The Journal of Open Source Software

scarlet2: Astronomical scene modeling in jax

Peter Melchior [1¶], **Charlotte Ward** [1], **Benjamin Remy** [1], **Matt L. Sampson** [1], **and Jared Siegel** [1]

1 Princeton University, United States ¶ Corresponding author

DOI: 10.xxxxxx/draft

Software

- Review
- Repository
- Archive

Editor: Ivelina Momcheva

Reviewers:

- @AlexandreAdam
- @Jammy2211

Submitted: 29 July 2025
Published: unpublished

Summary

Large astronomical imaging surveys now contain billions of celestial sources, from stars in our Galaxy to galaxies towards the edge of the observable Universe. Because of improvements in resolution and sensitivity of current and future observing instruments, the images reveal more information than ever before. But they are also more difficult to analyze because galaxies exhibit complex morphologies, which cannot be described by traditional parametric models. And because there are so many sources, they routinely overlap with each other, either due to physical interactions or due to their close alignment along the line of sight. To extract all information of interest and avoid biases from incorrect modeling assumptions, it is therefore necessary to simultaneously model full scenes comprising many sources instead of analyzing each source separately, and each of the source models may itself need to be composed of multiple, morphologically complex components.

Statement of need

`scarlet2` is a Python package for full-scene modeling in observational astronomy. It inherits modeling assumptions from `scarlet` (Melchior et al., 2018), namely that a scene comprises multiple sources, each source comprises multiple components, and each component is determined by a spectrum model and a morphology model, whose outer product represents the light emission in a sky region as a hyperspectral data cube (wavelength × height × width). `scarlet2` retains the object-oriented paradigm and many classes and functions from `scarlet`, but augments standard Python with the `jax` library (Bradbury et al., 2018) and the `equinox` package (Kidger & Garcia, 2021) for automatic differentiation and just-in-time compilation.

`scarlet2` acts as a flexible, modular, and extendable modeling language for celestial sources that combines parametric and non-parametric models to describe complex scenarios such as multi-source blending, strong-lensing systems, supernovae and their host galaxies, etc. As a modeling language, `scarlet2` is agnostic about the optimization or inference method the user wants to employ, but it also provides methods to optimize the likelihood function or sample from the posterior, which utilize the `optax` package (DeepMind et al., 2020) or the `numpyro` inference framework (Bingham et al., 2019; Phan et al., 2019), respectively. The likelihood of multiple observations (at different resolutions, wavelengths, or observing epochs) can be combined for a joint model of static and transient sources. To match the coordinates from different observations, `scarlet2` utilizes the `Astropy` package (Astropy Collaboration, 2013). `scarlet2` can also interface with deep learning methods. Besides being natively portable to GPUs, parameters can be specified with neural networks as data-driven priors, which helps break the degeneracies that arise when multiple components are fit simultaneously (Sampson et al., 2024).

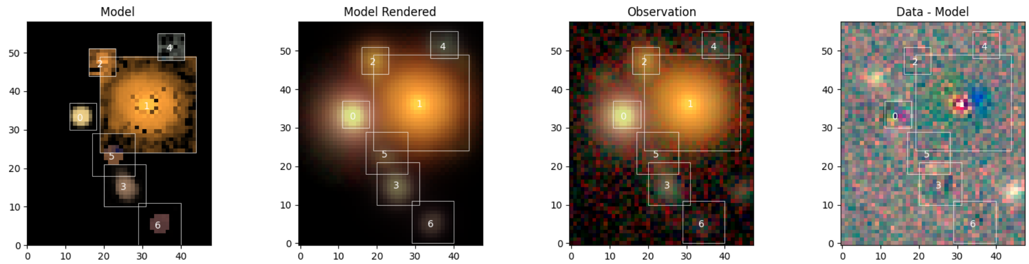

Figure 1: Scene with seven detected sources in multi-band images from the Hyper Suprime-Cam Subaru Strategic Program. Each source is modelled with a non-parametric spectrum and morphology (1st panel), the entire scene is then convolved with the telescope's point spread function (2nd panel) and compared to the observations (3rd panel). The residuals (4th panel) reveal the presence of previously undetected sources and source components (e.g. in the center of source #1).

To support the wide range of scientific studies that will be made with large sky surveys, `scarlet2` was designed with flexibility and ease of use in mind. Several publications have developed and demonstrated its capabilities, including modeling of interstellar dust embedded in distant galaxies (Siegel & Melchior, 2025) and of transient sources such as active galactic nuclei (Ward et al., 2025) and tidal disruption events (Yao et al., 2025). Future developments will integrate `scarlet2` into cloud-based science platforms, provide support for users to make effective modeling choices and to validate their inference results, and create a robust processing pipeline for joint pixel-level analyses of surveys from the Vera C. Rubin Observatory, the Euclid mission, the Nancy Grace Roman Space Telescope, and the La Silla Schmidt Southern Survey.

Acknowledgements

We acknowledge contributions from the LINCC Frameworks Incubator Program, in particular from software engineers Max West, Drew Oldag, and Sean McGuire, in adopting comprehensive software workflows through the Python Project Template (Oldag et al., 2024) and creating a user-focused recommendation and validation suite.

References

Astropy Collaboration. (2013). Astropy: A community Python package for astronomy. *Astronomy and Astrophysics*, *558*. https://doi.org/10.1051/0004-6361/201322068

Bingham, E., Chen, J. P., Jankowiak, M., Obermeyer, F., Pradhan, N., Karaletsos, T., Singh, R., Szerlip, P. A., Horsfall, P., & Goodman, N. D. (2019). Pyro: Deep universal probabilistic programming. *Journal of Machine Learning Research*, *20*, 28:1–28:6. http://jmlr.org/papers/v20/18-403.html

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/jax-ml/jax

DeepMind, Babuschkin, I., Baumli, K., Bell, A., Bhupatiraju, S., Bruce, J., Buchlovsky, P., Budden, D., Cai, T., Clark, A., Danihelka, I., Dedieu, A., Fantacci, C., Godwin, J., Jones, C., Hemsley, R., Hennigan, T., Hessel, M., Hou, S., … Viola, F. (2020). *The DeepMind JAX Ecosystem*. http://github.com/google-deepmind

Kidger, P., & Garcia, C. (2021). Equinox: Neural networks in JAX via callable PyTrees and filtered transformations. *Differentiable Programming Workshop at Neural Information Processing Systems 2021*.

Melchior, P., Moolekamp, F., Jerdee, M., Armstrong, R., Sun, A.-L., Bosch, J., & Lupton, R. (2018). SCARLET: Source separation in multi-band images by Constrained Matrix Factorization. *Astronomy and Computing*, *24*, 129. https://doi.org/10.1016/j.ascom.2018.07.001

Oldag, D., DeLucchi, M., Beebe, W., Branton, D., Campos, S., Chandler, C. O., Christofferson, C., Connolly, A., Kubica, J., Lynn, O., Malanchev, K., Malz, A. I., Mandelbaum, R., McGuire, S., & Wenneman, C. (2024). A Python Project Template for Healthy Scientific Software. *Research Notes of the American Astronomical Society*, *8*(5), 141. https://doi.org/10.3847/2515-5172/ad4da1

Phan, D., Pradhan, N., & Jankowiak, M. (2019). Composable effects for flexible and accelerated probabilistic programming in NumPyro. *arXiv Preprint arXiv:1912.11554*.

Sampson, M. L., Melchior, P., Ward, C., & Birmingham, S. (2024). Score-matching neural networks for improved multi-band source separation. *Astronomy and Computing*, *49*, 100875. https://doi.org/10.1016/j.ascom.2024.100875

Siegel, J. C., & Melchior, P. (2025). Spatially Resolved Galaxy–Dust Modeling with Coupled Data-driven Priors. *The Astrophysical Journal*, *986*(2), 212. https://doi.org/10.3847/1538-4357/add3f9

Ward, C., Melchior, P., Sampson, M. L., Burke, C. J., Siegel, J., Remy, B., Birmingham, S., Ramey, E., & Velzen, S. van. (2025). Disentangling transients and their host galaxies with scarlet2: A framework to forward model multi-epoch imaging. *Astronomy and Computing*, *51*, 100930. https://doi.org/10.1016/j.ascom.2025.100930

Yao, Y., Chornock, R., Ward, C., Hammerstein, E., Sfaradi, I., Margutti, R., Kelley, L. Z., Lu, W., Liu, C., Wise, J., Sollerman, J., Alexander, K. D., Bellm, E. C., Drake, A. J., Fremling, C., Gilfanov, M., Graham, M. J., Groom, S. L., Hinds, K. R., … Wold, A. (2025). A Massive Black Hole 0.8 kpc from the Host Nucleus Revealed by the Offset Tidal Disruption Event AT2024tvd. *The Astrophysical Journal Letters*, *985*(2), L48. https://doi.org/10.3847/2041-8213/add7de