

cortecs: A Python package for compressing opacities

Arjun B. Savel  ^{1,2}¶, Megan Bedell  ², and Eliza M.-R. Kempton  ¹

¹ Astronomy Department, University of Maryland, College Park, 4296 Stadium Dr., College Park, MD 207842 USA ² Flatiron Institute, Simons Foundation, 162 Fifth Avenue, New York, NY 10010, USA ¶ Corresponding author

DOI: [10.21105/joss.06104](https://doi.org/10.21105/joss.06104)

Software

- [Review](#) ↗
- [Repository](#) ↗
- [Archive](#) ↗

Editor: Ivelina Momcheva 

Reviewers:

- [@pscicluna](#)
- [@doriannblain](#)

Submitted: 03 November 2023

Published: 05 August 2024

License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC BY 4.0](#)).

Summary

The absorption and emission of light by exoplanet atmospheres encode details of atmospheric composition, temperature, and dynamics. Fundamentally, simulating these processes requires detailed knowledge of the opacity of gases within an atmosphere. When modeling broad wavelength ranges at high resolution, such opacity data, for even a single gas, can take up multiple gigabytes of system random-access memory (RAM). This aspect can be a limiting factor when considering the number of gases to include in a simulation, the sampling strategy used for inference, or even the architecture of the system used for calculations. Here, we present cortecs, a Python tool for compressing opacity data. cortecs provides flexible methods for fitting the temperature, pressure, and wavelength dependencies of opacity data and for evaluating the opacity with accelerated, GPU-friendly methods. The package is actively developed on GitHub (<https://github.com/arjunsavel/cortecs>), and it is available for download with pip and conda.

Statement of need

Observations with the latest high-resolution spectrographs (e.g., IGRINS / Gemini South, ESPRESSO / VLT, MAROON-X / Gemini North; Mace et al. (2018); Seifahrt et al. (2020); Pepe et al. (2021)) have motivated RAM-intensive simulations of exoplanet atmospheres at high spectral resolution. cortecs enables these simulations with more gases and on a broader range of computing architectures by compressing opacity data.

Broadly, generating a spectrum to compare against recent high-resolution data requires solving the radiative transfer equation over tens of thousands of wavelength points (e.g., Beltz et al., 2023; Gandhi et al., 2023; Line et al., 2021; Maguire et al., 2023; Prinoth et al., 2023; Savel et al., 2022; Wardenier et al., 2023). To decrease computational runtime, some codes have parallelized the problem on GPUs (e.g., Lee et al., 2022; Line et al., 2021). However, GPUs cannot in general hold large amounts of data in their video random-access memory (VRAM) (e.g., Ito et al., 2017); only the cutting-edge, most expensive GPUs are equipped with VRAM in excess of 30 GB (such as the NVIDIA A100 or H100). RAM and VRAM management is therefore a clear concern when producing high-resolution spectra.

How do we decrease the RAM footprint of these calculations? By far the largest contributor to the RAM footprint, at least as measured on disk, is the opacity data. For instance, the opacity data for a single gas species across the wavelength range of the Immersion GRating INfrared Spectrometer spectrograph (IGRINS, Mace et al., 2018) takes up 2.1 GB of non-volatile memory (i.e., the file size is 2.1 GB) at float64 precision and at a resolving power of 400,000 (as used in Line et al. (2021); with 39 temperature points and 18 pressure points, using, e.g., the Polyansky et al. (2018) water opacity tables). In many cases, not all wavelengths need to be loaded, e.g. if the user is down-sampling the resolution of their opacity function. Even so, it stands to reason that decreasing the amount of RAM/VRAM consumed by opacity data

would strongly decrease the total amount of RAM/VRAM consumed by the radiative transfer calculation.

One solution is to isolate redundancy: While the wavelength dependence of opacity is sharp for many gases, the temperature and pressure dependencies are generally smooth and similar across wavelengths (e.g., Barber et al., 2014; Coles et al., 2019; Polyansky et al., 2018). This feature implies that the opacity data should be compressible without significant loss of accuracy at the spectrum level.

While our benchmark case (see the Benchmark section below) demonstrates the applicability of cortecs to high-resolution opacity functions of molecular gases, the package is general and the compression/decompression steps of the package can be applied to any opacity data in HDF5 format that has pressure and temperature dependence, such as the opacity of neutral atoms or ions. Our benchmark only shows, however, that the amounts of error from our compression technique is reasonable in the spectra of exoplanet atmospheres at pressures greater than a microbar for a single composition. This caveat is important to note for a few reasons:

1. Based on error propagation, the error in the opacity function may be magnified in the spectrum based on the number of cells that are traced during radiative transfer. The number of spatial cells used to simulate exoplanet atmospheres (in our case, 100) is small enough that the cortecs error is not large at the spectrum level.
2. Exoplanet atmospheres are often modeled in hydrostatic equilibrium at pressures greater than a microbar (e.g., Barstow et al., 2020; Showman et al., 2020). When modeling atmospheres in hydrostatic equilibrium, the final spectrum essentially maps to the altitude at which the gas becomes optically thick. If cortecs-compressed opacities were used to model an optically thin gas over large path lengths, however, then smaller opacities would be more important. cortecs tends to perform worse at modeling opacity functions that jump from very low to very high opacities, so it may not perform optimally in these optically thin scenarios.
3. The program may perform poorly for opacity functions with sharp features in their temperature-pressure dependence (e.g., the Lyman series transitions of hydrogen, Kurucz, 2017). That is, the data may require so many parameters to be fit that the compression is no longer worthwhile.

Methods

cortecs seeks to compress redundant information by representing opacity data not as the opacity itself but as fits to the opacity as a function of temperature and pressure. We generally refer to this process as *compression* as opposed to *fitting* to emphasize that we do not seek to construct physically motivated, predictive, or comprehensible models of the opacity function. Rather, we simply seek representations of the opacity function that consume less RAM/VRAM. The compression methods we use are *lossy* — the original opacity data cannot be exactly recovered with our methods. We find that the loss of accuracy is tolerable for at least the hot Jupiter emission spectroscopy application (see Benchmark below).

We provide three methods of increasing complexity (and flexibility) for compressing and decompressing opacity: polynomial-fitting, principal components analysis (PCA, e.g., Jolliffe & Cadima, 2016) and neural networks (e.g., Alzubaidi et al., 2021). The default neural network architecture is a fully connected neural network; the user can specify the desired hyperparameters, such as number of layers, neurons per layer, and activation function. Alternatively, any keras model (Chollet, 2015) can be passed to the fitter. Each compression method is paired with a decompression method for evaluating opacity as a function of temperature, pressure, and wavelength. These decompression methods are tailored for GPUs and are accelerated with the JAX code transformation framework (Bradbury et al., 2018). An example of this reconstruction is shown in Figure 1. In the figure, opacities less than 10^{-60} are ignored. This is because, to become optically thick at a pressure of 1 bar and temperature of 1000 K, a column would need

to be nearly 10^{34} m long. Here we show a brief derivation of this. The length of the column, ds is $ds = \frac{\tau}{\alpha}$, where τ is the optical depth, and α is the absorption coefficient. Setting $\tau = 1$, we have $ds = \frac{1}{\alpha}$. The absorption coefficient is the product of the opacity and the density of the gas: $ds = \frac{1}{\kappa_\lambda \rho}$. Therefore, $ds = \frac{1}{\kappa_\lambda \rho}$. The density of the gas ρ is the pressure divided by the product of the temperature and the gas constant: $\rho = \frac{P}{k_B T \mu}$ for mean molecular weight μ . This leads to the final equation for the column length: $ds = \frac{k_B T \mu}{P \kappa_\lambda}$. For CO, the mean molecular weight is 28.01 g/mol. Plugging in, we arrive at $ds \approx 10^{34}$ m (roughly 10^{17} parsecs) for $\kappa_\lambda = 10^{-33}$ cm²/g, which is equivalent to roughly a cross-section of $\sigma_\lambda = 10^{-60}$ m².

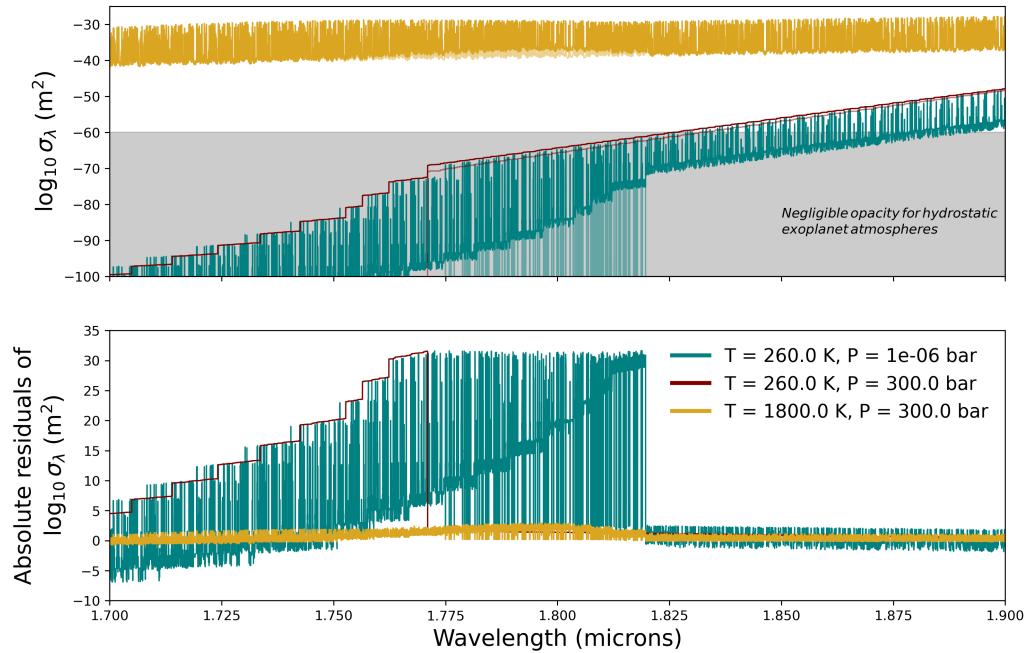


Figure 1: Top panel: The original opacity function of CO (Rothman et al., 2010) (solid lines) and its cortecs reconstruction (transparent lines) over a large wavelength range and at multiple temperatures and pressures. Bottom panel: the absolute residuals between the opacity function and its cortecs reconstruction. σ_λ is the opacity, in units of square meters. We cut off the opacity at 10^{-104} , explaining the shape of the residuals in teal and dark red. Note that opacities less than 10^{-60} are not generally relevant for the benchmark presented here; an opacity of $\sigma_\lambda = 10^{-60}$ would require a column nearly 10^{34} m long to become optically thick at a pressure of 1 bar and temperature of 1000 K.

Workflow

A typical workflow with cortecs involves the following steps:

1. Acquiring: Download opacity data from a source such as the ExoMol database (Tennyson et al., 2016) or the HITRAN database (Gordon et al., 2017).
2. Fitting: Compress the opacity data with cortecs's fit methods.
3. Saving: Save the compressed opacity data (the fitted parameters) to disk.
4. Loading: Load the compressed opacity data from disk in whatever program you're applying the data—e.g., within your radiative transfer code.
5. Decompressing: Evaluate the opacity with cortecs's eval methods.

The accuracy of these fits may or may not be suitable for a given application. It is important to test that the error incurred using cortecs does not impact the results of your application—for instance, by using the `cortecs.fit.metrics.calc_metrics` function to calculate the error

incurred by the compression and by calculating spectra with and without using cortecs-compressed opacities. We provide an example of such a benchmarking exercise below.

Benchmark: High-resolution retrieval of WASP-77Ab

As a proof of concept, we perform a parameter inference exercise (a “retrieval,” [Madhusudhan & Seager, 2009](#)) on the high-resolution thermal emission spectrum of the fiducial hot Jupiter WASP-77Ab ([August et al., 2023](#); [Line et al., 2021](#); [Mansfield et al., 2022](#)) as observed at IGRINS. The retrieval pairs PyMultiNest ([Buchner et al., 2014](#)) sampling with the CHIMERA radiative transfer code ([Line et al., 2013](#)), with opacity from H₂O ([Polyansky et al., 2018](#)), CO ([Rothman et al., 2010](#)), CH₄ ([Hargreaves et al., 2020](#)), NH₃ ([Coles et al., 2019](#)), HCN ([Barber et al., 2014](#)), H₂S ([Azzam et al., 2016](#)), and H₂ – H₂ collision-induced absorption ([Karman et al., 2019](#)). The non-compressed retrieval uses the data and retrieval framework from ([Line et al., 2021](#)), run in an upcoming article (Savel et al. 2024, submitted). For this experiment, we use the PCA-based compression scheme implemented in cortecs, preserving 2 principal components and their corresponding weights as a function for each wavelength as a lossy compression of the original opacity data.

Using cortecs, we compress the input opacity files by a factor of 13. These opacity data (as described earlier in the paper) were originally stored as 2.1 GB .h5 files containing 39 temperature points, 18 pressure points, and 373,260 wavelength points. The compressed opacity data are stored as a 143.1 MB .npz file, including the PCA coefficients and PCA vectors (which are reused for each wavelength point). These on-disk memory quotes are relatively faithful to the in-memory RAM footprint of the data when stored as numpy arrays (2.1 GB for the uncompressed data and 160 MB for the compressed data). Reading in the original files takes 1.1 ± 0.1 seconds, while reading in the compressed files takes 24.4 ± 0.3 ms. Accessing/evaluating a single opacity value takes 174.0 ± 0.5 ns for the uncompressed data and 789 ± 5 ns for the compressed data. All of these timing experiments are performed on a 2021 MacBook Pro with an Apple M1 Pro chip and 16 GB of RAM.

Importantly, we find that our compressed-opacity retrieval yields posterior distributions (as plotted by the corner package, [Foreman-Mackey, 2016](#)) and Bayesian evidences that are consistent with those from the benchmark retrieval using uncompressed opacity ([Figure 2](#)) within a comparable runtime. The two posterior distributions exhibit slightly different substructure, which we attribute to the compressed results requiring 10% more samples to converge (about 5 hours of extra runtime on a roughly 2 day-long calculation) and residual differences between the compressed and uncompressed opacities. The results from this exercise indicate that our compression/decompression scheme is accurate enough to be used in at least some high-resolution retrievals.

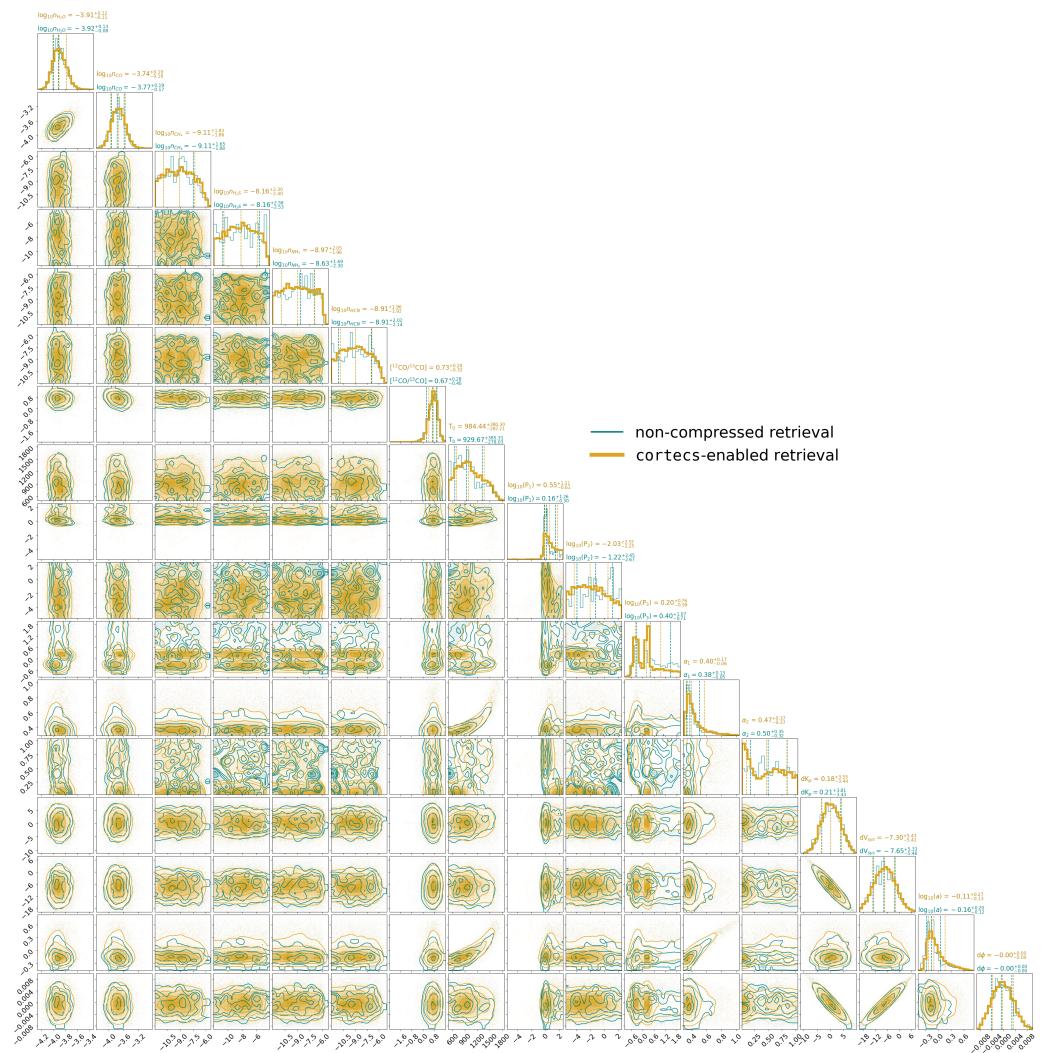


Figure 2: The posterior distributions for our baseline WASP-77Ab retrieval (teal) and our retrieval using opacities compressed by cortecs (gold).

Method	Compression factor	Median absolute deviation	Compression time (s)	Decompression time (s)
PCA	13	0.30	2.6×10^1	2.3×10^2
Polynomials	44	0.24	7.8×10^2	3.6×10^3
Neural network	9	2.6	1.4×10^7	3.6×10^4

Comparison of compression methods used for the full HITEMP CO line list ([Rothman et al., 2010](#)) over the IGRINS wavelength range at a resolving power of 250,000, cumulative for all data points. Note that the neural network compression performance and timings are only assessed at a single wavelength point and extrapolated over the full wavelength range.

Acknowledgements

A.B.S. and E.M-R.K. acknowledge support from the Heising-Simons Foundation. We thank Max Isi for helpful discussions.

References

- Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., & Farhan, L. (2021). Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, 8, 1–74. <https://doi.org/10.1186/s40537-021-00444-8>
- August, P. C., Bean, J. L., Zhang, M., Lunine, J., Xue, Q., Line, M., & Smith, P. C. B. (2023). Confirmation of Subsolar Metallicity for WASP-77Ab from JWST Thermal Emission Spectroscopy. 953(2), L24. <https://doi.org/10.3847/2041-8213/ace828>
- Azzam, A. A., Tennyson, J., Yurchenko, S. N., & Naumenko, O. V. (2016). ExoMol molecular line lists–XVI. The rotation–vibration spectrum of hot H₂S. *Monthly Notices of the Royal Astronomical Society*, 460(4), 4063–4074. <https://doi.org/10.1093/mnras/stw1133>
- Barber, R., Strange, J., Hill, C., Polyansky, O., Mellau, G. C., Yurchenko, S., & Tennyson, J. (2014). ExoMol line lists–III. An improved hot rotation–vibration line list for HCN and HNC. *Monthly Notices of the Royal Astronomical Society*, 437(2), 1828–1835. <https://doi.org/10.1093/mnras/stt2011>
- Barstow, J. K., Changeat, Q., Garland, R., Line, M. R., Rocchetto, M., & Waldmann, I. P. (2020). A comparison of exoplanet spectroscopic retrieval tools. *Monthly Notices of the Royal Astronomical Society*, 493(4), 4884–4909. <https://doi.org/10.1093/mnras/staa548>
- Beltz, H., Rauscher, E., Kempton, E. M.-R., Malsky, I., & Savel, A. B. (2023). Magnetic effects and 3D structure in theoretical high-resolution transmission spectra of ultrahot jupiters: The case of WASP-76b. *The Astronomical Journal*, 165(6), 257. <https://doi.org/10.3847/1538-3881/acd24d>
- Bradbury, J., Frostig, R., Hawkins, P., Johnson, M. J., Leary, C., Maclaurin, D., Necula, G., Paszke, A., VanderPlas, J., Wanderman-Milne, S., & Zhang, Q. (2018). JAX: Composable transformations of Python+NumPy programs (Version 0.3.13). <http://github.com/google/jax>
- Buchner, J., Georgakakis, A., Nandra, K., Hsu, L., Rangel, C., Brightman, M., Merloni, A., Salvato, M., Donley, J., & Kocevski, D. (2014). X-ray spectral modelling of the AGN obscuring region in the CDFS: Bayesian model selection and catalogue. *Astronomy & Astrophysics*, 564, A125. <https://doi.org/10.1051/0004-6361/201322971>
- Chollet, F. (2015). Keras. <https://keras.io>.
- Coles, P. A., Yurchenko, S. N., & Tennyson, J. (2019). ExoMol molecular line lists–XXXV. A rotation–vibration line list for hot ammonia. *Monthly Notices of the Royal Astronomical Society*, 490(4), 4638–4647. <https://doi.org/10.1093/mnras/stz2778>
- Foreman-Mackey, D. (2016). Corner.py: Scatterplot matrices in python. *The Journal of Open Source Software*, 1(2), 24. <https://doi.org/10.21105/joss.00024>
- Gandhi, S., Kesseli, A., Zhang, Y., Louca, A., Snellen, I., Brogi, M., Miguel, Y., Casasayas-Barris, N., Pelletier, S., Landman, R., & others. (2023). Retrieval survey of metals in six ultrahot jupiters: Trends in chemistry, rain-out, ionization, and atmospheric dynamics. *The Astronomical Journal*, 165(6), 242. <https://doi.org/10.3847/1538-3881/accd65>
- Gordon, I. E., Rothman, L. S., Hill, C., Kochanov, R. V., Tan, Y., Bernath, P. F., Birk, M., Boudon, V., Campargue, A., Chance, K., & others. (2017). The HITRAN2016 molecular spectroscopic database. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 203, 3–69. <https://doi.org/10.1016/j.jqsrt.2017.06.038>
- Hargreaves, R. J., Gordon, I. E., Rey, M., Nikitin, A. V., Tyuterev, V. G., Kochanov, R. V., & Rothman, L. S. (2020). An accurate, extensive, and practical line list of methane for the HITEMP database. *The Astrophysical Journal Supplement Series*, 247(2), 55.

<https://doi.org/10.3847/1538-4365/ab7a1a>

- Ito, Y., Matsumiya, R., & Endo, T. (2017). Ooc_cuDNN: Accommodating convolutional neural networks over GPU memory capacity. *2017 IEEE International Conference on Big Data (Big Data)*, 183–192. <https://doi.org/10.1109/BigData.2017.8257926>
- Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: A review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2065), 20150202. <https://doi.org/10.1098/rsta.2015.0202>
- Karman, T., Gordon, I. E., Der Avoird, A. van, Baranov, Y. I., Boulet, C., Drouin, B. J., Groenenboom, G. C., Gustafsson, M., Hartmann, J.-M., Kurucz, R. L., & others. (2019). Update of the HITRAN collision-induced absorption section. *Icarus*, 328, 160–175. <https://doi.org/10.1016/j.icarus.2019.02.034>
- Kurucz, R. L. (2017). Including all the lines: Data releases for spectra and opacities. *Canadian Journal of Physics*, 95(9), 825–827. <https://doi.org/10.1139/cjp-2016-0794>
- Lee, E. K., Wardenier, J. P., Printhon, B., Parmentier, V., Grimm, S. L., Baeyens, R., Carone, L., Christie, D., Deitrick, R., Kitzmann, D., & others. (2022). 3D radiative transfer for exoplanet atmospheres. gCMCRT: A GPU-accelerated MCRT code. *The Astrophysical Journal*, 929(2), 180. <https://doi.org/10.3847/1538-4357/ac61d6>
- Line, M. R., Brogi, M., Bean, J. L., Gandhi, S., Zalesky, J., Parmentier, V., Smith, P., Mace, G. N., Mansfield, M., Kempton, E. M.-R., & others. (2021). A solar c/o and sub-solar metallicity in a hot jupiter atmosphere. *Nature*, 598(7882), 580–584. <https://doi.org/10.1038/s41586-021-03912-6>
- Line, M. R., Wolf, A. S., Zhang, X., Knutson, H., Kammer, J. A., Ellison, E., Deroo, P., Crisp, D., & Yung, Y. L. (2013). A systematic retrieval analysis of secondary eclipse spectra. I. A comparison of atmospheric retrieval techniques. *The Astrophysical Journal*, 775(2), 137. <https://doi.org/10.1088/0004-637X/775/2/137>
- Mace, G., Sokal, K., Lee, J.-J., Oh, H., Park, C., Lee, H., Good, J., MacQueen, P., Oh, J. S., Kaplan, K., & others. (2018). IGRINS at the discovery channel telescope and gemini south. *Ground-Based and Airborne Instrumentation for Astronomy VII*, 10702, 204–221. <https://doi.org/10.1117/12.2312345>
- Madhusudhan, N., & Seager, S. (2009). A temperature and abundance retrieval method for exoplanet atmospheres. *The Astrophysical Journal*, 707(1), 24. <https://doi.org/10.1088/0004-637X/707/1/24>
- Maguire, C., Gibson, N. P., Nugroho, S. K., Ramkumar, S., Fortune, M., Merritt, S. R., & Mooij, E. de. (2023). High-resolution atmospheric retrievals of WASP-121b transmission spectroscopy with ESPRESSO: Consistent relative abundance constraints across multiple epochs and instruments. *Monthly Notices of the Royal Astronomical Society*, 519(1), 1030–1048. <https://doi.org/10.1093/mnras/stac3388>
- Mansfield, M., Wiser, L., Stevenson, K. B., Smith, P., Line, M. R., Bean, J. L., Fortney, J. J., Parmentier, V., Kempton, E. M.-R., Arcangeli, J., & others. (2022). Confirmation of water absorption in the thermal emission spectrum of the hot jupiter WASP-77Ab with HST/WFC3. *The Astronomical Journal*, 163(6), 261. <https://doi.org/10.3847/1538-3881/ac658f>
- Pepe, F., Cristiani, S., Rebolo, R., Santos, N., Dekker, H., Cabral, A., Di Marcantonio, P., Figueira, P., Curto, G. L., Lovis, C., & others. (2021). ESPRESSO at VLT-on-sky performance and first results. *Astronomy & Astrophysics*, 645, A96. <https://doi.org/10.1051/0004-6361/202038306>
- Polyansky, O. L., Kyuberis, A. A., Zobov, N. F., Tennyson, J., Yurchenko, S. N., & Lodi, L. (2018). ExoMol molecular line lists XXX: A complete high-accuracy line list for

- water. *Monthly Notices of the Royal Astronomical Society*, 480(2), 2597–2608. <https://doi.org/10.1093/mnras/sty1877>
- Prinoth, B., Hoeijmakers, H. J., Pelletier, S., Kitzmann, D., Morris, B. M., Seifahrt, A., Kasper, D., Korhonen, H. H., Burheim, M., Bean, J. L., Benneke, B., Borsato, N. W., Brady, M., Grimm, S. L., Luque, R., Stürmer, J., & Thorsbro, B. (2023). Time-resolved transmission spectroscopy of the ultra-hot Jupiter WASP-189 b. 678, A182. <https://doi.org/10.1051/0004-6361/202347262>
- Rothman, L. S., Gordon, I., Barber, R., Dothe, H., Gamache, R. R., Goldman, A., Perevalov, V., Tashkun, S., & Tennyson, J. (2010). HITEMP, the high-temperature molecular spectroscopic database. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 111(15), 2139–2150. <https://doi.org/10.1016/j.jqsrt.2010.05.001>
- Savel, A. B., Kempton, E. M.-R., Malik, M., Komacek, T. D., Bean, J. L., May, E. M., Stevenson, K. B., Mansfield, M., & Rauscher, E. (2022). No umbrella needed: Confronting the hypothesis of iron rain on WASP-76b with post-processed general circulation models. *The Astrophysical Journal*, 926(1), 85. <https://doi.org/10.3847/1538-4357/ac423f>
- Seifahrt, A., Bean, J. L., Stürmer, J., Kasper, D., Gers, L., Schwab, C., Zechmeister, M., Stefánsson, G., Montet, B., Dos Santos, L. A., & others. (2020). On-sky commissioning of MAROON-x: A new precision radial velocity spectrograph for gemini north. *Ground-Based and Airborne Instrumentation for Astronomy VIII*, 11447, 305–325. <https://doi.org/10.1117/12.2561564>
- Showman, A. P., Tan, X., & Parmentier, V. (2020). Atmospheric dynamics of hot giant planets and brown dwarfs. *Space Science Reviews*, 216, 1–83. <https://doi.org/10.1007/s11214-020-00758-8>
- Tennyson, J., Yurchenko, S. N., Al-Refaie, A. F., Barton, E. J., Chubb, K. L., Coles, P. A., Diamantopoulou, S., Gorman, M. N., Hill, C., Lam, A. Z., & others. (2016). The ExoMol database: Molecular line lists for exoplanet and other hot atmospheres. *Journal of Molecular Spectroscopy*, 327, 73–94. <https://doi.org/10.1016/j.jms.2016.05.002>
- Wardenier, J. P., Parmentier, V., Line, M. R., & Lee, E. K. H. (2023). 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. *Monthly Notices of the Royal Astronomical Society*, 525(4), 4942–4961. <https://doi.org/10.1093/mnras/stad2586>