JAXtronomy: A JAX port of lenstronomy

Alan Huang [1], **Simon Birrer** [1], **Natalie B. Hogg** [2], **Aymeric Galan** [3,4], **Daniel Gilman** [5], **Anowar J. Shajib** [5,6,7], and **Nan Zhang** [8]

1 Department of Physics and Astronomy, Stony Brook University, Stony Brook, NY 1794, USA **2** Laboratoire Univers et Particules de Montpellier, CNRS and Université de Montpellier (UMR-5299), 34095 Montpellier, France **3** Max-Planck-Institut für Astrophysik, Karl-Schwarzschild Straße 1, 85748 Garching, Germany **4** Technical University of Munich, TUM School of Natural Sciences, Physics Department, James-Franck-Straße 1, 85748 Garching, Germany **5** Department of Astronomy and Astrophysics, University of Chicago, Chicago, Illinois 60637, USA **6** Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA **7** Center for Astronomy, Space Science and Astrophysics, Independent University, Bangladesh, Dhaka 1229, Bangladesh **8** Department of Physics, University of Illinois, 1110 West Green St., Urbana, IL 61801, USA

Summary

Gravitational lensing is a phenomenon where light bends around massive objects, resulting in distorted images seen by an observer. Studying graviationally lensed objects can give us key insights into cosmology and astrophysics, such as constraints on the expansion rate of the universe and dark matter models.

Thus, we introduce JAXtronomy, a re-implementation of the gravitational lensing software package lenstronomy[1] (Birrer, 2021; Birrer & Amara, 2018) usingJAX[2]. JAX is a Python library that uses an accelerated linear algebra (XLA) compiler to improve the performance of computing software. Our core design principle of JAXtronomy is to maintain an identical API to that of lenstronomy.

The main JAX features utilized in JAXtronomy are just-in-time-compilation, which can lead to significant reductions in execution time, and automatic differentiation, which allows for the implementation of gradient-based algorithms that were previously impossible. Additionally, JAX allows code to be run on GPUs or parallelized across CPU cores, further boosting the performance of JAXtronomy.

Statement of need

lenstronomy has been widely applied to numerous science cases, with more than 200 publications making use of the software, and has an increasing number of dependent packages relying on features of lenstronomy. For instance, science cases directly involving lenstronomy include galaxy evolution studies using strong lensing (Anowar J. Shajib et al., 2021; Sheu et al., 2025; Tan et al., 2024) and detailed lens modeling for measuring the Hubble constant using time-delay cosmography by the TDCOSMO collaboration (Birrer, S. et al., 2020; Birrer, Simon & Treu, Tommaso, 2021; Collaboration et al., 2025; Gilman, D. et al., 2020; Millon, M. et al., 2020; Schmidt et al., 2025; A. J. Shajib et al., 2022; **?**).

Examples of packages dependent on lenstronomy for general-purpose lensing computations and image modelling include the dolphin package (Anowar J. Shajib et al., 2025) for automated lens modeling, the galight package (Ding et al., 2020) for galaxy morphology measurements,

---

[1] https://github.com/lenstronomy/lenstronomy
[2] https://github.com/jax-ml/jax

SLSim (Khadka et al, 2025, in prep) for simulating large populations of strong lenses, pyHalo (Gilman et al., 2019) and mejiro (Wedig et al., 2025) for simulating strong lenses with dark matter substructure, and PALTAS (Wagner-Carena et al., 2023) for neural network inference tasks.

In many of these applications, computational constraints are the key limiting factor for strong gravitational lensing science. For example, increased data quality and number of lenses to analyze makes lens modeling a computational bottleneck, and expensive ray-tracing through tens of thousands of dark matter substructures limit the amount of images that can be simulated, especially for the training of neural networks and simulation-based inferences. These ever-increasing computational costs have lead to the development of several JAX-accelerated and GPU-accelerated strong-lensing packages, such as gigalens (Gu et al., 2022), herculens (Galan et al., 2022), paltax (Wagner-Carena et al., 2024), GLaD (Wang et al., 2025), caustics[3] (Stone et al., 2024) and Google Research's jaxstronomy[4].

Why JAXtronomy?

JAXtronomy inherits a wide range of features from lenstronomy that are not offered by any of the aforementioned JAX-accelerated or GPU-accelerated software. Such features include lenstronomy's linear amplitude solver, which reduces the number of sampled parameters during lens modeling, as well as a variety of log likelihood functions and optional punishment terms. Furthermore, JAXtronomy aims to maintain an identical API to lenstronomy so that packages dependent on lenstronomy can transition seamlessly to JAXtronomy.

Improvements over lenstronomy in image simulation

The simulation of a lensed image comes in three main steps. The first step begins with a coordinate grid in the angles seen by the observer. These coordinates are ray-traced through the deflectors back to the source plane. This process requires the calculation of light ray deflection angles at each deflector. Second, the surface brightness of the source is calculated on the ray-traced coordinate grid. This produces a lensed image. Third, the lensed image gets convolved by the point spread function (PSF) originating from diffraction of the telescope optics and atmospheric turbulence. Due to the various choices in deflector mass profiles, light model profiles, grid size, and PSF kernel size, the overall runtime of the pipeline can vary significantly.

In the following sections, we outline the improvements in performance that JAXtronomy has over lenstronomy for each step in the pipeline. These performance benchmarks were run using an Intel(R) Xeon(R) Gold 6338 CPU @ 2.00GHz, an NVIDIA A100 GPU, and JAX version 0.7.0.

Deflection angle calculations

Each entry in the table indicates how much faster JAXtronomy is compared to lenstronomy at computing deflection angles for the corresponding deflector profile and grid size. Some comparisons vary significantly with values of function arguments, so a range is given rather than a number.

Deflector Profile	60x60 grid (cpu)	180x180 grid (cpu)	180x180 grid (gpu)
CONVERGENCE	0.4x	1.1x	0.5x
CSE	1.6x	2.6x	2.6x
EPL	5.1x - 15x	9.2x - 17x	37x - 120x

---

[3]https://github.com/Ciela-Institute/caustics
[4]https://github.com/google-research/google-research/tree/master/jaxstronomy

Deflector Profile	60x60 grid (cpu)	180x180 grid (cpu)	180x180 grid (gpu)
EPL (jax) vs EPL_NUMBA	1.4x	3.0x	13x
EPL_MULTIPOLE_M1M3M4	2.1x - 7x	6.4x - 13x	42x - 108x
HERNQUIST	2.0x	3.4x	5.8x
HERNQUIST_ELLIPSE_CSE	3.8x	5.4x	40x
MULTIPOLE	0.9x	1.0x	8.3x - 14x
MULTIPOLE_ELL	1.5x - 2.1x	2.0x - 2.8x	70x
NIE/SIE	0.5x	0.5x	2.0x
NFW	1.6x	3.3x	4.5x
NFW_ELLIPSE_CSE	4.1x	6.7x	31x
PJAFFE	1.0x	1.2x	2.8x
PJAFFE_ELLIPSE_POTENTIAL	1.4x	1.6x	3.1x
SHEAR	0.7x	2.0x	0.9x
SIS	1.4x	3.3x	2.0x
TNFW	2.4x	5.8x	7.5x

For small enough grid sizes, JAXtronomy computes deflection angles slower than `lenstronomy` when using certain deflector profiles. This is because function call overheads are significantly higher in JAX than in standard Python, so computations that are already fast in Python can end up slower in JAX. In these cases, the benefit of using JAX is to have automatic differentiation for lens modeling.

Flux calculations

An analogous table for the different light profiles is shown below. The MULTI_GAUSSIAN and MULTI_GAUSSIAN_ELLIPSE profiles include five GAUSSIAN and GAUSSIAN_ELLIPSE components, respectively, highlighting JAX's improved performance in sequential computations.

Light Profile	60x60 grid (cpu)	180x180 grid (cpu)	180x180 grid (gpu)
CORE_SERSIC	2.0x	6.7x	4.2x
GAUSSIAN	1.0x	2.5x	1.3x
GAUSSIAN_ELLIPSE	1.5x	3.6x	2.0x
MULTI_GAUSSIAN	3.7x	11x	7.8x
MULTI_GAUSSIAN_ELLIPSE	4.0x	13x	6.9x
SERSIC	1.0x	1.7x	3.9x
SERSIC_ELLIPSE	1.9x	5.7x	3.2x
SERSIC_ELLIPSE_Q_PHI	1.7x	5.5x	3.3x
SHAPELETS (n_max=6)	6.2x	3.4x	15x
SHAPELETS (n_max=10)	6.0x	4.5x	17x

FFT Convolution

We find that FFT convolution using JAX on CPU results in variable performance boosts or slowdowns compared to `lenstronomy` (which uses `scipy`'s FFT convolution). On a 60x60 grid, and kernel sizes ranging from 3 to 45, JAX on CPU ranges from being 1.1x to 2.9x faster than `lenstronomy`, with no obvious correlation to kernel size. On a 180x180 grid, and kernel sizes ranging from 9 to 135, JAXtronomy on CPU ranges from being 0.7x to 2.5x as fast as `lenstronomy`, with no obvious correlation to kernel size.

95 However, FFT convolution using JAX on GPU is significantly faster than `scipy`. On a 60x60
96 grid, and kernel sizes ranging from 3 to 45, JAX on GPU ranges from being 1.5x to 3.5x
97 faster than `lenstronomy`, with JAX performing better at higher kernel sizes. On a 180x180
98 grid, and kernel sizes ranging from 9 to 135, JAXtronomy on GPU is about 10x to 20x as fast
99 as `lenstronomy`, again with JAX performing better at higher kernel sizes.

Improvements over lenstronomy in lens modelling

101 The process of lens modelling involves finding best-fit parameters describing a lensed system from
102 real data. In `lenstronomy`, this typically involves a Particle Swarm Optimizer (PSO) (Kennedy
103 & Eberhart, 1995) for optimization and Monte Carlo Markov Chains for posterior sampling.
104 JAXtronomy retains these lens modelling algorithms from `lenstronomy` while benefitting from
105 the increased performance outlined above.

106 In the following table, we compare JAXtronomy's PSO performance to that of `lenstronomy`
107 when modeling a lens with an elliptical power-law mass profile, Sersic light profile, and a
108 quadruply-imaged point source. We use a 100x100 grid and a size 13 PSF kernel. These
109 benchmarks were performed using the same hardware as in the previous section.

Number of Particles	1 CPU core	64 CPU cores (parallelized)	GPU
64	4x	16x	5x
128	4x	18x	5.5x
256	4.7x	30x	9x
512	4.7x	34x	11x

110 Additionally, using JAX's autodifferentiation, we have implemented the L-BFGS gradient descent
111 algorithm from the `Optax`[5] library (DeepMind et al., 2020) for optimization. This is a significant
112 improvement over `lenstronomy`'s PSO, which does not have access to gradient information.
113 Due to the random nature of the PSO, we do not present a concrete comparison between
114 `lenstronomy` and JAXtronomy for how long it takes to find best-fit parameters. However, we
115 note that JAXtronomy can find a good fit within one minute, while `lenstronomy` can take
116 hours.

References

118 Birrer, S. (2021). Gravitational lensing formalism in a curved arc basis: A continuous
119 description of observables and degeneracies from the weak to the strong lensing regime.
120 *The Astrophysical Journal*, *919*(1), 38. https://doi.org/10.3847/1538-4357/ac1108

121 Birrer, S., & Amara, A. (2018). Lenstronomy: Multi-purpose gravitational lens modelling
122 software package. *Physics of the Dark Universe*, *22*, 189–201. https://doi.org/10.1016/j.
123 dark.2018.11.002

124 Birrer, Simon, & Treu, Tommaso. (2021). TDCOSMO - v. Strategies for precise and accurate
125 measurements of the hubble constant with strong lensing. *Astronomy & Astrophysics*, *649*,
126 A61. https://doi.org/10.1051/0004-6361/202039179

127 Birrer, S., Shajib, A. J., Galan, A., Millon, M., Treu, T., Agnello, A., Auger, M., Chen, G.
128 C.-F., Christensen, L., Collett, T., Courbin, F., Fassnacht, C. D., Koopmans, L. V. E.,
129 Marshall, P. J., Park, J.-W., Rusu, C. E., Sluse, D., Spiniello, C., Suyu, S. H., … Van de
130 Vyvere, L. (2020). TDCOSMO - IV. Hierarchical time-delay cosmography − joint inference
131 of the hubble constant and galaxy density profiles⋆. *Astronomy & Astrophysics*, *643*, A165.
132 https://doi.org/10.1051/0004-6361/202038861

[5]https://github.com/google-deepmind/optax

133 Collaboration, T., Birrer, S., Buckley-Geer, E. J., Cappellari, M., Courbin, F., Dux, F.,
134 Fassnacht, C. D., Frieman, J. A., Galan, A., Gilman, D., Huang, X.-Y., Knabel, S.,
135 Langeroodi, D., Lin, H., Millon, M., Morishita, T., Motta, V., Mozumdar, P., Paic, E., …
136 Wong, K. C. (2025). *TDCOSMO 2025: Cosmological constraints from strong lensing time*
137 *delays*. https://arxiv.org/abs/2506.03023

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

142 Ding, X., Silverman, J., Treu, T., Schulze, A., Schramm, M., Birrer, S., Park, D., Jahnke, K.,
143 Bennert, V. N., Kartaltepe, J. S., Koekemoer, A. M., Malkan, M. A., & Sanders, D. (2020).
144 The Mass Relations between Supermassive Black Holes and Their Host Galaxies at $1 < z$
145 < 2 HST-WFC3. *The Astrophysical Journal*, *888*(1), 37. https://doi.org/10.3847/1538-
146 4357/ab5b90

147 Galan, A., Vernardos, G., Peel, A., Courbin, F., & Starck, J.-L. (2022). Using wavelets
148 to capture deviations from smoothness in galaxy-scale strong lenses. *Astronomy &*
149 *Astrophysics*, *668*, A155. https://doi.org/10.1051/0004-6361/202244464

150 Gilman, D., Birrer, S., Nierenberg, A., Treu, T., Du, X., & Benson, A. (2019). Warm dark
151 matter chills out: Constraints on the halo mass function and the free-streaming length of
152 dark matter with eight quadruple-image strong gravitational lenses. *Monthly Notices of the*
153 *Royal Astronomical Society*, *491*(4), 6077–6101. https://doi.org/10.1093/mnras/stz3480

154 Gilman, D., Birrer, S., & Treu, T. (2020). TDCOSMO - III. Dark matter substructure meets
155 dark energy. The effects of (sub)halos on strong-lensing measurements of H0. *Astronomy*
156 *& Astrophysics*, *642*, A194. https://doi.org/10.1051/0004-6361/202038829

157 Gu, A., Huang, X., Sheu, W., Aldering, G., Bolton, A. S., Boone, K., Dey, A., Filipp, A.,
158 Jullo, E., Perlmutter, S., Rubin, D., Schlafly, E. F., Schlegel, D. J., Shu, Y., & Suyu, S. H.
159 (2022). GIGA-lens: Fast bayesian inference for strong gravitational lens modeling. *The*
160 *Astrophysical Journal*, *935*(1), 49. https://doi.org/10.3847/1538-4357/ac6de4

161 Kennedy, J., & Eberhart, R. (1995). Particle swarm optimization. *Proceedings of ICNN'95*
162 *- International Conference on Neural Networks*, *4*, 1942–1948 vol.4. https://doi.org/10.
163 1109/ICNN.1995.488968

164 Millon, M., Galan, A., Courbin, F., Treu, T., Suyu, S. H., Ding, X., Birrer, S., Chen, G. C.-F.,
165 Shajib, A. J., Sluse, D., Wong, K. C., Agnello, A., Auger, M. W., Buckley-Geer, E. J.,
166 Chan, J. H. H., Collett, T., Fassnacht, C. D., Hilbert, S., Koopmans, L. V. E., … Van
167 de Vyvere, L. (2020). TDCOSMO - i. An exploration of systematic uncertainties in the
168 inference of H0 from time-delay cosmography. *Astronomy & Astrophysics*, *639*, A101.
169 https://doi.org/10.1051/0004-6361/201937351

170 Schmidt, T., Treu, T., Birrer, S., Millon, M., Sluse, D., Galan, A., Shajib, A., Lemon, C.,
171 Dux, F., & Courbin, F. (2025). TDCOSMO. XVIII. Strong lens model and time-delay
172 predictions for J1721+8842, the first einstein zigzag lens. *Astronomy & Astrophysics*.
173 https://doi.org/10.1051/0004-6361/202449984

174 Shajib, Anowar J., Nihal, N. S., Tan, C. Y., Sahu, V., Birrer, S., Treu, T., & Frieman, J. (2025).
175 *dolphin: A fully automated forward modeling pipeline powered by artificial intelligence for*
176 *galaxy-scale strong lenses*. https://arxiv.org/abs/2503.22657

177 Shajib, Anowar J., Treu, T., Birrer, S., & Sonnenfeld, A. (2021). Dark matter haloes of
178 massive elliptical galaxies at z ∼ 0.2 are well described by the Navarro-Frenk-White
179 profile. *Monthly Notices of the Royal Astronomical Society*, *503*(2), 2380–2405. https:
180 //doi.org/10.1093/mnras/stab536

181 Shajib, A. J., Wong, K. C., Birrer, S., Suyu, S. H., Treu, T., Buckley-Geer, E. J., Lin, H., Rusu, C.

E., Poh, J., Palmese, A., Agnello, A., Auger-Williams, M. W., Galan, A., Schuldt, S., Sluse, D., Courbin, F., Frieman, J., & Millon, M. (2022). TDCOSMO. IX. Systematic comparison between lens modelling software programs: Time-delay prediction for WGD 2038−4008. *Astronomy & Astrophysics*, *667*, A123. https://doi.org/10.1051/0004-6361/202243401

Sheu, W., Shajib, A. J., Treu, T., Sonnenfeld, A., Birrer, S., Cappellari, M., Oldham, L. J., & Tan, C. Y. (2025). Project Dinos II: redshift evolution of dark and luminous matter density profiles in strong-lensing elliptical galaxies across $0.1 < z < 0.9$. *Monthly Notices of the Royal Astronomical Society*, *541*(1), 1–27. https://doi.org/10.1093/mnras/staf976

Stone, C., Adam, A., Coogan, A., Yantovski-Barth, M. J., Filipp, A., Setiawan, L., Core, C., Legin, R., Wilson, C., Barco, G. M., Hezaveh, Y., & Perreault-Levasseur, L. (2024). Caustics: A python package for accelerated strong gravitational lensing simulations. *Journal of Open Source Software*, *9*(103), 7081. https://doi.org/10.21105/joss.07081

Tan, C. Y., Shajib, A. J., Birrer, S., Sonnenfeld, A., Treu, T., Wells, P., Williams, D., Buckley-Geer, E. J., Drlica-Wagner, A., & Frieman, J. (2024). Project Dinos I: A joint lensing-dynamics constraint on the deviation from the power law in the mass profile of massive ellipticals. *Monthly Notices of the Royal Astronomical Society*, *530*(2), 1474–1505. https://doi.org/10.1093/mnras/stae884

Wagner-Carena, S., Aalbers, J., Birrer, S., Nadler, E. O., Darragh-Ford, E., Marshall, P. J., & Wechsler, R. H. (2023). From images to dark matter: End-to-end inference of substructure from hundreds of strong gravitational lenses. *The Astrophysical Journal*, *942*(2), 75. https://doi.org/10.3847/1538-4357/aca525

Wagner-Carena, S., Lee, J., Pennington, J., Aalbers, J., Birrer, S., & Wechsler, R. H. (2024). A strong gravitational lens is worth a thousand dark matter halos: Inference on small-scale structure using sequential methods. *The Astrophysical Journal*, *975*(2), 297. https://doi.org/10.3847/1538-4357/ad6e70

Wang, H., Suyu, S. H., Galan, A., Halkola, A., Cappellari, M., Shajib, A. J., & Cernetic, M. (2025). *GPU-accelerated gravitational lensing & dynamical (GLaD) modeling for cosmology and galaxies*. https://arxiv.org/abs/2504.01302

Wedig, B., Daylan, T., Birrer, S., Cyr-Racine, F.-Y., Dvorkin, C., Finkbeiner, D. P., Huang, A., Huang, X., Karthik, R., Khadka, N., Natarajan, P., Nierenberg, A. M., Peter, A. H. G., Pierel, J. D. R., Tang, X. T., & Wechsler, R. H. (2025). The roman view of strong gravitational lenses. *The Astrophysical Journal*, *986*(1), 42. https://doi.org/10.3847/1538-4357/adc24f