Scalable Bayesian Inference for Finding Strong Gravitational Lenses

Yash Patel
Department of Statistics
University of Michigan
yppatel@umich.edu

Jeffrey Regier
Department of Statistics
University of Michigan
regier@umich.edu

Abstract

Finding strong gravitational lenses in astronomical images allows us to assess cosmological theories and understand the large-scale structure of the universe. Previous works on lens detection do not quantify uncertainties in lens parameter estimates or scale to modern surveys. We present a fully amortized Bayesian procedure for lens detection that overcomes these limitations. Unlike traditional variational inference, in which training minimizes the reverse Kullback-Leibler (KL) divergence, our method is trained with an expected forward KL divergence. Using synthetic GalSim images and real Sloan Digital Sky Survey (SDSS) images, we demonstrate that amortized inference trained with the forward KL produces well-calibrated uncertainties in both lens detection and parameter estimation.

1 Introduction

Strong gravitational lensing events are widely used to validate and parameterize the ΛCDM model, the current concordance model in cosmology [1, 2, 3, 4, 5]. Despite extensive study, it remains challenging to efficiently detect strong lenses and to accurately estimate their characteristics. Because researchers anticipate that the upcoming Legacy Survey of Space and Time (LSST) will image roughly 10^5 lenses [6], such efficient detection is of interest.

Non-generative deep learning detectors are computationally efficient [7, 8, 9], but they sacrifice the accuracy and uncertainty quantification provided by fully generative models. Additionally, they do not cope well with blending, instances where *multiple* galaxies overlap visually. Handling blending is paramount, as it is anticipated that 62% of imaged galaxies in LSST will be blended [10]. Standalone deblenders have been developed [11, 12], but they lack probabilistic interpretability. Bayesian methods have been shown to address these deficiencies yet remain too computationally demanding to be run on large datasets. A recent work [13] employs a bespoke sampling process to improve upon Hamiltonian Monte Carlo yet still requires 105 seconds on four Nvidia A100 GPUs for a single lens. Further, previous works that used variational inference were trained with the reverse KL divergence, which is known to produce underdispersed posterior approximations [14]. A separate line of work has studied lens substructure [15, 16]; however, these works presume the identification of a lens.

We propose to detect strong lensing while fitting a generative model for deblending with an inference procedure that is scalable to modern astronomical surveys. Our method is amortized, supports calibrated uncertainty quantification, and is available at `https://github.com/prob-ml/bliss`.

2 Statistical Model

Astronomical images record radiation originating from light sources, such as stars and galaxies. Catalogs contain properties of these sources, such as their locations and fluxes. It is of interest to infer the posterior distributions for these properties in addition to those of strong gravitational lenses.

Machine Learning and the Physical Sciences workshop, NeurIPS 2022.

We propose the following generative model for this task, which extends the BLISS model [17, 18]. First, draw the number of imaged light sources from a Poisson process, $S \sim \text{Poisson}(\mu\eta)$, with μ denoting the average density of light sources per square degree of the image and η denoting the number of square degrees. Then, for each source $s = 1, ..., S$, the location and type of the source are

$$u_s \mid S \sim \text{Unif}([0, H] \times [0, W]) \qquad \text{and} \qquad a_s \mid S \sim \text{Bernoulli}(\rho_s), \tag{1}$$

where ρ_s is the proportion of sources that are stars, and $1 - \rho_s$ is the proportion that are galaxies. We use the star and galaxy flux models presented in [18], namely TruncatedPareto($f_{\min}, 0.5$) for stars and a bulge-and-disk model $\mathcal{G}$ for galaxies, parameterized by g_s. The full specification of g_s is given in Table 1. Additionally, if a source is a galaxy, whether it is lensed is indicated by

$$\gamma_s \mid (S, a_s = 0) \sim \text{Bernoulli}(\rho_\ell), \tag{2}$$

where ρ_ℓ is the proportion of galaxies that are lensed. We assume that all lensing events require a pair of galaxies (s, s'), with s acting as a lens and s' being lensed. The galaxy s' is initially rendered unlensed (with $g_{s'}$), followed by a resampling operation on a grid warped by the singular isothermal ellipsoid (SIE) lensing potential, parameterized by $r_\ell := (\theta_E, q_1, q_2, \theta_x, \theta_y)$, whose values are determined by interactions between s and s' [19]. Denoting the grid distortion as $\mathcal{D}_{r_\ell}$, a lens pair is rendered as

$$f_{0,s,s'} \mid (S, a_s = 0, a_{s'} = 0, \gamma_{s'} = 1, g_s, g_{s'}, r_\ell) = \mathcal{G}(g_s) + \mathcal{D}_{r_\ell}\left(\mathcal{G}(g_{s'})\right). \tag{3}$$

Denote the background photon contribution as ζ_n, the contribution from a source s to pixel n as $\lambda_{n,s}$, and the complete set of latent variables as z. See Table 1 for descriptions and priors of such parameters. Then, the number of photon arrivals observed at pixel n is

$$x_n \mid z \sim \text{Poisson}\left(\zeta_n + \sum_{s=1}^{S} \lambda_{n,s}\right). \tag{4}$$

3 Variational Inference

We aim to minimize the expected forward KL divergence to approximate the posterior distribution using forward amortized variance inference (FAVI) [20]. We thus aim to solve

$$\arg\min_{\varphi} \mathbb{E}_{(x,z)\sim p(z)p(x|z)} \left[\log(q_\varphi(z|x))\right]. \tag{5}$$

Because we employ the FAVI loss, we are not restricted to reparameterizable distributions. We thus use the following variational distribution:

$$q_\varphi(z|x) = q(S) \prod_{s=1}^{S} q(\ell_s|S) q(a_s|S) q(g_s|S, a_s) q(\gamma_s|S, a_s) q(r_{\ell,s}|S, a_s, \gamma_s). \tag{6}$$

Table 1 gives the distributional form of each factor. Each factor was approximated using a separate "encoder" neural network, one for each of the following tasks: source count estimation, source classification, galaxy parameter estimation, lens classification, and lens parameter estimation.

4 Results

All encoders were implemented in PyTorch [21] with standard CNN architectures and employed the tiling decomposition described in [18]. Each was trained separately on synthetic images from the generative model, with galaxies rendered using GalSim [22]. Optimization was done using Adam [23]. Training these models required five hours using eight Nvidia RTX 2080 Ti GPUs. This is a one-time cost: inference can be run on an arbitrary number of images thereafter without additional training. We used both synthetic data and images from SDSS for validation.

Name	Generative	Variational	Description
S	Poisson$(\mu\eta)$	Categorical	Number of sources
u	$\mathcal{U}([0,H]\times[0,W])$	$\log(u)\sim\mathcal{N}(\mu_u,\sigma_u^2)$	Location of source
a	Bernoulli(ρ_s)	Bernoulli(μ_{a_s})	Type of source
f_1	Pareto$(f_{\min},\alpha)$	$\log(f_1)\sim\mathcal{N}(\mu_{f_1},\sigma_{f_1}^2)$	Star flux
f_T	Pareto$(f_{\min},\alpha_f)$	$\log(f_T)\sim\mathcal{N}(\mu_{f_T},\sigma_{f_T}^2)$	Total galactic flux
d_p	$\mathcal{U}[0,1]$	$\operatorname{logit}(d_p)\sim\mathcal{N}(\mu_{d_p},\sigma_{d_p}^2)$	Disk flux proportion
β	$\mathcal{U}[0,2\pi]$	$\operatorname{logit}\left(\frac{\beta}{2\pi}\right)\sim\mathcal{N}(\mu_\beta,\sigma_\beta^2)$	Galaxy ellipse angular offset
d_q	$\mathcal{U}[0,1]$	$\operatorname{logit}(d_q)\sim\mathcal{N}(\mu_{d_q},\sigma_{d_q}^2)$	Disk minor-to-major axis ratio
b_q	$\mathcal{U}[0,1]$	$\operatorname{logit}(b_q)\sim\mathcal{N}(\mu_{b_q},\sigma_{b_q}^2)$	Bulge minor-to-major axis ratio
a_d	Gamma(α_d,β_d)	$\log(a_d)\sim\mathcal{N}(\mu_{a_d},\sigma_{a_d}^2)$	Major axis for the disk
a_b	Gamma(α_b,β_b)	$\log(a_b)\sim\mathcal{N}(\mu_{a_b},\sigma_{a_b}^2)$	Major axis for the bulge
f_0	Composite Sérsic	N/A	Galaxy flux
γ	Bernoulli(ρ_ℓ)	Bernoulli(μ_γ)	Indicator of lensing
θ_E	$\mathcal{U}[\theta_{E,\min},\theta_{E,\max}]$	$\log(\theta_E)\sim\mathcal{N}(\mu_{\theta_E},\sigma_{\theta_E}^2)$	Einstein radius
θ_x	$\mathcal{N}(0,1)$	$\theta_x\sim\mathcal{N}(\mu_{\theta_x},\sigma_{\theta_x}^2)$	Lens center x
θ_y	$\mathcal{N}(0,1)$	$\theta_y\sim\mathcal{N}(\mu_{\theta_y},\sigma_{\theta_y}^2)$	Lens center y
q_ℓ	$\mathcal{U}[0,1]$	N/A	Lens minor-to-major axis ratio
β_ℓ	$\mathcal{U}[-\pi/4,\pi/4]$	N/A	Lens angular offset
e_1	$\frac{1-q_\ell}{1+q_\ell}\cos(\beta_\ell)$	$\operatorname{logit}\left(\frac{e_1+1}{2}\right)\sim\mathcal{N}(\mu_{e_1},\sigma_{e_1}^2)$	Lens ellipticity (factor 1)
e_2	$\frac{1-q_\ell}{1+q_\ell}\sin(\beta_\ell)$	$\operatorname{logit}\left(\frac{e_2+1}{2}\right)\sim\mathcal{N}(\mu_{e_2},\sigma_{e_2}^2)$	Lens ellipticity (factor 2)

Table 1: Parameters for the generative model and variational distribution. The four partitions of the table respectively correspond to the detection, star, galaxy, and lens parameters.

4.1 Synthetic Images

The encoders were trained on data generated through the posited forward model, as shown in Figure 1. Post-training validation was also performed in a number of ways. In particular, Figure 1 also serves as a visual qualitative posterior check. To assess uncertainty calibration, discrete and continuous latent quantities were handled separately. Discrete quantities, namely galaxy and lens detection, were plotted with their outputted posterior probabilities against the empirical proportions. Continuous variable posterior calibration was assessed with coverage percentages for 90% Bayes credible intervals. Results in Figure 2 reveal well-calibrated posterior distributions for detection and parameter estimates for both galaxies and lenses. Understanding specific sources of calibration imperfections is of interest; one plausible cause stems from the limited expressivity of the encoders, owing to the fact the neural networks have a finite number of layers.

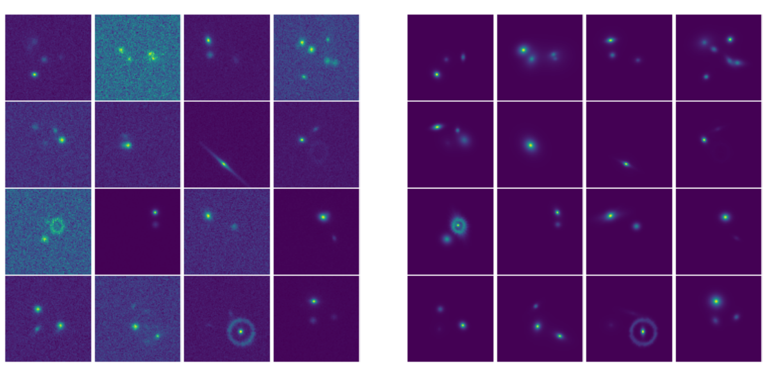
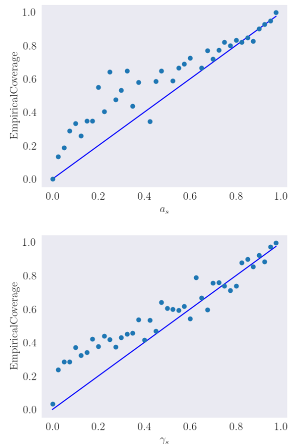


Figure 1: Synthetic images from our generative model. Each is normalized against the brightest object in the image. The left panel shows the original synthetic images and the right our reconstructions. Checking the similarity of the reconstructed images serves as an initial qualitative posterior check.



Name	Coverage	Name	Coverage
f_T	94.66 %	$f_{T,\ell}$	91.22%
d_p	89.62 %	$d_{p,\ell}$	87.79%
β_s	90.88 %	$\beta_{s,\ell}$	89.87%
d_q	88.11 %	$d_{q,\ell}$	89.05%
b_q	87.29 %	$b_{q,\ell}$	89.42%
a_d	87.24 %	$a_{d,\ell}$	89.24%
a_b	90.20 %	$a_{b,\ell}$	95.47%
θ_E	94.39 %	e_1	94.66%
θ_x	92.22 %	e_2	89.78%
θ_y	91.50 %		

Figure 2: Assessment of posterior calibration for detection and continuous parameters, respectively shown in the graph and table. The blue lines in the graphs represent the ideal calibrations. Bayes credible intervals were constructed for 90% coverage.

4.2 Sloan Digital Sky Survey (SDSS)

We additionally apply our model to two SDSS images, referencing annotations of lenses from [24]. We demonstrate successful detections in Figure 3, importantly achieved *without* false positives. The images were both 1489 $\times$ 2048 pixels and inference required just 25 seconds for each image.

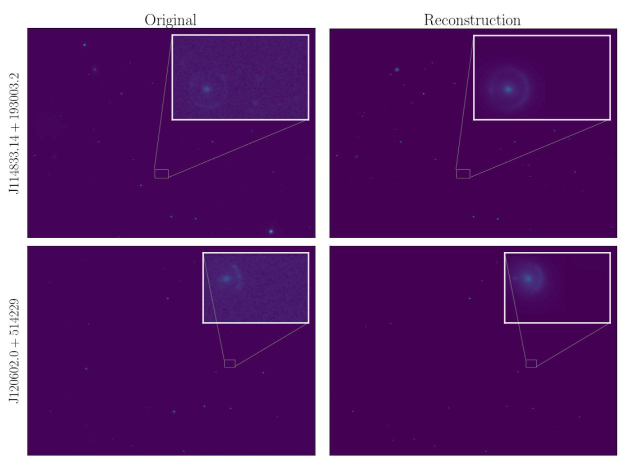


Figure 3: The left column shows the original images from SDSS and the right our reconstructions. The top row is the reconstruction pair for the field containing J114833.14+193003.2 and the bottom that for J120602.0+514229. The zoom box is to highlight the subregions containing the lenses.

5 Discussion

We have shown that amortized inference performed by FAVI efficiently detects strong lenses and estimates parameters in both synthetic and real data settings while providing well-calibrated uncertainty estimates. With this foundation, a number of extensions of this research are possible. One is the use of this approach to infer *weak* lensing events, whose manifestations in data are quite different. For this, several non-trivial adjustments would be necessary in both the generative and inference procedures. Characterization of dark matter substructures is also of great interest and would similarly require extensions to the SIE model employed here.

6 Impact Statement

This work builds on the use of machine learning to further astronomical and, more generally, scientific understanding, when addressing problems in which uncertainty quantification is a necessary component. Beyond direct application of the techniques presented here to the image data gathered in future astronomical surveys, particularly the recently launched JWST [25], the broader variational inference methodology we propose could be extended to future scientific queries.

References

[1] Daniel Gilman, Jo Bovy, Tommaso Treu, Anna Nierenberg, Simon Birrer, Andrew Benson, and Omid Sameie. Strong lensing signatures of self-interacting dark matter in low-mass haloes. *Monthly Notices of the Royal Astronomical Society*, 507(2):2432–2447, 2021.

[2] Yashar D Hezaveh, Neal Dalal, Daniel P Marrone, Yao-Yuan Mao, Warren Morningstar, Di Wen, Roger D Blandford, John E Carlstrom, Christopher D Fassnacht, Gilbert P Holder, et al. Detection of lensing substructure using ALMA observations of the dusty galaxy SDP.81. *The Astrophysical Journal*, 823(1):37, 2016.

[3] S Vegetti, LVE Koopmans, A Bolton, T Treu, and R Gavazzi. Detection of a dark substructure through gravitational imaging. *Monthly Notices of the Royal Astronomical Society*, 408(4):1969–1981, 2010.

[4] Simona Vegetti and LVE Koopmans. Statistics of mass substructure from strong gravitational lensing: quantifying the mass fraction and mass function. *Monthly Notices of the Royal Astronomical Society*, 400(3):1583–1592, 2009.

[5] David W Hogg and RD Blandford. The gravitational lens system B1422+231: dark matter, superluminal expansion and the Hubble constant. *Monthly Notices of the Royal Astronomical Society*, 268(4):889–893, 1994.

[6] Tonghua Liu, Shuo Cao, Jia Zhang, Marek Biesiada, Yuting Liu, and Yujie Lian. Testing the cosmic curvature at high redshifts: the combination of LSST strong lensing systems and quasars as new standard candles. *Monthly Notices of the Royal Astronomical Society*, 496(1):708–717, 2020.

[7] Yashar D Hezaveh, Laurence Perreault Levasseur, and Philip J Marshall. Fast automated analysis of strong gravitational lenses with convolutional neural networks. *Nature*, 548(7669):555–557, 2017.

[8] Johann Brehmer, Siddharth Mishra-Sharma, Joeri Hermans, Gilles Louppe, and Kyle Cranmer. Mining for dark matter substructure: Inferring subhalo population properties from strong lenses with machine learning. *The Astrophysical Journal*, 886(1):49, 2019.

[9] Stephon Alexander, Sergei Gleyzer, Evan McDonough, Michael W Toomey, and Emanuele Usai. Deep learning the morphology of dark matter substructure. *The Astrophysical Journal*, 893(1):15, 2020.

[10] Javier Sanchez, Ismael Mendoza, David P Kirkby, Patricia R Burchat, et al. Effects of overlapping sources on cosmic shear estimation: Statistical sensitivity and pixel-noise bias. *Journal of Cosmology and Astroparticle Physics*, 2021(07):043, 2021.

[11] Peter Melchior, Fred Moolekamp, Maximilian Jerdee, Robert Armstrong, A-L Sun, James Bosch, and Robert Lupton. Scarlet: Source separation in multi-band images by constrained matrix factorization. *Astronomy and Computing*, 24:129–142, 2018.

[12] François Lanusse, Rachel Mandelbaum, Siamak Ravanbakhsh, Chun-Liang Li, Peter Freeman, and Barnabás Póczos. Deep generative models for galaxy image simulations. *Monthly Notices of the Royal Astronomical Society*, 504(4):5543–5555, 2021.

[13] A Gu, X Huang, W Sheu, G Aldering, AS Bolton, K Boone, A Dey, A Filipp, E Jullo, S Perlmutter, et al. Giga-lens: Fast Bayesian inference for strong gravitational lens modeling. *arXiv preprint arXiv:2202.07663*, 2022.

[14] Kevin P. Murphy. *Probabilistic Machine Learning: Advanced Topics*. MIT Press, 2023.

[15] Marco Chianese, Adam Coogan, Paul Hofma, Sydney Otten, and Christoph Weniger. Differentiable strong lensing: uniting gravity and neural nets through differentiable probabilistic programming. *Monthly Notices of the Royal Astronomical Society*, 496(1):381–393, 2020.

[16] Siddharth Mishra-Sharma and Ge Yang. Strong lensing source reconstruction using continuous neural fields. *arXiv preprint arXiv:2206.14820*, 2022.

[17] Runjing Liu, Jon D McAuliffe, and Jeffrey Regier. Variational inference for deblending crowded starfields. *arXiv preprint arXiv:2102.02409*, 2021.

[18] Derek Hansen, Ismael Mendoza, Runjing Liu, Ziteng Pang, Zhe Zhao, Camille Avestruz, and Jeffrey Regier. Scalable Bayesian inference for detection and deblending in astronomical images. *arXiv preprint arXiv:2207.05642*, 2022.

[19] Ramesh Narayan and Matthias Bartelmann. Lectures on gravitational lensing. *arXiv preprint astro-ph/9606001*, 1996.

[20] Luca Ambrogioni, Umut Güçlü, Julia Berezutskaya, Eva Borne, Yağmur Güçlütürk, Max Hinne, Eric Maris, and Marcel Gerven. Forward amortized inference for likelihood-free variational marginalization. In *The 22nd International Conference on Artificial Intelligence and Statistics*, pages 777–786. PMLR, 2019.

[21] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. PyTorch: An imperative style, high-performance deep learning library. *Advances in Neural Information Processing Systems*, 32, 2019.

[22] Barnaby TP Rowe, Mike Jarvis, Rachel Mandelbaum, Gary M Bernstein, James Bosch, Melanie Simet, Joshua E Meyers, Tomasz Kacprzak, Reiko Nakajima, Joe Zuntz, et al. Galsim: The modular galaxy image simulation toolkit. *Astronomy and Computing*, 10:121–150, 2015.

[23] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[24] Zhong-Lue Wen, Jin-Lin Han, Xiang-Yang Xu, Yun-Ying Jiang, Zhi-Qing Guo, Peng-Fei Wang, and Feng-Shan Liu. Discovery of four gravitational lensing systems by clusters in the SDSS dr6. *Research in Astronomy and Astrophysics*, 9(1):5, 2009.

[25] Jonathan P Gardner, John C Mather, Mark Clampin, Rene Doyon, Matthew A Greenhouse, Heidi B Hammel, John B Hutchings, Peter Jakobsen, Simon J Lilly, Knox S Long, et al. The James Webb Space Telescope. *Space Science Reviews*, 123(4):485–606, 2006.

Checklist

The checklist follows the references. Please read the checklist guidelines carefully for information on how to answer these questions. For each question, change the default **[TODO]** to [Yes] , [No] , or [N/A] . You are strongly encouraged to include a **justification to your answer**, either by referencing the appropriate section of your paper or providing a brief inline description. For example:

- Did you include the license to the code and datasets? [Yes] See Section.
- Did you include the license to the code and datasets? [No] The code and the data are proprietary.
- Did you include the license to the code and datasets? [N/A]

Please do not modify the questions and only use the provided macros for your answers. Note that the Checklist section does not count towards the page limit. In your paper, please delete this instructions block and only keep the Checklist section heading above along with the questions/answers below.

1. For all authors...

(a) Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope? [Yes]
(b) Did you describe the limitations of your work? [Yes]
(c) Did you discuss any potential negative societal impacts of your work? [Yes]
(d) Have you read the ethics review guidelines and ensured that your paper conforms to them? [Yes]

2. If you are including theoretical results...
   (a) Did you state the full set of assumptions of all theoretical results? [N/A]
   (b) Did you include complete proofs of all theoretical results? [N/A]
3. If you ran experiments...
   (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [Yes]
   (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [Yes]
   (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [No]
   (d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [Yes]
4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...
   (a) If your work uses existing assets, did you cite the creators? [Yes]
   (b) Did you mention the license of the assets? [N/A]
   (c) Did you include any new assets either in the supplemental material or as a URL? [No]
   (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [N/A]
   (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [N/A]
5. If you used crowdsourcing or conducted research with human subjects...
   (a) Did you include the full text of instructions given to participants and screenshots, if applicable? [N/A]
   (b) Did you describe any potential participant risks, with links to Institutional Review Board (IRB) approvals, if applicable? [N/A]
   (c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? [N/A]