous retrouvons les paramètres de l'atmosphère de la planète, dans la région sondée par notre spectre de transmission, aux valeurs suivantes :  $\text{VMR}[\text{H}_2\text{O}] = -4.4^{+0.4}_{-0.4}$ , et  $P_{\text{cloud}} > \sim 0.2$  bar (nuages gris), qui sont tous deux cohérents avec les études précédentes sur cette planète. L'abondance d'eau obtenue est légèrement sous-solaire, bien que lorsque combinée avec les abondances de CO super-solaires précédemment récupérées par d'autres études, cela impliquerait un rapport C/O super-solaire. De plus, nous mesurons un décalage net vers le bleu du signal de la planète à  $-4.62^{+0.46}_{-0.44} \text{ km s}^{-1}$ , ce qui est plus grand que les valeurs obtenues dans plusieurs études précédentes. Il est peu probable qu'il en résulte uniquement des vents dans l'atmosphère de la planète, bien que cela pourrait s'expliquer par un signal de transit dominé par le limbe arrière (soir) de la planète. Ce grand décalage vers le bleu est observé dans toutes les différentes méthodes de détection/récupération qui ont été effectuées et dans chacun des deux transits indépendamment.

**Mots-clés :** Planètes et satellites : atmosphères — Planètes et satellites : individuels (HD 189733 b) — Méthodes : analyse de données — Techniques : spectroscopiques

**ABSTRACT.** We present the first exoplanet atmosphere detection made as part of the SPIRou Legacy Survey, a Large Observing Program of 300 nights exploiting the capabilities of SPIRou, the new near-infrared high-resolution ( $R \sim 70,000$ ) spectro-polarimeter installed on the

Canada-France-Hawaii Telescope (CFHT; 3.6-m). We observed two transits of HD 189733, an extensively studied hot Jupiter that is known to show prominent water vapor absorption in its transmission spectrum. When combining the two transits, we successfully detect the planet’s water vapor absorption at  $5.9\sigma$  using a cross-correlation  $t$ -test, or with a  $\Delta\text{BIC} > 10$  using a log-likelihood calculation. Using a Bayesian retrieval framework assuming a parametrized T-P profile atmosphere models, we constrain the planet atmosphere parameters, in the region probed by our transmission spectrum, to the following values :  $\text{VMR}[\text{H}_2\text{O}] = -4.4_{-0.4}^{+0.4}$ , and  $P_{\text{cloud}} > \sim 0.2$  bar (gray clouds), both of which are consistent with previous studies of this planet. Our retrieved water volume mixing ratio is slightly sub-solar although, combining it with the previously retrieved super-solar CO abundances from other studies would imply super-solar C/O ratio. We furthermore measure a net blue shift of the planet signal of  $-4.62_{-0.44}^{+0.46}\text{km s}^{-1}$ , which is somewhat larger than many previous measurements and unlikely to result solely from winds in the planet’s atmosphere, although it could possibly be explained by a transit signal dominated by the trailing limb of the planet. This large blue shift is observed in all the different detection/retrieval methods that were performed and in each of the two transits independently.

**Keywords :** Planets and satellites : atmospheres — Planets and satellites : individual (HD 189733 b) — Methods : data analysis — Techniques : spectroscopic

## 2.1. Introduction

The characterization of exoplanet atmospheres using transmission or emission spectroscopy has grown considerably since it was proposed by Seager & Sasselov (2000). The spectra that such techniques yield can be used to probe the state and composition of an exoplanet’s atmosphere. This provides insight into physical and chemical processes at play, which can then be interpreted through different formation pathways and evolutionary histories of the planet, usually by measuring the metallicity and the C/O ratio via their molecular abundances (e.g., Öberg et al. (2011); Pelletier et al. (2021)). The most successful observations to date have typically been obtained from space (e.g., Madhusudhan et al. 2014; Sing et al. 2016; Barstow et al. 2017; Pinhas et al. 2019; Welbanks et al. 2019), with the frequent use of the Hubble Space Telescope Wide Field Camera 3 (HST/WFC3) spectrometric and Spitzer Space Telescope photometric capabilities. The exquisite image and instrument stability enable the detection of subtle spectral differences induced by the planet’s atmosphere during its transit or eclipse and phase curve (which can be on the order of only a few tens of parts-per-million, ppm, e.g., Kreidberg et al. 2014). One caveat, however, is that these instruments usually have a limited wavelength coverage, which limits the simultaneous detection of water and carbon-bearing molecules (which prevents the computation of an accurate C/O ratio) and can sometimes lead to mixed results, e.g. the unclear nature of the strong spectral feature in the  $4.5\ \mu\text{m}$  Spitzer band in WASP-127 b

atmosphere, which could come from CO and/or CO<sub>2</sub> (Spake et al., 2021). Ground-based high dispersion spectroscopy (HDS) can achieve similar precision by resolving each line independently. The change in orbital radial velocity of the planet can be exploited to disentangle the signal of the planet from that of stellar and telluric signals (e.g. Snellen et al. 2010). In that case, the transit/emission spectrum is probed by looking for a signal (from either atomic or molecular species) that follows the planet radial velocity, which can be orders of magnitudes smaller than that of its star. Compared to low-resolution spectroscopy, HDS has the added benefit of being able to disentangle absorption by different molecules with overlapping bands (but non-overlapping lines, e.g. Giacobbe et al. 2021) and from different origins (e.g., stellar, planetary, and telluric) because the individual lines are resolved. Moreover, wind speeds and global dynamics of the planet’s atmosphere can be measured (through the Doppler shift and broadening of the planet’s absorption lines; e.g., Wyttenbach et al. 2015; Louden & Wheatley 2015; Flowers et al. 2019).

The firsts successful atmospheric characterizations from the ground were detections of sodium in the optical on HD 189733 b (Redfield et al. 2008; with the High Resolution Spectrograph,  $R \sim 60,000$ , on the Hobby-Eberly Telescope) and HD 209458 b (Snellen et al. 2008; with the High Dispersion Spectrograph,  $R \sim 45,000$ , on the Subaru Telescope; even though these results are inconsistent with the more recent ones from Casasayas-Barris et al. (2020)). The first near-infrared (NIR) HDS study was presented by Snellen et al. (2010), with the detection of carbon monoxide (CO) in the atmosphere of HD 209458 b via transmission spectroscopy using CRIRES ( $R \sim 100,000$ ; Kaeufl et al. 2004) installed at the VLT (8.2-m telescope). Many more such detections followed : H<sub>2</sub>O and CO for several transiting and non-transiting exoplanets (Brogi et al., 2012; Birkby et al., 2013; Brogi et al., 2016; Birkby et al., 2017; Webb et al., 2020), HCN for HD 209458 b (Hawker et al., 2018), and the effects of global atmospheric dynamics, such as day-to-night winds and/or eastward winds (jet streams) on HD 209458 b and HD 189733 b (Snellen et al., 2010; Brogi et al., 2016; Flowers et al., 2019; Beltz et al., 2021).

Multiple other high-resolution instruments have enabled additional interesting results again for transiting and non-transiting targets. A lot of them came from optical instruments (a non-exhaustive list includes HARPS, e.g., sodium detection on WASP-76 b, Seidel et al. 2019; HARPS-North, e.g., neutral Fe and Ti on the atmosphere of KELT-9 b, Hoeijmakers et al. 2018, ESPRESSO, e.g., the iron condensation on the nightside of WASP-76 b, Ehrenreich et al. 2020, and CARMENES VIS-channel, e.g., extended H $\alpha$  envelope detected on KELT-9 b Yan & Henning 2018), but the following list highlights some of the near-infrared ones. The NIRSPEC spectrograph (0.95–5.4  $\mu\text{m}$ ,  $R \simeq 25,000$ ; McLean et al. 1998) on the Keck II telescope led to the detection of several species, such as CO, H<sub>2</sub>O (e.g., Rodler et al., 2013; Lockwood et al., 2014; Piskorz et al., 2016, 2017, 2018; Buzard et al., 2020), and more recently metastable helium (He, at 1083 nm; e.g., Kirk et al. 2020). CARMENES (NIR-channel, 0.96–1.71  $\mu\text{m}$ ,

$R = 80,400$ , Calar Alto Observatory, 3.5-m; Quirrenbach et al. 2014) has also been very prolific in recent years. He absorption was detected on multiple targets (Allart et al., 2018, 2019; Salz et al., 2018; Alonso-Floriano et al., 2019; Palle et al., 2020), as well as H<sub>2</sub>O absorption in both HD 189733 b and HD 209458 b (Alonso-Floriano et al., 2019; Sánchez-López et al., 2019, 2020). Also, GIANO (0.95–2.45  $\mu\text{m}$ ,  $R = 50,000$ , Telescopio Nazionale Galileo, 3.6-m; Oliva et al. 2006) was able to confirm the HD 189733 b detection of H<sub>2</sub>O (Brogi et al., 2018) and of metastable He (Guilluy et al., 2020). Moreover, Guilluy et al. (2019) also detected H<sub>2</sub>O absorption in the atmosphere of the non-transiting HD 102195 b planet, and found evidence for the presence of methane (CH<sub>4</sub>) as well.

The *Spectro-Polarimètre InfraRouge* (SPIRou; Donati et al. 2020) is a new fiber-fed échelle spectro-polarimeter operating in the NIR domain installed on the CFHT. SPIRou was primarily designed to detect and characterize Earth-like planets in the habitable zone of low-mass stars via precise radial velocity (RV), down to  $\sim 1 \text{ m s}^{-1}$ , and to study the stellar magnetic fields using its polarimetry capabilities (e.g., Martioli et al., 2020; Klein et al., 2021). SPIRou saw its first light April 2018, and is the first instrument to have both a high spectral resolution ( $R \sim 70,000$ ) as well as such a broad continuous NIR spectral range ( $\sim 0.95\text{--}2.50 \mu\text{m}$ ), which provides an unparalleled richness of spectral information. Its wider spectral range gives access to more individual absorption features, thus making SPIRou well suited for exoplanet atmospheric characterization via the cross-correlation technique (see Deibert et al. 2021a; Pelletier et al. 2021). SPIRou also has a higher throughput than many slit spectrographs and its design provides a more stable line spread function.

In this paper we report the first results of the SPIRou Legacy Survey (SLS) on exoplanet atmosphere characterization for the planet HD 189733 b (Bouchy et al., 2005). Initiated at the start of SPIRou operations, the SLS is a CFHT Large Program of 300 telescope nights (PI : Jean-François Donati) whose main goals are to search for planets around M dwarfs using precision RV measurement, characterize the magnetic fields of young low-mass stars and their impact on star and planet formation, and probe the atmosphere of exoplanets using HDS. The planet targeted here, HD 189733 b, is one of the most studied hot Jupiters to date, and due to the brightness of its active host star (K2V;  $H = 5.59 \text{ mag}$ ), it offers great opportunities for in-depth analyses. The study of the orbital motion, Rossiter-McLaughlin effect (RME; Rossiter, 1924; McLaughlin, 1924) and magnetic field made with these SPIRou SLS observations of HD 189733 b are presented in Moutou et al. (2020).

Its atmosphere has been studied many times and its transmission spectrum has revealed water (e.g., Brogi et al., 2016, 2018; Alonso-Floriano et al., 2019), CO (de Kok et al., 2013a, and references therein), hydrogen (Lecavelier des Etangs et al., 2012; Bourrier et al., 2013, 2020), metastable helium (Salz et al., 2018; Guilluy et al., 2020), and sodium (Redfield et al., 2008;

Table 2.1. SPIRou observations of HD 189733

	Transit 1	Transit 2
UT Date	2018-09-22	2019-06-15
BJD (d) <sup>a</sup>	2458383.77	2458649.95
Texp (s) <sup>b</sup>	250	250
Seeing (") <sup>c</sup>	0.79	0.85
SNR <sup>d</sup>	259	224
<b>Number of exposures :</b>		
Before ingress	0	12
During transit	21	24
After egress	15	14
Total	36	50
Tot. observ. time (h) <sup>e</sup>	2.50	3.47

Note. — <sup>a</sup> Barycentric Julian date at the start of the transit sequence; <sup>b</sup> Exposure time of a single exposure; <sup>c</sup> Mean value of the seeing during the transit; <sup>d</sup> Mean SNR per pixel, per exposure, at 1.7  $\mu\text{m}$ ; <sup>e</sup> Total observing time in hours.

Jensen et al., 2011; Huitson et al., 2012; Wytttenbach et al., 2015). The rotation of the planet, consistent with being tidally locked, and evidence of winds have also been detected (e.g., Wytttenbach et al., 2015; Louden & Wheatley, 2015; Brogi et al., 2016; Flowers et al., 2019). Furthermore, Barstow (2020) studied the properties and location of clouds in its atmosphere, which have a substantial effect on retrieved abundances. Their results pointed toward a heterogeneous atmosphere, with small-particle aerosols covering at least 60% of the terminator region, reaching to low pressures, but no gray cloud.

This paper is organized as follows. In Section 2.2, we present the observational setup and describe the observational data along with the details of the reduction. In Section 2.3 we explain the telluric and stellar signal removal process, and also present the atmospheric models that are used for the cross-correlation analysis. Section 2.4 details the methods that we used to extract the planet’s atmosphere signal and present the associated results. In Section 2.5 we discuss our findings, and summarize our main results in Section 2.6.

## 2.2. Observations and data reduction

All data presented here were obtained with SPIRou (Donati et al., 2020). The spectrograph covers the  $Y$ ,  $J$ ,  $H$ , and  $K_s$  bands ( $\sim 0.95\text{--}2.50 \mu\text{m}$ ) simultaneously at a nominal resolving power of  $R = \lambda/\Delta\lambda \sim 70,000$  (with a sampling precision of  $\sim 2.3 \text{ km s}^{-1}$  per pixel), over 50 spectral orders. The fiber-fed spectrograph is bench-mounted in a vacuum tank to maximize its stability

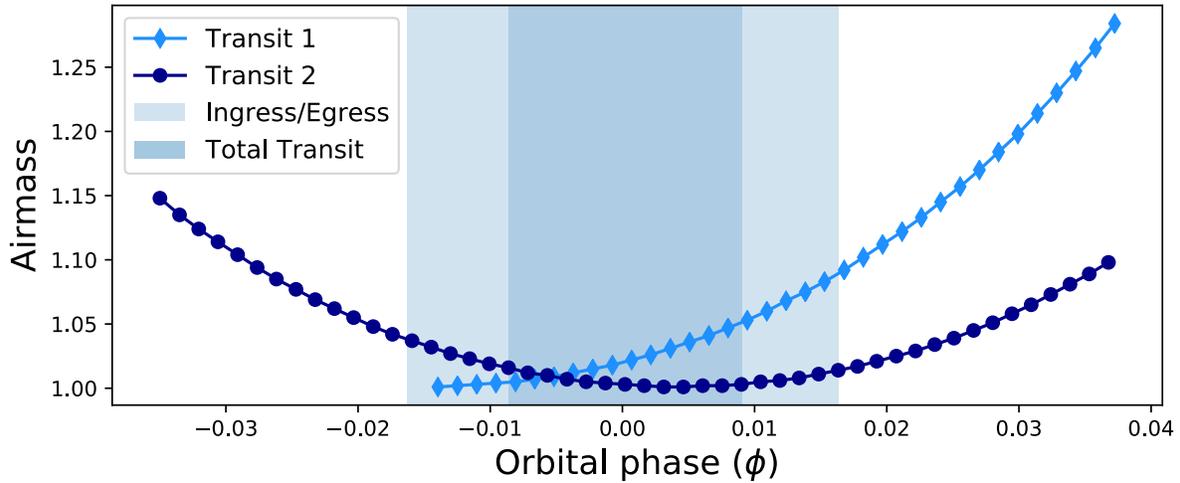


FIGURE 2.1 Airmass variation during the two SPIRou transit observations (light blue diamonds for Tr1 and dark blue circles for Tr2; the shaded area shows the span of the transit event).

and radial velocity precision. This further gives a much more stable line-spread function, greatly reducing the level of systematic errors (Artigau et al., 2014). SPIRou has a  $4096 \times 4096$ -pixel H4RG detector, with  $15 \mu\text{m}$ -wide pixels. The overall throughput is around 4 to 8% in the  $Y$  and  $J$  bands, while it increases to 10–12% in the  $H$  and  $K$  bands. There are two science fibers that monitor the object and one calibration fiber that can track a Fabry-Pérot etalon for simultaneous wavelength calibration.

Two transits of HD 189733 b were observed as part of the SLS. The first transit (hereafter Tr1) was observed on UT September 22, 2018, as part of SPIRou commissioning observations (later combined to SLS data), and the second (hereafter Tr2) on UT June 15, 2019. Both sets of observations were taken without moving the polarimeter optics (rhombs) to ensure the highest possible instrument stability, with the Fabry-Pérot in the calibration channel, and both with an exposure time of 250 s per individual exposure. Table 2.1 lists the parameters of the observations. The first data set consists of 2.5 h, divided into 36 exposures, where the first 21 are in transit and the remaining 15 are out-of-transit. Technical operations during this night prevented the start of the sequence early enough to observe the star before ingress. The second data set consists of 50 exposures in total, where 24 are in transit, 12 before, and 14 after transit, for a total of  $\sim 3.5$  h. Conditions were photometric for both transit sequences, with an average seeing of around  $0.82''$  as estimated from the guiding images. The airmass remained under 1.3 for the total duration of both observations (see Figure 2.1). The water column density was more stable during the second night with a range of 2.5–3.0 vs 2.6–3.6 mm  $\text{H}_2\text{O}$  for the first night. The signal-to-noise ratio (S/N) temporal mean per order and the spectral mean (over the  $H$ -band only) per exposure for

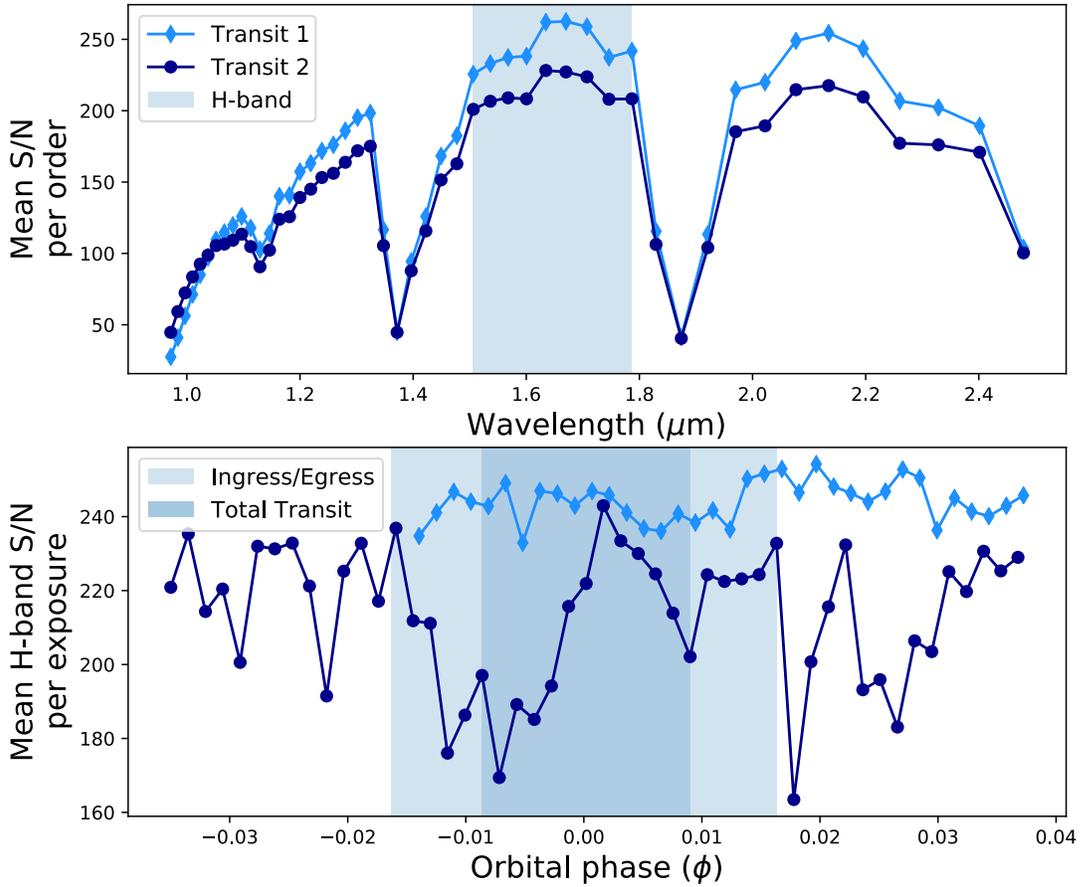


FIGURE 2.2 Temporal mean of the S/N per order (top panel; the shaded area highlights the orders in the  $H$ -band) and spectral mean of the S/N for  $H$ -band per exposure (bottom panel; the shaded area shows the span of the transit event) during SPIRou observations (light blue diamonds for Tr1 and dark blue circles for Tr2).

both transits are shown in Figure 2.2. The Tr2 S/N is much more variable, but has out-of-transit observations on both sides.

The data were reduced using `APER0` (A Pipeline to Reduce Observations; version 0.6.131; Cook et al., in prep.), the SPIRou data reduction software.

The first step is to pre-process the raw data. This removes certain detector effects (correcting for the top and bottom reference pixels, median filtering against the dark amplifiers and the  $1/f$  noise; Artigau et al. 2018). `APER0` then calibrates observations, correcting for the dark, flagging bad pixels (from a bad pixel map created from calibration flats and darks), removing

background scattered light (both locally and globally), and cleaning hot pixels (via interpolating over high-sigma outliers compared to their immediate neighbors).

The order positions are found and fit with polynomials and pixels are registered onto a common reference grid (using a master Fabry-Pérot, FP). In addition, the order geometry and the geometry of the detector (slicer shape, slicer tilt and other optics) are separated into changes across-order, along-order and an affine transformation matrix (essentially being characterized by a shift in  $dx$ ,  $dy$  and an A-B-C-D matrix, which can describe a translation, reflection, scale, rotation and or shear in the detector). These geometric changes are applied, along with the order polynomial fits in order to straighten the image.

The straightened image is then optimally extracted (Horne, 1986) and cosmic ray correction is applied. This produces an extracted 2D spectrum;  $E2DS$  of dimensions 4088 (4096 minus 8 reference pixels) by 49 orders (corresponding to physical orders of #79 in the blue to #31 in the red). The  $E2DS$  are flat and blaze corrected (the blaze is fitted with a simple *sinc* model) and a wavelength solution is available for each night (the wave calibration is done using both a hollow-cathode UNe lamp and the FP; Hobson et al. 2021). Thermal background is corrected with a dark calibration, scaled in amplitude to match the thermal background of the reddest orders.

As a final step, when the FP is used simultaneously in the calibration fiber, the contamination from the simultaneous FP is corrected for (using a set of calibrations with dark in the science fiber and the FP in the reference fiber to correct the contamination from the reference fiber into the science fiber). While the  $E2DS$  are produced for the two science fibers (A and B) and for the combined flux in the science fibers (AB), we only used the AB extraction as this is the relevant data product for non-polarimetric observations.

## 2.3. Analysis

### 2.3.1. Telluric absorption correction

A large portion of the NIR domain that SPIRou covers is affected by absorption in the Earth’s atmosphere (telluric absorption), which is resolved into thousands of individual narrow molecular lines at SPIRou’s resolution. This absorption is strongest and line densities are highest in between the  $YJHK$  photometric bandpasses, but many lines are nevertheless present with various strengths throughout the domain. It is crucial that these telluric absorption lines be precisely corrected for or masked out prior to seeking the subtle spectral signature of an exoplanet atmosphere seen in transit, especially as they arise from molecules also expected to be present in exoplanet atmospheres.

The spectra are corrected for telluric line contamination within APERO in a two-step process. First, the blaze normalized  $E2DS$  spectra are pre-cleaned by removing the best-fitting TAPAS

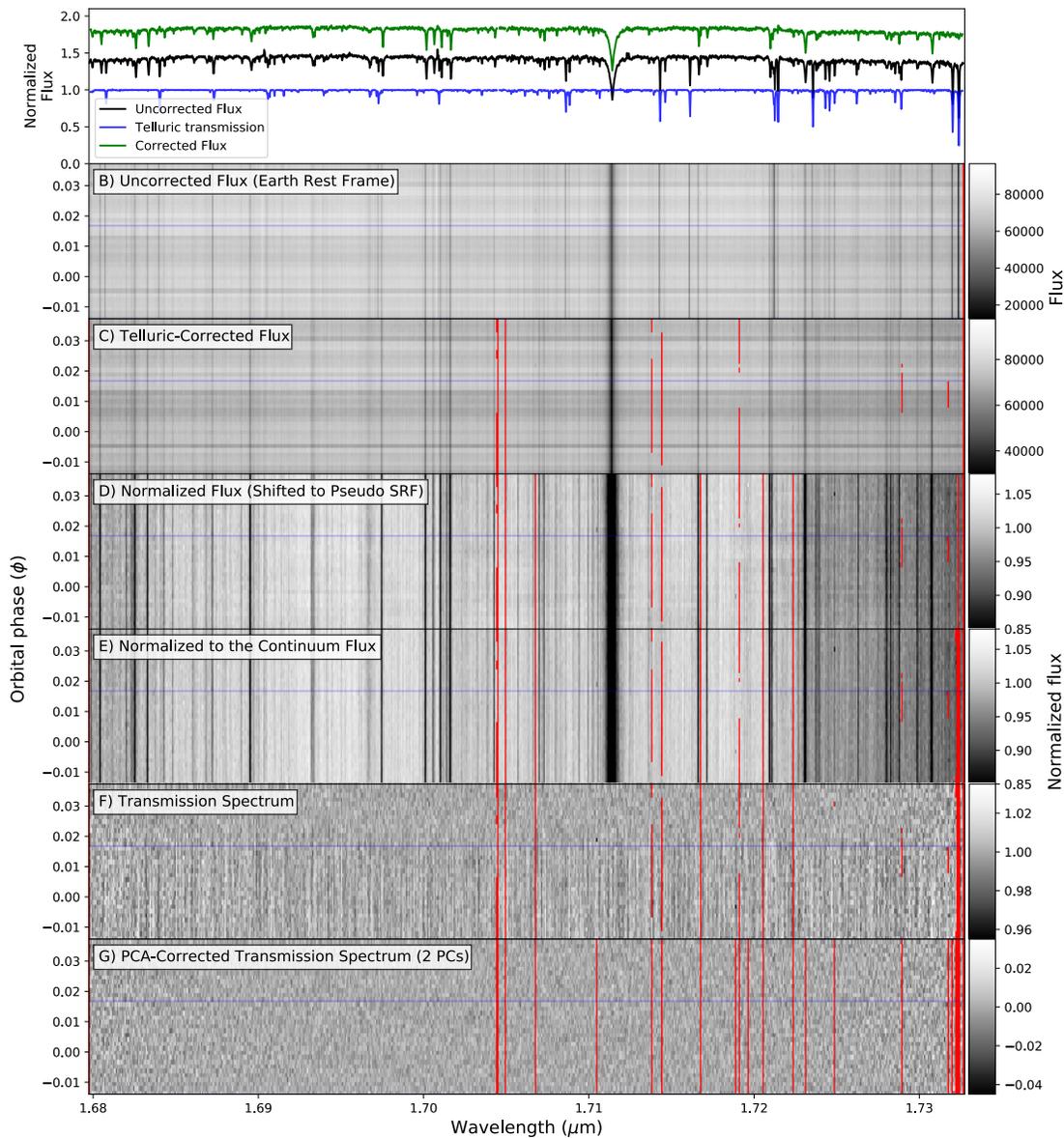


FIGURE 2.3 Analysis steps that are applied to the observed spectral time series of the first transit. Here the full order spanning  $1.6797$  to  $1.7327 \mu\text{m}$  is shown. *Panel A* : The uncorrected (black), the telluric-corrected (green) observed spectra and the reconstructed telluric transmission spectrum (blue) are shown with an offset to facilitate visibility. *Panel B* : Uncorrected spectra (counts normalized by the blaze function). *Panel C* : Telluric-corrected spectra. The masked pixels are shown in red; here the masking is done by APERO. *Panel D* : The high variance columns of the mean normalized spectra are masked, and then the spectra are co-aligned (shifted) in the pseudo-stellar rest frame. *Panel E* : The regions with deep telluric lines are masked. Every spectrum is normalized to the continuum level of the Master-Out spectrum. *Panel F* : Planetary transmission spectra, where each spectrum was divided by the Master-Out spectrum. *Panel G* : Final planetary transmission spectra corrected for the vertical residual structures using PCA. The blue line in panels B-G shows the egress position (mid-transit is at phase 0).

Table 2.2. Adopted parameters for the system HD 189733

Parameter	Symbol	Value	Unit	Reference
Stellar mass	$M_{\star}$	$0.806 \pm 0.048$	$M_{\odot}$	T08, B19
Stellar radius	$R_{\star}$	$0.756 \pm 0.018$	$R_{\odot}$	T08
Planet mass	$M_{\text{p}}$	$1.123 \pm 0.045$	$M_{\text{J}}$	T08
Planet radius	$R_{\text{p}}$	$1.138 \pm 0.027$	$R_{\text{J}}$	T08
Semi-amplitude	$K$	$0.205 \pm 0.006$	$\text{km s}^{-1}$	T08
Orbital semi-major axis	$a$	$0.03099^{+0.00060}_{-0.00063}$	AU	T08
Orbital period	$P_{\text{orb}}$	$2.218575123 \pm 0.000000057$	d	B19
Orbital eccentricity	$e$	$0.0028 \pm 0.0038$		B19
Orbital inclination	$i_{\text{p}}$	$85.712 \pm 0.036$	$^{\circ}$ (deg)	B19
Epoch of transit	$T_0$	$2453968.837031 \pm 0.000020$	BJD	B19
Transit duration	$T_{14}$	$1.8012 \pm 0.0016$	h	B19
Systemic velocity Tr1	$v_{\text{sys,Tr1}}$	$-2.59 \pm 0.21$	$\text{km s}^{-1}$	This work
Systemic velocity Tr2	$v_{\text{sys,Tr2}}$	$-2.76 \pm 0.21$	$\text{km s}^{-1}$	This work

Note. — (T08) Torres et al. 2008 and (B19) Baluev et al. 2019.

atmospheric transmission spectrum (Bertaux et al., 2014). This best-fit model is found by minimizing the cross-correlation signal between the TAPAS model and the telluric residuals, i.e. the data from which this same model was removed. Second, the telluric line residuals are corrected using the Principal Component Analysis (PCA) approach developed by Artigau et al. (2014) and implemented in *APER0* (Artigau et al., in prep.). The approach exploits the fact that the absorbance spectrum of Earth’s atmosphere can be expressed, in log-space, as a linear combination of absorbance spectra from different chemical species (chiefly  $\text{H}_2\text{O}$ ,  $\text{O}_2$ ,  $\text{CO}_2$  with minor  $\text{O}_3$ ,  $\text{CH}_4$  and  $\text{N}_2\text{O}$  features). As part of SPIRou night time calibrations, rapidly rotating A stars are observed and the derived telluric absorption from these observations is added to a library. The telluric correction spectrum is constructed through a linear combination (in log space) of the first 7 principal components (PCs) of the library of telluric absorptions, for each observed spectrum.

The most opaque regions in Earth’s atmosphere (saturated lines, or lines with transmission smaller than 10%) will let little to no light through, resulting in a poorly constrained/unreliable correction. The *APER0* pipeline performs the telluric absorption correction for lines with a transmission down to  $\sim 10\%$ , and deeper lines are masked out. The total reconstructed Earth transmission spectrum calculated by the pipeline (the product of the best-fit TAPAS model and the PCA reconstruction of the residuals) is one of the data product, allowing further masking if desired. An example of the telluric absorption reconstruction and correction is shown in Figure 2.3, panels A to C.

Sky emission lines are also removed by the `APER0` pipeline as part of the telluric correction process. The pipeline draws from a library of sky spectra taken at various times through the life of `SPIRou` and constructs a linear combination of the first 9 PCs of this library, again for every exposure individually, to reconstruct and subtract the sky emission lines in the observed spectrum. The final telluric corrected spectra are re-multiplied by the blaze to keep the level proportional to the photon counts.

In our analysis below, we found that further masking of deep telluric lines (i.e. further masking parts of the previously telluric corrected regions) improved the results, possibly because low-level telluric absorption residuals remain in the spectra. Using the reconstructed telluric spectrum to define our mask, we masked all telluric lines for which the transmission in the core is below 30% in Tr1 and 35% in Tr2, and for those lines, we also mask the wings by extending the mask until the transmission reaches 97% (Figure 2.3 E); these limits were empirically determined, based on an injection-recovery test. This step removes a total 8.6% and 10.6% of the spectral domain in Tr1 and Tr2, respectively. This masking is applied at step 3 in the following section.

### 2.3.2. Transmission spectrum construction

Starting from the telluric-corrected *E2DS* spectra, after division by the normalized blaze function to flatten them, we applied the following operations to construct the planet’s transmission spectra. These operations were applied on each order individually and separately for each transit sequence.

1) To remove bad pixels, the spectra are first normalized by dividing them by their median value (over wavelengths). Then, spectral pixels with a time variance (over the different spectra) more than  $4\sigma$  above the mean variance are masked. In total, this removes 0.6% and 0.5% of the pixels in Tr1 and Tr2, respectively.

2) All the spectra are Doppler shifted (via cubic spline interpolation that handles masked arrays) from the observer to a pseudo-stellar rest frame (SRF) by the opposite of the stellar radial velocity variation relative to the middle of the sequence,  $\Delta v_S(t)$ . The stellar radial velocity  $v_S(t)$  is itself defined as,

$$v_S(t) = v_{\text{bary}}(t) + v_{\text{sys}} + v_{\text{reflex}}(t), \quad (2.3.1)$$

where  $v_{\text{bary}}$  is the barycentric velocity of the observer (and in our case it is the barycentric Earth radial velocity, *BERV*),  $v_{\text{sys}}$  is the systemic radial velocity, and  $v_{\text{reflex}}(t) = K \sin[2\pi(\phi(t) + 0.5)]$  is the radial component of the reflex motion of the star induced by the planet (Keplerian), where  $K$  is the stellar radial velocity semi-amplitude (which we fixed; see Table 2.2), and  $\phi$  is the planet orbital phase ( $\phi = 0$  at mid-transit). We then define  $\Delta v_S(t)$  as follows :

$$\Delta v_S(t) = \Delta v_{\text{bary}}(t) + \Delta v_{\text{reflex}}(t), \quad (2.3.2)$$

where  $\Delta v_{\text{bary}}(t) = v_{\text{bary}}(t) - v_{\text{bary}}(t_{\text{mid.exp.}})$ , the difference between the barycentric velocity during the sequence ( $v_{\text{bary}}(t)$ ) and its value at the middle of the sequence ( $v_{\text{bary}}(t_{\text{mid.exp.}})$ ), and equivalently for the reflex motion term. This aligns the stellar lines across all spectra (see Figure 2.3 D), even though they are not exactly in the SRF, while minimizing the shifts applied to the data, hence minimizing the associated interpolation errors (the first half of the exposures are thus shifted by roughly the same amount as the second half, but in the opposite direction). In our case, the shifts applied are at most  $\sim 5\%$  and  $\sim 7\%$  of a SPIRou spectral pixel, for Tr1 and Tr2 respectively. Many of the HD 189733 system and orbital parameters are well known and the values that were used here are listed in Table 2.2. We measured  $v_{\text{sys}}$  directly from our data by computing the CCF of the telluric corrected spectra with a synthetic spectrum from a PHOENIX atmosphere model (Husser et al., 2013) with  $T_{\text{eff}} = 5100 \text{ K}$ ,  $\log g = 4.5$ , and  $Z = 0$ . We computed this CCF for all orders across the  $H$  band, measured the peak position, subtracted  $v_{\text{bary}}(t)$  and  $v_{\text{reflex}}(t)$ , then calculated the mean value by weighting by the S/N of the orders, and finally computed the mean over all spectra for each transit. We obtained values of  $-2.22 \pm 0.04 \text{ km s}^{-1}$  for Tr1, and  $-2.39 \pm 0.04 \text{ km s}^{-1}$  for Tr2. Then to determine the heliocentric radial velocity of the HD 189733 system from our observations, we subtracted  $0.67 \pm 0.04 \text{ km s}^{-1}$  to compensate for the gravitational redshift and  $-0.3 \pm 0.2 \text{ km s}^{-1}$  for the convective blueshift of a K2 star (Leão et al., 2019). The final  $v_{\text{sys}}$  values are reported in Table 2.2.

3) The additional telluric masking mentioned above is applied; this was not done earlier to limit interpolations on masked pixels (Figure 2.3 D to E).

4) The stellar spectral features are removed following the technique of Allart et al. (2017) : (a) A “Master-Out” spectrum is built by taking the mean of all the out-of-transit spectra (full-disk stellar spectrum). (b) All spectra are divided by a low-pass-filtered<sup>1</sup> version of their ratio with the Master-Out; this flattens out the continuum and removes the modal noise, thus reducing the correlation noise due to slopes in the spectra (Figure 2.3 E). (c) A second iteration of the Master-Out spectrum is calculated using the normalized out-of-transit spectra. (d) The normalized spectra are divided by this final Master-Out spectrum, which yields the planetary transmission spectra (see Figure 2.3 F), and sigma clipped at  $6\sigma$ , to prevent outliers to be included in the next step.

5) At this stage, we can see residual vertical features in the spectral time series 2D representation (Figure 2.3 F); these seem to appear because of the Master-Out is built with a subset of spectra and better represents the spectra closest to it in time (the out-of-transit spectra, in this case), and the residuals could be linked to systematic effects that vary with time. To remove

---

1. Median filter of width 51 pixels followed by convolution with a Gaussian kernel of width 5 pixels.

these residual features, we use a PCA approach. The PCs are built from the logarithm of the flux in the time dimension using each spectral pixel time series as a sample point (which is the transpose equivalent to building the PCs basis in the spectral space and having each exposure as a sample point). Each sample (spectral pixel time series) is then divided by the exponential of their reconstruction based on the firsts  $n$  chosen PCs. To determine the number of PCs to use, we performed an injection-recovery test for both analysis methods presented below, the cross-correlation function (CCF) and the log-likelihood. We injected the model at  $-K_P$  and the known  $v_{\text{sys}} + v_{\text{rad}}$ . This test indicates that removing 2 and 3 PCs yields the best detection significance of the injected signals for Tr1 and Tr2, respectively, and this is what we adopted for all analyses. The higher telluric fraction to be masked and the number of PCs to remove in the second transit could be linked to its lower data quality compared to Tr1.

6) Finally, any remaining outlier pixels are masked using sigma clipping (at  $3\sigma$ ) in the time dimension, and the mean of each spectrum is subtracted out, to keep a zero mean for the cross-correlation. This yields the final spectral planetary transmission values  $f_i$  (where  $i$  indexes over both time and wavelength), shown in Figure 2.3 G, that are used for the cross-correlation/log-likelihood mapping in Section 2.4.

During the analysis of the Tr2 data set, it was found that a significant change in systematic spectral noise occurred around passage of the target through the meridian, near airmass of 1. Similar noise patterns, always occurring as the target passes near the zenith, were also found in other SPIRou data sets (e.g. Pelletier et al. 2021). The cause of this is still being investigated, but is thought to be related to the azimuth angle of the telescope (TelAz). This effect is most important when the variation (or gradient) of the TelAz is maximal. See Pelletier et al. (2021) for a longer discussion of this effect. In our case, this effect had no impact on Tr1, as all data were taken after meridian crossing. For Tr2, we managed to work around this problem by splitting the data set in two subsets : before (27 exposures) and after (23 exposures) the meridian crossing. We applied all the analysis steps above to both subsets independently, and then we merged them back into a single sequence at the end.

### 2.3.3. Atmospheric Model

Obtaining information about the molecular content in the planet’s atmosphere requires the use of the cross-correlation method/log-likelihood mapping with high spectral resolution synthetic planetary transmission spectra. The models of HD 189733 b and associated transmission spectra that were used here were generated using the SCARLET framework (Self-Consistent Atmospheric Retrieval framework for Exoplanets; Benneke & Seager, 2012, 2013; Benneke, 2015; Benneke et al., 2019). SCARLET generates transmission spectra for a simulated planetary atmosphere in hydrostatic equilibrium by considering the molecular opacities at each pressure

layer. Included in the opacities are contributions from  $\text{H}_2\text{O}$ ,  $\text{H}_2$  broadening following Burrows & Volobuyev (2003), and collision-induced broadening from  $\text{H}_2/\text{H}_2$  as well as  $\text{H}_2/\text{He}$  collisions from Borysow (2002). The models are generated at a resolving power of  $R = 250,000$  using line-by-line radiative transfer and are later convolved to match the instrumental resolution of SPIRou. The resulting output is the transit depth as a function of wavelength, i.e.  $R_p^2(\lambda)/R_\star^2$ , which corresponds to the observed transmission spectrum calculated above, pending normalization by the continuum level.

The choice of the water lines list used in the models is important, given the preponderance of this molecule in the atmosphere and throughout the transmission spectrum of HD 189733 b. In this work, we adopted the POKAZATEL (Polyansky et al., 2018) water line list from Exomol (Tennyson et al., 2016). Most previous analyses used the HITEMP 2010 line list, but Gandhi et al. (2020) explicitly compared the two with HD 189733 b CRIRES data from Birkby et al. (2013) (thermal emission) and reported a strong agreement between the two, with a slightly higher signal from POKAZATEL. Similarly, Nugroho et al. (2021) observed a better detection significance for  $\text{H}_2\text{O}$  in WASP-33 b when using the POKAZATEL line list than when using HITEMP 2010, even though the signal is found at the same location. Webb et al. (2020) also compared these line lists in their study of the non-transiting HD 179949 b’s atmosphere and found consistent results, although their data favored HITEMP 2010 models.

In principle, any temperature-pressure (T-P) profile can be used to generate the atmospheric models, but here we adopted an analytical atmospheric T-P profile from Guillot (2010). For simplicity, we fixed three of the four parameters of the profile, namely  $\kappa_{\text{IR}}$  the atmospheric opacity in the IR wavelengths,  $\gamma$  the ratio between the optical and IR opacity, and  $T_{\text{int}}$  the planetary internal temperature, while keeping  $T_{\text{eq}}$ , the atmospheric equilibrium temperature, as a free parameter. We fixed the values to  $\kappa_{\text{IR}} = 10^{-1.5}$ ,  $\gamma = 10^{-0.85}$  and  $T_{\text{int}} = 100$  so that the shape of the resulting profile would roughly resemble those found in the literature for HD 189733 b, more specifically from Sing et al. (2016); Brogi & Line (2019). Both hazes and a gray cloud deck can be included in the models, but here we neglected the hazes contribution given the NIR wavelength range of SPIRou. We did, however, include a gray cloud deck contribution, characterized by its cloud top pressure  $P_{\text{cloud}}$  (bar), as this can have a large effect on the contrasts/depths of spectral lines. Rayleigh scattering is included by default even though its contribution is not significant in the NIR.

The models considered in our study are thus described by 3 free parameters : the water VMR, the temperature  $T_p$  (K) of the isothermal atmosphere profile, and the cloud top pressure  $P_{\text{cloud}}$  (bar). Once a model is generated, we convolve it to the resolving power of SPIRou, 70,000, and bin it to match the observed data pixel sampling.

## 2.4. Atmospheric Signal Extraction : Methods and results

Even with a high S/N SPIRou spectrum, individual planetary absorption lines are often weak and buried in the noise. To maximize the signal and the detection strength, we combine the signal from many lines via the cross-correlation and log-likelihood mapping techniques. This is why a maximum number of spectral lines is desired, which then justifies the need for a wide spectral range. In this work, we tested three specific approaches for detecting HD 189733 b’s signal and constraining its atmospheric parameters, but the models first need to be processed to better represent the data. In this section, we present how we process the models followed by the three approaches, i.e. the cross-correlation (and  $t$ -test), the log-likelihood mapping and MCMC retrieval.

### 2.4.1. Model processing

At this point in the analysis, the planetary atmosphere signal was affected by all the processing steps that were applied to the data. This mostly concerns the subtraction of the PCs, which are generally not orthogonal to the planet transmission spectra. Their subtraction (step 5 from Section 2.3.2) may remove part of the actual planet signal and introduce artefacts in the spectral time series, which can then bias the determination of the best parameters (velocities, abundances, temperature, etc.). We thus need to apply the same treatment to the model before comparing it to the data to ensure a better representation. So instead of using the models directly for the cross-correlation, we use a processed model transit sequence. We proceed as follows : we generate full synthetic transit sequences by injecting the model spectrum (described in Section 2.3.3) in a reconstruction of the observed data. This reconstructed signal is built by multiplying together (i) the spectral median of every spectrum (for every order; same as step 1 from Section 2.3.2), (ii) the master-out spectrum and (iii) the PCA reconstructed version of the transmission spectrum, all of which are (mostly) planet signal free. We inject the model at  $v_P(t)$ , the total planet radial velocity,

$$v_P(t) = v_{\text{bary}}(t) + v_{\text{sys}} + v_{\text{reflex,P}}(t) - v_{\text{reflex}}(t) + v_{\text{rad}}, \quad (2.4.1)$$

where  $v_{\text{reflex,P}}(t) = K_P \sin[2\pi(\phi(t))]$  is the radial part of orbital velocity of the planet and  $K_P$  is the planet’s radial velocity semi-amplitude, and  $v_{\text{rad}}$  is a constant additional velocity term to account for potential shifts. Since the spectra are in the pseudo-SRF (from step 2), here  $v_{\text{bary}}(t) = v_{\text{bary}}(t_{\text{mid.exp.}})$ . We can inject the models at different combinations of  $K_P$  and  $v_{\text{rad}}$ . Then, the relevant steps of the analysis are reapplied on the modeled transit sequence, i.e. 1, 5 and 6; step 2 is ignored since the reconstructed signal is already in the pseudo-SRF, so no

additional shifts are needed; step 3 is ignored since the spectra are already masked (from the reconstructed spectra) so no additional masks are applied (this also excludes the sigma clipping parts in steps 1 and 6); and step 4 is ignored since the spectra are already normalized to the Master-Out continuum, and the Master-Out stays the same. For step 5, we project the PCs obtained with the real observations onto the synthetic transit sequence, and remove this reconstruction from the synthetic sequence. The effect of this process is to replicate in the model spectra, to the extent possible, the subtraction of the planet signal that occurs when subtracting the PCs from the actual data. After this PCs subtraction, we then remove the mean light curve for each order, as was done on the observed data. This synthetic, PCs- and mean-subtracted time series ( $m_i$ ) is then used to calculate the cross-correlation and log-likelihood.

## 2.4.2. Cross-Correlation

The transmission model of the planet’s atmosphere can be cross-correlated with the observed planetary transmission spectra. If the model is adequate and the molecules in the model are present in the planet’s atmosphere, then there should be a peak in the cross-correlation function (CCF) at the expected radial velocity of the planet, and, in subsequent exposures that peak should shift according to the orbital velocity of the planet. From a sequence of several exposures in transit, we can then isolate the signal from the planet by combining the correlation signal that comes only from the right radial velocity in each exposure.

2.4.2.1. Algorithm. Based on the equations in Gibson et al. (2020) (and also used in Nugroho et al. 2021), the CCF can be written as,

$$\text{CCF}(\boldsymbol{\theta}) = \sum_{i=1}^N \frac{f_i \cdot m_i(\boldsymbol{\theta})}{\sigma_i^2}, \quad (2.4.2)$$

which is equivalent to a weighted CCF, where  $f_i$  are the observed values of the planetary transmission spectrum (described in Section 2.3.2) with associated uncertainties  $\sigma_i$ ,  $m_i$  are the model values (described in Sections 2.3.3 and 2.4.1), and  $\boldsymbol{\theta}$  is the model parameter vector, which includes the atmospheric model parameters and the applied orbital and systemic velocities (i.e. the Doppler shift at velocity  $v_P(t)$ ). The index  $i$  represents both time and wavelength, and the summation is done over  $N$  data points (number of unmasked pixels,  $N = 4564944$  for Tr1 and  $N = 6656359$  for Tr2). In practice, we first sum over wavelengths for each order, then over all orders, then over time. When summing over time, we apply a weighing to each spectrum according to a transit model of HD 189733 b, which takes into account that the planet’s signal is present only during the transit, while also being weaker during the ingress and egress. To build the transit model (transit depth at a given time), we use the equations from Mandel & Agol (2002), assuming a 4-parameters non-linear limb-darkening law (Claret, 2000) with theoretical

coefficients  $u_1 = 0.9488$ ,  $u_2 = -0.5850$ ,  $u_3 = 0.3856$ , and  $u_4 = -0.1318$ , based on 3D models of HD 189733 taken in Hayek et al. (2012) and valid for the H-band<sup>2</sup>. We note that using a variable limb-darkening law for the different parts of the spectrum would be more precise, but refrain from doing so for simplicity. In fact, going from a simple linear law to a non-linear one only had a minimal impact on the results. We used the ephemeris and system parameters listed in Table 2.2.

The uncertainties  $\sigma_i$  were determined by first calculating, for each spectral pixel, the standard deviation over time of the  $f_i$  values. However, to make sure we do not underestimate the noise by removing too many PCs, we take the standard deviation of  $f_i$  after the removal of only one PC, even though the final transmission spectra are corrected from three. This provides an empirical measure of the relative noise across the spectral pixels that captures not only the variance due to photon noise but also due to such effects as the telluric and background subtraction residuals, but it does not convey how the noise inherent to one spectrum compares to that of another. To capture this latter effect and include it in the  $\sigma_i$ , the dispersion values calculated above were multiplied, for each spectrum, by the ratio of the median relative photon noise of that spectrum divided by the median relative photon noise of all spectra. This is computed prior to normalization : the S/N variations across the night and the different orders are thus accounted for in this  $\sigma_i$  term, acting as a weight. The final uncertainty values  $\sigma_i$  thus reflect both temporal and spectral variability.

The CCF (equation 2.4.2) is calculated for every order of every spectrum in the transit sequence (time series) for an array of  $v_{\text{rad}}$  of size  $n_v$ . This gives a  $86 \times 49 \times n_v$  matrix (when combining the transits<sup>3</sup>) for a given  $K_P$  value and for each model tested. We then take the sum of the CCFs over all of the available SPIRou spectral orders, which gives a 2D cross-correlation map ( $86 \times n_v$ ) as a function of time and  $v_{\text{rad}}$ .

We observed that a few spectra, which happened to be those where the absolute value of the TelAz angle gradient was near its maximum value (at meridian passing), had noisier CCFs than the others. We chose to exclude the two spectra most affected by this, corresponding to exposures #26 and 27 of Tr2 (with TelAz angle gradient above  $18^\circ$ -per-exposure; Tr1 is not affected by this as the star did not cross the meridian). This exclusion was also applied to the analysis in the following sections.

From the total CCF as a function of the  $v_{\text{rad}}$  (again, for a given  $K_P$  and a given model), we compute the S/N by dividing the total CCF by its standard deviation, the latter being calculated by excluding the region around the peak ( $\pm 15 \text{ km s}^{-1}$ ). We expect the correlation peak to follow  $v_{\text{rad}} = 0 \text{ km s}^{-1}$  when  $K_P = K_{P,0} = 151.35 \text{ km s}^{-1}$ , the expected value (from the values in Table 2.2). A departure from 0 for  $v_{\text{rad}}$  could be indicative of high-altitude winds in the planet's

---

2. It was computed with the `BATMAN PYTHON` package (Kreidberg, 2015).

3. We combine the two transits by simply concatenating the two CCF time series, order-per-order. There are 36 spectra in Tr1 and 50 in Tr2, for a total of 86.

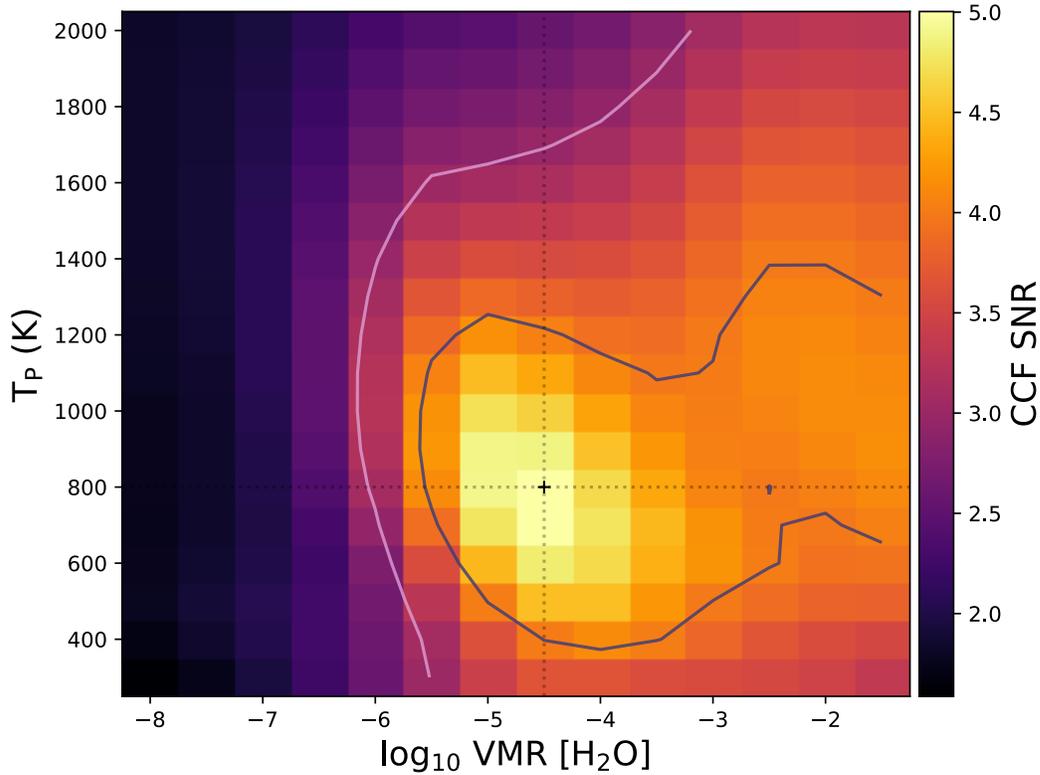


FIGURE 2.4 Maximum S/N of the CCF of different models as a function of  $\log_{10}$  VMR[ $\text{H}_2\text{O}$ ] and  $T_P$  for  $P_{\text{cloud}} = 10^{-0.5}$  bar,  $K_P = K_{P,0}$  and  $v_{\text{rad}} = v_{\text{peak}}$ , using the combined transits. The contour shows where the S/N drops by 1 and 2 from the maximum. A maximum S/N of 5.0 is found at  $\log_{10}$  VMR[ $\text{H}_2\text{O}$ ] = -4.5,  $T_P = 800$  K, and  $P_{\text{cloud}} = 10^{-0.5}$  bar.

atmosphere and other dynamic effects (such as planet rotation or a asymmetry between the eastward and westward planet limbs that are probed during the transit) that were not included in the model itself, causing a net blue shift in the observed planetary lines.

2.4.2.2. Results. We computed the cross-correlation of a grid of models containing only  $\text{H}_2\text{O}$  as their major constituent. In principle, the molecular or atomic species that are present in both the stellar and planetary atmospheres will introduce signal contamination via the RME when trying to isolate the planetary signal. However, we ignored this effect as water is not expected to be present in the atmosphere of a star like HD 189733. The explored parameter values on the grid were the following : volume-mixing ratios (VMRs) for water ranging from  $\log_{10}$  VMR[ $\text{H}_2\text{O}$ ] = -8 to -1.5 with steps of 0.5, equilibrium temperatures  $T_P$  (input to the Guillot T-P profile) ranging from 300 to 2000 K in steps of 100 K and gray cloud deck top pressures of  $\log_{10} P_{\text{cloud}}$  (bar) from -5 to 2, also in steps of 0.5. For this, the  $v_{\text{rad}}$  range goes from

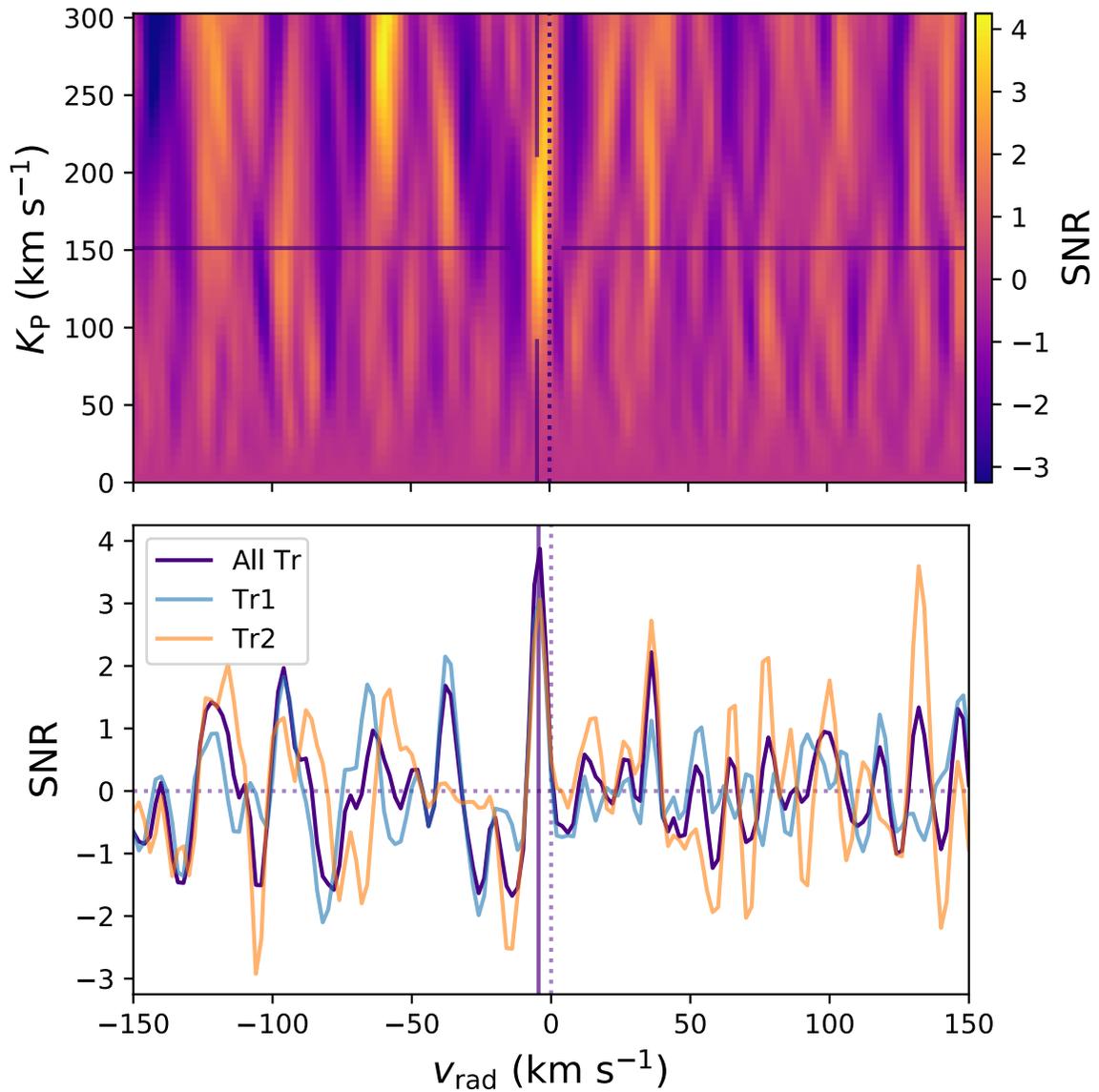


FIGURE 2.5 [Top panel] Color-coded S/N map of the co-added cross-correlation signal from both transits and the best-fit model ( $\log_{10} \text{VMR}[\text{H}_2\text{O}] = -4.5$ ,  $T_p = 800$  K, and  $P_{\text{cloud}} = 10^{-0.5}$  bar) as a function of the zero-phase planet radial velocity and the orbital radial velocity semi-amplitude  $K_P$ . [Bottom panel] Horizontal cut of S/N map at the known  $K_P$  of the planet. The peak is found at  $-4.5$   $\text{km s}^{-1}$  with a S/N of 4.0.

$-70$  to  $70$   $\text{km s}^{-1}$ , with  $2$   $\text{km s}^{-1}$  steps (71 steps in total; roughly the size of a SPIRou pixel,  $\simeq 2.3$   $\text{km s}^{-1}$ ), but  $K_P$  is fixed at  $K_{P,0}$ . When combining the two transits, we obtain the maximum CCF S/N grid shown in Figure 2.4. The best-fit model has  $\log_{10} \text{VMR}[\text{H}_2\text{O}] = -4.5$ ,  $T_p = 800$  K and  $P_{\text{cloud}} = 10^{-0.5}$  bar, at  $v_{\text{rad}} = -4.5$   $\text{km s}^{-1}$ .

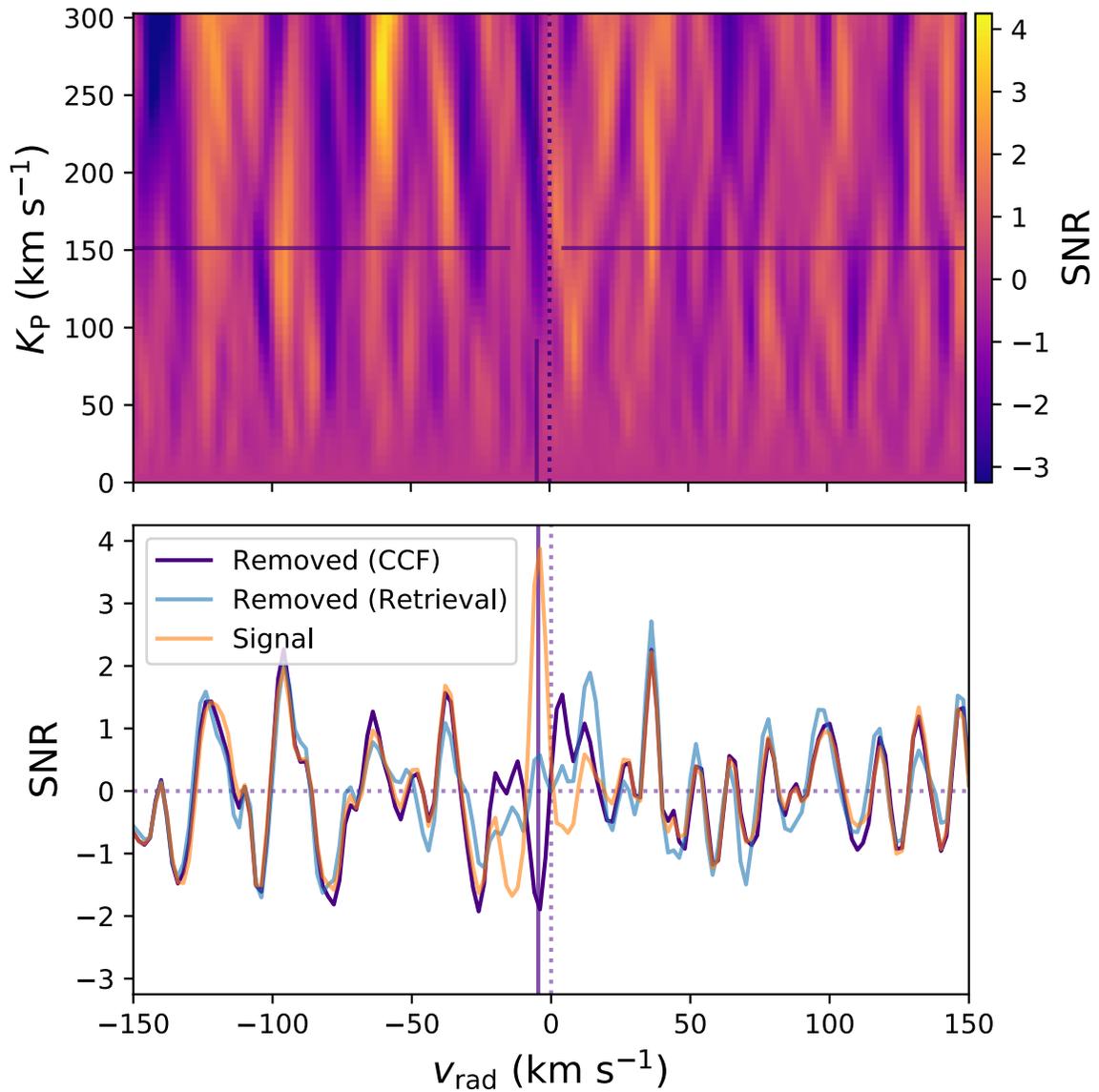


FIGURE 2.6 [Top panel] Same as Figure 2.5, but with the injected negative best-fit model from the CCF grid, with the same color scale. [Bottom panel] The original signal of the combined transits (orange curve, same as Figure 2.5) is reduced considerably when the CCF best-fit model is injected negatively in the data prior to the analysis (indigo curve). When looking at the negative injection of the best-fit model found in the retrieval with the  $\ln \mathcal{L}$  (see Section 2.4.4), we get a cleaner signal removal (blue curve). In both cases, the systematic noise structures away from the peak are nearly identical as before.

To obtain a better estimation of the detection significance of the best fit model, we repeated all calculations for different combinations of  $v_{\text{rad}}$  and  $K_P$  (yielding different projected planet radial velocities during the transit sequence). We extended the  $v_{\text{rad}}$  range from -150 to 150  $\text{km s}^{-1}$

(still with  $2 \text{ km s}^{-1}$  steps) to get a broader view of the map, while the  $K_P$  values range from 0 to  $2K_{P,0} \simeq 303 \text{ km s}^{-1}$ , with  $\sim 6 \text{ km s}^{-1}$  steps (51 steps in total). This gives us the total cross-correlation signal map as a function of  $K_P$  and  $v_{\text{rad}}$ . Again, to produce a S/N detection map, we divided the cross-correlation map by the standard deviation of the full map computed by excluding the region around the peak ( $\pm 15 \text{ km s}^{-1}$  in  $v_{\text{rad}}$  and  $\pm 70 \text{ km s}^{-1}$  in  $K_P$ ). We did not include the negative  $K_P$  region, but verified and confirmed that no spurious signal could be found at  $-K_{P,0}$ .

The full CCF map of the best-fit model is shown in Figure 2.5. We detect  $\text{H}_2\text{O}$  with an S/N = 4.0 at  $v_{\text{rad}} = -4.5 \text{ km s}^{-1}$  (at  $K_{P,0}$ ). We note that the position in  $K_P$  has large uncertainties ( $153^{+57}_{-37} \text{ km s}^{-1}$ ), but these are expected due to the small variation in orbital velocity of the planet during the transit. It is nonetheless still consistent with  $K_{P,0}$ . Moving forward, we assume  $K_P$  is well known and equal to  $K_{P,0}$ . Also, the peak is clearly blue-shifted from the planetary rest frame ( $v_{\text{rad}} = 0$ ). This is discussed in more details in Section 2.5.3. The CCFs of each individual spectrum using the best-fit model are shown in Figure 2.7.

When considering only Tr1 or only Tr2, we get a weaker detection : a peak S/N of  $\sim 3.2$  is observed at  $v_{\text{rad}} = -4.7 \text{ km s}^{-1}$  for Tr1, while for Tr2, we get a peak S/N of  $\sim 3.1$  at  $v_{\text{rad}} = -4.3 \text{ km s}^{-1}$ . These results, even though they are weaker detections on their own, are consistent within the uncertainties with one another and with the combined transits detection. This further supports the idea that our detected peak is physical as opposed to being a noise artefact. Again, the low S/N from Tr2 could partly be explained by the high variation in exposure S/N (see Figure 2.2).

To make sure that the atmospheric parameters that were retrieved by the CCF S/N method are reliable, an injection test was performed. By injecting a negative version of the best-fit model in the data, the original CCF peak should nicely disappear. We did such an injection in the telluric-corrected data using the best-fit model, applying  $v_{\text{rad}} = -4.5 \text{ km s}^{-1}$ , and reapplied all the analysis steps (from Section 2.3.2). We confirmed that when doing so the CCF peak signal indeed disappears, as can be seen in Figure 2.6. The same injection test was also verified for the  $\ln \mathcal{L}$  approach of the next section (not shown).

As a validation of our adopted approach of applying the data processing steps to the model spectra before calculating the CCF/ $\ln \mathcal{L}$ , we repeated our analysis, but using “unprocessed” models<sup>4</sup>, and then performed the same negative injection test. When the best-fit models found by the “unprocessed models”<sup>5</sup> analysis were negatively injected in the data before carrying out the analysis with unprocessed models, the detection peak could not be adequately removed

---

4. Only a median normalization was applied, but most importantly no PCs were removed to the synthetic sequence.

5. The best-fit parameters differed by a factor  $10^3$  in VMR[ $\text{H}_2\text{O}$ ], while having similar  $T_P$  compared to those based on the “processed models” analysis.

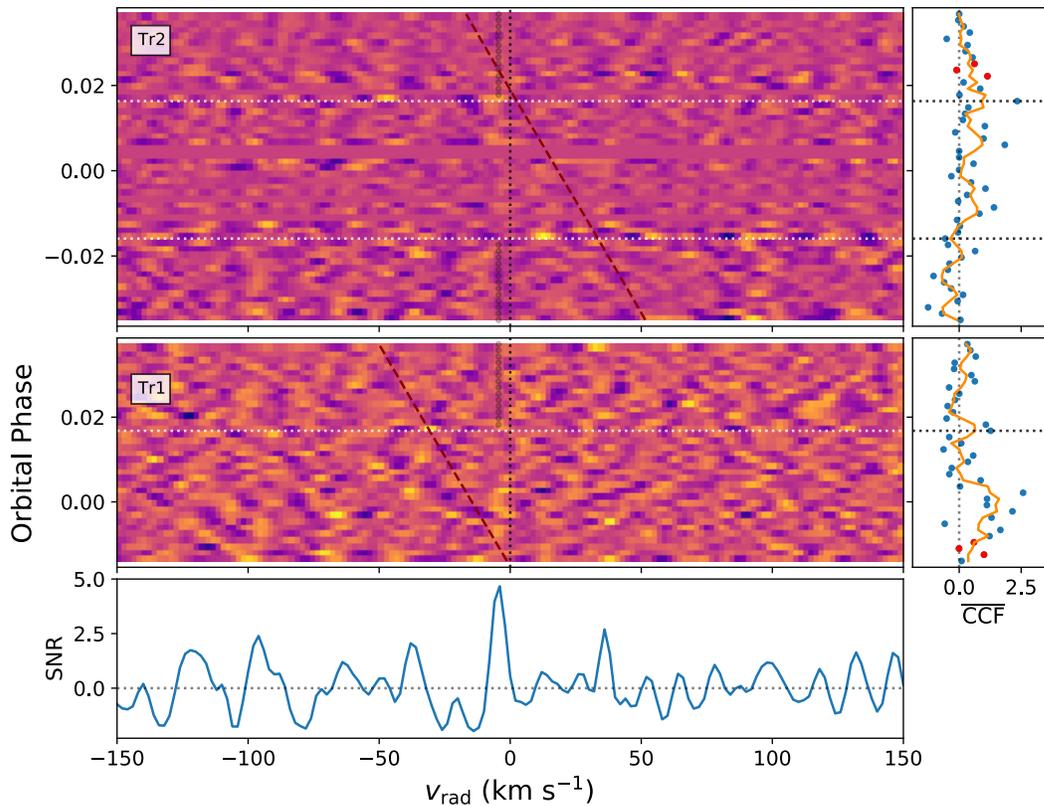


FIGURE 2.7 [Left] CCF (normalized by the dispersion of the region away from the peak) from individual spectra as a function of  $v_{\text{rad}}$  for Tr2 (top panel) and Tr1 (middle panel), in the planet rest frame. Red dashed lines show the position of the BERV, while whites (left) and black (right) dotted lines shows the ingress and egress positions. The vertical dotted lines show the planet path for  $v_{\text{rad}} = 0 \text{ km s}^{-1}$  and  $v_{\text{rad}} = v_{\text{peak}}$  (only shown out-of-transit to increase visibility during transit). The bottom panel shows the combined transits CCF S/N curve. [Right] Mean CCF for a 3-pixel column bin centered on  $v_{\text{peak}}$  (blue points) and the 3-exposures binned signal (orange line). The red dots show where the BERV crosses the peak position by less than  $2.3 \text{ km s}^{-1}$ . We can see that the detection is not significantly affected by these points.

(not shown here). This indicates that applying the data processing to the models yields best-fit parameters that are more representative of what is initially present in the data, and thus that this approach is to be preferred.

We looked at another statistical metric to differently quantify the significance of our  $\text{H}_2\text{O}$  detection. We performed a Student's  $t$ -test (Student, 1908), similarly to what was done in other studies before (e.g. Birkby et al., 2017; Brogi et al., 2018; Cabot et al., 2019; Alonso-Floriano

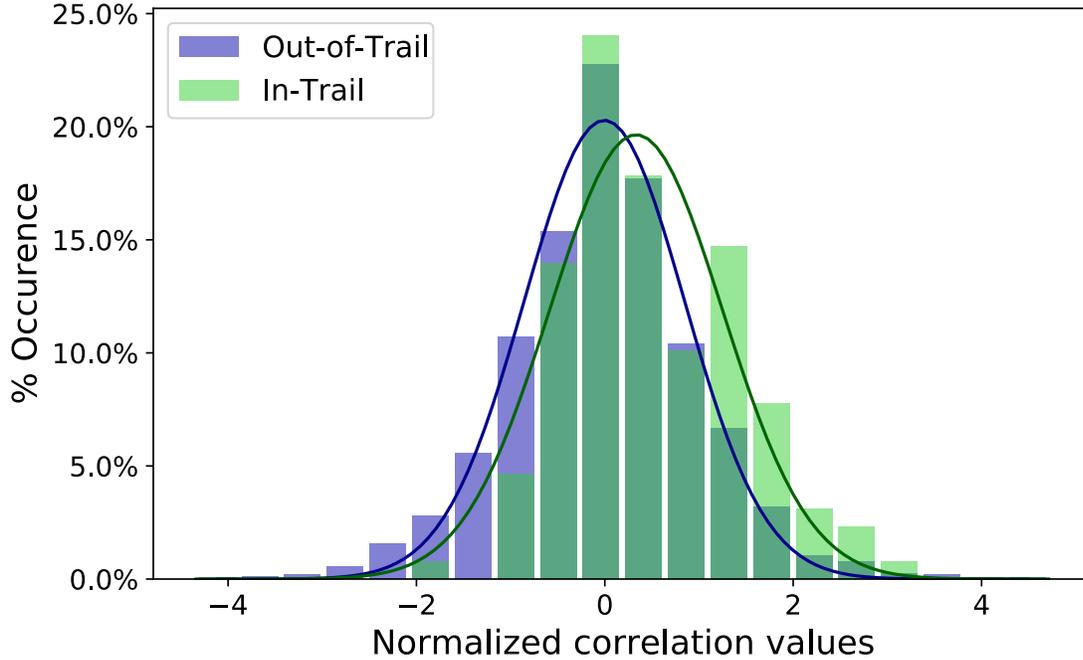


FIGURE 2.8 Generalized  $t$ -test results : Distributions of cross-correlation values, normalized by the dispersion of the out-of-trail regions, away from (out-of-trail, blue), and near (in-trail, green) the planet radial velocity (with  $K_{P,0}$  and  $v_{\text{rad}} = -4.5 \text{ km s}^{-1}$ ), with their associated best-fit Gaussian distributions, each with their corresponding mean and variance (blue and green lines, respectively). A detection of the transmission signal of the planet is expected to shift the in-trail distribution to higher correlation values, and this is what we observe, with the two distributions differing at the  $5.9\sigma$  level.

et al., 2019; Webb et al., 2020). The  $t$ -test verifies the null hypothesis that two Gaussian distributions have the same mean value. In our case, the two distributions to be compared are drawn from our correlation map. On one hand, we have the in-trail distribution of CCF values, i.e. where the planet signal should be, that we took to be correlation values from 3-pixels wide columns centered at  $v_{\text{rad}} = -4.5 \text{ km s}^{-1}$  (as done in Birkby et al. 2017 and suggested in Cabot et al. 2019), and on the other hand, we have the out-of-trail distribution, which includes the CCF values more than  $10 \text{ km s}^{-1}$  away from  $v_p(t)$ , where there should be no planet signal. The  $t$ -test then evaluates the likelihood that these two samples were drawn from the same distribution.

We computed this test using the in transit CCFs from Figure 2.5 with  $K_P = K_{P,0}$  and  $v_{\text{rad}} = v_{\text{peak}} = -4.5 \text{ km s}^{-1}$ . The results are shown in Figure 2.8, for the combined transits, and indicate that the in-trail distribution is different from the out-of-trail one at the level of  $5.9\sigma$ , which further supports our detection. As a sanity check, we also compared the in- and out-of-trail distributions for the out-of-transit CCFs. As expected, we found that both distributions/samples

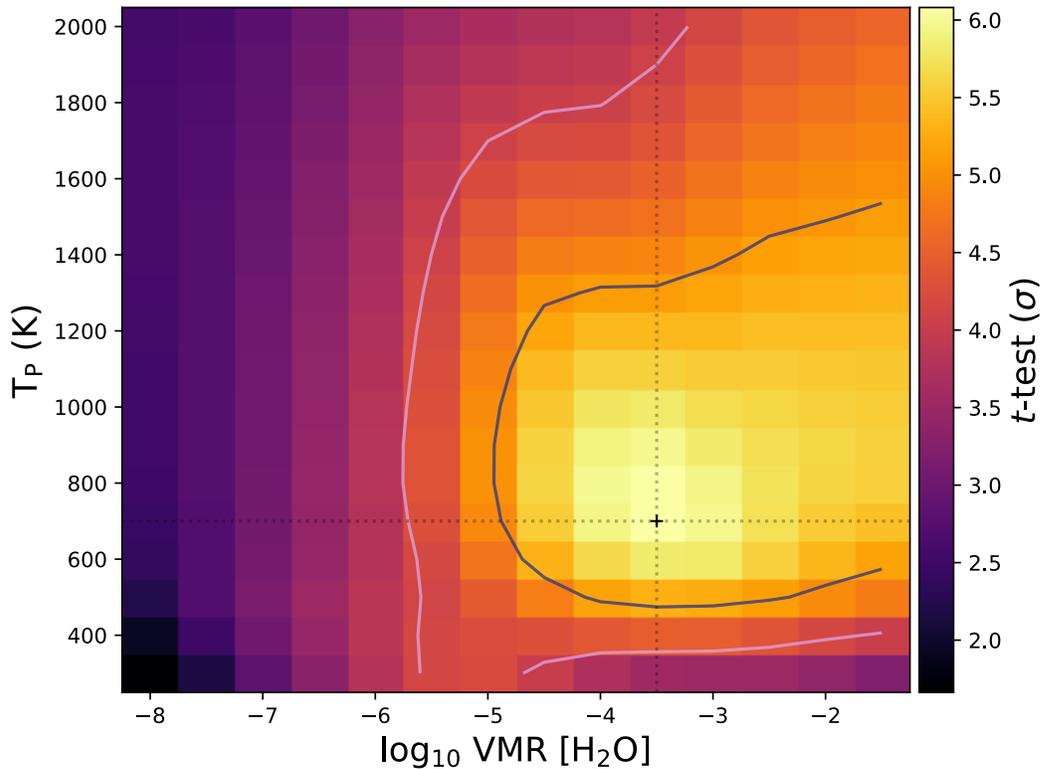


FIGURE 2.9 Same as Figure 2.4 but for the  $t$ -test value. The contour shows where  $\sigma$  drops by 1 and 2 from the maximum. A maximum  $6.1\sigma$  is found for  $\log_{10}\text{VMR}[\text{H}_2\text{O}] = -3.5$ ,  $T_{\text{P}} = 700\text{ K}$ , and  $P_{\text{cloud}} = 10^{-0.5}\text{ bar}$ .

had similar mean and variance (different at  $0.4\sigma$ ; not shown here). This Student  $t$ -test is complementary to the CCF S/N method as it uses a different approach to assess uncertainties from the underlying noise. The high  $t$ -test detection significance gives good confidence in our detection despite the somewhat modest CCF S/N mentioned above. As a mean of comparison, we also computed the  $t$ -test for the same CCF grid described above and the results are shown in Figure 2.9. The  $t$ -test yields similar  $T_{\text{P}}$  and  $P_{\text{cloud}}$ , but seems to favor overall higher VMRs.

### 2.4.3. Log-Likelihood Mapping Method

In addition to the CCF and  $t$ -test calculations, we also computed the log-likelihood values for the different models. We used the approach from Gibson et al. (2020) by introducing scaling factors  $\alpha$  and  $\beta$  for the model and noise, such that the model  $m_i$  becomes  $\alpha\beta m_i$ <sup>6</sup>, but we fixed

<sup>6</sup>.  $\alpha$  accounts for any scaling uncertainties of the model and  $\beta$  accounts for potential scaling uncertainties of the white noise.

$\alpha = 1$  as water is not expected to be present in an extended atmosphere for HD 189733 b (see also Brogi & Line 2019). By nulling the partial derivative of the standard  $\ln \mathcal{L}$  with respect to  $\beta$  (i.e., removing the dependency on  $\beta$ ), the log-likelihood function can be written as :

$$\ln \mathcal{L} = -\frac{N}{2} \ln \left[ \frac{1}{N} \left( \sum \frac{f_i^2}{\sigma_i^2} + \sum \frac{m_i^2}{\sigma_i^2} - 2 \sum \frac{f_i m_i}{\sigma_i^2} \right) \right], \quad (2.4.3)$$

where the summation is implied over  $i$  (both spectral pixels and time). By inspecting this equation, we can see that the first term is a constant (for a given data set, related to the variance of the data) and the last term is related to the CCF from above (eq. 2.4.2). The middle term is related to the variance of the model and introduces the main difference compared to the previous CCF approach. The effect of this term is to reduce the likelihood value for models with higher variances; so for comparable fits to the data, the lower variance (more conservative) model should be preferred.

2.4.3.1. Results. For every model in our grid (same as above), we computed the  $\ln \mathcal{L}$  for  $v_{\text{rad}}$  going from  $-70$  to  $+70 \text{ km s}^{-1}$  and  $K_{\text{P}} = K_{\text{P},0}$ . The highest  $\ln \mathcal{L}$  value occurs around  $\log_{10} \text{VMR}[\text{H}_2\text{O}] = -4.5$ ,  $T_{\text{P}} = 500 \text{ K}$ , and  $P_{\text{cloud}} = 1 \text{ bar}$ , with  $v_{\text{rad}} = -4.7 \text{ km s}^{-1}$ .

Then, to establish how the best-fit model fares compared to the others, we looked at the Bayesian Information Criterion (BIC).<sup>7</sup> In this formalism, the model with the lowest BIC is preferred (here, taken to be the best-fit model), and the evidence against certain models (with higher BIC) is usually described as “very strong” when  $\Delta \text{BIC} = 2 \Delta \ln \mathcal{L}$  is greater than 10 (Kass & Raftery, 1995). Figure 2.10 shows this quantity for the best-fit value of  $P_{\text{cloud}} = 1 \text{ bar}$ , as a function of  $T_{\text{P}}$  and  $\log_{10} \text{VMR}[\text{H}_2\text{O}]$ . Even though there is a degeneracy between the temperature and the VMR (as well as between the VMR and the cloud deck top pressure, better seen in Figure 2.11), we can still put some constraints on the parameters.

## 2.4.4. Markov Chain Monte Carlo

It quickly becomes time consuming and computationally expensive to compute the full log-likelihood maps for bigger model grids (either in sampling or number of parameters). Instead of computing the log-likelihood for the full grid, we used a Markov Chain Monte Carlo (MCMC) approach that allows us to compute a posterior distribution of each atmospheric parameter and obtain estimates of their uncertainties. The MCMC sampling was done using the `python` library `emcee` (Foreman-Mackey et al., 2013), which implements the affine-invariant ensemble sampler by Goodman & Weare (2010).

<sup>7</sup>.  $\text{BIC} = k \ln(n) - 2 * \ln \mathcal{L}$ ; where  $k$  is the number of parameters,  $n$  is the number of data points and  $\ln \mathcal{L}$  is our log-likelihood value for our model for each combination of parameters. For fixed values of  $k$  and  $n$ , the lowest BIC value is related to the highest  $\ln \mathcal{L}$ .

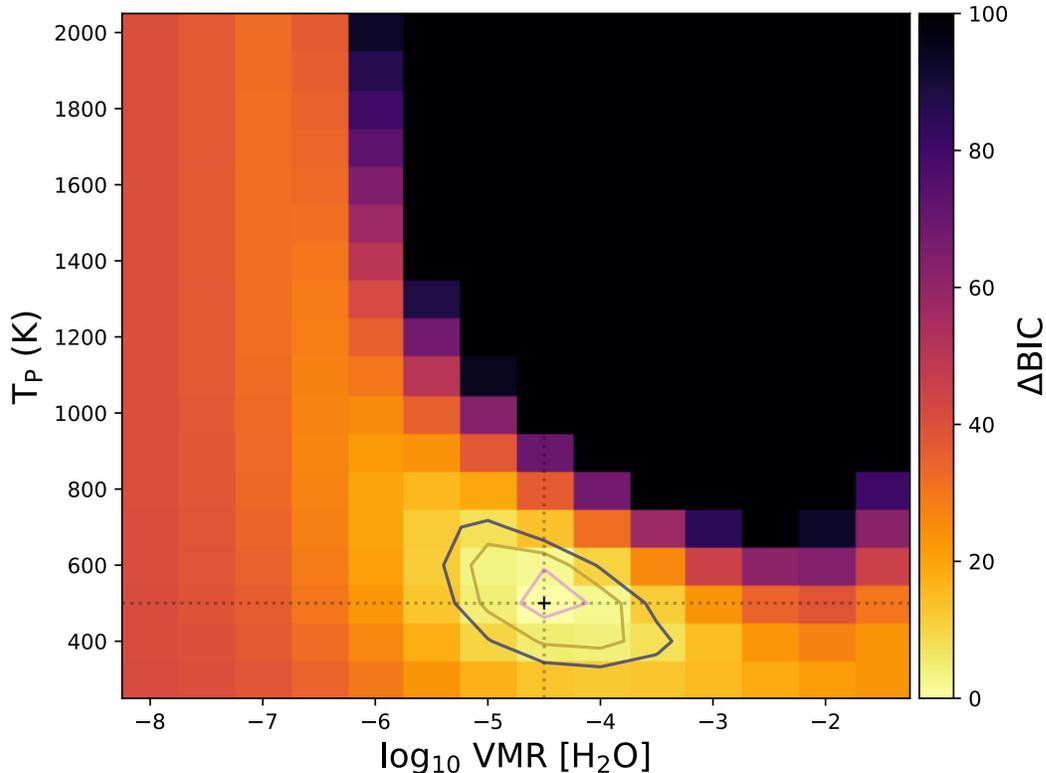


FIGURE 2.10 Change in BIC value from the peak,  $\Delta\text{BIC} = 2 \Delta \ln \mathcal{L}$ , for the different models in the grid, for a cloud deck top pressure of 1 bar, at  $K_{\text{P},0}$  and  $v_{\text{rad}} = v_{\text{peak}}$ . The contours shown increase by 2, 6, and 10 from the minimum value. For a increase in BIC of 2, 6, and 10, respectively, the best-fit model is typically regarded as being positively, strongly and very strongly favored compared to the other models.

To save time on the generation of models, we applied an N-d linear interpolation<sup>8</sup> over our grid for the parameters  $\text{VMR}[\text{H}_2\text{O}]$ ,  $T_{\text{P}}$  (K) and  $P_{\text{cloud}}$  (bar) dimensions. We are aware that these interpolated models are not as accurate,<sup>9</sup> nor are they self-consistent, but they offer a reasonable approximation to derive useful constraints on the atmosphere parameters.

Using the combined transits sequences, we ran 50 walkers for 4500 steps and varied five parameters :  $\log_{10}\text{VMR}[\text{H}_2\text{O}]$ ,  $T_{\text{P}}$ ,  $\log_{10}P_{\text{cloud}}$ ,  $K_{\text{P}}$  and  $v_{\text{rad}}$ . We chose to include  $K_{\text{P}}$  in the retrieval even though this value is well known, to see if it could be retrieved independently and how it would affect  $v_{\text{rad}}$ .

8. Using SciPy interpolation module `RegularGridInterpolator`, which only accepts regular grids, but may have uneven spacing.

9. As compared to models computed with exact parameters, worst-case interpolated models show differences of at most 62 ppm for the strongest lines, with the biggest errors coming from the cloud top pressure interpolation.

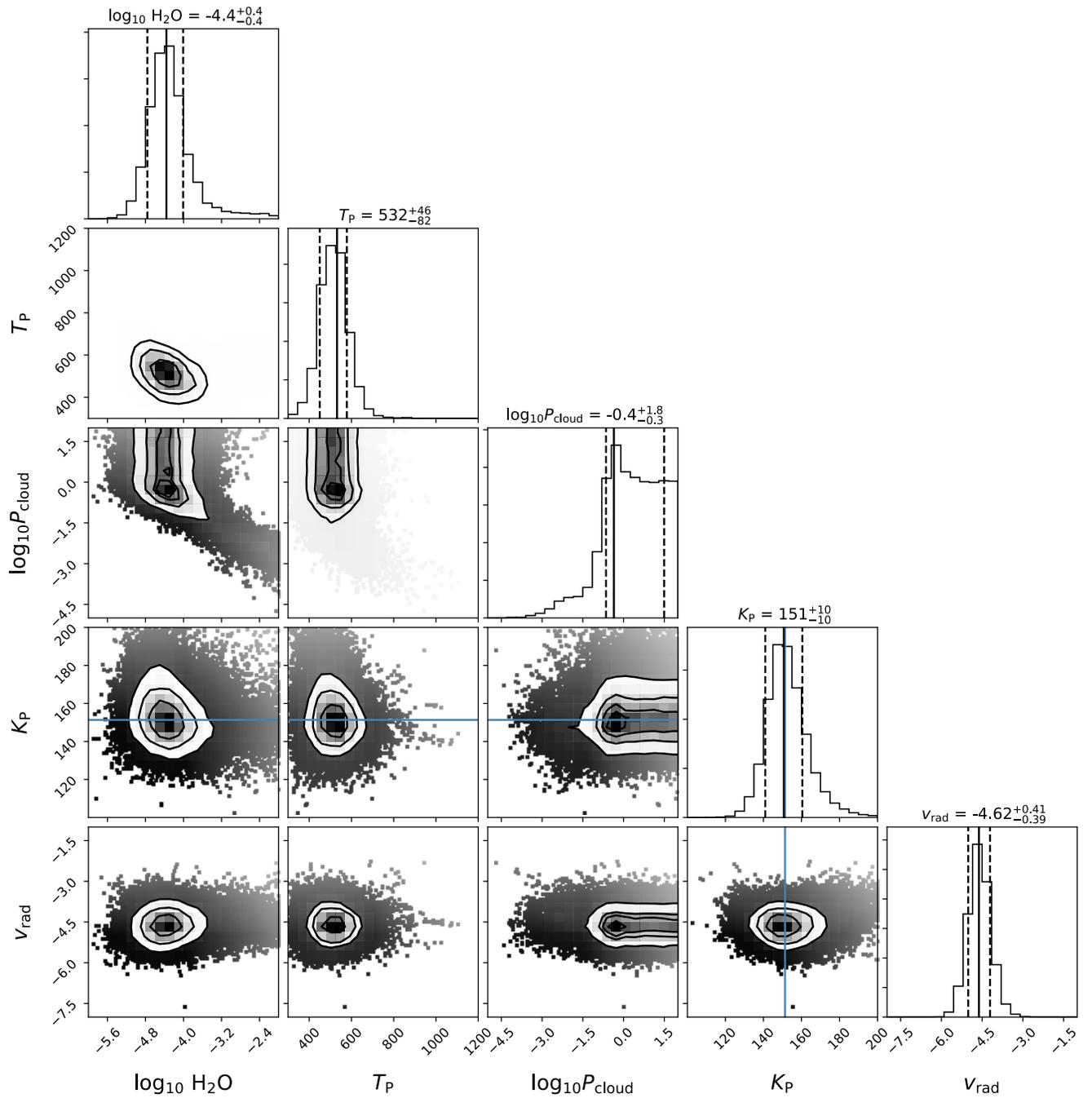


FIGURE 2.11 Posterior probability distributions for the parameters of the MCMC fit on the HD 189733 b data using an atmosphere with H<sub>2</sub>O and a gray cloud deck as described in Section 2.3.3. The blue line shows the position of  $K_{P,0}$ .

Table 2.3. MCMC Retrieval Parameter Priors and Results for HD 189733 b

Parameter	Uniform Prior	Results	Unit
$\log_{10} \text{H}_2\text{O}$	$\mathcal{U}(-8, -1.5)$	$-4.4^{+0.4}_{-0.4}$	
$T_{\text{P}}^a$	$\mathcal{U}(300, 2000)$	$532^{+46}_{-82}$	K
$\log_{10} P_{\text{cloud}}$	$\mathcal{U}(-5, 2)$	$-0.4^{+1.8}_{-0.3}$	bar
$K_{\text{P}}$	$\mathcal{U}(100, 200)$	$151^{+10}_{-10}$	$\text{km s}^{-1}$
$v_{\text{rad}}^b$	$\mathcal{U}(-20, 10)$	$-4.62^{+0.41}_{-0.39}$	$\text{km s}^{-1}$

Note. — The marginalized parameters from the likelihood analysis with  $\pm 1\sigma$  error. <sup>a</sup> The retrieved  $T_{\text{P}}$  is the equilibrium temperature input to the Guillot T-P and represents a temperature of  $727^{+63}_{-111}$  K at 1 bar. <sup>b</sup> The uncertainty shown in the last column results from the MCMC analysis only. The uncertainty of  $\pm 0.21 \text{ km s}^{-1}$  from  $v_{\text{sys}}$  must be added to it in quadrature to obtain the total uncertainty on  $v_{\text{rad}}$ .

We are discarding the first 1000 steps of each chain as burn-in<sup>10</sup>. The resulting posterior probability distributions of the five parameters are shown in Figure 2.11. The priors that were used and the resulting marginalized values for each parameter are listed in Table 2.3. The  $1\sigma$  uncertainties correspond to the range of parameters containing 68% of the MCMC samples.

We find relatively good constraints on the parameters. The favored  $T_{\text{P}}$  and VMR values are close to previous values in the literature (see Section 2.5.2), both from low- and high-resolution data. The resulting T-P profile is shown in Figure 2.12. The cloud top pressure is consistent with relatively deep clouds (at pressures around  $\gtrsim 0.2$  bar) (McCullough et al., 2014; Pinhas et al., 2019). However, the apparent degeneracy between the water abundance and cloud top pressure leads to probability at slightly higher cloud altitudes. This is not a detailed exploration of cloud modeling as in Barstow (2020), since we did not include the impact of having different cloud fractions (a thorough cloud analysis is beyond the scope of the paper), but our results seem to land somewhere in between their results with and without cloud fractions.

Nonetheless, the CCF peak disappear when the best-fit from this retrieval is injected negatively in the data, as shown in Figure 2.6, meaning that it represents well the planetary signal. The cleaner signal removal indicates that this method is probably better at identifying the best model.

## 2.5. Discussion

Our results demonstrate that water is detected in each of our two SPIRou transmission spectra of HD 189733 b (although more marginally in Tr2). Combining both data sets, the standard

10. This is where the chains were overall converged.

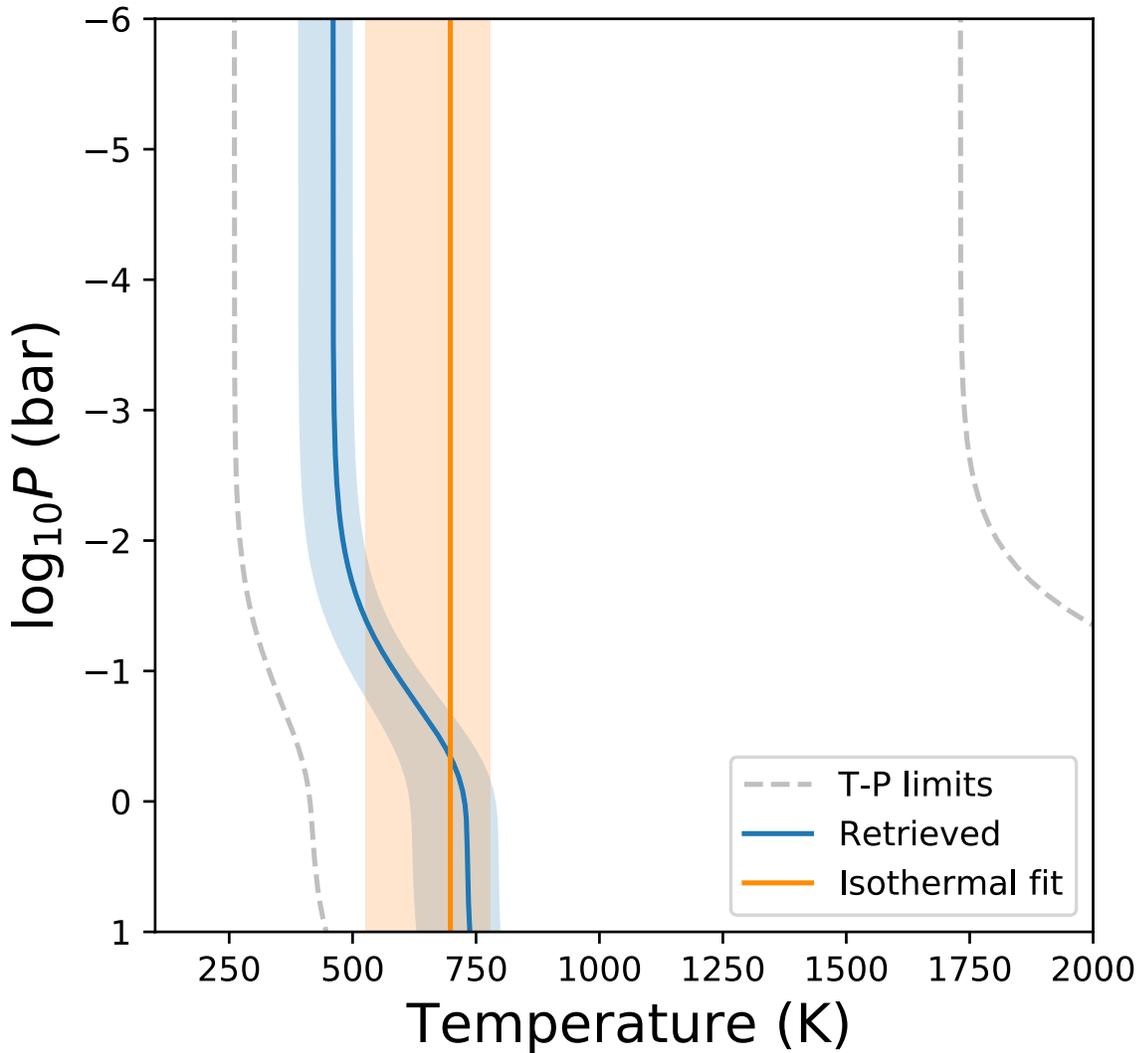


FIGURE 2.12 Retrieved Guillot T-P profile (blue) where the shaded region represents the  $1\sigma$  uncertainties, compared to the limits of the prior (gray dashed lines). The retrieved profile has a temperature of  $727^{+63}_{-111}$  K at 1 bar. Overplotted is the retrieved isothermal profile (orange), with  $T_P = 698^{+80}_{-172}$  K, that was obtained from a previous run of the MCMC analysis, that yielded nearly identical values for the other parameters.

cross-correlation method yields a detection with  $S/N = 4.0$ , based on the dispersion of values in the full CCF  $K_P$  vs  $v_{\text{rad}}$  map, or  $5.9\sigma$  according to the Student's  $t$ -test. The peak in the CCF map is located near the expected  $K_P$  and  $v_{\text{rad}}$ , adding confidence to our detection. Similarly, the log-likelihood approach yields a significant detection, with a  $\Delta\text{BIC} > 10$  compared to water-less

Table 2.4. Overview of the results from the model grid for the different methods, using their respective best-fit model.

Method	Significance <sup>a</sup>			Combined Best-fit Model Parameters			
	Tr1	Tr2	Combined	$v_{\text{rad}}$ (km s <sup>-1</sup> )	$\log_{10}\text{VMR}[\text{H}_2\text{O}]$	$T_{\text{P}}$ (K)	$\log_{10}P_{\text{cloud}}$ (bar)
CCF	4.4	3.0	5.0	-4.5	-4.5	800	-0.5
$t$ -test	5.4	4.0	6.1	-4.3	-3.5	700	-0.5
$\ln \mathcal{L}$ <sup>b</sup>	24	20	41	-4.7	-4.5	500	0

Note. — <sup>a</sup> The metrics used are SNR for the CCF method,  $\sigma$  for the  $t$ -test, and  $\Delta\text{BIC}$  for  $\ln \mathcal{L}$ . <sup>b</sup> The  $\Delta\text{BIC}$  value is computed between the best-fit model (i.e. with the largest  $\ln \mathcal{L}$  value) and a water-less unconstrained model (i.e. with  $\log_{10}\text{VMR}[\text{H}_2\text{O}] = -8$ ).

models. A comparison of the results obtained from the model grid for all methods is presented in Table 2.4.

While we obtain a convincing detection using all approaches – cross-correlation,  $t$ -test and log-likelihood – and retrieve similar/consistent best-fit model parameters, we observe some discrepancies for the “acceptable range” of models from each method (see Figures 2.4, 2.9 and 2.10). The main discrepancy is that using the CCF (or  $t$ -test, which is based on the CCF maps), models with a combination of higher temperature and higher  $\log_{10}\text{VMR}[\text{H}_2\text{O}]$  seem acceptable (upper-right region on the grid of Figure 2.4), while they are strongly disfavored in the log-likelihood analysis. The main difference between the two methods is that the log-likelihood takes into account not only the variance of the observed spectrum, but also the variance of the model, effectively penalizing models with a high variance, and thus preventing higher variance in the model to compensate for a weaker signal in the data. For that reason, we tend to lend more credence to this approach when it comes to constraining the atmosphere parameters. The CCF and  $t$ -test remain important, however, for confirming a detection of the planet atmosphere.

### 2.5.1. Effects of telluric residuals

As explained above, we masked spectral regions affected by deep telluric lines before computing the CCF and log-likelihood, where deep was defined here as lines reaching below 30 to 35% transmission in their core. Applying this mask provided a gain in CCF peak S/N <sup>11</sup>, even though the spectra were corrected for telluric absorption by `APER0`. This likely indicates that the telluric correction is not perfect and that a low level of residuals persists. Masking these residuals reduces the number of data points that can be used, but doing so still improves the

11. We focus this discussion on the CCF since this is one of the most common metric in the literature, but the same general behaviour was observed for the two other metrics, i.e.  $t$ -test and  $\ln \mathcal{L}$ , in what follows, except when specified.

significance of the detection. As an indication of the magnitude of the effect, if we do not mask deep telluric lines (beyond the masking of saturated lines done by `APER0`), the peak CCF S/N is decreased by  $\sim 2$  when removing a single PC in step 6 from Section 2.3.2 (but decreases by only  $\sim 0.6$  when removing three) as compared to the value with additional masking. This implies that there most likely is telluric residual contamination, but that removing three PCs is an effective way to subtract them.

The presence of telluric residuals would also affect the different orders to varying degrees, as the telluric lines are not spread uniformly – in density and strengths – across the orders. In our analysis above, we included all spectral orders<sup>12</sup> in our CCF and log-likelihood calculations, but in some previous studies, the spectral regions most affected by tellurics have sometimes been omitted from the calculations. We tested this to see the effect on the peak CCF S/N, and indeed found that masking some orders could lead to improvements. For instance, when masking 18% of the orders<sup>13</sup>, the peak of the 2D map can increase to 5.3 (from the value of 4.0 when including all orders), while its position does not change much. However, this cut did not improve the  $t$ -test nor  $\ln \mathcal{L}$  values and even led to smaller values. Also, for the model grids, the order cut moves the favored models to slightly higher temperatures.

Nonetheless, we did not adopt this approach for our analysis above, to avoid the risk that the specific choice of orders to omit would introduce a bias in the favored atmosphere model. We preferred to use as much as possible of the spectral range provided by `SPIROU`, at the cost of a somewhat lower detection S/N. Still, the retrieved S/N of 4.0 in CCF for the combined transits is quite smaller than previous detections, e.g. the water detection with an S/N of 6.6 from Alonso-Floriano et al. (2019), coming from a single transit using `CARMENES` data, with lower overall S/N, a smaller spectral range, but a slightly better spectral resolution. Our smaller S/N could be explained by multiple reasons. First, it could be due to the subjective character of the determination of the S/N with the map standard deviation (Cabot et al., 2019). Variables such as the extent on which the CCF map and its standard deviation are computed, and the  $v_{\text{rad}}$  and  $K_p$  sampling, can greatly change the results (our goes from  $\pm 150 \text{ km s}^{-1}$  with steps of  $2 \text{ km s}^{-1}$ , while theirs goes from  $\pm 65 \text{ km s}^{-1}$  with steps of  $1.3 \text{ km s}^{-1}$ ). Computing the S/N from the 2D map or a 1D array can also change the results. The  $t$ -test is also subject to arbitrariness through the RV sampling that can increase the sample size of both the in- and out-of-trail distributions (Cabot et al., 2019), but is usually preferred to determine detection significance (e.g., Birkby et al., 2017; Brogi et al., 2018). Second, the lower S/N could also be due to our conservative choice not to remove any spectral orders. Third, in our approach the CCF is computed with an injected and reprocessed modeled transit sequence; the reprocessing thus applied to the model can diminish

---

12. Again, except those heavily filled with saturated telluric lines that are automatically masked by `APER0`.

13. The orders centered at 1.48, from 1.75 to 1.97, plus 2.20 and  $2.26 \mu\text{m}$ , which correspond mostly orders between the H and K bands and a few others.

the overall line contrast, which would decrease the CCF (but will also be more representative of the processed signal buried in the observed data). Finally, the position of the tellurics relative to the planet signal can affect the results. If the tellurics are not properly corrected/removed, any residuals crossing the planetary path could erroneously increase the retrieved planet signal.

## 2.5.2. Retrieved Atmospheric Parameters

Our retrieval results indicate that the probed region of the atmosphere has a temperature that is much lower than the equilibrium temperature ( $T_P = 531_{-81}^{+47}$  K, equivalent to  $T = 727_{-111}^{+63}$  K at 1 bar, while  $T_{eq} = 1200$  K), a water abundance that is slightly sub-solar ( $\log_{10} \text{VMR}[\text{H}_2\text{O}] = -4.4_{-0.4}^{+0.4}$ , solar abundance would be around  $-3$  and  $-3.3$  for  $T_P \lesssim$  and  $> 1200$  K, respectively, Madhusudhan et al. 2014, see their Figure 3), and a mostly clear atmosphere (cloud deck at pressure  $\gtrsim 0.2$  bar). These results are in line with previous studies at both low and high spectral resolution. For instance, values of  $\log_{10} \text{VMR}[\text{H}_2\text{O}]$  between  $-3.3$  and  $-5$  for  $T_P$  between  $700$ – $800$  K, or fixed at the equilibrium value, have been reported (McCullough et al., 2014; Madhusudhan et al., 2014; Welbanks et al., 2019; Pinhas et al., 2019)<sup>14</sup> at low-resolution; and between  $-3$  and  $-5$  for models with an atmosphere temperature of around  $500$  K (from parameterized, non-isothermal T-P profiles) at high resolution (Birkby et al., 2013; Brogi et al., 2016, 2018; Alonso-Floriano et al., 2019; Brogi & Line, 2019). Our results are thus within the range of values found previously, and consistent with sub-solar abundances.

In Madhusudhan (2012), they list two interpretations for low inferred water abundances in hot Jupiters such as HD 189733 b. 1) The water measurement could be representative of the bulk oxygen abundance of the planet’s atmosphere, which would indicate a sub-stellar oxygen abundance, and thus, an overall low planetary atmosphere metallicity. This would be potential evidence for a water-poor formation scenario (with solar relative abundances of the elements, i.e.  $C/O \sim 0.5$ ). 2) For planets with  $T_{eq} \gtrsim 1200$  K, CO becomes a major reservoir for atmospheric oxygen. Therefore, our retrieved sub-stellar water abundance could instead be due to a super-stellar  $C/O$  ratio for HD 189733 b (in which case more than half of the oxygen would be locked in CO). This would potentially indicate a formation history dominated by high  $C/O$  gas beyond the ice-line, as opposed to solid accretion of oxygen-rich material (Cridland et al., 2019). We note, though, that this work focuses solely on the detection of water which prevents us from lifting the degeneracy between these two scenarios.

When looking at previous strong CO detections and their tendency towards slightly super-solar abundance (at around  $\log_{10}[\text{CO}] \simeq -3$ ; de Kok et al., 2013a; Brogi et al., 2016; Cabot et al., 2019; Flowers et al., 2019; Brogi & Line, 2019), combined with our retrieved sub-solar  $\text{H}_2\text{O}$  abundance, this is consistent with high  $C/O$  ratio (closer to 1 than solar). This seems to

---

14. Most of those results were obtained using a retrieval with a 6-parameter temperature-pressure (T-P) profile, as in Madhusudhan & Seager (2009).

be on trend with other hot Jupiters, such as HD 209458 b (Gandhi et al., 2020) and  $\tau$  Boo b (Pelletier et al., 2021, ; and references therein).

Recent studies based on 3D atmosphere models have shown that retrieval results based on 1D models, as have been used here, may be biased (Flowers et al., 2019; MacDonald et al., 2020; Beltz et al., 2021). For instance, MacDonald et al. (2020) analytically showed that a compositional difference between the morning and evening side of the terminator could lead to retrieved 1D uniform temperatures that are many hundreds of degrees colder than the real average terminator temperature. This could possibly explain why, similarly to the previously mentioned results, we retrieved a low  $T_P$  value. They also show that species distributed uniformly around the terminator (such as  $H_2O$ ) are biased towards higher abundances. In addition, the omission from our models of CO and other molecules likely to be present in the atmosphere of HD 189733 b could also bias the retrieved abundance of water (Brogi & Line, 2019; MacDonald et al., 2020). The retrieved parameters in the present work should thus be used with caution - although the magnitude of the aforementioned effects is likely to be small compared to our (relatively large) uncertainties in  $T_P$  and  $\log_{10} \text{VMR}[H_2O]$ .

### 2.5.3. Radial velocity offset

In many previous studies of HD 189733 b, a net blue-shift of the planet atmosphere absorption signal was observed and usually attributed to the presence of large-scale high-altitude day-to-night winds and the presence of eastward jets. Such high-altitude day-to-night winds can be partially accounted for by the  $v_{\text{rad}}$  term in equation 2.4.1 (when they are not already included in the atmosphere models used). Additionally, the planetary rotation/eastward jets should increase the blue shift velocity, due to the potential asymmetries between the morning and evening sides of the terminator. The combination of a hotter and more inflated blue-shifted evening terminator and a colder and less inflated morning red-shifted side would lead towards globally blue-shifted velocities (Flowers et al., 2019). For HD 189733 b, an overall (limb integrated) shift of around  $-1.5$  to  $-2 \text{ km s}^{-1}$  was obtained in several dynamical studies, with both low- and high-resolution (Louden & Wheatley, 2015; Brogi et al., 2016, 2018; Flowers et al., 2019). Larger shifts were also observed with values around  $-4 \text{ km s}^{-1}$  (Alonso-Floriano et al., 2019; Brogi & Line, 2019; Damiano et al., 2019) that are still in agreement with the other results, considering their larger individual uncertainties ( $\sim 1 - 2 \text{ km s}^{-1}$ ).

Here we measure a net blue-shift of  $-4.62_{-0.44}^{+0.46} \text{ km s}^{-1}$ , which is somewhat larger than what was observed in many previous studies, but still consistent with the results of Alonso-Floriano et al. and Damiano et al. within the uncertainties. We observe a similar blue-shift using both the CCF and log-likelihood approaches, and in each individual transit, which suggests that it is a

real feature of our data. It was also present and at this value regardless of the analysis parameters that were applied (eg. telluric fraction masking, number of PCs, etc.).

Our  $-4.6 \text{ km s}^{-1}$  measurement is compatible within  $1-2\sigma$  with the blueshift of  $-5.3_{-1.4}^{+1.0} \text{ km s}^{-1}$  observed for the trailing limb of the planet by Loudon & Wheatley (2015), which could imply that the signal we measure is dominated by the trailing (evening) limb. To investigate this, we checked if we could see a difference in the value of  $v_{\text{rad}}$  between the ingress and the egress in our data, as was done in Loudon & Wheatley (2015); Flowers et al. (2019). We indeed saw a signal that was more red shifted during ingress and more blue shifted during egress, which would support the above explanation, but the significance of this signal is too low to claim it with confidence. In addition, an ingress-egress difference could have biased our measured net  $v_{\text{rad}}$  as the ingress of both of our transits suffered in some way from technical problems. A small part of the ingress from Tr1 is missing, due to a late start of the observation, while there was a substantial drop in data S/N during the first half of Tr2. This imbalance between the ingress and egress coverage and data quality could have biased our net  $v_{\text{rad}}$  values toward a value more representative of the egress, i.e. more blue-shifted.

Our larger net blue shift could, perhaps, be explained by the different line list that we used to build our models. However, when computing the CCF or  $\ln \mathcal{L}$  with the best-fit model built with the HITEMP 2010 water line list (instead of Exomol), we get a virtually identical shift, but with a slightly smaller S/N. Also, the cross-correlation between two models with identical parameters but using two different water line lists (i.e. HITEMP and Exomol) shows a relative shift of only  $\sim 0.023 \text{ km s}^{-1}$ .

Taken at face value, a net blue-shift as high as we have measured, if coming from large-scale winds, would be hard to explain by current GCMs, based on the modeled CCF curves from Flowers et al. (2019) for different rotation speed. Including the contribution of the planet rotation can lead to higher values of net blue-shift, but for a synchronous rotation (period of 2.2 days, consistent with the results of Flowers et al. (2019)) this effect would not be enough : a CCF peak value between  $-2$  and  $-1 \text{ km s}^{-1}$  would be expected (see their Figure 12). A planet rotation period of  $\lesssim 1.30$  days (faster than the synchronous case, yielding an equatorial rotational velocity of  $\gtrsim 4.85 \text{ km s}^{-1}$  and a maximum wind speed of  $2.76 \text{ km s}^{-1}$ ) could lead to a net blue-shift closer to  $-5 \text{ km s}^{-1}$  when considering the rotation and winds together (last row from Figure 12 in Flowers et al. 2019). Such a fast rotation (faster than synchronous) should lead to a significant rotational broadening. Using the CCF approach with models that were rotationally broadened in a simple manner (Gaussian broadening), the narrowest profiles (no rotation) were always favored, as expected and pointed out in Brogi et al. (2016). On the other hand, using the log-likelihood approach with the same broadening yielded rotation speed of  $v_{\text{rot}} = 2.0_{-2.0}^{+1.5} \text{ km s}^{-1}$  which is broadly consistent with synchronous rotation (equatorial  $v_{\text{rot}} = 2.849 \text{ km s}^{-1}$ ; Flowers et al. 2019), but

higher rotation speeds were not excluded by the  $\ln \mathcal{L}$ . An overall smaller net blue shift is observed and decreases with higher rotation speed, but that is still too high to be explained from synchronous rotation. Nonetheless, this seems to indicate that rotational information is present in the data and could partially explain the high blue shift.

Finally, errors in the absolute SPIRou wavelength calibration are not expected to be larger than a few  $0.1 \text{ km s}^{-1}$  and are thereby not likely to be the cause of the high blue shift that we report.

## 2.6. Concluding remarks

We presented one of the first exoplanet atmosphere characterization data sets obtained with the SPIRou high-resolution near-infrared spectrograph recently mounted on the 3.6 m CFHT within the frame work of the Large Observing Program called the SPIRou Legacy Survey totalling 300 CFHT nights, and demonstrated that this instrument is capable of characterizing exoplanet atmospheres via transmission spectroscopy.

Our analysis of the data revealed a detection of water in the atmosphere of HD 189733 b at an S/N of 4.0 using the standard CCF method, and  $5.9\sigma$  from a Student’s  $t$ -test, and that a good atmosphere model selection can be achieved using the more recent log-likelihood mapping methods from Brogi & Line (2019) and Gibson et al. (2020). This allowed us to put constraints on the temperature, water volume mixing ratio, and gray cloud deck level of the planet atmosphere. Our results favour temperatures that are significantly lower than  $T_{\text{eq}}$  ( $T_{\text{P}} \simeq 532 \text{ K}$ , equivalent to  $T = 727_{-111}^{+63} \text{ K}$  at 1 bar), a sub-solar  $\log_{10} \text{VMR}[\text{H}_2\text{O}] \simeq -4.4$ , and a gray cloud deck at pressures of  $> 0.2 \text{ bar}$  - all within the ranges of values found previously in the literature. Moreover, the absorption signal of the planetary atmosphere is detected with a substantial blue-shift of  $-4.62_{-0.44}^{+0.46} \text{ km s}^{-1}$  - a value  $0.6\text{--}3 \text{ km s}^{-1}$  bluer than previous literature results. The cause of this difference remains unclear, but it seems unlikely that such a high blue-shift can arise only from winds in the planet’s atmosphere. The planet’s rotation and/or a signal dominated by the trailing limb of the planet could be at play here.

This analysis focused mainly on the detection of water, but much more remains to be done with these data. Given the spectral coverage of SPIRou, which extends to the end of the  $K$  band, it should be possible to probe for the presence of carbon monoxide in the atmosphere of HD 189733 b. This analysis will require an explicit treatment of the Rossiter-McLaughlin effect, as CO is also present in the stellar spectrum which will introduce time-varying CO features during transit. The SPIRou spectral range also covers the helium metastable triplet at  $1.083 \mu\text{m}$ , which probes the extended atmosphere. This absorption was previously detected for HD 189733 b (Salz et al., 2018; Guilluy et al., 2020) using CARMENES and GIARPS (GIANO

+ HARPS) data, respectively, but is also seen in SPIRou (Donati et al., 2020) and it will be interesting to further compare these results.

This work relied on 1D atmosphere models with an isothermal temperature structure. In the future, it will be interesting to analyze the data using more complex models, in particular; 3D models that can capture variations in composition or conditions across the planet limb, as well as global atmosphere dynamics. A free spectral retrieval, rather than a retrieval constrained on a model grid as done here, would also be beneficial.

Finally, we point out that the throughput of SPIRou in the  $Y$  and  $J$  band was recently increased by a factor of 2 and 1.6, respectively, following the replacement of its set of rhomboid prisms (see Donati et al. 2020). This should improve the overall quality of future data compared with what was presented here, and should consequently enable better characterization of planetary atmospheres. In particular, the significant increase in throughput in the blue should greatly aid probes of the  $1.083\ \mu\text{m}$  metastable helium line used to study the extended atmosphere. As improvements in the data reduction software of SPIRou are continually being made, including better correction of telluric absorption, the overall capabilities of SPIRou for characterizing exoplanet atmospheres are also expected to improve.

### *Acknowledgements*

The authors acknowledge financial support for this research from the Natural Science and Engineering Research Council of Canada (NSERC), the Institute for Research on Exoplanets (iREx) and the University of Montreal (UdeM). These results are based on observations obtained at the Canada-France-Hawaii Telescope (CFHT) which is operated from the summit of Maunakea by the National Research Council of Canada, the Institut National des Sciences de l'Univers of the Centre National de la Recherche Scientifique of France, and the University of Hawaii. The observations at the Canada-France-Hawaii Telescope were performed with care and respect from the summit of Maunakea which is a significant cultural and historic site. We thank the reviewer for her/his careful reading and feedback that helped us improve the quality of this manuscript. SP and ADB acknowledge funding from the Technologies for Exo-Planetary Science (TEPS) CREATE program. JFD acknowledges funding from the European Research Council under the H2020 research & innovation programme (grant #740651 NewWorlds). VB acknowledges that this work has been carried out in the frame of the National Centre for Competence in Research "PlanetS" supported by the Swiss National Science Foundation (SNSF). This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (project FOUR ACES grant agreement No 724427; project SPICE DUNE; grant agreement No 947634). XD, GH and EM acknowledge funding from the French National Research Agency (ANR) under contract number ANR-18-CE31-0019 (SPlaSH). XD acknowledges funding in the framework of the Investissements d'Avenir program (ANR-15-IDEX-02), through the funding of the "Origin of Life" project of the Univ. Grenoble-Alpes. BK acknowledges funding from

the European Research Council under the European Union's Horizon 2020 research and innovation programme (grant agreement #865624, GPRV). J.H.C.M. is supported in the form of a work contract funded by Fundação para a Ciência e a Tecnologia (FCT) with the reference DL 57/2016/CP1364/CT0007; and also supported from FCT through national funds and by FEDER-Fundo Europeu de Desenvolvimento Regional through COMPETE2020-Programa Operacional Competitividade e Internacionalização for these grants UIDB/04434/2020 & UIDP/04434/2020, PTDC/FIS-AST/32113/2017 & POCI-01-0145-FEDER-032113, PTDC/FIS-AST/28953/2017 & POCI-01-0145-FEDER-028953, PTDC/FIS-AST/29942/2017. NS acknowledges funding FCT - Fundação para a Ciência e a Tecnologia through national funds and by FEDER through COMPETE2020 - Programa Operacional Competitividade e Internacionalização by these grants : UID/FIS/04434/2019 ; UIDB/04434/2020 ; UIDP/04434/2020 ; PTDC/FIS-AST/32113/2017 & POCI-01-0145-FEDER-032113 ; PTDC/FIS-AST/28953/2017 & POCI-01-0145-FEDER-028953 ; PTDC/FIS-AST/28987/2017 & POCI-01-0145-FEDER-028987.

# Chapitre 3

---

## CO or no CO? Narrowing the CO abundance constraint and recovering the H<sub>2</sub>O detection in the atmosphere of WASP-127 b using SPIRou

Ce manuscrit a été soumis à la revue arbitrée *The Astronomical Journal*.

ANNE BOUCHER<sup>1</sup>, DAVID LAFRENIÈRE<sup>1</sup>, STEFAN PELLETIER<sup>1</sup>, ANTOINE DARVEAU-BERNIER<sup>1</sup>, MICHAEL RADICA<sup>1</sup>, ROMAIN ALLART<sup>1,\*</sup>, ÉTIENNE ARTIGAU<sup>1,2</sup>, NEIL J. COOK<sup>1,2</sup>, FLORIAN DEBRAS<sup>3</sup>, RENÉ DOYON<sup>1,2</sup>, ERIC GAIDOS<sup>4</sup>, BJÖRN BENNEKE<sup>1</sup>, CHARLES CADIEUX<sup>1</sup>, ANDRES CARMONA<sup>5</sup>, RYAN CLOUTIER<sup>6,y</sup>, PÍA CORTÉS-ZULETA<sup>7</sup>, NICOLAS B. COWAN<sup>8,9</sup>, XAVIER DELFOSSE<sup>5</sup>, JEAN-FRANÇOIS DONATI<sup>3</sup>, PASCAL FOUQUÉ<sup>3,10</sup>, THIERRY FORVEILLE<sup>5</sup>, KONSTANTIN GRANKIN<sup>11</sup>, GUILLAUME HÉBRARD<sup>12</sup>, JORGE H. C. MARTINS<sup>13</sup>, EDER MARTIOLI<sup>14,15</sup>, ADRIEN MASSON<sup>16</sup>, AND SANDRINE VINATIER<sup>16</sup>

<sup>1</sup> Institut de Recherche sur les Exoplanètes, Université de Montréal, Département de Physique, C.P. 6128 Succ. Centre-ville, Montréal, QC H3C 3J7, Canada

<sup>2</sup> Observatoire du Mont-Mégantic, Université de Montréal, Notre-Dame-des-Bois, J0B 2E0, Canada

<sup>3</sup> Université de Toulouse, CNRS, IRAP, 14 av. Belin, 31400 Toulouse, France

<sup>4</sup> Department of Earth Sciences, University of Hawai'i at Manoa, Honolulu, HI 96822 USA

<sup>5</sup> Université Grenoble Alpes, CNRS, IPAG, 38000 Grenoble, France

<sup>6</sup> Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA, 02138, USA

<sup>7</sup> Aix Marseille Univ, CNRS, CNES, LAM, Marseille, France

<sup>8</sup> Department of Earth & Planetary Sciences, McGill University, 3450 rue University, Montréal, QC H3A 0E8, Canada

<sup>9</sup> Department of Physics, McGill University, 3600 rue University, Montréal, QC H3A 2T8, Canada

<sup>10</sup> CFHT Corporation; 65-1238 Mamalahoa Highway, Kamuela, HI 96743, USA

<sup>11</sup> Crimean Astrophysical Observatory, Department of Stellar Physics, Nauchny, 298409, Crimea

<sup>12</sup> Institut d’astrophysique de Paris, UMR7095 CNRS, Université Pierre & Marie Curie, 98bis boulevard Arago, 75014 Paris,

France

<sup>13</sup> Instituto de Astrofísica e Ciências do Espaço, Universidade do Porto, CAUP, Rua das Estrelas, 4150-762 Porto, Portugal

<sup>14</sup> Laboratório Nacional de Astrofísica, Rua Estados Unidos 154, Itajubá, MG 37504-364, Brazil

<sup>15</sup> Institut d’astrophysique de Paris, UMR7095 CNRS, Sorbonne Université, 98 bis bd Arago, 75014 Paris, France

<sup>16</sup> LESIA, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, Université de Paris, 5 place Jules Janssen, 92195 Meudon, France

**RÉSUMÉ.** Des mesures précises des abondances chimiques dans les atmosphères planétaires sont nécessaires pour bien contraindre les histoires de formation et d’évolution des exoplanètes. Une étude récente de l’atmosphère de WASP-127 b, une sous-Saturne chaude gonflée qui orbite son étoile hôte de type solaire en 4.2 d, utilisant HST et Spitzer a révélé un spectre de transmission riche en structures spectrales avec un excès inhabituellement fort absorption dans la bande photométrique de *Spitzer* à  $4,5 \mu\text{m}$ . Cependant, comme la résolution et la couverture en longueur d’onde de ces instruments sont limitées, il n’a pas été possible de démêler la contribution du  $\text{CO}_2$  et/ou du CO qui induirait un tel signal, et des rapports C/O faible et élevé restaient équiprobables. Nous présentons ici les premières observations par spectroscopie de transmission haute résolution de cette planète dans le proche infrarouge ( $0,9\text{--}2,5 \mu\text{m}$ ), à l’aide de SPIRou au CFHT, dans le but de démêler le CO du  $\text{CO}_2$  grâce à la bande de CO à  $2,3 \mu\text{m}$ . Du  $\text{H}_2\text{O}$  est détecté à un niveau de  $\sim 5\sigma$  et un signal potentiel compatible avec le OH est détecté à  $2.8\sigma$ . À partir d’une analyse de récupération conjointe SPIRou + HST + *Spitzer*, nous avons pu mettre une limite supérieure à l’abondance de CO de  $\log_{10}[\text{CO}] < -4.3$ , excluant ainsi un scénario riche en CO, et estimer un  $\log_{10}[\text{CO}_2]$  de  $-4.6 \pm 0.4$ , qui est le niveau nécessaire pour expliquer l’absorption observée à  $4,5 \mu\text{m}$ . Nous également pu poser des contraintes sur l’abondance d’autres molécules principales porteuses de C, O et N. Nos résultats favorisent les scénarios en déséquilibre chimique, et conduisent à un  $\text{C/O} = 0.04^{+0.03}_{-0.03}$  qui correspond à une atmosphère fortement appauvrie en carbone. Nous discutons de certains scénarios de formation qui seraient en accord avec nos résultats quelque peu surprenants. Des observations supplémentaires à haute et basse résolution, en particulier avec JWST et surtout du côté jour, seront nécessaires pour confirmer ces résultats et améliorer encore notre compréhension de ce monde inhabituel.

**Mots-clés :** Planètes et satellites : atmosphères — Planètes et satellites : individuels (WASP-127 b) — Méthodes : analyse de données — Techniques : spectroscopiques

**ABSTRACT.** Precise measurements of chemical abundances in planetary atmospheres are necessary to constrain the formation and evolution histories of exoplanets. A recent study of the atmosphere of WASP-127 b, a close-in puffy sub-Saturn orbiting its solar-type host star in 4.2 d,

using HST and *Spitzer* revealed a feature-rich transmission spectrum with unusually strong excess absorption in the  $4.5\ \mu\text{m}$  *Spitzer* band. However, the limited resolution and wavelength coverage of these instruments could not disentangle the contribution of  $\text{CO}_2$  and/or  $\text{CO}$  that would induce such a signal, and both a low and a high C/O ratio were possible. Here, we present the first high-resolution transmission spectroscopy observations of this planet in the near-infrared ( $0.9\text{--}2.5\ \mu\text{m}$ ), using SPIRou at the CFHT, with the goal to disentangle  $\text{CO}$  from  $\text{CO}_2$  through the  $2.3\ \mu\text{m}$   $\text{CO}$  band.  $\text{H}_2\text{O}$  is detected at a  $t$ -test level of  $\sim 5\sigma$  and a tentative signal consistent with  $\text{OH}$  is observed at  $2.8\sigma$ . From a joint SPIRou + HST + *Spitzer* retrieval analysis, we were able to put an upper limit on the mixing ratio of  $\log_{10}[\text{CO}] < -4.3$ , thus ruling out a  $\text{CO}$ -rich scenario, and estimate a  $\log_{10}[\text{CO}_2]$  of  $-4.6 \pm 0.4$ , which is the level needed to match the excess absorption seen at  $4.5\ \mu\text{m}$ . We also place constraints on the abundances of other major C-, O-, and N-bearing molecules. Our results favour disequilibrium chemistry scenarios, and lead to a C/O ratio  $= 0.04^{+0.03}_{-0.03}$  that corresponds to a highly carbon-depleted atmosphere. We discuss some formation scenarios that would be in line with our somewhat surprising results. Additional observations at high and low-resolution, particularly with JWST and especially of the planet's emission from the day-side, will be needed to confirm these results and further improve our understanding of this unusual world.

**Keywords :** Planets and satellites : atmospheres — Planets and satellites : individual (WASP-127 b) — Methods : data analysis — Techniques : spectroscopic

### 3.1. Introduction

The atmospheric metallicity and carbon-to-oxygen ratios (C/O) of exoplanets are tracers of their atmospheric chemistry, their formation location, and their migration history (Madhusudhan et al., 2014; Madhusudhan, 2019). This stems from the fact that different volatile compounds (i.e.  $\text{H}_2\text{O}$ ,  $\text{CO}_2$ ,  $\text{CH}_4$ ,  $\text{CO}$ ,  $\text{NH}_3$ , etc.) condense at varying temperatures, and thus at different distances from the host star in the natal proto-planetary disk. This then introduces variations in the C/O ratio of the gaseous and condensed phases of matter which accrete onto a forming planet (Öberg et al., 2011). Therefore, the measurement of chemical abundances is a powerful tool to explore the origins and evolution pathways of exoplanets. Hot giant exoplanets with extended atmospheres are prime targets for these measurements thanks to their large sizes and scale heights relative to their host star, and high atmospheric temperature that keep most carbon and oxygen bearing molecules in vapour phase. This facilitates the remote detection of atomic and molecular species via transmission or emission spectroscopy. However, precise measurement of the C/O ratio remains a challenging task as it necessitates simultaneous constraints of all major C- and O-bearing molecules contained in the planet's atmosphere and an appropriate understanding of the chemistry (e.g., Tsai et al., 2021). The *Hubble Space Telescope* Wide Field

Camera 3 (HST/WFC3) G141 grism, widely used to study exoplanetary atmospheres (e.g., Sing et al., 2016; Welbanks et al., 2019; Pinhas et al., 2019, and references therein), provides access to the water-band feature at  $1.4\ \mu\text{m}$ , but not to CO or CO<sub>2</sub> — the bulk of whose spectral features lie further into the infrared. Complementary information could also be extracted from *Spitzer* photometry, but this is limited to a few wide band-pass photometric data points — rendering impossible the disentanglement of contributions from CO and CO<sub>2</sub>. Addressing this fundamental challenge requires additional observations which expand the limited wavelength range and spectral resolution of these instruments.

A first attempt at measuring the C/O by combining space- and ground-based observations was performed by Madhusudhan et al. (2011), where they inferred a carbon-rich atmosphere ( $C/O \geq 1$ ) in the day-side of the ultra-hot Jupiter WASP-12 b. This sparked a debate, wherein some contested this result (e.g., Cowan et al., 2012; Crossfield et al., 2012; Swain et al., 2013; Line et al., 2014; Kreidberg et al., 2015), while others further affirmed it (e.g., Föhning et al., 2013; Stevenson et al., 2014). Since then, analyses of multiple transit observations for other planets have been carried out to measure their C/O ratios. Most retrieved values have been closer to solar ( $C/O=0.54$ ) than to 1 (Line et al., 2014; Benneke, 2015). Due to the large uncertainties on many C/O measurements, it is also unclear whether the general trend, if there is any, points toward super-*stellar* values (i.e., when compared to host stars as opposed to the Sun), or indeed whether planetary C/O ratios are consistent with those of their host stars (Brewer et al., 2017). More recently, studies using ground-based high-resolution cross-correlation spectroscopy, with the combined constraints on H<sub>2</sub>O and CO, showed evidence that certain hot Jupiters (HD 209458 b, HD 189733 b, Tau Boötis b) have elevated C/O ratios (Brogi & Line, 2019; Gandhi et al., 2019; Jacobbe et al., 2021; Pelletier et al., 2021; Boucher et al., 2021) while others like WASP-77 A b have C/O ratios closer to solar (Line et al., 2021). As these various inferred atmospheric metallicities and C/O ratios could imply different formation conditions, more in depth measurements could reveal population trends on how and where in the protoplanetary disk giant planets form.

Here, we aim to measure the elemental abundances of the major molecules in the atmosphere of WASP-127 b to gain insight into its formation and migration history. We use ground-based high-dispersion spectroscopy (HDS) in the near-infrared (NIR), which has proven to be a powerful probe of the composition of exoplanetary atmospheres (e.g., Snellen et al., 2010; Birkby et al., 2013; Nugroho et al., 2017; Brogi et al., 2018; Alonso-Floriano et al., 2019; Boucher et al., 2021; Line et al., 2021).

This paper is organized as follows : the next section is an overview of existing studies of WASP-127 b and the remaining open science questions regarding this unusual system. In Section 3.3, we present the observational setup and data. In Section 3.4 we briefly describe the data

Tableau 3.1 Overview of previous detections on WASP-127 b

Studies	Instruments	H <sub>2</sub> O	Na	K	Hazes	Grey Clouds	Others
Palle et al. (2017a)	NOT/ALFOSC	...	Y (hints)	N	Rayleigh slope	MC/PC <sup>a</sup>	Hints of TiO, VO
Chen et al. (2018)	GTC/OSIRIS	Y (hints)	Y	Y	Y	MC/PC <sup>a</sup>	Li
Žák et al. (2019)	ESO/HARPS	...	Y	...	...	...	...
Welbanks et al. (2019)	GTC/OSIRIS	Y (weak)	Y	Y	...	...	...
Seidel et al. (2020b)	ESO/HARPS	...	N	...	...	...	...
Allart et al. (2020)	VLT/ESPRESSO	N	Y	N	...	Y	No Li or H
dos Santos et al. (2020)	GST/Phoenix	...	...	...	...	...	No He
Skaf et al. (2020)	HST/WFC3	Y	...	...	...	Y	FeH
Spake et al. (2021)	HST+ <i>Spitzer</i>	Y	Y	N	Y	N	CO <sub>2</sub> , but no Li

NOTES — Y for detected, N for not detected; <sup>a</sup> MC/PC : Mostly clear – Partly cloudy.

reduction, the telluric and stellar signal removal procedures. Section 3.5 details the atmospheric modeling as well as the planetary signal extraction methods, and Section 3.6 presents the associated results. In Section 3.7 we discuss our findings, and summarize our main results in Section 3.8.

## 3.2. The unusual WASP-127 system

WASP-127 b (Lam et al., 2017) is a highly-inflated sub-Saturn ( $0.165 \pm 0.021 M_{\text{Jup}}$ ; Seidel et al. 2020b) that orbits a bright solar-type star (G5V,  $H = 8.738$  mag) with an orbital period of 4.17 days (see Table 3.2 for the stellar and planetary parameters of this system). Its radius ( $1.31 \pm 0.03 R_{\text{Jup}}$ ; Seidel et al. 2020b) is anomalously large for a planet in such an old system ( $9.7 \pm 1.0$  Gyr; Allart et al. 2020) and places it on the edge of the short-period Neptune desert (Mazeh et al., 2016). The position of the star in a color-magnitude diagram indicates that it is moving off the main-sequence, and onto the RGB (Lam et al., 2017), and the associated increase in irradiation of the planet is thought to be the principal cause of the re-inflation of the planet’s radius (Lopez & Fortney, 2016; Hartman et al., 2016).

However, other mechanisms have been proposed to explain the puffiness of short-period giant planets, such as tidal heating (Bodenheimer et al., 2001, 2003), enhanced atmospheric opacity (Burrows et al., 2007), or Ohmic heating (Batygin et al., 2011; Thorngren & Fortney, 2018). Migration processes could have shaped the orbit through planet-planet interactions (e.g. Ford & Rasio, 2008) or through the Kozai-Lidov effect (Kozai, 1962; Lidov, 1962; Fabrycky & Tremaine, 2007; Bourrier et al., 2018). The latter could also explain the retrograde misaligned orbit of WASP-127 b ( $\lambda = -128_{-5}^{+6^\circ}$ ; Allart et al. 2020), which further hints at a unusual formation and evolution pathway for this planet. Photo-evaporation was also suggested to partly explain the highly inflated state of WASP-127 b (Owen & Lai, 2018), but this is probably not the case due to the relatively low XUV flux that the planet receives from its old host star (Chen et al., 2018; dos Santos et al., 2020).

With an equilibrium temperature of 1400 K, an extremely low density ( $\sim 0.09 \text{ g cm}^{-3}$ ) and a low mean molecular weight around  $2.3 \text{ g mol}^{-1}$  (Skaf et al., 2020), WASP-127 b is very puffy, with an atmospheric scale height of about 2100 km, making it an extremely favourable target for transit spectroscopy (the expected signal in transmission for one scale height is around 420 ppm; Allart et al. 2020). Contrary to other highly inflated exoplanets that show unexpectedly flat spectra (Libby-Roberts et al., 2020), previous atmospheric studies of WASP-127 b have yielded strong detections of atomic and molecular species, with ground-based observations in the visible at both low- and high-resolution being fruitful, and are summarized in Table 3.1. The overall consensus seem to agree with WASP-127 b having a partly cloudy atmosphere, with scattering hazes, water, and sodium.

The high-resolution ESPRESSO data by Allart et al. (2020) constrained the presence of a gray absorbing cloud deck to be between 0.3 and 0.5 mbar (roughly consistent with the values from Skaf et al. 2020 at  $\log_{10} P_{\text{cloud}} = 1.85^{+0.97}_{-0.66} \text{ Pa} \simeq -3.15 \text{ bar}$ ). These same observations also enabled the measurement of the weak signature of the Rossiter-McLaughlin effect (RME; Rossiter 1924; McLaughlin 1924), the  $v \sin(i_*)$  of the slowly rotating star, and the spin-orbit angle of the misaligned and retrograde planet.

The richness of spectral information in WASP-127 b’s atmosphere has been unambiguously confirmed by Spake et al. (2021, hereafter S21). They presented a combination of low-resolution spectroscopic transit observations with HST/WFC3 and the Space Telescope Imaging Spectrograph (STIS), as well as photometric observations with the Spitzer Space Telescope at 3.6 and  $4.5 \mu\text{m}$ . A retrieval analysis performed on their combined data revealed absorption from Na,  $\text{H}_2\text{O}$ , and  $\text{CO}_2$ — enabling the first firm constraints on WASP-127 b’s atmospheric metallicity and C/O ratio. Their analysis also revealed evidence for wavelength-dependent scattering from small-particle aerosols, however they did not detect K or Li, and found no evidence for a gray cloud deck. They tested both free<sup>1</sup>- and equilibrium-chemistry retrievals (with the frameworks ATMO, Tremblin et al. 2015, 2016, and NEMESIS, Irwin et al. 2008) with both favoring super-solar abundances of  $\text{CO}_2$  (compared to the expected value from a solar composition in chemical equilibrium, from an unusually strong absorption feature at  $4.5 \mu\text{m}$ ; see their Figure 18), indicative of a high metallicity (Lodders & Fegley, 2002; Öberg & Bergin, 2016). However, with only one photometric data point at  $4.5 \mu\text{m}$ , S21 were unable to disentangle the contributions of CO and  $\text{CO}_2$  in a model-independent way. Due to this degeneracy, they found conflicting C/O ratios depending on the retrieval type. On the one hand, in the chemical equilibrium retrieval they found a super-solar C/O (roughly between 0.8 and 0.9), with an expected abundance of  $\log_{10} \text{CO} \simeq -2$  (see their Figure 12). On the other hand, in their free-chemistry retrieval, they found a sub-solar C/O ratio (below 0.2) as the abundance of CO remains unconstrained by the data. They concluded that spectroscopy with JWST will allow for precise constraints of on the

---

1. All individual molecular abundances are free parameters.

Table 3.2. WASP-127 System Parameters

Stellar Parameters	Value
Spectral Type <sup>a</sup>	G5
<i>H</i> magnitude <sup>b</sup>	8.738 mag
Stellar mass ( $M_*$ ) <sup>c</sup>	$0.950 \pm 0.023 M_\odot$
Stellar radius ( $R_*$ ) <sup>c</sup>	$1.333 \pm 0.037 R_\odot$
Temperature ( $T_{\text{eff}}$ ) <sup>d</sup>	$5842 \pm 14$ K
Surface gravity ( $\log g$ ) <sup>d</sup>	$4.23 \pm 0.02$ cgs
Metallicity ([Fe/H]) <sup>d</sup>	$-0.19 \pm 0.01$
RV Semi-amplitude ( $K_*$ ) <sup>c</sup>	$21.51 \pm 2.78$ m s <sup>-1</sup>
$v \sin(i_*)$ <sup>d</sup>	$0.53^{+0.07}_{-0.05}$ km s <sup>-1</sup>
Age <sup>d</sup>	$9.7 \pm 1.0$ Gyr
Planet Parameters	Value
Planet mass ( $M_p$ ) <sup>c</sup>	$0.165 \pm 0.021 M_J$
Planet radius ( $R_p$ ) <sup>c</sup>	$1.311 \pm 0.025 R_J$
Orbital period ( $P$ ) <sup>c</sup>	4.178062(03) days
Epoch of transit ( $T_0$ ) <sup>c</sup>	2456776.62124(23) BJD
Transit duration ( $T_{14}$ ) <sup>c</sup>	0.181(37) days
Semi-major axis ( $a$ ) <sup>c</sup>	$0.04840^{+0.00136}_{-0.00095}$ AU
Inclination ( $i_p$ ) <sup>d</sup>	$87.85 \pm 0.35^\circ$
Eccentricity ( $e$ ) <sup>c</sup>	0
Planet RV semi-amplitude ( $K_{P,0}$ ) <sup>e</sup>	$129.95 \pm 1.2$ km s <sup>-1</sup>
Systemic velocity Tr1 ( $v_{\text{sys},1}$ ) <sup>f</sup>	$-8.92 \pm 0.06$ km s <sup>-1</sup>
Systemic velocity Tr2 ( $v_{\text{sys},2}$ ) <sup>f</sup>	$-8.92 \pm 0.06$ km s <sup>-1</sup>
Systemic velocity Tr3 ( $v_{\text{sys},3}$ ) <sup>f</sup>	$-8.97 \pm 0.09$ km s <sup>-1</sup>

References. — <sup>a</sup> Lam et al. (2017); <sup>b</sup> Cutri et al. (2003); <sup>c</sup> Seidel et al. (2020b); <sup>d</sup> Allart et al. (2020); <sup>e</sup>  $K_{P,0}$  was computed by sampling 5000 values of  $M_p$ ,  $M_*$ ,  $P$ , and  $i_p$  from a normal distribution using their mean value and uncertainties listed here.; <sup>f</sup> This work.

C/O ratio or WASP-127 b. Overall, a deeper investigation of this planet is needed to help shed light on how highly inflated planets form and evolve with time.

The previous strong detections of spectral features make WASP-127 b a prime target for ground-based studies at high-resolution in the NIR. With no such studies having as of yet been published (although with an expected signal of 800–1000 ppm of water in the NIR at high-resolution; Allart et al. 2020), we aim to bridge this gap with SPIRou (Donati et al., 2020) and lift the degeneracy between the two previously proposed scenarios by detecting and quantifying the abundances of the main constituents of its atmosphere with transmission spectroscopy.

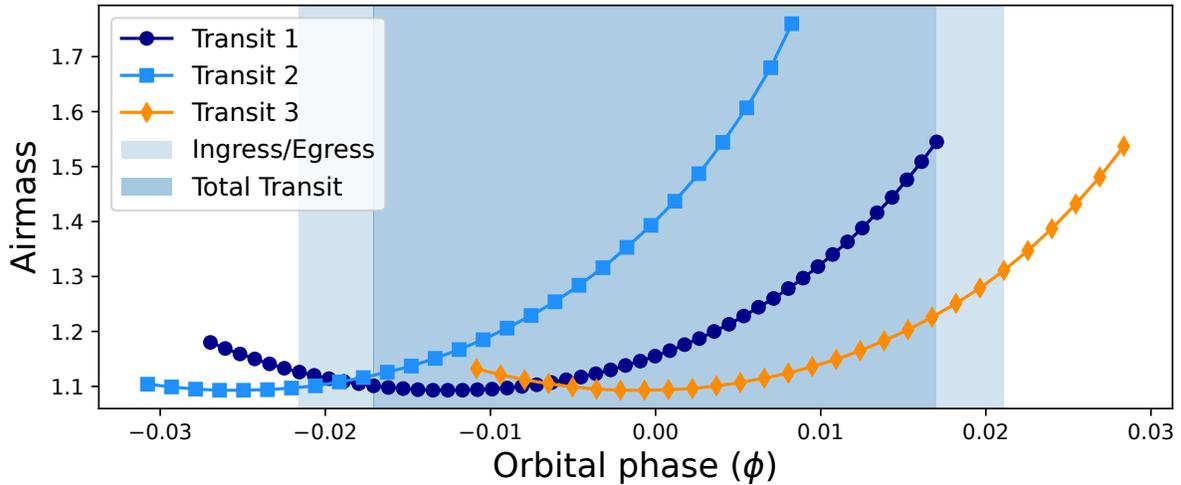


FIGURE 3.1 Airmass variation during the three SPIRou transit observations (dark blue circles for Transit 1, light blue squares for Transit 2 and orange diamonds for Transit 3). The shaded area shows the span of the transit event). All three transits were taken under relatively favorable airmass conditions ( $<1.8$ ).

### 3.3. Observations

In this work, we used the *Spectro-Polarimètre InfraRouge* (SPIRou; Donati et al. 2020), a fiber-fed échelle spectro-polarimeter installed at the Canada-France-Hawaii Telescope (CFHT, 3.6-m). SPIRou has a spectral resolution of  $R = 70,000$  (with 2 pixels per resolution element, yielding a sampling precision of  $\sim 2.3 \text{ km s}^{-1}$  per pixel) and a broad NIR spectral range covering the  $Y$ ,  $J$ ,  $H$  and  $K_s$  bands simultaneously ( $\sim 0.95\text{--}2.50 \mu\text{m}$ ). SPIRou splits the incoming target light into two orthogonal polarization states, each of which feeding its own science fiber (fibers A and B), allowing both states to be observed simultaneously. A third fiber (C) is fed by a calibration source. The large NIR spectral range of SPIRou gives us access to the absorption features of a multitude of the major molecular species present in the atmospheres of exoplanets such as  $\text{H}_2\text{O}$ ,  $\text{CH}_4$ ,  $\text{CO}_2$ ,  $\text{HCN}$ , and  $\text{NH}_3$ , but most importantly the  $2.3 \mu\text{m}$  CO band needed to differentiate the two scenarios presented in S21. The high-resolving power of SPIRou allows us to resolve and distinguish unique spectral line forests from different molecules (e.g., CO versus  $\text{CO}_2$ ), even if their bands overlap — and its suitability for characterizing exoplanet atmospheres has already been demonstrated by several studies (e.g., Pelletier et al., 2021; Boucher et al., 2021). Of all existing facilities in the world to date, only SPIRou, GIANO (Oliva et al., 2006), CRIRES+ (Follert et al., 2014), IGRINS (Park et al., 2014), iSHELL (Rayner et al., 2012), and NIRSPEC (McLean et al., 1998) cover the  $K$ -band and hence can lift the CO/ $\text{CO}_2$  degeneracy which has hampered previous data.

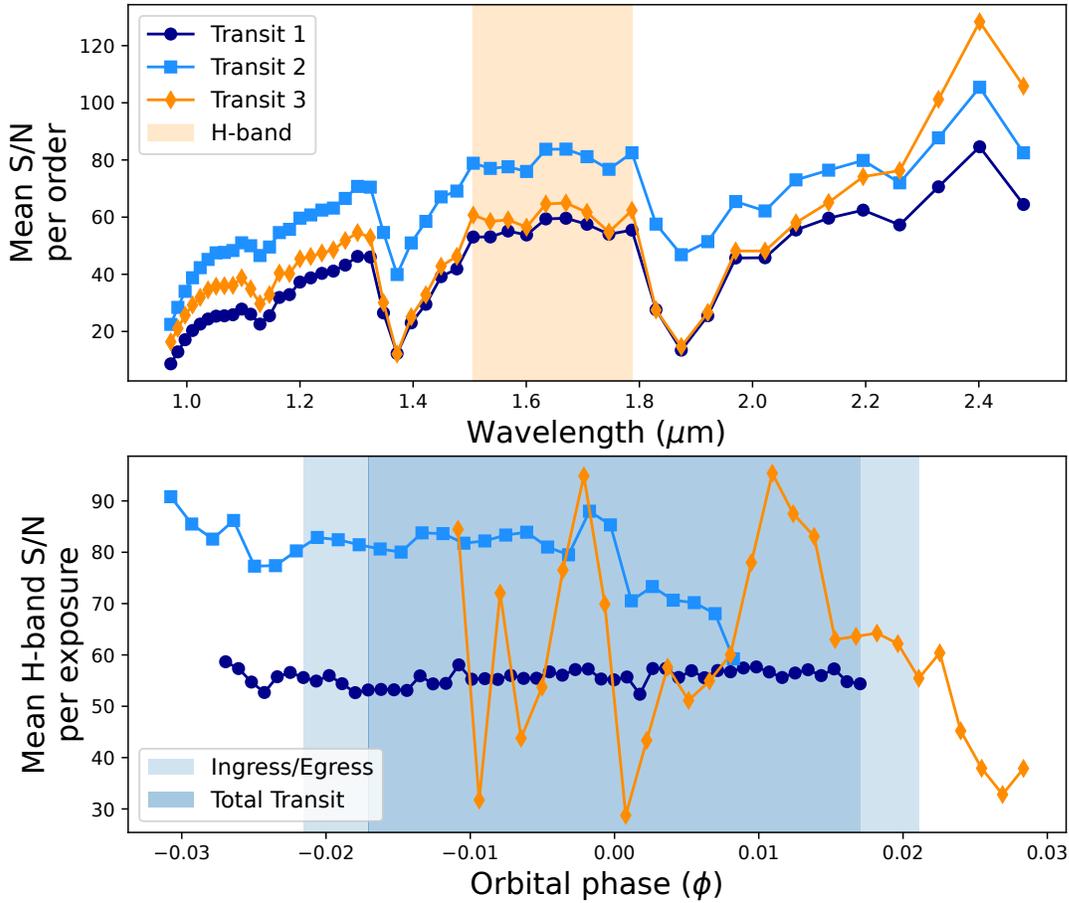


FIGURE 3.2 *Top panel* : Temporal mean of the S/N per order for each of our three SPIRou transits. The shaded area highlights the orders in the *H*-band. *Bottom panel* : Time evolution of the spectral mean of the S/N in the *H*-band per exposure, with the shaded area denoting the span of the transit event. Here we see that while the first two transits were taken under relatively stable conditions, the Transit 3 data shows a much higher variability due to clouds/fog.

Three partial transits of WASP-127 b were observed with SPIRou for a total of 11.94 hours of observing time. All sets of observations were taken without moving the polarimeter retarders (Fresnel rhombs) to ensure the highest possible instrument stability, and placing the Fabry-Pérot in the calibration channel to better track small relative spectral drifts.

The observations of the first transit (hereafter Tr1) were obtained on 2020 March 11, as part of the SPIRou Legacy Survey (SLS). Initiated at the start of SPIRou operations, the SLS is a CFHT Large Program of 310 telescope nights (PI : Jean-François Donati ; Donati et al. 2020) whose main goals are to search for planets around M dwarfs using precision RV measurements,

Table 3.3. SPIRou observations of WASP-127

Transit	Tr1	Tr2	Tr3
UT Date	2020-03-11	2021-03-22	2021-05-03
BJD (d) <sup>a</sup>	2458919.85	2459295.86	2459337.72
Texp (s) <sup>b</sup>	300	500	500
Seeing (") <sup>c</sup>	0.61–0.76	0.70–1.21	0.52–1.74
S/N <sup>d</sup>	57.5	81.0	61.7
<b>Number of exposures :</b>			
Before ingress	7	7	0
During transit	43	21	23
After egress	0	0	5
Total	50	28	28
Total observing time (h) <sup>e</sup>	4.18	3.86	3.90

Note. — <sup>a</sup> Barycentric Julian date at the start of the observing sequence; <sup>b</sup> Exposure time of a single exposure; <sup>c</sup> Range of values of the seeing during the transit; <sup>d</sup> Mean S/N per pixel, per exposure, at  $1.7 \mu\text{m}$ ; <sup>e</sup> Total observing time in hours.

characterize the magnetic fields of young low-mass stars and their impact on star and planet formation, and probe the atmosphere of exoplanets using HDS. The second and third transits (hereafter Tr2 and Tr3, respectively) were observed on 2021 March 22 and May 3 (Program 21AC02/21AF18, PI Boucher/Debras). Tr1 was observed with an exposure time of 300 s, but this was increased to 500 s for the following transits due to a lower than expected signal-to-noise ratio (S/N) seen in Tr1. The instrument rhombs were changed in August 2020 (between Tr1 and Tr2/Tr3), increasing the throughput in spectral bands *Y* and *J* by factors of 1.5 and 1.4, respectively, thus also improving the S/N in these bands for Tr2 and Tr3. An overview of the observation specifications are listed in Table 3.3. For each transit, the airmass curve is shown in Figure 3.1, while the S/N temporal mean per order and the spectral mean (over the *H*-band only) per exposure are shown in Figure 3.2. Sky conditions were photometric and dry for the first two observing sequences. However, fog caused Tr2’s last exposure to be only 373 s (instead of 500 s). Nevertheless, we retain it for our analysis. For the third transit, poor weather conditions (higher water column density and mean extinction that varied between 0.2 and up to 1.8 mag, with an increase toward the end of the sequence) led to highly variable data quality and S/N. Nonetheless, the data reduction pipeline and the estimation of uncertainties are reliable to allow the data to be used anyway.

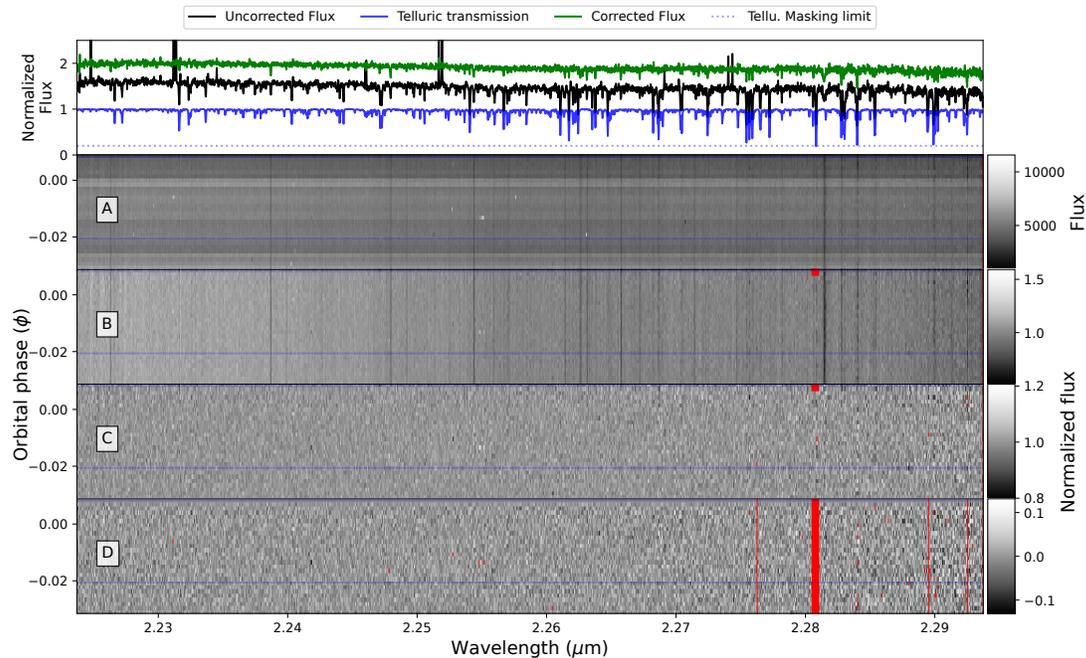


FIGURE 3.3 Analysis steps that are applied to observed SPIRou spectral time series observations. Here the full order covering 2.2237–2.2937  $\mu\text{m}$  of Tr2 is shown. The top panel shows the uncorrected (black), the telluric-corrected (green) spectra, and the reconstructed telluric transmission spectrum (blue), with an offset to facilitate visibility. Panel (A) : Telluric-corrected spectra. The masked pixels are shown in red ; here the masking is done by APERO. Panel (B) : Spectra normalized to the continuum level of the reference spectrum. Most of the bad pixels were corrected or masked and the spectra were shifted in the pseudo star rest frame. Panel (C) : Transmission spectra, where each spectrum was divided by the reference spectrum. Panel (D) : Final transmission spectra, corrected for the vertical residual structures using PCA (2 principal components removed). The total standard deviation of the transmission spectra time series is reduced by a factor 1.3–1.9 from step C to D, depending on the transit considered. The blue horizontal lines in panels (A)–(D) show the ingress position (mid-transit is at phase 0). The masked pixels are shown in red.

### 3.4. Data Reduction and Analysis

All data were reduced using *A Pipeline to Reduce Observations* (APER0 ; version 0.7.179 ; Cook et al., in prep.), the SPIRou data reduction software. APER0 performs all calibrations and pre-processing to remove dark, bad pixel, background and detector non-linearity corrections (Artigau et al., 2018), localization of the orders, geometric changes in the image plane, correction of the flat and blaze, hot pixel and cosmic ray correction, wavelength calibration (using both a hollow-cathode UNe lamp and the Fabry-Pérot etalon ; Hobson et al. 2021), and removal of diffuse light from the reference fiber leaking into the science channels (when a Fabry-Pérot is

used simultaneously in the calibration fiber). This is done using a combination of nightly calibrations and master calibrations. The result is an optimally extracted spectrum (Horne, 1986) of dimensions 4088 pixels (4096 minus 8 reference pixels) with 49 orders (one order, #80, is not extracted by the data reduction software), referred to as extracted 2D spectra; *E2DS*. While the *E2DS* are produced for the two science fibers individually (A and B) and for their combined flux (AB), we only used the AB extraction as this is the relevant data product for non-polarimetric observations.

### 3.4.1. Telluric absorption correction

APER0 also provides telluric-corrected versions of the spectra, as well as the reconstruction of the Earth’s transmittance (Figure 3.3, top panel). For our analysis, we use this data product to minimize contamination from telluric absorption (and sky emission). This step is essential to be able to recover the subtle signal of an exoplanet’s atmosphere, especially as telluric contamination arises from molecules also expected to be present in exoplanet atmospheres. The low number of observable transits of the WASP-127’s system (due to its low declination) prevented us from selecting optimal transits where the RV trail from the planet does not overlap with the telluric rest frame. Our observations thus have the potential to be contaminated by residual tellurics and their consideration and proper correction are even more crucial.

The telluric correction done in APER0 is a two-step process and is briefly outlined here (this also includes the sky emission correction). 1) The extracted spectra of both science targets and a large set of rapidly rotating hot stars are fitted with an Earth’s transmittance model from TAPAS (Bertaux et al., 2014) that leaves percent-level residuals. 2) From the ensemble of hot star observation, we derive a correction model for the residuals with three components for each pixel (optical depths for the H<sub>2</sub>O, the dry components, and a constant). This residual model is adjusted to each science observation according to the optical depth of each component from the TAPAS fit. The resulting correction leaves residuals at the level of the PCA-based method of Artigau et al. (2014), but has the advantage of a simplicity and that any spurious point in the data will result in a local error rather than affecting the transmission globally as for a PCA analysis. Finally, a reconstruction of the telluric spectrum is derived using the fitted TAPAS template and the residuals model, for each observed spectrum. The pipeline performs the telluric absorption correction for lines with a transmission down to  $\sim 10\%$  (i.e., with relative depths of 90% with respect to the continuum), with deeper lines being masked out, and additional masking being at the discretion of the user. Following an injection-recovery test, we observed that some of the data yielded slightly better detections with further masking. We thus additionally masked telluric lines (from the reconstructed telluric spectra) with core transmission below 20% and 50% depending on the depth of tellurics for a given night, and this mask was extended from the

core of those lines until their transmission in the wings reached 97%. No additional masking on Tr1 was found to be necessary from injection-recovery tests.

### 3.4.2. Transmission spectrum construction

To build the transmission spectra, we follow a similar process to Boucher et al. (2021), and apply the technique individually to each order (and each transit). We briefly summarize the process here.

- Bad pixels correction : The telluric-corrected *E2DS* spectra are blaze normalized (Figure 3.3 A), then the bad pixels are identified and corrected based on the method of Brogi et al. (2014), where the spectra are divided by their mean value (wavelength-wise, to bring everything to the same continuum level) and then each spectral pixel is divided by its mean in time, yielding a data residual map, i.e., a noise map. We take the absolute value of the map and fit a second order polynomial to each exposure. This recovers the noise floor, which follows the noisier borders of each order, and we subtract it from our noise map. Then, any pixels deviating by more than  $5\sigma$  from the flat noise map are flagged as bad pixels. For isolated pixels, they are corrected with a spline-interpolation, while groups of 2-3 pixels are linearly interpolated. Groups of four or more pixels are masked (less than 0.001% of the pixels are masked in this step for all transits).
- Stellar signal alignment : After the bad pixel corrections, we return to the wavelength-normalized spectra and Doppler shift each spectrum to a pseudo stellar rest frame (SRF), where the stellar lines are aligned, but not centered at zero velocity (Figure 3.3 B). We accomplish this by shifting only by the negative of the *variation* of the barycentric Earth RV and stellar orbital motion<sup>2</sup>. This RV variation from each exposure is compared to the value from the middle of each transit sequence. This means that the first half of the exposures are shifted by roughly the same amount as the second half, but in the opposite direction, therefore minimizing interpolation errors caused by shifts of large fractions of pixels. In this case, the shifts are at most 12% of a spectral pixel on the SPIRou detector. We then mask the telluric lines as described above.
- Reference spectrum construction and removal : Next, we build a reference spectrum representative of the stellar spectrum to remove its contribution from each observation and leave only the planetary signal. The usual procedure is to only use the out-of-transit exposures to construct this reference spectrum, thus minimizing any contamination by the planetary signal. However, given that our observations only include a small number of low-S/N exposures, and only cover either ingress or egress, a reference spectrum

---

2. Stellar orbital motion was included for completeness, even if it is negligible.

constructed in this way would not be a true representation of our entire observing sequence. For this reason, we construct the reference spectrum by taking the median of *all* spectra (in the pseudo SRF). Additionally, for Tr3, due to the low S/N of certain exposures, we build the reference spectrum using only those exposures with mean  $H$  band S/N  $\geq 40$ . These final reference spectra should not contain much residual planet signal as the latter move over roughly 4.3, 3.3 and 3.6 times the line full width at half maximum (FWHM) during the transit for Tr1, Tr2 and Tr3 respectively, compared to the quasi-stationary stellar signal. Dividing by the reference spectra leaves low-frequency variations, that can be corrected by dividing the spectra by a low-pass filtered<sup>3</sup> version of the spectra divided by the reference.

- We observed in some parts of some spectra that this ratio was far from 1.0, meaning that these parts are more poorly represented by the reference spectrum. We thus chose to mask the regions where this ratio was  $6\sigma$  away from the mean value, which removes around 0.6, 0.7 and 2.3% of the data points for transits 1, 2 and 3 respectively.
- The individual transmission spectra are obtained by dividing the continuum-normalized spectra by the reference spectrum (Figure 3.3 C). We then reapply a final sigma-clipping at  $6\sigma$  to remove any remaining outlier pixels.
- Systematic noise residuals correction : We used a PCA-based approach to remove any remaining pseudo-static signals (e.g., stellar and telluric residuals). We build the PC base in the time direction using the natural logarithm of the transmission spectra themselves, and then perform injection/recovery tests (at the negative  $K_{P,0}$ ) to determine the appropriate number of PCs to remove. A combination of the best retrieved  $t$ -test, CCF SNR and/or  $\Delta$  BIC are considered to make the selection, as these metrics do not always lead to the same conclusions. This results in the removal of 2, 2, and 5 PCs for Tr1, Tr2, and Tr3, respectively. The optimal number of necessary PCs seems to follow the data quality and S/N (i.e., Tr3 needs more PCs to better uncover an injected underlying signal). We tested using only a PCA to correct for all the stationary (or quasi-stationary) signal from the telluric uncorrected spectra in a single step. However, this single step approach yielded results that were not as good as the two steps approach that we adopted here (more telluric residuals were visible).
- Finally, we perform one last sigma clip of spectral pixels deviating by more than  $5\sigma$  in the time direction within each order<sup>4</sup>. We then remove the mean of each spectrum

---

3. Median filter of width 51 pixels followed by convolution with a Gaussian kernel of width 5 pixels.

4. Here, the removal of the noise floor, to account for the natural increase of noise near the order borders, was also applied.

(wavelength-wise) to keep a zero mean for the computation of the cross-correlation function (see Fig. 3.3 D). We note that the effects from the last few sigma clips are very minimal.

We measured  $v_{\text{sys}}$  directly from our data by computing the CCF of the telluric-corrected 1D spectra (*S1D*; another data product from `APER0`) with a synthetic spectrum from a PHOENIX atmospheric model (Husser et al., 2013) with  $T_{\text{eff}} = 5800 \text{ K}$ ,  $\log g = 4.0$ , and  $[M/H] = 0$ . We computed the CCF with the signal in *H* band of every *S1D*, weighted by the second derivative of the model (a proxy for the strength of the absorption lines), and measured the peak position with the bisector method. We then subtracted  $v_{\text{bary}}(t)$  and  $v_{\text{orb},*}(t)$ , (where  $v_{\text{bary}}(t)$  is the barycentric velocity of the observer — in our case it is the barycentric Earth RV, BERV — and  $v_{\text{orb},*}(t) = -K_* \sin[2\pi(\phi(t))]$  is the radial part of the orbital velocity of the star), and took the mean over all spectra of each transit to get the observed  $v_{\text{sys}}$  (the values are listed in Table 3.2). We computed the gravitational redshift to be at  $0.47^{+0.01}_{-0.02} \text{ km s}^{-1}$ , and estimated the convective blue shift to be  $-0.42 \pm 0.15 \text{ km s}^{-1}$  (Leão et al., 2019). As these values roughly compensate one another, we chose to simply ignore these effects in the determination of the heliocentric RV of the WASP-127 system.

## 3.5. Atmospheric Signal Extraction : Models and Methods

Even though the data are now cleaned of telluric and stellar signals, the individual planetary absorption lines are still buried within the remaining noise. We thus need a cross-correlation type analysis that combines the signal of all lines over the available spectral range to reveal the planetary atmospheric signal. In this section, we first present how WASP-127 b’s atmospheric models are generated and how we process them to better represent the data. Then, we present the methods that we tested to detect the atmosphere of WASP-127 b : cross-correlation, *t*-test, and log-likelihood mapping.

### 3.5.1. Atmospheric Model

We generated synthetic transmission spectra of WASP-127 b’s atmosphere using the open-source `petitRADTRANS` framework (PRT; Mollière et al., 2019; Mollière et al., 2020). PRT computes transmission and emission spectra of exoplanets with clear or cloudy atmospheres. It can produce low-resolution ( $R = 1,000$ ) or high-resolution ( $R = 10^6$ ) models by considering the molecular opacities at each pressure layer using either a correlated-k treatment or line-by-line radiative transfer, respectively. In this work we used 50 pressure layers between  $10^{-10}$  and

$10^2$  bar, log-uniformly spaced, and fixed the reference pressure to  $P_0 = 10$  mbar<sup>5</sup>, where the bulk of our transmission signal is expected to originate. The molecular opacities and associated line lists used in this work include <sup>6</sup> H<sub>2</sub>O, CO, CO<sub>2</sub>, and OH (Rothman et al., 2010), CH<sub>4</sub> (Yurchenko et al., 2020), HCN (Barber et al., 2014; Harris et al., 2006), NH<sub>3</sub> (Yurchenko et al., 2011), FeH (Bernath, 2020), TiO (see references in Mollière et al., 2019), and C<sub>2</sub>H<sub>2</sub> (Rothman et al., 2013). Out of these elements, only H<sub>2</sub>O, CO, CO<sub>2</sub>, CH<sub>4</sub>, NH<sub>3</sub>, and TiO are expected to have abundances larger than  $10^{-10}$  for a solar composition in chemical equilibrium. All these molecules have major features within SPIRou’s spectral range that could potentially be detected given a high enough volume mixing ratio (VMR). Also included are the absorption from H<sub>2</sub> broadening (Burrows & Volobuyev, 2003), collision-induced broadening from H<sub>2</sub>/H<sub>2</sub> and H<sub>2</sub>/He collisions (Borysow, 2002). The abundances of H<sub>2</sub> and He were fixed to 85 and 15% (i.e., solar-like), respectively. To save time on the generation of the high-resolution PRT models, we down-sampled the line-by-line opacities by a factor of 4, which gave us models at  $R = 250,000$ .

We adopted an analytical atmospheric temperature-pressure (T-P) profile from Guillot (2010), who derived a parametrized relation between temperature and optical depth valid for plane-parallel static grey atmospheres. For simplicity and to anchor the shape of the profile to that of S21, we fixed three of the four parameters of the profile, namely  $\kappa_{\text{IR}}$  the atmospheric opacity in the IR wavelengths,  $\gamma$  the ratio of the optical and IR opacities, and  $T_{\text{int}}$  the planetary internal temperature, while keeping  $T_{\text{eq}}$ , the atmospheric equilibrium temperature, as a free parameter<sup>7</sup>. We fixed the above parameters to  $\kappa_{\text{IR}} = 10^{-3} \text{ cm g}^{-1}$ ,  $\gamma = 10^{-1.5}$  and  $T_{\text{int}} = 500$  K, based on the values retrieved by S21 using ATMO, and do not change them for the entirety of this work.

We include a gray cloud deck whose contribution is characterized by its cloud-top pressure  $P_{\text{cloud}}$ . In order to model the chromatic absorption by aerosols ( $\sigma_{\text{haze}}$ ), we used a power law of the form :

$$\sigma_{\text{haze}}(\lambda) = S\sigma_0 \left( \frac{\lambda}{\lambda_0} \right)^\gamma, \quad (3.5.1)$$

where  $S$  is the empirical enhancement factor,  $\sigma_0$  is the scattering cross-section of molecular hydrogen at  $\lambda_0 = 0.35 \mu\text{m}$ , and  $\gamma$  sets the wavelength dependence of the scattering.

We explored the possibility that the planet’s signal is broadened. For this, we used a simple Gaussian broadening kernel, centered on zero and characterized by its FWHM to mimic, to first order, the effect of the planet’s rotation and atmospheric dynamics. Even if a Gaussian kernel

5. This is where the planetary radius is equal to  $R_p$ .

6. A complete list of the available opacities is given in the documentation of PRT at : <https://petitradtrans.readthedocs.io/en/latest/index.html>. Also, we added some high-resolution opacities to PRT manually (OH and FeH) following their instruction, and used the open-access DACE database, computed with HELIOS-K (Grimm & Heng, 2015; Grimm et al., 2021), to compute the opacity grids : <https://dace.unige.ch/opacityDatabase/>

7. We will interchange the references to  $T_{\text{eq}}$  and  $T_p$ .

might not be the best representation of the atmosphere in transit, it still gives a good approximation of the width of the average line profile. This kernel is convolved with the instrumental profile, a Gaussian kernel matching the resolving power of SPIRou. Once generated, the model is then convolved with this kernel and binned to match SPIRou’s sampling.

3.5.1.1. Model processing. At this point in the analysis, all the detrending steps applied to the data have not only affected the telluric and stellar signals, but also to some extent, that of the planetary atmosphere (Brogi & Line, 2019). This is mostly due to the subtraction of the PCs, which are generally not orthogonal to the planet’s transmission spectrum. Their subtraction (see Section 3.4.2) may warp part of the actual planet signal and introduce artifacts in the spectral time series which can then bias the determination of atmospheric parameters (velocities, abundances, temperature, etc.), although the pre- and post-PCA operations have no impact on the planetary transmission spectra. We thus apply the same treatment to the model before comparing it to the data to ensure a better representation. This method was first presented in Brogi & Line (2019), was successfully used in many studies and has since become a standard procedure for this type of analysis (e.g., Gandhi et al., 2019; Pelletier et al., 2021; Boucher et al., 2021; Kasper et al., 2021; Line et al., 2021; Gibson et al., 2022). We proceed as follows : we generate full synthetic transit sequences by injecting the model spectrum (described in Section 3.5.1) into a reconstruction of the observed data. This reconstructed signal is built by multiplying together (i) the reference spectrum, (ii) the spectral median of every spectrum (for every order; same as the first step from Section 3.4.2) and (iii) the PCA reconstructed version of the systematic noise removed from the transmission spectrum. This last component may contain a small fraction of the planetary signal, but this is not a significant problem as it is largely dominated by the systematic noise. We then inject the model at  $v_P(t)$ , the total planet RV,

$$v_P(t) = v_{\text{bary}}(t) + v_{\text{sys}} + v_{\text{orb,P}}(t) - v_{\text{orb,*}}(t) + v_{\text{rad}}, \quad (3.5.2)$$

where  $v_{\text{sys}}$  is the systemic RV ;  $v_{\text{orb,P}}(t) = K_P \sin[2\pi(\phi(t))]$  is the radial part of the orbital velocity of the planet ;  $K_P$  and  $K_*$  are the planet’s and star’s RV semi-amplitude, and  $v_{\text{rad}}$  is a constant additional velocity term to account for potential shifts. Note that here,  $v_{\text{bary}}(t) = v_{\text{bary}}(t_{\text{mid.exp.}})$ , the barycentric velocity at the middle of the sequence, because the spectra are in the pseudo-SRF (from the second step in Section 3.4.2). We can inject the models at different combinations of  $K_P$  and  $v_{\text{rad}}$  to produce different planetary paths in the velocity space.

Then, the relevant steps of the analysis are reapplied to the modeled “reconstructed” transit sequence. For the PC removal step, we use the PCs obtained with the real observations to reconstruct the synthetic transit sequence, and remove this reconstruction from the synthetic sequence. The effect of this process is to replicate in the model spectra, to the extent possible, any subtraction of the planet signal which occurs while subtracting the PCs from the actual data. After this

PC subtraction (done in log-space), we remove the mean spectrum for each order, as was done on the observed data. This synthetic, PC- and mean-subtracted model transmission time series ( $m_i$ ) is then used for the computation of the cross-correlation and log-likelihood.

### 3.5.2. Cross-Correlation and Log-Likelihood Mapping

To extract and maximize the planetary signal, we combine the signal of the many buried, but resolved absorption lines that are found over the whole spectral range of SPIRou with cross-correlation and similar approaches. The approach that we used to detect the planet signal and constrain the atmospheric parameters is described below.

3.5.2.1. Algorithm. Based on the equations in Gibson et al. (2020) (also used in Nugroho et al. 2020, 2021; Boucher et al. 2021; Gibson et al. 2022), we write the cross-correlation function (CCF) as

$$\text{CCF}(\boldsymbol{\theta}, \nu_P) = \sum_{i=1}^N \frac{f_i \cdot m_i(\boldsymbol{\theta}, \nu_P)}{\sigma_i^2}, \quad (3.5.3)$$

which is equivalent to a weighted CCF, where  $f_i$  are the transmission spectra (described in Section 3.4.2) with associated uncertainties  $\sigma_i$ ,  $m_i$  is the model (described in Sections 3.5.1 and 3.5.1.1), and  $\boldsymbol{\theta}$  is the model parameter vector, which includes the atmospheric model parameters, and generated at a given orbital solution  $\nu_P$ . The index  $i$  runs over all times and wavelengths in the data set, and the summation is done over  $N$  data points (total number of unmasked pixels).

The uncertainties  $\sigma_i$  were determined by first calculating, for each spectral pixel, the standard deviation over time of the  $f_i$  values. However, to make sure we do not underestimate the noise by removing too many PCs, we take the standard deviation of  $f_i$  after the removal of only one PC, even if the final transmission spectra are corrected from more. This provides an empirical measure of the relative noise across the spectral pixels which captures not only the variance due to photon noise but also due to such effects as the telluric and background subtraction residuals, but it does not convey how the noise inherent to one spectrum compares to that of another. To capture this latter effect and include it in the  $\sigma_i$ , the dispersion values calculated above were multiplied, for each spectrum, by the ratio of the median relative photon noise of that spectrum divided by the median relative photon noise of all spectra. This is computed prior to normalization : the S/N variations across the night and the different orders are thus accounted for in this  $\sigma_i$  term, acting as a weight. The final uncertainty values  $\sigma_i$  thus reflect both temporal and spectral variability. We compared with another method to compute the uncertainties to validate ours. For this, we used a  $\sigma$  map following the Gibson et al. (2020) method, i.e., by optimizing a Poisson noise function of the form  $\sigma_i = \sqrt{aF_i + b}$  with the residuals from removing the first 5 PCs. Since it yielded very similar results, we continued with our method.

The CCF (equation 3.5.3) is calculated for every order of every spectrum for an array of  $v_{\text{rad}}$  of size  $n_v$ . This gives a cube with size  $106 \times 49 \times n_v$  (when combining all three transits<sup>8</sup>) for a given  $K_p$  value in the modeled sequence and for each model tested. To combine everything into a single CCF, we first sum the above cube over orders, and then over time by applying a weight to each spectrum according to a transit model (transit depth at a given time), computed with the BATMAN package (Kreidberg, 2015). We used a non-linear limb darkening law from Claret (2000) with fixed coefficients  $u_1 = 0.5944$ ,  $u_2 = 0.0707$ ,  $u_3 = -0.1204$ , and  $u_4 = 0.0202$ , taken in S21, and valid for their white light-curve fit using WFC3+G141. The ephemeris and system parameters used are listed in Table 3.2.

The CCF equation above can then be used to compute the S/N and/or the  $t$ -test (see next Section) and estimate the level of detection of a given signal. However, to get a better model selection method, we start from equation 3.5.3 and then map it to a  $\beta$ -optimised likelihood function, as presented in Gibson et al. (2020)<sup>9</sup>, which is a more generalized form of the likelihood function in Brogi & Line (2019) :

$$\ln \mathcal{L} = -\frac{N}{2} \ln \left[ \frac{1}{N} \left( \sum \frac{f_i^2}{\sigma_i^2} + \sum \frac{m_i^2}{\sigma_i^2} - 2 \sum \frac{f_i m_i}{\sigma_i^2} \right) \right] \quad (3.5.4)$$

where the summation is implied over  $i$  (both spectral pixels and time). This equation can be written in a more compact form, i.e. :

$$\ln \mathcal{L} = -\frac{N}{2} \ln \frac{\chi^2}{N}, \quad (3.5.5)$$

by using the definition of the  $\chi^2$  :

$$\chi^2 = \sum \frac{(f_i - m_i)^2}{\sigma_i^2}. \quad (3.5.6)$$

3.5.2.2. Detection significance. There exist multiple methods to quantify the detection significance of a signal — here, we present the three that we computed : the S/N of the CCF peak, the  $t$ -test, and  $\Delta$  BIC.

First, the ‘‘S/N significance’’ is determined by dividing the total CCF (either the 2D  $K_p$  versus  $v_{\text{rad}}$  map, or the 1D version varying only with  $v_{\text{rad}}$ ) by its standard deviation, the latter being calculated by excluding the region around the peak ( $\pm 15 \text{ km s}^{-1}$  in the  $v_{\text{rad}}$  space and  $\pm 70 \text{ km s}^{-1}$  in the  $K_p$  space). This is a useful metric to quickly assess whether or not a detection is significant, but its value should be viewed as only approximate as it can vary highly with the  $v_{\text{rad}}$  and  $K_p$

8. We combine the transits by simply concatenating the three CCF time series, order-per-order. There are 50 spectra in Tr1 and 28 in Tr2 and Tr3, for a total of 106.

9. The additional scaling factor  $\alpha$ , which accounts for any scaling uncertainties of the model was set to 1 following (Brogi & Line, 2019), because we do not expect much signal to come from an extended exosphere. Also, this parameter is correlated with the VMR,  $T_p$ , and  $P_{\text{cloud}}$ , and including it just adds an additional degeneracy.

range and sampling, the extent of the excluded peak region, whether it is computed on the 2D map or the 1D CCF, etc. (Cabot et al., 2019).

The second is the Welch  $t$ -test (Welch, 1947), which has been used in many previous studies (e.g., Birkby et al., 2013; Birkby et al., 2017; Brogi et al., 2018; Cabot et al., 2019; Alonso-Floriano et al., 2019; Webb et al., 2020; Boucher et al., 2021; Giacobbe et al., 2021). The  $t$ -test verifies the null hypothesis that two Gaussian distributions have the same mean value. The Welch  $t$ -test is a generalisation of the Student test (Student, 1908), for which the two samples can have unequal variances and/or unequal sample sizes. In our case, the two distributions to be compared are drawn from our correlation map. On the one hand, we have the in-trail distribution of CCF values, that is, the CCF values within 3-pixels wide columns centered on the peak (i.e., the signal following the planet RV path; as done in Birkby et al. 2017 and suggested in Cabot et al. 2019). On the other hand, we have the out-of-trail distribution, which includes the CCF values more than  $10 \text{ km s}^{-1}$  away from  $v_p(t)$ , where there should be no planet signal. The  $t$ -test then evaluates the likelihood that these two samples were drawn from the same distribution. This approach is more robust against outliers than the CCF SNR, and provides a complementary assessment of the detection significance. However, this metric might still not reflect the true confidence in a molecular detection since, as pointed out in Cabot et al. (2019), it is not robust against oversampling in the velocity space, and thus somewhat arbitrary.

The third is the Bayesian Information Criterion (BIC)<sup>10</sup>, which we applied to our  $\ln \mathcal{L}$  results to establish how the best-fit model fares compared to the others. In this formalism, the model with the lowest BIC is preferred (here, taken to be the best-fit model), and the evidence against models with higher BIC is usually described as “very strong” when  $\Delta \text{BIC} = 2 \Delta \ln \mathcal{L}$  is greater than 10 (Kass & Raftery, 1995). In this work, we take the best-fit model to be the one with the highest  $\ln \mathcal{L}$  value, for a given set of parameters.

We note here that in all following calculations, we exclude the exposures where the planet velocity is within  $2.3 \text{ km s}^{-1}$  (1 pixel) of the BERV in order to prevent any potential telluric contamination of the planet signal. This leads to the removal of the last six and three exposures of Tr1 and Tr2 respectively, and the first three of Tr3.

### 3.6. Results

We first applied the CCF,  $t$ -test, and  $\ln \mathcal{L}$  analyses to our SPIRou WASP-127 b transmission data. We used models consisting of H, He, and one of the individual molecules that are usually present in the atmospheres of giant planets, namely  $\text{H}_2\text{O}$ ,  $\text{CO}$ ,  $\text{CO}_2$ ,  $\text{CH}_4$ ,  $\text{HCN}$ ,  $\text{NH}_3$ , and  $\text{C}_2\text{H}_2$  (Lodders & Fegley, 2002). For completeness, we also included FeH, TiO, and OH. We cycled

---

10.  $\text{BIC} = k \ln(n) - 2 \ln \mathcal{L}$ ; where  $k$  is the number of parameters,  $n$  is the number of data and  $\ln \mathcal{L}$  is our log-likelihood value for our model for each combination of parameters. For fixed values of  $k$  and  $n$ , the lowest BIC corresponds to the highest  $\ln \mathcal{L}$ .

through the aforementioned molecules one at a time, and for each case, we tested VMRs between  $10^{-9}$  and  $10^{-2}$ . We obtained a clear detection of  $\text{H}_2\text{O}$ , as well as a tentative detection of  $\text{OH}$ .

To further quantify the  $\text{H}_2\text{O}$  detection, we ran a series of tests on a grid of models across a broad range of  $\log_{10} \text{VMR}[\text{H}_2\text{O}]$ ,  $T_{\text{eq}}$ , and  $P_{\text{cloud}}$  to find the model which maximised the detection significance. According to the  $t$ -test, the best-fit model has  $\log \text{H}_2\text{O} = -4$ ,  $T_{\text{eq}} = 800 \text{ K}$  and  $P_{\text{cloud}} = 0.1 \text{ bar}$ . The CCF time-series in the planet rest frame, at  $K_{\text{P}} = K_{\text{P},0} = 129.95 \pm 1.2 \text{ km s}^{-1}$  (the expected position of the planet), is shown in Figure 3.4. The  $K_{\text{P}}$  versus  $v_{\text{rad}}$   $t$ -test map is shown in Figure 3.5. Note that to stay conservative, our  $t$ -test maps have been scaled such that there are no noise structures that are above the  $\pm 3\sigma$  level. On the CCF time-series (Figure 3.4), in Tr1 and Tr2, there appear to be some noise structures in the out-of-transit exposures at the same  $v_{\text{rad}}$  as the planetary signal that mimics an extended signal. This is most likely some residual noise from the data reduction that simply coincides with  $v_{\text{peak}}$ . Especially in Tr1, the first few exposures are much noisier than the rest and we can see similar amplitudes at other velocities. However, it seems to be present for any fraction of the tellurics that is masked, and any number of PC that are removed. These exposures also appear noisier than the rest in the CCF map of other non-detected molecules, such as  $\text{CO}_2$  and  $\text{OH}$ , which further suggest that some residuals are present. We found that removing a second-order polynomial in the time direction was able to remove those structures (while still keeping the signal). However, recomputing the coefficients for every tested model would be computationally too expensive. For this reason, we did not apply the polynomial removal step.

The  $t$ -test peak is found at  $K_{\text{P}} = 148_{-49}^{+37} \text{ km s}^{-1}$  and  $v_{\text{rad}} = -7.2_{-3.6}^{+4.2} \text{ km s}^{-1}$ , where the uncertainties corresponds to a drop of  $1\sigma$  from the peak value. Large error bars on  $K_{\text{P}}$  are expected the planet’s acceleration during transit is low, especially as we only have partial transit sequences. At  $K_{\text{P},0}$ , the signal is shifted at  $v_{\text{rad}} = -7.2_{-3.7}^{+4.3} \text{ km s}^{-1}$ , and the in-trail distribution is different from the out-of-trail one at the level of  $5.3\sigma$ , which highly supports our water detection. The CCF and  $\ln \mathcal{L}$  maps ( $K_{\text{P}}$  versus  $v_{\text{rad}}$  maps) are very similar, but are not shown to limit redundancy. The respective best-fit models for the CCF and  $\ln \mathcal{L}$  all lead to convincing detections even though their parameters are different, as also seen in Boucher et al. (2021). Following the same arguments, we will put more credence on the model selection of the  $\ln \mathcal{L}$ , which will be used in the next section (Section 3.6.1) for the full retrieval analysis.

In principle, the RME should be accounted for when molecular or atomic species are predicted to be present in both the stellar and planetary atmospheres. However, this effect is negligible since WASP-127 is a slow rotator ( $v \sin(i_*) = 0.53_{-0.05}^{+0.07} \text{ km s}^{-1}$ ; Allart et al. 2020). The resulting contamination signal would appear at  $K_{\text{P}} \simeq 4 \text{ km s}^{-1}$ , which is far away from  $K_{\text{P},0}$ .

We also observe a tentative  $\text{OH}$  signal, at a scaled level of  $2.8\sigma$ , but at a similar  $v_{\text{rad}}$  as the water signal ( $v_{\text{peak}} = -6.1 \text{ km s}^{-1}$ , at  $K_{\text{P},0}$ ), for a model only containing  $\text{OH}$  with a  $\text{VMR} = 10^{-4}$ . The other model parameters were set to be the same as for the best  $\text{H}_2\text{O}$  model. The

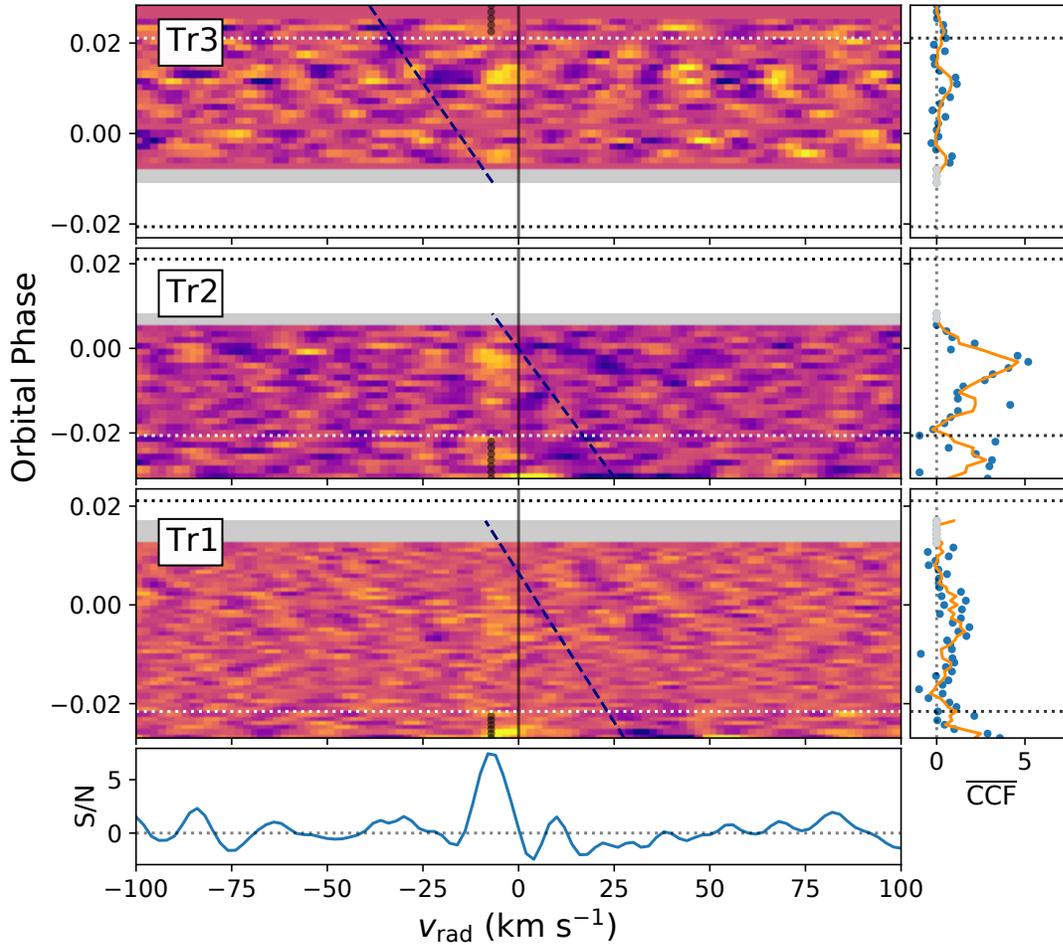


FIGURE 3.4 Planet rest-frame cross-correlation time series of the best-fit H<sub>2</sub>O model for the three SPIRou transits of WASP-127 b. (Left panels) Normalized CCFs from individual exposures as a function of  $v_{\text{rad}}$  for Tr3 (first panel), Tr2 (second panel) and Tr1 (third panel), in the planetary rest frame ( $K_p = K_{p,0}$ ). The navy dashed lines show the position of the BERV, while the horizontal dotted lines show the ingress and egress positions. The black vertical lines show the planet path for  $v_{\text{rad}} = 0 \text{ km s}^{-1}$  (full line) and  $v_{\text{rad}} = v_{\text{peak}}$  (dotted line — only shown out-of-transit to increase visibility during transit). The exposures where the BERV crosses the peak position by less than  $2.3 \text{ km s}^{-1}$  are excluded, and shown in grey. The bottom panel shows the 1D CCF S/N curve for the combined transits. (Right panels) Mean CCF for a 3-pixel column bin centered on  $v_{\text{peak}}$  (blue points) and the 3-exposures binned signal (orange line). The grey dots show the excluded exposures.

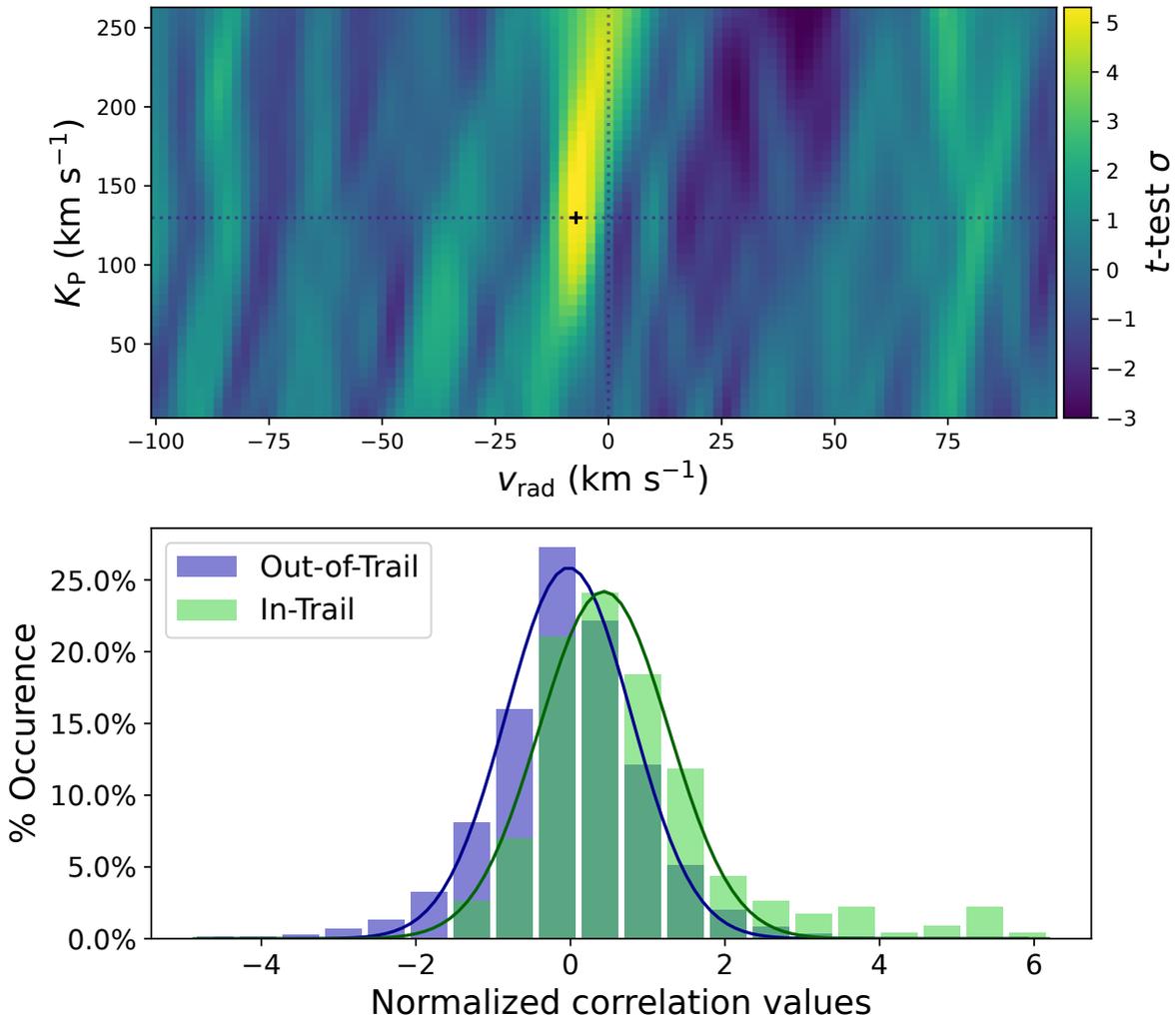


FIGURE 3.5 Welch  $t$ -test results : (Top panel) Scaled  $t$ -test detection significance as a function of  $K_P$  and  $v_{\text{rad}}$ . The H<sub>2</sub>O signal is detected near the expected  $K_P$  value, but blue-shifted (the black cross indicates the peak position at  $K_{P,0}$ ). (Bottom panel) Distributions of cross-correlation values, normalized by the dispersion of the out-of-trail regions, away from (out-of-trail, blue), and near (in-trail, green) the planet RV (with  $K_{P,0}$  and  $v_{\text{rad}} = -7.2^{+4.3}_{-3.7}$  km s<sup>-1</sup>), with their associated best-fit Gaussian distributions, each with their corresponding mean and variance (blue and green lines, respectively). A detection of the transmission signal of the planet is expected to shift the in-trail distribution to higher correlation values, and this is what we observe, with the two distributions differing at the  $5.3\sigma$  level.

resulting scaled  $t$ -test map is shown on Figure 3.6. This finding is discussed in more depth in Section 3.7.1.2.

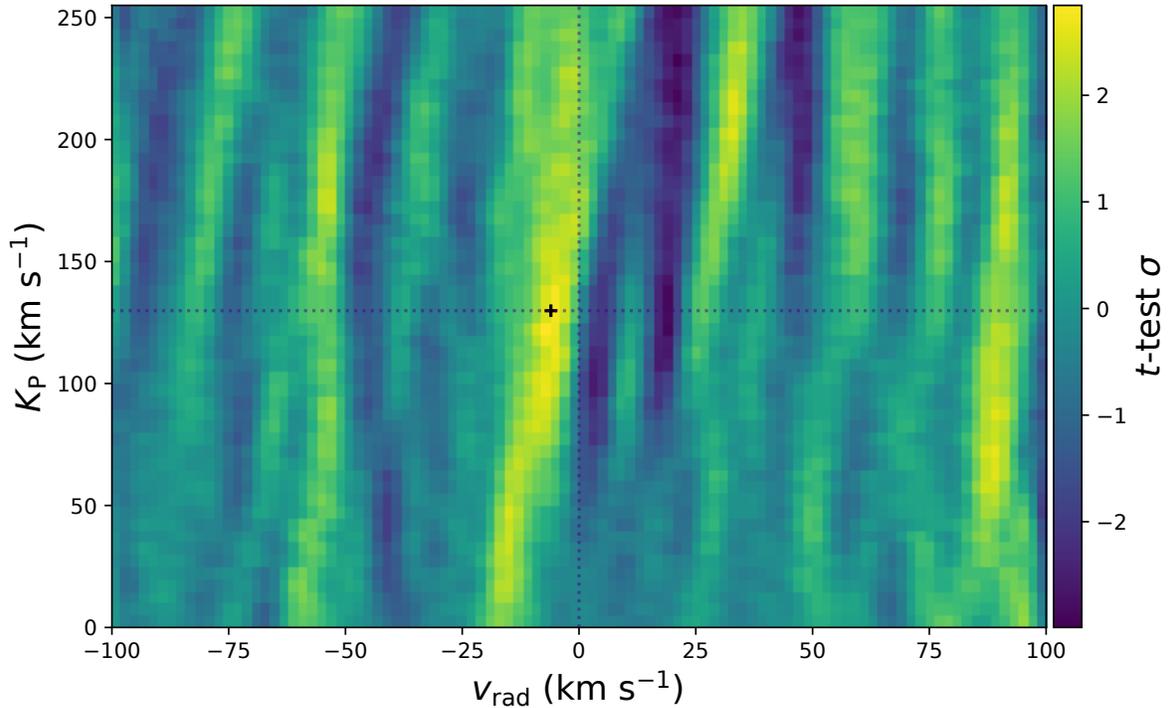


FIGURE 3.6 Same as Figure 3.5, but for an OH-only model with a  $\text{VMR} = 10^{-4}$ . The  $K_{P,0}$  peak of  $2.8\sigma$  is found at  $-6.1 \text{ km s}^{-1}$ , similar to the  $\text{H}_2\text{O}$  signal.

### 3.6.1. Full Retrievals

We next computed full, free-parameter atmospheric retrievals. We ran three types of retrievals : one using only the high-resolution SPIRou data (hereafter HRR, for High Resolution Retrieval), one using only the low-resolution HST WFC3, STIS and *Spitzer* data from S21 (hereafter LRR, for Low Resolution Retrieval), and finally, a joint retrieval by combining the low- and high-resolution data (hereafter JR, for Joint Retrieval). High-resolution models were generated from  $0.9$  to  $2.55 \mu\text{m}$  and low-resolution ( $R = 1000$ ) models from  $0.3$  to  $5 \mu\text{m}$ . Both model sets are generated from the same parameters values. We down-sampled and binned the pre-existing high-resolution models (for JR) or the low-resolution model (for LRR) to  $R = 75$  to match the HST WFC3 data, and  $R = 50$  for STIS. Finally, we binned the low-resolution model over the *Spitzer* IRAC1 (at  $3.6 \mu\text{m}$ ) and IRAC2 (at  $4.5 \mu\text{m}$ ) bandpass transmission functions (JR and LRR).

Following Brogi & Line (2019), we add the contribution of the low-dispersion spectroscopy (LDS) data to the log-likelihood using :

$$\ln \mathcal{L}_{\text{tot}} = \ln \mathcal{L}_{\text{HDS}} + \ln \mathcal{L}_{\text{LDS}}, \quad (3.6.1)$$

where  $\ln \mathcal{L}_{\text{HDS}}$  is given by equation 3.5.4, while  $\ln \mathcal{L}_{\text{LDS}}$  is given by :

$$\ln \mathcal{L}_{\text{LDS}} = -\frac{1}{2}\chi^2; \quad (3.6.2)$$

and the  $\chi^2$  summation is done on the low-resolution data points.

For the HRR, we included the white-light transit depth in the  $\ln \mathcal{L}$  computation. Namely, we made use of the down-sampled model, compared its mean transit depth to the mean of the WFC3 data, and added that to the  $\ln \mathcal{L}$ . This was done to anchor the transit depth to its observed value and to limit the exploration in the parameter space that would otherwise lead to completely different transit depths<sup>11</sup>, since all continuum information is lost in the analysis of the high-resolution data.

We included the opacity contribution from H<sub>2</sub>O, CO, CO<sub>2</sub>, FeH, CH<sub>4</sub>, HCN, NH<sub>3</sub>, C<sub>2</sub>H<sub>2</sub>, TiO, OH, as well as Na and K for completeness and to achieve a better fit to the STIS data. While most included species are not necessarily expected to be detected (based on the sensitivity of the SPIRou data, see Appendix 3.9), useful constraints on their abundances may still be obtained and ruleout certain chemical scenarios. This is also crucial to properly estimate elemental abundance ratios (C/O, C/H, O/H), and to limit the biases that would come from not including all the relevant molecules. We let  $T_{\text{eq}}$  and  $R_{\text{P}}$  vary, while the reference pressure is still fixed to  $P_0 = 10$  mbar. With the inclusion of the STIS data covering shorter wavelengths, we also include the two scattering parameters, i.e. the scattering index  $\gamma$  and the enhancement factor  $S$  from Equation 3.5.1, as free parameters. The seventeen model parameters are thus : the VMRs of H<sub>2</sub>O, CO, CO<sub>2</sub>, FeH, CH<sub>4</sub>, HCN, NH<sub>3</sub>, C<sub>2</sub>H<sub>2</sub>, TiO, OH, Na and K,  $T_{\text{P}}$  (K),  $\log_{10} P_{\text{cloud}}$  (bar),  $R_{\text{P}}$  ( $R_{\text{J}}$ ),  $\gamma$ , and  $S$ . To these, we added the following orbital and dynamical parameters :  $K_{\text{P}}$ ,  $v_{\text{rad}}$ , and FWHM, the latter being the width of the Gaussian broadening kernel. Several of these parameters will be sensitive only when either the low-resolution or the high-resolution data are included, but should not alter the results or lead to scenarios that have no physical interpretations when they are included even without sensitivity. Our atmospheric retrievals therefore include a total of 20 parameters.

We used a Markov Chain Monte Carlo (MCMC) to explore the parameter space and compute the posterior distribution of each parameter and estimate their uncertainties. The MCMC sampling was done using the python library `emcee` (Foreman-Mackey et al., 2013), which implements the affine-invariant ensemble sampler by Goodman & Weare (2010). For each of our runs, we combined the three transits sequences, included all orders, and followed 50 walkers until convergence, i.e. when we had at least 10 times the auto-correlation time in our amount of steps.

---

11. The exclusion of the white-light transit depth in the  $\ln \mathcal{L}$  calculation led to much higher radius and much lower temperature, both of which affect the line contrast in a correlated manner.

Once a steady state was reached, we ran 5200, 28000, and 5300 steps for the HRR, LRR, JR, respectively. The priors that were used are listed in Table 3.4. The priors for all parameters, except  $K_P$ , are uniform. The  $K_P$  prior was chosen to be a Gaussian centered on the previously measured value of  $K_P = 129.95 \text{ km s}^{-1}$ , and  $\sigma = 1.2 \text{ km s}^{-1}$ . Maintaining it as a free parameter allow us to capture the effect of its small uncertainty on the posterior distributions of the other parameters.

The resulting posterior probability distributions for our three types of retrieval (HRR, LRR and JR) and their associated median T-P profiles are shown in Figure 3.7. Their respective median parameters are also tabulated in Table 3.4, along with their  $1 \sigma$  uncertainties, corresponding to the range containing 68% of the MCMC samples, or their  $2 \sigma$  upper or lower limit (for non-detections, at 95.4%). Though all 20 parameters were included in all of the fits, we chose to remove some from the corner plots to improve visualization — namely those which showed less “relevant” non-detections and/or no correlation with other parameters. These parameters are TiO, Na, K,  $K_P$ ,  $v_{\text{rad}}$ , and FWHM. The best-fit models (using the set of parameters yielding the highest  $\ln \mathcal{L}$ , for each type) are shown in Figure 3.8, and all provide a generally good fit to the S21 data. With the highest  $\ln \mathcal{L}$  model parameters, we re-computed the scaled  $t$ -test and found a value of  $5.5\sigma$ ; indicating that the inclusion of the other molecules is indeed beneficial (compared to just having water).

3.6.1.1. Joint Retrieval. The differences between our three types of retrievals (HRR, LRR and JR) are discussed in Section 3.7.2, but from here onward, we will focus on the JR as it provides the best constraints on all parameters.

We detect a slightly sub-solar  $\text{H}_2\text{O}$  abundance of  $\log_{10}\text{H}_2\text{O} = -3.8 \pm 0.3$ , but a super-solar  $\text{CO}_2$  abundance of  $\log_{10}\text{CO}_2 = -4.6 \pm 0.4$  (both values being close to the S21 free-chemistry values)<sup>12</sup>. We note that our constraint on the  $\text{CO}_2$  abundance is mainly driven by the *Spitzer* IRAC2 data point since the  $\text{CO}_2$  upper limit from the HRR is incompatible within  $1\sigma$  to the VMR from the JR. We retrieve upper limits of  $\log_{10}\text{CO} \lesssim -4.3$  ( $2\sigma$  limit, which is sub-solar; discussed further in Section 3.7.1.1), and  $\log_{10}\text{FeH} \lesssim -8.7$  ( $2\sigma$ , discussed further in Section 3.7.1.3). The  $\text{CO}_2$  detection combined with the upper limit found for CO excludes the S21 chemical equilibrium scenario. Since a CO abundance greater than  $\log_{10}\text{CO} \simeq -4$  would have been detected with SPIRou only, not seeing it in both retrievals which used SPIRou (HRR and JR) means that the signal at  $4.5 \mu\text{m}$  comes (in large part) from  $\text{CO}_2$ .

We also obtain non-detections for  $\text{CH}_4$ ,  $\text{C}_2\text{H}_2$ , TiO and K (consistent with Palle et al. 2017b, Allart et al. 2020, and S21). The retrieval slightly favours models including some HCN and  $\text{NH}_3$  (discussed in Section 3.7.1.4), with their upper limits listed in Table 3.4. We recover the Na detection,  $\log_{10}\text{Na} = -6.9_{-0.9}^{+1.1}$ , which is roughly consistent with the retrieved value in S21 ( $\simeq$

12. The solar abundances for a chemical equilibrium `FastChem` model of WASP-127 b, with the retrieved JR  $R_P$  and T-P profile, are  $\log_{10}\text{H}_2\text{O} \simeq -3.4$ ,  $\log_{10}\text{CO} \simeq -3.4$ , and  $\log_{10}\text{CO}_2 \simeq -6.1$ .

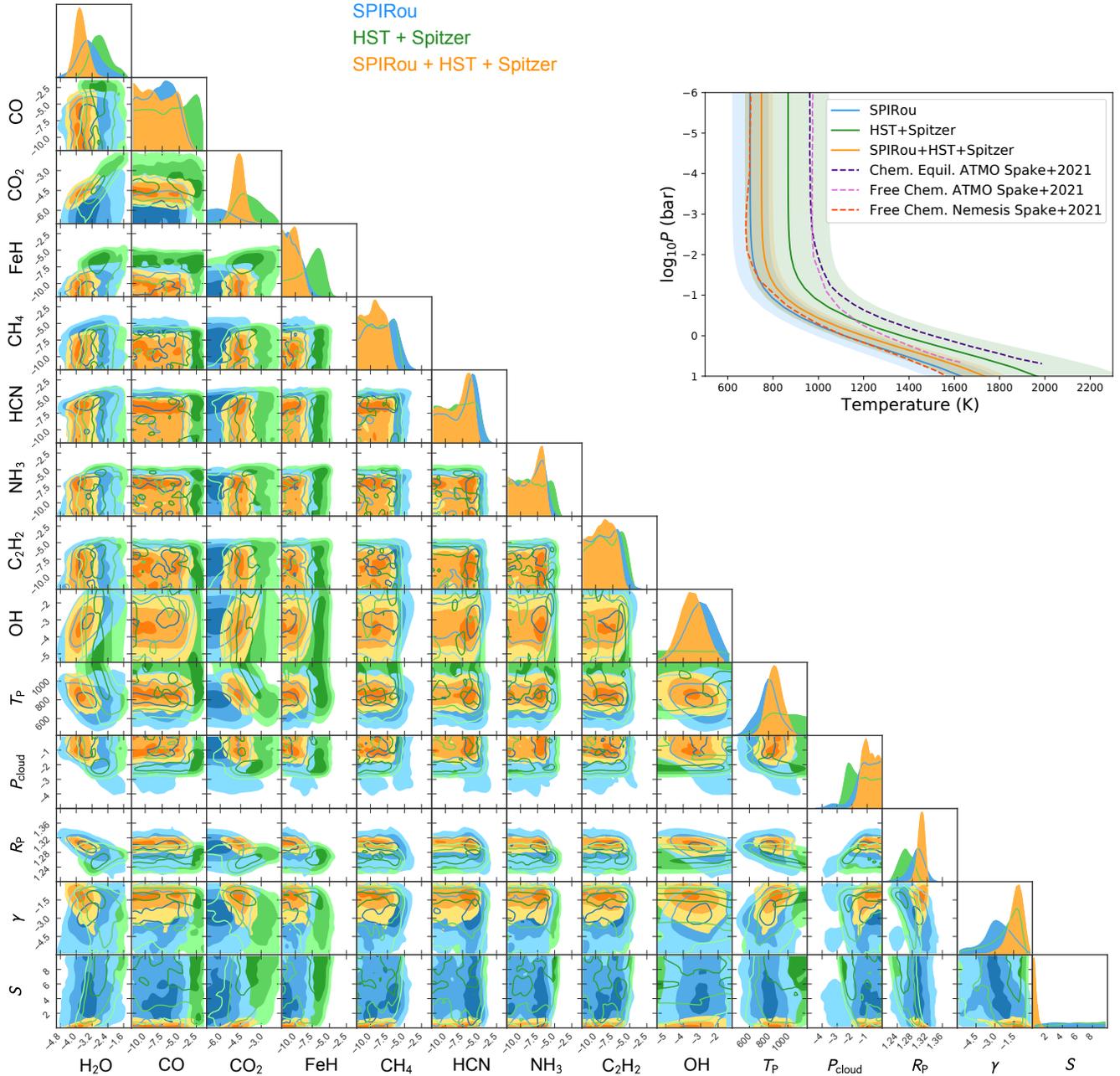


FIGURE 3.7 Constraints on the atmospheric properties retrieved with the three SPIRou transits and/or the S21 HST and *Spitzer* data of WASP-127 b. (Bottom left) Posterior probability distributions for the parameters of the MCMC fits using a parameterized T-P atmosphere with free chemical abundances, a gray cloud deck and hazes, and Gaussian rotational broadening. The results for the high-resolution only (HRR) retrieval are shown in blue, the low-resolution only (LRR) is in green, and the joint retrieval (JR) is shown in orange. The posterior distributions of TiO, Na, K,  $K_p$ ,  $v_{\text{rad}}$ , and FWHM are not shown to improve visibility, and because they do not affect our conclusions. (Top right) Median retrieved Guillot T-P profiles where the shaded region represents the  $1\sigma$  uncertainties, compared to the mean profiles retrieved in S21 (dashed lines). The median retrieved profiles have a temperature of  $1120^{+130}_{-114}$  K,  $1365^{+279}_{-292}$  K, and  $1195^{+78}_{-109}$  K, for HRR, LRR and JR respectively, at 1 bar. They are all consistent with the NEMESIS profile from S21, and within a few  $\sigma$  to the ATMO ones.

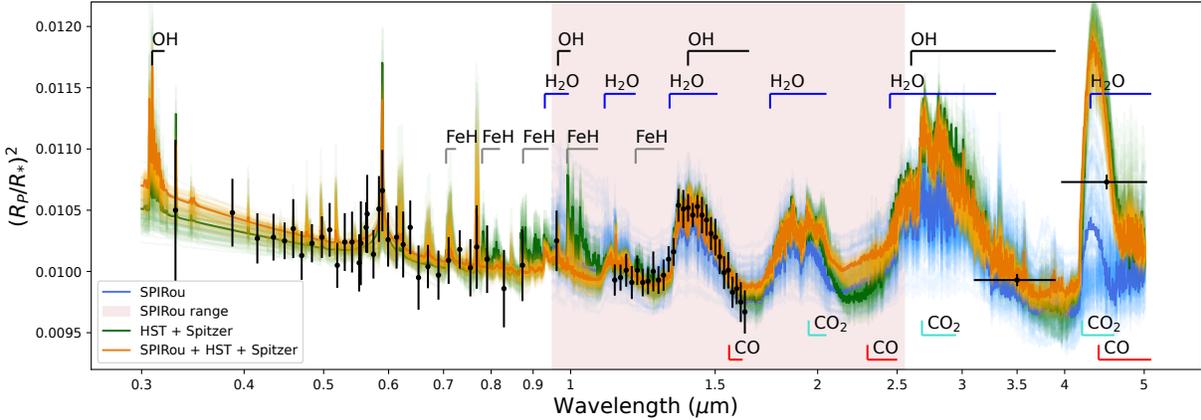


FIGURE 3.8 Best  $\ln \mathcal{L}$  (dark) and 100 randomly generated from MCMC sample models (light) of the transmission spectrum of WASP-127 b (with  $R = 1000$ ) using the HRR (blue), the LRR (green) and the JR parameters (orange), compared to the HST and *Spitzer* data from S21 (black). The red shaded region shows the wavelength range of SPIRou. The position of the absorption bands of the major molecules considered are shown.

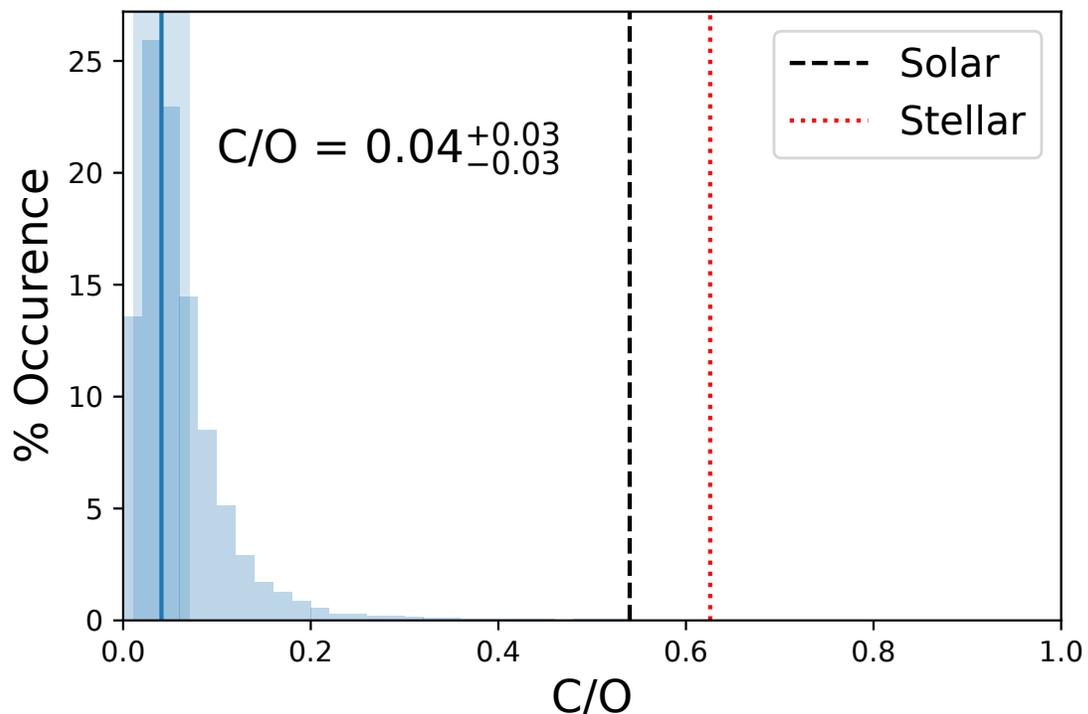


FIGURE 3.9 Posterior distribution of the C/O ratio for the JR retrieval. We find the maximum of the distribution at  $0.04^{+0.03}_{-0.03}$ . This is highly sub-solar/stellar, as seen from comparing to the computed values of solar (black dashed line) and the stellar-composition (red dotted line).

Table 3.4. MCMC Retrieval Parameter Priors and Results for WASP-127 b

Parameter	Priors	SPIRou only	HST+Spitzer only	SPIRou+HST+Spitzer	Unit
$\log_{10} \text{H}_2\text{O}$	$\mathcal{U}(-12, -1)$	$-3.30^{+0.49}_{-0.63}$	$-2.79^{+0.44}_{-0.52}$	$-3.78^{+0.29}_{-0.31}$	
$\log_{10} \text{CO}$	$\mathcal{U}(-12, -1)$	$< -4.42$	$< -1.68$	$< -4.33$	
$\log_{10} \text{CO}_2$	$\mathcal{U}(-12, -1)$	$-5.93^{+1.23}_{-1.01}$	$-3.96^{+0.99}_{-0.92}$	$-4.64^{+0.38}_{-0.42}$	
$\log_{10} \text{FeH}$	$\mathcal{U}(-12, -1)$	$< -8.03$	$< -5.11$	$< -8.67$	
$\log_{10} \text{CH}_4$	$\mathcal{U}(-12, -1)$	$< -5.39$	$< -5.97$	$< -7.09$	
$\log_{10} \text{HCN}$	$\mathcal{U}(-12, -1)$	$< -4.60$	$< -4.72$	$< -5.70$	
$\log_{10} \text{NH}_3$	$\mathcal{U}(-12, -1)$	$< -5.71$	$< -4.76$	$< -6.31$	
$\log_{10} \text{C}_2\text{H}_2$	$\mathcal{U}(-12, -1)$	$< -5.40$	$< -5.20$	$< -6.38$	
$\log_{10} \text{TiO}$	$\mathcal{U}(-12, -1)$	$< -5.03$	$< -8.56$	$< -9.65$	
$\log_{10} \text{OH}$	$\mathcal{U}(-12, -1)$	$-2.85^{+0.82}_{-0.57}$	$-6.82^{+2.03}_{-5.15}$	$-3.51^{+0.61}_{-0.58}$	
$\log_{10} \text{Na}$	$\mathcal{U}(-12, -1)$	...	$-5.02^{+1.57}_{-1.39}$	$-6.91^{+1.12}_{-0.90}$	
$\log_{10} \text{K}$	$\mathcal{U}(-12, -1)$	...	$< -5.15$	$< -6.93$	
$T_{\text{P}}$	$\mathcal{U}(400, 1500)$	$799^{+85}_{-124}$	$1017^{+182}_{-273}$	$860^{+63}_{-90}$	K
$\log_{10} P_{\text{cloud}}$	$\mathcal{U}(-6.0, 0.0)$	$> -2.79$	$> -2.65$	$> -1.59$	bar
$R_{\text{P}}$	$\mathcal{U}(1.0, 1.6)$	$1.29^{+0.02}_{-0.02}$	$1.26^{+0.02}_{-0.02}$	$1.30^{+0.01}_{-0.01}$	$R_{\text{J}}$
$K_{\text{P}}$	$\mathcal{G}(129.95, 1.2)$	$130.45^{+1.14}_{-1.31}$	$129.89^{+1.08}_{-1.29}$	$130.02^{+1.21}_{-1.12}$	km s <sup>-1</sup>
$v_{\text{rad}}$	$\mathcal{U}(-17.0, 3.0)$	$-7.75^{+0.63}_{-0.41}$	$-6.85^{+5.36}_{-7.97}$	$-7.73^{+0.54}_{-0.47}$	km s <sup>-1</sup>
FWHM	$\mathcal{U}(0, 20)$	$9.07^{+1.02}_{-1.31}$	...	$9.52^{+1.07}_{-1.24}$	km s <sup>-1</sup>
$\gamma$	$\mathcal{U}(-6.0, 0.0)$	$-2.27^{+1.37}_{-1.13}$	...	$-1.24^{+0.56}_{-0.44}$	
$S$	$\mathcal{U}(0.0, 10.0)$	$1.91^{+0.95}_{-1.73}$	...	$0.29^{+0.15}_{-0.26}$	
C/O	...	$0.01^{+0.01}_{-0.01}$	$0.04^{+0.04}_{-0.14}$	$0.04^{+0.03}_{-0.03}$	
[C/H] <sup>a</sup>	...	$-1.88^{+0.87}_{-0.68}$	$-0.12^{+0.92}_{-1.16}$	$-1.22^{+0.38}_{-0.41}$	
[O/H] <sup>a</sup>	...	$0.43^{+0.71}_{-0.54}$	$0.61^{+0.62}_{-0.71}$	$-0.17^{+0.34}_{-0.47}$	

Note. — The marginalized parameters from the likelihood analysis with their  $\pm 1\sigma$  errors, or their  $2\sigma$  upper or lower limits. <sup>a</sup> Abundances ratios compared to the solar value in  $\log_{10}$ .

$-7.0 \pm 0.4$ ). Finally, we obtain a surprisingly good constraint on the OH abundance at  $\log_{10} \text{OH} = -3.5 \pm 0.6$ , which will be discussed in more detail in Section 3.7.1.2.

The T-P profile (Figure 3.7) is consistent with the NEMESIS retrieval from S21, but at lower temperature than their ATMO retrieval. The cloud top pressure is located at  $P_{\text{cloud}} \gtrsim 0.025$  bar, which is significantly lower in altitude than what was found by Allart et al. (2020) and Skaf et al. (2020); both roughly around  $10^{-3}$  bar. For Skaf et al. (2020), the low-resolution of their observations could have hampered a precise determination of  $P_{\text{cloud}}$ , and for Allart et al. (2020), this discrepancy could come from the larger value of  $R_{\text{P}}$  that they used : a larger planet radius would need higher clouds to dampen a given signal amplitude. Still, the LRR  $P_{\text{cloud}}$  is in line with the literature values, which seems to indicate that it is driven by the low-resolution data.

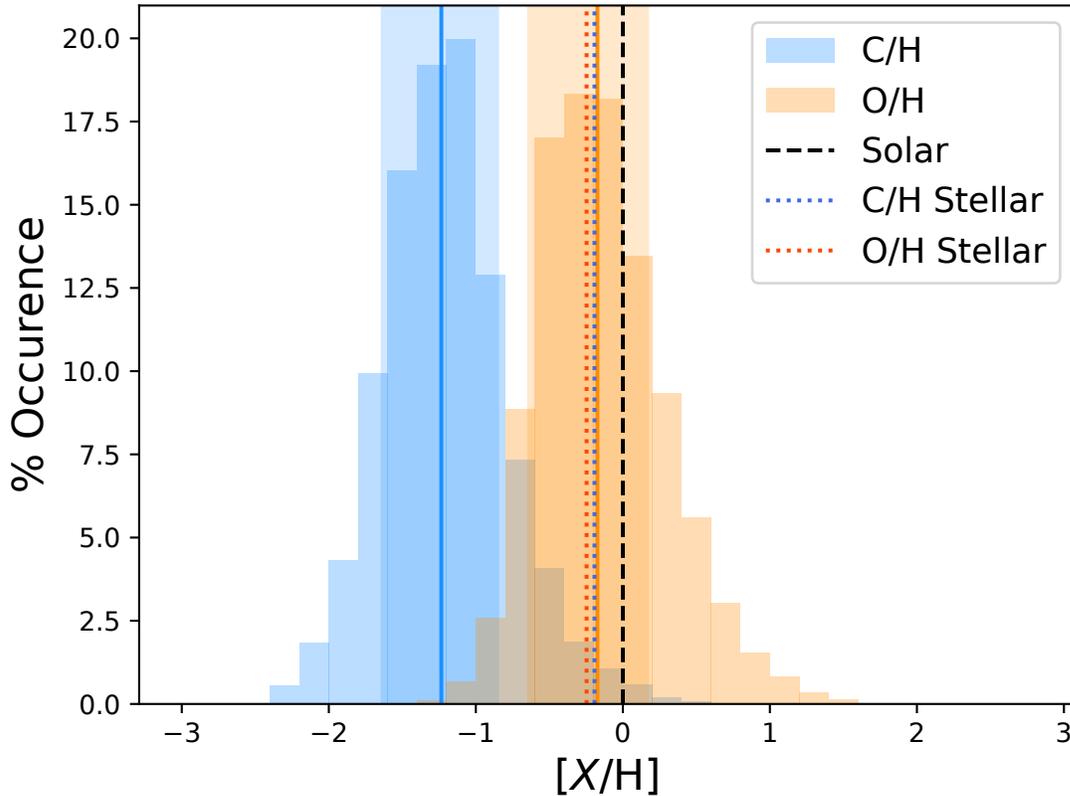


FIGURE 3.10 Distributions of the  $[C/H]$  (blue), and  $[O/H]$  (orange) ratios of the JR, i.e. the  $\log_{10}$  of abundance ratios compared to the solar values. The color-matched lines show the maximum of the distribution and the shaded region shows the  $1\sigma$  uncertainties. The black dashed and colored dotted (blue and red) lines shows the solar and stellar ratios, respectively (assuming that the stellar atomic C and O abundances follow  $[M/H] = -0.193$ ). Although both are sub-solar, we can see that  $O/H$  is close to stellar, while the  $C/H$  is highly sub-stellar.

Comparatively, in S21, they did not include a grey cloud deck contribution as they did not find evidence in the data to support its presence, but rather a stronger haze signal.

On that matter, we observe that our retrieved radius of  $R_p \simeq 1.31 R_J$  (for  $P_0 = 10$  mbar) is much smaller than the ATMO values in S21 ( $R_p$  between  $1.38$  and  $1.45 R_J$  at  $1$  bar, roughly equivalent to  $\sim 1.53$ – $1.61 R_J$  at  $P_0$ ), but is more consistent with their NEMESIS radius ( $1.30 \pm 0.2 R_J$ ), and also with that of Skaf et al. (2020) ( $R_p \simeq 1.16 R_J$  at  $10$  bar, equivalent to  $1.32 R_J$  at  $P_0$ ). It is also in line with the value from Seidel et al. (2020b), at  $1.30 \pm 0.2 R_J$ , but this is not surprising as most of our fixed parameters were taken from their work. Even though  $R_p$  is degenerate with the chemical abundance,  $P_{\text{cloud}}$ , and temperature, our constraint on the radius

is extremely tight due to the combination of the low- and high-resolution data, which will be discussed further in Section 3.7.2.

The retrieved RV semi-amplitude is  $K_p = 130.3 \pm 1.1 \text{ km s}^{-1}$ . Uncorrelated/non-degenerate posterior distributions are found for the shift of the planet signal,  $v_{\text{rad}} = -7.7^{+0.5}_{-0.5} \text{ km s}^{-1}$ , and the width of the “rotational” broadening Gaussian kernel,  $\text{FWHM} = 9.5^{+1.1}_{-1.2} \text{ km s}^{-1}$ . Such a large blue shift is unusual, but not unheard of and is discussed in more details in Section 3.7.3. A large width for the rotational broadening was seen in all the types of kernel that we used, and could imply a rapidly rotating planet, although its expected synchronous rotation rate is only  $\sim 1.6 \text{ km s}^{-1}$ , but super-rotation could amplify the broadening. A puffy atmosphere like WASP-127 b’s could have interesting dynamics and 3D models certainly could help shed light on these results.

Looking back at Figure 3.7, we see that the abundances of some molecules are correlated with  $P_{\text{cloud}}$ , which results in broader constraints on absolute abundances. However, the abundance posteriors of some molecules are correlated (especially  $\text{H}_2\text{O}$  and  $\text{CO}_2$ ) due to their similar dependancies on the continuum level, which means that their relative abundances can be accurately constrained (Gibson et al., 2022). The  $\text{H}_2\text{O}/\text{CO}_2$  ratio is constrained to  $\log_{10} \text{H}_2\text{O}/\text{CO}_2 = 0.9^{+0.2}_{-0.3}$ , i.e.,  $\text{H}_2\text{O}$  is between  $\sim 4$  and 13 times more abundant than  $\text{CO}_2$  in WASP-127 b’s atmosphere.

We can also combine the retrieved constraints of all carbon- and oxygen-bearing molecules to compute the atmospheric C/O ratio as :

$$\text{C/O}_{\text{gas}} = \frac{\text{CO} + \text{CO}_2 + \text{CH}_4 + \text{HCN} + 2\text{C}_2\text{H}_2}{\text{H}_2\text{O} + \text{CO} + 2\text{CO}_2 + \text{TiO} + \text{OH}}, \quad (3.6.3)$$

assuming that no other major C- or O- bearing molecules are present. We computed C/H and O/H in a similar fashion, but using  $\text{H} = 2\text{H}_2\text{O} + 4\text{CH}_4 + \text{HCN} + 3\text{NH}_3 + 2\text{C}_2\text{H}_2 + \text{OH} + 2(1 - \sum_i X_i)\text{H}_2$ , where the last term represents twice the total abundance of  $\text{H}_2$ , i.e. 0.85 scaled by the total minus the sum of the other molecules. For comparison, we generated `FastChem` chemical equilibrium models of planetary atmospheres with solar<sup>13</sup> and stellar compositions (where we assumed that the stellar C, N, and O abundances followed the estimated metallicity trend of  $[\text{M}/\text{H}] = -0.193$ ; Stassun et al. 2019), and using the retrieved JR T-P profile and  $R_p$ . We took the median abundance over the pressure region between  $10^2$  and  $10^{-6}$  bar, and then computed the atomic ratios with the same equation 3.6.3. From those abundances, we obtain solar and stellar composition C/O equal to 0.55 and 0.63, respectively, while the C/H and O/H are respectively  $2.73 \times 10^{-4}$ , and  $4.94 \times 10^{-4}$  for the solar composition mix, and  $1.75 \times 10^{-4}$ , and  $2.80 \times 10^{-4}$  for stellar composition. The resulting posterior of the C/O for WASP-127 b is shown in Figure 3.9,

---

13. The solar abundance from `FastChem` are based on Lodders & Palme (2009)

the C/H and O/H distributions are shown in Figure 3.10, and their best-fit values are listed in Table 3.4.

We obtain a highly sub-solar C/O of  $0.04_{-0.03}^{+0.03}$ , compared to the computed  $C/O_{\odot} = 0.55$  (for the solar composition mix, close to the common literature value of 0.54, as expected), and the computed stellar composition<sup>14</sup> one of  $C/O_{*} = 0.63$ . This indicates that the atmosphere of WASP-127 b is either highly oxygen-rich or carbon-depleted, compared to what it should be with solar or stellar composition. A solar value is rejected at more than  $16\sigma$ .

On the contrary, the retrieved C/H, and O/H being highly sub-stellar and roughly stellar, respectively, shows that the low C/O is due to C depletion rather than an O-rich environment. The choice of which molecules are included in these ratio calculations can significantly change the inferred results. Nonetheless, since we computed the expected planetary atmosphere abundance ratios for solar and stellar compositions in the same way as for our actual results, it should be a good means of comparison. The formation and evolution mechanisms that could reproduce such elemental abundance ratios in WASP-127 b’s atmosphere are explored in Section 3.7.4.

## 3.7. Discussion

The primary goal of this study was to differentiate between the two S21 scenarios — a goal we accomplished by showing that WASP-127 b does not have a CO abundance high enough to explain the S21 *Spitzer* results (expected to be around  $\log_{10}\text{CO} \simeq -1.9$  for the chemical equilibrium scenario), but rather has  $\log_{10}\text{CO} < -4.3$ . We were able to set a tight constraint on the C/O ratio, at  $C/O = 0.04_{-0.03}^{+0.03}$ , implying a severely carbon-depleted atmosphere. Finally, we detect water at a very high significance, and potentially detect a signal consistent with OH. These partially unexpected results should be put in context to be better understood. In this section, we will first discuss some of the molecular detections and non-detections that were obtained. A comparison between the different types of retrievals will follow. Then, we will explore the potential formation mechanisms of WASP-127 b. Finally, we will briefly review the observed blue-shift.

### 3.7.1. Molecular (non-)detections

3.7.1.1. CO and CO<sub>2</sub>. The non-detection of CO is surprising because, as stated in Spake et al. (2021), a scenario where all of the carbon is found in CO<sub>2</sub> and little to none in CO would be thermochemically implausible as no obvious non-equilibrium mechanism would deplete CO by many orders of magnitude while enhancing CO<sub>2</sub>. The relative abundance between CO<sub>2</sub> and CO has been explored in Heng & Lyons (2016). They conclude that CO<sub>2</sub> should rarely be dominant

---

14. To be more specific, these solar and stellar composition abundance ratios are computed using the same sub-sample of molecules that are used to compute our retrieved WASP-127 b values.

compared to CO ( $\text{CO}_2/\text{CO} \simeq 1$ ) and  $\text{H}_2\text{O}$  in hot and  $\text{H}_2$  dominated atmospheres, and that it can be dominant if the C abundance is enhanced by a few orders of magnitude compared to solar, which is not our case (see also Moses et al., 2013b; Line et al., 2014). As mentioned above, we get a strong constraint on  $\text{CO}_2$ , but this detection and its retrieved abundance ( $\sim -5$ ; which could still be sub-dominant to CO given its upper limit at  $\sim -4.3$ ) are highly reliant on the IRAC2  $4.5 \mu\text{m}$  point. The best-fit HRR model (SPIRou only) favors the inclusion of  $\text{CO}_2$ , but with a much smaller abundance (with a  $\log_{10}$  VMR  $\sim -6.95$ ), while the posterior distribution only offers an upper limit.

This is discussed further in Section 3.7.4, and some explanations are proposed to answer this surprising result.

3.7.1.2. OH. Our results indicate a marginal OH signal, just under the limit of significance according to our scaled  $t$ -test ( $2.4\sigma$ ). The OH signal is highly visible in Tr2, partially visible in Tr1, but not visible in Tr3 (which has relatively poorer data quality). Overall, we consider this detection as tentative and it will necessitate future confirmation. From a pure equilibrium chemistry standpoint, OH is not expected to be present on WASP-127 b as it would normally imply  $\text{H}_2\text{O}$  dissociation, and WASP-127 b’s equilibrium temperature of 1400 K is too low for such dissociation to occur ( $\text{H}_2\text{O}$  dissociation is expected to occur for  $T_{\text{dayside}} \gtrsim 2500 \text{ K}$ ; Parmentier et al. 2018). As our OH detection is uncertain, we note that the exclusion of OH from our analysis does not greatly affect the O/H and C/O abundances ratios :  $\log_{10}$  O/H decreases to a value of  $-3.90^{+0.29}_{-0.33} \simeq 0.25 \times \text{solar} \simeq 0.45 \times \text{stellar}$  (compared to its median value of  $-3.51^{+0.34}_{-0.47} \simeq 0.62 \times \text{solar} \simeq 1.1 \times \text{stellar}$ , when including the OH), becoming slightly sub-stellar rather than stellar, while the C/O increases to  $\simeq 0.11^{+0.05}_{-0.04}$ , which is still highly C-depleted.

We tested a few things to challenge the validity of the signal. We first looked for contamination from residual sky OH emission lines, but found no obvious signal following the BERV (either positive or negative). Plus, we already exclude the exposures where  $v_p$  crosses the BERV in all of our analyses. We then verified whether the signal could be an alias originating from cross-correlation between the different molecular absorption spectra, but did not see significant signals around  $0 \text{ km s}^{-1}$  with any other molecular specie that we included in our analyses. Then, we excluded the OH molecule from the best-fit JR model : the resulting CCF S/N increased ( $\sim 0.5 \text{ dex}$ ), but the  $t$ -test is decreased by  $\sim 0.8\sigma$ , and the  $\ln \mathcal{L}$  significantly dropped by  $\Delta \text{BIC} \simeq 25$ . This, again, seems to imply that the signal actually comes from OH, or at least, something with a similar absorption spectrum.

A priori, the presence of OH high in a planet’s atmosphere could be due to photochemistry. To check, we generated VULCAN models of WASP-127 b (Tsai et al., 2017)<sup>15</sup> that include photo-chemistry, and used the JR retrieved T-P profile. We observed that OH is present in small

<sup>15</sup>. VULCAN is an open-source photochemical kinetics Python code for exoplanetary atmospheres, in which FastChem is implemented.

abundances at  $P \gtrsim 1$  bar ( $\simeq 10^{-11}$ ), slightly decreases until  $P \simeq 0.1$  bar, but re-increases to peak again at values of  $\sim 10^{-6}$  around  $P \simeq 10^{-7}$  bar. These amounts are still too low to be detected anywhere in the atmosphere. If we artificially increase the temperature to  $\sim 1400$  K (more representative of what the day side *could* look like), the OH abundance increases to  $\sim 10^{-8}$  and  $10^{-5}$ , at  $P \simeq 1$  and  $10^{-7}$  bar, respectively. This abundance could potentially be detected (according to the injection-recovery tests, see appendix 3.9, Figure 3.11), but it is still lower than the observed abundance. We note that the stellar spectral energy distribution used to compute our VULCAN models might not be representative of the true spectrum of WASP-127, as we did not account for the fact that the star is leaving the main sequence. We simply used the same PHOENIX model as the one used to compute the systemic RV above. Also, we assumed a stellar composition for the planetary atmosphere, but another composition could be affected differently. The above calculations are rather limited and a more detailed study would be needed to address this possibility.

Other possibilities may explain the OH presence. For instance, if the day side were hot enough to thermally-dissociate some  $\text{H}_2\text{O}$ , winds could then transport the resulting OH to the terminator region where it would be visible to transit observations. Also, vertical mixing could bring up OH from deeper and hotter atmospheric layers and make it visible. Finally, we cannot exclude the possibility of a spurious signal that happens to match the OH line structures and land at the correct  $v_{\text{rad}}$ . Overall, strong disequilibrium processes are necessary to maintain the presence of OH.

Obtaining a more representative T-P profile of the day-side of WASP-127 b through emission spectroscopy, which will be possible with JWST, could help illuminate this puzzling result. Additional transmission observations at high-resolution would also be beneficial.

3.7.1.3. FeH. As briefly mentioned previously, Skaf et al. (2020) reported a tentative detection of FeH which motivated its inclusion in this work. They argued that their signal cannot be caused by clouds, but conclude that it is still possible that they observe another, yet unidentified, opacity source with absorption features resembling those from FeH. S21 also suggest that there may be other unknown and unresolved absorbers between  $0.8$  and  $1.2 \mu\text{m}$ . Additionally, Kesseli et al. (2020) found no statistical evidence of FeH in any of the 12 hot-Jupiters they analyzed using high-resolution CARMENES observations. They argued that the previous tentative FeH detections from Skaf et al. (2020) (and others) could be caused by their use of low-resolution spectra, and further insist that it makes it more difficult to distinguish species with overlapping opacities and differentiate them from the continuum opacity. They also show how the FeH condenses at temperatures below  $1800$  K, which then makes WASP-127 b much too cold to harbor FeH vapour, in an equilibrium standpoint. Our SPIRou-only injection-recovery tests indicate that if FeH was present with an abundance down to  $\sim 10^{-8}$ , we would have easily detected it

in our analysis (Appendix 3.9, Figure 3.11). From the JR results, we conclude that there is no evidence for FeH in the atmosphere WASP-127 b, at least in an abundance greater than  $\sim 10^{-8.7}$ .

3.7.1.4. N-bearing molecules. Our focused searches for HCN and NH<sub>3</sub> were unsuccessful. However, when analyzing their posterior distributions in the complete retrieval (Figure 3.7), their upper limits display peaks instead of plateaus like the other non-detections. These peaks are not considered proper detections, but since they are at, or close to, the detection limits for both species, we investigated the question further. Similarly to what we did for OH, we compared the best-fit model against one without HCN and NH<sub>3</sub>. When excluding both of these molecules, the decrease in CCF S/N,  $t$ -test and  $\ln \mathcal{L}$  are minimal, i.e.  $\sim 0.15$  dex,  $0.15\sigma$ , and  $\Delta\text{BIC} \simeq 4$ , respectively. The latter still falls in the “marginal” evidence regime (Kass & Raftery, 1995). Given the low significance of this signal, we still consider these non-detections, but do not exclude the possibility that HCN and NH<sub>3</sub> are present in the atmosphere of WASP-127 b. Further observations could also help validate this.

### 3.7.2. Comparison between the types of retrievals

The retrieved parameters across our three different retrievals are roughly consistent with one another, with, not surprisingly, the best constraints obtained with the JR. In all retrievals H<sub>2</sub>O is detected, but the LRR and HRR seem to be less constrained and suffer more from degeneracies with  $P_{\text{cloud}}$  and the scattering factor  $S$ . Likewise, CO<sub>2</sub> is convincingly detected in the LRR and JR, but is at the limit of detection in HRR, which was expected from our injection recovery tests (Appendix 3.9, Figure 3.11). From the CO posterior distribution of the LRR, we can see that the lack of independent spectral features for CO<sub>2</sub> pulls it towards high CO abundances — a result consistent with the S21 findings. This high CO region is excluded when the SPIRou data is considered thanks to the availability of the 2.3  $\mu\text{m}$  CO band. A similar result can be seen for FeH, where the LRR accepts higher abundance of  $\log_{10} \text{FeH} \sim -6.6^{+1.8}_{-2.3}$ . This is, again, compatible with the findings from S21 : they were able to put a constraint on FeH at  $\log_{10} \text{FeH} \sim -7.6$ , but did not claim a detection. This “high” abundance is most likely due to the large range in  $P_{\text{cloud}}$  allowed by, and the limited wavelength coverage of the WFC3 data, and the degeneracy with  $S$ . Looking at the LRR model in Figure 3.8 (in green), the FeH signal can be seen around 0.9–1  $\mu\text{m}$ , but no major features are captured by the STIS data points, which permits a certain amount of FeH to be present. Still, the presence of FeH is ruled-out by the SPIRou data.

The  $T_{\text{p}}$  and  $R_{\text{p}}$  values of the three retrievals are all very similar to one another. The inclusion of the white light transit depth in the HRR helps to better constrain the continuum level. We observed that by not including the transit depth, of all the retrieved parameters,  $T_{\text{p}}$  and  $R_{\text{p}}$  were the most affected, leading to much lower temperatures, but higher radius values. This led to a model that had an offset in transit depth, but when it was scaled to match the observed transit

depth, the overall structures in the HST WFC3 data were well reproduced (like the other models in Figure 3.8). This means that the HRR is good at extracting the information on the line contrast. The retrieved  $T_p$  from the HRR and JR give T-P profiles that are more in line with the S21 NEMESIS one, while the LRR seems to be more consistent with ATMO ones, even though all are consistent with one another within  $1-2\sigma$ . As expected, the orbital parameters ( $K_p$ ,  $v_{\text{rad}}$  and FWHM; not shown) are all nearly identical for HRR and JR.

The cloud top pressures inferred from the retrievals including the high-resolution data are consistent with one another, but look different from the LRR, which is broader and extends to smaller values (higher-altitude cloud deck). This arises from the known strong correlations between  $P_{\text{cloud}}$ , the absolute chemical abundances, the temperature, and also with the reference radius and pressure (e.g., Benneke & Seager, 2012; Griffith, 2014; Heng & Kitzmann, 2017), preventing us from putting tight constraints on the absolute abundances. At high-resolution, we benefit from having access to the relative line strengths above the continuum, but the removal of the continuum through data processing also causes additional degeneracies (e.g., Fisher & Heng, 2019; Brogi & Line, 2019; Gibson et al., 2020). Combining the continuum information from low-resolution with the line shape and contrast at high-resolution (and also relative strength of the broad spectral features; Sing 2018; Pino et al. 2018; Khalafinejad et al. 2021; Gibson et al. 2022), while also marginalizing over all affected parameters, allows for much better constraints. Overall, even though a combination of slightly higher-clouds and higher-abundances cannot be completely ruled out ( $\sim 0.01$  bar; in the HRR and JR), the low-clouds and lower-abundances seem to be preferred. A cloud level at  $\sim 10^{-3}$  bar is ruled out by our results.

Concerning the scattering properties, they all seem to agree on non-Rayleigh scattering, with  $\gamma$  values larger than  $-4$ , although the HRR is unsurprisingly less constrained. The constraint on the  $S$  factor highly benefits from the combination of the low- and high-resolution data, as none of them appears to be able to constrain it on their own.

Finally, all the C/O posterior distributions are consistent between the retrievals, but again, the best constraints come from the JR. They all favor a low C/O, with a carbon depleted atmosphere. The LRR allows for higher C/O (double peak distribution with principal peak at low C/O, not shown) as though it could be an hybrid combination of the two S21 scenarios. Our free chemistry results and theirs are not quite the same (not as constrained on the low-C/O side) likely because of the difference between our levels of clouds and hazes and theirs. Compared to the other retrievals, the LRR is consistent with stellar C/H and O/H within the  $1\sigma$  uncertainties.

This demonstrates how one can improve the determination of atmospheric parameters by combining low- and high-resolution data sets. The low-resolution allows to get the continuum and broadband information, while the high-resolution allows to probe the line shape and contrast. This grants the ability to lift some degeneracies that would arise from having either one or the other. Future atmospheric analysis will thus benefit from having both types of data.

### 3.7.3. Radial Velocity Offset

As transmission spectroscopy probes only the limb of a planet’s atmosphere at the terminator, it is naturally sensitive to day-to-night winds which result in a net blue shift of the absorption signal in transit (Snellen et al., 2010; Miller-Ricci Kempton & Rauscher, 2012). Here we measure a strong blue-shift of  $-7.7_{-0.5}^{+0.5}$  km s<sup>-1</sup> for WASP-127 b. This shift is seen in each of the three transits, independently of the analysis steps applied. It is hard, however, to conclude whether or not this shift results from global day-to-night winds or something else. Such large blue shifts have been observed in several targets. For example, the CCF signal in the  $K_P$  versus  $v_{\text{rad}}$  map of the ultra-hot Jupiter WASP-76 b (Ehrenreich et al., 2020) is highly blue shifted, but this is caused by the signal mostly coming from the trailing (evening) limb of the terminator because of non-uniformity of the atmosphere. Wardenier et al. (2021) investigated this further with 3D general circulation models (GCMs) and confirmed that their different models with and without drag also displayed a blue-shifted signal. For some models, it even displays a double peaked CCF. The blue-shift is intensified when they introduce an asymmetry in composition in the terminator limbs (i.e., a depleted abundance in the leading/morning limb compared to the trailing limb). Water might condensate on the night side and be depleted on the morning limb, while being relatively abundant in the evening limb, which would lead to a faint to invisible signal that is progressively more blue-shifted during ingress, and a blue shifted signal during total transit and egress. This, in addition to super-rotational winds could be a plausible explanation for this high blue shift. Even if WASP-127 b is not a ultra-hot Jupiter like WASP-76 b, its highly inflated state could have an impact on the dynamics. Similarly to the reasons mentioned by (Line et al., 2021) to explain the noticeable offset in their emission signal of WASP-77 Ab, the shift we detect could be caused by the propagation of the error in the mid-point, an unnoticed eccentricity, the rotation (which would preferentially blue-shift a day-side hot spot) and/or longitudinal temperature advection (in the form of eastward winds), or a combination of a few of these hypotheses. In that sense, a signal dominated by the hotter, more inflated, and more blue-shifted trailing limb, combined with a cloudy leading limb, an overall day-to-night wind, and potential errors on the mid-point, could potentially generate such a blue shift. 3D GCMs would be useful to see if and how such a large blue-shift could be explained (Flowers et al., 2019; Beltz et al., 2021; Wardenier et al., 2021; Savel et al., 2022). Nonetheless, Line et al. note that whatever the source of this shift, it should not affect the retrieval of the chemical abundance and/or the thermal profile because it is non-degenerate with those atmospheric parameters.

If the shift comes from an asymmetry in temperature and/or composition between the trailing and evening limbs, then our retrieved chemical abundances and temperature profile should be used as estimations only. MacDonald et al. (2020) showed that using 1D atmospheric models could lead to smaller retrieved temperature and smaller temperature gradient than reality, but

also to over-estimations of the chemical abundances (also Pluriel et al., 2020). It would thus be highly valuable to apply retrieval techniques using 3D models.

### 3.7.4. Implications for the Formation of WASP-127 b

In our analysis, we uncovered an unexpected combination of atmospheric C/O and X/H abundances.

Such low C/O values ( $\sim 0.04$ ) have rarely been seen. The free chemistry scenario from S21 yielded low C/O, but could not be confirmed at the time. Lavie et al. (2017) estimated the C/O of HR 8799 d and e to be zero, but that was later revised to a solar value by Mollière et al. (2020). MacDonald & Madhusudhan (2019) were able to put an upper limit of  $C/O < 0.33$  for the exo-Neptune HAT-P-26 b, with a super-solar metallicity, and suggested a formation history with significant planetesimal accretion. From static disk formation scenarios, our measured value of  $C/O = 0.04^{+0.03}_{-0.03}$  suggests an oxygen-rich environment, likely due to significant solid accretion. Conversely, our determined sub-solar C/H ratio would imply a formation scenario dominated by metal poor gas accretion. The combination of these two together is puzzling, as a formation scenario dominated by the accretion of icy solids would tend to elevate the C/H and especially the O/H ratio above what we find. Alternatively, a gas-dominated accretion scenario beyond the CO or CO<sub>2</sub> ice lines to explain the low C/H ratio would naturally result in a C/O ratio higher than what we find (e.g., Öberg et al., 2011). Our observed abundance ratios are thus hard to reconcile with the composition trends predicted by most existing planet formation models that rely on static disks (Madhusudhan, 2019).

Overall, three scenarios could explain what we observe : 1) either C is present, but in invisible phases, 2) C is depleted, but not O, or 3) O is added back to a metal poor gas. In scenario 1, the composition could be stellar, but where a large amount of C-bearing elements would be in forms that are not probed in our transmission spectra, such as condensed hydrocarbons or graphite phases. This would lead to under-estimations of both the C/H and C/O. In scenario 2, only the C is depleted, which is more difficult to explain through formation in static disks, since in such cases, a carbon depletion also implies an oxygen depletion. In scenario 3, both C and O would be originally depleted, from a metal poor gas accretion (e.g., beyond the CO ice line), and would then accrete O-rich material. The last two cases would lead to a low C/H and C/O. We explored the scenarios 2 and 3 more in depth, assuming that the C-bearing molecules are in vapour phase, and investigated composition predictions from studies that include the effects of evolving disk chemistry (e.g., Eistrup et al., 2018; Cevallos Soto et al., 2022).

Eistrup et al. (2018) employed an extensive kinetic chemistry gas-grain reaction network to evolve the abundances of chemical species in a proto-planetary disk over time, considering different irradiation levels and initial chemistries. One possible formation scenario resulting

from this would be that WASP-127 b formed near 10 AU, and after 5–7 Myr. At this position in space and time, Eistrup et al.’s models show a decrease in CO gas with an increase in frozen CO<sub>2</sub> (see their Figure 7c). This comes from the collision of CO gas with solid grains, and the subsequent fast reaction with OH to produce CO<sub>2</sub> on the grain surface. The OH radicals can be produced by the dissociation of H<sub>2</sub>O by cosmic rays, which would also diminish the local H<sub>2</sub>O abundance, but in this case it would be produced via O<sub>2</sub> freeze-out. Beyond the O<sub>2</sub> iceline, the O<sub>2</sub> freezes-out, is further hydrogenated on grains, is then dissociated into OH ice, and finally reacts with the CO to form CO<sub>2</sub>. This leads to a large drop in CO gas-phase beyond  $\sim 10$  AU after 5 Myr, and increase in CO<sub>2</sub> ices. The gaseous C/O, C/H, and O/H ratios from Eistrup et al. (2018) models (their Figure 9) at 5–7 Myrs and near 10 AU fit almost perfectly our retrieved values. Then and there, the bulk of the gas is highly C-depleted, and is also nearly O-depleted. Additional accretion of H<sub>2</sub>O and CO<sub>2</sub> ices (both still highly abundant in the disk, and available as building blocks for the core) with mixing in the envelope would help keep the O/H at stellar values (or slightly above, as observed), while lowering the already low gaseous C/O to near zero values (as observed). This scenario would imply that the disparity in CO and CO<sub>2</sub> was already present at the time of formation, and somehow has survived until today.

Late formation could be a plausible scenario given that the low-metallicity of WASP-127 implies a lower density of solids in the protoplanetary disk — meaning that the initial planet core growth is more difficult and thus takes longer in the core accretion scenario (Alessi & Pudritz, 2018). By the time it was ready to accrete gas in a runaway fashion, the disk could have been already partly dissipated (as demonstrated in Eistrup et al. 2018). Such a scenario could furthermore explain why the growth of WASP-127 b was truncated at a sub-Saturn mass instead of continuing to a super-Jupiter size — the requisite materials were simply no longer present in the disk to facilitate this runaway growth (Pollack et al., 1996).

More recently, Cevallos Soto et al. (2022) explored the impact on the planetary atmosphere composition from including the effects of radial pebble drift, gas advection, mass accretion and viscosity in their chemically evolving disks. Pebble accretion is an efficient way to build a core quickly, even with a lower amount of solids, and allows the accretion of substantial amount of gas before the disk disperses (Lambrechts & Johansen, 2012). We find that one of their scenarios could potentially match our results, even at times as early as 10<sup>5</sup> yrs. However, it would still require accretion of both gas and ices, with significant envelope-core mixing to match our observations (e.g., Ali-Dib, 2017; Madhusudhan et al., 2017; Öberg & Wordsworth, 2019).

The best match comes from the cases including both advection and pebble drift, with fixed mass accretion rate and pebble size (at 10<sup>-9</sup> M<sub>⊙</sub> yr<sup>-1</sup> and 0.3 cm; their case d). In this scenario, they show that radial pebble drift puts all the CO in gas-form, while inside  $\sim 10$  AU the H<sub>2</sub>O and CO<sub>2</sub> ices are enhanced. Thus, a formation with gas-dominated accretion of the metal-poor gas

beyond the CO ice line with subsequent migration and accretion of H<sub>2</sub>O and CO<sub>2</sub> ices offers a possible explanation for what we observe.

Also, the inclusion of pebble drift yields “valleys” in the gaseous C/O ratio (around 1, 10, and  $\sim 80$ ) AU, for case d). Specifically, at 10 AU the C/O of both the gas and ice phases is close to what we find for WASP-127 b. The individual elemental ratios, however, are either too low (ices), or too high (gases) compared to our retrieved values (the discrepancy being worst for C than O). This may point to the accretion of ices and gases as having sculpted the atmosphere of WASP-127 b. Subsequent sublimation of the ices in the envelope would result in chemical ratios that are coherent with our results. However, such a situation would not readily explain the overabundance of CO<sub>2</sub> compared to CO, considering the CO<sub>2</sub> rich ices would be offset by CO rich gas.

The CO<sub>2</sub>/CO tension could possibly be explained by post-formation evolution. Fleury et al. (2019) showed that irradiation of a H<sub>2</sub>-CO gas mixture with Ly $\alpha$  promotes very efficient photo-chemistry, leading to CO depletion while enhancing production of CO<sub>2</sub> or H<sub>2</sub>O, depending on the temperature. At temperatures below  $\sim 1200$  K, CO<sub>2</sub> is the main product. At higher temperatures, H<sub>2</sub>O becomes the dominant product, to the detriment of CO<sub>2</sub>. Their experiments also show that the production of CO<sub>2</sub> never reaches a plateau, meaning that given the right conditions this process could be kept on for a long time. From our results, WASP-127 b’s atmosphere is exactly at the right temperature to promote CO<sub>2</sub> production and CO destruction. It is unclear whether photo-chemistry could deplete the upper atmosphere of WASP-127 b of CO and bring the results of the scenario (d) of Cevallos Soto et al. (2022) in line with what we observe. Again, observations of the day side of WASP-127 b would be useful to verify the possibility of such a scenario.

### 3.8. Concluding remarks

This work presented the first high-resolution near-infrared transmission spectrum of WASP-127 b. We combined our high-resolution SPIRou data with existing HST and *Spitzer* data from Spake et al. (2021) to perform a joint high and low-resolution MCMC retrieval and place improved constraints on the atmospheric parameters of WASP-127 b.

Using the SPIRou data only, we clearly recovered the H<sub>2</sub>O detection at a level of  $\sim 5\sigma$  and found a tentative signal of OH ( $2.8\sigma$ ). The presence of OH would be surprising as it is not expected to exist at the temperature of the planet’s terminator probed by our observations. If the signal is real, OH must come from other parts of the atmosphere and be brought to the terminator, or it could arise from photo-chemistry. No other chemical element was retrieved with statistical significance when using only the SPIRou data. The detected signal is blue-shifted, which may indicate that it is dominated by the evening limb of the terminator, and could imply nightside

condensation of water. Our injection-recovery tests showed that we are nonetheless sufficiently sensitive to most of the tested elements to put useful upper limits on their abundance in the atmosphere of WASP-127 b. We performed a series of retrievals, at low- and/or high-resolution, to better compare the different results, and the limitations of each retrieval type. As expected, the parameters are better constrained with the joint retrieval, combining both low- and high-resolution observations, and our main results are based on that.

One important result of this work is the non-detection of CO, with an upper limit of  $\log_{10}\text{CO} < -4.3$ , which excludes the chemical equilibrium scenario from S21. This indicates that the strong signal from the  $4.5\ \mu\text{m}$  IRAC2 band mainly comes from  $\text{CO}_2$  absorption, the abundance of which we constrain to be  $\log_{10}\text{CO}_2 = -4.6 \pm 0.4$ . The presence of  $\text{CO}_2$  and lack of CO likely mean that non-equilibrium processes are at play in WASP-127 b’s atmosphere, and how this imbalance is maintained remains to be explained.

Our retrieved abundances yield severely sub-solar C/O ( $C/O = 0.04^{+0.03}_{-0.03}$ ), resulting from a sub-stellar C/H, but stellar O/H. Such disparity between the elemental abundances is unexpected when considering the more standard formation scenarios, assuming that all the major C-bearing elements are in gas phase rather than hidden in a condensed form. However, a late/slow formation scenario at around 10 AU within a chemically evolving disk, where the bulk of the accretion of the gas occurred between  $\sim 5$  and 7 Myrs, offers an interesting formation avenue. In such a scenario, photo-chemistry on  $\text{CO}_2$  grains in the disk causes CO gas depletion that can explain the discrepancy between our CO and  $\text{CO}_2$  abundances, but need to have survived until now. The expected composition in such models is consistent with our observations. This scenario could benefit from further accretion of O-rich material, such as  $\text{H}_2\text{O}$  or  $\text{CO}_2$  ices (from crossing their ice lines), that would be vaporised in the atmosphere to increase the O/H and lower the C/O. Similarly, when considering the effects of radial pebble drift, gas advection, mass accretion and viscosity (such as was done by Cevallos Soto et al. 2022), we find that a core accretion scenario, with accretion of both gas and ices beyond the  $\text{H}_2\text{O}$  and  $\text{CO}_2$  ice lines ( $\sim 10$  AU), with envelope-core mixing and further migration and enrichment through the ice lines, and subsequent CO depletion could potentially produce an atmosphere with the abundances we observed.

We showed that by doing joint retrievals, we are able to lift degeneracies that affect transmission spectroscopy at high or low-resolution, independently. Indeed, our joint retrieval results do not show much correlation between abundances,  $T_{\text{P}}$ ,  $P_{\text{cloud}}$ , and  $R_{\text{P}}$ , and favor a relatively clear atmosphere, with  $R_{\text{P}}$  being consistent with, and almost exactly the same as, the value from Seidel et al. (2020b).

We detect the planet signal at large blue shift of  $v_{\text{rad}}$  of  $-7.7\ \text{km s}^{-1}$ , and we infer a large “rotational” Gaussian broadening kernel. Strong winds, fast rotation and/or a signal dominated by the trailing limb of the planet could explain these results. A more detailed 3D global circulation

modeling of WASP-127 b’s atmosphere would be highly beneficial to explain these large blue shift and broadening, and better represent the overall dynamics at play.

Given that many things remain unclear about this fascinating planet, WASP-127 b is most certainly a prime target for future observations, from the ground at high-resolution and at lower resolution with JWST. More high-resolution observations would be useful, for instance, to confirm or rule out our tentative OH detection. Also, observations of the day side, through emission spectroscopy, would be particularly useful to constrain the T-P profile and chemistry of the hot hemisphere, as well as inform global circulation models. A more in-depth study of the stellar parameters is also needed to better anchor and compare to the current observations. The fact that the star is leaving the main-sequence could be a factor in explaining what we see here.

#### *Acknowledgements*

We thank the referee Paolo Giacobbe for the careful reading, the feedback, and suggestions that really helped to improve the quality of this manuscript. The authors acknowledge financial support for this research from the Natural Science and Engineering Research Council of Canada (NSERC), the Institute for Research on Exoplanets (iREx) and the University of Montreal. These results are based on observations obtained at the Canada-France-Hawaii Telescope (CFHT) which is operated from the summit of Mauna-kea by the National Research Council of Canada, the Institut National des Sciences de l’Univers of the Centre National de la Recherche Scientifique of France, and the University of Hawaii. The observations at the Canada-France-Hawaii Telescope were performed with care and respect from the summit of Mauna-kea which is a significant cultural and historic site. SP and ADB acknowledge funding from the Technologies for Exo-Planetary Science (TEPS) CREATE program. FD thanks the CNRS/INSU Programme National de Planétologie (PNP) and Programme National de Physique Stellaire (PNPS) for funding support. XD acknowledges funding from the ANR of France under contract number ANR18CE310019 (SPlaSH) and in the framework of the Investissements d’Avenir program (ANR-15-IDEX-02), through the funding of the “Origin of Life” project of the Grenoble-Alpes University. JFD acknowledges funding from the European Research Council (ERC) under the H2020 research & innovation programme (grant agreement #740651 NewWorlds). EG acknowledges support from NASA awards 80NSSC20K0957 and 80NSSC20K0251 (Exoplanets Research Program) TF is supported by the ANR in the framework of the Investissements d’Avenir program (ANR-15-IDEX-02), through the funding of the “Origin of Life” project of the Grenoble-Alpes University J.H.C.M. is supported in the form of a work contract funded by Fundação para a Ciência e Tecnologia (FCT) with the reference DL 57/2016/CP1364/CT0007; and also supported from FCT through national funds and by FEDER-Fundo Europeu de Desenvolvimento Regional through COMPETE2020-Programa Operacional Competitividade e Internacionalização for these grants UIDB/04434/2020 & UIDP/04434/2020, PTDC/FIS-AST/32113/2017 & POCI-01-0145-FEDER-032113, PTDC/FIS-AST/28953/2017 & POCI-01-0145-FEDER-028953, PTDC/FIS-AST/29942/2017.

J.H.C.M. is also supported by supported by FCT - Fundação para a Ciência e a Tecnologia through national funds and by FEDER through COMPETE2020 - Programa Operacional Competitividade e Internacionalização by these grants : UID/FIS/04434/2019; UIDB/04434/2020; UIDP/04434/2020; PTDC/FIS-AST/32113/2017 & POCI-01-0145-FEDER-032113; PTDC/FIS-AST/28953/2017 & POCI-01-0145-FEDER-028953; PTDC/FIS-AST/28987/2017 & POCI-01-0145-FEDER-028987.

We appreciate all the work done by the developers of the NumPy, SciPy, Matplotlib, and Astropy packages, which represent truly amazing tools and are at the base of this work (Jones et al., 2001; Hunter, 2007; Astropy Collaboration et al., 2013). We are also grateful to the developers of petitRADTRANS for making their code publicly available, which allowed to easily and quickly produce atmospheric models independently for our retrievals.

### 3.9. Appendix : Injection-Recovery tests

To assess the detection significance level of various molecules in our data, we performed an injection-recovery test on the same molecules included in the high-resolution detection analysis from Section 3.6.

We produced synthetic transmission spectra including all the above molecules at their chemical equilibrium abundances, along with hydrogen and helium. We computed mean chemical equilibrium abundances of each element with `FastChem` (Stock et al., 2018), using Spake et al.’s  $Z_{\text{trace}}$ , O, C and Na retrieved abundance and T-P profile (that we found equivalent to setting  $T_{\text{eq}} = 1150$  K using our Guillot profile with fixed  $\kappa_{\text{IR}}$ ,  $\gamma$  and  $T_{\text{int}}$ )<sup>16</sup>. Next, to test the detectability of a specific molecule, only its VMR was varied, with values ranging from  $10^{-9}$  to  $10^{-2}$ , while all the other molecules VMRs were held fixed at their equilibrium VMR value. For all models, we tested two cases : one with a clear atmosphere ( $P_{\text{cloud}} = 10^2$  bar, and no haze), and one with the scattering conditions from S21 (i.e., no grey clouds, but with  $\gamma = -1.50$  and  $S = e^{2.68} \simeq 14.6$ , the scattering parameters from equation 3.5.1); no rotational broadening was applied<sup>17</sup>.

We injected the models at  $K_{\text{P}}$  and a velocity of  $v_{\text{rad}} = 35.0 \text{ km s}^{-1}$ , away from the real signal to avoid contamination, in the three transits. For each molecule and VMR tested, we computed the  $\ln \mathcal{L}$  first with the model that was injected, at  $K_{\text{P}}$  and  $v_{\text{rad}}$ , and second with a model including the same abundances except for the tested molecule, which was set to zero. This way, we can see how the data is able to differentiate between models with and without the tested molecule, and at what level it is preferred. The  $\Delta \text{BIC}$  was then obtained by comparing these two values. From the  $\Delta \text{BIC}$  values at all VMRs tested, we used linear interpolation to determine the VMR value that would yield a  $\Delta \text{BIC}$  of 10, which was taken as our detection limit. The results are shown in Figure 3.11. Our limits are compared with the retrieved abundances from the two scenarios in S21 (chemical equilibrium and the free chemistry). We also added the retrieved abundances of the detected molecules from Skaf et al. (2020) for comparison. Although they are not shown here, the  $t$ -test metrics yielded similar detection limits, while the CCF yielded slightly worse limits.

The addition of the scattering hazes worsens the detection limits (as expected), sometimes by more than 2 orders of magnitude (the worst case being for TiO, where the bulk of its signal comes from the blue side of the SPIRou wavelength coverage, most affected by the hazes). In any case, the detectable CO abundance is lower than the expected chemical equilibrium scenario, meaning that we should be able to confirm or rule it out using our data. We should also be able to easily detect H<sub>2</sub>O, and could detect equilibrium abundances of CO<sub>2</sub> and TiO in a clear atmosphere. For the free chemistry values of S21, however, we can only expect to detect H<sub>2</sub>O and FeH (or any other molecules with abundances above these limits) : the CO<sub>2</sub> abundance is just at the limit of detection, meaning that it could be hard to detect it with high significance when working only with our SPIRou data.

---

16. Because of how we computed the chemical equilibrium abundances and their respective  $1 \sigma$  error bars (explained above), there might be slight discrepancies between S21’s abundances and ours, but the relative abundances stays the same, and it should not affect our conclusions.

17. We tested the impact of the rotational broadening on these limits, and found that the abundances were affected by less than an order of magnitude, meaning that our conclusions, especially for CO, still stand.

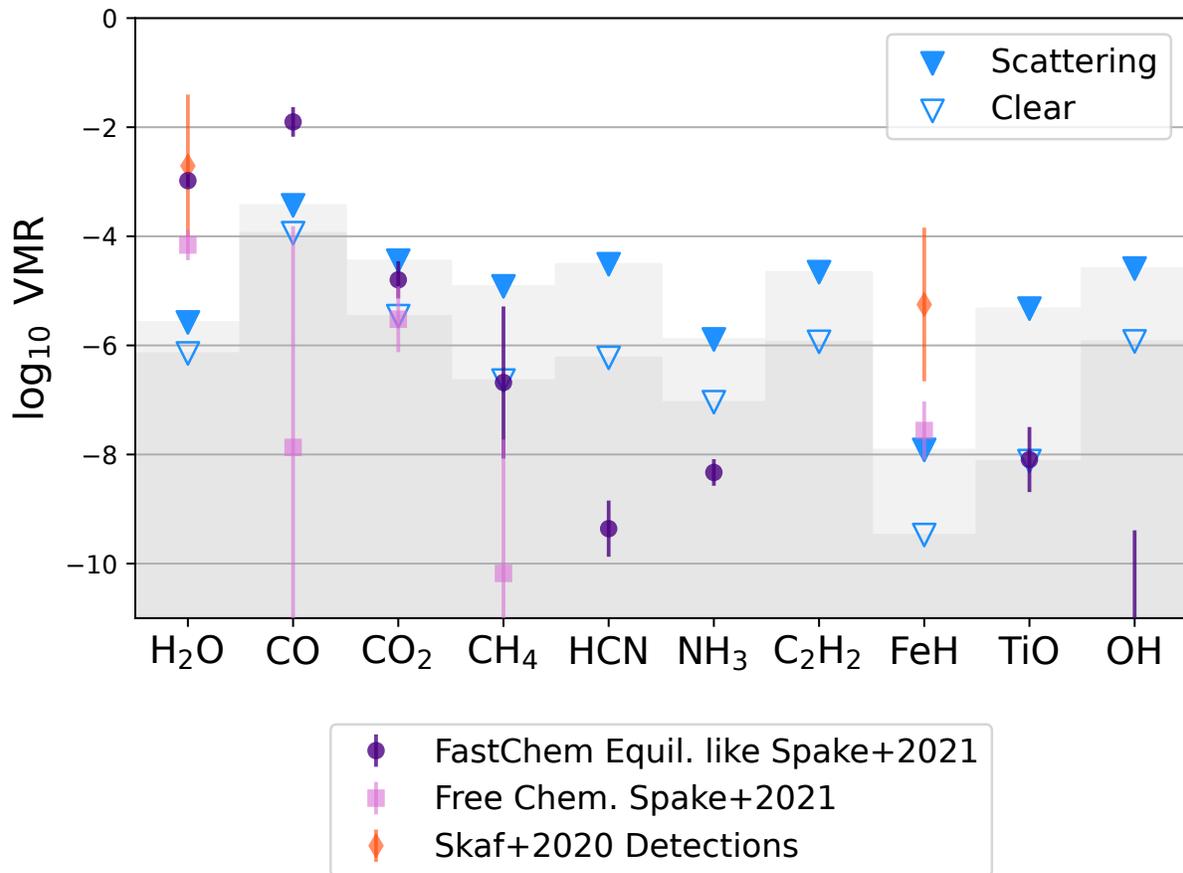


FIGURE 3.11 Detection limits of the abundances of various molecules in modeled atmospheres with and without hazes (filled and empty triangles), for the combined SPIRou transits and using the  $\Delta\text{BIC}$  significance metric. The abundances in the shaded grey regions are not expected to be detected with our SPIRou data. The limits are compared to the results from S21 : for the chemical equilibrium (dark purple circles) and free chemistry (light pink squares) scenarios ; as well as the abundances of detected molecules in Skaf et al. 2020 (orange diamonds). When the triangles are below the purple/pink/orange symbols, we can expect to detect the specific molecules, if it indeed follows the predictions from these scenarios. With the SPIRou data, we are thus very sensitive to H<sub>2</sub>O and FeH, we should be able to detect CO in the equilibrium chemistry case, and are at the limit of detection for CO<sub>2</sub>.



# Chapitre 4

---

## Conclusion

L'étude des exoplanètes n'a cessé de grimper en popularité et de s'approfondir depuis les trente dernières d'années. Cependant, encore beaucoup de concepts sont mal compris, particulièrement en lien avec la formation, l'évolution et la migration des planètes. C'est pourquoi les études détaillées de ces objets sont importantes. Il a été montré que la composition chimique permettait justement de retracer cet historique de formation, d'évolution et de migration des planètes. Les paramètres tels que le profil de température et la dynamique, permettent également d'en savoir plus sur l'état de l'atmosphère et des processus physiques et chimiques en courts. C'est ce qui a motivé l'étude présentée dans cette thèse, dont l'objectif était de caractériser l'atmosphère de Jupiter et sous-Saturne chaudes. Ces travaux de recherche ont été faits par spectroscopie de transit à partir des données dans l'infrarouge proche et à haute résolution de l'instrument SPIROU.

L'analyse de la Jupiter chaude HD 189733 b est une des premières analyses de caractérisation atmosphérique de SPIROU et la première du *SPIROU Legacy Survey*, validant ainsi sa capacité à être utilisé pour faire de la spectroscopie de transit. Nous avons décelé la présence d'eau et inféré une abondance légèrement sous-solaire. Une méthode de récupération par inférence bayésienne a permis de poser des contraintes sur les paramètres atmosphériques, tels que la température, le niveau de nuage, et le décalage en vitesse du signal. Même si cette méthode a été appliquée sur une grille de modèles interpolés, cela représente en fait la première application de cette technique sur des spectres de transmission à haute résolution. Les résultats obtenus (relativement faible abondance d'eau et une atmosphère claire de nuages) sont cohérents avec ceux de la littérature. Une comparaison des métriques de détections a permis de montrer les avantages et désavantages de chacune, confirmant que la log-vraisemblance est mieux adaptée pour faire de la sélection de modèles. Une injection négative/inverse du meilleur modèle retrouvé avec notre méthode de récupération a supprimé le signal observé quasi parfaitement, démontrant ainsi sa bonne représentation du signal atmosphérique de la planète.

Notre étude du spectre de transmission de WASP-127 b est la première à avoir été faite à haute résolution dans l'infrarouge proche, et la première à présenter une analyse de récupération complète pour un spectre de transmission. Nos données SPIROU, dont la sensibilité a été confirmée par tests d'injection, ont permis de différencier entre les scénarios à haut ou à faible C/O présentés dans Spake et al., et ce,

avant JWST. Une abondance de  $\log \text{CO} > -4.3$  (limite à  $2\sigma$ ) est exclue, et le C/O est estimé à 0.04, ce qui fait de WASP-127 b la première atmosphère ayant un aussi bas C/O, bien que beaucoup de modèles de formation tendent à prédire de fortes populations de planètes ayant de faibles C/O. Nous avons détecté la présence de  $\text{H}_2\text{O}$ , de  $\text{CO}_2$ , de Na, ainsi que la présence potentielle de OH, et mis des contraintes sur les autres molécules contenant du C, O, et N. Leur abondance estimée est de  $\log \text{H}_2\text{O} \simeq -3.8$ ,  $\log \text{CO}_2 \simeq -4.6$ ,  $\log \text{Na} \simeq -6.9$ , et  $\log \text{OH} \simeq -3.5$ . Bien que l'abondance du  $\text{CO}_2$  dépende fortement du point Spitzer à  $4.5 \mu\text{m}$ , l'analyse avec les données SPIRou seules semble également prédire la présence de  $\text{CO}_2$  avec une abondance à la limite de détection, ce qui renforce la crédibilité de cette détection. La détection inattendue d'un signal s'apparentant à du OH impliquerait des conditions extrêmes au côté jour de WASP-127 b. La sortie de la séquence principale de son étoile, augmentant le flux reçu par et le rayon de la planète, et/ou le mélange vertical pourraient être en cause. Si sa présence est confirmée, des études plus approfondies des processus hors équilibres et de l'influence de la photochimie seront bénéfiques à entreprendre. Des modélisations d'abondances chimiques plus spécifiques, suivant différentes voies de formation, et adaptées au cas plus exotique de WASP-127 b seront nécessaires pour mieux contraindre son historique de formation et d'évolution. Nos résultats ont d'ailleurs mis en lumière l'importance de prendre en considération l'évolution de la composition dans les disques. Cet exemple montre bien à quel point il est complexe de retrouver tout l'historique de formation d'une planète seulement avec sa composition chimique.

Dans les deux études, l'importance d'appliquer la même séquence de réduction aux modèles qu'aux données a été démontrée, menant ainsi à de meilleures comparaisons et représentations subséquentes des données par les modèles. Des décalages vers le bleu importants des signaux ont été observés. Ceux-ci peuvent être expliqués par la présence de forts vents globaux allant du côté jour au côté nuit, ou encore par un signal qui est dominé par le côté soir du terminateur. D'autres études ont observé des décalages de cet ordre de grandeur, et ceux-ci semblent donc réalistes. L'observation d'un plus grand nombre de transits permettrait d'étudier séparément et de comparer les spectres en transmission du limbe soir ou matin, ce qui pourrait confirmer l'origine du décalage.

Des analyses supplémentaires de ces ensembles de données permettraient d'en extraire davantage d'informations, surtout pour HD 189733 b. Par exemple, la présence connue du CO dans l'atmosphère de HD 189733 b, mais qui nécessite un traitement préalable de l'effet Rossiter-McLaughlin, ainsi que la présence d'hélium (Allart et al., 2022, en préparation). Une méthode de récupération complète appliquée à des données conjointes de SFD et SHD (étant donné la quantité incroyable d'observations qui ont été faites de ce système) permettrait d'extraire un inventaire plus complet des éléments chimiques présents dans son atmosphère, de mesurer ses ratios d'abondance plus précisément, et d'avoir une idée beaucoup plus claire de la nature de la faible abondance d'eau, ainsi que de son historique de formation.

Des données complémentaires à faible et à haute résolution de WASP-127 b, surtout du côté jour et en particulier avec JWST, seront nécessaires pour confirmer ou infirmer nos résultats. Cela permettrait, entre autres, d'établir le profil de température du côté jour et les conditions nécessaires pour la présence du OH et/ou d'aller chercher d'autres structures spectrales du  $\text{CO}_2$  sur une couverture spectrale plus large. De plus, des études plus détaillées sur le spectre de l'étoile même seront fortement utiles, d'abord pour

avoir une meilleure estimation du flux et pour extraire de meilleurs spectres de transmission, mais aussi pour mieux comprendre l'interaction étoile-planète et comment l'évolution de l'étoile hors de la séquence principale affecte l'atmosphère de la planète.

Nous avons montré que SPIRou est un instrument efficace pour faire de la spectroscopie de transit, et qu'il permet de contraindre les abondances moléculaires, le niveau de nuages, et la dynamique via le contraste et la forme des raies du signal planétaire. Nous avons démontré la complémentarité des observations à haute résolution à celles de basse résolution. De meilleures contraintes peuvent ainsi être placées sur l'ensemble des paramètres. Nous avons montré que la log-vraisemblance est adéquate pour faire la sélection de modèle, et que les méthodes de récupération, bien que computationnellement chères, représentent la façon la plus efficace d'explorer l'espace des paramètres. La prochaine grosse étape serait d'utiliser ces méthodes de récupération couplées à des GCMs, ou au moins de commencer avec des suites de modèles 2, ou 2.5D. Cela permettra d'enlever certaines approximations et hypothèses simples, et d'avoir une image plus complète et réaliste de la nature de ces objets.

L'optimisation de la correction des telluriques et de la réduction des données brutes (par exemple, via la soustraction du bruit de fond, surtout dans l'IRP, et la minimisation du bruit en général) sera toujours la bienvenue, surtout lorsqu'il sera question d'observer les signaux plus faibles des planètes de type terrestre, plus petites et froides. Le développement de métriques qui seraient moins arbitraires dans la mesure de l'intensité des détections serait fort profitable pour la communauté.

JWST viendra réellement révolutionner la spectroscopie de transmission et d'émission en SFD, étant donné son énorme aire collectrice, sa couverture spectrale exceptionnellement large dans l'IR, et à la précision incroyable des instruments à son bord. En particulier, l'instrument NIRISS permettra de faire de la spectroscopie sans fente sur un intervalle spectral de 0.6 à 2.8  $\mu\text{m}$ , et à une résolution spectrale moyenne de 700, soit près de  $5\times$  celle de HST/WFC3, mais sur un intervalle plus large. L'équipe de l'Institut de Recherche sur les Exoplanètes, avec son expertise de cet instrument en plus des instruments à haute résolution comme SPIRou et NIRISS, saura certainement faire bon usage de toutes ces données extraordinaires.

Tout comme cela a été le cas pour HST, on ne peut même pas s'imaginer toutes les merveilles qu'on va découvrir avec JWST et les futurs télescope géants, et quelles méthodes ingénieuses seront développées pour étudier ces mondes lointains et fascinants.



## Bibliographie

---

- Alessi, M., & Pudritz, R. E. 2018, *MNRAS*, 478, 2599
- Ali-Dib, M. 2017, *MNRAS*, 467, 2845
- Allart, R., Lovis, C., Pino, L., et al. 2017, *A&A*, 606, A144
- Allart, R., Bourrier, V., Lovis, C., et al. 2018, *Science*, 362, 1384
- . 2019, *A&A*, 623, A58
- Allart, R., Pino, L., Lovis, C., et al. 2020, *A&A*, 644, A155
- ALMA Partnership, Brogan, C. L., Pérez, L. M., et al. 2015, *ApJ*, 808, L3
- Alonso-Floriano, F. J., Snellen, I. A. G., Czesla, S., et al. 2019, *A&A*, 629, A110
- Alonso-Floriano, F. J., Sánchez-López, A., Snellen, I. A., et al. 2019, *Astronomy and Astrophysics*, 621, 1
- Andrews, S. M., & Birnstiel, T. 2018, in *Handbook of Exoplanets*, ed. H. J. Deeg & J. A. Belmonte, 136
- Andrews, S. M., Wilner, D. J., Zhu, Z., et al. 2016, *ApJ*, 820, L40
- Arcangeli, J., Désert, J.-M., Line, M. R., et al. 2018, *The Astrophysical Journal Letters*, 855, L30
- Armitage, P. J. 2007, arXiv e-prints, astro
- Artigau, É., Saint-Antoine, J., Lévesque, P.-L., et al. 2018, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 10709, High Energy, Optical, and Infrared Detectors for Astronomy VIII, ed. A. D. Holland & J. Beletic, 107091P
- Artigau, É., Kouach, D., Donati, J.-F., et al. 2014, in *Proc. SPIE*, Vol. 9147, Ground-based and Airborne Instrumentation for Astronomy V, 914715
- Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, *Annual Review of Astronomy and Astrophysics*, 47, 481, publisher : Annual Reviews
- Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, *A&A*, 558, A33
- Atreya, S. K., Crida, A., Guillot, T., et al. 2018, in *Saturn in the 21st Century*, *Cambridge Planetary Science* (Cambridge : Cambridge University Press), 5–43
- Baluev, R. V., Sokov, E. N., Jones, H. R. A., et al. 2019, *MNRAS*, 490, 1294
- Barber, R. J., Strange, J. K., Hill, C., et al. 2014, *Monthly Notices of the Royal Astronomical Society*, 437, 1828
- Barman, T. S., Konopacky, Q. M., Macintosh, B., & Marois, C. 2015, *ApJ*, 804, 61
- Barstow, J. K. 2020, *MNRAS*, 497, 4183
- Barstow, J. K., Aigrain, S., Irwin, P. G. J., & Sing, D. K. 2017, *ApJ*, 834, 50

Basri, G., & Brown, M. E. 2006, *Annual Review of Earth and Planetary Sciences*, 34, 193

Batygin, K., Stevenson, D. J., & Bodenheimer, P. H. 2011, *ApJ*, 738, 1

Baxter, C., Désert, J.-M., Parmentier, V., et al. 2020, *Astronomy & Astrophysics*, 639, A36, arXiv : 2007.15287

Beatty, T. G., Stevens, D. J., Collins, K. A., et al. 2017, *The Astronomical Journal*, 154, 25

Beltz, H., Rauscher, E., Brogi, M., & Kempton, E. M. R. 2021, *AJ*, 161, 1

Benneke, B. 2015, ArXiv e-prints, arXiv :1504.07655

Benneke, B., & Seager, S. 2012, *ApJ*, 753, 100

—. 2013, *ApJ*, 778, 153

Benneke, B., Knutson, H. A., Lothringer, J., et al. 2019, *Nature Astronomy*, 3, 813

Bernath, P. F. 2020, *J. Quant. Spec. Radiat. Transf.*, 240, 106687

Bertaux, J. L., Lallement, R., Ferron, S., Boonne, C., & Bodichon, R. 2014, *A&A*, 564, A46

Birkby, J. L., de Kok, R. J., Brogi, M., et al. 2013, *MNRAS*, 436, L35

Birkby, J. L., de Kok, R. J., Brogi, M., Schwarz, H., & Snellen, I. A. G. 2017, *The Astronomical Journal*, 153, 138

Bodenheimer, P., Grossman, A. S., Decamp, W. M., Marcy, G., & Pollack, J. B. 1980, *Icarus*, 41, 293

Bodenheimer, P., Hubickyj, O., & Lissauer, J. J. 2000, *Icarus*, 143, 2

Bodenheimer, P., Laughlin, G., & Lin, D. N. C. 2003, *ApJ*, 592, 555

Bodenheimer, P., Lin, D. N. C., & Mardling, R. A. 2001, *ApJ*, 548, 466

Bonfils, X., Gillon, M., Udry, S., et al. 2012, *A&A*, 546, A27

Booth, R. A., Clarke, C. J., Madhusudhan, N., & Ilee, J. D. 2017, *Monthly Notices of the Royal Astronomical Society*, 469, 3994, publisher : Oxford Academic

Borsa, F., Giacobbe, P., Bonomo, A. S., et al. 2022, arXiv e-prints, arXiv :2204.04948

Borysow, A. 2002, *A&A*, 390, 779

Boss, A. P. 1997, *Science*, 276, 1836

Boss, A. P. 2000, *ApJ*, 536, L101

Boss, A. P., Wetherill, G. W., & Haghighipour, N. 2002, *Icarus*, 156, 291

Boucher, A., Darveau-Bernier, A., Pelletier, S., et al. 2021, *AJ*, 162, 233

Bouchy, F., Udry, S., Mayor, M., et al. 2005, *A&A*, 444, L15

Bourrier, V., Lecavelier des Etangs, A., Ehrenreich, D., Tanaka, Y. A., & Vidotto, A. A. 2016, *A&A*, 591, A121

Bourrier, V., Lecavelier des Etangs, A., Dupuy, H., et al. 2013, *A&A*, 551, A63

Bourrier, V., Lecavelier des Etangs, A., Ehrenreich, D., et al. 2018, *A&A*, 620, A147

Bourrier, V., Wheatley, P. J., Lecavelier des Etangs, A., et al. 2020, *MNRAS*, 493, 559

Brewer, J. M., Fischer, D. A., & Madhusudhan, N. 2017, *The Astronomical Journal*, 153, 83, publisher : American Astronomical Society

Brogi, M., de Kok, R. J., Albrecht, S., et al. 2016, *ApJ*, 817, 106

Brogi, M., Giacobbe, P., Guilluy, G., et al. 2018, *A&A*, 615, A16

Brogi, M., Kok, R. J. d., Birkby, J. L., Schwarz, H., & Snellen, I. a. G. 2014, *Astronomy & Astrophysics*, 565, A124

Brogi, M., Line, M., Bean, J., Désert, J.-M., & Schwarz, H. 2017, *ApJ*, 839, L2

Brogi, M., & Line, M. R. 2019, *AJ*, 157, 114

Brogi, M., Snellen, I. A. G., de Kok, R. J., et al. 2012, *Nature*, 486, 502

Burrows, A., Hubeny, I., Budaj, J., Knutson, H. A., & Charbonneau, D. 2007, *ApJ*, 668, L171

Burrows, A., & Sharp, C. M. 1999, *The Astrophysical Journal*, 512, 843, publisher : IOP Publishing

Burrows, A., & Volobuyev, M. 2003, *ApJ*, 583, 985

Buzard, C., Finnerty, L., Piskorz, D., et al. 2020, *AJ*, 160, 1

Cabot, S. H. C., Madhusudhan, N., Hawker, G. A., & Gandhi, S. 2019, *MNRAS*, 482, 4422

Cabot, S. H. C., Bello-Arufe, A., Mendonça, J. M., et al. 2021, *AJ*, 162, 218

Cabrera, J., Csizmadia, S., Lehmann, H., et al. 2014, *ApJ*, 781, 18

Caldas, A., Leconte, J., Selsis, F., et al. 2019, *A&A*, 623, A161

Campbell, B., Walker, G. A. H., & Yang, S. 1988, *ApJ*, 331, 902

Carroll, B., & Ostlie, D. 2007, *An Introduction to Modern Astrophysics* (Pearson Addison-Wesley)

Casasayas-Barris, N., Pallé, E., Yan, F., et al. 2019, *Astronomy & Astrophysics*, 628, A9, publisher : EDP Sciences

Casasayas-Barris, N., Pallé, E., Yan, F., et al. 2020, *A&A*, 635, A206

Casasayas-Barris, N., Palle, E., Stangret, M., et al. 2021, *A&A*, 647, A26

Cevallos Soto, A., Tan, J. C., Hu, X., Hsu, C.-J., & Walsh, C. 2022, *arXiv e-prints*, arXiv :2202.02483

Charbonneau, D., Brown, T. M., Latham, D. W., & Mayor, M. 2000, *ApJ*, 529, L45

Charbonneau, D., Brown, T. M., Noyes, R. W., & Gilliland, R. L. 2002, *ApJ*, 568, 377

Chen, G., Casasayas-Barris, N., Pallé, E., et al. 2020, *Astronomy & Astrophysics*, 635, 171

Chen, G., Pallé, E., Nortmann, L., et al. 2017, *A&A*, 600, L11

Chen, G., Pallé, E., Welbanks, L., et al. 2018, *A&A*, 616, A145

Claret, A. 2000, *A&A*, 363, 1081

Collier Cameron, A., Guenther, E., Smalley, B., et al. 2010, *MNRAS*, 407, 507

Cont, D., Yan, F., Reiners, A., et al. 2021, *A&A*, 651, A33

—. 2022, *A&A*, 657, L2

Cowan, N. B., Machalek, P., Croll, B., et al. 2012, *ApJ*, 747, 82

Cridland, A. J., Dishoeck, E. F. v., Alessi, M., & Pudritz, R. E. 2020, *Astronomy & Astrophysics*, 642, A229, publisher : EDP Sciences

Cridland, A. J., van Dishoeck, E. F., Alessi, M., & Pudritz, R. E. 2019, *A&A*, 632, A63

Crossfield, I. J. M., Barman, T., Hansen, B. M. S., Tanaka, I., & Kodama, T. 2012, *The Astrophysical Journal*, 760, 140, publisher : IOP Publishing

Currie, T., Burrows, A., Madhusudhan, N., et al. 2013, *ApJ*, 776, 15

Cutri, R. M., Skrutskie, M. F., van Dyk, S., et al. 2003, *VizieR Online Data Catalog*, II/246

Damiano, M., Micela, G., & Tinetti, G. 2019, *ApJ*, 878, 153

de Kok, R., Brogi, M., Snellen, I., et al. 2013a, *Astronomy & Astrophysics*, 554, A82

de Kok, R. J., Birkby, J., Brogi, M., et al. 2013b, 150, 1

de Wit, J., & Seager, S. 2013, *Science*, 342, 1473

Deeg, H. J., & Belmonte, J. A. 2018, *Handbook of Exoplanets*, doi :10.1007/978-3-319-55333-7

Deibert, E. K., de Mooij, E. J. W., Jayawardhana, R., et al. 2021a, *AJ*, 161, 209

—. 2021b, *ApJ*, 919, L15

Deming, D., & Sheppard, K. 2017, *ApJ*, 841, L3

Désert, J.-M., Lecavelier des Etangs, A., Hébrard, G., et al. 2009, *ApJ*, 699, 478

Désert, J.-M., Sing, D., Vidal-Madjar, A., et al. 2011, *A&A*, 526, A12

Diamond-Lowe, H., Stevenson, K. B., Bean, J. L., Line, M. R., & Fortney, J. J. 2014, *ApJ*, 796, 66

Donati, J. F., Kouach, D., Moutou, C., et al. 2020, *MNRAS*, 498, 5684

dos Santos, L. A., Ehrenreich, D., Bourrier, V., et al. 2020, *A&A*, 640, A29

Ehrenreich, D., Bourrier, V., Wheatley, P. J., et al. 2015, *Nature*, 522, 459, number : 7557 Publisher : Nature Publishing Group

Ehrenreich, D., Lovis, C., Allart, R., et al. 2020, *Nature*, 580, 597

Eistrup, C., Walsh, C., & van Dishoeck, E. F. 2018, *A&A*, 613, A14

Evans, T. M., Sing, D. K., Kataria, T., et al. 2017, *Nature*, 548, 58

Fabrycky, D., & Tremaine, S. 2007, *ApJ*, 669, 1298

Falco, A., Zingales, T., Pluriel, W., & Leconte, J. 2022, *A&A*, 658, A41

Fisher, C., & Heng, K. 2018, *MNRAS*, 481, 4698

—. 2019, *ApJ*, 881, 25

Flagg, L., Johns-Krull, C. M., Nofi, L., et al. 2019, *ApJ*, 878, L37

Fleury, B., Gudipati, M. S., Henderson, B. L., & Swain, M. 2019, *ApJ*, 871, 158

Flowers, E., Brogi, M., Rauscher, E., Kempton, E. M. R., & Chiavassa, A. 2019, *AJ*, 157, 209

Follert, R., Dorn, R. J., Oliva, E., 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, 914719

Ford, E. B., & Rasio, F. A. 2008, *ApJ*, 686, 621

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, *PASP*, 125, 306

Fortney, J. J. 2005, *MNRAS*, 364, 649

Fortney, J. J., Marley, M. S., & Barnes, J. W. 2007, *ApJ*, 659, 1661

Fortney, J. J., Mordasini, C., Nettelmann, N., et al. 2013, *ApJ*, 775, 80

Föhring, D., Dhillon, V. S., Madhusudhan, N., et al. 2013, *Monthly Notices of the Royal Astronomical Society*, 435, 2268, publisher : Oxford Academic

Gandhi, S., & Jermyn, A. S. 2020, *MNRAS*, 499, 4984

Gandhi, S., Madhusudhan, N., Hawker, G., & Piette, A. 2019, *The Astronomical Journal*, 158, 228, publisher : American Astronomical Society

Gandhi, S., Brogi, M., Yurchenko, S. N., et al. 2020, *MNRAS*, 495, 224

Genest, F., Lafrenière, D., Boucher, A., et al. 2022, *The Astronomical Journal*, 163, 231

Giacobbe, P., Brogi, M., Gandhi, S., et al. 2021, *Nature*, 592, 205

- Gibson, N. P., Nugroho, S. K., Lothringer, J., Maguire, C., & Sing, D. K. 2022, MNRAS, doi :10.1093/mnras/stac091
- Gibson, N. P., Merritt, S., Nugroho, S. K., et al. 2020, MNRAS, 493, 2215
- Gillon, M., Triaud, A. H. M. J., Demory, B.-O., et al. 2017, Nature, 542, 456
- Goodman, J., & Weare, J. 2010, Communications in Applied Mathematics and Computational Science, 5, 65
- Griffith, C. A. 2014, Philosophical Transactions of the Royal Society of London Series A, 372, 20130086
- Grimm, S. L., & Heng, K. 2015, The Astrophysical Journal, 808, 182
- Grimm, S. L., Malik, M., Kitzmann, D., et al. 2021, The Astrophysical Journal Supplement Series, 253, 30, publisher : American Astronomical Society
- Guillot, T. 2010, A&A, 520, A27
- Guilluy, G., Sozzetti, A., Brogi, M., et al. 2019, A&A, 625, A107
- Guilluy, G., Andretta, V., Borsa, F., et al. 2020, A&A, 639, A49
- Harada, C. K., Kempton, E. M.-R., Rauscher, E., et al. 2021, The Astrophysical Journal, 909, 85, publisher : American Astronomical Society
- Harris, G. J., Tennyson, J., Kaminsky, B. M., Pavlenko, Y. V., & Jones, H. R. A. 2006, Monthly Notices of the Royal Astronomical Society, 367, 400
- Hartman, J. D., Bakos, G. Á., Bhatti, W., et al. 2016, The Astronomical Journal, 152, 182
- Hawker, G. A., Madhusudhan, N., Cabot, S. H. C., & Gandhi, S. 2018, ApJ, 863, L11
- Hayek, W., Sing, D., Pont, F., & Asplund, M. 2012, A&A, 539, A102
- Haynes, K., Mandell, A. M., Madhusudhan, N., Deming, D., & Knutson, H. 2015, ApJ, 806, 146
- Heng, K. 2016, ApJ, 826, L16
- Heng, K., & Kitzmann, D. 2017, MNRAS, 470, 2972
- Heng, K., & Lyons, J. R. 2016, ApJ, 817, 149
- Henry, G. W., Marcy, G. W., Butler, R. P., & Vogt, S. S. 2000, ApJ, 529, L41
- Herman, M. K., de Mooij, E. J. W., Nugroho, S. K., Gibson, N. P., & Jayawardhana, R. 2022, arXiv e-prints, arXiv :2203.11188
- Hobson, M. J., Bouchy, F., Cook, N. J., et al. 2021, A&A, 648, A48
- Hoeijmakers, H. J., Ehrenreich, D., Heng, K., et al. 2018, Nature, 560, 453
- Hoeijmakers, H. J., Ehrenreich, D., Kitzmann, D., et al. 2019, Astronomy & Astrophysics, 627, A165, arXiv : 1905.02096
- Hoeijmakers, H. J., Cabot, S. H. C., Zhao, L., et al. 2020, Astronomy & Astrophysics, 641, A120, publisher : EDP Sciences
- Horne, K. 1986, PASP, 98, 609
- Huitson, C. M., Sing, D. K., Vidal-Madjar, A., et al. 2012, MNRAS, 422, 2477
- Hunter, J. D. 2007, Computing in Science and Engineering, 9, 90
- Husser, T. O., Wende-von Berg, S., Dreizler, S., et al. 2013, A&A, 553, A6
- Ikoma, M., Nakazawa, K., & Emori, H. 2000, The Astrophysical Journal, 537, 1013, publisher : IOP Publishing

Irwin, P. G. J., Teanby, N. A., de Kok, R., et al. 2008, *J. Quant. Spec. Radiat. Transf.*, 109, 1136

Iyer, A. R., Swain, M. R., Zellem, R. T., et al. 2016, *ApJ*, 823, 109

Jensen, A. G., Redfield, S., Endl, M., et al. 2011, *ApJ*, 743, 203

Johansen, A., & Lambrechts, M. 2017, *Annual Review of Earth and Planetary Sciences*, 45, 359, *\_eprint* : <https://doi.org/10.1146/annurev-earth-063016-020226>

Johnson, J. A., Butler, R. P., Marcy, G. W., et al. 2007, *ApJ*, 670, 833

Jones, E. D., Oliphant, T. E., & Peterson, P. 2001, in *SciPy : Open Source Scientific Tools for Python* ; <http://www.scipy.org/>

Kaeufl, H.-U., Ballester, P., Biereichel, P., et al. 2004, in *Proc. SPIE*, Vol. 5492, *Ground-based Instrumentation for Astronomy*, ed. A. F. M. Moorwood & M. Iye, 1218–1227

Kasper, D., Bean, J. L., Line, M. R., et al. 2021, *ApJ*, 921, L18

Kass, R. E., & Raftery, A. E. 1995, *Journal of the American Statistical Association*, 90, 773

Keles, E., Kitzmann, D., Mallonn, M., et al. 2020, *Monthly Notices of the Royal Astronomical Society*, 498, 1023, publisher : Oxford Academic

Kempton, E. M.-R., & Bean, J. L. 2017, *ArXiv e-prints*, arXiv :1705.05847

Kesseli, A. Y., & Snellen, I. A. G. 2021, *ApJ*, 908, L17

Kesseli, A. Y., Snellen, I. A. G., Alonso-Floriano, F. J., Mollière, P., & Serindag, D. B. 2020, *AJ*, 160, 228

Khalafinejad, S., Molaverdikhani, K., Blečić, J., et al. 2021, *M. Lendl*, 15, 24

Kirk, J., Alam, M. K., López-Morales, M., & Zeng, L. 2020, *AJ*, 159, 115

Kitzmann, D., Patzer, A. B. C., von Paris, P., Godolt, M., & Rauer, H. 2011, *A&A*, 531, A62

Klein, B., Donati, J.-F., Moutou, C., et al. 2021, *MNRAS*, 502, 188

Knutson, H. A., Charbonneau, D., Allen, L. E., et al. 2007, *Nature*, 447, 183

Knutson, H. A., Madhusudhan, N., Cowan, N. B., et al. 2011, *ApJ*, 735, 27

Kozai, Y. 1962, *AJ*, 67, 591

Kratter, K. M., Murray-Clay, R. A., & Youdin, A. N. 2010, *ApJ*, 710, 1375

Kreidberg, L. 2015, *PASP*, 127, 1161

Kreidberg, L., Bean, J. L., Désert, J.-M., et al. 2014, *The Astrophysical Journal*, 793, L27

Kreidberg, L., Bean, J. L., Désert, J.-M., et al. 2014, *Nature*, 505, 69

Kreidberg, L., Line, M. R., Bean, J. L., et al. 2015, *ApJ*, 814, 66

Lam, K. W., Faedi, F., Brown, D. J., et al. 2017, *Astronomy and Astrophysics*, 599, 1

Lambrechts, M., & Johansen, A. 2012, *A&A*, 544, A32

Landman, R., Sánchez-López, A., Mollière, P., et al. 2021, *A&A*, 656, A119

Latham, D. W., Stefanik, R. P., Mazeh, T., Mayor, M., & Burki, G. 1989, *Nature*, 339, 38

Lavie, B., Mendonça, J. M., Mordasini, C., et al. 2017, *The Astronomical Journal*, 154, 91, publisher : American Astronomical Society

Leão, I. C., Pasquini, L., Ludwig, H. G., & de Medeiros, J. R. 2019, *MNRAS*, 483, 5026

Lecavelier des Etangs, A., & Lissauer, J. J. 2022, *New Astronomy Reviews*, 94, 101641

Lecavelier Des Etangs, A., Pont, F., Vidal-Madjar, A., & Sing, D. 2008a, *A&A*, 481, L83

Lecavelier Des Etangs, A., Vidal-Madjar, A., Désert, J. M., & Sing, D. 2008b, *A&A*, 485, 865

Lecavelier des Etangs, A., Vidal-Madjar, A., McConnell, J. C., & Hébrard, G. 2004, *A&A*, 418, L1

Lecavelier Des Etangs, A., Ehrenreich, D., Vidal-Madjar, A., et al. 2010, *A&A*, 514, A72

Lecavelier des Etangs, A., Bourrier, V., Wheatley, P. J., et al. 2012, *A&A*, 543, L4

Léger, A., Rouan, D., Schneider, J., et al. 2009, *A&A*, 506, 287

Libby-Roberts, J. E., Berta-Thompson, Z. K., Désert, J.-M., et al. 2020, *AJ*, 159, 57

Lidov, M. L. 1962, *Planet. Space Sci.*, 9, 719

Line, M. R., Fortney, J. J., Marley, M. S., & Sorahana, S. 2014, *ApJ*, 793, 33

Line, M. R., Knutson, H., Wolf, A. S., & Yung, Y. L. 2014, *The Astrophysical Journal*, 783, 70, publisher : IOP Publishing

Line, M. R., & Parmentier, V. 2016, *ApJ*, 820, 78

Line, M. R., Teske, J., Burningham, B., Fortney, J. J., & Marley, M. S. 2015, *ApJ*, 807, 183

Line, M. R., Marley, M. S., Liu, M. C., et al. 2017, *ApJ*, 848, 83

Line, M. R., Brogi, M., Bean, J. L., et al. 2021, *Nature*, 598, 580

Lockwood, A. C., Johnson, J. A., Bender, C. F., et al. 2014, *ApJ*, 783, L29

Lodders, K. 2004, *The Astrophysical Journal*, 611, 587, publisher : IOP Publishing

Lodders, K., & Fegley, B. 2002, *Icarus*, 155, 393

Lodders, K., & Palme, H. 2009, *Meteoritics and Planetary Science Supplement*, 72, 5154

Lopez, E. D., & Fortney, J. J. 2016, *The Astrophysical Journal*, 818, 4

Lothringer, J. D., Sing, D. K., Rustamkulov, Z., et al. 2022, *arXiv e-prints*, arXiv :2204.03639

Louden, T., & Wheatley, P. J. 2015, *Astrophysical Journal Letters*, 814, L24

MacDonald, R. J., Goyal, J. M., & Lewis, N. K. 2020, *ApJ*, 893, L43

MacDonald, R. J., & Madhusudhan, N. 2017, *MNRAS*, 469, 1979

MacDonald, R. J., & Madhusudhan, N. 2019, *Monthly Notices of the Royal Astronomical Society*, 486, 1292, publisher : Oxford Academic

Macintosh, B., Graham, J. R., Barman, T., et al. 2015, *Science*, doi :10.1126/science.aac5891

Madhusudhan, N. 2012, *ApJ*, 758, 36

Madhusudhan, N. 2019, *Annual Review of Astronomy and Astrophysics*, 57, 617

Madhusudhan, N., Agúndez, M., Moses, J. I., & Hu, Y. 2016, *Space Sci. Rev.*, 205, 285

Madhusudhan, N., Amin, M. A., & Kennedy, G. M. 2014, *The Astrophysical Journal*, 794, L12, publisher : IOP Publishing

Madhusudhan, N., Bitsch, B., Johansen, A., & Eriksson, L. 2017, *Monthly Notices of the Royal Astronomical Society*, 469, 4102

Madhusudhan, N., Knutson, H., Fortney, J. J., & Barman, T. 2014, *Protostars and Planets VI*, 739

Madhusudhan, N., Mousis, O., Johnson, T. V., & Lunine, J. I. 2011, *The Astrophysical Journal*, 743, 191, publisher : IOP Publishing

Madhusudhan, N., & Seager, S. 2009, *The Astrophysical Journal*, 707, 24

Madhusudhan, N., Harrington, J., Stevenson, K. B., et al. 2011, *Nature*, 469, 64

Mandel, K., & Agol, E. 2002, *ApJ*, 580, L171

Mansfield, M., Line, M. R., Bean, J. L., et al. 2021, *Nature Astronomy*, 5, 1224

Mansfield, M., Wisner, L., Stevenson, K. B., et al. 2022, arXiv e-prints, arXiv :2203.01463

Martoli, E., Hébrard, G., Moutou, C., et al. 2020, *A&A*, 641, L1

Matsuo, T., Shibai, H., Ootsubo, T., & Tamura, M. 2007, *ApJ*, 662, 1282

Matzner, C. D., & Levin, Y. 2005, *ApJ*, 628, 817

Mayor, M., & Queloz, D. 1995, *Nature*, 378, 355

Mayor, M., Marmier, M., Lovis, C., et al. 2011, arXiv e-prints, arXiv :1109.2497

Mazeh, T., Holczer, T., & Faigler, S. 2016, *A&A*, 589, A75

McCullough, P. R., Crouzet, N., Deming, D., & Madhusudhan, N. 2014, *ApJ*, 791, 55

McLaughlin, D. B. 1924, *ApJ*, 60, 22

McLean, I. S., Becklin, E. E., Bendiksen, O., et al. 1998, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 3354, *Infrared Astronomical Instrumentation*, ed. A. M. Fowler, 566–578

Mikal-Evans, T., Sing, D. K., Kataria, T., et al. 2020, arXiv :2005.09631 [astro-ph], arXiv : 2005.09631

Miller-Ricci, E., Seager, S., & Sasselov, D. 2009, *ApJ*, 690, 1056

Miller-Ricci Kempton, E., & Rauscher, E. 2012, *ApJ*, 751, 117

Mollière, P., Wardenier, J. P., van Boekel, R., et al. 2019, *A&A*, 627, A67

Mollière, P., Stolker, T., Lacour, S., et al. 2020, *Astronomy & Astrophysics*, doi :10.1051/0004-6361/202038325, publisher : EDP Sciences

Mordasini, C., Mollière, P., Dittkrist, K.-M., Jin, S., & Alibert, Y. 2015, *International Journal of Astrobiology*, 14, 201, publisher : Cambridge University Press

Mordasini, C., van Boekel, R., Mollière, P., Henning, T., & Benneke, B. 2016, *ApJ*, 832, 41

Moses, J. I., Madhusudhan, N., Visscher, C., & Freedman, R. S. 2013a, *The Astrophysical Journal*, 763, 25

Moses, J. I., Line, M. R., Visscher, C., et al. 2013b, *The Astrophysical Journal*, 777, 34

Moutou, C., Dalal, S., Donati, J. F., et al. 2020, *A&A*, 642, A72

Nakagawa, Y., Nakazawa, K., & Hayashi, C. 1981, *Icarus*, 45, 517

Nixon, M. C., & Madhusudhan, N. 2022, arXiv e-prints, arXiv :2201.03532

Nugroho, S. K., Gibson, N. P., de Mooij, E. J. W., et al. 2020, *MNRAS*, 496, 504

Nugroho, S. K., Kawahara, H., Masuda, K., et al. 2017, *The Astronomical Journal*, 154, 221

Nugroho, S. K., Kawahara, H., Gibson, N. P., et al. 2021, *ApJ*, 910, L9

Öberg, K. I., & Bergin, E. A. 2016, *ApJ*, 831, L19

Öberg, K. I., Murray-Clay, R., & Bergin, E. A. 2011, *ApJ*, 743, L16

Oklopčić, A., & Hirata, C. M. 2018, *ApJ*, 855, L11

Oklopčić, A. 2019, *ApJ*, 881, 133

Oliva, E., Origlia, L., Baffa, C., et al. 2006, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 6269, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, ed. I. S. McLean & M. Iye, 626919

Ormel, C. W., Dullemond, C. P., & Spaans, M. 2010, *Icarus*, 210, 507

- Owen, J. E. 2019, *Annual Review of Earth and Planetary Sciences*, 47, 67
- Owen, J. E., & Lai, D. 2018, *MNRAS*, 479, 5012
- Palle, E., Chen, G., Prieto-Arranz, J., et al. 2017a, *Astronomy & Astrophysics*, 602, L15
- . 2017b, *Astronomy & Astrophysics*, 602, L15
- Palle, E., Nortmann, L., Casasayas-Barris, N., et al. 2020, *A&A*, 638, A61
- Park, C., Jaffe, D. T., Yuk, I.-S., 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, 91471D
- Parmentier, V., Showman, A. P., & Fortney, J. J. 2021, *MNRAS*, 501, 78
- Parmentier, V., Line, M. R., Bean, J. L., et al. 2018, *Astronomy & Astrophysics*, 617, A110
- Pelletier, S., Benneke, B., Darveau-Bernier, A., et al. 2021, *AJ*, 162, 73
- Perryman, M. 2011, *The Exoplanet Handbook*
- Petigura, E. A., Howard, A. W., & Marcy, G. W. 2013, *Proceedings of the National Academy of Science*, 110, 19273
- Pinhas, A., & Madhusudhan, N. 2017, *ArXiv e-prints*, arXiv :1705.08893
- Pinhas, A., Madhusudhan, N., Gandhi, S., & MacDonald, R. 2019, *Monthly Notices of the Royal Astronomical Society*, 482, 1485
- Pinilla, P., Benisty, M., & Birnstiel, T. 2012a, *A&A*, 545, A81
- Pinilla, P., Birnstiel, T., Ricci, L., et al. 2012b, *A&A*, 538, A114
- Pino, L., Ehrenreich, D., Wyttenbach, A., et al. 2018, *A&A*, 612, A53
- Piskorz, D., Benneke, B., Crockett, N. R., et al. 2016, *ApJ*, 832, 131
- . 2017, *AJ*, 154, 78
- Piskorz, D., Buzard, C., Line, M. R., et al. 2018, *AJ*, 156, 133
- Pluriel, W., Leconte, J., Parmentier, V., et al. 2022, *A&A*, 658, A42
- Pluriel, W., Zingales, T., Leconte, J., & Parmentier, V. 2020, *A&A*, 636, A66
- Pollack, J. B., Hubickyj, O., Bodenheimer, P., et al. 1996, *Icarus*, 124, 62
- Polyansky, O. L., Kyuberis, A. A., Zobov, N. F., et al. 2018, *MNRAS*, 480, 2597
- Quirrenbach, A., Amado, P. J., Caballero, J. A., 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, 91471F
- Rackham, B. V., Apai, D., & Giampapa, M. S. 2018, *ApJ*, 853, 122
- . 2019, *AJ*, 157, 96
- Rafikov, R. R. 2004, *AJ*, 128, 1348
- . 2011, *ApJ*, 727, 86
- Rasmussen, K. C., Brogi, M., Rahman, F., et al. 2021, *arXiv e-prints*, arXiv :2108.12057
- Rauscher, E., & Kempton, E. M. R. 2014, *ApJ*, 790, 79
- Rayner, J., Bond, T., Bonnet, M., et al. 2012, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 8446, *Ground-based and Airborne Instrumentation for Astronomy IV*, ed. I. S. McLean, S. K. Ramsay, & H. Takami, 84462C

Redfield, S., Endl, M., Cochran, W. D., & Koesterke, L. 2008, *ApJ*, 673, L87

Rocchetto, M., Waldmann, I. P., Venot, O., Lagage, P. O., & Tinetti, G. 2016, *ApJ*, 833, 120

Rodler, F., Kürster, M., & Barnes, J. R. 2013, *MNRAS*, 432, 1980

Rodmann, J., Henning, T., Chandler, C. J., Mundy, L. G., & Wilner, D. J. 2006, *Astronomy & Astrophysics*, 446, 211, number : 1 Publisher : EDP Sciences

Rossiter, R. A. 1924, *ApJ*, 60, 15

Rothman, L. S., Gordon, I. E., Barber, R. J., et al. 2010, *J. Quant. Spec. Radiat. Transf.*, 111, 2139

Rothman, L. S., Gordon, I. E., Babikov, Y., et al. 2013, *J. Quant. Spec. Radiat. Transf.*, 130, 4

Saba, A., Tsiaras, A., Morvan, M., et al. 2021, arXiv e-prints, arXiv :2108.13721

Salz, M., Czesla, S., Schneider, P. C., et al. 2018, *A&A*, 620, A97

Sánchez-López, A., Alonso-Floriano, F. J., López-Puertas, M., et al. 2019, *A&A*, 630, A53

Sánchez-López, A., López-Puertas, M., Snellen, I. A. G., et al. 2020, *A&A*, 643, A24

Savel, A. B., Kempton, E. M. R., Malik, M., et al. 2022, *ApJ*, 926, 85

Schwarz, H., Brogi, M., Birkby, J., et al. 2013, European Planetary Science Congress 2013, held 8-13 September in London, UK., 8, EPSC2013

Schwarz, H., Brogi, M., de Kok, R., Birkby, J., & Snellen, I. 2015, *Astronomy & Astrophysics*, 576, A111

Schwieterman, E. W., Kiang, N. Y., Parenteau, M. N., et al. 2018, *Astrobiology*, 18, 663

Seager, S., & Deming, D. 2010, *ARA&A*, 48, 631

Seager, S., Deming, D., & Valenti, J. A. 2009, *Astrophysics and Space Science Proceedings*, 10, 123

Seager, S., Kuchner, M., Hier-Majumder, C. A., & Militzer, B. 2007, *ApJ*, 669, 1279

Seager, S., & Sasselov, D. D. 2000, *Astrophysical Journal*, 537, 916

Sedaghati, E., Boffin, H. M. J., MacDonald, R. J., et al. 2017, *Nature*, 549, 238, number : 7671 Publisher : Nature Publishing Group

Seidel, J. V., Ehrenreich, D., Pino, L., et al. 2020a, *A&A*, 633, A86

Seidel, J. V., Ehrenreich, D., Wyttenbach, A., et al. 2019, *A&A*, 623, A166

Seidel, J. V., Lendl, M., Bourrier, V., et al. 2020b, *A&A*, 643, A45

Seidel, J. V., Ehrenreich, D., Allart, R., et al. 2021, *A&A*, 653, A73

Shallue, C. J., & Vanderburg, A. 2018, *AJ*, 155, 94

Sheppard, K. B., Mandell, A. M., Tamburo, P., et al. 2017, *ApJ*, 850, L32

Showman, A. P., Fortney, J. J., Lewis, N. K., & Shabram, M. 2012, *The Astrophysical Journal*, 762, 24

Showman, A. P., Fortney, J. J., Lian, Y., et al. 2009, *ApJ*, 699, 564

Showman, A. P., & Polvani, L. M. 2011, *ApJ*, 738, 71

Sing, D. K. 2018, arXiv e-prints, arXiv :1804.07357

Sing, D. K., Vidal-Madjar, A., Désert, J. M., Lecavelier des Etangs, A., & Ballester, G. 2008a, *ApJ*, 686, 658

Sing, D. K., Vidal-Madjar, A., Lecavelier des Etangs, A., et al. 2008b, *ApJ*, 686, 667

Sing, D. K., Désert, J. M., Fortney, J. J., et al. 2011, *A&A*, 527, A73

Sing, D. K., Fortney, J. J., Nikolov, N., et al. 2016, *Nature*, 529, 59

Skaf, N., Bieger, M. F., Edwards, B., et al. 2020, *AJ*, 160, 109

Smette, A., Sana, H., Noll, S., et al. 2015, *A&A*, 576, A77

Snellen, I. A. G., Albrecht, S., de Mooij, E. J. W., & Le Poole, R. S. 2008, *A&A*, 487, 357

Snellen, I. A. G., Kok, R. J. D., Mooij, E. J. W. D., & Albrecht, S. 2010, *Nature*, 465, 1049

Soter, S. 2006, *AJ*, 132, 2513

Spake, J. J., Sing, D. K., Evans, T. M., et al. 2018, *Nature*, 557, 68

Spake, J. J., Sing, D. K., Wakeford, H. R., et al. 2021, *MNRAS*, 500, 4042

Stangret, M., Casasayas-Barris, N., Pallé, E., et al. 2020, arXiv :2003.04650

Stassun, K. G., Oelkers, R. J., Paegert, M., et al. 2019, *AJ*, 158, 138

Stevenson, K. B., Bean, J. L., Madhusudhan, N., & Harrington, J. 2014, *The Astrophysical Journal*, 791, 36, publisher : IOP Publishing

Stevenson, K. B., Désert, J.-M., Line, M. R., et al. 2014, *Science*, 346, 838

Stevenson, K. B., Line, M. R., Bean, J. L., et al. 2017, *The Astronomical Journal*, 153, 68, publisher : American Astronomical Society

Stock, J. W., Kitzmann, D., Patzer, A. B. C., & Sedlmayr, E. 2018, *MNRAS*, 479, 865

Student. 1908, *Biometrika*, 6, 1

Swain, M., Deroo, P., Tinetti, G., et al. 2013, *Icarus*, 225, 432

Sánchez-López, A., López-Puertas, M., Snellen, I. a. G., et al. 2020, *Astronomy & Astrophysics*, 643, A24, publisher : EDP Sciences

Tabernero, H. M., Zapatero Osorio, M. R., Allart, R., et al. 2021, *A&A*, 646, A158

Talens, G. J. J., Justesen, A. B., Albrecht, S., et al. 2018, *A&A*, 612, A57

Tennyson, J., Yurchenko, S. N., Al-Refaie, A. F., et al. 2016, *Journal of Molecular Spectroscopy*, 327, 73

Thorngren, D. P., & Fortney, J. J. 2018, *AJ*, 155, 214

Thorngren, D. P., Fortney, J. J., Murray-Clay, R. A., & Lopez, E. D. 2016, *The Astrophysical Journal*, 831, 64

Tinetti, G., Vidal-Madjar, A., Liang, M.-C., et al. 2007, *Nature*, 448, 169

Torres, G., Winn, J. N., & Holman, M. J. 2008, *ApJ*, 677, 1324

Tremblin, P., Amundsen, D. S., Chabrier, G., et al. 2016, *ApJ*, 817, L19

Tremblin, P., Amundsen, D. S., Mourier, P., et al. 2015, *ApJ*, 804, L17

Triaud, A. H. M. J. 2018, *Handbook of Exoplanets*, 1375

Triaud, A. H. M. J., Collier Cameron, A., Queloz, D., et al. 2010, *A&A*, 524, A25

Trilling, D. E., Lunine, J. I., & Benz, W. 2002, *A&A*, 394, 241

Tsai, S.-M., Lyons, J. R., Grosheintz, L., et al. 2017, *ApJS*, 228, 20

Tsai, S.-M., Malik, M., Kitzmann, D., et al. 2021, *ApJ*, 923, 264

Tsiaras, A., Waldmann, I. P., Zingales, T., et al. 2018, *AJ*, 155, 156

Vallisneri, M., Taylor, S. R., Simon, J., et al. 2020, *ApJ*, 893, 112

Vidal-Madjar, A., Lecavelier des Etangs, A., Désert, J.-M., et al. 2003, *Nature*, 422, 143

Vidal-Madjar, A., Désert, J. M., Lecavelier des Etangs, A., et al. 2004, *ApJ*, 604, L69

Vidal-Madjar, A., Sing, D. K., Lecavelier Des Etangs, A., et al. 2011, *A&A*, 527, A110

Žák, J., Kabáth, P., Boffin, H. M. J., Ivanov, V. D., & Skarka, M. 2019, *AJ*, 158, 120

- Wakeford, H. R., Sing, D. K., Deming, D., et al. 2017, *The Astronomical Journal*, 155, 29, publisher : American Astronomical Society
- Wallack, N. L., Knutson, H. A., & Deming, D. 2021, *AJ*, 162, 36
- Wardenier, J. P., Parmentier, V., Lee, E. K. H., Line, M. R., & Gharib-Nezhad, E. 2021, *MNRAS*, 506, 1258
- Webb, R. K., Brogi, M., Gandhi, S., et al. 2020, *MNRAS*, 494, 108
- Weidenschilling, S. J. 1980, *Icarus*, 44, 172
- Welbanks, L., & Madhusudhan, N. 2019, *The Astronomical Journal*, 157, 206, publisher : American Astronomical Society
- Welbanks, L., Madhusudhan, N., Allard, N. F., et al. 2019, *ApJ*, 887, L20
- Welbanks, L., Madhusudhan, N., Allard, N. F., et al. 2019, *The Astrophysical Journal*, 887, L20, publisher : American Astronomical Society
- Welch, B. L. 1947, *Biometrika*, 34, 28
- West, R. G., Gillen, E., Bayliss, D., et al. 2019, *MNRAS*, 486, 5094
- Windmark, F., Birnstiel, T., Güttler, C., et al. 2012, *Astronomy & Astrophysics*, 540, A73, publisher : EDP Sciences
- Wolszczan, A., & Frail, D. A. 1992, *Nature*, 355, 145
- Wytttenbach, A., Ehrenreich, D., Lovis, C., Udry, S., & Pepe, F. 2015, *A&A*, 577, A62
- Yan, F., & Henning, T. 2018, *Nature Astronomy*, 86
- Yan, F., Wytttenbach, A., Casasayas-Barris, N., et al. 2021, *A&A*, 645, A22
- Yan, F., Pallé, E., Reiners, A., et al. 2022a, arXiv e-prints, arXiv :2204.10158
- Yan, F., Reiners, A., Pallé, E., et al. 2022b, *A&A*, 659, A7
- Youdin, A., & Johansen, A. 2007, *ApJ*, 662, 613
- Yurchenko, S. N., Barber, R. J., & Tennyson, J. 2011, *Monthly Notices of the Royal Astronomical Society*, 413, 1828
- Yurchenko, S. N., Mellor, T. M., Freedman, R. S., & Tennyson, J. 2020, *Monthly Notices of the Royal Astronomical Society*, 496, 5282
- Zahnle, K., Marley, M. S., Freedman, R. S., Lodders, K., & Fortney, J. J. 2009, *ApJ*, 701, L20
- Zhang, J., Kempton, E. M.-R., & Rauscher, E. 2017, *The Astrophysical Journal*, 851, 84, publisher : American Astronomical Society
- Zhang, M., Knutson, H. A., Wang, L., Dai, F., & Barragán, O. 2022, *AJ*, 163, 67
- Zhang, M., Knutson, H. A., Kataria, T., et al. 2018, *The Astronomical Journal*, 155, 83, publisher : American Astronomical Society
- Zhang, Y., Snellen, I. A. G., Mollière, P., et al. 2020, *Astronomy & Astrophysics*, doi :10.1051/0004-6361/202038412, arXiv : 2009.05821
- Zhao, M., O'Rourke, J. G., Wright, J. T., et al. 2014, *ApJ*, 796, 115
- Zucker, S. 2003, *Monthly Notices of the Royal Astronomical Society*, 342, 1291, publisher : Oxford Academic

Öberg, K. I., & Wordsworth, R. 2019, *The Astronomical Journal*, 158, 194, publisher : American Astronomical Society

Žák, J., Kabáth, P., Boffin, H. M. J., Ivanov, V. D., & Skarka, M. 2019, *The Astronomical Journal*, 158, 120, publisher : IOP Publishing