Scalable Bayesian Inference for Finding Strong Gravitational Lenses

Yash Patel[1], Jeffrey Regier[1]

[1]Department of Statistics, University of Michigan

UNIVERSITY OF MICHIGAN

Introduction

Strong gravitational lensing events are widely used to validate and parameterize the ΛCDM model. 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, such efficient detection is of interest**.

Non-generative deep learning detectors are computationally efficient, 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**.

We propose to **detect strong lensing** while **fitting a generative model for deblending** with an inference procedure that is scalable to modern astronomical surveys.

Statistical Model

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**

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

where ρ_s is the proportion of sources that are stars, and $1 - \rho_s$ is the proportion that are galaxies. If a source is a galaxy, **whether it is lensed is indicated by**

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

where ρ_ℓ is the proportion of galaxies that are lensed. If galaxy s is lensed, it is rendered on a grid **warped by the singular isothermal ellipsoid (SIE) potential**, parameterized by $r_\ell := (\theta_E, q_1, q_2, \theta_x, \theta_y)$.

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. The **number of photon arrivals at pixel n is**

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

Variational Inference

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

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

We thus use the following variational distribution:

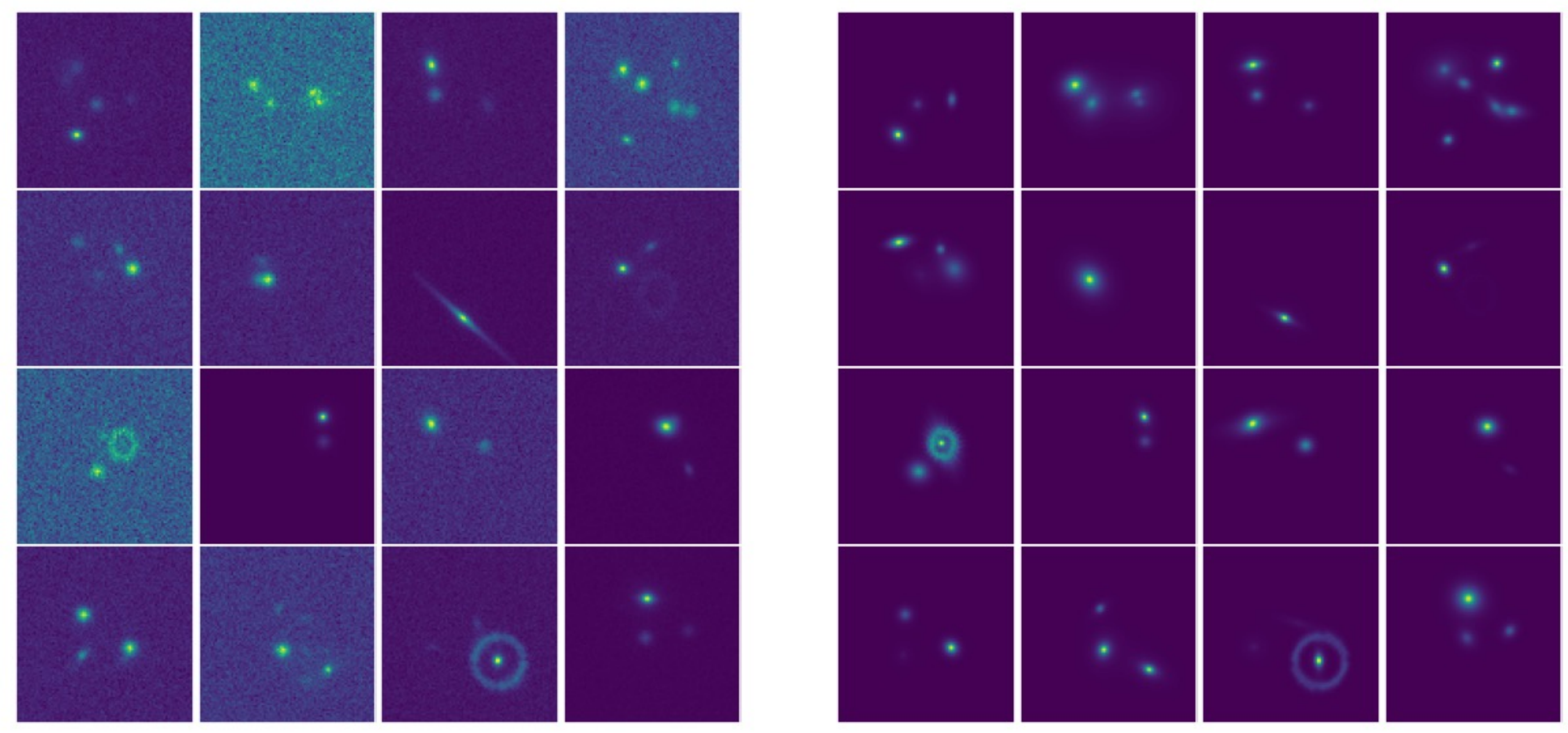
$$q_\varphi(z|x) = q(S) \prod_{s=1}^{S} q(\ell_s|S) q(a_s|S) q(g_s|S, a_s) q(\gamma_s|S, a_s) q(r_{\ell,s}|S, a_s, \gamma_s)$$

Each factor approximated using a separate "encoder" neural net.

Results

Task I: Synthetic Data

The figure below shows synthetic images from our generative model. 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.

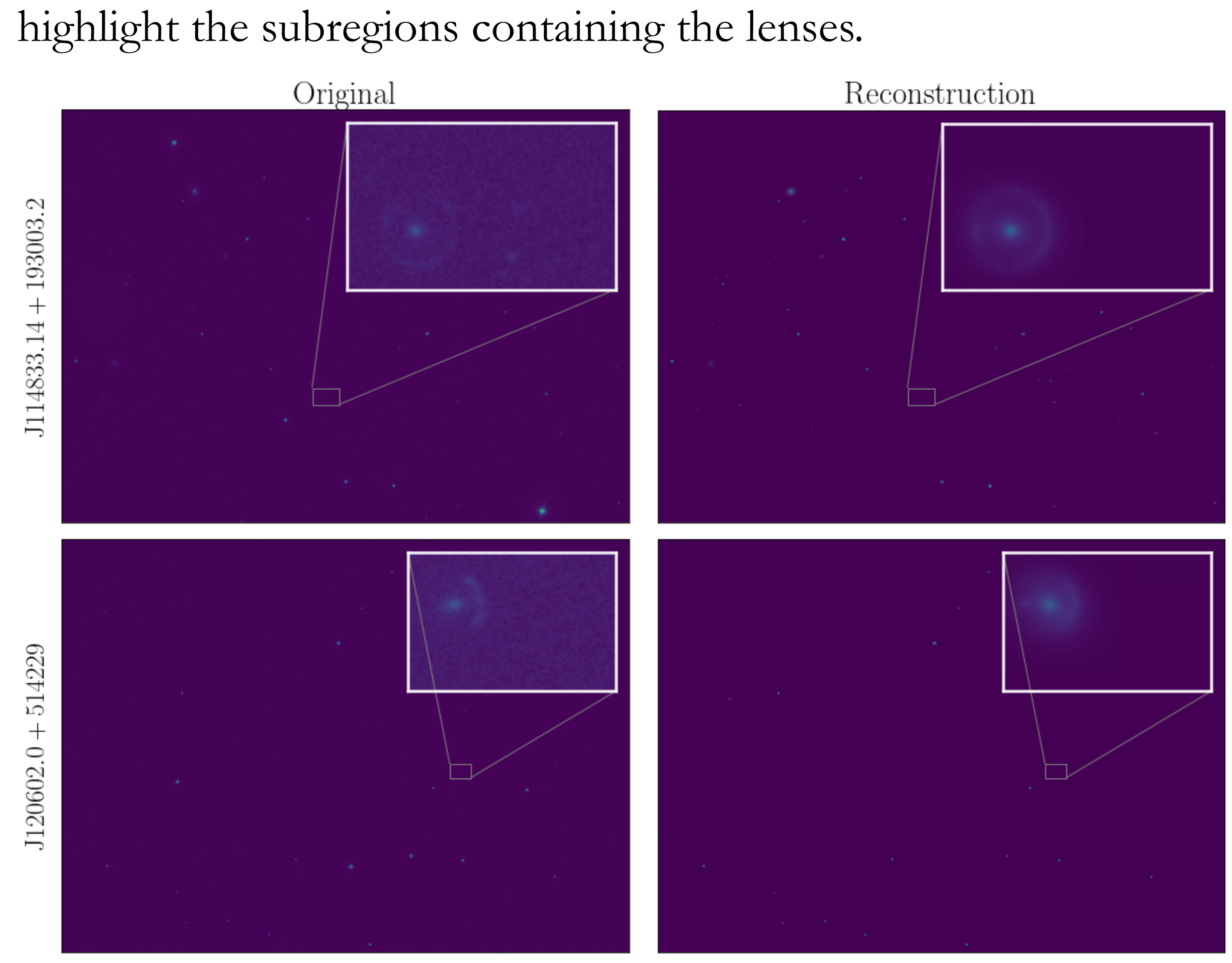


Uncertainty was **assessed with coverage percentages for 90% Bayes credible intervals.**

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 %		

Task II: Real Data

We additionally **apply our model to two SDSS images**. We **demonstrate successful detections** in the figure below, importantly achieved **without false positives**. The images were both **1489 × 2048 pixels and inference required just 25 seconds** for each image. The **left column shows the original images from SDSS and the right our reconstructions**. The zoom box is to highlight the subregions containing the lenses.



Conclusion

Several extensions of this research are possible. One is the use of this approach to infer weak lensing events. Characterization of dark matter substructures is also of great interest and would similarly require extensions to the SIE model employed here.