

Université de Montréal

Caractérisation d'atmosphères d'exoplanètes à haute  
résolution à l'aide de l'instrument SPIRou et  
développement de méthodes d'extraction  
spectrophotométriques pour le télescope spatial James  
Webb.

par

**Antoine Darveau-Bernier**

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

Orientation Astrophysique

3 Octobre 2022



# Université de Montréal

Faculté des arts et des sciences

---

Cette thèse intitulée

## Caractérisation d'atmosphères d'exoplanètes à haute résolution à l'aide de l'instrument SPIRou et développement de méthodes d'extraction spectrophotométriques pour le télescope spatial James Webb.

présentée par

**Antoine Darveau-Bernier**

a été évaluée par un jury composé des personnes suivantes :

*Yashar Hezaveh*

---

(président-rapporteur)

*René Doyon*

---

(directeur de recherche)

*David Lafrenière*

---

(codirecteur)

*Björn Benneke*

---

(membre du jury)

*Jean-Michel Désert*

---

(examineur externe)

*Oliver Sonnentag*

---

(représentant du doyen de la FESP)



# Résumé

---

L'étude des exoplanètes et de leur atmosphère a connu une croissance fulgurante dans les deux dernières décennies. Les observations spectrophotométriques à partir d'observatoires spatiaux comme Hubble ont permis d'apporter certaines contraintes sur les phénomènes physiques et la composition de leur atmosphère, notamment grâce à la spectroscopie d'éclipse. Ces découvertes concernent généralement les planètes les plus favorables à cette technique, dont font partie les Jupiter chaudes. Cependant, les conclusions tirées à partir de telles observations comportent leur lot de dégénérescences, causées par leur faible résolution spectrale, leur couverture restreinte en longueurs d'onde et leur précision photométrique limitée. Ces lacunes peuvent être corrigées en partie grâce à la complémentarité des spectrographes à haute résolution basés au sol ainsi qu'à l'aide du nouveau télescope spatial James Webb (JWST).

Cette thèse présente, en premier lieu, une des premières analyses combinées d'observations spectrophotométriques prises avec l'instrument *Wide Field Camera 3* de Hubble et d'observations à haute résolution avec l'instrument SPIRou (SpectroPolarimètre InfraRouge) du télescope Canada-France-Hawaï. Cette analyse avait pour cible le côté jour de la Jupiter ultra chaude WASP-33 b, la deuxième exoplanète la plus chaude connue à ce jour. Aux températures se retrouvant dans l'atmosphère de WASP-33 b, avoisinant les 3000 K, des molécules comme l'eau ne peuvent demeurer stables. Cependant, le CO, beaucoup plus résistant à la dissociation thermique, reste observable. Les données de SPIRou ont donc permis de confirmer la détection des raies d'émission du CO, en accord avec deux précédentes études. La combinaison avec les données de Hubble a aussi mené à l'obtention d'un premier estimé de son abondance avec un rapport de mélange volumique de  $\log_{10} \text{CO} = -4.07^{+1.51}_{-0.60}$ . De plus, cette analyse a pu améliorer les contraintes sur la structure verticale en température et confirmer la présence d'une stratosphère. Des limites supérieures sur d'autres molécules comme l'eau, le TiO et le VO ont aussi pu être établies.

En second lieu, un algorithme d'extraction spectrale intitulé ATOCA (Algorithme de Traitement d'Ordres ContAminés) est présenté. Celui-ci est dédié au mode d'observation SOSS (*Single Object Slitless Spectroscopy*) de l'instrument NIRISS (*Near InfraRed Imager and Slitless Spectrograph*), la contribution canadienne à JWST. Ce mode d'observation couvre

une plage de longueurs d'onde allant de 0,6 à 2,8  $\mu\text{m}$  simultanément grâce à la présence des deux premiers ordres de diffraction sur le détecteur. La nécessité d'un nouvel algorithme provient du fait que ces deux ordres générés par le « grisme » du mode SOSS se chevauchent sur une petite portion du détecteur. En conséquence, la région de l'ordre 1 couvrant les plus grandes longueurs d'onde (1,6–2,8  $\mu\text{m}$ ) est contaminée par l'ordre 2 couvrant l'intervalle entre 0,85 et 1,4  $\mu\text{m}$ . ATOCA permet donc de décontaminer chacun des ordres en construisant d'abord un modèle linéaire de chaque pixel individuel du détecteur, en fonction du flux incident. Ce flux peut ensuite être extrait simultanément pour les deux ordres en comparant le modèle aux observations et en solutionnant le système selon un principe de moindres carrés. Ces travaux ont pu montrer qu'il est possible de décontaminer en dessous de 10 ppm pour chaque spectre individuel.

**Mots clés :** Analyse de données astronomiques — Réduction de données — Spectrophotométrie — Atmosphères d'exoplanète — Télescopes spatiaux — Jupiter chaudes — Spectroscopie à haute résolution

# Abstract

---

In the last decades, the research on exoplanets and their atmosphere has grown phenomenally. Space based observatories with spectrophotometric capabilities like Hubble allowed to put some constraints on the physical processes occurring in exoplanets' atmosphere and their chemical composition. These discoveries concern mainly the hotter and larger planets, such as Hot Jupiters, which are the most favorable for atmospheric characterization. However, due to their low spectral resolution and their limited wavelength range and photometric accuracy, the scientific conclusions based on these observations can be degenerate. Some of these degeneracies can be lifted with the use of ground-based high-resolution spectrographs or the new James Webb Space Telescope (JWST).

On the one hand, this thesis present one of the first analysis combining Hubble's spectrophotometric data and high-resolution observations obtained with SPirou (SpectroPolarimètre InfraRouge) at the Canada-France-Hawai telescope. This analysis targeted the dayside of the Ultra Hot Jupiter WASP-33 b, the second-hottest exoplanet known to date. WASP-33 b atmosphere can reach temperatures high enough ( $\sim 3000$  K) to dissociate molecules such as water. However, CO, which is much more resistant to thermal dissociation, remains observable. SPIrou's observations allowed us to confirm the presence of CO emission lines in WASP-33 b emission spectrum, in agreement with two previous studies. With the addition of published Hubble data, we were able to push further and provide the first estimate CO abundance, with a volume mixing ratio of  $\log_{10} \text{CO} = -4.07_{-0.60}^{+1.51}$ .

On the other hand, this thesis propose a new spectral extraction algorithm called ATOCA (Algorithm to Treat Order Contamination) specifically designed for the SOSS (*Single Object Slitless Spectroscopy*) mode of the NIRISS instrument (*Near InfraRed Imager and Slitless Spectrograph*), the Canadian contribution to JWST. This observing mode covers the wavelength range spanning from 0.6 to 2.8  $\mu\text{m}$  simultaneously, due to the presence of the two first diffraction orders on NIRISS detector. The need for a new algorithm arises from the fact that these orders, originating from SOSS "grism", overlap on a small portion of the detector. Consequently, the region of order 1 covering the longest wavelengths (1.6–2.8  $\mu\text{m}$ ) is contaminated by the signal from order 2 between 0.85 and 1.4  $\mu\text{m}$ . Hence, ATOCA allows to decontaminate both orders by first building a linear model of each individual pixel of

the detector, with the incident flux as an independent variable. This flux is then extracted simultaneously for the two orders by comparing the model to the detector image and by solving the system for the best least square fit. This work has shown that ATOCA can reduce the contamination level below 10 ppm for each individual spectrum.

**Keywords** : Astronomy data analysis — Data reduction — Spectrophotometry — Exoplanet atmospheres — Space telescopes — Hot Jupiters — 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> .....	25
<b>Introduction</b> .....	27
<b>Chapitre 1. Atmosphères d'exoplanètes</b> .....	29
Types d'atmosphères .....	29
1.1. Structure .....	30
1.1.1. Hauteur d'échelle .....	31
1.1.2. Température d'équilibre .....	32
1.1.3. Profil de température .....	32
1.1.4. Conduction .....	33
1.1.5. Radiation .....	34
1.1.6. Convection .....	34
1.2. Chimie de l'atmosphère .....	36
1.2.1. Constituants principaux .....	36
1.2.2. Dissociation moléculaire & ionisation .....	38
1.2.3. Inversion de température .....	39
1.2.4. Déséquilibre chimique .....	41
1.2.5. Circulation atmosphérique .....	43
1.2.6. Présence de nuages .....	45
1.2.7. Rapport C/O et métallicité .....	47
1.3. Détermination du spectre de l'atmosphère .....	50

<b>Chapitre 2. Spectroscopie appliquée aux atmosphères d'exoplanètes .....</b>	<b>53</b>
2.1. Éléments de base .....	53
2.1.1. Éclipse primaire (transit) .....	53
2.1.2. Éclipse secondaire (éclipse).....	56
2.1.2.1. Sensibilité au profil de température .....	58
2.1.3. Comparaison entre éclipses primaire et secondaire.....	60
2.2. Spectrophotométrie à partir d'observatoires spatiaux.....	61
2.2.1. Hubble et <i>Spitzer</i> .....	63
2.2.2. Télescope spatial James Webb .....	66
2.3. Spectroscopie à haute résolution .....	68
2.3.1. Méthode de vitesse radiale .....	69
2.3.2. Extraction du signal planétaire à haute résolution.....	76
2.3.3. Détermination des propriétés de l'atmosphère .....	79
2.4. Obtention de spectres.....	81
2.4.1. Fonctionnement d'un spectrographe.....	81
2.4.2. Techniques d'extraction conventionnelles .....	84
<b>Chapitre 3. Probing the atmospheres of Ultra Hot Jupiters with CO : A combined study of WASP-33b with SPIRou &amp; HST.....</b>	<b>89</b>
Contribution de l'étudiant .....	89
ABSTRACT.....	91
3.1. Introduction .....	91
3.1.1. WASP-33b.....	94
3.2. Observations and Data Reduction.....	95
3.2.1. Observations.....	95
3.2.2. Data Reduction.....	97
3.3. Modelling and extraction of the planetary signal .....	99
3.3.1. Atmosphere models .....	99
3.3.1.1. Planetary models .....	99
3.3.1.2. Stellar model .....	101
3.3.2. Comparison with High Resolution data.....	101
3.3.3. Comparison with Low Resolution data .....	102

3.4. Results .....	103
3.4.1. Retrievals.....	103
3.4.2. Cross-correlations.....	109
3.5. Discussion .....	110
3.5.1. Comparison with previous observations .....	110
3.5.2. Implications .....	110
3.5.3. Further improvements.....	112
3.6. Conclusion.....	113
<b>Chapitre 4. ATOCA : an algorithm to treat order contamination.</b>	
<b>Application to the NIRISS SOSS mode.....</b>	<b>117</b>
Contribution de l'étudiant .....	118
ABSTRACT.....	120
4.1. Introduction .....	120
4.2. The SOSS Trace Overlap Problem .....	122
4.3. Description of the Extraction Method.....	127
4.3.1. The model.....	129
4.3.2. Solving for $\mathbf{f}$ .....	130
4.3.3. Regularization.....	131
4.3.4. Other considerations.....	132
4.4. Implementation .....	133
4.5. Validation .....	136
4.5.1. Simulations.....	136
4.5.2. Results.....	137
4.6. Future improvements .....	144
4.7. Conclusion.....	144
Appendix A : Minimization of the $\chi^2$ .....	147
Appendix B : Trapezoidal integration on a grid .....	147
Appendix C : Comparison to optimal extraction method.....	148
<b>Chapitre 5. Conclusion.....</b>	<b>151</b>

**Bibliographie..... 153**

## Liste des tableaux

---

1	WASP-33 System Parameters .....	93
2	SPIRou observations of WASP-33 .....	97
3	MCMC Retrieval Parameter Priors and Results .....	109



## Table des figures

---

1	Structure de l'atmosphère de Jupiter. ....	31
2	Structure de l'atmosphère de Pluton. ....	33
3	Sections efficaces de molécules typiques. ....	35
4	Zones convectives des Jupiter chaudes. ....	36
5	Opacité du $H^-$ . ....	37
6	Piège à froid. ....	39
7	Photochimie de WASP-33 b. ....	42
8	Modèles de circulation globale. ....	44
9	Structures en 3 dimensions. ....	46
10	Rapports d'abondances dans le disque proto-planétaire. ....	47
11	Relation entre la masse et la métallicité. ....	49
12	Dégénérescence du C/O et de la métallicité. ....	50
1	Éclipses primaire et secondaire. ....	54
2	Schématisation de la géométrie de l'éclipse primaire. ....	55
3	Spectre de transmission de HD209458b. ....	57
4	Spectre d'émission de HD 209458 b. ....	58
5	Spectres d'émission de WASP-121 b, côté jour et nuit. ....	59
6	Spectres de transmission de 10 Jupiter chaudes. ....	62
7	Spectres d'éclipse de 19 Jupiter chaudes. ....	64
8	Spectre de transmission avec NIRISS/SOSS. ....	66
9	Spectre de transmission avec NIRSpec. ....	67
10	Résolutions spectrale. ....	69
11	Technique de corrélation croisée (masque binaire). ....	70
12	Technique de corrélation croisée (patron). ....	71

13	Fonction de corrélation croisée. ....	72
14	Vitesse orbitale. ....	73
15	Vitesse radiale d'une exoplanète. ....	74
16	Étapes d'extraction du signal planétaire. ....	75
17	Série temporelle de fonctions de corrélation croisée. ....	77
18	Carte de détection de CCF. ....	78
19	Sensibilité du profil des raies. ....	80
20	Effet de la dispersion sur une image. ....	82
21	Schématisation d'un réseau de diffraction. ....	83
22	Schéma du grisme utilisé pour le mode SOSS de NIRISS. ....	84
23	Decteur du mode SOSS de NIRISS. ....	85
24	Mode de balayage spatial de WFC3. ....	86
1	Airmass variation. ....	95
2	Signal-to-noise ratios variation. ....	96
3	Reduction steps. ....	98
4	Posterior distribution. ....	104
5	Spectra and profiles distributions. ....	105
6	Spectra and profiles distributions (WFC3 only). ....	106
7	Cross-correlation function time series. ....	107
8	Welch <i>t</i> -test maps. ....	108
9	Comparison of dayside CO abundances. ....	111
1	Logo de ATOCA. ....	118
2	SOSS mode testing image. ....	122
3	Contamination factors. ....	123
4	Estimated transit contamination levels. ....	124
5	Example of ATOCA inputs and outputs. ....	128
6	Trapezoidal integration on an irregular grid. ....	132
7	Decontamination steps. ....	134
8	Decontamination on a single image. ....	138

9	Stability of decontamination.....	139
10	Decontaminated transmission spectrum. ....	140
11	Transmission spectrum without decontamination. ....	141
12	Comparison between detector image and ATOCA modeling. ....	141
13	Stability of decontamination of WASP-52 time series. ....	142



## Liste des sigles et des abréviations

---

ATOCA	Algorithme de Traitement d'Ordres Contaminés ou <i>Algorithm to Treat Order Contamination</i>
APERO	<i>A Pipeline to Reduce Observations</i>
BERV	Vitesse radiale du barycentre de la Terre, de l'anglais <i>Barycentric Earth Radial Velocity</i>
CCF	Fonction de corrélation croisée, de l'anglais <i>Cross-Correlation Function</i>
CFHT	Télescope Canada-France-Hawaï, de l'anglais <i>Canada-France-Hawaii Telescope</i>
CIA	Absorption induite par collision, de l'anglais <i>Collision Induced Absorption</i>
CV3	Phase de test de JWST <i>Cryogenic Vacuum 3</i>
HDS	Spectroscopie à haute dispersion, de l'anglais <i>High Dispersion Spectroscopy</i>

HRS	Spectroscopie à haute résolution, de l'anglais <i>High Resolution Spectroscopy</i>
HST	Télescope spatial Hubble, de l'anglais <i>Hubble Space Telescope</i>
IRAC	<i>InfraRed Array Camera</i>
JWST	Télescope spatial James Webb, de l'anglais <i>James Webb Space Telescope</i>
MMR	Rapport de mélange molaire, de l'anglais <i>Molar Mixing Ratio</i>
NICMOS	<i>Near Infrared Camera and Multi-Object Spectrometer</i>
NIRCam	<i>Near InfraRed Camera</i>
NIRISS	<i>Near InfraRed Imager and Slitless Spectrograph</i>
NIRSpec	<i>Near InfraRed Spectrograph</i>
MCMC	Méthode de Monte-Carlo par chaînes de Markov, de l'anglais <i>Markov chain Monte-Carlo</i>
MIRI	<i>Mid InfraRed Imager</i>

PCA	Analyse en composantes principales, de l'anglais <i>Principal Component Analysis</i>
PI	Scientifique principal, de l'anglais <i>Principal Investigator</i>
RCB	Frontière radiative-convective, de l'anglais <i>Radiative-Convective Boundary</i>
SDSS	<i>Sloan Digital Sky Survey</i>
SNR	Rapport signal sur bruit, de l'anglais <i>signal-to-noise ratios</i>
SOSS	<i>Single Object Slitless Spectroscopy</i>
SPIRou	SpectroPolarimètre InfraRouge
STARSHIPS	<i>Spectral Transmission And Radiation Search for High resolution Planet Signal</i>
STIS	<i>Space Telescope Imaging Spectrograph</i>
VMR	Rapport de mélange volumique, de l'anglais <i>Volume Mixing Ratio</i>

WASP	<i>Wide Angle Search for Planets</i>
WFC3	<i>Wide Field Camera 3</i>
WFSS	<i>Wide Field Slitless Spectroscopy</i>

*À Benoît et ton amour de la nature,  
À Khalil et ta fascination pour l'espace,  
Vous qui êtes maintenant parmi les étoiles.*



## Remerciements

---

Tout d'abord, j'aimerais remercier mes deux co-directeurs, René et David, pour leur confiance et pour m'avoir donné la chance de travailler sur ce projet. Merci René pour ton entrain, ta motivation, ta vision d'ensemble, pour les discussions hors sujet qui rendent toujours l'environnement de travail plus agréable. Merci David pour le temps précieux que tu m'as accordé, à moi et à tous tes étudiants. Merci pour ta minutie et ton souci du détail, en particulier lorsque vient le temps de relire, ligne par ligne, des codes (pas toujours bien écrits), ou alors pour la correction de manuscrits. Je me suis trouvé très chanceux de vous avoir tous les deux comme conseillers; vous formez un duo du tonnerre!

Ensuite, je voudrais dire merci à Anne (Boucher). J'ai adoré travailler avec toi pendant toutes ces années... et ça part de loin! J'admire ta capacité à tout connaître et ton organisation. D'ailleurs, une chance que tu étais là pour ça! J'admire aussi ta capacité à foncer, même si le travail à faire semble être une montagne et qu'il ne semble pas y avoir d'issue. Tu y vas et tu te rends au bout des choses. J'espère que tu as pu te convaincre maintenant que tu as une place en recherche, si le cœur t'en dit bien évidemment.

J'aimerais aussi remercier Loïc, pour toutes nos grandes réflexions sur le mode SOSS. Nos discussions m'ont beaucoup aidé à améliorer mon projet et mieux comprendre ce que j'étais en train de faire. Ta profonde compréhension de NIRISS est une ressource très précieuse pour moi et les autres étudiants. Sans oublier le fait que c'est un vrai plaisir de t'avoir comme collègue.

Merci à mes collègues, qui contribué à rendre ces années des plus agréables. En particulier, merci à Stefan pour nos discussions tard le soir à Roger-Gaudry, pour ton aide à écrire ces fameuses demandes de temps et à y saupoudrer un peu de ta magie qui leur donne un aspect extraordinaire. Merci aussi pour le lift mémorable entre Montréal et Toronto pour aller voir la finale des Raptors... ou plutôt la conférence TEPS.

Merci à mes amis pour tout. Un merci particulier à Olivier Gingras avec qui j'ai passé la majorité de mon temps à l'école. Entre autres, celui avec qui j'ai passé des nuits blanches pour compléter des rapports de laboratoire.

Ce travail n'aurait pas été possible sans l'aide de mes proches, en particulier mes parents, qui m'ont toujours encouragé dans ma soif de connaissances. Vous avez fait en sorte que toutes les chances soient de mon côté pour que je puisse poursuivre mes projets.

Je voudrais aussi remercier ma douce-moitié, Marilyn, qui a vécu toutes les étapes du doctorat avec moi, les joyeuses comme les plus difficiles. Tu m'as permis décrocher à des moments où j'en avais besoin.

# Introduction

---

Nous entamons actuellement un véritable âge d'or pour l'étude des exoplanètes. Des télescopes spatiaux tels que Kepler et TESS ont fait décupler leur population connue, totalisant à ce jour 5090 objets confirmés selon la *NASA Exoplanet Archive*<sup>1</sup>. Ceci, combiné à la précision de spectrographes à haute résolution comme HARPS, a su révéler une diversité inattendue de propriétés (densité, température, etc.). Ces avancées ont bouleversé notre compréhension de leurs processus de formation et ont soulevé maintes questions sur leur différentes compositions et les phénomènes physiques ayant cours dans leur atmosphère et leur intérieur.

Des observatoires spatiaux comme Hubble et *Spitzer* ainsi que d'autres situés sur Terre ont permis de mieux cerner plusieurs de ces questionnements et d'y apporter certaines pistes de solution. Malgré qu'ils n'aient pas été originalement conçus dans cette optique, des techniques novatrices permettant de caractériser l'atmosphère d'exoplanètes ont pu être développées et peaufinées grâce à ces télescopes.

L'avènement du télescope spatial James Webb (JWST) promet d'entretenir cette lancée. Sa capacité à sonder le spectre infrarouge avec une précision inégalée pourra révéler la présence de molécules tel que le dioxyde de carbone, le méthane et l'oxygène moléculaire qui, auparavant, étaient presque inaccessibles.

L'amélioration des techniques d'analyse appliquées à des spectrographes infrarouges à haute résolution basés au sol est aussi une avenue très prometteuse. Malgré les difficultés liées aux fluctuations de l'atmosphère terrestre et un signal plus faible que dans l'espace, ces observations permettront de confirmer indubitablement la présence de certaines molécules. Elles apporteront aussi une sensibilité différente à la structure et aux phénomènes relatifs à l'atmosphère des exoplanètes. Ainsi, ces techniques agiront en toute complémentarité avec les données provenant de télescopes spatiaux.

Cette thèse s'insère précisément dans ce contexte. Dans un premier temps, elle présente une analyse conjointe de données prise à partir du télescope spatial Hubble et du SpectroPolarimètre InfraRouge (SPIRou) ayant pour cible l'exoplanète WASP-33 b. Ces travaux

---

1. <https://exoplanetarchive.ipac.caltech.edu/>

ont pu améliorer les contraintes sur la structure de son atmosphère et sur certaines abondances moléculaires, en prenant avantage de l'information fournie par la signature spectrale du monoxyde de carbone. La combinaison de ces deux techniques d'observation différentes représente une des premières analyses de la sorte.

Dans un deuxième temps, cet ouvrage propose un nouvel algorithme de réduction de données spectroscopiques, dédié au mode *Single Object Slitless Spectroscopy* de l'instrument NIRISS (*Near InfraRed Imager and Slitless Spectrograph*), la contribution de l'agence spatiale canadienne à JWST. Celui-ci permettra de s'assurer d'une extraction fidèle de spectres à haute sensibilité photométrique, nécessaire à la bonne interprétation des données.

Mais avant tout, les chapitres 1 et 2 offrent une synthèse des notions nécessaires afin de mieux comprendre les articles constituant le coeur de cette thèse. Le premier traite des atmosphères d'exoplanètes alors que le second présente certaines méthodes permettant de les caractériser en se basant sur la spectroscopie.

# Chapitre 1

---

## Atmosphères d'exoplanètes

### Types d'atmosphères

Seager & Deming (2010) classent les exoplanètes en 5 catégories, selon leur atmosphère. La première catégorie inclue les atmosphères riches en hydrogène et en hélium. Celles-ci présentent généralement une composition semblable à celle des étoiles, du fait de leur masse assez élevée pour capturer le gaz présent dans le disque protoplanétaire (disque résiduel à la formation d'une étoile et lieu de formation des planètes). Jupiter ainsi que Saturne font partie de cette catégorie.

La deuxième catégorie inclue les planètes possédant une atmosphère gazeuse avec de l'hydrogène (principalement sous forme de  $H_2$ ) expulsé par la planète elle-même ; c'est le processus de dégazage. Le gaz ne provient pas dans ce cas-ci de la nébuleuse protoplanétaire. La planète devrait être assez froide et massive ( $\sim 10 - 30M_{\odot}$ ) pour pouvoir conserver son hydrogène. Une différence majeure avec la catégorie précédente est l'absence d'hélium. En effet, ce gaz ne peut se retrouver en assez grande quantité dans des particules rocheuses (en l'occurrence des silicates) participant à la formation d'une telle planète (Elkins-Tanton & Seager, 2008). Il n'existe pas de cas représentatif dans le système solaire. Par contre, certaines exoplanètes faisant partie des catégories des super-Terres ou des mini-Neptunes pourraient présenter de telles atmosphères.

Le troisième type d'exoplanète se résume à celles dont l'atmosphère est dominée par des gaz plus massifs émis par l'activité géologique, comme le  $CO_2$ . Malgré que ces processus de dégazage émettent aussi de l'hydrogène, celui-ci ne peut être conservé, l'exoplanète étant trop peu massive ou ayant une température trop élevée pour éviter qu'il ne s'échappe. Malgré que l'atmosphère de la Terre soit maintenant dominée par le diazote, celle-ci entre tout de même dans cette catégorie puisqu'elle était auparavant bel et bien dominée par le  $CO_2$ . Celui-ci est maintenant dissout dans les océans ou emprisonné dans la roche sédimentaire.

Les exoplanètes similaires à la Terre, mais fortement irradiées par leur étoile forment la quatrième catégorie. L'influence majeure de l'étoile susceptible d'évaporer les éléments volatiles laisse place à une atmosphère composée de silicates, sous forme de condensats ou même de vapeur, et d'autres éléments ayant un point de fusion élevé. Il existe certains candidats pour représenter cette catégorie, comme K2-141 b (Dang et al., 2021). L'avènement de nouveaux observatoires et instruments optimisés pour la caractérisation d'atmosphères d'exoplanètes devrait permettre de confirmer l'existence de ce type d'atmosphère dans les années à venir.

Finalement, la cinquième catégorie inclut les planètes sans atmosphère. Elles sont petites et trop peu massives ou trop fortement irradiées pour pouvoir garder quelque espèce gazeuse sous leur influence gravitationnelle. La planète Mercure en est un bon exemple.

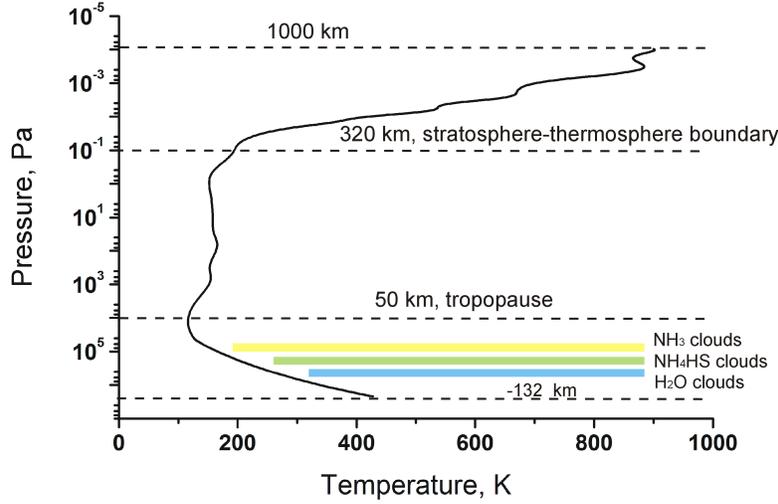
Étant donné la vaste gamme d'atmosphères d'exoplanètes, cet ouvrage se concentrera sur les atmosphères de Jupiters chaudes. Arborant une atmosphère dominée par l'hydrogène et l'hélium, elles appartiennent à la première catégorie d'atmosphères. Cependant, elles ont la particularité d'être fortement irradiées par leur étoile hôte, ce qui a pour effet de gonfler leur atmosphère. Elles ont aussi généralement des orbites très rapprochées de leur étoile et présentent donc de très courtes périodes de révolution ( $\lesssim 20$  jours).

Ces caractéristiques font des Jupiters chaudes les exoplanètes les plus favorables à être observées. Ce sont d'ailleurs des planètes de ce type qui firent l'objet de la première détection par la méthode de vitesse radiale (Mayor & Queloz, 1995) ainsi que la méthode du transit (Charbonneau et al., 2000; Henry et al., 2000). Dans les deux dernières décennies, leur étude a permis de perfectionner les techniques de caractérisation d'atmosphères. Ceci a aussi soulevé plusieurs questions quant à leur formation et la connexion qui existe entre leur atmosphère observable et leur composition.

Le chapitre qui suit traitera, en un premier temps, de la composition et de la structure des atmosphères de Jupiters chaudes. En un deuxième temps, il sera question des techniques permettant de les caractériser. En lien avec le premier article composant cette thèse, une attention particulière sera portée aux Jupiters ultra chaudes (UHJ), une sous-catégorie des Jupiters chaudes.

## 1.1. Structure

La structure des atmosphères d'exoplanètes est loin d'être uniforme. Il suffit de se fier à l'exemple que nous donne les planètes du système solaire pour se rendre compte de leur complexité (voir figure 1). Cette section a pour but de discuter de ces mécanismes et de leur influence sur l'atmosphère observable.



**Figure 1** – Relation entre la température et la pression (comparable à l’altitude) dans l’atmosphère de Jupiter. Les données proviennent de la sonde Galilée (Seiff et al., 1998).

### 1.1.1. Hauteur d’échelle

Tout d’abord, il est nécessaire de définir ce qu’est la hauteur d’échelle  $H$ . Celle-ci est dérivée à partir de l’équation d’équilibre hydrostatique,

$$\frac{dP}{dz} = -\rho g \quad (1.1.1)$$

décrivant la variation de la pression  $P$  par rapport à la hauteur  $z$  dans l’atmosphère. Celle-ci doit équilibrer l’équivalent de la force gravitationnelle donné par  $\rho g$ , où  $\rho$  est la densité et  $g$  est l’accélération gravitationnelle. En supposant un gaz parfait et une température constante, on obtient la relation suivante :

$$P(z) = P_0 \exp -z \frac{\mu g}{k_B T} = P_0 \exp \frac{-z}{H}, \quad (1.1.2)$$

où  $k_B$  est la constante de Boltzmann,  $T$  est la température et  $\mu$  est le poids moléculaire moyen. La hauteur d’échelle,

$$H \equiv \frac{k_B T}{\mu g}, \quad (1.1.3)$$

est une caractéristique locale d’un gaz puisque  $T$  n’est pas nécessairement constante partout dans l’atmosphère. Cependant, dans certaines parties, l’équation 1.1.2 est valide. On a donc une relation directe entre la pression et l’altitude. La hauteur d’échelle décrit aussi la distribution de la densité selon l’altitude dans le cas d’une atmosphère isotherme.

### 1.1.2. Température d'équilibre

La structure de l'atmosphère est en partie dictée par les conditions frontières. Celles-ci sont influencées par l'énergie provenant de l'étoile hôte (parties supérieures de l'atmosphère) et l'énergie émise par la planète. L'équilibre entre ces énergies est nécessaire (sinon la planète surchaufferait ou se refroidirait constamment), ce qui donne lieu au concept de température d'équilibre. On demande à ce que le flux absorbé ( $F_{\text{abs}}$ ) par la planète soit égal au flux émis ( $F_{\text{emi}}$ ). On a donc que

$$\underbrace{\beta F_p}_{F_{\text{emi}}} = \underbrace{(1 - A) F_{\star}}_{F_{\text{abs}}}, \quad (1.1.4)$$

où  $F_p$  est le flux de la planète,  $\beta$  est la fraction de la surface de la planète pouvant émettre du rayonnement,  $A$  est l'albédo de Bond<sup>1</sup> et  $F_{\star}$  est le flux de l'étoile reçu par la planète (pas nécessairement absorbé). Le facteur  $\beta$  est une manière de tenir compte de la possibilité de redistribuer l'énergie par la planète. En considérant les sources d'énergie internes à la planète comme négligeables (approximation justifiée pour des planètes de type Jupiter chaudes) et en utilisant l'approximation du corps noir pour déterminer  $F_p$  et  $F_{\star}$ , on arrive à l'équation suivante :

$$T_{\text{eq}} = T_{\star} \sqrt{\frac{R_{\star}}{2d}} \left( \frac{1 - A}{\beta} \right)^{1/4} \quad (1.1.5)$$

où  $T_{\star}$  et  $R_{\star}$  sont la température et le rayon de l'étoile hôte et  $d$  est la distance moyenne de la planète à son étoile.

Pour être plus fidèle à la réalité, on ne peut considérer simplement le flux émis par un corps noir et il faut utiliser des modèles existant pour celui de l'étoile. La condition d'équilibre est utilisée comme critère de cohérence lors de la modélisation du flux de la planète,  $F_{\star}$  faisant office de constante du problème. Certains modèles plus complexes dits auto-cohérents vont demander à ce que cet équilibre soit respecté à chaque couches de l'atmosphère. Le flux émis et absorbé à une couche donnée est alors influencé par les autres couches. Cette condition est appelé « équilibre radiatif ».

### 1.1.3. Profil de température

Le profil de température est influencé par les mécanismes de transport d'énergie. Plus le transport (ou la redistribution) se fait efficacement, plus la température aura tendance à s'uniformiser (selon l'altitude dans ce cas-ci). On caractérise cette efficacité par le gradient thermique,  $\frac{dT}{dz}$ .

Le profil de température de Jupiter est présenté à la figure 1 en tant qu'exemple. La pression peut y être vue intuitivement comme un indicateur de l'altitude. Les mécanismes

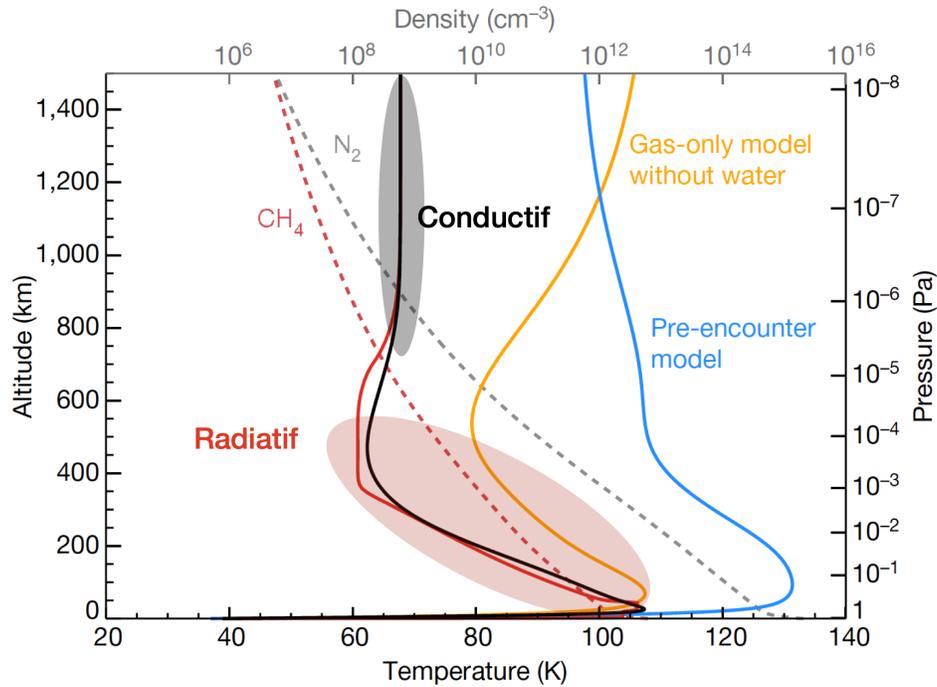
---

1. L'albédo de Bond mesure la fraction d'énergie réfléchiée par un objet, intégrée sur tout le spectre lumineux.

de transport d'énergie donnent lieu à trois différents régimes, définissant ainsi les limites de la troposphère, la stratosphère et la thermosphère (du bas vers le haut).

Il existe trois principaux mécanismes influençant la redistribution verticale de l'énergie, soit la conduction, la radiation et la convection. L'atmosphère s'adaptera selon les conditions données afin d'optimiser l'évacuation de l'énergie.

### 1.1.4. Conduction



**Figure 2** – Structure verticale de la température dans l'atmosphère de Pluton. Les données proviennent de la sonde *New Horizons*. La figure est tirée et adaptée de Zhang et al. (2017). Les courbes en tirets montrent les densités en nombre de molécules par cm<sup>3</sup>. Le profil mesuré est représenté par la courbe en noir alors que les courbes en couleur représentent différents modèles. La principale différence entre le modèle rouge et les autres est l'inclusion des particules aérosols dans le calcul d'équilibre radiatif.

La conduction a un effet négligeable pour la majorité des atmosphères de planètes (dans les régions qui nous intéressent), puisque celles-ci sont généralement très peu conductrices. Cependant, il existe certains cas où la conduction est un des mécanismes dominants, comme dans l'atmosphère de Pluton. En effet, le transport d'énergie y est principalement fait par radiation en dessous de 700 km puis par conduction. Ce changement de régime est attribué aux aérosols se formant à haute altitude et empêchant la lumière de s'échapper efficacement par radiation. Ces observations ont été faites par la sonde *New Horizons* lors de son survol (en anglais *Flyby*) de Pluton en juillet 2015. La Figure 2 adaptée de Zhang et al. (2017) montre

le profil de température qui a pu être mesuré. Cette étude a d'ailleurs montré que Pluton serait le seul objet planétaire du système solaire pour lequel la structure de l'atmosphère est dictée par les particules aérosols et non par le gaz.

### 1.1.5. Radiation

Le transport radiatif est, quant à lui, pratiquement toujours présent dans les atmosphères d'exoplanètes. L'énergie est redistribuée grâce aux photons émis qui sont réabsorbés un peu plus loin après une marche aléatoire. Le libre parcours moyen<sup>2</sup> des photons dépend de la capacité du milieu à interagir avec la matière, c'est-à-dire son opacité ainsi que de la densité du milieu. Ces quantités physiques sont mises en relation via l'équation

$$\chi_\nu \equiv k_\nu \rho, \quad (1.1.6)$$

où  $\rho$  est la densité,  $k_\nu$  est l'opacité et  $\chi_\nu$  est le coefficient d'extinction. Ce dernier est en fait l'inverse du libre parcours moyen ; la chance d'être arrêté est inversement proportionnelle à la distance qu'un photon peut franchir sans être arrêté. L'indice  $\nu$  dénote la dépendance selon la fréquence du photon. Des exemples de sections efficaces (quantité similaire à l'opacité) sont présentés à la figure 3 pour un cas typique de Jupiters chaudes. On peut voir entre autres la prédominance de l'absorption de l'eau (plus de détails à la section 1.2).

Une autre quantité à considérer est la profondeur optique, définie sous sa forme différentielle comme

$$d\tau_\nu \equiv \chi_\nu ds, \quad (1.1.7)$$

où  $ds$  est l'épaisseur que traverse un faisceau lumineux au sein d'un milieu dont l'extinction, à une fréquence  $\nu$ , est caractérisée par  $\chi_\nu$ . Ainsi, en intégrant  $ds$  sur une profondeur finie,  $\tau_\nu$  devient un indicateur de la quantité de matière absorbante. Par exemple, si une grande partie des photons peut traverser le milieu, on a  $\tau_\nu \ll 1$  et on dira que celui-ci est optiquement mince. Un milieu optiquement épais respecterait la condition  $\tau_\nu \gg 1$ .

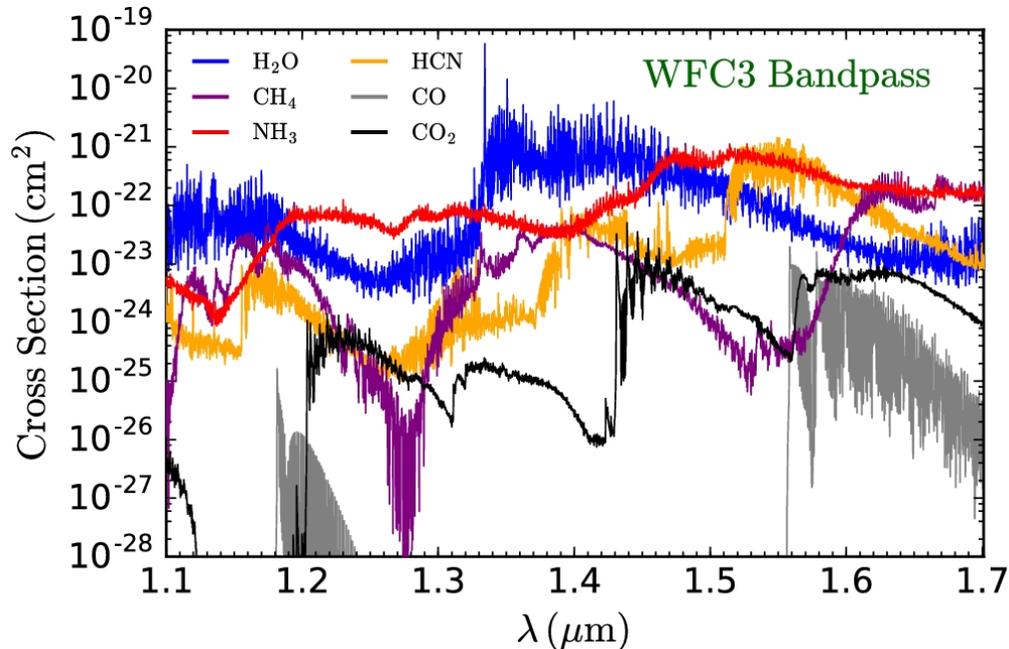
Dans le cas des couches optiquement épaisses de l'atmosphère, chacune est influencée par ses voisins ; la profondeur optique est trop grande pour communiquer plus loin. Ainsi, un gradient de température s'établit afin de respecter les conditions aux frontières via une équation de diffusion.

### 1.1.6. Convection

Dans certaines conditions (à des profondeurs optiques élevées), la radiation devient moins efficace pour retransmettre l'énergie que la convection. C'est ce qu'on appelle l'« instabilité convective ». Celle-ci est régie par le « critère de Schwarzschild » allant comme suit : si l'on considère une bulle de gaz et qu'on introduit une perturbation par rapport à sa position

---

2. Distance moyenne franchie par une particule sans interagir avec une autre.



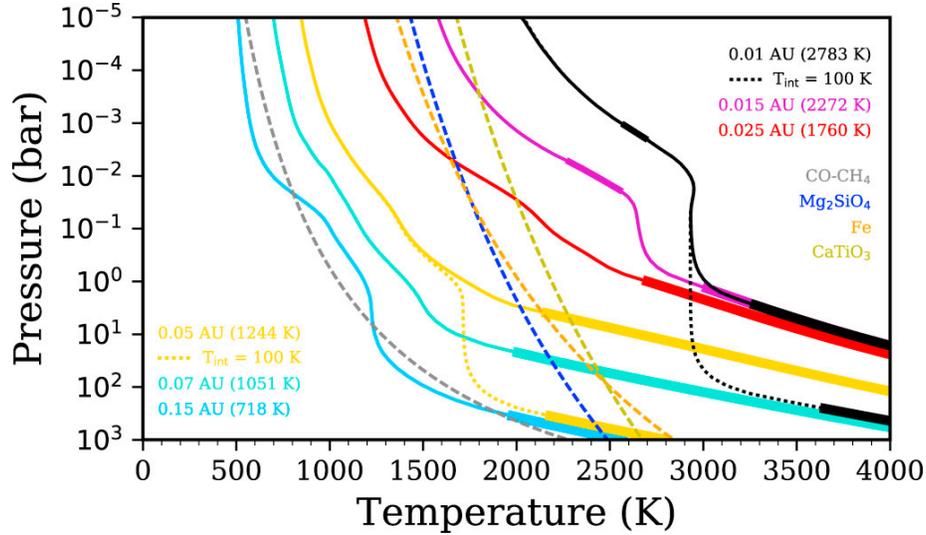
**Figure 3** – Sections efficaces (quantité similaire à l’absorption  $\kappa_\lambda$ ) de molécules détectables dans l’intervalle de longueur d’onde couvert par l’instrument WFC3 (*Wide Field Camera 3*) du télescope Hubble. Elles sont calculées à une température de 1400 K et 1 mbar, valeurs représentatives de la haute atmosphère de HD209458b, une Jupiter chaude. La figure est tirée de MacDonald & Madhusudhan (2017).

(verticale dans le cas d’une atmosphère), alors il y a instabilité lorsque celle-ci continue à dévier par rapport à la position initiale. Dans le cas adiabatique (sans échange d’énergie avec le milieu ambiant), le critère d’instabilité peut s’écrire comme suit :

$$\frac{dT}{dz} \geq \frac{-g}{c_P} \quad (1.1.8)$$

où  $g$  est la gravité et  $c_P$  est la chaleur spécifique à pression constante. Sous cette condition, la convection apparaîtra et diminuera le gradient de température. La description du phénomène est très complexe et se fait de manière phénoménologique. De plus, la convection influence la composition de chaque couche d’atmosphère, compliquant ainsi les calculs d’opacité nécessaires au transport par radiation.

Dans le cas des Jupiters chaudes, la zone convective se situe généralement à la base de leur atmosphère avec certaines exceptions aux températures plus élevées (voir figure 4). La limite entre cette zone convective (à la base de l’atmosphère) et la zone radiative (au pressions plus faibles) s’appelle la frontière radiative-convective ou RCB (de l’anglais *Radiative-Convective Boundary*). Celle-ci est influencée par la température intrinsèque de la planète ainsi que par le flux reçu de l’étoile hôte (ex : Sudarsky et al., 2003) : la première faisant monter la RCB aux pressions plus faibles et la seconde poussant la RCB plus profondément dans l’atmosphère



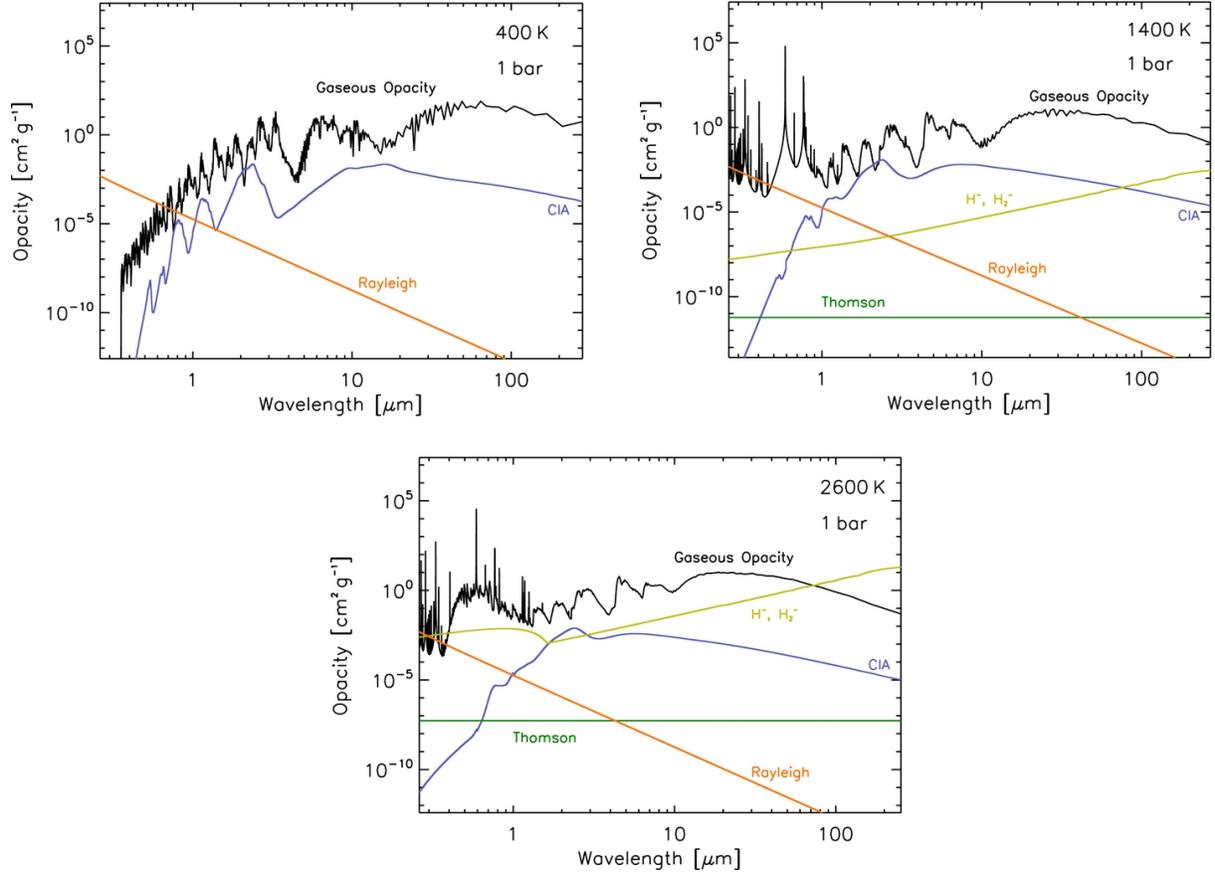
**Figure 4** – Profils de températures pour une sélection de modèles de Jupiter chaudes. La figure est tirée de Thorngren et al. (2019). Les courbes fines montrent les régions dominées par la radiation alors que les courbes plus épaisses désignent les zones convectives. Les différentes couleurs représentent des modèles supposant différentes  $T_{\text{eq}}$  et distances orbitales. Pour les atmosphères avec les températures les plus élevées (courbes mauve et noire), une seconde zone convective se forme à des pressions plus basses. Une solution alternative supposant  $T_{\text{int}} = 100$  K est montrée en traits pointillés pour le modèle à 1244 K (jaune) et 2783 K (noir). Autrement,  $T_{\text{int}}$  est déterminé empiriquement à partir d’observations sur un ensemble de Jupiter chaudes (Thorngren & Fortney, 2018). Les courbes en tirets désignent les courbes de condensation pour différentes espèces chimiques (voir section 1.2.6), à l’exception de la courbe en gris qui désigne la frontière CO-CH<sub>4</sub> tirée de Visscher (2012) (voir section 1.2).

(pressions élevées). La figure 4 montre d’ailleurs l’effet d’une température intrinsèque plus basse ( $T_{\text{int}} = 100$  K).

## 1.2. Chimie de l’atmosphère

### 1.2.1. Constituants principaux

Contrairement à Jupiter dont l’atmosphère détectable est composée principalement de méthane (CH<sub>4</sub>) et d’ammoniac (NH<sub>3</sub>, Bolton et al. 2017), les Jupiter chaudes donnent accès à d’autres molécules comme la vapeur d’eau, le CO et le CO<sub>2</sub>. Leur température d’équilibre élevée provoque des atmosphères plus ténues et permet aux molécules porteuses d’oxygène de monter dans l’atmosphère au lieu d’être confinées sous forme de nuages aux couches plus basses, comme dans le cas de Jupiter (Atreya et al., 2002). La figure 4 montre la frontière CO-CH<sub>4</sub> où le monoxyde de carbone remplace le méthane en tant que principale molécule



**Figure 5** – Contributions relatives de différentes espèces gazeuses à l’opacité. Les abondances sont calculées en supposant l’équilibre chimique (voir section 1.2.4) pour une pression de 1 bar et une température de 400 K (coin supérieur gauche), 1400 K (coin supérieur droit) et 2600 K (bas). L’opacité due à la diffusion de Thomson (vert), à la diffusion de Rayleigh (orange), au  $\text{H}^+$  combiné au  $\text{H}_2$  (jaune) et au CIA (bleu) y sont montrées de manière distinctive alors que celle de toutes les autres espèces (molécules, atomes, ions) sont incluses dans la courbe en noir. La figure est tirée de Freedman et al. (2014).

porteuse de carbone<sup>3</sup> (Lodders & Fegley, 2002; Visscher, 2012). Notons que pour des températures avoisinant cette limite, le méthane devrait être encore observable (Madhusudhan, 2012). Cependant, celui-ci n’est pas détecté dans certaines atmosphères où on s’attend à le retrouver, comme pour GJ 436 b (Stevenson et al., 2010), WASP-107 b (Kreidberg et al., 2018) et WASP-117 b (Carone et al., 2021). D’ailleurs il n’y a que très peu de détections de méthane (ex : HD 209458 b, Giacobbe et al. 2021), certaines (ex : Swain et al., 2008, 2009; Tinetti et al., 2010) ayant été remises en question (Gibson et al., 2011) et d’autres n’étant que marginales (Guilluy et al., 2019). Plusieurs théories ont tenté d’expliquer cet apparent manque de méthane ; certaines sont présentées à la section 1.2.4.

3. En supposant des abondances relatives semblables au Soleil.

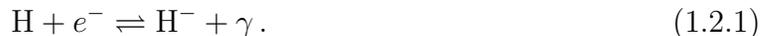
Contrairement au méthane, l'eau est une molécule très présente dans les atmosphères observables de Jupiter chaudes ainsi que certains alcalins, comme le sodium et le potassium (ex : Deming et al., 2013; Sing et al., 2016; Spake et al., 2021b).

Malgré que ce type d'exoplanète soit composé en grande majorité d'hydrogène (sous forme de  $H_2$ ) et d'hélium, ces éléments sont difficilement détectables puisqu'ils interagissent peu avec la lumière à ces températures. Leur présence est malgré tout dévoilée via des processus de collision de type  $H_2-H_2$  et  $He-H_2$  induisant une polarisation (une déclinaison du phénomène « d'absorption induite par collision » ou CIA pour *Collision Induced Absorption*).

### 1.2.2. Dissociation moléculaire & ionisation

Aux températures extrêmement élevées ( $\gtrsim 2200$  K) qu'atteignent les UHJ, l'hydrogène moléculaire commence à se dissocier en hydrogène atomique (H). La plupart des molécules seront d'ailleurs assujetties à ce phénomène appelé « dissociation moléculaire », comme l'eau, le  $CO_2$ , le TiO, le VO, etc. Ceci aura pour conséquence d'augmenter l'abondance relative d'atomes dans l'atmosphère et de diminuer l'importance des molécules dans l'atmosphère observable. Le CO est un cas particulier, car il n'est que très peu affecté par la dissociation à des températures  $\lesssim 3000$  K (Lodders & Fegley, 2002).

Parallèlement, les effets d'ionisation prendront aussi de l'ampleur, libérant ainsi des électrons ( $e^-$ ) dans l'atmosphère. Entre autres, les atomes comme le sodium et le potassium, souvent présents dans l'atmosphère de Jupiter chaudes, seront portés à s'ioniser. Les électrons ainsi libérés pourront se recombiner avec l'hydrogène atomique pour favoriser la production de l'ion  $H^-$  et d'un photon ( $\gamma$ ) via la réaction :



Inversement, comme la réaction l'indique, un photon avec une énergie suffisante (0,754 eV) pourra ioniser le  $H^-$ . Cette interaction de type lié-libre<sup>4</sup> créera une opacité continue qui s'étendra sur des longueurs d'onde plus petite que  $1,64 \mu m$  (correspondant à l'énergie minimale d'ionisation). Les interactions de type libre-libre<sup>5</sup>, aussi connues sous le nom de *bremsstrahlung*, contribueront elles aussi à l'opacité continue aux plus grandes longueurs d'onde. La figure 5 tirée de Freedman et al. (2014) montre bien la transition d'une opacité continue à l'autre. Le CIA, dominant pour les température plus basses, se fait progressivement remplacer par l'opacité du  $H^-$ . Dans l'encadré du bas, on peut aussi distinguer la contribution de l'opacité de type lié-libre du  $H^-$  par la coupure à  $1,64 \mu m$ .

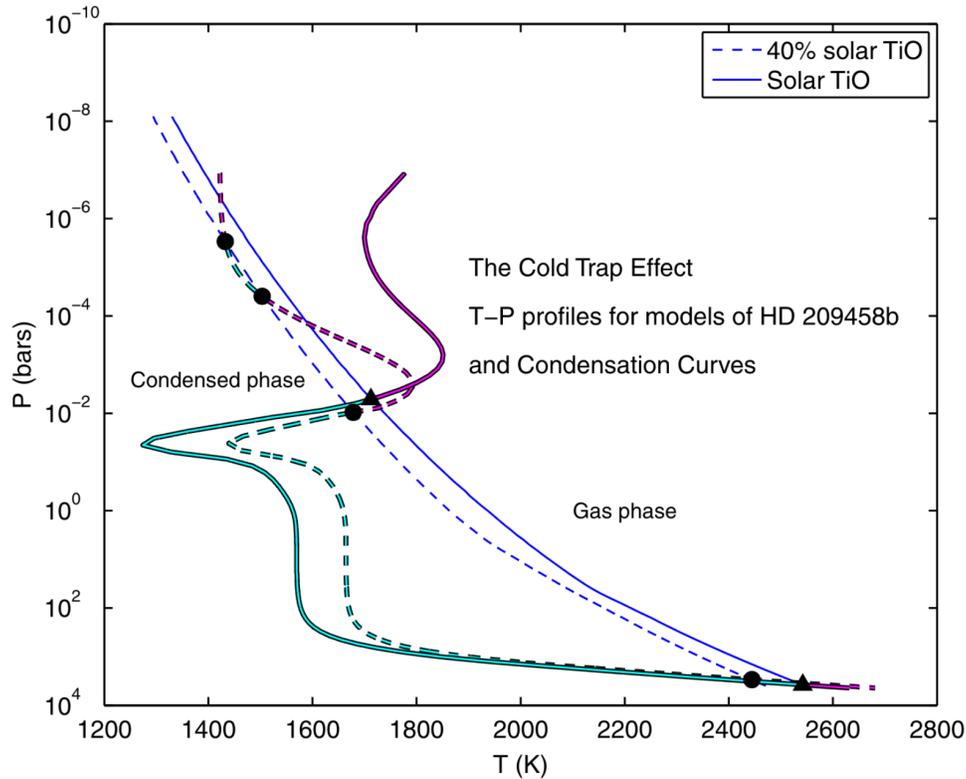
---

4. Le terme lié-libre fait référence à l'état de l'électron au cours de la réaction. Celui-ci passe d'un état lié (à un atome dans ce cas-ci) à un état libre.

5. Interaction entre un électron et un ion avoisinant, menant à la déviation de la trajectoire de l'électron et à l'absorption d'un photon.

Notons que le phénomène de dissociation moléculaire ainsi que l'ionisation sont tous deux sensibles à la pression et seront favorisés à de basses pressions. Les abondances relatives des molécules enclines à se dissocier seront donc corrélées aux profils de températures.

### 1.2.3. Inversion de température



**Figure 6** – Piège à froid du TiO (Spiegel et al., 2009). Les courbes de condensation du TiO sont présentées en bleu ; la molécule étant à l'état gazeux à droite de la courbe et solide à gauche. Les courbes passant de cyan à magenta représentent des profils de température pour des modèles de HD 209458 b. Pour toutes les courbes, deux métallicités différentes sont présentées, soit solaire (ligne continue) et 40% solaire (tirets). Les cercles et les triangles marquent les intersections entre les profils de température et les courbes de condensation.

Les inversions de température (parfois appelés stratosphères) sont un bon exemple de l'influence de la chimie sur la structure de l'atmosphère. Certains cas d'inversion sont dus à des processus photo-chimiques. Dans le cas de la Terre, par exemple, c'est le cycle ozone-oxygène qui en est responsable. Le dioxygène étant dissocié par les rayons ultra-violettes laisse place à la recombinaison rapide en ozone. Le cycle chimique impliqué produit de la chaleur et réchauffe la stratosphère.

L'inversion de température est une propriété très répandue dans d'autres atmosphères de planètes du système solaire, généralement à cause d'aérosols induits par le méthane. Or,

pour les Jupiter chaudes, le méthane est très peu présent (voir section 1.2) et ne peut donc pas occasionner d'inversion (Seager & Deming, 2010). D'autres espèces chimiques ont été proposées comme l'oxyde de titane (TiO) et l'oxyde de vanadium (VO) par Hubeny et al. (2003) et Fortney et al. (2008) (pour des températures  $\gtrsim 2000$  K) ainsi que des aérosols à base de soufre par Zahnle et al. (2009).

Cependant, les inversions de température firent l'objet de débats au sein de la communauté scientifique, en particulier concernant la Jupiter chaude HD 209458 b, une des exoplanètes les plus étudiées. Les observations qui, à l'origine, favorisaient une inversion (Burrows et al., 2007; Knutson et al., 2008) furent par la suite réfutées (Diamond-Lowe et al., 2014; Line et al., 2016). Ces contradictions étant généralement causées par les choix de modélisation et par le manque d'observations précises combiné à des méthodes de réductions de données différentes, à une sous-estimation des erreurs ou à de la variabilité astrophysique (plus de détails dans Hansen et al. 2014, Crossfield 2015 et dans la section 2).

Malgré cela, des inversions ont pu être confirmées dans la dernière décennie pour certaines UHJ, comme HAT-P-7 b (Christiansen et al., 2010; Wong et al., 2015; Mansfield et al., 2018), WASP-33 b (Haynes et al., 2015), WASP-121 b (Evans et al., 2017), WASP-18 b (Sheppard et al., 2017; Arcangeli et al., 2018) et WASP-76 b (Edwards, 2020). Pourtant, le TiO n'y a pas été détecté sans équivoque<sup>6</sup> et le VO a seulement été confirmé de manière robuste dans l'atmosphère de WASP-76 b (Pelletier et al., 2022).

Cette apparente lacune n'empêche pas d'expliquer les inversions observées dans le cas des UHJ. En effet, selon Lothringer et al. (2018), l'absorption par les métaux comme le fer et le magnésium (présents sous forme de vapeur), le SiO, différents hydrures (ex : FeH, MgH, SiH) ainsi que l'opacité du  $H^-$  sont suffisants. Ces espèces sont d'ailleurs parfois détectées à haute résolution, comme dans le cas de la UHJ KELT-9 b (Pino et al., 2020) ce qui a permis à la fois de confirmer l'inversion et de l'expliquer par le coefficient d'absorption du fer dans la partie visible du spectre. Le fait que ces planètes se trouvent généralement en orbite autour d'étoile chaudes contribue aussi aux inversions à cause de leur flux élevé en lumière bleue et UV. À cela s'ajoute la dissociation de l'eau aux basses pressions (ex : Arcangeli et al., 2018; Mansfield et al., 2021). En effet, l'eau étant un efficace agent de refroidissement, celle-ci ne peut pas refroidir les couches supérieures de l'atmosphère (à basse pression) où la molécule est dissociée, contribuant ainsi un déficit dans le refroidissement des couches supérieures. Pour la même raison, une atmosphère dans laquelle un haut rapport C/O (voir section 1.2.7) favoriserait le CO au détriment du  $H_2O$  pourrait aussi expliquer en partie une inversion de température (Mollière et al., 2015).

Néanmoins, le TiO et le VO sous forme gazeuse représentent actuellement l'hypothèse la plus répandue pour expliquer une inversion dans l'atmosphère des Jupiter chaudes (Hubeny et al., 2003; Burrows et al., 2007; Fortney et al., 2008), un peu plus froide que les UHJ. Baxter

---

6. Cas particulier pour WASP-33 b discuté à la section 3

et al. (2020) observent d’ailleurs une transition dans un échantillon de Jupiter chaudes se produisant à partir de  $T_{\text{eq}} \sim 1700$  K, ce qui correspond à la température où le TiO et le VO s’évaporent. En-dessous de cette température, le TiO et le VO pourraient subir l’effet de « piège à froid », ou *cold trap* en anglais. Ce principe est montré à la figure 6 tirée de Spiegel et al. (2009). En résumé, aux pressions plus basses, la température atteint un minimum local en dessous de la limite de condensation du TiO. Celui-ci devrait donc se retrouver à l’état solide et tomber plus profondément dans l’atmosphère sous forme de précipitations. Ainsi, malgré que les conditions aux altitudes plus élevées permettent la présence de TiO gazeux, celui-ci est « piégé » plus profondément dans l’atmosphère. Un mécanisme similaire peut aussi se produire horizontalement lorsqu’il existe des régions de température différentes sur la surface de la planète (Showman et al., 2009; Parmentier et al., 2013). D’autres phénomènes peuvent aussi causer un déficit en TiO et VO dans les hautes couches de l’atmosphère, comme la photo-dissociation causé par la radiation ultraviolette (suggéré par Knutson et al. 2010). Notons que ces phénomènes ne sont pas forcément applicables aux planètes les plus chaudes.

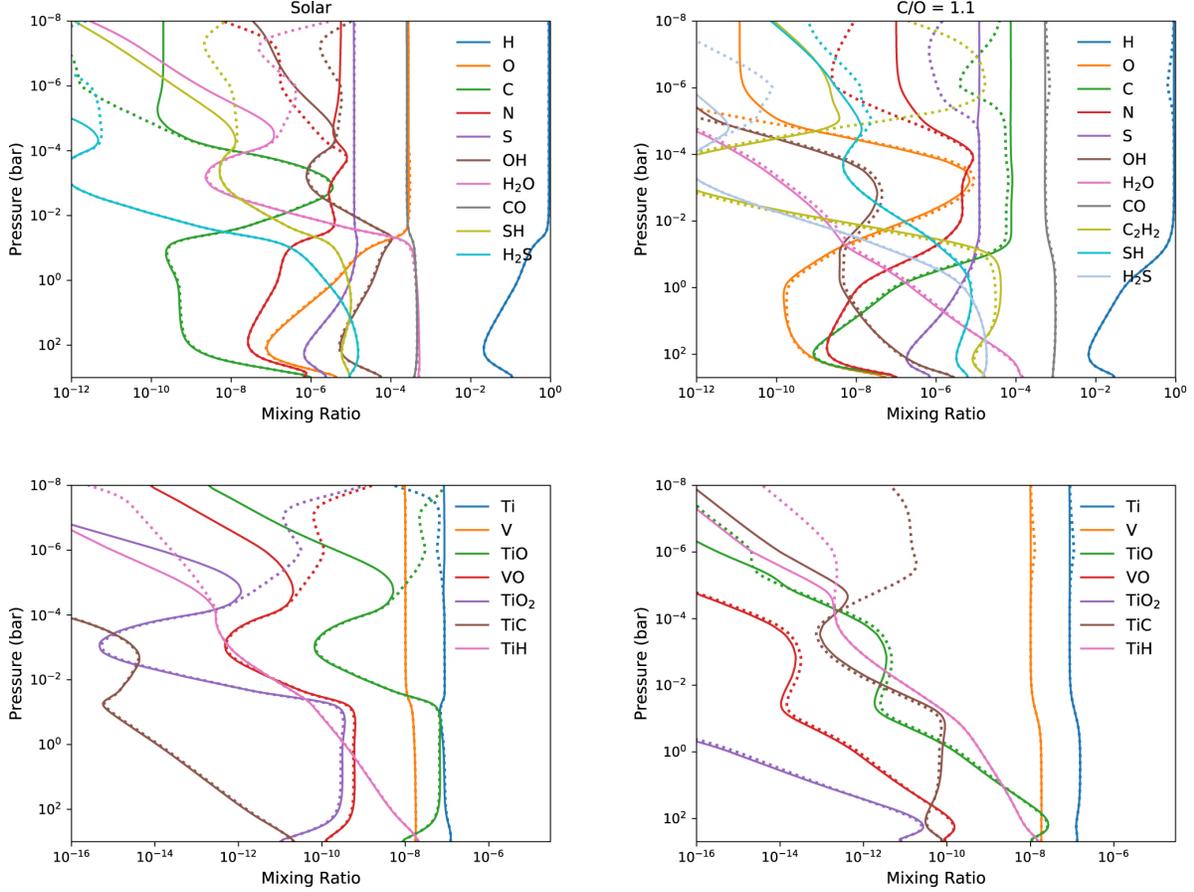
En outre, d’autres effets peuvent être invoqués pour contrer les effets de piège. Par exemple, des processus de mélange verticaux pourraient soulever les condensats aux altitudes où la température est assez élevée pour qu’ils se dissocient et libèrent le titane ou le vanadium. Ceci nous amène à considérer la notion de déséquilibre chimique.

#### 1.2.4. Déséquilibre chimique

L’essentiel de la théorie développée précédemment est basé sur la supposition que toutes les réactions chimiques considérées dans un système (par exemple dans chaque couche d’un modèle d’atmosphère) sont à l’équilibre, donc que chaque réaction est compensée par d’autres réactions, de manière à ce que les abondances relatives demeurent constantes. Cet état stationnaire, appelé équilibre chimique, peut être obtenu de manière pratique en minimisant l’énergie libre de Gibbs (ou enthalpie libre) à pression et température constante, en prenant compte des espèces chimiques susceptibles d’avoir une influence.

Cependant, il est peu probable que cet équilibre soit respecté de manière générale aux pressions sondées dans les atmosphères d’exoplanètes. En effet, plusieurs processus peuvent venir perturber cet équilibre local.

**Photochimie** – La lumière provenant de l’étoile hôte peut provoquer des réactions photochimiques dans les plus hautes couches de l’atmosphère. La figure 7 tirée de Tsai et al. (2021) présente cet effet pour le cas de la UHJ WASP-33 b en utilisant le code de modélisation VULCAN (Tsai et al., 2017). On peut constater que les molécules sont influencées principalement aux altitudes élevées (faibles pressions), avec des abondances plus faibles qu’attendu. À l’inverse, les espèces mono-atomiques sont favorisées puisque les molécules sont dissociées



**Figure 7** – Courbes d’abondances en rapport de mélange volumique (de l’anglais *Volume Mixing Ratio*) pour certaines espèces chimiques attendues dans l’atmosphère de WASP-33 b. Les courbes en pointillé indiquent les valeurs attendues à l’équilibre chimique, comparativement aux courbes continues qui tiennent compte des effets photochimiques. La figure est tirée de Tsai et al. (2021).

(pour le cas spécifique de WASP-33 b) à cause des températures élevées et du flux intense de l’étoile.

**Mélange vertical** – Les différentes couches verticales considérées dans les modèles 1D ne sont généralement pas cloisonnées. Par exemple, les zones convectives (voir figure 4) peuvent transporter des parcelles de gaz sur une distance considérable par le phénomène d’advection, déstabilisant ainsi l’équilibre des espèces chimiques. Des processus de diffusion peuvent aussi contribuer à mélanger les différentes couches, comme la diffusion moléculaire et la diffusion turbulente (*eddy diffusion*). La capacité de ces phénomènes à déstabiliser l’équilibre chimique dépend de leur temps caractéristique par rapport à la vitesse des réactions chimiques. Certains modèles de cinétique chimique (*kinetics models*) existent pour quantifier ces effets dans le cas des Jupiter chaudes, notamment Zahnle et al. (2009); Moses et al. (2011); Rimmer

& Helling (2016); Tsai et al. (2017). Pour reprendre l'exemple de la figure 7, WASP-33 b ne devrait pas être influencée par les effets de mélange verticaux. En effet, malgré que les courants de convection (verticaux) soient favorisés pour des planètes plus chaudes comme WASP-33 b (ex : Komacek et al., 2019; Tan & Komacek, 2019; Baxter et al., 2021), les vitesses de réactions chimiques le sont également et peuvent ainsi se stabiliser rapidement (Tsai et al., 2021). Par contre, pour les Jupiter chaudes typiques, ce mélange vertical devrait avoir une influence considérable sur les abondances.

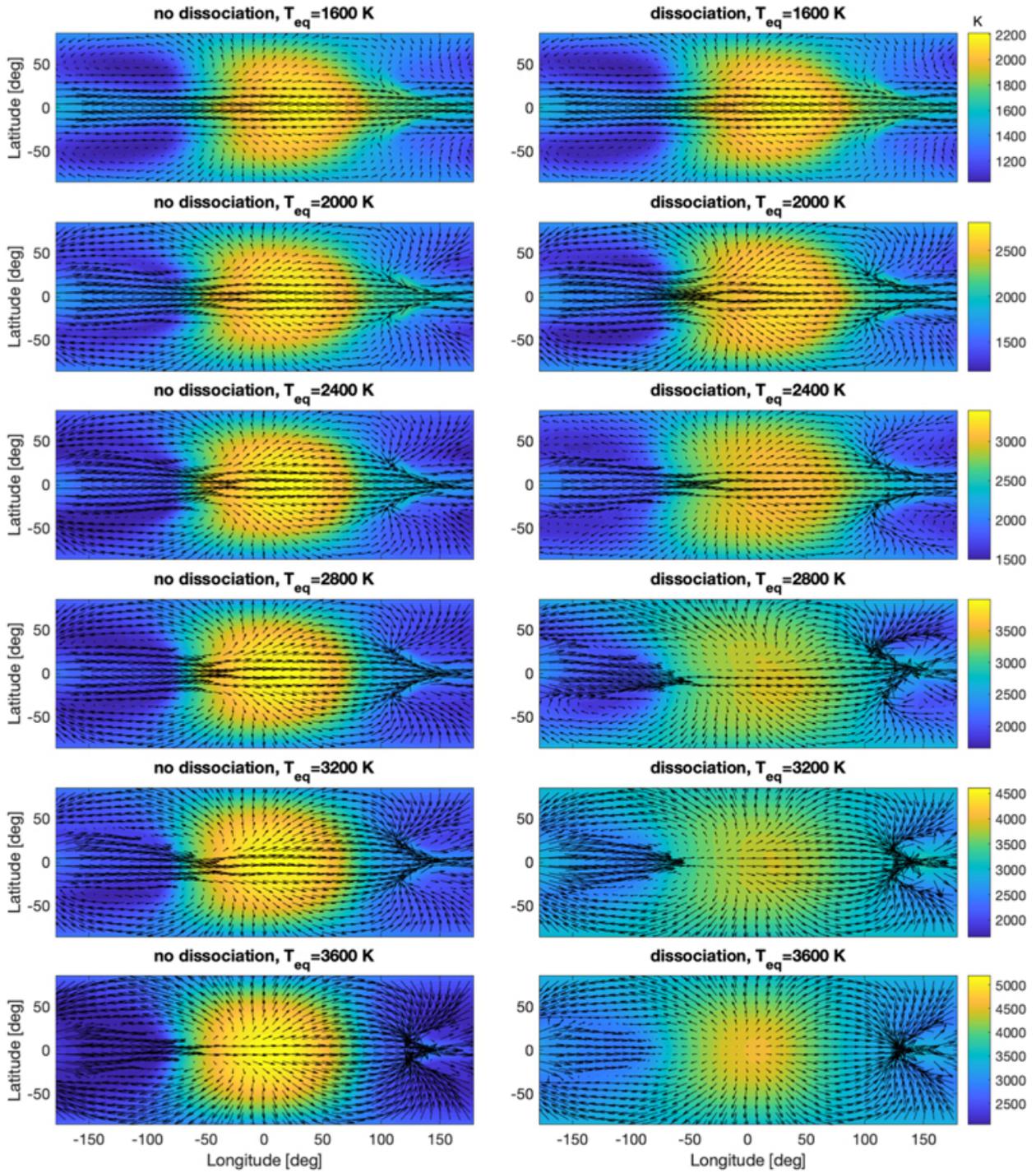
D'autres mécanismes peuvent affecter l'équilibre chimique comme la condensation combinée à des précipitations (impliqué dans les *cold traps* à la section 1.2.3). Le **mélange horizontal** présenté à la section suivante est aussi susceptible de perturber cet équilibre, en particulier dans le cas des UHJ où de grands contrastes de températures sont présents, causant de l'advection dynamique rapide (Baeyens et al., 2021).

### 1.2.5. Circulation atmosphérique

Les atmosphères d'exoplanètes ne sont pas uniformes ; les différentes bandes de Jupiter en témoignent. Les Jupiter chaudes n'y dérogent pas avec leur orbite très rapprochée de leur étoile. Deux conséquences majeures en découlent. Premièrement, l'étoile réchauffe la planète de manière à favoriser l'équateur au détriment des pôles. Deuxièmement, la proximité de ce type de planète les rend très susceptibles d'être en rotation synchrone, c'est-à-dire qu'elle montrent toujours la même face à leur étoile. Un gradient de température important s'établit donc entre le côté jour et le côté nuit, engendrant des processus de circulation atmosphérique.

Une panoplie de modèles ont été proposés par différents auteurs, comme Showman et al. (2009), Rauscher & Menou (2010), Heng et al. (2011) et Kataria et al. (2016), pour n'en nommer que quelques-uns. La plupart s'entendent pour dire que les Jupiter chaudes présenteraient peu de bandes (potentiellement une seule) contrairement aux géantes gazeuses du système solaire. Elles devraient aussi présenter un courant atmosphérique équatorial allant dans la même direction que la rotation de la planète (super-rotation).

De plus, à partir d'une température d'équilibre supérieure à  $\sim 1600$  K, les planètes semblent être peu efficaces pour redistribuer l'énergie du côté jour vers le côté nuit, menant à de grandes différences de températures. Cette tendance est causée en partie par la diminution des échelles de temps radiatives à des températures plus élevées, permettant ainsi de libérer de l'énergie avant qu'elle ne soit transférée au côté nuit par advection. L'apparition de nuage du côté nuit à partir de cette température semble aussi ralentir cette redistribution (Parmentier et al., 2021). Cet effet devrait aussi s'accroître suivant le niveau de radiation de l'étoile (ex : Perez-Becker & Showman, 2013; Komacek & Showman, 2016). Les observations semblent d'ailleurs corroborer ces prédictions pour la plupart des Jupiters chaudes, par



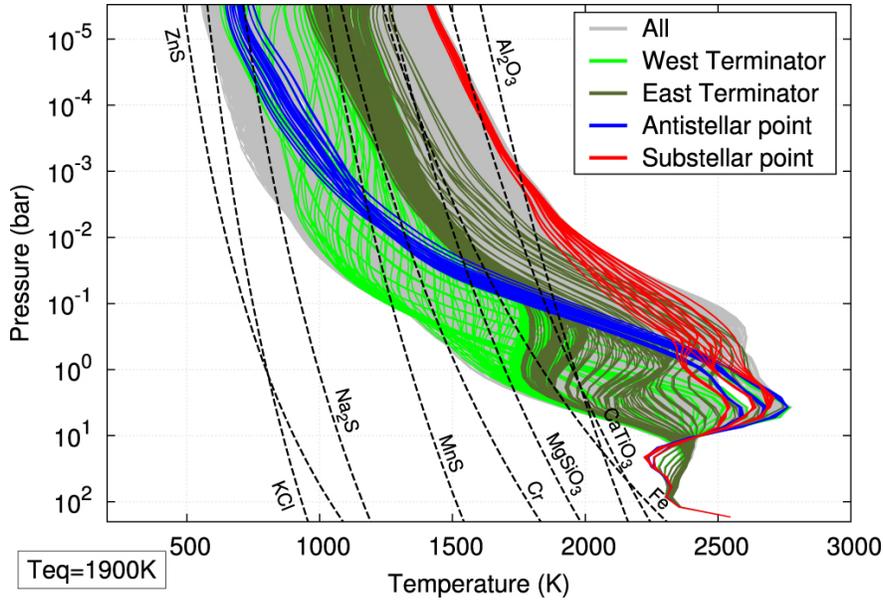
**Figure 8** – Exemples de modèles de circulation globale tiré de Tan & Komacek (2019). Chaque fenêtre représente une carte de la surface de la planète (à une pression de 70 mbar) modélisée en supposant différentes  $T_{\text{eq}}$  et en incluant ou non l’effet de la dissociation de l’hydrogène (voir 1.2.2). L’échelle de couleur indique la température locale et les flèches représentent la direction et la vitesse des vents.

exemple Schwartz et al. (2017) et Zhang et al. (2018). Dans le cas des UHJ, la recombinaison de l'hydrogène pourrait contribuer à réchauffer l'hémisphère nocturne (Bell & Cowan, 2018). En effet, l'hydrogène atomique, dissocié aux températures extrêmes présentes du côté jour (voir section 1.2.2), serait transporté vers le côté nuit (plus froid) où il se recombine en  $H_2$ . Or cette réaction exothermique libère une grande quantité d'énergie, réchauffant ainsi le côté nuit. Une réalisation d'un modèle tirée de Tan & Komacek (2019) présentée à la figure 8 illustre ces effets. Dans la colonne de gauche, on peut remarquer l'augmentation du contraste entre les deux hémisphères lorsque la température augmente. La colonne de droite montre l'efficacité de la recombinaison à redistribuer l'énergie pour des  $T_{eq}$  élevées. Cependant, d'autres phénomènes pourraient aussi accroître ces différences de température, comme la traînée magnétique (*magnétique drag*) due à l'augmentation du champ magnétique suite à l'ionisation aux températures très élevées (Beltz et al., 2022).

Par ailleurs, considérant leur hétérogénéité marquée, il devient difficile de concevoir ces atmosphères comme pouvant être décrites par un (seul) profil de température, à une dimension. La figure 9 montre la diversité de ces structures verticales pour une simulation de Jupiter chaude à 1900 K (Parmentier et al., 2016). Des variations considérables sont présentes entre les différentes régions en termes de température, mais aussi dans la forme des profils. Notons que le point anti-stellaire (à l'opposé du point substellaire) n'est pas toujours le plus froid, particulièrement en profondeur. Ceci est dû au fait que l'énergie est redistribuée vers le côté nuit par des vents allant de l'Ouest vers l'Est. C'est d'ailleurs ce qui explique que le terminateur Ouest soit une des régions les plus froides de la planète. La figure 9 montre aussi les courbes de condensation pour différentes espèces chimiques. Certaines d'entre elles (ex : MnS) croisent uniquement les profils de température de régions plus froides (point substellaire, terminateur ouest), menant au phénomène de piège à froid horizontal (similaire aux pièges à froid verticaux présentés à la section 1.2.4). Cette condensation peut aussi mener à la formation de nuage.

### 1.2.6. Présence de nuages

La formation de nuages a des conséquences considérables sur les spectres d'atmosphères, en particulier dans la spectroscopie de transit. Ceux-ci sont formés par la condensation de particules dans l'atmosphère et vont diffuser la lumière. Cependant, notre compréhension de la formation et de la chimie des nuages demeure lacunaire. La modélisation est donc difficile à faire en termes de la physique fondamentale (ex : Helling et al., 2008; Marley et al., 2013). Plusieurs vont mettre de côté les modèles prédictifs et considérer l'effet des nuages par le biais d'une paramétrisation (ex : Barstow et al., 2017; MacDonald & Madhusudhan, 2017). En général, la pression jusqu'à laquelle les nuages s'étendent (hauteur maximale) ainsi que la manière dont ils absorbent la lumière (type de diffusion) seront les paramètres de



**Figure 9** – Profils de température pour une Jupiter chaude à  $T_{eq} = 1900$  K (en gris : toutes longitudes et latitudes ; vert pâle : terminateur (Interface entre le côté jour et le côté nuit) à l’Ouest du point substellaire (Point où la surface de une planète fait directement face à son l’étoile) ; vert foncé : terminateur à l’Est du point substellaire ; bleu : à moins de  $20^\circ$  du point à l’opposé du point substellaire ; rouge : à moins de  $20^\circ$  du point substellaire). La figure est tirée de Parmentier et al. (2016) et les modèles de circulation globale proviennent de Showman et al. (2009). Les courbes de condensation de différentes espèces chimiques sont aussi présentées.

base. Certains modèles vont aussi inclure la grosseur des particules diffusives, la distribution verticale de celles-ci dans l’atmosphère, la pression à partir de laquelle se forment les nuages (hauteur de la base) et/ou la nature chimique des composantes (ex : Helling et al., 2017, 2019; Parmentier et al., 2016).

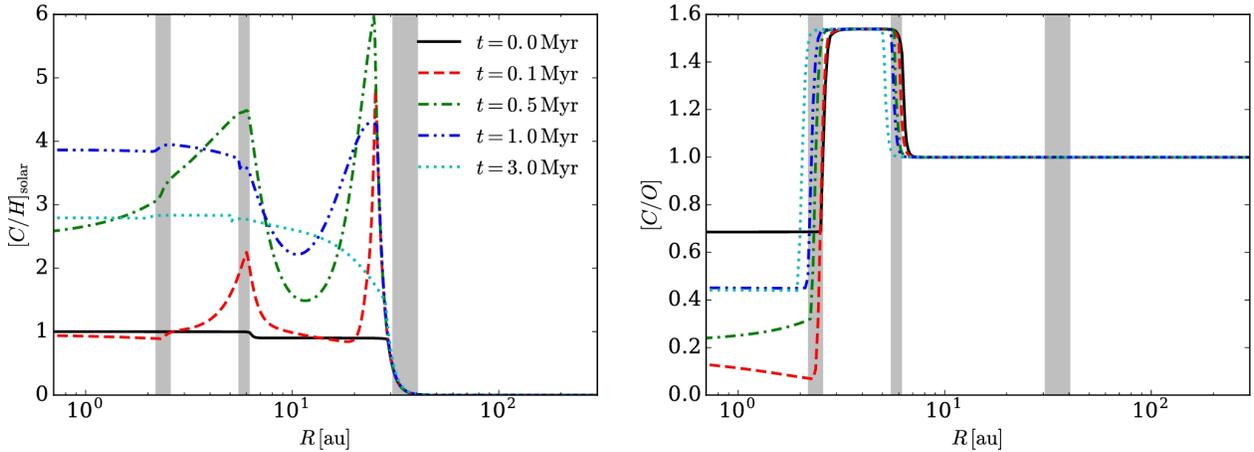
Comme vu précédemment, les changements dans l’état (solide, liquide ou gazeux) d’une molécule peuvent être caractérisés par des diagrammes de phase mettant en relation la température et la pression et sont directement reliés aux profils de température par des courbes de condensation. Ces dernières ne donnent cependant pas d’information sur la densité des nuages ainsi que leur étendue verticale, qui sont souvent laissées comme paramètres libres.

La manière dont la lumière est diffusée par ces condensats est modélisée par la théorie de Mie. Certains auteurs calculent directement les coefficients d’absorption à partir de la taille des particules (ex : Ackerman & Marley 2001; Parmentier et al. 2016). Cette technique donne plus d’information sur la nature des constituants. D’autres utilisent des formes plus générales de la théorie en paramétrant l’absorption. Deux types sont considérés, soit la diffusion (ou opacité) *grise* (c’est-à-dire uniforme selon la longueur d’onde  $\lambda$ ) et la diffusion de Rayleigh allant comme  $\lambda^{-4}$ . Le premier cas s’applique aux particules dont la taille est de l’ordre de  $\lambda$  ou plus grande. La diffusion de Rayleigh est le cas limite de la théorie de Mie où la taille des

particule est  $\ll \lambda$ . Les aérosols (particules solides) présents dans l’atmosphère peuvent aussi produire de la diffusion de type Rayleigh.

### 1.2.7. Rapport C/O et métallicité

Malgré qu’elle puisse devenir très complexe, la détermination de la composition globale de l’atmosphère est cruciale pour comprendre les processus de formation des exoplanètes, notamment via le rapport C/O et la métallicité<sup>7</sup>. Ceux-ci sont reliés aux processus de formation des exoplanètes (ex : Öberg et al., 2011; Madhusudhan et al., 2014a; Ali-Dib et al., 2014; Booth et al., 2017; Eistrup et al., 2018; Dash et al., 2022). En effet, comme le montre la figure 10, la distance où s’est formée la planète par rapport à son étoile influence les rapports d’abondance dans le gaz disponible. Ces rapports sont aussi modulés par les changements dans l’orbite d’une planète durant son évolution ainsi que par la manière dont la planète s’est formée (ex : régimes d’accrétion, instabilité gravitationnelle).



**Figure 10** – Rapports C/H (gauche) et C/O (droite) présents sous forme gazeuse dans le disque proto-planétaire en fonction de la distance par rapport à l’étoile hôte. Les différentes courbes montrent l’évolution temporelle de ces abondances relatives. Les zones ombragées (de gauche à droite respectivement) représentent les régions où le H<sub>2</sub>O, le CO<sub>2</sub> et le CO se condensent sous forme de glace, appauvrissant ainsi le gaz. La figure est tirée de Booth et al. (2017).

La détermination du rapport C/O peut se faire en énumérant les éléments après avoir déterminé les abondances de chacune des molécules détectables (ex : Lee et al., 2012; Line et al., 2013; Tremblin et al., 2015; Baudino et al., 2015; Mollière et al., 2015), ou alors en modélisant l’activité chimique globale dans l’atmosphère (ex : Fortney et al., 2008; Miller-Ricci et al., 2009; Benneke & Seager, 2013; Lothringer et al., 2018). La première méthode est la plus directe. Les abondances de chaque élément y sont traitées de manière indépendante comme paramètres libres. Cependant, elle peut être biaisée puisque l’atmosphère détectable

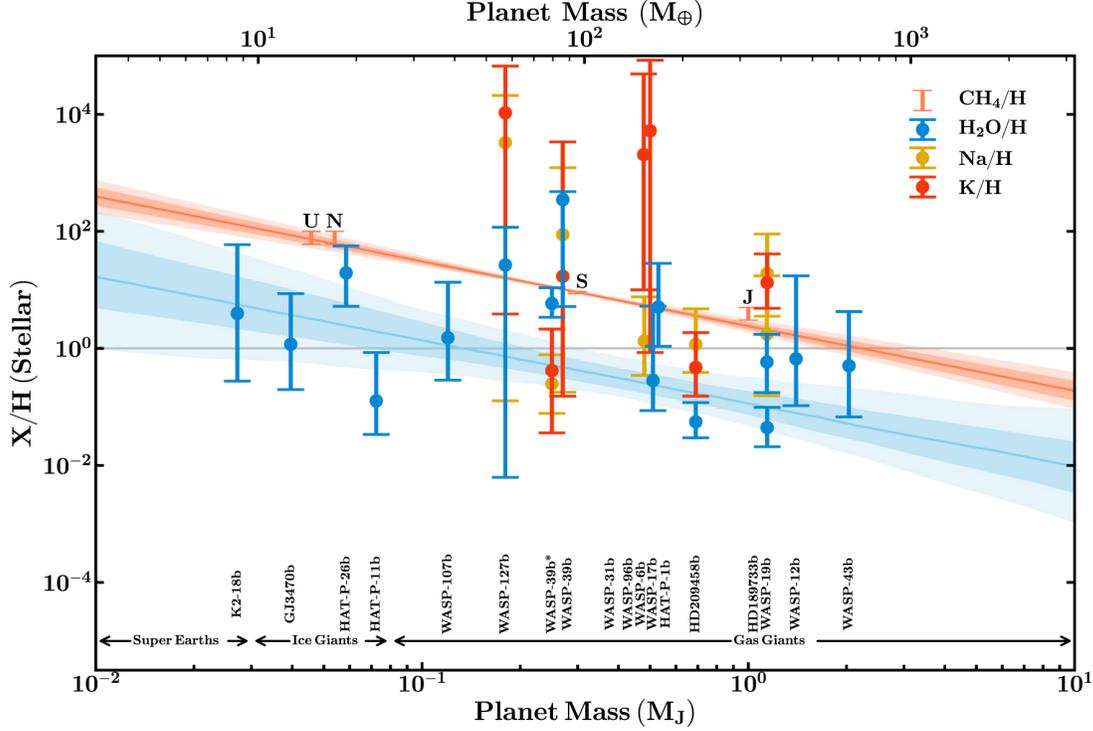
7. La métallicité est donnée par la fraction d’éléments qui ne sont pas de l’hydrogène ou de l’hélium.

n'est pas nécessairement représentative de la véritable composition. Comme dans le cas de Jupiter, certaines espèces peuvent se condenser profondément dans l'atmosphère, restant dissimulées. La seconde méthode (ex : Visscher et al., 2010; Kreidberg et al., 2015; Benneke, 2015) comporte aussi sa part de problèmes. En plus de son temps de calcul plus onéreux, elle nécessite une bonne compréhension des phénomènes physico-chimiques et suppose un équilibre chimique qui n'est pas nécessairement respecté. Cependant, elle a l'avantage de pouvoir contraindre le rapport C/O sans avoir à détecter de molécules porteuses de carbone ou d'oxygène. Ce genre d'analyse nécessite seulement deux paramètres libres (pour la chimie), soit le rapport C/O et la métallicité.

La détermination de la métallicité peut se faire de la même manière, c'est-à-dire en modélisant les processus chimiques dans l'atmosphère, ou de manière similaire à la méthode directe pour le rapport C/O. Cette dernière se résume à utiliser l'abondance d'une molécule bien contrainte comme proxy de la métallicité. En général l'eau est utilisée pour les Jupiter chaudes et le méthane pour les géantes du système solaire. Plus récemment, le CO a aussi été proposé (ex : Pelletier et al., 2021). Concrètement, l'abondance de la molécule adéquate est comparée avec l'abondance estimée pour la même composition que le soleil, mais ajustée à la métallicité de l'étoile et corrigée à la température correspondante. Welbanks et al. (2019) utilise cette méthode afin de comparer les géantes du système solaire avec les Jupiter chaudes. La figure 11 tirée du même article présente la relation qui semble exister entre la masse des planètes et leur métallicité, se manifestant par une loi de puissance.

Cependant, on sait qu'il existe une dégénérescence entre le rapport C/O et la détermination des différentes abondances (ex : Madhusudhan et al. 2014b; Benneke 2015; Kreidberg et al. 2015). Par exemple, un taux plus élevé d'oxygène peut favoriser la formation de H<sub>2</sub>O (l'espèce la plus susceptible d'abriter l'oxygène entre 300 et 3500 K, Madhusudhan et al. 2014b) pour une même métallicité, et donc accroître la signature de cette molécule. À l'inverse, un rapport C/O élevé favoriserait la formation de molécules à base de carbone au détriment de l'eau (Madhusudhan, 2012). Cet effet est illustré à la figure 12.

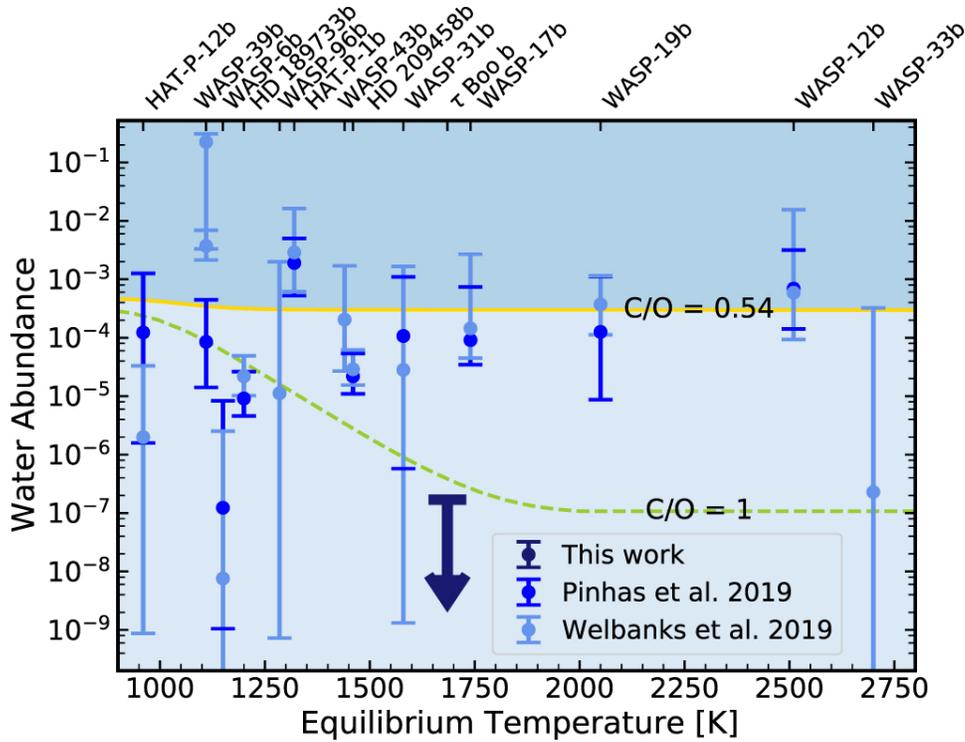
Les différentes méthodes d'analyse révèlent le lien étroit entre le rapport C/O et la métallicité. La modélisation n'est cependant pas toujours en accord avec les méthodes plus directes. Par exemple, dans le cas de WASP-12 b et HD 209458 b, Stevenson et al. (2014); Madhusudhan et al. (2014b); MacDonald & Madhusudhan (2017) (méthode directe) obtiennent un rapport C/O  $\gtrsim 1$  (ou une métallicité très faible) alors que la modélisation complète (Benneke, 2015; Kreidberg et al., 2015) favorise des rapports C/O  $< 1$  (donc faibles). Certaines contradictions dans une même méthode existent aussi (voir Line et al. 2016 comparé à MacDonald & Madhusudhan 2017). Celles-ci sont dues à d'autres considérations comme le traitement des nuages (voir section 1.2.6) ou des données de nature différente (transit VS éclipse; voir chapitre 2).



**Figure 11** – Relation entre la masse et la métallicité pour l’atmosphère des géantes du système solaire et celle de 17 exoplanètes. Le méthane ( $\text{CH}_4$ , couleur corail) est utilisé comme indicateur de la métallicité pour les planètes du système solaire (Atreya et al., 2018), alors que l’eau ( $\text{H}_2\text{O}$ , bleu), le sodium (Na, jaune) et le potassium (K, rouge) sont pour l’échantillon d’exoplanètes. Chaque valeur est normalisée par rapport à la valeur attendue en supposant les mêmes abondances que leur étoile hôte. Les droites, ainsi que les zones ombragées les entourant, représentent la meilleure solution et les intervalles de confiance ( $1\sigma$  et  $2\sigma$ ) pour une loi de puissance. La figure est tirée de Welbanks et al. (2019).

Pour réconcilier les techniques d’analyse, il faudra acquérir plus d’informations sur les abondances des différentes molécules présentes dans l’atmosphère des exoplanètes (ex : Line et al., 2013; Spake et al., 2021b). Dans le cas des Jupiters chaudes, l’oxygène est déjà bien contraint grâce à l’eau, ce qui n’est pas le cas pour le carbone (contenu majoritairement dans le  $\text{CO}$  ou le  $\text{CH}_4$ , Benneke 2015). À cet égard, l’apport du Télescope spatial James Webb et de la spectroscopie à haute résolution sera présenté dans le chapitre 2.

Le cas de Jupiter elle-même reste mal contraint. En effet, la sonde *Juno*, présentement en orbite autour de Jupiter, devait contraindre l’abondance d’eau dissimulée très profondément dans son atmosphère. Une limite supérieure avait été précédemment obtenue par la sonde *Galilée* (Atreya & Wong, 2005; Mousis et al., 2012). En sondant la région équatoriale, *Juno* a pu déterminer un rapport de mélange molaire (MMR, *Molar Mixing Ratio*) de  $2.5^{+2.2}_{-1.6} \times 10^3$  ppm, soit  $2.7^{2.4}_{-1.7}$  fois l’abondance solaire en oxygène (Li et al., 2020). Cependant, cette mesure ne peut pas exclure une absence d’eau (MMR= 0 ppm) à  $2\sigma$  et donc demeure une



**Figure 12** – Mesures de l’abondance d’eau (VMR) pour un échantillon de Jupiter chaudes en fonction de la température d’équilibre. Les valeurs sont tirées de Pinhas et al. (2019) et Welbanks et al. (2019), auxquelles la valeur mesurée pour  $\tau$  Boo b est ajoutée (*This work*; Pelletier et al. 2021). La valeur attendue pour deux différents C/O sont représentés par les courbes dorée (C/O = 0,54) et verte (C/O = 1).

limite supérieure. De plus, les auteurs indiquent que leur résultat n’est pas nécessairement représentatif de l’abondance globale sur Jupiter.

### 1.3. Détermination du spectre de l’atmosphère

Avec la théorie développée précédemment, il est maintenant possible de déterminer un spectre d’atmosphère. Pour les modèles en une dimension, la structure verticale est divisée en couches uniformes. Avec un profil en température donné (T-P), on peut calculer la densité ainsi que les différentes abondances des espèces chimiques (voir section 1.2). Les opacités dues à chaque molécules sont influencées par la température et la pression et peuvent donc être déterminées. Notons que les modèles plus complexes (auto-cohérents) entameront ensuite un processus itératif afin que l’équilibre radiatif soit respecté (ex : Burrows et al., 1997; Barman et al., 2001; Hubeny et al., 2003; Lothringer et al., 2018). Des corrections au profil de température sont alors apportées jusqu’à convergence du modèle (Lucy, 1964; Hauschildt et al., 2003). Les modèles plus simples ont cependant l’avantage d’être plus efficaces. De plus, les modèles complexes doivent introduire davantage de suppositions, comme l’équilibre

radiatif, équilibre chimique, la connaissance approfondie des réactions ayant lieu, etc., ce qui n'est pas nécessairement représentatif (ex : sections 1.2.4 et 1.2.5).

Une fois les opacités calculées, il est possible de déterminer l'intensité du flux émergent à chaque couches via l'équation de transfert radiatif. Une de ses déclinaisons s'écrit comme suit :

$$\frac{dI_\nu}{ds} = -\chi_\nu I_\nu + \epsilon_\nu. \quad (1.3.1)$$

Elle décrit l'influence du passage au travers d'un milieu d'épaisseur  $ds$  sur l'intensité ( $I_\nu$ ) d'un jet de lumière. La nouvelle variable  $\epsilon_\nu$  introduite ici est l'émissivité. Une couche de matière peut donc absorber et émettre de l'intensité. On peut réécrire l'équation en fonction de la profondeur optique, ce qui mène à

$$\frac{dI_\nu}{d\tau_\nu} = I_\nu + \frac{\epsilon_\nu}{\chi_\nu}. \quad (1.3.2)$$

Le dernier terme de l'équation est une fonction source parfois notée  $S_\nu = \frac{\epsilon_\nu}{\chi_\nu}$ . Plusieurs simplifications peuvent être faites dans le cas des exoplanètes. Celles-ci seront présentées au chapitre 2.

L'équation de transfert radiatif a longtemps été utilisée pour caractériser les atmosphères stellaires. Les différences principales dans le cas des exoplanètes sont la présence de nuages ainsi que l'influence de l'étoile hôte modifiant les conditions frontières (température d'équilibre).

Avec la fonction source donnée et indépendante de  $\tau_\nu$ , il est possible d'intégrer sur  $\tau_\nu$  pour obtenir la relation suivante,

$$I_\nu(\tau_\nu) = I_\nu(0)e^{-\tau_\nu} + S_\nu(1 - e^{-\tau_\nu}). \quad (1.3.3)$$

Nous y référerons dans le chapitre suivant.



# Chapitre 2

---

## Spectroscopie appliquée aux atmosphères d'exoplanètes

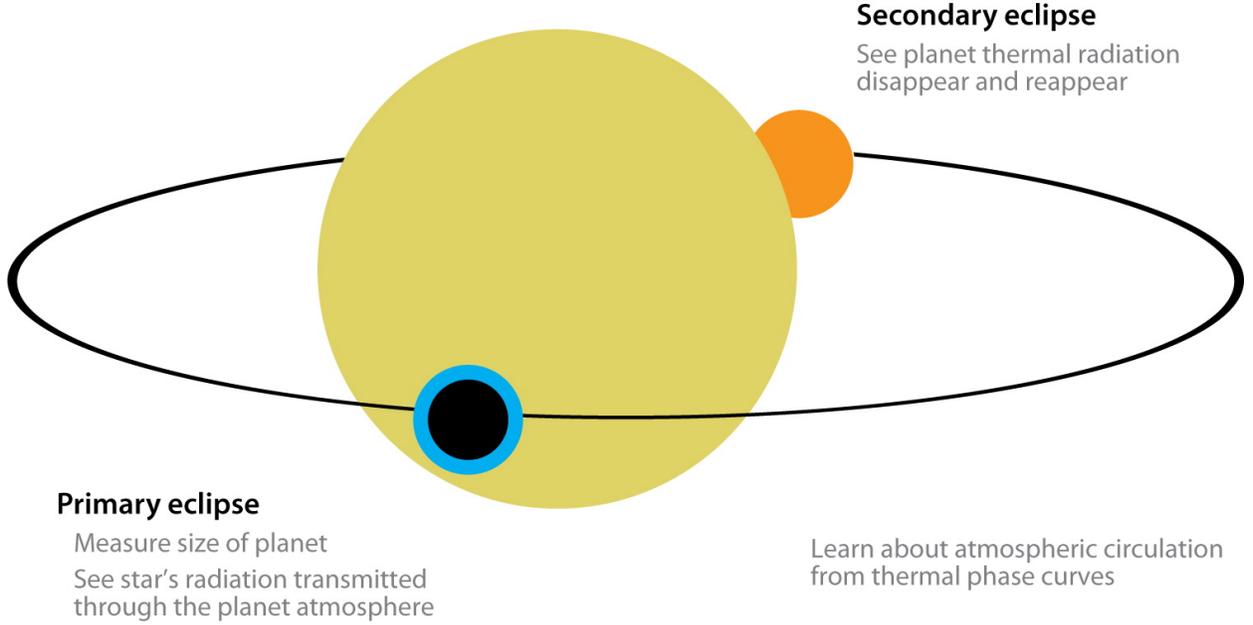
L'interaction entre la lumière et la matière permet entre autres de déterminer les propriétés chimiques de l'atmosphère des exoplanètes. Chaque espèce chimie possède une signature spectrale bien précise suite à l'émission ou l'absorption de photons causée par les transitions d'électrons d'un niveau d'énergie à un autre. Les transitions d'énergie de la rotation de la vibration de certaines molécules produisent aussi une telle signature spectrale. Ce ne sont pas toutes les composantes des exoplanètes qui seront détectables. Par exemple, celles-ci peuvent tout simplement être séquestrées trop profondément dans l'atmosphère pour être visibles. Les conditions physiques environnantes ont aussi un effet sur l'intensité et la forme particulière de ces signatures. Le terme « atmosphère détectable » sera utilisé pour désigner ce qui est visible et non ce qui compose en réalité l'atmosphère. Une bonne connaissance des phénomènes physiques et chimiques ayant lieu dans les atmosphères d'exoplanètes permet de faire le pont entre les deux.

### 2.1. Éléments de base

La spectroscopie de transit ou d'éclipse consiste à utiliser les variations de la lumière dues à l'occultation de l'étoile par la planète (éclipse primaire ou transit) ou de la planète par l'étoile (éclipse secondaire ou éclipse). Les deux situations sont représentées à la figure 1. En étudiant la modulation du spectre au cours d'un de ces événements, il est possible de départager le signal provenant de l'étoile et celui de la planète. On doit donc mesurer le flux à une certaine fréquence temporelle pour obtenir ce qu'on appelle une « courbe de lumière ».

#### 2.1.1. Éclipse primaire (transit)

La quantité de lumière bloquée par la planète, sans considérer l'atmosphère et l'assombrissement centre-bord, est simplement donnée par le rapport des aires :  $\pi R_p^2 / \pi R_\star^2$ . À cela,



**Figure 1** – Illustration des phénomènes d'éclipses primaire et secondaire tirée de Seager & Deming (2010). Coin supérieur droit : « *Éclipse secondaire, On peut voir la radiation due à la planète disparaître puis réapparaître* ». Coin inférieur gauche : « *Éclipse primaire, On peut mesurer la taille de la planète et on voit la radiation de l'étoile transmise à travers l'atmosphère de la planète* ». Coin inférieur droit : « *On peut apprendre à propos de la circulation atmosphérique par une courbe de phase thermique* ».

il faut ajouter l'atténuation de la lumière traversant l'atmosphère. L'intensité de lumière transmise est décrite par l'équation 1.3.3 présentée à la section 1.3. Dans le cas de l'éclipse primaire, la lumière provenant de l'étoile domine toute autre source de radiation, donc le dernier terme de l'équation disparaît. On a alors

$$I_{\lambda}(\tau_{\lambda}) = I_{\lambda}(0)e^{-\tau_{\lambda}}. \quad (2.1.1)$$

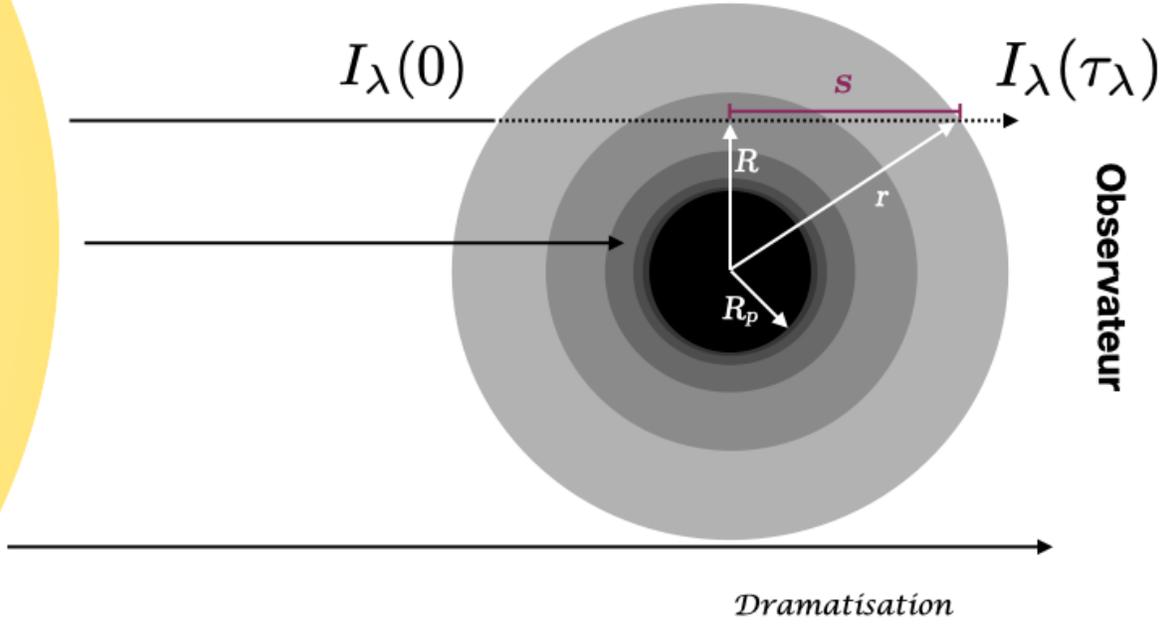
$I_{\lambda}(0)$  dans ce cas-ci est l'intensité en provenance de l'étoile. On cherche la proportion de lumière absorbée, donc  $1 - I_{\lambda}/I_{\lambda}(0)$ . Comme l'atmosphère n'est pas uniforme, on utilise le principe de stratification (comme illustré à la figure 2).

On peut calculer l'intensité absorbée dans un anneau d'atmosphère d'épaisseur infinitésimale  $dR$  et de rayon  $R$  par

$$(\text{aire de l'anneau}) \times (\text{proportion de lumière absorbée}) \approx 2\pi R dR \times (1 - e^{-\tau_{\lambda}(R)}). \quad (2.1.2)$$

La profondeur optique  $\tau_{\lambda}(R)$  doit être calculée en intégrant sur le parcours total  $2s$  (voir figure 2). Numériquement, il faut sommer sur toutes les couches traversées, donc

$$\tau_{\lambda}(R) = 2 \sum_n k_{\lambda n} \rho_n \Delta s_n. \quad (2.1.3)$$



**Figure 2** – Schéma de la géométrie d'un transit.  $R$  est la distance projetée par rapport au centre de la planète. Cette dernière est perpendiculaire à la ligne de visée (axe horizontal). Les différentes régions grises ombragées représentent les couches d'un modèle d'atmosphère à une dimension (verticale). Les modèles sont paramétrés selon la pression et non l'altitude, ce qui explique l'augmentation de l'épaisseur des strates (voir section 1.1.1).

Les termes  $k_{\lambda n}$  et  $\rho_n$  proviennent de l'équation 1.1.6 et sont déterminées par un modèle d'atmosphère. Le terme  $\Delta s_n$  est relié aux différentes couches par  $R$  et  $r$ . Ainsi, à une certaine longueur d'onde, la quantité totale de lumière bloquée est donnée par

$$\frac{\Delta f}{f} = \frac{R_p^2 + \int_{R_p}^{\infty} 2R(1 - e^{-\tau_{\lambda}(R)})dR}{R_{\star}^2}, \quad (2.1.4)$$

où  $f$  est le flux en dehors du transit et  $\Delta f$  est la variation de flux due au transit.  $\tau_{\lambda}$  est responsable de la dépendance selon la longueur d'onde, ce qui permet d'obtenir un spectre.

La figure 3 (partie gauche) présente des courbes de lumière à différentes pages de longueurs d'onde (bandes spectrales) obtenues avec le télescope Hubble. La partie de droite montre la variation de la profondeur des courbes de lumière (profondeur de transit). On peut voir la différence de la forme des transits (gauche) ainsi que dans la profondeur (droite). Pour une planète particulière, la forme de la courbe de lumière sera principalement influencée par l'« assombrissement centre-bord » de l'étoile hôte. En bref, les bords de l'étoile contribuent moins au flux que le centre due à la géométrie. En effet, l'angle entre la ligne de visée d'un

observateur et l'axe vertical (ou altitude) augmente près des extrémités. Par opposition, au centre de l'étoile, cet angle devient nul et les couches de l'atmosphère sont donc perpendiculaires à la ligne de visée. Ainsi, la contribution aux bords de l'étoile provient de couches plus élevées dans l'atmosphère et donc généralement plus froides (dans le cas d'une étoile typique). Ainsi, plus la planète sera près du centre de l'étoile, plus la proportion de lumière bloquée sera grande. Cet effet est plus important pour les courbes se rapprochant du violet puisque les longueurs d'onde courtes sont plus sensibles aux variations de température.

La partie de droite de la figure 3 montre la sensibilité des spectres de transmission par rapport aux nuages (voir section 1.2.6). Pour le modèle incluant la diffusion de Rayleigh (courbe bleue), on observe plus d'absorption aux longueurs d'onde plus courtes. En comparaison, pour le modèle considérant une opacité grise (courbe rouge) les bandes d'absorption d'eau situées dans l'infrarouge sont moins prononcées.

Finalement, on peut estimer l'intervalle des pressions sondées d'environ 4 à 5 hauteurs d'échelle ( $H$ ) en partant de l'équation 2.1.4 et en considérant une atmosphère isotherme (développement dans Griffith 2014). Ainsi, au regard de l'équation 1.1.3, une température élevée, un poids moléculaire moyen et une gravité de surface faible mènent à une variation du signal plus prononcée puisque les pressions sondées seront étalées sur des altitudes plus grandes. Dans le cas d'une Jupiter chaude, le spectre de transit devrait pouvoir atteindre des variations de  $\sim 0,1\%$  ce qui se rapproche de la valeur de  $0,03\%$  montrée à la figure 3 (droite).

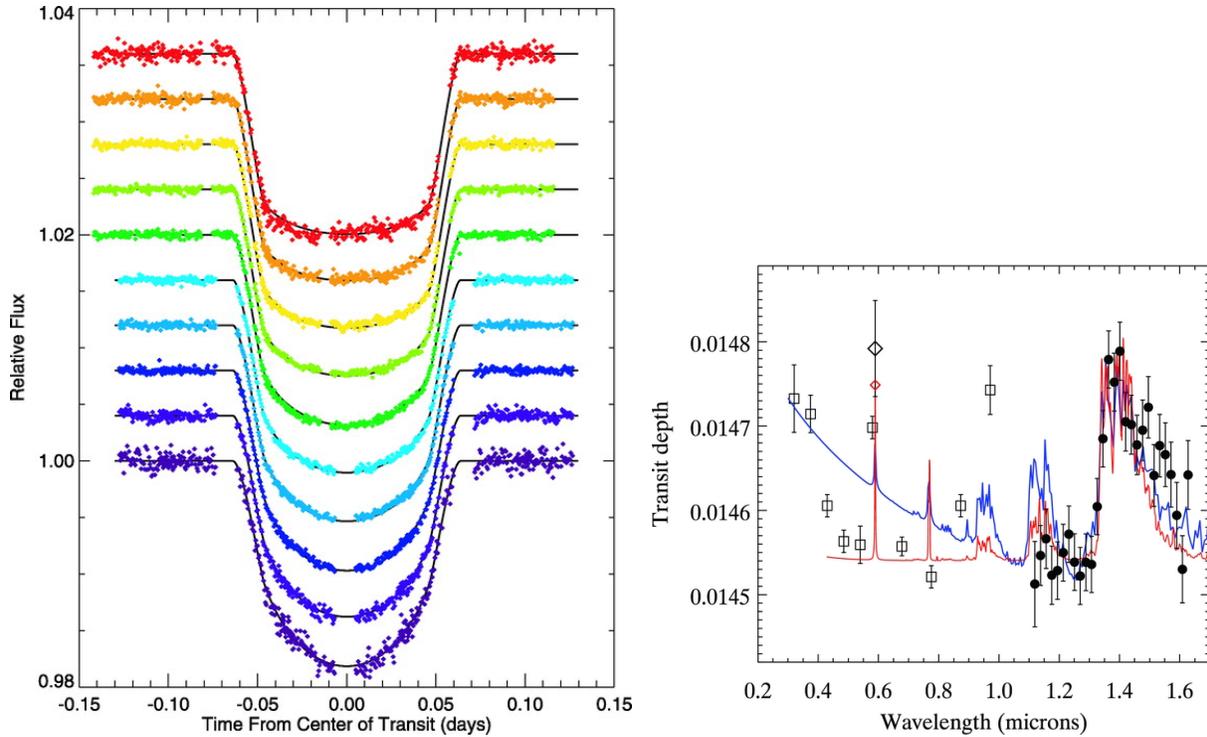
## 2.1.2. Éclipse secondaire (éclipse)

Durant l'éclipse secondaire, la lumière provenant de la planète est totalement bloquée par l'étoile tandis qu'on observe la somme du flux de la planète et de l'étoile en dehors de l'éclipse. Ainsi, la profondeur relative du signal de l'éclipse est donnée par

$$\frac{\Delta f}{f} = \frac{(f_\star + f_p) - f_\star}{f_\star} = \frac{f_p}{f_\star}. \quad (2.1.5)$$

Le cas de l'éclipse secondaire est bien différent de celui de l'éclipse primaire. Le spectre obtenu est en majeure partie dû à l'émission thermique de la planète elle-même pour des longueurs d'onde situées dans l'infrarouge. L'amplitude du signal est donc directement liée à la température de la planète et à son rayon. De plus, puisque la lumière reçue provient directement de la planète, celle-ci doit franchir un parcours beaucoup moins grand au travers des différentes couches, ce qui permet de sonder plus profondément dans l'atmosphère.

Finalement, dans le cas des Jupiters chaudes, la région de la planète faisant face à l'observateur est celle complètement illuminée par l'étoile, c'est-à-dire la région diurne. C'est un avantage puisque cette région est généralement beaucoup plus chaude (voir section 1.2.5) et



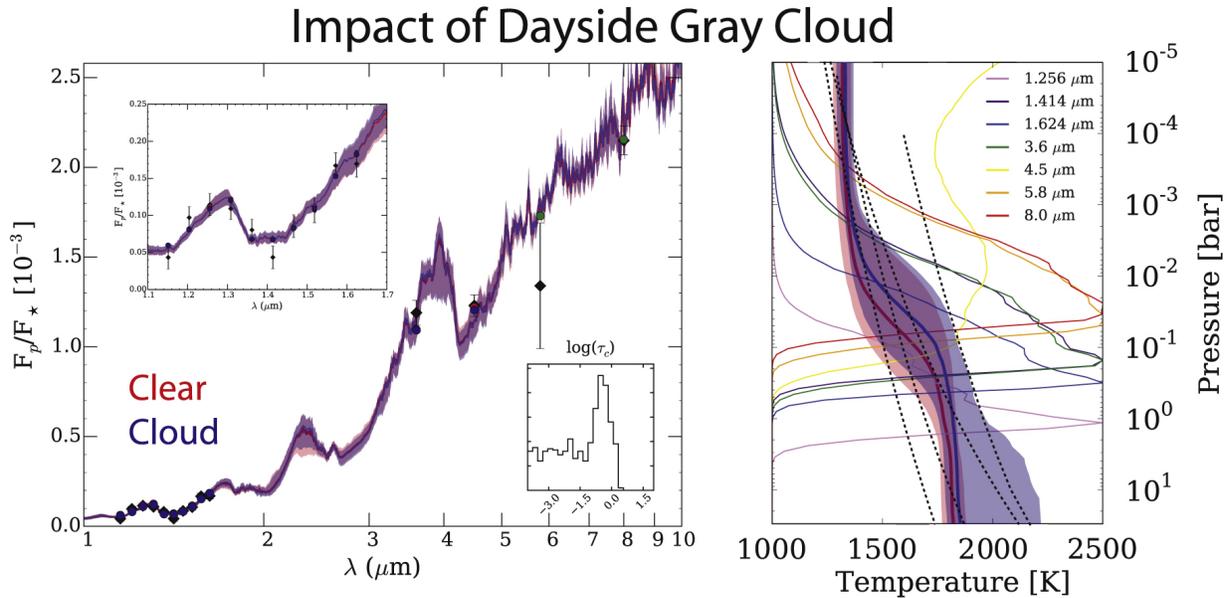
**Figure 3 – Gauche :** Figure tirée de Knutson et al. (2007a) présentant des courbes de lumières prises avec l'instrument STIS (*Space Telescope Imaging Spectrograph*) du télescope spatial Hubble pour l'exoplanète de type Jupiter chaude HD 209458 b. Dix bandes spectrales sont présentées avec les courbes théoriques calculées analytiquement pour chacune d'elle (Mandel & Agol, 2002). Chacune des courbes successives sont décalées de 0,004.

**Droite :** Spectre de transmission tiré de Deming et al. (2013) combinant les observations de plusieurs instruments pour HD 209458 b. Les carrés blancs présentent les données de Knutson et al. (2007a) réanalysées (figure de gauche), les cercles noirs présentent celles de l'instrument WFC3 (*Wide Field Camera 3*) de Hubble (Deming et al., 2013) et le diamant blanc est tiré de Charbonneau et al. (2002). Un modèle isotherme considérant des nuages avec une opacité grise (Burrows et al., 2007) est décrit par la courbe rouge. Celle en bleue (Dobbs-Dixon et al., 2012) est obtenue en considérant une opacité allant comme  $\lambda^{-4}$  de type Rayleigh, sans opacité grise additionnelle. Le diamant rouge est simplement donné par l'intégration du modèle en rouge sur la même bande spectrale que le diamant blanc.

elle émet donc plus intensément. Les conditions peuvent cependant être différentes de celles sondées par l'éclipse primaire.

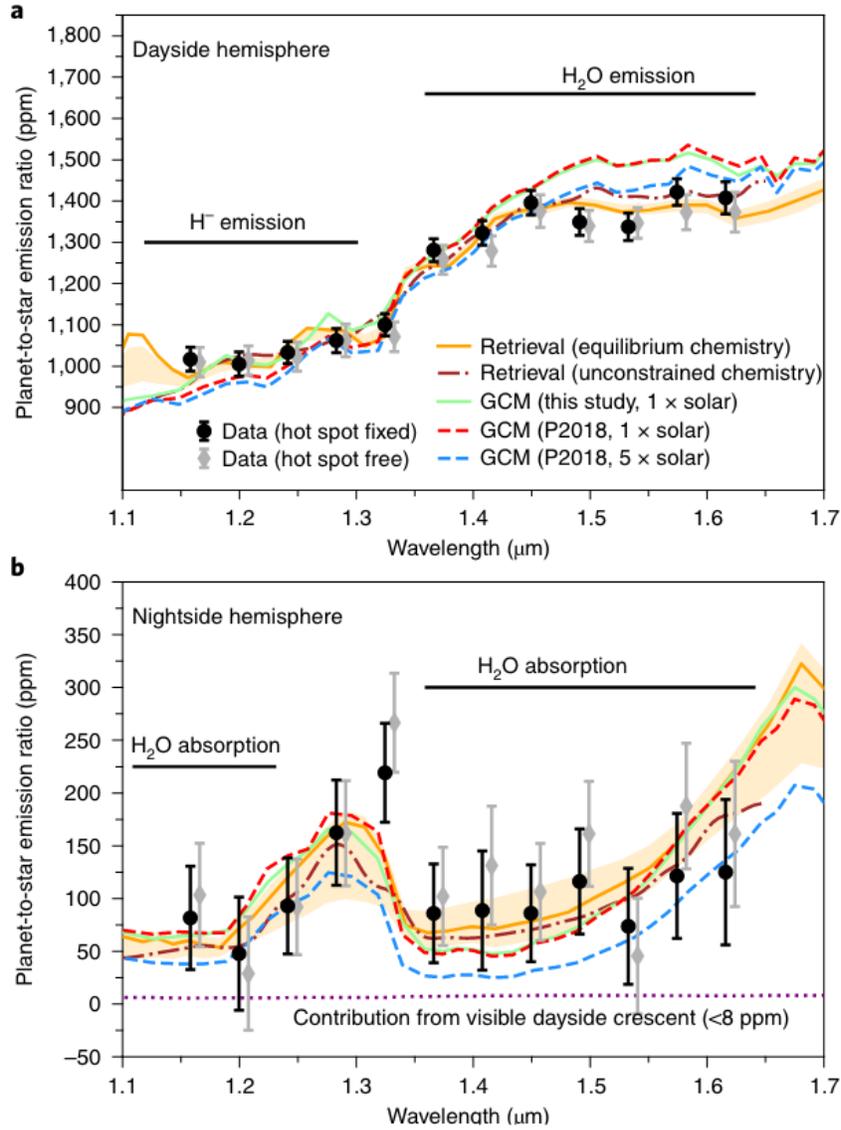
La figure 4 présente un exemple de spectre d'émission pour la même exoplanète qu'à la figure 3, HD 209458 b. On peut voir que l'amplitude du signal augmente en allant vers les longueurs d'onde plus élevées. Ceci est dû au fait que le flux de l'étoile décroît selon l'approximation de Rayleigh-Jeans ( $f_{\star} \propto \lambda^{-4}$ ) alors que celui de HD 209458 b atteint son maximum dans les environs de  $10 \mu\text{m}$ . Les variations dans le spectre sont du même ordre que pour le transit ( $\leq 0,01\%$ ), ce qui est vrai seulement dans certains cas comme celui de

HD 209458 b. Pour les planètes plus petites ou plus froides, par exemple, le spectre d'émission est généralement beaucoup plus difficile à détecter que celui en transmission.



**Figure 4** – Figure tirée de Line et al. (2016) présentant, à gauche, le spectre d'émission de HD 209458 b et, à droite, le profil de température. Pour les deux fenêtres, les 1000 meilleures solutions pigées aléatoirement dans un intervalle de confiance de  $1\text{-}\sigma$  sont présentées par les régions de couleur pâles. Les solutions médianes de celles-ci sont données par les courbes de la même couleur. La faible différence entre le modèle *Clear* sans nuages (en rouge) et celui *Cloudy* avec nuages (en bleu) montre que ceux-ci ont très peu d'impact pour un spectre d'émission. **Fenêtre de gauche** : Les données de  $1$  à  $2\ \mu\text{m}$  proviennent de Hubble/WFC3 et les 4 autres, de l'instrument IRAC (*InfraRed Array Camera*) à bord du télescope spatial *Spitzer* (Diamond-Lowe et al., 2014). Un agrandissement des données de WFC3 est montré dans le coin supérieur gauche. La distribution de probabilité de la profondeur optique  $\tau_c$  due aux nuages est présentée dans le coin inférieur droit. **Fenêtre de droite** : Les courbes en couleurs sont les fonctions de contributions (ou fonction de Chapman) à une bande passante donnée. Celles-ci indiquent comment chaque section de l'atmosphère contribue au spectre. Le profil de température déterminé dans les sections où cette fonction est pratiquement nulle sont donc une extrapolation due aux modèles. Les courbes de condensation (voir section 1.2.6) pour le  $\text{MgSiO}_3$ ,  $\text{MgSiO}_4$ , Fe, et TiO sont en pointillé.

2.1.2.1. Sensibilité au profil de température. La forme du spectre d'émission mérite aussi que l'on s'y attarde. En effet, dans le cas de HD 209458 b, la bande d'eau à  $1,4\ \mu\text{m}$  se trouve en absorption. Ceci indique un gradient de température allant en diminuant selon l'altitude, comme le profil de température de la fenêtre de droite (figure 4) le montre. Par contre, certaines Jupiters très chaudes (UHJ) vont plutôt arborer des bandes d'émission, comme c'est le cas pour le côté jour de WASP-121 b (figure 5 a). Une même planète peut aussi présenter des variations considérables selon les phases observées ; la fenêtre b montre la bande d'eau se



**Figure 5** – Spectres d’émission de WASP-121 b observés du côté jour et côté nuit. La figure est tirée de Mikal-Evans et al. (2022). Les données, représentées par les cercles noirs, proviennent de l’instrument WFC3. Les diamants gris montrent les résultats pour une méthode de réduction alternative. Différents modèles identifiés par les courbes de couleur sont comparés aux observations. Les bandes d’absorption dominantes sont identifiées par les traits noirs. La fenêtre du haut montre le spectre obtenu durant l’éclipse (côté jour) et celle du bas présente celui obtenu à une phase de 0,95, c’est-à-dire immédiatement avant le début du transit. On voit donc en **b** le côté nuit avec une très faible contribution d’un croissant de côté jour, pour lequel la contribution est indiquée par la ligne pointillée violette.

trouvant plutôt en absorption du côté nuit. Cette relation entre le gradient de température (dans la région de l’atmosphère sondée) et la forme des bandes (ou raies) peut s’expliquer de manière qualitative. Les variations de l’opacité selon la longueur d’onde permettent de sonder différentes couches de l’atmosphère. Hors, celles-ci vont émettre avec une intensité influencée

par la température locale ; une couche plus chaude aura un flux plus important qu’une couche plus froide. Ainsi, au centre d’une bande moléculaire, le flux observé proviendra des couches supérieures et inversement en dehors des bandes (donc couches inférieures). Pour le cas de HD 209458 b et le côté nuit de WASP-121 b, la température diminue en altitude, donc le flux sera moins élevé au centre de la bande d’eau. À l’opposé, pour le côté jour de WASP-121 b, le profil de température subit une inversion. La bande d’eau sondera donc une altitude avec une température plus élevée, se manifestant par une hausse du flux à ces longueurs d’onde.

### 2.1.3. Comparaison entre éclipses primaire et secondaire

Les deux formes d’éclipse ont une sensibilité différente aux propriétés de l’atmosphère et présentent chacune leurs dégénérescences. Leurs observations combinées permettent de lever certaines d’entre elles.

Premièrement, notons qu’à l’équation 2.1.4 (éclipse primaire) le rayon de la planète joue un rôle important. Ceci peut devenir problématique puisqu’il est difficile de différencier l’absorption due à la planète elle-même de celle de son atmosphère. En pratique, le rayon de la planète est associé à la pression à la base de l’atmosphère, par exemple 100 bar. Or, pour un rayon sur-estimé, le modèle présupera que les régions sondées auront une pression sur-évaluée, puisque la base de l’atmosphère sera surélevée. Le nombre d’absorbeurs devra donc être sous-évalué pour expliquer la faible amplitude du signal à ces pressions élevées. D’ailleurs, Griffith (2014) a démontré qu’une erreur de 1% sur le rayon d’une Jupiter chaude typique mène à des incertitudes d’un facteur 100 à 10 000 sur la dérivation des abondances relatives des espèces présentes dans l’atmosphère.

Deuxièmement, la géométrie particulière du transit force la lumière à parcourir une grande distance dans l’atmosphère. Une absorption relativement faible peut donc avoir malgré tout une épaisseur optique considérable (Fortney, 2005). Dans le cas d’une atmosphère sans nuage, c’est un avantage puisque certaines molécules ayant une faible signature optique seront détectables malgré cela. Cet effet est visible en comparant les spectres de transmission et d’émission. Par exemple, la bande d’absorption de l’eau à  $1,4 \mu\text{m}$  a une amplitude de  $\sim 0,025\%$  dans la figure 3 et de  $\sim 0.005\%$  dans la figure 4.

Cependant, dans le cas des nuages dont l’opacité est uniforme sur le spectre, la profondeur optique devient pratiquement infinie, donc complètement opaque. Ainsi, le rayon de la planète est pratiquement indiscernable de la région nuageuse ce qui amplifie l’incertitude sur la détermination des abondances relatives. Ce qui était un désavantage pour l’éclipse secondaire devient alors un avantage, puisque la géométrie amenuise l’effet des nuages. La figure 4 montre que la présence des nuages a très peu d’influence sur le spectre d’émission. Malgré que les courbes de condensation croisent les profils de température, leur profondeur optique,  $\tau_c$  se situe en dessous de 1 (voir encadré, figure gauche), ce qui peut être interprété comme un

milieu transparent. Le profil de température doit être un peu décalé en altitude, mais demeure largement à l'intérieur de la région de  $1\sigma$ . L'impact négligeable des nuages en émission est aussi dû au fait que la région sondée est beaucoup plus grande qu'en transmission (toute la face diurne de la planète). Or, comme les modèles de circulation atmosphérique le montrent (voir figure 9), l'atmosphère n'est pas homogène ; les nuages risquent alors d'être distribués de manière non-uniforme. Les couches plus profondes de l'atmosphère seraient donc accessibles.

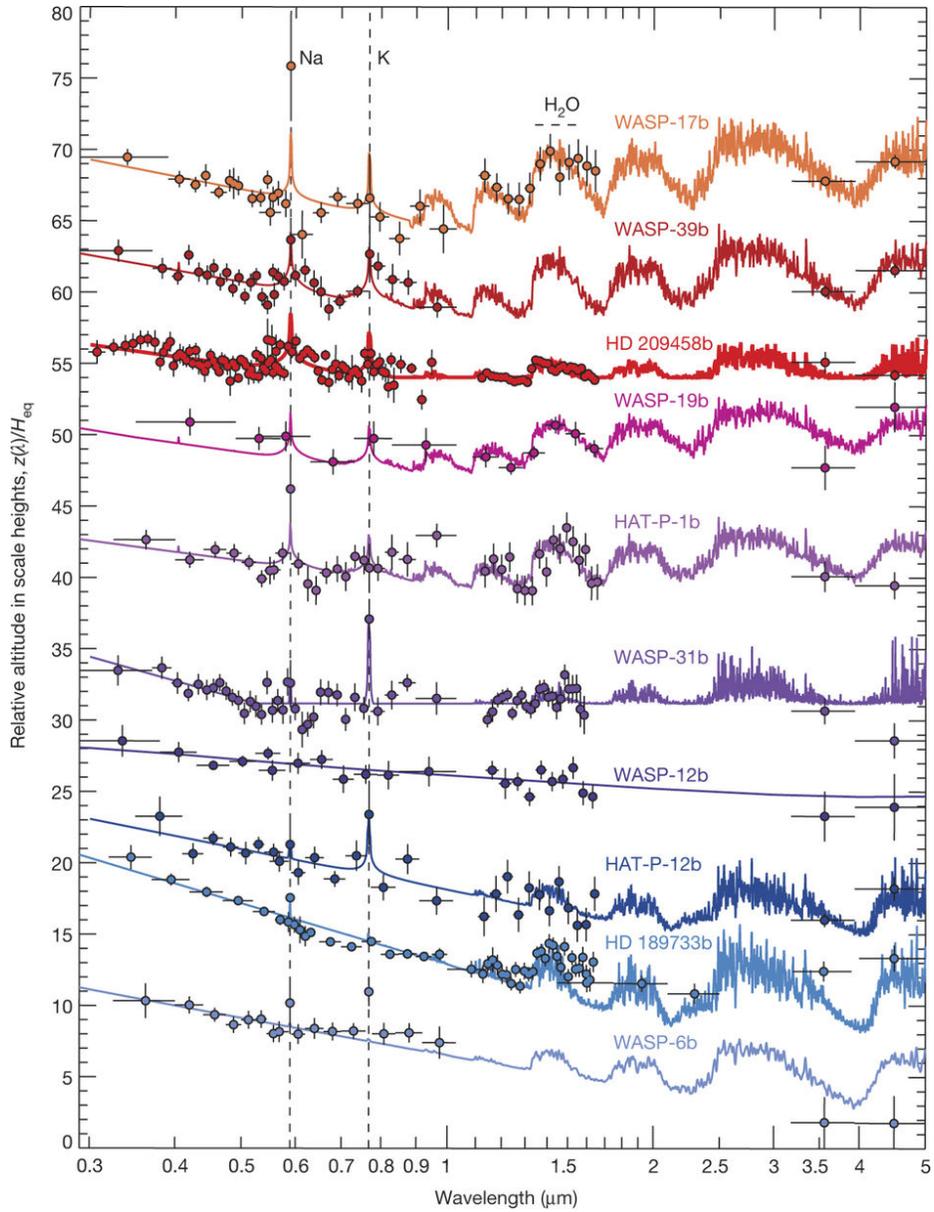
Troisièmement en ce qui a trait à la détermination des profils de température, l'éclipse secondaire est beaucoup plus performante. Encore une fois, la géométrie de l'éclipse primaire la rend sensible à une plus petite partie de l'atmosphère. C'est d'ailleurs une des raisons pour laquelle l'atmosphère isotherme est fréquemment utilisée comme approximation. Par contre, puisque les régions sondées en émission dépendent de la profondeur optique et donc de la concentration des absorbeurs à chacune des couches, il existe une dégénérescence entre la détermination des abondances et le profil de température. C'est d'ailleurs la principale cause de l'incertitude sur le profil de température à la figure 4.

L'analyse conjointe du spectre de transmission et d'émission permet de lever en partie ces dégénérescences. En effet, la sensibilité accrue du transit aux abondances des absorbeurs permet de mieux contraindre les profils de température obtenus avec l'occultation, alors que celle-ci permet de réduire l'erreur sur le rayon en contraignant les pressions sondées grâce au profil de température. Cette technique suppose cependant que les abondances observées en transit (donc dans le terminator) seront représentatives des celles observées en émission (donc du côté jour).

## 2.2. Spectrophotométrie à partir d'observatoires spatiaux

La spectrophotométrie à partir d'observatoires spatiaux occupe une place importante pour l'analyse de l'atmosphère d'exoplanètes. Ces observations offrent une stabilité exceptionnelle en se départissant des problèmes liés à l'atmosphère de la Terre comme la scintillation ou l'absorption tellurique. Cette dernière est particulièrement importante dans l'infrarouge où l'atmosphère de la Terre absorbe une grande partie de la lumière. On perd donc de larges plages d'observation. Les bandes d'absorption de l'eau vont d'ailleurs couper considérablement le flux observé, l'eau étant très présente dans notre propre atmosphère. Or, c'est une des molécules importante que l'on cherche à détecter. De plus, cette absorption peut varier en fonction des conditions d'observation.

Les observatoires spatiaux ont aussi l'avantage d'avoir moins de contraintes liées aux variations diurnes. *Spitzer*, par exemple, est en orbite héliocentrique et peut donc observer en continu pendant plusieurs jours. Hubble, quant à lui, fait le tour de la Terre périodiquement



**Figure 6** – Spectres de transmission de 10 Jupiter chaudes (Sing et al., 2016). Les modèles (Fortney et al., 2008) correspondant le mieux aux données sont représentés par les courbes de couleur. Les barres d’erreur indiquent l’incertitude de  $1\sigma$  sur la mesure (verticalement) et la bande spectrale (horizontalement). Les données de Hubble sont distribuées entre  $0.3$  et  $1.01 \mu\text{m}$  pour STIS et  $1.1$  à  $1.7 \mu\text{m}$  pour WFC3. Seulement les bandes à  $3.6$  et  $4.5 \mu\text{m}$  de *Spitzer*/IRAC sont utilisées dans ce cas-ci par souci d’uniformité. Les spectres sont décalés verticalement aux fins de la présentation. L’échelle verticale choisie ( $z(\lambda)/H_{eq}$ ) permet de comparer les spectres plus directement en donnant une indication des pressions sondées,  $H_{eq}$  étant la hauteur d’échelle à la température d’équilibre (voir sections 1.1.1 et 1.1.2) et  $z(\lambda)$  l’altitude apparente de l’atmosphère. Les bandes d’absorption du sodium, du potassium et de l’eau spécialement ciblées par les observations sont indiquées en haut de la figure.

toutes les 96 minutes environ. Pour observer des événements tels qu'un transit, il doit cumuler plusieurs orbites ; la fréquence des orbites est assez grande pour pouvoir échantillonner convenablement la majorité de ces événements.

La précision photométrique atteinte par ces instruments permet donc de mesurer précisément les variations de l'intensité du flux à chaque longueur d'onde par rapport à une valeur de référence (spectre hors éclipse ou transit). Ainsi, il est possible d'obtenir une mesure de la profondeur du signal qui peut être comparée directement aux équations 2.1.4 ou 2.1.5.

### 2.2.1. Hubble et *Spitzer*

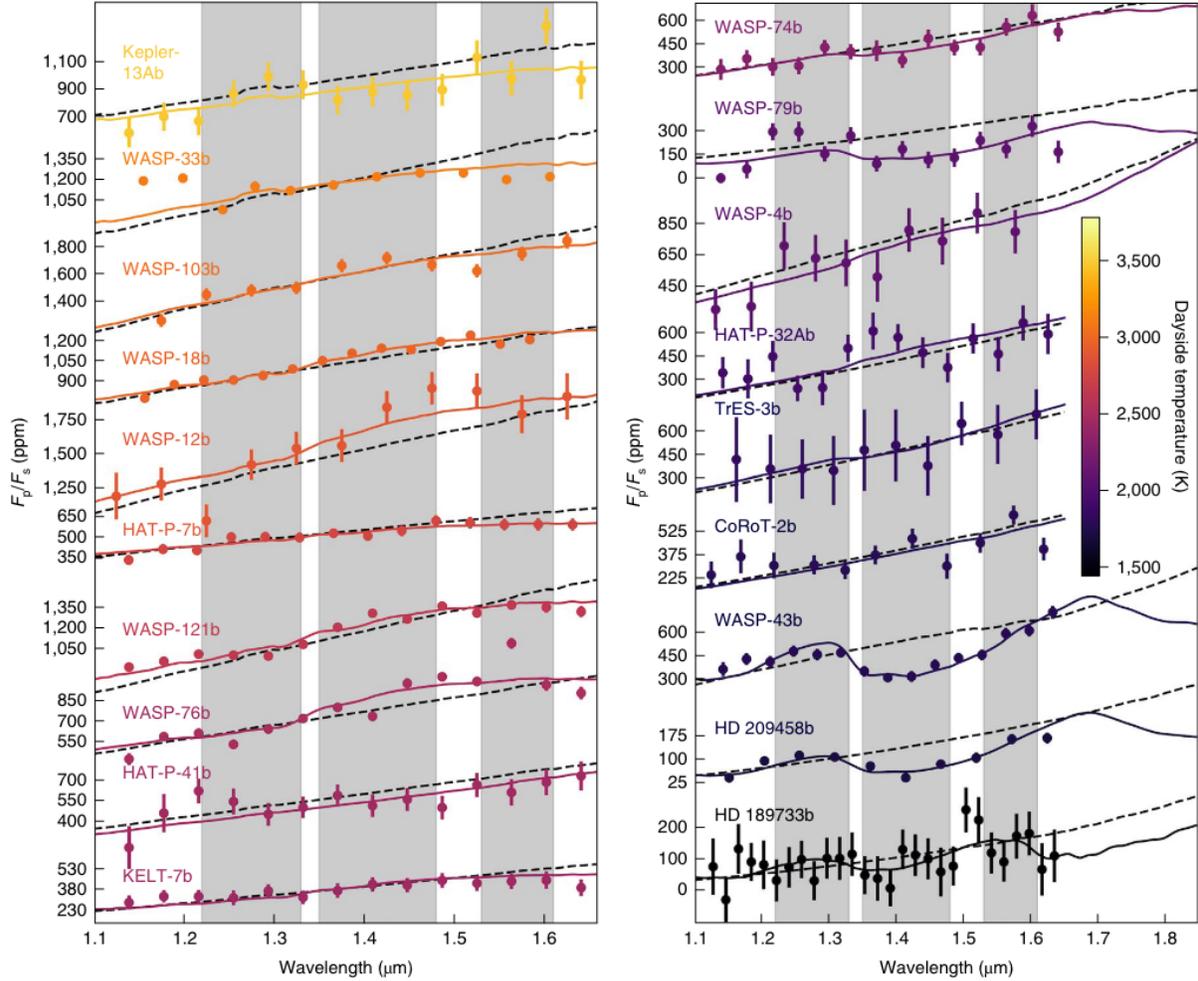
Le télescope spatial Hubble avec ses instruments WFC3 (*Wide Field Camera 3*) et STIS (*Spatial Telescope Imaging Spectrograph*) a démontré son efficacité au cours des dernières années à mesurer les spectres d'éclipses d'exoplanètes. La précision des spectres présentés aux figures 3, 4 et 6 en témoignent. Les principaux obstacles pour effectuer ces observations proviennent du fait que Hubble est en orbite rapproché autour de la Terre. Il est donc soumis à des différences de température considérables, selon qu'il se situe du côté jour ou du côté nuit. L'instrument NICMOS (*Near Infrared Camera and Multi-Object Spectrometer*) était utilisé avant WFC3 jusqu'à ce qu'il ne soit plus fonctionnel (2013). Cependant, les effets systématiques associés à cet instrument n'ont pas été assez bien caractérisés avant la fin de sa vie utile, plaçant l'analyse d'atmosphères d'exoplanètes à la limite de détection de cet instrument (Crouzet et al., 2012). C'est pourquoi les données de NICMOS sont très peu présentes dans la littérature subséquente.

Une fonctionnalité importante de WFC3 est le *spatial scanning mode* (McCullough & MacKenty, 2012). Elle est utilisée lors d'observations d'étoiles très brillantes et permet d'éviter la saturation des pixels en déplaçant le spectre sur le détecteur durant une acquisition d'image (sans trop affecter la résolution spectrale). Cette méthode a l'avantage de répartir le flux sur plusieurs pixels, ce qui diminue l'incertitude.

Cependant, les télescopes Hubble et *Spitzer* n'ont pas été conçus pour mesurer les atmosphères d'exoplanètes. Les biais observationnels causés par les détecteurs sont du même ordre que le signal recherché ( $\sim 0.01\%$  du signal de l'étoile pour les Jupiter chaudes). Dans le cas de l'instrument IRAC (*InfraRed Array Camera*) de *Spitzer*, les principaux problèmes proviennent du fait que les pixels du détecteur ont une sensibilité variable selon la position de l'étoile (Morales-Calderón et al., 2006) et qu'il est impossible de garder l'image de l'étoile exactement au même endroit. De plus la sensibilité du détecteur augmente en fonction du temps pour atteindre un plateau à cause de l'« emprisonnement de charge<sup>1</sup> ». Les investigations menées au cours des dernières années ont su réduire considérablement l'erreur sur les

---

1. Ce concept (Knutson et al., 2007b) se résume par l'idée que les électrons peuvent être absorbés par le détecteur sans contribuer au signal, à cause de certaines impuretés. Une fois que ces « prisons » sont pleines, le détecteur atteint sa pleine sensibilité.



**Figure 7** – Spectres d’émission de 19 Jupiter chaudes (Mansfield et al., 2021). Les cercles accompagnés de barres d’erreurs montrent les données prises par WFC3. La couleur représente la température du côté jour de la planète estimée à partir des données en comparant le flux observé ( $f_p$  de l’équation 2.1.5) avec celui d’un corps noir. Les courbes en couleur montrent le modèle offrant le meilleur accord avec les données alors que celles en tirets montrent l’équivalent pour un corps noir. La région ombragée centrale identifie la bande d’eau principale (à  $1,4 \mu\text{m}$ ) alors que les deux régions autour ( $1,22\text{--}1,33 \mu\text{m}$  et  $1,53\text{--}1,61 \mu\text{m}$ ) n’incluent pas (ou peu) de signal de l’eau.

données de IRAC. Malgré tout, leur robustesse est souvent remise en question (ex : Hansen et al. 2014) et des résultats substantiellement différents peuvent être obtenus par différents auteurs. Or, *Spitzer* offre les meilleures données disponibles dans la plage de longueurs d’onde qu’il couvre. L’analyse des atmosphères d’exoplanètes souffre donc d’une certaine dépendance aux techniques de réduction de données utilisées, en particulier pour la détermination des abondances chimiques et du profil de température. Dans le cas de HD 209458 b, par exemple, la réduction des données pouvait faire la différence entre un profil de température inversé ou non (Knutson et al., 2008; Diamond-Lowe et al., 2014; Evans et al., 2015). De plus, les

bandes spectrales sont très larges et chacune couvre le signal de différentes molécules comme le CO, le CO<sub>2</sub>, le H<sub>2</sub>O, le NH<sub>3</sub> et le CH<sub>4</sub> (ex : Pluriel et al., 2020; Baxter et al., 2021). Les paramètres orbitaux utilisés d’une visite à l’autre peuvent aussi biaiser les données (ex : Tsiaras et al., 2018; Yip et al., 2020). La variabilité due à l’étoile hôte est aussi une source d’erreur lorsque vient le temps de combiner des données de différentes époques (ex : Bruno et al., 2020). Finalement, tous ces biais complexifient la combinaison de données de différentes sources (STIS, WFC3, *Spitzer*) et peuvent avoir une influence considérable sur les propriétés physiques inférées par la comparaison avec des modèles (ex : Barstow et al., 2015).

Cette incohérence est moins présente dans le cas de Hubble puisque le flux est mieux distribué sur les pixels grâce au *spatial scanning mode*. Les variations de températures dues à l’orbite terrestre ont aussi un effet important, mais sont généralement bien modélisées (Stevenson & Fowler, 2019).

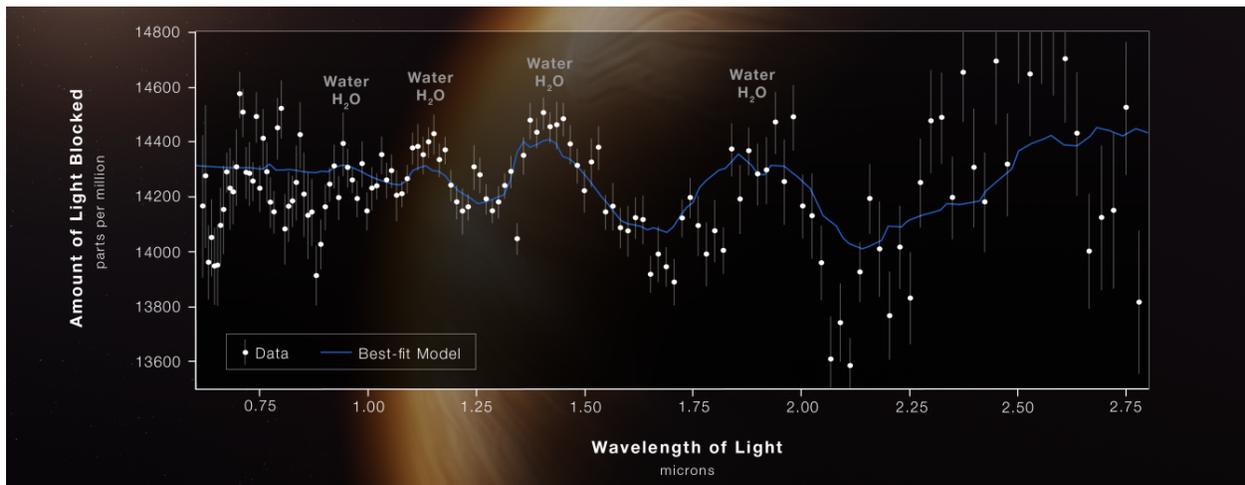
La figure 6 montre une sélection de transits de Jupiter chaudes obtenues à l’aide des télescopes spatiaux Hubble et *Spitzer*. On peut remarquer que les données de STIS sont utiles pour cibler les bandes d’absorption des alcalins (Na, K) et la pente due à la diffusion de Rayleigh (voir section 1.2.6). Bien que ce type d’absorption diminue selon  $\lambda^{-4}$ , la pente observée dans le spectre n’est pas nécessairement de cette ordre, comme le montre la diversité présente dans la figure 6. Elle est influencée, entre autres, par la distribution des particules dans l’atmosphère (voir la figure 1 de Barstow et al. 2017). Les données de WFC3 permettent quant à elles de cibler une des bandes d’absorption de l’eau. Celle-ci varie en amplitude d’un objet à l’autre, ce qui peut indiquer des abondances différentes. Comme mentionné à la section 2.1.1, cette variation peut aussi être causée par des nuages qui auraient pour effet de surélever la base du spectre. La présence de nuages à haute altitude peut donc être interprétée à tort comme une faible abondance d’eau. C’est ce qui est suspecté dans le cas de HD 209458 b, pour lequel l’absorption due à l’eau est très faible (Lecavelier Des Etangs et al., 2008).

Une étude similaire a aussi été produite par Mansfield et al. (2021) pour des spectres en émission. Les résultats pour 19 Jupiters chaudes est présenté à la figure 7. D’emblée, on peut remarquer que la bande d’eau à 1,4  $\mu\text{m}$  est beaucoup moins proéminente qu’en transmission, ce qui est attendu (comme souligné à la section 2.1.3). Cette bande se retrouve d’ailleurs sous différentes forme : en absorption (ex : WASP-43 b et HD 209458 b), en émission (ex : WASP-12 b et WASP-121 b) ou simplement absente (ex : WASP-18 b, WASP-103 b). Ceci met en évidence notamment la diversité de structures présentes dans les atmosphères de Jupiters chaudes. Les planètes pour lesquelles on ne peut pas distinguer la bande d’eau peuvent être expliquées par des profils isothermes, auquel cas le spectre devrait s’apparenter à celui d’un corps noir. Cet absence peut aussi être causée par une faible abondance d’eau ou alors par la présence d’autres absorbeurs comme le H<sup>-</sup>. Il est cependant difficile de départager ces deux scénarios, ce qui a mené dans certains cas à des conclusions très différentes sur la métallicité

et le rapport C/O. Dans le cas de WASP-18 b par exemple, Sheppard et al. (2017) obtinrent une métallicité très élevée ( $C/H = 283_{-138}^{+395}$  fois la valeur solaire) en négligeant la contribution du  $H^-$ . Or, une analyse subséquente (Arcangeli et al., 2018) montra qu'en l'incluant et en ajoutant l'effet de la dissociation de l'eau (voir section 1.2.2), il était possible d'expliquer les données avec des valeurs de métallicité et de C/O plus près des valeurs attendues.

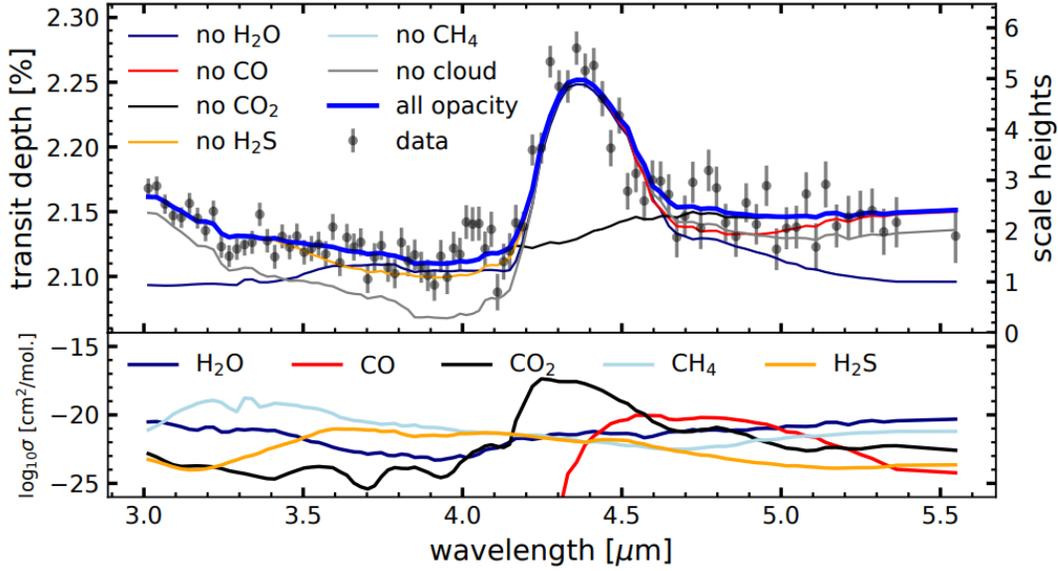
Le problème général provient principalement du manque de couverture en longueur d'onde pour pouvoir cibler d'autres bandes moléculaires (d'une même ou d'une différente molécule) et du manque de résolution spectrale (en particulier pour *Spitzer*). En effet, les valeurs obtenues pour la métallicité sont généralement basées sur une seule molécule, et parfois même sur seulement une ou deux des bandes photométriques *Spitzer*. Une plus large bande spectrale ainsi qu'une meilleure résolution permettrait notamment de lever (ou alléger) la dégénérescence entre les nuages et les abondances (ex : Barstow et al. 2013b) et celle entre le profil de température et les abondances (ex : Barstow et al. 2013a). C'est d'ailleurs un des objectifs du télescope spatial James Webb (JWST).

### 2.2.2. Télescope spatial James Webb



**Figure 8** – Premier spectre de transmission (publié) d'une exoplanète par JWST. Les données ont été prises avec le mode SOSS de l'instrument NIRISS et ciblent la géante gazeuse WASP-96 b. Le spectre présente plusieurs bandes d'absorption de l'eau identifiées dans la figure. Le modèle offrant la meilleure correspondance avec les observations est représenté par la courbe bleue. **Crédits : NASA, ESA, CSA, STScI**

Le télescope spatial James Webb (JWST) a été lancé le 25 décembre 2021. Sa mise en service est maintenant complétée et ce faisant, il est maintenant entièrement opérationnel. Ainsi, il offre à la communauté scientifique une vision plus large et précise des atmosphères d'exoplanètes grâce aux leçons tirées de Hubble et *Spitzer*. Afin d'éviter les grandes variations de températures influençant les détecteurs, JWST est positionné au point de Lagrange L2.



**Figure 9** – Première détection claire de CO<sub>2</sub> dans l’atmosphère d’une exoplanète. Les données ont été prises avec le mode BOTS de l’instrument NIRSpec et ciblent la géante gazeuse WASP-39 b. Le PRISM (prisme) a été choisi comme élément dispersif afin de couvrir une large bande simultanément (0,5–5,5  $\mu\text{m}$ ). Cependant, seulement la partie entre 3 et 5,5  $\mu\text{m}$  est présentée ici ; les longueurs d’onde plus courtes sont sujettes à la saturation et demandent une analyse plus approfondie. La bande de CO<sub>2</sub> à 4,3  $\mu\text{m}$  se démarque largement, avec un niveau de détection de  $26\sigma$ . Des modèles excluant différentes contributions sont présentés par les courbes en couleurs. Les sections efficaces associées aux molécules sont montrées dans la fenêtre du bas. La figure est tirée de The JWST Transiting Exoplanet Community Early Release Science Team et al. (2022).

De plus, un bouclier thermique lui permettra de demeurer extrêmement stable tout au long de sa vie utile en le refroidissant passivement pour le maintenir à des températures assez basses pour ne pas émettre significativement dans l’infrarouge moyen. Son miroir de 6.5 m permet une précision inégalée dans l’astronomie spatiale.

JWST accueille à son bord quatre instruments scientifiques, soit NIRISS (*Near InfraRed Imager and Slitless Spectrograph*), NIRSpec (*Near InfraRed Spectrograph*), NIRCам (*Near InfraRed Camera*) et MIRI (*Mid InfraRed Imager*). Chacun d’eux est capable d’étudier les atmosphères d’exoplanètes grâce à sa capacité spectroscopique (Beichman et al., 2014). Leur performance respective face à la spectroscopie de transit ou d’éclipse a été évaluée par Doyon et al. (2012), Ferruit et al. (2014); Batalha et al. (2015), Greene et al. (2007) et Kendrew et al. (2015). La couverture totale en longueurs d’onde obtenue en combinant les divers instruments s’étend de 0.6 à 28  $\mu\text{m}$ . Cependant, les erreurs systématiques augmentent considérablement après  $\sim 11 \mu\text{m}$  pour MIRI (l’instrument couvrant la plage de 5 à 28  $\mu\text{m}$ ).

NIRISS, NIRCам et MIRI offrent un mode de spectroscopie sans fente comme dans le cas de Hubble. NIRSpec est plus conventionnel avec un mode d’observation par spectroscopie

classique (avec fente). Dans le mode SOSS (*Single Object Slitless Spectroscopy*) de NIRISS la lumière est volontairement dé-focalisée afin de répartir le signal à chaque longueur d’onde sur  $\sim 25$  pixels (plus de détails au chapitre 4). Cette spécificité inspirée du *Spatial scanning mode* de WFC3 permet en particulier d’observer des cibles relativement brillantes (J mag  $\gtrsim 7,3^2$ ).

Parmi les meilleures observations à ce jour, certaines ont su atteindre un bruit de l’ordre de 30 ppm pour Hubble ( $R = \lambda/\delta\lambda \sim 70$ ) en combinant 12 transits (Kreidberg et al., 2014) et de 65 ppm pour *Spitzer* avec un seul transit (Knutson et al., 2009). Pour NIRISS (SOSS, 1–2.8  $\mu\text{m}$ ), NIRCcam (2.5–5  $\mu\text{m}$ ) et MIRI (5–11  $\mu\text{m}$ ), Greene et al. (2016) évalue le bruit instrumental associé à un seul transit à, respectivement, 20 ppm, 30 ppm et 50 ppm.

La toute première observation d’un transit avec JWST a d’ailleurs été présentée en juillet 2022, montrant le spectre de transmission de WASP-96 b pris avec NIRISS/SOSS. Celui-ci est présenté à la figure 8 et montre le potentiel de ce mode d’observation avec sa large couverture spectrale simultanée, donnant accès à plusieurs bandes moléculaires en une simple visite.

Plus récemment, la première détection de CO<sub>2</sub> sans ambiguïté a été faite grâce au prisme de NIRSpec, dans le cadre du program *Early Release Science* DD-ERS 1366 (Batalha et al., 2017; Bean et al., 2018). Cette découverte découle de l’observation du spectre de transmission de WASP-39 b, présenté à la figure 9. On peut y voir la signature proéminente du CO<sub>2</sub> à 4,3  $\mu\text{m}$ . On remarque aussi les différentes molécules auxquelles cet instrument donne accès. Notons qu’auparavant, les seules données photométriques précises disponibles, soit celles de *Spitzer* couvraient l’entièreté de ce domaine de longueurs d’onde avec seulement deux points.

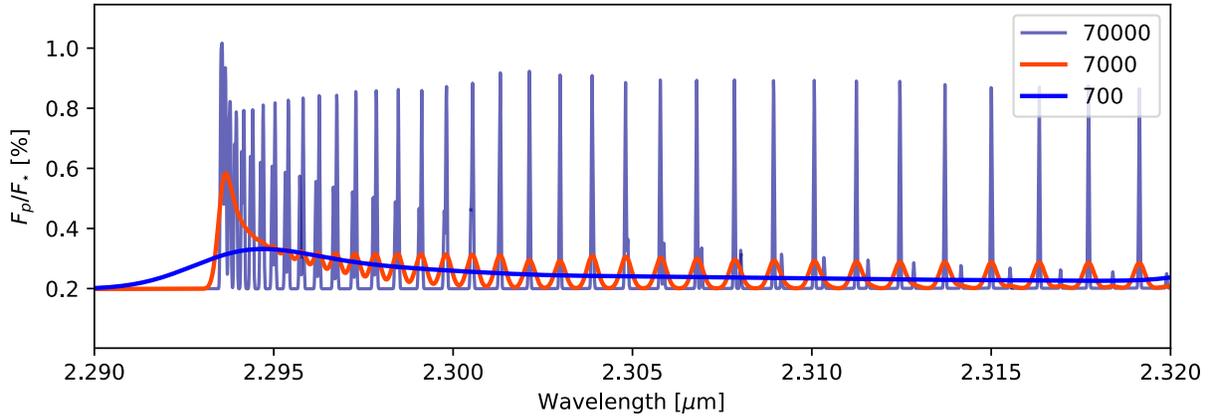
En somme, ces deux résultats illustrent la capacité inégalée de JWST à étudier l’atmosphère des exoplanètes. La sensibilité, la résolution spectrale ainsi que la couverture en longueurs d’onde permettront notamment de mieux contraindre les différentes abondances et la structure en température, levant en partie les dégénérescences affligeant les instruments précurseurs.

## 2.3. Spectroscopie à haute résolution

Il existe une méthode complémentaire aux observations spectrophotométriques permettant d’analyser l’atmosphère d’exoplanètes : la spectroscopie à haute résolution (HRS de l’anglais *High Resolution Spectroscopy*) ou haute dispersion (HDS de l’anglais *High Dispersion Spectroscopy*). Cette technique utilise des spectrographes de type « échelle », offrant une résolution spectrale relativement grande ( $\lambda/\Delta\lambda \gtrsim 50,000$ ) pour résoudre les raies individuelles qui forment les bandes moléculaires observées à plus faible résolution<sup>3</sup>. La figure 10 montre trois exemples de résolution appliquées à une des bandes moléculaires de CO. La

2. <https://jwst-docs.stsci.edu/jwst-near-infrared-imager-and-slitless-spectrograph/niriss-performance/niriss-bright-limits>

3. Toutes les figures montrant des spectres présentées dans les chapitres précédents sont considérées comme étant à basse résolution.



**Figure 10** – Bande moléculaire de CO à différentes résolutions spectrales, pour un modèle de spectre d’émission de WASP-33. Trois résolutions différentes sont montrées, soit  $\lambda/\Delta\lambda = 70\,000$ , 7000 et 700.

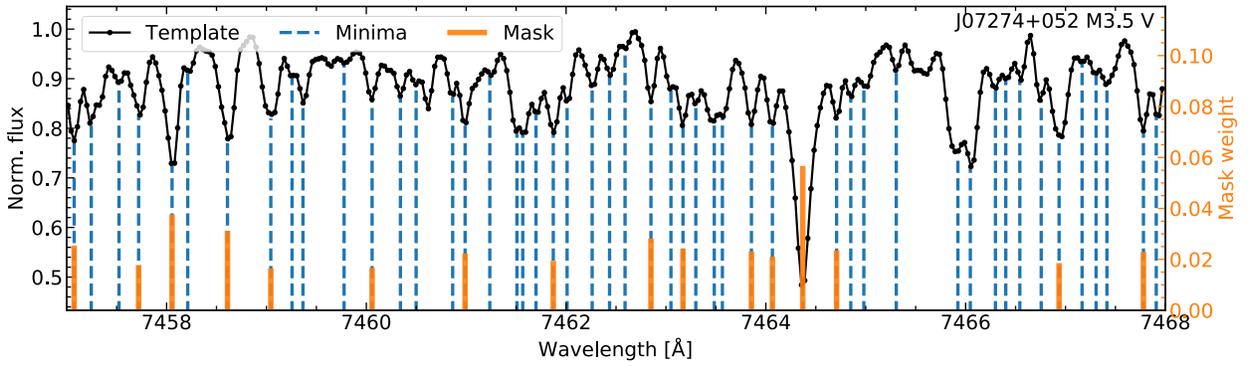
plus basse (700) correspond à celle du mode NIRISS/SOSS et, à l’opposé, celle de 70 000 correspond à l’instrument SPIRou dont il sera question dans le chapitre 3. Ainsi, les multiples raies composant la bande de CO ne représentent qu’une simple bosse à la résolution de NIRISS/SOSS. Ceci est vrai pour la majorité des molécules dans l’infrarouge, comme celles présentées à la figure 9, où chacune des sections efficaces (fenêtre du bas) dissimulent en réalité un riche ensemble de raies individuelles.

Cependant, ce gain en résolution se fait généralement aux dépens de la précision photométrique. En effet, les spectrographes de ce type sont généralement beaucoup plus gros et massifs. Il est donc présentement impossible de les intégrer à des observatoires spatiaux comme JWST ou Hubble. Ainsi, les effets liés à l’atmosphère terrestre viennent miner leur sensibilité. De plus, il est nécessaire d’isoler préalablement le flux de l’objet ciblé avant de pouvoir disperser la lumière, contrairement au cas des spectrographes sans fente comme NIRISS/SOSS ou Hubble/WFC3. Une fibre optique ou une fente est généralement utilisée à cet effet, créant ainsi ce qu’on appelle une « ouverture ». De ce fait, les décalages dans le positionnement de l’objet dus notamment à la scintillation atmosphérique vont engendrer des fluctuations additionnelles causées par les variations du flux entrant et sortant de l’ouverture. Il est donc impossible d’isoler le signal de l’atmosphère d’une planète en se basant sur une précision photométrique relative, comme montré précédemment (ex : section 2.2).

### 2.3.1. Méthode de vitesse radiale

La HDS profite plutôt du mouvement orbital de la planète autour de son étoile afin de départager les deux signaux. En effet, à ces résolutions, il est possible de mesurer le décalage de la lumière dû à l’effet Doppler en analysant les variations globales d’un ensemble de

raies spectrales. Le spectre d'un objet sera ainsi poussé vers le bleu lorsqu'il se rapproche d'un observateur le long de la ligne de visée (axe radial), et vers le rouge lorsqu'il s'éloigne. La technique de détection d'exoplanètes par vitesse radiale est basée sur ce phénomène et consiste à mesurer le mouvement infime d'une étoile autour du centre de masse du système. Celui-ci est généralement en dessous d'une centaine de mètres par seconde, ce qui correspond à un décalage en longueur d'onde équivalent à une petite fraction ( $\sim 1/1000$ ) de la largeur des raies observables. Cependant, en combinant l'information globale de l'ensemble du spectre en utilisant des techniques de corrélation croisée avec un masque binaire ou un patron du spectre de l'étoile (en anglais *template matching*), une telle précision devient atteignable.

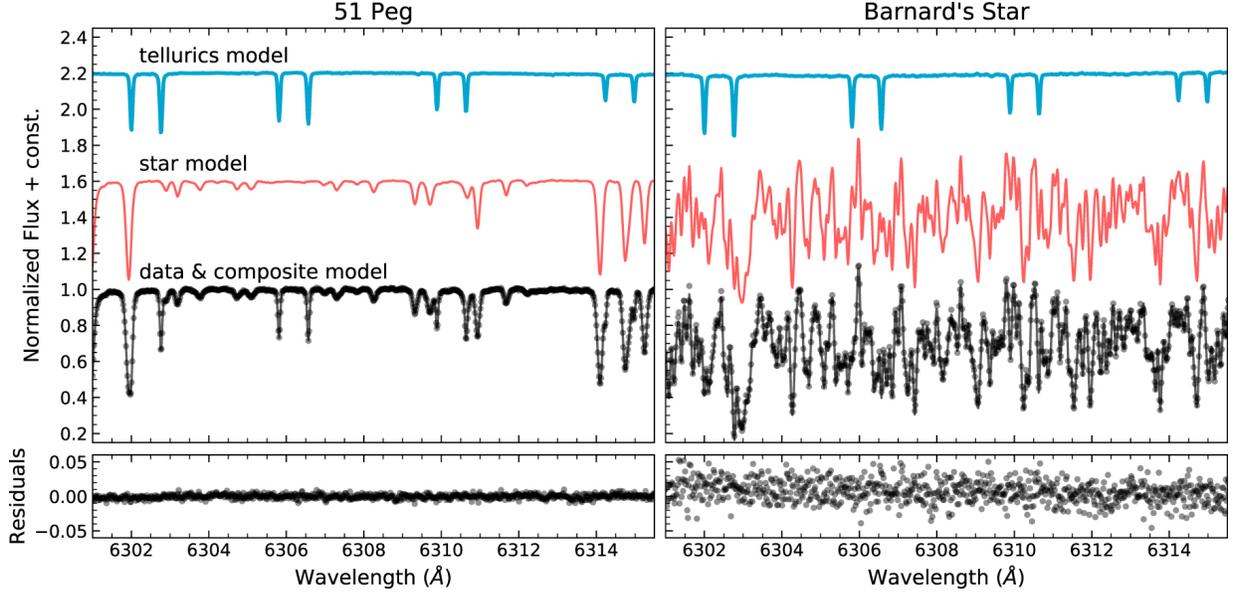


**Figure 11** – Exemple montrant la technique de corrélation croisée en utilisant un masque binaire (traits orangés) tiré de Lafarga et al. (2020). Ici, le masque est déterminé en identifiant les raies d'absorptions (traits bleus discontinus) à partir d'un patron du spectre de l'étoile (courbe noire) obtenu en combinant plusieurs observations. Les raies les plus proéminentes sont conservées et un poids leur est associé selon leur intensité (échelle sur l'axe de droite).

Les deux techniques sont représentées aux figures 11 et 12 (patron). Dans les deux situations, le principe est le même, soit d'abord de déterminer une fonction analytique, respectivement un masque binaire et un patron de l'étoile. Ensuite, cette fonction (ici notée  $m(\lambda)$ ) est comparée aux données en appliquant un décalage Doppler relativiste le long de l'axe des longueurs d'onde selon la transformation

$$\lambda \mapsto \lambda \sqrt{\frac{1 - \beta}{1 + \beta}}. \quad (2.3.1)$$

Ici,  $\lambda$  est la longueur d'onde et  $\beta = v/c$  où  $v$  est la vitesse de la cible et  $c$  est la vitesse de la lumière. Pour trouver la vitesse associée à un spectre, il suffit donc de trouver le décalage correspondant le mieux aux données. Pour ce faire, plusieurs méthodes mathématiques existent. La manière la plus classique consiste à déterminer une fonction de corrélation croisée (CCF de l'anglais *Cross-Correlation Function*) pour chaque observation d'un objet donné. La CCF



**Figure 12** – Exemple montrant la technique de corrélation croisée en utilisant un patron de l'étoile (courbes rouges) combiné à un modèle du spectre d'absorption tellurique (courbes bleues), constituant ainsi la courbe en noir. La figure, tirée de Bedell et al. (2019), montre la même région du spectre de l'étoile 51 Peg et de l'étoile de Barnard. Les modèles de l'étoile et des telluriques sont déterminés à même les données. Le résidu entre les données et le modèle offrant le meilleur accord est présenté dans la fenêtre du bas.

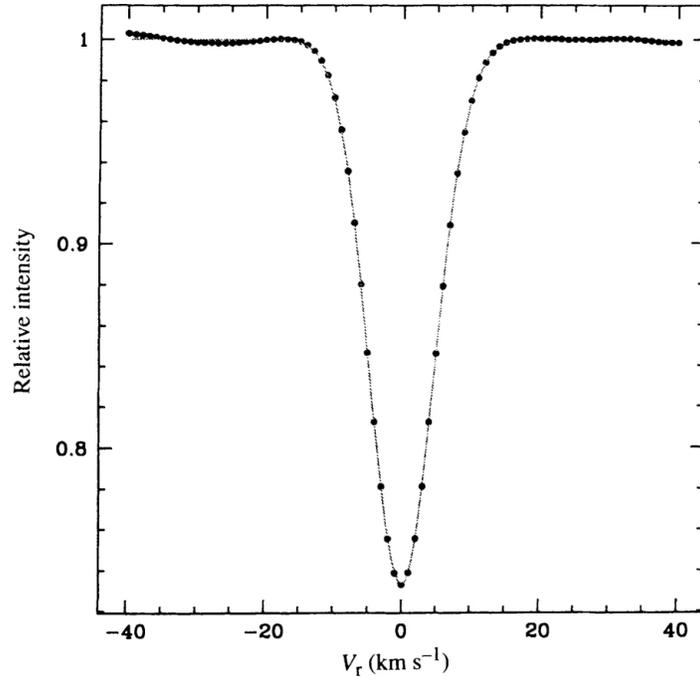
dépend donc de la vitesse  $v$  et s'exprime généralement sous la forme suivante :

$$\text{CCF}(v) = \sum_i w_i f_i m_i(v). \quad (2.3.2)$$

La somme est comprise sur les longueurs d'onde  $\lambda_i$  d'un spectre observé  $f_i$  (points noirs dans les figures 11 et 12),  $m_i$  représente la fonction  $m(\lambda)$  évaluée aux longueurs d'onde transformées selon l'équation 2.3.1 et  $w_i$  est un poids associé aux données, parfois donné par  $(\text{incertitude})_i^{-2}$  ou simplement l'unité. La composante radiale de la vitesse de l'étoile est alors donnée par le « centre » de la CCF, pouvant être déterminé de plusieurs manières (ex : fonction Gaussienne).

En suivant les variations de la vitesse radiale ainsi obtenue, il est possible de mesurer les paramètres de l'orbite de l'étoile autour du centre de masse d'un potentiel système planétaire. Ce mouvement orbital peut être relativement complexe. Cependant, dans le cas des Jupiter chaudes, les orbites sont généralement circulaires à cause des forces de marée intenses. L'évolution de la vitesse radiale en fonction du temps peut donc être paramétrée simplement par l'équation

$$v_{\text{orb}}(t) = -K_{\star} \sin 2\pi\phi(t), \quad (2.3.3)$$



**Figure 13** – Exemple de CCF pour l'étoile 51 Peg, tiré de Mayor & Queloz (1995). Chaque point montre la valeur de la CCF calculée à un décalage Doppler associé à une vitesse  $v$  (axe des abscisses). La vitesse mesurée est donnée par le centre d'une fonction Gaussienne offrant le meilleur accord avec la CCF (courbe en noir).

où  $\phi$  est la phase de la planète, définie comme

$$\phi \equiv \frac{t - t_0}{P} \pmod{1}. \quad (2.3.4)$$

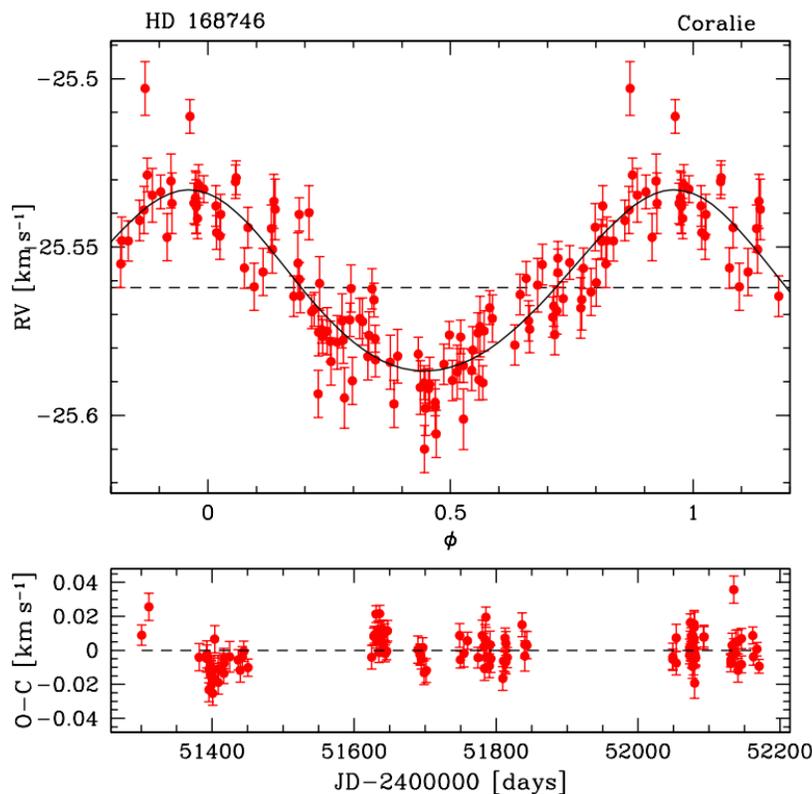
Le facteur  $K_\star$  est la demi-amplitude associée à l'étoile, donnée par

$$K_\star \equiv \left( \frac{2\pi G}{P} \right)^{1/3} \frac{M_p \sin i}{(M_\star + M_p)^{2/3}}. \quad (2.3.5)$$

Ici,  $G$  est la constante gravitationnelle,  $P$  est la période orbitale,  $M_\star$  et  $M_p$  sont les masses de l'étoile ainsi que de la planète et  $i$  est l'inclinaison du plan de l'orbite par rapport à un observateur. Il est donc possible de paramétrer une orbite circulaire avec seulement deux variables, soit  $K$  et  $P$ .

La phase  $\phi$  mérite une attention particulière. Celle-ci indique la position d'un objet le long de son orbite et se situe dans l'intervalle  $[0; 1[$ . Dans le cas général d'une orbite elliptique,  $t_0$  est défini comme le point où la planète est le plus près de son étoile, soit le périastre, donné par  $t_{\text{Peri}}$ . Cependant, dans une orbite circulaire, ce point ne peut être déterminé puisque la distance entre l'étoile et la planète demeure constante. Ainsi, dans ce cas particulier,  $t_0$  correspond au moment auquel la planète fait dos à l'observateur, c'est-à-dire lorsqu'elle

montre son côté nuit ou au centre du transit (pour un système éclipçant). Cette seconde définition de la phase est d'ailleurs souvent utilisée pour représenter des courbes de lumière, même pour des orbites elliptiques, auquel cas elle ne doit pas être confondue avec la phase orbitale de l'équation 2.3.3.

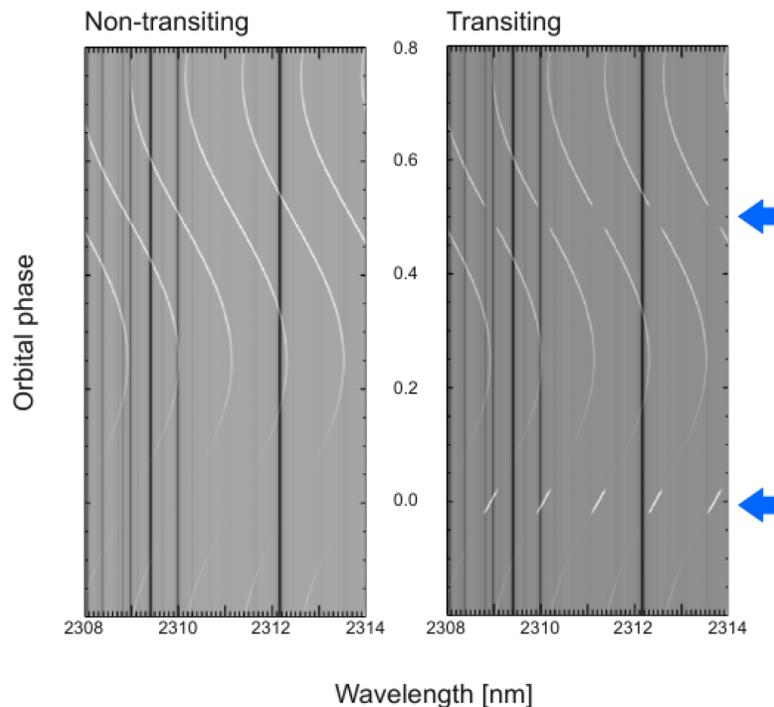


**Figure 14** – Exemple de courbe de vitesse radiale (RV) pour l'étoile HD 168746 tirée de Pepe et al. (2002). Les mesures prises avec l'instrument Coralie sont montrées en rouge avec leur incertitude. La fenêtre du haut montre la vitesse radiale en fonction de la phase. Le modèle d'orbite offrant le meilleur accord avec les données y est représenté par la courbe noire continue. La vitesse du centre de masse du système (vitesse systémique) est montrée par la droite horizontale discontinue. La fenêtre du bas présente le résidu (*Observed-Computed*) entre le modèle et les observations, en fonction du temps cette fois.

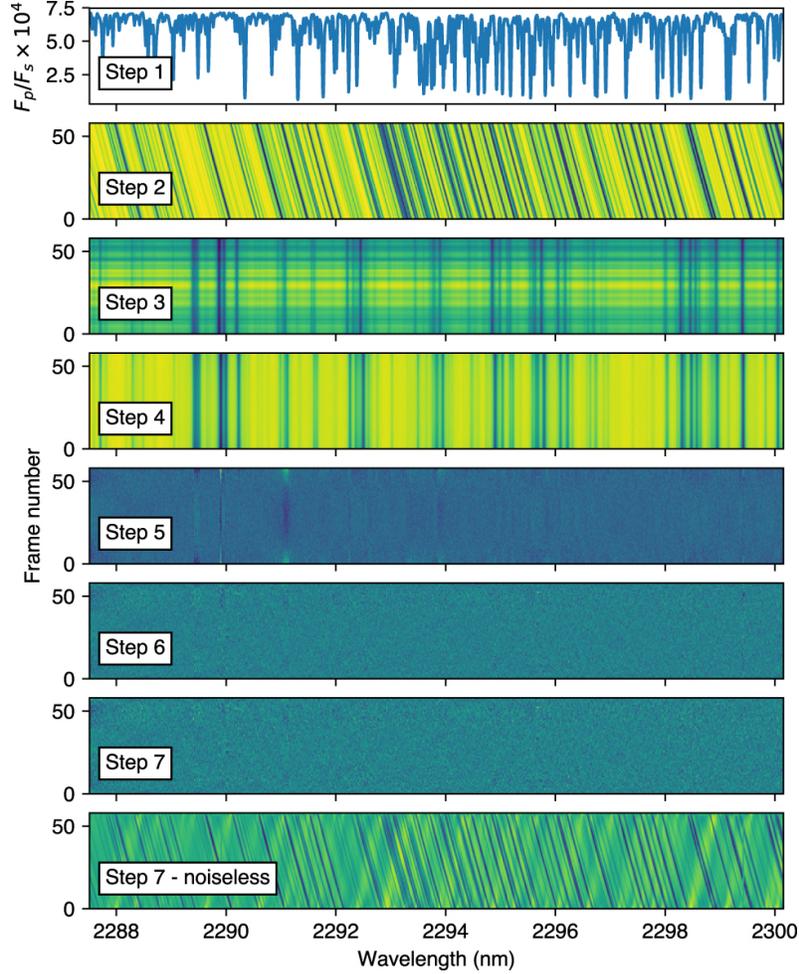
Un exemple de mouvement orbital quasi circulaire est montré à la figure 14. Une fois les données exprimées selon la phase, il devient possible de distinguer la fonction sinusoïdale. Notons ici que la phase orbitale est basée sur le périastre puisque le modèle favorise une faible excentricité (donc une orbite elliptique). Ainsi, la phase zéro se situe près du point où la planète est en quadrature et non au centre du transit (si présent). Ce dernier se situerait plutôt à  $\phi \approx 0,75$  dans cet exemple. Par ailleurs, une constante, soit la vitesse systémique ( $v_{\text{sys}}$ ), doit aussi être ajoutée au modèle de la vitesse orbitale afin de tenir compte du déplacement du système planétaire. Une autre composante n'est pas mentionnée dans la

figure, soit la vitesse radiale du barycentre de la Terre (BERV, de l'anglais *Barycentric Earth Radial Velocity*) due au mouvement orbital terrestre.

Ces principes utilisés pour caractériser le mouvement des étoiles peuvent aussi l'être dans le cas des exoplanètes, à la différence que, d'une part, le signal est beaucoup plus faible comparé à celui de l'étoile et que, d'autre part, le mouvement de la planète est considérablement plus grand. La figure 15 illustre bien la variation du spectre de la planète par rapport à celui de l'étoile. C'est cette caractéristique qui sera exploitée afin d'isoler le signal planétaire.



**Figure 15** – Schéma de la variation du signal atmosphérique d'une exoplanète par effet Doppler due à son orbite (tiré de Schwarz et al. 2013). L'axe vertical est donné par la phase orbitale, avec une valeur nulle correspondant au moment où le côté nuit de la planète fait face à l'observateur (transit dans un système éclipsant) et une phase de 0,5 correspondant au côté jour (éclipse secondaire). Les éclipses sont d'ailleurs identifiées par les flèches bleues. Le signal représentant des raies de la planète suit les courbes en blanc. En comparaison, celui de l'étoile et des raies telluriques est statique, comme le montre les raies sombres. Notons que le signal de la planète ne devrait être visible lors de réelles observations, étant donné qu'il est plusieurs ordres de grandeur moins important. L'effet Doppler est facilement discernable dans le cas de la planète, avec des raies décalées vers le rouge lorsqu'elle s'éloigne de l'observateur (phase entre 0 et 0,5) et vers le bleu lorsqu'elle se rapproche (phase  $> 0,5$ ). L'intensité du signal varie elle aussi, culminant lorsque le côté jour est entièrement visible dans le cas du système non-éclipsant (*non-transiting*). Dans l'autre cas, le signal disparaît durant l'éclipse secondaire et s'amplifie momentanément durant le transit.



**Figure 16** – Étapes de réduction de données à haute résolution pour isoler le signal planétaire. Dans cet exemple tiré de Brogi & Line (2019), les données sont simulées à partir de vraies observations de l’instrument CRIRES ciblant la planète HD 209458 b (Schwarz et al., 2015; Brogi et al., 2017). La fenêtre du haut (*Step 1*) montre le spectre d’émission synthétique qui sera injecté à différentes vitesses radiales (*Step 2*). La troisième fenêtre présente les données modélisées « brutes », incluant les variations du flux observé dues au changement de masse d’air et aux fluctuations dans l’atmosphère terrestre formant les stries horizontales. On peut aussi y distinguer des stries horizontales provenant de l’absorption tellurique ou de raies stellaires. À la quatrième fenêtre, chaque exposition a été normalisée par sa valeur médiane. Les expositions sont alors combinées afin d’obtenir un spectre de référence, composé principalement du signal de l’étoile hôte et de l’atmosphère terrestre. Celui-ci est ensuite utilisé afin de retirer le maximum de sa contribution à chaque exposition, laissant principalement des résidus de telluriques encore visibles (*Step 5*). Ces derniers sont corrigés en appliquant un traitement additionnel aux données (*Step 6*) différant selon les auteurs. La septième fenêtre ne montre pas d’étape de réduction supplémentaire ; elle ne fait qu’exclure les colonnes ayant une trop grande dispersion. Finalement, la huitième fenêtre (*Step 7 - noiseless*) illustre l’effet des étapes de réduction sur le spectre injecté (*Step 2*).

### 2.3.2. Extraction du signal planétaire à haute résolution

Puisque le signal de la planète est de plusieurs ordres de grandeur plus faible que la précision des spectres, un traitement spécifique doit donc être appliqué aux données. Le but étant de retirer les signaux dûs aux raies de l'étoile, à l'absorption telluriques et à de potentielles erreurs systématiques non-corrigées. La méthode la plus répandue consiste à prendre des observations en série temporelle continue. Ainsi, les étapes de réductions subséquentes seront basée sur l'hypothèse que les signaux autres que celui de la planète n'auront que d'infimes variations en longueur d'onde tout au long de la série d'observations.

Les étapes typiques de réduction sont présentées à la figure 16. On peut y remarquer d'emblée l'ampleur des décalages causés par le mouvement orbital de la planète (dans ce cas, HD 209458 b), avec des raies d'absorption se déplaçant sur plusieurs canaux de longueur d'onde (*Step 2*). Ensuite, l'étape 3 (*Step 3*) illustre la mauvaise précision photométrique des données, compensée en normalisant chaque spectre par rapport à sa valeur médiane (*Step 4*). Un patron du spectre de l'étoile et des telluriques peut alors être créé en combinant chaque observation, tout en supposant que leur signal est quasi-statique<sup>4</sup>, c'est-à-dire qu'elle ne varieront pratiquement pas en longueur d'onde (voir *Step 4*).

Le produit final de cette réduction (*Step 7*) est dominé par le bruit. Malgré cela, le signal de l'atmosphère de la planète demeure bel et bien présent, comme le montre la fenêtre du bas. Cependant, en la comparant avec la deuxième étape, on peut voir que le processus vient altérer ce signal. Ainsi, pour obtenir une comparaison fidèle entre les données réduites et un modèle, ce dernier doit être injecté dans les données brutes (*Step 3*) et être ensuite soumis au même processus.

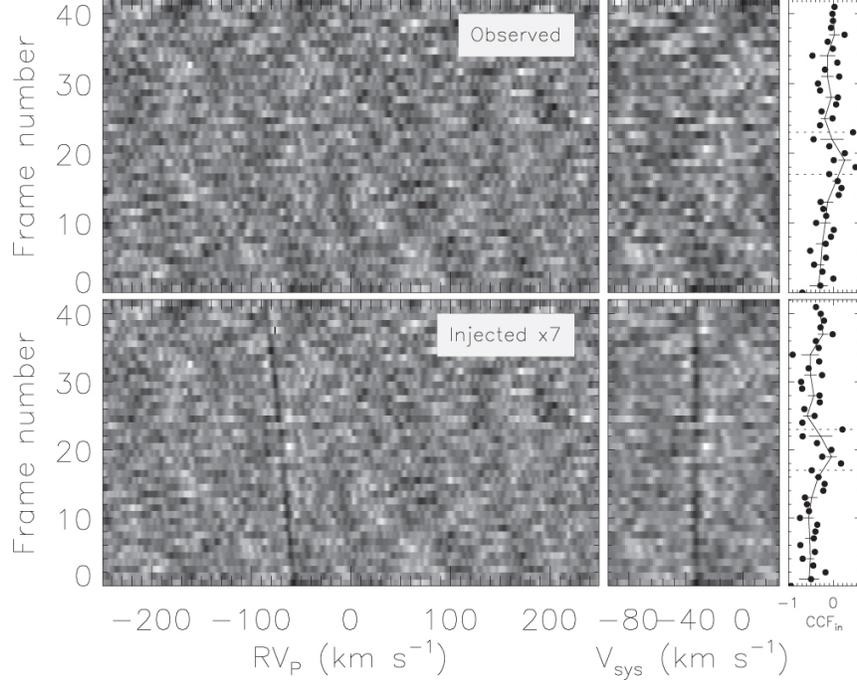
Cette dernière considération n'est cependant pas toujours prise en compte. En effet, celle-ci n'a été introduite que très récemment. Les travaux antérieurs se basaient donc sur les méthodes de vitesse radiale classiques en utilisant directement un patron de spectre d'émission ou un masque binaire. D'ailleurs, ces méthodes sont encore communément utilisées pour détecter la présence de molécules dans l'atmosphère d'exoplanètes, malgré qu'il soit plus difficile d'en faire une analyse quantitative<sup>5</sup>.

Dans les deux situations, la recherche du signal de l'atmosphère peut se faire de manière similaire. Dans le cas classique on calcule typiquement la CCF (équation 2.3.2) à chaque exposition. Comme l'illustre la figure 17, cette étape n'est généralement pas suffisante pour faire ressortir le signal de la planète au dessus du niveau de bruit. Cependant, en combinant chaque CCF dans le référentiel au repos de la planète (fenêtre du centre), il est possible

---

4. Certains travaux, dont l'article qui sera présenté au chapitre 3, prennent en compte le faible mouvement radial du spectre de l'étoile et font donc une distinction entre ce dernier et l'absorption tellurique ou d'autres erreurs systématiques.

5. Certaines méthodes ont été proposée afin de quantifier le degré d'accord pour différents modèles en se basant sur la CCF, comme Brogi et al. (2013), Brogi et al. (2016) et Pino et al. (2020)

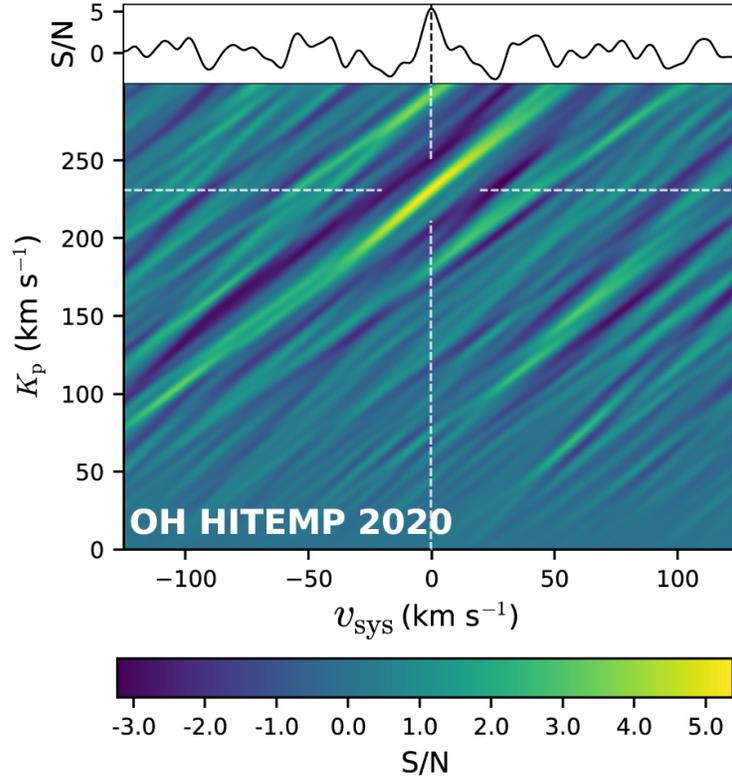


**Figure 17** – Série temporelle de CCF pour la planète 51 Peg b (figure tirée de Birkby et al. 2017). L’intensité des CCF est donnée par les tons de gris. Les observations sont présentées dans la rangée du haut de la figure et celles du bas montrent les mêmes observations avec le signal attendu de la planète injecté avec une amplitude 7 fois plus grande. Les fenêtres de gauche et au centre montrent respectivement les fonctions dans le référentiel de l’observateur et dans le référentiel corrigé pour le mouvement orbital de la planète ( $v_{\text{orb}}$ ). Pour le signal injecté, on peut clairement distinguer le signal de la planète. Dans le référentiel corrigé, celui-ci est aligné le long d’une droite centrée sur la vitesse systémique attendue ( $v_{\text{sys}} = -33,25 \text{ km s}^{-1}$ ). Les fenêtres de droite montrent la valeur moyenne pour trois pixels centrés autour de  $v_{\text{sys}} = -33 \text{ km s}^{-1}$  (points noirs). La ligne noire présente la même valeur regroupée sur 3 expositions.

d’améliorer une fois de plus le rapport signal sur bruit, pouvant mener à une détection de l’atmosphère. Cette procédure peut être répétée pour différentes valeurs du paramètre orbital  $K_P$ <sup>6</sup> afin de quantifier le niveau de détection d’un modèle d’atmosphère donné. Notons que la majorité des cas de détection obtenus à ce jour concernent des Jupiter chaudes avec des orbites circulaires. La demi-amplitude est donc suffisante pour les paramétrer, la période de révolution étant généralement bien contrainte par d’autres méthodes (ex : transit, vitesse radiale).

La figure 18 présente un exemple de carte de détection ainsi obtenue pour la UHJ WASP-33 b. Loin de la demi-amplitude attendue, les CCF combinées ne montrent aucun pic distinctif. Cependant, pour des valeurs de  $K_P$  près de  $230,9 \text{ km s}^{-1}$  (tirets blancs), les CCF

6.  $K_P$  est défini de la même manière qu’à l’équation 2.3.5 en interchangeant les indices « p » et « \* ». Il en va de même pour l’équation de la vitesse orbitale (2.3.3) à laquelle une phase de 0,5 doit être ajoutée, éliminant ainsi le signe négatif devant l’équation.



**Figure 18** – Carte de détection de CCF pour la molécule de OH dans le spectre d’émission de WASP-33b. La figure tirée de Nugroho et al. (2021) montre la CCF combinée pour différentes valeurs de  $K_p$ . Un pic de détection apparaît à la position  $(v_{\text{sys}}, K_p)$  attendue (tirets blancs). Le rapport signal sur bruit (S/N) représenté par l’échelle de couleur est obtenu en mesurant l’écart-type dans les régions excluant le pic. La fenêtre du haut présente une vue en coupe à la valeur de  $K_p$  attendue.

s’additionnent de manière constructive autour de  $v_{\text{sys}} = -0,3_{-5,6}^{+5,3} \text{ km s}^{-1}$ , en accord avec la valeur escomptée de  $-3 \text{ km s}^{-1}$ . L’intervalle de confiance de la détection est déterminé, dans ce cas-ci, par rapport à l’écart-type ( $\sigma$ ) mesuré loin du maximum. Cependant, il n’est pas statistiquement robuste et dépend de la région choisie pour estimer le niveau de bruit.

Il existe d’autres méthodes plus sophistiquées pour le quantifier, comme le test  $t$  de Welch. Celui-ci vise à quantifier le degré de séparation entre deux distributions de variances inégales. La distribution de valeurs de CCF près de la trace du signal planétaire (*in-trail*) est comparée à celle hors de la trace (*out-of-trail*). La trace fait référence à l’intervalle de vitesse où le signal de la planète est attendu, pour des paramètres orbitaux donnés, en l’occurrence  $K_p$ . Par exemple, à la figure 13, la distribution de CCF près de la trace serait construite

à partir des valeurs à la fenêtre de droite. Ainsi, le test  $t$  de Welch peut être calculé pour différents  $K_P$  et  $v_{\text{sys}}$  pour produire des cartes de détection similaire à la figure 18.

Finalement, notons que pour toutes les techniques de réduction mentionnées jusqu’ici, il est nécessaire de normaliser chaque spectre. Les équations 2.1.4 et 2.1.5 ne peuvent donc pas être utilisées tel quel afin d’être comparées aux données puisqu’elles sont basées sur une mesure photométrique de la variation temporelle du flux. Elles doivent donc être réécrites sous la forme d’un produit du flux de l’étoile. Pour un spectre d’émission, ce flux relatif est alors donné par

$$f_{\text{rel}} = \frac{f_{\star} + f_P}{f_{\star}} = 1 + \frac{f_P}{f_{\star}} = 1 + \left( \frac{\Delta f}{f} \right)_{\text{eclipse}}. \quad (2.3.6)$$

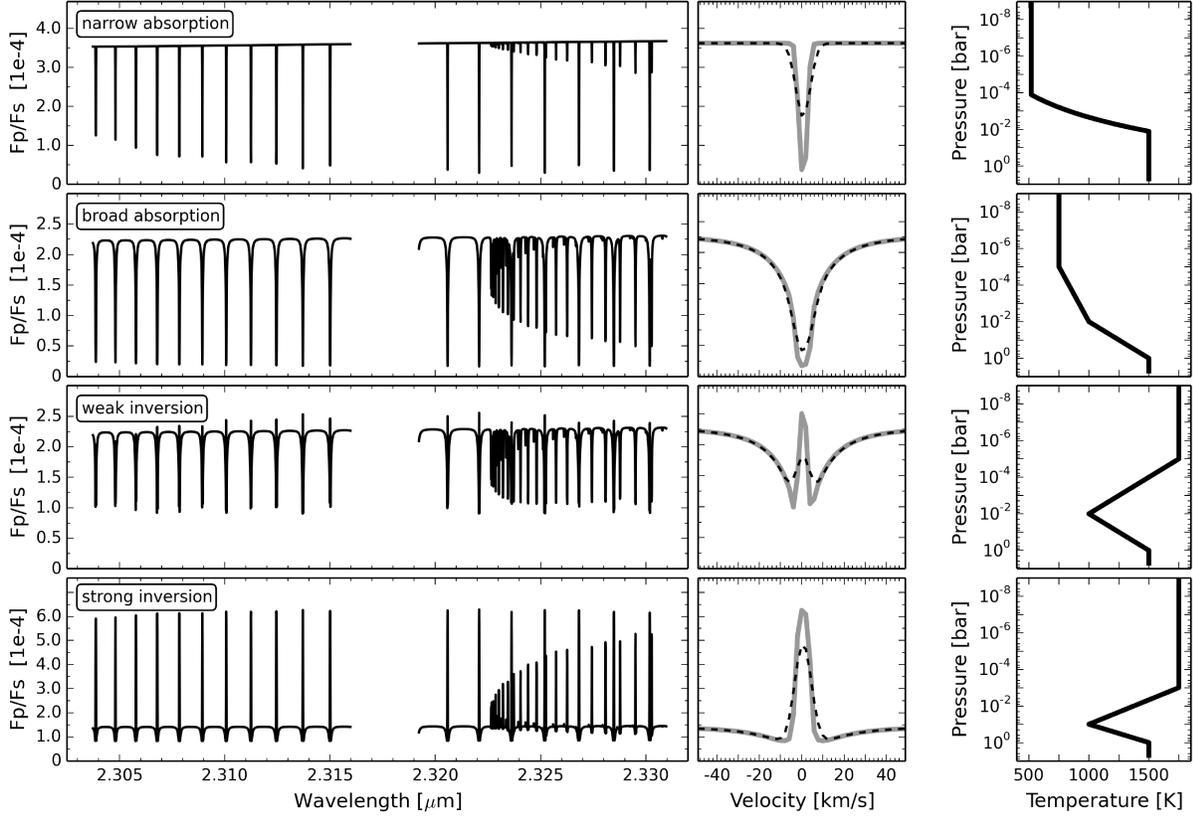
La dernière égalité met en relation la profondeur de l’éclipse de l’équation 2.1.5 et le flux relatif.

### 2.3.3. Détermination des propriétés de l’atmosphère

La HDS est complémentaire aux observations photométriques à plus basse résolution. En effet cette méthode est très sensible à la forme des raies moléculaires individuelles, qui elles-mêmes regorgent d’informations sur la structure et la composition des atmosphères.

Premièrement, lorsqu’il est possible de les résoudre, les raies moléculaires n’offrent aucune ambiguïté concernant la nature d’une espèce chimique. En effet, leur position relative forme un patron intrinsèque à chaque molécule, dicté par leur structure fondamentale. Pour reprendre l’exemple de la bande de CO montrée précédemment à la figure 10, celle-ci est formée à partir des modes de vibration-rotation de la molécule. On y retrouve un mode de vibration principal se déclinant en plusieurs sous-niveaux d’énergie basés sur les modes de rotations de la molécule. Chacun de ces modes (vibration et rotation) est déterminé de manière unique par la mécanique quantique. En comparaison, aux plus basses résolutions, les raies moléculaires se confondent pour former de larges bandes, perdant ainsi une partie de l’information.

Deuxièmement, l’allure des raies est fortement liée à la structure de l’atmosphère. En spectroscopie de transit, elle est notamment influencée par la distribution de la vitesse des vents dans le terminateur (ex : Louden & Wheatley, 2015; Flowers et al., 2019; Seidel et al., 2021) et par les nuages et aérosols (ex : Pino et al., 2018; Sánchez-López et al., 2020). En émission, c’est principalement la structure en température qui façonne les raies, comme le montre la figure 19. Dans ce contexte, la présence ou non d’une inversion est alors sans équivoque ; elle fait la différence entre une raie d’émission et une raie d’absorption. La troisième rangée de la figure montre aussi un exemple particulier de profil mixte (avec et sans inversion). Or, à basse résolution, il serait impossible d’avoir une aussi bonne définition sur la



**Figure 19** – Influence de la structure en température sur le profil des raies de CO. La figure, tirée de Schwarz et al. (2015), montre un exemple qualitativement différent dans chacune des quatre rangées, pour des modèles du spectre d’émission de HD 209458 b. Les scénarios présentent (de haut en bas) : un fort gradient de température menant à des raies d’absorption étroites (*narrow absorption*), un faible gradient résultant en une raie plus large (*broad absorption*) et deux types d’inversions, soit une plus douce (*weak inversion*) et une plus accentuée (*strong inversion*). La colonne de gauche montre le spectre obtenu à la pleine résolution du modèle, sur la plage de longueurs d’onde couverte par les deux détecteurs centraux de l’instrument CRIRES. La colonne du milieu montre une des raies de CO ( $\approx 2,3205 \mu\text{m}$ ) à la résolution du modèle (courbes en gris) et celle de l’instrument CRIRES (tirets noirs). Le profil de température utilisé est présenté à droite.

structure en température. Les observations favoriseraient alors l’un ou l’autre de ces scénarios (inversion ou non). Les deux signaux contraires pourraient même aller jusqu’à s’annuler pour produire un spectre sans traits caractéristiques (raies, bandes d’absorption).

Troisièmement, une résolution élevée permet d’explorer un plus grand intervalle de pressions. En effet, l’information prodiguée par le cœur des raies donne accès à des pressions plus faibles, en transmission et en émission. L’opacité d’une espèce chimique culminant à ces longueurs d’onde donne ainsi accès à des couches supérieures de l’atmosphère. De plus, la capacité à isoler certaines raies individuelles (ex : triplet d’Hélium métastable, Oklopčić

& Hirata 2018) peut mener à sonder des altitudes encore plus élevées comme l'exosphère et même à caractériser des atmosphères en évaporation (ex : Nortmann et al., 2018; Salz et al., 2018; Allart et al., 2019; Oklopčić, 2019).

Cependant, les observations à haute résolution comportent leur lot de complexité. Comme mentionné précédemment, elles nécessitent un traitement particulier afin de se départir d'effets nuisibles tels que l'absorption telluriques et les fluctuations atmosphériques. Ceci a pour effet d'éliminer l'information reliée au continuum planétaire, ajoutant ainsi des dégénérescences, notamment pour la détermination d'abondances absolues (en transmission) ou de la température effective (en émission), lors d'observation à haute résolution uniquement. De plus, certains éléments optiques peuvent engendrer une perte considérable de signal. La résolution aussi en est pour quelque-chose, puisque le flux est réparti sur des intervalles spectraux très fins. Cette technique est donc complémentaire aux observations à plus faible dispersion et leur combinaison promet d'enrichir considérablement notre compréhension de l'atmosphère des exoplanètes.

## 2.4. Obtention de spectres

Les méthodes d'analyse présentées dans les sections précédentes nécessitent des spectres extraits de manière adéquate. Dans la majorité des cas, ceux-ci sont obtenus à partir d'images bidimensionnelles sur lesquelles se retrouve la « trace » de l'objet ciblé. Or, le passage d'un à l'autre est une étape cruciale pour obtenir la précision nécessaire à la caractérisation d'atmosphères d'exoplanètes. Le chapitre qui suit présentera certains principes de base en spectroscopie et quelques techniques d'extraction de spectres seront ensuite présentées.

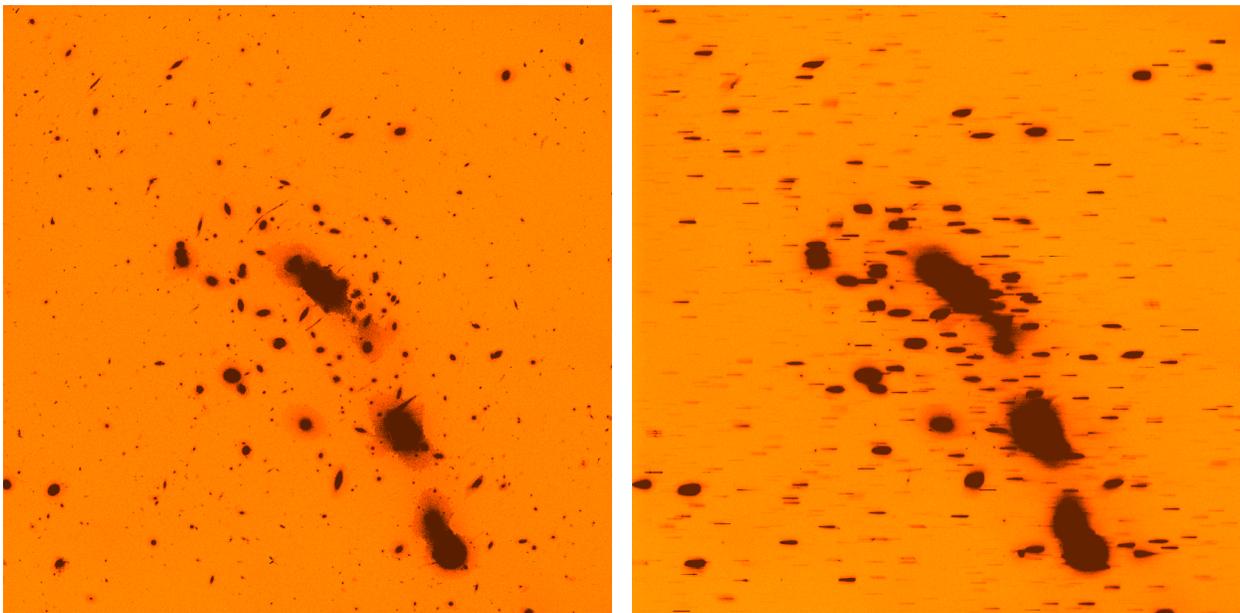
### 2.4.1. Fonctionnement d'un spectrographe

Afin d'obtenir un spectre, la lumière d'une cible isolée ou d'une portion du ciel observée doit être dispersée selon ses longueurs d'onde. La figure 20 montre l'effet d'un élément dispersif de type « grisme » sur l'image d'un amas de galaxies. Les points deviennent alors des lignes horizontales, étalant la lumière de chaque source présente le long de ce qui est appelé l'axe spectral. Ici, un filtre permettant de conserver uniquement une petite portion du spectre est ajouté afin de limiter le chevauchement d'un objet à l'autre. Cette simulation est basée sur le mode *Wide Field Slitless Spectroscopy* (WFSS) de l'instrument NIRISS, en référence à la spectroscopie dite « sans fente », par opposition à la spectroscopie conventionnelle. Cette dernière fait intervenir une fente perpendiculaire à l'axe spectral afin de bloquer le signal des autres sources. Dans l'exemple de la figure 20, on aurait donc une tranche verticale de l'image de droite. L'axe vertical, complémentaire à l'axe spectral, est aussi appelé l'axe spatial. Certains spectrographes vont même limiter la région observée en utilisant des

fibres optiques équivalent à quelques secondes d'arc de diamètre. Ceux-ci sont alors capables d'isoler le signal d'une source ponctuelle unique.

Le signal, isolé ou non, est alors transmis à un élément dispersif principal, parfois constitué d'un prisme, mais plus fréquemment d'un réseau de diffraction composé de fentes (en transmission) ou d'un miroir dentelé (en réflexion). Ces réseaux produisent un patron d'interférence donné par une enveloppe et des pics d'interférence, dont la position sera influencée par la longueur d'onde (plus fortement déviée pour les longueurs d'onde élevées). Chacun de ces pics, appelés ordre de diffraction, formera donc un spectre, excepté à l'ordre zéro, où la lumière n'est pas déviée. Le concept d'ordres de diffraction pour un réseau en transmission est imagé à la figure 21. Celle-ci montre notamment une augmentation de la résolution spectrale aux ordres plus élevés, se traduisant par un meilleur étalement des différentes couleurs. Les spectrographes de type échelle vont d'ailleurs cibler des ordres très élevés afin d'obtenir une grande dispersion spectrale.

Contrairement à ce que laisse croire le schéma, ces ordres sont portés à se chevaucher ; les longueurs d'onde plus grandes (rouges) empiétant sur les plus courtes (bleues) de l'ordre subséquent. Pour des résolutions modestes, il est possible d'utiliser un filtre pour limiter le domaine spectral (ex : figure 20). Cependant, il est impossible de faire de même aux

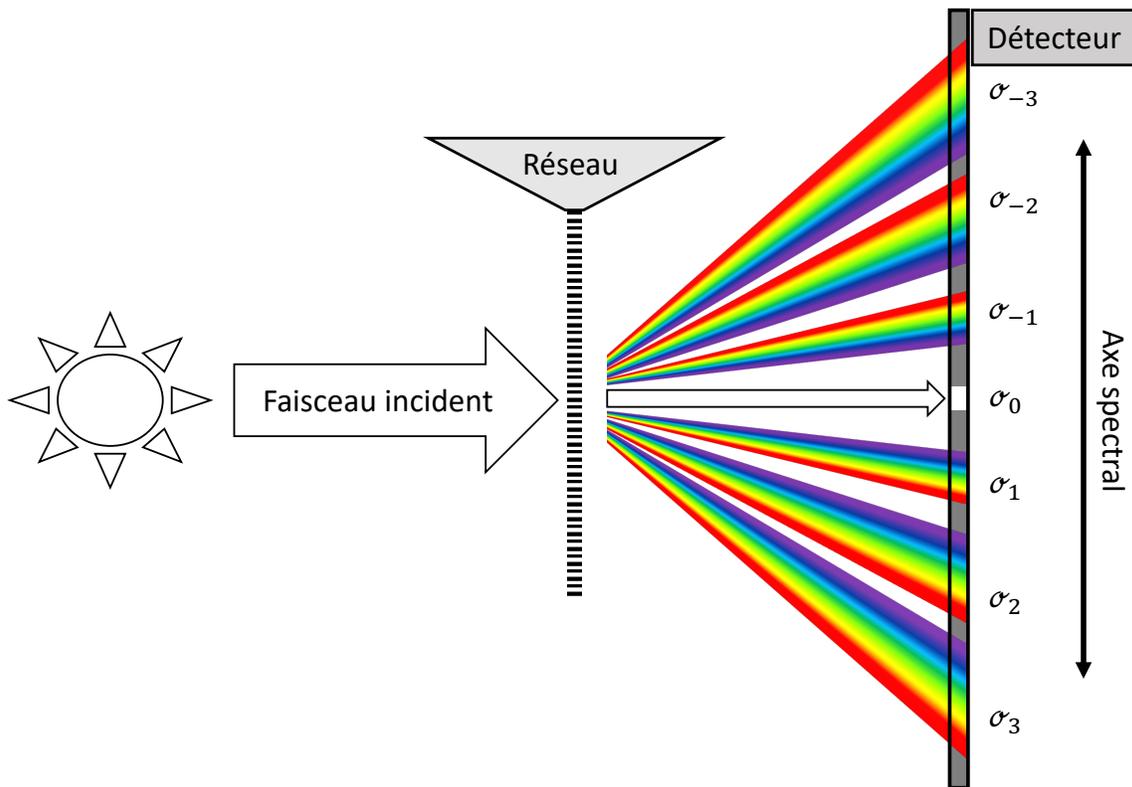


**Figure 20** — Simulation d'une observation de l'amas de galaxies MACS J0416.1-2403 prise avec le mode *Wide Field Slitless Spectroscopy* de l'instrument NIRISS en utilisant le filtre F115W. À droite, la même image est dispersée horizontalement à l'aide du grisme GR150R. La figure est tirée du *JWST User Documentation* (<https://jwst-docs.stsci.edu/jwst-near-infrared-imager-and-slitless-spectrograph/niriss-observing-modes/niriss-wide-field-slitless-spectroscopy>).

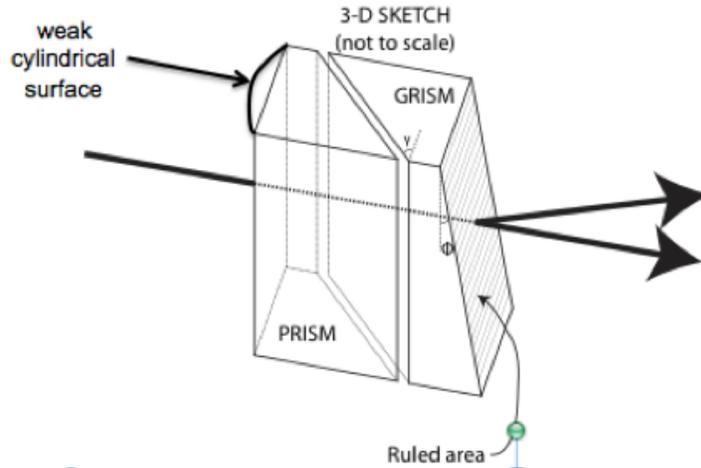
résolutions plus élevées sans fortement compromettre l'intervalle de longueur d'onde sondé. C'est pourquoi un second élément dispersif (généralement un prisme) est ajouté à l'optique de l'instrument, mais cette fois dans la direction perpendiculaire à l'axe spectral, donnant lieu à de la « dispersion croisée » (en anglais *cross-dispersion*). Cette direction est parfois appelée « axe spatial » dans le cas de spectrographes conventionnels ou sans fente. Cette appellation perd son sens pour les spectrographes à fibre optique où l'information spatiale est perdue une fois la cible isolée ; dans ce cas, on utilisera plus le terme plus général « axe de dispersion croisée ».

Dans cette optique, le mode SOSS de NIRISS utilise un grisme<sup>7</sup> composé d'un prisme dont une des faces est un réseau de diffraction, lequel suit un autre prisme dispersant dans la direction perpendiculaire. Ceci est représenté par le schéma à la figure 22. On peut y voir la lumière entrer d'abord par une face à courbure cylindrique permettant d'élargir le « profil spatial », soit la répartition du flux le long de l'axe spatial. Le grisme provoque ensuite l'effet

7. Le mot grisme provient de l'anglais *grism* formé de la combinaison de *grating* et *prism*



**Figure 21** – Schéma d'un réseau de diffraction en transmission. La lumière incidente de l'objet ciblé traverse un réseau de diffraction et se disperse selon plusieurs ordres, ici allant de -3 à 3. L'ordre zéro ne subit pas de diffraction, d'où le faisceau de lumière blanche observé sur le détecteur.



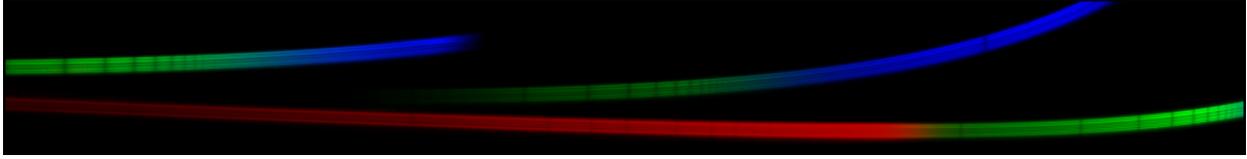
**Figure 22** – Schéma du grisme utilisé pour le mode SOSS de NIRISS. La lumière pénètre d’abord à travers un premier prisme ayant une face à courbure cylindrique permettant d’éta-ler la lumière sur plusieurs pixels. Elle est ensuite transmise à un grisme permettant une dispersion croisée jusqu’au détecteur de NIRISS. Le réseau de diffraction est situé sur la face externe du grisme, identifiée par *Ruled area*.

de dispersion croisée avec un réseau de diffraction sur la face sortante. Le résultat observé sur le détecteur de NIRISS est présenté à la figure 23. Les trois premiers ordres de diffraction  $y$  sont présentés. Notons que sans l’élément de dispersion croisée causant l’effet de courbure verticale, les traces de chaque ordres tomberaient alors l’une sur l’autre. Par ailleurs, malgré que l’ordre 2 s’atténue considérablement aux longueurs d’onde plus élevées (couleur verte), celui-ci ne disparaît pas complètement et vient contaminer la partie la plus rouge de l’ordre 1. Ce problème de chevauchement fait l’objet du second article (chapitre 4) composant cette thèse.

### 2.4.2. Techniques d’extraction conventionnelles

L’extraction « par boîte » (de Boer & Snijders, 1981) et l’extraction « optimale » (Horne, 1986) sont les deux techniques les plus répandues pour passer d’une trace sur un détecteur à un spectre à une dimension.

La première méthode (la plus simple) consiste à sommer les valeurs mesurées par les pixels le long de l’axe spatial. Celle-ci est fréquemment utilisée pour l’analyse de données prises avec l’instrument WFC3. La figure 24 présente un exemple de spectre obtenu pour cette caméra en utilisant le grisme G141. Pour distribuer le flux sur plusieurs pixels, le mode de balayage spatial est utilisé. Celui-ci consiste à décaler progressivement l’image de la cible perpendiculairement à l’axe spectral au cours d’une même observation. La stratégie est similaire dans le cas du mode SOSS de NIRISS, à la différence qu’elle est directement implantée dans l’optique de l’instrument, avec sa lentille cylindrique. Ainsi, il est possible d’observer des objets plus brillants sans saturer et de réduire l’incertitude, l’erreur de chaque

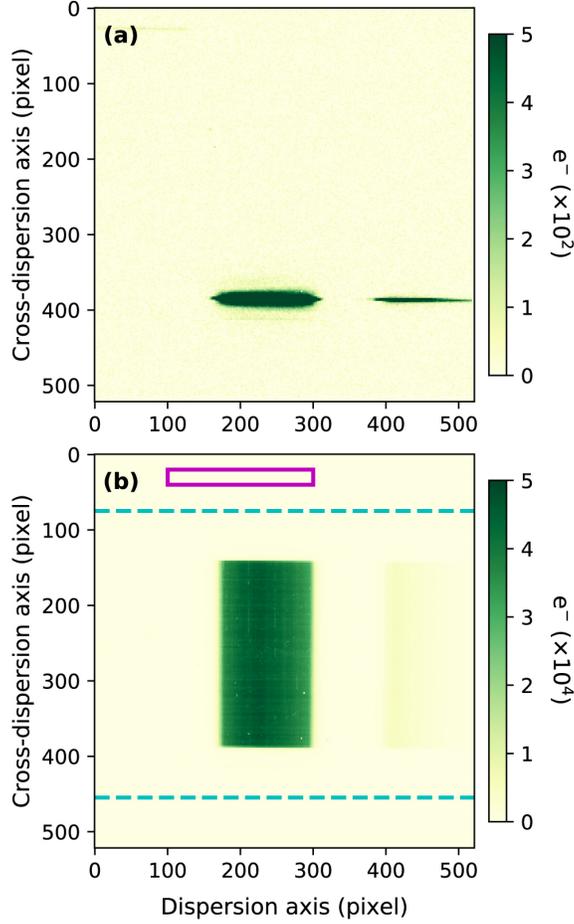


**Figure 23** – Image du détecteur de NIRISS dans le contexte du mode SOSS. Les couleurs sont une représentation de l’intervalle de longueurs d’onde infrarouges observées par le mode, transposées dans le spectre visible. On y trouve les trois premiers ordres de diffraction du grisme (voir figure 22). Le premier est situé au bas de l’image et traverse la totalité du détecteur. Il couvre les longueurs d’onde les plus longues (rouges). La partie bleue du second ordre commence en haut à droite du détecteur et s’atténue vers le centre du détecteur, aux longueurs d’onde correspondant au début de l’ordre 1. Le troisième ordre, situé en haut à gauche de l’image, couvre sensiblement le même intervalle que l’ordre 2. L’étalement des couleur sur l’axe horizontal (axe spectral) est le fruit de la dispersion du réseau de diffraction situé sur la face externe du grisme. La courbure des traces est créée par la dispersion croisée des prismes le long de l’axe spatial (vertical). La largeur des traces, quant à elle, est causée par la lentille cylindrique. **Crédits : NASA, CSA, and NIRISS team/Loic Albert, University of Montreal.**

pixel s’additionnant en quadrature. Dans l’exemple de la figure 24, le spectre est obtenu en additionnant les pixels le long de l’axe de dispersion croisée et situés à l’intérieur de l’ouverture délimitée par les tirets cyan. Le bruit de fond doit avoir été préalablement corrigé, ce qui est fait dans cet exemple en soustrayant la valeur médiane des pixels contenus dans la région encadrée en violet. Le résultat est une intensité totale mesurée en électrons en fonction des colonnes du détecteur. Pour obtenir un spectre à proprement dit, une calibration en longueurs d’onde est ensuite nécessaire. Cette étape ne sera pas traitée dans le présent ouvrage.

Cette technique revient à compter la quantité totale de photons ayant été convertis en électrons par les pixels. Elle ne tient cependant pas compte des pixels aberrants (mauvais pixels, rayons cosmiques, pixels chauds, etc.). Ceux-ci doivent donc être corrigés préalablement à l’extraction, ce qui peut occasionner des biais. De plus, elle ne fait aucune distinction entre les pixels ayant reçu ou non du signal provenant de la cible. Par exemple, dans le cas présenté à la figure 24, les pixels près des limites de l’ouverture sont inclus dans la somme alors qu’il ne reçoivent pas de flux. Pour des objets à haut rapport signal sur bruit, la contribution de ces pixels est négligeable. Inversement, pour un signal plus faible, il faut recourir à la méthode d’extraction optimale. C’est d’ailleurs le cas pour plusieurs spectrographes à haute résolution.

Cette seconde méthode consiste à pondérer chaque pixel selon un profil spatial afin de tenir compte de la distribution relative du flux le long de l’axe de dispersion croisée. Le



**Figure 24** – . Exemple d’intégration obtenue par l’instrument WFC3 avec le grisme G141 en mode de « balayage spatial » (*spatial scanning mode*), tiré de Mikal-Evans et al. (2021). (a) Première lecture non-destructrice du détecteur. Les traces du premier et du second ordre de diffraction apparaissent de gauche à droite respectivement. (b) Dernière lecture non-destructrice du détecteur. Le mouvement du télescope permet de répartir le flux dans la direction spatiale (*Cross-dispersion axis*). Les tirets cyan montrent l’intervalle de pixels utilisés pour l’extraction par boîte. L’encadré violet identifie la région sélectionnée pour déterminer le bruit de fond. L’échelle de couleurs est différente de (a), laissant l’ordre 2 plus difficile à distinguer.

spectre à la colonne  $k$  du détecteur est alors donné par l’équation

$$S_k = \sum_i \frac{D_i \cdot P_i / \sigma_i^2}{\sum_i P_i^2 / \sigma_i^2}, \quad (2.4.1)$$

où la somme est faite sur les pixels de cette colonne, étiquetés par les indices  $i$ . Le terme  $D$  représente l’intensité mesurée et  $\sigma$  est l’incertitude associée à chaque pixel. Le profil spatial est déterminé selon l’équation

$$P_i = \frac{P_i}{\sum_j P_j}, \quad (2.4.2)$$

où la somme se fait encore une fois sur la colonne. Le résultat consiste donc en une moyenne pondérée le long de l'axe spatial. Notons que ces équations supposent une correction préalable du bruit de fond<sup>8</sup> et d'autres sources d'erreurs, tout comme la méthode d'extraction par boîte.

Outre le fait qu'elle permette d'augmenter le rapport signal sur bruit pour des sources d'intensité faible ou modérée, l'extraction optimale a l'avantage de traiter explicitement les pixels aberrants. En effet, puisque le résultat est une moyenne pondérée, il est possible de simplement exclure certaines données en leur associant un poids nul.

Le principal désavantage de cette méthode est qu'elle requiert de connaître précisément le profil spatial. Or, il n'est pas simple de le déterminer, ce qui peut insérer des biais dans le spectre résultant. Pour éviter de telles erreurs, il est nécessaire d'obtenir une précision relative sur le profil spatial  $P$  beaucoup plus grande que celle des données  $D$ . Dans le cas d'observations à faible rapport signal sur bruit avec certains instruments, une forme analytique comme une gaussienne peut être présumée. Cependant, à des précisions plus élevées, l'allure du profil spatial peut différer de l'approximation analytique. D'autres approches se baseront plutôt sur les données elle-mêmes pour le déterminer, en partant du principe que le profil varie lentement selon les longueurs d'onde. Il est donc possible de combiner des colonnes adjacentes afin d'augmenter la précision sur le profil (ex : Horne, 1986; Marsh, 1989). Cette étape demeure complexe et semble introduire des erreurs dans certaines situations (ex : Rackham et al., 2017). Néanmoins, cette technique est aussi utilisée pour extraire des données de WFC3 comme celles présentées à la figure 24, avec des performances similaires (ex : Kreidberg et al., 2014).

Par ailleurs, ces deux méthodes sont basées sur l'hypothèse que l'axe de dispersion croisée est parfaitement perpendiculaire à l'axe spectral et aligné avec les coordonnées du détecteur. Ainsi, chaque colonne est supposée couvrir le même intervalle de longueurs d'onde et devrait donc recevoir le même flux sous-jacent. Or, cette contrainte n'est pas toujours respectée. C'est le cas pour plusieurs spectrographes dont le *Sloan Digital Sky Survey* (SDSS : York et al. 2000), NIRISS/SOSS et même WFC3/G141. À basse résolution (ex : WFC3), les conséquences sont généralement négligeables, mais dans d'autres contextes, elles peuvent mener à une mauvaise correction du ciel, notamment des raies d'émission OH (ex : Bolton & Schlegel, 2010).

De plus, les méthodes conventionnelles ne permettent pas de départager le flux de deux sources ou ordres de diffraction superposés. Pour le mode SOSS plus particulièrement, le flux extrait dans la partie rouge du premier ordre de diffraction (voir figure 23) contiendrait une petite proportion du flux de l'ordre 2. Les spectrographes sans fente comme NIRISS/SOSS

---

8. La méthode de Horne (1986) permet un traitement itératif du bruit de fond, appelé *sky background*, mais la soustraction préalablement à l'extraction est souvent préférée (ex : Arcangeli et al., 2018)

sont aussi sujets à la contamination provenant d'autres sources, ce qui complexifie l'extraction des spectres et ajoute des contraintes aux cibles observées.

Plusieurs méthodes découlant de l'extraction optimale ont été proposées dans la littérature afin de tenir compte de distorsions dans la solution en longueurs d'onde (ex : Bolton & Schlegel, 2010; Cornachione et al., 2019) ou de chevauchement entre différentes sources (ex : Khmil & Surdej, 2002; Hynes, 2002; Lucy & Walsh, 2003; Min et al., 2020). Cependant, aucune de ces méthodes ne tient compte du cas spécifique d'ordres de diffraction d'une même source, comme dans le cas du mode SOSS. Elles n'ont pas non plus été testées pour des observations spectrophotométriques en série temporelle, nécessaires à l'analyse de l'atmosphère d'exoplanètes. Finalement, dans le cas du SOSS, la forme particulière du profil spatial empêche l'application de certaines de ces techniques basées sur une paramétrisation analytique.

## Chapitre 3

---

# Probing the atmospheres of Ultra Hot Jupiters with CO : A combined study of WASP-33b with SPIRou & HST.

Ce manuscrit sera soumis à la revue arbitrée  
*Monthly Notices of the Royal Astronomical Society.*

ANTOINE DARVEAU-BERNIER<sup>1</sup>, ANNE BOUCHER<sup>1</sup>, STEFAN PELLETIER<sup>1</sup>, DAVID LAFRENIÈRE<sup>1</sup>, RENÉ DOYON<sup>1</sup>, LOUIS-PHILIPPE COULOMBE<sup>1</sup>, JAKE TAYLOR<sup>1</sup>, NICK COWAN<sup>2, 3</sup>, NEIL J. COOK<sup>1</sup>, ÉTIENNE ARTIGAU<sup>1</sup>, BJÖRN BENNEKE<sup>1</sup>, LUC BAZINET<sup>1</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> Department of Earth & Planetary Sciences, McGill University, 3450 rue University, Montréal, QC H3A 0E8, Canada

<sup>3</sup> Department of Physics, McGill University, 3600 rue University, Montréal, QC H3A 2T8, Canada

## Contribution de l'étudiant

Ce premier article vise à améliorer la contrainte sur la métallicité de la planète WASP-33 b en combinant la spectroscopie à haute résolution avec des données spectrophotométriques préexistantes. Pour ce faire, j'ai dû d'abord écrire les demandes de temps en tant que scientifique principal pour obtenir des observations avec le SpectroPolarimètre InfraRouge (SPIRou) au télescope Canada-France-Hawaï. Ces demandes ont nécessité le développement d'un code de simulation en langage Python permettant d'évaluer le niveau de détection de

l’atmosphère de WASP-33 b. L’écriture de ce code s’est fait en collaboration avec Anne Boucher et les modèles d’atmosphères ont été générés par Stefan Pelletier en utilisant SCARLET (Benneke, 2015).

Anne Boucher et moi avons donc dû implémenter toutes les étapes nécessaires à la modélisation du mouvement du spectre de la planète (ex : figure 15), sans oublier la contribution du signal de l’étoile et de l’atmosphère terrestre. Nous avons aussi implémenté tous les algorithmes liés à l’extraction et à l’analyse du signal planétaire (ex : figure 16), incluant une analyse en composantes principales (PCA : *Principal Component Analysis*) et le calcul de CCF. Ceci n’inclut pas les étapes préliminaires de réduction menant à des spectres unidimensionnels pour chaque ordre de diffraction, qui est plutôt faite par APERO, le logiciel de réduction de données de SPIROU et de correction des telluriques.

Le même code d’analyse a ensuite pu être transposé à l’analyse des données obtenues par SPIROU. Anne Boucher a ensuite poursuivi le développement en perfectionnant le code pour mener à la création de STARSHIPS<sup>1</sup> (Boucher et al., 2021). Mes contributions plus spécifiques dans la suite du développement incluent notamment l’ajout de noyaux de convolution pour traiter de la rotation de la planète en spectroscopie de transit ainsi que la plupart des facettes reliées à la spectroscopie en émission. Nos discussions avec Stefan Pelletier ont aussi contribuer au développement de STARSHIPS. J’ai aussi développé en parallèle un code nommé ExoFile<sup>2</sup> permettant de gérer les données provenant du Nasa Exoplanet Archive et d’obtenir des paramètres plus cohérents pour les planètes s’y retrouvant. Ceci a permis entre autres de faciliter la recherche des meilleures cibles, la planification des observations ainsi que la simulation des orbites (pour STARSHIPS). C’est d’ailleurs une version préliminaire de ExoFile qui m’a permis de sélectionner WASP-33 b en tant qu’une des cibles les plus favorables pour être observé en émission avec SPIROU.

J’ai accompli la majeure partie de l’analyse des données ainsi que l’écriture de l’article, avec les conseils et l’aide de Anne Boucher et Stefan Pelletier.

---

1. <https://github.com/boucherastro/starships>

2. <https://github.com/AntoineDarveau/exofile>

ABSTRACT. Dayside atmospheres of ultra-hot Jupiters have now puzzled astronomers for the better part of a decade. Many emission spectra obtained with the Wide Field Camera 3 of the Hubble Space Telescope exhibit very smooth features. This increases the level of difficulty in differentiating between continuum opacities like  $\text{H}^-$  and different molecular bands due to the relatively narrow instrument bandpass and low spectral resolution. Thus, retrievals of the temperature profiles and abundances generally relied significantly on model assumptions. High resolution spectroscopy constitutes a promising alternative to constraint the abundance of individual chemical species due to its sensitivity to individual lines. In this work, we present observations of the dayside of the UHJ WASP-33 b with the echelle spectrograph SPIRou (*SpectroPolarimètre InfraRouge*) at the Canada-France-Hawaii Telescope. We detect carbon monoxide in both WASP-33 b’s eastern and western dayside hemispheres. The molecular lines are detected in emission, indicating that WASP-33 b’s atmosphere is thermally inverted, consistent with previous studies. We further perform a combined retrieval analysis of SPIRou and WFC3 data and infer a slightly low CO volume mixing ratio of  $\log_{10} \text{CO} = -4.07^{+1.51}_{-0.60}$ , but still consistent with Solar- or Jupiter-like values. We also obtained an upper limit for water of  $\log_{10} \text{H}_2\text{O} < -3.09$  (at high pressures), which takes into account the effect of molecular dissociation using a parametrization of the temperature and pressure. Both SPIRou and WFC3 data are not sensitive to the presence of TiO below ( $\log_{10} \text{TiO} < -2.81$ ). This work emphasizes the potential of combining low- and high-resolution data to narrow our understanding of exoplanet’s atmospheres.

Keywords *Unified Astronomy Thesaurus concepts* : Exoplanet atmospheres (487) ; Hot Jupiters (753) ; Exoplanet atmospheric composition (2021)

### 3.1. Introduction

High dispersion spectroscopy has emerged in the last decade as a powerful new technique to characterize exoplanets atmospheres (e.g., Brogi et al., 2012; Birkby et al., 2017; Schwarz et al., 2013; de Kok et al., 2013). By taking advantage of the extreme wavelength precision of these instruments, it is possible to distinguish the rapid change in the planet’s radial velocity from the mostly static signal from its host star, from the telluric absorption and from some other potential systematics (Snellen et al., 2010). Also, at such high resolution ( $R > 50,000$ ), each individual spectral line of molecular bands can be resolved which can reveal information about the atmospheric vertical structure, the elemental abundances and even the distribution of winds in the planet’s atmosphere (e.g., Snellen et al., 2010; Louden & Wheatley, 2015; Flowers et al., 2019).

However, this technique comes with a cost. Indeed, ground-based high-resolution spectrographs generally do not have a high photometric precision. Hence they are less sensitive to

broadband features. This implies, among other things, that absolute quantities like the planet’s radius (in transmission) or the brightness temperature of the continuum (in emission) cannot be precisely determined. These instruments also generally bear lower signal-to-noise ratios (SNR) due to the fact that light is widely spread in wavelength. Finally, the cross-correlation techniques used to isolate the planetary signals are based on the assumption that the planet’s radial velocity changes rapidly compared to other signals. Therefore, the most adequate objects for these measurements need to bear a large number of strong spectral lines, to have short revolution periods and need to be bright.

In the context of emission spectroscopy of exoplanets, Ultra-Hot Jupiters (UHJ) are among the most favorable targets for atmospheric characterization. Indeed, this class of exoplanet presents the highest planet-to-star flux ratios and they orbit extremely close to their star, conferring them with very short orbital period and high temperatures. Moreover, at such high temperatures ( $T_{eq} > 2200$  K), nearly all of their constituents are expected to be in gaseous phase. However, previous spectrophotometric observations of secondary eclipses with the Wide Field Camera 3 (WFC3) onboard the Hubble Space Telescope (HST) have uncovered features (mainly the 1.4 water band) that are lower than expected (if not absent), and even planetary flux consistent with blackbody emission (e.g., Arcangeli et al., 2018; Pluriel et al., 2020; Cartier et al., 2016; Kreidberg et al., 2018; Mansfield et al., 2018; Evans et al., 2017; Haynes et al., 2015).

A part of the explanation comes from the fact that, in the context of emission spectroscopy, the line intensity of a particular species is highly dependant on its abundance and the temperature-pressure profile (TP-profile). For instance, an inversion of temperature, i.e., an increase of temperature at low pressures (high altitude), would cause emission lines whereas decreasing temperature gradient would translate into absorption lines. The strength of these features is determined by the interplay between the abundance, which influences the range of pressures probed, and the amplitude of the temperature variation within this range (e.g., Griffith, 2014). In the particular case of an isothermal profile, no spectral feature can be formed and the emission spectrum would resemble that of a blackbody, as highlighted by Mansfield et al. (2021).

On the other hand, some spectra have shown to deviate from blackbody emission, even if they were (almost) featureless (e.g., Evans et al., 2017; Haynes et al., 2015). This is mainly attributed to effects that occur at stellar-like temperatures : the opacity of the  $H^-$  ion and the thermal dissociation of some molecules (water in this case). At such high temperatures, some atomic species (Na, K, Ti and V) become ionized, which will increase the abundance of free electrons in the atmosphere. In parallel, many molecules like water and  $H_2$  will thermally dissociate, favoring the presence of atoms over molecules. These phenomena are both accentuated at lower pressures. Consequently, while the continuum opacity of  $H^-$  becomes

Table 1. WASP-33 System Parameters

Stellar Parameters	Value
Spectral Type <sup>a</sup>	A5
$H$ magnitude <sup>b</sup>	$7.516 \pm 0.024$ mag
Stellar mass ( $M_*$ ) <sup>c</sup>	$1.561^{+0.045}_{-0.079} M_\odot$
Stellar radius ( $R_*$ ) <sup>c</sup>	$1.509^{+0.016}_{-0.027} R_\odot$
Temperature ( $T_{\text{eff}}$ ) <sup>d</sup>	$7308 \pm 71$ K
Surface gravity ( $\log g$ ) <sup>d</sup>	4.5 cgs
Metallicity ( $[\text{Fe}/\text{H}]$ ) <sup>d</sup>	0.0
RV Semi-amplitude ( $K_*$ ) <sup>e</sup>	$460 \pm 110$ m s <sup>-1</sup>
$v \sin(i_*)$ <sup>f</sup>	$90 \pm 10$ km s <sup>-1</sup>
Planet Parameters	Value
Planet mass ( $M_p$ ) <sup>g</sup>	$3.28 \pm 0.73 M_J$
Planet radius ( $R_p$ ) <sup>g</sup>	$1.679^{+0.019}_{-0.030} R_J$
Orbital period ( $P$ ) <sup>h</sup>	1.219870(89) days
Epoch of transit ( $T_0$ ) <sup>h</sup>	2454163.223(67) BJD
Transit duration ( $T_{14}$ ) <sup>i</sup>	2.854(12) hours
Semi-major axis ( $a$ ) <sup>c</sup>	$0.02555 \pm 0.00017$ AU
Inclination ( $i_P$ ) <sup>d</sup>	$87.67 \pm 1.81^\circ$
Eccentricity ( $e$ ) <sup>g</sup>	0
Planet RV semi-amplitude ( $K_{P,0}$ ) <sup>e</sup>	$231 \pm 4$ km s <sup>-1</sup>
Systemic velocity ( $v_{\text{sys}}$ ) <sup>j</sup>	0 km s <sup>-1</sup>

References. — <sup>a</sup> Cannon & Pickering (1918); <sup>b</sup> Cutri et al. (2003); <sup>c</sup> Kovács et al. (2013); <sup>d</sup> Stassun et al. (2017); <sup>e</sup>  $K_*$  and  $K_{P,0}$  were computed by sampling 5000 values of  $M_P$ ,  $M_*$ ,  $P$ , and  $i_P$  from a normal distribution using their mean value and uncertainties listed in this table; <sup>f</sup> Cameron et al. (2010); <sup>g</sup> Turner et al. (2016), <sup>h</sup> Zhang et al. (2018); <sup>i</sup> Chakrabarty & Sengupta (2019); <sup>j</sup> No reference.

more important, the more the water feature will be attenuated, which can affect the signal seen in the WFC3 bandpass.

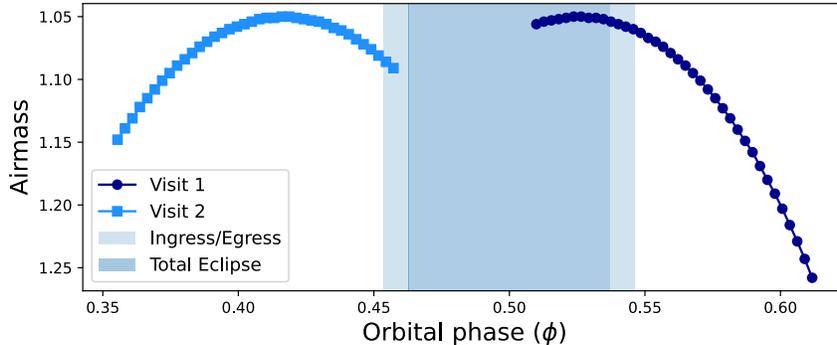
Although a satisfactory explanation exists for this deficiency of water features in emission spectra of UHJ (e.g. Parmentier et al., 2018), it still causes a big challenge when it comes to analysing their atmosphere. Indeed, the lack of distinct features like the water band increases the degeneracies in WFC3 observations due to its limited bandpass and resolution. Therefore, it increases the dependency on the modeling assumptions and impedes on the determination of the TP profile. High resolution spectroscopy can mitigate some of these degeneracies by resolving individual lines. Particularly, in the said bandpass, some molecules like OH and FeH (Kesseli et al., 2020) are less affected by molecular dissociation. Moreover, some high

resolution spectrographs cover different (and in some cases broader) spectral domains, giving access to atomic species at shorter wavelengths (e.g. Yan et al., 2022b; Pino et al., 2020) or, at longer wavelengths, to CO that needs higher temperatures to dissociate (Lodders & Fegley, 2002).

### 3.1.1. WASP-33b

In this study, we use a combination of new high-resolution observations and former spectrophotometric observations to analyse the atmosphere of the second hottest UHJ currently known, WASP-33 b (Cameron et al., 2010). Its eclipse spectrum was previously observed with WFC3 (Haynes et al., 2015) and the *Spitzer Space Telescope* (Deming et al., 2012), but the lack of distinct features prevented the TP-profile to be well constrained. Most of the conclusions were based on the stitching of *Spitzer* points, which is known to be a delicate task (e.g. Pluriel et al., 2020; Changeat & Edwards, 2021). A tentative detection of TiO was also proposed to explain the excess in blue-most part of the eclipse spectrum. This was later tentatively confirmed by the detection of TiO emission lines (supporting a temperature inversion), with observations taken with the High Dispersion Spectrograph on the Subaru telescope (Nugroho et al., 2017). However, the TiO signal presented a small discrepancy with the expected planet’s orbital semi-amplitude. In fact, this detection could not be reproduced by Herman et al. (2020), who were only able to establish a  $3\sigma$  upper limit of  $10^{-9}$  on the TiO volume mixing ratio (VMR). Moreover the initial detection significance was weakened and slightly offset in velocity space in a re-analysis of the Subaru data with a new line list (Serindag et al., 2021), leaving some doubts on both the presence of TiO and the detection of a temperature inversion.

Hydroxyl radical (OH) was also found in WASP-33 b (Nugroho et al., 2021), making it the very first detection of this molecule in an exoplanet’s atmosphere. This was all the more significant since it is a by-product of thermal dissociation of water. Other atomic species add up to the list of detections in the emission spectrum of WASP-33 b, i.e. Fe (Herman et al., 2022; Nugroho et al., 2020; Cont et al., 2021) and Si (Cont et al., 2022). All of these detections were consistent with emission lines, confirming the temperature inversion. However, the interpretation and abundance determination from the above detections are complicated for many reasons. Indeed, Fe and Si are also found in the stellar atmosphere which presents strong sinusoidal variations, being a  $\delta$ -scuti star (Herrero et al., 2011; von Essen et al., 2020), which add strong systematics and complicates the extraction of the planetary signal. Furthermore, the abundances of these species are sensitive to the local conditions and can vary a lot at different altitudes. This increases the level of degeneracy between the TP and abundance profiles. On the contrary, CO is less influenced by thermal dissociation due to its strong bounds (Lodders & Fegley, 2002). It is also generally well mixed



**Figure 1** – Airmass variation as a function of orbital phase for each visit on WASP-33 b. The shaded regions indicate the phases in total (dark blue) and partial (light blue) eclipse.

in exoplanets’ atmospheres, making it a prime molecule to probe the structure and dynamics in place. In fact, this molecule was recently discovered in WASP-33 b (Yan et al., 2022a; van Sluijs et al., 2022), but no attempt has been made yet to precisely estimate its abundance.

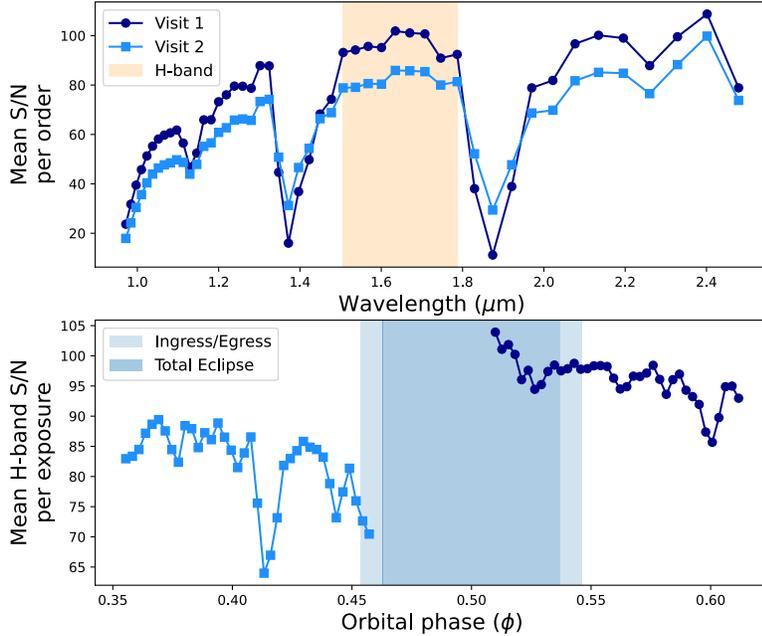
In this paper, we present the most precise constraint to date on the CO abundance and the TP-profile of the UHJ WASP-33 b. This is the result of the combination of high- and low-resolution observations of its emission spectrum. It also represents the third detection of CO and thus a firm confirmation of its presence in the dayside of this exoplanet. The new observations taken with the *SpectroPolarimètre InfraRouge* at the CFHT are presented in §3.2 along with the reduction steps. In §3.3, we describe the atmospheric modeling and the post-processing needed to properly isolate and characterize the planet’s emission spectrum. This is followed by the presentation of the new detection of the CO line and the analysis leading to the new constraints (§3.4). Finally, we compare this new result with previous detections and discuss some of its implications in §3.5.

## 3.2. Observations and Data Reduction

### 3.2.1. Observations

SPIRou (*SpectroPolarimètre InfraRouge*, Donati et al. 2020) is an infrared échelle spectrograph installed at the 3.6 meter Canada-France-Hawaii Telescope (CFHT). It covers a wavelength range from 0.95 to 2.50  $\mu\text{m}$  distributed into 49 orders of diffraction (from order 30 to 79 on the detector). This, combined with its nominal resolving power of  $\lambda/\Delta\lambda \approx 70,000$  allows SPIRou to simultaneously and efficiently probe for spectral lines of molecules such as CO, H<sub>2</sub>O, CH<sub>4</sub> and OH.

Two observations were taken under the program 19BC026 (PI Darveau-Bernier) during the nights of October 7 and November 4 2019. More details about the observations can be found in Table 2. The non-polarimetric mode of SPIRou was preferred to ensure the stability



**Figure 2** – Signal-to-noise ratios of each exposure on WASP-33 b. The top panel presents the SNR per spectral order with the H-band highlighted in light salmon pink. The bottom panel shows mean SNR per order in the H-band as a function of orbital phase. The rest of the description is the same as Fig. 1.

of the instrument optics. Hence, only the Stokes parameter I (intensity) was measured<sup>3</sup>. We opted to use the Fabry-Perot calibration (Hobson et al., 2021) in the reference channel to get the best radial velocity precision. Any potential contamination coming from the calibration channel can be corrected in the data reduction process as it would be static.

The observations were taken close to secondary eclipse in order to get the dayside of WASP-33 b, where we expect the planets emission signal to be the strongest. To optimize the (SNR) and to facilitate scheduling, the ingress and egress were taken during separate visits. The airmass stayed below 1.3 during all observations, as shown in Fig. 1, while the mean *H*-band SNR varied between 65 and 100 (see Fig. 2).

Better conditions on the first night, drop in signal in the middle of the second night exp time chosen to reach a high signal to noise, but separate enough to sample the radial velocity variations of WASP-33 b.

3. To obtain the other components, the angle of the rhomboid must be changed at each observations which has an impact on the observed spectra and complicates the removal of the non-planetary signal.

Table 2. SPIRou observations of WASP-33

Visit	1	2
UT Date	2019-10-07	2019-11-04
BJD (d) <sup>a</sup>	2458763.98	2459478.85
Integration time (s) <sup>b</sup>	267	267
Seeing (") <sup>c</sup>	0.49–0.66	0.76–0.91
S/N <sup>d</sup>	64.9	61.8
<b>Number of exposures :</b>		
Before ingress	0	36
During eclipse	14	2
After egress	24	0
Total	38	38
Total observing time (h) <sup>e</sup>	2.98	2.98

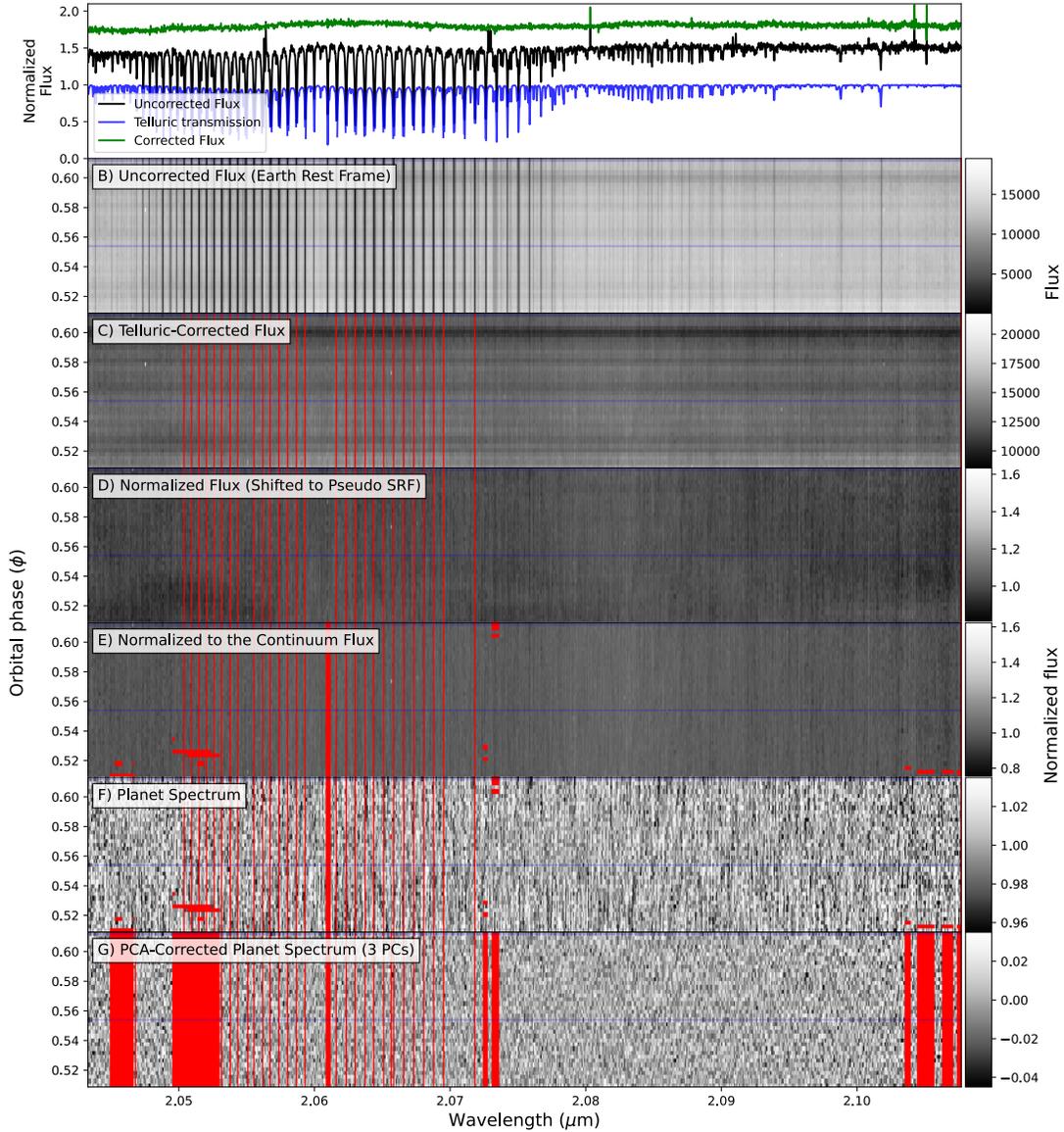
Note. — <sup>a</sup> Barycentric Julian date at the start of the observing sequence; <sup>b</sup> Exposure time of a single exposure; <sup>c</sup> Range of values of the seeing during the transit; <sup>d</sup> Mean S/N per pixel, per exposure, at  $1.7\ \mu\text{m}$ ; <sup>e</sup> Total observing time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

### 3.2.2. Data Reduction

The first part of the reduction was done with *APER0*, *A PipelinE to Reduce Observations*, (Cook et al., in prep.) version 0.7.232. *APER0* is a powerful tool to extract spectra from raw data of fiber-fed échelle spectrographs. Details on all the reduction steps are presented in Artigau et al. (2018). The result of this extraction is a wavelength calibrated spectrum for each spectral order, labeled E2DS, with a pixel sampling of 2.3 km/s.

*APER0* does not only extract the data, but also provides a correction for the telluric lines, which is a major concern in infrared spectroscopy. It uses the TAPAS Earth transmittance models (Bertaux et al., 2014) as a first estimate of the tellurics absorption lines. An additional correction with 3 degrees of freedom ( $\text{H}_2\text{O}$  optical depths, the dry components, and a constant), optimized using a library of rapidly rotating stars observations, is then applied on the residuals. The full process is described in more detail in (Cadieux et al., 2022).

This correction can remove most of the telluric lines (Artigau et al., 2014; Cadieux et al., 2022). However, to extract the planetary transmission spectra, a more thorough reduction is needed. We used *STARSHIPS* (Boucher et al., 2021, 2022), an open source framework specifically designed to extract and compare planetary signals from high resolution spectra (currently adapted for SPIRou data products). Reduction steps performed by *STARSHIPS* to clean the data are shown in Fig. 3 and summarized here. Firstly, each spectrum produced by *APER0* is divided beforehand by the order specific blaze function. Secondly, each order of each exposure from the telluric corrected spectra is normalized by its median. This puts each



**Figure 3** – Reduction steps to isolate the planetary signal for order 37 of the first visit. (Top panel) Examples of products from the APERO pipeline that are used in the current analysis. Each spectrum is normalized by its median and a vertical shift of 0.5 is applied between each of them. B) Uncorrected flux (APER0 product) as a function of WASP-33 b orbital phase. C) Telluric-corrected flux (APER0 product) as a function of orbital phase. The values in red represent the spectral pixel that are masked by the reduction steps. In the panel C and D, it marks the spectral regions that could not be corrected due to strong telluric absorption. D) Median-normalized flux in WASP-33 pseudo stellar rest frame. This steps highlights the low-frequency variations between exposures. E) Flux normalized to the continuum, exposure-by-exposure. This has the effect to remove most of the low order variations from step D) and to identify outliers (additional red regions). F) Isolated planetary spectrum : the master stellar spectrum has been removed from each exposure of the panel E. However, some residual vertical noise structures remain. G) PCA-corrected planetary spectra time serie, using two principal components. The vertical residual are removed and spectral channels containing masked values or exhibiting higher standard deviation are excluded.

observation at the same level (see panel D of Fig. 3). Each of them is also aligned on the stellar rest frame by applying a Doppler shift to remove the variation from the barycentric earth radial velocity (BERV). This allows to combine each exposure using a weighted mean, based on the order’s SNR, to create a reference spectrum of the star. Each epoch is reduced independently since the atmospheric conditions and nightly calibrations can change, which could potentially affect the wavelength solution; the remaining reduction is based on the assumption that systematic effects will change smoothly (in wavelength or in time) which will not hold for different visits. A low-pass filter is then built to remove the remaining modal noise (varying both in time and wavelength) from each exposure (Fig steps, panel E). The reference spectrum is then rebuilt and divided to all exposures to isolate the planetary signal. Finally, a principal component analysis is applied to correct for any quasi-static residuals.

### 3.3. Modelling and extraction of the planetary signal

At this stage, it is not possible to directly extract the planetary spectrum from the data, considering that its features are buried into the noise. However, we can benefit from the fact that the spectrum encloses a large number of spectral features to amplify their collective signal. One approach would be to perform a cross-correlation with a mask around all lines that are expected to be found at the planet’s atmospheric condition. The cross-correlation product would then consist in the mean planetary line. This approach stems from the field of stellar radial velocity measurements (e.g. Baranne et al., 1979; Queloz, 1995) and has the advantage of necessitating minimal assumptions on the spectrum. Nevertheless, it can be difficult to interpret, and any quantitative measurements like the molecular abundance or the structure of the atmosphere are hardly inferred (Brogi & Line, 2019). Moreover, spectral lines probing different altitudes may not have identical shapes (due their different opacity), making their combination not trivial.

Another approach was chosen in the present analysis. It derives from the more recent template matching technique that has proven to be more effective for the measurements of the radial velocity of cool stars, where some molecular lines are not well defined (e.g. Anglada-Escudé & Butler, 2012; Astudillo-Defru et al., 2017). In this context, a stellar template obtained by combining all observations is used to correlate with each single observation in an iterative process. However, concerning planetary emission spectroscopy, it is generally not possible to produce such template due to the extremely low SNR; a forward modeling approach is therefore needed.

#### 3.3.1. Atmosphere models

3.3.1.1. Planetary models. In this analysis, we used the open source `python` package `petitRADTRANS` (Mollière et al., 2019) to generate the planet’s atmosphere spectra, as done

in Boucher et al. (2022), with the difference that it is in emission instead of transmission. This package has the capability to generate high resolution spectra ( $R = 250,000$  which can be later downgraded to SPIRou’s resolving power) in a time small enough to perform a statistical analysis. It uses a line-by-line opacity treatment in the context of high resolution retrievals and it can also generate low-resolution spectra much more rapidly using the correlated-k method.

`petitRADTRANS` can be used as a free-chemistry forward-model in the sense that all the chemical species abundances and the structure of atmosphere must be specified as an input, as opposed to equilibrium chemistry models where more global parameters (e.g. metallicity, C/O ratio) are used to calculate the abundances by minimizing the Gibbs free energy. We choose to include species that have dominant features in the SPIRou bandpass and that could be present in WASP-33 b’s atmosphere, namely  $\text{H}_2\text{O}$ ,  $\text{CO}$ ,  $\text{CO}_2$ , (Rothman et al., 2010)  $\text{FeH}$  (Bernath, 2020),  $\text{C}_2\text{H}_2$  (Rothman et al., 2013),  $\text{VO}$ ,  $\text{TiO}$  (see references in Mollière et al. 2019),  $\text{HCN}$  (Barber et al., 2014) and  $\text{OH}$  (Rothman et al., 2010). For the continuum opacities, we only fit for  $\text{H}^-$  (Gray, 2008), but collision induced absorption from  $\text{H}_2\text{--H}_2$  (Burrows & Volobuyev, 2003) and  $\text{H}_2\text{--He}$  (Borysow, 2002) were taken into consideration. We also allowed a grey opacity to account for the presence of clouds. Unlike in transmission, the exact shape of spectral lines in the emission spectrum highly depends on the temperature-pressure (TP) profile. Therefore, contrarily to Boucher et al. (2022), we used a modified Guillot profile (Mollière et al., 2019) — instead of a standard Guillot profile (Guillot, 2010) — to allow for more complex line shapes (see Mollière, 2017) with the internal temperature ( $T_{\text{int}}$ ) fixed to 500 K. The other parameters are :  $\gamma$ , the ratio between the optical and infrared mean opacities ;  $\delta$ , the ratio between the mean infrared opacity  $\kappa_{IR}$  and the surface gravity  $g$  ;  $T_{\text{eq}}$ , the equilibrium temperature.  $P_{\text{trans}}$ , the transition pressure and  $\alpha$ , a scale factor, are added in the modified Guillot profile to parametrize the upper layers of the atmosphere. We also fixed  $R_p$ , the planet’s radius, with the value from Table 1. Given all this, we are left with a model comprising 17 free parameters : 10 chemical species, 1 for the grey cloud opacity and 6 for the temperature profile.

In the high temperature regimes present in WASP-33 b’s atmosphere, molecules typically found in hot-Jupiters (e.g.  $\text{H}_2\text{O}$ ,  $\text{CO}_2$ , etc.) will become thermally dissociated (e.g., Tsai et al., 2021). To include this effect in our forward models, we opted for the empirical parametrization presented in Parmentier et al. (2018), where the abundance profile  $A$  is a combination of a constant abundance,  $A_0$ , and dissociated abundance,  $A_d$ , given by the relation

$$\frac{1}{A} = \left( \frac{1}{\sqrt{A_0}} + \frac{1}{\sqrt{A_d}} \right)^2. \quad (3.3.1)$$

The term  $A_d$  is expressed by a power law of the pressure, following the equation

$$A_d = 10^{-\gamma} P^\alpha 10^{\frac{\beta}{T}}, \quad (3.3.2)$$

where  $\gamma$ ,  $\alpha$  and  $\beta$  are free parameters, specific to each species. We applied a slight modification to this prescription by adding a dependency to  $A_0$  to allow the modified profile to shift toward higher abundances. Only the parameter  $\gamma$  is changed by adding a dependency to  $A_0$  given by,

$$\gamma = \log_{10} \frac{A_0}{A_{\text{Ref}}} - \gamma_{\text{Ref}}. \quad (3.3.3)$$

The thermal dissociation was applied to all species for which the parameters were specified in Table 1 of Parmentier et al. (2018), namely  $\text{H}_2$ ,  $\text{H}_2\text{O}$ ,  $\text{TiO}$ ,  $\text{VO}$ ,  $\text{H}^-$  and free electrons. The dissociated components of these molecules will be in atomic form, which do not have significant opacity in SPIRou’s bandpass, so we did not account for these. However,  $\text{H}_2$  is a major constituent of the atmosphere and will have a significant impact on the continuum opacities. Thus, we considered that all the hydrogen produced by this reaction stayed in its atomic form and did not recombine with other constituents.

3.3.1.2. Stellar model. The measurement of the planet emission signal is relative to its host star’s spectrum, hence it needs to be properly modeled. To do so, we used Phoenix HiRes synthetic spectra (Husser et al., 2013) at the metallicity,  $\log g$  and effective temperature values listed in Table 1. We also took into account the rotation of the star by applying a rotation kernel to each spectra based on stellar  $v \sin(i_*)$ .

### 3.3.2. Comparison with High Resolution data

At high resolution, the comparison between the models and the reduced data must take into account the impact of the reduction process. The procedure described here is the same as Boucher et al. (2022), but adapted for emission spectroscopy. To do so, we first reconstruct a full time-series using the principal components that were removed from Step F of Fig. 3. Then, the planetary spectrum model is injected directly in the reconstructed spectra via the multiplicative term

$$\left(\frac{R_p}{R_*}\right)^2 \frac{F_p(v_p)}{F_*(v_*)} + 1, \quad (3.3.4)$$

where  $F_p(v_p)$  and  $F_*(v_*)$  are the planet and star emitted flux, Doppler-shifted at their respective observed radial velocity, while  $R_p$  and  $R_*$  are their respective radii. This assumes that the transformations of the spectrum leading to Step F are also multiplicative. The observed radial velocity of the planet is given by

$$v_p(t) = v_{\text{bary}}(t) + v_{\text{sys}} + v_{\text{orb,P}}(t) - v_{\text{orb,*}}(t), \quad (3.3.5)$$

where  $v_{\text{bary}}$  is the barycentric Earth radial velocity and  $v_{\text{sys}}$  is the systemic velocity (see Table 1). The two other components,  $v_{\text{orb,P}}$  and  $v_{\text{orb,*}}$  are the radial velocities due to the orbital movement of the planet and the star respectively. Both are given by the equation

$$v_{\text{orb}}(t) = \pm K \sin[2\pi(\phi(t))], \quad (3.3.6)$$

with the positive version associated to the planet and the negative, to the star.  $K$  is the semi-amplitude (for the planet or the star) listed in Table 1 and  $\phi$  is the orbital phase.

Finally, the PCA correction (Steps G) is re-applied on the injected spectra with the difference that the principal components are not re-computed (see Boucher et al. 2022 for more details). This will remove the continuum and some smoothly varying features (e.g. wings of the lines), while modifying the shape of the remaining features in the planetary spectrum.

This procedure is primordial to allow for proper model comparison. In this analysis, we opted for the  $\ln \mathcal{L}$  prescription from Gibson et al. (2020), going as follow :

$$\ln \mathcal{L}_{\text{HR}} = -\frac{N}{2} \ln \frac{1}{N} \sum \frac{(f_i - m_i)^2}{\sigma_i^2}, \quad (3.3.7)$$

where the index  $i$  represents each spectral pixel for every integration in time whereas  $f_i$  and  $m_i$  are the fully processed spectra and models. The subscript HR refers to High-Resolution. Note that contrarily to Brogi & Line (2019), no distinction is made between orders; the  $\ln \mathcal{L}$  include all orders from all exposures. The estimation of the noise,  $\sigma_i$ , is determined following the same procedure as described in Boucher et al. (2022). In brief, for a given night, the noise is estimated empirically by computing the standard deviation over time on the PCA-corrected spectra (see Fig. 3, panel G), after the removal of a single component. Then, to account for the variation in SNR, the result is scaled, exposure-wise, by the median relative photon noise at each exposure (normalized by the median relative photon noise of all exposures).

Note also that the summation can be expanded in three terms :

$$\sum \frac{f_i^2}{\sigma_i^2} + \sum \frac{m_i^2}{\sigma_i^2} - 2 \sum \frac{f_i m_i}{\sigma_i^2}, \quad (3.3.8)$$

with the last term being equivalent to a weighted cross-correlation function (hereafter CCF), which is often used in such high-resolution studies. It is also interesting to see that the first term only depends on the data and the second term only on the model. Hence, for an atmospheric model with fixed parameters and varying  $v_p$ , the variation of the  $\ln \mathcal{L}$  will be mainly influenced by the CCF term. By doing so for each individual integration, a CCF map can be generated. This technique, the CCF method, is generally used to estimate the level of detection of a planet atmospheric signal.

### 3.3.3. Comparison with Low Resolution data

In this analysis, we will also use eclipse measurements of WASP-33 b taken by Haynes et al. (2015) with HST/WFC3. These are spectrophotometric measurements, so the standard

prescription for the  $\ln \mathcal{L}$  can be used, i.e.

$$\ln \mathcal{L}_{\text{LR}} = - \sum \frac{(f_i - m_i)^2}{\sigma_i^2}. \quad (3.3.9)$$

Here, LR stands for Low-Resolution. The modeled and observed eclipse depth in the spectral bin  $i$  are given respectively by  $m_i$  and  $f_i$  with

$$m_i = \left( \frac{R_p}{R_\star} \right)^2 \left[ \frac{F_p}{F_\star} \right]_i. \quad (3.3.10)$$

## 3.4. Results

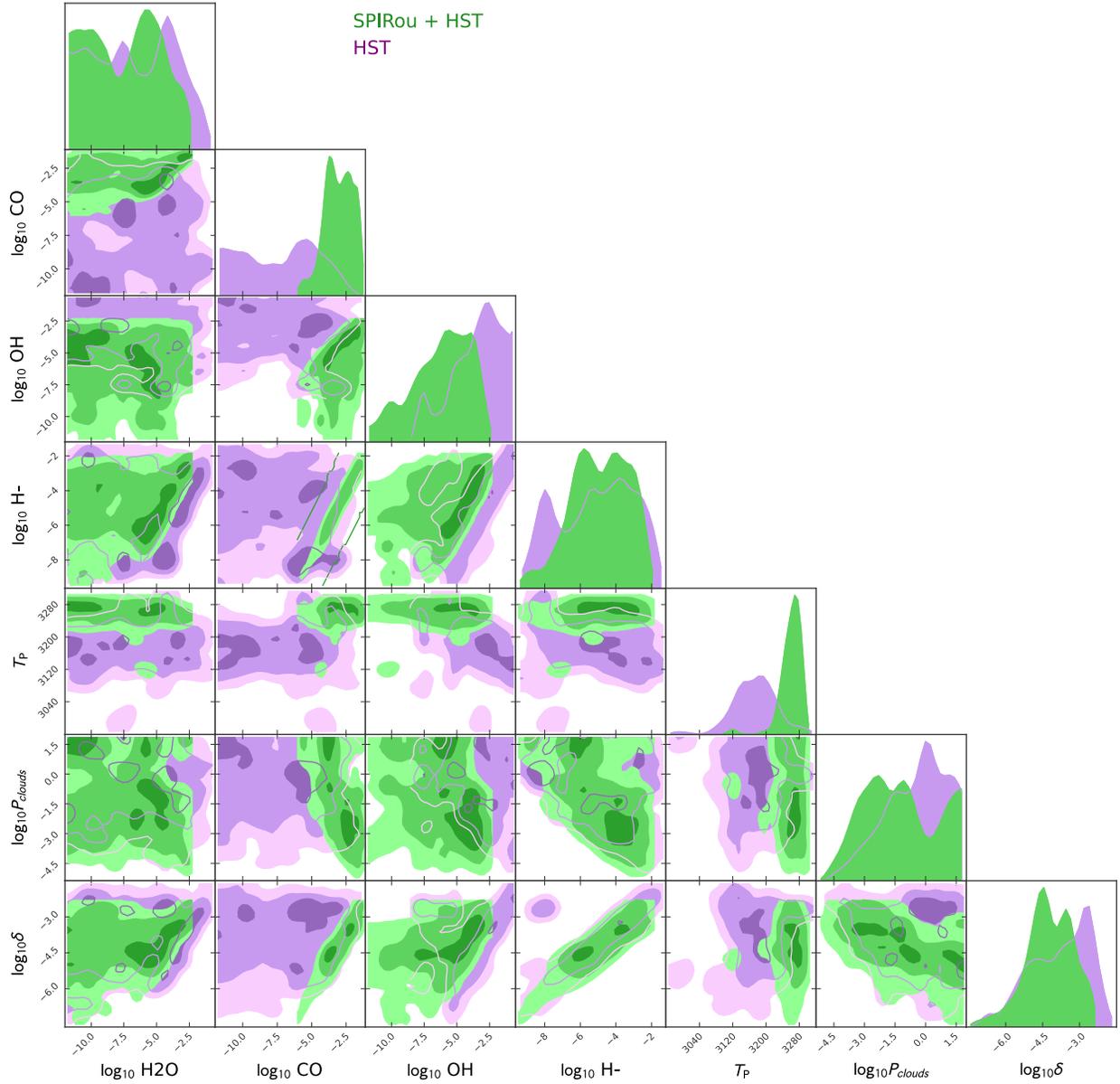
### 3.4.1. Retrievals

To initialize the search for planetary signal detection in SPIRou data with the CCF method, we first need a representative planet’s emission spectrum. We use the median TP-profile from a retrieval analysis on WFC3 eclipse data by Haynes et al. (2015) and compute an emission spectrum using `petitRADTRANS`. We tested for a model including only CO with an abundance that was producing strong lines without saturation ( $\text{VMR} \sim 10^{-4}$ ) and did the same for other molecules individually. This allowed to confirm the detection of CO with the CCF method (similarly to Fig. 8), consistent with the planetary orbital motion ( $K_p$  and  $v_{\text{sys}}$ ). We did not find any signal for  $\text{H}_2\text{O}$  nor any other species present in SPIRou’s bandpass. These preliminary steps ensured that the retrieval analysis could be initialized at  $K_p$  and  $v_{\text{sys}}$  where the planet’s atmospheric signal lies, considering the high numerical cost occurring at such large wavelength range and resolution.

In parallel, we performed another retrieval on WFC3 data alone, using `petitRADTRANS` (see §3.3) with the objective to narrow down the initialization of the atmospheric parameters, including the VMRs. It also constituted an opportunity to update the previous results given that our understanding of UHJ has evolved since their publication. More specifically, the opacity of  $\text{H}^-$  and the molecular dissociation were not taken into account, while both are expected to have a strong influence on the shape of the emission spectrum of WASP-33 b. The posterior distributions are shown in Fig. 4 (in purple).

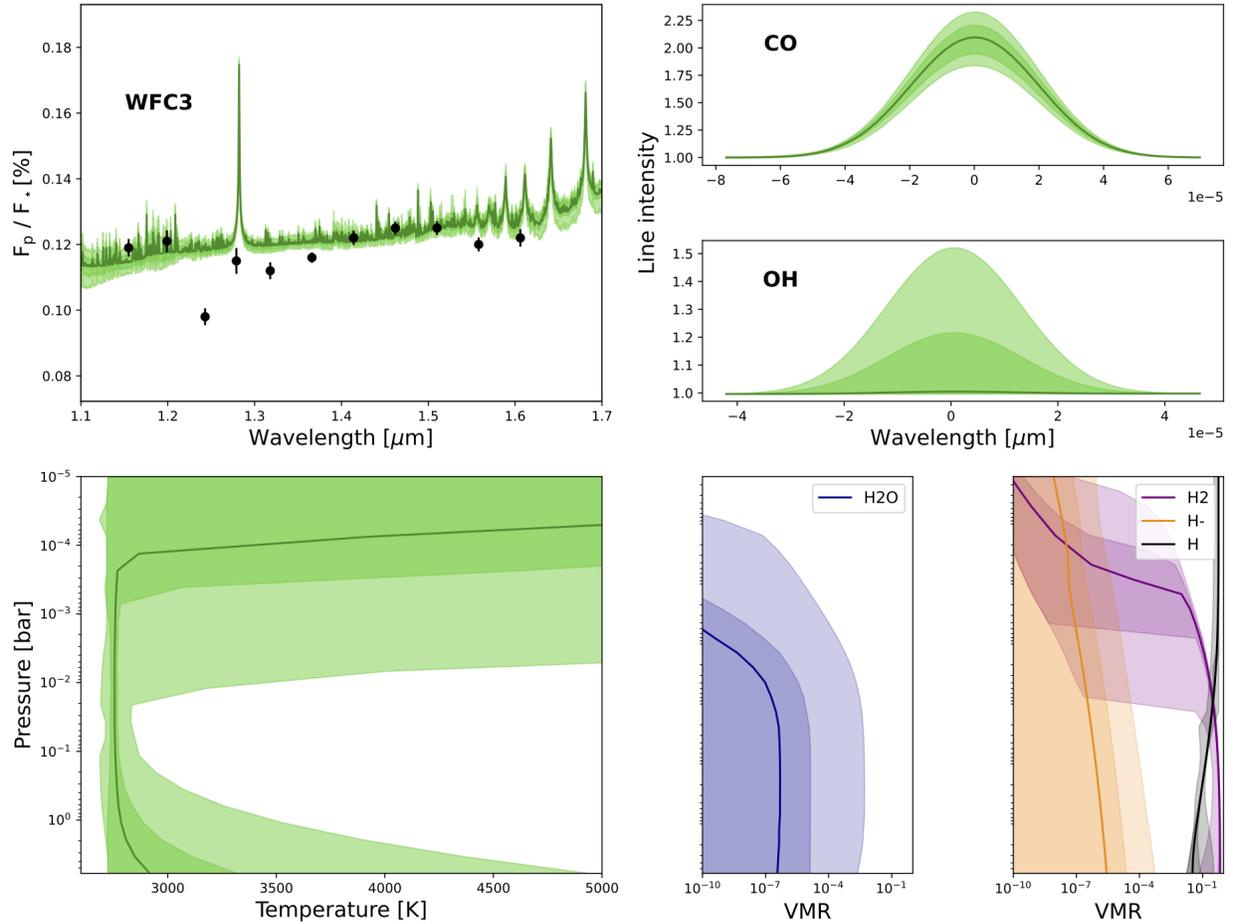
We then used this distribution to initialize the walkers from a joint analysis, using both the low resolution HST/WFC3 and our two high resolution SPIRou visits. The  $\ln \mathcal{L}$  was computed simply by summing the  $\ln \mathcal{L}_{\text{HR}}$  from Eq. 3.3.7 and  $\ln \mathcal{L}_{\text{LR}}$  from Eq. 3.3.9. The posterior distribution presented in Fig. 4 (in green) illustrates the improvement on the constraints when the high-resolution data are included.

For all the retrievals using `petitRADTRANS`, we used the `emcee` package (Foreman-Mackey et al., 2013) to explore the parameter space. This is a `python` implementation of a Markov Chain Monte Carlo (MCMC) using the Goodman and Weare’s algorithm (Goodman &



**Figure 4** – Posterior distributions of a sub-sample of parameters for the combined retrieval (green) of HST/WFC3 and SPIRou. The result of the retrieval using only HST/WFC3 are presented in purple. The contours represent the regions of confidence for 1, 2 and  $3\sigma$ , going from the darkest to the lightest color.

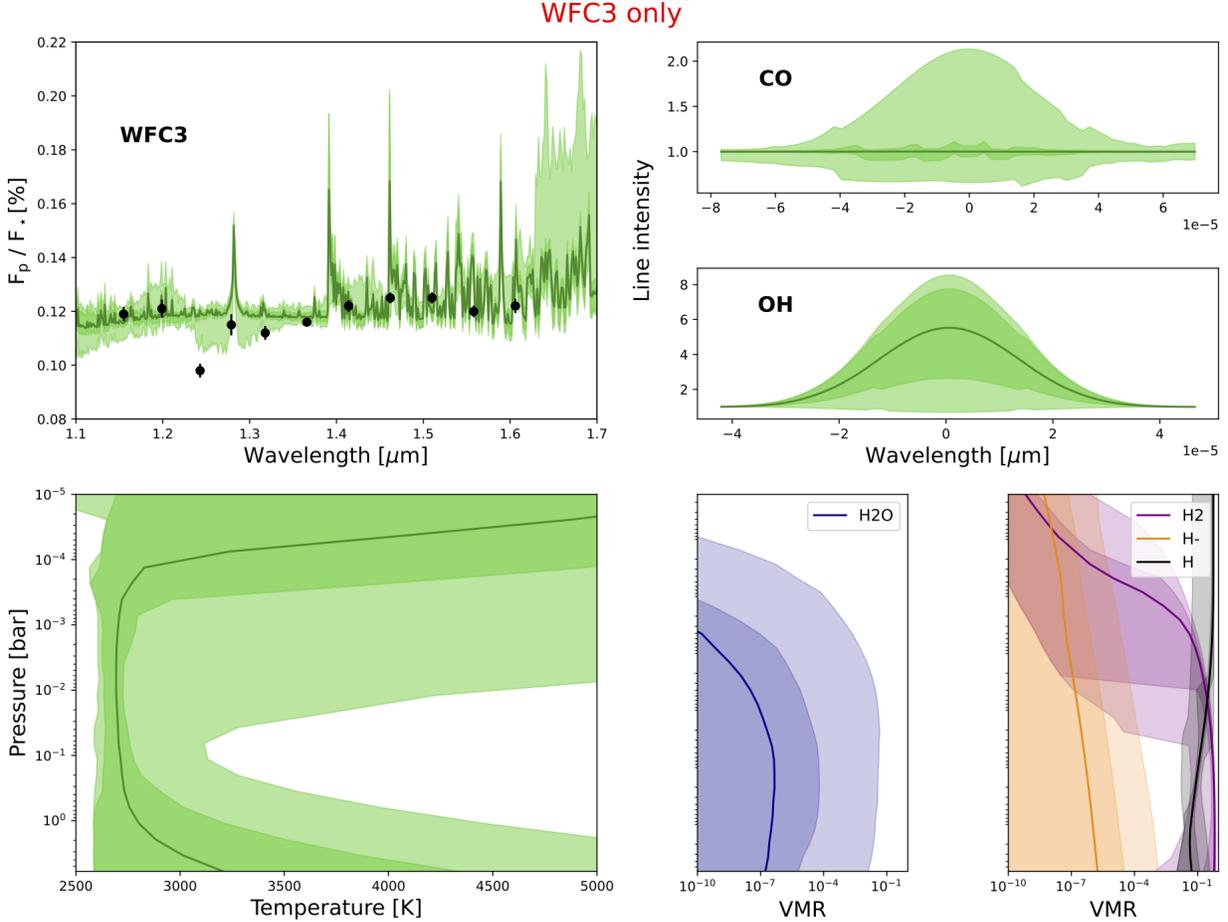
Weare, 2010). We used 50 walkers and waited until convergence (i.e., once the number of steps reach at least ten times the auto-correlation time for all parameters). The resulting posterior distributions are shown in Fig. 4 and the retrieved parameters values are listed in Table 3. Note that we did not assume a constant VMR for  $\text{H}_2\text{O}$ ,  $\text{TiO}$ ,  $\text{VO}$  and  $\text{H}^-$ , so the values in the table refer to the abundance in regions where it is not affected by dissociation



**Figure 5** – Distributions of spectra and profiles resulting from the combined retrieval of HST/WFC3 and SPIRou data. The top left panel presents the resulting distribution of eclipse spectra, at low resolution for visualisation purposes. The WFC3 observed data points are added for comparison. At the top right corner of the figure, a close-up of a selected major line for CO (top) and OH (bottom) is shown at the resolution of SPIRou (70,000). The bottom panels illustrate the distributions of TP-profiles (left) and VMR profiles (right). The VMR profiles are separated in two panels for clarity. Species with constant abundances are not shown here. TiO and VO are also omitted to improve visibility, but have similar profiles as the H<sub>2</sub>O one. For all panels, the solid line represents the median of a selection of samples at each wavelength or pressure and the shaded regions mark the 1- (dark) and 2- $\sigma$  (light) uncertainties.

(at higher pressures). Nevertheless, they both are upper limits (apart from H<sup>-</sup>) and the dissociation can only lower them as shown by the H<sub>2</sub>O profile in Fig. 5 (top-right panel).

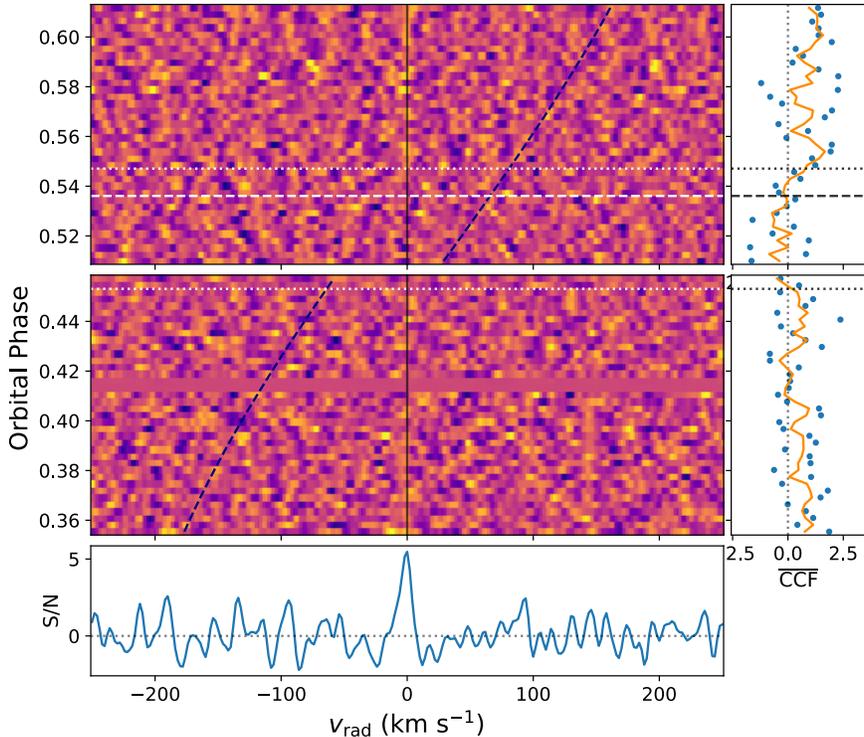
Both posterior distributions (for the HST/WFC3 only and the joint analysis) show a strong correlation between the H<sup>-</sup>, the OH and the parameter  $\delta$  of the TP-profile. This is due to their cohabitation in WFC3 bandpass and the degeneracy between the temperature



**Figure 6** – Distributions of spectra and profiles resulting from the retrieval using only HST/WFC3 observations. The description is the same as Fig. 5.

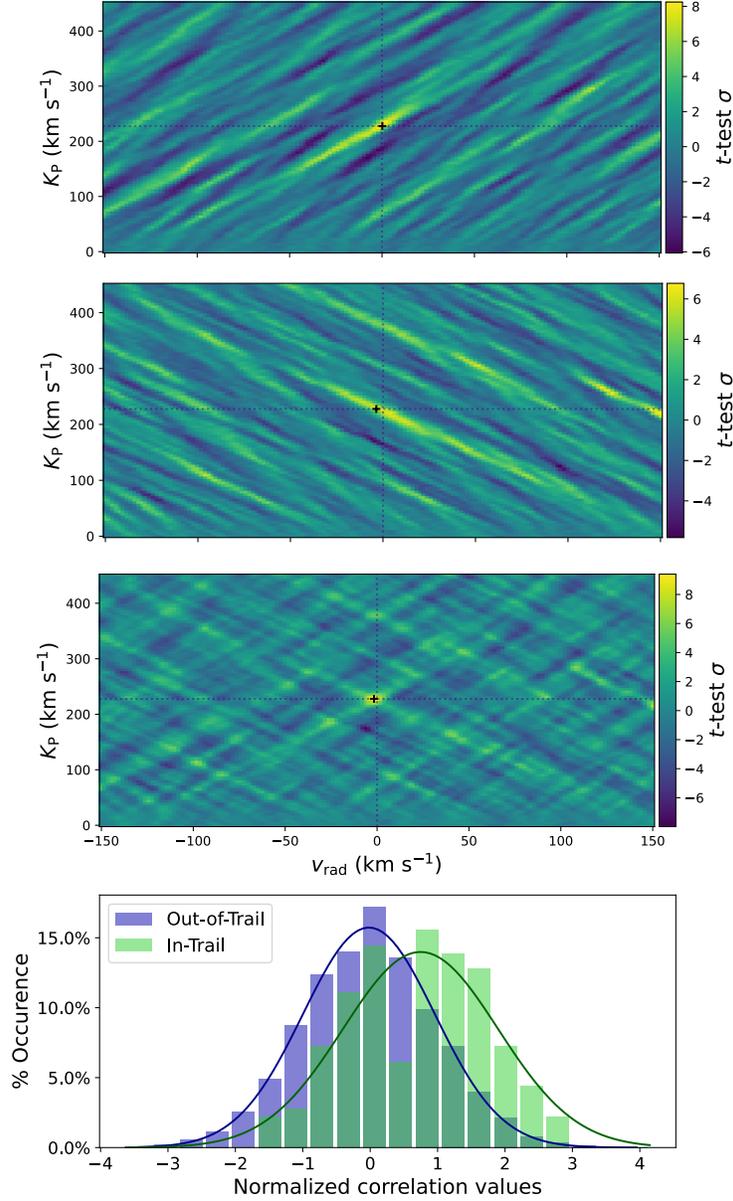
gradient, encoded by  $\delta$ , and the abundance. The same effect is seen for CO VMR in the combined retrieval.

The distribution of spectra and profiles retrieved for the HST+SPIRou case is shown in Fig. 5. The equivalent is shown in Fig. 6 for the retrieval using only WFC3 observations. It confirms that the addition of SPIRou data is able to improve the constraint on the TP-profile (between  $\sim 3 \times 10^{-1}$  and  $\sim 10^{-3}$  bar) while remaining consistent with the low-resolution observations. For the combined retrieval, all realisations of the TP-profile included a thermal inversion, which was not the case for WFC3 alone. This inversion can also be inferred directly from the CO line shape clearly seen as emission lines in Fig. 5. In comparison, the OH line profiles (Fig. 5, bottom right) demonstrate that the OH lower boundary is not driven by the high-resolution data since realizations without OH lines are not excluded. In fact, SPIRou’s data eliminate the models with higher OH abundance, which was favored by WFC3 as highlighted by the OH marginalized distribution in Fig. 4 and in Fig. 6 by the single OH line and the strong lines in the redmost part of low-resolution (WFC3) panel. Finally, the most



**Figure 7** – Cross-correlation function time-series in WASP-33 b rest frame. The normalized CCFs computed at each exposures are shown in colors in the left panels as a function of radial velocity in the planetary rest frame. The vertical axis represents the orbital phase corresponding to each integration. The black dashed line indicates the BERV, where any remaining telluric contamination should be found. The first and last contact are represented by the white dotted lines. The top panel includes observation during total eclipse, which are located below the white dashed line. The bottom panel shows the sum of CCF along all exposures, normalized by the standard deviation away from the detection peak. The right panels present the mean value of 3 spectral pixels around  $v_{rad} = 0 \text{ km s}^{-1}$  (blue dots) and the binned value for 3 exposures (orange line) as a function of phase. The horizontal black lines are the continuity of the white lines in the left panels.

relevant abundance profiles illustrate the effect of molecular dissociation; the  $\text{H}_2$  abundance decrease is counterbalanced by an increase of atomic H and the  $\text{H}_2\text{O}$  is practically absent at pressures lower than  $10^{-3}$ . Note that the profiles for  $\text{H}_2$  and H are entirely driven by the TP-profiles using the parametrization presented in §3.3. Concerning the  $\text{H}^-$  abundance profile, we can see that, despite the parametrization from Parmentier et al. (2018), it could be well represented by a constant abundance. Although not shown in Fig. 5 (top right), TiO and VO profiles show the same/similar behavior as  $\text{H}_2\text{O}$ , with the difference mainly at high pressures. Note also that the pressures between  $10^{-1}$  and  $10^0$  are barely probed (and not probed at higher pressures). The same goes for the lowest pressures ( $P \lesssim 10^{-4}$ ), where the



**Figure 8** – Welch  $t$ -test results for the best-fit CO model. The first three panels show the scaled values of the Welch  $t$ -test for the two individual visits (first and second panels) and their combination (third panel). The maps present the significance as a function of the model’s orbital semi-amplitude and systemic radial velocity. The dotted lines indicates the expected values from references in Table 1 and the cross indicates the maximum significance. The bottom panel presents the distribution of CCF values for the combined visits, near and away from the planet radial velocity — labeled respectively as “in trail” and “out of trail”, respectively. The “in trail” region comprises velocities of  $\pm 3$  km/s around the expected value. The two curves represents the best fit obtained for a normal distribution.

Table 3. MCMC Retrieval Parameter Priors and Results

Parameter	Priors	HST only	SPIRou+HST	Unit
$\log_{10} \text{H}_2\text{O}$	$\mathcal{U}(-12, -1)$	$< -1.88$	$< -3.09$	
$\log_{10} \text{CO}$	$\mathcal{U}(-12, -1)$	$< -3.05$	$-4.07^{+1.51}_{-0.60}$	
$\log_{10} \text{CO}_2$	$\mathcal{U}(-12, -1)$	$< -0.50$	$< -3.21$	
$\log_{10} \text{FeH}$	$\mathcal{U}(-12, -1)$	$< -6.80$	$< -7.49$	
$\log_{10} \text{TiO}$	$\mathcal{U}(-12, -1)$	$< -3.66$	$< -2.81$	
$\log_{10} \text{VO}$	$\mathcal{U}(-12, -1)$	$< -2.94$	$< -2.40$	
$\log_{10} \text{C}_2\text{H}_2$	$\mathcal{U}(-12, -1)$	$< -3.99$	$< -2.35$	
$\log_{10} \text{HCN}$	$\mathcal{U}(-12, -1)$	$< -3.46$	$< -3.54$	
$\log_{10} \text{OH}$	$\mathcal{U}(-12, -1)$	$-2.49^{+1.99}_{-1.72}$	$-5.02^{+1.67}_{-2.66}$	
$\log_{10} \text{H}^-$	$\mathcal{U}(-12, -1)$	$-3.40^{+1.66}_{-2.89}$	$-3.73^{+0.90}_{-2.41}$	
$T_{\text{P}}$	$\mathcal{U}(3000, 3500)$	$3170^{+57}_{-36}$	$3266^{+21}_{-18}$	K
$\log_{10} P_{\text{cloud}}$	$\mathcal{U}(-6.0, 2.0)$	$> -3.36$	$> -3.93$	bar
$K_{\text{P}}$	$\mathcal{U}(195, 255)$	...	$226.09^{+0.96}_{-1.06}$	$\text{km s}^{-1}$
$v_{\text{rad}}$	$\mathcal{U}(-30, 30)$	...	$-0.66^{+0.52}_{-0.51}$	$\text{km s}^{-1}$
$\log_{10} \delta$	$\mathcal{U}(-10.0, -1.0)$	$-2.93^{+0.75}_{-1.63}$	$-4.38^{+1.19}_{-0.59}$	
$\gamma$	$\mathcal{U}(0, 4)$	$1.86^{+0.91}_{-0.51}$	$> 1.53$	
$\log_{10} P_{\text{trans}}$	$\mathcal{U}(-8.0, 0.0)$	$< -2.01$	$< -3.28$	bar
$\alpha$	$\mathcal{U}(-1.0, 1.0)$	...	...	

Note. — The marginalized parameters from the likelihood analysis with their  $\pm 1\sigma$  errors, or their  $2\sigma$  upper or lower limits. The abundances of the different species refer to the VMR.

profile is driven by the parameters  $P_{\text{trans}}$  and  $\gamma$  which are not constrained (see Table 3). This translates into large uncertainties on the TP-profile at these pressures.

The Figures 5 and 6 also show the comparison between the retrieved spectra and WFC3 data. The models cannot reproduce some of the observed spectral bins, especially around the hydrogen stellar line at  $1.28 \mu\text{m}$ . As mentioned in Haynes et al. (2015), this is likely attributable to the variations of the hydrogen line due to the  $\delta$ -Scuti nature of the star. These effects could not be corrected without access to other hydrogen lines.

### 3.4.2. Cross-correlations

Now that the best fit model is determined, it is possible to use a cross-correlation analysis as an alternative way to quantify the detection significance of the CO signal. We used the best fit model to generate a high-resolution planetary emission spectrum and compared it to SPIRou’s observations of each nights, following the procedure described in §3.3.2. The values of the cross-correlation term (see Eq. 3.3.8) as a function of phase are presented in Fig. 7. The planetary signal is distinguishable in the first visit (top panel), starting to appear during

egress — this is clearer when looking at the CCF peak value (right panel). The signal in the second visit (bottom panel) is less apparent in an exposure-by-exposure basis, but the peak values are globally positives outside of the eclipse. Moreover, when combined together, they add constructively to result in a SNR= 5 detection peak at the expected systemic velocity.

We also put the detection through the Welch  $t$ -test (Welch, 1947), by splitting the CCF values in two populations (around and away from the planetary velocity). The test quantifies level of disparity between the two distributions, assuming a normal distribution. This is repeated for a range of  $K_P$  and  $v_{\text{sys}}$ , leading to Fig. 8. Both visits show a clear detection, peaking almost exactly at the expected orbital parameters. Their co-addition also results in a well constraint signal due to the complementarity of the egress and ingress in velocity space. Note that this detection is entirely driven by the CO lines, since no other molecules were detected in SPIRou’s observations.

## 3.5. Discussion

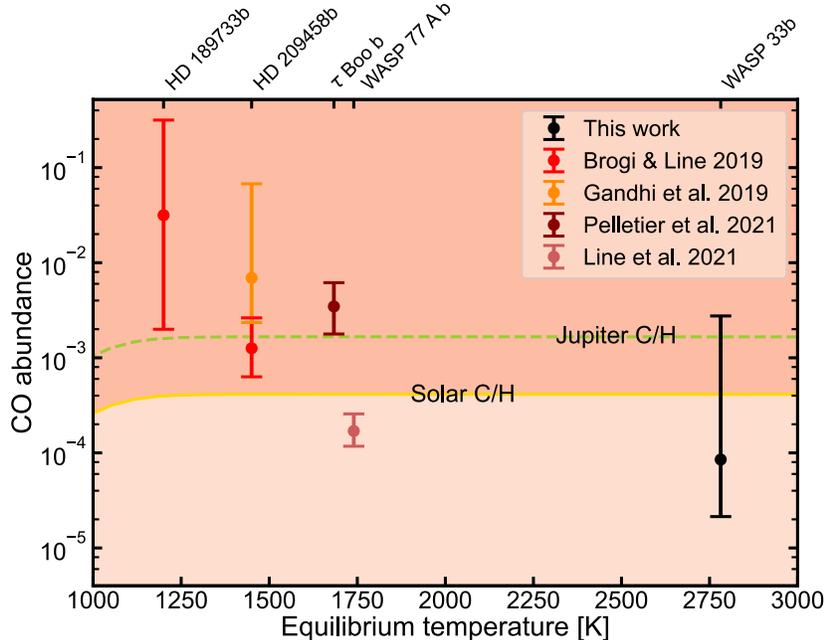
### 3.5.1. Comparison with previous observations

Our results are also in line with the previous detections of CO with GIANO (Yan et al., 2022a) and ARIES (van Sluijs et al., 2022). They all found a signal in emission, implying a temperature inversion. Further comparison are hardly done considering that no constraints were provided for the CO abundance, although an attempt was made by van Sluijs et al. (2022) to estimate the TP-profile. The latter is consistent with the profile retrieved here. None of these studies found evidence for OH nor H<sub>2</sub>O, which is also in agreement with our CCF analysis. This does not question the detection of OH from Nugroho et al. (2021) since we do not have the sensitivity to exclude the presence of OH emission lines, as shown in Fig. 5. In fact, the combined retrieval favors the presence of OH, however, this is primarily driven by the WFC3 data. The retrieved upper limits of the other molecules are also consistent with the non-detections found in the literature.

We do not explore the phase variations highlighted by van Sluijs et al. (2022) due to our limited SNR. Nevertheless, our observations do not exclude a potential phase variation. In fact, the SNR is stronger in the egress than the ingress, as shown by Fig. 7 and 8. However, it is not clear if it can be attributed to the poorer conditions during the night or to actual phase variation. We also did not fit for phase-dependent parameters which is beyond the scope of this paper.

### 3.5.2. Implications

This work reinforces the use of CO as a good probe of atmospheric structure, especially in the context of UHJ. Indeed, it does not dissociate as much as other species like H<sub>2</sub>O,



**Figure 9** – Comparison between direct measurements of the CO abundance in the dayside of exoplanets. This figure has been updated from Pelletier et al. (2021) Fig. 11. The gold solid line corresponds to the CO abundance that would be expected with a Solar C/H ratio, assuming chemical equilibrium at the planet equilibrium temperatures. This line marks the boundary between the light and dark area covering respectively the sub- and super-solar parameter space. The green dashed line is analog to the gold line for Jupiter’s C/H ratio.

TiO, VO (in the pressures typically probed) and it is not influenced as much by the local conditions as atomic species or other molecules like hybrids (e.g. OH, FeH). Moreover, it is not found in early-type host stars, making it less prone to contamination as opposed to other atomic species (e.g. Cont et al., 2022; Cont et al., 2021; Nugroho et al., 2020; Herman et al., 2022), especially in the situation where the star is pulsing. Finally, rapidly rotating stars are particularly well designated for high-resolution spectroscopy because their broadened lines will be more easily distinguished from the narrower planetary lines. In this regard, CO represent a good benchmark to compare UHJ and, by extension, Hot Jupiters. Fig. 9 presents the comparison between WASP-33 b and the few Hot-Jupiters with direct CO measurements from high-resolution spectroscopy, i.e., HD 189733 b (Brogi & Line, 2019), HD 209458 b (Brogi & Line, 2019; Gandhi et al., 2019),  $\tau$  Boo b (Pelletier et al., 2021), WASP-77 A b (Line et al., 2021). In contrast with the super-solar constraints for the three coldest of these planets, our measurement favors a slightly sub-solar CO abundance, similarly to WASP-77 A b. This suggest a diversity in the atmosphere composition of Hot-Jupiters. However, the precision of our measurement limits further conclusions, given that it is still consistent within  $1\sigma$  with a Solar and Jupiter C/H.

The temperature inversion that we confirmed by the addition of SPIRou data also agrees with model predictions. Furthermore, it is not in contradiction with our upper limit on the TiO and VO, which were suggested as UV absorbers responsible for inverted profiles (e.g., Hubeny et al., 2003; Fortney et al., 2008). Indeed, at such high temperatures, the combination of  $\text{H}^-$ , the absorption by atomic metals or other species and the flux of early type host stars is enough to cause it (Lothringer et al., 2018). Water dissociation could also contribute to the temperature inversion (Mansfield et al., 2021).

This study also leans toward an implicit constraint on  $\text{H}^-$  and OH, based on the combination of WFC3 and SPIRou data. It is in line with the current consensus that  $\text{H}^-$  should have a major effect in UHJ atmospheres, especially in the WFC3 bandpass (e.g., Lothringer et al., 2018; Arcangeli et al., 2018). It also emphasizes the potential of combining high-resolution spectroscopy techniques based on the planet’s orbital motion with low resolution spectrophotometric data. The former being sensitive to temperature gradients and able to identify individual molecules, whereas the latter being more sensitive to the continuum and the ratio between the planet and the star. However, note that the abundance retrieved on  $\text{H}^-$  opacity is influenced by the abundance of free electrons which was fixed here to  $\text{VMR}_{e^-} = 10^{-6}$ . Nevertheless, this does not question its presence in the observed emission spectrum of WASP-33 b. The constraint on OH is also expected as an indicator of molecular dissociation.

We did not attempt to determine the C/O ratio since a considerable part of these species is found in atomic form, which we can’t constrain. Despite that, CO constitutes one of the main carbon bearers of the atmosphere at these temperatures (e.g., Zhang et al., 2018; Tsai et al., 2021), while the other molecules containing oxygen are more prone to dissociation. Note also that the absence of water is not necessarily attributable to a high C/O, contrarily to Pelletier et al. (2021).

### 3.5.3. Further improvements

The parametrization used for the abundance profiles represents a balance between free- and equilibrium-chemistry. However, in this case, the coefficient ( $\alpha$ ,  $\gamma$  and  $\beta$ ) were predetermined on equilibrium chemistry calculations. These could be left as free parameters for the most relevant molecules, like water. An alternative solution would be to use allow for combination of a constant abundance and the predetermined parametrized profiles, using a weighting factor. This has the advantage to only require one additional variable instead of three. Furthermore, the inclusion of *Spitzer* data in the analysis might improve even more the constraints on the TP-profile and some abundances, like  $\text{CO}_2$ .

### 3.6. Conclusion

In this work, we extracted the emission spectrum of WASP-33 b from two nights of high-resolution observations, targeting the eclipse ingress and egress separately, using SPIRou at the CFHT. We were able to confirm the detection of CO in emission, implying the presence of a thermal inversion on the day-side of WASP-33 b. We also combined these high resolution data with spectrophotometric observations from HST/WFC3 in a statistical analysis. This led to a first estimate of the CO VMR of  $\log_{10} \text{CO} = -4.07^{+1.51}_{-0.60}$  on WASP-33b, and to constraints on the TP-profile. We also obtained an upper limit on the water abundance, with  $\log_{10} \text{H}_2\text{O} < -3.09$ .

We emphasize that CO is well suited to probe and compare atmospheres of Hot-Jupiters since it is less influenced by the local conditions, as opposed to  $\text{H}_2\text{O}$ ,  $\text{CO}_2$  or OH. Simulations with global circulation models also predict that it should be well mixed within the atmosphere, allowing for potential phase dependent comparisons. This could be achieved by further observations with SPIRou.

This study highlights the advantages of combining high-resolution ground-based spectroscopy with low-resolution spectro-photometric observations. With the advent of the James Webb Space Telescope (JWST), direct comparison (in the same bandpass) between these two methods is now possible for a diversity of molecules like CO. In this regard, SPIRou is well positioned given that it shares the wavelength domain of the NIRISS (Near Infrared Imager and Slitless Spectrograph ; Doyon et al., in prep) SOSS (Single Object Slitless Spectroscopy ; Albert et al., in prep) mode.

Acknowledgments. This project was undertaken with the financial support of the Canadian Space Agency (CSA-ASC) and the *Fonds de Recherche du Québec en Nature et Technologies* (FRQNT). A.D.B. and S.P. thank the Technologies for Exo-Planetary Science (TEPS) CREATE program, without whom this research would not be possible. The authors also acknowledge financial and social support of 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 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. This paper makes use of data from the first public release of the WASP data (Butters et al., 2010) as provided by the WASP consortium and services at the NASA Exoplanet Archive, which is

operated by the California Institute of Technology, under contract with the National Aeronautics and Space Administration under the Exoplanet Exploration Program. This research has made use of the NASA Exoplanet Archive, which is operated by the California Institute of Technology, under contract with the National Aeronautics and Space Administration under the Exoplanet Exploration Program.

*Software :*

- CFHT (SPIRou)
- HST (WFC3)
- Exoplanet Archive

*Facilities :*

- `astropy` (Astropy Collaboration et al., 2013)
- `scipy` (Virtanen et al., 2020)
- `ipython` (Pérez & Granger, 2007)
- `matplotlib` (Hunter, 2007)
- `numpy` (Harris et al., 2020)
- `ExoFile` (Darveau-Bernier et al., in prep. <https://github.com/AntoineDarveau/exofile>)
- `pygtc` (Bocquet & Carter, 2016)) `APER0` (Cook et al., in prep.)



# Chapitre 4

---

## ATOCA : an algorithm to treat order contamination. Application to the NIRISS SOSS mode.

Ce manuscrit a été publié dans la revue arbitrée *Publication of the Astronomical Society of the Pacific* le 28 septembre 2022 (# 2022PASP..134i4502D).

ANTOINE DARVEAU-BERNIER<sup>1</sup>, LOÏC ALBERT<sup>1</sup>, GEERT JAN TALENS<sup>2</sup> DAVID  
LAFRENIÈRE<sup>1</sup> MICHAEL RADICA<sup>1</sup> RENÉ DOYON<sup>1</sup> NEIL J. COOK<sup>1</sup> JASON F. ROWE<sup>3, 1</sup>  
ROMAIN ALLART<sup>1</sup> ÉTIENNE ARTIGAU<sup>1</sup> BJÖRN BENNEKE<sup>1</sup> NICOLAS COWAN<sup>4, 5</sup> LISA  
DANG<sup>5, 1</sup> NÉSTOR ESPINOZA<sup>6, 7</sup> DOUG JOHNSTONE<sup>8, 9</sup> LISA KALTENEGGER<sup>10</sup> OLIVIA  
LIM<sup>1</sup> TYLER PAULY<sup>6</sup> STEFAN PELLETIER<sup>1</sup> CAROLINE PIAULET<sup>1</sup> ARPITA ROY<sup>6, 7</sup>  
PIERRE-ALEXIS ROY<sup>1</sup> JARED SPLINTER<sup>4</sup> JAKE TAYLOR<sup>1</sup> JAKE D. TURNER<sup>11</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> Department of Astrophysical Sciences, Princeton University, 4 Ivy Lane, Princeton, NJ 08544, USA

<sup>3</sup> Bishop's University, 2600 College Street, Sherbrooke, QC J1M 1Z7, Canada

<sup>4</sup> Department of Earth & Planetary Sciences, McGill University, 3450 rue University, Montréal, QC H3A  
0E8, Canada

<sup>5</sup> Department of Physics, McGill University, 3600 rue University, Montréal, QC H3A 2T8, Canada

<sup>6</sup> Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA

<sup>7</sup> Department of Physics & Astronomy, Johns Hopkins University, 3400 N Charles St, Baltimore, MD  
21218, USA

<sup>8</sup> NRC Herzberg Astronomy and Astrophysics, 5071 West Saanich Rd, Victoria, BC, V9E 2E7, Canada

<sup>9</sup> Department of Physics and Astronomy, University of Victoria, Victoria, BC, V8P 5C2, Canada

<sup>10</sup> Department of Astronomy and Carl Sagan Institute, Cornell University, 302 Space Sciences Building,  
Ithaca, NY 14853, USA

<sup>11</sup> Department of Astronomy and Carl Sagan Institute, Cornell University, Ithaca, NY 14853, USA

## Contribution de l'étudiant



**Figure 1** – Logo de ATOCA. Crédits : Anne Boucher

Ce second article concerne le développement d'un algorithme d'extraction de données spectrophotométriques dédié au mode SOSS de NIRISS. Celui-ci s'intitule ATOCA, pour « Algorithme de Traitement d'Ordres ContAminés » ou, en anglais, *Algorithm to Treat Order ContAmination*. Le mot « atoca », d'origine autochtone, est un québécoisisme pour désigner la canneberge, représentée par le logo illustré à la figure 1.

L'idée à la base de l'algorithme a été proposée originalement par David Lafrenière. J'ai dû l'adapter pour mieux répondre aux spécificités de NIRISS/SOSS, comme les résolutions différentes entre les ordres, l'inclinaison de l'axe spectral et la précision requise pour l'analyse d'atmosphères d'exoplanètes.

J'ai donc développé l'entièreté de ATOCA. Une fois testé et validé, j'ai pu l'implémenter dans l'algorithme officiel de réduction de JWST<sup>1</sup> avec l'aide de Geert Jan Talens et Tyler Pauly. Mes discussions avec Loïc Albert et David Lafrenière ont grandement contribué à la conception et à l'amélioration de ATOCA.

J'ai écrit le manuscrit et produit les figures (excepté 2) de l'article. Les co-auteurs, principalement Loïc Albert, Geert Jan Talens, Michael Radica, David Lafrenière et Jason Rowe, ont apporté plusieurs commentaires et suggestions qui ont permis de perfectionner le résultat final.

Les codes développés durant cette thèse sont basés sur des bibliothèques plus générales, incluant `numpy`, `scipy`, `astropy`, etc.

1. <https://github.com/spacetelescope/jwst>

Finalement, il va sans dire que mes conversations récurrentes avec mes codirecteurs, David Lafrenière et René Doyon, ont eu un impact majeur sur la réussite de ce projet.

ABSTRACT. After a successful launch, the James Webb Space Telescope is preparing to undertake one of its principal mission objectives, the characterization of the atmospheres of exoplanets. The Single Object Slitless Spectroscopy (SOSS) mode of the Near Infrared Imager and Slitless Spectrograph (NIRISS) is the only observing mode that has been specifically designed for this objective. It features a wide simultaneous spectral range (0.6–2.8  $\mu\text{m}$ ) through two spectral diffraction orders. However, due to mechanical constraints, these two orders overlap slightly over a short range, potentially introducing a “contamination” signal in the extracted spectrum. We show that for a typical box extraction, this contaminating signal amounts to 1% or less over the 1.6–2.8  $\mu\text{m}$  range (order 1), and up to 1% over the 0.85–0.95  $\mu\text{m}$  range (order 2). For observations of exoplanet atmospheres (transits, eclipses or phase curves) where only temporal variations in flux matter, the contamination signal typically biases the results by order of 1% of the planetary atmosphere spectral features strength. To address this problem, we developed the Algorithm to Treat Order Contamination (ATOCA). By constructing a linear model of each pixel on the detector, treating the underlying incident spectrum as a free variable, ATOCA is able to perform a simultaneous extraction of both orders. We show that, given appropriate estimates of the spatial trace profiles, the throughputs, the wavelength solutions, as well as the spectral resolution kernels for each order, it is possible to obtain an extracted spectrum accurate to within 10 ppm over the full spectral range.

Keywords *Unified Astronomy Thesaurus concepts* : Spectrophotometry (1556); Space telescopes (1547); Exoplanet atmospheres (487); Astronomy data reduction (1861); Transmission spectroscopy (2133); Astronomy data modeling (1859);

## 4.1. Introduction

One of the key observing modes of the Near Infrared and Slitless Spectrograph (NIRISS, Doyon et al., in prep) onboard the James Webb Space Telescope (JWST) is the Single Object Slitless Spectroscopy (SOSS) mode (Albert et al., in prep). It enables time-series spectroscopy in the 0.6–2.8  $\mu\text{m}$  range for bright targets, which is of particular use for exoplanet transit spectroscopy. Indeed, simulations have demonstrated SOSS as a key mode to use with JWST on the brightest exoplanet targets (Greene et al., 2016; Batalha & Line, 2017; Louie et al., 2018; Schlawin et al., 2018) and it has been selected by multiple Cycle 1 programs, including the Early Release Science program DD-ERS 1366 (Batalha et al., 2017; Bean et al., 2018), as well as many Guaranteed Time Observations, e.g., GTO 1201 (Lafreniere, 2017) and the General Observer Programs 1935 (Kempton et al., 2021), 2062 (Mayo et al., 2021), 2113 (Espinoza et al., 2021b), 2589 (Lim et al., 2021), 2594 (Spake et al., 2021a) and 2722 (Madhusudhan et al., 2021). SOSS uses the GR700XD cross-dispersion grating prism (Doyon et al., in prep) in the pupil wheel of NIRISS to produce a series of three spectral traces :

order 1 ( $0.83\text{--}2.8\ \mu\text{m}$ ), order 2 ( $0.6\text{--}1.4\ \mu\text{m}$ ) and order 3 ( $0.6\text{--}0.95\ \mu\text{m}$ ). In practice, order 3 does not warrant much consideration due to its faint signal and the fact that it does not increase the wavelength domain. A slight (22 pixel wide) defocus along the spatial axis is purposely included to enable observations of bright targets without saturating the detector pixels. Mechanical constraints in the thickness of the GR700XD element at the design phase prevented the first and second orders from being fully separated, resulting in an overlap by about half the trace width towards the red wavelength ends of the traces (See Figure 2)

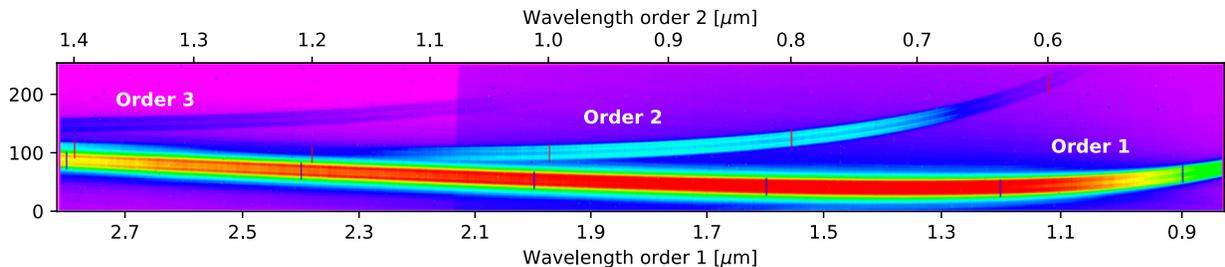
As a result, established methods for spectrum extraction cannot be applied directly to the regions affected by contamination. Typically, at high signal-to-noise, a box-extraction method (de Boer & Snijders, 1981) is preferred, which is performed by simply summing over the spatial axis all pixels located within a fixed-width aperture. This technique has been utilized in many space-based relative spectral measurements of exoplanets to date, for example : transits, eclipses or phase curves (e.g., Deming et al., 2013; Wakeford et al., 2013; Sing et al., 2015; Evans et al., 2017) as well as in many ground-based observations (e.g., Jordán et al., 2013; Diamond-Lowe et al., 2018). The main advantages are the fact that a box extraction is easy to implement and that it is less prone to modeling errors. On the other hand, at lower signal-to-noise, an optimal extraction (Horne, 1986; Robertson, 1986; Marsh, 1989) is often better. Indeed, by weighting the pixels according to their relative contribution to the signal, a better precision can be reached. This requires the determination of a spatial profile, which is a delicate task that can introduce biases in the resulting spectra (Horne, 1986; Jordán et al., 2013). Nevertheless, it is still used in the exoplanet community to perform spectrophotometric measurements from space (e.g., Kreidberg et al., 2014; Stevenson & Fowler, 2019) or from the ground (e.g., Berta et al., 2011; Stevenson et al., 2014), with comparable results. However, these two methods have no mechanism to distinguish between contributions from overlapping traces from different sources or diffraction orders.

Yet, the challenge of extracting spectra from blended sources is not unprecedented. It was needed notably in the context of long-slit spectroscopy for several science applications, such as the observation of galactic nuclei (Lucy & Walsh, 2003) or crowded star fields (Hynes, 2002). In fact, a task as common as a simple sky subtraction is by itself a type of decontamination. Hence, various techniques have been proposed over the past two decades (e.g., Hynes, 2002; Khmil & Surdej, 2002; Lucy & Walsh, 2003; Bolton & Schlegel, 2010). However, due to the particularity of the NIRISS SOSS mode and the precision it requires, it was necessary to develop a dedicated algorithm.

In this article, we present ATOCA ; an algorithm designed to properly decontaminate and extract overlapping orders. Though the methods that make up the ATOCA algorithm can be applied generally to the problem of extracting overlapping spectral orders, we focus here on the NIRISS/SOSS mode of JWST. Proper extraction of SOSS observations was our primary motivation for creating ATOCA, and the algorithm has been made part of the official JWST

pipeline<sup>2</sup> – the data management system (DMS) – as part of the stage 2 spectral extraction step.

The article is divided as follow : Section 4.2 presents an estimate of the level of contamination that is expected with the NIRISS/SOSS mode. Then, the algorithm and its implementation are presented in sections 4.3 and 4.4, followed by section 4.5 where we evaluate the performance of the decontamination.

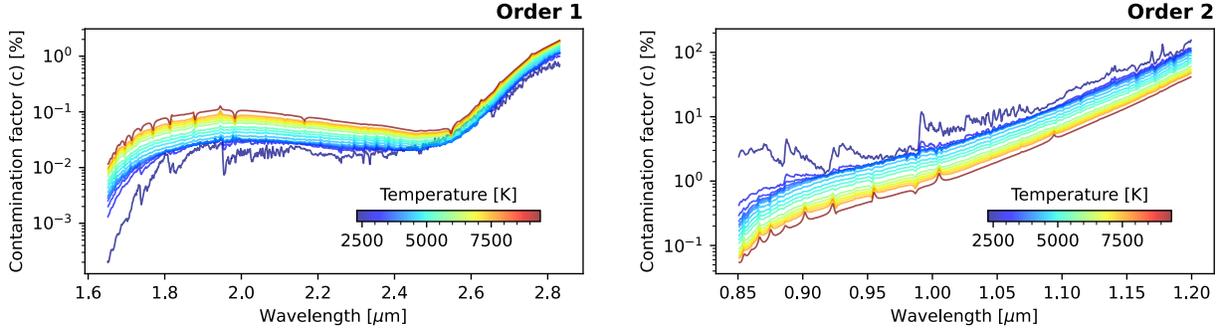


**Figure 2** – High signal-to-noise co-added image of multiple SOSS mode observations on a tungsten lamp. This image was obtained during cryogenic vacuum testing (CV3) of the telescope and instruments. Order 1 is the most apparent feature extending from  $0.83 \mu\text{m}$  on the right to  $2.8 \mu\text{m}$  on the left portion of the image. The second order can be seen starting at  $0.6 \mu\text{m}$  at the top right part of the image and extends out to  $1.4 \mu\text{m}$ . At approximately  $1.1 \mu\text{m}$  the second order is significantly blended with the first order. The faint third order can be observed above the 2nd order. The overlap between the spectral orders 1 and 2 on the left side of the image complicates the spectrum extraction and motivates this paper. Since the order 2 covers shorter wavelengths than order 1, this problem should be even more striking in actual astrophysical targets which are warmer ( $T \geq 3000 \text{ K}$ ) than the tungsten filament used in the laboratory ( $1500 \text{ K}$ ). The vertical lines were added to mark the wavelengths, from right to left, at  $0.9, 1.2, 1.6, 2.0, 2.4, 2.8 \mu\text{m}$  for order 1 (in black) and at  $0.6, 0.8, 1.0, 1.2, 1.4 \mu\text{m}$  for order 2 (in red).

## 4.2. The SOSS Trace Overlap Problem

The optics of the SOSS mode were designed with strong mechanical constraints, one of which, the total thickness of the GR700XD element, prevented the cross-dispersing prism from being sufficiently inclined to cleanly separate spectral orders 1 and 2 (see Albert et al. in prep). As a result, the red end of order 2 ( $\lambda \geq 1.1 \mu\text{m}$ ) partially overlaps with that of order 1 ( $\lambda \geq 2.2 \mu\text{m}$ ) (see Figure 2). This cross contamination of the signals is a major issue during spectrum extraction and will bias results using the simple aperture-based methods discussed in Section 4.1. The amount of contamination can be characterized by measuring the contaminating signal present in the extraction region,  $F_{\text{contam}}$ , allowing for the definition

2. <https://github.com/spacetelescope/jwst>



**Figure 3** – Contamination factors (see equation 4.2.5) for a range of stellar effective temperatures (color-coded). These values hold for a standard box extraction, using a 25 pixel wide aperture. The wavelength domain is not entirely shown here; the first order (left panel) is virtually uncontaminated below  $1.6 \mu\text{m}$  whereas the second order (right panel) contamination levels increase exponentially at longer wavelengths ( $> 1.1 \mu\text{m}$ ).

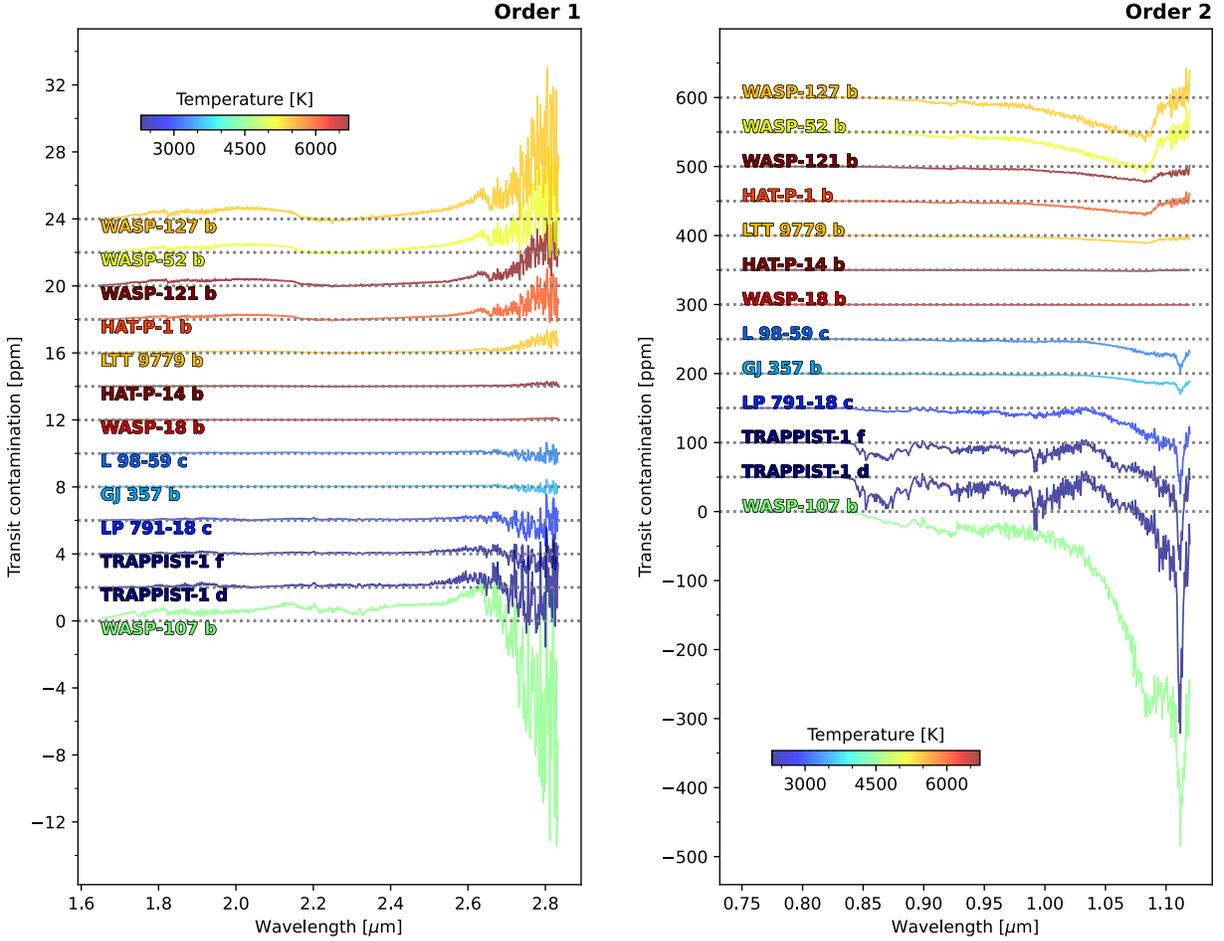
of a contamination factor,  $c_{order}$ , via the following ratio :

$$c_{order} = \frac{F_{contam}}{F_{order}}. \quad (4.2.1)$$

Here,  $F_{order}$  is the flux that would be extracted from the targeted order if there was no contamination. For example, for the first order,  $F_{contam}$  is given by the signal from the second order that is overlapping with the region of interest. Based on laboratory measurements of the blaze function, detector efficiency, and end-to-end throughput of the telescope and NIRISS instrument, it is possible to simulate this contamination factor for any given extraction method.

Figure 3 shows the contamination factors,  $c$ , for a standard box extraction using a 25 pixel wide aperture. The simulations were made using ATOCA and are described in Section 4.5.1. The aperture width was determined by minimizing the standard deviation of the extracted flat white light-curve (i.e. perfectly stable star spectrum without any transit) over the contaminated band-pass ( $2.0\text{--}2.8 \mu\text{m}$ ). This is equivalent to the commonly used procedure where the standard deviation of the fitted white light-curve is minimized (e.g. Wakeford et al., 2013; Sing et al., 2015). To get a realistic estimation, we had to consider the expected jitter on the telescope pointing  $\sim 5 \text{ mas}$  (see Albert et al., in prep.) which has the effect to increase the width of the required aperture. The incident fluxes at different stellar effective temperatures were modeled using PHOENIX HiRes synthetic spectra (Husser et al., 2013). Note also that the contamination factor is presented in % to make a distinction with the transit depth (which is typically in ppm), both being relative quantities.

For the first diffraction order, most of the affected wavelength range exhibits levels of contamination below 0.1 %, until  $2.6 \mu\text{m}$ , after which  $c_1$  increases exponentially up to 2 %.



**Figure 4** – Expected contamination levels during transit for a diverse sample of exoplanets. Each target is shifted vertically by 2 ppm for the first order (left panel) and by 50 ppm for the second order (right panel). A horizontal dotted line is drawn for each model to mark its zero point. The color of the lines refers to the effective temperature of the host star, which influences the level of contamination. As in Figure 3, these estimates were made assuming a 25-pixel-wide box extraction and the wavelength domain was limited for the same reasons.

The stellar effective temperature also has a significant effect, since the second order covers shorter wavelengths ( $0.6 \leq \lambda \leq 1.4 \mu\text{m}$ ) than the first order ( $0.83 \leq \lambda \leq 2.8 \mu\text{m}$ ), hence a star with a stronger relative flux contribution at short wavelengths will be more affected. The second order, on the other hand, suffers more drastically from contamination, reaching levels above 100%. Fortunately, the wavelength domain affected ( $\lambda \geq 0.85 \mu\text{m}$ ) is shared with order 1 (and is within the region where order 1 does not suffer from contamination), so very little information is lost. The wavelengths towards the blue end of the spectrum ( $\lambda \leq 0.85 \mu\text{m}$ ) and complementary to order 1 correspond to the part of the detector where

the orders' spatial positions deviate from one another, creating a drop in the contamination levels.

The above discussion holds for any absolute flux measurement, but what is the impact for the intended application of the SOSS mode — exoplanet time-series — whose measurements are relative.

The fluxes measured in order 1 will be a combination of the true flux in that order,  $F_1$ , contaminated by some flux from order 2,  $F_{\text{contam}}$ . Assuming a simplified top-hat model for an exoplanet transit (no limb darkening, instantaneous ingress and egress, non-grazing) this flux can be calculated, for cases in and out of transit, via :

$$F_{out} = F_1 + F_{\text{contam}} \quad (4.2.2)$$

$$F_{in} = (1 - d - \delta_1(\lambda_1))F_1 + (1 - d - \delta_2(\lambda_2))F_{\text{contam}} \quad (4.2.3)$$

where  $F_{out}$  and  $F_{in}$  are respectively the mean flux outside of transit and during transit measured by an extraction around the first order's trace,  $d$  is the transit depth due to the opaque planet (i.e., without considering an atmosphere) and  $\delta_1(\lambda_1)$  and  $\delta_2(\lambda_2)$  are the wavelength-dependent transit depths due to the planet's atmosphere for orders 1 and 2, respectively.  $\lambda_1$  and  $\lambda_2$  are the wavelength solutions at each order, which are both a function of the column position,  $x$ , such that  $\lambda_1(x)$  and  $\lambda_2(x)$ . The transit depth measured on a contaminated trace is, by definition :

$$D = 1 - F_{in}/F_{out}. \quad (4.2.4)$$

Recalling the order contamination factor from equation 4.2.1, then the transit depth can be written :

$$D = d + \delta_1(\lambda_1) + \frac{c_1}{1 + c_1} (\delta_2(\lambda_2) - \delta_1(\lambda_1)). \quad (4.2.5)$$

In the case where there is no chromatic variation in the atmospheric signal (i.e., a flat transmission spectrum),  $\delta_2(\lambda_2) = \delta_1(\lambda_1)$  so  $D = d + \delta_1(\lambda_1)$ . Therefore, contamination has no bearing on the retrieved transit depth. In other words, the second order contaminating signal changes by exactly the same relative amount during transit as the first order signal.

In the case where the atmospheric signal is different at the two overlapping wavelengths, then the difference

$$\Delta(x) = \delta_2(\lambda_2(x)) - \delta_1(\lambda_1(x)) \quad (4.2.6)$$

modulated by  $c_1/(1 + c_1)$  will affect the transit depth (i.e., the last term in equation 4.2.5). To make a distinction with the contamination factor  $c$ , we will use the name “transit

contamination” to refer to this last term of equation 4.2.5. Generally,  $\Delta$  will be about the same order of magnitude as  $\delta_1$  and  $\delta_2$ , so a good estimation can be drawn simply from the contamination term. Moreover, for order 1,  $c$  will be small, hence  $c/(c + 1) \approx c$ . So, concerning the first order’s relative measurement, the spurious signal can be approximated to less than 1% (i.e.,  $c_1$ ) of the chromatic contribution of the transit signal,  $\delta_1$ . For example, if we take a hypothetical transmission spectrum with a spectral feature for the first order of  $\delta_1(2.7 \mu\text{m}) = 300 \text{ ppm}$  above the mean transit depth  $d$ . Let’s also assume that there is a spectral feature from the second order at the corresponding columns (see Figure 2) of  $\delta_2(1.35 \mu\text{m}) = -200 \text{ ppm}$ , i.e., 200 ppm below the mean transit depth. This would result in a difference of  $\Delta(x) = 500 \text{ ppm}$  and the resulting contamination signal will be around 1.5 ppm, considering a contamination factor of  $c_1 \approx 3\%$  (see Figure 3).

Nevertheless, to fully grasp the importance of this effect, we computed the resulting contamination signal in transmission for a variety of exoplanets, most of them being part of the NIRISS Exploration of the Atmospheric diversity of Transiting exoplanets (NEAT) GTO program (Lafreniere, 2017). The results are shown in Figure 4. The transit models for each planet were produced using the SCARLET atmosphere framework (Benneke & Seager, 2012, 2013; Benneke, 2015) assuming, for simplicity, cloud-free atmospheres with solar elemental abundances and chemical equilibrium. These assumptions generally lead to stronger signals, hence upper limits on the estimates. For the first order, the transit contamination is constrained below 8 ppm for all targets except WASP-107 b, for which it reaches almost  $-12 \text{ ppm}$ . This corresponds to  $\approx 1\%$  of the planet atmospheric signal, as expected. For example, the transmission spectrum of WASP-107 b presents spectroscopic variations around 2500 ppm (see Figure 10), which would lead to an expected transit contamination signal of  $\sim 25 \text{ ppm}$ , not far from 12 ppm value. On the other hand, the second order is much more affected, with levels around 100 ppm or more in the longer wavelength range shown in Figure 4. This contamination comes from the wings of the first order’s spatial profile. The longer wavelengths are not presented since the second order becomes almost completely diluted into the first order (see Figure 2). On the other side, below  $0.8 \mu\text{m}$ , where the second order contributes unique wavelength coverage, the transit contamination seems to vanish. However, this drastic drop in contamination is attributable to the simulations, as discussed in Section 4.5.2.

Whilst the systematic error on the transit signal may seem small, it must not be taken lightly and should be prevented using the ATOCA extraction method presented in the next section. We also want to emphasize that this is a systematic error, and not a randomly distributed source of noise such as shot noise. These estimations could also be worsened by any other relative signals that depends on wavelength; like limb darkening and stellar contamination from unocculted regions (e.g. Rackham et al., 2018; Genest et al., 2022). Moreover, the examples presented here assume that the trace shape is perfectly stable within a whole time-series and that the trace position is varying within the expectation, i.e., following a

random normal distribution with a dispersion of 5 mas. In the context of real observations, these assumptions may not be true due to the finite pointing precision of the Fine Guidance Sensor (FGS) and possible variations in the point spread function (breathing effects, wavefront variations, etc.). Therefore, to obtain the most stable transmission spectrum, one may need to increase the width of the aperture (along the spatial axis) used for the box extraction (e.g., Diamond-Lowe et al., 2018; Mikal-Evans et al., 2021) in order to minimize the variation due to the signal moving in and out of the aperture, hence increase the contribution of the contaminating order. Furthermore, even in the context of standard extractions, ATOCA will help to calibrate and extract the one-dimensional spectra. The contamination would also need to be properly characterized by identifying the contribution of each order to ensure the reliability of any results. Finally, extraction using ATOCA will ensure that science applications needing absolute flux calibration can be realized with SOSS.

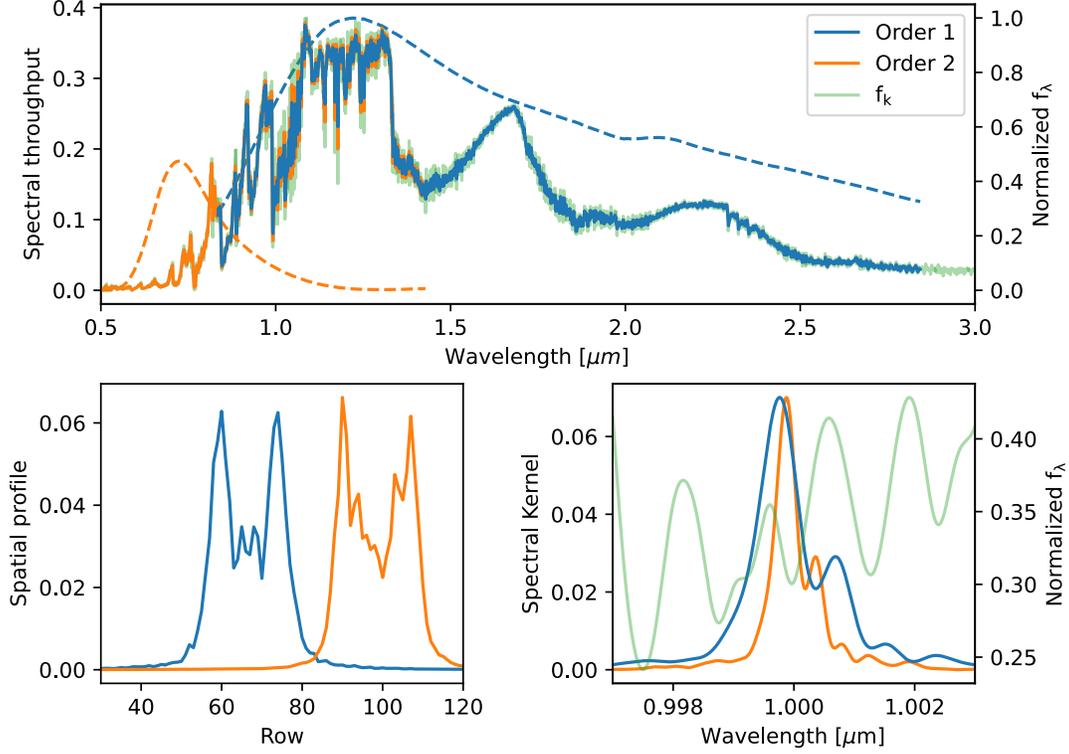
### 4.3. Description of the Extraction Method

The core idea behind ATOCA is to determine the underlying flux by fitting each order directly and simultaneously on the detector image, pixel by pixel. To do so, we first need to establish a linear model of each individual pixel with the flux as independent variable. More precisely, we need to discretize the flux by evaluating it on a wavelength grid. Each of the nodes (or elements) of the resulting flux array is an independent variable. Then, by minimizing the  $\chi^2$  with respect to each of these nodes, we are able to express the flux as the solution of a linear system, which ultimately enables us to explicitly extract it. Hence, no forward modelling of the flux is needed for an extraction. However, to be accurate, it requires a thorough knowledge of the detector’s properties. The formalism of the algorithm is described below.

For the following equations, each pixel that will be used for the fit will be labeled by the index  $i$  and the total number of relevant pixels is given by  $N_i$ . There is no need to account for the two-dimensional nature of the detector with additional indices. Each diffraction order also needs to be identified, here using the index  $n$ . Now, to determine the flux falling on each pixel, we need for each order : 1) the wavelength solution, 2) the spectral throughput<sup>3</sup>, and 3) the spatial throughput. From the wavelength solution, we can define the central wavelength of a pixel  $i$  at order  $n$  as  $\lambda_{ni}$ . It will also be useful to define  $\lambda_{ni}^+$  and  $\lambda_{ni}^-$ ; the wavelengths at the pixel borders in the spectral direction, and  $\Delta\lambda_{ni} = \lambda_{ni}^+ - \lambda_{ni}^-$ ; the pixel spectral coverage. The spectral and spatial throughputs are given respectively by the function  $T_n(\lambda)$  and the constant  $P_{ni}$ . Take notice that the former depends on the wavelength but not explicitly on the pixel, whereas the opposite is true for the latter. Finally, let the spectral flux density of the target incident upon the spectrograph be  $f(\lambda)$ .

---

3. Throughput is defined here as an end-to-end wavelength-dependent transmission (detected flux divided by the flux impinging the telescope).



**Figure 5** – Spectrum and associated spectral profiles for a 2300 K star, along with contributions from the first (blue) and second (orange) orders. The top panel shows the spectral throughputs (dashed lines) and the flux density of a 2300 K PHOENIX spectrum downgraded at each order’s resolution and at a higher resolution representative of the underlying flux (in green). The bottom left panel displays the spatial profiles along column 500, where the overlap occurs. The corresponding wavelength from Figure 2 is  $\sim 2.31 \mu\text{m}$  for order 1 and  $\sim 1.17 \mu\text{m}$  for order 2. At the bottom right, an example of the convolution kernels centered at  $1 \mu\text{m}$  is shown. The green curve is the underlying flux (same as top panel).

It is also important to consider that the flux density is seen by each order at a different resolution, so for an order  $n$ , we need an additional input : 4) the spectral resolution kernel  $\kappa(\tilde{\lambda}, \lambda)$ . It specifies the convolved flux  $\tilde{f}$  in relation with an incident flux through the equation

$$\tilde{f}_n(\tilde{\lambda}) = \int_0^{\infty} \kappa_n(\tilde{\lambda}, \lambda) f(\lambda) d\lambda. \quad (4.3.1)$$

Figure 5 presents some visualizations of the aforementioned quantities. The difference between the two orders after convolution with the resolution kernels becomes apparent in the overlapping wavelength range, where the first order (blue curve) is not superimposed perfectly on the second order (orange curve).

### 4.3.1. The model

With all this in place, we are now able to define a model of the detector. The number of photo-electrons detected by pixel  $i$  can be represented, up to a multiplicative constant, as

$$\begin{aligned} M_i &= \sum_n \int_{\lambda_{ni}^-}^{\lambda_{ni}^+} P_{ni} T_n(\lambda) \tilde{f}_n(\lambda) \lambda d\lambda \\ &= \sum_n \int_{\lambda_{ni}^-}^{\lambda_{ni}^+} a_{ni}(\lambda) \tilde{f}_n(\lambda) d\lambda, \end{aligned} \quad (4.3.2)$$

where  $a_{ni}(\lambda)$  accounts for all the coefficients that are not the flux. Note that the summation is made over all orders  $n$  that contribute to the signal measured by a given pixel  $i$ . In the case of the NIRISS/SOSS mode, the index  $n$  covers only the first and second orders; the third order is not considered since it does not cover the same pixels. To translate this model into a numerical form, we can define a grid where  $f$  is projected, labeled by the index  $k$  so that  $f(\lambda_k) = f_k$  and  $\Delta\lambda_k = \lambda_{k+1} - \lambda_k$ . The length of the discretized grid would then be given by  $N_k$ , so that  $1 \leq k \leq N_k$ . Similarly,  $N_{\tilde{k}}$  is the length of the convolved flux  $\tilde{f}$ , so that  $1 \leq \tilde{k} \leq N_{\tilde{k}}$ . We also need a numerical form of this integral. There are multiple ways to do this, but for ATOCA, we use the trapezoidal method on a specified grid as illustrated in Figure 6. The details of this method are in Section 4.7. Independently of the chosen integration technique, the numerical form of the integral will look like,

$$M_i = \sum_n \sum_{\tilde{k}} w_{in\tilde{k}} a_{in\tilde{k}} \tilde{f}_{n\tilde{k}}, \quad (4.3.3)$$

with  $w_{in\tilde{k}}$  given by the integration method. To link the diffraction orders, we want to write these equations according to the underlying flux  $f(\lambda)$ , following equation 4.3.1. In the numerical form,

$$\tilde{f}_{n\tilde{k}} = \sum_k \kappa_{n\tilde{k}k} f_k \quad (4.3.4)$$

with  $\kappa_{n\tilde{k}k}$  being the coefficients of the convolution kernel at order  $n$ . The numerical integration method as well as the kernel are comprised in them.

Finally, we have that

$$M_i = \sum_n \sum_{\tilde{k}} w_{in\tilde{k}} a_{in\tilde{k}} \sum_k \kappa_{n\tilde{k}k} f_k. \quad (4.3.5)$$

This equation can be written in a more intuitive matrix form as

$$\begin{pmatrix} M \end{pmatrix}_{N_i} = \sum_n \begin{pmatrix} w_n a_n \end{pmatrix}_{N_i \times N_{\tilde{k}}} \begin{pmatrix} \kappa_n \end{pmatrix}_{N_{\tilde{k}} \times N_k} \begin{pmatrix} f \end{pmatrix}_{N_k}. \quad (4.3.6)$$

To simplify the notation again, we can put all the coefficients for each order in single matrices  $\mathbf{b}_n$  with dimensions  $N_i \times N_k$ ,

$$\begin{pmatrix} M \end{pmatrix}_{N_i} = \sum_n \begin{pmatrix} b_n \end{pmatrix}_{N_i \times N_k} \begin{pmatrix} f \end{pmatrix}_{N_k}, \quad (4.3.7)$$

and add them together in one matrix  $\mathbf{B}$  to have a final model of each valid pixel given by

$$\boxed{\mathbf{M}_{N_i} = \mathbf{B}_{N_i \times N_k} \mathbf{f}_{N_k}}. \quad (4.3.8)$$

This result is one of the main utilities of ATOCA, which is a linear model of the full NI-RISS/SOSS detector. One could use it to generate quick simulations, given a model of the incident flux.

### 4.3.2. Solving for $\mathbf{f}$

Now that we have a model of the intensity at each relevant pixel, we can link their individual measured intensity  $D_i$ , with  $f_i$  by fitting directly the pixel model on the detector using a  $\chi^2$  minimization. Given the following equation,

$$\chi^2 = \left\| \frac{\mathbf{D} - \mathbf{M}}{\sigma} \right\|^2. \quad (4.3.9)$$

with  $\mathbf{D}$  being the array of measured intensities on each pixels  $i$  and  $\sigma$  the array of their uncertainties, then the best solution can be found by imposing

$$\frac{d\chi^2}{df_k} = 0. \quad (4.3.10)$$

The detailed calculations found in the Appendix 4.7 lead to the following system of equations :

$$\begin{pmatrix} \mathbf{B} \\ \sigma \end{pmatrix}_{N_k \times N_i}^T \begin{pmatrix} \mathbf{D} \\ \sigma \end{pmatrix}_{N_i} = \begin{pmatrix} \mathbf{B} \\ \sigma \end{pmatrix}_{N_k \times N_i}^T \begin{pmatrix} \mathbf{B} \\ \sigma \end{pmatrix}_{N_i \times N_k} \mathbf{f}_{N_k}, \quad (4.3.11)$$

where

$$\begin{pmatrix} \mathbf{B} \\ \sigma \end{pmatrix}_{N_i \times N_k} = \text{diag} \left( \frac{1}{\sigma} \right)_{N_i \times N_i} \mathbf{B}_{N_i \times N_k}. \quad (4.3.12)$$

This system can now be solved for  $\mathbf{f}$ . A comparison with the optimal extraction method (Horne, 1986) is presented in Appendix 4.7.

To precisely estimate the integral representing each pixel, an oversampled numerical grid is required (see Figure 6 and paragraph *Grid sampling* of Section 4.3.4). Hence, for a given pixel, the solution of equation 4.3.11 would be highly degenerate. In many situations, the system will still be invertible since many pixels can cover slightly different wavelength ranges, but the solutions will then be extremely unstable. This is an ill-conditioned system, where a slight change in the observation vector  $\mathbf{D}$  could cause large differences in the solution  $\mathbf{f}$ . To circumvent this problem, the system needs to be regularized.

### 4.3.3. Regularization

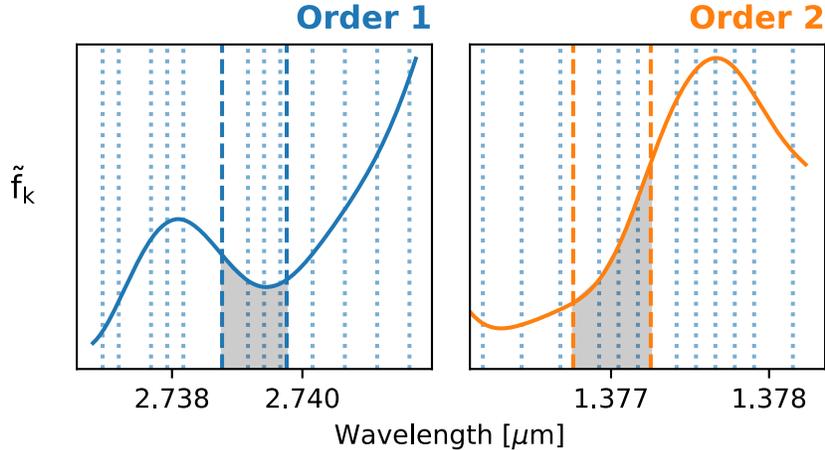
There are multiple ways to perform regularization. The one chosen here is Tikhonov regularization (Tikhonov, 1963). It is also referred as Phillip-Twomey’s regularization (Phillips, 1962) and ridge-regression (Horel, 1962). This technique has been used in astrophysics in a variety of similar situations requiring the inversion of an integral equation (e.g., Kunasz et al., 1973; Thompson, 1990). In fact, it is part of some advanced spectral extraction methods for fiber-fed spectrographs that were identified among the most effective methods (Min et al., 2020). It is also used in the context of spline interpolation of noisy data (Green & Silverman, 1993; Hastie & Tibshirani, 1990). The main idea is to add a regularization term to the linear  $\chi^2$  (equation 4.3.9), yielding the following equation :

$$\chi_{\text{Reg}}^2 = \chi^2 + \alpha \|\Gamma \mathbf{f}\|^2 . \quad (4.3.13)$$

Here,  $\alpha$  is a Lagrange multiplier and  $\Gamma$  is a linear operator that adds a “cost” depending on the nature of the solution. Generally, it is used to favor smoother solutions, and hence reducing the level of overfitting. We can obtain the new solution by minimizing the system in the same fashion as before, differentiating with respect to  $f_k$ , with the following result :

$$\begin{pmatrix} \mathbf{B} \\ \sigma \end{pmatrix}^T \begin{pmatrix} \mathbf{D} \\ \sigma \end{pmatrix} = \left[ \begin{pmatrix} \mathbf{B} \\ \sigma \end{pmatrix}^T \begin{pmatrix} \mathbf{B} \\ \sigma \end{pmatrix} - \alpha \Gamma \right] \mathbf{f}_{N_k} . \quad (4.3.14)$$

In this case, the Tikhonov matrix  $\Gamma$  was chosen to be the first derivative operator as in Li et al. (2015) or Piskunov et al. (2021). The choice of  $\alpha$  can be done in many different ways. Usually, the general idea is to find a good balance between the regularization term and the  $\chi^2$  which can be done using the L-curve criterion (e.g., Hansen, 1992) or generalized cross-validation techniques, GCV (Golub et al., 1979; Wahba, 1977). The L-curve technique has the advantage of being robust, especially against correlated noise (Hansen & O’Leary, 1993) compared to the GCV. However, it tends to over-smooth the solution, which is not optimal. As for the GCV, it is too computationally intensive in the context of large-scale problems. However, in the present situation, the objective is not the direct product of the regularization, i.e., the underlying flux, but rather the modeling of the pixel after re-integration (this will be discussed in sections 4.3.4 paragraph *Proper output* and 4.4). Thus, it is not necessary to find the most physically accurate solution. In fact, what is needed is the smoothest solution that, once re-projected on the detector, fits the observations within the uncertainties. Consequently, we defined custom criteria to determine  $\alpha$  with the objective to keep the dependency of the solution’s sensitivity to the scaling factor lower than the expected noise. Mathematically, this comes back to defining a threshold on the derivative of the  $\chi^2$  with respect to  $\log \alpha$ . The same convergence criterion was used by Khmil & Surdej (2002) in a similar context.



**Figure 6** – Example of the trapezoidal integration on a specific grid. The shaded grey region represents the area under the curve for a trapezoidal integration over the wavelength coverage of a given pixel delimited by the vertical dashed lines. The vertical dotted lines indicate the irregular grid used to do the integration. The flux at each order’s resolution is given by the solid lines.

#### 4.3.4. Other considerations

Grid sampling. One important aspect of the technique is the choice of the wavelength grid. Since the simulated pixels are the results of numerical integrations, they are subject to computational errors. One way to define the grid would be to use a grid representative of the native pixel sampling and to oversample each interval of this grid by a certain factor. This will however create unnecessarily large systems and increase the computation time. Moreover, for the regularization method to be well-behaved, it is better to have errors of the same order for each node. Indeed, as highlighted by Puetter et al. (2005), the dynamical range needs to be constrained, lest some regions will be overfitted and others underfitted. Given all this, we opted for an irregular grid designed to make the magnitude of the integration error between subsequent nodes more uniform. To estimate the integration error on each node, we compared the result of a trapezoidal integration (as it is done in ATOCA) with a more precise Romberg’s integration. The intervals with an estimated error higher than a specified tolerance were oversampled by a factor of two iteratively until the tolerance was satisfied. This method only requires an estimate of the function to be integrated, which can be given by a user or directly estimated from the data. An example of an uneven grid is shown in Figure 6, represented by the vertical dotted lines.

Proper output. As mentioned before, the underlying flux  $\mathbf{f}$  is not the end product of an extraction. Indeed, since it has a resolution higher than both orders on the detector,  $\mathbf{f}$  will be degenerate. In fact, solving for  $\mathbf{f}$  is a deconvolution, which is subject to instabilities or artifacts (Bolton & Schlegel, 2010). Thus, one will have to reconvolve the result to get rid

of these effects. In this case, the underlying flux has to be integrated on bins representing a pixel grid. This can be done by invoking equation 4.3.8 and reconstructing the detector, which can be used to assess the quality of the detector modeling by examining the residuals. It could also be used to rebuild each order independently using the  $\mathbf{b}_n$  matrices from equation 4.3.7. Furthermore, it is possible to get a one-dimensional spectrum by re-integrating on a grid representative of the pixel sampling. In fact, this is equivalent to reconstructing a single row of the detector. In this context, without any dimension in the cross-dispersion axis, the spatial profile would not be relevant anymore and its value should be set to unity.

Wavelength distortions. In a variety of situations, the wavelength solution will not be constant along the axis perpendicular to the dispersion. Generally, this can be due to differential refraction in the Earth’s atmosphere and imperfect spectrograph optics (Horne, 1986). This effect is seen in many spectrographs (e.g. Bolton & Schlegel, 2010; Piskunov et al., 2021). It can also be caused by observing techniques like the spatial-scan mode of the *Hubble Space Telescope*’s Wide Field Camera 3 (Deming et al., 2013). In the case of the NIRISS SOSS mode, a tilt is present due to the grism configuration (Albert et al., in prep.). Standard extraction procedures will either neglect this distortion (Sing et al., 2015) or re-sample the detector image by interpolating along the dispersion axis (e.g., Kreidberg et al., 2014). ATOCA has the advantage of implicitly accounting for this distortion by treating each pixel individually, and using the full 2D wavelength solution.

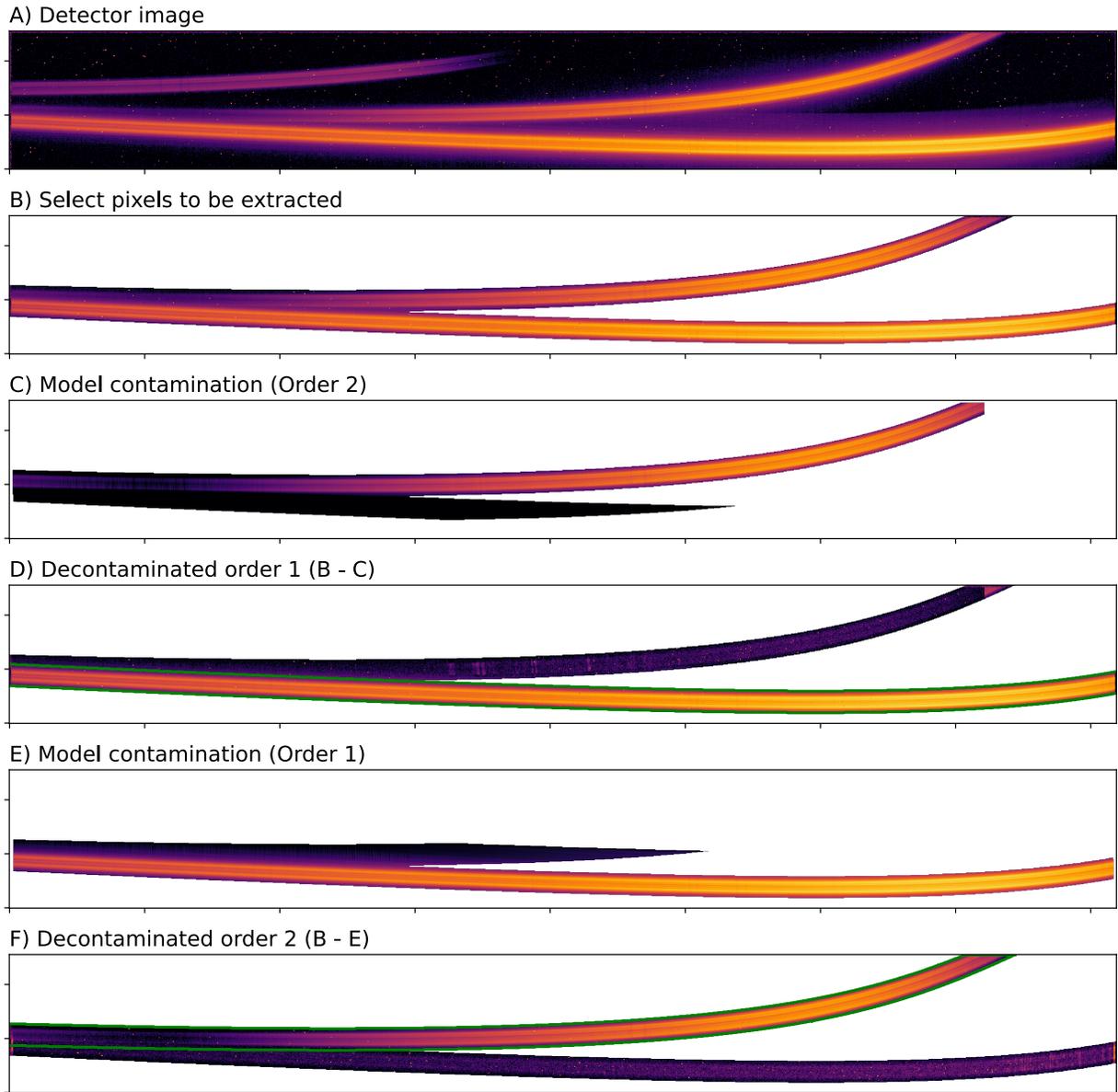
## 4.4. Implementation

ATOCA has been implemented in `python 3.8` as an option for the spectral extraction of NIRISS/SOSS observations in the JWST Data Management System (DMS). It is part of the `Extract1dStep`<sup>4</sup> of the `spec2pipeline` calibration step. A description of the inputs and an example Jupyter Notebook can be found at [https://github.com/AntoineDarveau/atoca\\_demo](https://github.com/AntoineDarveau/atoca_demo).

However, the end product is not the spectrum extracted with this method, as described in the Section 4.3.4 paragraph *Proper output*. Indeed, one caveat regarding the product of an extraction with ATOCA is the dependency on the accuracy of the model. Just like the optimal extraction, the method needs a relative precision on the spatial profile much better than the corresponding data in order to make an unbiased spectral extraction (Horne, 1986). Moreover, inconsistencies between the spectral orders in the wavelength solution, the throughput or the resolution kernels could result in an intermediate solution between the two orders that will not satisfy the expected accuracy. However, in the case of contaminated spectra, it is possible to circumvent this problem by reconstructing the trace of each individual order with the  $b_n$  matrices of equation 4.3.7, which are then used to decontaminate the detector image; the

---

4. [https://jwst-pipeline.readthedocs.io/en/latest/jwst/extract\\_1d/arguments.html](https://jwst-pipeline.readthedocs.io/en/latest/jwst/extract_1d/arguments.html)



**Figure 7** – Decontamination steps for the first and second spectral orders. From top to bottom : A) processed image of the detector with only signal from the targeted object, B) selection of relevant pixels, C) reconstructed image of the contaminating orders after a fit using ATOCA, D) decontaminated image of the first order after subtraction of the contaminating order ; the data is now ready for a standard box extraction around first order (delimited by the green lines). Panels E and F are the same steps as C and D, but the extraction of the second order. The color scale is logarithmic, with the lowest intensity in black and the highest in yellow. The pixels that are not considered in the analysis are in white.

process is shown in Figure 7. This allows for a more classical extraction, like a box extraction, to be performed afterwards on each decontaminated trace. This technique has the advantage of reducing the dependency on the accuracy of the model. For instance, as shown in Figure 7, the modeling will only affect contaminated columns of the first order, leaving the rest untouched. This technique was preferred in the context of NIRISS/SOSS mode, although the ATOCA spectra are also available as a byproduct of the detector fit. Another output of the ATOCA is the model of each order (1st and 2nd) that were used for the decontamination. This could be useful to assess the quality of the fit by simply looking at the residuals. It would also be required to quantify the level of contamination (see Section 4.2). Furthermore, this model of the detector can be used to assign values to bad pixels located inside the aperture without the need for a separate outlier correction routine.

As shown in Figure 7, only the well-behaved pixels that will be extracted are considered in the application of ATOCA. This is in order to reduce the computational time and to make sure that the detector fit is not biased by superfluous regions of the detector. However, even with pixel selection, the size of the matrix  $B$  in equation 4.7.5 is still considerable, with a size of  $N_{pixel} \times N_k$ . For example, in the realistic scenario of a 40-pixel aperture and a reasonable oversampling of the wavelength grid (tolerance of  $10^{-3}$  per pixel), the dimension of the matrix will be around  $150,000 \times 6000 = 9 \times 10^8$ . To overcome this problem, we took advantage of the fact that most elements of the matrix are null. Indeed, each order (or source) modeled by the matrix  $B$  is represented as a block-diagonal that can be shifted with respect to the main diagonal. This enabled us to use the `scipy` (Virtanen et al., 2020) sparse matrices and drastically lessen the computational time and the amount of memory needed.

The implementation required the addition of new reference files to the Calibration Reference Data System (CRDS)<sup>5</sup> used by the DMS, since ATOCA needs, for each order, the two-dimensional wavelength solutions, the spectral resolution kernels, the spatial profiles, and the spectral throughputs. The one-dimensional wavelength solutions and the throughputs are already a product of the standard calibrations, so the two-dimensional wavelength maps can be created simply by applying the tilt to the existing solution. The convolution kernels were determined from monochromatic point spread functions generated with `WebbPSF`. The resulting images were rectified to remove the tilt and summed over the spatial axis to keep only the spectral dependency. The most demanding input is still the determination of the spatial profile, but techniques typically used for optimal extractions could be applied directly to most of the spectral ranges, except for the overlapping parts. An algorithm to estimate the spatial profiles of both orders within the contaminated region is currently in development (Radica et al., 2022), and will be made available to complement ATOCA before the release of Cycle 1 data.

---

5. <https://jwst-crds.stsci.edu/>

As mentioned in section 4.3.4, ATOCA needs an estimate of the underlying flux  $f_k$  to generate an oversampled grid. In the context of the NIRISS/SOSS mode, this is done by extracting the underlying flux  $f_k$  for each order separately with only the most contaminated pixels masked. These rough extractions are done over a grid at native pixel sampling, which ensures the stability of the solution and removes the need for any regularization. Precision will be lost in this process and the contamination will not be treated correctly, but it is sufficient to generate the estimate. The latter is also used to provide a first estimate of the regularization factor, which is refined afterwards as described in section 4.3.3. This step can take some time depending of the precision needed for the oversampled grid and the number of relevant pixels. Fortunately, it only needs to be performed once for a given time series, since the underlying spectrum will not vary enough to justify a different level of regularization.

In realistic observations, the position of the trace will change slightly between visits due to the angle of the pupil wheel and variation in the target acquisition, which will alter the wavelength solution as well as the spatial profile. These changes will effectively take the form of a rotation and a spatial shift of these reference files and are implemented as input parameters (rotation and translation) captured by the keyword `soss_transform = [x, y, theta]`. These parameters can be either specified by the user or determined within `Extract1dStep` by fitting the measured trace centroid.

## 4.5. Validation

### 4.5.1. Simulations

Two types of simulations were used in the context of this work : simulations produced with ATOCA itself (ATOCA simulations), and those produced by the instrument development team (IDT simulations). The first type takes advantage of the ATOCA capability to directly model the detector image using equation 4.7.5, given an input spectrum. Hence it is possible to use it to generate simplistic simulations to validate the internal consistency of the method. To make sure that the numerical precision was not an issue, the wavelength grid used for the simulation was oversampled to limit the integration error over each pixel below  $\sim 1$  ppm. We only included photon and background noise, so neither bad pixels,  $1/f$  noise, nor cosmic ray hits were taken into account. The reference files were based on the best current knowledge of the NIRISS/SOSS mode, as described in section 4.4. We used the throughput and the one-dimensional wavelength solution estimated or measured in the lab by the NIRISS instrument development team. The spatial profiles were determined from the IDT simulations

The IDT simulations (see Albert et al, in prep) are made by distributing an incident flux directly on an oversampled image using a trace as wide as a single oversampled pixel. This signal is then convolved in two dimensions with a grid of monochromatic kernels from

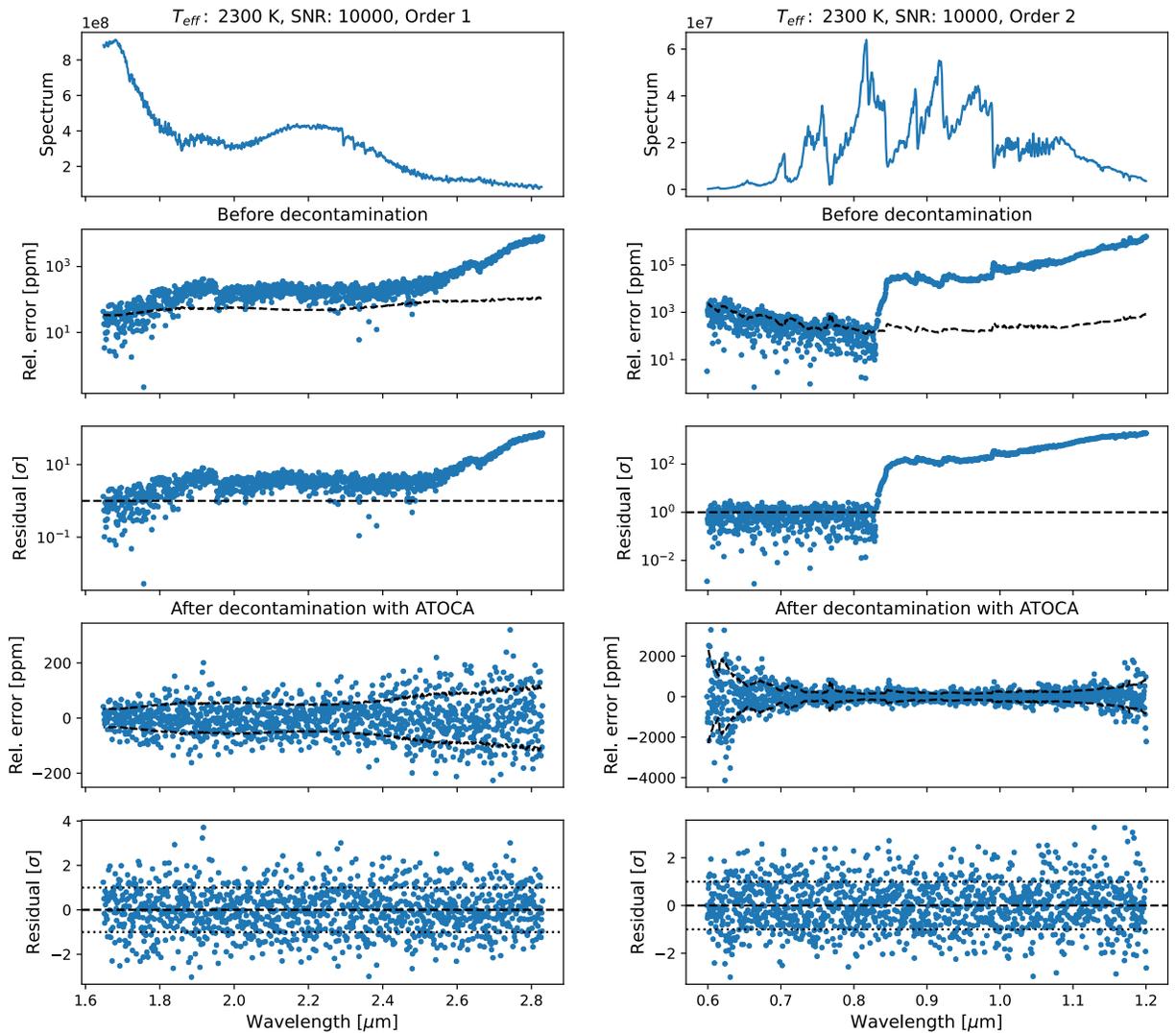
WebbPSF and re-sampled at the native pixel resolution. These simulations were used to test the versatility and robustness of ATOCA on more realistic simulations. They also include the  $1/f$  noise, the effects of flat-fielding and bad pixels. In all situations, the incident stellar fluxes are taken from high-resolution PHOENIX synthetic spectra (Husser et al., 2013) and the transit models from SCARLET (Benneke, 2015).

## 4.5.2. Results

For the first series of validations, the simulations were made with ATOCA itself to assess the performances of the decontamination and internal consistency. For each decontamination, the uneven wavelength grid was built using the procedure described in Section 4.3.4 paragraph *Grid sampling*. The relative tolerance had to be less than  $10^{-3}$ , a level of precision that implies a reasonable grid length, and hence moderate computational time. We also do not expect the precision of a single pixel to get higher than this, which corresponds to a SNR of 1000.

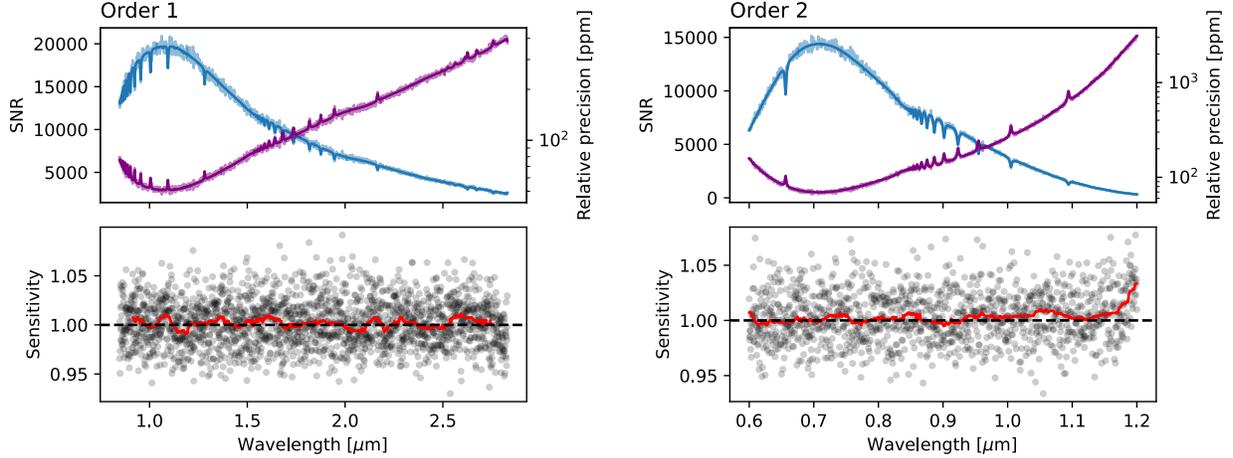
A first test was done on a single exposure by comparing each extraction to an equivalent simulation free of contamination. Different maximum pixel signal-to-noise ratios were tested, ranging from 200 to 10,000. This verifies our ability to decontaminate observations with the most extreme levels of precision planned for the NIRISS/SOSS mode. An example of the results is shown in Figure 8 for a star with an effective temperature of 2300 K and for a SNR of 10,000. The decontamination performance proved to be very effective at removing the contamination from the second order. In the case presented in Figure 8, ATOCA was able to go from a contamination of  $\sim 10000$  ppm, equivalent to a hundred times the expected uncertainties, to virtually no contamination. The residuals fall within the expected uncertainties ( $<100$  ppm) and seem free of any correlated noise. The same conclusions hold for all stellar temperatures. It is also interesting to note the clear cut around  $0.85 \mu\text{m}$  in the contamination levels (before decontamination). This is an artefact of the simulations since the monochromatic kernels used for the two-dimensional convolution only cover 128 native pixels. Thus, there is a threshold at 69 rows around the center of the trace where the wings of the spatial profile are not modeled. In the context of real observations, the contamination levels should extend below  $0.85 \mu\text{m}$ , while continuing to decrease.

The second and third tests were inspired by the commissioning programs COM/NIRISS 1091 (Martel et al., 2020) and COM/NIRISS 1541 Espinoza et al. (2021a). The former is a flat time series observation comprising 876 integrations on the standard A1V star, BD+601753. This will quantify the stability of the decontamination on multiple integrations for representative SNRs (177 at maximum pixel). The results are presented in Figure 9. The flat sensitivity spectrum shows that the decontamination of multiple integrations is stable. The combined spectrum reaches a precision of less than 100 ppm at best and between 150 and



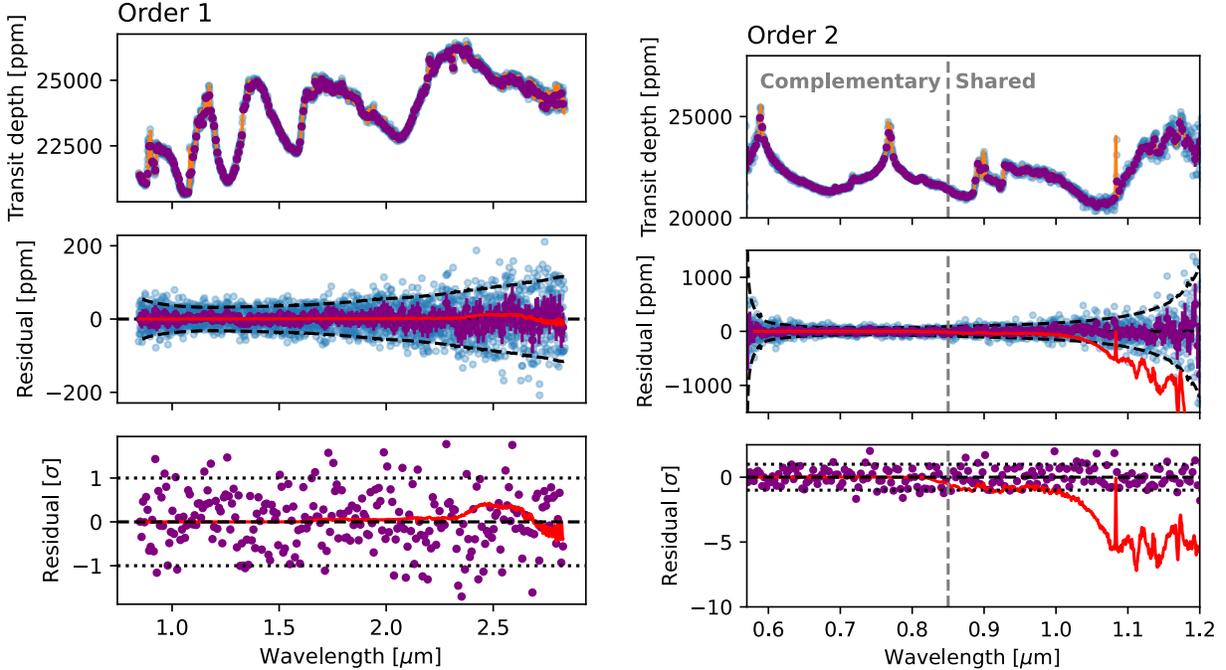
**Figure 8** – Decontamination on a single image. The extracted spectra are shown in the top panel. The four other panels show the extraction residuals before and after application of ATOCA, at two different scales. The dashed curves correspond to the expected  $1\text{-}\sigma$  uncertainties (in absolute value for the panels 2 and 3). For the bottom panel, the  $1\text{-}\sigma$  thresholds are marked by a dotted line. The simulation is the equivalent of a co-added image with a signal-to-noise ratio of 10,000 for a 2300 K star.

400 ppm in the regions subject to contamination. The measured standard deviation along the time series is in good agreement with the expected uncertainties in the entire spectral range, as evidenced by the sensitivity curve. Note that these results only accounts for photon noise; in realistic observations, other sources of noise like  $1/f$ , jitter or other detector effects might become dominant at certain wavelengths.



**Figure 9** – Stability of the decontaminated spectra within a time series. The top panel displays the expected signal-to-noise ratio (opaque blue) and the corresponding precision (opaque purple) of the time-series observation of BD+601753. To determine the measured signal-to-noise and precision (light blue and light purple), the standard deviation for each spectral channel is computed using the box-extracted spectra of the 876 integrations. The sensitivity shown at the bottom panel (light black dots) is obtained by taking the ratio of the measured (light purple) and expected (opaque purple) scatter for each spectral channel. A smoothed version is represented by the red curve, using a median filter of 81 pixels, equivalent to  $\sim 0.8 \mu\text{m}$  for order 1 and  $\sim 0.4 \mu\text{m}$  for order 2.

The latter program consists of another time-series observation of an expected featureless transit, evaluating the precision of relative measurements. Our simulation was based on the primary target of this program, the massive hot-Jupiter HAT-P-14 b, which was simplified to a step-transit (no limb darkening, instantaneous ingress and egress) with a flat transmission spectrum. We also neglected the effect of non-linearity in the ramps and forced a signal-to-noise ratio of 400 per pixel at maximum. This is above the capability of a SOSS-mode single integration, but it allows to push the decontamination at higher levels. The same framework was also applied to a transit spectrum of an exoplanet similar to WASP-107 b in a fourth verification (same transmission spectrum and same star, but different magnitude). In this case however, the signal-to-noise was artificially increased to 1000 for the same reason as mentioned above. Both validations led to the same conclusions, so the results of the flat transit are not shown here. The WASP-107 b-like transmission spectrum is presented in Figure 10. It confirms that the procedure can reach the expected precision and accuracy on relative measurements. Even with a relative tolerance of  $10^{-3}$  for ATOCA’s wavelength grid (see Section 4.3.4 paragraph *Grid sampling*), the combination of all extractions reaches an accuracy of less than 100 ppm in the order 1 contaminated region with, again, no evidence of systematic bias. The performance is even more obvious when it comes to the second order, where the contamination reaches levels of 1000 ppm. For comparison, Figure 11 presents the

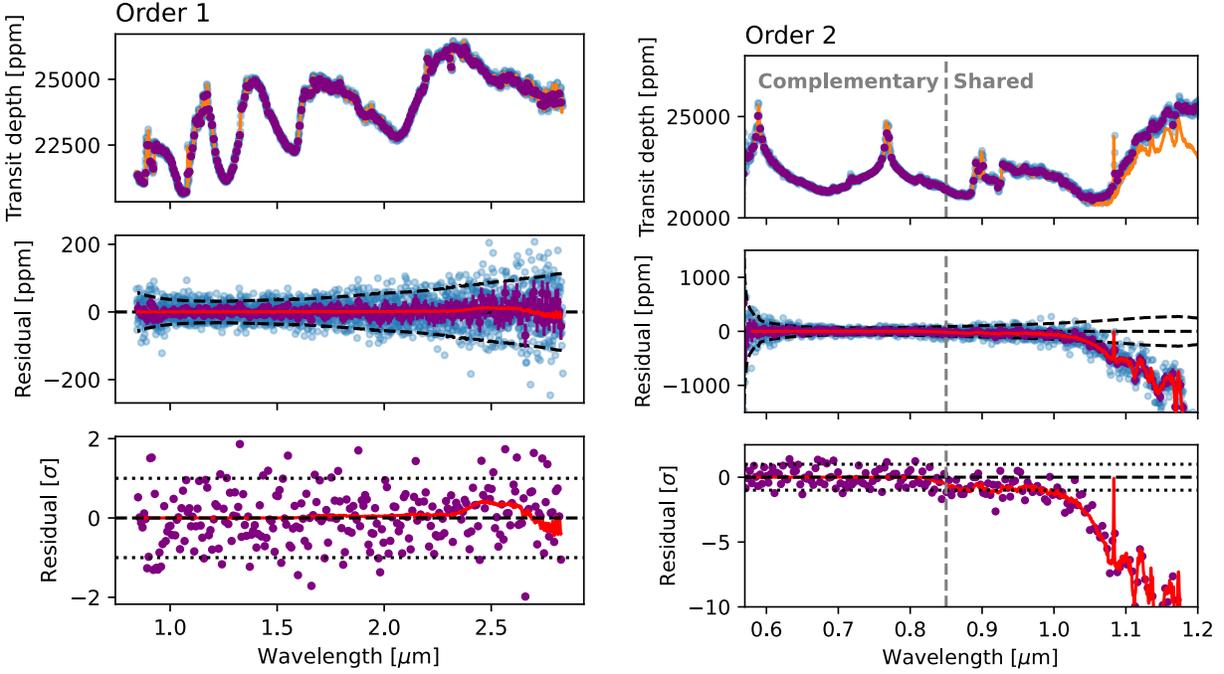


**Figure 10** – Decontaminated transmission spectrum for a target similar to WASP-107 b. The transmission spectrum as well as the residual in ppm and scaled to the uncertainties  $\sigma$  are presented by the three panels. In each of them, the results are shown at native sampling in blue and with 8-pixel bins in purple. The dashed black lines indicate the expected value or the  $1\sigma$  uncertainties. The estimation of the contamination from equation 4.2.5 is plotted in red. In the first panel, the input transmission spectrum is also shown in orange. The vertical gray dashed line delimit the wavelength range from the second order that is complementary to (or shared with) the first order.

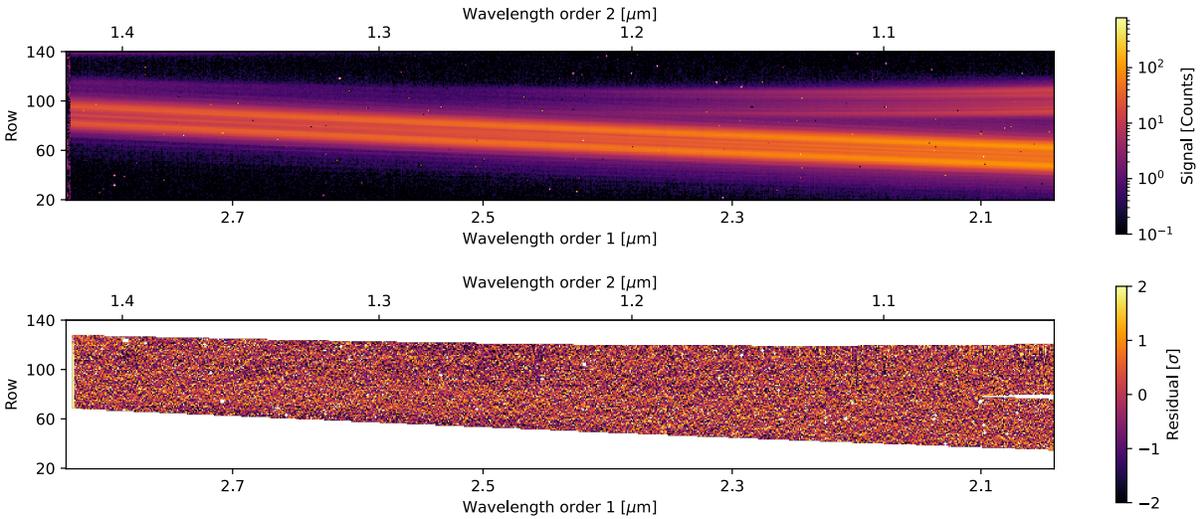
results of the WASP-107 b time series without any decontamination. We can see that the polluting signal follows the expected curve (red) taken from equation 4.2.5. It also confirms that the levels of contamination for the first order are small compared to the actual signal. The result of the flat transit is not shown here since it leads to the same conclusions.

Based on these four tests, some remarks can be made. First, it is interesting to note that they entail qualitatively different high-resolution PHOENIX synthetic spectra (Husser et al., 2013) at  $T_{eff} = 9400$  K (BD+601753),  $T_{eff} = 6700$  K (HAT-P-14),  $T_{eff} = 4500$  K (WASP-107) and  $T_{eff} = 2300$  K; the hottest spectrum showing well-defined absorption features on a smooth continuum and the coldest one containing features with noise-like behaviour. This is an assessment of the robustness of ATOCA regarding the nature of the underlying spectrum  $f_k$ .

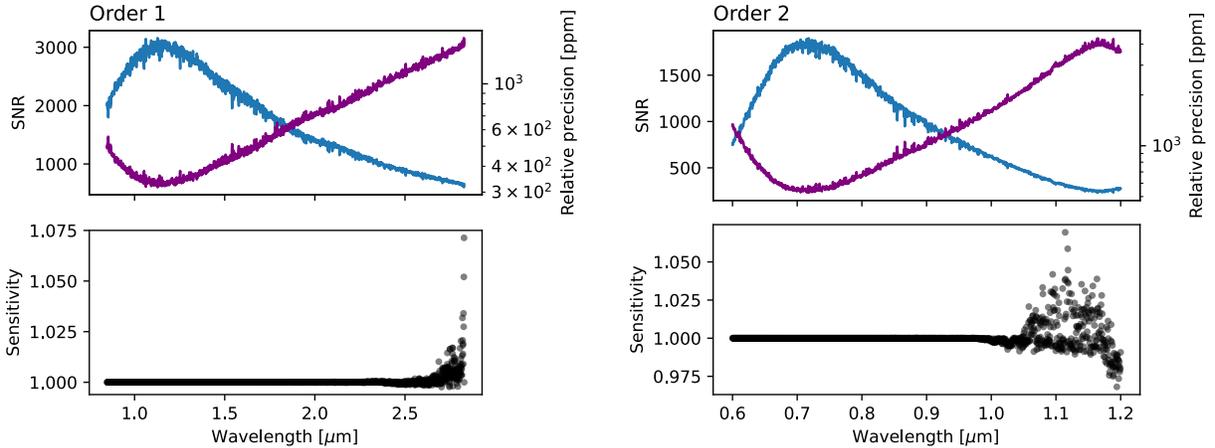
Second, the tests were initially run assuming that the spatial profiles, the wavelength solutions, and the throughputs were exactly known, resulting in errors consistent with the expected noise limits for each application of ATOCA. However, the same tests were repeated



**Figure 11** – Transmission spectrum for a target similar to WASP-107 b, without decontamination. Same description as in figure 10.



**Figure 12** – Comparison between the ATOCA modeling and the detector image of a single exposure. The top panel presents a close up of the region of the detector where the overlap occurs, corresponding to the lower left corner of Figure 2. The residual between the model extracted by ATOCA and the simulated detector image is shown in the bottom panel. The color scale is in units of  $\sigma$ , the pixels uncertainty. The pixels that were not used for the fit (e.g., bad pixels or background) are in white.



**Figure 13** – Stability of the decontaminated extraction of the simulated WASP-52 time series. The top panel displays the expected signal-to-noise ratio (blue) and the corresponding precision (purple) of the time-series observation of WASP-52 along the 103 integrations. The sensitivity shown at the bottom panel is given by the ratio of the measured scatter before and after decontamination.

with the reference files slightly shifted from their nominal values to confirm the robustness of the decontamination by applying a rotation and spatial shift (see end of Section 4.4). We found no evidence that it affected the extraction for reasonable values, i.e., within the expected precision of the reference files. More precisely, we tested for shift in the dispersion direction up to 0.5 pixel, for shift in the spatial axis up to 0.1 pixel and for rotations up to 0.01 deg.

Third, based on the apparent agreement between the expected and measured transit contamination seen in Figure 11, it would seem that the correction for contamination could be made after a standard extraction, directly on the one-dimensional spectra using equation 4.2.5. However, these examples used an idealized transit, with a perfectly stable stellar spectrum and without considering any limb darkening. Further analysis should be done before making any conclusion on this possible alternative to correct for order contamination.

ATOCA was also applied on the realistic time series simulation from the IDT to test the robustness of the algorithm. We present here an example on WASP-52. In this case, the tests were designed to assess the quality of integration-by-integration decontamination as well as the stability of the decontamination. Only the stellar spectrum was included since adding a transit would only add complexity to the interpretation of the results, without bringing additional information. The time series comprises a total of 103 integrations with a signal-to-noise ratio per pixel reaching up to  $\sim 250$ . Contrary to the more simplistic simulations, we did not have access to each individual order, which is more representative of the context of real observations. Therefore, we used the residual of the full detector model (combined

orders) for individual integrations as an indicator of the quality of the decontamination. The logic being that, if the model is able to represent correctly the overlapping region as well as the pixels covering the wavelength domain shared between both orders, then we can be confident that the overall model is accurate. Figure 12 presents the residuals for a single integration, given by the equation

$$\text{residual} = \frac{\text{observation} - \text{model}}{\text{uncertainty}}. \quad (4.5.1)$$

In this situation, since the input spectrum is perfectly stable, the uncertainties could be determined empirically using the standard deviation of each pixel throughout the full time series. The result is consistent with Gaussian noise and there is no evidence of any correlated features.

Concerning the stability of the decontamination, we had to do a similar comparison as in Figure 9. This is done to evaluate if an imperfect knowledge of the reference files would introduce additional scatter in the decontaminated spectra. However, this time, the spectrum extracted with the standard technique could not be used directly to avoid a possible contribution from the bad pixel modeling. Instead, we used the standard deviation of the pixels along all integrations, as it was done above to estimate the uncertainties, and then computed the summation in quadrature with the weight specific to the extraction method. In this manner, the bad pixels are not included in the summation. This was done for the time series before and after decontamination. The results are presented in Figure 13. The sensitivity increases slightly at longer wavelengths where the contamination is at its peak, but it remains contained below 5% for practically all of the domain of both spectral orders. This means that the precision, which is around 1000 ppm in the current example as shown in the top panel, would differ by only 50 ppm.

Based on these two tests, we can conclude that ATOCA can model the detector image within the uncertainties and at a low cost in terms of noise. This shows again that the algorithm is robust to inexact reference files.

On a different note, it is important to mention that the quality of the modeling is strongly influenced by the intertwining between both orders in their shared wavelength domain. This effect is accentuated in regions where the signal from both orders is strong, in which case a poor representation of the relation between the two can lead to over- and under-estimation. It can also be compensated with lower regularization factors, i.e., over-fitting. This was seen in many situations where the reference files were biased on purpose, as well as in the realistic simulations. This effect can be overcome by a proper estimation of the reference files. Thankfully, the regions that require higher precision are practically free of contamination, so the calibrations of the spatial profiles, the wavelength solutions and the throughput are relatively straightforward. Conversely, in the overlapping region, the throughput from the

second order drops considerably, leaving the solution of the underlying flux dominated by the corresponding wavelength from the first order. This means that the model of the order 2 in the overlapping region is more permissive.

## 4.6. Future improvements

Hyper-parameters. The choice of regularization parameters could benefit from further improvements. Indeed, the current criterion could lead to unstable solutions, which would be very effective at modeling each valid pixel of the detector, but mediocre when it comes to accurately estimating any pixels that are not included in the fit. The latter objective could be achieved using other criteria. For example, non-exhaustive cross-validation techniques (e.g., k-fold, Monte-Carlo) would be a judicious choice since their primary objective is to be able to simulate a set of data points that are voluntarily excluded from the fit.

Choice of Tikhonov matrix. The injection-recovery tests on our simulations pointed towards comparable performances with the first and second derivative operator. The former was preferred due to its slightly lower complexity. However, the latter could end up being a more appropriate choice. Indeed, the second derivative is used in spline interpolation on noisy data to smooth out the solutions, which is not far from the problem we are facing here. It would also be a more physical explanation, given that the finite resolution of observations enforces a smooth solution with only small variations of the second derivative.

Background fitting. One forthcoming challenge with the NIRISS/SOSS mode is the background subtraction. While the spatial spread of the trace profile is very effective to improve the precision of the measurements, it also greatly reduces the number of pixels available to measure and remove the background contribution. This problem is even more concerning in the SUBSTRIP96 observing mode, where the two traces cover the entire range of rows for some columns. ATOCA could circumvent this problem by directly including the background in the fitting, adding the parameters needed to model the background at each column to the solution vector.

## 4.7. Conclusion

In this work, we presented an alternative spectral extraction method to solve the overlap problem pertaining to the NIRISS/SOSS mode. We first characterized the extent of the contamination for the first and second diffraction orders. It was found that for absolute measurements, the levels were kept below 0.1% for most of the wavelength domain of the first order, except for wavelengths greater than  $2.6 \mu\text{m}$  where they reach 1%. For the second order, the effect is much more important, but concerns mainly the wavelength range already covered by order 1. For relative measurements, for which the SOSS mode is specifically designed, the same levels of contamination are expected, but only on the chromatic differences of the signal

(e.g., the difference in transit depth). This means that the systematic error due to the overlap should not be the dominant source of noise in the first diffraction order, although one should always assess its importance.

Nevertheless, it is still important to provide a way to disentangle each order’s contribution to at least quantify the contamination, but also to allow a proper extraction for any absolute measurements or scenarios where the relative contamination becomes non-negligible. Consequently, we developed ATOCA, an algorithm that enables the modeling and extraction of overlapping orders (or sources), and decontamination the detector image. We showed that, given reasonable estimates of the spatial profiles, the wavelength solutions, and the spectral throughputs, ATOCA was able to decontaminate the data up to the required precision. We also characterized the robustness of the decontamination by introducing errors in the reference files and by applying it to realistic simulations. A first version of the algorithm is available in the JWST official pipeline. A development version is also available on github<sup>6</sup>.

ATOCA is a promising technique to disentangle the contributions of overlapping spectral traces. Its framework might be transferable to other contexts, like field decontamination or multi-object slitless spectroscopy. ATOCA might also provide a powerful alternative to manage distorted wavelength solutions. All this potential has yet to be vetted throughout real observations which should come soon with the upcoming commissioning of NIRISS/SOSS.

Acknowledgments. The name of the algorithm, ATOCA, is a word used in North-America French to designate the cranberry fruit. It was borrowed from the native American languages, possibly from Algonquin spoken by nations living in the area of the present-day Wisconsin (Fitzpatrick, 2013) or from the Wandat word *atokha* (TLFQ, 1998). In this regard, we want to acknowledge the pivotal contribution of the First Nations to the North American French culture. We want to thank Anne Boucher for the design of the ATOCA logo. This project was undertaken with the financial support of the Canadian Space Agency (CSA-ASC) and the *Fonds de Recherche du Québec en Nature et Technologies* (FRQNT). We would also like to thank the Space Telescope Science Institute (STScI) for their trust and help during the implementation process. A.D.B., C.P. and S.P. wants to thank the Technologies for Exo-Planetary Science (TEPS) CREATE program, without whom this research would not be possible. The authors also acknowledge financial and social support of the Institute for Research on Exoplanets (iREx) and the University of Montreal. MR would like to acknowledge funding from FRQNT, as well as the National Sciences and Engineering Research Council of Canada (NSERC). C.P. acknowledges financial support by the NSERC Vanier Scholarship. D.J. is supported by NRC Canada and by an NSERC Discovery Grant. Support for J.D.T. was provided by NASA through the NASA Hubble Fellowship grant # HST-HF2-51495.001-A awarded by the Space Telescope Science Institute, which is operated by the Association of

---

6. <https://github.com/AntoineDarveau/jwst>

Universities for Research in Astronomy, Incorporated, under NASA contract NAS5-26555. R. A. is a Trottier Postdoctoral Fellow and acknowledges support from the Trottier Family Foundation

*Software :*

- WebbPSF (Perrin et al., 2015)
- JWST Data Management System <https://jwst-pipeline.readthedocs.io/en/latest/index.html>
- `scipy` (Virtanen et al., 2020)
- `ipython` (Pérez & Granger, 2007)
- `matplotlib` (Hunter, 2007)
- `numpy` (Harris et al., 2020)

## Appendix A : Minimization of the $\chi^2$

From equation 4.3.9, we can compute the derivative of the  $\chi^2$  :

$$\begin{aligned} \frac{d\chi^2}{df_k} &= \frac{d}{df_k} \sum_i \frac{1}{2} \left( \frac{D_i - M_i}{\sigma_i} \right)^2 \\ &= \sum_i - \left( \frac{D_i - M_i}{\sigma_i} \right) \frac{d}{df_k} \frac{M_i}{\sigma_i} \\ &= \sum_i - \left( \frac{D_i - \sum_{k'} B_{ik'} f_{k'}}{\sigma_i} \right) \frac{d}{df_k} \frac{1}{\sigma_i} \sum_{k''} B_{ik''} f_{k''}. \end{aligned} \quad (4.7.1)$$

However, since the  $B_{ik''}$  are simple coefficients, we have

$$\frac{d}{df_k} \sum_{k''} B_{ik''} f_{k''} = \sum_{k''} B_{ik''} \frac{df_{k''}}{df_k} = \sum_{k''} B_{ik''} \delta_{kk''} = B_{ik}, \quad (4.7.2)$$

to finally end up with

$$\frac{d\chi^2}{df_k} = \sum_i - \left( \frac{D_i - \sum_{k'} B_{ik'} f_{k'}}{\sigma_i} \right) \frac{B_{ik}}{\sigma_i}. \quad (4.7.3)$$

By setting this equation equal to zero, we have

$$0 = \sum_i - \frac{D_i}{\sigma_i} \frac{B_{ik}}{\sigma_i} + \sum_i \sum_{k'} \frac{B_{ik}}{\sigma_i} \frac{B_{ik'}}{\sigma_i} f_{k'}, \quad (4.7.4)$$

or

$$\sum_i \frac{D_i}{\sigma_i} \frac{B_{ik}}{\sigma_i} = \sum_i \sum_{k'} \frac{B_{ik}}{\sigma_i} \frac{B_{ik'}}{\sigma_i} f_{k'}. \quad (4.7.5)$$

## Appendix B : Trapezoidal integration on a grid

Consider the following integral :

$$M_i = \sum_n \int_{\lambda_{ni}^-}^{\lambda_{ni}^+} P_{ni} T_n(\lambda) \tilde{f}_n(\lambda) \lambda d\lambda = \sum_n \int_{\lambda_{ni}^-}^{\lambda_{ni}^+} \tilde{g}_{ni}(\lambda) d\lambda, \quad (4.7.6)$$

with  $\tilde{g}$  being the integrand. We can approximate the integral in a discrete form with the trapezoidal method,

$$\begin{aligned} M_i \simeq \sum_n \left( \frac{\tilde{g}_{ni}(\lambda_{ni}^-) + \tilde{g}_{niL_{ni}}}{2} (\lambda_{L_{ni}} - \lambda_{ni}^-) \right. \\ \left. + \sum_{\tilde{k}=L_{ni}+1}^{H_{ni}-1} \frac{\tilde{g}_{ni\tilde{k}} + \tilde{g}_{ni(\tilde{k}+1)}}{2} \Delta\lambda_{\tilde{k}} \right. \\ \left. + \frac{\tilde{g}_{niH_{ni}} + \tilde{g}_{ni}(\lambda_{ni}^+)}{2} (\lambda_{ni}^+ - \lambda_{H_{ni}}) \right), \end{aligned} \quad (4.7.7)$$

where  $L_{ni}$  and  $H_{ni}$  are respectively the lowest and highest  $\tilde{k}$  located into the pixel, so that  $L_{ni}$  is the smallest  $\tilde{k}$  for which  $\lambda_{ni}^- \leq \lambda_{\tilde{k}}$  and  $H_{ni}$  is the biggest  $\tilde{k}$  for which  $\lambda_{\tilde{k}} \leq \lambda_{ni}^+$ . Now, we

need to express  $\tilde{g}_{ni}(\lambda_{ni}^-)$  and  $\tilde{g}_{ni}(\lambda_{ni}^+)$  with the  $\tilde{g}_{ni\tilde{k}}$ . To do so, we can use linear interpolation to write

$$\begin{aligned}\tilde{g}(\lambda_i^-) &= \frac{\lambda_{L_{ni}} - \lambda_i^-}{\Delta\lambda_{L_{ni}-1}} \tilde{g}_{L_{ni}-1} + \frac{\lambda_i^- - \lambda_{L_{ni}-1}}{\Delta\lambda_{L_{ni}-1}} \tilde{g}_{L_{ni}} \\ \tilde{g}(\lambda_i^+) &= \frac{\lambda_{H_{ni}+1} - \lambda_i^+}{\Delta\lambda_{H_{ni}}} \tilde{g}_{H_{ni}} + \frac{\lambda_i^+ - \lambda_{H_{ni}}}{\Delta\lambda_{H_{ni}}} \tilde{g}_{H_{ni}+1}.\end{aligned}\quad (4.7.8)$$

Thus, this approximation will be permitted for pixels where  $\lambda_i^- \geq \lambda_{\tilde{k}=1}$  and  $\lambda_i^+ \leq \lambda_{\tilde{k}=N_{\tilde{k}}}$ . We can now substitute equation 4.7.8 into equation 4.7.7 and, after some algebra, we get

$$\begin{aligned}M_i &= \sum_n \frac{1}{2} \left( \frac{(\lambda_{L_{ni}} - \lambda_i^-)^2}{\Delta\lambda_{L_{ni}-1}} \tilde{g}_{L_{ni}-1} \right. \\ &\quad + \left( \lambda_{L_{ni}+1} - \lambda_i^- + \frac{(\lambda_i^- - \lambda_{L_{ni}-1})(\lambda_{L_{ni}} - \lambda_i^-)}{\Delta\lambda_{L_{ni}-1}} \right) \tilde{g}_{L_{ni}} \\ &\quad + \sum_{\tilde{k}=L_{ni}+1}^{H_{ni}-1} (\Delta\lambda_{\tilde{k}} + \Delta\lambda_{\tilde{k}-1}) \tilde{g}_{\tilde{k}} \\ &\quad + \left( \lambda_i^+ - \lambda_{H_{ni}-1} + \frac{(\lambda_i^+ - \lambda_{H_{ni}})(\lambda_{H_{ni}+1} - \lambda_i^+)}{\Delta\lambda_{H_{ni}}} \right) \tilde{g}_{H_{ni}} \\ &\quad \left. + \frac{(\lambda_i^+ - \lambda_{H_{ni}})^2}{\Delta\lambda_{H_{ni}}} \tilde{g}_{H_{ni}+1} \right).\end{aligned}\quad (4.7.9)$$

This has the form needed to compute the  $w_{ni\tilde{k}}$  in the equation

$$M_i = \sum_n \sum_{\tilde{k}=L_{ni}-1}^{H_{ni}+1} w_{ni\tilde{k}} \tilde{g}_{\tilde{k}}. \quad (4.7.10)$$

## Appendix C : Comparison to optimal extraction method

As highlighted in Horne (1986), the optimal extraction can be seen as the best correspondence (in the sense of a  $\chi^2$  minimization) between the sky-subtracted data and a model of the detector. The latter is a function of the incident flux expressed as a scaling of the spatial profile. This is also the main idea of ATOCA and its relation with the optimal extraction method in the absence of contamination is demonstrated in the following lines.

Let's consider the situation where there is no tilt in the wavelength solution and only one order is present so there is no contamination. Then, the integration weights,  $w_{i\tilde{k}}$  (see 4.3.3), will only depend on the spectral axis. Moreover, by invoking the mean value theorem, the integral from equation 4.3.2 becomes

$$M_i = (\lambda_i^+ - \lambda_i^-) \cdot P_i \cdot T(\lambda_i^*) \cdot \tilde{f}(\lambda_i^*) \cdot \lambda_i^* \quad (4.7.11)$$

where  $\lambda_i^* \in [\lambda_i^-, \lambda_i^+]$ . Thus, by choosing the integration grid to be the  $\lambda_i^*$ , the factors from equation 4.3.3 are given by  $w_{ik} = (\lambda_i^+ - \lambda_i^-)$ ,  $a_k = \lambda_i^* \cdot T(\lambda_i^*)$  and  $\tilde{f}_k = \tilde{f}(\lambda_i^*)$ . Furthermore,

since the wavelength solution is fixed along the spatial axis, the index  $i$  can be removed from the notation, except for  $P_i$ . This leads to the model of a pixel given exactly by

$$M_i = P_i w_k a_k \tilde{f}_k. \quad (4.7.12)$$

Finally, after  $\chi^2$  minimization with the procedure described in Section 4.7, the best solution for  $\tilde{f}_k$  is given by a linear equation where each row is given by

$$\sum_i \frac{D_i \cdot P_i}{\sigma_i^2} = \sum_i \frac{P_i^2}{\sigma_i^2} w_k a_k \tilde{f}_k = \sum_i \frac{P_i^2}{\sigma_i^2} S_k, \quad (4.7.13)$$

where  $S_k = w_k a_k \tilde{f}_k$  is the integrated signal over a pixel, which corresponds to the results of an optimal extraction. Each row of this linear system is independent, so the solution is directly obtained by re-arranging the terms, resulting in the equation

$$S_k = \sum_i \frac{D_i \cdot P_i / \sigma_i^2}{\sum_i P_i^2 / \sigma_i^2}, \quad (4.7.14)$$

which is the exact definition of the optimal extraction.

This means that the determination of the spatial profile can be inspired by the optimal extraction. Moreover, considering that the product of ATOCA is similar to one of an optimal extraction method, it thus constitutes an alternative to optimal extraction that can account for a distorted wavelength solution. However, in this scenario, the integral will not be solved exactly for each pixel, leaving the extraction subject to numerical integration errors. Yet, these errors can be mitigated with a good choice of integration method or an oversampled grid.



# Chapitre 5

---

## Conclusion

Notre compréhension de l’atmosphère d’exoplanètes s’est beaucoup développée dans les dernières décennies, en particulier concernant celle des Jupiters chaudes. Les observations faites à partir des télescopes spatiaux Hubble et *Spitzer* ont permis notamment d’étudier le côté jour de ce type d’exoplanètes. Cependant, les analyses qui s’ensuivirent ont pu mettre en évidence les limites de ces instruments. Celles-ci découlent entre autres de la couverture en longueurs d’onde limitée de la caméra WFC3 de Hubble et des difficultés liées à l’extraction des données de *Spitzer*. De manière générale, leur faible résolution spectrale est aussi un enjeu, avec seulement deux points photométriques pour *Spitzer* couvrant une bande passante très large et une résolution inférieure à  $\sim 100$  pour WFC3. En ce qui concerne les Jupiters ultra chaudes ( $T_{eq} > 2200$  K), ces contraintes observationnelles ont un impact majeur sur la détermination de leurs propriétés physiques comme l’abondance relative des espèces chimiques présentes dans leur atmosphère ainsi que la structure en température. Cependant, il est possible de repousser ces limites en utilisant des instruments ayant une résolution spectrale beaucoup plus élevée, ou alors en ayant recours au tout nouveau télescope spatial James Webb, similaire à Hubble, mais avec une meilleure sensibilité, une résolution spectrale accrue et un domaine de longueurs d’onde beaucoup plus étendu. C’est dans ce contexte que s’insère cette thèse.

D’une part, nous avons présenté une des premières analyses conjointes de données spectrophotométriques de WFC3 et d’observations à haute résolution, en ciblant la face diurne de la Jupiter ultra chaude WASP-33 b. Pour ce faire, nous avons dû obtenir des données à l’aide de l’instrument SPIRou, ayant une résolution de  $\lambda/\Delta\lambda = 70\,000$  et une couverture spectrale allant de  $0.95$  to  $2.50\ \mu\text{m}$ . L’ajout de ces nouvelles données a permis d’établir une première contrainte sur l’abondance de CO ( $\log \text{VMR}_{\text{CO}} = -4.07^{+1.51}_{-0.60}$ ) en élargissant l’intervalle de longueurs d’onde observées et en détectant indéniablement la signature spectrale de cette molécule. Cette détection, en conformité à deux études précédentes, permet aussi de confirmer la présence d’une inversion de température révélée par la forme des raies de CO

en émission (et non en absorption). La sensibilité au CO apportée par SPIRou est d’autant plus utile que la signature spectrale de l’eau ciblée par WFC3 est quasiment absente du spectre d’émission de WASP-33 b à cause de la dissociation thermique. L’analyse combinée a aussi pu améliorer les contraintes sur le profil de température et déterminer des limites supérieures sur l’abondance d’autres molécules comme le H<sub>2</sub>O ( $\text{VMR}_{\text{H}_2\text{O}} < 10^{-3.09}$ ), le CO<sub>2</sub> ( $\text{VMR}_{\text{CO}_2} < 10^{-3.21}$ ), le TiO ( $\text{VMR}_{\text{TiO}} < 10^{-2.81}$ ) et le VO ( $\text{VMR}_{\text{VO}} < 10^{-2.40}$ ). Les données de WFC3 ont aussi permis d’apporter des contraintes (relativement larges) sur l’abondance de OH ( $\log \text{VMR}_{\text{OH}} = -5.02^{+1.67}_{-2.66}$ ) et de H<sup>-</sup> ( $\log \text{VMR}_{\text{H}^-} = -3.73^{+0.90}_{-2.41}$ ). Ces travaux illustrent bien le potentiel et la complémentarité de la spectroscopie à haute résolution faite à l’aide d’instruments au sol.

D’autre part, nous avons développé l’algorithme d’extraction ATOCA et l’avons appliqué au mode SOSS de l’instrument NIRISS de JWST. Ce mode d’observation spectrophotométrique a été conçu pour étudier l’atmosphère des exoplanètes entre 0,6 et 2,8  $\mu\text{m}$  avec une grande précision et une résolution de ( $\lambda/\Delta\lambda \approx 700$ ). Ces caractéristiques permettront de lever certaines des dégénérescences propres à WFC3. Cependant, les deux principaux ordres de diffraction produits par le grisme de NIRISS se superposent partiellement entre environ 1,6 et 2,8  $\mu\text{m}$  pour l’ordre 1 et entre 0,85 et 1,4  $\mu\text{m}$  pour l’ordre 2. Nous avons donc d’abord estimé le niveau de contamination qui pourrait découler d’une extraction conventionnelle par boîte. Dans les cas d’un spectre individuel, celui-ci peut atteindre jusqu’à 1% pour l’ordre 1. Il en va de même pour l’ordre 2 dans la région spectrale d’intérêt située entre 0,85 et 0,95  $\mu\text{m}$ . Dans le cas de mesures relatives prises en série temporelle tel que la spectroscopie de transit, les effets de cette contamination s’annulent en partie. En effet, la composante constante d’un spectre à l’autre disparaît, ne laissant qu’une contamination relative. Par exemple, dans le cas d’un spectre de transmission, la contamination sera du même ordre que pour une mesure absolue ( $\sim 1\%$ ), mais en s’appliquant plutôt sur le spectre de transmission. Ensuite, nous avons détaillé ATOCA et présenté les performances de la décontamination en l’appliquant sur des simulations d’observations du mode SOSS de NIRISS. Nous avons montré que la décontamination était efficace et n’introduisait pas d’erreurs systématiques pour une précision en dessous de 10 ppm.

Cette thèse est représentative d’une nouvelle ère concernant l’étude des exoplanètes, où de nouvelles techniques d’observation et d’analyse ainsi que la mise en service d’instruments dédiés ultra performants offriront une vision toute neuve sur leur atmosphère. Il sera donc possible de mieux comprendre les phénomènes s’y produisant, d’obtenir de meilleures contraintes sur leur composition et, ultimement, de mieux comprendre les processus de formation menant à cette grande diversité d’exoplanètes.

# Bibliographie

---

- Ackerman, A. S., & Marley, M. S. 2001, *ApJ*, 556, 872
- Ali-Dib, M., Mousis, O., Petit, J.-M., & Lunine, J. I. 2014, *ApJ*, 785, 125
- Allart, R., Bourrier, V., Lovis, C., et al. 2019, *A&A*, 623, A58
- Anglada-Escudé, G., & Butler, R. P. 2012, *ApJS*, 200, 15
- Arcangeli, J., Desert, J.-M., Line, M. R., et al. 2018, *The Astrophysical Journal*, 855, L30, arXiv : 1801.02489
- 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
- Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, *A&A*, 558, A33
- Astudillo-Defru, N., Forveille, T., Bonfils, X., et al. 2017, *A&A*, 602, A88
- Atreya, S. K., Crida, A., Guillot, T., et al. 2018, in *Saturn in the 21st Century*, ed. K. H. Baines, F. M. Flasar, N. Krupp, & T. Stallard (Cambridge University Press), 5–43
- Atreya, S. K., Mahaffy, P. R., Niemann, H. B., & Owen, T. C. 2002, *Highlights of Astronomy*, 12, 597
- Atreya, S. K., & Wong, A.-S. 2005, *Space Sci. Rev.*, 116, 121
- Baeyens, R., Decin, L., Carone, L., et al. 2021, *MNRAS*, 505, 5603
- Baranne, A., Mayor, M., & Poncet, J. L. 1979, *Vistas in Astronomy*, 23, 279
- Barber, R. J., Strange, J. K., Hill, C., et al. 2014, *Monthly Notices of the Royal Astronomical Society*, 437, 1828
- Barman, T. S., Hauschildt, P. H., & Allard, F. 2001, *ApJ*, 556, 885
- Barstow, J. K., Aigrain, S., Irwin, P. G. J., et al. 2013a, *MNRAS*, 430, 1188
- Barstow, J. K., Aigrain, S., Irwin, P. G. J., Fletcher, L. N., & Lee, J.-M. 2013b, *MNRAS*, 434, 2616
- Barstow, J. K., Aigrain, S., Irwin, P. G. J., Kendrew, S., & Fletcher, L. N. 2015, *Monthly Notices of the Royal Astronomical Society*, 448, 2546
- Barstow, J. K., Aigrain, S., Irwin, P. G. J., & Sing, D. K. 2017, *ApJ*, 834, 50

Batalha, N., Kalirai, J., Lunine, J., Clampin, M., & Lindler, D. 2015, ArXiv e-prints

Batalha, N., Bean, J. L., Stevenson, K. B., et al. 2017, The Transiting Exoplanet Community Early Release Science Program, JWST Proposal ID 1366. Cycle 0 Early Release Science

Batalha, N. E., & Line, M. R. 2017, *AJ*, 153, 151

Baudino, J. L., Bézard, B., Boccaletti, A., et al. 2015, *A&A*, 582, A83

Baxter, C., Désert, J.-M., Parmentier, V., et al. 2020, *A&A*, 639, A36

Baxter, C., Désert, J.-M., Tsai, S.-M., et al. 2021, *Astronomy & Astrophysics*, 648, A127

Bean, J. L., Stevenson, K. B., Batalha, N. M., et al. 2018, *PASP*, 130, 114402

Bedell, M., Hogg, D. W., Foreman-Mackey, D., Montet, B. T., & Luger, R. 2019, *AJ*, 158, 164

Beichman, C., Benneke, B., Knutson, H., et al. 2014, *PASP*, 126, 1134

Bell, T. J., & Cowan, N. B. 2018, *ApJ*, 857, L20

Beltz, H., Rauscher, E., Roman, M. T., & Guilliat, A. 2022, *AJ*, 163, 35

Benneke, B. 2015, ArXiv e-prints

Benneke, B. 2015, arXiv :1504.07655 [astro-ph], arXiv : 1504.07655

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

Benneke, B., & Seager, S. 2013, *ApJ*, 778, 153

Benneke, B., & Seager, S. 2013, *ApJ*, 778, 153

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

Berta, Z. K., Charbonneau, D., Bean, J., et al. 2011, *The Astrophysical Journal*, 736, 12

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

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

Bocquet, S., & Carter, F. W. 2016, *The Journal of Open Source Software*, 1

Bolton, A. S., & Schlegel, D. J. 2010, *Publications of the Astronomical Society of the Pacific*, 122, 248, aDS Bibcode : 2010PASP..122..248B

Bolton, S. J., Adriani, A., Adumitroaie, V., et al. 2017, *Science*, 356, 821

Booth, R. A., Clarke, C. J., Madhusudhan, N., & Ilee, J. D. 2017, *MNRAS*, 469, 3994

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

Boucher, A., Pelletier, S., Darveau-Bernier, A., Lafrenière, D., & Artigau, e. 2022, *MNRAS*, 498

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

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

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

Brogi, M., Snellen, I. A. G., de Kok, R. J., et al. 2013, *ApJ*, 767, 27

- Bruno, G., Lewis, N. K., Alam, M. K., et al. 2020, *Monthly Notices of the Royal Astronomical Society*, 491, 5361
- Burrows, A., Hubeny, I., Budaj, J., Knutson, H. A., & Charbonneau, D. 2007, *ApJ*, 668, L171
- Burrows, A., & Volobuyev, M. 2003, *ApJ*, 583, 985
- Burrows, A., Marley, M., Hubbard, W. B., et al. 1997, *ApJ*, 491, 856
- Butters, O. W., West, R. G., Anderson, D. R., et al. 2010, *A&A*, 520, L10
- Cadieux, C., Doyon, R., Plotnykov, M., et al. 2022, *AJ*, 164, 96
- Cameron, A. C., Guenther, E., Smalley, B., et al. 2010, *Monthly Notices of the Royal Astronomical Society*, 407, 507
- Cannon, A. J., & Pickering, E. C. 1918, *Annals of Harvard College Observatory*, 91, 1
- Carone, L., Mollière, P., Zhou, Y., et al. 2021, *A&A*, 646, A168
- Cartier, K. M. S., Beatty, T. G., Zhao, M., et al. 2016, *The Astronomical Journal*, 153, 34
- Chakrabarty, A., & Sengupta, S. 2019, *The Astronomical Journal*, 158, 39
- Changeat, Q., & Edwards, B. 2021, *The Astrophysical Journal Letters*, 907, L22
- 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
- Christiansen, J. L., Ballard, S., Charbonneau, D., et al. 2010, *ApJ*, 710, 97
- Cont, D., Yan, F., Reiners, A., et al. 2021, *Astronomy & Astrophysics*, 651, A33
- Cont, D., Yan, F., Reiners, A., et al. 2022, *A&A*, 657, L2
- Cornachione, M. A., Bolton, A. S., Eastman, J. D., et al. 2019, *Publications of the Astronomical Society of the Pacific*, 131, 124503, aDS Bibcode : 2019PASP..13114503C
- Crossfield, I. J. M. 2015, *PASP*, 127, 941
- Crouzet, N., McCullough, P. R., Burke, C., & Long, D. 2012, *ApJ*, 761, 7
- Cutri, R. M., Skrutskie, M. F., van Dyk, S., et al. 2003, *VizieR Online Data Catalog*, II/246
- Dang, L., Cowan, N. B., Hammond, M., et al. 2021, *A Hell of a Phase Curve : Mapping the Surface and Atmosphere of a Lava Planet K2-141b*, JWST Proposal. Cycle 1, ID. #2347
- Dash, S., Majumdar, L., Willacy, K., et al. 2022, *The Astrophysical Journal*, 932, 20
- de Boer, K. S., & Snijders, M. A. J. 1981, *IUE ESA Newsletter*, 14, 154
- de Kok, R., Brogi, M., Snellen, I., et al. 2013, *Astronomy & Astrophysics*, 554, A82
- Deming, D., Fraine, J. D., Sada, P. V., et al. 2012, *The Astrophysical Journal*, 754, 106
- Deming, D., Wilkins, A., McCullough, P., et al. 2013, *ApJ*, 774, 95
- Deming, D., Wilkins, A., McCullough, P., et al. 2013, *The Astrophysical Journal*, 774, 95, aDS Bibcode : 2013ApJ...774...95D
- Diamond-Lowe, H., Berta-Thompson, Z., Charbonneau, D., & Kempton, E. M. R. 2018, *AJ*, 156, 42
- Diamond-Lowe, H., Stevenson, K. B., Bean, J. L., Line, M. R., & Fortney, J. J. 2014, *ApJ*, 796, 66

- Dobbs-Dixon, I., Agol, E., & Burrows, A. 2012, *ApJ*, 751, 87
- Donati, J. F., Kouach, D., Moutou, C., et al. 2020, *MNRAS*, 498, 5684
- Doyon, R., Hutchings, J. B., Beaulieu, M., et al. 2012, in *Proc. SPIE*, Vol. 8442, *Space Telescopes and Instrumentation 2012 : Optical, Infrared, and Millimeter Wave*, 84422R
- Edwards, B. 2020, *The Astronomical Journal*, 14
- Eistrup, C., Walsh, C., & van Dishoeck, E. F. 2018, *A&A*, 613, A14
- Elkins-Tanton, L. T., & Seager, S. 2008, *ApJ*, 685, 1237
- Espinoza, N., Loic, A., Goudfrooij, P., et al. 2021a, *NIRISS Sensitivity and Stability for Transiting Exoplanet Observations*, JWST Proposal. Cycle 0
- Espinoza, N., Baeyens, R., Bello-Arufe, A., et al. 2021b, *Exploring the morning and evening limbs of a transiting exoplanet*, JWST Proposal. Cycle 1, ID. #2113
- Evans, T. M., Aigrain, S., Gibson, N., et al. 2015, *MNRAS*, 451, 680
- Evans, T. M., Sing, D. K., Kataria, T., et al. 2017, *Nature*, 548, 58
- Ferruit, P., Birkmann, S., Böker, T., et al. 2014, in *Proc. SPIE*, Vol. 9143, *Space Telescopes and Instrumentation 2014 : Optical, Infrared, and Millimeter Wave*, 91430A
- Fitzpatrick, S. 2013, *Cranberry*
- Flowers, E., Brogi, M., Rauscher, E., Kempton, E. M. R., & Chiavassa, A. 2019, *AJ*, 157, 209
- Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, *PASP*, 125, 306
- Fortney, J. J. 2005, *MNRAS*, 364, 649
- Fortney, J. J., Lodders, K., Marley, M. S., & Freedman, R. S. 2008, *ApJ*, 678, 1419
- Freedman, R. S., Lustig-Yaeger, J., Fortney, J. J., et al. 2014, *The Astrophysical Journal Supplement Series*, 214, 25
- Freedman, R. S., Lustig-Yaeger, J., Fortney, J. J., et al. 2014, *ApJS*, 214, 25
- Gandhi, S., Madhusudhan, N., Hawker, G., & Piette, A. 2019, *AJ*, 158, 228
- Genest, F., Lafrenière, D., Boucher, A., et al. 2022, *arXiv e-prints*, arXiv :2205.09859
- Giacobbe, P., Brogi, M., Gandhi, S., et al. 2021, *Nature*, 592, 205
- Gibson, N. P., Pont, F., & Aigrain, S. 2011, *MNRAS*, 411, 2199
- Gibson, N. P., Merritt, S., Nugroho, S. K., et al. 2020, *Monthly Notices of the Royal Astronomical Society*, 493, 2215
- Golub, G. H., Heath, M., & Wahba, G. 1979, *Technometrics*, 21, 215
- Goodman, J., & Weare, J. 2010, *Communications in Applied Mathematics and Computational Science*, 5, 65
- Gray, D. F. 2008, *The Observation and Analysis of Stellar Photospheres* (Cambridge University Press)
- Green, P. J., & Silverman, B. W. 1993, *Nonparametric regression and generalized linear models : a roughness penalty approach* (Crc Press)

- Greene, T., Beichman, C., Eisenstein, D., et al. 2007, in Proc. SPIE, Vol. 6693, Techniques and Instrumentation for Detection of Exoplanets III, 66930G
- Greene, T. P., Line, M. R., Montero, C., et al. 2016, ApJ, 817, 17
- Griffith, C. A. 2014, Philosophical Transactions of the Royal Society of London Series A, 372, 20130086
- Griffith, C. A. 2014, Philosophical Transactions of the Royal Society A : Mathematical, Physical and Engineering Sciences, 372, 20130086
- Guillot, T. 2010, A&A, 520, A27
- Guilluy, G., Sozzetti, A., Brogi, M., et al. 2019, A&A, 625, A107
- Hansen, C. J., Schwartz, J. C., & Cowan, N. B. 2014, MNRAS, 444, 3632
- Hansen, P. C. 1992, SIAM review, 34, 561
- Hansen, P. C., & O'Leary, D. P. 1993, SIAM journal on scientific computing, 14, 1487
- Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Nature, 585, 357
- Hastie, T. J., & Tibshirani, R. 1990, Statist. Sci., 35
- Hauschildt, P. H., Barman, T. S., Baron, E., & Allard, F. 2003, in Astronomical Society of the Pacific Conference Series, Vol. 288, Stellar Atmosphere Modeling, ed. I. Hubeny, D. Mihalas, & K. Werner, 227
- Haynes, K., Mandell, A. M., Madhusudhan, N., Deming, D., & Knutson, H. 2015, The Astrophysical Journal, 806, 146
- Helling, C., Tootill, D., Woitke, P., & Lee, E. 2017, A&A, 603, A123
- Helling, C., Ackerman, A., Allard, F., et al. 2008, MNRAS, 391, 1854
- Helling, C., Iro, N., Corrales, L., et al. 2019, A&A, 631, A79
- Heng, K., Menou, K., & Phillipps, P. J. 2011, MNRAS, 413, 2380
- 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, The Astronomical Journal, 163, 248
- Herman, M. K., Mooij, E. J. W. d., Jayawardhana, R., & Brogi, M. 2020, The Astronomical Journal, 160, 93
- Herrero, E., Morales, J. C., Ribas, I., & Naves, R. 2011, A&A, 526, L10
- Hobson, M. J., Bouchy, F., Cook, N. J., et al. 2021, A&A, 648, A48
- Horel, A. 1962, Chem. Eng. Progress., 58, 54
- Horne, K. 1986, PASP, 98, 609
- Hubeny, I., Burrows, A., & Sudarsky, D. 2003, ApJ, 594, 1011
- Hunter, J. D. 2007, Computing in Science & Engineering, 9, 90
- Husser, T. O., Wende-von Berg, S., Dreizler, S., et al. 2013, A&A, 553, A6
- Hynes, R. I. 2002, A&A, 382, 752
- Hynes, R. I. 2002, Astronomy and Astrophysics, 382, 752
- Jordán, A., Espinoza, N., Rabus, M., et al. 2013, ApJ, 778, 184

- Kataria, T., Sing, D. K., Lewis, N. K., et al. 2016, *ApJ*, 821, 9
- Kempton, E. M. R., Bean, J. L., Deming, D., et al. 2021, Unshrouding the Sub-Neptune Population : The Case of TOI-421b, JWST Proposal. Cycle 1, ID. #1935
- Kendrew, S., Scheithauer, S., Bouchet, P., et al. 2015, *PASP*, 127, 623
- Kesseli, A. Y., Snellen, I. A. G., Alonso-Floriano, F. J., Mollière, P., & Serindag, D. B. 2020, *The Astronomical Journal*, 160, 228
- Khmil, S. V., & Surdej, J. 2002, *Astronomy & Astrophysics*, 387, 347
- Knutson, H. A., Charbonneau, D., Allen, L. E., Burrows, A., & Megeath, S. T. 2008, *The Astrophysical Journal*, 673, 526, arXiv :0709.3984 [astro-ph]
- Knutson, H. A., Charbonneau, D., Cowan, N. B., et al. 2009, *ApJ*, 703, 769
- Knutson, H. A., Charbonneau, D., Noyes, R. W., Brown, T. M., & Gilliland, R. L. 2007a, *ApJ*, 655, 564
- Knutson, H. A., Howard, A. W., & Isaacson, H. 2010, *ApJ*, 720, 1569
- Knutson, H. A., Charbonneau, D., Allen, L. E., et al. 2007b, *Nature*, 447, 183
- Komacek, T. D., & Showman, A. P. 2016, *ApJ*, 821, 16
- Komacek, T. D., Showman, A. P., & Parmentier, V. 2019, *ApJ*, 881, 152
- Kovács, G., Kovács, T., Hartman, J. D., et al. 2013, *Astronomy & Astrophysics*, 553, A44, arXiv : 1205.5060
- Kreidberg, L., Line, M. R., Thorngren, D., Morley, C. V., & Stevenson, K. B. 2018, *ApJ*, 858, L6
- Kreidberg, L., Bean, J. L., Désert, J.-M., et al. 2014, *ApJ*, 793, L27
- Kreidberg, L., Bean, J. L., Désert, J.-M., et al. 2014, *Nature*, 505, 69, arXiv :1401.0022 [astro-ph]
- 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
- Kreidberg, L., Line, M. R., Parmentier, V., et al. 2018, *The Astronomical Journal*, 156, 17
- Kunasz, C. V., Jefferies, J. T., & White, O. R. 1973, *A&A*, 28, 15
- Lafarga, M., Ribas, I., Lovis, C., et al. 2020, *A&A*, 636, A36
- Lafreniere, D. 2017, NIRISS Exploration of the Atmospheric diversity of Transiting exoplanets (NEAT), JWST Proposal. Cycle 1, ID. #1201
- Lecavelier Des Etangs, A., Vidal-Madjar, A., Désert, J. M., & Sing, D. 2008, *A&A*, 485, 865
- Lee, J. M., Fletcher, L. N., & Irwin, P. G. J. 2012, *MNRAS*, 420, 170
- Li, C., Ingersoll, A., Bolton, S., et al. 2020, *Nature Astronomy*, 4, 609
- Li, G., Zhang, H., & Bai, Z. 2015, arXiv e-prints, arXiv :1506.06188
- Lim, O., Albert, L., Artigau, E., et al. 2021, Atmospheric reconnaissance of the TRAPPIST-1 planets, JWST Proposal. Cycle 1, ID. #2589
- Line, M. R., Wolf, A. S., Zhang, X., et al. 2013, *ApJ*, 775, 137
- Line, M. R., Stevenson, K. B., Bean, J., et al. 2016, *AJ*, 152, 203

Line, M. R., Brogi, M., Bean, J. L., et al. 2021, *Nature*, 598, 580

Lodders, K., & Fegley, B. 2002, *Icarus*, 155, 393

Lodders, K., & Fegley, B. 2002, *Icarus*, 155, 393

Lothringer, J. D., Barman, T., & Koskinen, T. 2018, *The Astrophysical Journal*, 866, 27

Louden, T., & Wheatley, P. J. 2015, *ApJ*, 814, L24

Louie, D. R., Deming, D., Albert, L., et al. 2018, *PASP*, 130, 044401

Lucy, L. B. 1964, *SAO Special Report*, 167, 93

Lucy, L. B., & Walsh, J. R. 2003, *The Astronomical Journal*, 125, 2266, aDS Bibcode : 2003AJ....125.2266L

MacDonald, R. J., & Madhusudhan, N. 2017, *MNRAS*, 469, 1979

Madhusudhan, N. 2012, *ApJ*, 758, 36

Madhusudhan, N., Amin, M. A., & Kennedy, G. M. 2014a, *ApJ*, 794, L12

Madhusudhan, N., Constantinou, S., Moses, J. I., Piette, A., & Sarkar, S. 2021, *Chemical Disequilibrium in a Temperate sub-Neptune*, JWST Proposal. Cycle 1, ID. #2722

Madhusudhan, N., Crouzet, N., McCullough, P. R., Deming, D., & Hedges, C. 2014b, *ApJ*, 791, L9

Mandel, K., & Agol, E. 2002, *ApJ*, 580, L171

Mansfield, M., Bean, J. L., Line, M. R., et al. 2018, *The Astronomical Journal*, 156, 10

Mansfield, M., Line, M. R., Bean, J. L., et al. 2021, *Nature Astronomy*, 5, 1224

Marley, M. S., Ackerman, A. S., Cuzzi, J. N., & Kitzmann, D. 2013, in *Comparative Climatology of Terrestrial Planets*, ed. S. J. Mackwell, A. A. Simon-Miller, J. W. Harder, & M. A. Bullock (University of Arizona Press), 367–392

Marsh, T. R. 1989, *PASP*, 101, 1032

Martel, A., Filippazzo, J., Volk, K., et al. 2020, *NIRISS GR700XD Flux Calibration*, JWST Proposal. Cycle 0

Mayo, A., Dressing, C., Fortenbach, C., et al. 2021, *Transmission Spectroscopy of the Super-Neptune WASP-166b*, JWST Proposal. Cycle 1, ID. #2062

Mayor, M., & Queloz, D. 1995, *Nature*, 378, 355

McCullough, P., & MacKenty, J. 2012, *Considerations for using Spatial Scans with WFC3*, Tech. rep., Space Telescope Science Institute

Mikal-Evans, T., Crossfield, I. J. M., Benneke, B., et al. 2021, *The Astronomical Journal*, 161, 18, aDS Bibcode : 2021AJ....161...18M

Mikal-Evans, T., Sing, D. K., Barstow, J. K., et al. 2022, *Nature Astronomy*, 6, 471

Miller-Ricci, E., Seager, S., & Sasselov, D. 2009, *ApJ*, 690, 1056

Min, L., Guang-wei, L., Ke, L., et al. 2020, *Chinese Astronomy and Astrophysics*, 44, 399

Mollière, P., van Boekel, R., Dullemond, C., Henning, T., & Mordasini, C. 2015, *ApJ*, 813, 47

Mollière, P. M. 2017, PhD thesis, Ruprecht-Karls University of Heidelberg, Germany

- Mollière, P., Wardenier, J. P., van Boekel, R., et al. 2019, *Astronomy & Astrophysics*, 627, A67
- Morales-Calderón, M., Stauffer, J. R., Kirkpatrick, J. D., et al. 2006, *ApJ*, 653, 1454
- Moses, J. I., Visscher, C., Fortney, J. J., et al. 2011, *ApJ*, 737, 15
- Mousis, O., Lunine, J. I., Madhusudhan, N., & Johnson, T. V. 2012, *ApJ*, 751, L7
- Nortmann, L., Pallé, E., Salz, M., et al. 2018, *Science*, 362, 1388
- Nugroho, S. K., Gibson, N. P., de Mooij, E. J. W., et al. 2020, *The Astrophysical Journal*, 898, L31
- Nugroho, S. K., Kawahara, H., Masuda, K., et al. 2017, *The Astronomical Journal*, 154, 221, arXiv : 1710.05276
- Nugroho, S. K., Kawahara, H., Gibson, N. P., et al. 2021, *The Astrophysical Journal Letters*, 910, L9
- Öberg, K. I., Murray-Clay, R., & Bergin, E. A. 2011, *ApJ*, 743, L16
- Oklopčić, A. 2019, *ApJ*, 881, 133
- Oklopčić, A., & Hirata, C. M. 2018, *ApJ*, 855, L11
- Parmentier, V., Fortney, J. J., Showman, A. P., Morley, C., & Marley, M. S. 2016, *ApJ*, 828, 22
- Parmentier, V., Showman, A. P., & Fortney, J. J. 2021, *MNRAS*, 501, 78
- Parmentier, V., Showman, A. P., & Lian, Y. 2013, *A&A*, 558, A91
- 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, arXiv :2105.10513 [astro-ph], arXiv : 2105.10513
- Pelletier, S., Benneke, B., Bean, J., et al. 2022, in *Bulletin of the American Astronomical Society*, Vol. 54, 100.04
- Pepe, F., Mayor, M., Galland, F., et al. 2002, *A&A*, 388, 632
- Pérez, F., & Granger, B. E. 2007, *Computing in Science and Engineering*, 9, 21
- Perez-Becker, D., & Showman, A. P. 2013, *ApJ*, 776, 134
- Perrin, M. D., Long, J., Sivaramakrishnan, A., et al. 2015, *WebbPSF : James Webb Space Telescope PSF Simulation Tool*, *Astrophysics Source Code Library*, record ascl :1504.007
- Phillips, D. L. 1962, *J. ACM*, 9, 84–97
- Pinhas, A., Madhusudhan, N., Gandhi, S., & MacDonald, R. 2019, *MNRAS*, 482, 1485
- Pino, L., Ehrenreich, D., Allart, R., et al. 2018, *A&A*, 619, A3
- Pino, L., Désert, J.-M., Brogi, M., et al. 2020, *The Astrophysical Journal*, 894, L27
- Piskunov, N., Wehrhahn, A., & Marquart, T. 2021, *Astronomy and Astrophysics*, 646, A32
- Pluriel, W., Whiteford, N., Edwards, B., et al. 2020, arXiv :2006.14199 [astro-ph], arXiv : 2006.14199
- Puetter, R., Gosnell, T., & Yahil, A. 2005, *Annual Review of Astronomy and Astrophysics*, 43, 139

- Queloz, D. 1995, in *New Developments in Array Technology and Applications*, ed. A. G. D. Philip, K. Janes, & A. R. Upgren, Vol. 167, 221
- Rackham, B., Espinoza, N., Apai, D., et al. 2017, *ApJ*, 834, 151
- Rackham, B. V., Apai, D., & Giampapa, M. S. 2018, *ApJ*, 853, 122
- Radica, M., Albert, L., Taylor, J., et al. 2022, arXiv e-prints, arXiv :2207.05136
- Rauscher, E., & Menou, K. 2010, *ApJ*, 714, 1334
- Rimmer, P. B., & Helling, C. 2016, *ApJS*, 224, 9
- Robertson, J. G. 1986, *Publications of the Astronomical Society of the Pacific*, 98, 1220, aDS  
Bibcode : 1986PASP...98.1220R
- 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
- Salz, M., Czesla, S., Schneider, P. C., et al. 2018, *A&A*, 620, A97
- Sánchez-López, A., López-Puertas, M., Snellen, I. A. G., et al. 2020, *A&A*, 643, A24
- Schlawin, E., Greene, T. P., Line, M., Fortney, J. J., & Rieke, M. 2018, *AJ*, 156, 40
- Schwartz, J. C., Kashner, Z., Jovmir, D., & Cowan, N. B. 2017, *ApJ*, 850, 154
- Schwarz, H., Brogi, M., Birkby, J., et al. 2013, 2
- Schwarz, H., Brogi, M., de Kok, R., Birkby, J., & Snellen, I. 2015, *A&A*, 576, A111
- Seager, S., & Deming, D. 2010, *ARA&A*, 48, 631
- Seidel, J. V., Ehrenreich, D., Allart, R., et al. 2021, *A&A*, 653, A73
- Seiff, A., Kirk, D. B., Knight, T. C. D., et al. 1998, *J. Geophys. Res.*, 103, 22857
- Serindag, D. B., Nugroho, S. K., Mollière, P., et al. 2021, *Astronomy & Astrophysics*, 645, A90
- Sheppard, K., Mandell, A. M., Tamburo, P., et al. 2017, *The Astrophysical Journal*, 850, L32, arXiv : 1711.10491
- Showman, A. P., Fortney, J. J., Lian, Y., et al. 2009, *ApJ*, 699, 564
- Sing, D. K., Wakeford, H. R., Showman, A. P., et al. 2015, *Monthly Notices of the Royal Astronomical Society*, 446, 2428
- Sing, D. K., Fortney, J. J., Nikolov, N., et al. 2016, *Nature*, 529, 59
- Snellen, I., de Kok, R., de Mooij, E., et al. 2010, *Proceedings of the International Astronomical Union*, 6, 208
- Spake, J., Rustamkulov, Z., Benneke, B., et al. 2021a, *The twin paradox : assessing planetary radius evolution with a CH<sub>4</sub> thermometer*, JWST Proposal. Cycle 1, ID. #2594
- Spake, J. J., Sing, D. K., Wakeford, H. R., et al. 2021b, *MNRAS*, 500, 4042
- Spiegel, D. S., Silverio, K., & Burrows, A. 2009, *ApJ*, 699, 1487
- Stassun, K. G., Collins, K. A., & Gaudi, B. S. 2017, *AJ*, 153, 136
- Stevenson, K. B., Bean, J. L., Madhusudhan, N., & Harrington, J. 2014, *ApJ*, 791, 36

- Stevenson, K. B., Bean, J. L., Seifahrt, A., et al. 2014, *The Astronomical Journal*, 147, 161
- Stevenson, K. B., & Fowler, J. 2019, *Analyzing Eight Years of Transiting Exoplanet Observations Using WFC3's Spatial Scan Monitor*, Instrument Science Report WFC3 2019-12, 16 pages
- Stevenson, K. B., Harrington, J., Nymeyer, S., et al. 2010, *Nature*, 464, 1161
- Sudarsky, D., Burrows, A., & Hubeny, I. 2003, *ApJ*, 588, 1121
- Swain, M. R., Vasisht, G., & Tinetti, G. 2008, *Nature*, 452, 329
- Swain, M. R., Tinetti, G., Vasisht, G., et al. 2009, *ApJ*, 704, 1616
- Tan, X., & Komacek, T. D. 2019, *The Astrophysical Journal*, 886, 26
- The JWST Transiting Exoplanet Community Early Release Science Team, Ahrer, E.-M., Alderson, L., et al. 2022, arXiv e-prints, arXiv :2208.11692
- Thompson, A. M. 1990, *A&A*, 240, 209
- Thorngren, D., Gao, P., & Fortney, J. J. 2019, *ApJ*, 884, L6
- Thorngren, D. P., & Fortney, J. J. 2018, *AJ*, 155, 214
- Tikhonov, A. N. 1963, in *Doklady Akademii Nauk*, Vol. 151, Russian Academy of Sciences, 501–504
- Tinetti, G., Deroo, P., Swain, M. R., et al. 2010, *ApJ*, 712, L139
- TLFQ. 1998, in *Dictionnaire historique du français québécois* (Québec : Presses de l'Université Laval)
- Tremblin, P., Amundsen, D. S., Mourier, P., et al. 2015, *ApJ*, 804, L17
- Tsai, S.-M., Lyons, J. R., Grosheintz, L., et al. 2017, *ApJS*, 228, 20
- Tsai, S.-M., Malik, M., Kitzmann, D., et al. 2021, *The Astrophysical Journal*, 923, 264
- Tsiaras, A., Waldmann, I. P., Zingales, T., et al. 2018, *The Astronomical Journal*, 155, 156
- Turner, J. D., Pearson, K. A., Biddle, L. I., et al. 2016, *MNRAS*, 459, 789
- van Sluijs, L., Birkby, J. L., Lothringer, J., et al. 2022, arXiv :2203.13234 [astro-ph], arXiv : 2203.13234
- Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, *Nature Methods*, 17, 261
- Visscher, C. 2012, *ApJ*, 757, 5
- Visscher, C., Moses, J. I., & Saslow, S. A. 2010, *Icarus*, 209, 602
- von Essen, C., Mallonn, M., Borre, C. C., et al. 2020, *Astronomy & Astrophysics*, 639, A34
- Wahba, G. 1977, *SIAM journal on numerical analysis*, 14, 651
- Wakeford, H. R., Sing, D. K., Deming, D., et al. 2013, *Monthly Notices of the Royal Astronomical Society*, 435, 3481
- Welbanks, L., Madhusudhan, N., Allard, N. F., et al. 2019, *ApJ*, 887, L20
- Welch, B. L. 1947, *Biometrika*, 34, 28
- Wong, I., Knutson, H. A., Lewis, N. K., et al. 2015, *ApJ*, 811, 122
- Yan, F., Pallé, E., Reiners, A., et al. 2022a, *Astronomy & Astrophysics*, 661, L6
- Yan, F., Reiners, A., Pallé, E., et al. 2022b, *Astronomy & Astrophysics*, 659, A7

- Yip, K. H., Tsiaras, A., Waldmann, I. P., & Tinetti, G. 2020, *The Astronomical Journal*, 160, 171
- York, D. G., Adelman, J., Anderson, John E., J., et al. 2000, *AJ*, 120, 1579
- Zahnle, K., Marley, M. S., Freedman, R. S., Lodders, K., & Fortney, J. J. 2009, *ApJ*, 701, L20
- Zhang, M., Knutson, H. A., Kataria, T., et al. 2018, *The Astronomical Journal*, 155, 83
- Zhang, X., Strobel, D. F., & Imanaka, H. 2017, *Nature*, 551, 352