The Journal of Open Source Software

STCTM: a forward modeling and retrieval framework for stellar contamination and stellar spectra

Caroline Piaulet-Ghorayeb [1,2]

1 Department of Astronomy & Astrophysics, University of Chicago, 5640 South Ellis Avenue, Chicago, IL 60637, USA **2** E. Margaret Burbidge Prize Postdoctoral Fellow

DOI: 10.xxxxxx/draft

Software

- Review
- Repository
- Archive

Editor: Romain Thomas

Reviewers:

- @MartianColonist
- @nenasedk

Submitted: 05 May 2025
Published: unpublished

Summary

Transmission spectroscopy is a key avenue for the near-term study of small-planet atmospheres and the most promising method when it comes to searching for atmospheres on temperate rocky worlds, which are often too cold for planetary emission to be detectable. At the same time, the small planets that are most amenable for such atmospheric probes orbit small M dwarf stars. This "M-dwarf opportunity" has encountered a major challenge because of late-type stars' magnetic activity, which lead to the formation of spots and faculae at their surface. If inhomogeneously distributed throughout the photosphere, this phenomenon can give rise to "stellar contamination", or the transit light source effect (TLSE). Specifically, the TLSE describes the fact that spectral contrasts between bright and dark spots at the stellar surface outside of the transit chord can leave wavelength-dependent imprints in transmission spectra that may be mistaken for planetary atmosphere absorption.

As the field becomes increasingly ambitious in the search for signs of even thin atmospheres on small exoplanets, the TLSE is becoming a limiting factor, and it becomes imperative to develop robust inference methods to disentangle planetary and stellar contributions to the observed spectra. Here, I present `stctm`, the STellar ConTamination Modeling framework, a flexible Bayesian retrieval framework to model the impact of the TLSE on any exoplanet transmission spectrum, and infer the range of stellar surface parameters that are compatible with the observations in the absence of any planetary contribution. With the `exotune` sub-module, users can also perform retrievals directly on out-of-transit stellar spectra in order to place data-driven priors on the extent to which the TLSE can impact any planet's transmission spectrum. The input data formats, stellar models, and fitted parameters are easily tunable using human-readable files and the code is fully parallelized to enable fast inferences.

Statement of need

The interpretation of high-precision exoplanet transmission spectra from facilities such as the Hubble Space Telescope (HST) and the James Webb Space Telescope (JWST) is increasingly dependent on a robust accounting for the effects of stellar contamination, particularly for small planets orbiting small stars. Despite a growing awareness of the Transit Light Source Effect (TLSE; (Rackham et al., 2018; TRAPPIST-1 JWST Community Initiative et al., 2024)), the community currently lacks flexible, open-source tools that allow for robust modeling and retrieval of stellar contamination signatures. Further, uncertainties in stellar models motivate flexible implementations with reproducible model setups supporting any user-specified stellar model source, such as PHOENIX or SPHINX model grids (Husser et al., 2013; Iyer et al., 2023).

While some forward models have been developed to simulate the impact of stellar heterogeneity on transmission spectra, these tools are either not publicly available, computationally intractable due to their serial-mode-only implementation, part of much larger codes that require more advanced user training (e.g. atmospheric retrievals), or not designed for inference. Further, the community lacks frameworks that enable to retrieve stellar surface properties from both observed planetary transmission spectra and out-of-transit stellar spectra in a Bayesian context. This gap limits our ability to quantify uncertainties in exoplanet atmospheric properties and to test the robustness of atmospheric detections.

`stctm` addresses this need by providing an open-source, modular, and user-friendly framework. It allows users to model a wide range of stellar surface configurations leveraging any spectral models, and to infer which stellar parameters could explain observations without invoking planetary absorption. It also supports retrievals on out-of-transit stellar spectra to independently assess the extent of potential stellar contamination by the host star. By enabling flexible, fast, and reproducible inference of the TLSE, `stctm` empowers the community to critically assess the reliability of exoplanet atmosphere detections.

Main features of the code

The user inputs are communicated to the code via an input `.toml` file for both TLSE retrievals and inferences from out-of-transit stellar spectra. The code follows similar phases for both types of retrievals:

- Reading in and parsing of the inputs (data file, stellar models file, saving options, MCMC fit setup)
- Running the MCMC fit
- Post-processing to create diagnostic plots, record model comparison and goodness-of-fit metrics, produce publication-ready figures, and store sample spectra and parameters for post-processing and publication support and reproducibility (e.g. Zenodo)

`exotune` retrievals on out-of-transit stellar spectra have an additional (optional) pre-processing step, allowing users to:

- start from a full time-series of spectra as the input (e.g. the output from Stage 3 of the `Eureka!`(Bell et al., 2022) pipeline which is widely used in the community), rather than simply a pre-computed stellar spectrum
- exclude certain time intervals/exposures (e.g. containing the transit) and wavelength ranges (e.g. saturated regions) when computing the median spectrum to perform the retrieval over

For `exotune` retrievals, an error inflation parameter can be leveraged to account for the often-large mismatch between the data and stellar models.

Documentation

The full documentation for `stctm` with installation, testing instructions, and real-data example retrievals on transmission spectra and out-of-transit stellar spectra are available at https://stctm.readthedocs.io/. A description of `stctm` can be found in several of the early papers that employed it (see next section).

Uses of STCTM in the literature

`stctm` has been applied widely to the interpretation of transmission spectra of rocky planets and small sub-Neptunes, including in (Ahrer et al., 2025; Lim et al., 2023; Piaulet-Ghorayeb et al., 2024, 2025; Radica et al., 2025; Roy et al., 2023).

Future Developments

The latest version of `stctm` at the time of writing (v2.1.1) supports MCMC retrievals on transmission spectra and on out-of-transit stellar spectra (`exotune`), and provides model comparison statistics, model and parameter samples as well as publication-ready figures. Future versions will expand on these functionalities to include user-friendly scripts tailored to post-processing only for an already-run retrieval (creating custom plots), as well as a Nested Sampling alternative for the retrievals. Users are encouraged to propose or contribute any other features.

Similar Tools

Here are a few open-source codes that offer functionalities focused on retrievals of the TLSE or on out-of-transit stellar spectra:

- Generic atmospheric retrievals (including TLSE-only retrievals on transmission spectra): `POSEIDON` (MacDonald & Madhusudhan, 2024)
- Retrievals on out-of-transit stellar spectra (Nested Sampling, serial run mode only): `StellarFit` (Radica et al., 2025)

Acknowledgements

CPG thanks Warrick Ball for pre-review suggestions that improved the code documentation and ease of set-up for the user.`stctm` relies heavily on other Python libraries which include `numpy` (Harris et al., 2020), `scipy` (Virtanen et al., 2020), `astropy` (Astropy Collaboration et al., 2013, 2018), `matplotlib` (Hunter, 2007), `pandas`(team, 2020), `emcee`(Foreman-Mackey et al., 2013), `corner` (Foreman-Mackey, 2016), and `pysynphot` (Horne, 2013). Users are also strongly encouraged to use `msg`(Townsend & Lopez, 2023) to obtain the grids of stellar models used in the inference step.

CPG also acknowledges support from the E. Margaret Burbidge Prize Postdoctoral Fellowship from the Brinson Foundation. She thanks R. MacDonald, O. Lim, and M. Radica for helpful conversations that helped shape `stctm`.

References

Ahrer, E.-M., Radica, M., Piaulet-Ghorayeb, C., Raul, E., Wiser, L. S., Welbanks, L., Acuna, L., Allart, R., Coulombe, L.-P., Louca, A. J., MacDonald, R. J., Saidel, M., Evans-Soma, T. M., Benneke, B., Christie, D., Beatty, T. G., Cadieux, C., Cloutier, R., Doyon, R., … Schlichting, H. E. (2025). Escaping helium and a highly muted spectrum suggest a metal-enriched atmosphere on sub-neptune GJ3090b from JWST transit spectroscopy. *arXiv e-Prints*, arXiv:2504.20428. https://doi.org/10.48550/arXiv.2504.20428

Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., Günther, H. M., Lim, P. L., Crawford, S. M., Conseil, S., Shupe, D. L., Craig, M. W., Dencheva, N., Ginsburg, A., VanderPlas, J. T., Bradley, L. D., Pérez-Suárez, D., de Val-Borro, M., Aldcroft, T. L., Cruz, K. L., Robitaille, T. P., Tollerud, E. J., … Astropy Contributors. (2018). The Astropy Project: Building an Open-science Project and Status of the v2.0 Core Package. *The Astronomical Journal*, *156*(3), 123. https://doi.org/10.3847/1538-3881/aabc4f

Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., Greenfield, P., Droettboom, M., Bray, E., Aldcroft, T., Davis, M., Ginsburg, A., Price-Whelan, A. M., Kerzendorf, W. E., Conley, A., Crighton, N., Barbary, K., Muna, D., Ferguson, H., Grollier, F., Parikh, M. M., Nair, P. H., … Streicher, O. (2013). Astropy: A community Python package for astronomy. *Astronomy & Astrophysics*, *558*, A33. https://doi.org/10.1051/0004-6361/201322068

Bell, T., Ahrer, E.-M., Brande, J., Carter, A., Feinstein, A. D., Caloca, G., Mansfield, M., Zieba, S., Piaulet, C., Benneke, B., Filippazzo, J., May, E., Roy, P.-A., Kreidberg, L., & Stevenson, K. (2022). Eureka!: An End-to-End Pipeline for JWST Time-Series Observations. *Journal of Open Source Software*, *7*(79), 4503. https://doi.org/10.21105/joss.04503

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

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. (2013). Emcee: The MCMC hammer. *Publications of the Astronomical Society of the Pacific*, *125*(925), 306. https://doi.org/10.1086/670067

Harris, C. R., Millman, K. J., Walt, S. J. van der, Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., Kerkwijk, M. H. van, Brett, M., Haldane, A., Río, J. F. del, Wiebe, M., Peterson, P., … Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, *585*(7825), 357–362. https://doi.org/10.1038/s41586-020-2649-2

Horne, K. (2013). *Pysynphot: Synthetic photometry software*. https://ascl.net/1303.023; Astrophysics Source Code Library.

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*, *9*(3), 90–95. https://doi.org/10.1109/MCSE.2007.55

Husser, T.-O., Wende-von Berg, S., Dreizler, S., Homeier, D., Reiners, A., Barman, T., & Hauschildt, P. H. (2013). A new extensive library of PHOENIX stellar atmospheres and synthetic spectra. *Astronomy & Astrophysics*, *553*, A6. https://doi.org/10.1051/0004-6361/201219058

Iyer, A. R., Line, M. R., Muirhead, P. S., Fortney, J. J., & Gharib-Nezhad, E. (2023). The SPHINX m-dwarf spectral grid. I. Benchmarking new model atmospheres to derive fundamental m-dwarf properties. *The Astrophysical Journal*, *944*(1), 41. https://doi.org/10.3847/1538-4357/acabc2

Lim, O., Benneke, B., Doyon, R., MacDonald, R. J., Piaulet, C., Artigau, É., Coulombe, L.-P., Radica, M., L'Heureux, A., Albert, L., Rackham, B. V., de Wit, J., Salhi, S., Roy, P.-A., Flagg, L., Fournier-Tondreau, M., Taylor, J., Cook, N. J., Lafrenière, D., … Darveau-Bernier, A. (2023). Atmospheric Reconnaissance of TRAPPIST-1 b with JWST/NIRISS: Evidence for Strong Stellar Contamination in the Transmission Spectra. *The Astrophysical Journal Letters*, *955*(1), L22. https://doi.org/10.3847/2041-8213/acf7c4

MacDonald, R. J., & Madhusudhan, N. (2024). *POSEIDON: Multidimensional atmospheric retrieval of exoplanet spectra*. https://doi.org/10.21105/joss.04873

Piaulet-Ghorayeb, C., Benneke, B., Radica, M., Raul, E., Coulombe, L.-P., Ahrer, E.-M., Kubyshkina, D., Howard, W. S., Krissansen-Totton, J., MacDonald, R. J., Roy, P.-A., Louca, A., Christie, D., Fournier-Tondreau, M., Allart, R., Miguel, Y., Schlichting, H. E., Welbanks, L., Cadieux, C., … Knutson, H. A. (2024). JWST/NIRISS reveals the water-rich "steam world" atmosphere of GJ 9827 d. *The Astrophysical Journal Letters*, *974*(1), L10. https://doi.org/10.3847/2041-8213/ad6f00

Piaulet-Ghorayeb, C., Benneke, B., Turbet, M., Moore, K., Roy, P.-A., Lim, O., Doyon, R., Fauchez, T. J., Albert, L., Radica, M., Coulombe, L.-P., Lafrenière, D., Cowan, N. B., Belzile, D., Musfirat, K., Kaur, M., L'Heureux, A., Johnstone, D., MacDonald, R. J., … Turner, J. D. (2025). Strict limits on potential secondary atmospheres on the temperate rocky exo-earth TRAPPIST-1 d. *arXiv e-Prints*. https://ui.adsabs.harvard.edu/abs/2024ApJ...974L..10P

Rackham, B. V., Apai, D., & Giampapa, M. S. (2018). The transit light source effect: False spectral features and incorrect densities for m-dwarf transiting planets. *The Astrophysical Journal*, *853*(2), 122. https://doi.org/10.3847/1538-4357/aaa08c

Radica, M., Piaulet-Ghorayeb, C., Taylor, J., Coulombe, L.-P., Benneke, B., Albert, L., Artigau, É., Cowan, N. B., Doyon, R., Lafrenière, D., L'Heureux, A., & Lim, O. (2025). Promise and peril: Stellar contamination and strict limits on the atmosphere composition of TRAPPIST-1 c from JWST NIRISS transmission spectra. *The Astrophysical Journal Letters*, *979*(1), L5. https://doi.org/10.3847/2041-8213/ada381

Roy, P.-A., Benneke, B., Piaulet, C., Gully-Santiago, M. A., Crossfield, I. J. M., Morley, C. V., Kreidberg, L., Mikal-Evans, T., Brande, J., Delisle, S., Greene, T. P., Hardegree-Ullman, K. K., Barman, T., Christiansen, J. L., Dragomir, D., Fortney, J. J., Howard, A. W., Kosiarek, M. R., & Lothringer, J. D. (2023). Water absorption in the transmission spectrum of the water world candidate GJ 9827 d. *The Astrophysical Journal Letters*, *954*(2), L52. https://doi.org/10.3847/2041-8213/acebf0

team, T. pandas development. (2020). Pandas-dev/pandas: pandas. *Zenodo*. https://doi.org/10.5281/zenodo.3509134

Townsend, R., & Lopez, A. (2023). MSG: A software package for interpolating stellar spectra in pre-calculated grids. *The Journal of Open Source Software*, *8*(81), 4602. https://doi.org/10.21105/joss.04602

TRAPPIST-1 JWST Community Initiative, Wit, J. de, Doyon, R., Rackham, B. V., Lim, O., Ducrot, E., Kreidberg, L., Benneke, B., Ribas, I., Berardo, D., Niraula, P., Iyer, A., Shapiro, A., Kostogryz, N., Witzke, V., Gillon, M., Agol, E., Meadows, V., Burgasser, A. J., … Way, M. J. (2024). A roadmap for the atmospheric characterization of terrestrial exoplanets with JWST. *Nature Astronomy*, *8*, 810–818. https://doi.org/10.1038/s41550-024-02298-5

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., … SciPy 1.0 Contributors. (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, *17*, 261–272. https://doi.org/10.1038/s41592-019-0686-2