exoplanet: Gradient-based probabilistic inference for exoplanet data & other astronomical time series

Daniel Foreman-Mackey[1], Rodrigo Luger[1,2], Eric Agol[3,2], Thomas Barclay[4], Luke G. Bouma[5], Timothy D. Brandt[6], Ian Czekala[7,8,9,10], Trevor J. David[1,11], Jiayin Dong[7,8], Emily A. Gilbert[12], Tyler A. Gordon[3], Christina Hedges[13,14], Daniel R. Hey[15,16], Brett M. Morris[17], Adrian M. Price-Whelan[1], and Arjun B. Savel[18]

1 Center for Computational Astrophysics, Flatiron Institute, New York, NY, USA **2** Virtual Planetary Laboratory, University of Washington, Seattle, WA, USA **3** Department of Astronomy, University of Washington, University of Washington, Seattle, WA, USA **4** Center for Space Sciences and Technology, University of Maryland, Baltimore County, Baltimore, MD, USA **5** Department of Astrophysical Sciences, Princeton University, Princeton, NJ, USA **6** Department of Physics, University of California, Santa Barbara, Santa Barbara, CA, USA **7** Department of Astronomy and Astrophysics, The Pennsylvania State University, University Park, PA, USA **8** Center for Exoplanets and Habitable Worlds, The Pennsylvania State University, University Park, PA, USA **9** Center for Astrostatistics, The Pennsylvania State University, University Park, PA, USA **10** Institute for Computational and Data Sciences, The Pennsylvania State University, University Park, PA, USA **11** Department of Astrophysics, American Museum of Natural History, New York, NY, USA **12** Department of Astronomy and Astrophysics, University of Chicago, Chicago, IL, USA **13** NASA Ames Research Center, Moffett Field, CA, USA **14** Bay Area Environmental Research Institute, Moffett Field, CA, USA **15** Sydney Institute for Astronomy, School of Physics, University of Sydney, Camperdown, New South Wales, Australia **16** Stellar Astrophysics Centre, Department of Physics and Astronomy, Aarhus University, Aarhus, Denmark **17** Center for Space and Habitability, University of Bern, Bern, Switzerland **18** Department of Astronomy, University of Maryland, College Park, MD, USA

Summary

exoplanet is a toolkit for probabilistic modeling of astronomical time series data, with a focus on observations of exoplanets, using PyMC3 (Salvatier et al., 2016). PyMC3 is a flexible and high-performance model building language and inference engine that scales well to problems with a large number of parameters. exoplanet extends PyMC3's modeling language to support many of the custom functions and probability distributions required when fitting exoplanet datasets or other astronomical time series.

While it has been used for other applications, such as the study of stellar variability (e.g., Gillen et al., 2020; Medina et al., 2020), the primary purpose of exoplanet is the characterization of exoplanets (e.g., Gilbert et al., 2020; Plavchan et al., 2020) or multiple star systems (e.g., Czekala et al., 2021) using time series photometry, astrometry, and/or radial velocity. In particular, the typical use case would be to use one or more of these datasets to place constraints on the physical and orbital parameters of the system, such as planet mass or orbital period, while simultaneously taking into account the effects of stellar variability.

Statement of need

Time domain astronomy is a priority of the observational astronomical community, with huge survey datasets currently available and more forthcoming. Within this research domain, there is significant investment into the discovery and characterization of exoplanets, planets orbiting stars other than our Sun. These datasets are large (on the scale of hundreds of thousands

of observations per star from space-based observatories such as *Kepler* and *TESS*), and the research questions are becoming more ambitious (in terms of both the computational cost of the physical models and the flexibility of these models). The packages in the *exoplanet* ecosystem are designed to enable rigorous probabilistic inference with these large datasets and high-dimensional models by providing a high-performance and well-tested infrastructure for integrating these models with modern modeling frameworks such as `PyMC3`. Since its initial release at the end of 2018, `exoplanet` has been widely used, with 64 citations of the Zenodo record (Foreman-Mackey et al., 2021) so far.

The *exoplanet* software ecosystem

Besides the primary `exoplanet` package, the *exoplanet* ecosystem of projects includes several other libraries. This paper describes, and is the primary reference for, this full suite of packages. The following provides a short description of each library within this ecosystem and discusses how they are related.

- `exoplanet`[1] is the primary library, and it includes implementations of many special functions required for exoplanet data analysis. These include the spherical geometry for computing orbits, some exoplanet-specific distributions for eccentricity (Kipping, 2013a; Van Eylen et al., 2019) and limb darkening (Kipping, 2013b), and exposure-time integrated limb darkened transit light curves.
- `exoplanet-core`[2] provides efficient, well-tested, and differentiable implementations of all of the exoplanet-specific operations that must be compiled for performance. These include an efficient solver for Kepler's equation (based on the algorithm proposed by Raposo-Pulido & Peláez, 2017) and limb darkened transit light curves (Agol et al., 2020). Besides the implementation for `PyMC3` and `Theano`, `exoplanet-core` includes implementations in `numpy` (Harris et al., 2020) and `jax` (Bradbury et al., 2018).
- `celerite2`[3], is an updated implementation of the *celerite* algorithm[4] (Foreman-Mackey, 2018; Foreman-Mackey et al., 2017) for scalable Gaussian Process regression for time series data. Like `exoplanet-core`, `celerite2` includes support for `numpy`, `jax`, and `PyMC3`, as well as some recent generalizations of the *celerite* algorithm (Gordon et al., 2020).
- `pymc3-ext`[5], includes a set of helper functions to make `PyMC3` more amenable to the typical astronomical data analysis workflow. For example, it provides a tuning schedule for `PyMC3`'s sampler (based on the method used by the `Stan` project Carpenter et al., 2017) that provides better performance on models with correlated parameters.
- `rebound-pymc3`[6] provides an interface between *REBOUND* (Rein & Liu, 2012), *REBOUNDx* (Tamayo et al., 2020), and `PyMC3` to enable inference with full N-body orbit integration.

Documentation & case studies

The main documentation page for the *exoplanet* libraries lives at docs.exoplanet.codes where it is hosted on ReadTheDocs. The tutorials included with the documentation are automatically executed on every push or pull request to the GitHub repository, with the goal of ensuring that the tutorials are always compatible with the current version of the code. The `celerite2` project has its own documentation page at celerite2.readthedocs.io, with tutorials that are similarly automatically executed.

---

[1] https://github.com/exoplanet-dev/exoplanet
[2] https://github.com/exoplanet-dev/exoplanet-core
[3] https://celerite2.readthedocs.io
[4] https://celerite.readthedocs.io
[5] https://github.com/exoplanet-dev/pymc3-ext
[6] https://github.com/exoplanet-dev/rebound-pymc3

Alongside these documentation pages, there is a parallel "Case Studies" website at gallery.exoplanet.codes that includes more detailed example use cases for `exoplanet` and the other libraries described here. Like the tutorials on the documentation page, these case studies are automatically executed using GitHub Actions, but at lower cadence (once a week and when a new release of the `exoplanet` library is made) since the runtime is much longer.

Similar tools

There is a rich ecosystem of tooling available for inference with models such as the ones supported by `exoplanet`. Each of these tools has its own set of strengths and limitations and we will not make a detailed comparison here, but it is worth listing some of these tools and situating `exoplanet` in this context.

Some of the most popular tools in this space include (and note that this is far from a comprehensive list!) EXOFAST (J. Eastman et al., 2013; J. D. Eastman et al., 2019), `radvel` (Fulton et al., 2018), `juliet` (Espinoza et al., 2019), `exostriker` (Trifonov, 2019), PYANETI (Barragán et al., 2019), `allesfitter` (Günther & Daylan, 2021), and `orbitize` (Blunt et al., 2020). These packages all focus on providing a high-level interface for designing models and then executing a fit. `exoplanet`, however, is designed to be lower level and more conceptually similar to tools like `batman` (Kreidberg, 2015), `PyTransit` (Parviainen, 2015), `ellc` (Maxted, 2016), `starry` (Luger et al., 2019), or `Limbdark.jl` (Agol et al., 2020), which provide the building blocks for evaluating the models required for inference with exoplanet datasets. In fact, several of the higher-level packages listed above include these lower-level libraries as dependencies, and our hope is that `exoplanet` could provide the backend for future high-level libraries.

As emphasized in the title of this paper, the main selling point of `exoplanet` when compared to other tools in this space is that it supports differentiation of all components of the model and is designed to integrate seamlessly with the `aesara` (Willard et al., 2021; formerly known as `Theano`, Theano Development Team, 2016) automatic differentiation framework used by `PyMC3`. This allows the use of modern inference algorithms such as No U-Turn Sampling (Hoffman & Gelman, 2014) or Automatic Differentiation Variational Inference (Kucukelbir et al., 2017). These algorithms can have some computational and conceptual advantages over inference methods that do not use gradients, especially for high-dimensional models. The computation of gradients is also useful for model optimization; this is necessary when, say, searching for new exoplanets, mapping out degeneracies or multiple modes of a posterior, or estimating uncertainties from a Hessian. Care has been taken to provide gradients which are numerically stable, and more accurate and faster to evaluate than finite-difference gradients which can be subject to significant numerical errors and require 2N computations of a model with N free parameters.

Acknowledgements

References

Agol, E., Luger, R., & Foreman-Mackey, D. (2020). Analytic Planetary Transit Light Curves and Derivatives for Stars with Polynomial Limb Darkening. *The Astronomical Journal*, *159*(3), 123. https://doi.org/10.3847/1538-3881/ab4fee

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*, 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

Barragán, O., Gandolfi, D., & Antoniciello, G. (2019). PYANETI: a fast and powerful software suite for multiplanet radial velocity and transit fitting. *Monthly Notices of the Royal Astronomical Society*, *482*, 1017–1030. https://doi.org/10.1093/mnras/sty2472

Blunt, S., Wang, J. J., Angelo, I., Ngo, H., Cody, D., De Rosa, R. J., Graham, J. R., Hirsch, L., Nagpal, V., Nielsen, E. L., Pearce, L., Rice, M., & Tejada, R. (2020). orbitize!: A Comprehensive Orbit-fitting Software Package for the High-contrast Imaging Community. *The Astronomical Journal*, *159*(3), 89. https://doi.org/10.3847/1538-3881/ab6663

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 0.2.5) [Computer software]. http://github.com/google/jax

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., & Riddell, A. (2017). Stan: A Probabilistic Programming Language. *Journal of Statistical Software*, *76*(1), 1–32. https://doi.org/10.18637/jss.v076.i01

Czekala, I., Ribas, Á., Cuello, N., Chiang, E., Macías, E., Duchêne, G., Andrews, S. M., & Espaillat, C. C. (2021). A Coplanar Circumbinary Protoplanetary Disk in the TWA 3 Triple M Dwarf System. *The Astrophysical Journal*, *912*(1), 6. https://doi.org/10.3847/1538-4357/abebe3

Eastman, J. D., Rodriguez, J. E., Agol, E., Stassun, K. G., Beatty, T. G., Vanderburg, A., Gaudi, B. S., Collins, K. A., & Luger, R. (2019). EXOFASTv2: A public, generalized, publication-quality exoplanet modeling code. *arXiv e-Prints*, arXiv:1907.09480. http://arxiv.org/abs/1907.09480

Eastman, J., Gaudi, B. S., & Agol, E. (2013). EXOFAST: A Fast Exoplanetary Fitting Suite in IDL. *Publications of the Astronomical Society of the Pacific*, *125*(923), 83. https://doi.org/10.1086/669497

Espinoza, N., Kossakowski, D., & Brahm, R. (2019). juliet: a versatile modelling tool for transiting and non-transiting exoplanetary systems. *Monthly Notices of the Royal Astronomical Society*, *490*(2), 2262–2283. https://doi.org/10.1093/mnras/stz2688

Foreman-Mackey, D. (2018). Scalable Backpropagation for Gaussian Processes using Celerite. *Research Notes of the American Astronomical Society*, *2*(1), 31. https://doi.org/10.3847/2515-5172/aaaf6c

Foreman-Mackey, D., Agol, E., Ambikasaran, S., & Angus, R. (2017). Fast and Scalable Gaussian Process Modeling with Applications to Astronomical Time Series. *The Astronomical Journal*, *154*, 220. https://doi.org/10.3847/1538-3881/aa9332

Foreman-Mackey, D., Savel, A., Luger, R., Czekala, I., Agol, E., Price-Whelan, A., Hedges, C., Gilbert, E., Barclay, T., Bouma, L., & Brandt, T. D. (2021). *exoplanet-dev/exoplanet* (Version 0.4.5) [Computer software]. Zenodo. https://doi.org/10.5281/zenodo.1998447

Fulton, B. J., Petigura, E. A., Blunt, S., & Sinukoff, E. (2018). RadVel: The Radial Velocity Modeling Toolkit. *Publications of the Astronomical Society of the Pacific*, *130*(986), 044504. https://doi.org/10.1088/1538-3873/aaaaa8

Gilbert, E. A., Barclay, T., Schlieder, J. E., Quintana, E. V., Hord, B. J., Kostov, V. B., Lopez, E. D., Rowe, J. F., Hoffman, K., Walkowicz, L. M., Silverstein, M. L., Rodriguez, J. E., Vanderburg, A., Suissa, G., Airapetian, V. S., Clement, M. S., Raymond, S. N., Mann, A. W., Kruse, E., … Winters, J. G. (2020). The First Habitable-zone Earth-sized Planet from TESS. I. Validation of the TOI-700 System. *The Astronomical Journal*, *160*(3), 116. https://doi.org/10.3847/1538-3881/aba4b2

Gillen, E., Briegal, J. T., Hodgkin, S. T., Foreman-Mackey, D., Van Leeuwen, F., Jackman, J. A. G., McCormac, J., West, R. G., Queloz, D., Bayliss, D., Goad, M. R., Watson, C. A., Wheatley, P. J., Belardi, C., Burleigh, M. R., Casewell, S. L., Jenkins, J. S., Raynard, L., Smith, A. M. S., … Vines, J. I. (2020). NGTS clusters survey - I. Rotation in the young benchmark open cluster Blanco 1. *Monthly Notices of the Royal Astronomical Society*, *492*(1), 1008–1024. https://doi.org/10.1093/mnras/stz3251

Gordon, T. A., Agol, E., & Foreman-Mackey, D. (2020). A Fast, Two-dimensional Gaussian Process Method Based on Celerite: Applications to Transiting Exoplanet Discovery and Characterization. *The Astronomical Journal*, *160*(5), 240. https://doi.org/10.3847/1538-3881/abbc16

Günther, M. N., & Daylan, T. (2021). Allesfitter: Flexible Star and Exoplanet Inference from Photometry and Radial Velocity. *The Astrophysical Journal Supplement Series*, *254*(1), 13. https://doi.org/10.3847/1538-4365/abe70e

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

Hoffman, M. D., & Gelman, A. (2014). The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, *15*(1), 1593–1623. https://jmlr.org/papers/v15/hoffman14a.html

Kipping, D. M. (2013a). Parametrizing the exoplanet eccentricity distribution with the beta distribution. *Monthly Notices of the Royal Astronomical Society*, *434*, L51–L55. https://doi.org/10.1093/mnrasl/slt075

Kipping, D. M. (2013b). Efficient, uninformative sampling of limb darkening coefficients for two-parameter laws. *Monthly Notices of the Royal Astronomical Society*, *435*, 2152–2160. https://doi.org/10.1093/mnras/stt1435

Kreidberg, L. (2015). batman: BAsic Transit Model cAlculatioN in Python. *Publications of the Astronomical Society of the Pacific*, *127*(957), 1161. https://doi.org/10.1086/683602

Kucukelbir, A., Tran, D., Ranganath, R., Gelman, A., & Blei, D. M. (2017). Automatic Differentiation Variational Inference. *Journal of Machine Learning Research*, *18*(14), 1–45. http://jmlr.org/papers/v18/16-107.html

Kumar, R., Carroll, C., Hartikainen, A., & Martin, O. A. (2019). ArviZ a unified library for exploratory analysis of Bayesian models in Python. *The Journal of Open Source Software.* https://doi.org/10.21105/joss.01143

Luger, R., Agol, E., Foreman-Mackey, D., Fleming, D. P., Lustig-Yaeger, J., & Deitrick, R. (2019). starry: Analytic Occultation Light Curves. *The Astronomical Journal*, *157*, 64. https://doi.org/10.3847/1538-3881/aae8e5

Maxted, P. F. L. (2016). ellc: A fast, flexible light curve model for detached eclipsing binary stars and transiting exoplanets. *Astronomy & Astrophysics*, *591*, A111. https://doi.org/10.1051/0004-6361/201628579

Medina, A. A., Winters, J. G., Irwin, J. M., & Charbonneau, D. (2020). Flare Rates, Rotation Periods, and Spectroscopic Activity Indicators of a Volume-complete Sample of Mid- to Late-M Dwarfs within 15 pc. *The Astrophysical Journal*, *905*(2), 107. https://doi.org/10.3847/1538-4357/abc686

Parviainen, H. (2015). PYTRANSIT: fast and easy exoplanet transit modelling in PYTHON. *Monthly Notices of the Royal Astronomical Society*, *450*(3), 3233–3238. https://doi.org/10.1093/mnras/stv894

Plavchan, P., Barclay, T., Gagné, J., Gao, P., Cale, B., Matzko, W., Dragomir, D., Quinn, S., Feliz, D., Stassun, K., Crossfield, I. J. M., Berardo, D. A., Latham, D. W., Tieu, B., Anglada-Escudé, G., Ricker, G., Vanderspek, R., Seager, S., Winn, J. N., … Zilberman, P. (2020). A planet within the debris disk around the pre-main-sequence star AU Microscopii. *Nature*, *582*(7813), 497–500. https://doi.org/10.1038/s41586-020-2400-z

Raposo-Pulido, V., & Peláez, J. (2017). An efficient code to solve the Kepler equation. Elliptic case. *Monthly Notices of the Royal Astronomical Society*, *467*(2), 1702–1713. https://doi.org/10.1093/mnras/stx138

Rein, H., & Liu, S.-F. (2012). REBOUND: an open-source multi-purpose N-body code for collisional dynamics. *Astronomy & Astrophysics*, *537*, A128. https://doi.org/10.1051/0004-6361/201118085

Salvatier, J., Wiecki, T. V., & Fonnesbeck, C. (2016). Probabilistic programming in python using PyMC3. *PeerJ Computer Science*, *2*, e55. https://doi.org/10.7717/peerj-cs.55

Tamayo, D., Rein, H., Shi, P., & Hernandez, D. M. (2020). REBOUNDx: a library for adding conservative and dissipative forces to otherwise symplectic N-body integrations. *Monthly Notices of the Royal Astronomical Society*, *491*(2), 2885–2901. https://doi.org/10.1093/mnras/stz2870

Theano Development Team. (2016). Theano: A Python framework for fast computation of mathematical expressions. *arXiv e-Prints*, *abs/1605.02688*. http://arxiv.org/abs/1605.02688

Trifonov, T. (2019). *The Exo-Striker: Transit and radial velocity interactive fitting tool for orbital analysis and N-body simulations* (p. ascl:1906.004).

Van Eylen, V., Albrecht, S., Huang, X., MacDonald, M. G., Dawson, R. I., Cai, M. X., Foreman-Mackey, D., Lundkvist, M. S., Silva Aguirre, V., Snellen, I., & Winn, J. N. (2019). The Orbital Eccentricity of Small Planet Systems. *The Astronomical Journal*, *157*(2), 61. https://doi.org/10.3847/1538-3881/aaf22f

Willard, B. T., Osthege, M., Ho, G., Vieira, R., Wiecki, T., Foreman-Mackey, D., Chaudhari, K., Legrand, N., Kumar, R., Lao, J., Abril-Pla, O., Fonnesbeck, C., Goldman, R. P., & Gorelli, M. (2021). *pymc-devs/aesara* (Version 2.0.7) [Computer software]. Zenodo. https://doi.org/10.5281/zenodo.4695331