



Likelihood Free Parameter Estimation for Kilonovae Light Curves

Malina Desai

A3D3

June 27, 2024

Outline

01

Motivation

02

Data

03

ML Methods

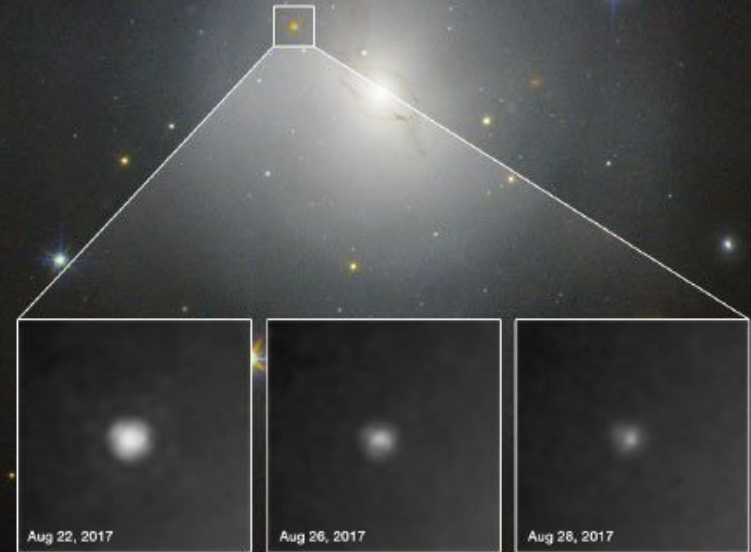
04

Results

What are Kilonovae?

- EM radiation from mergers of compact binaries: two neutron stars (BNS), or a neutron star with a black hole (NSBH)
- Short-lived events that are difficult to detect
- Important to our understanding of heavy element synthesis, nuclear equation of state

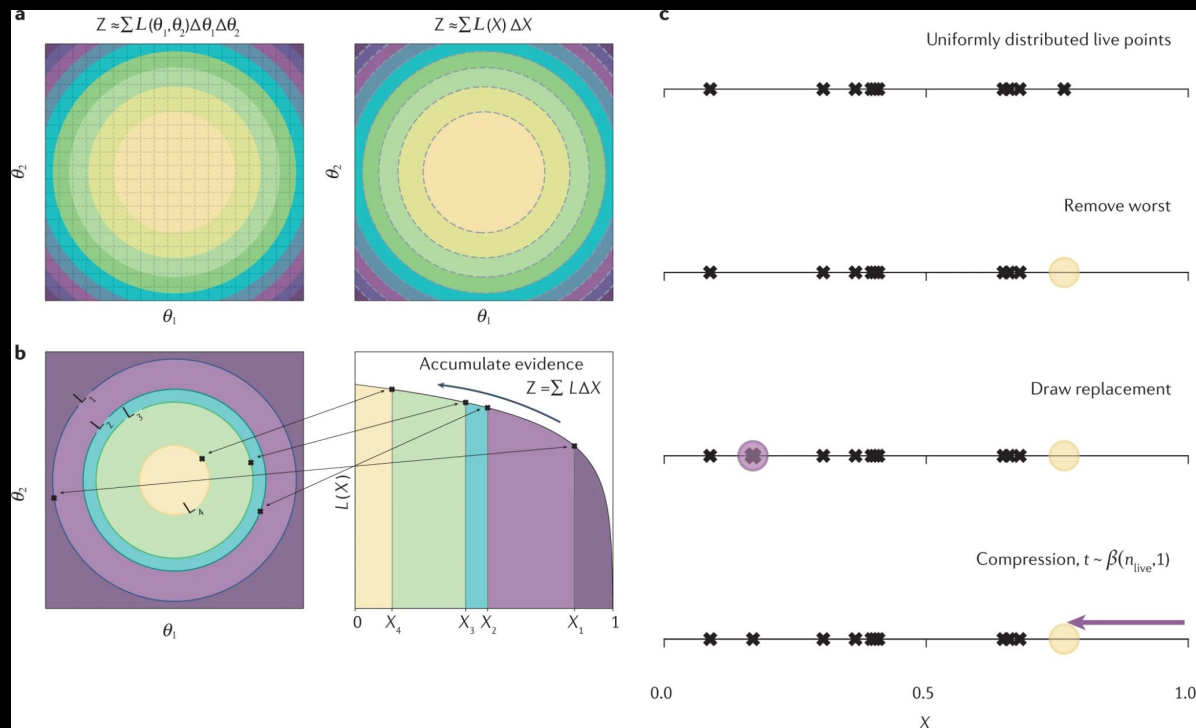
AT2017gfo¹



¹Abbott, B. P. et al., PRL, 119, 161101

How is current parameter estimation done?

- Through Bayesian inference
- We use PyMultiNest to perform nested sampling, resulting in a posterior distribution²
- Right: nested sampling schematic³



²Buchner, J. et al., 2014 A&A.

³Ashton, G., Bernstein, N., Buchner, J. et al. *Nat Rev Methods Primers* 2, 39 (2022).

Why improve parameter estimation?

01

Avoid computationally expensive likelihood calculation at each inference step

02

Increase in signal complexity results in increased computing cost

03

Real-time inference is transferable to other parameter estimation problems

Intrinsic Parameters of the Signal

01

Mass

02

Velocity

03

Lanthanide Fraction

Extrinsic Parameters of the Signal

01

Arrival Time

02

Luminosity Distance

Data Generation



- Using the Kasen 2017 hydrodynamical grid model⁴
- Simulate light curves as seen by ZTF in the g,r,i bands using the NMMA package⁵
- encode non-detections as ZTF limiting magnitude 22.0 in all bands

Data Processing

- ensure all light curves have the same number of points
- limit data to a minimum of 8 detections across all bands

Challenge: Data generation is time-consuming and heavy

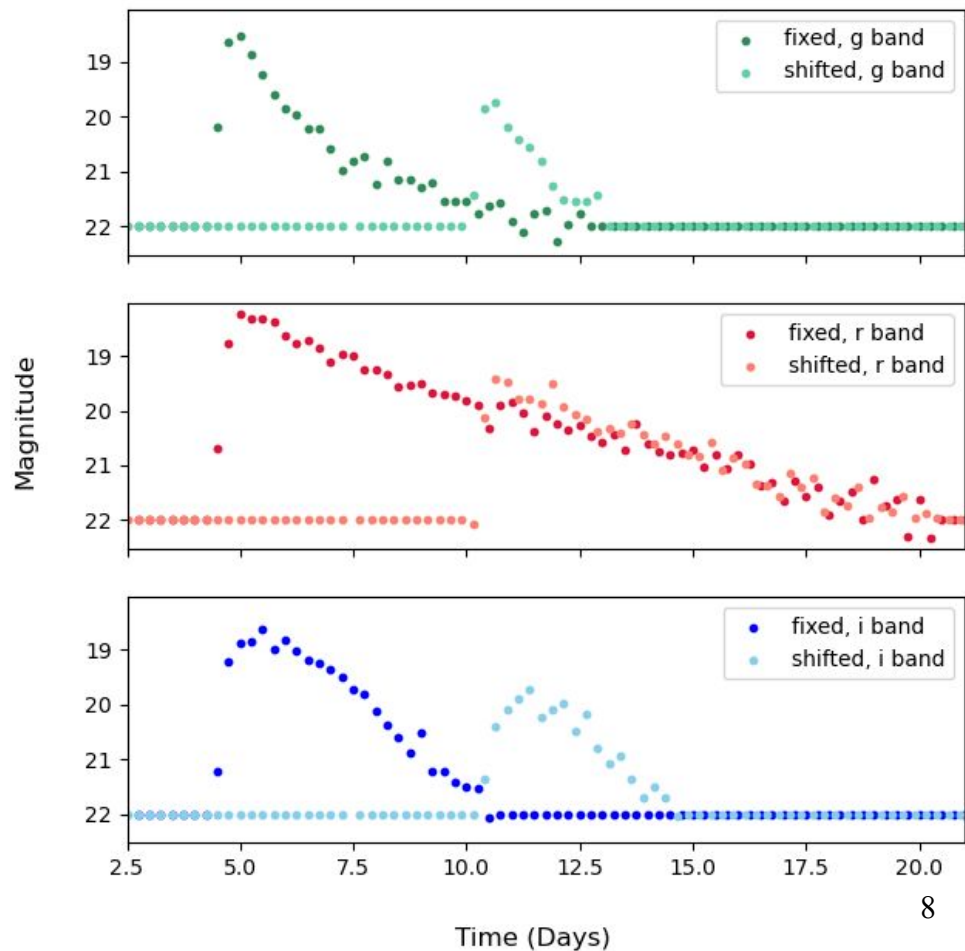
⁴Kasen et al. Nature, 551(7678):80–84, Nov. 2017.

⁵Pang et al. Nature, 14(1):8352, Dec. 2023.

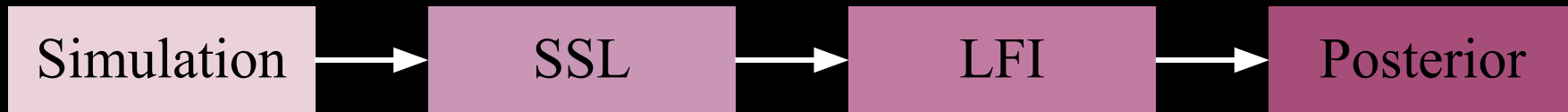
Fixed vs. Shifted

- Unique noise instance for each light curve
- All fixed data generated at $t=0$ days, $d_L = 50$ Mpc
- Right: fixed and shifted ($t=+5.65$ days, $d_L = 86.3$ Mpc) light curves shown for all 3 ZTF bands for the same injection parameters

Light Curve for $\log_{10}(M_{ej}): -1.86$, $\log_{10}(V_{ej}): -1.03$,
 $\log_{10}(X_{lan}): -7.97$



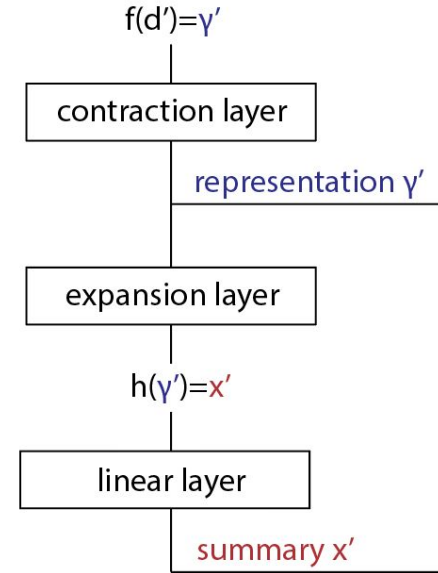
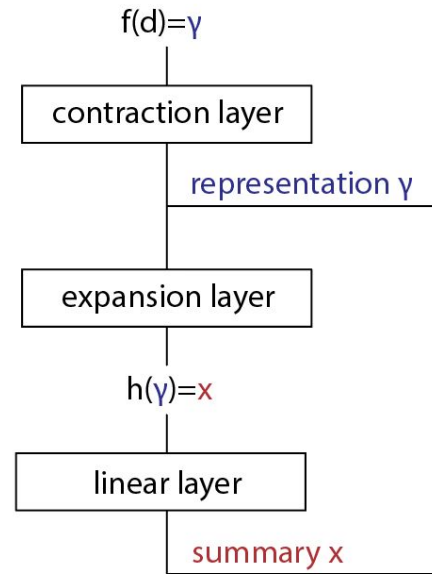
Implementing Likelihood Free Inference (LFI)



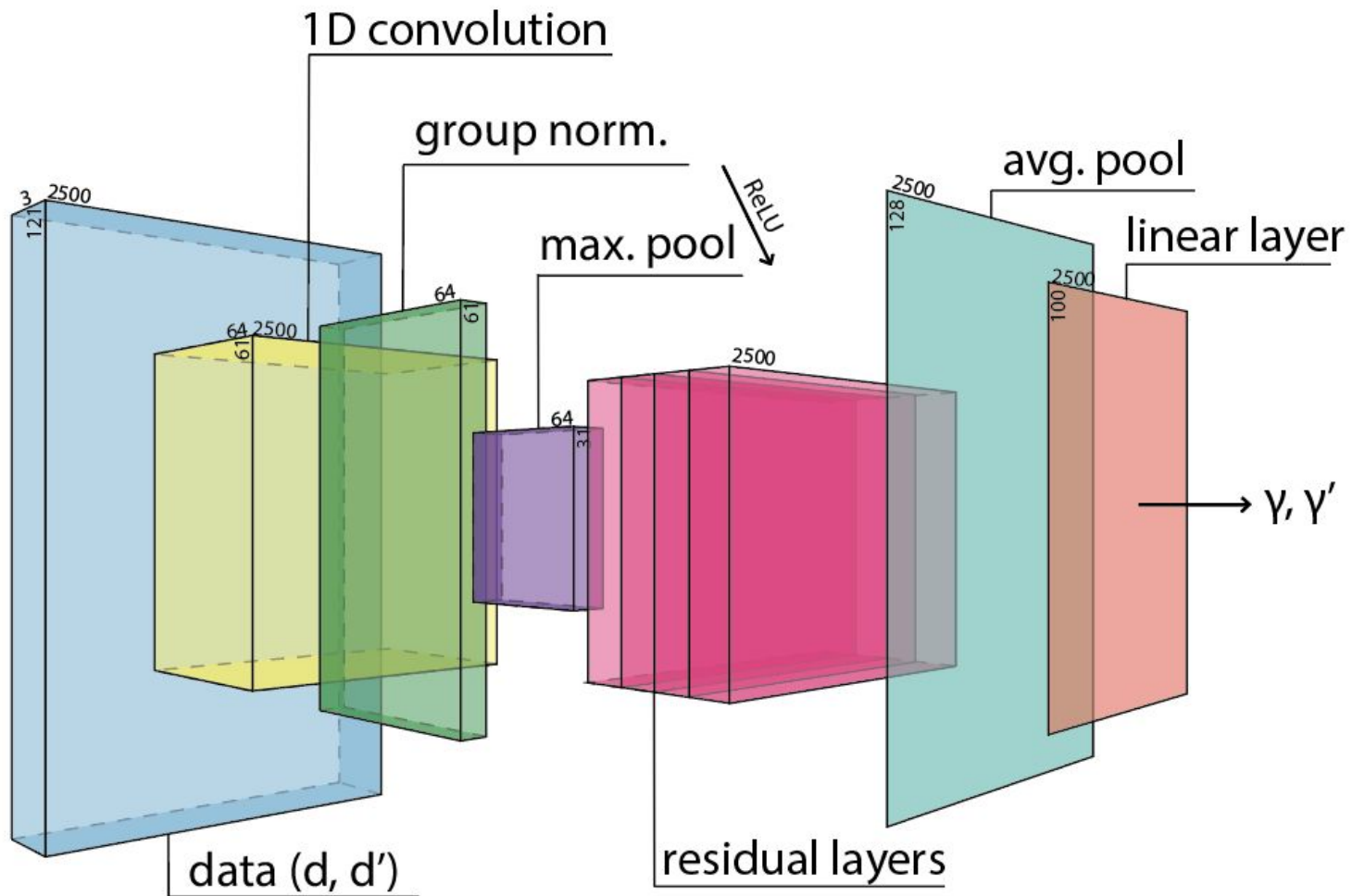
- Avoid likelihood calculation
- We use self-supervised learning (SSL), where the model generates labels for the data, to create a representation
- We use normalizing flows as our LFI to produce a posterior distribution

Data Summary Γ

$$\Gamma \equiv h \circ f$$



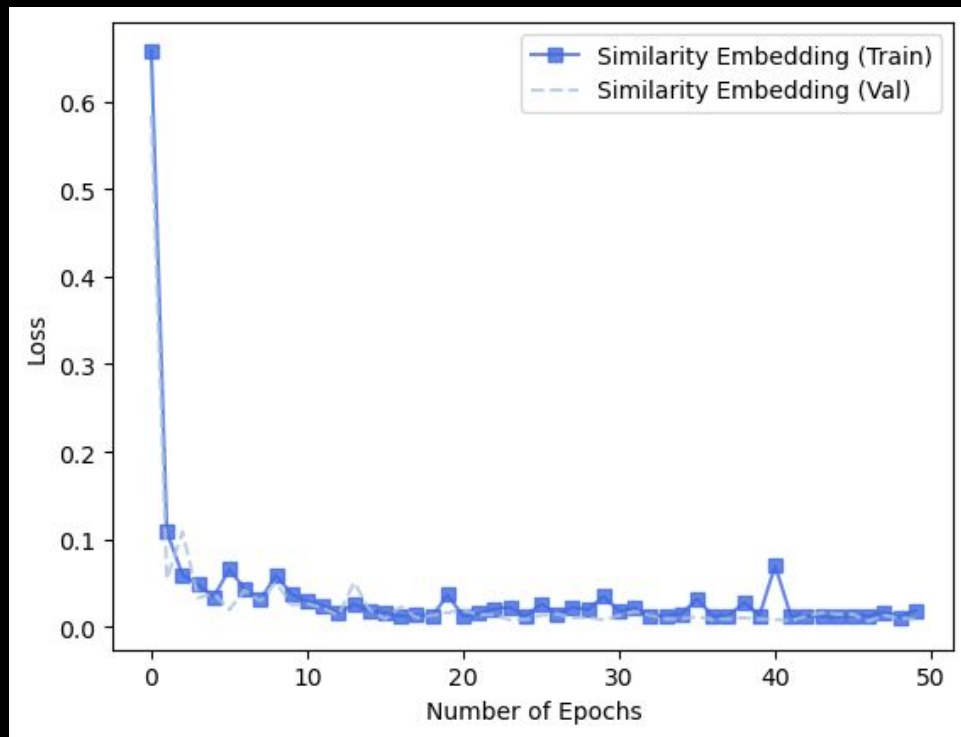
$$f(d, d') =$$



Variance-Invariance-Covariance Regularization Loss

$$\mathcal{L} = \text{Var} + \text{Invar} + \text{Covar}$$

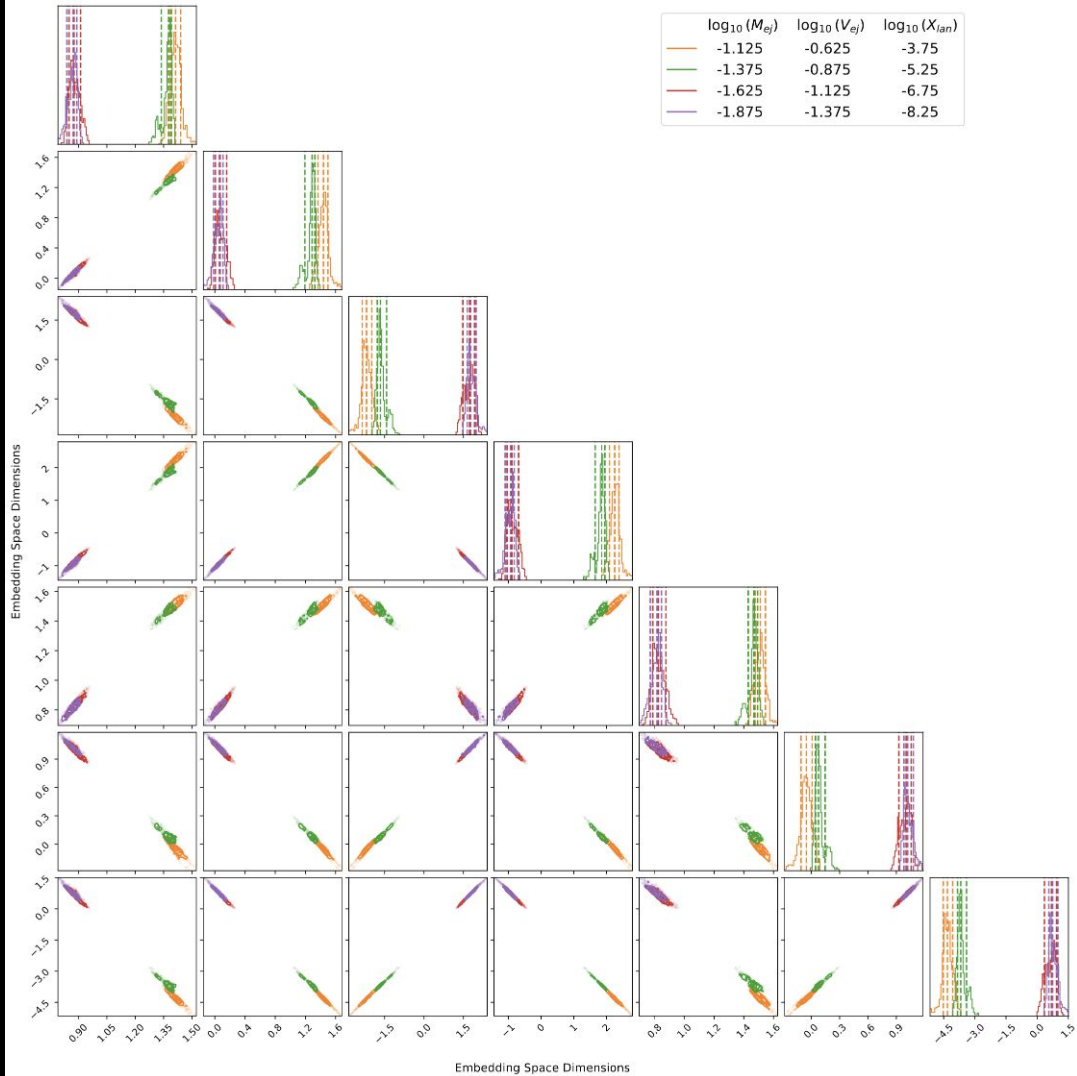
- Use a weighted loss function composed of 3 parts⁶
- Invariance is captured through MSE
- Right: training loss for the embedding
- 20 mins to train on GPU



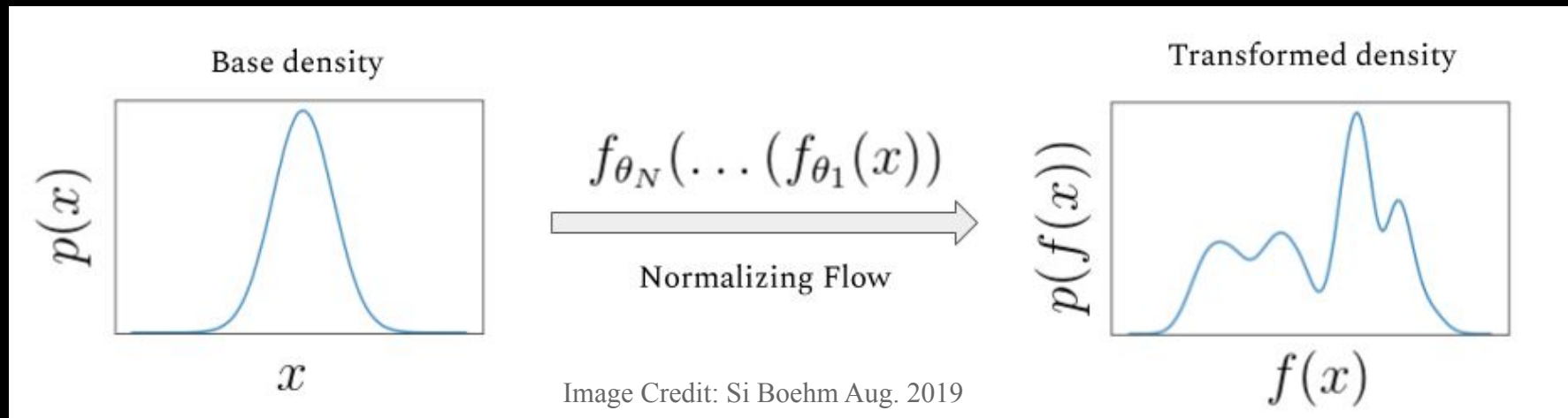
⁶A. Bardes, J. Ponce, and Y. LeCun., ICLR 2022.

Similarity Embedding

- Parameters are embedded in a 7-dimensional space to summarize data
- Dimensionality and size of network determined by hyperparameter tuning
- Right: displays clustering of similar injection parameters regardless of t and d_L



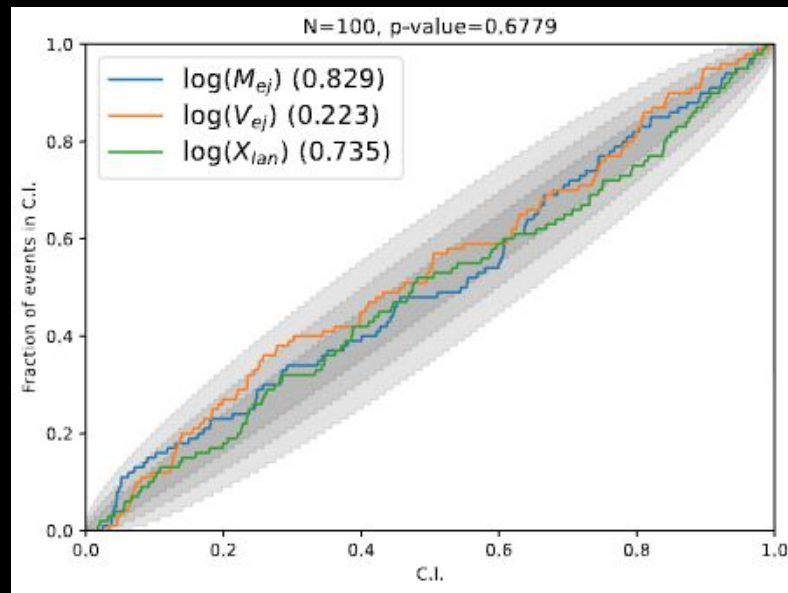
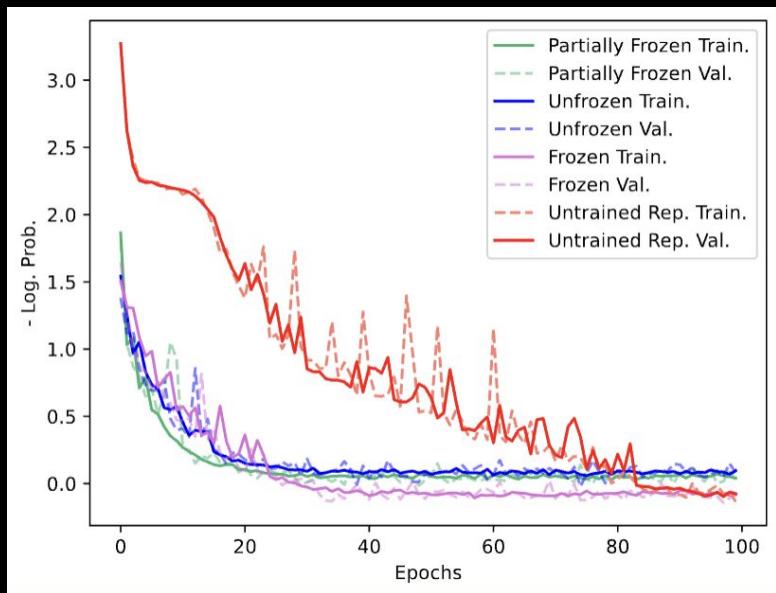
Normalizing Flows



- Transforms standard normal distribution to a complex sampling distribution⁷
- We use Masked Affine Autoregressive Transforms
- Not informed of the time and distance shifts

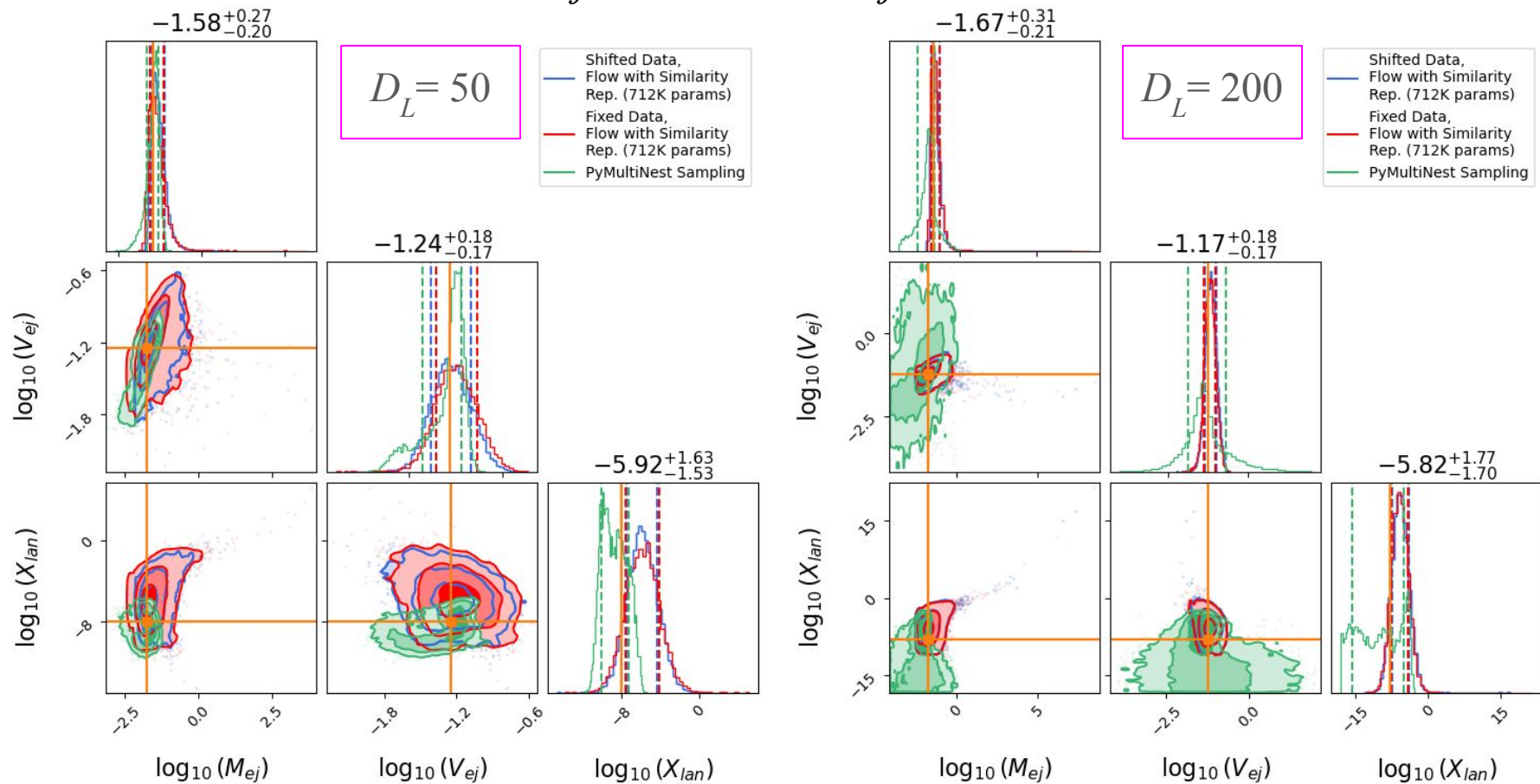
⁷Papamakarios et al. Journal of Machine Learning Research 22, Mar. 2021.

Loss Curve and PP Plot



- Left: Training with the embedding converges faster
- Training takes ~50 mins
- Right: Test sampling remains within 3 standard deviations

Results for $M_{ej} = -1.75$, $V_{ej} = -0.75$, $X_{lan} = -8$



Broader Applications

- Similar infrastructure can be used for sky localization of gravitational waves, other physics problems
- Works very well for simple waveforms such as simple harmonic oscillator and sine-gaussian pulse⁸

Life Here





Questions?

mmdesai@mit.edu