

Development, maintenance, and post-processing tools for astrophysical spectroscopic databases

Jingxin Zhang

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

of

University College London.

Primary supervisor: **Jonathan Tennyson**

Secondary supervisor: **Sergey Yurchenko**

Department of Physics & Astronomy

University College London

April 7, 2025

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

Abstract

Molecular spectra are critical for characterizing the atmospheres of exoplanets and cool stars, providing crucial information into their composition, temperature, and physical conditions. This thesis focuses on the development, maintenance, and post-processing tools for astrophysical spectroscopic databases, with an emphasis on enhancing the accessibility and usability of high-resolution molecular line lists.

The ExoMol database provides extensive molecular line lists for high-temperature conditions relevant to exoplanetary and stellar atmospheres, which requires the maintenance and continuous updates. A newly structured file format has been implemented to improve data organization and retrieval efficiency.

This work presents PyExoCross, a Python adaptation of the Fortran-based ExoCross program, designed for post-processing huge molecular line lists. PyExoCross facilitates the computation of local thermodynamic equilibrium (LTE) and non-local thermodynamic equilibrium (non-LTE) stick spectra, cross sections, partition functions, specific heats, cooling functions, lifetimes, and oscillator strengths, supporting ExoMol, HITRAN, and HITEMP databases.

The ExoMolHR tools is introduced as an empirical high-resolution molecular spectrum calculator, enabling users to generate spectra for specific temperature and wavenumber/wavelength ranges. This web-based database allows users to download and visualize spectral data efficiently, bridging the gap between raw molecular line lists and practical astrophysical applications.

The advancements presented in this thesis significantly improve the computational tools available for analyzing molecular spectra in astrophysics, supporting ongoing and future exoplanetary and stellar atmosphere studies.

Impact Statement

The advancements in molecular spectroscopic databases and post-processing tools presented in this thesis have a significant impact on astrophysical research, particularly in the characterization of exoplanetary and stellar atmospheres. By improving the accessibility, accuracy, and computational efficiency of high-resolution molecular line lists, this work enhances the ability of researchers to interpret observational spectra from ground- and space-based telescopes.

The ExoMol line lists have been used in various laser-based techniques, including laser absorption spectroscopy, laser induced breakdown spectroscopy, laser fragmentation, and laser induced fluorescence. Additionally, they have been applied in plasma stoichiometry, fusion plasma research, gas sensor design, as well as in studies of combustion and explosions. The maintenance of the ExoMol database, including its new file structure, ensure the long-term reliability and usability of molecular data critical for atmospheric modeling.

The introduction of PyExoCross, a Python adaptation of ExoCross, streamlines the post-processing of extensive molecular line lists, making it easier to generate local thermodynamic equilibrium (LTE) and non-local thermodynamic equilibrium (non-LTE) spectra, cross sections, and thermodynamic properties. This facilitates more efficient simulations of astrophysical environments, benefiting both theoretical studies and observational data analysis.

The ExoMolHR tools provides a user-friendly platform for generating high-resolution molecular spectra, enabling researchers to directly access and visualize spectral data tailored to specific conditions. This enhances the integration of molecular data into atmospheric retrieval models, improving the accuracy of exoplanet

and cool star atmosphere characterization.

This research contributes to the broader astrophysical community, supporting future discoveries in planetary science, stellar astrophysics, and atmospheric modeling. The tools and methodologies developed in this thesis will remain valuable resources for the ongoing exploration of the universe.

Research Paper Declaration

Publications relevant to this thesis

- **Chapter 5**, **6**, and **7**

  Jingxin Zhang, Jonathan Tennyson, and Sergei N. Yurchenko. PyExoCross: a Python program for generating spectra and cross-sections from molecular line lists. *RAS Techniques and Instruments*, **3**(1):257–287, 2024. doi: 10.1016/j.jqsrt.2024.109083

- **Chapter 9**

  Jingxin Zhang, Christian Hill, Jonathan Tennyson, and Sergei N. Yurchenko. ExoMolHR: A Relational Database of Empirical High-resolution Molecular Spectra. *The Astrophysical Journal Supplement Series*, **276**(2):67, 2025. doi: 10.3847/1538-4365/ada288

- **Chapter 3** and **4**

  Jonathan Tennyson, Sergei N. Yurchenko, **Jingxin Zhang**, Charles A. Bowesman, Ryan P. Brady, Jeanna Buldyreva, Katy L. Chubb, Robert R. Gamache, Maire N. Gorman, Elizabeth R. Guest, Christian Hill, Kyriaki Kefala, A. E. Lynas-Gray, Thomas M. Mellor, Laura K. McKemmish, Georgi B. Mitev, Irina I. Mizus, Alec Owens, Zhijian Peng, Armando N. Perri, Marco Pezzella, Oleg L. Polyansky, Qianwei Qu, Mikhail Semenov, Oleksiy Smola, Andrei Solokov, Wilfrid Somogyi, Apoorva Upadhyay, Samuel O. M. Wright, and Nikolai F. Zobov. The 2024 release of the ExoMol database: molecular line lists for exoplanet and other hot atmospheres. *J. Quant. Spectrosc. Radiat. Transf.*, **326**:109083, 2024. doi: 10.1016/j.jqsrt.2024.109083

- **Chapter 3** and **4**

  Jonathan Tennyson, Marco Pezzella, **Jingxin Zhang**, and Sergei N. Yurchenko. Data structures for photoabsorption within the ExoMol project. *RAS Techniques and Instruments*, **2**(1):231–237, 2023. doi: 10.1093/rasti/rzad014

Other publications

- Sergei N. Yurchenko, Charles A. Bowesman, Ryan P. Brady, Elizabeth R. Guest, Kyriaki Kefala, Georgi B. Mitev, Alec Owens, Armando N. Perri, Marco Pezzella, Oleksiy Smola, Andrei Sokolov, **Jingxin Zhang**, and Jonathan Tennyson. ExoMol line lists – LX. Molecular line list for the ammonia isotopologue $^{15}NH_3$. *Monthly Notices of the Royal Astronomical Society*, **533**(3):3442–3456, 2024. doi:10.1093/mnras/stae1849

- Georgi B. Mitev, Charles A. Bowesman, **Jingxin Zhang**, Sergei N. Yurchenko, and Jonathan Tennyson. ExoMol line lists – LXI. A trihybrid line list for rovibronic transitions of the hydroxyl radical (OH). *Monthly Notices of the Royal Astronomical Society*, **536**(4):3401–3420, 2025. doi: 10.1093/mnras/stae2803

- Georgi B. Mitev, Marco Pezzella, Charles A. Bowesman, **Jingxin Zhang**, Sergei N. Yurchenko, and Jonathan Tennyson. ExoMol Photodissociation Cross Sections – II: Continuum Absorption and Predissociation Spectra for the Hydroxyl Radical. *Monthly Notices of the Royal Astronomical Society*, 2025. (Submitted)

Acknowledgements

I would like to express my deepest gratitude to my supervisors, Prof. Jonathan Tennyson and Prof. Sergey Yurchenko, for their invaluable time, expertise, and seemingly endless patience throughout my PhD journey. Without their guidance and encouragement, this endeavor would not have been possible. I am also grateful to my fellow PhD students and academic colleagues: Charles Bowesman, Ryan Brady, Bridgette Cooper, Elizabeth Guest, Kyriaki Kefala, Tony Lynas-Gray, Thomas Mellor, Georgi Mitev, Alec Owens, Armando Perri, Marco Pezzella, Oleg Polyansky, Qianwei Qu, Mikhail Semenov, Oleksiy Smola, Andrei Solokov, Wilfrid Somogyi, and Apoorva Upadhyay, for their continuous help, advice, companionship and support.

I am especially grateful to Prof. Jonathan Tennyson for his tremendous support in both my academic and personal life. He often shared his knowledge of physics with me, helping me bridge the gaps in my background as a mathematics undergraduate. I sincerely appreciate his meticulous guidance on my papers, tirelessly reviewing and refining them. He arranged for me to co-supervise master students as a second advisor, which provided me with valuable experience in mentoring others on research projects. He provided me the opportunities to attend academic conferences in several countries. These experiences allowed me to connect with researchers from diverse backgrounds, engage in valuable academic exchanges, explore different landscapes, and immerse myself in various cultures.

My heartfelt thanks also go to Prof. Sergey Yurchenko for his significant contributions to my coding projects. He patiently explained the underlying mathematical and physical logic and derivations behind many computational formulas, often

dedicating a great deal of time to ensure my understanding. His rigorous academic approach has profoundly influenced me, and I deeply appreciate the meticulous feedback he provided to refine my research papers.

My special thanks to Christian Hill, who assisted me with the database website and patiently answered my coding-related queries. I am incredibly thankful for his invitation to the International Atomic Energy Agency (IAEA) consultancy meeting, which was a truly special and unforgettable experience for me. I would also like to thank Qianwei Qu, who shared invaluable PhD experiences and advice with me, and Charles Bowesman, who disscussed with me about non-LTE theory and provided useful suggestions for optimization in my code. My sincere appreciation goes to Elizabeth Guest, Kyriaki Kefala, Armando Perri, and Marco Pezzella, who often helped me with trivial little matters encountered in life.

Living in the student accommodation during my PhD was a unique experience, and I want to thank all the roommates and friends I met over the years. Every September, I had the chance to meet new roommates and fellow residents, and although our time together was often short-lived, their companionship made my life fulfilling and colorful. I also extend my gratitude to my long-distance friends who, despite time zone differences, stayed connected and shared our life updates with each other.

Above all, I am most grateful to my parents for their unwavering support, encouragement and companionship. Finally, I would like to extend my appreciation to everyone at UCL who has been part of my journey.

Contents

1 Introduction **29**

1.1 Background and motivation 29

1.2 Objectives of the thesis 30

1.3 Structure of the thesis 32

1.4 Significance and contributions 33

1.5 Data availability 34

2 Spectra of Exoplanets **35**

2.1 Resolving power 35

2.2 Transit spectroscopy 36

2.3 High-resolution cross-correlation spectroscopy 38

2.4 Direct imaging spectroscopy 42

2.5 Data quality 42

2.6 Astrophysically important molecular databases 43

2.6.1 VAMDC portal 43

2.6.2 ExoMol database 44

2.6.3 TheoReTS database 44

2.6.4 NASA Ames database 45

2.6.5 Kurucz database 46

2.6.6 GEISA database 47

2.6.7 HITRAN and HITEMP databases 48

2.6.8 JPL database 49

2.6.9 CDMS database 49

2.6.10 CASDA24 database 50
2.6.11 VALD3 database 51
2.6.12 BASECOL database 52
2.6.13 W@DIS information system 52
2.6.14 CDSD, NOSD, NDSD, and ASD databases 53
2.6.15 LEIDEN photodissociation and photoionization cross-section database 55
2.7 Summary 56

3 EXOMOL and HITRAN Databases Overview 58
3.1 Quantum numbers 58
3.2 EXOMOL Database overview 61
3.2.1 Explicit use of the EXOMOL database 62
3.2.2 Data products from the EXOMOL database 63
3.3 EXOMOL file structure and data format 64
3.3.1 Master file 66
3.3.2 Definition file 67
3.3.3 Line list files 69
3.3.3.1 States file 69
3.3.3.2 Transitions file 70
3.4 EXOMOL API 72
3.4.1 EXOMOL DOI on ZENODO 72
3.4.2 Searching for data through the API 72
3.4.3 Accessing the data via the API 73
3.5 HITRAN Database overview 73
3.6 HITRAN file structure and data format 74
3.7 HITRAN API 74
3.8 Summary 74

4 Updating and Expanding the EXOMOL Database 76
4.1 Data maintenance 76

4.2 New line lists and new versions . . . 77
4.2.1 Versioning and naming . . . 77
4.2.2 Steps for updating a new line list for each isotopologue . . . 78
4.2.3 Update master file . . . 79
4.2.4 Create definition file . . . 79
4.3 New JSON format files . . . 81
4.4 Upload new ExoMol files . . . 81
4.4.1 Steps for uploading files to the ExoWeb server . . . 81
4.4.2 Steps for uploading files to the ExoMol website . . . 85
4.5 Dealing with UV processes . . . 86
4.6 Summary . . . 87

5 Preprocessing ExoMol Data 88
5.1 Processing the master file . . . 88
5.2 Processing the definition file . . . 89
5.3 Processing the states and transitions files . . . 89
5.4 Processing the partition function files . . . 91
5.5 Challenges for processing line lists . . . 91
5.6 Summary . . . 93

6 Post-processing ExoMol and HITRAN Data: PyExoCross 94
6.1 ExoCross overview . . . 94
6.2 Introduction . . . 95
6.3 Main functionality . . . 97
6.3.1 Partition functions . . . 97
6.3.2 Specific heats . . . 100
6.3.3 Cooling functions . . . 103
6.3.4 Radiative lifetimes . . . 105
6.3.5 Oscillator strengths . . . 107
6.3.6 Line intensities and stick spectra . . . 109
6.3.6.1 Absorption coefficient . . . 109

6.3.6.2 Emission coefficient . . . 110
6.3.6.3 Stick spectra . . . 110
6.3.6.4 Intensity thresholds . . . 111
6.3.7 Cross sections . . . 112
6.3.7.1 Grids of wavenumber and wavelength . . . 113
6.3.7.2 Absorption and emission cross sections . . . 114
6.3.7.3 A comparative analysis of program run times with PyExoCross, ExoCross, and RADIS . . . 115
6.4 Filters . . . 119
6.4.1 Uncertainty filters . . . 119
6.4.2 Intensity threshold filters . . . 120
6.4.3 Quantum number filters . . . 120
6.5 Plotting options . . . 120
6.6 Data format conversion . . . 121
6.6.1 Convert data format from ExoMol to HITRAN . . . 121
6.6.2 Convert data format from HITRAN to ExoMol . . . 124
6.6.2.1 Construct the ExoMol states file (.states) . . . 124
6.6.2.2 Construct the ExoMol transitions file (.trans) 125
6.6.2.3 Constructing an ExoMol air broadening file (.air) . . . 126
6.6.2.4 Constructing an ExoMol self broadening file (.self) . . . 127
6.7 Additional functionality . . . 128
6.8 Optimize by parallelization and vectorization . . . 128
6.9 Validation of the PyExoCross . . . 130
6.10 Calculation protocol . . . 130
6.11 Units . . . 131
6.12 Summary . . . 132

7 Line broadening and line profiles 134

7.1 Wing cut-off . . . 135

7.2 Broadening parameters . . . 135
7.3 Doppler and Gaussian profile . . . 136
7.4 Lorentzian profile . . . 137
7.4.1 Extracting broadening parameters from `.broad` . . . 138
7.4.2 Mixtures of broadeners . . . 138
7.4.3 Lifetime broadening . . . 139
7.5 Voigt profile . . . 141
7.5.1 SciPy Voigt profile . . . 142
7.5.2 SciPy wofz Voigt profile . . . 143
7.5.3 Humlíček Voigt profile . . . 144
7.5.4 Pseudo-Voigt profile . . . 145
7.5.4.1 Thompson pseudo-Voigt profile . . . 145
7.5.4.2 Kielkopf pseudo-Voigt profile . . . 146
7.5.4.3 Olivero pseudo-Voigt profile . . . 146
7.5.4.4 Liu-Lin pseudo-Voigt profile . . . 146
7.5.4.5 Rocco pseudo-Voigt profile . . . 147
7.6 Binned line profile . . . 147
7.6.1 Binned Doppler and Gaussian profile . . . 148
7.6.2 Binned Lorentzian profile . . . 148
7.6.3 Binned Voigt profile . . . 150
7.7 Line profile comparison . . . 151
7.7.1 A comparative analysis of the sampling and binned line profile methodologies . . . 151
7.7.2 A comparative analysis of line profile errors . . . 152
7.7.3 A comparative analysis of program run times with different line profiles . . . 155
7.8 Summary . . . 156

8 Simulating Non-LTE Spectra 157

8.1 Introduction . . . 157
8.2 Populations and densities . . . 158

8.2.1 LTE populations using the Boltzmann distributions 159
8.2.2 Non-LTE populations using two-temperature Treanor distributions 160
8.2.3 Non-LTE populations using custom vibronic densities . . . 160
8.3 LTE and non-LTE absorption and emission intensities 161
8.3.1 LTE and non-LTE absorption intensities 162
8.3.1.1 Non-LTE absorption intensities using two-temperature Treanor distribution 162
8.3.1.2 Non-LTE absorption intensities using custom vibrational (vibronic) density 162
8.3.1.3 Non-LTE absorption intensities using custom rovibrational (rovibronic) population 163
8.3.2 LTE and non-LTE emission intensities 163
8.3.2.1 Non-LTE emission intensities using two-temperature Treanor distribution 163
8.3.2.2 Non-LTE emssion intensities using custom vibrational (vibronic) density 164
8.3.2.3 Non-LTE emission intensities using custom rovibrational (rovibronic) population 164
8.4 Results and discussion 164
8.5 Summary 166

9 A High-Resolution ExoMol Database: ExoMolHR 172
9.1 Introduction 173
9.2 MARVEL database overview 174
9.3 ExoMolHR database overview 175
9.4 Methodology and algorithm 176
9.5 File structure and features 179
9.6 Use cases 184
9.7 Application programming interface (API) 185
9.8 Comparisons and discussion 185

9.9 Database summary and sources 190
9.10 Summary 190

10 Conclusion **195**
10.1 Summary of the thesis 195
10.2 Contributions and impact 196
10.3 Future work 197
10.3.1 Major databases update 197
10.3.2 Automatic maintenance of the ExoMol database 198
10.3.3 Continuous expansion of the ExoMol database with atomic data 199
10.3.4 Improve PyExoCross 199
10.3.5 Link to other data products 200

Appendices **201**

A API structure in JSON format **201**

B JSON format master file **203**

C JSON format definition file **205**

Bibliography **209**

List of Figures

2.1 Upper panel: Transiting phase curve of the hot Jupiter HD 189733 b observed at 4.5 μm with IRAC/Spitzer, showing atmospheric limb absorption during primary transit and day-side emission before secondary eclipse. Lower panel: Schematic of a transiting exoplanet, with the color bar indicating the day side (red) and night side (black). Figure is from Pluriel [1]. 38

2.2 The high-resolution cross-correlation spectroscopy (HRCCS) technique. Upper panel: The illumination of a non transiting planet throughout its orbit and highlights when its spectrum is red- or blueshifting. At high-resolution, the spectrum of planet is resolved into a dense forest of individual lines in a pattern that is unique to each molecule. These lines trace out the radial velocity curve of the planet (ΔRV$\sim$km/s), allowing it to be robustly disentangled from the essentially stationary stellar lines (ΔRV$\sim$m/s) and the contaminating spectral lines from telluric features of Earth. The maximum rate of change of velocity is reached around superior and inferior conjunction. At the quadrature points highlighted in the top panel, the spectral lines of the close in planet are maximally separated in velocity from the host star and tellurics, but the change in Doppler shift of the planet lines is small here making it difficult to disentangle them. The planet can be observed via its thermally emitted spectrum, or by the spectrum it reflects from the host star. Figure is from Birkby [2]. 39

3.1 Summary of the current ExoMol data structures [3]. New JSON format file types are highlighted in colored and will be introduced in Chapter 4. 65

4.1 Flowchart for updating line list files to the ExoMol website. 82

4.2 The steps for creating new line list files and updating existing line list files to the ExoMol website. 86

5.1 The relationship between the wavenumber range in filenames and the input parameters. 93

6.1 Partition functions of $^{1}H_{2}^{16}O$, $^{12}C^{16}O_2$, $^{32}S^{16}O_2$, $^{1}H^{12}C^{14}N$, $^{12}C^{16}O$ and $^{14}N^{16}O$ with temperatures from 0 K to 5000 K from the ExoMol database. 99

6.2 Specific heats of $^{1}H_{2}^{16}O$, $^{12}C^{16}O_2$, $^{32}S^{16}O_2$, $^{1}H^{12}C^{14}N$, $^{12}C^{16}O$ and $^{14}N^{16}O$ with temperatures from 200 K to high temperature or T_{max} from the ExoMol database. 102

6.3 The temperature-dependent cooling functions of $^{14}N^{16}O$, $^{27}Al^{16}O$, $^{27}Al^{1}H$ and $^{24}Mg^{1}H$ with temperatures from 0 K to 5000 K from the ExoMol database. 104

6.4 Upper panel: Lifetime of $^{14}N^{16}O$, $^{24}Mg^{1}H$ and $^{27}Al^{1}H$ using the ExoMol line lists. Lower panel: State-dependent lifetime in range of 10^{-4} s - 10^{-1} s of $^{27}Al^{1}H$ from the ExoMol database. The colors range from purple ($J = 0$) to yellow ($J = 80$). 106

6.5 Oscillator strength of $^{14}N^{16}O$, $^{24}Mg^{16}O$, and $^{27}Al^{1}H$ from the ExoMol line lists. 108

6.6 Stick spectra: Absorption intensities of $^{24}Mg^{1}H$ at $T = 10$, 300, 1000, 3000, and 5000 K simulated using the ExoMol line list of dataset XAB [4]. Upper panel: Intensities in wavenumber range $[0, 30\,000]$ cm^{-1}. Lower panel: Intensities zoom at $[19\,200, 19\,300]$ cm^{-1}. 112

6.7 Cross-section: Absorption spectrum of $^{24}Mg^{1}H$ at $T = 10$, 300, 1000, 3000, and 5000 K simulated using the ExoMol line list of the dataset XAB [4] broadened using a Gaussian line profile with its HWHM $\alpha_D = 3$ cm^{-1}. Upper panel: Cross-sections in wavenumber range $[0, 30\,000]$ cm^{-1}. Lower panel: Cross-sections zoom for $[9200, 9600]$ cm^{-1}. 116

6.8 Cross-section: Absorption spectrum of $^{1}H_2{}^{16}O$ at $T = 10$, 300, 1000, 3000, and 5000 K simulated using the ExoMol line list of the dataset POKAZATEL [5] broadened using a SciPy Voigt line profile with the bin size 0.1 cm^{-1} at $P = 1$ bar, HWHM mixed with H_2 and He broadeners and also applied uncertainty and threshold filters (0.01 cm^{-1} and 10^{-30} cm molecule^{-1}). Upper panel: Cross-sections in wavenumber range $[0, 26\,000]$ cm^{-1}. Lower panel: Cross-sections zoom for $[10\,300, 10\,330]$ cm^{-1}. 117

6.9 Cross-section: Absorption spectrum of $^{1}H_2{}^{16}O$, $^{12}C^{16}O_2$, and $^{32}S^{16}O_2$, at $T = 300$ and 3000 K simulated using the ExoMol line list of the dataset POKAZATEL [5], UCL-4000 [6], and ExoAmes [7] broadened using a SciPy Voigt line profile with bin size 0.1 cm^{-1} at $P = 1$ bar, HWHM mixed with H_2 and He broadeners and also applied uncertainty and threshold filters (0.01 cm^{-1} and 10^{-30} cm molecule^{-1}). Upper panel: Cross-sections in wavenumber range $[0, 15\,500]$ cm^{-1}. Lower panel: Cross-sections zoom for $[2300, 2500]$ cm^{-1}. 118

6.10 Absolute absorption line intensities of different electronic states $B'^2\Sigma^+$ - $X^2\Sigma^+$, $A^2\Pi$ - $X^2\Sigma^+$, and $X^2\Sigma^+$ - $X^2\Sigma^+$ bands of the ExoMol $^{24}Mg^{1}H$ simulated at $T = 300$ K. 121

6.11 Calculation protocol of the PyExoCross program. T and P refer to temperature and pressure separately. 132

7.1 Illustration of Gaussian, Lorentzian, Voigt profiles, and the convolution of Gaussian and Lorentzian profiles. 134

7.2 Photodissociation absorption spectrum of $^{16}O^{1}H$ at $T = 10$, 300, 1000, 3000, and 5000 K simulated using the EXOMOL line list of the dataset MYTHOS [8] broadened using a Voigt line profile with the bin size 0.1 cm^{-1} at $P = 1$ bar in wavenumber range $[0, 131\,900]$ cm^{-1}. 140

7.3 Comparison of the sampling and binned methods for different line profiles is shown for a single normalized transition with a wing cutoff of 100 cm^{-1} and a bin size of 0.125 cm^{-1} (8 points per cm^{-1}) at a temperature of $T = 300$ K and pressure $P = 1$ bar. Left panel: Doppler profile.Middle panel: Lorentzian profile. Right panel: Voigt profile. 152

7.4 Line profile errors $\varepsilon_{\text{area}}$ (Eq. (7.51)) and $\varepsilon_{\text{sampling}}$ (Eq. (7.52)) are analyzed for different profiles representing the same single normalized transition ($\tilde{\nu}_{if} = 4064.574702$ cm^{-1}), including Gaussian ($\alpha_{\text{G}} = 0.1$), Lorentzian (with γ_{L} estimated by $\gamma_{\text{ref}} = 0.07$ and $n_L = 0.5$), and Voigt profiles. These profiles are evaluated over a large wavenumber range of $[0, 8000]$ cm^{-1} at different temperatures T, pressures P, and bin sizes $\Delta\tilde{\nu}$. Top left: $\varepsilon_{\text{area}}$ for different temperatures T at $P = 1$ bar, with $N_{\text{points}} = 80\,001$ grid points (bin size $\Delta\tilde{\nu} = 0.1$ cm^{-1}) without a wing cutoff. Top right: $\varepsilon_{\text{area}}$ for different pressures P at $T = 300$ K, with $\Delta\tilde{\nu} = 0.1$ cm^{-1} without a wing cutoff. Bottom left: $\varepsilon_{\text{area}}$ for different numbers of grid points N_{points} (bin sizes $\Delta\tilde{\nu}$ cm^{-1}) at $T = 300$ K, $P = 1$ bar, without a wing cutoff. Bottom right: $\varepsilon_{\text{sampling}}$ for different numbers of grid points N_{points} (bin sizes $\Delta\tilde{\nu}$ cm^{-1}) at $T = 300$ K, $P = 1$ bar, with a wing cutoff of 100 cm^{-1}. 154

8.1 Comparison the equations of LTE and non-LTE absorption and emission intensities. 161

8.2 Absorption stick spectra at a pressure of 1 bar for CN [9] in LTE at $T = 1000$ K and non-LTE at equilibrium temperature ($T_{\text{vib}} = 1000$ K), with a rotational temperature of $T_{\text{rot}} = 296$ K for the non-LTE case in the $[0, 10\,000]$ cm^{-1} region (top display) and zoomed in $[3200, 4200]$ cm^{-1} region (bottom display). 166

8.3 Emission stick spectra at a pressure of 1 bar for CN [9] in LTE at $T = 1000$ K and non-LTE at equilibrium temperature ($T_{\text{vib}} = 1000$ K), with a rotational temperature of $T_{\text{rot}} = 296$ K for the non-LTE case in the $[0, 10\,000]$ cm^{-1} region (top display) and zoomed in $[3200, 4200]$ cm^{-1} region (bottom display). 167

8.4 Absorption cross sections at a pressure of 1 bar for CN [9] in LTE at $T = 1000$ K and non-LTE at equilibrium temperature ($T_{\text{vib}} = 1000$ K), with a rotational temperature of $T_{\text{rot}} = 296$ K for the non-LTE case in the $[0, 10\,000]$ cm^{-1} region (top display) and the relative differences (bottom display). 168

8.5 Absorption cross sections at a pressure of 1 bar for CN [9] in LTE at $T = 1000$ K and non-LTE at equilibrium temperature ($T_{\text{vib}} = 1000$ K), with a rotational temperature of $T_{\text{rot}} = 296$ K for the non-LTE case in the $[3000, 5000]$ cm^{-1} region (top display) and the relative differences (bottom display). 169

8.6 Emission cross sections at a pressure of 1 bar for CN [9] in LTE at $T = 1000$ K and non-LTE at equilibrium temperature ($T_{\text{vib}} = 1000$ K), with a rotational temperature of $T_{\text{rot}} = 296$ K for the non-LTE case in the $[0, 10\,000]$ cm^{-1} region (top display) and the relative differences (bottom display). 170

8.7 Emission cross sections at a pressure of 1 bar for CN [9] in LTE at $T = 1000$ K and non-LTE at equilibrium temperature ($T_{\text{vib}} = 1000$ K), with a rotational temperature of $T_{\text{rot}} = 296$ K for the non-LTE case in the $[3000, 5000]$ cm^{-1} region (top display) and the relative differences (bottom display). 171

9.1 Process flowchart for ExoMolHR database generation. 178

9.2 Four steps for getting the intensity file from the ExoMolHR website. 183

9.3 A comparison of the line intensities (cm/molecule) from the ExoMol, ExoMolHR (top), and HITRAN (bottom) databases at $T = 296$ K for $^{1}H_{2}{}^{16}O$, $^{15}N^{1}H_{3}$, $^{12}C^{1}H_{4}$, and $^{32}S^{16}O_{2}$. In the upper panels, the highlighted points represent the intensities provided by ExoMolHR database, while the darker background illustrates the full coverage offered by the POKAZATEL, CoYuTe-15, MM, and ExoAmes line lists for $^{1}H_{2}{}^{16}O$, $^{15}N^{1}H_{3}$, $^{12}C^{1}H_{4}$, and $^{32}S^{16}O_{2}$, respectively. 187

9.4 A comparison of the line intensities (cm/molecule) from the ExoMolHR (top) and ExoMol (bottom) databases at temperatures $T = 1000$ and 2000 K for $^{1}H_{2}{}^{16}O$, $^{48}Ti^{16}O$, and $^{51}V^{16}O$. 188

9.5 A comparison of the densities of wavenumbers (left) and line intensities (middle) from the ExoMol, ExoMolHR, and HITRAN databases at temperature $T = 296$ K for the molecules $^{1}H_{2}{}^{16}O$, $^{14}N^{1}H_{3}$, $^{12}C^{1}H_{4}$, and $^{32}S^{16}O_{2}$. The line lists from the ExoMol database provide full coverage, using the POKAZATEL, CoYuTe, MM, and ExoAmes line lists for $^{1}H_{2}{}^{16}O$, $^{14}N^{1}H_{3}$, $^{12}C^{1}H_{4}$, and $^{32}S^{16}O_{2}$, respectively. The right hand panels present 2D plots of the density distributions between wavenumbers and line intensities (cm/molecule) for the ExoMolHR database at $T = 296$ K. 189

List of Tables

2.1 Overview of the line lists for 9 molecules in the THEORETS database, classified as suitable for hot or cold conditions. 45

2.2 Overview of the line lists for 6 molecules in the NASA AMES database, classified as suitable for hot or cold conditions. 46

2.3 Overview of CASDA24 database with the list of molecules, their HITRAN ID, and corresponding database names. 50

2.4 Related molecules of CDSD NOSD, NDSD, and ASD databases . 55

2.5 Comparative overview of astrophysically important molecular databases (EXOMOL, THEORETS, NASA AMES, KURUCZ, GEISA, HITRAN, HITEMP, JPL, CDMS, CASDA24, VALD3, BASECOL, W@DIS, CDSD, NOSD, NDSD, ASD, and LEIDEN) used for astronomical studies. 57

3.1 The EXOMOL database releases comparison. 61

3.2 Summary of data provided by the EXOMOL Database. 65

3.3 Specification of EXOMOL file types. The file types highlighted in colours are mentioned in the PYEXOCROSS and EXOMOLHR projects. Their data structures are described in the following sections. All these file types are included in data maintenance. Note that EXOCROSS/PYEXOCROSS generates additional file types such as `.os` for oscillator strengths but these files are not distributed via the EXOMOL website or database. 66

3.4 Format of the EXOMOL master file, `exomol.all`; each entry starts on a new line. 67

3.5 Format of the EXOMOL definition file, <ISO-SLUG>__<DATASET>.def; each entry starts on a new line. 68

3.6 Format of the EXOMOL states file, <ISO-SLUG>__<DATASET>.states; each entry starts on a new line. 70

3.7 Extract from the states file of the $^{24}Mg^{1}H$ line list in the EXOMOL database, 24Mg-1H__XAB.states.bz2. 70

3.8 Format of the EXOMOL transitions file, <ISO-SLUG>__<DATASET>.trans; each entry starts on a new line. 71

3.9 Extract from the transitions file of the $^{24}Mg^{1}H$ line list in the EXOMOL database, 24Mg-1H__XAB.trans.bz2. 71

3.10 HITRAN line-by-line format in 2004 Edition [10] (160-character record). 75

3.11 Description of the parameters presented in the 160-character records in the line-by-line section of the HITRAN database [10]. . . 75

4.1 InChI and InChIKey for the isotopologues of molecule HCl. 80

4.2 Keywords used in the JSON format master file (exomol.all.json). 82

4.3 Keywords used in the JSON format definition file (<ISOTOPOLOGUE>__<DATASET>.def.json). 83

6.1 Specification of the .pf partition function file format. 99

6.2 Extract from the .pf partition function file generated from the $^{27}Al^{1}H$ AloHa [11] EXOMOL line lists. 99

6.3 Specification of the .cp specific heat file format. 102

6.4 Extract from the .cp specific heat file generated from the $^{27}Al^{1}H$ AloHa [11] EXOMOL line lists. 102

6.5 Specification of the .cf cooling function file format. 104

6.6 Extract from the .cf cooling function file generated from the $^{27}Al^{1}H$ AloHa [11] EXOMOL line lists. 104

6.7 Specification of the .os oscillator strength file format. 108

6.8 Extract from the `.os` oscillator strength file generated from the $^{14}N^{16}O$ XABC [12] ExoMol line lists. 108

6.9 Specification of the mandatory part of the `.stick` stick spectra file with the extra columns for the optional quantum numbers. 111

6.10 Extract from the `.stick` stick spectra file generated from the emission $^{12}C^{16}O_2$ UCL-4000 [6] ExoMol line list. 111

6.11 Specification of the `.cross` cross-section file format. 115

6.12 Extract from the `.cross` cross-section file generated from the $^{1}H_2{}^{16}O$ POKAZATEL [5] ExoMol line lists. 115

6.13 Specification of the file types related to the PyExoCross (contents in brackets are optional). 122

6.14 The uncertainty codes used by the HITRAN database [13]. 123

6.15 Conversion of the quantum number values between the ExoMol and HITRAN database. 123

6.16 Specification of the mandatory part of the `.broad` broadening file format. 127

6.17 Extract from the `.broad` broadening file of the H_2 broadening file `14N-1H3__H2.broad` [14]. 127

6.18 Default units used in the PyExoCross program. 133

7.1 Example of program run times on 32 cores, in s, for computing cross sections using different line profiles at a temperature of $T = 300$ K, pressures $P = 1$ bar, and a wing cutoff of 25 cm^{-1} for the ExoMol H_2O POKAZATEL line list [5], over the wavenumber range $[20000, 22000]$ cm^{-1}. The left panel compares the run times for different numbers of transitions $(10^4, 10^6, 10^8)$ with a bin size of 0.1 cm^{-1} (20001 grid points) without any filters. The middle panel compares the run times for for different numbers of transitions $(10^4, 10^6, 10^8)$ with a bin size of 0.1 cm^{-1} (20001 grid points) using uncertainty and threshold filters. The right four columns compare the run times for different bin sizes $(0.1, 0.01, 0.005, 0.001$ $cm^{-1})$ and corresponding numbers of grid points $(20001, 200001, 400001, 2000001)$ using 10^8 transitions with uncertainty and threshold filters. 155

8.1 Summary of the three cases considered for spectral simulation. . . . 159

8.2 Specification of the mandatory part of the `.nlte` non-LTE stick spectra file with the extra columns for the optional quantum numbers. 165

8.3 Extract from the `.nlte` non-LTE stick spectra file generated from the emission $^{12}C^{14}N$ Trihybrid [9] ExoMol line list. 165

9.1 Specification of the output `.csv` ExoMolHR file format. 180

9.2 Extract from the ExoMolHR file of $^1H_2{}^{16}O$ at temperature $T =$ 296.0 K (`yyyymmddhhmmss__1H2-16O__296K.csv`). 180

9.3 Quantum number labels for each isotopologue in the ExoMolHR database. 181

9.4 The formats and descriptions corresponding to quantum number labels. 182

9.5 Summary of lines with low uncertainties and high resolutions extracted from the ExoMol database and used to form the ExoMolHR database. 192

List of Listings

1 Python code for calculating the physical and astronomical constants. 133

2 A sample of the Python code for extracting broadeners from the `.broad` broadening file. PyExoCross program reads the `.broad` file and stores the data in a `DataFrame` named `broad_df`. Simultaneously, the program generates another `DataFrame`, `states_trans_df`, by merging the states file (`.states`) with the transitions file (`.trans`), using the upper and lower state IDs. The columns `Jpp` and `J"` represent the lower J quantum numbers of the transitions, while `id_broad` indicates the corresponding index in the `broad_df`. The parameters `gamma_L` and `n_air` in the code correspond to γ_{ref} (the reference HWHM) and n_{air} (the temperature exponent for air-broadening). 139

3 A sample Python code snippet demonstrating the use of the `SciPy` package to calculate the Voigt profile [15]. $\mathrm{d}\nu = \tilde{\nu} - \tilde{\nu}_{fi}$ represents the difference from the line centre ($\tilde{\nu}_{fi}$). The parameter `sigma` is the standard deviation σ_{G} of the Gaussian profile, while `gamma` is the Lorentzian profile half-width for pressure broadening γ_{L}. 142

4 A method for calculating the Voigt profile through the convolution of Gaussian and Lorentzian profiles, using the `SciPy` Python library, specifically the `scipy.special.voigt_function`. . . 142

5 A sample of the Python code used to compute the Voigt profile [15, 16] with the real part of the Faddeeva function. 144

A.1 The structure of API in JSON format used for search queries using the `HTTP GET` request. 201

B.1 JSON format master file `exomol.all.json`. 203

C.1 JSON format definition (`.def.json`) file of $^{1}H_{2}{}^{16}O$ POKAZATEL dataset. 205

Chapter 1

Introduction

1.1 Background and motivation

Molecular spectroscopy plays a crucial role in numerous scientific and technological domains, including atmospheric science, astrophysics, and remote sensing. The accurate modeling and interpretation of molecular spectra require comprehensive spectroscopic databases, which provide detailed molecular line lists and parameters for a wide range of species. These databases facilitate the study of planetary atmospheres, stellar and exoplanetary environments, and laboratory spectroscopy. Among the most prominent molecular databases are HITRAN (High-Resolution Transmission Molecular Absorption Database) [10, 13] and ExoMol [17], which contain vast amounts of spectral data essential for high-resolution spectroscopic applications.

With the increasing demand for high-accuracy and high-resolution spectroscopic data, continuous efforts are required to update, expand, and refine existing molecular databases. The ExoMol project is dedicated to providing comprehensive molecular line lists specifically for modeling exoplanetary and other hot atmospheres [18]. Likewise, the HITRAN database has been a widely used database of atmospheric and astrophysical spectroscopy for decades, with periodic updates ensuring its relevance for contemporary research and applications [10, 13].

Despite the remarkable progress made in the field of molecular spectroscopy, there remain significant challenges in processing and managing large spectroscopic

datasets, improving the accuracy and coverage of molecular parameters, and ensuring seamless compatibility between different databases. One of the major issues in the current state of spectroscopic databases is the need for continuous updates and improvements in line list completeness, transition intensities, and data accessibility for various astrophysical and atmospheric applications. Given the importance of accurate spectroscopic data for exoplanet characterization, atmospheric modeling, and high-resolution spectral analyses, there is a pressing need for a systematic approach to improving these databases.

The motivation for this research stems from the need to bridge existing gaps by developing efficient methodologies for data processing, updating, and expanding the ExoMol database, as well as introducing new computational tools for post-processing and analyzing molecular spectroscopic data. This work aims to contribute to the continuous advancement of spectroscopic databases by ensuring they are not only comprehensive but also more accurate, accessible, and optimized for practical applications in planetary and astrophysical research with the functionality needed by the users community.

By addressing key computational and methodological challenges in spectroscopic data handling, this thesis seeks to improve the overall quality and usability of molecular spectroscopic databases. Specifically, it focuses on the processing, post-processing, and expansion of such databases, with particular emphasis on ExoMol and HITRAN. The thesis presents novel methods and computational tools designed to enhance the accuracy, accessibility, and applicability of these databases across various scientific domains.

1.2 Objectives of the thesis

The primary objective of this thesis is to introduce and review key molecular spectroscopic databases, with a particular emphasis on ExoMol and HITRAN. In order to facilitate the effective utilization of these databases, this research aims to develop and apply advanced data processing methodologies that enable the efficient handling of large spectroscopic datasets. By doing so, it ensures that researchers can

extract and apply meaningful and high-precision data from these vast repositories.

Another important goal of this work is to update and expand the EXOMOL database through the generation of new high-resolution molecular line lists. This involves not only the incorporation of additional molecular species but also the improvement of existing datasets to enhance their accuracy and applicability. These updates will support a wide range of scientific investigations, from planetary atmosphere studies to astrophysical spectral analysis.

In addition, this thesis seeks to implement and refine PYEXOCROSS, a Python-based tool designed for the post-processing of spectroscopic data. This tool provides functionalities such as the calculation of partition functions, specific heats, cooling functions, radiative lifetimes, stick spectra, and cross-sections, which are essential for accurate spectral modeling. The development of PYEXOCROSS is crucial for enabling researchers to efficiently interpret and manipulate spectral data in a manner that aligns with their specific research needs.

An integral part of this research is the analysis and refinement of line profile models, including Gaussian, Doppler, Lorentzian, and Voigt profiles. These models play a vital role in the accurate simulation of spectral features and, consequently, their improvement is essential for achieving high-fidelity spectral analysis. By carefully analyzing and optimizing these profiles, this thesis aims to enhance the accuracy of spectroscopic simulations used in astrophysical and atmospheric studies.

Furthermore, the thesis explores the simulation of non-local thermodynamic equilibrium (non-LTE) spectra, which are critical for understanding the behavior of molecular systems in environments that deviate from thermodynamic equilibrium. The study of non-LTE conditions allows for a more precise characterization of exoplanetary atmospheres, stellar envelopes, and other astrophysical phenomena where standard local thermodynamic equilibrium (LTE) assumptions are not applicable.

Finally, this research introduces EXOMOLHR, a high-resolution extension of the EXOMOL database. The motivation behind EXOMOLHR is to provide a dataset that is tailored for applications requiring extremely fine spectral details, such as ex-

oplanet atmospheric modeling and high-resolution laboratory spectroscopy. The development of this new database ensures that researchers have access to spectroscopic data of the highest possible resolution.

1.3 Structure of the thesis

This thesis is structured in a way that ensures a comprehensive understanding of the development, processing, and application of spectroscopic databases. Chapter 2 provides an overview of various molecular spectroscopic databases, including ExoMol, HITRAN, HITEMP, and other databases accessed via the Virtual Atomic and Molecular Data Centre (VAMDC) portal [19, 20], by discussing their structure, content, and specific use cases in the field of molecular spectroscopy. Chapter 3 is dedicated to a more detailed exploration of the ExoMol and HITRAN databases, analyzing their respective file structures, data formats, and the application programming interface (API) functionalities to demonstrate their significance and usability in scientific research.

Following the introduction to these databases, Chapter 4 describes the updates and expansions made to the ExoMol database. It details the generation of new molecular line lists, modifications in formatting master and definition files in JavaScript Object Notation (JSON) [21] format, and the procedures involved in uploading these files to ExoMol repositories to enhance their accessibility and scientific utility. Chapter 5 delves into the methodologies employed for preprocessing spectroscopic data. This includes techniques for handling master files, state and transition files, and partition function data, ensuring that the vast amounts of information contained within these databases are organized and processed effectively.

Chapter 6 introduces PyExoCross, a Python-based post-processing tool designed for working with spectroscopic data. This chapter explains its core functionalities which is computing molecular properties, such as partition functions, specific heats, cooling functions, radiative lifetimes, oscillator strengths, stick spectra, and cross-sections, highlighting its role in refining and analyzing spectral data. PyExoCross also provides uncertainty, intensity threshold, and quantum num-

ber filters. The data format of ExoMol and HITRAN can be inter converted by PyExoCross. In addition, PyExoCross offers extra tools to automate batch downloads of line list files from the ExoMol database. Chapter 7 focuses on the theoretical aspects of line broadening and the mathematical models used for line profiles, including Gaussian, Doppler, Lorentzian, and Voigt line profiles, which are essential for accurate spectral simulations when calculating the cross sections. Chapter 8 investigates the simulation of non-LTE spectra, emphasizing the importance of non-LTE effects in astrophysical and atmospheric environments. It presents methodologies for computing non-LTE population distributions and their impact on spectral line intensities.

Chapter 9 introduces ExoMolHR, an empirical high-resolution extension of the ExoMol database, which is designed for high-temperature molecular line lists. This chapter describes its structure and methodologies. The ExoMolHR database is designed for improved spectroscopic accuracy and applicability.

Finally, Chapter 10 presents a summary of the key findings and contributions of this research, along with a discussion of potential future directions for the improvement and expansion of molecular spectroscopic databases. This concluding section outlines prospective updates, refinements in data processing methodologies, and the broader implications of this work for the scientific community.

1.4 Significance and contributions

This research contributes to the field of molecular spectroscopy by enhancing the ExoMol database with updated and expanded molecular line lists. It also develops advanced data processing and post-processing tools to efficiently handle large spectroscopic datasets while improving the accuracy of spectroscopic simulations through refined line profile modeling. Furthermore, it enables non-LTE spectral simulations for astrophysical and atmospheric applications and introduces ExoMolHR, a new high-resolution spectroscopic database tailored for exoplanetary and astrophysical studies. By achieving these objectives, this thesis aims to support the pursuit of more accurate and reliable spectroscopic data for planetary,

stellar, and atmospheric research and much more, see Section 3.2.1.

1.5 Data availability

The programs, spectroscopic models and line lists mentioned in this thesis are available online. PyExoCross is stet open-source program for variational calculations of spectra of both atoms and molecules, which is available at https://github.com/ExoMol/PyExoCross. PyExoCross manual (https://pyexocross.readthedocs.io/) provides the supplementary data, provides comprehensive working instructions and use case examples. The high temperature molecular line lists [22] can be freely accessible through the ExoMol database website at https://www.exomol.com/. The high-resolution molecular database, ExoMolHR, is available at https://www.exomol.com/exomolhr/. The high-resolution transmission molecular absorption database HITRAN, which is a comprehensive collection of spectroscopic parameters [10, 13], is available at https://hitran.org.

Chapter 2

Spectra of Exoplanets

Recent observational campaigns using space-based telescopes like the Hubble Space Telescope (HST) and the James Webb Space Telescope (JWST), as well as ground-based instruments, have significantly advanced our ability to characterize exoplanetary atmospheres [23]. Future missions, such as the Ariel space telescope [24], will further advance transit spectroscopy, providing unprecedented details into the chemical and thermal structures of a wide variety of exoplanets. This chapter provides an overview of spectroscopic methods and widely used astrophysically important molecular databases relevant to these research efforts.

2.1 Resolving power

Resolving power (R) is a measure that defines the ability to spectroscopically distinguish closely spaced spectral features. It is mathematically defined as:

$$R = \frac{\lambda}{\Delta\lambda} \tag{2.1}$$

where λ is the observed wavelength and $\Delta\lambda$ is the smallest wavelength difference that can be resolved. It is crucial for extracting detailed molecular and atmospheric information from spectroscopic observations by separating individual molecular and atomic lines, minimizing contamination from overlapping features, and accurately measuring line broadening effects related to temperature, pressure, and turbulence [25].

A higher resolving power means finer spectral details can be distinguished,

which is essential for studying complex astrophysical and molecular processes. High resolving power is essential in exoplanet atmosphere characterization, stellar spectroscopy, and interstellar medium studies, where it enables the detection of molecular species, precise velocity measurements through Doppler shifts, and the refinement of atmospheric models [25].

In exoplanet research, high-resolution spectroscopy (HRS) plays a key role in transmission and emission spectroscopy, allowing planetary spectral features to be disentangled from stellar and telluric contamination while resolving molecular line profiles that encode information about atmospheric temperature, pressure, and dynamics. Additionally, resolving power enhances sensitivity to weak spectral lines, improves telluric correction, and supports the development of advanced retrieval techniques. Recent advancements in HRS, as discussed by Kuhn et al. [25], have significantly improved spectral retrieval techniques, highlighting the importance of high resolving power in characterizing exoplanetary atmospheres with unprecedented precision. Achieving high resolving power is critical for advancing astrophysical and planetary research, as it enables precise spectroscopic measurements that deepen our understanding of atmospheric processes, planetary evolution, and potential biosignatures in exoplanetary atmospheres.

2.2 Transit spectroscopy

Transit spectroscopy is a powerful observational technique that provides critical insights into the atmospheric composition of exoplanets. It requires a forward atmospheric model to interpret the results. When an exoplanet passes in front of its host star, a small fraction of the stellar light is filtered through the planetary atmosphere, allowing for the detection of specific absorption features associated with molecular and atomic species [26, 27]. By comparing the stellar spectrum during transit to the out-of-transit spectrum, it is possible to infer the presence of various atmospheric constituents, including water vapor, methane, carbon dioxide, and other gases relevant to planetary climate and habitability.

The observation of the phase curve of an exoplanet (i.e., the time-dependent

change in the brightness of a planet as seen from Earth during one orbital period) is the most straightforward way to probe the planet's longitudinal structure [1]. The brightness of the planet is determined by the combined emitted and reflected light in the observational wavelength. Giant exoplanets are compelling targets for transmission spectroscopy. Their atmospheres can be studied by analyzing starlight filtered through them during the primary transit and the thermal emission during the secondary transit [1]. Phase curve observations provide insights into the atmospheric properties of exoplanets, such as thermal emission from the day side and absorption features from the limb during transit.

In the study of exoplanetary atmospheres, secondary eclipse observations offer important insights into thermal emission from the planet. During a secondary eclipse, when the planet passes behind its host star, direct measurement of the emitted radiation is possible, free from the interference of starlight [28]. This observation allows for the determination of temperature and provides complementary information to primary transit measurements, which mainly focus on the composition of the planet [29]. Analyzing thermal emission during secondary eclipses helps to develop a more complete understanding of exoplanetary atmospheric properties, including thermal structure and the presence of molecular species.

Observations of exoplanetary transmission and emission spectra, recorded during primary transit and secondary eclipse, complement one another. The transmission spectra are primarily sensitive to the planet's composition, with minimal dependence on the thermal profile, whereas the emission spectra provide insights into both the planetary composition and thermal characteristics [30]. Figure 2.1 illustrates the phase curve of a transiting planet (HD 189733 b) observed at 4.5 μm with IRAC/Spitzer, highlighting the primary transit and secondary eclipse [1]. The red and black regions indicate the day and night sides of the planet, respectively, emphasizing which atmospheric regions are observed throughout the phase curve.

One of the main challenges in transit spectroscopy is achieving sufficient spectral resolution and signal-to-noise ratio to accurately discern atmospheric features. Advanced statistical techniques, such as retrieval algorithms and machine learning

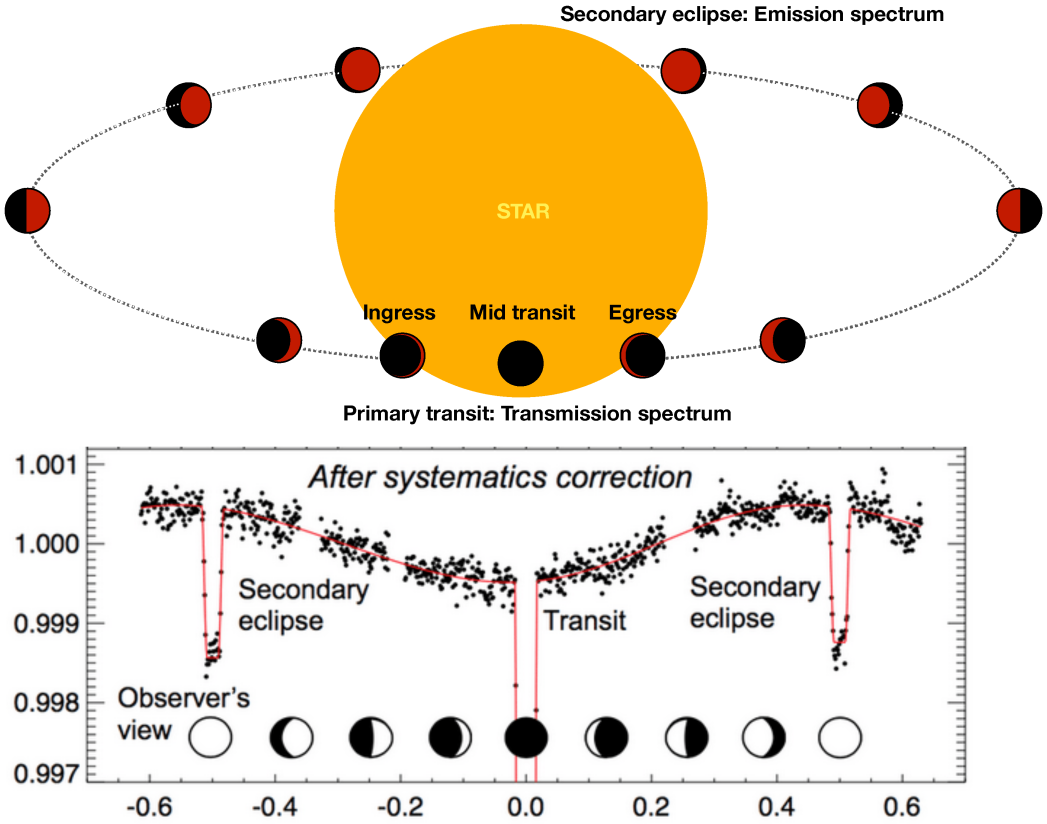


Figure 2.1: Upper panel: Transiting phase curve of the hot Jupiter HD 189733 b observed at 4.5 μm with IRAC/Spitzer, showing atmospheric limb absorption during primary transit and day-side emission before secondary eclipse. Lower panel: Schematic of a transiting exoplanet, with the color bar indicating the day side (red) and night side (black). Figure is from Pluriel [1].

approaches, are increasingly being employed to extract robust constraints on atmospheric parameters.

2.3 High-resolution cross-correlation spectroscopy

High-Resolution Spectroscopy (HRS) plays a crucial role in disentangling planetary spectral features from stellar and telluric contamination. The HRS technique allows for the extraction of a planet's radial velocity curve by resolving its spectrum into a dense forest of molecular lines, enabling the separation of planetary signals from stationary stellar and telluric lines, see Figure 2.2 [2].

High-Resolution Cross-Correlation Spectroscopy (HRCCS) is an important method for probing the composition of exoplanetary atmospheres. High accuracy is required for high-resolution cross-correlation spectroscopy given the high-

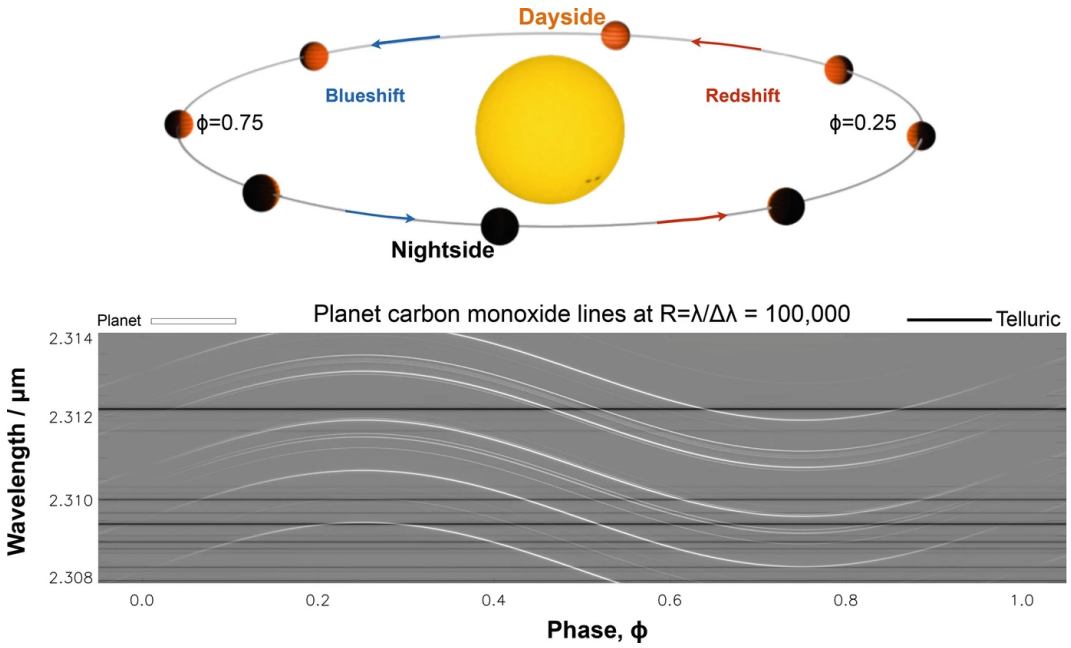


Figure 2.2: The high-resolution cross-correlation spectroscopy (HRCCS) technique. Upper panel: The illumination of a non transiting planet throughout its orbit and highlights when its spectrum is red- or blueshifting. At high-resolution, the spectrum of planet is resolved into a dense forest of individual lines in a pattern that is unique to each molecule. These lines trace out the radial velocity curve of the planet (ΔRV$\sim$km/s), allowing it to be robustly disentangled from the essentially stationary stellar lines (ΔRV$\sim$m/s) and the contaminating spectral lines from telluric features of Earth. The maximum rate of change of velocity is reached around superior and inferior conjunction. At the quadrature points highlighted in the top panel, the spectral lines of the close in planet are maximally separated in velocity from the host star and tellurics, but the change in Doppler shift of the planet lines is small here making it difficult to disentangle them. The planet can be observed via its thermally emitted spectrum, or by the spectrum it reflects from the host star. Figure is from Birkby [2].

resolution observations [31]. In my research, high resolving power is defined as $R > 100\,000$ while high accuracy corresponds to uncertainties of ≤ 0.01 cm^{-1}. By analyzing the minute changes in spectral lines caused by the motion of an exoplanet relative to its host star, this technique allows for the direct detection of molecular species and the measurement of wind patterns within exoplanetary atmospheres [32, 33].

In exoplanet research, Doppler shifts observed in both stars and their orbiting planets are crucial for detecting and characterizing exoplanetary systems. The Doppler shift of a star, measured through the radial velocity method, is a well-

established technique for detecting exoplanets. As a planet orbits its host star, the gravitational influence induces periodic motion in the star, causing its spectral lines to shift due to the Doppler effect [34]. This method was pivotal in the discovery of 51 Pegasi b [34], the first confirmed exoplanet around a main-sequence star, through high-resolution cross-correlation spectroscopy [35]. While more challenging to observe, detecting the Doppler shift in the spectral lines of a planet itself can provide direct measurements of its orbital velocity and atmospheric composition [32]. High-resolution spectroscopy enables the identification of specific molecular species and wind patterns in exoplanetary atmospheres. For example, observations of HD 209458b [32] using this technique successfully detected carbon monoxide (CO) in the atmosphere and constrained the velocity of the planet.

Cross-correlation spectroscopy method enhances signal detection by matching observed spectra against theoretical molecular templates. This technique has been particularly effective in identifying key atmospheric species, such as water vapor, carbon monoxide, and titanium oxide, even in planets with high-altitude clouds that obscure transmission spectra [36, 37]. By combining high-resolution spectroscopy with Doppler shift measurements, researchers can map atmospheric circulation patterns, detect equatorial jet streams, and study the effects of planetary rotation and weather phenomena.

The orbital velocity of a planet is governed by Kepler's laws, and the farther a planet is from its star, the slower it moves. Specifically, for a planet in orbit, the velocity (v) is related to the semi-major axis (a) of the orbit by the equation:

$$v \propto \frac{1}{\sqrt{a}} \tag{2.2}$$

In the context of high-resolution cross-correlation, the speed at which a planet moves affects the amplitude of the radial velocity (RV) signal. The Doppler shift is the change in the observed wavelength of light from a moving object, and for planets, this shift corresponds to their radial velocity (motion toward or away from the observer) [38]. For planets closer to their stars, the larger orbital velocities result in a greater Doppler shift, making it easier to detect and measure this motion

using high-resolution spectroscopy. On the other hand, planets farther from their stars, with slower velocities, will produce smaller Doppler shifts, making it harder to give strong molecular detections with the same precision. Therefore, for high-resolution cross-correlation observations, a planet cannot be too far from its star, as its slower orbital velocity would result in a smaller Doppler shift, making it challenging to disentangle the planetary motion from the noise, especially when dealing with high-resolution spectra where tiny shifts are crucial [39].

Accuracy is a crucial factor in high-resolution cross-correlation spectroscopy, requiring stable and well-calibrated instruments. The next generation of spectrographs, such as ESPRESSO on the Very Large Telescope (VLT) and future instruments on the Extremely Large Telescope (ELT), will push the boundaries of this technique, enabling more precise measurements of exoplanetary atmospheres [40]. Additionally, advancements in data reduction and machine learning algorithms will further improve the accuracy and sensitivity of molecular detections [41]. A phase mask of spectral lines is used to perform cross-correlation, further enhancing the detection of exoplanetary signals in high-resolution spectroscopy.

In high-resolution cross-correlation spectroscopy (HRCCS), the detection of molecular species in exoplanetary atmospheres relies on an accurate molecular line list to serve as a reference template. The observed spectrum, which contains the combined light from the star, planet, and atmosphere of Earth, is cross-correlated with this line list to extract the faint planetary signal. The accuracy of the line list is crucial, as discrepancies in line positions, strengths, or broadening can significantly impact the reliability of detections [18]. Therefore, high-resolution and high-precision molecular data are essential for maximizing the sensitivity of HRCCS and accurately identifying atmospheric constituents. The ExoMolHR (see Chapter 9) line lists provide high-resolution molecular spectroscopic data essential for this process, offering precise theoretical and experimental line positions for a wide range of molecules relevant to exoplanetary atmospheres.

2.4 Direct imaging spectroscopy

Direct imaging is a cutting-edge method for studying exoplanetary spectra that involves capturing the light directly emitted or reflected by an exoplanet, rather than relying on indirect measurements through transit or radial velocity techniques [42]. This approach is particularly valuable for characterizing young, self-luminous exoplanets and those orbiting at large angular separations from their host stars [43].

The primary challenge in direct imaging spectroscopy is the overwhelming brightness of the host star, which can obscure the much fainter planetary signal. To overcome this, sophisticated high-contrast imaging techniques, such as adaptive optics, coronagraphy, and starshade technology, are employed to suppress stellar light and enhance the visibility of planetary spectra [44].

The combination of direct imaging spectroscopy with complementary methods, such as transit and Doppler shift spectroscopy, offers a comprehensive approach to understanding the diverse atmospheric properties of exoplanets. As observational techniques and instrumentation continue to improve, direct imaging will play an increasingly vital role in expanding our knowledge of planetary systems and enhancing our understanding of planet formation and evolution.

2.5 Data quality

High accurate molecular line lists play an important role in scientific research fields, including atmospheric modeling, astrophysical observations, and remote sensing. The quality of molecular line lists directly influences the reliability of simulations and data analyses. The sensitivity of the cross-correlation method is highly dependent on the precision of the line positions in the template spectrum [45]. If the wavelength positions of these lines are inaccurate, the cross-correlation process may fail to add their contributions constructively, potentially undermining the results [45]. Uncertainties in line positions, intensities, and other spectroscopic parameters can lead to significant errors in spectral interpretation [46]. Tennyson and Yurchenko [47] emphasize that the accuracy of line lists affects the detection and characterization of molecules in exoplanetary atmospheres. They discuss ad-

vancements in improving line list accuracy by incorporating empirical energy levels derived through the MARVEL(Measured Active Rotational-Vibrational Energy Levels) database. This approach enhances the reliability of spectral models used in exoplanet research.

2.6 Astrophysically important molecular databases

Several databases provide extensive spectroscopic data essential for astrophysical research, particularly in the study of exoplanetary and stellar atmospheres. This chapter provides a comparative overview of several prominent molecular databases (see Table 2.5), including ExoMol [3, 17, 48], TheoReTS [49], NASA Ames [50], Kurucz [51], and GEISA [52], as well as other databases HITRAN [53], HITEMP [54, 55], JPL [56], CDMS [57, 58], CaSDa24 [59, 60], VALD3 [61], BASECOL [62], W@DIS information system [63–65], and CDSD [66–68] which are integrated into the VAMDC portal [19, 20], and Leiden photodissociation and photoionization cross-section database [69]. Additionally, three more high-temperature databanks, NOSD [70], NDSD [71], and ASD [72], are related to the CDSD database. In Table 2.5, primary data refers to information directly generated by this database, while secondary data consists of information sourced from external databases.

2.6.1 VAMDC portal

The Virtual Atomic and Molecular Data Centre (VAMDC) [19, 20] is a European Union-funded initiative that integrates atomic and molecular data from various sources into a secure, well-documented, and interoperable e-science platform. The VAMDC portal is freely accessible at http://www.vamdc.eu. It provides standardized access to existing databases, facilitating research in fields such as astrophysics, atmospheric science, and plasma physics. By integrating multiple databases, the VAMDC portal enables efficient retrieval of spectroscopic and collisional data, supporting research and computational modeling. Its open-access framework enhances data accessibility, making it a vital resource for the global scientific community.

2.6.2 ExoMol database

The ExoMol database [3, 17, 48] specializes in molecular line lists for exoplanetary and other high-temperature atmospheres. Its data originate from theoretical computations using advanced spectroscopic models, often refined by fitting to experimental measurements.

The ExoMol database 2024 release contains line lists for 91 molecules and 225 isotopologues, with data applicable to temperatures up to 5000 K or higher for certain species. The ExoMol database includes a diverse range of molecular species categorized into metal and other hydrides, metal and other oxides, ions, larger molecules, other diatomics, and triatomic molecules. This extensive coverage ensures that ExoMol provides high-accuracy line lists for a broad spectrum of molecular species relevant to exoplanetary and high-temperature astrophysical environments. The database covers the infrared region but extends to a broader spectral range, including the submillimeter domain. The ExoMol database provides extensive resources for modeling and simulating the spectra of hot astronomical objects. The data can be obtained from its website https://www.exomol.com/.

Additionally, the MoLLIST project [73] has generated line lists, primarily for diatomic species, which are accessible through the ExoMol website. These line lists are predominantly constructed by integrating experimental wavelength measurements with dipole moments computed using advanced electronic structure methods, similar to those employed by ExoMol.

The ExoMol database is exceptionally large, containing extensive spectroscopic data for molecular transitions. Due to its sheer size and high-resolution datasets, it is not included in the VAMDC portal, which primarily hosts more compact spectroscopic databases optimized for efficient querying and retrieval. Instead, ExoMol is managed independently to accommodate its extensive data volume and specialized computational requirements.

2.6.3 TheoReTS database

The Theoretical Reims-Tomsk Spectral (TheoReTS) database [49] computes polyatomic molecules, with its data from ab initio calculations of potential energy

and dipole moment surfaces. These surfaces, which describe the molecular interactions and electronic behavior, serve as essential inputs for solving the Schrödinger equation, enabling the accurate determination of energy levels, transition frequencies, and spectral intensities. The computational approach ensures a high degree of accuracy, making the database a reliable source for theoretical spectral data. It uses variational methods to determine vibration-rotation energy levels and transition probabilities.

Table 2.1: Overview of the line lists for 9 molecules in the THEORETS database, classified as suitable for hot or cold conditions.

Hot or Cold \ Molecules	C_2H_4	CF_4	CH_3F	CH_4	GeH_4	NF_3	PH_3	SF_6	SiH_4
Hot	✓	✓		✓					
Cold			✓		✓	✓	✓	✓	✓

The resulting spectral information is systematically organized and integrated into the database, providing comprehensive coverage of various molecular species. The database includes spectroscopic data for 21 isotopologues from 9 molecules: C_2H_4, CF_4, CH_3F, CH_4, GeH_4, NF_3, PH_3, SF_6, and SiH_4 (see Table 2.1). THEORETS database covers the infrared spectral region, with a particular emphasis on wavelengths ranging from 4.0 μm to 50.0 μm. Additionally, the database provides tools for spectral simulations, including absorption coefficients and cross-sections, incorporating various line shapes and apparatus functions. More information and access to the data can be found on its official website https://theorets.tsu.ru.

2.6.4 NASA AMES database

The NASA AMES database [50] focuses on Ames molecular spectroscopic data for astrophysical and atmospheric studies. This database is constructed using a combination of high-accuracy ab initio quantum mechanical calculations and empirical refinements to provide precise spectroscopic data for astrophysical and atmospheric studies. Rovibrational infrared line lists are generated using exact quantum mechanical methods within the Born-Oppenheimer approximation. High-level ab initio calculations, primarily employing the coupled-cluster method with single, double, and perturbative triple excitations CCSD(T), are performed in conjunction

with augmented correlation-consistent polarized valence quadruple-zeta (aug-cc-pVTZ) basis sets. These computations yield highly accurate potential energy surfaces (PES) and dipole moment surfaces (DMS), which are then used to derive line positions and intensities. To enhance the reliability of theoretical predictions, the PES and DMS are refined by incorporating high-resolution experimental spectroscopic data.

The NASA AMES database provides rovibrational IR line lists, energy levels, potential energy surfaces, and dipole moment surfaces for 6 molecules: CO_2, CS_2, N_2O, NH_3, OCS, and SO_2 (see Table 2.2). These datasets are generated with stringent intensity thresholds to ensure their relevance for high-resolution spectral analyses. Through this rigorous computational and empirical approach, the NASA AMES database provides accurate and comprehensive spectroscopic data essential for interpreting molecular spectra in astrophysical and atmospheric environments. The data are accessible through their website https://huang.seti.org/.

Table 2.2: Overview of the line lists for 6 molecules in the NASA AMES database, classified as suitable for hot or cold conditions.

Hot or Cold \ Molecules	CO_2	CS_2	N_2O	NH_3	OCS	SO_2
Hot	✓		✓		✓	
Cold		✓		✓		✓

2.6.5 KURUCZ database

The KURUCZ database [51] provides extensive atomic and molecular line lists, primarily for studies of stellar atmospheres. Its molecular data originate from semi-empirical calculations that combine experimental measurements with theoretical models. While the database includes line lists for various diatomic and polyatomic molecules, these molecular datasets are generally considered less precise compared to the other specialized databases discussed in this chapter.

The KURUCZ database focuses on atomic spectral data in the ultraviolet (UV) and visible regions, covering wavelengths from 0.25 μm to 1.0 μm. In terms of molecular coverage, the database lists 48 molecules; however, only 12 of these

contain substantial spectroscopic data, including AlH, AlO, C_2, CaH, CaO, CH, CN, CO, CrH, FeH, H_2, H_3^+, H_2O, MgH, MgO, NaH, OH, OH^+, SiH, TiH, TiO, and VO. The database offers a vast compilation of atomic and some molecular data relevant to stellar spectroscopy. The KURUCZ database can be freely accessed from http://kurucz.harvard.edu.

2.6.6 GEISA database

Established in 1974 by the Atmospheric Radiation Analysis group at the Laboratoire de Météorologie Dynamique (LMD) in France, Gestion et Etude des Informations Spectroscopiques Atmosphériques: Management and Study of Atmospheric Spectroscopic Information (GEISA) database [52] is designed to support accurate and efficient atmospheric radiative transfer modeling. It is the reference database for several current or planned Thermal and Short-Wave InfraRed (TIR and SWIR) space missions, including Infrared Atmospheric Sounding Interferometer (IASI), IASI New Generation (IASI-NG), Carbon Dioxide Monitoring Mission (MicroCarb), and MEthane Remote sensing LIdar missioN (Merlin). The 2020 edition of the GEISA database is freely available at https://geisa.aeris-data.fr, which involves 58 molecules (145 isotopic species), totaling 6 746 987 entries, and covers a spectral range from 10^{-6} to 35 877 cm^{-1}. This extensive compilation is accessible through the AERIS data and services center, offering powerful graphical tools and convenient data searching, filtering, and plotting capabilities.

The GEISA compiles, evaluates, and curates spectroscopic data from both experimental measurements and theoretical calculations. The database undergoes rigorous validation processes to ensure accuracy and consistency, incorporating data from laboratory spectroscopic studies, quantum mechanical calculations, and prior spectroscopic databases. The line lists include parameters such as wavenumber, intensity, broadening coefficients, and temperature dependence, all essential for modeling molecular absorption and emission in diverse atmospheric conditions. Advanced computational techniques are employed to process and format the data, making it accessible for scientific and operational use. Additionally, the GEISA database is periodically updated to integrate new spectroscopic data and enhance its

usability through improved search and visualization tools.

The primary aim of GEISA is to provide a comprehensive and high-accuracy spectroscopic database for atmospheric and planetary applications. The database is designed to support radiative transfer calculations, enabling precise modeling of atmospheric absorption and emission spectra. By continuously updating and expanding its molecular line lists, the GEISA database ensures the availability of high-resolution spectral data crucial for climate studies, remote sensing, and astrophysical applications.

2.6.7 HITRAN and HITEMP databases

For temperate planets with Earth-like temperatures, the HITRAN database [13] serves as the primary reference for molecular spectroscopic data in atmospheric and astrophysical studies. Accessible at https://hitran.org/, HITRAN has been developed over the past 50 years [13, 74] and provides extensive, high-accuracy spectroscopic data. It includes line lists and line broadening parameters for 55 small molecules present in atmosphere of Earth, as well as cross sections data for larger species and collision-induced absorption effects. A key strength of HITRAN is its comprehensive treatment of air-broadening effects for Earth, essential for precise atmospheric modeling. Further details about the HITRAN database are provided in Chapter 3.

While HITRAN does not aim to be comprehensive at high temperatures, efforts are underway to expand its applicability through the HITEMP database [54, 55], with contributions to its latest release in progress. The HITEMP database focuses on high-temperature spectroscopic data for molecular species. It serves as an extension of the HITRAN database, specifically designed to provide spectroscopic data at elevated temperatures, making it suitable for studies of exoplanetary and stellar atmospheres. The database includes spectroscopic data for 8 molecules: H_2O, CO_2, N_2O, CO, CH_4, NO, NO_2, and OH, with data applicable for temperatures reaching 6000 K.

HITEMP specializes in high-temperature infrared spectroscopy, covering a wavelength range from 1 μm to 100 μm, particularly for molecules such as H_2O and

CO_2. The database provides downloadable datasets, including detailed information on molecular line lists and spectroscopic parameters. More information and access to the data can be found on the official website https://hitran.org/hitemp/.

2.6.8 JPL database

Jet Propulsion Laboratory Molecular Spectroscopy Database (JPL) [56] is an extensive resource primarily used in radio astronomy, astrophysics, and atmospheric sciences. The database compiles data from laboratory measurements, theoretical calculations, and previous spectroscopic compilations, ensuring accuracy for microwave, submillimeter, and far-infrared rotational transitions.

It provides spectral line parameters for over 750 molecules, including their isotopologues and vibrationally excited states. The JPL database catalog covers species such as H_2O, CO, NH_3, CH_4, HCN, SO_2, and complex organic molecules, making it invaluable for interstellar medium (ISM) studies and planetary atmospheres. The methodology combines experimental microwave spectroscopy with quantum chemical calculations to predict additional transitions. The database spans a wavelength range from millimeter to far-infrared regions (20 GHz to several THz), with data accessible through the JPL website https://spec.jpl.nasa.gov/home.html.

2.6.9 CDMS database

The Cologne Database for Molecular Spectroscopy (CDMS) [57, 58] is a widely used resource in radio astronomy and interstellar medium studies, providing extensive spectral line catalogs for over 750 molecular species, including complex organic molecules. The database compiles data from high-resolution rotational spectroscopy experiments and quantum chemical calculations, ensuring reliable spectral line assignments.

Key molecules include CO, H_2O, NH_3, CH_3OH, HCN, SO, and many complex organic molecules relevant to astrochemistry. The CDMS database covers a spectral range from microwave to far-infrared wavelengths (frequencies from 100 MHz to several THz) and provides transition frequencies, Einstein A-coefficients, and partition functions. The CDMS database is accessible via the CDMS website

https://cdms.astro.uni-koeln.de, supporting advanced searches and data retrieval.

2.6.10 CASDA24 database

The Calculated Spectroscopic Databases 2024 (CASDA24) [59, 60] is the latest iteration of high-resolution molecular spectroscopy databases developed by the Laboratoire Interdisciplinaire Carnot de Bourgogne (ICB) at the Université de Bourgogne. This comprehensive cluster comprises ten independent databases (see Table 2.3), each dedicated to a specific molecule, integrating both updated and new datasets since previous releases.

Table 2.3: Overview of CASDA24 database with the list of molecules, their HITRAN ID, and corresponding database names.

Molecule	HITRAN ID	Database Name	Wavenumber Range
CH_4	06	MECASDA	$[0.000, 6772.680]$ cm^{-1}
CH_3Cl	24	CHMECASDA	$[657.935, 1145.886]$ cm^{-1}
SF_6	30	SHECASDA	$[319.583, 1496.228]$ cm^{-1}
C_2H_4	38	ECASDA	$[620.016, 3299.977]$ cm^{-1}
CF_4	42	TFMECASDA	$[0.924, 1329.697]$ cm^{-1}
GeH_4	52	GECASDA	$[648.358, 2270.201]$ cm^{-1}
RuO_4	–	RUCASDA	$[889.987, 948.254]$ cm^{-1}
SiF_4	–	TFSICASDA	$[352.538, 1060.680]$ cm^{-1}
UF_6	–	UHECASDA	$[616.438, 638.916]$ cm^{-1}
SiH_4	–	SICASDA	$[0.023, 1180.165]$ cm^{-1}

The CASDA24 database offers detailed lists of calculated spectral lines, eigenvectors, energy levels, and additional relevant information derived from the most recent experimental spectral analyses. These resources are essential for researchers engaged in high-resolution spectroscopy, providing accurate molecular data for various scientific applications.

The data provided in the CASDA24 database are generated through meticulous computational procedures. Effective Hamiltonian (EH) models are employed, with parameters finely tuned by fitting to assigned experimental lines. Dipole moment and polarizability parameters are similarly optimized, ensuring that the calculated spectral properties closely match observed data. These parameters result from the latest spectroscopic analyses published by the ICB group and close collaborators in

peer-reviewed journals. Additionally, the database includes purely calculated line broadening coefficients, offering comprehensive datasets for various applications. All analytical calculations are performed using specialized programs developed by the Dijon group, with significant collaboration from the Reims group, ensuring robustness and reliability in the data provided.

The CASDA24 database serves as a vital tool for the scientific community, facilitating advancements in molecular spectroscopy and related fields. Researchers can access the database through its website https://vamdc.icb.cnrs.fr/index.php.

2.6.11 VALD3 database

The Vienna Atomic Line Database (VALD3) [61] is a collection of atomic and molecular transition parameters relevant to astronomers for stellar spectroscopic studies. Its data originate from a compilation of various sources, offering evaluated and recommended values for spectral lines. It provides specialized tools for selecting line subsets tailored to various astrophysical applications, including line identification, spectroscopic observation planning, chemical composition analysis, radial velocity measurements, and model atmosphere calculations [61].

The database covers a wide spectral range, with atomic transitions predominantly in the ultraviolet (UV) and visible regions, as well as some near-infrared (IR) data. VALD3 grants users access to tools for extracting and combining data based on user-defined quality assessments. In terms of molecular coverage, the database includes spectroscopic data for 9 molecules: C_2, CH, CN, CO, OH, MgH, NH, SiH, and TiO. TiO [75] data from multiple computational sources were added in 2012, contributing to the study of molecular absorption in stellar atmospheres. Additionally, C_2 [76], CH [77], CN [78], CO [79], OH [80], MgH [81, 82] (partly from EXOMOL database), and SiH [83] line lists were incorporated from KURUCZ database (see Section 2.6.5), with the original data compiled into two line lists for each molecule, one for the lines with observed energy levels, and another one for the lines with predicted energy levels. More recently, NH [84, 85] data were added in 2021, further strengthening the role of VALD3 database as a critical resource for modeling molecular transitions in astronomical spectra.

The dataset contains approximately 1.2 million spectral lines with accurately measured wavelengths spanning from 20 Å to 200 μm, along with over 200 million lines with predicted positions, compiled into 180 files that map to the original data sources. The data can be achieved from its website https://vald.astro.uu.se.

2.6.12 BASECOL database

BASECOL [62, 86] is a ro-vibrational collisional excitation database and utilities. BASECOL specializes in collisional rate coefficients for molecular interactions with atoms, ions, molecules, and electrons, crucial for astrophysical and atmospheric modeling.

The database provides ro-vibrational excitation and de-excitation rates for species such as H_2O, CO, H_2, NH_3, and CH_4 molecules and their isotopologues. The data are generated using a combination of quantum scattering calculations and experimental measurements at temperatures from 1 K to 3000 K, making it essential for cold interstellar clouds, protoplanetary disks, and hot astrophysical environments. The database provides state-to-state rate coefficients and cross sections, allowing precise modeling of radiative transfer in various astrophysical media. The BASECOL database is accessible on website https://basecol.vamdc.eu, offering a user-friendly interface for retrieving collisional data.

2.6.13 W@DIS information system

The Water internet Accessible Distributed Information System (W@DIS) [64, 65] is a dedicated repository for water (H_2O) and its isotopologues, providing comprehensive data for atmospheric, planetary, and astrophysical applications. The database compiles experimental and theoretical results on energy levels, transition frequencies, and line profiles across the infrared, microwave, and far-infrared regions. W@DIS contains data for $H_2{}^{16}O$, $H_2{}^{17}O$, $H_2{}^{18}O$, HDO, $HD^{18}O$, and D_2O, covering temperatures from 10 K to over 3000 K, making it suitable for modeling terrestrial and extraterrestrial water vapor. The spectral range extends from submillimeter to near-infrared wavelengths, supporting applications in planetary atmospheres, exoplanet characterization, and interstellar chemistry. Data are sourced

from high-resolution laboratory spectroscopy and variational quantum mechanical calculations, ensuring high accuracy. W@DIS is freely available through the website https://wadis.saga.iao.ru.

2.6.14 CDSD, NOSD, NDSD, and ASD databases

The Carbon Dioxide Spectroscopic Databank (CDSD) provides comprehensive calculated line lists for carbon dioxide (CO_2) essential for atmospheric, planetary, and astrophysical applications at different reference temperatures. These line lists include detailed spectroscopic parameters such as line positions, line intensities, lower-state energies, air-broadening coefficients, self-pressure-induced broadening coefficients, temperature dependence exponents for air broadening, and full line labeling. Two primary versions of the CDSD databank, CDSD-296 and CDSD-1000, cater to different temperature conditions and spectral ranges.

The CDSD-296 databank [67] focuses on terrestrial atmospheric applications, providing over 419 600 spectral lines from 5.9 cm^{-1} to 12 784.1 cm^{-1}. This databank was generated with an intensity cutoff of 10^{-30} cm/molecule at a reference temperature of 296 K. It is particularly useful for atmospheric physics, including the monitoring of carbon dioxide in atmosphere of Earth, as well as in planetary studies, such as investigations of the atmospheres of Mars and Venus. Additionally, the databank is relevant to astrophysics and radiative transfer modeling in combustion processes. A portion of CDSD-296 is incorporated into the HITRAN database. The spectral data within CDSD-296 were derived using an effective Hamiltonian (EH) and effective dipole moment (EDM) parameters fitted to experimental data obtained by the research group or compiled from peer-reviewed literature. The line-shape parameters were obtained from recent semi-empirical calculations, ensuring a high level of accuracy in modeling atmospheric absorption and emission spectra. The CDSD-296 databank is available at https://ftp.iao.ru/pub/CDSD-296/.

The CDSD-1000 database [66] is optimized for high-temperature environments, such as Venusian and Martian atmospheres, combustion studies, and exoplanet research, containing 3 950 500 spectral lines from 257 cm^{-1} to 9648 cm^{-1}, with data computed for 1000 K. This dataset was generated with an intensity cut-

off of 10^{-27} cm/molecule at a temperature of 1000 K. It is particularly relevant for high-temperature environments, making it suitable for studying atmospheric processes involving carbon dioxide, planetary atmospheres, and radiative transfer in combustion processes. Unlike CDSD-296, the CDSD-1000 databank has been incorporated into the HITEMP database, which is designed for modeling high-temperature spectra. As with CDSD-296, the spectroscopic data in CDSD-1000 were computed using effective Hamiltonian (EH) and effective dipole moment (EDM) parameters, refined against experimental data either obtained by the research group or sourced from peer-reviewed literature. The line-shape parameters in CDSD-1000 also originate from the latest semi-empirical calculations, further ensuring the reliability of the dataset for various scientific applications. The CDSD-1000 databank can be download from https://ftp.iao.ru/pub/CDSD-1000/.

Both versions of the CDSD databank were developed using the effective operator method, employing global weighted fits of spectroscopic parameters, including those of the effective Hamiltonian (EH) and effective dipole moment (EDM) operators to observed data compiled from the literature. They provide critical spectroscopic data for diverse research fields, supporting accurate modeling of radiative transfer processes in planetary and astrophysical environments.

There are some related high-temperature databases which can expand on CDSD database, see Table 2.4. The free Nitrous Oxide Spectroscopic Databank (NOSD-1000 [70], website: https://ftp.iao.ru/pub/LTS/NOSD-1000/) and Nitrogen Dioxide Spectroscopic Databank (NDSD-1000 [71], website: https://ftp.iao.ru/pub/NDSD/) provide additional high-accuracy spectroscopic data for CO_2 isotopologues, supporting diverse applications in atmospheric and astrophysical modeling [87]. The Acetylene Spectroscopic Databank (ASD-1000 [72], website: https://ftp.iao.ru/pub/ASD1000/) focuses on acetylene spectroscopy, providing high-resolution and high-temperature spectral line data. They are all linear molecules treated using similar effective Hamiltonian (EH) formalism. This is particularly valuable for planetary atmosphere studies and laboratory spectroscopy, where acetylene plays a key role in radiative transfer and chemical processes.

Table 2.4: Related molecules of CDSD NOSD, NDSD, and ASD databases

Database	Databank Name	Molecule
CDSD	Carbon Dioxide Spectroscopic Databank	CO_2
NOSD	Nitrous Oxide Spectroscopic Databank	N_2
NDSD	Nitrogen Dioxide Spectroscopic Databank	NO_2
ASD	Acetylene Spectroscopic Databank	C_2H_2

2.6.15 LEIDEN photodissociation and photoionization cross-section database

Vacuum-ultraviolet (VUV) photons play a crucial role in driving chemical processes in space, making it essential to accurately determine photodissociation and photoionization rates across different radiation fields. These rates are derived from the cross sections of individual atoms and molecules, which are compiled in the LEIDEN VUV photodissociation and photoionization cross-section database [69]. Over the past decade, advances in both experimental and theoretical methods have led to the determination of numerous new cross sections, particularly for the photoionization of radicals. To incorporate these advancements, the LEIDEN database has been updated with more recent data and expanded to include several additional astrophysically relevant species.

The cross sections from the database have been used to compute photodissociation and photoionization rates in various radiation fields, including cosmic-ray-induced VUV fluxes. Additionally, the attenuation of these rates in shielded environments has been calculated as a function of dust, molecular and atomic hydrogen, atomic carbon, and self-shielding column densities. The degree to which different shielding mechanisms influence the rates depends on the specific atom or molecule, as well as the dust absorbance properties. The latest update to the LEIDEN VUV cross-section database includes 14 newly added molecular species and revisions for 16 previously cataloged entries. This update is accompanied by a concise review of the underlying physical processes, with a particular focus on photoionization mechanisms that were not extensively discussed in earlier database revisions. All updated data are publicly accessible through the LEIDEN VUV cross-

section database https://home.strw.leidenuniv.nl/ ewine/photo/, ensuring continued support for astrophysical and spectroscopic studies.

2.7 Summary

This chapter provides a comprehensive overview of different spectroscopic techniques and the role of spectral databases in modern astrophysical and atmospheric research. It begins by discussing transit spectroscopy, which measures the absorption features of exoplanet atmospheres during transits, providing insights into their composition and temperature. High-resolution cross-correlation spectroscopy enables precise measurement of radial velocity to study exoplanet characteristics and stellar motion. Direct imaging spectroscopy offers a method to observe exoplanets by capturing their light directly, bypassing the effects of stellar glare.

In this chapter, an overview of various astrophysically important molecular databases is provided. These databases play a crucial role in providing accurate line lists, molecular data, and high-temperature spectral information, essential for interpreting observational data. The databases highlighted include ExoMol, TheoReTS, NASA Ames, Kurucz, GEISA, HITRAN, HITEMP, JPL, CDMS, CaSDa24, VALD3, BASECOL, W@DIS, CDSD, NOSD, NDSD, ASD, and Leiden, each contributing unique data for molecular, atomic, and atmospheric studies. The chapter emphasizes the importance of these databases in supporting research across exoplanet studies, stellar atmospheres, and atmospheric sciences, facilitating high-precision spectroscopy for both theoretical and observational applications.

Table 2.5: Comparative overview of astrophysically important molecular databases (ExoMol, TheoReTS, NASA Ames, Kurucz, GEISA, HITRAN, HITEMP, JPL, CDMS, CaSDa24, VALD3, BASECOL, W@DIS, CDSD, NOSD, NDSD, ASD, and Leiden) used for astronomical studies.

Database	Primary Focus	Primary	Secondary	$N_{\rm mol}$	Temperature Range	Wavenumber Range	Wavelength Range
ExoMol	Exoplanet and hot atmosphere spectroscopy	✓	✓	91	$[0,10000]$ K	$[0,50000]$ cm^{-1}	$[0.1,100]$ μm
TheoReTS	High-temperature spectra of polyatomic molecules	✓		9	≤ 2000 K	$[200,25000]$ cm^{-1}	$[4,50]$ μm
NASA Ames	High-temperature Ames molecular spectroscopic data	✓		6	≤ 3000 K	$[0,30000]$ cm^{-1}	$[2.5,25]$ μm
Kurucz	Stellar atmospheres various diatomic and polyatomic molecules		✓	12	UV-visible	$[10000,40000]$ cm^{-1}	$[0.25,1]$ μm
GEISA	Spectroscopic parameters for radiative transfer simulations	✓	✓	58	≤ 3000 K	$[0,35877.03]$ cm^{-1}	$[0.279,100]$ μm
HITRAN	Atmospheric molecular absorption data		✓	55	≤ 800 K	$[0,25000]$ cm^{-1}	$[0.4,100]$ μm
HITEMP	High-temperature molecular spectroscopy		✓	8	≤ 6000 K	$[100,10000]$ cm^{-1}	$[1,100]$ μm
JPL	Rotational transitions of molecules in astrophysics		✓	136	≤ 1000 K	Microwave, submillimeter	mm to sub-mm
CDMS	Molecular spectroscopy for interstellar medium studies		✓	332	≤ 1000 K	Microwave, IR	mm to sub-mm
CaSDa24	High-resolution spectra of small molecules	✓		10	≤ 2000 K	$[0,6772.68]$ cm^{-1}	$[1.47,100]$ μm
VALD3	Molecular data for Stellar spectroscopy		✓	8	$[1000,250000]$ K	$[10000,50000]$ cm^{-1}	$[0.1,1]$ μm
BASECOL	Collisional excitation and relaxation data		✓		$[1,3000]$ K		
W@DIS	Water (H_2O) spectroscopic data system		✓	1	$[10,3000]$ K		
CDSD	Carbon dioxide (CO_2) spectroscopic data	✓		1	≤ 3000 K	$[5.9,12784]$ cm^{-1}	$[0.78,1695]$ μm
NOSD	Nitrous oxide (N_2O) spectroscopic data	✓		1	≤ 1500 K	$[260,8310]$ cm^{-1}	$[1.2,38.5]$ μm
NDSD	Nitrogen dioxide (NO_2) spectroscopic data	✓		1	≤ 1500 K	$[466,4776]$ cm^{-1}	$[2,21.5]$ μm
ASD	Acetylene (C_2H_2) spectroscopic data	✓		1	≤ 1500 K	$[3,10000]$ cm^{-1}	$[1,3333.4]$ μm
Leiden	VUV photodissociation and photoionization cross sections	✓	✓	89	≤ 20000 K	$[50000,1000000]$ cm^{-1}	$0.01,0.2$ μm

$N_{\rm mol}$: The number of molecules included in the database.

Chapter 3

ExoMol and HITRAN Databases Overview

This chapter aims to provide a comprehensive overview of the ExoMol [17] and HITRAN [13] databases, with the details of their data structures (see Section 3.3 and 3.6) and application program interface (API, see Section 3.4 and 3.7). The HITEMP database, which is a high-temperature molecular spectroscopic database associated with the HITRAN database, shares the same format as the HITRAN database [54, 55].

A notable feature of ExoMol, which it has in common with HITEMP, TheoReTS [49], and NASA Ames [50], is that their line lists are specifically created for research related to high-temperature molecules. These line lists are typically massive; for instance, the ExoMol database encompasses over 10^{12} transitions. Consequently, it becomes essential to have efficient and adaptable methods for post-processing this extensive data, enabling its transformation into formats such as cross-sections that are suitable for applications in astronomical modeling and spectroscopic research.

3.1 Quantum numbers

Quantum number cases, labels, and formats are essential for accurately describing molecular states in spectroscopic databases and computational models. They provide a standardized way to represent quantum states across different theo-

retical and experimental datasets, ensuring consistency and interoperability. Different quantum number cases correspond to various coupling schemes used to describe angular momenta in molecules, which is crucial for correctly interpreting transitions and interactions. Labels help uniquely identify quantum states, avoiding ambiguities in line lists and facilitating comparisons between different datasets. Standardized formats ensure that spectroscopic data can be efficiently exchanged, processed, and applied in simulations, making them indispensable for high-precision spectroscopic studies.

Quantum numbers play a crucial role in defining molecular states within spectroscopic databases. In the ExoMol database [17], they are an essential component of the states (`.states`) file, providing a detailed characterization of molecular energy levels. Similarly, in the HITRAN database [13], quantum numbers are incorporated within the parameter (`.par`) file, ensuring precise identification and classification of spectral transitions. These quantum numbers facilitate accurate modeling of molecular spectra, enabling their application in atmospheric, astrophysical, and laboratory studies. Their inclusion in both databases enhances the reliability and consistency of spectroscopic data, supporting a wide range of scientific investigations.

Quantum numbers are important for characterizing molecular energy levels and transitions. These quantum numbers are categorized into exact quantum numbers, which are fundamental and strictly conserved, and approximate quantum numbers, which serve as useful labels to describe molecular states but may not be rigorously conserved.

Exact quantum numbers in the ExoMol database correspond to fundamental physical symmetries of the molecule and are strictly conserved during transitions. For example:

- Total angular momentum (J): Represents the total angular momentum of the molecule, including rotational and electronic contributions (but excluding nuclear spin).

- Parity ($+/-$): Denotes the symmetry of the wavefunction under inversion ($+$ for

symmetric, − for antisymmetric).

- Symmetry index: Describes the overall symmetry of the state within the molecular point group, which is particularly important for distinguishing states in symmetric molecules. This index is derived from the irreducible representations of the appropriate symmetry group (e.g., A_1, A_2, B_1, B_2 in C_{2v} symmetry).

Approximate quantum numbers in the ExoMol database provide additional descriptors for molecular states but may not be strictly conserved due to interactions such as spin-orbit or Coriolis coupling. For example:

- Vibrational quantum numbers (v): Represent vibrational excitations of different normal modes.

- Electronic state: Indicate the electronic configuration, often labeled as $X^1\Sigma^+$, $A^1\Pi$, $B^2\Sigma^+$, etc.

- Vibronic states: A combination of vibrational and electronic states, which accounts for vibronic coupling effects.

- Rotational quantum numbers (e.g., K, K_a, and K_c): Used to describe rotational motion, particularly in asymmetric top molecules:

  – K: The projection of the rotational angular momentum along the principal axis in symmetric tops.

  – K_a and K_c: The pseudo-quantum numbers used for asymmetric top molecules, referring to the prolate and oblate principal axes, respectively.

Quantum numbers play a critical role in line broadening by influencing how molecular transitions interact with their environment. The rotational quantum number (J) affects pressure broadening, as higher J states experience different collisional interactions. Vibrational (v) and electronic state labels determine energy level spacing, impacting both Doppler and pressure broadening. Parity and symmetry influence selection rules, affecting transition probabilities and collisional effects.

Additionally, rotational substructure quantum numbers (K, K_a, and K_c) in asymmetric tops modify line shapes, while hyperfine structure can further split and alter line profiles. Understanding these quantum properties is essential for accurately modeling spectral line shapes in astrophysical and atmospheric studies.

3.2 ExoMol Database overview

The ExoMol project was designed in 2011 to provide molecular line lists for interpreting spectra, modelling and simulating the atmospheres of hot exoplanets, cool stars, brown dwarfs, and other hot astronomical objects [18]. This project is supported by comprehensive and extensive high-temperature molecular transition line lists pertinent to spectroscopic studies [3, 17, 48]. It specifically focuses on providing results for elevated temperatures which are insufficiently addressed by existing databases such as HITRAN [88] or GEISA [89]. The methodology of the ExoMol procedure integrates ab initio calculations with empirical laboratory data [90, 91]. Variational nuclear motion calculations are employed to produce line lists [90, 91], which are freely accessible through the ExoMol database website at https://www.exomol.com/. The ExoMol line lists are systematically published as a series in the journal Monthly Notices of the Royal Astronomical Society and summarised in four-yearly release papers. The updates information regarding its 2016 [3], 2020 [48], and the current 2024 [17] releases is presented in Table 3.1.

Table 3.1: The ExoMol database releases comparison.

ExoMol database release	Molecule number	Isotopologue number	Data types number	File types number	Transitions number
2016 release	35	65		12	
2020 release	80	190	11	15	10^{11}
2024 release	91	224	14	18	10^{12}

File types and data types are defined in Table 3.3 and Table 3.2, respectively. The number of molecules and isotopologues can be accessed from the master file (see Section 3.3.1) online https://www.exomol.com/db/exomol.all.

The ExoMol database serves as the foundational database for the following four chapters: database updates and maintenance (see Chapter 4), PyExoCross

program (see Chapter 6), non-local thermodynamic equilibrium (non-LTE) calculation program (see Chapter 8), and the EXOMOLHR database (see Chapter 9).

3.2.1 Explicit use of the EXOMOL database

Since the EXOMOL 2020 release, EXOMOL line lists have been widely used in the modeling of gas giant atmospheres, including sub-Neptunes [92–107], Neptune-like exoplanets [108–112], so-called Hycean planets [113–115], Saturn-like exoplanets [116–129], young Jupiters [130], cool Jupiters [131], warm Jupiters [97, 132–136], hot Jupiters [99, 137–202] and ultra-hot Jupiters [203–239]. They have also been employed in characterising the atmospheres of terrestrial planets [240–250], particularly hot rocky exoplanets [251–259], those containing volatiles [260] and those undergoing photoevaporation [261, 262]. Within our Solar System, they have been used to study the Venusian atmosphere [263, 264], volcanic plumes on Io [265] and comets [266]. In all types of planetary atmospheres, they have been applied in the modeling of clouds and hazes [92, 96, 124, 141, 172, 174, 185, 267–274] and photochemistry [275].

EXOMOL line lists have been used to assess the detection capabilities of both current and future space-based telescopes [97, 249, 276–284] and have frequently been integrated into various radiative transfer, forward modeling, spectral synthesis, and opacity codes [165, 285–295]. They have also been used to study the detectability of biosignatures [296] and prebiosignatures [246]. In star research, EXOMOL line lists have also contributed to the study of molecular species in the atmospheres of main-sequence dwarfs [297–303], giants [304–306], brown dwarfs [130, 271, 273, 307–325], eruptive Young Stellar Objects [326], in stellar winds from Oxygen-rich [327–329] and Carbon-rich [330, 331] Asymptotic Giant Branch stars, in cool Carbon stars [332, 333], in supergiants [327], in stellar merger remnants [334, 335] and modelling supernova ejecta [336]. They have also been used in the investigation of molecular species in the interstellar medium (ISM) [337, 338] and in studying the isotope exchange rates of molecules crucial to these environments [339].

Beyond astronomy, EXOMOL line lists have also been used in laser absorption

spectroscopy [340–343], laser induced breakdown spectroscopy [344, 345], laser fragmentation [346], laser induced fluorescence [347], plasma stoichiometry [348], fusion plasmas [349], the design of gas sensors [350], combustion [351] and explosions [352]. EXOMOL line lists are frequently used as benchmarks for *ab initio* calculations [353–358] and for comparing line strengths [359–366].

3.2.2 Data products from the EXOMOL database

The EXOMOL database is a comprehensive resource for molecular line lists, widely used in exoplanetary and stellar atmospheric studies. Over the years, several data products have been developed to expand its applications and enhance its usability for specific scientific purposes.

EXOMOLOP is an extension of the EXOMOL database that provides precomputed opacities for a range of molecules relevant to high-temperature exoplanetary atmospheres. Chubb et al. [367] introduced the EXOMOLOP database, which computed opacity cross sections and k-tables for all molecules based on EXOMOL line lists. By using EXOMOL line lists, EXOMOLOP ensures high-precision opacity calculations essential for modeling radiative transfer in exoplanetary and stellar environments. The availability of such data facilitates accurate modeling of atmospheric compositions and thermal structures.

Lifetimes DataBase (LIDB) [368] is a specialized database that provides molecular vibrational and vibronic state radiative lifetimes. It enables accurate modeling of radiative effects in low-temperature plasma environments by offering vibrational and electronic state-resolved radiative lifetimes. LIDB database also includes partial lifetimes, which help identify dominant decay channels for molecules. Currently, the database includes datasets for 36 molecules derived from EXOMOL line lists. The data is freely accessible through https://www.exomol.com/lidb, where users can dynamically explore molecular datasets or use an API for data retrieval. LIDB database is linked to the Quantemol DataBase (QDB) [369], further enhancing its integration into plasma modeling applications. Future expansions of LIDB database will incorporate more molecules, important isotopologues, and neutral and singly-charged atomic species.

Additionally, I created a EXOMOL High-Resolution database (EXOMOLHR), specifically designed to offer high-resolution (resolving power $R > 100000$) line lists tailored for applications requiring accurate spectral details, such as exoplanet atmospheric characterization and laboratory spectroscopy. This new database contains refined molecular data that enable accurate spectroscopic modeling of exoplanetary and stellar atmospheres, supporting the next generation of observational missions and high-resolution spectroscopic studies. The data is available at https://www.exomol.com/exomolhr. Chapter 9 will introduce the EXOMOLHR database in details.

Other data produces include specific heats [370], lifetimes [371], and cooling functions [3] are detailed in Chapter 6.

3.3 EXOMOL file structure and data format

The EXOMOL data structure, initially defined by Tennyson et al. [372], has undergone substantial enhancements as outlined in the subsequent work [3]. Nevertheless, the states and transitions files, which were integral to the original specification, have retained their original foundational structures. Consequently, utilities developed for the previous format will remain operational [3].

Table 3.2 presents a summary of the various data types available in the EXOMOL database. Table 3.3 provides a summary of the file types associated with each isotopologue within the revised EXOMOL database. A detailed enumeration of these files follows below. Note that EXOCROSS [373] and PYEXOCROSS generate additional file types such as `.os` for oscillator strengths and `.stick` for stick spectra, however, these files are not distributed via the EXOMOL website and database. Figure 3.1 illustrates a schematic representation of the file structures of the EXOMOL database.

New JSON format master and definition files have been implemented in the EXOMOL database and will be described details in Chapter 4. The following Section 3.3.1 and 3.3.2 will only discusses the old text version.

Table 3.2: Summary of data provided by the EXOMOL Database.

Data type
Line lists
Pressure broadening coefficients
Absorption cross sections
Partition functions
Cooling functions
Specific heat - heat capacity
Examples of ExoCross input files
Temperature dependent super-lines (histograms)
Photo-absorption continuum cross section*
Photo-dissociation cross sections including VUV absorption*
Temperature and pressure dependent opacities
Spectroscopic Models

* Denotes a new or updated data type.

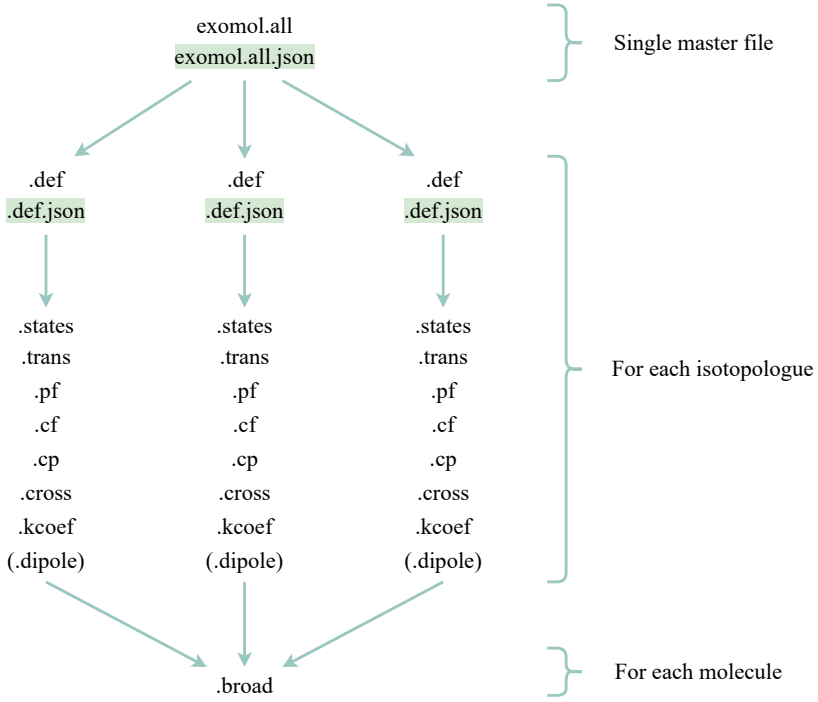


Figure 3.1: Summary of the current EXOMOL data structures [3]. New JSON format file types are highlighted in colored and will be introduced in Chapter 4.

Table 3.3: Specification of EXOMOL file types. The file types highlighted in colours are mentioned in the PYEXOCROSS and EXOMOLHR projects. Their data structures are described in the following sections. All these file types are included in data maintenance. Note that EXOCROSS/PYEXOCROSS generates additional file types such as `.os` for oscillator strengths but these files are not distributed via the EXOMOL website or database.

File extension	$N_{\rm files}$	File DSname	Contents
`exomol.all`	1	Master	Single file defining contents of the EXOMOL database
`exomol.all.json`	1	Master	Single file defining contents of the EXOMOL database in JSON format
`.def`	$N_{\rm tot}$	Definition	Defines contents of other files for each isotopologue
`.def.json`	$N_{\rm tot}$	Definition	Defines contents of other files for each isotopologue in JSON format
`.states`	$N_{\rm tot}$	States	Energy levels, quantum numbers, (uncertainties, lifetimes, Landé g-factors)
`.trans`	[a]	Transitions	Einstein A coefficients, (wavenumber)
`.broad`	$N_{\rm mol}$	Broadening	Parameters for pressure-dependent line profiles
`.cross`	[b]	cross-sections	Temperature or temperature and pressure-dependent cross-sections
`.pf`	$N_{\rm tot}$	Partition function	Temperature-dependent partition function
`.cf`	$N_{\rm tot}$	Cooling function	Temperature-dependent cooling function
`.cp`	$N_{\rm tot}$	Specific heat	Temperature-dependent specific heat
`.kcoef`	[c]	$k-$coefficients	Temperature and pressure-dependent k-coefficients
`.dipoles`	$N_{\rm tot}$	Dipoles	Transition dipoles including phases
`.super`	[d]	Super-lines	Temperature dependent super-lines (histograms) on a wavenumber grid
`.nm`	[f]	VUV cross-sections	Temperature and pressure dependent VUV cross-sections (wavelength, nm)
`.fits`, `.h5`, `.kta`	[g]	Opacities	Temperature and pressure dependent opacities for radiative-transfer applications
`.pdef`	$N_{\rm tot}$	Definition	Defines contents of `.photo` files for each isotopologue
`.cont`	$N_{\rm iso}$	Continuum	Continuum contribution to the photoabsorption
`.photo`	[e]	Phtot-dissociation	Temperature dependent photo-dissociation cross-sections
`.overview`	$N_{\rm mol}$	Overview	Overview of datasets available
`.readme`	$N_{\rm iso}$	Readme	Specifies data formats
`.model`	$N_{\rm iso}$	Model	Model specification

$N_{\rm files}$ Total number of possible files.
$N_{\rm mol}$ Number of molecules in the database.
$N_{\rm tot}$ It is the sum of $N_{\rm iso}$ for the $N_{\rm mol}$ molecules in the database.
$N_{\rm iso}$ Number of isotopologues considered for the given molecule.
[a] There are $N_{\rm tot}$ sets of `.trans` files.
But for molecules with large numbers of transitions the .trans files are subdivided into wavenumber regions.
[b] There are $N_{\rm cross}$ sets of `.cross` files for isotopologue.
[c] There are $N_{\rm kcoef}$ sets of `.kcoef` files for each isotopoplogue.
[d] There are N_T sets of T-dependent super-lines.
[e] There are N_T sets of T-dependent photo-dissociation cross-sections.
[f] There are N_{VUV} sets of VUV cross-sections.
[g] Set of opacity files in the format native to specific radiative-transfer programs.

3.3.1 Master file

A master file (`exomol.all`) contains an overview of the contents of the EXOMOL database which can be obtained at https://www.exomol.com/db/exomol.all. This master file points towards the files that store the actual data [3].

The master file is designed to facilitate machine-based searches, thereby allowing users to efficiently verify the current contents of the database and determine the latest update dates for its entries [3]. It is organized as a list of molecules, each accompanied by its corresponding isotopologues with their recommended datasets only. The first column of the master file is the details of contents and the second column is their corresponding descriptions. The actual EXOMOL data files are stored

according to isotopologue. Table 3.4 delineates the format of the master file. New JSON format master file will be described in Section 4.2.3.

Table 3.4: Format of the EXOMOL master file, `exomol.all`; each entry starts on a new line.

Field	Fortran format	C format	Description
Header information			
`ID`	`A13`	`%13s`	Always the ASCII string '`exomol.def`'
$V_{\rm all}$	`I8`	`%8d`	Version number, format YYYYMMDD, recording last update of the whole database
$N_{\rm mol}$	`I4`	`%4d`	Number of molecules in the database
Followed by $N_{\rm mol}$ entries of the form			
Molecule information			
$N_{\rm name}$	`I3`	`%3d`	Number of molecule names listed
Followed by $N_{\rm name}$ entries of the form			
`MolName`	`A27`	`%27s`	(Common) name of the molecule
`MolFormula`	`A27`	`%27s`	Molecule chemical formula
$N_{\rm iso}$	`I4`	`%4d`	Number of isotopologues considered
Followed by $N_{\rm iso}$ entries of the form			
Isotopologue information			
`IsoKey`	`A27`	`%27s`	Inchi key of the isotopologue
`IsoFormula`	`A27`	`%27s`	Isotopologue chemical formula
`Iso-slug`	`A160`	`%160s`	Isotopologue slug identifier, see text for details
`DSName`	`A10`	`%10s`	Isotopologue dataset name
$V_{\rm ds}$	`I8`	`%8d`	Version number with format YYYYMMDD
Metadata			Free format

3.3.2 Definition file

The definition file `.def` contains pertinent information regarding the information of the EXOMOL database for a specific isotopologue, as well as guidelines on the usage of the data [3]. Similar to the master file, the `.def` file structures information with the first column listing content details and the second column providing their corresponding descriptions. The `.def` files are named in format `<ISO-SLUG>__<DATASET>.def`. The structure is delineated in Table 3.5 and can be categorized into various sections (emphasized in bold) that each pertain to distinct characteristics of the molecule in the EXOMOL format. Each section may encompass several subsections that elaborate on further aspects of that section. Consequently, the file format exhibits a hierarchical tree structure. EXOMOL database has a new JSON (JavaScript Object Notation [21]) format `.def.json` (`<ISOTOPOLOGUE>__<DATASET>.def.json`) file, more details see Section 4.3.

Table 3.5: Format of the EXOMOL definition file, `<ISO-SLUG>__<DATASET>.def`; each entry starts on a new line.

Field	Fortran format	C format	Description
Header information			
`ID`	`A10`	`%10s`	Always the ASCII string '`exomol.def`'
`IsoFormula`	`A27`	`%27s`	Isotopologue chemical formula
`Iso-slug`	`A160`	`%160s`	Isotopologue identifier, see text for details
`DSName`	`A10`	`%10s`	Isotopologue dataset name
$V_{\rm ds}$	`I8`	`%8d`	Version number with format YYYYMMDD
`MolKey`	`A27`	`%27s`	Standard inchi key of the molecule
$N_{\rm atom}$	`I4`	`%4d`	Number of atoms
Atom definition (the following 2 lines occur $N_{\rm atom}$ times)			
$I_{\rm atom}$	`I3`	`%3d`	Isotope number
Atom	`A3`	`%3s`	Element symbol
Isotopologue information			
m_{Da} m_{kg}	`F12.6,1X,ES14.8`	`%12.6f %14.8e`	Isotopologue mass in Da and kg
$I_{\rm sym}$	`A6`	`%6s`	Molecular symmetry Group (if $N_{\rm atom} = 2$ then C or D)
$N_{\rm irrep}$	`I4`	`%4d`	Number of irreducible representations
Symmetry definitions (the following 3 lines occur $N_{\rm irrep}$ times)			
$I_{\rm irrep}$	`I3`	`%3d`	Irreducible representation ID
Symmetry	`A6`	`%6s`	Irreducible representation label
$g_{\rm ns}$	`I3`	`%3d`	Nuclear spin degeneracy
EXOMOL information			
$T_{\rm max}$	`F8.2`	`%8.2f`	Maximum temperature of the line list
$N_{\rm broad}$	`I3`	`%3d`	No. of pressure broadeners available
$D_{\rm avail}$	`I2`	`%2d`	Dipole availability (1 = Yes 0 = No)
$N_{\rm cross}$	`I3`	`%3d`	No. of cross section files available
$N_{\rm kcoef}$	`I3`	`%3d`	No. of k-coefficient files available
State file information			
$\tau_{\rm avail}$	`I2`	`%2d`	Flag denoting lifetime availability (1 = Yes 0 = No)
$g_{\rm avail}$	`I2`	`%2d`	Flag denoting Landé g-factor availability (1 = Yes 0 = No)
$N_{\rm states}$	`I10`	`%10d`	No. of states in `.states` file
$N_{\rm cases}$	`I3`	`%3d`	No. of quanta cases
Quanta case definition (the following line occurs $N_{\rm cases}$ times)			
Case Label	`A8`	`%8s`	Label of the quanta case
$N_{\rm quanta}$	`I3`	`%3d`	No. of quanta defined
Quanta definition (the following 3 lines occur $N_{\rm quanta}$ times)			
Quanta Label	`A3`	`%3s`	
$F_{\rm Fortran}$ $F_{\rm C}$	`A8,1x,A8`	`%8s %8s`	Fortran and C format for quanta
Description	`A40`	`%40s`	Short description of quanta
Trans file definition			
$N_{\rm trans}$	`I15`	`%15d`	Total number of transitions
$N_{\rm files}$	`I4`	`%4d`	Number of transition files
$\tilde{\nu}_{\rm max}$	`F8.2`	`%8.2f`	Maximum wavenumber (in cm^{-1})
$E_{\rm lower}$	`F8.2`	`%8.2f`	Higher energy with complete set of transitions (in cm^{-1})
Partition function information			
$T^Q_{\rm max}$	`F8.2`	`%8.2f`	Maximum temperature of partition functions
$T_{\rm step}$	`F5.2`	`%5.2f`	Temperature step size
$C_{\rm avail}$	`I2`	`%2d`	Cooling function availability (1 = Yes 0 = No)
Pressure broadening information (default values)			
γ_D	`F6.4`	`%6.4f`	Default value of Lorentzian half-width for all lines (in cm^{-1}/bar)
n_D	`F5.3`	`%5.3f`	Default value of the temperature exponent for all lines
Pressure broadening information ($N_{\rm broad} > 0$, the following 6 lines occur $N_{\rm broad}$ times)			
Broadener Label	`A8`	`%8s`	Label for particular broadener
Broadener File Name	`A20`	`%20s`	File name of particular broadener
$J_{\rm max}$	`I7/F7.1`	`%7d/%7.1f`	Maximum J'' for which pressure broadened parameters are provided
γ_L	`F6.4`	`%6.4f`	Value of Lorentzian half-width for lines with $J' > J_{\rm max}$ (in cm^{-1}/bar)
n_L	`F5.3`	`%5.3f`	Value of the temperature exponent for lines with $J' > J_{\rm max}$
N_Q	`I4`	`%4d`	Number of defined quantum number sets
Quantum sets (the following 3 lines occur N_Q times)			
Set Code	`A2`	`%2s`	A code that defines this set of quantum numbers
$N_{\rm lines}$	`I6`	`%6d`	No. of lines in the broad that contain this code
$N_{\rm Quanta}$	`I4`	`%4d`	No. of quantum numbers defined
Quantum numbers (the following line occurs $N_{\rm Quanta}$ times)			
Quantum Label	`A4`	`%4s`	
State file extra information			
$U_{\rm avail}$	`F12.6`	`%12.6f`	Uncertainty availability (1 = Yes 0 = No)

3.3.3 Line list files

The fundamental structure of the EXOMOL database consists of an extensive enumeration of the molecular states involved [3]. Irrespective of the specific molecule, the configuration of any EXOMOL state file can be ascertained from its corresponding definition file [3]. This section will introduce the line list files which are the cores of the EXOMOL data structure. Line list files includes the states file `.states` and transitions file `.trans`. The `.states` file contains the energy levels along with their corresponding quantum numbers, while the `.trans` file provides the Einstein A-coefficients which are associated with upper and lower state indices (counting numbers) which point to the corresponding the `.states` file.

3.3.3.1 States file

The states files `.states` provides the state ID, which should match the line number, and is used in the `.trans` file to identify the initial and final states. Following this, the state energy, degeneracy, and total angular momentum are specified. In most cases, the total angular momentum is given by the quantum number J, but for hyperfine resolved line lists, this is F where $F = |\underline{J} + \underline{I}|$ and I is the nuclear spin [17]. The uncertainty that was introduced in the EXOMOL 2020 release paper [48] has become a mandatory component of the recommended EXOMOL line lists, although it remains unavailable for older line lists and the majority of non-EXOMOL lists provided by the database. This uncertainty is crucial for high-resolution studies and can inform the future experimental priorities. EXOMOLHR database only includes the states with low uncertainty, more details see Chapter 9. Additionally, the state lifetime τ is now a required feature for recommended EXOMOL line lists; its definition has been broadened to encompass the lifetime derived from both radiative decay [371] and predissociation effects [374].

The `.states` files are compressed using `.bz2` format. The PYEXOCROSS and EXOMOLHR projects allow programs to read and process this format without the necessity of pre-decompressing the `.bz2` files. A comprehensive specification of the mandatory section of the states file is presented in Table 3.6. A sample extract from a `.states` file is given in Table 3.7.

Table 3.6: Format of the EXOMOL states file, `<ISO-SLUG>__<DATASET>.states`; each entry starts on a new line.

Field	Fortran format	C format	Description
Mandatory fields			
ID	`I12`	`%12d`	State ID (state identifier)
$\tilde{E}$	`F12.6`	`%12.6f`	Recommended state energy in cm^{-1}
$g_{\rm tot}$	`I6`	`%6d`	State degeneracy
J	`I7/F7.1`	`%7d/%7.1f`	Total angular momentum quantum number, J or F (integer/half-integer)
ΔE	`F12.6`	`%12.6f`	Energy uncertainty in cm^{-1}
τ	`ES12.4`	`%12.4e`	State lifetime (aggregated radiative and predissociative lifetimes) in s
Optional fields			
(g)	`F10.6`	`%10.6f`	Landé g-factor (optional)
(QN)	See `.def` file		State quantum numbers, may be several columns (optional)
(Label)	`A2`	`%2s`	Abbreviation giving source of state energy (optional)
($\tilde{E}_{\rm cal}$)	`F12.6`	`%12.6f`	Calculated state energy in cm^{-1} (optional)

ID: a non-negative integer index, starting from 1.
J: total angular momentum quantum, excluding nuclear spin.
The optional fields are in order if $U_{\rm avail} = 1$, $\text{Life}_{\rm avail} = 1$, $g_{\rm avail} = 1$.
Quanta Cases in order of definition.
Fortran format: J integer: `(I12,1x,F12.6,1x,I6,I7,1x,ES12.4,1x,F10.6)`
or J half-integer: `(I12,1x,F12.6,1x,I6,F7.1,1x,ES12.4,1x,F10.6)`.

Table 3.7: Extract from the states file of the ^{24}Mg^{1}H line list in the EXOMOL database, `24Mg-1H__XAB.states.bz2`.

ID	$\tilde{E}$	$g_{\rm tot}$	J	ΔE	τ	g	$+/-$	e/f	ElecState	v	Λ	Σ	Ω	Label	$\tilde{E}_{\rm cal}$
34	24986.629171	4	0.5	2.507500	5.4398E-08	-0.000637	+	e	A(2PI)	4	1	-0.5	0.5	Ca	24986.629171
35	25163.713092	4	0.5	0.022010	2.0005E-07	2.002233	+	e	B'(2SIGMA+)	4	0	0.5	0.5	Ma	25163.694359
36	25876.111435	4	0.5	0.101011	3.1000E-07	2.002217	+	e	B'(2SIGMA+)	5	0	0.5	0.5	Ma	25876.060619
37	26224.114885	4	0.5	3.007500	5.8534E-08	-0.000612	+	e	A(2PI)	5	1	-0.5	0.5	Ca	26224.114885
38	26564.751435	4	0.5	0.112871	3.3290E-07	2.002249	+	e	B'(2SIGMA+)	6	0	0.5	0.5	Ma	26564.687216
39	27228.719704	4	0.5	4.007500	3.5659E-07	2.002009	+	e	B'(2SIGMA+)	7	0	0.5	0.5	Ca	27228.719704
40	27373.978859	4	0.5	3.507500	6.5066E-08	-0.000434	+	e	A(2PI)	6	1	-0.5	0.5	Ca	27373.978859
41	27867.308628	4	0.5	4.507500	5.7407E-07	2.002278	+	e	B'(2SIGMA+)	8	0	0.5	0.5	Ca	27867.308628

ID: State counting number.
$\tilde{E}$: State energy in cm^{-1}.
$g_{\rm tot}$: State degeneracy.
J: Total angular momentum.
ΔE: Energy uncertainty in cm^{-1}.
τ: Lifetime in s.
g: Landé g-factor.
$+/-$: Total parity.
e/f: Rotationless parity.
ElecState: Electronic state; state term value using PYVALEM format [375, 376].
v: State vibrational quantum number.
Λ: Projection of the electronic angular momentum.
Σ: Projection of the electronic spin.
Ω: $\Omega = \Lambda + \Sigma$, projection of the total angular momentum.
Label: Source label of state energy; 'Ma' for MARVEL and 'Ca' for calculated.
$\tilde{E}_{\rm cal}$: Original (calculated) state energy in cm^{-1}.

3.3.3.2 Transitions file

The transitions file `.trans` has a simple structure, given in Table 3.8. Table 3.9 is a sample extract from a `.trans` file of ^{24}Mg^{1}H line list. Two pointers, i and f, reference rows in the `.states` file to identify the upper and lower states, along with more details that characterizing these states [3]. The `.trans` file may be quite extensive, and there is an optional fourth column, $\tilde{\nu}_{fi}$, which provides the transition wavenumber/frequency. If this column is not included, the wavenumber

should be calculated from the states file as the difference between the upper and lower energy levels [3] and actually we recommend doing it this way in all cases.

Table 3.8: Format of the EXOMOL transitions file, `<ISO-SLUG>__<DATASET>.trans`; each entry starts on a new line.

Field	Fortran format	C format	Description
f	`I12`	`%12d`	Upper state ID
i	`I12`	`%12d`	Lower state ID
A_{fi}	`ES10.4`	`%10.4e`	Einstein A coefficient in s^{-1}
$(\tilde{\nu}_{fi})$	`F15.6`	`%15.6f`	Transition wavenumber in cm^{-1} (optional)

i means initial state and f means final state.
Fortran format: `(I12,1x,I12,1x,ES10.4,1x,F15.6)`.

Table 3.9: Extract from the transitions file of the $^{24}Mg^{1}H$ line list in the EXOMOL database, `24Mg-1H__XAB.trans.bz2`.

f	i	A_{fi}	$\tilde{\nu}_{fi}$
6848	6733	8.9488E-06	53.783064
2020	2135	1.1292E-04	54.413400
2020	1905	2.2579E-02	54.480361
2250	2135	2.2821E-02	54.483636
5699	5814	8.8812E-08	54.756859

f: Upper state counting number.
i: Lower state counting number.
A_{fi}: Einstein A-coefficient in s^{-1}.
$\tilde{\nu}_{fi}$: Transition wavenumber in cm^{-1} (optional).

For large line lists in the EXOMOL database, a huge number of transitions are provided as a series of files which are named by wavenumber regions. To facilitate these files use, these `.trans` files are frequently split into specific wavenumber regions. Only for small line lists, the `.trans` files provide the wavenumbers as the fourth column, and even then, they are still recalculated in my programs using the energies from the `.states` file, as they might not incorporate updates made to the `.states` file due to various corrections. Additionally, the `.trans` files are compressed using the `.bz2` format, but can be read and processed by the programs mentioned in this thesis directly and automatically without predecompressing.

3.4 ExoMol API

The general approach of ExoMol database is molecule-by-molecule: a comprehensive line list is generated and accessible for each specific molecule [372]. The line list for each isotopologue is stored as a separate data structure, which can be accessed either directly or through the application programming interface (API). The API of the ExoMol database allows users to extract data directly between computers instead of manually downloading datasets from the website.

3.4.1 ExoMol DOI on Zenodo

Zenodo is an open-access repository developed by CERN and OpenAIRE that allows researchers to store, share, and preserve scientific data, software, and publications. Each upload is assigned a Digital Object Identifier (DOI), making datasets citable, trackable, and easily referenced in scientific publications. Zenodooffers immediate availability, with no waiting time for uploads to go online, and DOIs are registered within seconds. It also supports both open and restricted access, allowing researchers to control data sharing, such as limiting access to sensitive information. With built-in versioning, users can update datasets while maintaining a clear record of previous versions.

The ExoMol database uses Zenodo to upload and archive its molecular line lists and spectroscopic data, ensuring long-term accessibility and citation through DOIs. By assigning DOIs to datasets, ExoMol enables researchers to reference specific versions of data in publications, promoting reproducibility and proper attribution. Zenodo also provides metadata indexing, making ExoMol data discoverable and easily accessible to the broader scientific community. The ExoMol community on Zenodo is freely accessible from https://zenodo.org/communities/exomol.

3.4.2 Searching for data through the API

The ExoMol database provides a comprehensive master file (exomol.all) that catalogs all available molecules and their isotopologues. For a specified molecule or list of isotopologues, the ExoMol database can be queried to

retrieve recommended datasets. This search can be conducted via the API, which supports queries using the HTTP GET request method outlined below. The structure of a JSON API file is illustrated in Listing A.1. To retrieve all datasets and associated files for a specific molecule, the following query can be used: http://exomol.com/api?molecule=<MOLECULE>. For example, http://exomol.com/api?molecule=H2O.

3.4.3 Accessing the data via the API

The EXOMOL website (https://www.exomol.com) hosts all datasets which are available for viewing and downloading. However, manually downloading multiple datasets directly from the website is inefficient. The organized structure of the EXOMOL database file system enables automatic batch downloads directly through software or by using tools like `wget` and `curl` [48]. The API structures the EXOMOL data resource files into URLs with the form http://www.exomol.com/db/<MOLECULE>/<ISO-SLUG>/<DATASET>/<FILENAME>, where <FILENAME> is structured as `<ISO-SLUG>__<DATASET>.<DATATYPE>`. For example, the compressed states file of the POKAZATEL line list for $^{1}H_{2}{}^{16}O$ can be freely obtainable at the URL https://www.exomol.com/db/H2O/1H2-16O/POKAZATEL/1H2-16O__POKAZATEL.states.bz2.

3.5 HITRAN Database overview

HITRAN is an acronym for the high-resolution transmission molecular absorption database, which is a comprehensive collection of spectroscopic parameters [10, 13]. HITRAN database serves as a crucial resource for various computational tools, enabling the prediction and simulation of light transmission and emission in atmospheric environments. HITRAN database includes 138 isotopologues from 55 molecules with over 10^7 transitions [13]. The HITRAN database is available at https://hitran.org.

3.6 HITRAN file structure and data format

Table 3.10 provides the detailed format of the parameters associated with each spectral line, corresponding to a single record in the database [10]. The record from its 2004 edition has a total length of 160 characters [10]. The parameters presented in Table 3.10 are described in Table 3.11.

3.7 HITRAN API

The HITRAN Application Programming Interface (HAPI) [377] is a Python-based set of routines designed to provide remote access to the functionality and data available on HITRANonline (https://hitran.org). Currently, HAPI supports downloading, filtering, and processing line-by-line transition data. Its primary purpose is to enhance the main site's capabilities, particularly for calculating spectra using various line shapes, including the flexible Hartmann-Tran (HT) profile [378], with optional consideration of instrumental functions. Each API feature is implemented as a Python function that accepts arguments specifying the task parameters.

3.8 Summary

This chapter introduces the ExoMol and HITRAN databases, two key resources for molecular spectroscopic data. ExoMol focuses on high-temperature line lists for exoplanetary and stellar atmospheres, providing extensive theoretical calculations for a wide range of molecules. In contrast, HITRAN is a widely used reference database for terrestrial and planetary atmospheres, offering high-precision experimental and semi-empirical data. It highlights the complementary nature of these databases, with ExoMol catering to extreme astrophysical conditions and HITRAN providing accurate data for atmospheric science.

Table 3.10: HITRAN line-by-line format in 2004 Edition [10] (160-character record).

Parameter	M	I	ν	S	A	γ_{air}	γ_{self}	E''	n_{air}	δ_{air}	V'	V''	Q'	Q''	I_{err}	I_{ref}	*	g'	g''
Field length	2	1	12	10	10	5	5	10	4	8	15	15	15	15	6	12	1	7	7
Fortran Format	I2	I1	F12.6	E10.3	E10.3	F5.4	F5.4	F10.4	F4.2	F8.6	A15	A15	A15	A15	6I1	6I2	A1	F7.1	F7.1
C Format	%2d	%1d	%12.6f	%10.3e	%10.3e	%5.4f	%5.4f	%10.4f	%4.2f	%8.6f	%15s	%15s	%15s	%15s	%6d	%12d	%1s	%7.1f	%7.1f

Table 3.11: Description of the parameters presented in the 160-character records in the line-by-line section of the HITRAN database [10].

Parameter	Meaning	Field length	Type	Comments or units
M	Molecule number	2	Integer	HITRAN chronological assignment
I	Isotopologue number	1	Integer	Ordering within a molecule by terrestrial abundance
ν	Vacuum wavenumber	12	Real	cm^{-1}
S	Intensity	10	Real	$\text{cm}^{-1}/(\text{molecule cm}^{-2})$ at standard 296 K
A	Einstein A-coefficient	10	Real	s^{-1}
γ_{air}	Air-broadened half-width	5	Real	HWHM at 296 K (in cm^{-1} atm^{-1})
γ_{self}	Self-broadened half-width	5	Real	HWHM at 296 K (in cm^{-1} atm^{-1})
E''	Lower-state energy	10	Real	cm^{-1}
n_{air}	Temperature-dependent exponent for γ_{air}	4	Real	Unitless, with $\gamma_{\text{air}}(T) = \gamma_{\text{air}}(T_0)(T_0/T)^{n_{\text{air}}}$
δ_{air}	Air pressure-induced line shift	8	Real	cm^{-1} atm^{-1} at 296 K
V'	Upper-state global “quanta”	15	Hollerith	See Table 3 in [10]
V''	Lower-state global “quanta”	15	Hollerith	See Table 3 in [10]
Q'	Upper-state local “quanta”	15	Hollerith	See Table 3 in [10]
Q''	Upper-state local “quanta”	15	Hollerith	See Table 3 in [10]
I_{err}	Uncertainty indices	6	Integer	Accuracy for 6 critical parameters
I_{ref}	Reference indices	12	Integer	References for 3/6 critical parameters
*	Flag	1	Character	Availability of program and data for the case of line mixing
g'	The statistical weight of the upper state	7	Real	See details in [379]
g''	The statistical weight of the lower state	7	Real	See details in [379]

Chapter 4

Updating and Expanding the ExoMol Database

This chapter outlines the process of updating new line lists in the ExoMol database (see Section 4.2), introduces the new ExoMol data structure (JavaScript Object Notation (JSON) [21] format) for master and definition files (see Section 4.3), and describes the procedure for uploading new line lists to the ExoMol website (see Section 4.4).

4.1 Data maintenance

The primary objectives of the database maintenance project include the uploading, updating, verification, and generation of ExoMol database files, with plans to eventually automate all of these processes.

Step 1: Review new line list files: Begin by checking the newly obtained line list files to get some line list file information for following steps, see more details in Section 4.2.2.

- In the states file, check if the lifetime column is existed. If it is missing, use PyExoCross to compute the lifetimes.
- Verify whether the partition function file exists. If it does not, use PyExoCross to calculate the partition functions.
- The file size of the states, transitions, and partition function files.

Step 2: Incorporate data into master file: Add the new line list information to the master file (see Section 4.2.3).

Step 3: Generate or modify definition files:

- Create a definition file for the new line lists (see Section 4.2.4).
- Alternatively, update the existing definition file if applicable.

Step 4: Update ExoMol website: Finally, update the ExoMol website with the latest information (see Section 4.4) and the Zenodo of ExoMol.

4.2 New line lists and new versions

The ExoMol database of molecular line lists plays a crucial role in providing essential spectroscopic data for the investigation of hot atmospheres [380]. Spectroscopy, as a underpinning technique, finds extensive applications across various scientific fields beyond just astronomy and geophysics [380]. The ongoing computation of additional line lists remains a key activity, driven by the high demand for such data [380]. Therefore, the regular updating of the ExoMol database is of paramount importance.

4.2.1 Versioning and naming

The version number of a line list corresponds to the date when the line list files were last updated, which includes the addition of new `.states` and/or `.trans` files. The version number follows the format `YYYYMMDD`, where `YYYY` represents the year, `MM` the month, and `DD` the day. For each dataset in the master file, the version number is linked to the corresponding definition (`.def`) file. The version number of the master file should always reflect the most recent update across all the datasets. Additionally, when a new line list is added or updated, its name must be updated in both the `.def` and the master file.

The use of version numbers for line lists is essential to ensure clarity and proper tracking of updates. Users can easily determine whether a line list has been updated and identify the most recent version of the data. This system helps maintain con-

sistency, supports reproducibility, and ensures that researchers are using the latest available datasets in their analyses.

4.2.2 Steps for updating a new line list for each isotopologue

Some of the steps are used for creating or updating the master and definition files. And some information is used to update the descriptions on the ExoMol website. The PyExoCross program repository is freely available on GitHub: https://github.com/Beryl-Jingxin/PyExoCross. And the Jupyter Notebook codes can be obtained from https://github.com/Beryl-Jingxin/PyExoCross/tree/main/ExoMol_format.

Step 1: Check the states file (`.states`).

- Use a Jupyter Notebook code in PyExoCross program to verify that the format of the line list is satisfies ExoMol standards.
- For the `.states` file, ensure sorting by state ID, and if lifetimes are missing, use the PyExoCross program to calculate lifetimes and add them into the `.states` file.
- If the electronic states column exists, make sure its format is satisfy IAEA standard, otherwise, change its format with PyValem [375, 376]. For example, convert format from X1sigma+ to X(1SIGMA+).
- Get the number of states, the maximum energy, and the file size of the states file.
- Read related paper to get the column names which include the quantum number labels.

Step 2: Check the transitions file (`.trans`).

- For the `.trans` file, sort by frequency or wavenumber if available. Otherwise, use the PyExoCross program to compute the wavenumbers first, and then reorder the `.trans` file accordingly.
- Get the number of transitions and the maximum wavenumber.

- Get the number and the file size of `.trans` files.

Step 3: Upload the line list files, including states (`.states`), transitions (`.trans`), generated partition function (`.pf`), and other associated files, to the ExoMol server (ExoWeb).

Step 4: Incorporate the new line list information into the master file (`exomol.all` and `exomol.all.json`) and update the master file on the ExoWeb server. The details of the master file is shown in Table 3.4 and Table 4.2.

Step 5: Create `.def` and `.def.json` files by extracting information from the relevant paper and using the provided Jupyter Notebook codes, then upload the definition files to the ExoWeb server.

Step 6: Update the line list information on the ExoMol website.

4.2.3 Update master file

To facilitate the use of the API and provide complete metadata, the database includes a master file (`exomol.all`), freely accessible at https://www.exomol.com/db/exomol.all. This master file references a definition file for each recommended line list. It is important to note that `.def` and `.def.json` files are available for each isotopologue for which ExoMol provides a line list.

The master file itself is assigned a version number in the `YYYYMMDD` format, which corresponds to the most recent update of a definition file. The version dates for individual line lists are specified within the master file. Add the new line list information to the master file in ExoMol format, as specified in Table 3.4. In the appendix, an extract from the current master file in JSON format (`exomol.all.json`) is given in Listing B.1.

4.2.4 Create definition file

If a definition file for the same molecule and isotopologue already exists for other datasets, modify it based on the existing file. The data structure of the definition is described in Table 3.5. Additional details are provided in following lists.

1. Version number in format `YYYYMMDD`

   Use the date when the line list was last updated in the ExoMol database as the version number.

2. InChI and InChIKey of isotopologue

   The International Chemical Identifier (InChI) is a standardized notation developed by the International Union of Pure and Applied Chemistry (IUPAC) [381], with significant contributions from the U.S. National Institute of Standards and Technology (NIST) [382] and the InChI Trust [383]. InChI is a non-proprietary, open-source chemical identifier designed to serve as an IUPAC-approved and endorsed standard for representing chemical structures [384]. The InChIKey is an encoded version of hash codes and compact chemical identifier derived from a source InChI string, supplemented with convenient "flag symbols" for easier use [384]. InChIKey always consists a fixed 27-character string. This compact format makes it a more convenient identifier for internet searches and database indexing. The first block of the InChIKey remains consistent for the same molecular skeleton. The second block represents variations such as isotopic substitutions, stereoconfiguration changes, tautomeric states, and coordination bonding [384].

   The InChIKey can be obtained from the NIST website https://webbook.nist.gov/chemistry/inchi-ser/ by using the InChI string (IUPAC International Identifier). On the NIST website, search using the "IUPAC Identifier" and enter the InChI string. Press "Search" to retrieve the IUPAC Standard InChIKey (note: this should be the InChIKey of the isotopologue). The process for determining the InChI is demonstrated by the following examples in Table 4.1.

Table 4.1: InChI and InChIKey for the isotopologues of molecule HCl.

Isotopologue	Iso-slug	IsoFormula	InChI	InChIKey
$^{1}H^{35}Cl$	1H-35Cl	(1H)(35Cl)	1S/ClH/h1H	VEXZGXHMUGYJMC-UHFFFAOYSA-N
$^{1}H^{37}Cl$	1H-37Cl	(1H)(37Cl)	1S/ClH/h1H/i1+2	VEXZGXHMUGYJMC-NJFSPNSNSA-N
$^{2}H^{35}Cl$	2H-35Cl	(2H)(35Cl)	1S/ClH/h1H/i/hD	VEXZGXHMUGYJMC-DYCDLGHISA-N
$^{2}H^{37}Cl$	2H-37Cl	(2H)(37Cl)	1S/ClH/h1H/i1+2/hD	VEXZGXHMUGYJMC-DQGQKLTASA-N

$^{2}H^{35}Cl$ and $^{2}H^{37}Cl$ are written as $D^{35}Cl$ and $D^{37}Cl$, respectively.

3. Isotopologue mass (Da) and (kg)

   The isotopologue mass is the sum over all isotopic masses. Isotopic mass (in Dalton) can be obtained from the Wikipedia page by searching for "Isotopes of <ISOTOPE_NAME>". For example, searching for "Isotopes of oxygen" provides the isotopic masses for various nuclides such as ^{16}O, ^{17}O, and ^{18}O. UNICONVERTERS.NET website (https://www.unitconverters.net/weight-and-mass/dalton-to-kilogram.htm) can convert unit from Dalton to kilograms.

4.3 New JSON format files

JSON (JavaScript Object Notation) [21] is a lightweight, human-readable data-interchange format [17]. The EXOMOL database newly offers data in JSON format, including master (`exomol.all.json`) and definition (`<ISOTOPOLOGUE>__<DATASET>.def.json`) files. In the appendix, an extract from the current JSON format master file (`exomol.all.json`) is provided in Listing B.1, with keywords defined in Table 4.2. The JSON format master file is available on https://www.exomol.com/db/exomol.all. Listing C.1 displays the JSON format definition file (`.def.json`) for the POKAZATEL $H_2{}^{16}O$ line list [5]. The structure of JSON keywords for the `.def` file is provided in Table 4.3.

4.4 Upload new EXOMOL files

To update the EXOMOL database with new or updated line lists, there are two main processes: uploading line list files to the EXOWEB server and then update the information of the line lists on the EXOMOL website, see Figure 4.1.

4.4.1 Steps for uploading files to the EXOWEB server

Step 1: Accessing the EXOWEB server and file path:

(1) Access the theory cluster (NEWT server):

ssh0.theory.phys.ucl.ac.uk

$ ssh username@ssh0.theory.phys.ucl.ac.uk

(2) Access the EXOWEB server from the NEWT server:

exoweb.projects.phys.ucl.ac.uk

Table 4.2: Keywords used in the JSON format master file (`exomol.all.json`).

Keywords	Meaning
ID	Always the string `"EXOMOL.master"`
version	Version number of master file with format `YYYYMMDD`
num_molecules	Number of molecules in the database
num_isotopologues	Number of isotopologues in the database
num_datasets	Number of datasets in the database
num_species	Number of individual datasets in the database
molecules	List of molecules
| formula∗	Molecule chemical formula
| | num_molecule_names	Number of molecule names listed
| | molecule_names	List of the molecule names
| | num_isotopologues	Number of isotopologues considered
| | linelist	Line list
| | | inchikey	InChIKey of isotopologue
| | | iso_slug	Iso-slug
| | | iso_formula	Formula of isotopologue
| | | dataset_name	Isotopologue dataset name
| | | version	Version number of dataset with format `YYYYMMDD`

*: The keywords with ∗ are available.
Note: The total number of datasets is smaller than the number of individual datasets due to two reasons. Firstly, some individual datasets across different molecules share the same dataset name. Secondly, different isotopologues of the same molecule may also have identical dataset names.

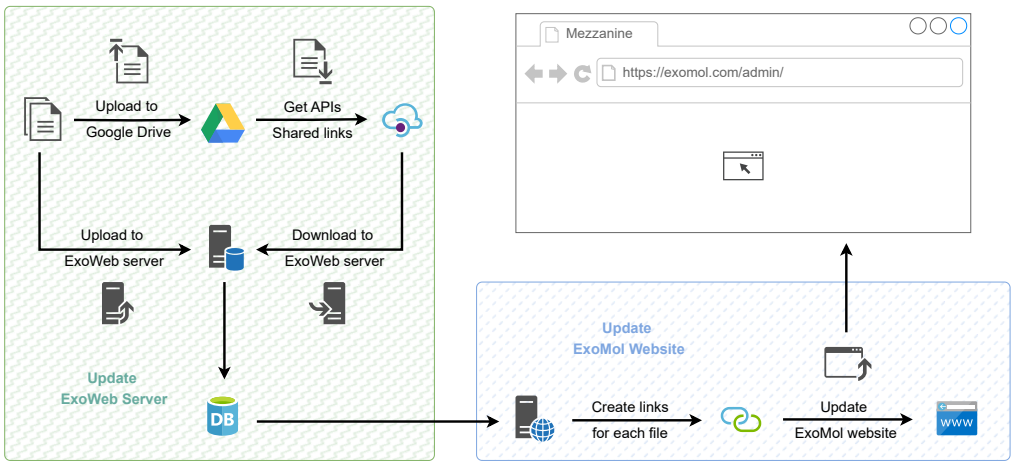


Figure 4.1: Flowchart for updating line list files to the EXOMOL website.

[username@newt ~]$ ssh username@exoweb.projects.phys.ucl.ac.uk

(3) Access the data directory:

[username@exoweb ~]$ cd /mnt/data/exomol/exomol3_data

Table 4.3: Keywords used in the JSON format definition file (`<ISOTOPOLOGUE>__<DATASET>.def.json`).

Keywords	Meaning
isotopologue	Isotopologue section
| iso_formula	Formula of isotopologue
| iso_slug	Iso-slug
| inchikey	InChIKey of isotopologue
| inchi	InChI of isotopologue
| mass_in_Da	Mass of isotopologue (sum of atom masses) in Da
| point_group	Symmetry group
atoms	Atoms section
| number_of_atoms	Number of atoms
| element	Element symbol
| | atom*	In format: Element symbol : Isotope number
irreducible_representations	Irreducible representation section
| label*	In format: Irreducible representation label : Nuclear spin degeneracy
dataset	Dataset line list files section
| name	Isotopologue dataset name
| version	Version number with format `YYYYMMDD`
| doi	DOI of dataset paper
| max_temperature	Maximum temperature of linelist
| num_pressure_broadeners	Number of pressure broadeners available
| nxsec_files	Number of cross section files available
| nkcoeff_files	Number of k-coefficient files available
| dipole_available	Dipole availability (true, false, or null)
| cooling_function_available	Cooling function availability (true, false, or null)
| specific_heat_available	Specific heat availability (true, false, or null)
| continuum	Photo-absorption continuum cross sections availability (true, false, or null)
| predis	Photo-dissociation cross sections availability (true, false, or null)
| states	States file part
| | number_of_states	Number of states in `.states` file
| | max_energy	Maximum energy (in cm^{-1}) in the states file
| | uncertainty_description	Uncertainty description
| | uncertainties_available	Uncertainty availability (true, false, or null)
| | lifetime_available	Lifetime availability (true, false, or null)
| | lande_g_available	Lande g-factor availability (true, false, or null)
| | num_quanta	Number of quanta defined
| | states_file_fields	Columns list in the states file
| | | name	Column name in `.states` file. There are three formats: 1. Main column name 2. Quantum case label : Quantum label 3. Auxiliary : Auxiliary title
| | | desc	Description of column in the states file
| | | ffmt	Fortran format
| | | cfmt	C format
| transitions	Transitions file(s) part
| | number_of_transitions	Total number of transitions in `.trans` file(s)
| | number_of_transition_files	Number of `.trans` file(s)
| | max_wavenumber	Maximum wavenumber (in cm^{-1})
partition_function	Partition function file section
| max_partition_function_temperature	Maximum temperature of partition function
| partition_function_step_size	Step size of temperature
broad	Broadening files section
| default_Lorentzian_half-width	Default value of Lorentzian half-width for all lines (in cm^{-1}/bar)
| default_temperature_exponent	Default value of temperature exponent for all lines
| label*	A label for a particular broadener
| | filename	Filename of particular broadened
| | max_J	Maximum J for which pressure broadening parameters provided
| | Lorentzian_half_width	Value of Lorentzian half-width for J"$> J_{\max}$
| | temperature_exponent	Value of temperature exponent for lines with J"$> J_{\max}$
| | num_quantum_number_sets	Number of defined quantum number sets
| | quantum_number_sets	Defined quantum number sets
| | | code	A code that defines this set of quantum numbers
| | | num_lines	Number of lines in the broad that contain this code
| | | num_quantum_numbers	Number of quantum numbers defined
| | | quantum_numbers	Defined quantum number(s)

*: The keywords with $*$ are available.
true, false, or null: true=yes, false=no, null=undefinded.

In this path, you can find folders for all molecules. The file path for the line list files follows the format: `<MOLECULE>/<ISO-SLUG>/<DATASET>/<LINE_LIST_FILE>`.

Step 2: Uploading a new created definition file (small file) to the EXOWEB server:

(1) Upload the newly created `.def` and `.def.json` files to GitHub.

(2) Open the `.def` and `.def.json` files on the GitHub website, select "Raw". Note: Use the "Raw" link, not the GitHub webpage link. Then copy the links starting with https://raw.githubusercontent.com/....

(3) On the EXOWEB server, download the `.def` and `.def.json` files from the GitHub website using the command:

$ wget https://raw.githubusercontent.com/...

Step 3: Uploading new `.states`, `.trans`, `.pf`, or their compressed files (large files) to the EXOWEB server:

The following are the detailed steps for downloading files from Google Drive using the command line API, assuming the files are privately shared and require identity authentication.

(1) Save the large line list files to Google Drive.

(2) Obtain the File ID: On Google Drive, right-click (or directly click) the file which is required to be uploaded to the EXOWEB server, select "Get Share Link", and extract the file ID from the link (`https://drive.google.com/open?id=XXXX`), where `XXXX` is the file ID and will be used in the following steps.

(3) Obtain the OAuth token:

1). Access the OAuth 2.0 Playground (https://developers.google.com/oauthplayground/).

2). In the "Step 1 Select & authorize APIs" dropdown menu, select "Drive API" and then choose "https://www.googleapis.com/auth/drive.readonly" to authorize using the scope.

3). Click the button “Authorize APIs” and then select “Exchange authorization code for tokens”. Get “Access token” and copy this “Access token” for use in subsequent steps.

(4) Download the Files: Before using the download command on the ExoWeb server, replace `XXXX` with the file ID, `YYYYY` with the “Access token”, and `ZZZZZ` with the saved file name (including the suffix, such as “myfile.zip”). In the ExoWeb server (Linux system), open the terminal and enter the following command:

$ curl -H ”Authorization: Bearer YYYYY” https://www.googleapis .com/drive/v3/files/XXXXX?alt=media -o ZZZZZ

Step 4: Updating or replacing the existing master file (`exomol.all`) on the ExoWeb server (Linux system):

(1) Direct Update: Edit the master file directly using the command:

$ vim exomol.all

Press the “Insert” key to enter edit mode. After editing, press “Esc” key to exit edit mode. Press “Shift” and “:” together, then type “wq” to save the file and exit automatically.

(2) Upload a new master file: The steps are similar to uploading a new definition file (small file).

Step 5: Updating Zenodo of the ExoMol line lists, opacities, and other data products, such as ExoMolHR line lists.

4.4.2 Steps for uploading files to the ExoMol website

Use the Site Administration portal at https://exomol.com/admin/ to add new modules, such as molecules, isotopologues, data sets, links, and more. To do this, click the ‘Add xxx’ button located in the top right corner, as illustrated in Figure 4.2. When entering the information, refer to existing samples for guidance on the details.

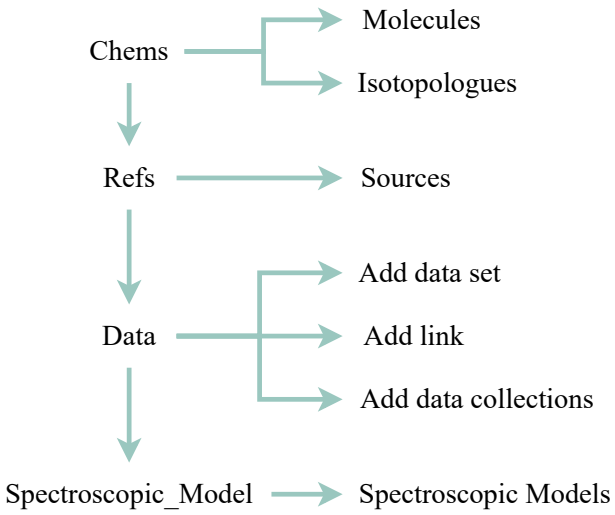


Figure 4.2: The steps for creating new line list files and updating existing line list files to the ExoMol website.

4.5 Dealing with UV processes

The ExoMol database is beginning to play a significant role in modeling UV processes in exoplanetary and stellar atmospheres [17]. UV radiation is crucial for understanding the photochemistry and photodissociation of molecules in these environments. The ExoMol database provides high accurate spectroscopic data for a variety of molecular species, including UV absorption cross sections, which are essential for simulating how molecules interact with UV light.

The database provides the calculation of line lists for UV transitions, which are critical for studying the dissociation, ionization, and chemical evolution of molecular species exposed to UV radiation [17]. These line lists, derived from advanced quantum mechanical models, include absorption data for molecules commonly found in astrophysical environments, enabling the study of their UV photochemistry. By offering these detailed UV data, the ExoMol database allows researchers to accurately simulate the effects of UV radiation on molecular species, advancing our understanding of the role UV processes play in the atmospheres of exoplanets and other celestial objects.

PyExoCross applies Gaussian [385] and lifetime (combined with Loretzian

or Voigt) [386] broadening to make stick spectra to be smooth, then obtains the photodissociation cross sections, more details are given in Section 7.4.3. This can be achieved using user-provided input parameters, allowing for flexible and customized computations.

4.6 Summary

The new JSON format master and definition files for ExoMol database make the conversion of the data to a dictionary with code more convenient for users to obtain to information in program. The continuous computation of additional line lists remains a crucial activity due to the high demand for line list data. Consequently, the regular updating and maintenance of the ExoMol database is essential.

Chapter 5

Preprocessing ExoMol Data

Inputs for PyExoCross (Chapter 6) and ExoMolHR (Chapter 9) require automatic preprocessing data of the ExoMol database with structures that include the master (`exomol.all`), definition (`.def` and `.def.json`), states (`.states`), transitions (`.trans`), and partition function (`.pf`) before the calculations of spectra properties. This preprocessing involves extracting and integrating data from multiple files to be subsequent processed to calculate physical quantities. The methods on data processing discussed in this section are implemented in the PyExoCross and ExoMolHR programs. These two programs operate independently and do not directly interact with each other. However, their underlying principles of data preprocessing methods are similar.

5.1 Processing the master file

To gather the required information, a program requires to read the master file (`exomol.all`) and retrieves details from the first column (information column) using keywords located in the second column (description column). The count and list of molecules, isotopologues, and datasets are crucial for subsequent processes. This process should check both unique and duplicate elements, allowing the program to automatically name, search, download, and save files efficiently.

The latest release of the ExoMol database contains 91 unique molecules, 225 isotopologues, and 107 distinct dataset names. However, the total count of isotopologues and dataset names appears as 270 due to duplication in the master file, where

each dataset name is listed under its corresponding isotopologue. This higher number results from the fact that many molecules have multiple isotopologues, and some isotopologues are linked to more than one dataset. As a result, the total number of species recorded in the ExoMol database is also 270.

To correctly associate molecules with their isotopologues and dataset names, the program first determines the number of isotopologues per molecule. It then duplicates molecules as needed to match the total of 270 entries. This expanded list allows the program to systematically download or read all `.def` files for uncertainty verification.

5.2 Processing the definition file

The `wget` command can be used with the API to download all `.def` files. It is necessary to ensure that the number of `.def` files matches the number of datasets. Reading the `.def` file follows the same process as reading the master file: extract keywords from the second column (description column) to retrieve the corresponding details in the first column (information column). The availabilities of lifetime, Landé g-factor, and uncertainty is essential for processing states files, as these are optional columns in states file. Furthermore, if uncertainty is available, the program should record this `.def` with the molecule, isotopologue (iso-slug and isotopologue formula), and dataset names for the ExoMolHR program.

The API URL for `.def` file is structured with the molecule, iso-slug, and dataset in the following format: http://www.exomol.com/db/<MOLECULE>/<ISO-SLUG>/<DATASET>/<ISO-SLUG>__<DATASET>.def. For instance, the URL of API for the $^{1}H_{2}{}^{16}O$ POKAZATEL `.def` file is http://www.exomol.com/db/H2O/1H2-16O/POKAZATEL/1H2-16O__POKAZATEL.def.

5.3 Processing the states and transitions files

Step 1: Read `.states` and `.trans` files from the exoweb server (ExoMol database cluster) if is permitted or download the files from the ExoMol website by program automatically with `wget` command.

Step 2: To download the corresponding `.states` file and `.trans` files more efficiently, use the API for each molecule which already contained the URLs of compressed `.states` file and `.trans` files. The URL format of API for each molecule line list is https://exomol.com/api/?molecule=<MOLECULE>&datatype=linelist. For example, the URL of API for molecule H_2O is https://exomol.com/api/?molecule=H2O&datatype=linelist.

Step 3: Extract the URLs for compressed `.states.bz2` and `.trans.bz2` files of all selected molecules and isotopologues and record these URLs into a text file named as `api__urls.txt`. All these URLs are in the format of http://www.exomol.com/db/<MOLECULE>/<ISO-SLUG>/<DATASET>/<FILENAME>.bz2 and consist of the molecule, iso-slug, datasets names and wavenumbers, hence, these URLs are crucial for following steps in other Python scripts. <FILENAME> for states files are in format `<MOLECULE>__<ISO-SLUG>__<DATASET>.states.bz2` (e.g.https://www.exomol.com/db/NO/14N-16O/XABC/14N-16O__XABC.states.bz2). Transitions filenames are in format `<MOLECULE>__<ISO-SLUG>__<DATASET>.trans.bz2` (e.g. https://www.exomol.com/db/NO/14N-16O/XABC/14N-16O__XABC.trans.bz2). For large line lists, transitions filenames use the wavenumber range to split the whole transitions into smaller segments, and their are in format `<MOLECULE>__<ISO-SLUG>__<DATASET>__<LOWERWAVENUMBER-UPPERWAVENUMBER>.trans.bz2` (e.g. https://www.exomol.com/db/H2O/1H2-16O/POKAZATEL/1H2-16O__POKAZATEL__00000-00100.trans.bz2).

Step 4: Use the command `wget` instead of the Python script to download compressed `.bz2` files with their APIs into the corresponding folders which are named in the format <PATH>/www.exomol.com/db/<MOLECULE>/<ISO-SLUG>/<DATASET>. For example, /home/username/data/www.exomol.com/db/H2O/1H2-16O/POKAZATEL. The `wget` command makes the downloading process more efficient and can be run in the background. Download files with command in the terminal:

```
wget -d -r -i /save_path/api__urls.txt
```

5.4 Processing the partition function files

The programs offer two approaches for accessing the partition function: either reading from a locally downloaded partition function file or directly retrieving it from the ExoMol website without downloading. The second method is recommended and uses an API-enabled web crawler. The URLs of API for partition function is in the format http://www.exomol.com/db/<MOLECULE>/<ISO-SLUG>/<DATASET>/<ISO-SLUG>__<DATASET>.pf. For instance, http://www.exomol.com/db/NH3/14N-1H3/CoYuTe/14N-1H3__CoYuTe.pf.

5.5 Challenges for processing line lists

I. Processing line lists (stored in `.states` and `.trans` files) involves linking these two file types. This is done by extracting rows from the `.states` file where the state IDs match the upper and lower state IDs in the `.trans` file. If the `.trans` file lacks a fourth column for wavenumbers, the wavenumbers should be calculated using the state energy from the `.states` file corresponding to the upper and lower state IDs.

II. In some `.states` files, state IDs may be discontinuous; the program can reindex these files, assigning new sequential indices as state ID values. To enhance efficiency, both the PyExoCross and ExoMolHR projects use `pandas.DataFrame.isin` instead of using two nested "for" loops (one for the upper state ID and one for the lower state ID).

III. When the programs apply the uncertainty, intensity, and quantum number filters, I use `pandas.DataFrame.isin` again.

IV. Line list filenames follow three different formats. The latest transitions files are named in two specific formats:

   i. Filenames are constructed using the isotopologue and dataset names, ending with the extension `.trans.bz2` (e.g. `14N-16O__XABC.trans.bz2`).

ii. Filenames are formatted to include the isotopologue, dataset, and the wavenumber range (xxxxx-yyyyy), ending with the extension `.trans.bz2` (e.g. `1H2-16O__POKAZATEL__00400-00500.trans.bz2`).

iii. The older version transition files follow the naming format of the latest files but include an extra version number indicated by `v*` (e.g. `14N-16O__XABC_v2.trans.bz2`).

After splitting the filenames, the program determines the final number N_{split} of segments.

Case 1. The first format filename only remains the dataset name (e.g. XABC) and $N_{\text{split}} = 1$.

Case 2. The second format filename only remains the wavenumber range (e.g. 00400-00500) and $N_{\text{split}} = 1$.

Case 3. The third format filename remains two parts: dataset name and version number (e.g. XABC and v2) and $N_{\text{split}} = 2$.

If $N_{\text{split}} = 1$, the program reads these filenames directly. Otherwise, if $N_{\text{split}} = 2$, the program will ignore this file and proceeds to the next file.

V. For molecules with numerous `.trans` files named by wavenumber range, the programs processing files in parallel to enhance efficiency and reduce runtime. In the PyExoCross and ExoMolHR projects, users specify a wavenumber range as an input parameter, allowing the programs to read only the `.trans` files within this specified range rather than processing all files. Users need only input the lower and upper bounds of the interval $[\nu_{\text{min}}, \nu_{\text{max}}]$, rather than providing all filenames as a long list. The programs can split filenames to obtain the left wavenumber values ν_{left} and right wavenumber values ν_{right}. If ν_{left} and ν_{right} meet specific conditions, the corresponding files will be processed, as illustrated in Figure 5.1. The relationships between these wavenumber values

are defined by the following equations:

$$\begin{gathered} \nu_{\rm left} \leq \nu_{\rm min} \leq \nu_{\rm right} \\ \nu_{\rm left} \geq \nu_{\rm min} \quad \text{and} \quad \nu_{\rm max} \leq \nu_{\rm right} \\ \nu_{\rm left} \leq \nu_{\rm max} \leq \nu_{\rm right} \end{gathered} \tag{5.1}$$

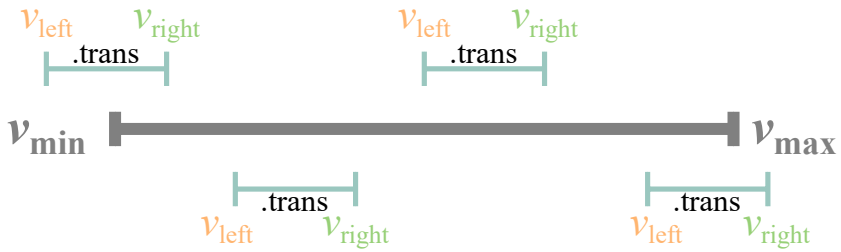


Figure 5.1: The relationship between the wavenumber range in filenames and the input parameters.

5.6 Summary

This chapter focuses on the basic task of my projects which is the data processing from the ExoMol database, which includes various structures such as the master file (`exomol.all`), definition file (`.def`), states file (`.states`), transitions file (`.trans`), and partition function file (`.pf`). The data processing methods discussed are implemented in both PyExoCross and ExoMolHR programs.

Chapter 6

Post-processing ExoMol and HITRAN Data: PyExoCross

ExoMol offers the post-processing capabilities tools via the Fortran program ExoCross [373] (GitHub: https://github.com/ExoMol/ExoCross) and the Python program PyExoCross (GitHub: https://github.com/ExoMol/PyExoCross). Example jobs for both ExoCross and PyExoCross are available on their respective GitHub repositories.

PyExoCross manual (https://pyexocross.readthedocs.io/), which is given to the supplementary data, provides comprehensive working instructions and use case examples.

6.1 ExoCross overview

ExoCross [373], written in Fortran, is efficient for processing large datasets but somewhat limited in flexibility. For instance, incorporating portions of the current ExoCross code as utilities within other applications is challenging. PyExoCross was developed to address this limitation.

ExoCross provides a number of functions, including generating partition functions, specific heats, cooling function, state-resolved radiative lifetimes, non-local thermodynamic equilibrium (non-LTE) spectra, pressure- and temperature-dependent cross sections, electric dipole, electric quadrupole and magnetic dipole spectra [17]. ExoCross can read and output data in both ExoMol and HITRAN

formats, as well as in SPECTRA (http://spectra.iao.ru/), PHOENIX [387], and other output formats for stellar modeling codes.

6.2 Introduction

PYEXOCROSS is a Python-based adaptation of the EXOCROSS [373] Fortran application, developed for post-processing large molecular line lists produced by the EXOMOL project [18] and other similar initiatives like the HITRAN [13, 388] and HITEMP [54, 55] databases.

PYEXOCROSS generates spectra and spectral properties based on underlying spectroscopic data and molecular line lists, including absorption and emission stick spectra (see Section 6.3.6.3), temperature- and pressure-dependent absorption and emission cross sections (see Section 6.3.7), temperature-dependent cooling functions (see Section 6.3.3), state-dependent radiative lifetimes (see Section 6.3.4), oscillator strengths (see Section 6.3.5) and thermodynamic properties such as partition functions (see Section 6.3.1) and specific heats (see Section 6.3.2). PYEXOCROSS calculates cross sections using four line profiles (see Chapter 7): Doppler, Gaussian, Lorentzian, and Voigt line profiles, available in both sampling and binned methods. Multiple options for Voigt profile calculations are provided, tested for speed and accuracy. PYEXOCROSS supports importing and exporting line lists in the EXOMOL, HITRAN, and HITEMP data formats, as well as conversions between the EXOMOL and HITRAN data formats (see Section 6.6).

The PYEXOCROSS program outputs results in the EXOMOL format, which is compatible with the effective Hamiltonian spectral simulation software PGOPHER [389]. The initial line list formats provided are the EXOMOL and HITRAN/HITEMP formats, both known for their extensive wavelength coverage and highly developed data structures. Additionally, PYEXOCROSS includes extra tools to automate batch downloads of line list files from the EXOMOL database.

The primary practical challenge lies in processing the huge amounts of data associated with extremely large datasets for polyatomic molecular hot line lists. Many molecules in the EXOMOL database have multiple transition files. For instance,

the parent isotopologue $^{23}Na^{16}O^{1}H$ of the molecule NaOH has an OYT5 dataset [390], which includes over 400 billion transitions distributed across 90 compressed transition files, with a total size of 463 GB. To overcome this issue, implementing automation and batch data processing are crucial for enhancing efficiency and minimizing errors stemming from operational faults. The PyExoCross program tackles the challenge of handling tens of billions of lines and optimizing for high throughput by employing advanced techniques such as efficient parallelization and vectorization, see Section 6.8.

Opacity in exoplanetary atmospheres is critical for modeling their temperature profiles and radiative transfer. Models like ÆSOPUS 2.1 [391] extend low-temperature opacities to higher pressures, improving atmospheric simulations. The EXOPLINES database [207] provides molecular absorption cross-sections for brown dwarfs and giant exoplanets, aiding in spectroscopic data interpretation. RAPOC [294] offers efficient opacity data for different conditions, helping integrate opacities into atmospheric models. These advancements enhance the understanding of exoplanet and brown dwarf atmospheres. Furthermore, ExoMol and HELIOS-K [392] databases provide extensive opacity data for exoplanetary atmospheres. ExoMolOP[367], part of the ExoMol database, contains high-temperature opacity data for various molecules, crucial for modeling the spectral signatures of exoplanets (Chubb et al., 2021). HELIOS-K [393] complements this by offering opacities across a range of temperatures, supporting more accurate atmospheric models. These resources are integral for refining models of exoplanetary climates, advancing our understanding of their radiative transfer properties.

Several similar programs are available for post-processing line lists, including the HITRAN interface HAPI [377, 394], SPECTRAPLOT.COM [395], and SPECTRA [396, 397]. Nonetheless, these programs would encounter difficulties in processing the extensive line lists necessary for conducting simulations of atmospheres at elevated temperatures. RADIS [398], a fast GPU-accelerated line-by-line computational tool for analyzing infrared absorption and emission spectra under both equilibrium and non-local thermodynamic equilibrium (non-LTE) con-

ditions, is specifically developed for this application. This tool is compatible with ExoMol[17, 18], HITRAN [13, 388] and HITEMP [54, 55] datasets. PyExoCross is designed to import input line lists in the formats utilized by ExoMol [3, 18] and HITRAN [13], subsequently providing output in the ExoMol format. Future enhancements to PyExoCross will include support for additional database formats.

6.3 Main functionality

This section presents an overview of the primary functionalities of the PyExoCross program, which encompasses partition functions, specific heats, cooling functions, radiative lifetimes, oscillator strengths, intensities, stick spectra, and cross sections. Additionally, the program offers supplementary functionalities such as filters, data format conversion, batch downloading of line list files, extraction of uncertainty datasets, and verification of updates to the ExoMol database. The methodology for calculating line profiles relevant to cross sections will be detailed in the next Chapter 7.

PyExoCross provides more flexible quantum number filters, allowing selections based on both labels and values. It also supports data format conversion between the ExoMol and HITRAN data formats, and this feature is not available in ExoCross. PyExoCross can calculate oscillator strengths, but the published ExoCross version does not support this capability [373]. When reading line list files, PyExoCross can directly process compressed `.bz2` files, whereas ExoCross requires decompression before use. Additionally, ExoCross needs to specify all line list filenames in the input file. While PyExoCross only requires defining the wavenumber range, allowing the program to automatically select the relevant line list files.

6.3.1 Partition functions

In physics, the partition function describes the statistical properties of a system in thermodynamic equilibrium [399]. It is a function of thermodynamic state variables, such as temperature and volume. Many aggregate thermodynamic vari-

ables, including total energy, free energy, entropy, and pressure, can be expressed in terms of the partition function or its derivatives. The partition function itself is dimensionless. In astrophysics and molecular spectroscopy, partition functions are important for computing equilibrium constants and modeling spectra In astrophysics and molecular spectroscopy, partition functions are important for computing equilibrium constants and modeling spectra [18]. While they are essential for accurately describing hot molecular systems, their precise determination at elevated temperatures remains challenging [400, 401].

The partition function plays a fundamental role in the calculation of various spectroscopic properties. In PyExoCross, it is necessary for determining specific heats, cooling functions, spectral intensities, and cross sections.

Comprehensive compilations of partition functions for key astrophysical species have been provided by Irwin, Sauval and Tatum, Barklem and Collet, Gamache et al. [402–405]. The partition function, denoted as $Q(T)$, is mathematically defined as the sum over states:

$$Q(T) = \sum_n g_n^{\text{tot}} e^{-c_2 \tilde{E}_n / T}. \tag{6.1}$$

The total degeneracy $g_n^{\text{tot}} = g_n^{\text{ns}}(2J_n + 1)$, incorporates the nuclear-spin statistical weight factor g_n^{ns} and the corresponding total angular momentum J_n. Notably, both ExoMol and HITRAN follow the physicists generally convention, applying the full nuclear spin degeneracy for g_n^{ns}, as discussed in [406]. The second radiation constant $c_2 = hc/k_B$ (in cm·K), is defined using the Planck constant h (in erg·s), the speed of light c (in cm/s), and the Boltzmann constant k_B (in erg/K). The energy term value is expressed as $\tilde{E}_n = E_n/hc$ (in cm^{-1}) and T denotes the temperature (K). PyExoCross offers the capability to calculate the molecular partition function as described by Eq. (6.1), using the energy term values contained within the relevant `.states` file of the specified line list across a range of temperatures.

In spectroscopic applications requiring a partition function $Q(T)$, it is recommended to retrieve it directly from an ExoMol partition function (`.pf`) file. These files often provide partition functions with greater accuracy compared to those ob-

Table 6.1: Specification of the `.pf` partition function file format.

Field	Fortran Format	C Format	Description
T	`F8.1`	`%8.1f`	Temperature in K
$Q(T)$	`F15.4`	`%15.4f`	Partition function (dimensionless)

Fortran format: `(F8.1,1x,F15.4).`

Table 6.2: Extract from the `.pf` partition function file generated from the $^{27}Al^{1}H$ AloHa [11] ExoMol line lists.

T	$Q(T)$
101.0	137.9430
102.0	139.2695
103.0	140.5960
104.0	141.9225
105.0	143.2491
106.0	144.5758
107.0	145.9024
108.0	147.2292
109.0	148.5560
110.0	149.8827

T: Temperature in K.
$Q(T)$: Partition function (dimensionless).

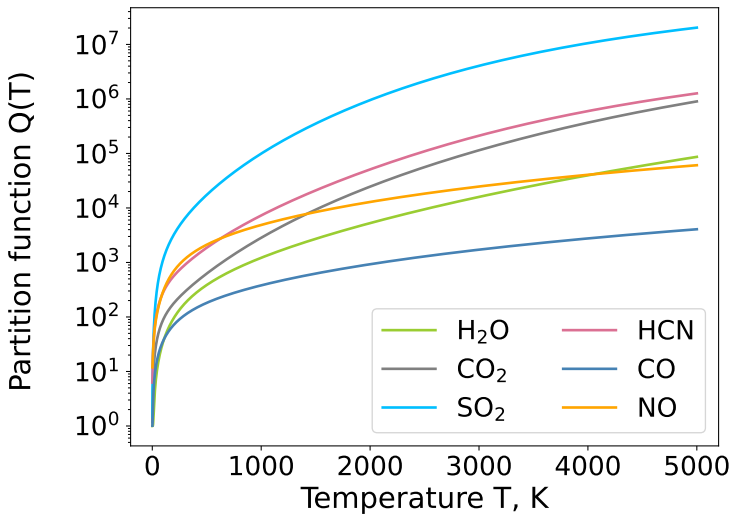


Figure 6.1: Partition functions of $^{1}H_2^{16}O$, $^{12}C^{16}O_2$, $^{32}S^{16}O_2$, $^{1}H^{12}C^{14}N$, $^{12}C^{16}O$ and $^{14}N^{16}O$ with temperatures from 0 K to 5000 K from the ExoMol database.

tained from a straightforward summation over the states in the `.states` file. However, if a `.pf` file is unavailable, $Q(T)$ for a specific temperature can be calculated using Eq. (6.1) as part of the spectral simulation process.

Table 6.1 outlines the format of the partition function file, while Table 6.2 provides an illustrative example. Figure 6.1 presents examples of partition functions computed using the ExoMol line lists over a temperature range of 0 K to 5000 K. The examples here and throughout are based on the following line lists: POKAZATEL dataset of $^{1}H_{2}{}^{16}O$ [5], UCL-4000 dataset of $^{12}C^{16}O_{2}$ [6], ExoAmes dataset of $^{32}S^{16}O_{2}$ [7], Harris dataset of $^{1}H^{12}C^{14}N$ [407], Li2015 dataset of $^{12}C^{16}O$ [408, 409] and XABC dataset of $^{14}N^{16}O$ [12].

PyExoCross offers multiple methods for obtaining the partition function: it can be retrieved from a downloaded local ExoMol `.pf` file or directly crawling from the ExoMol website, computed using the ExoMol `.states` file or accessed via the HITRAN website. It is worth noting that a recent study by Wang et al. [370] indicates that, in many cases, the maximum temperature specified by the HITRAN database for a converged partition function [405] may be overestimated.

6.3.2 Specific heats

In thermodynamics, the specific heat capacity of a substance is defined as the amount of heat required to raise the temperature of one unit mass of the substance by one degree [410]. Also referred to as massic heat capacity or simply specific heat, it represents the heat capacity of a given sample. More formally, it quantifies the thermal energy needed to achieve a unit temperature increase per unit mass, making it a fundamental property in characterizing the thermal behavior of materials [411]. C_p we considered here is the specific heat at constant pressure. Specific heat is important in spectroscopic studies as it characterizes how molecules distribute thermal energy among vibrational, rotational, and translational degrees of freedom [412]. This distribution influences the population of energy levels, which in turn affects line intensities in absorption and emission spectra. Accurate modeling of specific heat is therefore essential for understanding thermal emission, radiative cooling, and interpreting molecular spectra in astrophysical environments.

The molar specific heat at constant pressure, denoted as C_p (in units of J K^{-1} mol^{-1}), for a specified temperature, can be determined using the molecular partition function as described by [413]. This calculation is typically performed on a defined grid of temperatures of 1 K.

$$C_p(T) = R\left[\frac{Q''}{Q} - \left(\frac{Q'}{Q}\right)^2\right] + \frac{5R}{2}. \tag{6.2}$$

The second term corresponds to the translational contribution, as noted by [48]. R represents the molar gas constant, which is equal to 8.31446261815324 J K^{-1} mol^{-1}. The partition function $Q(T)$ is defined in Eq. (6.1). The first and second moments Q' and Q'' of the partition function $Q(T)$ can be evaluated directly from the values associated with the energy terms.

$$\begin{aligned} Q'(T) &= T\frac{\mathrm{d}Q}{\mathrm{d}T} \\ &= \sum_n g_n^{\mathrm{tot}} \left(\frac{c_2\tilde{E}_n}{T}\right) \exp\left(-\frac{c_2\tilde{E}_n}{T}\right), \\ Q''(T) &= T^2\frac{\mathrm{d}^2Q}{\mathrm{d}T^2} + 2Q' \\ &= \sum_n g_n^{\mathrm{tot}} \left(\frac{c_2\tilde{E}_n}{T}\right)^2 \exp\left(-\frac{c_2\tilde{E}_n}{T}\right). \end{aligned} \tag{6.3}$$

PyExoCross uses these equations in conjunction with the molecular `.states` file to calculate specific heats across a grid of temperatures. Table 6.3 delineates the specific heat file format, while Table 6.4 provides an illustrative example of the specific heat file for AlH. Additionally, Figure 6.2 presents the specific heats over the temperature interval of $T \in [200, 10000]$ K with the ExoMol line lists: POKAZATEL dataset of $^{1}H_2{}^{16}O$ [5], UCL-4000 dataset of $^{12}C^{16}O_2$ [6], ExoAmes dataset of $^{32}S^{16}O_2$ [414], Harris dataset of $^{1}H^{12}C^{14}N$ [407], Li2015 dataset of $^{12}C^{16}O$ [408, 409] and XABC dataset of $^{14}N^{16}O$ [12].

Following the work of Wang et al. [370], a nearly comprehensive set of specific heats is now accessible on the ExoMol website.

Table 6.3: Specification of the `.cp` specific heat file format.

Field	Fortran Format	C Format	Description
T	`F8.1`	`%8.1f`	Temperature in K
$C_p(T)$	`F15.4`	`%15.4f`	Specific heat (dimensionless)

Fortran format: `(F8.1,1x,F15.4).`

Table 6.4: Extract from the `.cp` specific heat file generated from the $^{27}Al^{1}H$ AloHa [11] ExoMol line lists.

T	$C_p(T)$
201.0	29.1545
202.0	29.1552
203.0	29.1560
204.0	29.1567
205.0	29.1575
206.0	29.1583
207.0	29.1591
208.0	29.1599
209.0	29.1608
210.0	29.1617

T: Temperature in K.
$C_p(T)$: Specific heat (J K^{-1} mol^{-1}).

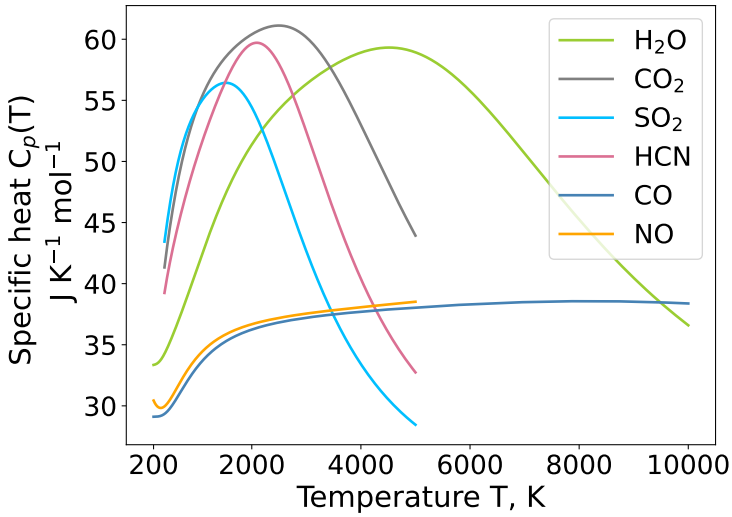


Figure 6.2: Specific heats of $^{1}H_2{}^{16}O$, $^{12}C^{16}O_2$, $^{32}S^{16}O_2$, $^{1}H^{12}C^{14}N$, $^{12}C^{16}O$ and $^{14}N^{16}O$ with temperatures from 200 K to high temperature or T_{max} from the ExoMol database.

6.3.3 Cooling functions

The cooling function is a temperature-dependent quantity that describes the rate at which a gas cools (loses energy) through radiation [415]. It is important in understanding the thermal balance of astrophysical gases, as it influences the emission spectra. In particular, the cooling function governs how gases cool in various environments, such as nebulae, star-forming regions, and stellar atmospheres. A precise understanding of the cooling function allows for accurate modeling of spectral line emissions, which is essential for interpreting the thermal properties and dynamics of these gases. Cooling functions are important across a wide range of astronomical environments, from the early Universe to the interstellar medium, star formation, and challenges related to planetary stability [416–420].

The temperature-dependent cooling function, $W(T)$, is given in erg (s molecule str)$^{-1}$ or equivalently 10^{-7} Watts (molecule str)$^{-1}$, represents the total energy emitted per unit solid angle (measured in steradians) by a molecule at a given temperature T [421], as determined by its emissivity [422]. $W(T)$ can be computed as

$$W(T) = \frac{1}{4\pi Q(T)} \sum_{f,i} A_{fi} hc \tilde{\nu}_{fi} g'_f e^{-c_2 \tilde{E}'_f / T}. \tag{6.4}$$

In Eq. (6.4), $Q(T)$ represents the partition function at temperature T, while g'_f and $\tilde{E}'_f$ denote the upper state degeneracy and energy, respectively, both of which can be extracted from the `.states` file. The parameters A_{fi} and $\tilde{\nu}_{fi}$ refer to the Einstein A-coefficient (s^{-1}) and the transition wavenumber (cm^{-1}), available in the `.trans` file (see Tables 3.8 and 3.9). To compute $W(T)$ over a grid of temperatures, PyExoCross requires the `.states` file, `.trans` files, and partition function file. Although some smaller `.trans` files include wavenumbers as column 4, it is critical to calculate these wavenumber values using the upper and lower state term energies from the `.states` file for accuracy, as specified by:

$$\tilde{\nu}_{fi} = \tilde{E}'_f - \tilde{E}''_i. \tag{6.5}$$

Table 6.5: Specification of the `.cf` cooling function file format.

Field	Fortran Format	C Format	Description
T	`F8.1`	`%8.1f`	Temperature in K
$W(T)$	`ES14.8`	`%14.8E`	Cooling function in erg (s molecule str)$^{-1}$

Fortran format: `(F8.1,1x,ES14.8).`

Table 6.6: Extract from the `.cf` cooling function file generated from the $^{27}Al^{1}H$ AloHa [11] ExoMol line lists.

T	$C_p(T)$
301.0	2.27535005E-15
302.0	2.33450567E-15
303.0	2.39479692E-15
304.0	2.45623840E-15
305.0	2.51884119E-15
306.0	2.58261937E-15
307.0	2.64758583E-15
308.0	2.71375278E-15
309.0	2.78113306E-15
310.0	2.84973952E-15

T: Temperature in K.
$W(T)$: Cooling function in erg (s molecule str)$^{-1}$.

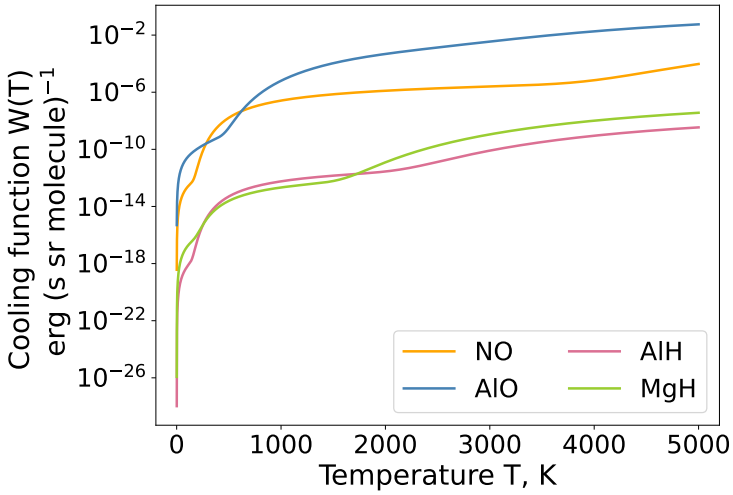


Figure 6.3: The temperature-dependent cooling functions of $^{14}N^{16}O$, $^{27}Al^{16}O$, $^{27}Al^{1}H$ and $^{24}Mg^{1}H$ with temperatures from 0 K to 5000 K from the ExoMol database.

Table 6.5 and Table 6.6 provide the format and an example of a cooling function file, respectively. Figure 6.3 illustrates the temperature-dependent cooling functions for XABC dataset of $^{14}N^{16}O$ [12], ATP dataset of $^{27}Al^{16}O$ [423], AloHa dataset of $^{27}Al^{1}H$ [11], and XAB dataset of $^{24}Mg^{1}H$ [4], derived from the ExoMol database.

6.3.4 Radiative lifetimes

The thermalization of a species or a specific state of that species is determined by the critical density, which is the ratio of the radiative lifetime of the state to the collisional excitation rate to that state [371]. Below the critical density, radiative processes become dominant. Radiative lifetimes influence the determination of the critical density and serve as a competing mechanism to collisional thermalization in various environments, including plasmas such as the interstellar medium. Lifetime is also used to test dipoles in ExoMol spectral models.

The radiative lifetime, τ_i (in s) of state i can be computed by taking the reciprocal of the sum of the Einstein A-coefficients A_{fi} (in s^{-1}) corresponding to transitions where $\tilde{E}_j < \tilde{E}_i$, as outlined in the `.trans` file(s) (refer to Table 3.9) [371]:

$$\tau_i = \frac{1}{\sum_f A_{fi}}. \tag{6.6}$$

The calculation of the state-dependent lifetime involves a systematic three-step process. Initially, transitions are categorized according to upper state classifications using the `pandas.DataFrame.groupby` function [424]. Each upper state, denoted as f, is associated with one or multiple Einstein A-coefficients, represented as A_{fi}. Secondly, the sum of the Einstein A-coefficients for each upper state is calculated. Finally, the inverse of each sum is computed. In principle, the parameter τ_i can be determined for every state i present in the `.states` file. However, in practice, the `.trans` files are not consistently complete. Consequently, in such cases, the lifetime is regarded as undefined and is represented as `NaN`. Certain lowest energy states, especially in highly symmetric molecules such as methane or H_3^+,

exhibit infinite lifetimes, which are denoted by the symbol "inf". The incorporation of lifetimes into the (recommended) `.states` files is being standardized within ExoMol database, as part of the broader effort to generalize the ExoMol data format to accommodate predissociation lifetimes [374].

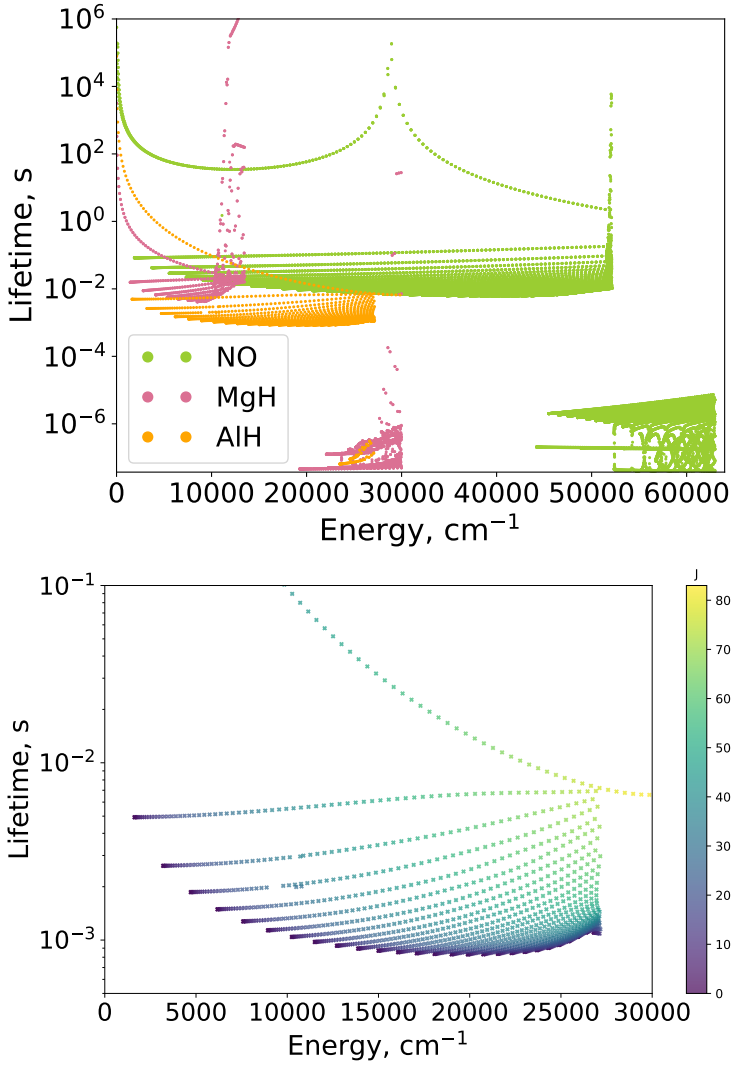


Figure 6.4: Upper panel: Lifetime of ^{14}N^{16}O, ^{24}Mg^{1}H and ^{27}Al^{1}H using the ExoMol line lists. Lower panel: State-dependent lifetime in range of 10^{-4} s - 10^{-1} s of ^{27}Al^{1}H from the ExoMol database. The colors range from purple ($J = 0$) to yellow ($J = 80$).

PyExoCross incorporates lifetimes into the `.states` file following the ExoMol data format. Additionally, it offers users the option to save the states file either as an uncompressed `.states` file or as a compressed `.states.bz2` file. Figure 6.4 (upper panel) displays state-dependent lifetimes for ^{14}N^{16}O [12],

^{24}Mg^{1}H [4], and ^{27}Al^{1}H [11] from the ExoMol database, while the lower panel focuses on lifetimes calculated from the ^{27}Al^{1}H AloHa line list. These lifetimes are fully state-specific.

Recently, a lifetime database (LiDB) was developed based on the ExoMol database for plasma modeling, which organizes lifetimes and decay pathways for vibronic states by aggregating over rotational and fine-structure states [368].

6.3.5 Oscillator strengths

The oscillator strength is a dimensionless parameter used in spectroscopy to quantify the probability of electromagnetic radiation being emitted or absorbed during transitions between the energy levels of an atom or molecule [425]. It is widely used by atomic physicists hence it is also commonly used by the stellar community to represent atomic spectra. The oscillator strength, denoted as f, quantifies the fractional contribution of effective electrons involved in the transition and is directly associated with the line intensity [426].

The statistical weight of an energy level, represented by the symbol g, serves as an indicator of the probability of a transition occurring between two energy levels in an atom or ion. The weighted oscillator strength, denoted as gf, is defined as:

$$gf = \frac{g'_{\rm tot} A_{fi}}{(c\tilde{\nu}_{fi})^2}, \tag{6.7}$$

where $g'_{\rm tot} = g_{\rm ns}(2J'+1)$ represents the degeneracy of the upper state with upper state angular momentum J'. A_{fi} and $\tilde{\nu}_{fi}$ denote the Einstein A-coefficient (s^{-1}) and the transition wavenumber (cm^{-1}), respectively, while c is the velocity of light (cm/s). The actual oscillator strength is defined as

$$f = gf/g''_{\rm tot}, \tag{6.8}$$

where $g''_{\rm tot}$ represents the degeneracy of the lower state. Table 6.7 shows the file format for oscillator strengths, while Table 6.8 provides an example. Figure 6.5 illustrates the oscillator strengths for ^{14}N^{16}O [12], ^{24}Mg^{1}H [4], and ^{27}Al^{1}H [11], as obtained from the ExoMol database.

Table 6.7: Specification of the `.os` oscillator strength file format.

Field	Fortran Format	C Format	Description
f	`I12`	`%12d`	Upper state degeneracy
i	`I12`	`%12d`	Lower state degeneracy
gf	`ES10.4`	`%10.4E`	Oscillator strength
$(\tilde{\nu}_i)$	`F12.6`	`%12.6f`	Central bin wavenumber, cm^{-1} (optional)

Fortran format: `(I12,1x,I12,1x,ES10.4,1x,F12.6).`

Table 6.8: Extract from the `.os` oscillator strength file generated from the ^{14}N^{16}O XABC [12] ExoMol line lists.

f	i	gf	$\tilde{\nu}_i$
2237	2604	2.3938E-29	32406.776335
16945	16990	2.6529E-30	32406.778880
11519	11563	9.3397E-30	32406.843732
15241	14758	4.6102E-26	32406.846700
15986	15778	1.0011E-25	32406.866875
13921	13689	9.0563E-31	32406.888211

f: Upper state ID.
i: Lower state ID.
gf: Oscillator strength.
$\tilde{\nu}_i$: Central bin wavenumber, cm^{-1} (optional).

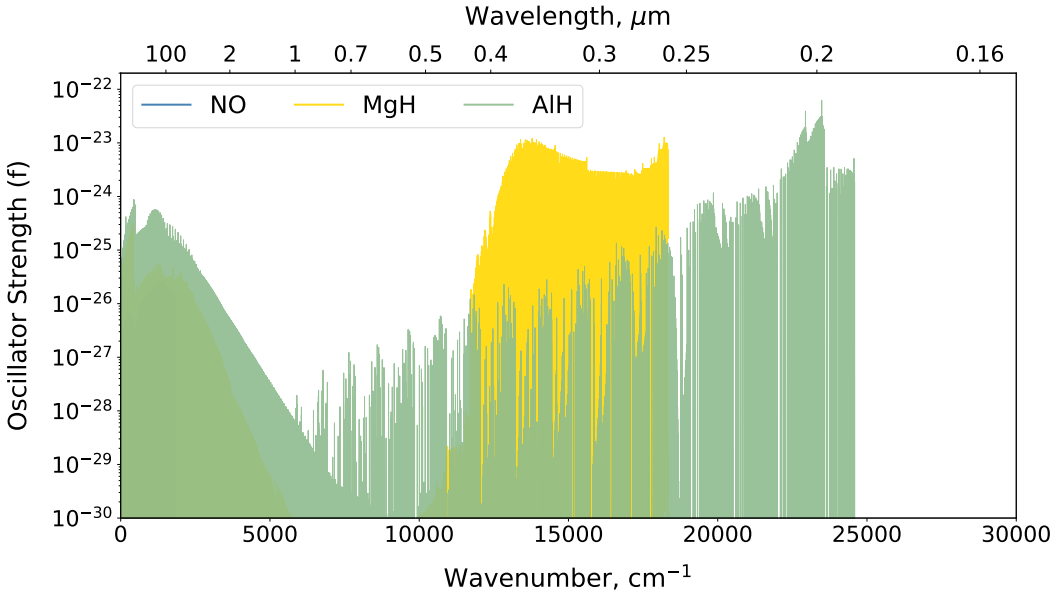


Figure 6.5: Oscillator strength of ^{14}N^{16}O, ^{24}Mg^{16}O, and ^{27}Al^{1}H from the ExoMol line lists.

6.3.6 Line intensities and stick spectra

Measuring the absolute intensity of a spectral line enables the estimation of the density of an atomic and molecular species. Since the line intensity reflects the population of the initial state, applying an appropriate equilibrium model that relates excited state and ground state densities allows for the determination of the total species density [427]. This approach is particularly valuable for determining the concentrations of plasma impurities and for evaluating relative abundances of constituent atoms and molecules in various environments [427].

This section discusses spectral line intensities. PyExoCross is capable of visualizing these intensities as stick spectra, as illustrated in Figure 6.6.

6.3.6.1 Absorption coefficient

The absorption coefficient, also referred to as the absorption line intensity I_{fi} (cm molecule^{-1}) and the line strength $S(T)$ [373]. The line intensities at a general temperature T can be calculated at a given temperature T using the following expression [373, 428]:

$$I_{f\leftarrow i} = \frac{g'_f A_{fi}}{8\pi c \tilde{\nu}_{fi}^2} \frac{e^{-c_2 \tilde{E}''_i/T}\left(1-e^{-c_2 \tilde{\nu}_{fi}/T}\right)}{Q(T)}. \tag{6.9}$$

Here, g'_f denotes the degeneracy of the upper state. $\tilde{E}''_i$ is the lower state energy. A_{fi} represents the Einstein A-coefficient (s^{-1}) and $\tilde{\nu}_{fi}$ is the transition wavenumber (cm^{-1}). $c_2 = hc/k_B$ is the second radiation constant (cm K), where h is the Planck constant (erg s), c is the velocity of light (cm/s), and k_B is the Boltzmann constant (erg/K). T represents the temperature in K. The temperature-dependent partition function, $Q(T)$, can be accessed from the partition function file (`.pf`).

When computing $I_{f\leftarrow i}$ using PyExoCross, the energy term values are obtained from the ExoMol `.states` file, which is also used to calculate $\tilde{\nu}_{fi}^2$. The Einstein A-coefficients are read from the ExoMol `.trans` file. The partition function is either retrieved from the partition function file or dynamically computed using the energy term values and Eq. (6.1), similar to the procedure for cooling functions.

6.3.6.2 Emission coefficient

The emission coefficient, also referred to as the emission line intensity or emissivity ε, is expressed in units of erg (s molecule sr)$^{-1}$, equivalent to 10^{-7} Watts (molecule sr)$^{-1}$ [373].

$$\varepsilon_{i\leftarrow f} = \frac{g'_f hcA_{fi}\tilde{\nu}_{fi}}{4\pi}\frac{e^{-c_2\tilde{E}'_f/T}}{Q(T)}, \tag{6.10}$$

where g'_f represents the upper state degeneracy, and $\tilde{E}'_f$ denotes the upper state energy term value. Similar to the calculation of absorption coefficients, the energy term values $\tilde{E}_f$ are obtained from the ExoMol `.states` file, which is also used to compute $\tilde{\nu}_{fi}$. The Einstein A-coefficients are sourced from the `.trans` file, while the partition function $Q(T)$ is either retrieved from the partition function file or computed using Eq. (6.1).

In the process of calculating the emission coefficient for the HITRAN database, the line lists file in HITRAN format (`.par`) exclusively supplies the energy of the lower state, denoted as E_{lower} (or E''_i). Consequently, the energy of the upper state must be ascertained using Eq. (6.5), while the partition function is obtained from the relevant partition function file.

It is important to note that the definitions of line intensities provided in Eqs. (6.9) and (6.10) assume 100 per cent isotopic abundance. Unlike HITRAN and HITEMP, these values are not scaled to account for isotopic fractional abundances. Therefore, for many applications, the results should be adjusted by multiplying them with the corresponding fractional abundance.

6.3.6.3 Stick spectra

Stick spectra consist of lists of frequencies and absorption or emission line intensities, calculated using Eqs. (6.9) and (6.10), respectively, along with the required quantum numbers for the upper and lower states [373]. PyExoCross allows the application of filters based on uncertainties, thresholds, and quantum numbers when generating stick spectra.

Table 6.9 outlines the essential components of the stick spectra file (`.stick`),

while Table 6.10 provides an example of the stick spectra output file. The spectra are represented graphically as vertical lines, referred to as "sticks", where the height of each line corresponds to its intensity (refer to Figure 6.6).

Table 6.9: Specification of the mandatory part of the `.stick` stick spectra file with the extra columns for the optional quantum numbers.

Field	Fortran format	C format	Description
$\tilde{\nu}$	`F12.6`	`%12.6f`	Frequency (wavenumber) in cm^{-1}
I	`ES14.8`	`%14.8E`	Intensity in cm $\mathrm{molecule}^{-1}$
J'	`I7/F7.1`	`%7d/%7.1f`	J-quantum number (integer/half-integer) for upper state
E'	`F12.4`	`%12.4f`	Upper state energy in cm^{-1}
J''	`I7/F7.1`	`%7d/%7.1f`	J-quantum number (integer/half-integer) for lower state
E''	`F12.4`	`%12.4f`	lower state energy in cm^{-1}
(QN$'$)			Extra quantum numbers for upper state, any format (optional)
(QN$''$)			Extra quantum numbers for lower state, any format (optional)

J: Total angular momentum quantum, excluding nuclear spin.

Table 6.10: Extract from the `.stick` stick spectra file generated from the emission $^{12}C^{16}O_2$ UCL-4000 [6] ExoMol line list.

$\tilde{\nu}_{fi}$	I_{fi}	J'	$\tilde{E}'$	J''	$\tilde{E}''$	G'_{tot}	e/f$'$	n'_1	n'_2	l'_2	n'_3	G''_{tot}	e/f$''$	n''_1	n''_2	l''_2	n''_3
2.14097930E+03	7.19331927E-42	53	16885.7781	52	14744.7988	A2	e	2	4	0	5	A1	e	0	0	0	6
2.14098491E+03	5.34838919E-49	57	17722.9753	56	15581.9904	A1	f	2	1	1	5	A2	f	0	15	11	2
2.14098618E+03	1.14302612E-40	53	17614.1173	53	15473.1311	A2	e	2	1	1	6	A1	f	0	7	3	5
2.14098939E+03	3.59170840E-44	31	15958.9262	31	13817.9368	A2	e	1	1	1	6	A1	f	0	14	6	2
2.14100560E+03	4.79864781E-46	46	16670.4951	45	14529.4895	A1	e	1	1	1	7	A2	e	0	18	8	1
2.14100571E+03	1.83277877E-38	47	17381.6735	46	15240.6678	A2	e	2	1	1	6	A1	e	0	7	3	5
2.14103232E+03	4.72220312E-35	51	12350.8118	50	10209.7795	A2	e	1	9	3	2	A1	e	0	0	0	4
2.14104025E+03	2.03144556E-43	71	16348.4763	71	14207.4361	A2	e	2	1	1	5	A1	f	0	15	11	1
2.14105413E+03	1.20843979E-41	64	13857.7367	64	11716.6826	A1	e	1	1	1	5	A2	f	0	12	10	1
2.14105796E+03	1.98268957E-44	48	16465.9930	49	14324.9351	A2	f	1	1	1	6	A1	f	0	17	7	1

$\tilde{\nu}_{fi}$: Absorption frequency (wavenumber) (in cm^{-1}).
I_{fi}: Absorption intensity (in cm $\mathrm{molecule}^{-1}$).
J' and J'': Total angular momentum quantum number for upper and lower state.
$\tilde{E}'$ and $\tilde{E}''$: Term value (in cm^{-1}) for upper and lower state.
G'_{tot} and G''_{tot}: Total symmetry for upper and lower state.
e/f$'$ and e/f$''$: Rotationless parity for upper and lower state.
n'_1 and n''_1: Normal mode quantum number 1 for upper and lower state.
n'_2 and n''_2: Normal mode quantum number 2 for upper and lower state.
l'_2 and l''_2: Vibrational angular momentum for upper and lower state.
n'_3 and n''_3: Normal mode quantum number 3 for upper and lower state.

6.3.6.4 Intensity thresholds

PyExoCross offers an intensity threshold filter that allows for the exclusion of weak lines. Applying an intensity threshold can enhance the efficiency of cross-section calculations or reduce the output generated by stick spectra computations. By default, the intensity threshold in PyExoCross is set to 10^{-30} cm $\mathrm{molecule}^{-1}$, although this value can be modified in the input file `.inp`. More intensity threshold filter details can be found in Section 6.4.2.

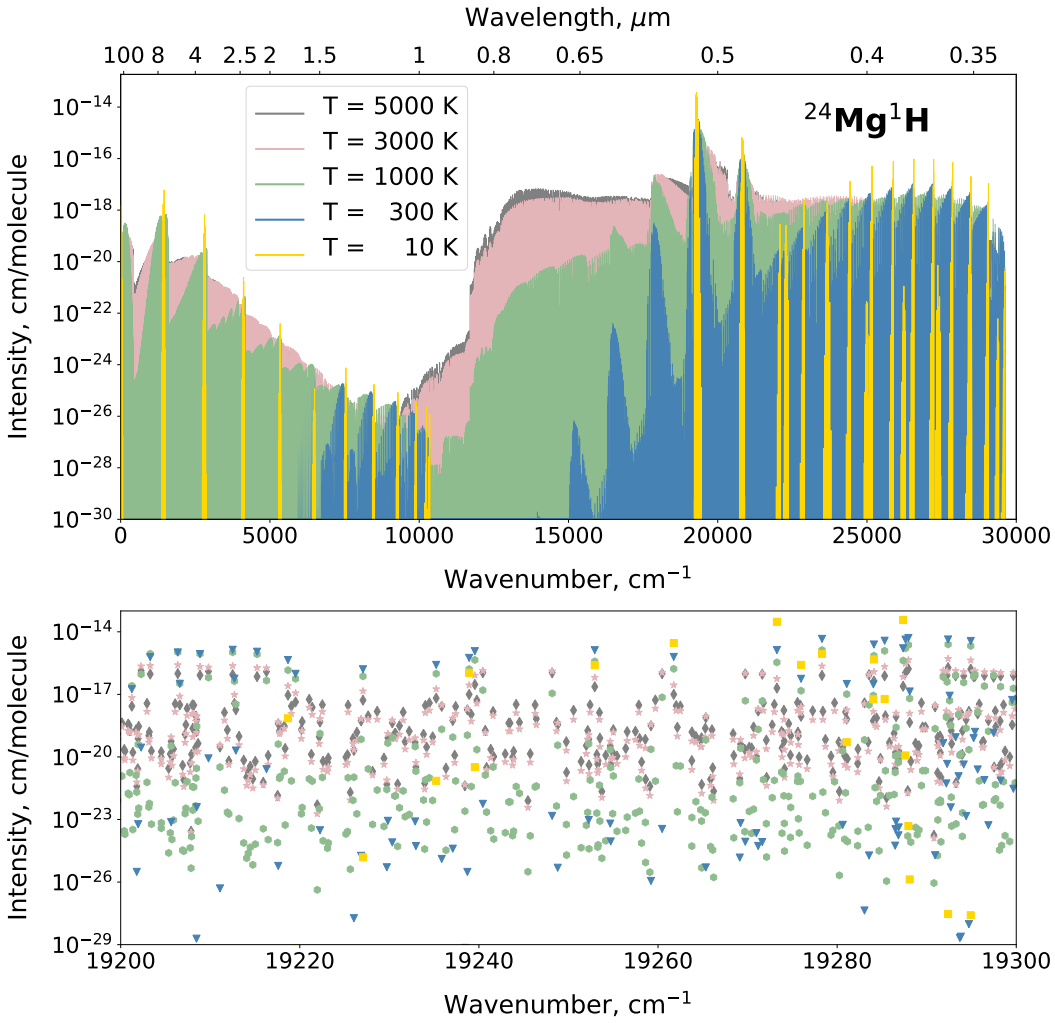


Figure 6.6: Stick spectra: Absorption intensities of ^{24}Mg^{1}H at $T = 10$, 300, 1000, 3000, and 5000 K simulated using the ExoMol line list of dataset XAB [4]. Upper panel: Intensities in wavenumber range $[0, 30\,000]$ cm^{-1}. Lower panel: Intensities zoom at $[19\,200, 19\,300]$ cm^{-1}.

6.3.7 Cross sections

In physics, the cross section quantifies the probability of a specific interaction occurring during the collision of two particles. Cross sections has units of area. The absorption and emission cross sections are the measures of the probability of absorption and emission processes, respectively.

PyExoCross calculates cross-sections (σ_{ab} for absorption or σ_{em} for emission) based on the specified wavenumber or wavelength range, temperature, and pressure, using line intensities (absorption coefficient I_{fi} from Eq. (6.9) or emission coefficient ε_{if} from Eq. (6.10)) and line profiles ($f_{\tilde{\nu}}(\tilde{\nu})$, discussed further in

Chapter 7). The program also supports uncertainty, threshold, and quantum number filters for cross-section computations (see Section 6.4).

6.3.7.1 Grids of wavenumber and wavelength

PyExoCross employs uniform wavenumber (cm^{-1}) and wavelength (μm) grids for cross sections calculations.

The grids for a given wavenumber range $[\tilde{\nu}_{min}, \tilde{\nu}_{max}]$ cm^{-1} can be defined using two methods:

i. Grid points number: Specifying the total number of grid points N_{points}, so $N_{points}-1$ represents the number of grid intervals. Therefore, the wavenumber bin size is determined as:

$$\Delta\tilde{\nu} = \frac{\tilde{\nu}_{max} - \tilde{\nu}_{min}}{N_{points} - 1}. \tag{6.11}$$

To ensure that $\Delta\tilde{\nu}$ is a "round" value, the number of grid points is typically chosen as an odd number [373].

ii. Grid bin size: Defining the size of the grid bin $\Delta\tilde{\nu}$. Then the number of the grid points can be defined as:

$$N_{points} = \frac{\tilde{\nu}_{max} - \tilde{\nu}_{min}}{\Delta\tilde{\nu}} + 1. \tag{6.12}$$

The wavelength grids generated by a wavelength range $[\tilde{\lambda}_{min}, \tilde{\lambda}_{max}]$ μm can be similarly defined by the above two methods with grid points number N_{points} or grid bin size $\Delta\tilde{\lambda}$. The relationship between wavenumber (in cm^{-1}) and wavelength (in μm) can be defined as:

$$\tilde{\nu}_{min} = \frac{10000}{\tilde{\lambda}_{max}} \quad and \quad \tilde{\nu}_{max} = \frac{10000}{\tilde{\lambda}_{min}}. \tag{6.13}$$

i. Grid points number: Specifying the total number of grid points N_{points}, the wavelength bin size is calculated as:

$$\Delta\tilde{\lambda} = \frac{\tilde{\lambda}_{max} - \tilde{\lambda}_{min}}{N_{points} - 1}. \tag{6.14}$$

ii. Grid bin size: When defining the size of the grid bin $\Delta\tilde{\lambda}$, the total number of grid points can be determined as:

$$N_{\text{points}} = \frac{\tilde{\lambda}_{\text{max}} - \tilde{\lambda}_{\text{min}}}{\Delta\tilde{\lambda}} + 1. \tag{6.15}$$

6.3.7.2 Absorption and emission cross sections

The absorption cross section can be interpreted as the effective area surrounding an absorbing atom or molecule that captures energy from an incident light beam. It varies for transitions between different energy levels, reflecting their differing probabilities of absorption.

The absorption cross-section σ_{ab} at a wavenumber $\tilde{\nu}$ for a transition $f \leftarrow i$, characterized by its intensity I_{fi} (as defined in Eq. (6.9)) and line profile $f_{\tilde{\nu}_{fi}}$, is expressed as:

$$I_{fi} = \int_{-\infty}^{\infty} \sigma_{fi}(\tilde{\nu})\mathrm{d}\tilde{\nu} \Rightarrow \sigma_{ab}(\tilde{\nu}) = I_{fi} f_{\tilde{\nu}_{fi}}(\tilde{\nu}). \tag{6.16}$$

The emission cross section is a measure of the probability that a molecule in an excited state emits a photon at a specific wavenumber during a radiative transition. It represents the effective area over which emission occurs and is essential for modeling radiative energy loss in gases and plasmas.

Similarly, the emission cross-section σ_{em} at a wavenumber $\tilde{\nu}$ for a transition $i \rightarrow f$ is determined using the emission coefficient ε_{if} (as defined in Eq. (6.10)) and the line profile $f_{\tilde{\nu}_{fi}}$:

$$\varepsilon_{if} = \int_{-\infty}^{\infty} \varepsilon_{if}(\tilde{\nu})\mathrm{d}\tilde{\nu} \Rightarrow \sigma_{em}(\tilde{\nu}) = \varepsilon_{if} f_{\tilde{\nu}_{if}}(\tilde{\nu}). \tag{6.17}$$

Table 6.11 shows the format of the cross section file `.cross` of the ExoMol database, while Table 6.12 provides an example excerpt from a cross section file. Figures 6.7 and 6.8 illustrate examples of ^{24}Mg^{1}H XAB [4] and ^{1}H$_2$^{16}O POKAZATEL [5] cross sections at $P = 1$ bar and $T = 10$, 300, 1000, 3000, and 5000 K, simulated using the ExoMol line lists. Additionally, Figure 6.9 compares the cross sections of ^{1}H$_2$^{16}O POKAZATEL [5], ^{12}C^{16}O$_2$ UCL-4000 [6], and ^{32}S^{16}O$_2$

ExoAmes [7] at $T = 300$ and 3000 K from the EXOMOL line list.

Table 6.11: Specification of the `.cross` cross-section file format.

Field	Fortran format	C format	Description
$\tilde{\nu}_i$	`F12.6`	`%12.6f`	Central bin wavenumber, cm^{-1}
σ_i	`ES14.8`	`%14.8E`	Absorption cross section, cm^2 $\mathrm{molecule}^{-1}$

Fortran format: `(F12.6,1x,ES14.8)`.

Table 6.12: Extract from the `.cross` cross-section file generated from the $^1H_2{}^{16}O$ POKAZATEL [5] EXOMOL line lists.

$\tilde{\nu}$ (cm^{-1})	σ_i (cm^2 $\mathrm{molecule}^{-1}$)
5000.000000	8.03246668E-25
5000.100000	1.84420150E-24
5000.200000	2.00318098E-23
5000.300000	3.84642766E-24
5000.400000	8.77939177E-25
5000.500000	4.33350935E-25
5000.600000	2.91607795E-25
5000.700000	2.30487176E-25

$\tilde{\nu}$: Central bin wavenumber, cm^{-1}.
σ_i: Absorption cross section, cm^2 $\mathrm{molecule}^{-1}$.

6.3.7.3 A comparative analysis of program run times with PYEXOCROSS, EXOCROSS, and RADIS

I tested the run times of PYEXOCROSS, EXOCROSS, and RADIS programs on 32 cores. PYEXOCROSS efficiently used the cores, taking 371 s to read the compressed `.bz2` files, process data, calculate absorption cross sections, and save results for H_2O at a temperature of $T = 3000$ K, pressure $P = 1$ bar, with a wing cutoff $= 25$ cm^{-1}, uncertainty ≤ 0.01 cm^{-1}, and threshold $\geq 10^{-40}$ cm/molecule, over the wavenumber range from 0 cm^{-1} to 41 200 cm^{-1} using the SciPy Voigt line profile with the EXOMOL $^1H_2{}^{16}O$ POKAZATEL line list [5]. The Fortran program EXOCROSS took 3600 s for the entire process with the uncompressed files under the same conditions, though only about 300 s were spent on actual data processing, with the majority of the time consumed by reading and unpacking the files.

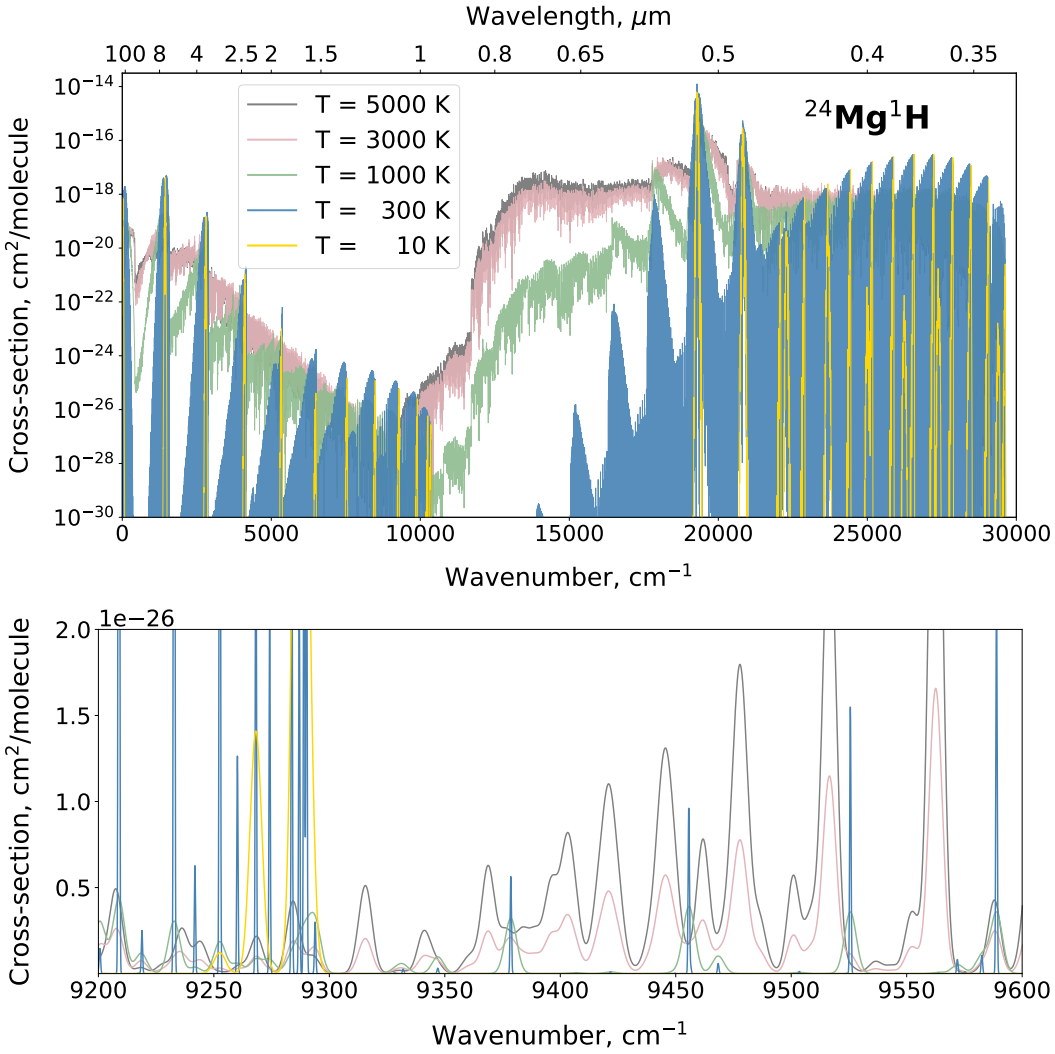


Figure 6.7: Cross-section: Absorption spectrum of ^{24}Mg^{1}H at $T = 10$, 300, 1000, 3000, and 5000 K simulated using the ExoMol line list of the dataset XAB [4] broadened using a Gaussian line profile with its HWHM $\alpha_{\mathrm{D}} = 3$ cm^{-1}. Upper panel: Cross-sections in wavenumber range $[0, 30000]$ cm^{-1}. Lower panel: Cross-sections zoom for $[9200, 9600]$ cm^{-1}.

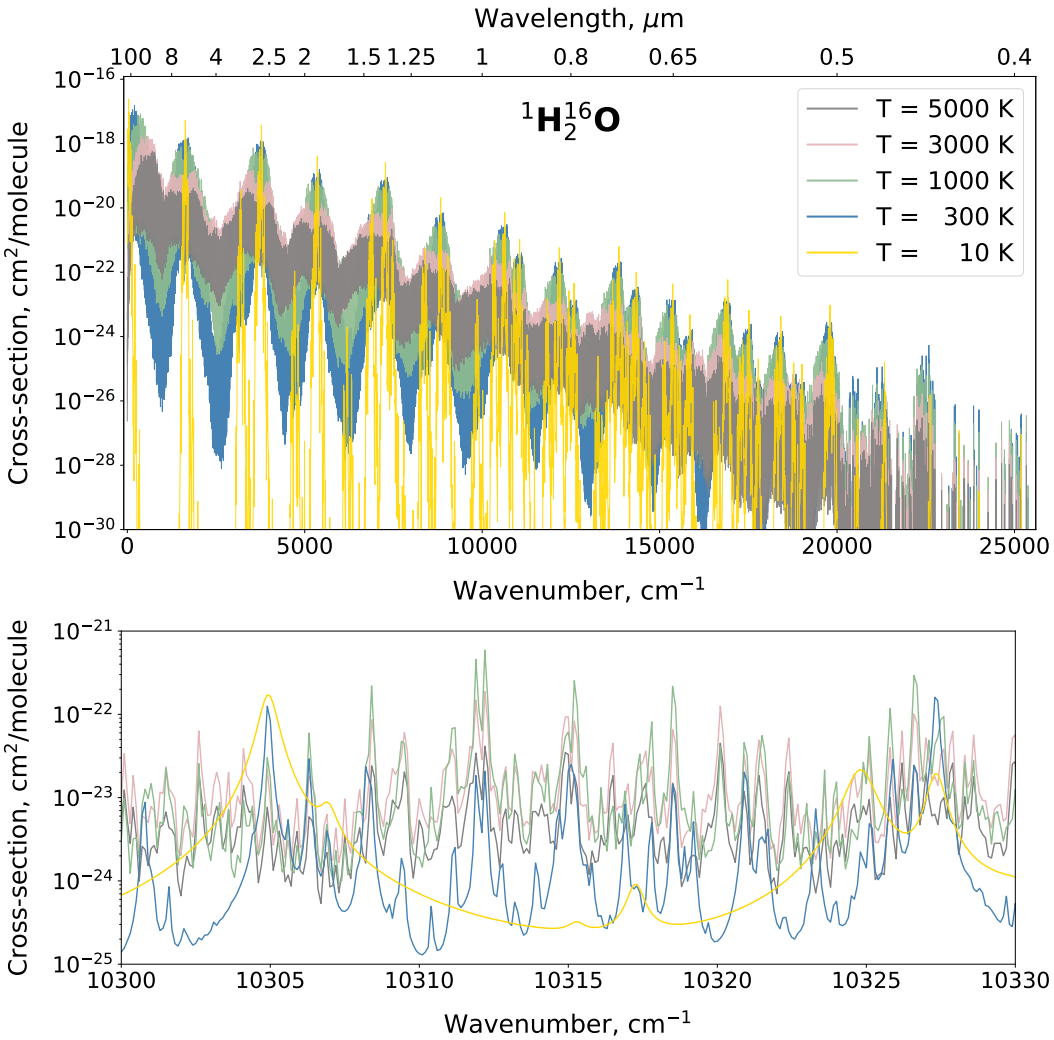


Figure 6.8: Cross-section: Absorption spectrum of $^1H_2{}^{16}O$ at $T = 10, 300, 1000, 3000$, and 5000 K simulated using the ExoMol line list of the dataset POKAZATEL [5] broadened using a SciPy Voigt line profile with the bin size 0.1 cm^{-1} at $P = 1$ bar, HWHM mixed with H_2 and He broadeners and also applied uncertainty and threshold filters (0.01 cm^{-1} and 10^{-30} cm molecule^{-1}). Upper panel: Cross-sections in wavenumber range $[0, 26000]$ cm^{-1}. Lower panel: Cross-sections zoom for $[10300, 10330]$ cm^{-1}.

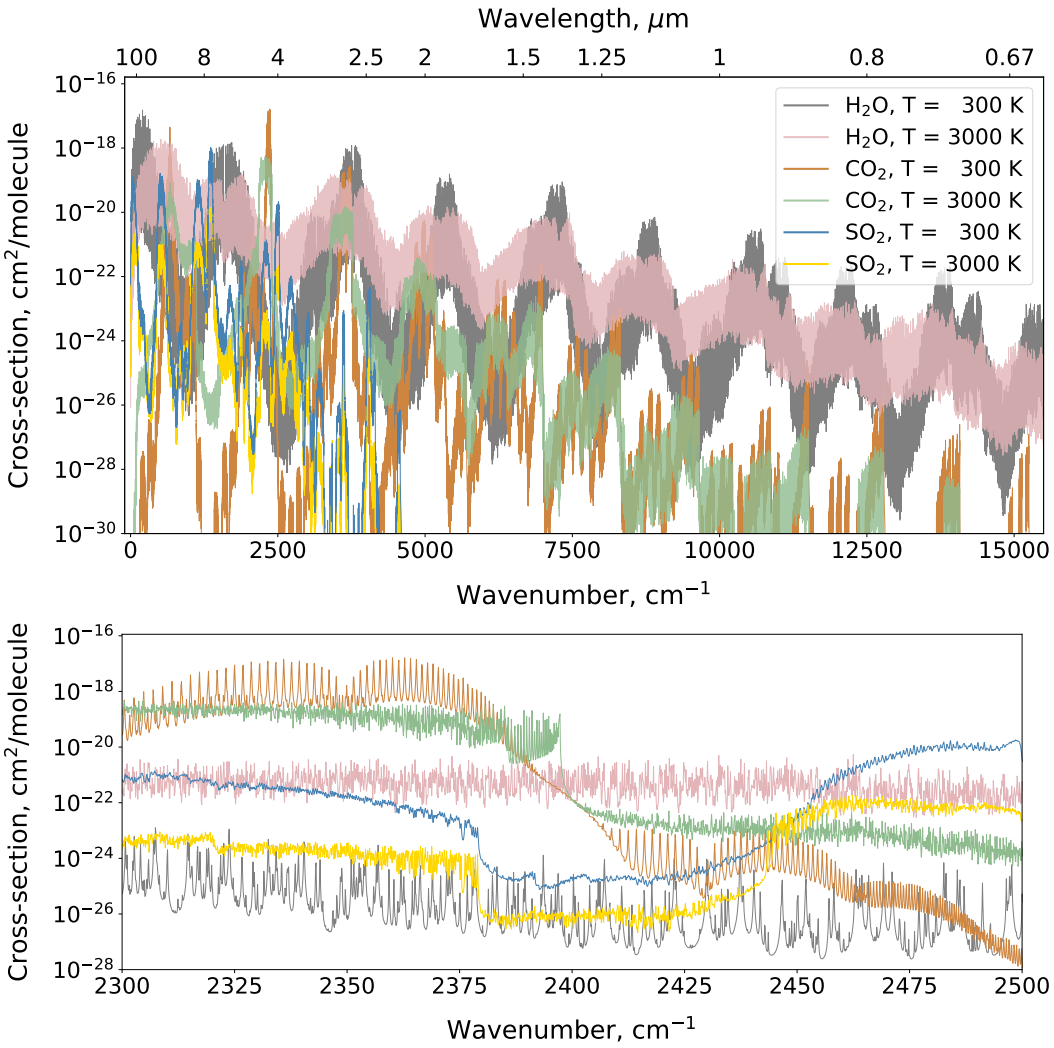


Figure 6.9: Cross-section: Absorption spectrum of $^{1}H_2{}^{16}O$, $^{12}C^{16}O_2$, and $^{32}S^{16}O_2$, at $T =$ 300 and 3000 K simulated using the ExoMol line list of the dataset POKAZATEL [5], UCL-4000 [6], and ExoAmes [7] broadened using a SciPy Voigt line profile with bin size 0.1 cm^{-1} at $P = 1$ bar, HWHM mixed with H_2 and He broadeners and also applied uncertainty and threshold filters (0.01 cm^{-1} and 10^{-30} cm molecule^{-1}). Upper panel: Cross-sections in wavenumber range $[0, 15\,500]$ cm^{-1}. Lower panel: Cross-sections zoom for $[2300, 2500]$ cm^{-1}.

RADIS downloads ExoMol compressed `.bz2` files, converts them into `HDF5` files, and saves both the original and converted files locally for future calculations. Both RADIS and PyExoCross programs I used to calculate the absorption cross sections of the ExoMol $^{1}H_{2}{}^{16}O$ POKAZATEL line list in the wavenumber range [20000,20500] cm^{-1} at a temperature of $T = 300$ K, pressure $P = 1$ bar, a threshold $\geq 10^{-40}$ cm/molecule, and without a wing cutoff using the Voigt line profile. In comparison, PyExoCross took 27 s to read the compressed `.bz2` files, process data, calculate cross sections, and save results. RADIS took 25 s (excluding the time for downloading, decompressing, and converting files) to read `HDF5` files, process data, and calculate cross sections.

6.4 Filters

PyExoCross allows users to apply filters based on the information provided in the `.states` file. These filters include uncertainty, intensity threshold, quantum number label and value, enabling tailored calculations as needed.

6.4.1 Uncertainty filters

An uncertainty filter can be applied in PyExoCross for tasks such as computing stick spectra, cross sections, or format conversions. This filter focuses on the accuracy of individual transition wavenumbers, allowing users to set a maximum acceptable uncertainty value, typically reflecting the upper bound of selected uncertainties.

In laboratory spectroscopy, uncertainties are measured in cm^{-1}. Values below 0.01 cm^{-1} are generally considered highly accurate, while even uncertainties up to 0.05 cm^{-1} are deemed reliable. In contrast, astronomers often describe spectral resolution using the resolving power, $R = \frac{\lambda}{\Delta\lambda}$, where λ is the wavelength, or in practice equivalently, $R \approx \frac{\tilde{\nu}}{\Delta\tilde{\nu}}$ for wavenumbers. Resolving powers R exceeding 100000 are classified as high resolution, though for some astronomy fields, $R >$ 50000 may suffice, see Chapter 2. These considerations allow PyExoCross to be adapted for diverse research needs.

6.4.2 Intensity threshold filters

An intensity threshold filter can be employed in PyExoCross during the computation of stick spectra, cross sections, or when executing format conversions. This filter ensures that only lines with stronger intensities which exceed a specified threshold are processed, allowing users to focus on stronger transitions. By default, the intensity threshold is set to $I > 10^{-30}$ cm molecule^{-1}.

Using an intensity threshold is particularly beneficial for handling molecules with large datasets, as it can significantly reduce computation times. However, it is important to note that excluding weak transitions with $I < 10^{-30}$ cm molecule^{-1} may lead to noticeable changes in the results, especially for applications sensitive to cumulative effects of low-intensity lines [429]. Thus, users should carefully consider the impact of such filters based on the specific requirements of their studies.

6.4.3 Quantum number filters

A quantum number filter can be applied in the calculation of stick spectra and cross sections, allowing the selection of transitions based on specific upper and lower quantum numbers. It is possible to filter by quantum number labels and specify upper and lower quantum number values by extracting some entries from the `.states` file. This feature enables users to focus on transitions that meet particular quantum state criteria.

Additionally, users can specify which quantum numbers to include in the output and define the desired printing format. By applying such filters, researchers can analyze spectra in relation to different quantum labels, facilitating a more targeted and detailed investigation of molecular transitions. Figure 6.10 illustrates the stick spectra of ^{24}Mg^{1}H XAB [4], with the electronic states used as the quantum number filter.

6.5 Plotting options

PyExoCross offers the optional functionality to plot visual figures of stick spectra and cross sections. Users can customize the bottom limit of the y-axis to define the minimum data coordinate displayed in these plots. For the plots of both

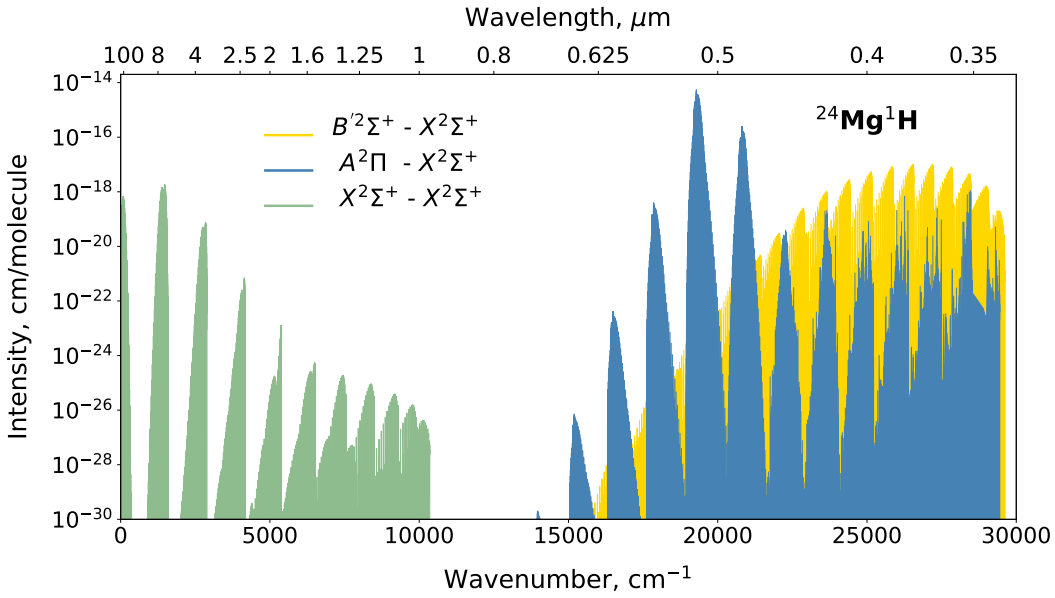


Figure 6.10: Absolute absorption line intensities of different electronic states $B'^2\Sigma^+$ - $X^2\Sigma^+$, $A^2\Pi$ - $X^2\Sigma^+$, and $X^2\Sigma^+$ - $X^2\Sigma^+$ bands of the ExoMol ^{24}Mg^{1}H simulated at $T = 300$ K.

stick spectra and cross sections, the default y-axis lower limit is set to 10^{-30} cm molecule^{-1}, but this value can be adjusted based on user preferences or specific requirements of the analysis.

6.6 Data format conversion

PyExoCross supports input data in both ExoMol and HITRAN formats and provides a flexible conversion of data format between the ExoMol and HITRAN databases. This flexibility facilitates seamless integration and analysis across different datasets. Table 6.13 provides an overview of all file types available in the ExoMol database, while the specific file types supported by PyExoCross are described in detail in the preceding sections.

6.6.1 Convert data format from ExoMol to HITRAN

PyExoCross facilitates the conversion of data from the ExoMol format to the HITRAN format (refer to Tables 3.10 and 3.11). This process involves several steps:

Step 1: Load the `.states` file from the ExoMol database.

Table 6.13: Specification of the file types related to the PyExoCross (contents in brackets are optional).

File extension	PyExoCross available	N_{files}	File DSname	Contents
		ExoMol data format		
`.all`	Input	1	Master	Single file defining contents of the ExoMol database.
`.def`	Input	1	Definition	Defines contents of other files for each isotopologue.
`.states`	Input/Output	1	States	Energy level, quantum number, lifetime, (Landé g-factor, uncertainty).
`.trans`	Input/Output	[a]	Transitions	Einstein A coefficient, (wavenumber).
`.broad`	Input/Output	[b]	Broadening	Parameters for pressure-dependent line profile.
`.pf`	Input/Output	1	Partition function	Temperature-dependent partition function, (cooling function).
`.cp`	Output	1	Specific heat	Temperature-dependent specific heat.
`.cf`	Output	1	Cooling function	Temperature-dependent cooling function.
`.os`	Output	1	Oscillator strength	Oscillator strength, (wavenumber).
`.cross`	Output	[c]	Cross sections	Temperature or temperature and pressure-dependent cross section.
		HITRAN data format		
`.par`	Input/Output	1	Line list file	HITRAN line list file.

N_{files} total number of possible files for each species.
[a] There is a single `.trans` file, but for molecules with large numbers of transitions, it is subdivided by wavenumber regions.
[c] There are N_{cross} sets of `.cross` files for each species.
[c] There are N_{broad} sets of `.broad` files for each species.

Step 2: Process the ExoMol `.trans` files in chunks.

Step 3: Read the partition function file from ExoMol database.

Step 4: Calculate the line intensities for each transition while applying the specified filters of uncertainty, intensity thresholds, and quantum numbers.

Step 5: Convert the exact values of the energy level uncertainties to be the HITRAN uncertainty error codes, see Table 6.14.

Step 6: Output the transition descriptors in HITRAN format, including error codes corresponding to the uncertainties and other relevant details.

To perform the format conversion, users have to provide the HITRAN molecule and isotopologue IDs. If the molecule is present in the HITRAN database, PyExoCross retrieves the necessary isotopic abundance from HITRANonline (https://hitran.org). Otherwise, it defaults to using an abundance of 1 for intensity calculations.

PyExoCross extracts the Einstein A-coefficients from the ExoMol `.trans` files, along with the lower state energy and the total degeneracies of the upper and lower states from the ExoMol `.states` file. Additionally, it supports retrieving γ_{self} and n_{air} values from the `.self` and `.air` broadening files

if available. If these files are absent, default values of $\gamma_{\rm self} = 0.07$ cm^{-1}/atm and $n_{\rm air} = 0.5$ cm^{-1}/atm are applied.

When processing error codes, PyExoCross only considers the first two error codes in HITRAN format, which correspond to the uncertainties in line position and intensity. The remaining four integer fields are set to zero. The error codes for line position uncertainties are detailed in Table 6.14. For intensity error codes, a default value of "0" is assigned to all states since the ExoMol database does not define these values. Table 6.15 provides further details on how quantum number formats are converted.

Table 6.14: The uncertainty codes used by the HITRAN database [13].

Uncertainty code	Absolute energy level uncertainty (ΔE) range (in cm^{-1})
0	$1 \leq \Delta E$ or Unreported
1	$0.1 \leq \Delta E < 1$
2	$0.01 \leq \Delta E < 0.1$
3	$0.001 \leq \Delta E < 0.01$
4	$0.0001 \leq \Delta E < 0.001$
5	$0.00001 \leq \Delta E < 0.0001$
6	$0.000001 \leq \Delta E < 0.00001$
7	$0.0000001 \leq \Delta E < 0.000001$
8	$0.00000001 \leq \Delta E < 0.0000001$
9	$0.000000001 \leq \Delta E < 0.00000001$

Table 6.15: Conversion of the quantum number values between the ExoMol and HITRAN database.

ExoMol database	HITRAN database			
Description	Total symmetry	Vibrational symmetry	Rotational Symmetry	Rotational parity (0 or 1)
Label	$G_{\rm tot}$	$G_{\rm vib}$	$G_{\rm rot}$	τ_i
0				s
1	A1′	A1′	A1′	a
2	A2′	A2′	A2′	
3	E′	E′	E′	
4	A1″	A1″	A1″	
5	A2″	A2″	A2″	
6	E″	E″	E″	

PyExoCross offers enhanced flexibility, allowing users to select preferred quantum number labels and their formats for storage in the HITRAN format. Users can specify global and local quantum number labels and their corresponding formats in the input file, enabling customization of quantum number representations to com-

ply with the 60-character limit of the HITRAN database format. If the total number of characters provided for quantum numbers is less than 60, PyExoCross automatically adjusts the formatting to ensure the output meets the exact 60-character requirement.

Currently, PyExoCross does not support the inclusion of air pressure-induced line shifts (δ_{air}), reference indices (I_{ref}), or their associated uncertainty flags. These fields are left blank in output file.

6.6.2 Convert data format from HITRAN to ExoMol

In principle, PyExoCross can convert HITRAN line list files (`.par`) into the ExoMol format line list files (`.states` and `.trans`). However, this process is not straightforward due to the lack of guaranteed self-consistency between the tabulated lower-state energies and transitions in the HITRAN data model. The primary challenge lies in constructing the ExoMol `.states` file and defining the associated state IDs. To address this, PyExoCross follows a series of steps to facilitate the conversion, including saving the converted ExoMol format states file (`.states`), transitions file (`.trans`), and broadening file (`.broad`). Additionally, PyExoCross also calculates the partition function file (`.pf`) based one the converted ExoMol format `.states` file.

6.6.2.1 Construct the ExoMol states file (`.states`)

The conversion process from HITRAN `.par` file to ExoMol format `.states` file involves the following steps:

Step 1: Read HITRAN Line List: Load the HITRAN `.par` file.

Step 2: Calculate upper state energies: For each line, compute the upper state energy ($\tilde{E}'_i$) using the wavenumber ($\tilde{\nu}_j$) and lower state energy ($\tilde{E}''_j$) based on Eq. (6.5).

Step 3: Convert error codes: Transform the uncertainty error codes into the upper bounds of their respective absolute energy level uncertainty ranges.

Step 4: Process quantum numbers: Extract quantum numbers, including angular momentum values F' and F'', from the HITRAN data.

Step 5: Split and organize data: Separate the HITRAN data into upper and lower state data sets, referred to as `DataFrames` in Python. Each `DataFrame` includes columns for Einstein A-coefficients (A), statistical weights (g), uncertainties, state energies ($\tilde{E}$), global and local quantum numbers. These two `DataFrames` are then concatenated, ensuring consistent the same column structures.

Step 6: Merge levels: Create a new `DataFrame` by merging levels with the same statistical weights ($g_{\rm tot}$), state energies ($\tilde{E}$), uncertainties, and quantum numbers. During this process, retain the lines with the highest uncertainty when multiple entries share the same quantum numbers and statistical weights.

Step 7: Assign state IDs: Sort the data by increasing state energies and assign continuous, incrementing integers as the state IDs.

Step 8: Generate ExoMol states file (`.states`): Create the ExoMol states file (`.states`) in the ExoMol data format. Each entry includes the state ID (i), state energy ($\tilde{E}$), total degeneracy ($g_{\rm tot}$), angular momentum (J or F), uncertainties (ΔE), and the associated quantum numbers, as detailed in Table 3.6.

6.6.2.2 Construct the ExoMol transitions file (`.trans`)

The steps for converting a HITRAN `.par` file into an ExoMol format `.trans` file are as follows:

Step 1: Reconstruct state IDs: For each line in the HITRAN `DataFrame`, determine the corresponding upper and lower state IDs by matching with the previously constructed state `DataFrame`.

Step 2: Calculate transition frequencies: Compute the transition frequencies using the energy difference between the upper and lower states, see Eq. (6.5).

Step 3: Generate ExoMol transitions file (`.trans`): Write the ExoMol transitions file (`.trans`), including the upper and lower state IDs, Einstein A-coefficient (A), and the transition wavenumber (ν), following the format specified in Table 3.8.

6.6.2.3 Constructing an ExoMol air broadening file (`.air`)

The concept of the "diet" algorithm comes from HITRAN [430]. This approach is similar to how doctors customize diets for individual patients, considering factors like metabolism and activity level [430]. In this context, "diet" refers to a tailored method for adding air-broadened half-widths of water-vapor lines to the HITRAN database, optimizing data based on specific transition characteristics [430].

The J-diet refers to the distribution of rotational quantum states (J) in a molecule, which is essential for understanding molecular transitions [17]. In both the ExoMol and HITRAN databases, the J-diet describes how rotational states are populated under specific conditions, influenced by temperature and pressure [13, 17]. This distribution affects line strengths and intensities in spectroscopic data, with higher J-states corresponding to higher energy levels. Accurate modeling of the J-diet is crucial for simulations of molecular spectra, particularly in exoplanet and stellar atmosphere studies.

The conversion process from HITRAN `.par` file to ExoMol format `.air` file includes the following steps:

Step 1: Combine data and remove duplicates: Merge the initial HITRAN `DataFrame` with the γ_{air}, n_{air}, and J'' columns, then drop duplicate rows to ensure unique entries.

Step 2: Generate air broadening file (`.air`): Populate "a0" into the first column, labeled "code". Save code, γ_{air}, n_{air}, and J'' columns into an ExoMol air broadening file (`.air`) following the ExoMol data format (see Table 6.16). This approach assumes a relatively simplified a0 J-diet for line broadening, as described by [14].

6.6.2.4 Constructing an ExoMol self broadening file (`.self`)

The procedure for converting a HITRAN `.par` file into the ExoMol format `.self` file encompasses several steps:

Step 1: Combine data and remove duplicates: Merge the initial HITRAN `DataFrame` with the γ_{self}, n_{air}, and J'' columns, then remove duplicate rows to ensure unique entries.

Step 2: Generate self broadening file (`.self`): Fill "a0" into the first column, labeled "code". Save code, γ_{self}, n_{air}, and J'' columns into an ExoMol self broadening file (`.self`) following the ExoMol data format. For more details on the line broadening codes, refer to Table 6.16.

Table 6.16: Specification of the mandatory part of the `.broad` broadening file format.

Field	Fortran format	C format	Description
code	`A2`	`%2s`	Code identifying quantum number set following J
γ_{ref}	`F6.4`	`%6.4f`	Lorentzian half-width at reference temperature and pressure in $\text{cm}^{-1}/\text{bar}$
n_{L}	`F6.3`	`%6.3f`	Temperature exponent
J"	`I7/F7.1`	`%7d/%7.1f`	Lower J-quantum number integer/half-integer

Fortran format: J integer: `(A2,1x,F6.4,1x,F6.3,1x,I7)`
or J half-integer: `(A2,1x,F6.4,1x,F6.3,1x,F7.1)`.

Table 6.17: Extract from the `.broad` broadening file of the H_2 broadening file `14N-1H3__H2.broad` [14].

Code	γ_{ref}	n_{L}	J''
a0	0.0908	0.583	0
a0	0.0935	0.559	1
a0	0.0926	0.527	2
a0	0.0850	0.508	3
a0	0.0800	0.529	4
a0	0.0768	0.506	5
a0	0.0728	0.501	6
a0	0.0698	0.506	7
a0	0.0671	0.514	8
a0	0.0655	0.500	9

Code: Code identifying quantum number set following J.
γ_{ref}: Lorentzian half-width at reference temperature and pressure in $\text{cm}^{-1}/\text{bar}$.
n_{L}: Temperature exponent.
J'': Lower J-quantum number integer/half-integer.

PyExoCross is also designed to convert data from the standard 160-character HITRAN format into the ExoMol format. However, due to the limitations discussed earlier, the resulting files should be carefully reviewed. With the inclusion of additional broadening parameters in the HITRAN database, such as those for H_2, He, CO_2, and H_2O, users can replace the air and self-broadening columns with these expanded data as needed. The quantum numbers are formatted in accordance with Table S1 and Table S2 from the HITRAN 2020 supplementary material [431]. Since the quantum numbers are written in the same format as in the HITRAN `.par` file, there is no need for further specification or adjustment of their format during conversion.

6.7 Additional functionality

PyExoCross also offers additional features to enhance user convenience:

1. Locating line lists with uncertainties: Identifies line lists in the ExoMol database where uncertainties are specified in the `.state` file.

2. Checking for updates: Compares the current line lists with the latest master file version or the version previously saved by the user, ensuring access to the most up-to-date data.

3. Streamlined file downloads: Allows users to bulk download states and transitions files directly and automatically via the ExoMol API, eliminating the need to manually download files individually.

6.8 Optimize by parallelization and vectorization

Parallelization and vectorization are key techniques for optimizing performance in Python. Parallelization splits tasks into smaller ones that run simultaneously on multiple CPU cores, improving efficiency for independent computations. Libraries like `multiprocessing` [432], `concurrent.futures` [433], and `Dask` [434] facilitate this. Vectorization, on the other hand, applies operations to entire arrays at once, avoiding explicit loops and leveraging optimized low-level

implementations like `NumPy`'s C-based operations [435]. This speeds up numerical tasks significantly. While parallelization is useful for CPU-bound tasks, vectorization is ideal for data processing and mathematical operations.

`Pandarallel` [436] is a Python library designed to accelerate computations in `Pandas` [424] by enabling parallel processing across multiple CPU cores.

`concurrent.futures` [433] library in Python provides the `ThreadPoolExecutor` wrapper, which abstracts the process of thread creation, execution, and termination, thereby simplifying multi-threaded programming.

`NumPy` [435] and `Pandas` [424] enable vectorization by performing operations on entire arrays or `Series` rather than iterating over individual elements. `NumPy` provides the underlying array-processing capabilities, while `Pandas` leverages these optimizations for efficient data manipulation.

`NumExpr` [437], a high-performance numerical expression evaluator for `NumPy`, further enhances computational efficiency. To accelerate mathematical function evaluations, `NumExpr` leverages the optimized Vector Mathematical Function Library (VML). This enables faster computation of vectorized mathematical functions, such as sine, exponential, and square root, by operating on contiguous memory-stored vectors. Compared to using `NumPy` alone, `NumExpr` achieves this performance boost through several key optimizations, particularly when processing large arrays. At first, it executes expressions as bytecode within a virtual machine, fully utilizing the CPU's branch prediction capabilities, making it faster than traditional Python expressions. Secondly, it employs SIMD (Single Instruction, Multiple Data) technology to significantly improve computational efficiency by applying the same operation to multiple data points simultaneously. Thirdly, `NumExpr` supports multi-core parallel computing, decomposing tasks into multiple subtasks that run concurrently across multiple CPU cores, further enhancing execution speed. Finally, efficient memory usage is another advantage of `NumExpr`. Unlike `NumPy`, which generates intermediate arrays during computations, `NumExpr` loads only small portions of data as needed. This approach significantly reduces memory consumption and enhances overall performance.

6.9 Validation of the PyExoCross

PyExoCross has been systematically tested across multiple molecules using different states file formats from the ExoMol database and some molecular line list files with HITRAN standard format from the HITRAN database. These tests ensure that the program can handle diverse file structures and resolve potential issues arising from format variations. To validate its accuracy, PyExoCross compares computed results for the same molecules across the ExoMol and HITRAN databases, ensuring consistency.

For properties such as partition functions, which are available in both databases, PyExoCross results are benchmarked against existing values to verify correctness. Cooling functions and specific heats, which are only available for a limited number of molecules in the ExoMol database, are validated against these existing datasets. Additionally, for properties implemented in ExoCross, PyExoCross performs direct comparisons to confirm its computational accuracy.

To assess the reliability of data format conversions between the ExoMol and HITRAN databases, PyExoCross generates spectra and other properties from both original and converted files, comparing the results to verify both format integrity and computational consistency. These comprehensive validation steps ensure the stability and accuracy of PyExoCross in processing spectroscopic data.

6.10 Calculation protocol

The PyExoCross program processes line lists with the following steps and the workflow can be seen in Figure 6.11).

Step 1: Populate the input file with the required data and settings to be processed by PyExoCross.

Step 2: Download data from the ExoMol and/or HITRAN websites if not already available locally.

Step 3: Read line list files: ExoMol (`.states`, `.trans`, `.pf`, `.broad`) or HITRAN (`.par`) files along with partition function data.

Step 4: Compute the partition function (if required).

Step 5: Compute specific heat (if required).

Step 6: Compute cooling function (if required).

Step 7: Compute radiative lifetime (if required).

Step 8: Compute oscillator strengths (if required).

Step 9: Apply filters based on uncertainties or quantum numbers to intensities (if specified).

Step 10: Compute line intensities (absorption or emission coefficients, LTE or non-LTE) with or without intensity thresholds filter (if specified).

Step 11: Compute line profiles (if required).

Step 12: Calculate cross sections on a specified grid of wavenumbers or wavelengths (if applicable).

Step 13: Save outputs such as partition functions, specific heats, cooling functions, radiative lifetimes, oscillator strengths, intensities, and cross sections to separate files.

Step 14: Plot stick spectra or cross sections (if requested).

Step 15: Provide a report on the program runtime.

6.11 Units

The default labels and units used in the PyExoCross program are detailed in Table 6.18. The physical and astronomical constants applied in the calculations are provided in Listing 1. These constants are sourced from `astropy.constants`, a sub-package of `Astropy`, which is a widely-used community Python package designed for astronomy [438].

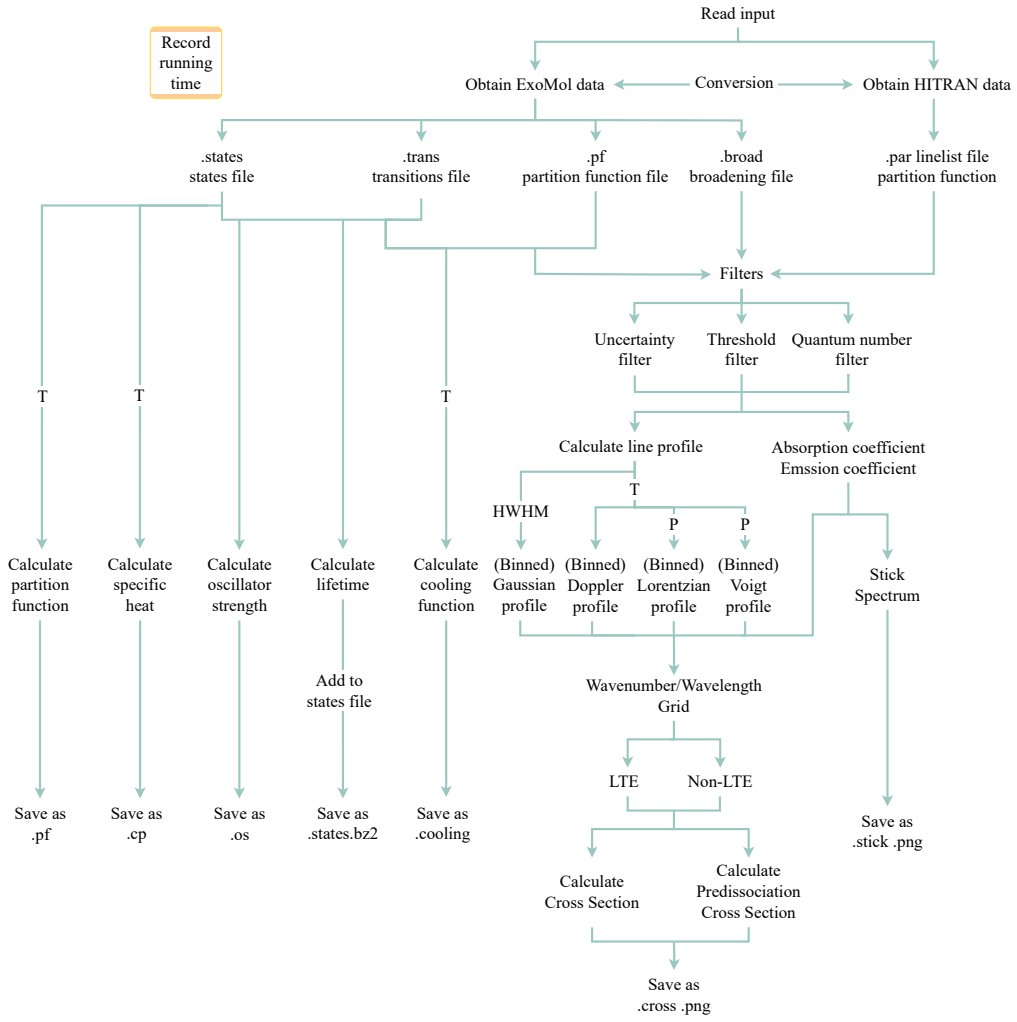


Figure 6.11: Calculation protocol of the PyExoCross program. *T* and *P* refer to temperature and pressure separately.

It is worth noting that some parameters utilized by PyExoCross are expressed in cgs units (e.g. erg) rather than SI units, reflecting conventions commonly used in certain domains of astrophysical research.

6.12 Summary

This chapter introduced PyExoCross a new Python-based program designed to calculate the spectroscopic properties of molecules using spectral line lists. As the Python version of the Fortran-based ExoCross [373], PyExoCross offers limited CPU parallelization but currently lacks GPU support.

Table 6.18: Default units used in the PyExoCross program.

Quantity	Label	Units
Partition function	$Q(T)$	none
Specific heat	$C_p(T)$	J K^{-1} mol^{-1}
Cooling function	$W(T)$	erg (s molecule str)$^{-1}$
Radiative lifetime	τ_i	s
Oscillator strength	f	none
Molar mass	M	Dalton
Wavenumber	ν	cm^{-1}
Wavelength	λ	μm
Temperature	T	K
Pressure	P	bar
Absorption coefficient	I_{fi}	cm molecule^{-1}
Emission coefficient	ε_{if}	erg (s molecule str)$^{-1}$
Absorption cross sections	$\sigma_{\rm ab}$	cm^2molecule^{-1}
Emission cross sections	$\sigma_{\rm em}$	erg cm (s molecule str)$^{-1}$

```
import astropy.constants as ac
Tref = 296.0                        # Reference temperature is 296 K
Pref = 1.0                          # Reference pressure is 1 bar
N_A = ac.N_A.value                  # Avogadro number (1/mol)
h = ac.h.to('erg s').value          # Planck's constant (erg s)
c = ac.c.to('cm/s').value           # Velocity of light (cm/s)
kB = ac.k_B.to('erg/K').value       # Boltzmann's const (erg/K)
R = ac.R.to('J / (K mol)').value    # Molar gas constant (J/(K mol))
c2 = h * c / kB                     # Second radiation constant (cm K)
```

Listing 1: Python code for calculating the physical and astronomical constants.

PyExoCross uses spectral line lists to compute various properties such as partition functions, specific heats, cooling functions, radiative lifetimes, oscillator strengths, stick spectra, and cross sections. Currently, PyExoCross supports data ingestion from the ExoMol, HITRAN, and HITEMP databases, with plans to expand functionality and support for additional data formats and databases.

Chapter 7

Line broadening and line profiles

An individual spectral line at high resolution is characterized by three key parameters: transition frequency, integrated intensity, and line profile, see Figure 7.1. The line profile reflects the effects of line broadening, which depend on temperature and pressure and are critical for many practical applications.

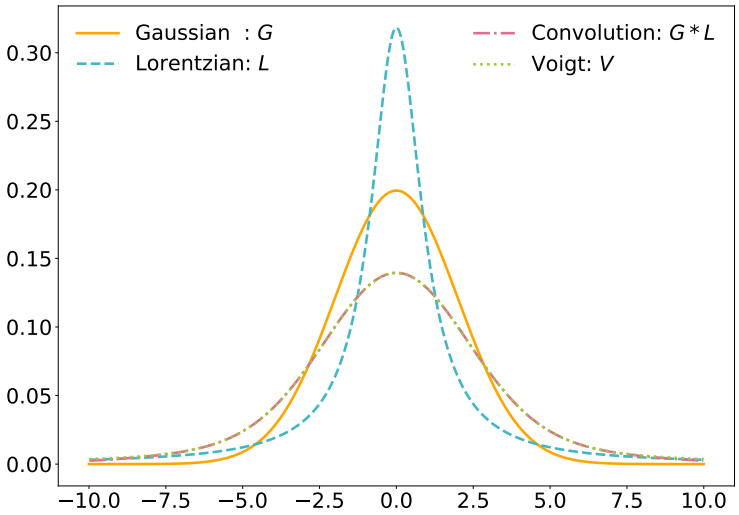


Figure 7.1: Illustration of Gaussian, Lorentzian, Voigt profiles, and the convolution of Gaussian and Lorentzian profiles.

Temperature-dependent Doppler broadening is represented by a Doppler line profile in the form of a Gaussian function. Pressure broadening, on the other hand, is modeled using a Lorentzian line profile. The Voigt profile, a convolution of the Gaussian and Lorentzian functions, provides a more comprehensive representation that accounts for both Doppler and pressure broadening effects [439]. Lifetime ef-

fects may also need to be considered, however, these contribute a Lorentzian component and do not alter the fundamental discussion of line shapes [374].

Line profiles, $f_{\tilde{\nu}}(\tilde{\nu})$, are integrable functions with areas normalized to unity:

$$\int_{-\infty}^{\infty} f_{\tilde{\nu}}(\tilde{\nu})\mathrm{d}\nu = 1. \tag{7.1}$$

PyExoCross provides four options for line profiles: Doppler, generalized Gaussian, Lorentzian, and Voigt profiles. While the Doppler profile is a specific Gaussian with its line width is determined by the mass of the species, temperature, and line position. In contrast, the generalized Gaussian profile allows users to directly specify the half-width at half-maximum (α_{G}). Section 7.5.1 explores seven different methods for computing Voigt profiles, offering insights into their implementation and applications.

7.1 Wing cut-off

Since line profiles theoretically have infinite spread, practical calculations often apply frequency (or wavelength) cutoffs to limit the calculation region around the line centre only and account for the fact that Voigt and Lorentzian functions typically do not accurately represent the far wings of a line profile [373]. Wing cutoffs not only influence computational efficiency and cross-section accuracy but are also essential when incorporating functions like the water continuum. This approach has been recently recommended for exoplanet studies by Anisman et al. [440], and this issue is explored in detail by Gharib-Nezhad et al. [441]. The Voigt profile is not accurate far from the line centre. By default, PyExoCross applies a wing cutoff of 25 cm^{-1}, consistent with HITRAN, HITEMP, and most water continuum representations. While users can modify this value, the default aligns with the recommendations outlined by Gharib-Nezhad et al. [441].

7.2 Broadening parameters

Pressure broadening in the ExoMol database follows the ExoMol diet approach [14], with line broadening parameters stored in `.broad` files. The broad-

eners available for a given species are specified in the corresponding definition file (`.def` and `.def.json`). Users can supply multiple broadening files with suitable suffixes (e.g. `.H2O`, `.CO2`). PyExoCross applies these suffixes to read the relevant files and, if necessary, process mixtures of broadeners, as detailed in Section 7.4.2. PyExoCross can also process user-provided broadening files, provided they use the ExoMol `.broad` file format, which is described in Table 6.16, with specifications detailed in Table 6.17. Currently, the ExoMol database includes `.broad` files for several broadeners, such as air, Ar, CH_4, CO, CO_2, H_2, H_2O, He, N_2, NH_3, NO, and O_2, with the list continuously expanding. The HITRAN database provides air and self-broadening data for all species, along with broadening parameters for He, H_2, H_2O, and CO_2 [442–446] for a small number of key species.

7.3 Doppler and Gaussian profile

The Gaussian line profile is defined as

$$f_{\tilde{\nu}_{fi},\alpha_{\mathrm{G}}}(\tilde{\nu}) = \sqrt{\frac{\ln 2}{\pi}}\frac{1}{\alpha_{\mathrm{G}}}\exp\left(-\frac{(\tilde{\nu}-\tilde{\nu}_{fi})^2\ln 2}{\alpha_{\mathrm{G}}^2}\right), \tag{7.2}$$

where α_{G} is the half-width at half-maximum (HWHM) and $\tilde{\nu}_{fi}$ is the line centre. α_{G} is related to the Gaussian standard deviation σ_{G} as [447]:

$$\sigma_{\mathrm{G}} = \frac{\alpha_{\mathrm{G}}}{\sqrt{2\ln 2}} \tag{7.3}$$

and can be specified in the PyExoCross input file.

The Doppler profile is based on the Gaussian shape specified in Eq. (7.2) with its half-width at half-maximum (HWHM) α_{D} given by [447]

$$\alpha_{\mathrm{D}} = \sqrt{\frac{2N_A k_B T\ln 2}{M}}\frac{\tilde{\nu}_{fi}}{c}, \tag{7.4}$$

which depends on the temperature, molecular mass, and line position. It is determined by key physical constants and parameters, including:

- k_B: Boltzmann constant (erg K).
- c: Speed of light (cm/s).
- T: Temperature (K).
- M: Molar mass of the isotopologue (Da = g mol^{-1}) in grams [442], which can be obtained from the HITRANONLINE website or the definition file (`.def` or `.def.json`) provided in the EXOMOL database.
- N_A: Avogadro's number (mol^{-1}). N_A/M can be replaced by $\frac{1}{1000m}$ where m is the isotopologue mass (in kg), available from the EXOMOL definition file.

It is noteworthy that for both the Doppler and generalized Gaussian profiles, the half-width at half-maximum (HWHM) is independent of the upper and lower molecular states.

PYEXOCROSS calculates the Gaussian profile with the Doppler HWHM $\alpha_{\rm D}$ obtained from the input file provided by users. The Doppler profile depends on the calculated temperature-dependent Doppler broadening $\alpha_{\rm D}$ (see Eq (7.4)).

7.4 Lorentzian profile

The Lorentzian profile refers to the Cauchy distribution, which has the following probability density function (PDF) [448, 449]:

$$f^{L}_{\tilde{\nu}_{fi},\gamma_{\rm L}}(\tilde{\nu}) = \frac{1}{\pi}\frac{\gamma_{\rm L}}{\left(\tilde{\nu}-\tilde{\nu}_{fi}\right)^2+\gamma_{\rm L}^2}, \tag{7.5}$$

where $\tilde{\nu}_{fi}$ represents the centroid in wavenumbers that indicates the peak position of the distribution, while γ_L denotes the Lorentzian half-width at half-maximum (HWHM). In PYEXOCROSS, it is typically assumed in a form that depends on temperature and pressure.

$$\gamma_{\rm L} = \gamma_{\rm ref}\left(\frac{T_{\rm ref}}{T}\right)^{n_{\rm L}}\frac{P}{P_{\rm ref}}. \tag{7.6}$$

The Lorentzian profile is centered at the transition frequency, denoted as $\tilde{\nu}_{fi}$. As in the HITRAN convention, a reference temperature $T_{\rm ref} = 296$ K and a reference pressure $P_{\rm ref} = 1$ bar are used [450]. The key broadening parameters are the reference half-width at half-maximum (HWHM) $\gamma_{\rm ref}$ and the temperature exponent $n_{\rm L}$. These parameters depend on the specified broadeners in the EXOMOL database (see Sections 7.2, 7.4.1). For HITRAN data, $\gamma_{\rm air}$ and $\gamma_{\rm self}$ are used to compute $\gamma_{\rm ref}$ based on user-defined mixing ratios (see Section 7.4.2). Similarly, $n_{\rm L}$ in the EXOMOL database corresponds to $n_{\rm air}$ in HITRAN line lists. In general, both $\gamma_{\rm ref}$ and $n_{\rm L}$ depend on the upper and lower states.

Users must specify the broadeners and their respective broadener mixture ratios in the `.inp` input file for PYEXOCROSS to correctly process the line broadening parameters. Users have the option to directly input the half-width at half-maximum (HWHM) $\gamma_{\rm L}$ of the Lorentzian profile as a constant value for calculations which bypasses the need to compute $\gamma_{\rm L}$ from the derived line-broadening parameters provided in the EXOMOL database files.

7.4.1 Extracting broadening parameters from `.broad`

The calculations of Lorentzian and Voigt broadening require data from both the states file (`.states`) and the broadening file (`.broad`). When the value of J'' in the `.states` file exceeds $J''_{\rm max}$ in the `.broad` file, the broadening parameters $\gamma_{\rm L}$ and $n_{\rm L}$ corresponding to $J''_{\rm max}$ are used as a default. For cases where J'' is a half-integer, the parameters are taken from the `.broad` file corresponding to the index value $J'' - \frac{1}{2}$. This indexing-based approach is particularly efficient when implemented in Python. An example of Python code for extracting broadening parameters from a `.broad` file is provided in Listing 2.

7.4.2 Mixtures of broadeners

In PYEXOCROSS users can incorporate multiple broadeners or specify mixtures of broadeners for their calculations. When dealing with mixtures, the fractional abundance (X_i) of each broadener i must be specified in the input file. The the overall value of half-width at half-maximum (HWHM) $\gamma_{\rm L}$, for the mixture is then

```
def extract_broad(broad_df, states_trans_df):
    max_broad_J = max(broad_df['Jpp'])
    Jpp = states_trans_df['J"'].values
    Jpp[Jpp > max_broad_J] = max_broad_J
    id_broad = (states_trans_df['J"']-0.1).round(0).astype('int32').values
    gamma_L = broad_df['gamma_L'][id_broad]
    n_air = broad_df['n_air'][id_broad]
    return(gamma_L, n_air)
```

Listing 2: A sample of the Python code for extracting broadeners from the `.broad` broadening file. PyExoCross program reads the `.broad` file and stores the data in a `DataFrame` named `broad_df`. Simultaneously, the program generates another `DataFrame`, `states_trans_df`, by merging the states file (`.states`) with the transitions file (`.trans`), using the upper and lower state IDs. The columns `Jpp` and `J"` represent the lower J quantum numbers of the transitions, while `id_broad` indicates the corresponding index in the `broad_df`. The parameters `gamma_L` and `n_air` in the code correspond to γ_{ref} (the reference HWHM) and n_{air} (the temperature exponent for air-broadening).

computed as the weighted sum of the individual broadening contributions γ_i^{L} from each broadener, as described by [373]:

$$\gamma_{\text{L}} = \sum_i \gamma_i^{\text{L}} X_i \quad \text{and} \quad \sum_i X_i = 1. \tag{7.7}$$

This approach allows for precise modeling of line profiles by considering the contributions of multiple broadening species, each weighted by its relative abundance. The flexibility of this feature enables PyExoCross to handle complex atmospheric and astrophysical scenarios involving diverse broadening conditions.

7.4.3 Lifetime broadening

The natural lifetime of a state, in principle, contributes to the spectral line profile, alongside temperature-dependent Doppler broadening and pressure-dependent collision broadening. However, in most cases, the contribution of natural lifetime broadening is negligible and thus often ignored in practical applications. Nevertheless, for predissociating states, lifetime broadening can have a significant impact [11, 451]. Lifetime broadening can be incorporated into the Voigt profile calculations when needed [374].

Similar to pressure broadening, lifetime broadening (γ_τ) also results in a

Lorentzian line shape, with its HWHM expressed as:

$$\gamma_\tau = \frac{\hbar}{2\tau hc}, \tag{7.8}$$

where τ is the natural lifetime (s) of the state. h represents Planck's constant (erg s), $\hbar = h/(2\pi)$ is the reduced Planck constant, and c is the speed of light (cm/s). This approach provides a straightforward means of including lifetime effects in line profile modeling when necessary.

The total Lorentzian component of the line profile, γ_{tot}, is obtained by summing the pressure broadening (γ_p) and lifetime broadening (γ_τ):

$$\gamma_{\text{tot}} = \gamma_p + \gamma_\tau \tag{7.9}$$

In practical applications, γ_{tot} serves as the total half-width at half-maximum (HWHM) for the Lorentzian profile. Since the Voigt profile, which combines Doppler and Lorentzian components, is typically used, incorporating γ_{tot} adds no

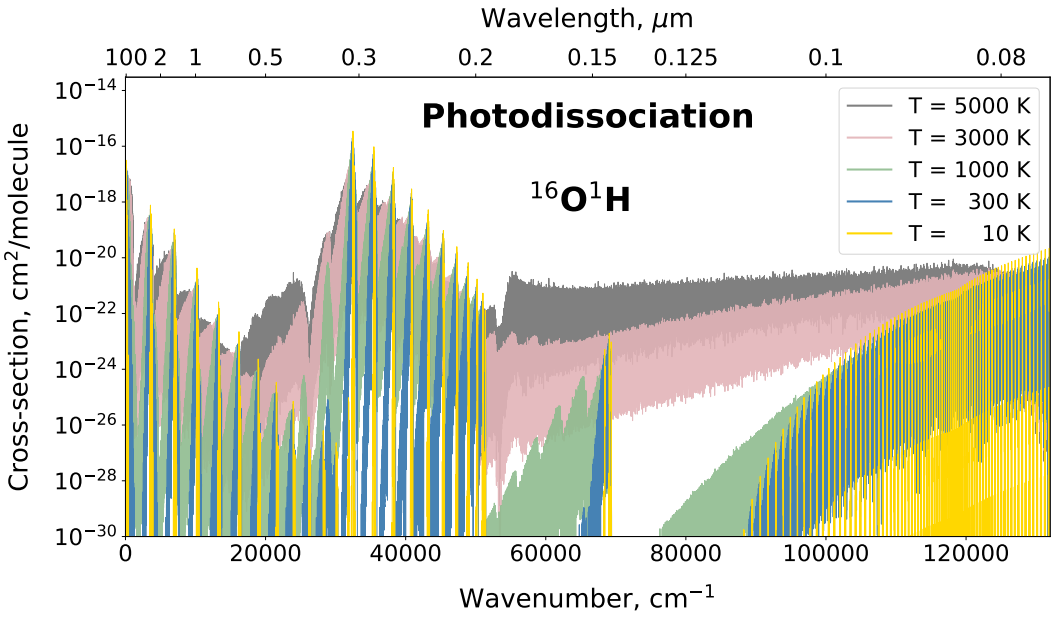


Figure 7.2: Photodissociation absorption spectrum of ^{16}O^{1}H at $T = 10$, 300, 1000, 3000, and 5000 K simulated using the ExoMol line list of the dataset MYTHOS [8] broadened using a Voigt line profile with the bin size 0.1 cm^{-1} at $P = 1$ bar in wavenumber range $[0, 131\,900]$ cm^{-1}.

computational complexity to the calculations. This approach allows for routine inclusion of lifetime broadening effects without requiring additional considerations for phenomena like predissociation.

PyExoCross uses lifetime broadening to calculate photodissociation cross sections. Figure 7.2 is an example of photodissociation cross sections at different temperatures for ExoMol $^{16}O^{1}H$ line lists [8].

7.5 Voigt profile

The Voigt profile is the result of convolving the Gaussian and Lorentzian profiles. The Voigt profile accurately represents the combined effects of Doppler and collisional broadening, making it indispensable for high-resolution spectroscopic applications. It can be expressed mathematically as [452, 453]:

$$\begin{aligned} f^{V}_{\tilde{\nu}_{fi},\sigma_{\mathrm{G}},\gamma_{\mathrm{L}}}(\tilde{\nu}) &= f^{G}_{\tilde{\nu}_{fi},\sigma_{\mathrm{G}}}(\tilde{\nu}) \otimes f^{\mathrm{L}}_{\tilde{\nu}_{fi},\gamma_{\mathrm{L}}}(\tilde{\nu}) \\ &= \int_{-\infty}^{+\infty} \mathrm{d}\tilde{\nu}' f^{G}(\tilde{\nu}-\tilde{\nu}',\sigma_{\mathrm{G}}) \times f^{\mathrm{L}}(\tilde{\nu}'-\tilde{\nu}_{fi},\gamma_{\mathrm{L}}) \\ &= \frac{y}{\pi^{3/2}\sqrt{2}\sigma_{\mathrm{G}}} \int_{-\infty}^{+\infty} \frac{e^{-t^2}}{(x-t)^2+y^2} \mathrm{d}t, \end{aligned} \tag{7.10}$$

where x represents the wavenumber scale normalized by the Gaussian standard deviation, while y denotes the ratio of the Lorentzian half-width at half-maximum (HWHM) to the Gaussian standard deviation.

$$x = \frac{\tilde{\nu}-\tilde{\nu}_{fi}}{\sqrt{2}\sigma_{\mathrm{G}}} \qquad \text{and} \qquad y = \frac{\gamma_{\mathrm{L}}}{\sqrt{2}\sigma_{\mathrm{G}}}. \tag{7.11}$$

$f^{G}_{\tilde{\nu}_{fi},\sigma_{\mathrm{G}}}(\tilde{\nu})$ is the Gaussian profile and $f^{\mathrm{L}}_{\tilde{\nu}_{fi},\gamma_{\mathrm{L}}}(\tilde{\nu})$ is the Lorentzian profile. σ_{G} is the Gaussian standard deviation, and γ_{L} corresponds to the Lorentzian HWHM. These parameters are functions of molecular properties, temperature, and pressure, as described in earlier sections (see Section 7.3 and 7.4). The central wavenumber or frequency, $\tilde{\nu}_{fi}$, is determined by the energy difference associated with the transition. This parameter defines the line centre and serves as a reference point for calculating

the shape of profile [454].

PyExoCross provides eight distinct methods for computing the Voigt profile, five of which are classified as pseudo-Voigt approximations. These nine methods offer a range of computational strategies, balancing precision and efficiency, as detailed below.

7.5.1 SciPy Voigt profile

The Python open-source library `SciPy` [15] provides a direct and efficient implementation for calculating the Voigt profile. A sample implementation using `SciPy` is shown in Listing 3.

```
from scipy.special import voigt_profile
def scipy_Voigt_profile(dv, sigma, gamma):
    return voigt_profile(dv, sigma, gamma)
```

Listing 3: A sample Python code snippet demonstrating the use of the `SciPy` package to calculate the Voigt profile [15]. $\mathrm{d}\nu = \tilde{\nu} - \tilde{\nu}_{fi}$ represents the difference from the line centre ($\tilde{\nu}_{fi}$). The parameter `sigma` is the standard deviation σ_{G} of the Gaussian profile, while `gamma` is the Lorentzian profile half-width for pressure broadening γ_{L}.

The `scipy.special.voigt_profile` function uses the `scipy.signal.convolve` method to compute the Voigt profile by convolving a one-dimensional Normal distribution (Gaussian profile) with standard deviation σ_{G} and a one-dimensional Cauchy distribution (Lorentzian profile) characterized by the half-width at half-maximum (HWHM) γ_{L} [15], see Eq. (7.10). Below is the Python implementation using `scipy.signal.convolve` to perform the convolution, as shown in Listing 4.

```
from scipy.signal import convolve
convolved = dx * convolve(Lorentzian_profile,
                          Gaussian_profile,
                          mode="same")
```

Listing 4: A method for calculating the Voigt profile through the convolution of Gaussian and Lorentzian profiles, using the `SciPy` Python library, specifically the `scipy.special.voigt_function`.

7.5.2 SciPy wofz Voigt profile

The Voigt profile can be computed using the open-source Python library `SciPy`, as shown in Listing 5. The `SciPy` function `scipy.special.wofz` provides a calculation of the Voigt profile with parameters σ_{G} and γ_{L}, as defined by

$$f^{V}_{\tilde{\nu}_{fi},\sigma_{\mathrm{G}},\gamma_{\mathrm{L}}}(\tilde{\nu}) = \frac{1}{\sqrt{2\pi}\sigma_{\mathrm{G}}} K(x,y), \tag{7.12}$$

where x and y are defined in Eq. (7.11). $K(x,y)$ represents the Voigt function satisfying $\int K \mathrm{d}x = \sqrt{\pi}$.

$$K(x,y) = \frac{y}{\pi} \int_{-\infty}^{+\infty} \frac{e^{-t^2}}{(x-t)^2 + y^2} \mathrm{d}t. \tag{7.13}$$

$K(x,y)$ comprises the real part of the complex error function, also known as the Faddeeva function $\omega(z)$ [455].

$$\omega(z) \equiv K(x,y) + iL(x,y) = \frac{i}{\pi} \int_{-\infty}^{+\infty} \frac{e^{-t^2}}{z-t} \mathrm{d}t. \tag{7.14}$$

The imaginary Voigt function, $L(x,y)$, is also of significant importance for various applications and should be carefully considered in relevant analyses [456]. $z = x + iy$ is a complex argument and $i = \sqrt{-1}$.

$$z = x + iy = \frac{\left(\nu - \tilde{\nu}_{fi} + i\gamma_{\mathrm{L}}\right)}{\sqrt{2}\sigma_{\mathrm{G}}}. \tag{7.15}$$

$$f^{V}_{\tilde{\nu}_{fi},\sigma_{\mathrm{G}},\gamma_{\mathrm{L}}}(\tilde{\nu}) = \frac{\mathbf{Re}\left[\omega\left(z\right)\right]}{\sigma_{\mathrm{G}}\sqrt{2\pi}}. \tag{7.16}$$

In Eq. (7.15), the limiting cases of $\sigma_{\mathrm{G}} = 0$ and $\gamma_{\mathrm{L}} = 0$ simplify Eq. (7.16) to the Lorentzian profile and Gaussian profile, respectively.

The `wofz` function in `scipy.special.wofz` can be computed with the complementary error function $\mathrm{erf}(z)$, as described by Pierluissi et al. [457]:

$$\omega(z) = e^{-z^2}\left(1 - \mathrm{erf}(-iz)\right). \tag{7.17}$$

```
import numpy as np
from scipy.special import wofz
def scipy_wofz_Voigt_profile(dv, sigma, gamma):
    z = (dv + 1j * gamma) / sigma / np.sqrt(2)
    voigt = (np.real(wofz(z)) / sigma / np.sqrt(2 * np.pi))
    return voigt
```

Listing 5: A sample of the Python code used to compute the Voigt profile [15, 16] with the real part of the Faddeeva function.

7.5.3 Humlíček Voigt profile

Humlíček's algorithm [458] is widely applied for approximating the Voigt profile function. This approach provides a means of calculating the complex probability function $\omega(z)$, where $z = x + iy$ (see Eqs. (7.10, 7.14, and 7.15)). Here, x represents the wavenumber scale normalized by the Doppler HWHM α_{D}, and y denotes the ratio of the Lorentzian HWHM γ_{L} to the Doppler HWHM α_{D}.

$$x = \frac{\left(\tilde{\nu} - \tilde{\nu}_{fi}\right)\sqrt{\ln 2}}{\alpha_{\mathrm{D}}} \qquad \text{and} \qquad y = \frac{\gamma_{\mathrm{L}}\sqrt{\ln 2}}{\alpha_{\mathrm{D}}}. \tag{7.18}$$

Humlíček's w4 algorithm uses a rational approximation technique to minimize the relative error of the imaginary part along the real axis [459]. This method allows for the precise evaluation of both the real and imaginary components of $\omega(z)$ with a high degree of relative accuracy [458–460].

$$\omega(z) = \begin{cases} \text{Region 1. If } s \geq 15 \\ \dfrac{t}{(0.5+t^2)\sqrt{\pi}} \\ \text{Region 2. If } 5.5 \leq s < 15 \\ \dfrac{t\left(1.4104739589+\frac{u}{\sqrt{\pi}}\right)}{0.75+u(3+u)} \\ \text{Region 3. If } s < 5.5 \text{ and } y \geq 0.195\,|x| - 0.176 \\ \dfrac{16.4955+t(20.20933+t(11.96482+t(3.778987+0.5642236t)))}{16.4955+t(38.82363+t(39.27121+t(21.69274+t(6.699398+t))))} \\ \text{Region 4. If } s < 5.5 \text{ and } y < 0.195\,|x| - 0.176 \\ e^{u} - \dfrac{t(36183.31-u(3321.99-u(1540.787-u(219.031-u(35.7668-u(1.320522-u0.56419))))))}{32066.6-u(24322.8-u(9022.23-u(2186.18-u(364.219-u(61.5704-u(1.84144-u))))))} \end{cases} \tag{7.19}$$

The complex function is defined as $t = y - xi$, with $s = |x| + y$ and $u = t^2$. Zaghloul [461] and Hui et al. [462] introduced a simplified and more efficient approximation method for evaluating the Faddeeva function, characterized by partitioning into six distinct regions for improved accuracy.

The Humlíček Voigt line profile is subsequently expressed as [458, 459]:

$$f^V_{\tilde{\nu}_{fi},\alpha_D,\gamma_L}(\tilde{\nu}) = \frac{\mathbf{Re}\left[\omega(z)\right]\sqrt{\ln 2}}{\pi\alpha_D}. \tag{7.20}$$

7.5.4 Pseudo-Voigt profile

PyExoCross offers five pseudo-Voigt line profiles; however, these are not the recommended algorithms. Due to their lower accuracy and slower runtime compared to other Voigt profiles, pseudo-Voigt profiles are not included in the comparison presented in Section 7.7.

The pseudo-Voigt function $V(\mathrm{d}\nu, \gamma_V)$ is widely applied for modeling experimental spectral line shapes [463]. Serving as an approximation to the Voigt profile $V(x)$, the pseudo-Voigt profile is given by a linear combination of a Gaussian curve $G(x)$ and a Lorentzian curve $L(x)$, rather than their convolution [463]. The normalized pseudo-Voigt profile $V(\mathrm{d}\nu, \gamma_V)$ is mathematically defined by the following equation [464, 465]:

$$V(\mathrm{d}\nu, \gamma_V) = \eta L(\mathrm{d}\nu, \alpha_D) + (1-\eta)G(\mathrm{d}\nu, \gamma_L) \quad \text{with} \quad 0 < \eta < 1, \tag{7.21}$$

where $\mathrm{d}\nu = \tilde{\nu} - \tilde{\nu}_{fi}$ represents the frequency difference, while α_D and γ_L denote HWHM of the Gaussian and Lorentzian profiles, respectively.

The accuracy of the pseudo-Voigt profile is influenced by the methods used to determine the parameter η and the Voigt profile HWHM γ_V.

7.5.4.1 Thompson pseudo-Voigt profile

For the Thompson pseudo-Voigt profile, η is defined as [464]:

$$\eta \approx 1.36603(\gamma_L/\gamma_V) - 0.47719(\gamma_L/\gamma_V)^2 + 0.11116(\gamma_L/\gamma_V)^3. \tag{7.22}$$

The half-width at half-maximum (HWHM) γ_V for the Thompson pseudo-Voigt profile can be derived from the corresponding Gaussian and Lorentzian HWHMs using the following expression [464]:

$$\gamma_V = \left[\alpha_D^5 + 2.69269\alpha_D^4\gamma_L + 2.42843\alpha_D^3\gamma_L^2 + 4.47163\alpha_D^2\gamma_L^3 + 0.07842\alpha_D\gamma_L^4 + \gamma_L^5\right]^{1/5}. \quad (7.23)$$

7.5.4.2 Kielkopf pseudo-Voigt profile

Kielkopf [466] introduced a weighted-sum approach to approximate the value of γ_V for the pseudo-Voigt profile, using Eqs. (7.21 and 7.22). This method is further discussed in Schreier [467].

$$\begin{aligned}\gamma_V &\approx \frac{1}{4}\left[(1+0.099\ln 2)\gamma_L + \sqrt{(1-0.099\ln 2)^2\gamma_L^2 + 4\ln 2\alpha_D^2}\right] \\ &\approx 0.5346\gamma_L + \sqrt{0.2166f\gamma_L^2 + \alpha_D^2}.\end{aligned} \quad (7.24)$$

7.5.4.3 Olivero pseudo-Voigt profile

Olivero and Longbothum [468] provided an alternative formulation to Eqs. (7.21 and 7.22), where HWHM γ_V is approximated by

$$\gamma_V \approx \left[1 - 0.18121(1-d^2) - (0.023665e^{0.6d} + 0.00418e^{-1.9d})\sin(\pi d)\right](\alpha_D + \gamma_L). \quad (7.25)$$

d is a dimensionless parameter related to α_D and γ_L [468].

$$d = \frac{\gamma_L - \alpha_D}{\gamma_L + \alpha_D}. \quad (7.26)$$

7.5.4.4 Liu-Lin pseudo-Voigt profile

Liu et al. [469] introduced a parameter d, which is related to the Gaussian and Lorentzian half-widths (α_D and γ_L), as defined in Eq. (7.26), to propose a new conversion for η originally presented by Kielkopf [466] (see Eq. (7.22)). The half-width of the Voigt profile remains consistent with Eq. (7.25). The Liu-Lin pseudo-

Voigt profile is determined using Eqs.(7.21, 7.25, 7.26, and 7.27).

$$\eta \approx 0.68188 + 0.61293d - 0.18384d^2 - 0.11568d^3. \tag{7.27}$$

7.5.4.5 Rocco pseudo-Voigt profile

Di Rocco and Cruzado [470] provided a theoretically justified approach for calculating γ_V and η in the pseudo-Voigt methodology, defining these parameters as follows:

$$\gamma_V = \omega_G b_{1/2}(y), \tag{7.28}$$

$$\eta \approx \frac{\sqrt{\pi} V_y(0)\gamma_V - \sqrt{\ln 2}}{\sqrt{\pi} V_y(0)\gamma_V(1 - \sqrt{\pi \ln 2})}. \tag{7.29}$$

Here,

$$\omega_G = \frac{\alpha_D}{\sqrt{\ln 2}}, \tag{7.30}$$

$$b_{1/2}(y) = y + \sqrt{\ln 2} e^{-0.6055y + 0.0718y^2 - 0.0049y^3 + 0.000136y^4}. \tag{7.31}$$

$V_y(0)$ is the Voigt profile in Eq. (7.10) at $x = 0$ is definded as:

$$V_y(0) = \frac{e^{y^2}(\Phi_c(y))}{\sqrt{\pi}\omega_G}, \tag{7.32}$$

where,

$$\Phi_c(y) = 1 - \mathrm{erf}(y) \quad \text{and} \quad y = \frac{\gamma_L}{\omega_G}. \tag{7.33}$$

7.6 Binned line profile

All the aforementioned methods use a sampling algorithm to evaluate the Voigt profile at a set of frequency wavenumbers $\tilde{\nu}_i$. However, a limitation of these methods is the potential for under-sampling the opacity (absorption or emission) at low sampling resolution, as discussed by Hill et al. [471]. To avoid this issue and guarantee that the full opacity is accurately preserved, a binned approach can be implemented. This approach involves defining the analytically integrated (averaged) cross sections within a bin $[\tilde{\nu}_k - \Delta\tilde{\nu}/2, \tilde{\nu}_k + \Delta\tilde{\nu}/2]$ as a sum of the contributions

from each spectral line [471] as given by:

$$\begin{aligned}\bar{\sigma}_k^{\mathrm{ab},fi} &= \frac{I_{fi}}{\Delta\tilde{\nu}}\int_{\tilde{\nu}_k-\Delta\tilde{\nu}/2}^{\tilde{\nu}_k+\Delta\tilde{\nu}/2} f_{\tilde{\nu}_{fi}}(\tilde{\nu})\mathrm{d}\tilde{\nu},\\ \bar{\sigma}_k^{\mathrm{em},if} &= \frac{\varepsilon_{if}}{\Delta\tilde{\nu}}\int_{\tilde{\nu}_k-\Delta\tilde{\nu}/2}^{\tilde{\nu}_k+\Delta\tilde{\nu}/2} f_{\tilde{\nu}_{if}}(\tilde{\nu})\mathrm{d}\tilde{\nu},\end{aligned} \tag{7.34}$$

where $\tilde{\nu}_{fi}$ is the line centre. I_{fi} and ε_{if} are the line intensities (absorption and emission) from Eqs. (6.9 and 6.10).

The related functions for the binned cross sections for the Doppler, Gaussian, Lorentzian and Voigt profiles are specified below.

7.6.1 Binned Doppler and Gaussian profile

Both the Doppler and Gaussian profiles use the Gaussian function. Using the property that the integral of the Gaussian function has a solvable analytical form, the error function (erf) can be applied:

$$\int e^{-x^2}\mathrm{d}x = \frac{\sqrt{\pi}}{2}\mathrm{erf}(x). \tag{7.35}$$

Therefore, the analytical integral can be applied to the Gaussian profile [471]:

$$\begin{aligned}\bar{\sigma}_k^{fi} &= \frac{I_{fi}}{\Delta\tilde{\nu}}\int_{\tilde{\nu}_k-\Delta\tilde{\nu}/2}^{\tilde{\nu}_k+\Delta\tilde{\nu}/2} f^G_{\tilde{\nu}_{fi}}(\tilde{\nu})\mathrm{d}\tilde{\nu}\\ &= \frac{I_{fi}}{2\Delta\tilde{\nu}}\left[\mathrm{erf}\left(x^+_{k,fi}\right)-\mathrm{erf}\left(x^-_{k,fi}\right)\right].\end{aligned} \tag{7.36}$$

$x^{\pm}_{k,fi}$ represent the scaled boundaries of the frequency bin k, centered at $\tilde{\nu}_k$ and related to the line center $\tilde{\nu}_{fi}$, defined as [471]:

$$x^{\pm}_{k,fi} = \frac{\sqrt{\ln 2}}{\alpha_{\mathrm{D}}}\left[\tilde{\nu}_k-\tilde{\nu}_{fi}\pm\frac{\Delta\tilde{\nu}}{2}\right]. \tag{7.37}$$

7.6.2 Binned Lorentzian profile

The integral of the Lorentzian function has an analytical solution which can be expressed as:

$$\int\frac{\mathrm{d}x}{x^2+\gamma^2} = \frac{1}{\gamma}\arctan\left(\frac{x}{\gamma}\right)+C. \tag{7.38}$$

When calculating profiles in the wing regions far from the line centre, the normalization correction factor approaches unity. To efficiently solve the far wings and reduce excessive computation time, I truncate the profiles at a certain distance $\Delta\tilde{\nu}/2$ from the line centre [472, 473]. Consequently, PyExoCross offers users two methods for computing binned Lorentzian and Voigt profiles, with or without applying a wing cutoff filter.

The analytical binned integral approach can similarly be applied to the Lorentzian profile $f^{\mathrm{L}}_{\tilde{\nu}_{fi}}(\tilde{\nu})$ as follows [373]:

$$\begin{aligned}\bar{\sigma}_k^{fi} &= \frac{I_{fi}}{\Delta\tilde{\nu}} \int_{\tilde{\nu}_k-\Delta\tilde{\nu}/2}^{\tilde{\nu}_k+\Delta\tilde{\nu}/2} f^{\mathrm{L}}_{\tilde{\nu}_{fi}}(\tilde{\nu})\mathrm{d}\tilde{\nu} \\ &= \frac{I_{fi}}{\pi\Delta\tilde{\nu}} \left[\arctan\left(y^{+}_{k,fi}\right) - \arctan\left(y^{-}_{k,fi}\right)\right].\end{aligned} \tag{7.39}$$

$\bar{\sigma}_k^{fi}$ represents the averaged cross section for bin k, centred at $\tilde{\nu}_k$ over the interval $\Delta\tilde{\nu}$. $\tilde{\nu}_{fi}$ denotes the centre of the spectral line. $y^{\pm}_{k,fi}$ can be defined as:

$$y^{\pm}_{k,fi} = \frac{\tilde{\nu}_k - \tilde{\nu}_{fi} \pm \Delta\tilde{\nu}/2}{\gamma_{\mathrm{L}}}. \tag{7.40}$$

Since line profiles are typically cut off at a distance $\Delta\tilde{\nu}/2$ from the line centre, it is common practice to preserve the total line strength by re-normalizing its area to unity [473]. To achieve this, Eq. (7.39) should be multiplied by the correction factor $b_{\mathrm{corr}} \geq 1$, which is given by:

$$b_{\mathrm{corr}} = \frac{\pi}{\arctan\left(y_{\mathrm{end},fi}\right) - \arctan\left(y_{\mathrm{start},fi}\right)}, \tag{7.41}$$

where $y_{\mathrm{start},fi}$ and $y_{\mathrm{end},fi}$ are defined as:

$$y_{\mathrm{start},fi} = \frac{\tilde{\nu}_{\mathrm{start}} - \tilde{\nu}_{fi}}{\gamma_{\mathrm{L}}} \quad \text{and} \quad y_{\mathrm{end},fi} = \frac{\tilde{\nu}_{\mathrm{end}} - \tilde{\nu}_{fi}}{\gamma_{\mathrm{L}}}. \tag{7.42}$$

7.6.3 Binned Voigt profile

Gauss-Hermite quadrature is a variant of Gaussian quadrature used in numerical analysis to approximate the values of integrals [474]:

$$\int_{-\infty}^{+\infty} e^{-t^2} f(t) \approx \sum_{m=1}^{n} \omega_m f(t_m), \tag{7.43}$$

where n denotes the number of used sample points. t_m are the roots of the Hermite polynomial $H_n(t)$ $(m = 1,2,...,n)$ (physicists' version). The Gauss-Hermite quadrature weights ω_m are provided by Abramowitz and Stegun [474].

$$\omega_m = \frac{2^{n-1} n! \sqrt{\pi}}{n^2 [H_{n-1}(t_m)]^2}. \tag{7.44}$$

In combination with the Gaussian and Lorentzian profiles, Humlíček's algorithm [458] is used as the Gauss-Hermite quadrature to develop a formulation of a similar integral method for the Voigt profile [373]. This method incorporates the half-width at half-maximum (HWHM) of the Doppler (and Gaussian) profile α_{D} and the HWHM of the Lorentzian profile γ_{L} (refer to Eq. (7.10)).

$$\begin{aligned} f^V_{\tilde{\nu}_{fi},\alpha_{\mathrm{D}},\gamma_{\mathrm{L}}}(\tilde{\nu}) &= \frac{y \ln 2}{\pi^{3/2} \alpha_{\mathrm{D}}^2} \int_{-\infty}^{+\infty} \frac{e^{-t^2}}{(x-t)^2 + y^2} \mathrm{d}t \\ &= \frac{y \ln 2}{\pi^{3/2} \alpha_{\mathrm{D}}^2} \sum_{m=1}^{N_{\text{G-H}}} \frac{\omega_m^{\text{G-H}}}{(x-t_m)^2 + y^2}. \end{aligned} \tag{7.45}$$

t_m are the Gauss-Hermite quadrature points (roots) and $\omega_m^{\text{G-H}}$ are the corresponding Gauss-Hermite quadrature weights. By applying Eq. (7.39) to form area-conserved integrals, the binned Voigt profile is derived as follows [373]:

$$\bar{\sigma}_k^{fi} = \frac{I_{fi}}{\pi^{3/2} \Delta\tilde{\nu}} \sum_{m=1}^{N_{\text{G-H}}} \omega_m^{\text{G-H}} \left[\arctan\left(z^+_{k,fi,m}\right) - \arctan\left(z^-_{k,fi,m}\right) \right]. \tag{7.46}$$

Typically, $N_{\text{G-H}} = 20$ is used as the number of Gauss-Hermite points. The scaled

limits $z^{\pm}_{k,fi,m}$ for bin k, centred at the grid point $\tilde{\nu}_k$, are defined as:

$$z^{\pm}_{k,fi,m} = \frac{\tilde{\nu}_k - \tilde{\nu}_{fi} - t_m \sigma_{\mathrm{G}} \pm \Delta\tilde{\nu}/2}{\gamma_{\mathrm{L}}}. \tag{7.47}$$

This expression can be further adjusted to account for the missing area due to the wing cutoff, similar to the Lorentzian example mentioned earlier. In this case, Eq. (7.46) $\bar{\sigma}^{fi}_k$ should be multiplied by the normalization correction factor b_{corr}, which is given by:

$$b_{\mathrm{corr}} = \frac{\pi}{\arctan\left(z_{\mathrm{end},fi,m}\right) - \arctan\left(z_{\mathrm{start},fi,m}\right)}, \tag{7.48}$$

where

$$z_{\mathrm{start},fi,m} = \frac{\tilde{\nu}_{\mathrm{start}} - \tilde{\nu}_{fi} - t_m \sigma_{\mathrm{G}}}{\gamma_{\mathrm{L}}} \quad \text{and} \quad z_{\mathrm{end},fi,m} = \frac{\tilde{\nu}_{\mathrm{end}} - \tilde{\nu}_{fi} - t_m \sigma_{\mathrm{G}}}{\gamma_{\mathrm{L}}}. \tag{7.49}$$

7.7 Line profile comparison

7.7.1 A comparative analysis of the sampling and binned line profile methodologies

When the resolution of cross sections (i.e., the number of points per wavenumber interval) is low, the results of cross section calculations become highly sensitive to the method used for profile evaluation—whether through sampling or binning. Sampling methods are designed to accurately reproduce cross sections at specific sampling points. In contrast, binned methods guarantee that the integrated cross section within a given bin is preserved, regardless of the number of grid points or integration intervals. Accordingly, binned methods are particularly beneficial when conserving the total flux, absorption, or emission coefficients is essential. On the other hand, sampling methods are preferable when the accuracy of cross sections at specific frequencies is critical.

Figure 7.3 illustrates the comparison of the Doppler, Lorentzian, and Voigt line profiles calculated using both sampling and binned methods for a single normalized

transition in the wavenumber range [4064, 4065] cm^{-1} at a temperature of $T = 300$ K and a pressure of $P = 1$ bar. These calculations use a bin size of 0.125 cm^{-1} (8 points per cm^{-1}) with a wing cutoff of 100 cm^{-1}. The continuous curves represent the exact line profiles, while the discrete points correspond to the numerically calculated values. In the sampling approach, the line profile passes precisely through each calculated point, preserving the shape and resolution of the profile. In contrast, the binning method adjusts the point heights—either above or below the true curve, such that the total integrated area (normalized to unity) remains conserved, as follows:

$$\int_{-\infty}^{\infty} f_{\tilde{\nu}_{fi}}(\tilde{\nu})d\nu \equiv \sum_i f_{\tilde{\nu}_{fi}}(\tilde{\nu}_i)\Delta\nu, \tag{7.50}$$

where the rectangular integration rule is used. This guarantees the conservation of total line strength (or integrated cross section), even when the profile shape is approximated.

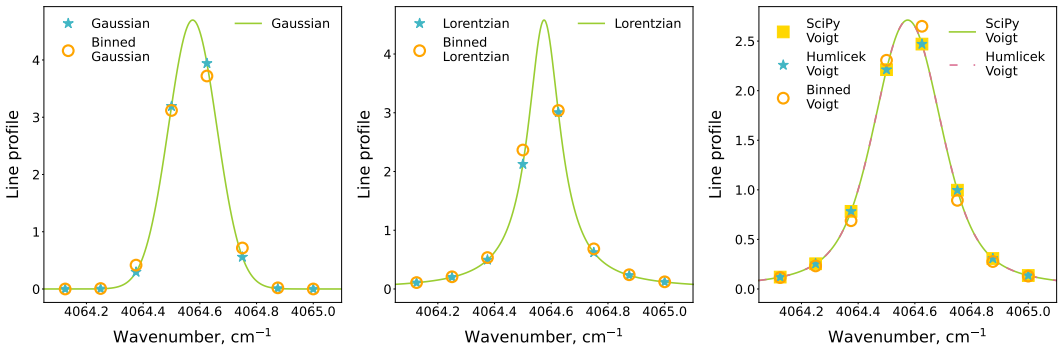


Figure 7.3: Comparison of the sampling and binned methods for different line profiles is shown for a single normalized transition with a wing cutoff of 100 cm^{-1} and a bin size of 0.125 cm^{-1} (8 points per cm^{-1}) at a temperature of $T = 300$ K and pressure $P = 1$ bar. Left panel: Doppler profile.Middle panel: Lorentzian profile. Right panel: Voigt profile.

7.7.2 A comparative analysis of line profile errors

The profile normalization error, defined as the difference of the integrated area of a normalized line profile from unity, is expressed as:

$$\varepsilon_{\text{area}} = 1 - \sum f_{\tilde{\nu}_{fi}}(\tilde{\nu})\Delta\tilde{\nu}. \tag{7.51}$$

For the binned methods, $\varepsilon_{\text{area}}$ is zero by construction, while for the sampling methods, it depends on the number of sampling points. Selecting an appropriate number of sampling points is crucial in radiative transfer applications [475]. Undersampling can result in the underestimation of opacity, while using too many points can significantly slow down the calculations.

The profile sampling error can be defined as a root-mean-square error (RMSE) $\varepsilon_{\text{sampling}}$, representing the deviation of the binned method profiles from the sampling method profiles at the sampling points, as expressed by:

$$\varepsilon_{\text{sampling}} = \text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(f_{\text{sampling}}(\tilde{\nu}_i) - f_{\text{binned}}(\tilde{\nu}_i)\right)^2}, \tag{7.52}$$

where N represents the number of sampling points. The sampling error $\varepsilon_{\text{sampling}}$ should be zero for the sampling methods by construction (within the limits of numerical accuracy).

Figure 7.4 provides a detailed analysis of the line profile errors $\varepsilon_{\text{area}}$ and $\varepsilon_{\text{sampling}}$ for a single normalized transition at $\tilde{\nu}_{if} = 4064.574702$ cm^{-1}, considering various temperatures T, pressures P, numbers of points N_{points}, and bin sizes $\Delta\tilde{\nu}$, over a broad wavenumber range of [0, 8000] cm^{-1}. All sampling methods exhibit negligible $\varepsilon_{\text{sampling}}$, while all binned methods are expected to achieve perfect $\varepsilon_{\text{sampling}}$, at least within numerical error limits.

In cross-comparisons, the Gaussian line profile consistently proves to be the most accurate sampling line profile for preserving the profile area, even with a smaller number of points, due to its well-defined mathematical properties and smooth distribution. The SciPy Voigt and Humlíček Voigt profiles offer similar levels of accuracy. For constant pressure and grid number, the uncertainty in the line profiles decreases with increasing temperature, as shown in Figure 7.4 (first panel). Conversely, line profiles have a loss of accuracy as the pressure or grid number increases, as illustrated in the second and third panels. The binned Lorentzian and Voigt profiles tend to degrade at higher pressures due to the non-negligible wing cutoff (25 cm^{-1}) of their wings beyond the interval $[0, 8000]$ cm^{-1}.

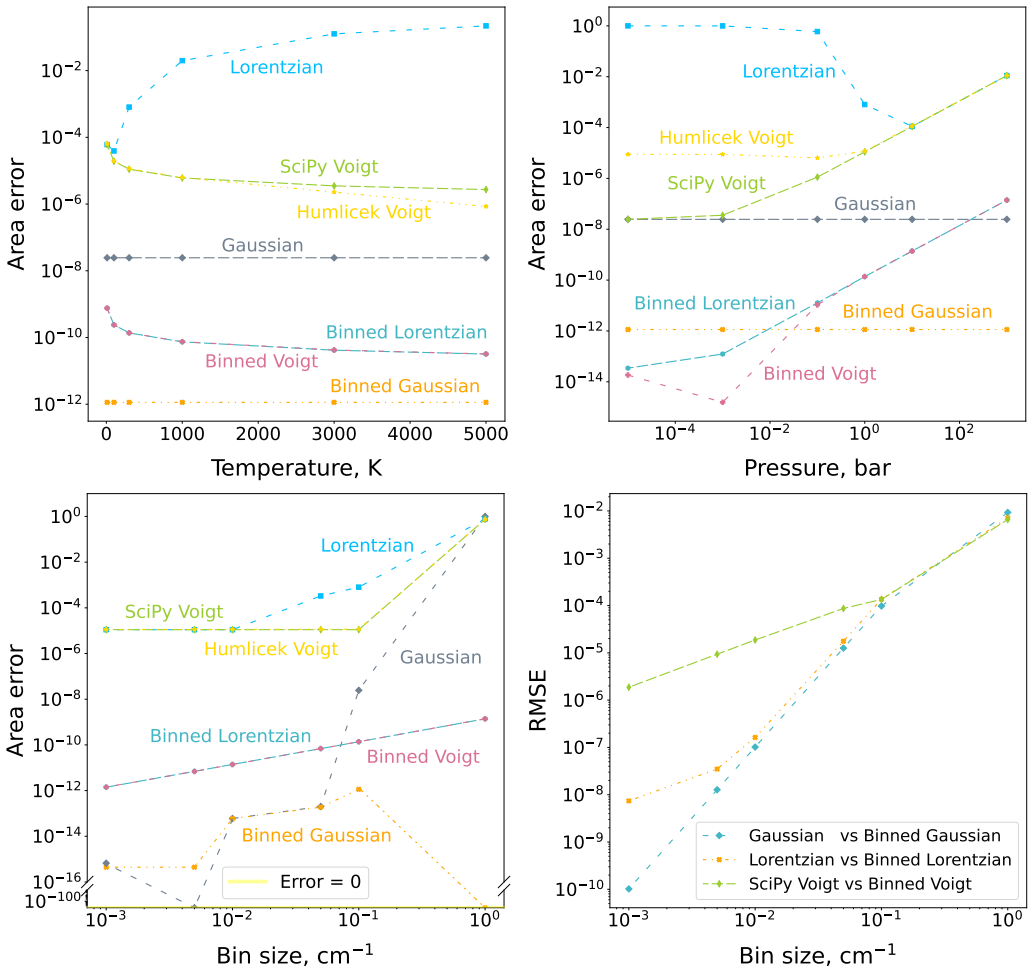


Figure 7.4: Line profile errors $\varepsilon_{\text{area}}$ (Eq. (7.51)) and $\varepsilon_{\text{sampling}}$ (Eq. (7.52)) are analyzed for different profiles representing the same single normalized transition ($\tilde{\nu}_{if} = 4064.574702\ \text{cm}^{-1}$), including Gaussian ($\alpha_{\text{G}} = 0.1$), Lorentzian (with γ_{L} estimated by $\gamma_{\text{ref}} = 0.07$ and $n_L = 0.5$), and Voigt profiles. These profiles are evaluated over a large wavenumber range of $[0, 8000]\ \text{cm}^{-1}$ at different temperatures T, pressures P, and bin sizes $\Delta\tilde{\nu}$. Top left: $\varepsilon_{\text{area}}$ for different temperatures T at $P = 1$ bar, with $N_{\text{points}} = 80\,001$ grid points (bin size $\Delta\tilde{\nu} = 0.1\ \text{cm}^{-1}$) without a wing cutoff. Top right: $\varepsilon_{\text{area}}$ for different pressures P at $T = 300$ K, with $\Delta\tilde{\nu} = 0.1\ \text{cm}^{-1}$ without a wing cutoff. Bottom left: $\varepsilon_{\text{area}}$ for different numbers of grid points N_{points} (bin sizes $\Delta\tilde{\nu}\ \text{cm}^{-1}$) at $T = 300$ K, $P = 1$ bar, without a wing cutoff. Bottom right: $\varepsilon_{\text{sampling}}$ for different numbers of grid points N_{points} (bin sizes $\Delta\tilde{\nu}\ \text{cm}^{-1}$) at $T = 300$ K, $P = 1$ bar, with a wing cutoff of $100\ \text{cm}^{-1}$.

The area errors and sampling errors $\varepsilon_{\text{sampling}}$ of different line profiles are displayed in the third and fourth panels of Figure 7.4. While the sampling profiles are nearly perfect and improve with an increased number of grid points, the binned profiles can introduce substantial errors, particularly at low resolution. Generally, in radiative transfer applications, direct sampling line profiles are preferred, especially

for coarser grids [373].

Pseudo-Voigt profiles are excluded from this analysis due to their relatively low accuracy.

7.7.3 A comparative analysis of program run times with different line profiles

To compare the program running times for computing cross sections with different line profiles, Table 7.1 provides examples of cross section calculations for different numbers of transitions, bin sizes (number of grid points), and filters used.

Table 7.1: Example of program run times on 32 cores, in s, for computing cross sections using different line profiles at a temperature of $T = 300$ K, pressures $P = 1$ bar, and a wing cutoff of 25 cm^{-1} for the ExoMol H_2O POKAZATEL line list [5], over the wavenumber range $[20000, 22000]$ cm^{-1}. The left panel compares the run times for different numbers of transitions $(10^4, 10^6, 10^8)$ with a bin size of 0.1 cm^{-1} (20001 grid points) without any filters. The middle panel compares the run times for for different numbers of transitions $(10^4, 10^6, 10^8)$ with a bin size of 0.1 cm^{-1} (20001 grid points) using uncertainty and threshold filters. The right four columns compare the run times for different bin sizes $(0.1, 0.01, 0.005, 0.001$ cm$^{-1})$ and corresponding numbers of grid points $(20001, 200001, 400001, 2000001)$ using 10^8 transitions with uncertainty and threshold filters.

Line profiles	Uncertainty filter		None	Uncertainty filter		0.01 cm^{-1}	Uncertainty filter		0.01 cm^{-1}
	Threshold filter		None	Threshold filter		10^{-30} cm molecule^{-1}	Threshold filter		10^{-30} cm molecule^{-1}
	Bin size $\Delta\tilde{\nu}$ =		0.1 cm^{-1}	Bin size $\Delta\tilde{\nu}$ =		0.1 cm^{-1}	$N_{\text{transitions}}$ =		10^8
	N_{points} =		20001	N_{points} =		20001	N_{points} =		$2000/\Delta\tilde{\nu}+1$
	$N_{\text{transitions}}$ =			$N_{\text{transitions}}$ =			Bin size (cm^{-1}) $\Delta\tilde{\nu}$ =		
	10^4	10^6	10^8	10^4	10^6	10^8	0.01	0.005	0.001
Doppler	2.70	6.79	54.63	3.01	2.64	8.00	7.66	9.42	13.43
Binned Doppler	4.37	10.52	114.23	3.98	4.00	10.12	9.45	12.65	19.66
Gaussian	2.84	7.43	52.53	3.74	3.85	9.37	8.72	11.09	15.00
Binned Gaussian	4.42	10.85	119.77	3.85	3.96	10.08	9.39	12.28	20.03
Lorentzian	2.83	7.39	51.69	3.87	3.97	9.47	8.77	10.46	14.15
Binned Lorentzian	4.14	9.16	61.24	3.95	3.95	9.67	9.24	11.31	14.73
SciPy Voigt	2.99	9.77	83.19	3.87	3.91	9.34	8.85	10.83	13.21
SciPy Wofz Voigt	3.10	10.26	121.01	3.80	3.97	9.71	9.09	11.73	17.27
Humlíček Voigt	3.24	10.69	161.25	3.90	4.05	10.15	12.62	20.03	46.62
Thompson pseudo-Voigt	2.99	9.89	85.86	3.80	3.90	10.14	9.27	11.63	19.85
Kielkopf pseudo-Voigt	4.01	10.15	85.98	3.82	3.94	10.08	9.57	13.09	19.72
Olivero pseudo-Voigt	3.93	9.93	88.55	3.81	3.89	10.01	9.58	12.48	20.44
Liu-Lin pseudo-Voigt	3.98	10.03	85.20	3.88	3.99	10.18	9.40	14.02	19.63
Rocco pseudo-Voigt	3.97	10.09	85.47	3.97	3.98	9.83	9.56	13.08	20.16
Binned Voigt	6.19	13.80	657.68	3.95	4.35	10.32	22.38	38.78	107.06

From Table 7.1, it is evident that the Lorentzian profile is the fastest, followed by the Doppler, Gaussian, and SciPy Voigt profiles. Pseudo-Voigt profile calculations use `NumExpr` [437], a fast numerical expression evaluator for `NumPy` [435], for data processing. `NumExpr` is designed to accelerate array operations and reduce memory usage compared to same calculations performed directly in Python

[437]. Additionally, its multi-threaded capabilities allow `NumExpr` to leverage all available CPU cores, typically resulting in significant performance improvements compared to `NumPy` [437]. The Humlíček Voigt profile, commonly used as an approximation method, runs relatively slowly. Due to the increased number of steps in the binned profile calculations, these methods require more time than the sampling methods.

The SciPy wofz Voigt, Humlíček Voigt, and Rocco pseudo-Voigt profiles offer similar accuracy but are significantly slower than the SciPy Voigt profile. Binned Voigt profiles provide higher accuracy than standard sampling Voigt profiles but costs considerably more computational time.

In summary, considering its higher accuracy and relatively short run times, the SciPy Voigt profile is recommended for calculating cross sections.

7.8 Summary

A high-resolution spectral line is characterized by three primary parameters: transition frequency, integrated intensity, and line profile, with the latter reflecting temperature- and pressure-dependent line broadening effects critical for practical applications. Temperature-dependent Doppler broadening results in a Gaussian and Doppler profile, while pressure broadening is modeled using a Lorentzian profile. The Voigt profile, a convolution of Gaussian and Lorentzian functions, accounts for both types of broadening. Lifetime effects contribute a Lorentzian component without altering the basic discussion of line shapes. Line profiles are integrable functions with normalized areas.

The PyExoCross tool supports various line profiles, including Doppler, generalized Gaussian, Lorentzian, and Voigt profiles, with the latter further explored through different computational methods in this chapter.

Chapter 8

Simulating Non-LTE Spectra

PyExoCross is designed to generate molecular cross sections (absorption or emission) across a grid of temperatures and pressures, using various line profiles (see Chapter 7) under both local thermodynamic equilibrium (LTE) (see Chapter 6) and non-local thermodynamic equilibrium (non-LTE) conditions. This chapter highlights new features added to PyExoCross, which allows for the calculation of non-LTE absorption and emission spectra intensities, in addition to cross sections.

8.1 Introduction

Previous studies on the retrieval of exoplanet atmospheres have generally assumed that the atmospheric species are in local thermodynamic equilibrium (LTE). It is well established by observation of the atmospheres of the Earth and other Solar System planets [476], including Venus [477], Mars [478], and the gas giants [479] that non-local thermodynamic equilibrium (non-LTE) effects can occur in planetary atmospheres, leading to spectral features that differ from those predicted by LTE assumptions [480]. Non-LTE phenomena are also detected in stellar atmospheres, comets [481], and the interstellar medium [482, 483]. The study of non-LTE spectroscopy dates back to Milne [484] and has since been developed through numerous significant contributions [485–492].

Maser (Microwave Amplification by Stimulated Emission of Radiation) emission is a well-documented astrophysical phenomenon that occurs under non-local

thermodynamic equilibrium (non-LTE) conditions. Unlike thermal radiation, which follows the Boltzmann distribution in LTE, maser emission arises due to population inversion, leading to stimulated amplification of radiation at specific wavelengths [493, 494]. This mechanism is commonly observed in astrophysical environments such as star-forming regions, circumstellar envelopes of evolved stars, and the nuclear regions of galaxies [495, 496]. Understanding maser emission in non-LTE conditions remains an active area of research, with ongoing efforts to refine models and incorporate additional physical mechanisms.

Certain physical features of some exoplanet atmospheres can drive molecules to non-LTE states, such as:

- In high altitude shock regions at the boundaries of rapidly moving jet streams, molecules tend to lose energy primarily in their rotational mode rather than their vibrational mode, resulting in non-LTE state distributions.
- In regions of the atmosphere subjected to intense stellar radiation, stellar pumping selectively energizes the vibrational modes of molecules, pushing them into non-LTE states.

These non-LTE molecular states generate emission and transmission spectra that significantly deviate from their LTE counterparts, and these differences can be detected under suitable conditions.

8.2 Populations and densities

PyExoCross can simulate non-LTE spectra using a simple two-temperature approach which is between rotational and vibrational (vibronic) temperatures when computing intensities, partition functions, or other temperature-dependent properties [373]. Different electronic, vibronic, or vibrational bands can be individually simulated using an efficient filtering scheme based on quantum numbers.

For this purpose, the total energy is approximated as the sum of the vibrational (or vibronic) and rotational contributions [497].

$$\tilde{E}_{v,J,k} = \tilde{E}_{v}^{\mathrm{vib}} + \tilde{E}_{J,k}^{v,\mathrm{rot}}. \tag{8.1}$$

- v: Generic vibrational (vibronic) quantum numbers.
- J: Total angular momentum.
- k: Rotational quantum numbers.

The pure vibronic contributions are considered as the corresponding energy values at $J = 0$ for integer spins, $J = 1/2$ for non-integer (half-integer) spins, or the lowest J permitted by the symmetry of the electronic term and parity, typically associated with the lowest states (usually "+" or "e") [373]. In some cases, certain spectral lines contain NaN values for specific quantum numbers in the `.states` file. The program PyExoCross typically ignores these lines during processing.

The rotational contribution is then simply given by:

$$\tilde{E}^{v,\mathrm{rot}}_{J,k} = \tilde{E}_{v,J,k} - \tilde{E}^{\mathrm{vib}}_{v}. \tag{8.2}$$

The spectral simulations are conducted under three different cases, as summarized in Table 8.1.

Table 8.1: Summary of the three cases considered for spectral simulation.

Case	Distribution	T_{rot}	T_{vib}	Relation	Section	Reference
LTE	Boltzmann distribution	✓	✓	$T = T_{\mathrm{rot}} = T_{\mathrm{vib}}$	Section 8.2.1	[498]
Non-LTE	Treanor distribution	✓	✓	$T_{\mathrm{rot}} \neq T_{\mathrm{vib}}$	Section 8.2.2	[497]
Non-LTE	Vibronic distribution	✓	✗		Section 8.2.3	[499]

8.2.1 LTE populations using the Boltzmann distributions

For the LTE case, the rotational and vibrational temperatures are identical, and the populations of molecular states follow the Boltzmann distributions at a given temperature T, as given by [480]:

$$F_{J,v,k}(T) = \frac{g^{\mathrm{tot}}_{J,v,k} e^{-c_2 \tilde{E}_{J,v,k}/T}}{Q(T)}, \tag{8.3}$$

where $g^{\mathrm{tot}}_{J,v,k}$ is the total degeneracy and $\tilde{E}_{J,v,k} = E_{J,v,k}/hc$ is the state energy term value (cm^{-1}). $c_2 = hc/k_B$ is the second radiation constant (cm K) and k_B is the

Boltzmann constant (erg/K). T is the temparture in K and $Q(T)$ is the temperature-dependent partition function. The LTE partition function $Q(T)$ is defined as a sum over states.

$$Q(T) = \sum_n g^{\text{tot}}_{J,v,k} e^{-c_2 \tilde{E}_{J,v,k}/T}. \tag{8.4}$$

8.2.2 Non-LTE populations using two-temperature Treanor distributions

To extend this framework to account for both vibrational and rotational temperatures, I use a simplified bi-temperature Treanor approximation model [500] and investigate the spectral signatures of different vibrational bands contributing to a generic non-LTE spectrum. In the Treanor approximation, the vibrational populations are modeled using a vibrational Boltzmann distribution at a distinct temperature T_{vib}. The total non-LTE population of a given state is then represented as the product of two Boltzmann distributions:

$$F_{J,v,k}(T_{\text{vib}}, T_{\text{rot}}) = \frac{g^{\text{tot}}_{J,v,k} e^{-c_2 \tilde{E}^{\text{vib}}_v / T_{\text{vib}}} e^{-c_2 \tilde{E}^{v,\text{rot}}_{J,k} / T_{\text{rot}}}}{Q(T_{\text{vib}}, T_{\text{rot}})}. \tag{8.5}$$

This non-LTE partition function $Q(T_{\text{vib}}, T_{\text{rot}})$ depends on two different temperatures T_{vib} and T_{rot}.

$$Q(T_{\text{vib}}, T_{\text{rot}}) = \sum_n g^{\text{tot}}_{J,v,k} e^{-c_2 \tilde{E}^{\text{vib}}_v / T_{\text{vib}}} e^{-c_2 \tilde{E}^{v,\text{rot}}_{J,k} / T_{\text{rot}}}. \tag{8.6}$$

8.2.3 Non-LTE populations using custom vibronic densities

In this case, non-LTE conditions involve the use of Vibronic distribution and rotational temperature. The vibrational densities n_{vib} can be applied as directly incorporated into the calculations through the custom provided `.states` file. The sum of the non-LTE vibronic population fraction (vibrational densities) n_{vib} is normalized to one.

$$\sum_i n^{\text{vib}}_i = 1. \tag{8.7}$$

The rotational states are assumed to be populated according to a Boltzmann distribution corresponding to a rotational temperature T_{rot}, which also represents the kinetic temperature of the environment [501]. The total population of J, v, and k is then approximated as the product of these distributions [480].

$$F_{J,v,k}(T_{\text{rot}}) = \frac{g^{\text{tot}}_{J,v,k} n_{\text{vib}} e^{-c_2 \tilde{E}^{v,\text{rot}}_{J,k}/T_{\text{rot}}}}{Q(T_{\text{rot}})}. \tag{8.8}$$

In this case, the non-LTE partition function is defined with the population density when runs over all J, v, and k.

$$Q(T_{\text{rot}}) = \sum_n g^{\text{tot}}_{J,v,k} n_{\text{vib}} e^{-c_2 \tilde{E}^{v,\text{rot}}_{J,k}/T_{\text{rot}}}. \tag{8.9}$$

8.3 LTE and non-LTE absorption and emission intensities

This section provides a more detailed presentation of the equations used to calculate absorption and emission intensities in both LTE and non-LTE conditions.

	Absorption intensity		*Emission intensity*
LTE	$I_{f\leftarrow i} = \frac{g'_{J,v,k} A_{fi}}{8\pi c \tilde{\nu}^2_{fi}} \frac{e^{-c_2 \tilde{E}''_{J,v,k}/T}\left(1 - e^{-c_2 \tilde{\nu}_{fi}/T}\right)}{Q(T)}$		$\varepsilon_{i\leftarrow f} = \frac{g'_{J,v,k} hc A_{fi} \tilde{\nu}_{fi}}{4\pi} \frac{e^{-c_2 \tilde{E}'_{J,v,k}/T}}{Q(T)}$
	$= \frac{g'_{J,v,k} A_{fi}}{8\pi c \tilde{\nu}^2_{fi}} \frac{1}{g''_{J,v,k}} F_{J'',v'',k''}(T)\left(1 - e^{-c_2 \tilde{\nu}_{fi}/T}\right)$		$= \frac{hc}{4\pi} g'_{J,v,k} A_{fi} \tilde{\nu}_{fi} F_{J',v',k'}(T)$
Non-LTE	$I_{f\leftarrow i} = \frac{g'_{J,v,k} A_{fi}}{8\pi c \tilde{\nu}^2_{fi}} \frac{e^{-c_2 \tilde{E}''^{\text{vib}}_v/T_{\text{vib}}} e^{-c_2 \tilde{E}''^{v,\text{rot}}_{J,k}/T_{\text{rot}}}\left(1 - e^{-c_2 \tilde{\nu}_{fi}/T}\right)}{Q(T_{\text{vib}}, T_{\text{rot}})}$		$\varepsilon_{i\leftarrow f} = \frac{hc}{4\pi} g'_{J,v,k} A_{fi} \tilde{\nu}_{fi} \frac{e^{-c_2 \tilde{E}'^{\text{vib}}_v/T_{\text{vib}}} e^{-c_2 \tilde{E}'^{v,\text{rot}}_{J,k}/T_{\text{rot}}}}{Q(T_{\text{vib}}, T_{\text{rot}})}$
	$I_{f\leftarrow i} = \frac{g'_{J,v,k} A_{fi}}{8\pi c \tilde{\nu}^2_{fi}} \frac{n''_{\text{vib}} e^{-c_2 \tilde{E}''^{v,\text{rot}}_{J,k}/T_{\text{rot}}}\left(1 - e^{-c_2 \tilde{\nu}_{fi}/T}\right)}{Q(T_{\text{rot}})}$		$\varepsilon_{i\leftarrow f} = \frac{hc}{4\pi} g'_{J,v,k} A_{fi} \tilde{\nu}_{fi} \frac{n'_{\text{vib}} e^{-c_2 \tilde{E}'^{v,\text{rot}}_{J,k}/T_{\text{rot}}}}{Q(T_{\text{rot}})}$
	$I_{f\leftarrow i} = \frac{g'_{J,v,k} A_{fi}}{8\pi c \tilde{\nu}^2_{fi}} \frac{1}{g''_{J,v,k}} F_{J'',v'',k''}(T_{\text{vib}}, T_{\text{rot}})\left(1 - e^{-c_2 \tilde{\nu}_{fi}/T}\right)$		$\varepsilon_{i\leftarrow f} = \frac{hc}{4\pi} A_{fi} \tilde{\nu}_{fi} F_{J',v',k'}(T_{\text{vib}}, T_{\text{rot}})$

Figure 8.1: Comparison the equations of LTE and non-LTE absorption and emission intensities.

It elaborates on the specific mathematical formulations for each case, highlighting the differences in the treatment of molecular energy state populations, temperature distributions, and the factors influencing radiative transfer in these two regimes. Figure 8.1 illustrates a comparison of the LTE and non-LTE equations for absorption and emission intensities, with color highlighting to clearly show the parts that can be replaced or modified between these cases.

8.3.1 LTE and non-LTE absorption intensities

To simulate LTE absorption spectra, the Boltzmann distributions are applied to the expression for absorption line intensities [373, 428].

$$\begin{aligned} I_{f\leftarrow i} &= \frac{g'_{J,v,k}A_{fi}}{8\pi c\tilde{\nu}_{fi}^2}\frac{e^{-c_2\tilde{E}''_{J,v,k}/T}(1-e^{-c_2\tilde{\nu}_{fi}/T})}{Q(T)}, \\ &= \frac{g'_{J,v,k}A_{fi}}{8\pi c\tilde{\nu}_{fi}^2}\frac{1}{g''_{J,v,k}}F_{J'',v'',k''}(T)(1-e^{-c_2\tilde{\nu}_{fi}/T}). \end{aligned} \tag{8.10}$$

$g'_{J,v,k}$ and $g''_{J,v,k}$ represent the total degeneracies of the upper and lower states, respectively. $F_{J'',v'',k''}(T)$ is expressed as Eq. (8.3) for lower state and can also be extracted from the custom offered `.states` file.

8.3.1.1 Non-LTE absorption intensities using two-temperature Treanor distribution

In the two-temperature Treanor distribution model, its partition function $Q(T_{\text{vib}}, T_{\text{rot}})$ can be determined using Eq. (8.6). The upper state total degeneracy $g'_{J,v,k}$ and the lower state energy $\tilde{E}''$ are used in the following equation.

$$I_{f\leftarrow i} = \frac{g'_{J,v,k}A_{fi}}{8\pi c\tilde{\nu}_{fi}^2}\frac{e^{-c_2\tilde{E}''^{\text{vib}}_{v}/T_{\text{vib}}}e^{-c_2\tilde{E}''^{\text{v,rot}}_{j,k}/T_{\text{rot}}}(1-e^{-c_2\tilde{\nu}_{fi}/T})}{Q(T_{\text{vib}}, T_{\text{rot}})}. \tag{8.11}$$

8.3.1.2 Non-LTE absorption intensities using custom vibrational (vibronic) density

This method applies a rotational temperature T_{rot}. The total degeneracy $g'_{J,v,k}$ corresponds to the upper state, while the energy and the custom vibrational (vi-

bronic) density n''_{vib}, as defined in Eq. (8.7), corresponds to the lower states. The rotational temperature-dependent partition function $Q(T_{\text{rot}})$ is defined as Eq. (8.9).

$$I_{f\leftarrow i} = \frac{g'_{J,v,k}A_{fi}}{8\pi c\tilde{\nu}_{fi}^2}\frac{n''_{\text{vib}}e^{-c_2\tilde{E}''^{\text{v,rot}}_{j,k}/T_{\text{rot}}}(1-e^{-c_2\tilde{\nu}_{fi}/T})}{Q(T_{\text{rot}})}. \tag{8.12}$$

8.3.1.3 Non-LTE absorption intensities using custom rovibrational (rovibronic) population

The non-LTE custom rovibrational (rovibronic) population $F_{J'',v'',k''}(T_{\text{vib}},T_{\text{rot}})$ for the lower state is used in this method which is definded in Eq. (8.5). Additionally, the total degeneracies for both the upper and lower states, $g'_{J,v,k}$ and $g''_{J,v,k}$, are incorporated.

$$I_{f\leftarrow i} = \frac{g'_{J,v,k}A_{fi}}{8\pi c\tilde{\nu}_{fi}^2}\frac{1}{g''_{J,v,k}}F_{J'',v'',k''}(T_{\text{vib}},T_{\text{rot}})(1-e^{-c_2\tilde{\nu}_{fi}/T}). \tag{8.13}$$

8.3.2 LTE and non-LTE emission intensities

To simulate LTE emission spectra, the Boltzmann distributions are applied to the expression for emission line intensities [373].

$$\begin{aligned}\varepsilon_{i\leftarrow f} &= \frac{g'_{j,v,k}hcA_{fi}\tilde{\nu}_{fi}}{4\pi}\frac{e^{-c_2\tilde{E}'_{j,v,k}/T}}{Q(T)},\\ &= \frac{g'_{j,v,k}hcA_{fi}\tilde{\nu}_{fi}}{4\pi}F_{J',v',k'}(T).\end{aligned} \tag{8.14}$$

$g'_{J,v,k}$ and $g''_{J,v,k}$ represent the total degeneracies of the upper and lower states, respectively. $F_{J',v',k'}(T)$ is expressed as Eq. (8.3) for upper state and can also be extracted from the custom offered `.states` file.

8.3.2.1 Non-LTE emission intensities using two-temperature Treanor distribution

In the two-temperature Treanor distribution model, its partition function $Q(T_{\text{vib}},T_{\text{rot}})$ can be determined using Eq. (8.6). The upper state total degeneracy

$g'_{J,v,k}$ and energy $\tilde{E}'$ are used in this model.

$$\varepsilon_{i\leftarrow f} = \frac{hc}{4\pi} g'_{J,v,k} A_{fi} \tilde{\nu}_{fi} \frac{e^{-c_2 \tilde{E}'^{\text{vib}}_v / T_{\text{vib}}} e^{-c_2 \tilde{E}'^{\text{v,rot}}_{j,k} / T_{\text{rot}}}}{Q(T_{\text{vib}}, T_{\text{rot}})}. \tag{8.15}$$

8.3.2.2 Non-LTE emssion intensities using custom vibrational (vibronic) density

Here, total degeneracy $g'_{J,v,k}$, energy $\tilde{E}'$, and custom vibrational (vibronic) density n'_{vib} (see Eq. (8.7)) are all corresponded to the upper states. The rotational temperature-dependent partition function $Q(T_{\text{rot}})$ is determined as Eq. (8.9).

$$\varepsilon_{i\leftarrow f} = \frac{hc}{4\pi} g'_{J,v,k} A_{fi} \tilde{\nu}_{fi} \frac{n'_{\text{vib}} e^{-c_2 \tilde{E}'^{\text{v,rot}}_{j,k} / T_{\text{rot}}}}{Q(T_{\text{rot}})}. \tag{8.16}$$

8.3.2.3 Non-LTE emission intensities using custom rovibrational (rovibronic) population

The upper state custom rovibrational (rovibronic) population $F_{J',v',k'}(T_{\text{vib}}, T_{\text{rot}})$ can be obtained from a user-provided `.states` file.

$$\varepsilon_{i\leftarrow f} = \frac{hc}{4\pi} A_{fi} \tilde{\nu}_{fi} F_{J',v',k'}(T_{\text{vib}}, T_{\text{rot}}). \tag{8.17}$$

8.4 Results and discussion

Table 8.2 shows the essential components of the non-LTE stick spectra file (`.nlte`), while Table 8.3 provides an example of the non-LTE absorption stick spectra output file for CN Trihybrid [9] ExoMol line list.

Figure 8.2 and Figure 8.3 present comparisons of absorption and emsission intensities for ExoMol CN line list Trihybrid [9] under LTE and non-LTE conditions at an equilibrium temperature of $T_{\text{vib}} = 1000$ K and a pressure of 1 bar. For the non-LTE case, a rotational temperature of 296 K is assumed. This scenario represents a significant temperature contrast, selected to illustrate the potential non-LTE effects that may arise in exoplanet atmospheres.

Table 8.2: Specification of the mandatory part of the `.nlte` non-LTE stick spectra file with the extra columns for the optional quantum numbers.

Field	Fortran format	C format	Description
$\tilde{\nu}$	`F12.6`	`%12.6f`	Frequency (wavenumber) in cm^{-1}
$I_{\rm NLTE}$	`ES14.8`	`%14.8E`	Non-LTE intensity in cm molecule^{-1}
J'	`I7/F7.1`	`%7d/%7.1f`	J-quantum number (integer/half-integer) for upper state
J''	`I7/F7.1`	`%7d/%7.1f`	J-quantum number (integer/half-integer) for lower state
(QN$'$)			Extra quantum numbers for upper state, any format (optional)
(QN$''$)			Extra quantum numbers for lower state, any format (optional)

J: Total angular momentum quantum, excluding nuclear spin.

Table 8.3: Extract from the `.nlte` non-LTE stick spectra file generated from the emission ^{12}C^{14}N Trihybrid [9] ExoMol line list.

$\tilde{\nu}_{fi}$	$I_{\rm NLTE}$	J'	J''	e/f$'$	v'	ElecState$'$	e/f$''$	v''	ElecState$''$
7844.210400	4.50771601E-25	18.5	18.5	f	6	X(2SIGMA+)	e	2	X(2SIGMA+)
7844.214095	8.75139200E-29	32.5	31.5	e	6	X(2SIGMA+)	e	2	X(2SIGMA+)
7844.218880	1.92683803E-33	64.5	64.5	f	2	A(2PI)	e	2	X(2SIGMA+)
7844.219980	7.36224253E-32	8.5	8.5	f	11	A(2PI)	e	6	A(2PI)
7844.227094	5.63734119E-27	31.5	30.5	f	6	X(2SIGMA+)	f	2	X(2SIGMA+)
7844.243430	1.41189130E-45	57.5	58.5	e	2	B(2SIGMA+)	e	12	X(2SIGMA+)
7844.248580	6.41842242E-37	9.5	9.5	e	4	B(2SIGMA+)	f	14	X(2SIGMA+)
7844.292230	2.27096176E-34	21.5	21.5	e	3	B(2SIGMA+)	f	9	A(2PI)
7844.310100	1.23184821E-51	77.5	76.5	e	11	A(2PI)	e	6	A(2PI)
7844.338050	1.19603430E-25	19.5	18.5	e	6	X(2SIGMA+)	e	2	X(2SIGMA+)

$\tilde{\nu}_{fi}$: Absorption frequency (wavenumber) (in cm^{-1}).
$I_{\rm NLTE}$: Non-LTE absorption intensity (in cm molecule^{-1}).
J' and J'': Total angular momentum quantum number for upper and lower state.
e/f$'$ and e/f$''$: Rotationless parity for upper and lower state.
v' and v'': State vibrational quantum number for upper and lower state.
ElecState$'$ and ElecState$''$: Notation of the electronic state for upper and lower state.

The non-LTE cross sections were generated using the ExoMol CN line list Trihybrid [9] at 1 bar of pressure. Figure 8.4 and Figure 8.6 compare the absorption and emission cross sections under LTE and non-LTE conditions, calculated using a Gaussian line profile with a Doppler half-width at half-maximum (HWHM) of 0.5. Figure 8.5 and Figure 8.7 provide a magnified view of the selected regions. The equilibrium temperature is set to $T_{\rm vib} = 1000$ K, with a pressure of 1 bar. For the non-LTE case, a rotational temperature of $T_{\rm rot} = 296$ K is assumed. This scenario, characterized by a large temperature contrast, is chosen to highlight the potential non-LTE effects that may occur in exoplanet atmospheres.

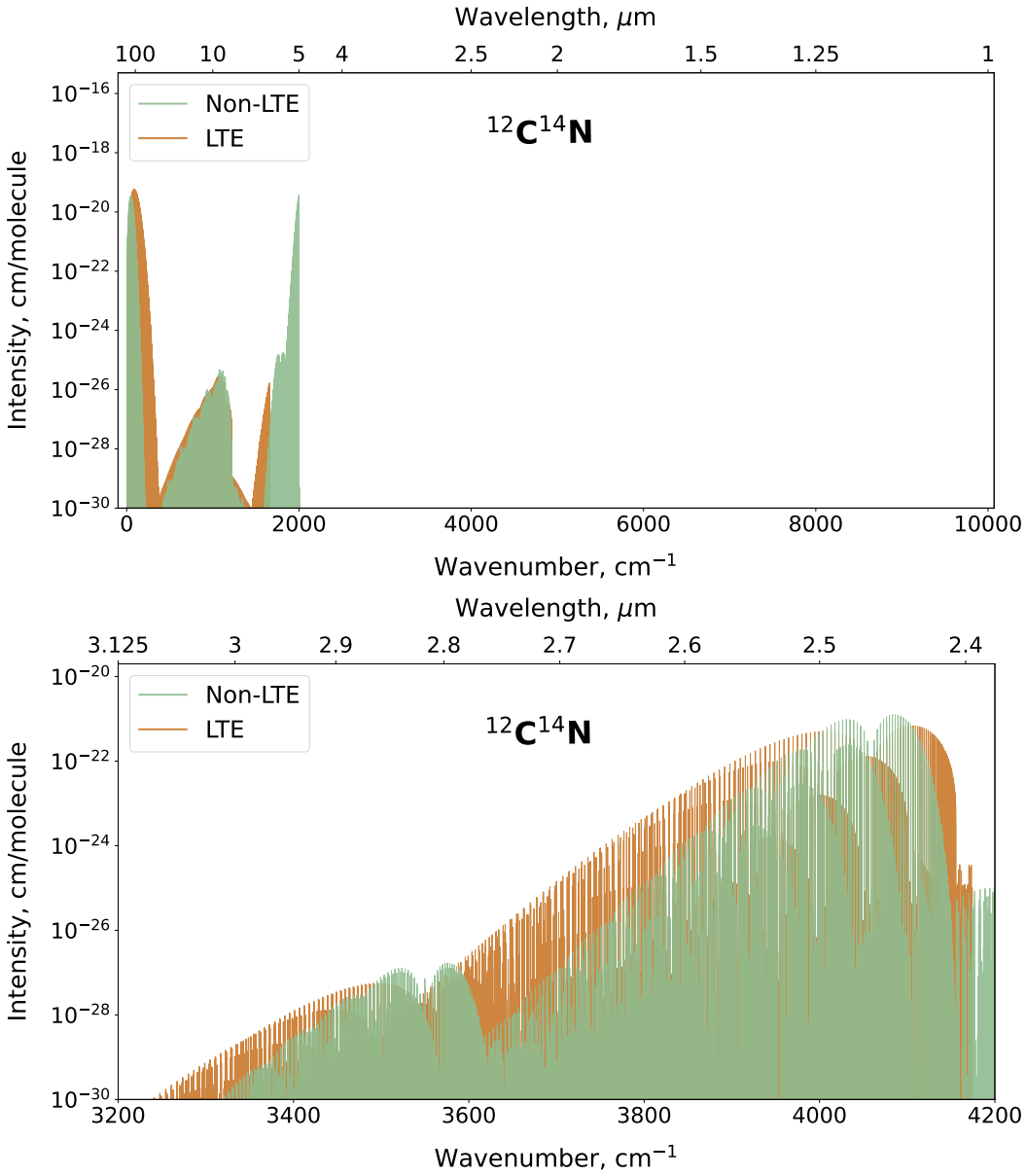


Figure 8.2: Absorption stick spectra at a pressure of 1 bar for CN [9] in LTE at $T = 1000$ K and non-LTE at equilibrium temperature ($T_{\text{vib}} = 1000$ K), with a rotational temperature of $T_{\text{rot}} = 296$ K for the non-LTE case in the $[0, 10000]$ cm^{-1} region (top display) and zoomed in $[3200, 4200]$ cm^{-1} region (bottom display).

8.5 Summary

Non-local thermodynamic equilibrium (non-LTE) calculations are essential in plasma modeling to accurately describe systems where the assumption of local thermodynamic equilibrium does not hold. In such conditions, the populations of atomic and molecular energy levels deviate from the Boltzmann distribution, neces-

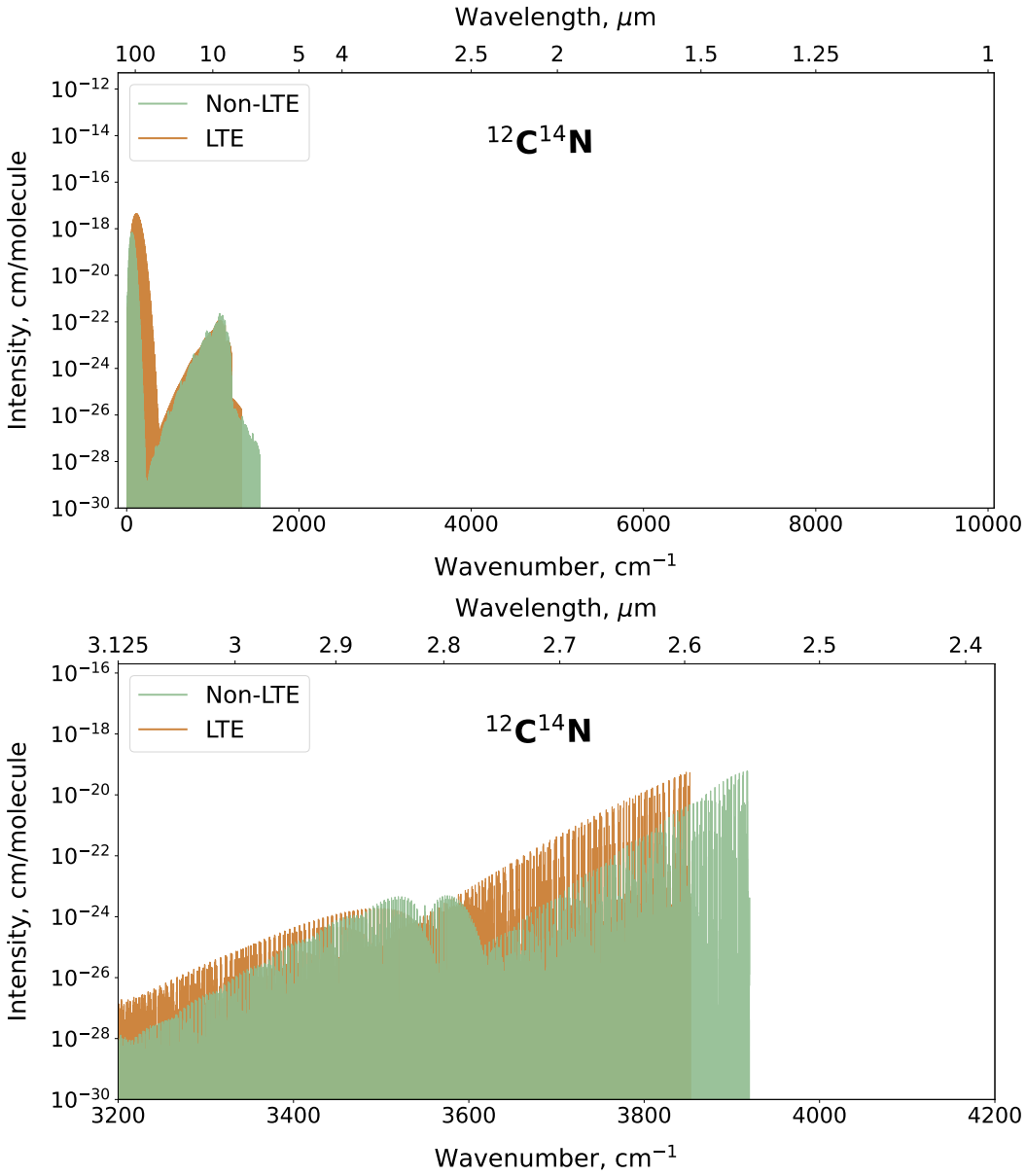


Figure 8.3: Emission stick spectra at a pressure of 1 bar for CN [9] in LTE at $T = 1000$ K and non-LTE at equilibrium temperature ($T_{\text{vib}} = 1000$ K), with a rotational temperature of $T_{\text{rot}} = 296$ K for the non-LTE case in the $[0, 10000]$ cm^{-1} region (top display) and zoomed in $[3200, 4200]$ cm^{-1} region (bottom display).

sitating detailed non-LTE modeling to predict plasma behavior accurately.

The Quantemol Database (QDB) (https://quantemol.com/quantemol-db/) provides plasma chemistries and reaction data, facilitating the development of non-LTE models by offering validated chemistry sets and cross-sectional data for various plasma processes. This resource supports the creation of detailed non-LTE models

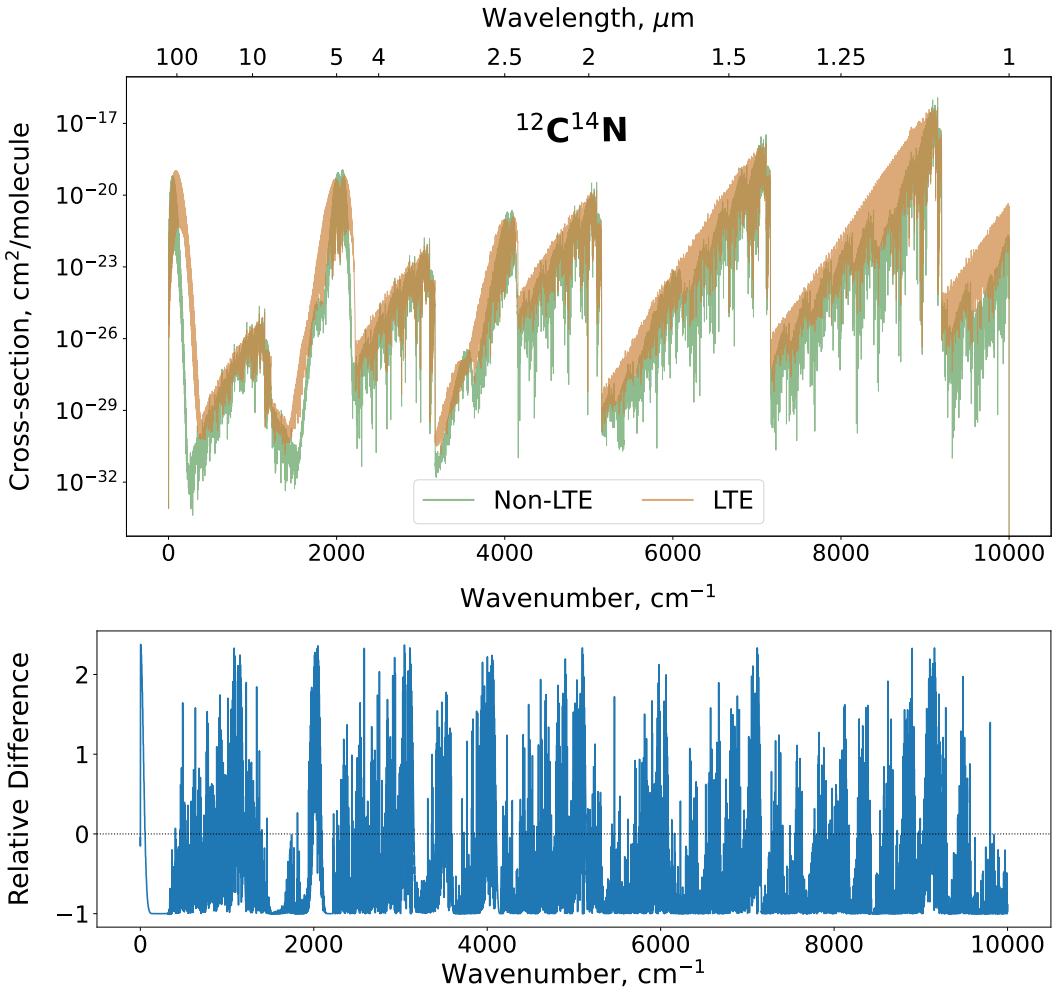


Figure 8.4: Absorption cross sections at a pressure of 1 bar for CN [9] in LTE at $T = 1000$ K and non-LTE at equilibrium temperature ($T_{\text{vib}} = 1000$ K), with a rotational temperature of $T_{\text{rot}} = 296$ K for the non-LTE case in the $[0, 10000]$ cm^{-1} region (top display) and the relative differences (bottom display).

by supplying necessary reaction rates and mechanisms [369], but obtaining sufficient state-dependent collisional/reaction data has proved problematic.

In the context of maser emissions, non-LTE modeling is crucial as masers operate under conditions where population inversion occurs, leading to stimulated emission. Accurate modeling of these phenomena requires incorporating non-LTE effects to simulate the vibronic distributions and rotational temperatures that influence maser activity. Therefore, integrating non-LTE calculations with resources like QDB enhances the accuracy of plasma models, particularly in applications involving maser emissions, by providing detailed insights into the underlying physical and

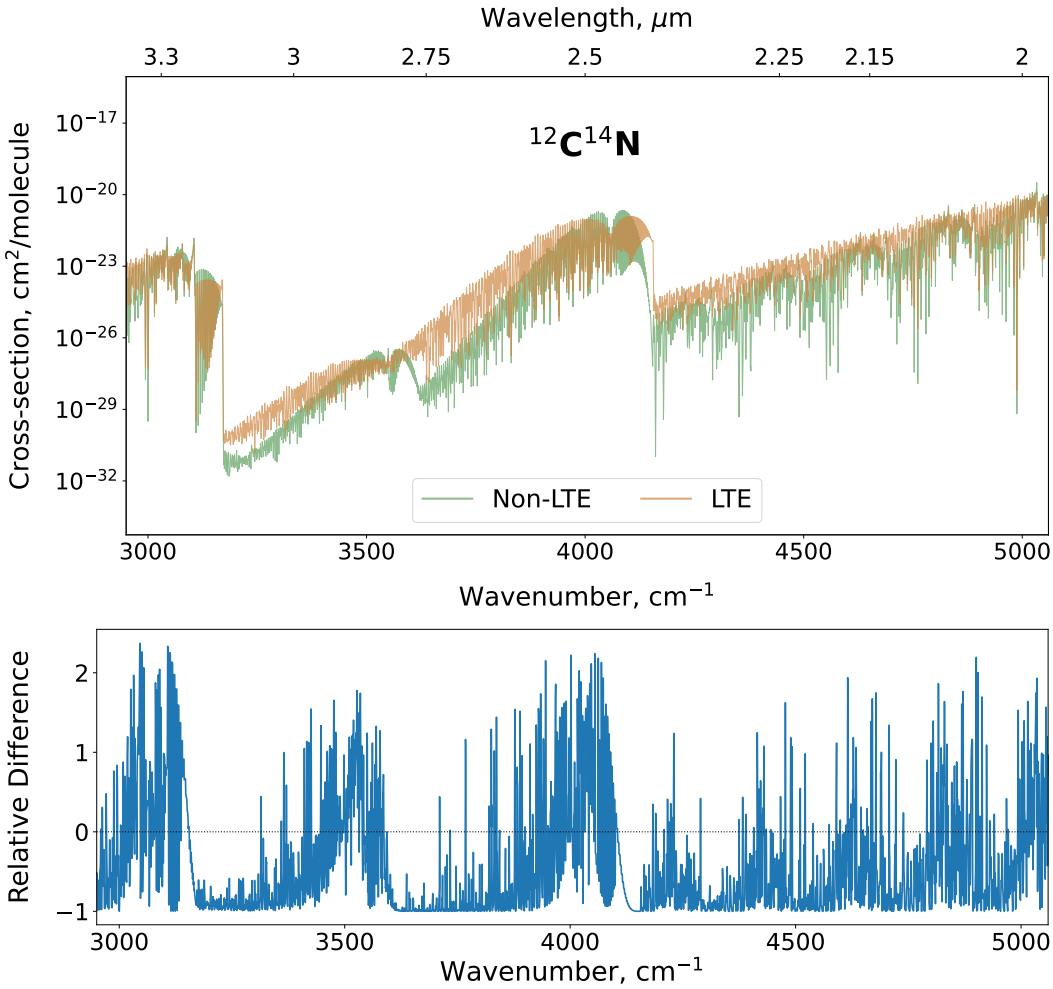


Figure 8.5: Absorption cross sections at a pressure of 1 bar for CN [9] in LTE at $T = 1000$ K and non-LTE at equilibrium temperature ($T_{\text{vib}} = 1000$ K), with a rotational temperature of $T_{\text{rot}} = 296$ K for the non-LTE case in the $[3000, 5000]$ cm^{-1} region (top display) and the relative differences (bottom display).

chemical processes.

Non-LTE effects are common in planetary atmospheres and can arise in various scenarios. Instead of examining each case individually, which may not always be feasible, the model presented here characterizes the overall non-LTE effect through a bi-temperature parametrization, a method widely used in such cases. By treating rotational and vibrational temperatures as separate parameters in a two-temperature Boltzmann model, this approach provides an efficient and robust way to incorporate an additional degree of freedom for modeling non-LTE spectral effects. Moreover, the functional form of the temperature-dependent term can be further explored. This

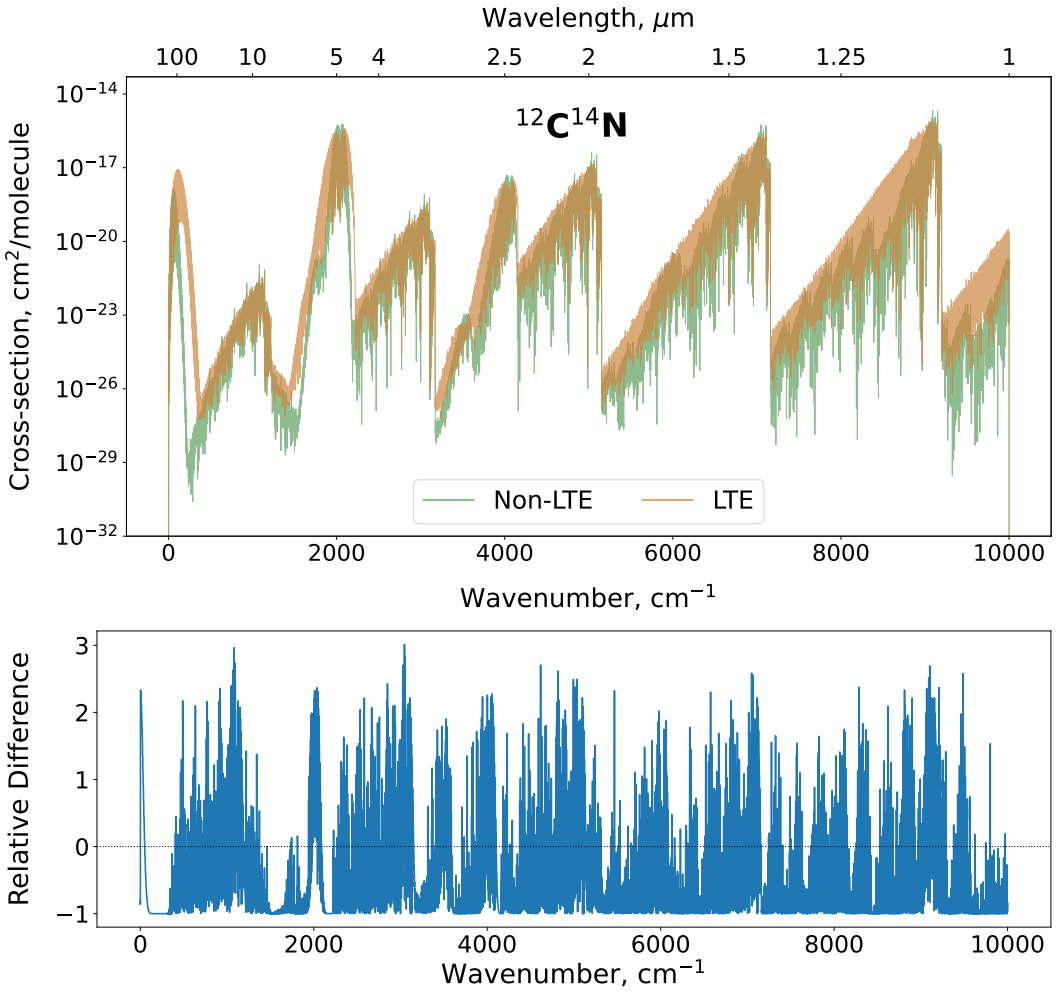


Figure 8.6: Emission cross sections at a pressure of 1 bar for CN [9] in LTE at $T = 1000$ K and non-LTE at equilibrium temperature ($T_{\text{vib}} = 1000$ K), with a rotational temperature of $T_{\text{rot}} = 296$ K for the non-LTE case in the $[0, 10000]$ cm^{-1} region (top display) and the relative differences (bottom display).

enables the determination of non-LTE absorption characteristics for molecules using the two-temperature model, without the need for explicitly calculating non-LTE population densities, which would require detailed a priori knowledge of specific physical mechanisms within atmospheric scenarios—knowledge that is often not feasible in the context of exoplanetary studies. However, users can still provide pre-existing vibrational density or rotational population data for the program to directly read and use in the calculations.

PyExoCross offers three methods for computing non-LTE stick spectra and cross-sections: the two-temperature Treanor distribution, a custom vibrational den-

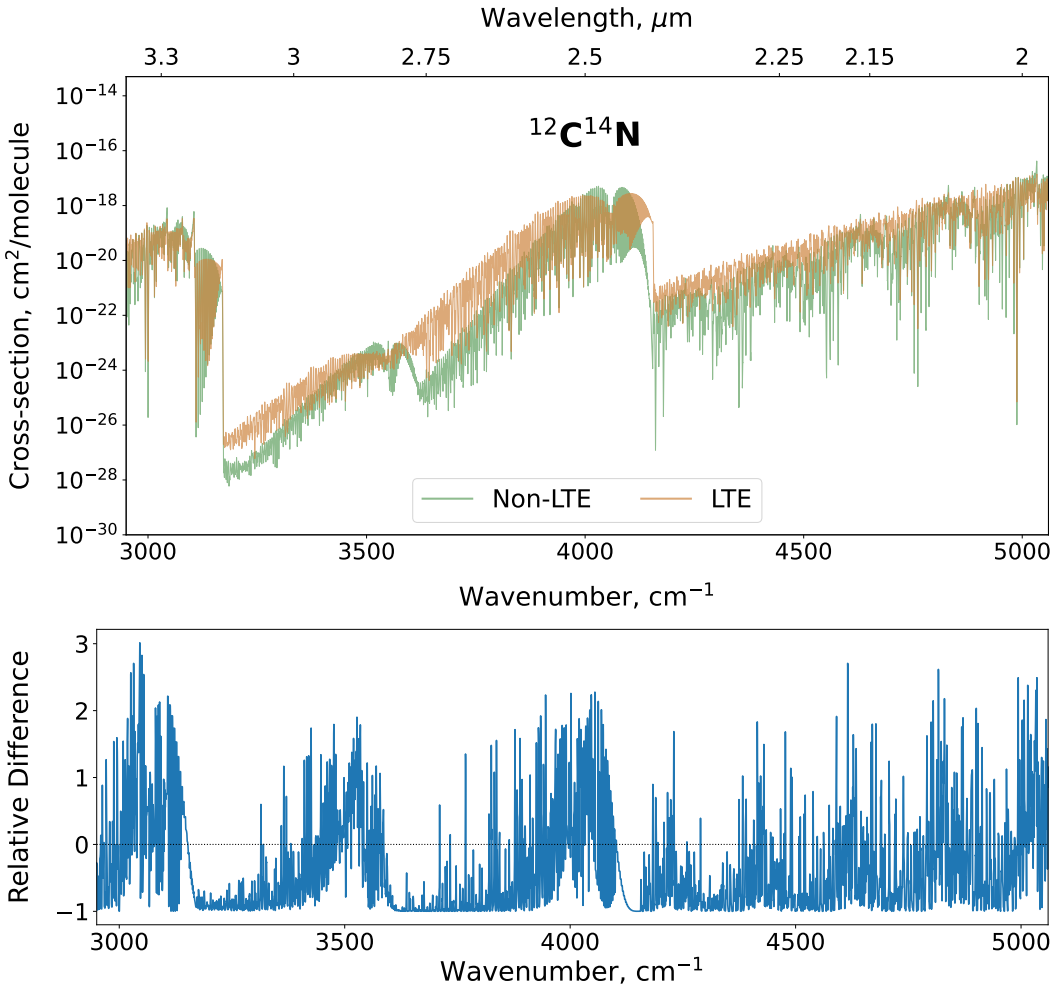


Figure 8.7: Emission cross sections at a pressure of 1 bar for CN [9] in LTE at $T = 1000$ K and non-LTE at equilibrium temperature ($T_{\rm vib} = 1000$ K), with a rotational temperature of $T_{\rm rot} = 296$ K for the non-LTE case in the $[3000, 5000]$ cm^{-1} region (top display) and the relative differences (bottom display).

sity, and a custom rotational population. The density and population provided by users are treated as input files. Currently, PyExoCross supports this functionality exclusively for data in the ExoMol format. Future versions will consider incorporating support for the HITRAN format.

Chapter 9

A High-Resolution ExoMol Database: ExoMolHR

ExoMolHR is an empirical high-resolution molecular spectrum database designed for high-temperature molecular line lists available from the ExoMol database. Where available, uncertainties in the recommended ExoMol datasets are used to select highly accurate spectral lines, which mainly rely on empirical energy levels derived through the Measured Active Rotation Vibration Energy Levels (MARVEL) database [502, 503]. This approach systematically improves the energy and transition data provided by the ExoMol database [17].

The ExoMolHR database provides a collection of high-resolution molecular transitions which can be interactively interrogated. This database offers line positions with calculated intensities for user-specified wavenumber/wavelength ranges and temperatures. Users can plot spectra directly on the freely accessible ExoMolHR website at https://www.exomol.com/exomolhr/ or download the data as CSV files. Additionally, cross sections can be computed using the Python program PyExoCross. The ExoMolHR database currently includes 24 307 135 spectral lines for 58 isotopologues from 33 molecules, with these numbers expected to rapidly grow as more MARVEL studies are completed and the ExoMol database continues to be updated and its line position accuracy improved.

9.1 Introduction

Tennyson and Yurchenko [18] initiated the ExoMol database to provide molecular line lists for exoplanet and other hot atmospheres. The project originally prioritized the completeness of these line lists over their accuracy [429]. In recent years, however, experimental data have been increasingly incorporated to improve the accuracy of the line positions provided by the database, as demonstrated in McKemmish et al. [504]. Consequently, the most recent ExoMol data release [17] claims 41 molecules with highly accurate transition wavenumbers for at least one isotopologue. The molecules included in the ExoMol database are selected based on their astronomical importance, and the data cover extensive temperature ranges for studies. The inclusion of high accuracy data marks a significant advance; nonetheless, the ExoMol database is huge, encompassing around 10^{13} transitions, far exceeding the size of other databases discussed earlier in Chapter 2. The high accuracy lines represent only a very small fraction of the total, currently 24 307 135 transitions. Due to its extensive volume, the ExoMol database is impractical for detailed spectroscopic analysis such as initial line assignments. This work aims to facilitate easy access to the high-resolution molecular data within the ExoMol database, which can be used for the assignment and analysis of strong spectral lines.

ExoMol has been systematically working to improve the accuracy of the energy levels and, consequently, the transition frequencies using the Measured Active Rotation Vibration Energy Levels (MARVEL) database [502, 503]. Several recent MARVEL studies have focused on species such as C_2 [505], AlO [423], VO [506], H_3^+ and its deuterated isotopologues [507], formaldehyde [508, 509], and methane [510]. Recent research by McKemmish et al. [504] demonstrates how MARVEL can be applied and extended to expand the number of accurately determined energy levels and transitions. ExoMol uses the accurate energy levels generated by MARVEL and related techniques to improve their data by replacing calculated energy levels with empirically determined ones. It is noteworthy that this technique often results in ExoMolHR providing many more high accuracy transitions than those directly observed in laboratory settings, as illustrated in studies like Al-Derzi

et al. [508] or Bowesman et al. [506]. For example, in the MARVEL database, a total of 16596 non-redundant transitions from 43 sources were processed, of which 16403 were successfully validated, yielding 5029 empirical energy levels for $H_2{}^{12}C^{16}O$ with statistically well-defined uncertainties [508]. In comparison, the EXOMOLHR database provides 317729 newly computed high-accuracy transitions, significantly expanding the available spectroscopic data for this molecule.

9.2 MARVEL database overview

The Measured Active Rotational-Vibrational Energy Levels (MARVEL) [502] is a high-accuracy spectroscopic database that provides empirical energy levels for molecules based on experimentally measured transition frequencies. It uses an inversion algorithm to process and validate spectroscopic data from various sources, ensuring consistency and statistical reliability. The MARVEL methodology constructs a connected spectroscopic network, where redundant and inconsistent measurements are identified and resolved to derive optimized empirical energy levels with well-defined uncertainties.

MARVEL is particularly useful for validating and improving theoretical spectroscopic models. By providing high-precision empirical energy levels, it serves as a critical reference for quantum chemistry calculations and astrophysical applications. Its data are widely used to benchmark and refine line lists, ensuring greater accuracy in spectroscopic modeling for atmospheric, planetary, and astrophysical studies.

A well-known challenge in generating hot molecular line lists for astronomical modeling is the trade-off between completeness and accuracy. While theoretical line lists aim to be comprehensive by including all possible transitions, their accuracy depends on the underlying computational methods. To address this, EXOMOL incorporates MARVEL data to enhance the precision of its spectroscopic models. By integrating high-accuracy empirical energy levels from MARVEL, EXOMOL can improve the accuracy of key transitions while maintaining the completeness of its computed line lists Tennyson [511]. Preconditioning the experimental transitions

is essential for obtaining reliable MARVEL energy levels with reliable and as low as possible uncertainties [502]. The MARVEL database is publicly accessible at MARVEL ONLINE 2.0 (http://kkrk.chem.elte.hu/marvelonline/) [512].

The EXOMOL database includes line lists that have been converted from the MARVEL database, referred to as "MARVELised" line lists with inclusion of uncertainties. The advantage of using the MARVEL database as a data source is that it provides predicted transitions with experimental accuracy, offers data for theoretical models, and improves transition frequencies, which are important for high-resolution exoplanet observations. The EXOMOLHR database also incorporates these "MARVELised" molecular data.

9.3 EXOMOLHR database overview

The EXOMOLHR database primarily stores empirical high-resolution spectral lines extracted from the EXOMOL database. Although line strength is not a criterion in the extraction process, the extracted lines are predominantly the stronger ones.

The 2024 release of the EXOMOL database provides spectroscopic data for 91 molecules and 225 isotopologues [17]. While the majority of these datasets (68 molecules) were generated in-house by members of the EXOMOL project, the database also hosts datasets for 36 molecules generated either at University College London (UCL) for other purposes or externally. Many of these hosted datasets include high accuracy transition data, such as those from Bernath's MolLIST project [73], which are based on high-resolution laboratory measurements. EXOMOLHR requires explicit uncertainties to be present in the data for inclusion. Currently, explicit uncertainties are only available for three external datasets produced by McKemmish's group at the University of New South Wales (UNSW), specifically for CN [9], ZrO [513], and NH [514].

The EXOMOL database provides definition files in both its native format `.def` and the JSON (JavaScript Object Notation, Pezoa et al. [21]) format `.def.json` (`<ISOTOPOLOGUE>__<DATASET>.def.json`). These files contain detailed information on molecular characteristics, uncertainty availability, quantum number

labels and formats, among other details, to facilitate the interpretation and automated processing of states and transitions. The generation of the EXOMOLHR database requires the information from these definition files.

Tables 3.6, 3.8, and 6.1 illustrate the structure of the states (`.states`), transitions (`.trans`), and partition function (`.pf`) files, respectively. It is worth mentioning that the energy level uncertainty column (ΔE in Table 3.6) is a relatively new feature to the EXOMOL database, and not all datasets contain uncertainties. EXOMOLHR requires the presence of uncertainties; thus, datasets without an ΔE column, as indicated in the EXOMOL `.def.json` file, are excluded. These excluded datasets are substantially ones that have not been updated using MARVEL energy levels and therefore do not contain any high accuracy data essential for EXOMOLHR.

The EXOMOLHR database focuses on providing high accuracy line positions for high-resolution studies, such as line identification or simulations of high-resolution spectra [17]. Initially, reduced line lists are constructed by extracting energy levels and corresponding transitions from the EXOMOL database with high accuracy ($\Delta E \leq 0.01$ cm^{-1} on energy levels). Transitions that can be predicted with high resolution ($R > 100\,000$) are then stored, forming the core of the database. For each transition, EXOMOLHR provides the frequency, frequency uncertainty, Einstein A-coefficient, intensity (at a user-specified temperature), lower energy, total level degeneracy, angular momentum, and quantum numbers for the upper and lower states. Users can use the interactive web interface to set parameters for filtering wavenumber/wavelength ranges, extracting low transitions uncertainty lines, and selecting lines with strong intensities, with the option to store and download the generated results.

9.4 Methodology and algorithm

The EXOMOLHR database is specifically designed to provide highly accurate line positions for high-resolution studies, including spectral line identification and the simulation of high-resolution spectra. Figure 9.1 illustrates the generation pro-

cess of the EXOMOLHR database, which involves several steps and requires data to be obtained separately from the EXOMOL website and its database. Initially, data for all molecules and isotopologues are scraped from the EXOMOL website. For each isotopologue, the recommended dataset is extracted, and the corresponding JSON format definition file (`.def.json`) is processed, focusing only on datasets with explicit uncertainties. Secondly, the line list files are processed to obtain the labels and formats of the quantum numbers from the EXOMOL JSON format definition files (`.def.json`). Many molecules have multiple sets of quantum numbers in the states files. TROVE quantum numbers where given, were ignored. This ensures the quantum numbers retained give a complete description of each state. Subsequently, lines from the `.states` and `.trans` files of the EXOMOL database are extracted based on their high accuracy (low uncertainty). Specifically, energy levels with an uncertainty $\Delta E \leq 0.01$ cm^{-1} in the `.states` file are selected for processing into transition wavenumbers.

The resolving power, R, is defined in terms of the wavelength λ and its uncertainty $\Delta\lambda$. In this context, the transition wavenumber $\tilde{\nu}$ and its uncertainty $\Delta\tilde{\nu}$ are used to calculate the resolving power R as follows:

$$R = \frac{\lambda}{\Delta\lambda} \approx \frac{\tilde{\nu}}{\Delta\tilde{\nu}}. \tag{9.1}$$

The transition wavenumber (frequency) $\tilde{\nu}$ is defined as the difference between the upper state energy $\tilde{E}'$ and the lower state energy $\tilde{E}''$.

$$\tilde{\nu} = \tilde{E}' - \tilde{E}'', \tag{9.2}$$

$$\Delta\tilde{\nu} = \sqrt{(\Delta E')^2 + (\Delta E'')^2}. \tag{9.3}$$

$\Delta E'$ and $\Delta E''$ represent the uncertainties of the upper and lower state energy levels, respectively. Finally, I extract high-resolution and high accuracy transitions that are determined to correspond to a resolving power of $R > 100000$ and transition uncertainty $\Delta\tilde{\nu} \leq 0.01$ cm^{-1}.

The EXOMOLHR database provides absorption line intensities S (or I_{fi}, in cm

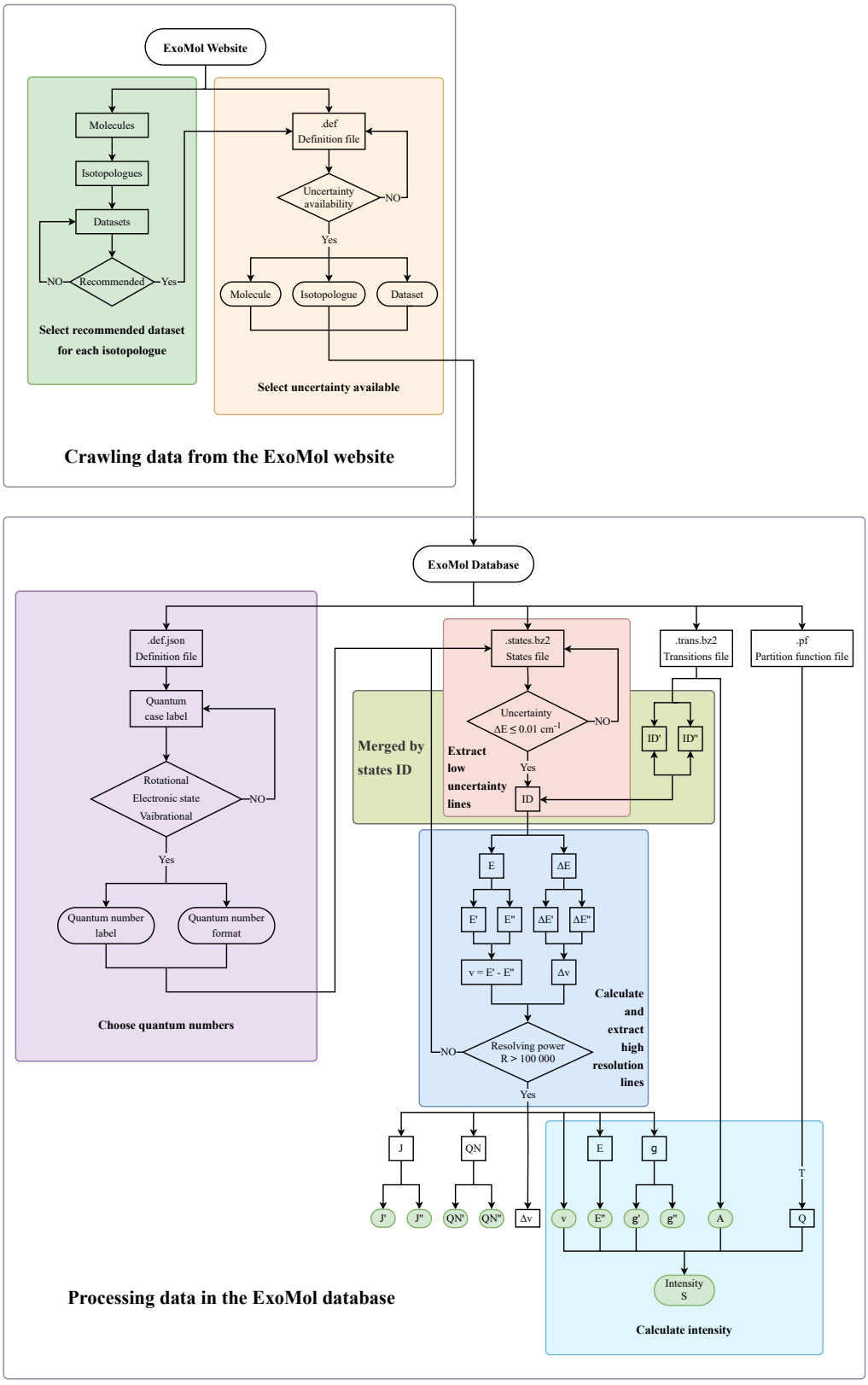


Figure 9.1: Process flowchart for EXOMOLHR database generation.

molecule^{-1}) at different temperatures T (in K) specified by users, using the Eq. (6.9) [373, 428]. In Figure 9.1, the lower blue section illustrates the steps for calculating absorption line intensities and the parameters required. Here, g' represents the upper state degeneracy and $\tilde{E}''$ is the lower state energy in the `.states` file. A is the Einstein A-coefficient (s^{-1}) in the `.trans` file. ν is the transition wavenumber (cm^{-1}), which can either be obtained from the `.trans` file for smaller datasets or calculated using Eq. (9.2). Q is the temperature-dependent partition function, provided by the ExoMol partition function file (`.pf`). The parameters highlighted with a green background represent the values stored in the ExoMolHR result files, which are provided for users to download.

9.5 File structure and features

Table 9.1 specifies the file structure of the ExoMolHR database. Table 9.2 provides a sample `.csv` file from the ExoMolHR database for $^{1}H_2{}^{16}O$ at temperature $T = 296$ K. Transitions form the core of the ExoMolHR database, including frequency (which can be filtered for a user-specified wavenumber range), transition uncertainty, Einstein A-coefficient, intensity (at a user-specified temperature), lower state energy level, total degeneracy, angular momentum, and quantum numbers for the upper and lower states. The output `.csv` files are not directly available on the website; however, their contents, with an additional intensity column for the specified temperature, can be extracted simply by requesting data for the desired wavenumber or wavelength range for the species of interest.

The ExoMol states files, especially those for polyatomic molecules, often include multiple sets of quantum numbers. However, for simplicity, the ExoMolHR database provides only a single set of quantum numbers for each state, typically based on a standard, normal mode description of vibrational motion. If alternative quantum number descriptions are required, they can be obtained from the corresponding `.states` file.

The ExoMolHR files are named in format `<yyyymmddhhmmss>__<ISO-SLUG>__<T>K.csv`, where `<yyyymmddhhmmss>` is a 14-digit character string

representing a timestamp in the format of year (`yyyy`), month (`mm`), day (`dd`), hour (`hh`), minute (`mm`), and second (`ss`). `<T>` denotes the temperature in units of K. The EXOMOLHR database calculates line intensities at user-specified temperatures, filtered by wavenumber or wavelength range and minimum intensity (threshold). The downloaded data is provided in `CSV` format, with column names as specified in Table 9.1 used as the header.

Table 9.1: Specification of the output `.csv` EXOMOLHR file format.

Field	Fortran Format	C Format	Description
$\tilde{\nu}$ or $\tilde{\lambda}$	`F12.6`	`%12.6f`	Transition wavenumber/frequency in cm^{-1} or wavelength in μm
$\Delta\nu$	`F12.6`	`%12.6f`	Uncertainty in the transition wavenumber in cm^{-1}
A	`ES10.4`	`%10.4E`	Einstein A coefficient in s^{-1}
S	`ES10.4`	`%10.4E`	Absorption intensity at the user specified temperature in cm/molecule
E''	`F12.6`	`%12.6f`	Lower state energy in cm^{-1}
g'	`I6`	`%6d`	Total state degeneracy for upper state
g''	`I6`	`%6d`	Total state degeneracy for lower state
J'	`I7/F7.1`	`%7d/%7.1f`	Total angular momentum quantum number (integer/half-integer) for upper state
J''	`I7/F7.1`	`%7d/%7.1f`	Total angular momentum quantum number (integer/half-integer) for lower state
(QN$'$)	See `.def` file	See `.def` file	Quantum number for upper state (optional)
(QN$''$)	See `.def` file	See `.def` file	Quantum number for lower state (optional)

Quantum number labels and formats are defined in the EXOMOL definition file (`.def.json`).

Table 9.2: Extract from the EXOMOLHR file of $^{1}H_2{}^{16}O$ at temperature $T = 296.0$ K (`yyyymmddhhmmss__1H2-16O__296K.csv`).

$\tilde{\nu}$	$\Delta\nu$	A	S	$\tilde{E}''$	g'	g''	J'	J''	$\Gamma'_{\rm rve}$	v'_1	v'_2	v'_3	K'_a	K'_c	$\Gamma''_{\rm rve}$	v''_1	v''_2	v''_3	K''_a	K''_c
500.018142	0.000029	2.3792E−03	1.3746E−30	4195.818039	15	13	7	6	A2	1	0	0	5	3	A1	0	0	1	1	6
500.035515	0.000009	1.7621E−01	2.3697E−24	2248.063135	27	27	13	13	A1	0	0	0	6	8	A2	0	0	0	3	11
500.436638	0.000023	3.8196E−04	1.3245E−30	4052.836634	45	45	7	7	B1	0	0	1	3	5	B2	0	2	0	3	4
500.621556	0.000028	9.2079E−02	3.4260E−23	1774.750359	75	75	12	12	B2	0	0	0	5	8	B1	0	0	0	2	11
500.661272	0.000012	2.6278E−04	2.3620E−30	4006.071011	93	99	15	16	B1	0	1	0	2	13	B2	0	0	0	7	10
500.878899	0.000050	1.9196E−01	1.0828E−24	2426.195175	27	27	13	13	A2	0	0	0	7	7	A1	0	0	0	4	10
501.573101	0.000071	5.8196E+01	2.2409E−22	2471.254912	23	21	11	10	A1	0	0	0	10	2	A2	0	0	0	9	1
501.573504	0.000071	5.8196E+01	6.7226E−22	2471.254532	69	63	11	10	B1	0	0	0	10	1	B2	0	0	0	9	2
502.064969	0.000043	6.5476E−03	2.4083E−30	4257.786653	13	11	6	5	A2	0	0	1	6	1	A1	1	0	0	4	2

yyyymmddhhmmss: Filename is structured with 14 digits of created date and time, iso-slug, and temperature.
yyyymmdd: Date with year, month, and day.
hhmmss: Time with hour, minute, and second.
$\tilde{\nu}$: Transition wavenumber/frequency in cm^{-1}.
$\Delta\nu$: Unceratinty in the transition wavenumber in cm^{-1}.
A: Einstein A-coefficient in s^{-1}.
S: Absorption intensity in cm/molecule.
$\tilde{E}''$: Lower state energy in cm^{-1}.
g' and g'': Total state degeneracy for upper and lower state.
J' and J'': Total angular momentum rotational quantum number for upper and lower state, excluding nuclear spin.
$\Gamma'_{\rm rve}$ and $\Gamma''_{\rm rve}$: Rovibrational symmetry label for upper and lower state.
v'_1 and v''_1: ν_1 symmetric stretch quantum number for upper and lower state.
v'_2 and v''_2: ν_2 bend quantum number for upper and lower state.
v'_3 and v''_3: ν_3 asymmetric stretch quantum number for upper and lower state.
K'_a and K''_a: K_a rotational quantum number for upper and lower state.
K'_c and K''_c: K_c rotational quantum number for upper and lower state.

Table 9.3 presents the quantum number labels and the format of the quantum number J for each isotopologue in the EXOMOLHR database. Table 9.4 provides the formats and descriptions of quantum numbers for various quantum number labels. The electronic state term values in both the EXOMOL and EXOMOLHR

databases are defined by PyVALEM format. PyVALEM [375, 376] is a Python package developed for parsing, validating, manipulating, and interpreting chemical formulas, quantum states, and labels for atoms, ions, and small molecules. For example, the electronic state $X^1\Sigma^+$ is recorded as `X(1SIGMA+)`. Sometimes, due to differences in conventions used for quantum number notation across various molecules in the ExoMol definition files (`.def.json`), variations or inconsistencies in quantum number labeling may occur. For example, quantum numbers that are conceptually similar may be denoted by different symbols such as v_i, n_i, or m_i, depending on the molecule. Conversely, the same label may represent different types of quantum numbers in different definition files. These variations have been carefully considered when I interpret and process the data.

Table 9.3: Quantum number labels for each isotopologue in the ExoMolHR database.

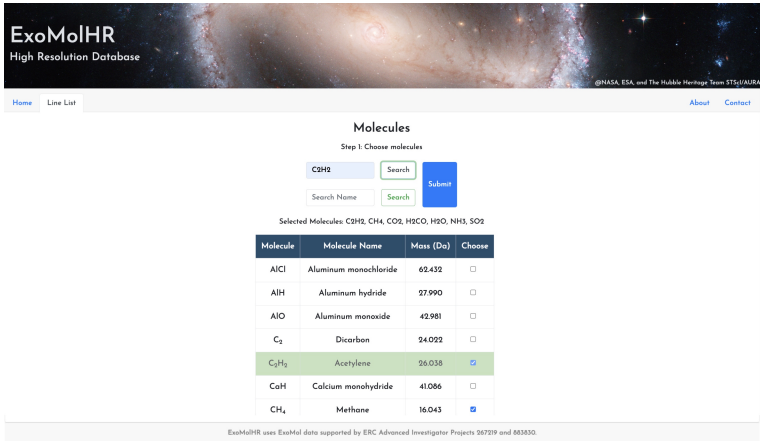
Molecule	Isotopologue	Quantum number label
AlCl	$^{27}Al^{35}Cl$, $^{27}Al^{37}Cl$	+/− e/f ElecState v Λ Σ Ω
AlH	$^{27}Al^{1}H$	
AlO	$^{26}Al^{16}O$, $^{27}Al^{16}O$, $^{27}Al^{17}O$, $^{27}Al^{18}O$	
BeH	$^{9}Be^{1}H$, $^{9}Be^{2}H$	
C_2	$^{12}C_2$	
CaH	$^{40}Ca^{1}H$	
CN	$^{12}C^{14}N$	
MgH	$^{24}Mg^{1}H$, $^{25}Mg^{1}H$, $^{26}Mg^{1}H$	
NO	$^{14}N^{16}O$	
PN	$^{31}P^{14}N$	
SiN	$^{28}Si^{14}N$, $^{28}Si^{15}N$, $^{29}Si^{14}N$, $^{30}Si^{14}N$	
SiO	$^{28}Si^{16}O$	
TiO	$^{47}Ti^{16}O$	
VO	$^{51}V^{16}O$	
YO	$^{89}Y^{16}O$, $^{89}Y^{17}O$, $^{89}Y^{18}O$	
ZrO	$^{90}Zr^{16}O$, $^{91}Zr^{16}O$, $^{92}Zr^{16}O$, $^{93}Zr^{16}O$, $^{94}Zr^{16}O$, $^{96}Zr^{16}O$	
SO	$^{32}S^{16}O$	+/− ElecState v Λ Σ Ω
LiOH	$^{6}Li^{16}O^{1}H$, $^{7}Li^{16}O^{1}H$	+/− e/f ElecState v_1 v_2 l_2 v_3
NH	$^{14}N^{1}H$, $^{14}N^{2}H$, $^{15}N^{1}H$, $^{15}N^{2}H$	+/− e/f ElecState v Λ N F1/F2/F3
CaOH	$^{40}Ca^{16}O^{1}H$	+/− e/f N ElecState L v_1 v_2 l_2 v_3 Ω F1/F2/F3
SO_2	$^{32}S^{16}O_2$	+/− K_a K_c v_1 v_2 v_3
H_2O	$^{1}H_2{}^{16}O$	G_{rve} v_1 v_2 v_3 K_a K_c
H_2S	$^{1}H_2{}^{32}S$	
H_3^+	$^{1}H_3^+$, $^{2}H_3^+$	e/f G_{rve} N Isomer v_1 v_2 l_2 G U K
	$^{1}H_2{}^{2}H^+$, $^{2}H_2{}^{1}H^+$	+/− G_{rve} N Isomer v_1 v_2 v_3 K_a K_c
NH_3	$^{14}N^{1}H_3$	+/− G_{tot} N n_1 n_2 n_3 n_4 l_3 l_4 τ_i K G_{rot}
	$^{15}N^{1}H_3$	G_{tot} n_1 n_2 n_3 l_3 n_4 l_4 G_{vib} K G_{rot}
C_2H_2	$^{12}C_2{}^{1}H_2$	G_{tot} K e/f G_{rot}
N_2O	$^{14}N_2{}^{16}O$	G_{tot} n_1 $n_2 lin$ l_2 n_3 P N
H_3O^+	$^{1}H_3{}^{16}O^+$	G_{tot} n_1 n_2 n_3 l_3 n_4 l_4 K G_{rot}
CH_4	$^{12}C^{1}H_4$	G_{tot} n_1 n_2 l_2 n_3 l_3 M_3 n_4 l_4 M_4 n
CO_2	$^{12}C^{16}O_2$	G_{tot} e/f n_1 n_2 l_2 n_3 m_1 m_2 m_3 m_4 m_5
H_2CO	$^{1}H_2{}^{12}C^{16}O$	G_{tot} v_1 v_2 v_3 v_4 v_5 v_6 K_a K_c
H_2CS	$^{1}H_2{}^{12}C^{32}S$	G_{tot} v_1 v_2 v_3 v_4 v_5 v_6 K_a K_c
OCS	$^{16}O^{12}C^{32}S$	G_{tot} v_1 v_2 l_2 v_3 e/f G_{vib} G_{rot} Coef n_1 n_2 n_3

Electronic state term values are defined using PyVALEM format [375, 376].

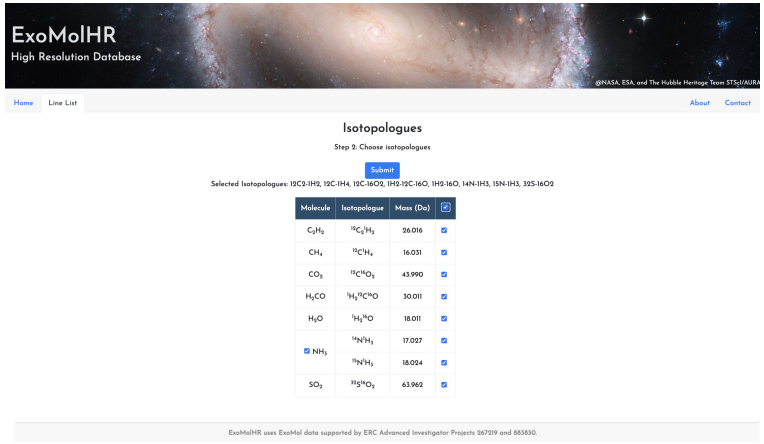
Table 9.4: The formats and descriptions corresponding to quantum number labels.

Label	Fortran Format	C Format	Description
J	I7/F7.1	%7d/%7.1f	Total angular momentum
	I7	%7d	AlCl, AlH, C_2, C_2H_2, CH_4, CO_2, H_2CO, H_2CS, H_2O, H_2S, H_3^+ H_3O^+, LiOH, N_2O, NH, NH_3, OCS, PN, SiO, SO, SO_2, TiO, VO
	F7.1	%7.1f	AlO, BeH, CN, CaH, CaOH, MgH, NO, SiN, YO, ZrO
$+/-$	A1	%1s	Total parity
e/f	A1	%1s	Rotationless parity
ElecState	A12	%12s	Electronic term value
v	I3	%3d	Vibrational quantum number
$\lvert\Lambda\rvert$	I3	%3d	Absolute value of the projection of electronic angular momentum
$\lvert\Sigma\rvert$	I3	%3d	Absolute value of the projection of the electronic spin
$\lvert\Omega\rvert$	I5/F5.1	%5d/%5.1f	Absolute value of the projection of the total angular momentum
	I5	%5d	AlCl, AlH, C_2, TiO, ZrO
	F5.1	%5.1f	AlO, BeH, CN, CaH, CaOH, MgH, NO, PN, SO, SiN, SiO, VO, YO
G_{tot}	A3	%3s	Total symmetry
G_{rot}	A3	%3s	Symmetry of rotational contribution
G_{rve}	A3	%3s	Rovibrational symmetry label for upper and lower state
G_{vib}	A3	%3s	Symmetry of vibrational contribution
N	I5	%5d	Vibrational state ID or rotational angular momentum
L	I3	%3d	Vibronic angular momentum quantum number
P	I2	%2d	Polyad number
F	I2	%2d	Fine structure counting number
F1/F2/F3	I2	%2d	Spin components F1 and F3
U	A1	%1s	U-notation of Watson [515]
G	I2	%2d	Absolute value of $k - l_2$
K	I2	%2d	State projection of the Watson [515] rotational quantum number
K_a	I2	%2d	K_a rotational quantum number
K_c	I2	%2d	K_c rotational quantum number
M_3, M_4	I2	%2d	Multiplicity index quantum numbers
v_1,v_2,v_3,v_4,v_5,v_6	I2	%2d	v_i Local mode vibrational quantum numbers
l_2,l_3,l_4	I2	%2d	Normal mode vibrational angular momentum quantum number
n_1,n_2,n_3,n_4	I2	%2d	Normal mode vibrational quantum numbers
m_1,m_2,m_3,m_4,m_5	I2	%2d	HITRAN normal mode quantum number [13]
n	I2	%2d	Polyad or rotational counting number
n_2lin	I2	%2d	Normal mode linear molecule bending quantum number
τ_i	I1	%1d	Inversion parity (0 or 1)
Isomer	A1	%1s	Nuclear spin isomer
Coef	F5.2	%5.2f	Coefficient with the largest contribution to the ($J = 0$) contracted set

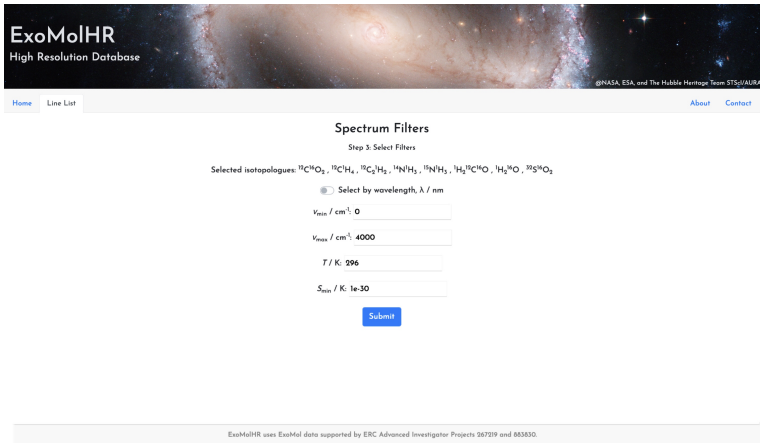
Electronic state term values are defined using the PyValem format [375, 376]. The label descriptions provided in the table are for reference purposes only and may sometimes vary slightly in their usage. More precise and detailed descriptions can be found in the definition file (`.def.json`) for the recommended dataset of each isotopologue. The meaning and more details of each label can be further understood by referring to the notes after the example tables in the corresponding line list papers.



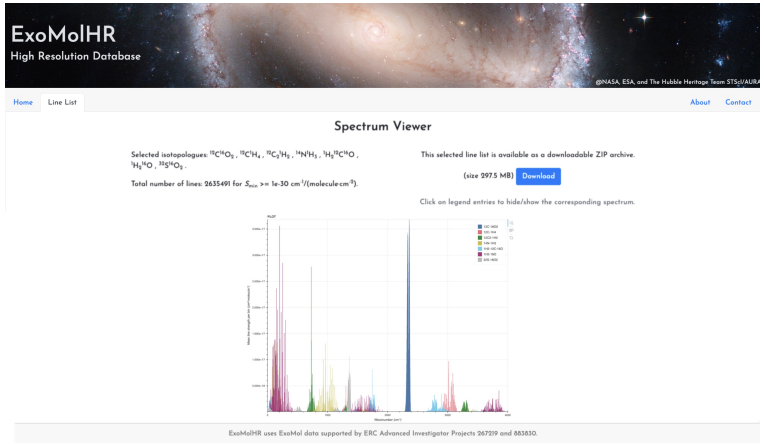
(a) Step 1. Choose molecules



(b) Step 2. Choose isotopologues



(c) Step 3. Select filters



(d) Step 4. Spectrum viewer

Figure 9.2: Four steps for getting the intensity file from the ExoMolHR website.

9.6 Use cases

The EXOMOLHR website offers a download link to `.csv` files including the calculated results and plots the stick spectra for one or multiple isotopologues. Users can follow these steps to obtain the final results. Figure 9.2 displays screenshots of these four steps.

Step 1. Molecules: Users can select one or multiple molecules from the form. Two search boxes are available for searching by molecule name and formula.

Step 2. Isotopologues: Similarly, users can choose one or more isotopologues, with an option to select all isotopologues by clicking the "select all" button.

Step 3. Spectral filters: Users can define the wavenumber range by specifying $\nu_{\rm min}$ and $\nu_{\rm max}$ in cm^{-1} or the wavelength range using $\lambda_{\rm min}$ and $\lambda_{\rm max}$ in μm. Users can also set the temperature T in K and the minimum intensity (threshold) $S_{\rm min}$ in cm/molecule.

Step 4. Spectrum viewer: Finally, users can get a plot of intensities calculated at the given temperature for one or multiple isotopologues, displayed in different colours for the requested wavenumber range. The webpage also provides the number of lines and the file size of the total result files. Users can download a compressed `.zip` file, with result files named separately for each isotopologue. The file structure is detailed in Section 9.5.

It is important to note that the EXOMOLHR website only provides stick spectra plots. The EXOMOL database offers a set of pressure broadening parameters [14, 516, 517], which can be combined with the EXOMOLHR `.csv` file to generate temperature- and pressure-dependent cross sections. The Python program PYEXOCROSS plans to extract intensities from the EXOMOLHR database along with broadening parameters from the EXOMOL database to calculate the cross sections. However, for most applications, it is advisable to use the original EXOMOL line list to ensure the inclusion of all spectral lines within the region of interest for this process.

9.7 Application programming interface (API)

Users can use the application programming interface (API) to access the spectrum viewer webpage directly. The API format is `https://www.exomol.com/exomolhr/get-data/?numin=<`$\nu_{\rm min}$`>&numax=<`$\nu_{\rm max}$`>&T=<`T`>&Smin=<`$S_{\rm min}$`>&iso=<ISO-SLUG`$_{\rm i}$`>`. To request data for multiple isotopologues, multiple `&iso=<ISO-SLUG`$_{\rm i}$`>` for each isotopologue should be appended, such as `&iso=<ISO-SLUG`$_1$`>&iso=<ISO-SLUG`$_2$`>`. For example, to request EXOMOLHR files with intensities calculated at a temperature T = 296 K and an intensity threshold $S_{\rm min} = 10^{-30}$ cm/molecule for ^{1}H$_2$^{16}O, ^{12}C^{16}O$_2$, and ^{32}S^{16}O$_2$ in the wavenumber range from $\nu_{\rm min} = 0$ to $\nu_{\rm max}$ = 1000 cm^{-1}, the API request would be https://www.exomol.com/exomolhr/get-data/?numin=0&numax=1000&T=296&Smin=1e-30&iso=1H2-16O&iso=12C-1H4&iso=32S-16O2. The automatically generated file name lists the species in the order of the original request.

The EXOMOLHR compressed `.zip` file is named using a 14-digit timestamp format (`yyyymmddhhmmss`) representing the year, month, day, hour, minute, and second of creation date and time. For example, https://www.exomol.com/exomolhr/get-data/download/?archive_name=20240907222904.zip. Users cannot download files directly through the API. Instead, to download the result files, users must click the download button on the spectrum viewer webpage. This approach is generally more user-friendly, allowing users to visualize the data and review file details before downloading, which helps prevent incorrect downloads.

9.8 Comparisons and discussion

Fifteen of the 33 species currently considered in EXOMOLHR database are also present in the HITRAN database [13]. To compare the results from EXOMOLHR with its initial source database, EXOMOL, and the high-resolution database HITRAN, the data in the following figures are calculated by the Python program PYEXOCROSS. To emulate HITRAN, only those lines with intensities greater than 10^{-30} cm/molecule at T = 296 K are considered. As shown, EX-

OMOLHR generally provides more complete coverage than HITRAN, with high resolution and low uncertainty.

Figure 9.3 presents a comparison of the intensities for H_2O, NH_3, CH_4, and SO_2 obtained from the 2024 release of the EXOMOL [17], EXOMOLHR, and the 2020 release of HITRAN [13] databases. These intensities (cm/molecule) are calculated at temperature $T = 296$ K, covering the full wavenumber range in cm^{-1}. The line lists used from the EXOMOLHR and EXOMOL databases include POKAZATEL of $^{1}H_2{}^{16}O$ [5], CoYuTe-15 of $^{15}N^{1}H_3$ [518], MM of $^{12}C^{1}H_4$ [519], and ExoAmes of $^{32}S^{16}O_2$ [7]. The distribution of EXOMOLHR and HITRAN points compared to EXOMOL points indicates that high-resolution spectra are generally concentrated in the low-wavenumber range at room temperature. The comparison between EXOMOLHR and HITRAN shows that high-resolution spectra with low uncertainty are predominantly distributed in the low-wavenumber region.

Figure 9.4 presents a comparative analysis of the intensities (cm/molecule) for H_2O [5] from the EXOMOL and EXOMOLHR databases at temperatures $T = 1000$ K and 2000 K, focusing on the wavenumber range of $[19\,200, 20\,200]$ cm^{-1} and providing a detailed examination within the narrower interval of $[19\,800, 19\,900]$ cm^{-1}. Similar comparisons are provided for $^{48}Ti^{16}O$ [504, 520] and $^{51}V^{16}O$ [504, 521], which are not included in the HITRAN database. The comparisons indicate that for TiO and VO, the EXOMOLHR coverage is patchier than for the better studied HITRAN molecules. Improving the high-resolution line coverage for these species will require new high-resolution spectroscopic measurements.

To analyze the distribution of high-resolution line lists across wavenumber ranges and intensities, Figure 9.5 uses $^{1}H_2{}^{16}O$, $^{14}N^{1}H_3$, $^{12}C^{1}H_4$, and $^{32}S^{16}O_2$ as examples to plot the density of points. These density plots are produced using a Python package `seaborn.kdeplot` [522], which facilitates the visualization of univariate or bivariate distributions through kernel density estimation (KDE). A kernel density estimate (KDE) plot is a method for visualizing the distribution of observations in a dataset, similar to a histogram [522]. It represents the data as a continuous probability density curve across one or more dimensions [522].

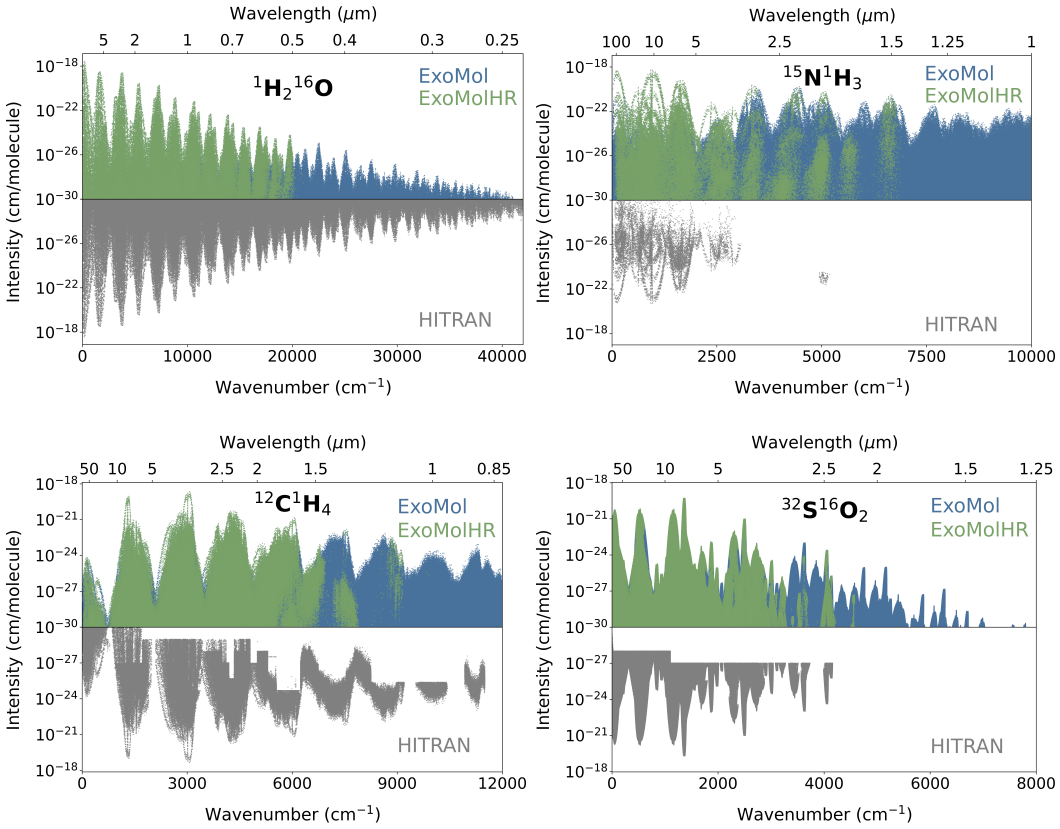


Figure 9.3: A comparison of the line intensities (cm/molecule) from the EXOMOL, EXOMOLHR (top), and HITRAN (bottom) databases at $T = 296$ K for $^{1}H_{2}{}^{16}O$, $^{15}N^{1}H_{3}$, $^{12}C^{1}H_{4}$, and $^{32}S^{16}O_{2}$. In the upper panels, the highlighted points represent the intensities provided by EXOMOLHR database, while the darker background illustrates the full coverage offered by the POKAZATEL, CoYuTe-15, MM, and ExoAmes line lists for $^{1}H_{2}{}^{16}O$, $^{15}N^{1}H_{3}$, $^{12}C^{1}H_{4}$, and $^{32}S^{16}O_{2}$, respectively.

In Figure 9.5, its left hand panels show the distribution density of spectral line points from Figure 9.3 across the wavenumber ranges for the EXOMOL, EXOMOLHR, and HITRAN databases at room temperature. It shows that, compared to the comprehensive dataset provided by the EXOMOL database, the spectra from high-resolution databases EXOMOLHR and HITRAN are predominantly concentrated in the lower wavenumber region. The middle panels demonstrates that after filtering out low-resolution and high-uncertainty line lists from the EXOMOL database, there is a noticeable reduction in the density of low-intensity spectral lines in the EXOMOLHR database. It is worth noting that omitting such transitions can significantly affect opacity [429]. The right hand panels offer a more intuitive

visualization of the density distribution of high-resolution spectra by illustrating the relationship between wavenumber and intensity in two-dimensional plots. These plots indicate a higher concentration of points in the lower wavenumber and lower intensity regions, with fewer high-resolution and higher-intensity spectra distributed in the higher wavenumber region.

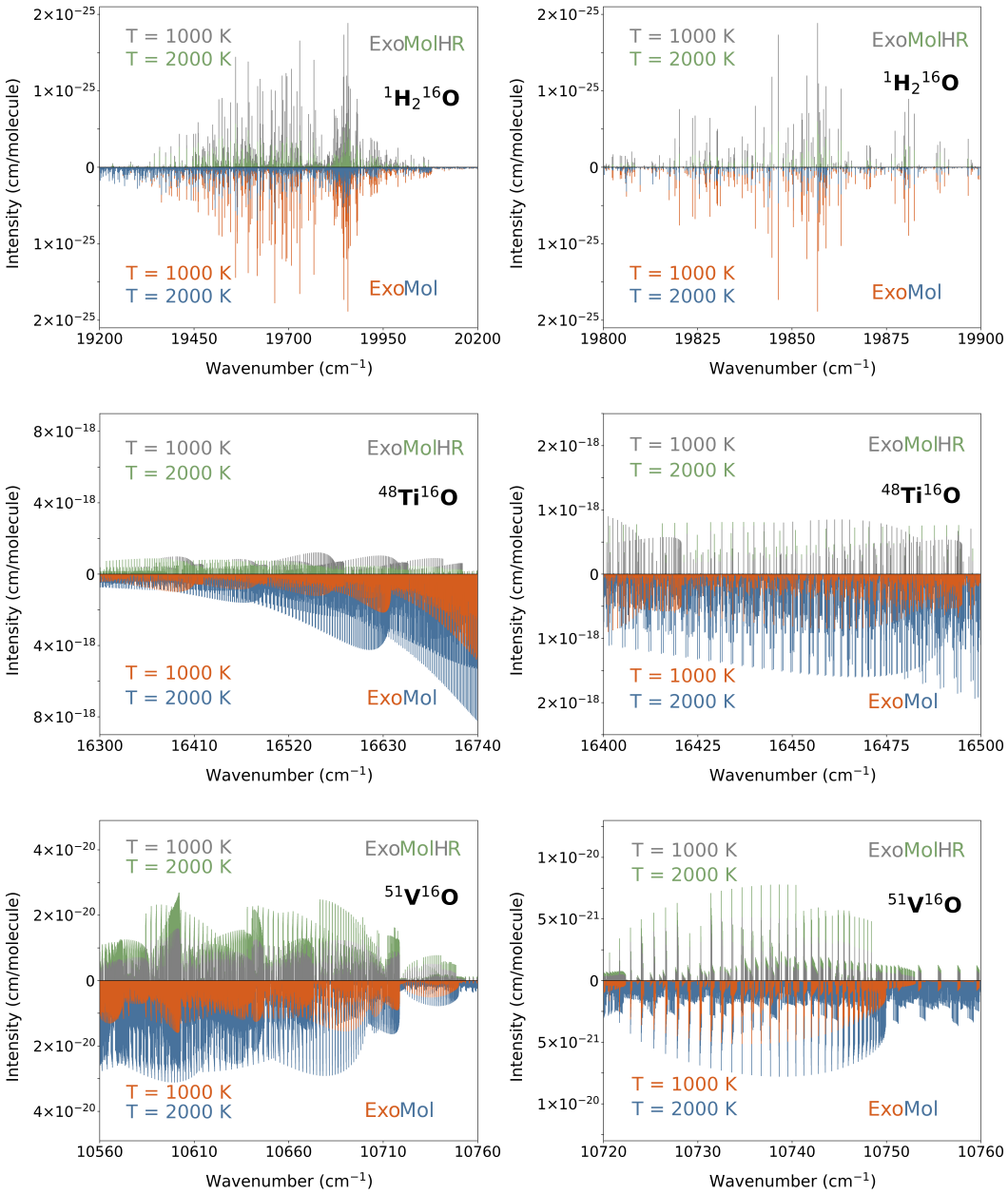


Figure 9.4: A comparison of the line intensities (cm/molecule) from the ExoMolHR (top) and ExoMol (bottom) databases at temperatures $T = 1000$ and 2000 K for $^{1}H_{2}{}^{16}O$, $^{48}Ti^{16}O$, and $^{51}V^{16}O$.

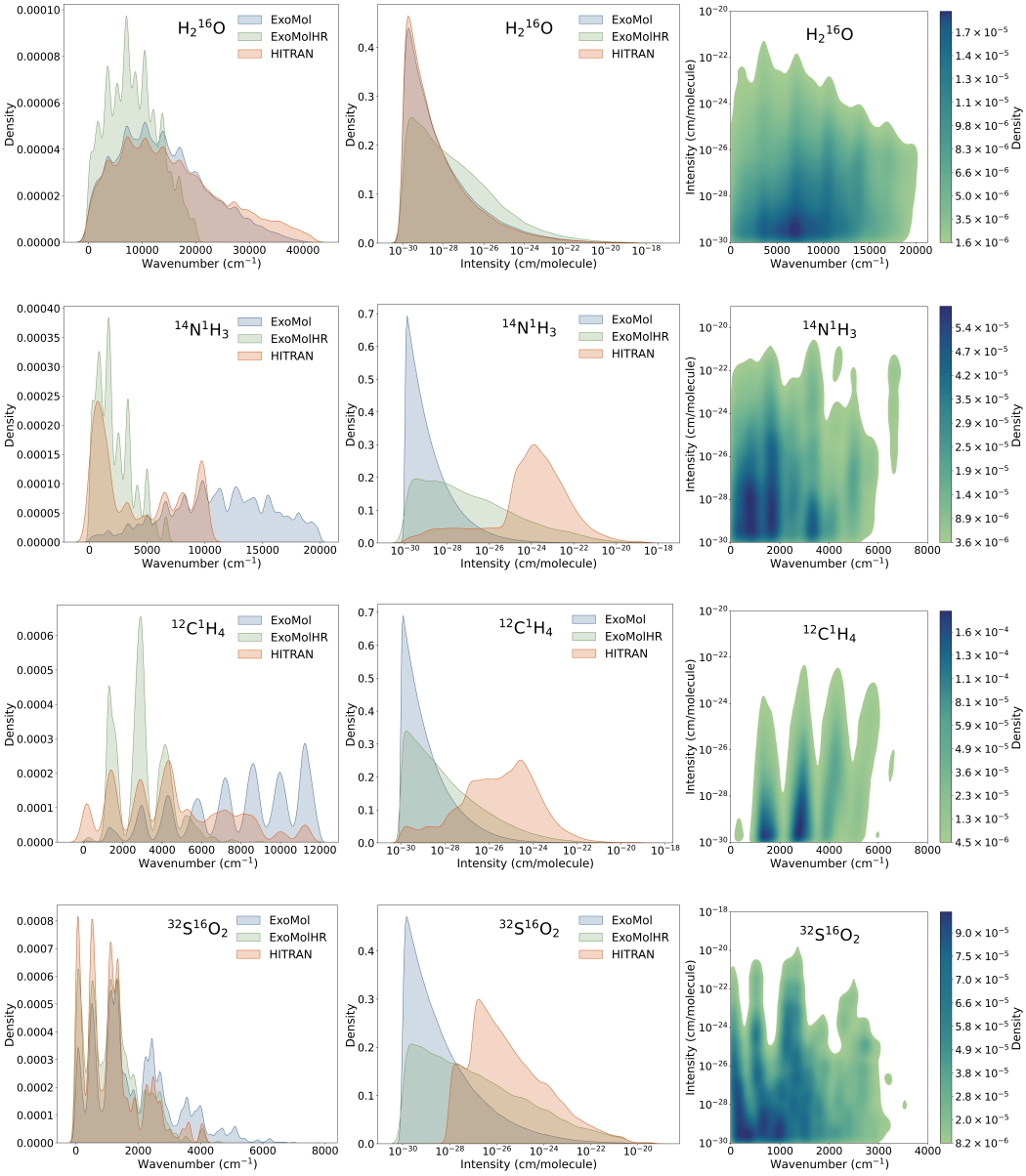


Figure 9.5: A comparison of the densities of wavenumbers (left) and line intensities (middle) from the EXOMOL, EXOMOLHR, and HITRAN databases at temperature $T = 296$ K for the molecules $^{1}H_2{}^{16}O$, $^{14}N^{1}H_3$, $^{12}C^{1}H_4$, and $^{32}S^{16}O_2$. The line lists from the EXOMOL database provide full coverage, using the POKAZATEL, CoYuTe, MM, and ExoAmes line lists for $^{1}H_2{}^{16}O$, $^{14}N^{1}H_3$, $^{12}C^{1}H_4$, and $^{32}S^{16}O_2$, respectively. The right hand panels present 2D plots of the density distributions between wavenumbers and line intensities (cm/molecule) for the EXOMOLHR database at $T = 296$ K.

9.9 Database summary and sources

Table 9.5 presents a list of molecules, isotopologues, datasets, and their corresponding versions (update dates) along with data sources. The table compares the number of initial states in the ExoMol database, as specified in the `.states` file, with the number of states having lower uncertainty ($\Delta E \leq 0.01$ cm^{-1}) used in the ExoMolHR database. Additionally, Table 9.5 also compares the total number of transitions and file size for each dataset in the ExoMol database with the final line list number and file size in the ExoMolHR database.

9.10 Summary

ExoMolHR is a newly developed high-resolution molecular spectroscopic database, derived from the ExoMol database, which is designed to provide molecular line lists for modeling exoplanet atmospheres at high temperatures. The ExoMolHR database achieves high resolution by filtering out states with high energy uncertainties. All transitions which can be predicted at high resolution $R > 100\,000$ are subsequently extracted to form the high-resolution line lists.

The ExoMolHR database includes 24 307 135 transitions associated with 238 263 states, determined with low uncertainties ($\Delta\nu \leq 0.01$ cm^{-1}), from 727 841 563 143 transitions and 129 237 384 states of the corresponding datasets within the ExoMol database. However, this process will inevitably result in missing lines in the ExoMolHR database. Users seeking completeness or wishing to check for missing lines are encouraged to use the original ExoMol line lists with the spectral generator programs, such as ExoCross [373] or PyExoCross.

The data in the ExoMolHR database are sourced from the ExoMol database, with low uncertainty line lists computed using data from the MARVEL database. The ExoMolHR database offers 58 isotopologues from 33 molecules, with the numbers expected to grow as more MARVEL studies are completed.

Virtual Atomicand Molecular Data Centre (VAMDC) [523] is an interoperable e-infrastructure that offers the international research community access to a wide array of atomic and molecular (A&M) data, aggregated from various A&M databases

accessible via this portal and accompanying user software. Additionally, VAMDC aims to serve as a comprehensive dissemination platform for A&M data providers and compilers, enhancing the visibility and accessibility of their work. Currently, VAMDC hosts 39 databases, including the High-Resolution Transmission Molecular Absorption (HITRAN) database [53], the Cologne Database for Molecular Spectroscopy (CDMS) [58], the Jet Propulsion Laboratory (JPL) database [56], the National Institute of Standards and Technology (NIST) Atomic Spectra database [524, 525], and the Vienna Atomic Line Database (VALD) [61]. Integrating the complete EXOMOL database into the VAMDC data center presents significant challenges due to its size. However, the EXOMOLHR database has been developed to extract high-resolution and high-accuracy molecular line list data from the EXOMOL database, thereby reducing the size of the initial EXOMOL database for record integration and making the high quality data easy to access.

Table 9.5: Summary of lines with low uncertainties and high resolutions extracted from the ExoMol database and used to form the ExoMolHR database.

ID_m	Molecule	ID_i	Isotopologue	Dataset	Version	N_{states}	N_{files}	N_{trans}	$Size_{trans}$	$N_{HRstates}$	$N_{HRlines}$	$Size_{HR}$	Source
1	AlCl	1	$^{27}Al^{35}Cl$	YNAT	20221231	65 869	1	4 722 048	47.5 MB	41	101	15.3 KB	[526]
		2	$^{27}Al^{37}Cl$	YNAT	20221231	67 507	1	5 748 704	57.7 MB	41	121	18.4 KB	
2	AlH	3	$^{27}Al^{1}H$	AloHa	20240307	1 364	1	29 725	152.9 KB	135	692	104.2 KB	[11]
3	AlO	4	$^{26}Al^{16}O$	ATP	20210622	93 350	1	4 866 540	53.6 MB	4 783	143 197	21.0 MB	[423]
		5	$^{27}Al^{16}O$	ATP	20210622	94 862	1	4 945 580	31.1 MB	4 980	149 577	22.0 MB	
		6	$^{27}Al^{17}O$	ATP	20210622	96 350	1	5 148 996	56.7 MB	4 787	142 905	21.0 MB	
		7	$^{27}Al^{18}O$	ATP	20210622	98 269	1	5 365 592	59.1 MB	4 799	142 976	21.0 MB	
4	BeH	8	$^{9}Be^{1}H$	Darby-Lewis	20240710	14 950	1	592 308	6.1 MB	132	507	76.4 KB	[527]
		9	$^{9}Be^{2}H$	Darby-Lewis	20240710	14 950	1	689 466	7.2 MB	104	310	46.8 KB	
5	C_2	10	$^{12}C_2$	8states	20200628	44 189	1	6 080 920	68.7 MB	8 376	445 682	65.5 MB	[505]
6	C_2H_2	11	$^{12}C_2{}^{1}H_2$	aCeTY	20220918	5 160 803	100	4 347 381 911	96.1 GB	8 898	473 850	47.9 MB	[528]
7	CH_4	12	$^{12}C^{1}H_4$	MM	20240113	9 155 208	121	50 395 644 806	439.0 GB	21 021	7 649 736	760.2 MB	[519]
8	CN	13	$^{12}C^{14}N$	Trihybrid	20210526	28 004	1	2 285 103	25.8 MB	4 833	244 808	36.0 MB	[9]
9	CO_2	14	$^{12}C^{16}O_2$	UCL-4000	20200630	3 480 477	20	2 557 549 946	21.7 GB	18 881	2 600 218	362.0 MB	[6]
10	CaH	15	$^{40}Ca^{1}H$	XAB	20220211	6 825	1	293 151	2.3 MB	1 165	12 341	1.8 MB	[4]
11	CaOH	16	$^{40}Ca^{16}O^{1}H$	OYT6	20230523	3 187 522	18	23 384 729 495	202.0 GB	1 424	12 984	1.6 MB	[529]
12	H_2CO	17	$^{1}H_2{}^{12}C^{16}O$	AYTY	20230430	10 297 025	100	12 688 112 669	111.1 GB	4 813	317 729	41.2 MB	[530]
13	H_2CS	18	$^{1}H_2{}^{12}C^{32}S$	MOTY	20221231	52 292 454	8	43 561 116 660	399.5 GB	3 625	72 218	6.6 MB	[531]
14	H_2O	19	$^{1}H_2{}^{16}O$	POKAZATEL	20230621	810 269	412	5 745 071 340	48.1 GB	14 395	3 520 554	396.2 MB	[5]
15	H_2S	20	$^{1}H_2{}^{32}S$	AYT2	20220918	220 631	35	115 032 941	873.0 MB	7 314	1 084 520	122.0 MB	[532]
16	H_3O^+	21	$^{1}H_3{}^{16}O^+$	eXeL	20200707	1 173 114	100	2 089 331 073	10.1 GB	232	1 785	251.3 KB	[533]
17	H_3^+	22	$^{1}H_2{}^{2}H^+$	ST	20230123	33 330	1	22 164 810	128.2 MB	109	646	87.2 KB	[507]
		23	$^{1}H_3^+$	MiZATeP	20230123	158 721	1	127 542 657	1.7 GB	994	13 606	1.9 MB	
		24	$^{2}H_2{}^{1}H^+$	MiZo	20221221	369 500	32	2 290 235 000	14.8 GB	115	683	92.2 KB	
		25	$^{2}H_3^+$	MiZo	20230124	37 410	21	36 078 183	170.3 MB	115	225	32.0 KB	

Continued on the next page

Table 9.5: (Continued)

ID_m	Molecule	ID_i	Isotopologue	Dataset	Version	N_{states}	N_{files}	N_{trans}	$Size_{trans}$	$N_{HRstates}$	$N_{HRlines}$	$Size_{HR}$	Source
18	LiOH	26	$^{6}Li^{16}O^{1}H$	OYT7	20231001	192 412	5	294 573 305	2.3 GB	255	840	113.3 KB	[534]
		27	$^{7}Li^{16}O^{1}H$	OYT7	20230529	203 762	5	331 274 717	2.6 GB	240	749	101.1 KB	
19	MgH	28	$^{24}Mg^{1}H$	XAB	20220211	3 148	1	88 575	953.2 KB	237	2 462	370.4 KB	[4]
		29	$^{25}Mg^{1}H$	XAB	20220211	3 156	1	88 776	955.6 KB	548	5 850	879.9 KB	
		30	$^{26}Mg^{1}H$	XAB	20220211	3 160	1	88 891	956.5 KB	537	5 339	803.1 KB	
20	N_2O	31	$^{14}N_2^{16}O$	TYM	20240527	1 759 068	11	1 360 351 722	11.0 GB	17 018	3 459 640	309.9 MB	[535]
21	NH	32	$^{14}N^{1}H$	kNigHt	20240301	4 076	1	327 014	1.6 MB	1 030	26 131	3.7 MB	[514]
		33	$^{14}N^{2}H$	kNigHt	20240301	7 406	1	778 105	3.7 MB	118	943	136.4 KB	
		34	$^{15}N^{1}H$	kNigHt	20240301	4 089	1	327 877	1.6 MB	118	943	136.4 KB	
		35	$^{15}N^{2}H$	kNigHt	20240301	7 465	1	785 940	3.8 MB	118	943	136.4 KB	
22	NH_3	36	$^{14}N^{1}H_3$	CoYuTe	20200730	5 095 730	200	16 941 637 250	142.9 GB	4 720	412 149	64.4 MB	[536]
		37	$^{15}N^{1}H_3$	CoYuTe-15	20240808	12 699 617	10	929 795 249	7.3 GB	2 699	148 651	22.0 MB	[518]
23	NO	38	$^{14}N^{16}O$	XABC	20210422	30 811	1	4 596 666	96.7 MB	3 044	106 711	15.7 MB	[12]
24	OCS	39	$^{16}O^{12}C^{32}S$	OYT8	20240425	2 399 110	10	2 527 364 150	26.4 GB	5 198	279 273	44.7 MB	[537]
25	PN	40	$^{31}P^{14}N$	PaiN	20240505	30 327	1	1 333 445	15.6 MB	32	44	6.8 KB	[538]
26	SO	41	$^{32}S^{16}O$	SOLIS	20230914	84 114	1	7 086 100	81.6 MB	536	2 501	366.5 KB	[539]
27	SO_2	42	$^{32}S^{16}O_2$	ExoAmes	20170131	3 270 271	80	1 300 000 000	25.4 GB	14 924	1 504 495	163.6 MB	[7]
28	SiN	43	$^{28}Si^{14}N$	SiNfull	20220809	131 936	1	43 646 806	506.9 MB	99	670	100.9 KB	[540]
		44	$^{28}Si^{15}N$	SiNfull	20220809	133 460	1	44 816 182	520.9 MB	56	464	69.9 KB	
		45	$^{29}Si^{14}N$	SiNfull	20220809	132 335	1	43 946 969	510.6 MB	56	464	69.9 KB	
		46	$^{30}Si^{14}N$	SiNfull	20220809	132 706	1	44 223 730	513.9 MB	56	464	69.9 KB	
29	SiO	47	$^{28}Si^{16}O$	SiOUVenIR	20211105	174 250	1	91 395 763	1.1 GB	911	8 729	1.3 MB	[541]
30	TiO	48	$^{48}Ti^{16}O$	Toto	20210825	301 245	1	58 983 952	689.5 MB	8 725	499 775	73.4 MB	[520]
31	VO	49	$^{51}V^{16}O$	HyVO	20231218	3 410 598	90	58 904 173 243	297.9 GB	7 043	635 722	68.1 MB	[521]

Continued on the next page

Table 9.5: (Continued)

ID_m	Molecule	ID_i	Isotopologue	Dataset	Version	N_{states}	N_{files}	N_{trans}	$Size_{trans}$	$N_{HRstates}$	$N_{HRlines}$	$Size_{HR}$	Source
32	YO	50	$^{89}Y^{16}O$	BRYTS	20230916	173 621	1	60 678 140	719.1 MB	28	25	3.9 KB	[542]
		51	$^{89}Y^{17}O$	BRYTS	20230916	182 598	1	62 448 157	740.9 MB	28	25	3.9 KB	
		52	$^{89}Y^{18}O$	BRYTS	20230916	182 547	1	64 164 605	761.9 MB	28	25	3.9 KB	
33	ZrO	53	$^{90}Zr^{16}O$	ZorrO	20230713	227 118	1	47 662 773	369.3 MB	5 313	145 317	21.3 MB	[513]
		54	$^{91}Zr^{16}O$	ZorrO	20230713	227 118	1	47 748 501	370.1 MB	1 058	5 164	776.8 KB	
		55	$^{92}Zr^{16}O$	ZorrO	20230713	227 124	1	47 830 250	370.8 MB	1 058	5 164	776.8 KB	
		56	$^{93}Zr^{16}O$	ZorrO	20230713	227 126	1	47 928 979	371.6 MB	1 058	5 164	776.8 KB	
		57	$^{94}Zr^{16}O$	ZorrO	20230713	227 128	1	47 994 352	372.1 MB	1 058	5 164	776.8 KB	
		58	$^{96}Zr^{16}O$	ZorrO	20230713	227 134	1	48 136 388	373.3 MB	1 058	5 164	776.8 KB	

ID_m: Molecule counting number.
ID_i: Isotopologue counting number.
Verison: ExoMol dataset update date in the format `YYYYMMDD`.
N_{states}: Number of states in states file.
N_{files}: Number of transitions files.
N_{trans}: Number of transitions in transitions file(s).
$Size_{trans}$: Total file size of bzip2 compressed transitions file(s).
$N_{HRstates}$: Number of states with uncertainties ≤ 0.01 cm^{-1}.
$N_{HRlines}$: Number of lines in ExoMolHR dataset with $R > 100\,000$.
$Size_{HR}$: File size of uncompressed ExoMolHR dataset.
Source: Data source reference of the ExoMol line list.

Chapter 10

Conclusion

This thesis has demonstrated the importance of continuously improving molecular spectroscopic databases to support advancements in astrophysics, atmospheric science, and laboratory spectroscopy. By refining data processing methodologies, introducing new computational tools, and expanding high-resolution spectral databases, this research provides essential contributions to the field.

The developments presented here will facilitate more accurate spectroscopic analyses, supporting future research in planetary atmospheres, exoplanetary studies, and high-resolution observational spectroscopy. The methodologies and tools established in this thesis set a foundation for further improvements in spectroscopic data management, ensuring that future studies in molecular spectroscopy are equipped with the most accurate and comprehensive resources available.

10.1 Summary of the thesis

This thesis has focused on advancing the field of molecular spectroscopy by enhancing the accuracy, accessibility, and usability of spectroscopic databases, particularly EXOMOL database. Through systematic data processing methodologies, novel computational tools, and high-resolution spectral modeling techniques, significant improvements have been made in the management and application of spectroscopic data. The research has addressed key challenges in handling large spectroscopic datasets, updating and expanding molecular line lists, improving line profile modeling, and simulating non-LTE spectral conditions.

A key contribution of this work has been the expansion of the ExoMol database, incorporating new and improved molecular line lists to ensure comprehensive coverage of important spectroscopic parameters. Additionally, the development and refinement of PyExoCross have provided a powerful post-processing tool capable of calculating molecular properties such as partition functions, specific heats, cooling functions, radiative lifetimes, oscillator strengths, stick spectra, line intensities, and cross sections. This tool has streamlined data processing and enhanced the ability to convert between different spectroscopic formats, particularly between ExoMol and HITRAN data structures.

The research has also contributed to the refinement of line profile models, including Gaussian, Doppler, Lorentzian, and Voigt profiles, improving the accuracy of cross-section calculations. Furthermore, by simulating non-LTE spectra, the thesis has provided valuable insights into the behavior of molecular systems in astrophysical and atmospheric environments where LTE assumptions are not valid. Another major achievement of this work is the introduction of ExoMolHR, a high-resolution extension of ExoMol, designed to support cutting-edge applications in exoplanetary and astrophysical research.

10.2 Contributions and impact

The work presented in this thesis makes several critical contributions to the field of molecular spectroscopy.

A major component of my work involved the maintenance and extension of the ExoMol database, where I contributed to the updating molecular line lists to the ExoMol website and generating new file structures. This task required careful processing the master, definition, states, and transitions files, ensuring that new or revised line lists were consistent with the database standards. Through this effort, I helped update the database for applications in exoplanetary and astrophysical research.

To facilitate high-precision spectral modeling, I designed and implemented PyExoCross, a Python program adapted from the original Fortran program Ex-

oCross. PyExoCross allows for flexible and efficient computation of a wide range of spectral properties including partition functions, specific heats, cooling functions, radiative lifetimes, oscillator strengths. It also calculates both absorption and emission intensities as well as cross sections, under both LTE and non-LTE conditions. PyExoCross significantly increases the usability of databases like ExoMol and HITRAN. These capabilities allow users to model spectra with high precision across various thermodynamic environments and excitation regimes.

Another significant contribution is the development of a new high-resolution spectroscopic database, ExoMolHR, which is designed to meet the growing demand for accurate line-by-line data in high-resolution exoplanet spectroscopy. This database is based on the ExoMol database by extracting high accuracy and high-resolution line lists and calculating line intensities with custom specified parameters. To ensure accessibility and usability of this new resource, I also developed a website for the ExoMolHR database. This platform provides a structured and accessible interface and data source for supporting high-resolution spectral analysis and atmospheric retrievals. The development of ExoMolHR database provides an important high-resolution resource for the study of exoplanetary atmospheres, enhancing the precision of spectral analysis in observational astrophysics.

These contributions aim to support the spectral analysis in astrophysical studies by providing calculating tools and high-quality molecular data for the interpretation of astronomical and laboratory spectra.

10.3 Future work

While this research has made significant advancements, there remain areas for future exploration and improvement.

10.3.1 Major databases update

One potential avenue is the continuous updating and expansion of the ExoMol database with additional molecular species and refined spectroscopic parameters. This includes incorporating new high-accuracy ab initio calculations and experimental data to further improve the reliability of these databases. The `.def` file

for new EXOMOL line list updates is currently created and maintained manually.

The ongoing development of EXOMOLHR opens new possibilities for high-resolution spectroscopy in exoplanetary science, allowing for more detailed atmospheric characterization of exoplanets. Expanding this database with more molecular species and finer spectral resolution will be crucial for upcoming space missions and ground-based observatories aiming to detect and characterize exoplanetary atmospheres.

10.3.2 Automatic maintenance of the EXOMOL database

The maintenance of the EXOMOL database is a tedious and repetitive task. To improve the efficiency of database management, it is essential to automate this process using Python scripts. Each time a new line list file is added, the program must extract molecular or atomic information from the line list, such as the number of states and transitions, and whether the file includes uncertainty, lifetimes, or Landé g-factors. However, the quantum number labels and formats in the states file must be retrieved from the corresponding published paper, posing a challenge for full automation. Currently, this information is manually extracted from the literature and incorporated into the definition files. When an existing line list file is updated, the program must determine whether modifications to the definition file are necessary by checking the version date recorded in the file. If needed, the version date and relevant information must be updated accordingly.

One of the challenges in generating definition files is the automated retrieval of InChI and InChIKey. In many cases, the InChIKey cannot be directly obtained by searching the NIST website. Instead, it often requires manual intervention, where the InChI is first determined based on prior knowledge and then converted into an InChIKey through the NIST website, see Section 4.2.4. Automating this process with a program is particularly challenging due to the lack of a straightforward computational method for consistently obtaining accurate results.

Additionally, the master file requires automatic detection of updates. This can be achieved by iterating through the version numbers in all definition files to identify entries that need to be updated or newly added.

10.3.3 Continuous expansion of the EXOMOL database with atomic data

The atomic data will be available on EXOMOL website in the future. While EXOMOL has traditionally focused on molecular line lists, the inclusion of atomic data has been explicitly requested by database users. Its inclusion significantly expand its applicability, especially for modeling high-temperature astrophysical and exoplanetary environments where both atomic and molecular species play crucial roles. The introduction of high-accuracy atomic line lists will be particularly beneficial for studying stellar spectra, brown dwarfs, and in particular hot exoplanetary atmospheres. Future research should focus on developing a framework for incorporating atomic transition data into the ExoMol infrastructure, ensuring consistency with existing molecular datasets and maintaining high standards of accuracy and completeness.

10.3.4 Improve PYEXOCROSS

An important direction is the further development of computational tools such as PYEXOCROSS, extending its functionalities to support additional spectroscopic calculations and data visualization techniques. Some CPU parallelization is available, at present there is no GPU support.

Future developments for PYEXOCROSS include k-table generation, opacity production, and a library of fractional isotopic abundances to compute intensities for different isotopologue distributions. The program will also enhance its capabilities by supporting more molecular spectral databases and facilitating data type and format conversions across various databases. Additionally, improvements in line profile modeling, particularly in non-LTE conditions, could further enhance the accuracy of spectroscopic simulations. Future work could also focus on applying machine learning techniques to improve the efficiency of data processing and spectral analysis, particularly for large datasets.

10.3.5 Link to other data products

Another important future direction is improving the integration of EXOMOL with other spectroscopic databases and data products. Modern astrophysical and atmospheric research relies on multiple sources of spectroscopic data. Creating systematic links between EXOMOL and these resources will enable researchers to seamlessly access and compare data across different platforms. This effort could involve the development of cross-database query tools, improved metadata standardization, and automated data format conversions. By strengthening these connections, EXOMOL will become a more versatile and comprehensive resource for the scientific community, facilitating more accurate and integrated modeling of complex spectral environments.

Appendix A

API structure in JSON format

Listing A.1: The structure of API in JSON format used for search queries using the `HTTP GET` request.

```
{
    "<ISOFORMULA>": {
        "molecule": "<MOLECULE>",
        "mass /u": <MASS>,
        "energylevels": {
            "data type": "other States files"
        },
        "linelist": {
             "data type": "line list",
             "<DATASET>": {
                "description": "<DATASET-DESCRIPTION>",
                "recommended": false,
                "files": [
                    {
                        "url": "<URL-TRANS>",
            "size": <SIZE>,
                        "description": "<TRANS-FILE-DESCRIPTION>"
                    },
                    ...
                    {
                        "url": "<URL-STATES>",
            "size": <SIZE>,
                        "description": "<STATES-FILE-DESCRIPTION>"
                    }
                ]
        },
            ...
            "<DATASET>": {
                "description": "<DATASET-DESCRIPTION>",
                "recommended": true,
                "files": [
                    {
                        "url": "<URL-TRANS>",
            "size": <SIZE>,
                        "description": "<TRANS-FILE-DESCRIPTION>"
                    },
                    ...
                    {
                        "url": "<URL-STATES>",
            "size": <SIZE>,
```

```
                        "description": "<STATES-FILE-DESCRIPTION>"
                }
                ]
            }
        },
        "partitionfunction": {
        "data type": "partition function",
        "<DATASET>": {
                "description": "<DATASET-DESCRIPTION>",
                "recommended": false,
                "files": [
                {
                        "url": "<URL-PF>",
            "size": <SIZE>,
                        "description": "<PF-FILE-DESCRIPTION>"
                    }
            ]
            },
            ...
        "<DATASET>": {
                "description": "<DATASET-DESCRIPTION>",
                "recommended": true,
                "files": [
                {
                        "url": "<URL-PF>",
            "size": <SIZE>,
                        "description": "<PF-FILE-DESCRIPTION>"
                    }
                ]
            }
        }
    }
    ...
    "<ISOFORMULA>": {
        "molecule": "<MOLECULE>",
        "mass /u": <MASS>,
        "energylevels": {
            "data type": "other States files"
        },
        "linelist": {
        ...
        },
        "partitionfunction": {
        ...
        }
    }
}
```

Appendix B

JSON format master file

Listing B.1: JSON format master file `exomol.all.json`.


```
{
    "ID": "EXOMOL.master",
    "version": 20241023,
    "num_molecules": 91,
    "num_isotopologues": 225,
    "num_datasets": 107,
    "num_species": 270,
    "molecules": {
        "H2O": {
            "num_molecule_names": 11,
            "molecule_names": [
                "Water",
                "Oxidane",
                "Hydrogen oxide",
                "Dihydrogen monoxide",
                "Hydrogen monoxide",
                "Dihydrogen oxide",
                "Hydrogen hydroxide",
                "Hydric acid",
                "Hydrohydroxic acid",
                "Hydroxic acid",
                "Hydrol"
            ],
            "num_isotopologues": 7,
            "linelist": [
                {
                    "inchikey": "XLYOFNOQVPJJNP-UHFFFAOYSA-N",
                    "iso_slug": "1H2-16O",
                    "iso_formula": "(1H)2(16O)",
                    "dataset_name": "POKAZATEL",
                    "version": 20180501
                },
                {
                    "inchikey": "XLYOFNOQVPJJNP-OUBTZVSYSA-N",
                    "iso_slug": "1H2-17O",
                    "iso_formula": "(1H)2(17O)",
                    "dataset_name": "HotWat78",
                    "version": 20161222
                },
                {
                    "inchikey": "XLYOFNOQVPJJNP-NJFSPNSNSA-N",
```

```
                    "iso_slug": "1H2-18O",
                    "iso_formula": "(1H)2(18O)",
                    "dataset_name": "HotWat78",
                    "version": 20161222
                },
                {
                    "inchikey": "XLYOFNOQVPJJNP-DYCDLGHISA-N",
                    "iso_slug": "1H-2H-16O",
                    "iso_formula": "(1H)(2H)(16O)",
                    "dataset_name": "VTT",
                    "version": 20160726
                },
                {
                    "inchikey": "XLYOFNOQVPJJNP-DYCDLGHISA-N",
                    "iso_slug": "1H-2H-16O",
                    "iso_formula": "(1H)(2H)(16O)",
                    "dataset_name": "Hewitt",
                    "version": 20161222
                },
                {
                    "inchikey": "XLYOFNOQVPJJNP-DYCDLGHISA-N",
                    "iso_slug": "1H-2H-16O",
                    "iso_formula": "(1H)(2H)(16O)",
                    "dataset_name": "TDB",
                    "version": 20240603
                },
                {
                    "inchikey": "XLYOFNOQVPJJNP-ZSJDYOACSA-N",
                    "iso_slug": "2H2-16O",
                    "iso_formula": "(2H)2(16O)",
                    "dataset_name": "Hewitt",
                    "version": 20161222
                }
            ]
        }
    }
}
```

Appendix C

JSON format definition file

Listing C.1: JSON format definition (`.def.json`) file of $^{1}H_{2}{}^{16}O$ POKAZATEL dataset.


```
{
    "isotopologue": {
        "iso_formula": "(1H)2(16O)",
        "iso_slug": "1H2-16O",
        "inchikey": "XLYOFNOQVPJJNP-UHFFFAOYSA-N",
        "inchi": "1S/H2O/h1H2",
        "mass_in_Da": 18.010565,
        "point_group": "C2v"
    },
    "atoms": {
        "number_of_atoms": 3,
        "element": {
            "H": 1,
            "O": 16
        }
    },
    "nuclear_spin_degeneracy": {
        "A1": 1,
        "A2": 1,
        "B1": 3,
        "B2": 3
    },
    "dataset": {
        "name": "POKAZATEL",
        "version": 20230621,
        "doi": "10.1093/mnras/sty1877",
        "max_temperature": 6000.0,
        "num_pressure_broadeners": 4,
        "nxsec_files": 0,
        "nkcoeff_files": 0,
        "dipole_available": false,
        "cooling_function_available": false,
        "specific_heat_available": true,
        "continuum": null,
        "predis": null,
        "states": {
            "number_of_states": 810269,
            "max_energy": 41200.00,
            "uncertainties_available": true,
            "lifetime_available": false,
            "lande_g_available": false,
```

```
"num_quanta": 6,
"states_file_fields": [
    {
        "name": "ID",
        "desc": "Unique integer identifier for the energy level",
        "ffmt": "I12",
        "cfmt": "%12d"
    },
    {
        "name": "E",
        "desc": "State energy in cm-1",
        "ffmt": "F12.6",
        "cfmt": "%12.6f"
    },
    {
        "name": "gtot",
        "desc": "Total energy level degeneracy",
        "ffmt": "I6",
        "cfmt": "%6d"
    },
    {
        "name": "J",
        "desc":
            "Total rotational quantum number, excluding nuclear spin",
        "ffmt": "I7",
        "cfmt": "%7d"
    },
    {
        "name": "unc",
        "desc": "Energy uncertainty in cm-1",
        "ffmt": "F12.6",
        "cfmt": "%12.6f"
    },
    {
        "name": "nltcs:Ka",
        "ffmt": "I2",
        "cfmt": "%2d",
        "desc": "Ka rotational quantum number"
    },
    {
        "name": "nltcs:Kc",
        "ffmt": "I2",
        "cfmt": "%2d",
        "desc": "Kc rotational quantum number"
    },
    {
        "name": "nltcs:v1",
        "ffmt": "I2",
        "cfmt": "%2d",
        "desc": "v1 symmetric stretch quantum number(3)"
    },
    {
        "name": "nltcs:v2",
        "ffmt": "I2",
        "cfmt": "%2d",
        "desc": "v2 bend quantum number"
    },
    {
        "name": "nltcs:v3",
        "ffmt": "I2",
        "cfmt": "%2d",
        "desc": "v3 asymmetrics stretch quantum number"
    },
```

```
            {
                "name": "nltcs:Grve",
                "ffmt": "A3",
                "cfmt": "%3s",
                "desc": "Rovibrational symmetry label"
            },
            {
                "name": "Auxiliary:Ecal",
                "desc": "Energy in cm-1 from variational spectroscopic model",
                "ffmt": "F12.6",
                "cfmt": "%12.6f"
            },
            {
                "name": "Auxiliary:SourceType",
                "ffmt": "A2",
                "cfmt": "%2s",
                "desc": "Indicates if the value is from MARVEL (Ma) or
                    calculated (Ca)"
            }
        ]
    },
    "transitions": {
        "number_of_transitions": 5745071340,
        "number_of_transition_files": 412,
        "max_wavenumber": 41200.00
    }
},
"partition_function": {
    "max_partition_function_temperature": 10000.0,
    "partition_function_step_size": 1.0
},
"broad": {
    "default_Lorentzian_half-width": 0.07,
    "default_temperature_exponent": 0.5,
    "H2": {
        "filename": "1H2-16O__H2.broad",
        "max_J": 50,
        "Lorentzian_half_width": 0.0209,
        "temperature_exponent": 0.027,
        "quantum_number_sets": [
            {
                "code": "a3",
                "num_lines": 23731,
                "quantum_numbers": [
                    "J'",
                    "Ka\"",
                    "Ka'"
                ]
            },
            {
                "code": "a1",
                "num_lines": 100,
                "quantum_numbers": [
                    "J'"
                ]
            },
            {
                "code": "a0",
                "num_lines": 51,
                "quantum_numbers": []
            }
        ]
    },
```

```
        "He": {
            "filename": "1H2-16O__He.broad",
            "max_J": 50,
            "Lorentzian_half_width": 0.0042,
            "temperature_exponent": 0.02,
            "quantum_number_sets": [
                {
                    "code": "b1",
                    "num_lines": 253,
                    "quantum_numbers": [
                        "J'",
                        "Ka\"",
                        "Ka'",
                        "Kc\"",
                        "Kc'",
                        "v1\"",
                        "v2\"",
                        "v3\"",
                        "v1'",
                        "v2'",
                        "v3'"
                    ]
                },
                {
                    "code": "a3",
                    "num_lines": 23731,
                    "quantum_numbers": [
                        "J'",
                        "Ka\"",
                        "Ka'"
                    ]
                },
                {
                    "code": "a1",
                    "num_lines": 100,
                    "quantum_numbers": [
                        "J'"
                    ]
                },
                {
                    "code": "a0",
                    "num_lines": 51,
                    "quantum_numbers": []
                }
            ]
        }
    }
}
```

Bibliography

[1] William Pluriel. Hot Exoplanetary Atmospheres in 3D. *Remote Sensing*, 15 (3):635, 2023. ISSN 2072-4292. doi: 10.3390/rs15030635.

[2] Jayne L. Birkby. *Spectroscopic Direct Detection of Exoplanets*, page 1485–1508. Springer International Publishing, 2018. ISBN 9783319553337. doi: 10.1007/978-3-319-55333-7_16.

[3] Jonathan Tennyson, Sergei N. Yurchenko, Ahmed F. Al-Refaie, Emma J. Barton, Katy L. Chubb, Phillip A. Coles, S. Diamantopoulou, Maire N. Gorman, Christian Hill, Aden Z. Lam, Lorenzo Lodi, Laura K. McKemmish, Yueqi Na, Alec Owens, Oleg L. Polyansky, T Rivlin, Clara Sousa-Silva, Daniel S. Underwood, Andrey Yachmenev, and Emil Zak. The ExoMol database: molecular line lists for exoplanet and other hot atmospheres. *J. Mol. Spectrosc.*, 327:73–94, 2016. doi: 10.1016/j.jms.2016.05.002.

[4] A. Owens, S. Dooley, L. McLaughlin, B. Tan, G. Zhang, S. N. Yurchenko, and J. Tennyson. ExoMol line lists – XLV. Rovibronic molecular line lists of calcium monohydride (CaH) and magnesium monohydride (MgH). *MNRAS*, 511:5448–5461, 2022. doi: 10.1093/mnras/stac371.

[5] O. L. Polyansky, A. A. Kyuberis, N. F. Zobov, J. Tennyson, S. N. Yurchenko, and L. Lodi. ExoMol molecular line lists XXX: a complete high-accuracy line list for water. *MNRAS*, 480:2597–2608, 2018. doi: 10.1093/mnras/sty1877.

[6] S. N. Yurchenko, Thomas M. Mellor, Richard S. Freedman, and J. Tennyson.

ExoMol molecular line lists XXXIX: Ro-vibrational molecular line list for CO_2. *MNRAS*, 496:5282–5291, 2020. doi: 10.1093/mnras/staa1874.

[7] D. S. Underwood, J. Tennyson, S. N. Yurchenko, Xinchuan Huang, David W. Schwenke, Timothy J. Lee, S. Clausen, and A. Fateev. ExoMol line lists XIV: A line list for hot SO_2. *MNRAS*, 459:3890–3899, 2016. doi: 10.1093/mnras/stw849.

[8] G. B. Mitev, M Pezzella, C A Bowesman, Jingxin Zhang, S. N. Yurchenko, and J. Tennyson. ExoMol Photodissociation Cross Sections – II: Continuum Absorption and Predissociation Spectra for the Hydroxyl Radical. *MNRAS*, 2025.

[9] Anna-Maree Syme and Laura K. McKemmish. Full spectroscopic model and trihybrid experimental-perturbative-variational line list for CN. *MNRAS*, 505:4383–4395, 2021. doi: 10.1093/mnras/stab1551.

[10] L S Rothman, D Jacquemart, A Barbe, D C Benner, M Birk, L R Brown, M R Carleer, C Chackerian, K Chance, L H Coudert, V Dana, V M Devi, J-M Flaud, R R Gamache, A Goldman, J-M Hartmann, K W Jucks, A G Maki, J-Y Mandin, S T Massie, J Orphal, A Perrin, C P Rinsland, M A H Smith, J Tennyson, R N Tolchenov, R A Toth, J Vander Auwera, P Varanasi, and G Wagner. The *HITRAN* 2004 molecular spectroscopic database. *J. Quant. Spectrosc. Radiat. Transf.*, 96:139–204, 2005. doi: 10.1016/j.jqsrt.2004.10.008.

[11] S. N. Yurchenko, W. Szajna, R. Hakalla, M. Semenov, A Sokolov, J. Tennyson, Y. Pavlenko, and M. R. Schmidt. ExoMol line lists – LIV. Empirical line lists for AlH and AlD and Emission spectroscopy of AlD in $A\ ^1\Pi$ ($v = 0, 1, 2$). *MNRAS*, 527:9736–9756, 2024. doi: 10.1093/mnras/stad3802.

[12] Qianwei Qu, Sergei N. Yurchenko, and Jonathan Tennyson. ExoMol molecular line lists – XLII: Rovibronic molecular line list for the low-lying states of NO. *MNRAS*, 504:5768–5777, 2021. doi: 10.1093/mnras/stab1154.

[13] I. E. Gordon, L. S. Rothman, R. J. Hargreaves, R. Hashemi, E. V. Karlovets, F. M. Skinner, E. K. Conway, C. Hill, R. V. Kochanov, Y. Tan, P. Wcisło, A. A. Finenko, K. Nelson, P. F. Bernath, M. Birk, V. Boudon, A. Campargue, K. V. Chance, A. Coustenis, B. J. Drouin, J.–M. Flaud, R. R. Gamache, J. T. Hodges, D. Jacquemart, E. J. Mlawer, A. V. Nikitin, V. I. Perevalov, M. Rotger, J. Tennyson, G. C. Toon, H. Tran, V. G. Tyuterev, E. M. Adkins, A. Baker, A. Barbe, E. Canè, A. G. Császár, A. Dudaryonok, O. Egorov, A. J. Fleisher, H. Fleurbaey, A. Foltynowicz, T. Furtenbacher, J. J. Harrison, J.–M. Hartmann, V.–M. Horneman, X. Huang, T. Karman, J. Karns, S. Kassi, I. Kleiner, V. Kofman, F. Kwabia–Tchana, N. N. Lavrentieva, T. J. Lee, D. A. Long, A. A. Lukashevskaya, O. M. Lyulin, V. Yu. Makhnev, W. Matt, S. T. Massie, M. Melosso, S. N. Mikhailenko, D. Mondelain, H. S. P. Müller, O. V. Naumenko, A. Perrin, O. L. Polyansky, E. Raddaoui, P. L. Raston, Z. D. Reed, M. Rey, C. Richard, R. Tóbiás, I. Sadiek, D. W. Schwenke, E. Starikova, K. Sung, F. Tamassia, S. A. Tashkun, J. Vander Auwera, I. A. Vasilenko, A. A. Vigasin, G. L. Villanueva, B. Vispoel, G. Wagner, A. Yachmenev, and S. N. Yurchenko. The HITRAN2020 molecular spectroscopic database. *J. Quant. Spectrosc. Radiat. Transf.*, 277:107949, 2022. doi: 10.1016/j.jqsrt.2021.107949.

[14] E. J. Barton, C. Hill, M. Czurylo, Heng-Ying Li, A Hyslop, S. N. Yurchenko, and J. Tennyson. The ExoMol diet: H_2 and He line-broadening parameters. *J. Quant. Spectrosc. Radiat. Transf.*, 203:490–495, 2017. doi: 10.1016/j.jqsrt.2017.01.028.

[15] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald,

Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020. doi: 10.1038/s41592-019-0686-2.

[16] Christian Hill. *Learning scientific programming with Python*. Cambridge University Press, 2020. doi: 10.1017/9781108778039.

[17] Jonathan Tennyson, Sergei N. Yurchenko, Jingxin Zhang, Charles A. Bowesman, Ryan P. Brady, Jeanna Buldyreva, Katy L. Chubb, Robert R. Gamache, Maire N. Gorman, Elizabeth R. Guest, Christian Hill, Kyriaki Kefala, A. E. Lynas-Gray, Thomas M. Mellor, Laura K. McKemmish, Georgi B. Mitev, Irina I. Mizus, Alec Owens, Zhijian Peng, Armando N. Perri, Marco Pezzella, Oleg L. Polyansky, Qianwei Qu, Mikhail Semenov, Oleksiy Smola, Andrei Solokov, Wilfrid Somogyi, Apoorva Upadhyay, Samuel O. M. Wright, and Nikolai F. Zobov. The 2024 release of the ExoMol database: molecular line lists for exoplanet and other hot atmospheres. *J. Quant. Spectrosc. Radiat. Transf.*, 326:109083, 2024. doi: 10.1016/j.jqsrt.2024.109083.

[18] J. Tennyson and S. N. Yurchenko. ExoMol: molecular line lists for exoplanet and other atmospheres. *MNRAS*, 425:21–33, 2012. doi: 10.1111/j.1365-2966.2012.21440.x.

[19] Carlo Maria Zwölf and Nicolas Moreau. The VAMDC species database. *The European Physical Journal D*, 78(7), 2024. ISSN 1434-6079. doi: 10.1140/epjd/s10053-024-00863-1.

[20] M.L. Dubernet, V. Boudon, J.L. Culhane, M.S. Dimitrijevic, A.Z. Fazliev, C. Joblin, F. Kupka, G. Leto, P. Le Sidaner, P.A. Loboda, H.E. Mason, N.J. Mason, C. Mendoza, G. Mulas, T.J. Millar, L.A. Nuñez, V.I. Perevalov, N. Piskunov, Y. Ralchenko, G. Rixon, L.S. Rothman, E. Roueff, T.A. Ryabchikova, A. Ryabtsev, S. Sahal-Bréchot, B. Schmitt, S. Schlemmer, J. Tennyson, V.G. Tyuterev, N.A. Walton, V. Wakelam, and C.J. Zeippen. Vir-

tual atomic and molecular data centre. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 111(15):2151–2159, 2010. ISSN 0022-4073. doi: 10.1016/j.jqsrt.2010.05.004.

[21] Felipe Pezoa, Juan L Reutter, Fernando Suarez, Martín Ugarte, and Domagoj Vrgoč. Foundations of JSON schema. In *Proceedings of the 25th International Conference on World Wide Web*, pages 263–273. International World Wide Web Conferences Steering Committee, 2016. doi: 10.1145/2872427.2883029.

[22] J. Tennyson and S. N. Yurchenko. The ExoMol atlas of molecular opacities. *Atoms*, 6:26, 2018. doi: 10.3390/atoms6020026.

[23] Björn Benneke, Ian Wong, Caroline Piaulet, Heather A. Knutson, Joshua Lothringer, Caroline V. Morley, Ian J. M. Crossfield, Peter Gao, Thomas P. Greene, Courtney Dressing, Diana Dragomir, Andrew W. Howard, Peter R. McCullough, Eliza M.-R. Kempton, Jonathan J. Fortney, and Jonathan Fraine. Water Vapor and Clouds on the Habitable-zone Sub-Neptune Exoplanet K2-18b. *The Astrophysical Journal Letters*, 887(1):L14, 2019. ISSN 2041-8213. doi: 10.3847/2041-8213/ab59dc.

[24] Giovanna Tinetti, Pierre Drossart, Paul Eccleston, Paul Hartogh, Astrid Heske, Jérémy Leconte, Giusi Micela, Marc Ollivier, Göran Pilbratt, Ludovic Puig, Diego Turrini, Bart Vandenbussche, Paulina Wolkenberg, Jean-Philippe Beaulieu, Lars A. Buchave, Martin Ferus, Matt Griffin, Manuel Guedel, Kay Justtanont, Pierre-Olivier Lagage, Pedro Machado, Giuseppe Malaguti, Michiel Min, Hans Ulrik Nørgaard-Nielsen, Mirek Rataj, Tom Ray, Ignasi Ribas, Mark Swain, Robert Szabo, Stephanie Werner, Joanna Barstow, Matt Burleigh, James Cho, Vincent Coudé du Foresto, Athena Coustenis, Leen Decin, Therese Encrenaz, Marina Galand, Michael Gillon, Ravit Helled, Juan Carlos Morales, Antonio García Muñoz, Andrea Moneti, Isabella Pagano, Enzo Pascale, Giuseppe Piccioni, David Pinfield, Subhajit Sarkar, Franck Selsis, Jonathan Tennyson, Amaury Triaud, Olivia

Venot, Ingo Waldmann, David Waltham, Gillian Wright, Jerome Amiaux, Jean-Louis Auguères, Michel Berthé, Naidu Bezawada, Georgia Bishop, Neil Bowles, Deirdre Coffey, Josep Colomé, Martin Crook, Pierre-Elie Crouzet, Vania Da Peppo, Isabel Escudero Sanz, Mauro Focardi, Martin Frericks, Tom Hunt, Ralf Kohley, Kevin Middleton, Gianluca Morgante, Roland Ottensamer, Emanuele Pace, Chris Pearson, Richard Stamper, Kate Symonds, Miriam Rengel, Etienne Renotte, Peter Ade, Laura Affer, Christophe Alard, Nicole Allard, Francesca Altieri, Yves André, Claudio Arena, Ioannis Argyriou, Alan Aylward, Cristian Baccani, Gaspar Bakos, Marek Banaszkiewicz, Mike Barlow, Virginie Batista, Giancarlo Bellucci, Serena Benatti, Pernelle Bernardi, Bruno Bézard, Maria Blecka, Emeline Bolmont, Bertrand Bonfond, Rosaria Bonito, Aldo S. Bonomo, John Robert Brucato, Allan Sacha Brun, Ian Bryson, Waldemar Bujwan, Sarah Casewell, Bejamin Charnay, Cesare Cecchi Pestellini, Guo Chen, Angela Ciaravella, Riccardo Claudi, Rodolphe Clédassou, Mario Damasso, Mario Damiano, Camilla Danielski, Pieter Deroo, Anna Maria Di Giorgio, Carsten Dominik, Vanessa Doublier, Simon Doyle, René Doyon, Benjamin Drummond, Bastien Duong, Stephen Eales, Billy Edwards, Maria Farina, Ettore Flaccomio, Leigh Fletcher, François Forget, Steve Fossey, Markus Fränz, Yuka Fujii, Álvaro García-Piquer, Walter Gear, Hervé Geoffray, Jean Claude Gérard, Lluis Gesa, H. Gomez, Rafał Graczyk, Caitlin Griffith, Denis Grodent, Mario Giuseppe Guarcello, Jacques Gustin, Keiko Hamano, Peter Hargrave, Yann Hello, Kevin Heng, Enrique Herrero, Allan Hornstrup, Benoit Hubert, Shigeru Ida, Masahiro Ikoma, Nicolas Iro, Patrick Irwin, Christopher Jarchow, Jean Jaubert, Hugh Jones, Queyrel Julien, Shingo Kameda, Franz Kerschbaum, Pierre Kervella, Tommi Koskinen, Matthijs Krijger, Norbert Krupp, Marina Lafarga, Federico Landini, Emanuel Lellouch, Giuseppe Leto, A. Luntzer, Theresa Rank-Lüftinger, Antonio Maggio, Jesus Maldonado, Jean-Pierre Maillard, Urs Mall, Jean-Baptiste Marquette, Stephane Mathis, Pierre Maxted, Taro Matsuo, Alexander Medvedev, Yamila Miguel,

Vincent Minier, Giuseppe Morello, Alessandro Mura, Norio Narita, Valerio Nascimbeni, N. Nguyen Tong, Vladimiro Noce, Fabrizio Oliva, Enric Palle, Paul Palmer, Maurizio Pancrazzi, Andreas Papageorgiou, Vivien Parmentier, Manuel Perger, Antonino Petralia, Stefano Pezzuto, Ray Pierrehumbert, Ignazio Pillitteri, Giampaolo Piotto, Giampaolo Pisano, Loredana Prisinzano, Aikaterini Radioti, Jean-Michel Réess, Ladislav Rezac, Marco Rocchetto, Albert Rosich, Nicoletta Sanna, Alexandre Santerne, Giorgio Savini, Gaetano Scandariato, Bruno Sicardy, Carles Sierra, Giuseppe Sindoni, Konrad Skup, Ignas Snellen, Mateusz Sobiecki, Lauriane Soret, Alessandro Sozzetti, A. Stiepen, Antoine Strugarek, Jake Taylor, William Taylor, Luca Terenzi, Marcell Tessenyi, Angelos Tsiaras, C. Tucker, Diana Valencia, Gautam Vasisht, Allona Vazan, Francesc Vilardell, Sabrine Vinatier, Serena Viti, Rens Waters, Piotr Wawer, Anna Wawrzaszek, Anthony Whitworth, Yuk L. Yung, Sergey N. Yurchenko, María Rosa Zapatero Osorio, Robert Zellem, Tiziano Zingales, and Frans Zwart. A chemical survey of exoplanets with ARIEL. *Exp. Astron.*, 46:135–209, 2018. doi: 10.1007/s10686-018-9598-x.

[25] Jonas Kuhn, Nicole Bobrowski, Guillaume Boudoire, Sergio Calabrese, Giovanni Giuffrida, Marco Liuzzo, Katcho Karume, Dario Tedesco, Thomas Wagner, and Ulrich Platt. High-spectral-resolution Fabry-Perot interferometers overcome fundamental limitations of present volcanic gas remote sensing techniques. *Frontiers In Earth Science*, 11:1039093, 2023. doi: 10.3389/feart.2023.1039093.

[26] Sara Seager. *Exoplanet Atmospheres: Physical Processes*. Princeton University Press, 2010. ISBN 9780691119144. doi: 10.2307/j.ctvcm4gvv.

[27] Nikku Madhusudhan. Exoplanetary Atmospheres: Key Insights, Challenges, and Prospects. *Annual Review of Astronomy and Astrophysics*, 57(1):617–663, 2019. ISSN 1545-4282. doi: 10.1146/annurev-astro-081817-051846.

[28] Michael R Line, Aaron S Wolf, Xi Zhang, Heather Knutson, Joshua A Kammer, Elias Ellison, Pieter Deroo, Dave Crisp, and Yuk L Yung. A system-

atic retrieval analysis of secondary eclipse spectra. I. A comparison of atmospheric retrieval techniques. *The Astrophysical Journal*, 775(2):137, 2013. doi: 10.1088/0004-637X/775/2/137.

[29] Michael R Line, Heather Knutson, Aaron S Wolf, and Yuk L Yung. A systematic retrieval analysis of secondary eclipse spectra. II. A uniform analysis of nine planets and their C to O ratios. *The Astrophysical Journal*, 783(2): 70, 2014. doi: 10.1088/0004-637X/783/2/70.

[30] Caitlin A Griffith. Disentangling degenerate solutions from primary transit and secondary eclipse spectroscopy of exoplanets. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 372(2014):20130086, 2014. doi: 10.1098/rsta.2013.0086.

[31] Siddharth Gandhi, Matteo Brogi, Sergei N Yurchenko, Jonathan Tennyson, Phillip A Coles, Rebecca K Webb, Jayne L Birkby, Gloria Guilluy, George A Hawker, Nikku Madhusudhan, Aldo S Bonomo, and Alessandro Sozzetti. Molecular cross-sections for high-resolution spectroscopy of super-Earths, warm Neptunes, and hot Jupiters. *Monthly Notices of the Royal Astronomical Society*, 495(1):224–237, 2020. ISSN 1365-2966. doi: 10.1093/mnras/staa981.

[32] Ignas A. G. Snellen, Remco J. de Kok, Ernst J. W. de Mooij, and Simon Albrecht. The orbital motion, absolute mass and high-altitude winds of exoplanet HD 209458b. *Nature*, 465(7301):1049–1051, 2010. ISSN 1476-4687. doi: 10.1038/nature09111.

[33] Matteo Brogi, Ignas A. G. Snellen, Remco J. de Kok, Simon Albrecht, Jayne Birkby, and Ernst J. W. de Mooij. The signature of orbital motion from the dayside of the planet τ Boötis b. *Nature*, 486(7404):502–504, 2012. ISSN 1476-4687. doi: 10.1038/nature11161.

[34] Michel Mayor and Didier Queloz. A Jupiter-mass companion to a solar-

type star. *Nature*, 378(6555):355–359, 1995. ISSN 1476-4687. doi: 10.1038/378355a0.

[35] E. F. Spring, J. L. Birkby, L. Pino, R. Alonso, S. Hoyer, M. E. Young, P. R. T. Coelho, D. Nespral, and M. López-Morales. Black Mirror: The impact of rotational broadening on the search for reflected light from 51 Pegasi b with high resolution spectroscopy. *Astronomy & Astrophysics*, 659:A121, 2022. ISSN 1432-0746. doi: 10.1051/0004-6361/202142314.

[36] M. Brogi, R. J. de Kok, J. L. Birkby, H. Schwarz, and I. A. G. Snellen. Carbon monoxide and water vapor in the atmosphere of the non-transiting exoplanet HD 179949 b. *Astronomy & Astrophysics*, 565:A124, 2014. ISSN 1432-0746. doi: 10.1051/0004-6361/201423537.

[37] Elyar Sedaghati, Henri M. J. Boffin, Ryan J. MacDonald, Siddharth Gandhi, Nikku Madhusudhan, Neale P. Gibson, Mahmoudreza Oshagh, Antonio Claret, and Heike Rauer. Detection of titanium oxide in the atmosphere of a hot Jupiter. *Nature*, 549(7671):238–241, 2017. ISSN 1476-4687. doi: 10.1038/nature23651.

[38] Patrick Gaulme, François-Xavier Schmider, and Ivan Gonçalves. Measuring planetary atmospheric dynamics with Doppler spectroscopy. *Astronomy & Astrophysics*, 617:A41, 2018. ISSN 1432-0746. doi: 10.1051/0004-6361/201832868.

[39] George A. Hawker and Ian R. Parry. High-resolution spectroscopy and high contrast imaging with the ELT: looking for O_2 in Proxima b. *Monthly Notices of the Royal Astronomical Society*, 484(4):4855–4864, 2019. ISSN 1365-2966. doi: 10.1093/mnras/stz323.

[40] C. Lovis, I. Snellen, D. Mouillet, F. Pepe, F. Wildi, N. Astudillo-Defru, J.-L. Beuzit, X. Bonfils, A. Cheetham, U. Conod, X. Delfosse, D. Ehrenreich, P. Figueira, T. Forveille, J. H. C. Martins, S. P. Quanz, N. C. Santos, H.-M. Schmid, D. Ségransan, and S. Udry. Atmospheric characterization of

Proxima b by coupling the SPHERE high-contrast imager to the ESPRESSO spectrograph. *Astronomy & Astrophysics*, 599:A16, 2017. ISSN 1432-0746. doi: 10.1051/0004-6361/201629682.

[41] Emily O. Garvin, Markus J. Bonse, Jean Hayoz, Gabriele Cugno, Jonas Spiller, Polychronis A. Patapis, Dominique Petit dit de la Roche, Rakesh Nath-Ranga, Olivier Absil, Nicolai F. Meinshausen, and Sascha P. Quanz. Machine learning for exoplanet detection in high-contrast spectroscopy: Revealing exoplanets by leveraging hidden molecular signatures in cross-correlated spectra with convolutional neural networks. *Astronomy & Astrophysics*, 689:A143, 2024. ISSN 1432-0746. doi: 10.1051/0004-6361/202449149.

[42] Brendan P. Bowler. Imaging Extrasolar Giant Planets. *Publications of the Astronomical Society of the Pacific*, 128(968):102001, 2016. ISSN 1538-3873. doi: 10.1088/1538-3873/128/968/102001.

[43] M. Houllé, A. Vigan, A. Carlotti, É. Choquet, F. Cantalloube, M. W. Phillips, J.-F. Sauvage, N. Schwartz, G. P. P. L. Otten, I. Baraffe, A. Emsenhuber, and C. Mordasini. Direct imaging and spectroscopy of exoplanets with the ELT/HARMONI high-contrast module. *Astronomy & Astrophysics*, 652: A67, 2021. ISSN 1432-0746. doi: 10.1051/0004-6361/202140479.

[44] Christian Marois, Bruce Macintosh, Travis Barman, B. Zuckerman, Inseok Song, Jennifer Patience, David Lafrenière, and René Doyon. Direct Imaging of Multiple Planets Orbiting the Star HR 8799. *Science*, 322(5906): 1348–1352, 2008. ISSN 1095-9203. doi: 10.1126/science.1166585.

[45] H. J. Hoeijmakers, R. J. de Kok, I. A. G. Snellen, M. Brogi, J. L. Birkby, and H. Schwarz. A search for TiO in the optical high-resolution transmission spectrum of HD 209458b: Hindrance due to inaccuracies in the line database. *A&A*, 575:A20, 2015. doi: 10.1051/0004-6361/201424794.

[46] Robert J. Hargreaves, Iouli E. Gordon, Michael Rey, Andrei V. Nikitin,

Vladimir G. Tyuterev, Roman V. Kochanov, and Laurence S. Rothman. An accurate, extensive, and practical line list of methane for the HITEMP database. *ApJS*, 247:55, 2020. doi: 10.3847/1538-4365/ab7a1a.

[47] J. Tennyson and S. N. Yurchenko. High Accuracy Molecular Line Lists for Studies of Exoplanets and Other Hot Atmospheres. *Frontiers Astron. Space Sci.*, 8:795040, 2022. doi: 10.3389/fspas.2021.795040.

[48] J Tennyson, S N Yurchenko, A. F. Al-Refaie, V. H. J. Clark, K. L. Chubb, E. K. Conway, A. Dewan, M. N. Gorman, C. Hill, A. E. Lynas-Gray, T. Mellor, L. K. McKemmish, A. Owens, O. L. Polyansky, M. Semenov, W. Somogyi, G. Tinetti, A. Upadhyay, I. Waldmann, Y. Wang, S. Wright, and O. P. Yurchenko. The 2020 release of the ExoMol database: Molecular line lists for exoplanet and other hot atmospheres. *J. Quant. Spectrosc. Radiat. Transf.*, 255:107228, 2020. doi: 10.1016/j.jqsrt.2020.107228.

[49] Michaël Rey, Andrei V. Nikitin, Yurii L. Babikov, and Vladimir G. Tyuterev. TheoReTS – an information system for theoretical spectra based on variational predictions from molecular potential energy and dipole moment surfaces. *J. Mol. Spectrosc.*, 327:138–158, 2016. doi: 10.1016/j.jms.2016.04.006.

[50] Xinchuan Huang, David W. Schwenke, and Timothy J. Lee. What It Takes to Compute Highly Accurate Rovibrational Line Lists for Use in Astrochemistry. *Acc. Chem. Res.*, 54:1311–1321, 2021. doi: 10.1021/acs.accounts.0c00624.

[51] Robert L. Kurucz. Including all the lines. *Can. J. Phys.*, 89:417–428, 2011. doi: 10.1139/p10-104.

[52] T. Delahaye, R. Armante, N.A. Scott, N. Jacquinet-Husson, A. Chédin, L. Crépeau, C. Crevoisier, V. Douet, A. Perrin, A. Barbe, V. Boudon, A. Campargue, L.H. Coudert, V. Ebert, J.-M. Flaud, R.R. Gamache, D. Jacquemart, A. Jolly, F. Kwabia Tchana, A. Kyuberis, G. Li, O.M. Lyulin, L. Manceron,

S. Mikhailenko, N. Moazzen-Ahmadi, H.S.P. Müller, O.V. Naumenko, A. Nikitin, V.I Perevalov, C. Richard, E. Starikova, S.A. Tashkun, Vl.G. Tyuterev, J. Vander Auwera, B. Vispoel, A. Yachmenev, and S. Yurchenko. The 2020 edition of the GEISA spectroscopic database. *J. Mol. Spectrosc.*, 380:111510, 2021. doi: 10.1016/j.jms.2021.111510.

[53] I. E. Gordon, L. S. Rothman, R. J. Hargreaves, R. Hashemi, E. V. Karlovets, F. M. Skinner, E. K. Conway, C. Hill, R. V. Kochanov, Y. Tan, P. Wcisło, A. A. Finenko, K. Nelson, P. F. Bernath, M. Birk, V. Boudon, A. Campargue, K. V. Chance, A. Coustenis, B. J. Drouin, J.-M. Flaud, R. R. Gamache, J. T. Hodges, D. Jacquemart, E. J. Mlawer, A. V. Nikitin, V. I. Perevalov, M. Rotger, J. Tennyson, G. C. Toon, H. Tran, V. G. Tyuterev, E. M. Adkins, A. Baker, A. Barbe, E. Canè, A. G. Császár, A. Dudaryonok, O. Egorov, A. J. Fleisher, H. Fleurbaey, A. Foltynowicz, T. Furtenbacher, J. J. Harrison, J.-M. Hartmann, V.-M. Horneman, X. Huang, T. Karman, J. Karns, S. Kassi, I. Kleiner, V. Kofman, F. KwabiaÔÇôTchana, N. N. Lavrentieva, T. J. Lee, D. A. Long, A. A. Lukashevskaya, O. M. Lyulin, V. Yu. Makhnev, W. Matt, S. T. Massie, M. Melosso, S. N. Mikhailenko, D. Mondelain, H. S. P. Müller, O. V. Naumenko, A. Perrin, O. L. Polyansky, E. Raddaoui, P. L. Raston, Z. D. Reed, M. Rey, C. Richard, R. Tóbiás, I. Sadiek, D. W. Schwenke, E. Starikova, K. Sung, F. Tamassia, S. A. Tashkun, J. Vander Auwera, I. A. Vasilenko, A. A. Vigasin, G. L. Villanueva, B. Vispoel, G. Wagner, A. Yachmenev, and S. N. Yurchenko. The HITRAN2020 molecular spectroscopic database. *J. Quant. Spectrosc. Radiat. Transf.*, 277:107949, 2022. doi: 10.1016/j.jqsrt.2021.107949.

[54] L.S. Rothman, I.E. Gordon, R.J. Barber, H. Dothe, R.R. Gamache, A. Goldman, V.I. Perevalov, S.A. Tashkun, and J. Tennyson. HITEMP, the high-temperature molecular spectroscopic database. *J. Quant. Spectrosc. Radiat. Transf.*, 111:2139 – 2150, 2010. doi: 10.1016/j.jqsrt.2010.05.001.

[55] R. J. Hargreaves, I. E. Gordon, L. S. Rothman, S. A. Tashkun, V. I. Perevalov,

A. A. Lukashevskaya, S. N. Yurchenko, J. Tennyson, and H. S. P. Müller. Spectroscopic line parameters of NO, NO_2, and N_2O for the HITEMP database. *J. Quant. Spectrosc. Radiat. Transf.*, 232:35 – 53, 2019. doi: 10.1016/j.jqsrt.2019.04.040.

[56] H M Pickett, R L Poynter, E A Cohen, M L Delitsky, J C Pearson, and H S P Müller. Submillimeter, millimeter, and microwave spectral line catalog. *J. Quant. Spectrosc. Radiat. Transf.*, 60:883–890, 1998.

[57] Holger SP Müller, CP Endres, J Stutzki, and S Schlemmer. The CDMS view on molecular data needs of Herschel, SOFIA, and ALMA. In *AIP Conference Proceedings*, volume 1545, pages 96–109. American Institute of Physics, 2013. doi: 10.1063/1.4815844.

[58] C. P. Endres, S. Schlemmer, P. Schilke, J. Stutzki, and H. S. P. Müller. The Cologne Database for Molecular Spectroscopy, CDMS, in the Virtual Atomic and Molecular Data Centre, VAMDC. *J. Mol. Spectrosc.*, 327:95–104, 2016. doi: 10.1016/j.jms.2016.03.005.

[59] C. Richard, O. Ben Fathallah, P. Hardy, R. Kamel, M. Merkulova, M. Rotger, O.N. Ulenikov, and V. Boudon. CaSDa24: Latest updates to the Dijon calculated spectroscopic databases. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 327:109127, 2024. ISSN 0022-4073. doi: 10.1016/j.jqsrt.2024.109127.

[60] C. Richard, V. Boudon, and M. Rotger. Calculated spectroscopic databases for the VAMDC portal: New molecules and improvements. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 251:107096, 2020. ISSN 0022-4073. doi: 10.1016/j.jqsrt.2020.107096.

[61] T. Ryabchikova, N. Piskunov, R. L. Kurucz, H. C. Stempels, U. Heiter, Yu Pakhomov, and P. S. Barklem. A major upgrade of the VALD database. *Phys. Scr.*, 90:054005, 2015. doi: 10.1088/0031-8949/90/5/054005.

[62] M. L. Dubernet, C. Boursier, O. Denis-Alpizar, Y. A. Ba, N. Moreau, C. M. Zwölf, M. A. Amor, D. Babikov, N. Balakrishnan, C. Balança, M. Ben Khalifa, A. Bergeat, C. T. Bop, L. Cabrera-González, C. Cárdenas, A. Chefai, P. J. Dagdigian, F. Dayou, S. Demes, B. Desrousseaux, F. Dumouchel, A. Faure, R. C. Forrey, J. Franz, R. M. García-Vázquez, F. Gianturco, A. Godard Palluet, L. González-Sánchez, G. C. Groenenboom, P. Halvick, K. Hammami, F. Khadri, Y. Kalugina, I. Kleiner, J. Kłos, F. Lique, J. Loreau, B. Mandal, B. Mant, S. Marinakis, D. Ndaw, P. Pirlot Jankowiak, T. Price, E. Quintas-Sánchez, R. Ramachandran, E. Sahnoun, C. Santander, P. C. Stancil, T. Stoecklin, J. Tennyson, F. Tonolo, R. Urzúa-Leiva, B. Yang, E. Yurtsever, and M. Żółtowski. BASECOL2023 scientific content. *Astronomy & Astrophysics*, 683:A40, 2024. ISSN 1432-0746. doi: 10.1051/0004-6361/202348233.

[63] A. D. Bykov, A. Z. Fazliev, N. N. Filippov, A. V. Kozodoev, A. I. Privezentsev, L. N. Sinitsa, M. V. Tonkov, and M. Yu. Tretyakov. Distributed information system on atmospheric spectroscopy. In *15th Symposium on High-Resolution Molecular Spectroscopy*, volume 6580, pages 228–239. Proc. SPIE, 2006. doi: 10.1117/12.724944.

[64] A. G. Császár, A. Z. Fazliev, and J. Tennyson. W@DIS - Prototype information system for systematization of spectral data of water. In *20th Colloquium on High Resolution Molecular Spectroscopy, Dijon, France*, pages 270–271, 2007.

[65] A. Akhlyostin, Z. Apanovich, A. Fazliev, A. Kozodoev, N. Lavrentiev, A. Privezentsev, O. Rodimova, S. a, A. G. Császár, and J. Tennyson. The current status of the W@DIS information system. In *22nd International Symposium on Atmospheric and Ocean Optics: Atmospheric Physics*, volume 10035, pages 96–108. Proc. SPIE, 2016. doi: 10.1117/12.2249235.

[66] S.A. Tashkun, V.I. Perevalov, J-L. Teffo, A.D. Bykov, and N.N. Lavrentieva. CDSD-1000, the high-temperature carbon dioxide spectroscopic data-

bank. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 82(1–4): 165–196, 2003. ISSN 0022-4073. doi: 10.1016/s0022-4073(03)00152-3.

[67] S.A. Tashkun, V.I. Perevalov, R.R. Gamache, and J. Lamouroux. CDSD-296, high-resolution carbon dioxide spectroscopic databank: An update. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 228:124–131, 2019. ISSN 0022-4073. doi: 10.1016/j.jqsrt.2019.03.001.

[68] João Vargas, Bruno Lopez, and Mário Lino da Silva. CDSDv: A compact database for the modeling of high-temperature CO_2 radiation. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 245:106848, 2020. ISSN 0022-4073. doi: 10.1016/j.jqsrt.2020.106848.

[69] H. R. Hrodmarsson and E. F. van Dishoeck. Photodissociation and photoionization of molecules of astronomical interest updates to the leiden photodissociation and photoionization cross section database. *A&A*, 675:A25, 2023. doi: 10.1051/0004-6361/202346645.

[70] S. A. Tashkun, V. I. Perevalov, and N. N. Lavrentieva. NOSD-1000, the high-temperature nitrous oxide spectroscopic databank. *J. Quant. Spectrosc. Radiat. Transf.*, 177:43–48, 2016. doi: 10.1016/j.jqsrt.2015.11.014.

[71] A. A. Lukashevskaya, N. N. Lavrentieva, A. C. Dudaryonok, and V. I. Perevalov. NDSD-1000: High-resolution, high-temperature Nitrogen Dioxide Spectroscopic Databank. *J. Quant. Spectrosc. Radiat. Transf.*, 184:205–217, 2016. doi: 10.1016/j.jqsrt.2016.07.014.

[72] O. M. Lyulin and V. I. Perevalov. ASD-1000: high-resolution, high-temperature acetylene spectroscopic databank. *J. Quant. Spectrosc. Radiat. Transf.*, 201:94–103, 2017. doi: 10.1016/j.jqsrt.2017.06.032.

[73] Peter F Bernath. MoLLIST: Molecular Line Lists, Intensities and Spectra. *J. Quant. Spectrosc. Radiat. Transf.*, 240:106687, 2020. doi: 10.1016/j.jqsrt.2019.106687.

[74] R. A. McClatchey, W. S. Benedict, S. A. Clough, D. E. Burch, R. F. Calfee, K. Fox, L. S. Rothman, and J. S. Garing. AFCRL Atmospheric Absorption Line Parameters Compilation. *Optical Physics Laboratory*, (434), 1973.

[75] B. Plez, 2012. (P2012).

[76] James S. A. Brooke, Peter F. Bernath, Timothy W. Schmidt, and George B. Bacskay. Line strengths and updated molecular constants for the C_2 Swan system. *J. Quant. Spectrosc. Radiat. Transf.*, 124:11 – 20, 2013. doi: 10.1016/j.jqsrt.2013.02.025.

[77] T. Masseron, B. Plez, S. Van Eck, R. Colin, I. Daoutidis, M. Godefroid, P.-F. Coheur, P. Bernath, A. Jorissen, and N. Christlieb. CH in stellar atmospheres: an extensive linelist. *A&A*, 571:A47, 2014. doi: 10.1051/0004-6361/201423956.

[78] James S. A. Brooke, Ram S. Ram, Colin M. Western, Gang Li, David W. Schwenke, and Peter F. Bernath. Einstein A coefficients and oscillator strengths for the A $^2\Pi$ – X $^2\Sigma^+$ (red) and B $^2\Sigma^+$ – X $^2\Sigma^+$ violet systems and rovibrational transitions in the X $^2\Sigma^+$ state of CN. *ApJS*, 210:23, 2014. doi: 10.1088/0067-0049/210/2/23.

[79] R. L. Kurucz. Robert L. Kurucz on-line database of molecular line lists, CO A-X, X-X transitions, 2010. (KCO).

[80] R. L. Kurucz. Robert L. Kurucz on-line database of molecular line lists, OH A-X, X-X transitions, 2007. (KOH).

[81] M. T. I. Ibrahim, D. Alatoom, T. Furtenbacher, A. G. Császár, S. N. Yurchenko, A. A. A. Azzam, and J. Tennyson. MARVEL analysis of high-resolution rovibrational spectra of $^{13}C^{16}O_2$. *J. Comput. Chem.*, 45:969–984, 2024. doi: 10.1002/jcc.27266.

[82] Ehsan GharibNezhad, Alireza Shayesteh, and Peter F. Bernath. Einstein A coefficients for rovibronic lines of the $A\,^2\Pi \rightarrow X\,^2\Sigma^+$ and $B'\,^2\Sigma^+ \rightarrow X\,^2\Sigma^+$

transitions of MgH. *MNRAS*, 432:2043–2047, 2013. doi: 10.1093/mnras/stt510.

[83] R. L. Kurucz. Robert L. Kurucz on-line database of molecular line lists, SiH A-X transitions, 2010. (KSiH).

[84] James S. A. Brooke, Peter F. Bernath, and Colin M. Western. Note: Improved line strengths of rovibrational and rotational transitions within the X $^3\Sigma^-$ ground state of NH. *J. Chem. Phys.*, 143:026101, 2015. doi: 10.1063/1.4923422.

[85] Anton M. Fernando, Peter F. Bernath, James N. Hodges, and Thomas Masseron. A new linelist for the A $^3\Pi$ – X $^3\Sigma^-$ transition of the NH free radical. *J. Quant. Spectrosc. Radiat. Transf.*, 217:29–34, 2018. doi: 10.1016/j.jqsrt.2018.05.021.

[86] M-L Dubernet, M. H. Alexander, Y. A. Ba, N. Balakrishnan, C. Balanca, C. Ceccarelli, J. Cernicharo, F. Daniel, F. Dayou, M. Doronin, F. Dumouchel, A. Faure, N. Feautrier, D. R. Flower, A. Grosjean, P. Halvick, J. Klos, F. Lique, G. C. McBane, S. Marinakis, N. Moreau, R. Moszynski, D. A. Neufeld, E. Roueff, P. Schilke, A. Spielfiedel, P. C. Stancil, T. Stoecklin, J. Tennyson, B. Yang, A-M. Vasserot, and L. Wiesenfeld. BASECOL2012: A collisional database repository and web service within the Virtual Atomic and Molecular Data Centre (VAMDC). *A&A*, 553:A50, 2013. doi: 10.1051/0004-6361/201220630.

[87] G. G. Matvienko, V. I. Perevalov, Yu. N. Ponomarev, L. N. Sinitsa, and V. N. Cherepanov. High-Resolution Molecular Spectroscopy in Tomsk: Establishment, Development, and Current Status. *Russian Physics Journal*, 59(4): 490–501, 2016. ISSN 1573-9228. doi: 10.1007/s11182-016-0799-3.

[88] I. E. Gordon, L. S. Rothman, R. J. Hargreaves, R. Hashemi, E. V. Karlovets, F. M. Skinner, E. K. Conway, C. Hill, R. V. Kochanov, Y. Tan, P. Wcisło,

A. A. Finenko, K. Nelson, P. F. Bernath, M. Birk, V. Boudon, A. Campargue, K. V. Chance, A. Coustenis, B. J. Drouin, J.-M. Flaud, R. R. Gamache, J. T. Hodges, D. Jacquemart, E. J. Mlawer, A. V. Nikitin, V. I. Perevalov, M. Rotger, J. Tennyson, G. C. Toon, H. Tran, V. G. Tyuterev, E. M. Adkins, A. Baker, A. Barbe, E. Canè, A. G. Császár, A. Dudaryonok, O. Egorov, A. J. Fleisher, H. Fleurbaey, A. Foltynowicz, T. Furtenbacher, J. J. Harrison, J.-M. Hartmann, V.-M. Horneman, X. Huang, T. Karman, J. Karns, S. Kassi, I. Kleiner, V. Kofman, F. Kwabia–Tchana, N. N. Lavrentieva, T. J. Lee, D. A. Long, A. A. Lukashevskaya, O. M. Lyulin, V. Yu. Makhnev, W. Matt, S. T. Massie, M. Melosso, S. N. Mikhailenko, D. Mondelain, H. S. P. Müller, O. V. Naumenko, A. Perrin, O. L. Polyansky, E. Raddaoui, P. L. Raston, Z. D. Reed, M. Rey, C. Richard, R. Tóbiás, I. Sadiek, D. W. Schwenke, E. Starikova, K. Sung, F. Tamassia, S. A. Tashkun, J. Vander Auwera, I. A. Vasilenko, A. A. Vigasin, G. L. Villanueva, B. Vispoel, G. Wagner, A. Yachmenev, and S. N. Yurchenko. The HITRAN2020 molecular spectroscopic database. *J. Quant. Spectrosc. Radiat. Transf.*, 277:107949, 2022. doi: 10.1016/j.jqsrt.2021.107949.

[89] Jonathan Tennyson, Lorenzo Lodi, Laura K McKemmish, and Sergei N Yurchenko. The ab initio calculation of spectra of open shell diatomic molecules. *J. Phys. B-At. Mol. Opt. Phys.*, 49:102001, 2016. doi: 10.1088/0953-4075/49/10/102001.

[90] Jonathan Tennyson. Accurate variational calculations for line lists to model the vibration rotation spectra of hot astrophysical atmospheres. *WIREs Comput. Mol. Sci.*, 2:698–715, 2012. doi: 10.1002/wcms.94.

[91] J. Tennyson and S. N. Yurchenko. The ExoMol project: Software for computing molecular line lists. *Int. J. Quantum Chem.*, 117:92–103, 2017. doi: 10.1002/qua.25190.

[92] Callie E. Hood, Jonathan J. Fortney, Michael R. Line, Emily C. Martin, Caroline V. Morley, Jayne L. Birkby, Zafar Rustamkulov, Roxana E. Lupu, and

Richard S. Freedman. Prospects for Characterizing the Haziest Sub-Neptune Exoplanets with High-resolution Spectroscopy. *AJ.*, 160:198, 2020. doi: 10.3847/1538-3881/abb46b.

[93] Anjali A. A. Piette and Nikku Madhusudhan. On the Temperature Profiles and Emission Spectra of Mini-Neptune Atmospheres. *ApJ*, 904:154, 2020. doi: 10.3847/1538-4357/abbfb1.

[94] Shang-Min Tsai, Hamish Innes, Tim Lichtenberg, Jake Taylor, Matej Malik, Katy Chubb, and Raymond Pierrehumbert. Inferring Shallow Surfaces on Sub-Neptune Exoplanets with JWST. *ApJL*, 922:L27, 2021. doi: 10.3847/2041-8213/ac399a.

[95] D. Blain, B. Charnay, and B. Bézard. 1D atmospheric study of the temperate sub-Neptune K2-18b. *A&A*, 646:A15, 2021. doi: 10.1051/0004-6361/202039072.

[96] B. Charnay, D. Blain, B. Bezard, J. Leconte, M. Turbet, and A. Falco. Formation and dynamics of water clouds on temperate sub-Neptunes: the example of K2-18b. *A&A*, 646, 2021. doi: 10.1051/0004-6361/202039525.

[97] Danny Gasman, Michiel Min, and Katy L. Chubb. Investigating the detectability of hydrocarbons in exoplanet atmospheres with JWST. *A&A*, 659: A114, 2022. doi: 10.1051/0004-6361/202141468.

[98] Natasha E. Batalha, Angie Wolfgang, Johanna Teske, Munazza K. Alam, Lili Alderson, Natalie M. Batalha, Mercedes López-Morales, and Hannah R. Wakeford. Importance of Sample Selection in Exoplanet-atmosphere Population Studies. *AJ.*, 165:14, 2022. doi: 10.3847/1538-3881/ac9f45.

[99] Maria Zamyatina, Eric Hébrard, Benjamin Drummond, Nathan J Mayne, James Manners, Duncan A Christie, Pascal Tremblin, David K Sing, and Krisztian Kohary. Observability of signatures of transport-induced chemistry in clear atmospheres of hot gas giant exoplanets. *MNRAS*, 519:3129–3153, 2022. doi: 10.1093/mnras/stac3432.

[100] Savvas Constantinou and Nikku Madhusudhan. Characterizing atmospheres of cloudy temperate mini-Neptunes with JWST. *MNRAS*, 514:2073–2091, 2022. doi: 10.1093/mnras/stac1277.

[101] Andrea Guzmán-Mesa, Daniel Kitzmann, Christoph Mordasini, and Kevin Heng. Chemical diversity of the atmospheres and interiors of sub-Neptunes: a case study of GJ 436 b. *MNRAS*, 513:4015–4036, 2022. doi: 10.1093/mnras/stac1066.

[102] N Lowson, G Zhou, C X Huang, D J Wright, B Edwards, E Nabbie, A Venner, S N Quinn, K A Collins, E Gillen, M Battley, A Triaud, C Hellier, S Seager, J N Winn, J M Jenkins, B Wohler, A Shporer, R P Schwarz, F Murgas, E Pallé, D R Anderson, R G West, R A Wittenmyer, B P Bowler, J Horner, S R Kane, J Kielkopf, P Plavchan, H Zhang, T Fairnington, J Okumura, M W Mengel, and B C Addison. Two mini-Neptunes transiting the adolescent K-star HIP 113103 confirmed with TESS and CHEOPS. *MNRAS*, 527: 1146–1162, 2023. doi: 10.1093/mnras/stad2756.

[103] Peter Gao, Anjali A. A. Piette, Maria E. Steinrueck, Matthew C. Nixon, Michael Zhang, Eliza M.-R. Kempton, Jacob L. Bean, Emily Rauscher, Vivien Parmentier, Natasha E. Batalha, Arjun B. Savel, Kenneth E. Arnold, Michael T. Roman, Isaac Malsky, and Jake Taylor. The Hazy and Metal-rich Atmosphere of GJ 1214 b Constrained by Near- and Mid-infrared Transmission Spectroscopy. *ApJ*, 951:96, 2023. doi: 10.3847/1538-4357/acd16f.

[104] Eliza M. R. Kempton, Michael Zhang, Jacob L. Bean, Maria E. Steinrueck, Anjali A. A. Piette, Vivien Parmentier, Isaac Malsky, Michael T. Roman, Emily Rauscher, Peter Gao, Taylor J. Bell, Qiao Xue, Jake Taylor, Arjun B. Savel, Kenneth E. Arnold, Matthew C. Nixon, Kevin B. Stevenson, Megan Mansfield, Sarah Kendrew, Sebastian Zieba, Elsa Ducrot, Achrene Dyrek, Pierre-Olivier Lagage, Keivan G. Stassun, Gregory W. Henry, Travis Barman, Roxana Lupu, Matej Malik, Tiffany Kataria, Jegug Ih, Guangwei Fu, Luis Welbanks, and Peter McGill. A reflective, metal-rich atmosphere for

GJ 1214b from its JWST phase curve. *Nature*, 620:67–71, 2023. doi: 10.1038/s41586-023-06159-5.

[105] Thomas Mikal-Evans, Nikku Madhusudhan, Jason Dittmann, Maximilian N. Günther, Luis Welbanks, Vincent Van Eylen, Ian J. M. Crossfield, Tansu Daylan, and Laura Kreidberg. Hubble Space Telescope Transmission Spectroscopy for the Temperate Sub-Neptune TOI-270 d: A Possible Hydrogen-rich Atmosphere Containing Water Vapor. *AJ.*, 165:84, 2023. doi: 10.3847/1538-3881/aca90b.

[106] Holmberg, Måns and Madhusudhan, Nikku. Possible Hycean conditions in the sub-Neptune TOI-270 d. *A&A*, 683:L2, 2024. doi: 10.1051/0004-6361/202348238.

[107] Spandan Dash, Matteo Brogi, Siddharth Gandhi, Marina Lafarga, Annabella Meech, Aaron Bello-Arufe, and Peter J. Wheatley. Constraints on atmospheric water abundance and cloud deck pressure in the warm Neptune GJ 3470 b via CARMENES transmission spectroscopy. *arXiv e-prints*, 2024.

[108] Andrea Guzmán-Mesa, Daniel Kitzmann, Chloe Fisher, Adam J. Burgasser, H. Jens Hoeijmakers, Pablo Márquez-Neila, Simon L. Grimm, Avi M. Mandell, Raphael Sznitman, and Kevin Heng. Information Content of JWST NIRSpec Transmission Spectra of Warm Neptunes. *AJ.*, 160:15, 2020. doi: 10.3847/1538-3881/ab9176.

[109] Gloria Guilluy, Amélie Gressier, Sam Wright, Alexandre Santerne, Adam Yassin Jaziri, Billy Edwards, Quentin Changeat, Darius Modirrousta-Galian, Nour Skaf, Ahmed Al-Refaie, Robin Baeyens, Michelle Fabienne Bieger, Doriann Blain, Flavien Kiefer, Mario Morvan, Lorenzo V. Mugnai, William Pluriel, Mathilde Poveda, Tiziano Zingales, Niall Whiteford, Kai Hou Yip, Benjamin Charnay, Jérémy Leconte, Pierre Drossart, Alessandro Sozzetti, Emmanuel Marcq, Angelos Tsiaras, Olivia Venot, Ingo Waldmann, and Jean-Philippe Beaulieu. ARES IV: Probing the Atmospheres

of the Two Warm Small Planets HD 106315c and HD 3167c with the HST/WFC3 Camera*. *AJ.*, 161:19, 2020. doi: 10.3847/1538-3881/abc3c8.

[110] Napaporn A-thano, Supachai Awiphan, Ing-Guey Jiang, Eamonn Kerins, Akshay Priyadarshi, Iain McDonald, Yogesh C. Joshi, Thansuda Chulikorn, Joshua J. C. Hayes, Stephen Charles, Chung-Kai Huang, Ronnakrit Rattanamala, Li-Chin Yeh, and Vik S Dhillon. Revisiting the Transit Timing and Atmosphere Characterization of the Neptune-mass Planet HAT-P-26 b. *AJ.*, 166:223, 2023. doi: 10.3847/1538-3881/acfeea.

[111] Billy Edwards, Quentin Changeat, Angelos Tsiaras, Andrew Allan, Patrick Behr, Simone R. Hagey, Michael D. Himes, Sushuang Ma, Keivan G. Stassun, Luis Thomas, Alexandra Thompson, Aaron Boley, Luke Booth, Jeroen Bouwman, Kevin France, Nataliea Lowson, Annabella Meech, Caprice L. Phillips, Aline A. Vidotto, Kai Hou Yip, Michelle Bieger, Amelie Gressier, Estelle Janin, Ing-Guey Jiang, Pietro Leonardi, Subhajit Sarkar, Nour Skaf, Jake Taylor, Ming Yang, and Derek Ward-Thompson. Characterizing a World Within the Hot-Neptune Desert: Transit Observations of LTT 9779 b with the Hubble Space Telescope/WFC3. *AJ.*, 166:158, 2023. doi: 10.3847/1538-3881/acea77.

[112] Michael Radica, Louis-Philippe Coulombe, Jake Taylor, Loic Albert, Romain Allart, Björn Benneke, Nicolas B. Cowan, Lisa Dang, David Lafrenière, Daniel Thorngren, Étienne Artigau, René Doyon, Laura Flagg, Doug Johnstone, Stefan Pelletier, and Pierre-Alexis Roy. Muted Features in the JWST NIRISS Transmission Spectrum of Hot Neptune LTT 9779b. *ApJL*, 962(1):L20, 2024. doi: 10.3847/2041-8213/ad20e4.

[113] Nikku Madhusudhan, Anjali A. A. Piette, and Savvas Constantinou. Habitability and Biosignatures of Hycean Worlds. *ApJ*, 918:1, 2021. doi: 10.3847/1538-4357/abfd9c.

[114] Nikku Madhusudhan, Subhajit Sarkar, Savvas Constantinou, Måns Holm-

berg, Anjali A. A. Piette, and Julianne I. Moses. Carbon-bearing Molecules in a Possible Hycean Atmosphere. *ApJL*, 956:L13, 2023. doi: 10.3847/2041-8213/acf577.

[115] Eliza M.-R. Kempton, Madeline Lessard, Matej Malik, Leslie A. Rogers, Kate E. Futrowsky, Jegug Ih, Nadejda Marounina, and Carlos E. Mu noz Romero. Where are the Water Worlds?: Self-consistent Models of Water-rich Exoplanet Atmospheres. *ApJ*, 953(1):57, 2023. doi: 10.3847/1538-4357/ace10d.

[116] Jessica J Spake, David K Sing, Hannah R Wakeford, Nikolay Nikolov, Thomas Mikal-Evans, Drake Deming, Joanna K Barstow, David R Anderson, Aarynn L Carter, Michael Gillon, Jayesh M Goyal, Guillaume Hebrard, Coel Hellier, Tiffany Kataria, Kristine W F Lam, A H M J Triaud, and Peter J Wheatley. Abundance measurements of H_2O and carbon-bearing species in the atmosphere of WASP-127b confirm its supersolar metallicity. *MNRAS*, 500:4042–4064, 10 2020. doi: 10.1093/mnras/staa3116.

[117] Lara O. Anisman, Billy Edwards, Quentin Changeat, Olivia Venot, Ahmed F. Al-Refaie, Angelos Tsiaras, and Giovanna Tinetti. WASP-117 b: An Eccentric Hot Saturn as a Future Complex Chemistry Laboratory. *AJ.*, 160:233, 2020. doi: 10.3847/1538-3881/abb9b0.

[118] Q. Changeat, B. Edwards, A. F. Al-Refaie, M. Morvan, A. Tsiaras, I. P. Waldmann, and G. Tinetti. KELT-11 b: Abundances of Water and Constraints on Carbon-bearing Molecules from the Hubble Transmission Spectrum. *AJ.*, 160:260, 2020. doi: 10.3847/1538-3881/abbe12.

[119] Knicole D. Colón, Laura Kreidberg, Luis Welbanks, Michael R. Line, Nikku Madhusudhan, Thomas Beatty, Patrick Tamburo, Kevin B. Stevenson, Avi Mandell, Joseph E. Rodriguez, Thomas Barclay, Eric D. Lopez, Keivan G. Stassun, Daniel Angerhausen, Jonathan J. Fortney, David J. James, Joshua Pepper, John P. Ahlers, Peter Plavchan, Supachai Awiphan, Cliff Kot-

nik, Kim K. McLeod, Gabriel Murawski, Heena Chotani, Danny LeBrun, William Matzko, David Rea, Monica Vidaurri, Scott Webster, James K. Williams, Leafia Sheraden Cox, Nicole Tan, and Emily A. Gilbert. An Unusual Transmission Spectrum for the Sub-Saturn KELT-11b Suggestive of a Subsolar Water Abundance. *AJ.*, 160:280, 2020. doi: 10.3847/1538-3881/abc1e9.

[120] M. C. Maimone, M. Brogi, A. Chiavassa, M. E. van den Ancker, C. F. Manara, J. Leconte, S. Gandhi, and W. Pluriel. Detecting H_2O with CRIRES+: WASP-20b. *A&A*, 667:A106, 2022. doi: 10.1051/0004-6361/202244383.

[121] Eva-Maria Ahrer, Lili Alderson, Natalie M. Batalha, Natasha E. Batalha, Jacob L. Bean, Thomas G. Beatty, Taylor J. Bell, Bjorn Benneke, Zachory K. Berta-Thompson, Aarynn L. Carter, Ian J. M. Crossfield, Nestor Espinoza, Adina D. Feinstein, Jonathan J. Fortney, Neale P. Gibson, Jayesh M. Goyal, Eliza M-R Kempton, James Kirk, Laura Kreidberg, Mercedes Lopez-Morales, Michael R. Line, Joshua D. Lothringer, Sarah E. Moran, Sagnick Mukherjee, Kazumasa Ohno, Vivien Parmentier, Caroline Piaulet, Zafar Rustamkulov, Everett Schlawin, David K. Sing, Kevin B. Stevenson, Hannah R. Wakeford, Natalie H. Allen, Stephan M. Birkmann, Jonathan Brande, Nicolas Crouzet, Patricio E. Cubillos, Mario Damiano, Jean-Michel Desert, Peter Gao, Joseph Harrington, Renyu Hu, Sarah Kendrew, Heather A. Knutson, Pierre-Olivier Lagage, Jeremy Leconte, Monika Lendl, Ryan J. MacDonald, E. M. May, Yamila Miguel, Karan Molaverdikhani, Julianne Moses, I, Catriona Anne Murray, Molly Nehring, Nikolay K. Nikolov, D. J. M. Petit Dit de la Roche, Michael Radica, Pierre-Alexis Roy, Keivan G. Stassun, Jake Taylor, William C. Waalkes, Patcharapol Wachiraphan, Luis Welbanks, Peter J. Wheatley, Keshav Aggarwal, Munazza K. Alam, Agnibha Banerjee, Joanna K. Barstow, Jasmina Blecic, S. L. Casewell, Quentin Changeat, K. L. Chubb, Knicole D. Colon, Louis-Philippe Coulombe, Tansu Daylan, Miguel De Val-Borro, Leen Decin, Leonardo A. Dos Santos, Laura Flagg, Kevin France, Guangwei Fu, A. Garcia Munoz, John E. Gizis, Ana Glid-

den, David Grant, Kevin Heng, Thomas Henning, Yu-Cian Hong, Julie Inglis, Nicolas Iro, Tiffany Kataria, Thaddeus D. Komacek, Jessica E. Krick, Elspeth K. H. Lee, Nikole K. Lewis, Jorge Lillo-Box, Jacob Lustig-Yaeger, Luigi Mancini, Avi M. Mandell, Megan Mansfield, Mark S. Marley, Thomas Mikal-Evans, Giuseppe Morello, Matthew C. Nixon, Kevin Ortiz Ceballos, Anjali A. A. Piette, Diana Powell, Benjamin Rackham, V, Lakeisha Ramos-Rosado, Emily Rauscher, Seth Redfield, Laura K. Rogers, Michael T. Roman, Gael M. Roudier, Nicholas Scarsdale, Evgenya L. Shkolnik, John Southworth, Jessica J. Spake, Maria E. Steinrueck, Xianyu Tan, Johanna K. Teske, Pascal Tremblin, Shang-Min Tsai, Gregory S. Tucker, Jake D. Turner, Jeff A. Valenti, Olivia Venot, Ingo P. Waldmann, Nicole L. Wallack, Xi Zhang, and Sebastian Zieba. Identification of carbon dioxide in an exoplanet atmosphere. *Nature*, 614:649–652, 2023. doi: 10.1038/s41586-022-05269-w.

[122] Nikolay K Nikolov, David K Sing, Jessica J Spake, Barry Smalley, Jayesh M Goyal, Thomas Mikal-Evans, Hannah R Wakeford, Zafar Rustamkulov, Drake Deming, Jonathan J Fortney, Aarynn Carter, Neale P Gibson, and Nathan J Mayne. Solar-to-supersolar sodium and oxygen absolute abundances for a 'hot Saturn' orbiting a metal-rich star. *MNRAS*, 515:3037–3058, 2022. doi: 10.1093/mnras/stac1530.

[123] Shang-Min Tsai, Elspeth K. H. Lee, Diana Powell, Peter Gao, Xi Zhang, Julianne Moses, Eric Hébrard, Olivia Venot, Vivien Parmentier, Sean Jordan, Renyu Hu, Munazza K. Alam, Lili Alderson, Natalie M. Batalha, Jacob L. Bean, Björn Benneke, Carver J. Bierson, Ryan P. Brady, Ludmila Carone, Aarynn L. Carter, Katy L. Chubb, Julie Inglis, Jérémy Leconte, Michael Line, Mercedes López-Morales, Yamila Miguel, Karan Molaverdikhani, Zafar Rustamkulov, David K. Sing, Kevin B. Stevenson, Hannah R. Wakeford, Jeehyun Yang, Keshav Aggarwal, Robin Baeyens, Saugata Barat, Miguel de Val-Borro, Tansu Daylan, Jonathan J. Fortney, Kevin France, Jayesh M. Goyal, David Grant, James Kirk, Laura Kreidberg, Amy Louca, Sarah E.

Moran, Sagnick Mukherjee, Evert Nasedkin, Kazumasa Ohno, Benjamin V. Rackham, Seth Redfield, Jake Taylor, Pascal Tremblin, Channon Visscher, Nicole L. Wallack, Luis Welbanks, Allison Youngblood, Eva-Maria Ahrer, Natasha E. Batalha, Patrick Behr, Zachory K. Berta-Thompson, Jasmina Blecic, S. L. Casewell, Ian J. M. Crossfield, Nicolas Crouzet, Patricio E. Cubillos, Leen Decin, Jean-Michel Désert, Adina D. Feinstein, Neale P. Gibson, Joseph Harrington, Kevin Heng, Thomas Henning, Eliza M.-R. Kempton, Jessica Krick, Pierre-Olivier Lagage, Monika Lendl, Joshua D. Lothringer, Megan Mansfield, N. J. Mayne, Thomas Mikal-Evans, Enric Palle, Everett Schlawin, Oliver Shorttle, Peter J. Wheatley, and Sergei N. Yurchenko. Photochemically produced SO_2 in the atmosphere of WASP-39b. *Nature*, 617: 483–487, 2023. doi: 10.1038/s41586-023-05902-2.

[124] D. Samra, Ch. Helling, K. L. Chubb, M. Min, L. Carone, and A. D. Schneider. Clouds form on the hot Saturn JWST ERO target WASP-96b. *A&A*, 669: 17, 2023. doi: 10.1051/0004-6361/202244939.

[125] Michael Radica, Luis Welbanks, Néstor Espinoza, Jake Taylor, Louis-Philippe Coulombe, Adina D Feinstein, Jayesh Goyal, Nicholas Scarsdale, Loïc Albert, Priyanka Baghel, Jacob L Bean, Jasmina Blecic, David Lafreni‘ere, Ryan J MacDonald, Maria Zamyatina, Romain Allart, Étienne Artigau, Natasha E Batalha, Neil James Cook, Nicolas B Cowan, Lisa Dang, René Doyon, Marylou Fournier-Tondreau, Doug Johnstone, Michael R Line, Sarah E Moran, Sagnick Mukherjee, Stefan Pelletier, Pierre-Alexis Roy, Geert Jan Talens, Joseph Filippazzo, Klaus Pontoppidan, and Kevin Volk. Awesome SOSS: transmission spectroscopy of WASP-96b with NIRISS/-SOSS. *MNRAS*, 524:835–856, 2023. doi: 10.1093/mnras/stad1762.

[126] Lili Alderson, Hannah R. Wakeford, Munazza K. Alam, Natasha E. Batalha, Joshua D. Lothringer, Jea Adams Redai, Saugata Barat, Jonathan Brande, Mario Damiano, Tansu Daylan, Nestor Espinoza, Laura Flagg, Jayesh M. Goyal, David Grant, Renyu Hu, Julie Inglis, Elspeth K. H. Lee, Thomas

Mikal-Evans, Lakeisha Ramos-Rosado, Pierre-Alexis Roy, Nicole L. Wallack, Natalie M. Batalha, Jacob L. Bean, Bjorn Benneke, Zachory K. Berta-Thompson, Aarynn L. Carter, Quentin Changeat, Knicole D. Colon, Ian J. M. Crossfield, Jean-Michel Desert, Daniel Foreman-Mackey, Neale P. Gibson, Laura Kreidberg, Michael R. Line, Mercedes Lopez-Morales, Karan Molaverdikhani, Sarah E. Moran, Giuseppe Morello, Julianne I. Moses, Sagnick Mukherjee, Everett Schlawin, David K. Sing, Kevin B. Stevenson, Jake Taylor, Keshav Aggarwal, Eva-Maria Ahrer, Natalie H. Allen, Joanna K. Barstow, Taylor J. Bell, Jasmina Blecic, Sarah L. Casewell, Katy L. Chubb, Nicolas Crouzet, Patricio E. Cubillos, Leen Decin, Adina D. Feinstein, Joanthan J. Fortney, Joseph Harrington, Kevin Heng, Nicolas Iro, Eliza M. R. Kempton, James Kirk, Heather A. Knutson, Jessica Krick, Jeremy Leconte, Monika Lendl, Ryan J. MacDonald, Luigi Mancini, Megan Mansfield, Erin M. May, Nathan J. Mayne, Yamila Miguel, Nikolay K. Nikolov, Kazumasa Ohno, Enric Palle, Vivien Parmentier, Dominique J. M. Petit Dit de la Roche, Caroline Piaulet, Diana Powell, Benjamin V. Rackham, Seth Redfield, Laura K. Rogers, Zafar Rustamkulov, Xianyu Tan, P. Tremblin, Shang-Min Tsai, Jake D. Turner, Miguel de Val-Borro, Olivia Venot, Luis Welbanks, Peter J. Wheatley, and Xi Zhang. Early Release Science of the exoplanet WASP-39b with JWST NIRSpec G395H. *Nature*, 614:664–669, 2023. doi: 10.1038/s41586-022-05591-3.

[127] Marylou Fournier-Tondreau, Ryan J MacDonald, Michael Radica, David Lafrenière, Luis Welbanks, Caroline Piaulet, Louis-Philippe Coulombe, Romain Allart, Kim Morel, Étienne Artigau, Loïc Albert, Olivia Lim, René Doyon, Björn Benneke, Jason F Rowe, Antoine Darveau-Bernier, Nicolas B Cowan, Nikole K Lewis, Neil J Cook, Laura Flagg, Frédéric Genest, Stefan Pelletier, Doug Johnstone, Lisa Dang, Lisa Kaltenegger, Jake Taylor, and Jake D Turner. Near-infrared transmission spectroscopy of HAT-P-18 b with NIRISS: Disentangling planetary and stellar features in the era of JWST. *MNRAS*, 528:3354–3377, 2023. doi: 10.1093/mnras/stad3813.

[128] Jake Taylor, Michael Radica, Luis Welbanks, Ryan J MacDonald, Jasmina Blecic, Maria Zamyatina, Alexander Roth, Jacob L Bean, Vivien Parmentier, Louis-Philippe Coulombe, Adina D Feinstein, Néstor Espinoza, Björn Benneke, David Lafreni‘ere, René Doyon, and Eva-Maria Ahrer. Awesome SOSS: atmospheric characterization of WASP-96b using the JWST early release observations. *MNRAS*, 524:817–834, 2023. doi: 10.1093/mnras/stad1547.

[129] Anne Boucher, David Lafreniére, Stefan Pelletier, Antoine Darveau-Bernier, Michael Radica, Romain Allart, étienne Artigau, Neil J Cook, Florian Debras, René Doyon, Eric Gaidos, Björn Benneke, Charles Cadieux, Andres Carmona, Ryan Cloutier, Pía Cortés-Zuleta, Nicolas B Cowan, Xavier Delfosse, Jean-François Donati, Pascal Fouqué, Thierry Forveille, Konstantin Grankin, Guillaume Hébrard, Jorge H C Martins, Eder Martioli, Adrien Masson, and Sandrine Vinatier. CO or no CO? Narrowing the CO abundance constraint and recovering the H2O detection in the atmosphere of WASP-127 b using SPIRou. *MNRAS*, 522:5062–5083, 2023. doi: 10.1093/mnras/stad1247.

[130] Niall Whiteford, Alistair Glasse, Katy L Chubb, Daniel Kitzmann, Shrishmoy Ray, Mark W Phillips, Beth A Biller, Paul I Palmer, Ken Rice, Ingo P Waldmann, Quentin Changeat, Nour Skaf, Jason Wang, Billy Edwards, and Ahmed Al-Refaie. Retrieval study of cool, directly imaged exoplanet 51 Eri b. *MNRAS*, 525:1375–1400, 2023. doi: 10.1093/mnras/stad670.

[131] Siyi Xu, Hannah Diamond-Lowe, Ryan J. MacDonald, Andrew Vanderburg, Simon Blouin, P. Dufour, Peter Gao, Laura Kreidberg, S. K. Leggett, Andrew W. Mann, Caroline V. Morley, Andrew W. Stephens, Christopher E. O’Connor, Pa Chia Thao, and Nikole K. Lewis. Gemini/GMOS Transmission Spectroscopy of the Grazing Planet Candidate WD 1856+534 b. *AJ.*, 162:296, 2021. doi: 10.3847/1538-3881/ac2d26.

[132] G. Guilluy, P. Giacobbe, I Carleo, P. E. Cubillos, A. Sozzetti, A. S. Bonomo,

M. Brogi, S. Gandhi, L. Fossati, V Nascimbeni, D. Turrini, E. Schisano, F. Borsa, A. F. Lanza, L. Mancini, A. Maggio, L. Malavolta, G. Micela, L. Pino, M. Rainer, A. Bignamini, R. Claudi, R. Cosentino, E. Covino, S. Desidera, A. Fiorenzano, A. Harutyunyan, V Lorenzi, C. Knapic, E. Molinari, E. Pacetti, I Pagano, M. Pedani, G. Piotto, and E. Poretti. The GAPS Programme at TNG XXXVIII. Five molecules in the atmosphere of the warm giant planet WASP-69b detected at high spectral resolution. *A&A*, 665:A104, 2022. doi: 10.1051/0004-6361/202243854.

[133] Zhoujian Zhang, Paul Mollière, Keith Hawkins, Catherine Manea, Jonathan J. Fortney, Caroline V. Morley, Andrew Skemer, Mark S. Marley, Brendan P. Bowler, Aarynn L. Carter, Kyle Franson, Zachary G. Maas, and Christopher Sneden. ELemental abundances of Planets and brown dwarfs Imaged around Stars (ELPIS). I. Potential Metal Enrichment of the Exoplanet AF Lep b and a Novel Retrieval Approach for Cloudy Self-luminous Atmospheres. *AJ.*, 166:198, 2023. doi: 10.3847/1538-3881/acf768.

[134] Ilaria Carleo, Paolo Giacobbe, Gloria Guilluy, Patricio E. Cubillos, Aldo S. Bonomo, Alessandro Sozzetti, Matteo Brogi, Siddharth Gandhi, Luca Fossati, Diego Turrini, Katia Biazzo, Francesco Borsa, Antonino F. Lanza, Luca Malavolta, Antonio Maggio, Luigi Mancini, Giusi Micela, Lorenzo Pino, Ennio Poretti, Monica Rainer, Gaetano Scandariato, Eugenio Schisano, Gloria Andreuzzi, Andrea Bignamini, Rosario Cosentino, Aldo Fiorenzano, Avet Harutyunyan, Emilio Molinari, Marco Pedani, Seth Redfield, and Hristo Stoev. The GAPS Programme at TNG XXXIX. Multiple Molecular Species in the Atmosphere of the Warm Giant Planet WASP-80 b Unveiled at High Resolution with tIANO-B *. *A&A*, 164:101, 2022. doi: 10.3847/1538-3881/ac80bf.

[135] Taylor J. Bell, Luis Welbanks, Everett Schlawin, Michael R. Line, Jonathan J. Fortney, Thomas P. Greene, Kazumasa Ohno, Vivien Parmentier, Emily Rauscher, Thomas G. Beatty, Sagnick Mukherjee, Lindsey S. Wiser,

Martha L. Boyer, Marcia J. Rieke, and John A. Stansberry. Methane throughout the atmosphere of the warm exoplanet WASP-80b. *Nature*, 623:709–712, 2023. doi: 10.1038/s41586-023-06687-0.

[136] Shubham Kanodia, Suvrath Mahadevan, Jessica Libby-Roberts, Gudmundur Stefansson, Caleb I. Canas, Anjali A. A. Piette, Alan Boss, Johanna Teske, John Chambers, Greg Zeimann, Andrew Monson, Paul Robertson, Joe P. Ninan, Andrea S. J. Lin, Chad F. Bender, William D. Cochran, Scott A. Diddams, Arvind F. Gupta, Samuel Halverson, Suzanne Hawley, Henry A. Kobulnicky, Andrew J. Metcalf, Brock A. Parker, Luke Powers, Lawrence W. Ramsey, Arpita Roy, Christian Schwab, Tera N. Swaby, Ryan C. Terrien, and John Wisniewski. TOI-5205b: A Short-period Jovian Planet Transiting a Mid-M Dwarf. *AJ.*, 165:120, 2023. doi: 10.3847/1538-3881/acabce.

[137] Michael Zhang, Yayaati Chachan, Eliza M.-R. Kempton, Heather A. Knutson, and Wenjun (Happy) Chang. PLATON II: New Capabilities and a Comprehensive Retrieval on HD 189733b Transit and Eclipse Data. *ApJ*, 899:27, 2020. doi: 10.3847/1538-4357/aba1e6.

[138] Anjali A. A. Piette, Nikku Madhusudhan, Laura K. McKemmish, Siddharth Gandhi, Thomas Masseron, and Luis Welbanks. Assessing spectra and thermal inversions due to TiO in hot Jupiter atmospheres. *MNRAS*, 496:3870–3886, 2020. doi: 10.1093/mnras/staa1592.

[139] Kai Hou Yip, Quentin Changeat, Billy Edwards, Mario Morvan, Katy L. Chubb, Angelos Tsiaras, Ingo P. Waldmann, and Giovanna Tinetti. On the Compatibility of Ground-based and Space-based Data: WASP-96 b, an Example*. *AJ.*, 161:4, 2020. doi: 10.3847/1538-3881/abc179.

[140] M. Lendl, Sz. Csizmadia, A. Deline, L. Fossati, D. Kitzmann, K. Heng, S. Hoyer, S. Salmon, W. Benz, C. Broeg, D. Ehrenreich, A. Fortier, D. Queloz, A. Bonfanti, A. Brandeker, A. Collier Cameron, L. Delrez, A. Garcia Muñoz, M. J. Hooton, P. F. L. Maxted, B. M. Morris, V. Van Groo-

tel, T. G. Wilson, Y. Alibert, R. Alonso, J. Asquier, T. Bandy, T. Bárczy, D. Barrado, S. C. C. Barros, W. Baumjohann, M. Beck, T. Beck, A. Bekkelien, M. Bergomi, N. Billot, F. Biondi, X. Bonfils, V. Bourrier, M. D. Busch, J. Cabrera, V. Cessa, S. Charnoz, B. Chazelas, C. Corral Van Damme, M. B. Davies, M. Deleuil, O. D. S. Demangeon, B. O. Demory, A. Erikson, J. Farinato, M. Fridlund, D. Futyan, D. Gandolfi, M. Gillon, P. Guterman, J. Hasiba, E. Hernandez, K. G. Isaak, L. Kiss, T. Kuntzer, A. Lecavelier des Etangs, T. Lüftinger, J. Laskar, C. Lovis, D. Magrin, L. Malvasio, L. Marafatto, H. Michaelis, M. Munari, V. Nascimbeni, G. Olofsson, H. Ottacher, R. Ottensamer, I. Pagano, E. Pallé, G. Peter, D. Piazza, G. Piotto, D. Pollacco, F. Ratti, H. Rauer, R. Ragazzoni, N. Rando, I. Ribas, M. Rieder, R. Rohlfs, F. Safa, N. C. Santos, G. Scandariato, D. Ségransan, A. E. Simon, V. Singh, A. M. S. Smith, M. Sordet, S. G. Sousa, M. Steller, Gy. M. Szabó, N. Thomas, M. Tschentscher, S. Udry, V. Viotto, I. Walter, N. A. Walton, F. Wildi, and D. Wolter. The hot dayside and asymmetric transit of WASP-189 b seen by CHEOPS. *A&A*, 643:A94, 2020. doi: 10.1051/0004-6361/202038677.

[141] Brianna I. Lacy and Adam Burrows. JWST Transit Spectra. I. Exploring Potential Biases and Opportunities in Retrievals of Tidally Locked Hot Jupiters with Clouds and Hazes. *ApJ*, 905:131, 2020. doi: 10.3847/1538-4357/abc01c.

[142] G. Chen, N. Casasayas-Barris, E. Palle, L. Welbanks, N. Madhusudhan, R. Luque, and F. Murgas. Detection of Na in WASP-21b's lower and upper atmosphere. *A&A*, 642:A54, 2020. doi: 10.1051/0004-6361/202038661.

[143] Kielan K. Wilcomb, Quinn M. Konopacky, Travis S. Barman, Christopher A. Theissen, Jean-Baptiste Ruffio, Laci Brock, Bruce Macintosh, and Christian Marois. Moderate-resolution K-band Spectroscopy of Substellar Companion κ Andromedae b. *AJ.*, 160:207, 2020. doi: 10.3847/1538-3881/abb9b1.

[144] Paolo Giacobbe, Matteo Brogi, Siddharth Gandhi, Patricio E. Cubillos, Aldo S. Bonomo, Alessandro Sozzetti, Luca Fossati, Gloria Guilluy, Ilaria

Carleo, Monica Rainer, Avet Harutyunyan, Francesco Borsa, Lorenzo Pino, Valerio Nascimbeni, Serena Benatti, Katia Biazzo, Andrea Bignamini, Katy L. Chubb, Riccardo Claudi, Rosario Cosentino, Elvira Covino, Mario Damasso, Silvano Desidera, Aldo F. M. Fiorenzano, Adriano Ghedina, Antonino F. Lanza, Giuseppe Leto, Antonio Maggio, Luca Malavolta, Jesus Maldonado, Giuseppina Micela, Emilio Molinari, Isabella Pagano, Marco Pedani, Giampaolo Piotto, Ennio Poretti, Gaetano Scandariato, Sergei N. Yurchenko, Daniela Fantinel, Alberto Galli, Marcello Lodi, Nicoletta Sanna, and Andrea Tozzi. Five carbon- and nitrogen-bearing species in a hot giant planet's atmosphere. *Nature*, 592:205–208, 2021. doi: 10.1038/s41586-021-03381-x.

[145] Casasayas-Barris, N., Palle, E., Stangret, M., Bourrier, V., Tabernero, H. M., Yan, F., Borsa, F., Allart, R., Zapatero Osorio, M. R., Lovis, C., Sousa, S. G., Chen, G., Oshagh, M., Santos, N. C., Pepe, F., Rebolo, R., Molaro, P., Cristiani, S., Adibekyan, V., Alibert, Y., Allende Prieto, C., Bouchy, F., Demangeon, O. D. S., Di Marcantonio, P., D´Odorico, V., Ehrenreich, D., Figueira, P., Génova Santos, R., González Hernández, J. I., Lavie, B., Lillo-Box, J., Lo Curto, G., Martins, C. J. A. P., Mehner, A., Micela, G., Nunes, N. J., Poretti, E., Sozzetti, A., Suárez Mascareño, A., and Udry, S. The atmosphere of HD 209458b seen with ESPRESSO - No detectable planetary absorptions at high resolution. *A&A*, 647:A26, 2021. doi: 10.1051/0004-6361/202039539.

[146] Dilovan B. Serindag, Ignas A. G. Snellen, and Paul Mollière. Measuring titanium isotope ratios in exoplanet atmospheres. *A&A*, 655:A69, 2021. doi: 10.1051/0004-6361/202141941.

[147] Kyle B. Sheppard, Luis Welbanks, Avi M. Mandell, Nikku Madhusudhan, Nikolay Nikolov, Drake Deming, Gregory W. Henry, Michael H. Williamson, David K. Sing, Mercedes López-Morales, Jegug Ih, Jorge Sanz-Forcada, Panayotis Lavvas, Gilda E. Ballester, Thomas M. Evans, Antonio García

Muñoz, and Leonardo A. dos Santos. The Hubble PanCET Program: A Metal-rich Atmosphere for the Inflated Hot Jupiter HAT-P-41b. *ApJ*, 161: 51, 2021. doi: 10.3847/1538-3881/abc8f4.

[148] Marrick Braam, Floris F. S. van der Tak, Katy L. Chubb, and Michiel Min. Evidence for chromium hydride in the atmosphere of hot Jupiter WASP-31b. *A&A*, 646:A17, 2021. doi: 10.1051/0004-6361/202039509.

[149] Q. Changeat, A. F. Al-Refaie, B. Edwards, I. P. Waldmann, and G. Tinetti. An exploration of model degeneracies with a unified phase curve retrieval analysis: The light and dark sides of WASP-43 b. *ApJ*, 913:73, 2021. doi: 10.3847/1538-4357/abf2bb.

[150] Alexander D. Rathcke, Ryan J. MacDonald, Joanna K. Barstow, Jayesh M. Goyal, Mercedes Lopez-Morales, Jo ao M. Mendonça, Jorge Sanz-Forcada, Gregory W. Henry, David K. Sing, Munazza K. Alam, Nikole K. Lewis, Katy L. Chubb, Jake Taylor, Nikolay Nikolov, and Lars A. Buchhave. HST PanCET Program: A Complete Near-UV to Infrared Transmission Spectrum for the Hot Jupiter WASP-79b. *AJ.*, 162:138, 2021. doi: 10.3847/1538-3881/ac0e99.

[151] Masato Ishizuka, Hajime Kawahara, Stevanus K. Nugroho, Yui Kawashima, Teruyuki Hirano, and Motohide Tamura. Neutral Metals in the Atmosphere of HD 149026b. *AJ.*, 161:153, 2021. doi: 10.3847/1538-3881/abdb25.

[152] Patricio E. Cubillos, Dylan Keating, Nicolas B. Cowan, Johanna M. Vos, Ben Burningham, Marie Ygouf, Theodora Karalidi, Yifan Zhou, and Eileen C. Gonzales. Longitudinally Resolved Spectral Retrieval (ReSpect) of WASP-43b. *ApJ*, 915:45, 2021. doi: 10.3847/1538-4357/abfe14.

[153] Elspeth K. H. Lee, Vivien Parmentier, Mark Hammond, Simon L. Grimm, Daniel Kitzmann, Xianyu Tan, Shang-Min Tsai, and Raymond T. Pierrehumbert. Simulating gas giant exoplanet atmospheres with Exo-FMS: comparing

semigrey, picket fence, and correlated-*k* radiative-transfer schemes. *MNRAS*, 506:2695–2711, 2021. doi: 10.1093/mnras/stab1851.

[154] Robin Baeyens, Leen Decin, Ludmila Carone, Olivia Venot, Marcelino Agúndez, and Paul Molli‘ere. Grid of pseudo-2D chemistry models for tidally locked exoplanets - I. The role of vertical and horizontal mixing. *MNRAS*, 505:5603–5653, 2021. doi: 10.1093/mnras/stab1310.

[155] Jean-Baptiste Ruffio, Quinn M. Konopacky, Travis Barman, Bruce Macintosh, Kielan K. W. Hoch, Robert J. De Rosa, Jason J. Wang, Ian Czekala, and Christian Marois. Deep Exploration of the Planets HR 8799 b, c, and d with Moderate-resolution Spectroscopy. *AJ.*, 162:290, 2021. doi: 10.3847/1538-3881/ac273a.

[156] Elyar Sedaghati, Ryan J MacDonald, Núria Casasayas-Barris, H Jens Hoeijmakers, Henri M J Boffin, Florian Rodler, Rafael Brahm, Matías Jones, Alejandro Sánchez-López, Ilaria Carleo, Pedro Figueira, Andrea Mehner, and Manuel López-Puertas. A spectral survey of WASP-19b with ESPRESSO. *MNRAS*, 505:435–458, 2021. doi: 10.1093/mnras/stab1164.

[157] Michael R. Line, Matteo Brogi, Jacob L. Bean, Siddharth Gandhi, Joseph Zalesky, Vivien Parmentier, Peter Smith, Gregory N. Mace, Megan Mansfield, Eliza M-R Kempton, Jonathan J. Fortney, Evgenya Shkolnik, Jennifer Patience, Emily Rauscher, Jean-Michel Desert, and Joost P. Wardenier. A solar C/O and sub-solar metallicity in a hot Jupiter atmosphere. *Nature*, 598: 580+, 2021. doi: 10.1038/s41586-021-03912-6.

[158] Jason J. Wang, Jean-Baptiste Ruffio, Evan Morris, Jacques-Robert Delorme, Nemanja Jovanovic, Jacklyn Pezzato, Daniel Echeverri, Luke Finnerty, Callie Hood, J. J. Zanazzi, Marta L. Bryan, Charlotte Z. Bond, Sylvain Cetre, Emily C. Martin, Dimitri Mawet, Andy Skemer, Ashley Baker, Jerry W. Xuan, J. Kent Wallace, Ji Wang, Randall Bartos, Geoffrey A. Blake, Andy Boden, Cam Buzard, Benjamin Calvin, Mark Chun, Greg Dopp-

mann, Trent J. Dupuy, Gaspard Duchêne, Y. Katherina Feng, Michael P. Fitzgerald, Jonathan Fortney, Richard S. Freedman, Heather Knutson, Quinn Konopacky, Scott Lilley, Michael C. Liu, Ronald Lopez, Roxana Lupu, Mark S. Marley, Tiffany Meshkat, Brittany Miles, Maxwell Millar-Blanchaer, Sam Ragland, Arpita Roy, Garreth Ruane, Ben Sappey, Tobias Schofield, Lauren Weiss, Edward Wetherell, Peter Wizinowich, and Marie Ygouf. Detection and Bulk Properties of the HR 8799 Planets with High-resolution Spectroscopy. *AJ.*, 162:148, 2021. doi: 10.3847/1538-3881/ac1349.

[159] Luis Welbanks and Nikku Madhusudhan. On Atmospheric Retrievals of Exoplanets with Inhomogeneous Terminators. *ApJ*, 933:79, 2022. doi: 10.3847/1538-4357/ac6df1.

[160] Katy L. Chubb and Michiel Min. Exoplanet Atmosphere Retrievals in 3D Using Phase Curve Data with ARCiS: Application to WASP-43b. *A&A*, 665: A2, 2022. doi: 10.1051/0004-6361/202142800.

[161] Arianna Saba, Angelos Tsiaras, Mario Morvan, Alexandra Thompson, Quentin Changeat, Billy Edwards, Andrew Jolly, Ingo Waldmann, and Giovanna Tinetti. The Transmission Spectrum of WASP-17 b From the Optical to the Near-infrared Wavelengths: Combining STIS, WFC3, and IRAC Data Sets. *AJ.*, 164:2, 2022. doi: 10.3847/1538-3881/ac6c01.

[162] Q. Changeat, B. Edwards, A. F. Al-Refaie, A. Tsiaras, J. W. Skinner, J. Y. K. Cho, K. H. Yip, L. Anisman, M. Ikoma, M. F. Bieger, O. Venot, S. Shibata, I. P. Waldmann, and G. Tinetti. Five Key Exoplanet Questions Answered via the Analysis of 25 Hot-Jupiter Atmospheres in Eclipse. *ApJS*, 260:3, 2022. doi: 10.3847/1538-4365/ac5cc2.

[163] Aaron David Schneider, Ludmila Carone, Leen Decin, Uffe Gråe Jørgensen, Paul Mollière, Robin Baeyens, Sven Kiefer, and Christiane Helling. Exploring the deep atmospheres of HD 209458b and WASP-43b using a non-

gray general circulation model. *A&A*, 664:A56, 2022. doi: 10.1051/0004-6361/202142728.

[164] Aaron David Schneider, Ludmila Carone, Leen Decin, Uffe Gråe Jørgensen, and Christiane Helling. No evidence for radius inflation in hot Jupiters from vertical advection of heat. *A&A*, 666:L11, 2022. doi: 10.1051/0004-6361/202244797.

[165] A. F. Al-Refaie, Q. Changeat, O. Venot, I. P. Waldmann, and G. Tinetti. A Comparison of Chemical Models of Exoplanet Atmospheres Enabled by TauREx 3.1. *ApJ*, 932:123, 2022. doi: 10.3847/1538-4357/ac6dcd.

[166] David Kasper, Jacob L. Bean, Michael R. Line, Andreas Seifahrt, Madison T. Brady, Joshua Lothringer, Lorenzo Pino, Guangwei Fu, Stefan Pelletier, Julian Stürmer, Björn Benneke, Matteo Brogi, and Jean-Michel Désert. Unifying High- and Low-resolution Observations to Constrain the Dayside Atmosphere of KELT-20b/MASCARA-2b. *AJ.*, 165(1):7, 2022. doi: 10.3847/1538-3881/ac9f40.

[167] Rebecca K Webb, Siddharth Gandhi, Matteo Brogi, Jayne L Birkby, Ernst de Mooij, Ignas Snellen, and Yapeng Zhang. Water observed in the atmosphere of τ Boötis Ab with CARMENES/CAHA. *MNRAS*, 514:4160–4172, 2022. doi: 10.1093/mnras/stac1512.

[168] Russell Deitrick, Kevin Heng, Urs Schroffenegger, Daniel Kitzmann, Simon L Grimm, Matej Malik, João M Mendonça, and Brett M Morris. The THOR + HELIOS general circulation model: multiwavelength radiative transfer with accurate scattering by clouds/hazes. *MNRAS*, 512:3759–3787, 2022. doi: 10.1093/mnras/stac680.

[169] Martin Blažek, Petr Kabáth, Anjali A A Piette, Nikku Madhusudhan, Marek Skarka, Ján Šubjak, David R Anderson, Henri M J Boffin, Claudio C Cáceres, Neale P Gibson, Sergio Hoyer, Valentin D Ivanov, and Patricio M Rojo. Con-

straints on TESS albedos for five hot Jupiters. *MNRAS*, 513:3444–3457, 2022. doi: 10.1093/mnras/stac992.

[170] William Pluriel, Jeremy Leconte, Vivien Parmentier, Tiziano Zingales, Aurelien Falco, Franck Selsis, and Pascal Borde. Toward a multidimensional analysis of transmission spectroscopy II. Day-night-induced biases in retrievals from hot to ultrahot Jupiters. *A&A*, 658:A42, 2022. doi: 10.1051/0004-6361/202141943.

[171] Matthew C. Nixon and Nikku Madhusudhan. Aura-3D: A Three-dimensional Atmospheric Retrieval Framework for Exoplanet Transmission Spectra. *ApJ*, 935:73, 2022. doi: 10.3847/1538-4357/ac7c09.

[172] Petros Spyratos, Nikolay K. Nikolov, Savvas Constantinou, John Southworth, Nikku Madhusudhan, Elyar Sedaghati, David Ehrenreich, and Luigi Mancini. A precise blue-optical transmission spectrum from the ground: evidence for haze in the atmosphere of WASP-74b. *MNRAS*, 521:2163–2180, 2023. doi: 10.1093/mnras/stad637.

[173] Luc Bazinet, Stefan Pelletier, Björn Benneke, Ricardo Salinas, and Gregory N. Mace. A Subsolar Metallicity on the Ultra-short-period Planet HIP 65Ab. *AJ.*, 167(5):206, 2024. doi: 10.3847/1538-3881/ad3071.

[174] Alexander D Rathcke, Lars A Buchhave, João M Mendonça, David K Sing, Mercedes López-Morales, Munazza K Alam, Gregory W Henry, Nikolay K Nikolov, A García Muñoz, Thomas Mikal-Evans, Hannah R Wakeford, Leonardo A Dos Santos, and Vinesh Maguire Rajpaul. HST PanCET programme: a flat optical transmission spectrum for the Hot Jupiter WASP-101b. *MNRAS*, 522:582–594, 2023. doi: 10.1093/mnras/stad1010.

[175] Fang Wang, Quentin Changeat, Giovanna Tinetti, Diego Turrini, and Sam O. M. Wright. Constraining the atmospheric elements in hot Jupiters with Ariel. *MNRAS*, 523:4365–4380, 2023. doi: 10.1093/mnras/stad1721.

[176] P. E. Cubillos, L. Fossati, T. Koskinen, C. Huang, A. G. Sreejith, K. France, P. Wilson Cauley, and C. A. Haswell. The Hubble/STIS near-ultraviolet transmission spectrum of HD 189733 b. *A&A*, 671:A170, 2023. doi: 10.1051/0004-6361/202245064.

[177] Aaron Bello-Arufe, Heather A. Knutson, Joao M. Mendonca, Michael M. Zhang, Samuel H. C. Cabot, Alexander D. Rathcke, Ana Ulla, Shreyas Vissapragada, and Lars A. Buchhave. Transmission Spectroscopy of the Lowest-density Gas Giant: Metals and a Potential Extended Outflow in HAT-P-67b. *AJ.*, 166:69, 2023. doi: 10.3847/1538-3881/acd935.

[178] Katy L Chubb, Daphne M Stam, Christiane Helling, Dominic Samra, and Ludmila Carone. Modelling reflected polarized light from close-in giant exoplanet WASP-96b using PolHEx (Polarization of hot exoplanets). *MNRAS*, 527:4955–4982, 2023. doi: 10.1093/mnras/stad3413.

[179] Z. Rustamkulov, D. K. Sing, S. Mukherjee, E. M. May, J. Kirk, E. Schlawin, M. R. Line, C. Piaulet, A. L. Carter, N. E. Batalha, J. M. Goyal, M. López-Morales, J. D. Lothringer, R. J. MacDonald, S. E. Moran, K. B. Stevenson, H. R. Wakeford, N. Espinoza, J. L. Bean, N. M. Batalha, B. Benneke, Z. K. Berta-Thompson, I. J. M. Crossfield, P. Gao, L. Kreidberg, D. K. Powell, P. E. Cubillos, N. P. Gibson, J. Leconte, K. Molaverdikhani, N. K. Nikolov, V. Parmentier, P. Roy, J. Taylor, J. D. Turner, P. J. Wheatley, K. Aggarwal, E. Ahrer, M. K. Alam, L. Alderson, N. H. Allen, A. Banerjee, S. Barat, D. Barrado, J. K. Barstow, T. J. Bell, J. Blecic, J. Brande, S. Casewell, Q. Changeat, K. L. Chubb, N. Crouzet, T. Daylan, L. Decin, J. Désert, T. Mikal-Evans, A. D. Feinstein, L. Flagg, J. J. Fortney, J. Harrington, K. Heng, Y. Hong, R. Hu, N. Iro, T. Kataria, E. M.-R. Kempton, J. Krick, M. Lendl, J. Lillo-Box, A. Louca, J. Lustig-Yaeger, L. Mancini, M. Mansfield, N. J. Mayne, Y. Miguel, G. Morello, K. Ohno, E. Palle, D. J. M. Petit dit de la Roche, B. V. Rackham, M. Radica, L. Ramos-Rosado, S. Redfield, L. K. Rogers, E. L. Shkolnik, J. Southworth, J. Teske, P. Tremblin, G. S. Tucker, O. Venot,

W. C. Waalkes, L. Welbanks, X. Zhang, and S. Zieba. Early Release Science of the exoplanet WASP-39b with JWST NIRSpec PRISM. *Nature*, 614: 659–663, 2023. doi: 10.1038/s41586-022-05677-y.

[180] Amy J. Louca, Yamila Miguel, and Daria Kubyshkina. Metallicity and Spectral Evolution of WASP 39b: The Limited Role of Hydrodynamic Escape. *ApJL*, 956:L19, 2023. doi: 10.3847/2041-8213/acfaec.

[181] Agnibha Banerjee, Joanna K Barstow, Carole A Haswell, and Stephen R Lewis. Effect of centrifugal force on transmission spectroscopy of exoplanet atmospheres. *MNRAS*, 523:L64–L68, 2023. doi: 10.1093/mnrasl/slad058.

[182] Ian J. M. Crossfield. Volatile-to-sulfur Ratios Can Recover a Gas Giant's Accretion History. *ApJL*, 952:L18, 2023. doi: 10.3847/2041-8213/ace35f.

[183] Savvas Constantinou, Nikku Madhusudhan, and Siddharth Gandhi. Early Insights for Atmospheric Retrievals of Exoplanets Using JWST Transit Spectroscopy. *ApJL*, 943:L10, 2023. doi: 10.3847/2041-8213/acaead.

[184] Adina D. Feinstein, Michael Radica, Luis Welbanks, Catriona Anne Murray, Kazumasa Ohno, Louis-Philippe Coulombe, Néstor Espinoza, Jacob L. Bean, Johanna K. Teske, Björn Benneke, Michael R. Line, Zafar Rustamkulov, Arianna Saba, Angelos Tsiaras, Joanna K. Barstow, Jonathan J. Fortney, Peter Gao, Heather A. Knutson, Ryan J. MacDonald, Thomas Mikal-Evans, Benjamin V. Rackham, Jake Taylor, Vivien Parmentier, Natalie M. Batalha, Zachory K. Berta-Thompson, Aarynn L. Carter, Quentin Changeat, Leonardo A. dos Santos, Neale P. Gibson, Jayesh M. Goyal, Laura Kreidberg, Mercedes López-Morales, Joshua D. Lothringer, Yamila Miguel, Karan Molaverdikhani, Sarah E. Moran, Giuseppe Morello, Sagnick Mukherjee, David K. Sing, Kevin B. Stevenson, Hannah R. Wakeford, Eva-Maria Ahrer, Munazza K. Alam, Lili Alderson, Natalie H. Allen, Natasha E. Batalha, Taylor J. Bell, Jasmina Blecic, Jonathan Brande, Claudio Caceres, S. L. Casewell, Katy L. Chubb, Ian J. M. Crossfield, Nicolas

Crouzet, Patricio E. Cubillos, Leen Decin, Jean-Michel Désert, Joseph Harrington, Kevin Heng, Thomas Henning, Nicolas Iro, Eliza M.-R. Kempton, Sarah Kendrew, James Kirk, Jessica Krick, Pierre-Olivier Lagage, Monika Lendl, Luigi Mancini, Megan Mansfield, E. M. May, N. J. Mayne, Nikolay K. Nikolov, Enric Palle, Dominique J. M. Petit dit de la Roche, Caroline Piaulet, Diana Powell, Seth Redfield, Laura K. Rogers, Michael T. Roman, Pierre-Alexis Roy, Matthew C. Nixon, Everett Schlawin, Xianyu Tan, P. Tremblin, Jake D. Turner, Olivia Venot, William C. Waalkes, Peter J. Wheatley, and Xi Zhang. Early Release Science of the exoplanet WASP-39b with JWST NIRISS. *Nature*, 614:670–675, 2023. doi: 10.1038/s41586-022-05674-1.

[185] David Grant, Nikole K. Lewis, Hannah R. Wakeford, Natasha E. Batalha, Ana Glidden, Jayesh Goyal, Elijah Mullens, Ryan J. MacDonald, Erin M. May, Sara Seager, Kevin B. Stevenson, Jeff A. Valenti, Channon Visscher, Lili Alderson, Natalie H. Allen, Caleb I. Ca nas, Knicole Colón, Mark Clampin, Néstor Espinoza, Amélie Gressier, Jingcheng Huang, Zifan Lin, Douglas Long, Dana R. Louie, Maria Pe na Guerrero, Sukrit Ranjan, Kristin S. Sotzen, Daniel Valentine, Jay Anderson, William O. Balmer, Andrea Bellini, Kielan K. W. Hoch, Jens Kammerer, Mattia Libralato, C. Matt Mountain, Marshall D. Perrin, Laurent Pueyo, Emily Rickman, Isabel Rebollido, Sangmo Tony Sohn, Roeland P. van der Marel, and Laura L. Watkins. JWST-TST DREAMS: Quartz Clouds in the Atmosphere of WASP-17b. *ApJL*, 956:L32, 2023. doi: 10.3847/2041-8213/acfc3b.

[186] Siddharth Gandhi, Sam de Regt, Ignas Snellen, Yapeng Zhang, Benson Rugers, Niels van Leur, and Quincy Bosschaart. JWST Measurements of ^{13}C, ^{18}O, and ^{17}O in the Atmosphere of Super-Jupiter VHS 1256 b. *ApJL*, 957:L36, 2023. doi: 10.3847/2041-8213/ad07e2.

[187] Juan C Zapata Trujillo, Maria M Pettyjohn, and Laura K McKemmish. High-throughput quantum chemistry: empowering the search for molecular candi-

dates behind unknown spectral signatures in exoplanetary atmospheres. *MNRAS*, 524:361–376, 2023. doi: 10.1093/mnras/stad1717.

[188] Luis Welbanks, Peter McGill, Michael Line, and Nikku Madhusudhan. On the Application of Bayesian Leave-one-out Cross-validation to Exoplanet Atmospheric Analysis. *AJ.*, 165:112, 2023. doi: 10.3847/1538-3881/acab67.

[189] Jake Taylor and Vivien Parmentier. Another look at the dayside spectra of WASP-43b and HD 209458b: Are there scattering clouds? *MNRAS*, 526: 2133–2140, 2023. doi: 10.1093/mnras/stad2287.

[190] Yi-Xian Chen, Adam Burrows, Ankan Sur, and Roberto Tejada Arevalo. Jupiter atmospheric models and outer boundary conditions for giant planet evolutionary calculations. *ApJ*, 957:36, 2023. doi: 10.3847/1538-4357/acf456.

[191] J. Hayoz, G. Cugno, S. P. Quanz, P. Patapis, E. Alei, M. J. Bonse, F. A. Dannert, E. O. Garvin, T. D. Gebhard, B. S. Konrad, and L. F. Sartori. Incorporating medium-resolution spectroscopy of close-in directly imaged exoplanets into atmospheric retrievals via cross-correlation. *A&A*, 678:22, 2023. doi: 10.1051/0004-6361/202245752.

[192] Jingxuan Yang, Patrick G J Irwin, and Joanna K Barstow. Testing 2D temperature models in Bayesian retrievals of atmospheric properties from hot Jupiter phase curves. *MNRAS*, 525:5146–5167, 2023. doi: 10.1093/mnras/stad2555.

[193] Connor J. Cheverall, Nikku Madhusudhan, and Mans Holmberg. Robustness measures for molecular detections using high-resolution transmission spectroscopy of exoplanets. *MNRAS*, 522:661–677, 2023. doi: 10.1093/mnras/stad648.

[194] Eva-Maria Ahrer, Kevin B. Stevenson, Megan Mansfield, Sarah E. Moran, Jonathan Brande, Giuseppe Morello, Catriona A. Murray, Nikolay K.

Nikolov, Dominique J. M. Petit Dit de la Roche, Everett Schlawin, Peter J. Wheatley, Sebastian Zieba, Natasha E. Batalha, Mario Damiano, Jayesh M. Goyal, Monika Lendl, Joshua D. Lothringer, Sagnick Mukherjee, Kazumasa Ohno, Natalie M. Batalha, Matthew P. Battley, Jacob L. Bean, Thomas G. Beatty, Bjorn Benneke, Zachory K. Berta-Thompson, Aarynn L. Carter, Patricio E. Cubillos, Tansu Daylan, Nestor Espinoza, Peter Gao, Neale P. Gibson, Samuel Gill, Joseph Harrington, Renyu Hu, Laura Kreidberg, Nikole K. Lewis, Michael R. Line, Mercedes Lopez-Morales, Vivien Parmentier, Diana K. Powell, David K. Sing, Shang-Min Tsai, Hannah R. Wakeford, Luis Welbanks, Munazza K. Alam, Lili Alderson, Natalie H. Allen, David R. Anderson, Joanna K. Barstow, Daniel Bayliss, Taylor J. Bell, Jasmina Blecic, Edward M. Bryant, Matthew R. Burleigh, Ludmila Carone, S. L. Casewell, Quentin Changeat, Katy L. Chubb, Ian J. M. Crossfield, Nicolas Crouzet, Leen Decin, Jean-Michel Desert, Adina D. Feinstein, Laura Flagg, Jonathan J. Fortney, John E. Gizis, Kevin Heng, Nicolas Iro, Eliza M. R. Kempton, Sarah Kendrew, James Kirk, Heather A. Knutson, Thaddeus D. Komacek, Pierre-Olivier Lagage, Jeremy Leconte, Jacob Lustig-Yaeger, Ryan J. MacDonald, Luigi Mancini, E. M. May, N. J. Mayne, Yamila Miguel, Thomas Mikal-Evans, Karan Molaverdikhani, Enric Palle, Caroline Piaulet, Benjamin V. Rackham, Seth Redfield, Laura K. Rogers, Pierre-Alexis Roy, Zafar Rustamkulov, Evgenya L. Shkolnik, Kristin S. Sotzen, Jake Taylor, P. Tremblin, Gregory S. Tucker, Jake D. Turner, Miguel de Val-Borro, Olivia Venot, and Xi Zhang. Early Release Science of the exoplanet WASP-39b with JWST NIRCam. *Nature*, 614:653–658, 2023. doi: 10.1038/s41586-022-05590-4.

[195] F. Lesjak, L. Nortmann, F. Yan, D. Cont, A. Reiners, N. Piskunov, A. Hatzes, L. Boldt-Christmas, S. Czesla, U. Heiter, O. Kochukhov, A. Lavail, E. Nagel, A. D. Rains, M. Rengel, F. Rodler, U. Seemann, and D. Shulyak. Retrieval of the dayside atmosphere of WASP-43b with CRIRES$^+$. *A&A*, 678:A23, 2023. doi: 10.1051/0004-6361/202347151.

[196] Billy Edwards and Quentin Changeat. Measuring Tracers of Planet Formation in the Atmosphere of WASP-77A b: Substellar O/H and C/H Ratios, with a Stellar C/O Ratio and a Potentially Superstellar Ti/H Ratio. *ApJL*, 962, 2024. doi: 10.3847/2041-8213/ad2000.

[197] Peter C. B. Smith, Michael R. Line, Jacob L. Bean, Matteo Brogi, Prune August, Luis Welbanks, Jean-Michel Desert, Jonathan Lunine, Jorge Sanchez, Megan Mansfield, Lorenzo Pino, Emily Rauscher, Eliza Kempton, Joseph Zalesky, and Martin Fowler. A Combined Ground-based and JWST Atmospheric Retrieval Analysis: Both IGRINS and NIRSpec Agree that the Atmosphere of WASP-77A b Is Metal-poor. *AJ.*, 167:110, 2024. doi: 10.3847/1538-3881/ad17bf.

[198] Aaron David Schneider, Paul Mollière, Gilles Louppe, Ludmila Carone, Uffe Gråe Jørgensen, Leen Decin, and Christiane Helling. Harnessing machine learning for accurate treatment of overlapping opacity species in general circulation models. *A&A*, 682:A79, 2024. doi: 10.1051/0004-6361/202348221.

[199] Taylor J. Bell, Nicolas Crouzet, Patricio E. Cubillos, Laura Kreidberg, Anjali A. A. Piette, Michael T. Roman, Joanna K. Barstow, Jasmina Blecic, Ludmila Carone, Louis-Philippe Coulombe, Elsa Ducrot, Mark Hammond, Jo ao M. Mendonça, Julianne I. Moses, Vivien Parmentier, Kevin B. Stevenson, Lucas Teinturier, Michael Zhang, Natalie M. Batalha, Jacob L. Bean, Björn Benneke, Benjamin Charnay, Katy L. Chubb, Brice-Olivier Demory, Peter Gao, Elspeth K. H. Lee, Mercedes López-Morales, Giuseppe Morello, Emily Rauscher, David K. Sing, Xianyu Tan, Olivia Venot, Hannah R. Wakeford, Keshav Aggarwal, Eva-Maria Ahrer, Munazza K. Alam, Robin Baeyens, David Barrado, Claudio Caceres, Aarynn L. Carter, Sarah L. Casewell, Ryan C. Challener, Ian J. M. Crossfield, Leen Decin, Jean-Michel Désert, Ian Dobbs-Dixon, Achr'ene Dyrek, Néstor Espinoza, Adina D. Feinstein, Neale P. Gibson, Joseph Harrington, Christiane Helling, Renyu

Hu, Nicolas Iro, Eliza M. R. Kempton, Sarah Kendrew, Thaddeus D. Komacek, Jessica Krick, Pierre-Olivier Lagage, Jérémy Leconte, Monika Lendl, Neil T. Lewis, Joshua D. Lothringer, Isaac Malsky, Luigi Mancini, Megan Mansfield, Nathan J. Mayne, Thomas Mikal-Evans, Karan Molaverdikhani, Nikolay K. Nikolov, Matthew C. Nixon, Enric Palle, Dominique J. M. Petit dit de la Roche, Caroline Piaulet, Diana Powell, Benjamin V. Rackham, Aaron D. Schneider, Maria E. Steinrueck, Jake Taylor, Luis Welbanks, Sergei N. Yurchenko, Xi Zhang, and Sebastian Zieba. Nightside clouds and disequilibrium chemistry on the hot Jupiter WASP-43b. *arXiv e-prints*, 2024.

[200] Diana Powell, Adina D. Feinstein, Elspeth K. H. Lee, Michael Zhang, Shang-Min Tsai, Jake Taylor, James Kirk, Taylor Bell, Joanna K. Barstow, Peter Gao, Jacob L. Bean, Jasmina Blecic, Katy L. Chubb, Ian J. M. Crossfield, Sean Jordan, Daniel Kitzmann, Sarah E. Moran, Giuseppe Morello, Julianne I. Moses, Luis Welbanks, Jeehyun Yang, Xi Zhang, Eva-Maria Ahrer, Aaron Bello-Arufe, Jonathan Brande, S. L. Casewell, Nicolas Crouzet, Patricio E. Cubillos, Brice-Olivier Demory, Achrène Dyrek, Laura Flagg, Renyu Hu, Julie Inglis, Kathryn D. Jones, Laura Kreidberg, Mercedes López-Morales, Pierre-Olivier Lagage, Erik A. Meier Valdés, Yamila Miguel, Vivien Parmentier, Anjali A. A. Piette, Benjamin V. Rackham, Michael Radica, Seth Redfield, Kevin B. Stevenson, Hannah R. Wakeford, Keshav Aggarwal, Munazza K. Alam, Natalie M. Batalha, Natasha E. Batalha, Björn Benneke, Zach K. Berta-Thompson, Ryan P. Brady, Claudio Caceres, Aarynn L. Carter, Jean-Michel Désert, Joseph Harrington, Nicolas Iro, Michael R. Line, Joshua D. Lothringer, Ryan J. MacDonald, Luigi Mancini, Karan Molaverdikhani, Sagnick Mukherjee, Matthew C. Nixon, Apurva V. Oza, Enric Palle, Zafar Rustamkulov, David K. Sing, Maria E. Steinrueck, Olivia Venot, Peter J. Wheatley, and Sergei N. Yurchenko. Sulphur dioxide in the mid-infrared transmission spectrum of WASP-39b. *Nature*, 626:979–983, 2024. doi: 10.1038/s41586-024-07040-9.

[201] Kai Hou Yip, Quentin Changeat, Ahmed Al-Refaie, and Ingo P. Wald-

mann. To Sample or Not to Sample: Retrieving Exoplanetary Spectra with Variational Inference and Normalizing Flows. *ApJ*, 961:30, 2024. doi: 10.3847/1538-4357/ad063f.

[202] R. Landman, T. Stolker, I. A. G. Snellen, J. Costes, S. de Regt, Y. Zhang, S. Gandhi, P. Molliere, A. Kesseli, A. Vigan, and A. Sanchez-Lopez. β pictoris b through the eyes of the upgraded crires plus atmospheric composition, spin rotation, and radial velocity. *A&A*, 682:A48, 2024. doi: 10.1051/0004-6361/202347846.

[203] Jamie Wilson, Neale P Gibson, Nikolay Nikolov, Savvas Constantinou, Nikku Madhusudhan, Jayesh Goyal, Joanna K Barstow, Aarynn L Carter, Ernst J W de Mooij, Benjamin Drummond, Thomas Mikal-Evans, Christiane Helling, Nathan J Mayne, and David K Sing. Ground-based transmission spectroscopy with FORS2: A featureless optical transmission spectrum and detection of H_2O for the ultra-hot Jupiter WASP-103b. *MNRAS*, 497: 5155–5170, 2020. doi: 10.1093/mnras/staa2307.

[204] Siddharth Gandhi and Adam S Jermyn. Coupled day-night models of exoplanetary atmospheres. *MNRAS*, 499:4984–5003, 2020. doi: 10.1093/mnras/staa3143.

[205] H. M. Tabernero, M. R. Zapatero Osorio, R. Allart, F. Borsa, N. Casasayas-Barris, O. Demangeon, D. Ehrenreich, J. Lillo-Box, C. Lovis, E. Palle, S. G. Sousa, R. Rebolo, N. C. Santos, F. Pepe, S. Cristiani, V Adibekyan, C. Allende Prieto, Yann Alibert, S. C. C. Barros, F. Bouchy, V Bourrier, V D'Odorico, X. Dumusque, J. P. Faria, P. Figueira, R. Genova Santos, J. Gonzalez Hernandez, I, S. Hojjatpanah, G. Lo Curto, B. Lavie, C. J. A. P. Martins, J. H. C. Martins, A. Mehner, G. Micela, P. Molaro, N. J. Nunes, E. Poretti, J. Seidel, V, A. Sozzetti, A. Suarez Mascareno, S. Udry, M. Aliverti, M. Affolter, D. Alves, M. Amate, G. Avila, T. Bandy, W. Benz, A. Bianco, C. Broeg, A. Cabral, P. Conconi, J. Coelho, C. Cumani, S. Deiries, H. Dekker, B. Delabre, A. Fragoso, M. Genoni, L. Genolet, I Hughes,

J. Knudstrup, F. Kerber, M. Landoni, J. L. Lizon, C. Maire, A. Manescau, P. Di Marcantonio, D. Megevand, M. Monteiro, M. Moschetti, E. Mueller, A. Modigliani, L. Oggioni, A. Oliveira, G. Pariani, L. Pasquini, J. L. Rasilla, E. Redaelli, M. Riva, S. Santana-Tschudi, P. Santin, P. Santos, A. Segovia, D. Sosnowska, P. Spano, F. Tenegi, O. Iwert, A. Zanutta, and F. Zerbi. ESPRESSO high-resolution transmission spectroscopy of WASP-76 b. *A&A*, 646, 2021. doi: 10.1051/0004-6361/202039511.

[206] Erin M. May, Thaddeus D. Komacek, Kevin B. Stevenson, Eliza M.-R. Kempton, Jacob L. Bean, Matej Malik, Jegug Ih, Megan Mansfield, Arjun B. Savel, Drake Deming, Jean-Michel Desert, Y. Katherina Feng, Jonathan J. Fortney, Tiffany Kataria, Nikole Lewis, Caroline Morley, Emily Rauscher, and Adam Showman. Spitzer Phase-curve Observations and Circulation Models of the Inflated Ultrahot Jupiter WASP-76b. *AJ.*, 162:158, 2021. doi: 10.3847/1538-3881/ac0e30.

[207] Ehsan Gharib-Nezhad, Aishwarya R. Iyer, Michael R. Line, Richard S. Freedman, Mark S. Marley, and Natasha E. Batalha. EXOPLINES: Molecular Absorption Cross-section Database for Brown Dwarf and Giant Exoplanet Atmospheres. *ApJS*, 254:34, 2021. doi: 10.3847/1538-4365/abf504.

[208] M. Stangret, E. Pallé, N. Casasayas-Barris, M. Oshagh, A. Bello-Arufe, R. Luque, V. Nascimbeni, F. Yan, J. Orell-Miquel, D. Sicilia, L. Malavolta, B. C. Addison, L. A. Buchhave, A. S. Bonomo, F. Borsa, S. H. C. Cabot, M. Cecconi, D. A. Fischer, A. Harutyunyan, J. M. Mendonça, G. Nowak, H. Parviainen, A. Sozzetti, and R. Tronsgaard. The obliquity and atmosphere of the ultra-hot Jupiter TOI-1431b (MASCARA-5b): A misaligned orbit and no signs of atomic or molecular absorptions. *A&A*, 654, 2021. doi: 10.1051/0004-6361/202040100.

[209] David Kasper, Jacob L. Bean, Michael R. Line, Andreas Seifahrt, Julian Stürmer, Lorenzo Pino, Jean-Michel Désert, and Matteo Brogi. Confirmation of Iron Emission Lines and Nondetection of TiO on the Dayside of

KELT-9b with MAROON-X. *ApJL*, 921(1):L18, 2021. doi: 10.3847/2041-8213/ac30e1.

[210] A. Roth, B. Drummond, E. Hebrard, P. Tremblin, J. Goyal, and N. Mayne. Pseudo-2D modelling of heat redistribution through H_2 thermal dissociation/recombination: consequences for ultra-hot Jupiters. *MNRAS*, 505:4515–4530, 2021. doi: 10.1093/mnras/stab1256.

[211] Dilovan B. Serindag, Stevanus K. Nugroho, Paul Mollière, Ernst J. W. de Mooij, Neale P. Gibson, and Ignas A. G. Snellen. Is TiO emission present in the ultra-hot Jupiter WASP-33b? A reassessment using the improved ExoMol TOTO line list. *A&A*, 645:A90, 2021. doi: 10.1051/0004-6361/202039135.

[212] James Kirk, Benjamin V. Rackham, Ryan J. MacDonald, Mercedes López-Morales, Néstor Espinoza, Monika Lendl, Jamie Wilson, David J. Osip, Peter J. Wheatley, Ian Skillen, Dániel Apai, Alex Bixel, Neale P. Gibson, Andrés Jordán, Nikole K. Lewis, Tom Louden, Chima D. McGruder, Nikolay Nikolov, Florian Rodler, and Ian C. Weaver. ACCESS and LRG-BEASTS: A Precise New Optical Transmission Spectrum of the Ultrahot Jupiter WASP-103b. *AJ.*, 162:34, 2021. doi: 10.3847/1538-3881/abfcd2.

[213] Quentin Changeat and Billy Edwards. The hubble WFC3 emission spectrum of the extremely hot jupiter KELT-9b. *ApJL*, 907:L22, 2021. doi: 10.3847/2041-8213/abd84f.

[214] Elspeth K. H. Lee, Bibiana Prinoth, Daniel Kitzmann, Shang-Min Tsai, Jens Hoeijmakers, Nicholas W. Borsato, and Kevin Heng. The Mantis Network II: examining the 3D high-resolution observable properties of the UHJs WASP-121b and WASP-189b through GCM modelling. *MNRAS*, 517:240–256, 2022. doi: 10.1093/mnras/stac2246.

[215] Bibiana Prinoth, H. Jens Hoeijmakers, Daniel Kitzmann, Elin Sandvik, Julia Seidel, V, Monika Lendl, Nicholas W. Borsato, Brian Thorsbro, David R.

Anderson, David Barrado, Kateryna Kravchenko, Romain Allart, Vincent Bourrier, Heather M. Cegla, David Ehrenreich, Chloe Fisher, Christophe Lovis, Andrea Guzman-Mesa, Simon Grimm, Matthew Hooton, Brett M. Morris, Maria Oreshenko, Lorenzo Pino, and Kevin Heng. Titanium oxide and chemical inhomogeneity in the atmosphere of the exoplanet WASP-189 b. *Nature Astr.*, 6:449–457, 2022. doi: 10.1038/s41550-021-01581-z.

[216] Thomas Mikal-Evans, David K. Sing, Joanna K. Barstow, Tiffany Kataria, Jayesh Goyal, Nikole Lewis, Jake Taylor, Nathan J. Mayne, Tansu Daylan, Hannah R. Wakeford, Mark S. Marley, and Jessica J. Spake. Diurnal variations in the stratosphere of the ultrahot giant exoplanet WASP-121b. *Nat. Astron.*, 6:471 – 479, 2022. doi: 10.1038/s41550-021-01592-w.

[217] Arjun B. Savel, Eliza M.-R. Kempton, Matej Malik, Thaddeus D. Komacek, Jacob L. Bean, Erin M. May, Kevin B. Stevenson, Megan Mansfield, and Emily Rauscher. No Umbrella Needed: Confronting the Hypothesis of Iron Rain on WASP-76b with Post-processed General Circulation Models. *AJ.*, 926:85, 2022. doi: 10.3847/1538-4357/ac423f.

[218] Aaron Bello-Arufe, Samuel H. C. Cabot, Joao M. Mendonca, Lars A. Buchhave, and Alexander D. Rathcke. Mining the Ultrahot Skies of HAT-P-70b: Detection of a Profusion of Neutral and Ionized Species. *AJ.*, 163:96, 2022. doi: 10.3847/1538-3881/ac402e.

[219] Quentin Changeat. On Spectroscopic Phase-curve Retrievals: H2 Dissociation and Thermal Inversion in the Atmosphere of the Ultrahot Jupiter WASP-103 b. *AJ.*, 163:106, 2022. doi: 10.3847/1538-3881/ac4475.

[220] S. de Regt, A. Y. Kesseli, I. A. G. Snellen, S. R. Merritt, and K. L. Chubb. A quantitative assessment of the VO line list: Inaccuracies hamper high-resolution VO detections in exoplanet atmospheres. *A&A*, 661:A109, 2022. doi: 10.1051/0004-6361/202142683.

[221] Måns Holmberg and Nikku Madhusudhan. A First Look at CRIRES+: Performance Assessment and Exoplanet Spectroscopy. *AJ.*, 164:79, 2022. doi: 10.3847/1538-3881/ac77eb.

[222] Luke Finnerty, Tobias Schofield, Ben Sappey, Jerry W. Xuan, Jean-Baptiste Ruffio, Jason J. Wang, Jacques-Robert Delorme, Geoffrey A. Blake, Cam Buzard, Michael P. Fitzgerald, Ashley Baker, Randall Bartos, Charlotte Z. Bond, Benjamin Calvin, Sylvain Cetre, Greg Doppmann, Daniel Echeverri, Nemanja Jovanovic, Joshua Liberman, Ronald A. López, Emily C. Martin, Dimitri Mawet, Evan Morris, Jacklyn Pezzato, Caprice L. Phillips, Sam Ragland, Andrew Skemer, Taylor Venenciano, J. Kent Wallace, Nicole L. Wallack, Ji Wang, and Peter Wizinowich. Keck Planet Imager and Characterizer Emission Spectroscopy of WASP-33b. *AJ.*, 166:31, 2023. doi: 10.3847/1538-3881/acda91.

[223] Yuanheng Yang, Guo Chen, Songhu Wang, and Fei Yan. High-resolution Transmission Spectroscopy of Ultrahot Jupiter WASP-33b with NEID. *AJ.*, 167:36, 2023. doi: 10.3847/1538-3881/ad10a3.

[224] Nataliea Lowson, George Zhou, Duncan J. J. Wright, Chelsea X. X. Huang, Joao M. Mendonca, Samuel H. C. Cabot, Christa Pudmenzky, Robert A. A. Wittenmyer, David W. W. Latham, Allyson Bieryla, Gilbert A. A. Esquerdo, Perry Berlind, and Michael L. L. Calkins. Multiepoch Detections of the Extended Atmosphere and Transmission Spectra of KELT-9b with a 1.5 m Telescope. *AJ.*, 165, 2023. doi: 10.3847/1538-3881/acacf3.

[225] Joost P. Wardenier, Vivien Parmentier, Michael R. Line, and Elspeth K. H. Lee. Modelling the effect of 3D temperature and chemistry on the cross-correlation signal of transiting ultra-hot Jupiters: a study of five chemical species on WASP-76b. *MNRAS*, 525:4942–4961, 2023. doi: 10.1093/mnras/stad2586.

[226] Louis-Philippe Coulombe, Bjorn Benneke, Ryan Challener, Anjali A. A.

Piette, Lindsey S. Wiser, Megan Mansfield, Ryan J. MacDonald, Hayley Beltz, Adina D. Feinstein, Michael Radica, Arjun B. Savel, Leonardo A. Dos Santos, Jacob L. Bean, Vivien Parmentier, Ian Wong, Emily Rauscher, Thaddeus D. Komacek, Eliza M. R. Kempton, Xianyu Tan, Mark Hammond, Neil T. Lewis, Michael R. Line, Elspeth K. H. Lee, Hinna Shivkumar, Ian J. M. Crossfield, Matthew C. Nixon, Benjamin V. Rackham, Hannah R. Wakeford, Luis Welbanks, Xi Zhang, Natalie M. Batalha, Zachory K. Berta-Thompson, Quentin Changeat, Jean-Michel Desert, Nestor Espinoza, Jayesh M. Goyal, Joseph Harrington, Heather A. Knutson, Laura Kreidberg, Mercedes Lopez-Morales, Avi Shporer, David K. Sing, Kevin B. Stevenson, Keshav Aggarwal, Eva-Maria Ahrer, Munazza K. Alam, Taylor J. Bell, Jasmina Blecic, Claudio Caceres, Aarynn L. Carter, Sarah L. Casewell, Nicolas Crouzet, Patricio E. Cubillos, Leen Decin, Jonathan J. Fortney, Neale P. Gibson, Kevin Heng, Thomas Henning, Nicolas Iro, Sarah Kendrew, Pierre-Olivier Lagage, Jeremy Leconte, Monika Lendl, Joshua D. Lothringer, Luigi Mancini, Thomas Mikal-Evans, Karan Molaverdikhani, Nikolay K. Nikolov, Kazumasa Ohno, Enric Palle, Caroline Piaulet, Seth Redfield, Pierre-Alexis Roy, Shang-Min Tsai, Olivia Venot, and Peter J. Wheatley. A broadband thermal emission spectrum of the ultra-hot Jupiter WASP-18b. *Nature*, 620: 292–298, 2023. doi: 10.1038/s41586-023-06230-1.

[227] Marshall C. Johnson, Ji Wang, Anusha Pai Asnodkar, Aldo S. Bonomo, B. Scott Gaudi, Thomas Henning, Ilya Ilyin, Engin Keles, Luca Malavolta, Matthias Mallonn, Karan Molaverdikhani, Valerio Nascimbeni, Jennifer Patience, Katja Poppenhaeger, Gaetano Scandariato, Everett Schlawin, Evgenya Shkolnik, Daniela Sicilia, Alessandro Sozzetti, Klaus G. Strassmeier, Christian Veillet, and Fei Yan. The PEPSI Exoplanet Transit Survey (PETS). II. A Deep Search for Thermal Inversion Agents in KELT-20 b/MASCARA-2 b with Emission and Transmission Spectroscopy. *AJ.*, 165, 2023. doi: 10.3847/1538-3881/acb7e2.

[228] Sydney Petz, Marshall C Johnson, Anusha Pai Asnodkar, Ji Wang, B Scott

Gaudi, Thomas Henning, Engin Keles, Karan Molaverdikhani, Katja Poppenhaeger, Gaetano Scandariato, Evgenya K Shkolnik, Daniela Sicilia, Klaus G Strassmeier, and Fei Yan. The PEPSI Exoplanet Transit Survey (PETS) - IV. Assessing the atmospheric chemistry of KELT-20b. *MNRAS*, 527:7079–7092, 2023. doi: 10.1093/mnras/stad3481.

[229] D. Kitzmann, H. J. Hoeijmakers, S. L. Grimm, N. W. Borsato, A. Lueber, and B. Prinoth. The Mantis network I. A standard grid of templates and masks for cross-correlation analyses of ultra-hot Jupiter transmission spectra. *A&A*, 669:A113, 2023. doi: 10.1051/0004-6361/202142969.

[230] Stefan Pelletier, Björn Benneke, Mohamad Ali-Dib, Bibiana Prinoth, David Kasper, Andreas Seifahrt, Jacob L. Bean, Florian Debras, Baptiste Klein, Luc Bazinet, H. Jens Hoeijmakers, Aurora Y. Kesseli, Olivia Lim, Andres Carmona, Lorenzo Pino, Núria Casasayas-Barris, Thea Hood, and Julian Stürmer. Vanadium oxide and a sharp onset of cold-trapping on a giant exoplanet. *Nature*, 619:491–494, 2023. doi: 10.1038/s41586-023-06134-0.

[231] Yaqing Shi, Wei Wang, Gang Zhao, Meng Zhai, Guo Chen, Zewen Jiang, Qinglin Ouyang, Thomas Henning, Jingkun Zhao, Nicolas Crouzet, Roy van Boekel, and Lisa Esteves. Thermal emission from the hot Jupiter WASP-103 b in J and Ks bands. *MNRAS*, 522:1491–1503, 2023. doi: 10.1093/mnras/stad891.

[232] Andrew Ridden-Harper, Ernst de Mooij, Ray Jayawardhana, Neale Gibson, Raine Karjalainen, and Marie Karjalainen. High-resolution Emission Spectroscopy of the Ultrahot Jupiter KELT-9b: Little Variation in Day- and Nightside Emission Line Contrasts. *AJ.*, 165(5):211, 2023. doi: 10.3847/1538-3881/acc654.

[233] Siddharth Gandhi, Aurora Kesseli, Yapeng Zhang, Amy Louca, Ignas Snellen, Matteo Brogi, Yamila Miguel, Núria Casasayas-Barris, Stefan Pelletier, Rico Landman, Cathal Maguire, and Neale P. Gibson. Retrieval Survey

of Metals in Six Ultrahot Jupiters: Trends in Chemistry, Rain-out, Ionization, and Atmospheric Dynamics. *AJ.*, 165(6):242, 2023. doi: 10.3847/1538-3881/accd65.

[234] Matteo Brogi, Vanessa Emeka-Okafor, Michael R. Line, Siddharth Gandhi, Lorenzo Pino, Eliza M.-R. Kempton, Emily Rauscher, Vivien Parmentier, Jacob L. Bean, Gregory N. Mace, Nicolas B. Cowan, Evgenya Shkolnik, Joost P. Wardenier, Megan Mansfield, Luis Welbanks, Peter Smith, Jonathan J. Fortney, Jayne L. Birkby, Joseph A. Zalesky, Lisa Dang, Jennifer Patience, and Jean-Michel Désert. The Roasting Marshmallows Program with IGRINS on Gemini South I: Composition and Climate of the Ultrahot Jupiter WASP-18 b. *AJ.*, 165:91, 2023. doi: 10.3847/1538-3881/acaf5c.

[235] B. Prinoth, H. J. Hoeijmakers, S. Pelletier, D. Kitzmann, B. M. Morris, A. Seifahrt, D. Kasper, H. H. Korhonen, M. Burheim, J. L. Bean, B. Benneke, N. W. Borsato, M. Brady, S. L. Grimm, R. Luque, J. Stuermer, and B. Thorsbro. Time-resolved transmission spectroscopy of the ultra-hot Jupiter WASP-189 b. *A&A*, 678, 2023. doi: 10.1051/0004-6361/202347262.

[236] Lennart van Sluijs, Jayne L Birkby, Joshua Lothringer, Elspeth K H Lee, Ian J M Crossfield, Vivien Parmentier, Matteo Brogi, Craig Kulesa, Don McCarthy, and David Charbonneau. Carbon monoxide emission lines reveal an inverted atmosphere in the ultra hot Jupiter WASP-33 b consistent with an eastward hot spot. *MNRAS*, 522:2145–2170, 2023. doi: 10.1093/mnras/stad1103.

[237] Q. Changeat, J. W. Skinner, J. y-k. Cho, J. Nättilä, I. P. Waldmann, A. F. Al-Refaie, A. Dyrek, B. Edwards, T. Mikal-Evans, M. Joshua, G. Morello, N. Skaf, A. Tsiaras, O. Venot, and K. H. Yip. Is the Atmosphere of the Ultrahot Jupiter WASP-121 b Variable? *ApJS*, 270, 2024. doi: 10.3847/1538-4365/ad1191.

[238] Luke Finnerty, Jerry W. Xuan, Yinzi Xin, Joshua Liberman, Tobias Schofield,

Michael P. Fitzgerald, Shubh Agrawal, Ashley Baker, Randall Bartos, Geoffrey A. Blake, Benjamin Calvin, Sylvain Cetre, Jacques-Robert Delorme, Greg Doppmann, Daniel Echeverri, Chih-Chun Hsu, Nemanja Jovanovic, Ronald A. López, Emily C. Martin, Dimitri Mawet, Evan Morris, Jacklyn Pezzato, Jean-Baptiste Ruffio, Ben Sappey, Andrew Skemer, Taylor Venenciano, J. Kent Wallace, Nicole L. Wallack, Jason J. Wang, and Ji Wang. Atmospheric Metallicity and C/O of HD 189733 b from High-resolution Spectroscopy. *AJ.*, 167:43, 2024. doi: 10.3847/1538-3881/ad1180.

[239] Cathal Maguire, Neale P. Gibson, Stevanus K. Nugroho, Mark Fortune, Swaetha Ramkumar, Siddharth Gandhi, and Ernst de Mooij. High-resolution atmospheric retrievals of WASP-76b transmission spectroscopy with ESPRESSO: Monitoring limb asymmetries across multiple transits. *arXiv e-prints*, 2024. doi: 10.48550/arXiv.2404.10463.

[240] Billy Edwards, Quentin Changeat, Mayuko Mori, Lara O. Anisman, Mario Morvan, Kai Hou Yip, Angelos Tsiaras, Ahmed Al-Refaie, Ingo Waldmann, and Giovanna Tinetti. Hubble WFC3 Spectroscopy of the Habitable-zone Super-Earth LHS 1140 b. *AJ.*, 161:44, 2020. doi: 10.3847/1538-3881/abc6a5.

[241] Anjali A A Piette, Nikku Madhusudhan, and Avi M Mandell. HyDRo: atmospheric retrieval of rocky exoplanets in thermal emission. *MNRAS*, 511: 2565–2584, 2021. doi: 10.1093/mnras/stab3612.

[242] Lorenzo V. Mugnai, Darius Modirrousta-Galian, Billy Edwards, Quentin Changeat, Jeroen Bouwman, Giuseppe Morello, Ahmed Al-Refaie, Robin Baeyens, Michelle Fabienne Bieger, Doriann Blain, Amélie Gressier, Gloria Guilluy, Yassin Jaziri, Flavien Kiefer, Mario Morvan, William Pluriel, Mathilde Poveda, Nour Skaf, Niall Whiteford, Sam Wright, Kai Hou Yip, Tiziano Zingales, Benjamin Charnay, Pierre Drossart, Jérémy Leconte, Olivia Venot, Ingo Waldmann, and Jean-Philippe Beaulieu. ARES. V. No

Evidence For Molecular Absorption in the HST WFC3 Spectrum of GJ 1132 b. *AJ.*, 161:284, 2021. doi: 10.3847/1538-3881/abf3c3.

[243] A. Gressier, M. Mori, Q. Changeat, B. Edwards, J. P. Beaulieu, E. Marcq, and B. Charnay. Near-infrared transmission spectrum of TRAPPIST-1 h using *Hubble* WFC3 G141 observations. *A&A*, 658, 2022. doi: 10.1051/0004-6361/202142140.

[244] Eleonora Alei, Bjoern S. Konrad, Daniel Angerhausen, John Lee Grenfell, Paul Molliere, Sascha P. Quanz, Sarah Rugheimer, Fabian Wunderlich, and LIFE Collaboration. Large Interferometer For Exoplanets (LIFE) V. Diagnostic potential of a mid-infrared space interferometer for studying Earth analogs. *A&A*, 665, 2022. doi: 10.1051/0004-6361/202243760.

[245] Dan J. Bower, Kaustubh Hakim, Paolo A. Sossi, and Patrick Sanan. Retention of Water in Terrestrial Magma Oceans and Carbon-rich Early Atmospheres. *Planet. Sci. J.*, 3:93, 2022. doi: 10.3847/PSJ/ac5fb1.

[246] A. B. Claringbold, P. B. Rimmer, S. Rugheimer, and O. Shorttle. Prebiosignature Molecules Can Be Detected in Temperate Exoplanet Atmospheres with JWST. *AJ.*, 166, 2023. doi: 10.3847/1538-3881/acdacc.

[247] Hannah Diamond-Lowe, Jo ao M. Mendonça, David Charbonneau, and Lars A. Buchhave. Ground-based Optical Transmission Spectroscopy of the Nearby Terrestrial Exoplanet LTT 1445Ab. *AJ.*, 165:169, 2023. doi: 10.3847/1538-3881/acbf39.

[248] Li Zhou, Bo Ma, Yong-Hao Wang, and Yi-Nan Zhu. Hubble WFC3 Spectroscopy of the Terrestrial Planets L 98-59 c and d: No Evidence for a Clear Hydrogen Dominated Primary Atmosphere. *Res. Astron. Astrophys.*, 23: 025011, 2023. doi: 10.1088/1674-4527/acaceb.

[249] Caprice L Phillips, Ji Wang, Billy Edwards, Romy Rodríguez Martínez, Anusha Pai Asnodkar, and B Scott Gaudi. Exploring the potential of Twinkle

to unveil the nature of LTT 1445 Ab. *MNRAS*, 526:2251–2264, 2023. doi: 10.1093/mnras/stad2822.

[250] Sarah E. Moran, Kevin B. Stevenson, David K. Sing, Ryan J. MacDonald, James Kirk, Jacob Lustig-Yaeger, Sarah Peacock, L. C. Mayorga, Katherine A. Bennett, Mercedes López-Morales, E. M. May, Zafar Rustamkulov, Jeff A. Valenti, Jéa I. Adams Redai, Munazza K. Alam, Natasha E. Batalha, Guangwei Fu, Junellie Gonzalez-Quiles, Alicia N. Highland, Ethan Kruse, Joshua D. Lothringer, Kevin N. Ortiz Ceballos, Kristin S. Sotzen, and Hannah R. Wakeford. High Tide or Riptide on the Cosmic Shoreline? A Water-rich Atmosphere or Stellar Contamination for the Warm Super-Earth GJ 486b from JWST Observations. *ApJL*, 948:L11, 2023. doi: 10.3847/2041-8213/accb9c.

[251] Li Zhou, Bo Ma, Yonghao Wang, and Yinan Zhu. Hubble WFC3 Spectroscopy of the Rocky Planet L 98-59 b: No Evidence for a Cloud-free Primordial Atmosphere. *AJ.*, 164, 2022. doi: 10.3847/1538-3881/ac8fe9.

[252] M. Zilinskas, C. P. A. van Buchem, Y. Miguel, A. Louca, R. Lupu, S. Zieba, and W. van Westrenen. Observability of evaporating lava worlds star. *A&A*, 661:A126, 2022. doi: 10.1051/0004-6361/202142984.

[253] Emily A. Whittaker, Matej Malik, Jegug Ih, Eliza M.-R. Kempton, Megan Mansfield, Jacob L. Bean, Edwin S. Kite, Daniel D. B. Koll, Timothy W. Cronin, and Renyu Hu. The Detectability of Rocky Planet Surface and Atmosphere Composition with the JWST: The Case of LHS 3844b. *AJ.*, 164: 258, 2022. doi: 10.3847/1538-3881/ac9ab3.

[254] Ian J. M. Crossfield, Matej Malik, Michelle L. Hill, Stephen R. Kane, Bradford Foley, Alex S. Polanski, David Coria, Jonathan Brande, Yanzhe Zhang, Katherine Wienke, Laura Kreidberg, Nicolas B. Cowan, Diana Dragomir, Varoujan Gorjian, Thomas Mikal-Evans, Björn Benneke, Jessie L. Christiansen, Drake Deming, and Farisa Y. Morales. GJ 1252b: A Hot Terrestrial

Super-Earth with No Atmosphere. *ApJL*, 937, 2022. doi: 10.3847/2041-8213/ac886b.

[255] S. Zieba, M. Zilinskas, L. Kreidberg, T. G. Nguyen, Y. Miguel, N. B. Cowan, R. Pierrehumbert, L. Carone, L. Dang, M. Hammond, T. Louden, R. Lupu, L. Malavolta, and K. B. Stevenson. K2 and *Spitzer* phase curves of the rocky ultra-short-period planet k2-141 b hint at a tenuous rock vapor atmosphere. *A&A*, 664, 2022. doi: 10.1051/0004-6361/202142912.

[256] L. J. Janssen, P. Woitke, O. Herbort, M. Min, K. L. Chubb, Ch. Helling, and L. Carone. The sulfur species in hot rocky exoplanet atmospheres. *ASTRONOMISCHE NACHRICHTEN*, 344, 2023. doi: 10.1002/asna.20230075.

[257] Anjali A. A. Piette, Peter Gao, Kara Brugman, Anat Shahar, Tim Lichtenberg, Francesca Miozzi, and Peter Driscoll. Rocky Planet or Water World? Observability of Low-density Lava World Atmospheres. *ApJ*, 954, 2023. doi: 10.3847/1538-4357/acdef2.

[258] Kevin Heng. The Transient Outgassed Atmosphere of 55 Cancri e. *ApJL*, 956:L20, 2023. doi: 10.3847/2041-8213/acfe05.

[259] Andrew Ridden-Harper, Stevanus K. Nugroho, Laura Flagg, Ray Jayawardhana, Jake D. Turner, Ernst de Mooij, Ryan MacDonald, Emily Deibert, Motohide Tamura, Takayuki Kotani, Teruyuki Hirano, Masayuki Kuzuhara, Masashi Omiya, and Nobuhiko Kusakabe. High-resolution Transmission Spectroscopy of the Terrestrial Exoplanet GJ 486b. *AJ.*, 165:170, 2023. doi: 10.3847/1538-3881/acbd39.

[260] M. Zilinskas, Y. Miguel, C. P. A. van Buchem, and I. A. G. Snellen. Observability of silicates in volatile atmospheres of super-Earths and sub-Neptunes Exploring the edge of the evaporation desert. *A&A*, 671, 2023. doi: 10.1051/0004-6361/202245521.

[261] Richard A. Booth, James E. Owen, and Matthaus Schulik. Dust formation

in the outflows of catastrophically evaporating planets. *MNRAS*, 518:1761–1775, 2023. doi: 10.1093/mnras/stac3121.

[262] Darryl Z. Seligman, Adina D. Feinstein, Dong Lai, Luis Welbanks, Aster G. Taylor, Juliette Becker, Fred C. Adams, Marvin Morgan, and Jennifer B. Bergner. Potential Melting of Extrasolar Planets by Tidal Dissipation. *ApJ*, 961:22, 2024. doi: 10.3847/1538-4357/ad0b82.

[263] Jane S Greaves, Paul B Rimmer, Anita M S Richards, Janusz J Petkowski, William Bains, Sukrit Ranjan, Sara Seager, David L Clements, Clara Sousa Silva, and Helen J Fraser. Low levels of sulphur dioxide contamination of Venusian phosphine spectra. *MNRAS*, 514(2):2994–3001, 2022. doi: 10.1093/mnras/stac1438.

[264] Joao A. Dias, Pedro Machado, and Jose Ribeiro. From Atmospheric Evolution to the Search of Species of Astrobiological Interest in the Solar System-Case Studies Using the Planetary Spectrum Generator. *Atmosphere*, 13(3), 2022. doi: 10.3390/atmos13030461.

[265] Ryoichi Koga, Tatsuya Suzuki, Fuminori Tsuchiya, Takeshi Sakanoi, and Yasuhiro Hirahara. ALMA Observation of SO2 Gas Originating from Io's Volcanic Plume and Lava Areas. *ApJ*, 907:L6, 2021. doi: 10.3847/2041-8213/abd39f.

[266] P. Cambianica, G. Cremonese, G. Munaretto, M. T. Capria, M. Fulle, W. Boschin, L. Di Fabrizio, and A. Harutyunyan. A high-spectral-resolution catalog of emission lines in the visible spectrum of comet C/2020 F3 (NEOWISE). *A&A*, 656:A160, DEC 16 2021. doi: 10.1051/0004-6361/202140309.

[267] Brianna I. Lacy and Adam Burrows. JWST Transit Spectra. II. Constraining Aerosol Species, Particle-size Distributions, Temperature, and Metallicity for Cloudy Exoplanets. *ApJ*, 904:25, 2020. doi: 10.3847/1538-4357/abbc6c.

[268] Luis Welbanks and Nikku Madhusudhan. Aurora: A Generalized Retrieval Framework for Exoplanetary Transmission Spectra. *ApJ*, page 114, 2021. doi: 10.3847/1538-4357/abee94.

[269] P. Lavvas and A. Arfaux. Impact of photochemical hazes and gases on exoplanet atmospheric thermal structure. *MNRAS*, 502:5643–5657, 2021. doi: 10.1093/mnras/stab456.

[270] Sagnick Mukherjee, Natasha E. Batalha, and Mark S. Marley. Cloud Parameterizations and their Effect on Retrievals of Exoplanet Reflection Spectroscopy. *ApJ*, 910:158, 2021. doi: 10.3847/1538-4357/abe53b.

[271] Maxence Lefèvre, Xianyu Tan, Elspeth K. H. Lee, and R. T. Pierrehumbert. Cloud-convection Feedback in Brown Dwarf Atmospheres. *ApJ*, 929:153, 2022. doi: 10.3847/1538-4357/ac5e2d.

[272] C. Jiang, G. Chen, E. Palle, F. Murgas, H. Parviainen, and Y. Ma. Featureless transmission spectra of 12 giant exoplanets observed by GTC/OSIRIS. *A&A*, 675, 2023. doi: 10.1051/0004-6361/202346091.

[273] Brianna Lacy and Adam Burrows. Self-consistent Models of Y Dwarf Atmospheres with Water Clouds and Disequilibrium Chemistry. *ApJ*, 950:8, 2023. doi: 10.3847/1538-4357/acc8cb.

[274] Sushuang Ma, Yuichi Ito, Ahmed Faris Al-Refaie, Quentin Changeat, Billy Edwards, and Giovanna Tinetti. YunMa: Enabling Spectral Retrievals of Exoplanetary Clouds. *ApJ*, 957(2):104, 2023. doi: 10.3847/1538-4357/acf8ca.

[275] Robin Baeyens, Thomas Konings, Olivia Venot, Ludmila Carone, and Leen Decin. Grid of pseudo-2D chemistry models for tidally locked exoplanets - II. The role of photochemistry. *MNRAS*, 512:4877–4892, 2022. doi: 10.1093/mnras/stac809.

[276] Billy Edwards and Ian Stotesbury. Terminus: A Versatile Simulator

for Space-based Telescopes. *AJ.*, 161:266, 2021. doi: 10.3847/1538-3881/abdf4d.

[277] Lorenzo V. Mugnai, Ahmed Al-Refaie, Andrea Bocchieri, Quentin Changeat, Enzo Pascale, and Giovanna Tinetti. Alfnoor: Assessing the Information Content of Ariel's Low-resolution Spectra with Planetary Population Studies. *AJ.*, 162:288, 2021. doi: 10.3847/1538-3881/ac2e92.

[278] Benjamin Charnay, Joao M. Mendonca, Laura Kreidberg, Nicolas B. Cowan, Jake Taylor, Taylor J. Bell, Olivier Demangeon, Billy Edwards, Carole A. Haswell, Giuseppe Morello, Lorenzo V. Mugnai, Enzo Pascale, Giovanna Tinetti, Pascal Tremblin, and Robert T. Zellem. A survey of exoplanet phase curves with Ariel. *Exp. Astron.*, 53:417–446, 2022. doi: 10.1007/s10686-021-09715-x.

[279] Gloria Guilluy, Alessandro Sozzetti, Paolo Giacobbe, Aldo S. Bonomo, and Giuseppina Micela. On the synergy between Ariel and ground-based high-resolution spectroscopy. *Exp. Astron.*, 53:655–677, 2022. doi: 10.1007/s10686-021-09824-7.

[280] Billy Edwards and Giovanna Tinetti. The Ariel Target List: The Impact of TESS and the Potential for Characterizing Multiple Planets within a System. *AJ.*, 164:15, 2022. doi: 10.3847/1538-3881/ac6bf9.

[281] Joanna K. Barstow, Quentin Changeat, Katy L. Chubb, Patricio E. Cubillos, Billy Edwards, Ryan J. MacDonald, Michiel Min, and Ingo P. Waldmann. A retrieval challenge exercise for the Ariel mission. *Exp. Astron.*, 53:447–471, 2022. doi: 10.1007/s10686-021-09821-w.

[282] Billy Edwards, Quentin Changeat, Angelos Tsiaras, Kai Hou Yip, Ahmed F. Al-Refaie, Lara Anisman, Michelle F. Bieger, Amelie Gressier, Sho Shibata, Nour Skaf, Jeroen Bouwman, James Y-K. Cho, Masahiro Ikoma, Olivia Venot, Ingo Waldmann, Pierre-Olivier Lagage, and Giovanna Tinetti. Exploring the Ability of Hubble Space Telescope WFC3 G141 to Uncover Trends

in Populations of Exoplanet Atmospheres through a Homogeneous Transmission Survey of 70 Gaseous Planets. *ApJS*, 269, 2023. doi: 10.3847/1538-4365/ac9f1a.

[283] Andrea Bocchieri, Lorenzo V. Mugnai, Enzo Pascale, Quentin Changeat, and Giovanna Tinetti. Detecting molecules in Ariel low resolution transmission spectra. *Exp. Astron.*, 56:605–644, 2023. doi: 10.1007/s10686-023-09911-x.

[284] Q. Changeat, Y. Ito, A. F. Al-Refaie, K. H. Yip, and T. Lueftinger. Toward Atmospheric Retrievals of Panchromatic Light Curves: ExPLOR-ing Generalized Inversion Techniques for Transiting Exoplanets with JWST and Ariel. *AJ.*, 167(5):195, 2024. doi: 10.3847/1538-3881/ad3032.

[285] MacDonald, Ryan J. and Lewis, Nikole K. Trident: A rapid 3d radiative-transfer model for exoplanet transmission spectra. *ApJ*, 929:20, 2022. doi: 10.3847/1538-4357/ac47fe.

[286] I. P. Waldmann, M. Rocchetto, G. Tinetti, E. J. Barton, S. N. Yurchenko, and J. Tennyson. τ-REx II: Retrieval of Emission Spectra. *ApJ*, 813:13, 2015.

[287] A. F. Al-Refaie, Q. Changeat, I. P. Waldmann, and G. Tinetti. TauREx 3: A Fast, Dynamic, and Extendable Framework for Retrievals. *ApJ*, 917:37, 2021. doi: 10.3847/1538-4357/ac0252.

[288] Joseph Harrington, Michael D. Himes, Patricio E. Cubillos, Jasmina Blecic, Patricio M. Rojo, Ryan C. Challener, Nate B. Lust, M. Oliver Bowman, Sarah D. Blumenthal, Ian Dobbs-Dixon, Andrew S. D. Foster, Austin J. Foster, M. R. Green, Thomas J. Loredo, Kathleen J. McIntyre, Madison M. Stemm, and David C. Wright. An Open-source Bayesian Atmospheric Radiative Transfer (BART) Code. I. Design, Tests, and Application to Exoplanet HD 189733b. *Planet. Sci. J.*, 3:80, 2022. doi: 10.3847/PSJ/ac3513.

[289] Patricio E. Cubillos, Joseph Harrington, Jasmina Blecic, Michael D. Himes, Patricio M. Rojo, Thomas J. Loredo, Nate B. Lust, Ryan C. Challener,

Austin J. Foster, Madison M. Stemm, Andrew S. D. Foster, and Sarah D. Blumenthal. An Open-source Bayesian Atmospheric Radiative Transfer (BART) Code. II. The Transit Radiative Transfer Module and Retrieval of HAT-P-11b. *Planet. Sci. J.*, 3:81, 2022. doi: 10.3847/PSJ/ac348b.

[290] Elspeth K. H. Lee, Joost P. Wardenier, Bibiana Prinoth, Vivien Parmentier, Simon L. Grimm, Robin Baeyens, Ludmila Carone, Duncan Christie, Russell Deitrick, Daniel Kitzmann, Nathan Mayne, Michael Roman, and Brian Thorsbro. 3D Radiative Transfer for Exoplanet Atmospheres. gCMCRT: A GPU-accelerated MCRT Code. *ApJ*, 929:180, 2022. doi: 10.3847/1538-4357/ac61d6.

[291] Patricio E Cubillos and Jasmina Blecic. The PYRAT BAY framework for exoplanet atmospheric modelling: a population study of Hubble/WFC3 transmission spectra. *MNRAS*, 505:2675–2702, 2021. doi: 10.1093/mnras/stab1405.

[292] Geronimo L. Villanueva, Thomas J. Fauchez, Vincent Kofman, Eleonora Alei, Elspeth K. H. Lee, Estelle Janin, Michael D. Himes, Jeremy Leconte, Michaela Leung, Sara Faggi, Mei Ting Mak, Denis E. Sergeev, Thea Kozakis, James Manners, Nathan Mayne, Edward W. Schwieterman, Alex R. Howe, and Natasha Batalha. Modeling Atmospheric Lines by the Exoplanet Community (MALBEC) Version 1.0: A CUISINES Radiative Transfer Intercomparison Project. *Planet Space Sci.*, 5:64, 2024. doi: 10.3847/PSJ/ad2681.

[293] Michiel Min, Chris W. Ormel, Katy Chubb, Christiane Helling, and Yui Kawashima. The ARCiS framework for exoplanet atmospheres - Modelling philosophy and retrieval. *A&A*, 642:A28, 2020. doi: 10.1051/0004-6361/201937377.

[294] Lorenzo V. Mugnai and Darius Modirrousta-Galian. RAPOC: The Rosseland and Planck opacity converter A user-friendly and fast opacity program for

Python. *Exp. Astron.*, 55:521–539, 2023. doi: 10.1007/s10686-022-09869-2.

[295] Adam J. Wheeler, Andrew R. Casey, and Matthew W. Abruzzo. Korg: Fitting, Model Atmosphere Interpolation, and Brackett Lines. *AJ.*, 167:83, 2024. doi: 10.3847/1538-3881/ad19cc.

[296] Zhuchang Zhan, Sara Seager, Janusz Jurand Petkowski, Clara Sousa-Silva, Sukrit Ranjan, Jingcheng Huang, and William Bains. Assessment of Isoprene as a Possible Biosignature Gas in Exoplanets with Anoxic Atmospheres. *Astrobiology*, 21:765–792, 2021. doi: 10.1089/ast.2019.2146.

[297] Hiroyuki Tako Ishikawa, Wako Aoki, Takayuki Kotani, Masayuki Kuzuhara, Masashi Omiya, Ansgar Reiners, and Mathias Zechmeister. Elemental abundances of M dwarfs based on high-resolution near-infrared spectra: Verification by binary systems. *PASJ*, 72:102, 2020. doi: 10.1093/pasj/psaa101.

[298] G. Contursi, P. de Laverny, A. Recio-Blanco, and P. A. Palicio. GSP-spec line list for the parametrisation of *Gaia*-RVS stellar spectra. *A&A*, 654, 2021. doi: 10.1051/0004-6361/202140912.

[299] E. Marfil, H. M. Tabernero, D. Montes, J. A. Caballero, F. J. Lazaro, J. Gonzalez Hernandez, I, E. Nagel, V. M. Passegger, A. Schweitzer, I Ribas, A. Reiners, A. Quirrenbach, P. J. Amado, C. Cifuentes, M. Cortes-Contreras, S. Dreizler, C. Duque-Arribas, D. Galadi-Enriquez, Th Henning, S. Jeffers, V, A. Kaminski, M. Kuerster, M. Lafarga, A. Lopez-Gallifa, J. C. Morales, Y. Shan, and M. Zechmeister. The CARMENES search for exoplanets around M dwarfs Stellar atmospheric parameters of target stars with STEPARSYN. *A&A*, 656, 2021. doi: 10.1051/0004-6361/202141980.

[300] Yu P. Lyubchyk, Ya Pavlenko, V, O. K. Lyubchyk, and H. R. A. Jones. Bands of NaH lines in Spectra of Late Type Stars. *Kinemat. Phys. Celest.*, 38:159–165, 2022. doi: 10.3103/S0884591322030059.

[301] Hugh R. A. Jones, Yakiv Pavlenko, Yuri Lyubchik, Mike Bessell, Nicole Allard, and David J. Pinfield. A blue depression in the optical spectra of M dwarfs. *MNRAS*, 523:1297–1309, 2023. doi: 10.1093/mnras/stad1391.

[302] P I Cristofari, J-F Donati, C P Folsom, T Masseron, P Fouqué, C Moutou, E Artigau, A Carmona, P Petit, X Delfosse, E Martioli, and the SLS consortium. Constraining atmospheric parameters and surface magnetic fields with ZeeTurbo: an application to SPIRou spectra. *MNRAS*, 522:1342–1357, 2023. doi: 10.1093/mnras/stad865.

[303] Jerry W. Xuan, Jason Wang, Luke Finnerty, Katelyn Horstman, Simon Grimm, Anne E. Peck, Eric Nielsen, Heather A. Knutson, Dimitri Mawet, Howard Isaacson, Andrew W. Howard, Michael C. Liu, Sam Walker, Mark W. Phillips, Geoffrey A. Blake, Jean-Baptiste Ruffio, Yapeng Zhang, Julie Inglis, Nicole L. Wallack, Aniket Sanghi, Erica J. Gonzales, Fei Dai, Ashley Baker, Randall Bartos, Charlotte Z. Bond, Marta L. Bryan, Benjamin Calvin, Sylvain Cetre, Jacques-Robert Delorme, Greg Doppmann, Daniel Echeverri, Michael P. Fitzgerald, Nemanja Jovanovic, Joshua Liberman, Ronald A. Lopez, Emily C. Martin, Evan Morris, Jacklyn Pezzato, Garreth Ruane, Ben Sappey, Tobias Schofield, Andrew Skemer, Taylor Venenciano, J. Kent Wallace, Ji Wang, Peter Wizinowich, Yinzi Xin, Shubh Agrawal, Clarissa R. Do, Chih-Chun Hsu, and Caprice L. Phillips. Validation of Elemental and Isotopic Abundances in Late-M Spectral Types with the Benchmark HIP 55507 AB System. *ApJ*, 962, 2024. doi: 10.3847/1538-4357/ad1243.

[304] Y V Pavlenko, A Evans, D P K Banerjee, T R Geballe, U Munari, R D Gehrz, C E Woodward, and S Starrfield. Isotopic ratios in the red giant component of the recurrent nova T Coronae Borealis. *MNRAS*, 498:4853–4863, 2020. doi: 10.1093/mnras/staa2658.

[305] Danny Gasman, Ioannis Argyriou, G. C. Sloan, Bernhard Aringer, Javier Alvarez-Marquez, Ori Fox, Alistair Glasse, Adrian Glauser, Olivia C. Jones,

Kay Justtanont, Patrick J. Kavanagh, Pamela Klaassen, Alvaro Labiano, Kirsten Larson, David R. Law, Michael Mueller, Omnarayani Nayak, Alberto Noriega-Crespo, Polychronis Patapis, Pierre Royer, and Bart Vandenbussche. JWST MIRI/MRS in-flight absolute flux calibration and tailored fringe correction for unresolved sources. *A&A*, 673, 2023. doi: 10.1051/0004-6361/202245633.

[306] Christian R. Hayes, Kim A. Venn, Fletcher Waller, Jaclyn Jensen, Alan W. Mcconnachie, John Pazder, Federico Sestito, Andre Anthony, Gabriella Baker, John Bassett, Joao Bento, Trystyn Berg, Gregory Burley, Jurek Brzeski, Scott Case, Edward Chapin, Timothy Chin, Eric Chisholm, Vladimir Churilov, Adam Densmore, Ruben Diaz, Jennifer Dunn, Michael Edgar, Tony Farrell, Veronica Firpo, Joeleff Fitzsimmons, Juan Font-Serra, Javier Fuentes, Colin Ganton, Manuel Gomez-Jimenez, Tim Hardy, David Henderson, Alexis Hill, Brian Hoff, Michael Ireland, Venu Kalari, Neal Kelly, Urs Klauser, Yuriy Kondrat, Kathleen Labrie, Sam Lambert, Lance Luvaul, Jon Lawrence, Jordan Lothrop, G. Scott Macdonald, Slavko Mali, Steve Margheim, Richard Mcdermid, Helen Mcgregor, Bryan Miller, Felipe Miranda, Rolf Muller, Jon Nielsen, Ryan Norbury, Oliver Oberdorf, Naveen Pai, Gabriel Perez, Pablo Prado, Ian Price, Carlos Quiroz, Vladimir Reshetov, Gordon Robertson, Roque Ruiz-Carmona, Ricardo Salinas, Kim M. Sebo, Andrew Sheinis, Matthew Shetrone, Keith Shortridge, Katherine Silversides, Karleyne Silva, Chris Simpson, Greg Smith, Kei Szeto, Julia Tims, Eduardo Toro, Cristian Urrutia, Sudharshan Venkatesan, Lewis Waller, Ivan Wevers, Ramunas Wierzbicki, Marc White, Peter Young, and Ross Zhelem. GHOST Commissioning Science Results: Identifying a New Chemically Peculiar Star in Reticulum II. *ApJ*, 955, 2023. doi: 10.3847/1538-4357/acebc0.

[307] Timothy D. Brandt, Trent J. Dupuy, Brendan P. Bowler, Daniella C. Bardalez Gagliuffi, Jacqueline Faherty, G. Mirek Brandt, and Daniel Michalik. A Dynamical Mass of 70±5 M_{Jup} for Gliese 229B, the First T Dwarf. *AJ.*, 160: 196, 2020. doi: 10.3847/1538-3881/abb45e.

[308] Anjali A A Piette and Nikku Madhusudhan. Considerations for atmospheric retrieval of high-precision brown dwarf spectra. *MNRAS*, 497:5136–5154, 2020. doi: 10.1093/mnras/staa2289.

[309] Aaron M. Meisner, Adam C. Schneider, Adam J. Burgasser, Federico Marocco, Michael R. Line, Jacqueline K. Faherty, J. Davy Kirkpatrick, Dan Caselden, Marc J. Kuchner, Christopher R. Gelino, Jonathan Gagne, Christopher Theissen, Roman Gerasimov, Christian Aganze, Chih-chun Hsu, John P. Wisniewski, Sarah L. Casewell, Daniella C. Bardalez Gagliuffi, Sarah E. Logsdon, Peter R. M. Eisenhardt, Katelyn Allers, John H. Debes, Michaela B. Allen, Nikolaj Stevnbak Andersen, Sam Goodman, Leopold Gramaize, David W. Martin, Arttu Sainio, Michael C. Cushing, and The Backyard Worlds Planet 9 Collaboration. New Candidate Extreme T Subdwarfs from the Backyard Worlds: Planet 9 Citizen Science Project. *ApJ*, 915, 2021. doi: 10.3847/1538-4357/ac013c.

[310] C. von Essen, M. Mallonn, A. Piette, N. B. Cowan, N. Madhusudhan, E. Agol, V. Antoci, K. Poppenhaeger, K. G. Stassun, S. Khalafinejad, and G. Tautvaisiene. TESS unveils the optical phase curve of KELT-1b - Thermal emission and ellipsoidal variation from the brown dwarf companion along with the stellar activity. *A&A*, 648:A71, 2021. doi: 10.1051/0004-6361/202038524.

[311] Sagnick Mukherjee, Jonathan J. Fortney, Rebecca Jensen-Clem, Xianyu Tan, Mark S. Marley, and Natasha E. Batalha. Modeling Polarization Signals from Cloudy Brown Dwarfs Luhman 16 A and B in Three Dimensions. *AJ.*, 923: 113, 2021. doi: 10.3847/1538-4357/ac2d92.

[312] Mark S. Marley, Didier Saumon, Channon Visscher, Roxana Lupu, Richard Freedman, Caroline Morley, Jonathan J. Fortney, Christopher Seay, Adam J. R. W. Smith, D. J. Teal, and Ruoyan Wang. The Sonora Brown Dwarf Atmosphere and Evolution Models. I. Model Description and Application

to Cloudless Atmospheres in Rainout Chemical Equilibrium. *ApJ*, 920:85, 2021. doi: 10.3847/1538-4357/ac141d.

[313] Michael C. Cushing, Adam C. Schneider, J. Davy Kirkpatrick, Caroline V. Morley, Mark S. Marley, Christopher R. Gelino, Gregory N. Mace, Edward L. Wright, Peter R. Eisenhardt, Michael F. Skrutskie, and Kenneth A. Marsh. An Improved Near-infrared Spectrum of the Archetype Y Dwarf WISEP J182831.08+265037.8. *ApJ*, 920:20, 2021. doi: 10.3847/1538-4357/ac12cb.

[314] Anna Lueber, Daniel Kitzmann, Brendan P. Bowler, Adam J. Burgasser, and Kevin Heng. Retrieval Study of Brown Dwarfs across the L-T Sequence. *ApJ*, 930:136, 2022. doi: 10.3847/1538-4357/ac63b9.

[315] Megan E Tannock, Stanimir Metchev, Callie E Hood, Gregory N Mace, Jonathan J Fortney, Caroline V Morley, Daniel T Jaffe, and Roxana Lupu. A 1.46-2.48 μm spectroscopic atlas of a T6 dwarf (1060 K) atmosphere with IGRINS: First detections of H_2S and H_2, and verification of H_2O, CH_4, and NH_3 line lists. *MNRAS*, 514, 2022. doi: 10.1093/mnras/stac1412.

[316] Jerry W. Xuan, Jason Wang, Jean-Baptiste Ruffio, Heather Knutson, Dimitri Mawet, Paul Molli‘ere, Jared Kolecki, Arthur Vigan, Sagnick Mukherjee, Nicole Wallack, Ji Wang, Ashley Baker, Randall Bartos, Geoffrey A. Blake, Charlotte Z. Bond, Marta Bryan, Benjamin Calvin, Sylvain Cetre, Mark Chun, Jacques-Robert Delorme, Greg Doppmann, Daniel Echeverri, Luke Finnerty, Michael P. Fitzgerald, Katelyn Horstman, Julie Inglis, Nemanja Jovanovic, Ronald López, Emily C. Martin, Evan Morris, Jacklyn Pezzato, Sam Ragland, Bin Ren, Garreth Ruane, Ben Sappey, Tobias Schofield, Andrew Skemer, Taylor Venenciano, J. Kent Wallace, and Peter Wizinowich. A Clear View of a Cloudy Brown Dwarf Companion from High-resolution Spectroscopy. *ApJ*, 937:54, 2022. doi: 10.3847/1538-4357/ac8673.

[317] Melanie J. Rowland, Caroline V. Morley, and Michael R. Line. Toward Ro-

bust Atmospheric Retrieval on Cloudy L Dwarfs: the Impact of Thermal and Abundance Profile Assumptions. *ApJ*, 947:6, 2023. doi: 10.3847/1538-4357/acbb07.

[318] David Barrado, Paul Mollière, Polychronis Patapis, Michiel Min, Pascal Tremblin, Francisco Ardevol Martinez, Niall Whiteford, Malavika Vasist, Ioannis Argyriou, Matthias Samland, Pierre-Olivier Lagage, Leen Decin, Rens Waters, Thomas Henning, María Morales-Calderón, Manuel Guedel, Bart Vandenbussche, Olivier Absil, Pierre Baudoz, Anthony Boccaletti, Jeroen Bouwman, Christophe Cossou, Alain Coulais, Nicolas Crouzet, René Gastaud, Alistair Glasse, Adrian M. Glauser, Inga Kamp, Sarah Kendrew, Oliver Krause, Fred Lahuis, Michael Mueller, Göran Olofsson, John Pye, Daniel Rouan, Pierre Royer, Silvia Scheithauer, Ingo Waldmann, Luis Colina, Ewine F. van Dishoeck, Tom Ray, Göran Östlin, and Gillian Wright. $^{15}NH_3$ in the atmosphere of a cool brown dwarf. *Nature*, 624:263–266, 2023. doi: 10.1038/s41586-023-06813-y.

[319] Sagnick Mukherjee, Natasha E. Batalha, Jonathan J. Fortney, and Mark S. Marley. PICASO 3.0: A One-dimensional Climate Model for Giant Planets and Brown Dwarfs. *ApJ*, 942:71, 2023. doi: 10.3847/1538-4357/ac9f48.

[320] Callie E. Hood, Jonathan J. Fortney, Michael R. Line, and Jacqueline K. Faherty. Brown Dwarf Retrievals on FIRE!: Atmospheric Constraints and Lessons Learned from High Signal-to-noise Medium-resolution Spectroscopy of a T9 Dwarf. *ApJ*, 953:170, 2023. doi: 10.3847/1538-4357/ace32e.

[321] Kyle Franson, Brendan P. Bowler, Mariangela Bonavita, Timothy D. Brandt, Minghan Chen, Matthias Samland, Zhoujian Zhang, Anna Lueber, Kevin Heng, Daniel Kitzmann, Trevor Wolf, Brandon A. Jones, Quang H. Tran, Daniella C. Bardalez Gagliuffi, Beth Biller, Jeffrey Chilcote, Justin R. Crepp, Trent J. Dupuy, Jacqueline Faherty, Clémence Fontanive, Tyler D. Groff, Raffaele Gratton, Olivier Guyon, Rebecca Jensen-Clem, Nemanja Jovanovic,

N. Jeremy Kasdin, Julien Lozi, Eugene A. Magnier, Koraljka Mužić, Aniket Sanghi, and Christopher A. Theissen. Astrometric Accelerations as Dynamical Beacons: Discovery and Characterization of HIP 21152 B, the First T-dwarf Companion in the Hyades*. *AJ.*, 165:39, 2023. doi: 10.3847/1538-3881/aca408.

[322] Elspeth K H Lee, Xianyu Tan, and Shang-Min Tsai. Dynamically coupled kinetic chemistry in brown dwarf atmospheres - I. Performing global scale kinetic modelling. *MNRAS*, 523:4477–4491, 2023. doi: 10.1093/mnras/stad1715.

[323] Sagnick Mukherjee, Jonathan J. Fortney, Caroline V. Morley, Natasha E. Batalha, Mark S. Marley, Theodora Karalidi, Channon Visscher, Roxana Lupu, Richard Freedman, and Ehsan Gharib-Nezhad. The Sonora Substellar Atmosphere Models. IV. Elf Owl: Atmospheric Mixing and Chemical Disequilibrium with Varying Metallicity and C/O Ratios. *ApJ*, 963:73, 2024. doi: 10.3847/1538-4357/ad18c2.

[324] Adam J. Burgasser, Rachel Bezanson, Ivo Labbe, Gabriel Brammer, Sam E. Cutler, Lukas J. Furtak, Jenny E. Greene, Roman Gerasimov, Joel Leja, Richard Pan, Sedona H. Price, Bingjie Wang, John R. Weaver, Katherine E. Whitaker, Seiji Fujimoto, Vasily Kokorev, Pratika Dayal, Themiya Nanayakkara, Christina C. Williams, Danilo Marchesini, Adi Zitrin, and Pieter van Dokkum. UNCOVER: JWST Spectroscopy of Three Cold Brown Dwarfs at Kiloparsec-scale Distances. *ApJ*, 962:177, 2024. doi: 10.3847/1538-4357/ad206f.

[325] Ben W. P. Lew, Thomas Roellig, Natasha E. Batalha, Michael Line, Thomas Greene, Sagnick Murkherjee, Richard Freedman, Michael Meyer, Charles Beichman, Catarina Alves de Oliveira, Matthew De Furio, Doug Johnstone, Alexandra Z. Greenbaum, Mark Marley, Jonathan J. Fortney, Erick T. Young, Jarron Leisenring, Martha Boyer, Klaus Hodapp, Karl Misselt, John Stansberry, and Marcia Rieke. High-precision Atmospheric Characterization of

a Y Dwarf with JWST NIRSpec G395H Spectroscopy: Isotopologue, C/O Ratio, Metallicity, and the Abundances of Six Molecular Species. *AJ.*, 167: 237, 2024. doi: 10.3847/1538-3881/ad3425.

[326] Zhen Guo, P. W. Lucas, R. Kurtev, J. Borissova, C. Contreras Peña, S. N. Yurchenko, L. C. Smith, D. Minniti, R. K. Saito, A. Bayo, M. Catelan, J. Alonso-García, A. Caratti o Garatti, C. Morris, D. Froebrich, J. Tennyson, K. Maucó, A. Aguayo, N. Miller, and H. D. S. Muthu. Spectroscopic confirmation of high-amplitude eruptive YSOs and dipping giants from the VVV survey. *MNRAS*, 528:1769–1788, 2024. doi: 10.1093/mnras/stad3700.

[327] T. Danilovich, C. A. Gottlieb, L. Decin, A. M. S. Richards, K. L. K. Lee, T. Kaminski, N. A. Patel, K. H. Young, and K. M. Menten. Rotational Spectra of Vibrationally Excited AlO and TiO in Oxygen-rich Stars. *AJ.*, 904:110, 2020. doi: 10.3847/1538-4357/abc079.

[328] A. Coenegrachts, T. Danilovich, F. De Ceuster, and L. Decin. The unusual 3D distribution of NaCl around the asymptotic giant branch star IK Tau. *A&A*, 678, 2023. doi: 10.1051/0004-6361/202346116.

[329] C. Sandin, L. Mattsson, K. L. Chubb, M. Ergon, and P. M. Weilbacher. Three-component modelling of O-rich AGB star winds I. Effects of drift using forsterite. *A&A*, 677, 2023. doi: 10.1051/0004-6361/202345841.

[330] Giada Pastorelli, Paola Marigo, Léo Girardi, Bernhard Aringer, Yang Chen, Stefano Rubele, Michele Trabucchi, Sara Bladh, Martha L Boyer, Alessandro Bressan, Julianne J Dalcanton, Martin A T Groenewegen, Thomas Lebzelter, Nami Mowlavi, Katy L Chubb, Maria-Rosa L Cioni, Richard de Grijs, Valentin D Ivanov, Ambra Nanni, Jacco Th van Loon, and Simone Zaggia. Constraining the thermally pulsing asymptotic giant branch phase with resolved stellar populations in the Large Magellanic Cloud. *MNRAS*, 498: 3283–3301, 2020. doi: 10.1093/mnras/staa2565.

[331] K. Eriksson, S. Hofner, and B. Aringer. Synthetic photometry for carbon-rich

giants V. Effects of grain-size-dependent dust opacities. *A&A*, 673, 2023. doi: 10.1051/0004-6361/202245206.

[332] V Karambelkar, M. M. Kasliwal, P. Tisserand, G. C. Clayton, C. L. Crawford, S. G. Anand, T. R. Geballe, and E. Montiel. R Coronae Borealis and dustless hydrogen-deficient carbon stars likely have different oxygen isotope ratios. *A&A*, 667, 2022. doi: 10.1051/0004-6361/202142918.

[333] D. A. García-Hernández, N. Kameswara Rao, David L. Lambert, K. Eriksson, A. B. S. Reddy, and Thomas Masseron. The Carbon Star DY Persei May Be a Cool R Coronae Borealis Variable. *ApJ*, 948:15, 2023. doi: 10.3847/1538-4357/acc574.

[334] Muhammad Zain Mobeen, Tomasz Kaminski, Alexis Matter, Markus Wittkowski, and Claudia Paladini. The mid-infrared environment of the stellar merger remnant V838 Monocerotis. *A&A*, 655:A100, 2021. doi: 10.1051/0004-6361/202142297.

[335] T. Steinmetz, T. Kaminski, M. Schmidt, and A. Kiljan. A bipolar structure and shocks surrounding the stellar-merger remnant V1309 Scorpii. *A&A*, 682, 2024. doi: 10.1051/0004-6361/202347818.

[336] S. Liljegren, A. Jerkstrand, P. S. Barklem, G. Nyman, R. Brady, and S. N. Yurchenko. The molecular chemistry of Type Ibc supernovae and diagnostic potential with the James Webb Space Telescope. *A&A*, 674:A184, 2023. doi: 10.1051/0004-6361/202243491.

[337] Sarah Nickerson, Naseem Rangwala, Sean W. J. Colgan, Curtis DeWitt, Xinchuan Huang, Kinsuk Acharyya, Maria Drozdovskaya, Ryan C. Fortenberry, Eric Herbst, and Timothy J. Lee. The First Mid-infrared Detection of HNC in the Interstellar Medium: Probing the Extreme Environment toward the Orion Hot Core. *ApJ*, 907(1):51, 2021. doi: 10.3847/1538-4357/abca36.

[338] Sarah Nickerson, Naseem Rangwala, Sean W. J. Colgan, Curtis DeWitt, Jose. S. S. Monzon, Xinchuan Huang, Kinsuk Acharyya, Maria. N. N. Droz-

dovskaya, Ryan. C. C. Fortenberry, Eric Herbst, and Timothy. J. J. Lee. The Mid-infrared Molecular Inventory toward Orion IRc2. *ApJ*, 945:26, 2023. doi: 10.3847/1538-4357/aca6e8.

[339] K. P. Bowen, P-M Hillenbrand, J. Lievin, D. W. Savin, and X. Urbain. Dynamics of the isotope exchange reaction of d with h_3^+, h_2d^+, and d_2h^+. *J. Chem. Phys.*, 154, 2021. doi: 10.1063/5.0038434.

[340] Anthony C. LoCurto, Michael A. Welch, Travis R. Sippel, and James B. Michael. High-speed visible supercontinuum laser absorption spectroscopy of metal oxides. *Opt. Lett.*, 46:3288–3291, 2021. doi: 10.1364/OL.428456.

[341] K. A. Daniel, C. M. Murzyn, D. J. Allen, K. P. Lynch, C. R. Downing, and J. L. Wagner. Coaxial laser absorption and optical emission spectroscopy of high-pressure aluminum monoxide. *Opt. Lett.*, 47(9):2350–2353, 2022. doi: 10.1364/OL.456342.

[342] C. M. Murzyn, D. J. Allen, A. N. Baca, M. L. Ching, and R. T. Marinis. Tunable infrared laser absorption spectroscopy of aluminum monoxide A $^2\Pi_i$ - X $^2\Sigma^+$. *J. Quant. Spectrosc. Radiat. Transf.*, 279:108029, 2022. doi: 10.1016/j.jqsrt.2021.108029.

[343] Jonathan J. Gilvey, Morgan D. Ruesch, Kyle A. Daniel, Charley R. Downing, Kyle P. Lynch, Justin L. Wagner, and Christopher S. Goldenstein. Quantum-cascade-laser-absorption-spectroscopy diagnostic for temperature, pressure, and NO $X^2\Pi_{1/2}$ at 500 kHz in shock-heated air at elevated pressures. *Appl. Optics*, 62:A12–A24, 2023. doi: 10.1364/AO.464623.

[344] A. Abraham, P.T. Lynch, and N. Glumac. On using ab initio calibration to fit temperature from AlO B-X emission. *Proceedings of the Combustion Institute*, 39:1249–1257, 2023. doi: 10.1016/j.proci.2022.09.050.

[345] Jens Riedel, Josefin Hufgard, and Yi You. LIBS at high duty-cycles: effect of repetition rate and temporal width on the excitation laser pulses. *Frontiers in Physics*, 11:1241533, 2023. doi: 10.3389/fphy.2023.1241533.

[346] S. M. Bobrovnikov, E. V. Gorlov, and V. I. Zharkov. Efficiency of Laser Excitation of PO Photofragments of Organophosphates. *Atmos. Ocean. Opt.*, 35:329–340, 2022. doi: 10.1134/S1024856022040017.

[347] Sergei M. Bobrovnikov, Evgeny V. Gorlov, and Viktor I. Zharkov. Laser-induced fluorescence of PO-photofragments of dimethyl methylphosphonate. *Appl. Optics*, 61:6322–6329, 2022. doi: 10.1364/AO.456005.

[348] Jonathan Merten, Erin Nicholas, Shawnda Ethridge, Hannah Bariola, Shealyn Chestnut, Anna Anders, Jackie Brees, and Mary Foster. Following laser-induced plasma stoichiometry with atomic absorption spectroscopy. *Spectroc. Acta Pt. A-Molec. Biomolec. Spectr.*, 200:106600, 2023. doi: 10.1016/j.sab.2022.106600.

[349] E. Pawelec, D. Borodin, S. Brezinsek, T. Dittmar, D. Douai, D. Mazur, A. Meigs, A. Shaw, B. Thomas, JET Contributors, and EUROfusion Tokamak Exploitation. Internal energy distributions of beh, bed, and bet molecules created during chemically assisted physical sputtering in jet tokamak plasma. *Phys. Plasmas*, 31:042516, 2024. doi: 10.1063/5.0199084.

[350] Melody Yekta, Mohammad Ali Zanjanchi, and Hossein Roohi. Adsorption of hcn, hnc and ch¡sub¿3¡/sub¿cn toxic gases on the m-doped (m = cr, fe, ni and zn) ganns: A dft-d study. *COLLOIDS AND SURFACES A-PHYSICOCHEMICAL AND ENGINEERING ASPECTS*, 684:133120, 2024. doi: 10.1016/j.colsurfa.2023.133120.

[351] Morgan D. Ruesch, Austin J. McDonald, Garrett C. Mathews, Steven F. Son, and Christopher S. Goldenstein. Characterization of the influence of aluminum particle size on the temperature of composite-propellant flames using CO absorption and AlO emission spectroscopy. *Proc. Combust. Inst.*, 38: 4365–4372, 2021. doi: 10.1016/j.proci.2020.06.163.

[352] Alexander Fateev. From optimising waste incineration to internal combustion

engines and defence applications. *Johnson Matthey Technol. Rev.*, 67:25–35, 2023. doi: 10.1595/205651323x16643556587827.

[353] Zhi Qin, Tianrui Bai, and Linhua Liu. Rovibrationally Resolved Photodissociation of AlH via Excited Electronic States. *ApJ*, 917:87, 2021. doi: 10.3847/1538-4357/ac06d1.

[354] Ramon S. da Silva, Maikel Y. Ballester, Laiz R. Ventura, and Carlos E. Fellows. Theoretical study of the spin-orbit coupling in the $X^2\Pi$ state of NO. *Chem. Phys. Lett.*, 780:138896, 2021. doi: 10.1016/j.cplett.2021.138896.

[355] Mark C. Zammit, Jeffery A. Leiding, James Colgan, Wesley Even, Christopher J. Fontes, and Eddy Timmermans. A comprehensive study of the radiative properties of NO-a first step toward a complete air opacity. *J. Phys. B-At. Mol. Opt. Phys.*, 55, 2022. doi: 10.1088/1361-6455/ac8213.

[356] Andrea Pietropolli Charmet, Paolo Stoppa, Alessandra De Lorenzi, Mattia Melosso, Andre Achilli, Luca Dore, Cristina Puzzarini, Elisabetta Cane, and Filippo Tamassia. Computational, rotational and ro-vibrational experimental investigation of monodeuterated chloromethane. *J. Quant. Spectrosc. Radiat. Transf.*, 305, 2023. doi: 10.1016/j.jqsrt.2023.108624.

[357] Tianrui Bai, Zhi Qin, and Linhua Liu. Ultraviolet spectroscopy of AlO from first principle. *J. Quant. Spectrosc. Radiat. Transf.*, 302:108587, 2023. doi: 10.1016/j.jqsrt.2023.108587.

[358] A. M. Velasco and C. Lavín. Absorption Line Oscillator Strengths for the $C^2\Pi(0)$-$A^2\Sigma^+(0)$ Band of Nitric Oxide. *ApJ*, 954:144, 2023. doi: 10.3847/1538-4357/acebcb.

[359] O. N. Ulenikov, O. V. Gromova, E. S. Bekhtereva, N. I. Raspopova, E. A. Sklyarova, C. Sydow, K. Berezkin, C. Maul, and S. Bauerecker. Line strengths, widths and shifts analysis of the $2\nu_2$, $\nu_2 + \nu_4$ and $2\nu_4$ bands in $^{28}SiH_4$, $^{29}SiH_4$ and $^{30}SiH_4$. *J. Quant. Spectrosc. Radiat. Transf.*, 270, 2021. doi: 10.1016/j.jqsrt.2021.107683.

[360] Dmytro Shapko, Petr Dohnal, Stepan Roucka, Liliia Uvarova, Miroslava Kassayova, Radek Plasil, and Juraj Glosik. Cavity ring-down spectroscopy study of neon assisted recombination of H_3^+ ions with electrons. *J. Mol. Spectrosc.*, 378:111450, 2021. doi: 10.1016/j.jms.2021.111450.

[361] Aigars Znotins, Florian Grussie, Andreas Wolf, Xavier Urbain, and Holger Kreckel. An approach for multi-color action spectroscopy of highly excited states of H_3^+. *J. Mol. Spectrosc.*, 378:111476, 2021. doi: 10.1016/j.jms.2021.111476.

[362] Aleksandra Foltynowicz, Lucile Rutkowski, Isak Silander, Alexandra C. Johansson, Vinicius Silva de Oliveira, Ove Axner, Grzegorz Sobon, Tadeusz Martynkien, Pawel Mergo, and Kevin K. Lehmann. Measurement and assignment of double-resonance transitions to the 8900-9100-cm^{-1} levels of methane. *Phys. Rev. A*, 103, 2021. doi: 10.1103/PhysRevA.103.022810.

[363] C. Lavín and A. M. Velasco. Transition Energies and Line Oscillator Strengths of the $C^2\Pi(0)$-$X^2\Pi(0–6)$ Absorption Bands of Nitric Oxide. A Theoretical Study. *ApJ*, 941:29, 2022. doi: 10.3847/1538-4357/ac9e56.

[364] A. Perrin, L. Manceron, R. Armante, F. Kwabia-Tchana, P. Roy, and G.C. Toon. First investigation of the ν_1, ν_1-ν_9, ν_1+ν_9, and ν_1+ν_7 absorption bands of nitric acid ($H^{14}N^{16}O_3$) at 3551.766 cm^{-1}, 3092.708 cm^{-1}, 4006.974 cm^{-1}, and 4127.782 cm^{-1}, respectively. *J. Mol. Spectrosc.*, 392:111741, 2023. doi: 10.1016/j.jms.2023.111741.

[365] A. Perrin, L. Manceron, R. Armante, F. Kwabia-Tchana, P. Roy, and G.C. Toon. Part 2: Validation for line lists generated for nitric acid ($H^{14}N^{16}O_3$) for the ν_1 band and its first two associated hot bands (ν_1+ν_9-ν_9, ν_1+ν_7-ν_7) in the 2.8 μm region, the ν_1-ν_9, ν_1+ν_9 and ν_1+ν_7 bands at 3.2 μm, 2.5 μm and 2.4 μm, respectively. *J. Mol. Spectrosc.*, 392:111740, 2023. doi: 10.1016/j.jms.2023.111740.

[366] I. A. Vasilenko, O. V. Naumenko, and V.-M. Horneman. Expert List of Absorption Lines of the $^{32}S^{16}O_2$ Molecule in the 0-4200 cm^{-1} Spectral Region. *Atmos. Ocean. Opt.*, 36:199–206, 2023. doi: 10.1134/S102485602303020X.

[367] K. L. Chubb, M. Rocchetto, S. N. Yurchenko, M. Min, I. Waldmann, J. K. Barstow, P. Molliére, A. F. Al-Refaie, M. Phillips, and J. Tennyson. The ExoMolOP Database: Cross-sections and *k*-tables for Molecules of Interest in High-Temperature Exoplanet Atmospheres. *A&A*, 646:A21, 2020. doi: 10.1051/0004-6361/202038350.

[368] Alec Owens, Tiantian He, Martin Hanicinec, Christian Hill, Sebastian Mohr, and J. Tennyson. LiDB: Database of molecular vibronic state radiative lifetimes for plasma processes. *Plasma Sources Sci. Technol.*, 32:085015, 2023. doi: 10.1088/1361-6595/aceeb0.

[369] J Tennyson, S. Mohr, M. Hanicinec, A. Dzarasova, C. Smith, S. Waddington, Bingqing Liu, L. L. Alves, K. Bartschat, A. Bogaerts, S. U. Engelmann, T. Gans, A. R. Gibson, S. Hamaguchi, K. R. Hamilton, C. Hill, D. O'Connell, S. Rauf, K. van 't Veer, and O. Zatsarinny. The 2021 release of the Quantemol database (QDB) of plasma chemistries and reactions. *Plasma Sources Sci. Technol.*, 31:095020, 2022. doi: 10.1088/1361-6595/ac907e.

[370] Rong Wang, U. Balciunaite, Juncai Chen, Cheng Yuan, A. Owens, and J. Tennyson. NASA polynomial representation of molecular specific heats. *J. Quant. Spectrosc. Radiat. Transf.*, 306:108617, 2023. doi: 10.1016/j.jqsrt.2023.108617.

[371] J. Tennyson, K. Hulme, O. K. Naim, and S. N. Yurchenko. Radiative lifetimes and cooling functions for astrophysically important molecules. *J. Phys. B-At. Mol. Opt. Phys.*, 49:044002, 2016. doi: 10.1088/0953-4075/49/4/044002.

[372] J. Tennyson, C. Hill, and S. N. Yurchenko. Data structures for ExoMol: Molecular line lists for exoplanet and other atmospheres. In *6^{th} inter-*

national conference on atomic and molecular data and their applications ICAMDATA-2012, volume 1545 of *AIP Conference Proceedings*, pages 186–195. AIP, New York, 2013. doi: 10.1063/1.4815853.

[373] Sergei N. Yurchenko, Ahmed F. Al-Refaie, and Jonathan Tennyson. ExoCross: A general program for generating spectra from molecular line lists. *A&A*, 614:A131, 2018. doi: 10.1051/0004-6361/201732531.

[374] J. Tennyson, M. Pezzella, Jingxin Zhang, and S. N. Yurchenko. Data structures for photoadsorption within the ExoMol project. *RASTI*, 2:231–237, 2023. doi: 10.1093/rasti/rzad014.

[375] C Hill and M Hanicinec. A Python package for managing simple chemical species and states. https://github.com/xnx/pyvalem, 2024. version 2.7.1.

[376] C Hill, Dipti, M Hanicinec, and et al. PyValem: a machine-readable notation and Python library for atomic and molecular structures and states. Manuscript in preparation, 2024.

[377] Roman V Kochanov, IE Gordon, LS Rothman, P Wcisło, C Hill, and JS Wilzewski. HITRAN Application Programming Interface (HAPI): A comprehensive approach to working with spectroscopic data. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 177:15–30, 2016. doi: 10.1016/j.jqsrt.2016.03.005.

[378] N.H. Ngo, D. Lisak, H. Tran, and J.-M. Hartmann. An isolated line-shape model to go beyond the Voigt profile in spectroscopic databases and radiative transfer codes. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 129:89–100, 2013. ISSN 0022-4073. doi: 10.1016/j.jqsrt.2013.05.034.

[379] M Simeckova, D Jacquemart, L S Rothman, R R Gamache, and A Goldman. Einstein a-coefficients and statistical weights for molecular absorption transitions in the hitran database. *J. Quant. Spectrosc. Radiat. Transf.*, 98: 130–155, 2006. ISSN 0022-4073. doi: 10.1016/j.jqrst.2005.07.003.

[380] Jonathan Tennyson and Sergei N Yurchenko. ExoMol at 10. *Astron. Geophys.*, 62:6.16–6.21, 2021. doi: 10.1093/astrogeo/atab102.

[381] International Union of Pure and Applied Chemistry. http://www.iupac.org/, 1919. Accessed: 2025-01-25.

[382] National Institute of Standards and Technology. http://www.nist.gov/, 1901. Accessed: 2025-01-25.

[383] InChI Trust. http://www.inchi-trust.org/, 2009. Accessed: 2025-01-25.

[384] Stephen R. Heller, Alan McNaught, Igor Pletnev, Stephen Stein, and Dmitrii Tchekhovskoi. InChI, the IUPAC International Chemical Identifier. *J. Cheminfo.*, 7:23, 2015. doi: 10.1186/s13321-015-0068-4.

[385] M. Pezzella, S. N. Yurchenko, and J. Tennyson. A method for calculating temperature-dependent photodissociaiton cross sections and rates. *Phys. Chem. Chem. Phys.*, 23:16390–16400, 2021. doi: 10.1039/D1CP02162A.

[386] Qianwei Qu, S. N. Yurchenko, and J. Tennyson. An empirical spectroscopic model for eleven electronic states of VO. *J. Mol. Spectrosc.*, 391:111733, 2023. doi: 10.1016/j.jms.2022.111733.

[387] D Jack, P H Hauschildt, and E Baron. Time-dependent radiative transfer with PHOENIX. *A&A*, 502:1043–1049, 2009.

[388] L. S. Rothman, I. E. Gordon, Y. Babikov, A. Barbe, D. Chris Benner, P. F. Bernath, M. Birk, L. Bizzocchi, V. Boudon, L. R. Brown, A. Campargue, K. Chance, E. A. Cohen, L. H. Coudert, V. M. Devi, B. J. Drouin, A. Fayt, J.-M. Flaud, R. R. Gamache, J. J. Harrison, J.-M. Hartmann, C. Hill, J. T. Hodges, D. Jacquemart, A. Jolly, J. Lamouroux, R. J. Le Roy, G. Li, D. A. Long, O. M. Lyulin, C. J. Mackie, S. T. Massie, S. Mikhailenko, H. S. P. Müller, O. V. Naumenko, A. V. Nikitin, J. Orphal, V. Perevalov, A. Perrin, E. R. Polovtseva, C. Richard, M. A. H. Smith, E. Starikova, K. Sung, S. Tashkun, J. Tennyson, G. C. Toon, Vl. G. Tyuterev, and G. Wagner. The

HITRAN 2012 molecular spectroscopic database. *J. Quant. Spectrosc. Radiat. Transf.*, 130:4 – 50, 2013. doi: 10.1016/jqsrt.2013.07.002.

[389] Colin M. Western. PGOPHER: A program for simulating rotational, vibrational and electronic spectra . *J. Quant. Spectrosc. Radiat. Transf.*, 186:221–242, 2017. doi: 10.1016/j.jqsrt.2016.04.010.

[390] A. Owens, J. Tennyson, and S. N. Yurchenko. ExoMol line lists – XLI. High-temperature molecular line lists for the alkali metal hydroxides KOH and NaOH. *MNRAS*, 502:1128–1135, 2021. doi: 10.1093/mnras/staa4041.

[391] Paola Marigo, Francesco Addari, Diego Bossini, Alessandro Bressan, Guglielmo Costa, Leo Girardi, Giada Pastorelli, Michele Trabucchi, and Guglielmo Volpato. ÆSOPUS 2.1: Low-temperature Opacities Extended to High Pressure. *ASTROPHYSICAL JOURNAL*, 976(1), NOV 1 2024. ISSN 0004-637X. doi: 10.3847/1538-4357/ad7b27.

[392] Simon L. Grimm, Matej Malik, Daniel Kitzmann, Andrea Guzmán-Mesa, H. Jens Hoeijmakers, Chloe Fisher, Jo ao M. Mendonça, Sergey N. Yurchenko, Jonathan Tennyson, Robert L. Kurucz, and Kevin Heng. HELIOS-K 2.0 and an Open-source Opacity Database for Exoplanetary Atmospheres. *ApJS*, 253:30, 2021. doi: 10.3847/1538-4365/abd773.

[393] Simon L. Grimm, Matej Malik, Daniel Kitzmann, Andrea Guzmán-Mesa, H. Jens Hoeijmakers, Chloe Fisher, João M. Mendonça, Sergey N. Yurchenko, Jonathan Tennyson, Fabien Alesina, Nicolas Buchschacher, Julien Burnier, Damien Segransan, Robert L. Kurucz, and Kevin Heng. HELIOS-K 2.0 Opacity Calculator and Open-source Opacity Database for Exoplanetary Atmospheres. *ApJS*, 253:30, 2021. doi: 10.3847/1538-4365/abd773.

[394] Roman V Kochanov, Iouli E Gordon, Laurence S Rothman, Yan Tan, Joshua Karns, Wyatt Matt, and Christian Hill. Updates and Current Status of

the HITRAN Application Programming Interface (HAPI). In *73rd International Symposium on Molecular Spectroscopy*, page MI07, 2018. doi: 10.15278/isms.2018.MI07.

[395] Christopher S Goldenstein, Victor A Miller, R Mitchell Spearrin, and Christopher L Strand. SpectraPlot. com: Integrated spectroscopic modeling of atomic and molecular gases. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 200:249–257, 2017. doi: 10.1016/j.jqsrt.2017.06.007.

[396] J. Tennyson and S. Miller. A program suite for the calculation of ro-vibrational spectra of triatomic molecules. *Comput. Phys. Commun.*, 55: 149–175, 1989.

[397] J. Tennyson, S. Miller, and C. R. Le Sueur. TRIATOM: a suite for basis set calculations of ro-vibrational spectra of triatomic molecules. *Comput. Phys. Commun.*, 75:339–364, 1993.

[398] D.C.M. van den Bekerom and E. Pannier. A discrete integral transform for rapid spectral synthesis. *J. Quant. Spectrosc. Radiat. Transf.*, 261:107476, 2021. doi: 10.1016/j.jqsrt.2020.107476.

[399] Kerson Huang. *Statistical Mechanics*. John Wiley & Sons, 2 edition, 1987. ISBN 0-471-81518-7.

[400] L. Neale and J. Tennyson. A high temperature partition function for H_3^+. *ApJ*, 454:L169–L173, 1995. doi: 10.1086/309789.

[401] C. Sousa-Silva, N. Hesketh, S. N. Yurchenko, C. Hill, and J. Tennyson. High Temperature partition functions and thermodynamic data for ammonia and phosphine. *J. Quant. Spectrosc. Radiat. Transf.*, 142:66–74, 2014. doi: 10.1016/j.jqsrt2014.03.012.

[402] A W Irwin. Polynomial partition-function approximations of 344 atomic species and molecular-species. *ApJS*, 45(4):621–633, 1981.

[403] A. J Sauval and J. B. Tatum. A set of partition functions and equilibrium constants for 300 diatomic molecules of astrophysical interest. *ApJS*, 56: 193–209, 1984. doi: 10.1086/190980.

[404] P. S. Barklem and R. Collet. Partition functions and equilibrium constants for diatomic molecules and atoms of astrophysical interest. *A&A*, 588:A96, 2016. doi: 10.1051/0004-6361/201526961.

[405] R. R. Gamache, C. Roller, E. Lopes, I. E. Gordon, L. S. Rothman, O. L. Polyansky, N. F. Zobov, A. A. Kyuberis, J. Tennyson, S N Yurchenko, A. G. Császár, T. Furtenbacher, Xinchuan Huang, David W. Schwenke, Timothy J. Lee, B J Drouin, S. A. Tashkun, V. I. Perevalov, and R V. Kochanov. Total Internal Partition Sums for 167 isotopologues of 53 molecules important in planetary atmospheres: application to HITRAN2016 and beyond. *J. Quant. Spectrosc. Radiat. Transf.*, 203:70–87, 2017. doi: 10.1016/j.jqsrt.2017.03.045.

[406] Ya. V. Pavlenko, S. N. Yurchenko, and Jonathan Tennyson. Analysis of first overtone bands of isotopologues of CO and SiO in stellar spectra. *A&A*, 633: A52, 2020.

[407] R J Barber, J K Strange, C Hill, O L Polyansky, G Ch Mellau, S N Yurchenko, and Jonathan Tennyson. ExoMol line lists – III. An improved hot rotation-vibration line list for HCN and HNC. *MNRAS*, 437:1828–1835, 2014. doi: 10.1093/mnras/stt2011.

[408] Gang Li, Iouli E. Gordon, Laurence S. Rothman, Yan Tan, Shui-Ming Hu, Samir Kassi, Alain Campargue, and Emile S. Medvedev. Rovibrational line lists for nine isotopologues of the CO molecule in the $X\ ^1\Sigma^+$ ground electronic state. *ApJS*, 216:15, 2015. doi: 10.1088/0067-0049/216/1/15.

[409] W. Somogyi, S. N. Yurchenko, and A. Yachmenev. Calculation of electric quadrupole linestrengths for diatomic molecules: Application to the H_2,

CO, HF, and O_2 molecules. *J. Chem. Phys.*, 155(21):214303, 2021. doi: 10.1063/5.0063256.

[410] R.A. EDWARDS. *The Measurement of Heat Quantity*, page 103–124. Elsevier, 1970. ISBN 9780080134314. doi: 10.1016/b978-0-08-013431-4.50011-8.

[411] David Halliday, Robert Resnick, and Jearl Walker. *Fundamentals of Physics*. John Wiley & Sons, 2013.

[412] Franz Schwabl. *Statistical mechanics*. Springer Science & Business Media, 2006.

[413] Jan M L Martin, Jean-Pierre François, and Renaat Gijbels. First principles computation of thermochemical properties beyond the harmonic approximation. i. method and application to the water molecule and its isotopomers. *J. Chem. Phys.*, 96:7633, 1992.

[414] R. Tóbiás, T. Furtenbacher, A. G. Császár, O. V. Naumenko, J. Tennyson, J.-M. Flaud, P. Kumard, and B. Poirier. Critical Evaluation of Measured Rotational-Vibrational Transitions of Four Sulphur Isotopologues of $S^{16}O_2$. *J. Quant. Spectrosc. Radiat. Transf.*, 208:152–163, 2018. doi: 10.1016/j.jqsrt.2018.01.006.

[415] Ralph S. Sutherland and M. A. Dopita. Cooling functions for low-density astrophysical plasmas. *The Astrophysical Journal Supplement Series*, 88: 253–327, 1993. ISSN 1538-4365. doi: 10.1086/191823.

[416] Ye Wang, Gary J Ferland, ML Lykins, RL Porter, Peter AM van Hoof, and RJR Williams. Radiative cooling II: effects of density and metallicity. *Monthly Notices of the Royal Astronomical Society*, 440(4):3100–3112, 2014. doi: 10.1093/mnras/stu514.

[417] J-C Gérard, Leonardos Gkouvelis, Bertrand Bonfond, G R Gladstone, A Mura, A Adriani, Denis Grodent, V Hue, and T K Greathouse. H3+ cool-

ing in the Jovian aurora: Juno remote sensing observations and modeling. *Icarus*, 389:115261, 2023. doi: 10.1016/j.icarus.2022.115261.

[418] T Giang Nguyen, Nicolas B Cowan, Raymond T Pierrehumbert, Roxana E Lupu, and John E Moores. The impact of ultraviolet heating and cooling on the dynamics and observability of lava planet atmospheres. *Monthly Notices of the Royal Astronomical Society*, 513(4):6125–6133, 2022. doi: 10.1093/mnras/stac1331.

[419] C. M. Coppola, L. Lodi, and J. Tennyson. Radiative cooling functions for primordial molecules. *MNRAS*, 415:487–493, 2011.

[420] S. Miller, T. Stallard, J. Tennyson, and H. Melin. Cooling by H_3^+ emission. *J. Phys. Chem. A*, 117:9770–9777, 2013.

[421] V. V. Melnikov, S. N. Yurchenko, J. Tennyson, and P. Jensen. Radiative cooling of H_3O^+ and its deuterated isotopologues. *Phys. Chem. Chem. Phys.*, 18:26268–26274, 2016. doi: 10.1039/C6CP04661D.

[422] Liesl Neale, Steven Miller, and Jonathan Tennyson. Spectroscopic properties of the H 3+ molecule: a new calculated line list. *The Astrophysical Journal*, 464:516, 1996. doi: 10.1086/177341.

[423] C. A. Bowesman, Meiyin Shuai, Sergei N. Yurchenko, and Jonathan Tennyson. A high resolution line list for AlO. *MNRAS*, 508:3181–3193, 2021. doi: 10.1093/mnras/stab2525.

[424] Wes McKinney. Data Structures for Statistical Computing in Python. In *Proceedings of the 9th Python in Science Conference*, volume 445, pages 51–56. Austin, TX, 2010. doi: 10.25080/Majora-92bf1922-00a.

[425] Wolfgang Demtröder. Laser spectroscopy: Basic concepts and instrumentation. *NASA STI/Recon Technical Report A*, 82:12273, 1981. doi: 10.1007/978-3-662-08257-7.

[426] RV Orloski and AG Trigueiros. Weighted oscillator strengths and lifetimes for the Si VIII spectrum. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 69(1):13–21, 2001. doi: 10.1016/S0022-4073(00)00049-2.

[427] Clifford W. Mendel and Edl Schamiloglu. *Plasma Diagnostic Techniques*, page 391–400. Elsevier, 2003. ISBN 9780122274107. doi: 10.1016/b0-12-227410-5/00583-4.

[428] U. G. Jørgensen, M. Larsson, A. Iwamae, and B. Yu. Line intensities for CH and their application to stellar atmospheres. *A&A*, 315:204, 1996.

[429] S. N. Yurchenko, J. Tennyson, Jeremy Bailey, Morgan D J Hollis, and Giovanna Tinetti. Spectrum of hot methane in astronomical objects using a comprehensive computed line list. *Proc. Nat. Acad. Sci.*, 111:9379–9383, 2014. doi: 10.1073/pnas.1324219111.

[430] Iouli E. Gordon, Laurence S. Rothman, Robert R. Gamache, David Jacquemart, Chris Boone, Peter F. Bernath, Mark W. Shephard, Jennifer S. Delamere, and Shepard A. Clough. Current updates of the water-vapor line list in HITRAN: A new "Diet" for air-broadened half-widths. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 108(3):389–402, 2007. ISSN 0022-4073. doi: 10.1016/j.jqsrt.2007.06.009.

[431] E. V. Karlovets, I. E. Gordon, L. S. Rothman, R. Hashemi, R. J. Hargreaves, G. Toon, A. Campargue, V. I. Perevalov, P. Čermak, M. Birk, G. Wagner, J. T. Hodges, J. Tennyson, and S. N. Yurchenko. The update of the line positions and intensities in the line list of carbon dioxide for the HITRAN2020 spectroscopic database. *J. Quant. Spectrosc. Radiat. Transf.*, 276:107896, 2021. doi: 10.1016/j.jqsrt.2021.107896.

[432] Michael M. McKerns, Leif Strand, Tim Sullivan, Alta Fang, and Michael A. G. Aivazis. Building a Framework for Predictive Science, 2012.

[433] *concurrent.futures*, 2009. Version 3.4.0.

[434] Dask Development Team. *Dask: Library for dynamic task scheduling*, 2016.

[435] Charles R. Harris, K. Jarrod Millman, St'efan J. van der Walt, Ralf Gommers, Pauli Virtanen, David Cournapeau, Eric Wieser, Julian Taylor, Sebastian Berg, Nathaniel J. Smith, Robert Kern, Matti Picus, Stephan Hoyer, Marten H. van Kerkwijk, Matthew Brett, Allan Haldane, Jaime Fern'andez del R'ıo, Mark Wiebe, Pearu Peterson, Pierre G'erard-Marchant, Kevin Sheppard, Tyler Reddy, Warren Weckesser, Hameer Abbasi, Christoph Gohlke, and Travis E. Oliphant. Array programming with NumPy. *Nature*, 585(7825):357–362, 2020. doi: 10.1038/s41586-020-2649-2.

[436] Pandarallel. https://github.com/nalepae/pandarallel, 2019. Version 1.6.5.

[437] D Cooke, T Hochberg, F Alted, I Vilata, M Wiebe, G de Menten, A Valentino, and RA McLeod. NumExpr: Fast numerical expression evaluator for NumPy, 2009.

[438] Thomas P Robitaille, Erik J Tollerud, Perry Greenfield, Michael Droettboom, Erik Bray, Tom Aldcroft, Matt Davis, Adam Ginsburg, Adrian M Price-Whelan, Wolfgang E Kerzendorf, et al. Astropy: A community Python package for astronomy. *Astronomy & Astrophysics*, 558:A33, 2013. doi: 10.1051/0004-6361/201322068.

[439] Franz Schreier. Computational aspects of speed-dependent Voigt profiles. *J. Quant. Spectrosc. Radiat. Transf.*, 187:44 – 53, 2017. doi: 10.1016/j.jqsrt.2016.08.009.

[440] L. Anisman, K. L. Chubb, J. Elsey, A. Al-Refaie, Q. Changeat, S. N. Yurchenko, J. Tennyson, and G. Tinetti. Cross-sections for heavy atmospheres: H_2O continuum. *J. Quant. Spectrosc. Radiat. Transf.*, 278:108013, 2022. doi: 10.1016/j.jqsrt.2021.108013.

[441] Ehsan (Sam) Gharib-Nezhad, Natasha E Batalha, Katy Chubb, Richard Freedman, Iouli E Gordon, Robert R Gamache, Robert J Hargreaves,

Nikole K Lewis, Jonathan Tennyson, and Sergei N Yurchenko. The impact of spectral line wing cut-off: Recommended standard method with application to MAESTRO opacity database. *RASTI*, 3:44–55, 2024. doi: 10.1093/rasti/rzad058.

[442] NH Ngo, Daniel Lisak, H Tran, and J-M Hartmann. An isolated line-shape model to go beyond the Voigt profile in spectroscopic databases and radiative transfer codes. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 129:89–100, 2013. doi: 10.1016/j.jqsrt.2013.05.034.

[443] Jonas S. Wilzewski, Iouli E. Gordon, Roman V. Kochanov, Christian Hill, and Laurence S. Rothman. H_2, He, and CO_2 line-broadening coefficients, pressure shifts and temperature-dependence exponents for the HITRAN database. Part 1: SO_2, NH_3, HF, HCl, OCS and C_2H_2. *J. Quant. Spectrosc. Radiat. Transf.*, 168:193–206, 2016. doi: 10.1016/j.jqsrt.2015.09.003.

[444] Y. Tan, R. Kochanov, V, L. S. Rothman, and I. E. Gordon. Introduction of Water-Vapor Broadening Parameters and Their Temperature-Dependent Exponents Into the HITRAN Database: Part I-CO2, N2O, CO, CH4, O-2, NH3, and H2S. *J. Quant. Spectrosc. Radiat. Transf.*, 124:11580–11594, 2019. doi: 10.1029/2019JD030929.

[445] P. Wcisło, F. Thibault, N. Stolarczyk, H.Jóźwiak, M. Słowinski, M. Gancewski, K. Stankiewicz, M. Konefał, S. Kassi, A. Campargue, Y. Tan, J. Wang, K. Patkowski, R. Ciuryło, D. Lisak, R. Kochanov, L.S. Rothman, and I.E. Gordon. The first comprehensive dataset of beyond-Voigt line-shape parameters from ab initio quantum scattering calculations for the HITRAN database: He-perturbed H_2 case study. *J. Quant. Spectrosc. Radiat. Transf.*, 260:107477, 2021. doi: 10.1016/j.jqsrt.2020.107477.

[446] Yan Tan, Frances M. Skinner, Shanelle Samuels, Robert J. Hargreaves, Robab Hashemi, and Iouli E. Gordon. H_2, He, and CO_2 Pressure-induced Parameters for the HITRAN Database. II. Line Lists of CO_2, N_2O, CO, SO_2,

OH, OCS, H_2CO, HCN, PH_3, H_2S, and GeH_4. *ApJS*, 262:40, 2022. doi: 10.3847/1538-4365/ac83a6.

[447] Gordon Leslie Squires. *Practical physics*. Cambridge university press, 2001. doi: 10.1017/CBO9781139164498.

[448] William Feller. *An introduction to probability theory and its applications, Volume 2*, page 704. New York: John Wiley & Sons Inc., 2 edition, 1971. doi: 10.1137/1014119.

[449] Norman L Johnson, Samuel Kotz, and Narayanaswamy Balakrishnan. *Continuous univariate distributions, volume 1*, chapter 16. New York: John Wiley & Sons Inc., 1994.

[450] Franz Schreier and Philipp Hochstaffl. Computational aspects of speed-dependent Voigt and Rautian profiles. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 258:107385, 2021. doi: 10.1016/j.jqsrt.2020.107385.

[451] Y. Pavlenko, J. Tennyson, S. N. Yurchenko, Mirek R. Schmidt, H. R. A. Jones, Y. Lyubchik, and A. Suárez Mascareño. AlH lines in the blue spectrum of Proxima Centauri. *MNRAS*, 516:5655–5673, 2022. doi: 10.1093/mnras/stac2588.

[452] BH Armstrong. Spectrum line profiles: the Voigt function. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 7(1):61–88, 1967. doi: 10.1016/0022-4073(67)90057-X.

[453] Floyd Herbert. Spectrum line profiles: a generalized Voigt function including collisional narrowing. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 14(9):943–951, 1974. doi: 10.1016/0022-4073(74)90021-1.

[454] Franz Schreier. The Voigt and complex error function: Humlíček's rational approximation generalized. *MNRAS*, 479(3):3068–3075, 2018. doi: 10.1093/mnras/sty1680.

[455] V. N. Faddeeva, N. M. Terent'ev, V. A. Fok, D. G. Fry, and B. A. Hons. *Tables of values of the function* $\omega(z) = \exp(-z^2)(1 + 2i/\sqrt{\pi}\int_0^2 \exp(t^2)\mathrm{d}t)$ *for complex argument*. Pergamon Press Oxford, 1961.

[456] Mofreh R Zaghloul. A simple reform for treating the loss of accuracy of Humlicek's W4 algorithm near the real axis. *arXiv preprint arXiv:1505.05596*, 2015. doi: 10.48550/arXiv.1505.05596.

[457] Joseph H Pierluissi, Peter C Vanderwood, and Richard B Gomez. Fast calculational algorithm for the Voigt profile. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 18(5):555–558, 1977. doi: 10.1016/0022-4073(77)90056-5.

[458] Josef Humlíček. An efficient method for evaluation of the complex probability function: the Voigt function and its derivatives. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 21(4):309–313, 1979. doi: 10.1016/0022-4073(79)90062-1.

[459] Josef Humlíček. Optimized computation of the Voigt and complex probability functions. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 27(4):437–444, 1982. doi: 10.1016/0022-4073(82)90078-4.

[460] M Kuntz. A new implementation of the Humlicek algorithm for the calculation of the Voigt profile function. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 57(6):819–824, 1997. doi: 10.1016/S0022-4073(96)00162-8.

[461] Mofreh R. Zaghloul. Algorithm 985: Simple, Efficient, and Relatively Accurate Approximation for the Evaluation of the Faddeyeva Function. *ACM Trans. Math. Softw.*, 44:22, aug 2017. doi: 10.1145/3119904.

[462] A K Hui, B H Armstrong, and A A Wray. Rapid computation of the Voigt and complex error functions. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 19(5):509–516, 1978. doi: 10.1016/0022-4073(78)90019-5.

[463] Takashi Ida, M Ando, and H Toraya. Extended pseudo-Voigt function for approximating the Voigt profile. *Journal of Applied Crystallography*, 33(6): 1311–1316, 2000. doi: 10.1107/S0021889800010219.

[464] P Thompson, DE Cox, and JB Hastings. Rietveld refinement of Debye-Scherrer synchrotron X-ray data from Al2O3. *Journal of Applied Crystallography*, 20(2):79–83, 1987. doi: 10.1107/S0021889887087090.

[465] F Sánchez-Bajo and FL Cumbrera. The use of the pseudo-Voigt function in the variance method of X-ray line-broadening analysis. *Journal of Applied Crystallography*, 30(4):427–430, 1997. doi: 10.1107/s0021889896015464.

[466] John F Kielkopf. New approximation to the Voigt function with applications to spectral-line profile analysis. *JOSA*, 63(8):987–995, 1973. doi: 10.1364/JOSA.63.000987.

[467] Franz Schreier. Optimized implementations of rational approximations for the Voigt and complex error function. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 112(6):1010–1025, 2011. doi: 10.1016/j.jqsrt.2010.12.010.

[468] John J Olivero and RL Longbothum. Empirical fits to the Voigt line width: A brief review. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 17(2):233–236, 1977. doi: 10.1016/0022-4073(77)90161-3.

[469] Yuyan Liu, Jieli Lin, Guangming Huang, Yuanqing Guo, and Chuanxi Duan. Simple empirical analytical approximation to the Voigt profile. *JOSA B*, 18 (5):666–672, 2001. doi: 10.1364/JOSAB.18.000666.

[470] Héctor Oscar Di Rocco and Alicia Cruzado. The Voigt profile as a sum of a Gaussian and a Lorentzian functions, when the weight coefficient depends only on the widths ratio. *Acta Physica Polonica A*, 122:666–669, 2012. doi: 10.12693/APhysPolA.122.666.

[471] C. Hill, S. N. Yurchenko, and J. Tennyson. Temperature-dependent molecular absorption cross sections for exoplanets and other atmospheres. *Icarus*, 226: 1673–1677, 2013. doi: 10.1016/j.icarus.2012.07.028.

[472] Jingxin Zhang, J. Tennyson, and Sergei N. Yurchenko. PyExoCross: a Python program for generating spectra and cross-sections from molecular line lists. *RASTI*, 3:257–287, 2024. doi: 10.1093/rasti/rzae016.

[473] C. M. Sharp and A. Burrows. Atomic and molecular opacities for brown dwarf and giant planet atmospheres. *ApJS*, 168:140–166, 2007. doi: 10.1086/508708.

[474] M Abramowitz and IA Stegun. Handbook of Mathematical Functions 10th Printing with Corrections. *Natl. Bur. Stand. Appl. Math. Ser*, 55, 1972.

[475] K. L. Chubb, M. Rocchetto, S. N. Yurchenko, M. Min, I. Waldmann, J. K. Barstow, P. Molliere, A. F. Al-Refaie, M. Phillips, and J. Tennyson. The ExoMolOP Database: Cross-sections and K-tables for Molecules of Interest in High-Temperature Exoplanet Atmospheres. *A&A*, 646:A21, 2021. doi: 10.1051/0004-6361/202038350.

[476] Manuel López-Puertas and Fredric William Taylor. *Non-LTE radiative transfer in the Atmosphere*. World Scientific, 2001. doi: 10.1142/4650.

[477] M.A. López-Valverde, Pierre Drossart, Robert Carlson, R. Mehlman, and Maarten Roos-Serote. Non-LTE infrared observations at Venus: From NIMS/Galileo to VIRTIS/Venus Express. *Planetary and Space Science*, 55(12): 1757–1771, 2007. ISSN 0032-0633. doi: 10.1016/j.pss.2007.01.008.

[478] M.A. López-Valverde, M. López-Puertas, J.J. López-Moreno, V. Formisano, D. Grassi, A. Maturilli, E. Lellouch, and P. Drossart. Analysis of non-LTE emissions at in the Martian atmosphere as observed by PFS/Mars Express and SWS/ISO. *Planetary and Space Science*, 53(10):1079–1087, 2005. ISSN 0032-0633. doi: 10.1016/j.pss.2005.03.007.

[479] Sang Joon Kim, Chae Kyung Sim, Jin Ho, Thomas R. Geballe, Yuk L. Yung, Steve Miller, and Yong Ha Kim. Hot CH4 in the polar regions of Jupiter. *Icarus*, 257:217–220, 2015. ISSN 0019-1035. doi: 10.1016/j.icarus.2015.05.008.

[480] Sam O M Wright, Ingo Waldmann, and Sergei N Yurchenko. Non-local thermal equilibrium spectra of atmospheric molecules for exoplanets. *MNRAS*, 512:2911–2924, 2022. doi: 10.1093/mnras/stac654.

[481] H. A. Weaver and M. J. Mumma. Infrared molecular emissions from comets. *The Astrophysical Journal*, 276:782, 1984. ISSN 1538-4357. doi: 10.1086/161664.

[482] F. F. S. van der Tak, J. H. Black, F. L. Schöier, D. J. Jansen, and E. F. van Dishoeck. A computer program for fast non-LTE analysis of interstellar line spectra: With diagnostic plots to interpret observed line intensity ratios. *Astronomy & Astrophysics*, 468(2):627–635, 2007. ISSN 1432-0746. doi: 10.1051/0004-6361:20066820.

[483] F. Lique, F. F. S. van der Tak, J. Kłos, J. Bulthuis, and M. H. Alexander. The importance of non-LTE models for the interpretation of observations of interstellar NO. *Astronomy & Astrophysics*, 493(2):557–563, 2008. ISSN 1432-0746. doi: 10.1051/0004-6361:200810453.

[484] Edward Arthur Milne. *Handbuch der Astrophysik: Band III / Zweite Hälfte Grundlagen der Astrophysik Dritter Teil*. Springer Berlin Heidelberg, 1930. ISBN 9783642907067. doi: 10.1007/978-3-642-90706-7.

[485] A. R. Curtis and R. M. Goody. Thermal radiation in the upper atmosphere. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 236(1205):193–206, 1956. ISSN 2053-9169. doi: 10.1098/rspa.1956.0128.

[486] J. T. Houghton. Absorption and emission by carbon-dioxide in the meso-

sphere. *Quarterly Journal of the Royal Meteorological Society*, 95(403): 1–20, 1969. ISSN 1477-870X. doi: 10.1002/qj.49709540302.

[487] William R. Kuhn and Julius London. Infrared Radiative Cooling in the Middle Atmosphere (30–110 km). *Journal of the Atmospheric Sciences*, 26(2):189–204, 1969. ISSN 1520-0469. doi: 10.1175/1520-0469(1969)026¡0189:IRCITM¿2.0.CO;2.

[488] Robert E. Dickinson. Infrared Radiative Heating and Cooling in the Venusian Mesosphere.I: Global Mean Radiative Equilibrium. *Journal of the Atmospheric Sciences*, 29(8):1531–1556, 1972. ISSN 1520-0469. doi: 10.1175/1520-0469(1972)029¡1531:IRHACI¿2.0.CO;2.

[489] John B. Kumer and Thomas C. James. CO2(001) and N2vibrational temperatures in the $50 \lesssim z \lesssim 130$ km altitude range. *Journal of Geophysical Research*, 79(4):638–648, 1974. ISSN 0148-0227. doi: 10.1029/ja079i004p00638.

[490] M López-Puertas, R Rodrigo, A Molina, and F.W Taylor. A non-LTE radiative transfer model for infrared bands in the middle atmosphere. I. Theoretical basis and application to CO2 15 μm bands. *Journal of Atmospheric and Terrestrial Physics*, 48(8):729–748, 1986. ISSN 0021-9169. doi: 10.1016/0021-9169(86)90022-x.

[491] M. López-Puertas, R. Rodrigo, J.J. López-Moreno, and F.W. Taylor. A non-LTE radiative transfer model for infrared bands in the middle atmosphere. II. CO2 (2.7 and 4.3 μm) and water vapour (6.3 μm) bands and N2(1) and O2(1) vibrational levels. *Journal of Atmospheric and Terrestrial Physics*, 48 (8):749–764, 1986. ISSN 0021-9169. doi: 10.1016/0021-9169(86)90023-1.

[492] Peter P. Wintersteiner, Richard H. Picard, Ramesh D. Sharma, Jeremy R. Winick, and Robert A. Joseph. Line-by-line radiative excitation model for the non-equilibrium atmosphere: Application to CO2 15-μm emission. *Journal of Geophysical Research: Atmospheres*, 97(D16):18083–18117, 1992. ISSN 0148-0227. doi: 10.1029/92jd01494.

[493] Moshe Elitzur. *Astronomical Masers*. Springer Netherlands, 1992. ISBN 9789401123945. doi: 10.1007/978-94-011-2394-5.

[494] Malcolm Gray. *Maser Sources in Astrophysics*. Cambridge University Press, 2012. ISBN 9780511977534. doi: 10.1017/CBO9780511977534.

[495] Mark J. Reid and James M. Moran. Masers. *Annual Review of Astronomy and Astrophysics*, 19(1):231–276, 1981. ISSN 1545-4282. doi: 10.1146/annurev.aa.19.090181.001311.

[496] E. M. L. Humphreys. Submillimeter and millimeter masers. *Proceedings of the International Astronomical Union*, 3(S242):471–480, 2007. doi: 10.1017/S1743921307013622.

[497] C. E. Treanor, J. W. Rich, and R. G. Rehm. Vibrational Relaxation of Anharmonic Oscillators with Exchange-Dominated Collisions. *J. Chem. Phys.*, 48 (4):1798–1807, 1968. doi: 10.1063/1.1668914.

[498] Lev Davidovich Landau and Evgenii Mikhailovich Lifshitz. *Statistical Physics: Volume 5*, volume 5. Elsevier, 2013.

[499] Michael J. Berry. Golden rule calculation of reaction product vibronic state distributions. *Chem. Phys. Lett.*, 27(1):73 – 77, 1974. ISSN 0009-2614. doi: https://doi.org/10.1016/0009-2614(74)80447-1.

[500] C. E. Treanor, J. W. Rich, and R. G. Rehm. Vibrational Relaxation of Anharmonic Oscillators with Exchange-Dominated Collisions. *The Journal of Chemical Physics*, 48(4):1798–1807, 1968. ISSN 1089-7690. doi: 10.1063/1.1668914.

[501] M López-Puertas and F W Taylor. *Non-LTE Radiative Transfer in the Atmosphere*. WORLD SCIENTIFIC, 2001. ISBN 9789812811493. doi: 10.1142/4650.

[502] T. Furtenbacher, A. G. Császár, and J. Tennyson. MARVEL: measured active rotational-vibrational energy levels. *J. Mol. Spectrosc.*, 245:115–125, 2007. doi: 10.1016/j.jms.2007.07.005.

[503] J. Tennyson, T. Furtenbacher, S. N. Yurchenko, and A. G. Császár. Empirical rovibrational energy levels for nitrous oxide. *J. Quant. Spectrosc. Radiat. Transf.*, 316:108902, 2024. doi: 10.1016/j.jqsrt.2024.108902.

[504] L. K. McKemmish, C. A. Bowesman, K. Kefala, A. N. Perri, A. M. Syme, S. N. Yurchenko, and J. Tennyson. A hybrid approach to generating diatomic line lists for high resolution studies of exoplanets and other hot astronomical objects: Updates to ExoMol MgO, VO and TiO line lists. *RASTI*, 3:565–583, 2024. doi: 10.1093/rasti/rzae037.

[505] L. K. McKemmish, A. M. Syme, J. Borsovszky, S. N. Yurchenko, J Tennyson, T. Furtenbacher, and A. G. Császár. Incorportating new experiments in diatomic spectral databases: An update to the $^{12}C_2$ MARVEL database and ExoMol line list of $^{12}C_2$. *MNRAS*, 497:1081–1097, 2020. doi: 10.1093/mnras/staa1954.

[506] C. A. Bowesman, H. Akbari, S. Hopkins, S. N. Yurchenko, and J. Tennyson. Fine and hyperfine resolved empirical energy levels for VO. *J. Quant. Spectrosc. Radiat. Transf.*, 289:108295, 2022. doi: 10.1016/j.jqsrt.2022.108295.

[507] C. A. Bowesman, I. I. Mizus, N. F. Zobov, O. L. Polyansky, J. Sarka, B. Poirier, M. Pezzella, S. N. Yurchenko, and J. Tennyson. ExoMol line lists - L: High-resolution line lists of H_3^+, H_2D^+, D_2H^+ and D_3^+. *MNRAS*, 519:6333–6348, 2023. doi: 10.1093/mnras/stad050.

[508] A. R. Al-Derzi, S. N. Yurchenko, J. Tennyson, M. Melosso, Ningjing Jiang, C. Puzzarini, L. Dore, T. Furtenbacher, R. Tobias, and A G. Császár. An improved rovibrational linelist of formaldehyde, $H_2\,^{12}C^{16}O$. *J. Quant. Spectrosc. Radiat. Transf.*, 266:107563, 2021. doi: 10.1016/j.jqsrt.2021.107563.

[509] M. Germann, A. Hjalten, J. Tennyson, S. N. Yurchenko, I. E. Gordon, C. Pett, I. Silander, K. Krzempek, A. Hudzikowski, A. Gluszek, G. Sobon, and A. Foltynowicz. Optical frequency comb Fourier transform spectroscopy of formaldehyde in the 1250 to 1390 cm^{-1} range: experimental line list and MARVEL analysis. *J. Quant. Spectrosc. Radiat. Transf.*, 312:108782, 2024. doi: 10.1016/j.jqsrt.2023.108782.

[510] K. Kefala, V. Boudon, S. N. Yurchenko, and J. Tennyson. Empirical rovibrational energy levels for methane. *J. Quant. Spectrosc. Radiat. Transf.*, 316: 108897, 2024. doi: 10.1016/j.jqsrt.2024.108897.

[511] J. Tennyson. The ExoMol project: Molecular opacity calculations at University College London. In Claudio Mendoza, Sylvaine Turck-Chiéze, and James Colgan, editors, *Workshop on Astrophysical Opacities*, volume 515 of *ASP*, pages 137–144. Astronomical Society of the Pacific, 2018.

[512] T Furtenbacher. Marvelonline: Marvel2.0, 2020.

[513] Armando N. Perri, Fadia Taher, and Laura K. McKemmish. Full spectroscopic model and trihybrid experimental-perturbative-variational line list for ZrO. *MNRAS*, 524:4631–4641, 2023. doi: 10.1093/mnras/stad2103.

[514] Armando N. Perri and Laura K. McKemmish. Full spectroscopic model and trihybrid experimental-perturbative-variational line list for NH. *MNRAS*, 531 (3):3023–3033, 2024. doi: 10.1093/mnras/stae1340.

[515] J. K. G. Watson. Higher-order vibration-rotation energies of the X_3 molecule. *J. Mol. Spectrosc.*, 103:350–363, 1984. doi: 10.1016/0022-2852(84)90062-6.

[516] E. R. Guest, J. Tennyson, and S. N. Yurchenko. Modelling the Rotational Dependence of Line Broadening using Machine Learning. *J. Mol. Spectrosc.*, 401:111901, 2024. doi: 10.1016/j.jms.2024.111901.

[517] J. Buldyreva, K. Stehlin, S. N. Yurchenko, E. R. Guest, and J. Tennyson. Semi-classical estimates of pressure-induced linewidths for infrared absorption by hot (exo)planetary atmospheres. *ApJS*, 276:23, 2025. doi: 10.3847/1538-4365/ad9b19.

[518] S. N. Yurchenko, Charles A. Bowesman, Ryan P. Brady, Elizabeth R. Guest, Kyraki Kefala, Georgi B. Mitev, Alec Owens, Armando N. Perri, Marco Pezzella, Oleksiy Smola, Andrei Solokov, Jingxin Zhang, and J Tennyson. ExoMol Line Lists – LX: Molecular line list for the ammonia isotopologue $^{15}NH_3$. *MNRAS*, 533:3442–3456, 2024. doi: 10.1093/mnras/stae1849.

[519] S. N. Yurchenko, A. Owens, K. Kefala, and J. Tennyson. ExoMol line lists – LVII. High accuracy ro-vibrational line list for methane (CH_4). *MNRAS*, 528:3719–3729, 2024. doi: 10.1093/mnras/stae148.

[520] L. K. McKemmish, T. Masseron, J Hoeijmakers, V. V. Pérez-Mesa, S. L. Grimm, S. N. Yurchenko, and J. Tennyson. ExoMol Molecular line lists – XXXIII. The spectrum of Titanium Oxide. *MNRAS*, 488:2836–2854, 2019. doi: 10.1093/mnras/stz1818.

[521] C. A. Bowesman, Qianwei Qu, L. K. McKemmish, S. N. Yurchenko, and J. Tennyson. ExoMol line lists – LV: Hyperfine-resolved molecular line list for vanadium monoxide ($^{51}V^{16}O$). *MNRAS*, 529:1321–1332, 2024. doi: 10.1093/mnras/stae542.

[522] Michael L. Waskom. seaborn: statistical data visualization. *Journal of Open Source Software*, 6(60):3021, 2021. doi: 10.21105/joss.03021.

[523] VAMDC Consortium. https://www.vamdc.org, 2009. Accessed: 2025-01-25.

[524] A. Kramida, Yu. Ralchenko, J. Reader, and and NIST ASD Team. Development of NIST Atomic Databases and Online Tools. *Atoms*, 8:56, 2020. doi: 10.3390/atoms8030056.

[525] A. Kramida, Yu. Ralchenko, J. Reader, and and NIST ASD Team. NIST Atomic Spectra Database (ver. 5.3), [Online]. Available: `http://physics.nist.gov/asd` [2017, April 2]. National Institute of Standards and Technology, Gaithersburg, MD., 2015.

[526] S. N. Yurchenko, E. Nogué, A. A. A. Azzam, and J. Tennyson. ExoMol line lists – IL. Rovibronic spectrum of aluminium monochloride (AlCl). *MNRAS*, 520:5183–5191, 2023. doi: 10.1093/mnras/stac3757.

[527] D. Darby-Lewis, J. Tennyson, S. N. Yurchenko, and K. D. Lawson. Vibrational resolved electron impact electronic excitation of BeH. *J. Phys. B-At. Mol. Opt. Phys.*, 53:135202, 2020. doi: 10.1088/1361-6455/ab87e9.

[528] K L Chubb, J. Tennyson, and S. N. Yurchenko. ExoMol Molecular linelists – XXXVII: spectra of acetylene. *MNRAS*, 493:1531–1545, 2020. doi: 10.1093/mnras/staa229.

[529] A. Owens, A. Mitrushchenkov, S. N. Yurchenko, and J. Tennyson. ExoMol line lists – XLVII. Rovibronic molecular line list of the calcium monohydroxide radical (CaOH). *MNRAS*, 516:3995–4002, 2022. doi: 10.1093/mnras/stac2462.

[530] A. F. Al-Refaie, S. N. Yurchenko, A. Yachmenev, and J. Tennyson. ExoMol line lists - VIII: A variationally computed line list for hot formaldehyde. *MNRAS*, 448:1704–1714, 2015. doi: 10.1093/mnras/stv091.

[531] T Mellor, A. Owens, S. N. Yurchenko, and J. Tennyson. ExoMol line lists – LI. Rvibronic molecular line list for thioformaldehyde (H_2CS). *MNRAS*, 520:1997–2008, 2022. doi: 10.1093/mnras/stad111.

[532] A. A. A. Azzam, S. N. Yurchenko, J. Tennyson, and O. V. Naumenko. ExoMol line lists XVI: A Hot Line List for H_2S. *MNRAS*, 460:4063–4074, 2016. doi: 10.1093/mnras/stw1133.

[533] S. N. Yurchenko, J. Tennyson, Steve Miller, V. V. Melnikov, J. O'Donoghue, and L. Moore. ExoMol molecular line lists XL: Ro-vibrational molecular line list for the hydronium ion (H_3O^+). *MNRAS*, 497:2340–2351, 2020. doi: 10.1093/mnras/staa2034.

[534] Alec Owens, Sam O M Wright, Yakiv Pavlenko, Alexander Mitrushchenkov, Jacek Koput, Sergei N Yurchenko, and Jonathan Tennyson. ExoMol line lists - LI. Molecular line list for lithium hydroxide (LiOH). *MNRAS*, 527: 731–738, 2024. doi: 10.1093/mnras/stad3226.

[535] S. N. Yurchenko, T Mellor, and J Tennyson. ExoMol line lists – LIX. High-temperature line list for N_2O. *MNRAS*, 534:1364–1375, 2024. doi: 10.1093/mnras/stae2201.

[536] P. A. Coles, Sergei N. Yurchenko, and Jonathan Tennyson. ExoMol molecular line lists XXXV: a rotation-vibration line list for hot ammonia. *MNRAS*, 490:4638–4647, 2019. doi: 10.1093/mnras/stz2778.

[537] A. Owens, , S. N. Yurchenko, and J. Tennyson. ExoMol line lists - LVIII. High-temperature molecular line list of carbonyl sulphide (OCS). *MNRAS*, 530:4004–4015, 2024. doi: 10.1093/mnras/stae1110.

[538] Mikhail Semenov, Nayla El-Kork, Sergei N. Yurchenko, and Jonathan Tennyson. ExoMol Line Lists – LXIV: Empirical rovibronic spectra of phosphorous mononitride (PN) covering the IR and UV regions. *MNRAS*, 536: 714–727, 2025. doi: 10.1093/mnras/stae2610.

[539] R. P. Brady, S. N. Yurchenko, J. Tennyson, and G.-S. Kim. ExoMol line lists – LVI. The SO line list, MARVEL analysis of experimental transition data and refinement of the spectroscopic model. *MNRAS*, 527:6675–6690, 2024. doi: 10.1093/mnras/stad3508.

[540] M. Semenov, J. Tennyson, and S. N. Yurchenko. ExoMol line lists – XLVI: Empirical rovibronic spectra of silicon mononitrate (SiN) covering 6 lowest

electronic states and 4 isotopologues. *MNRAS*, 516:1158–1169, 2022. doi: 10.1093/mnras/stac2004.

[541] S. N. Yurchenko, J. Tennyson, A.-M. Syme, A. Y. Adam, V. H. J. Clark, B. Cooper, C. P. Dobney, S. T. E. Donnelly, M. N. Gorman, A. E. Lynas-Gray, T. Meltzer, A. Owens, Q. Qu, M. Semenov, W. Somogyi, A. Upadhyay, S. Wright, and J. C. Zapata Trujillo. ExoMol line lists – XLIV. Infrared and ultraviolet line list for silicon monoxide (^{28}Si^{16}O). *MNRAS*, 510:903–919, 2022. doi: 10.1093/mnras/stab3267.

[542] S. N. Yurchenko, R. P. Brady, J. Tennyson, A. N. Smirnov, O. A. Vasilyev, and V. G. Solomonik. ExoMol line lists – LIII. Empirical Rovibronic spectra Yttrium Oxide (YO). *MNRAS*, 527:4899–4912, 2024. doi: 10.1093/mnras/stad3225.