Amortized variational inference for supernovae light curves

Alexis Sánchez, Guillermo Cabrera, Pablo Huijse, Francisco Förster

ALeRCE
Automatic Learning for the Rapid Classification of Events

INSTITUTO MILENIO DE ASTROFÍSICA

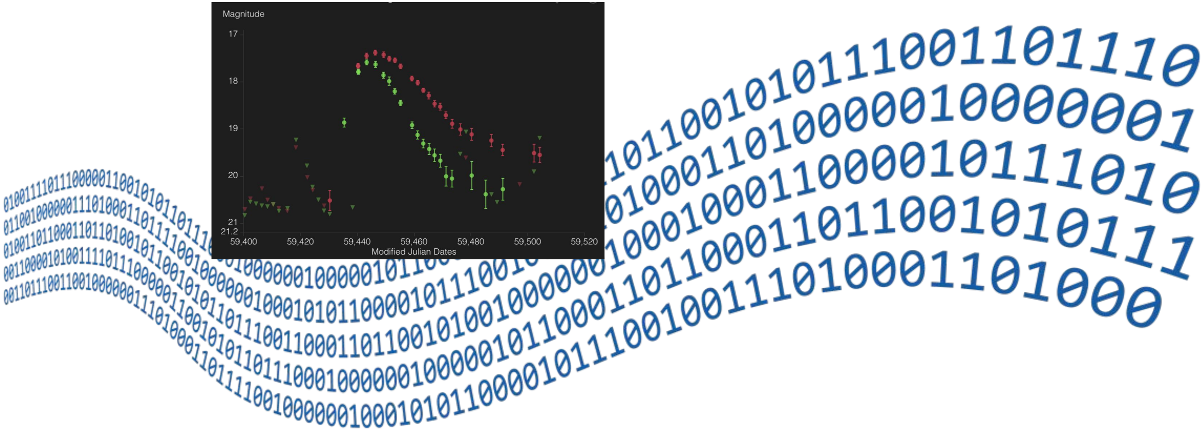
LIBERTAS CAPITUR
UNIVERSIDAD AUSTRAL DE CHILE

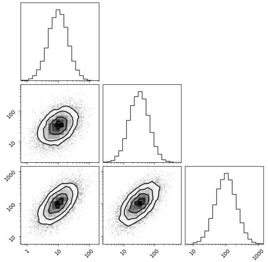
Introduction

Supernovae **characterization** is key to **understanding** their nature.

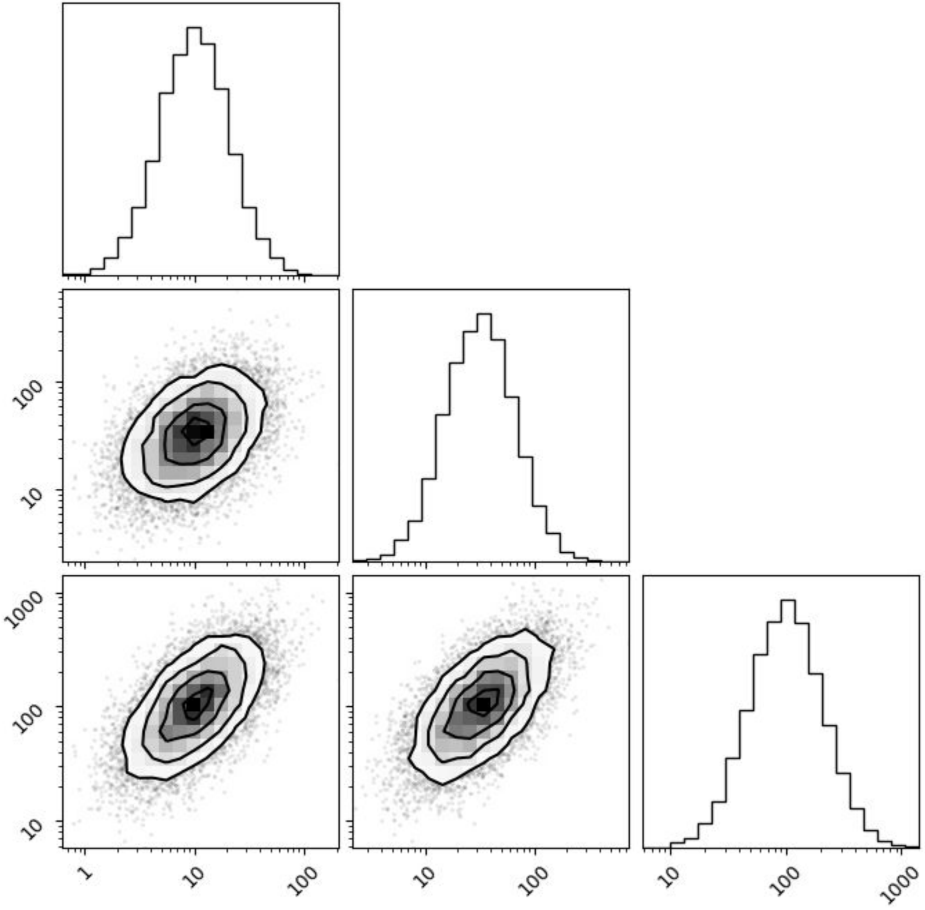
Reliable estimates can be obtained with **parametric models** and **Bayesian inference**.

But Bayesian methods are **unfeasible** for **real-time analysis** of alert streams.





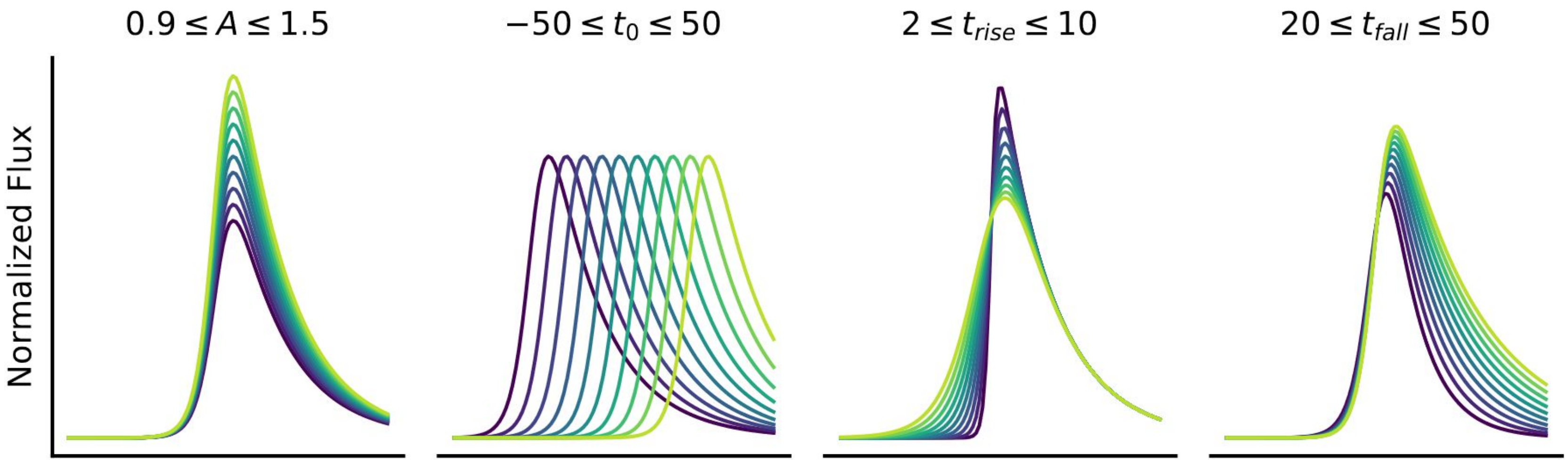
We are interested in the posterior distribution of the parameters



Supernova parametric model

We consider a model with 4 parameters for type Ia supernovae (Bazin+2009)

$$A\,\frac{e^{-(t-t_0)/t_{fall}}}{1+e^{-(t-t_0)/t_{rise}}}$$



Inference

Ideally, we would use Markov Chain Monte Carlo (MCMC) for Bayesian Inference

- However this is not ideal for real time inference or large amounts of data.

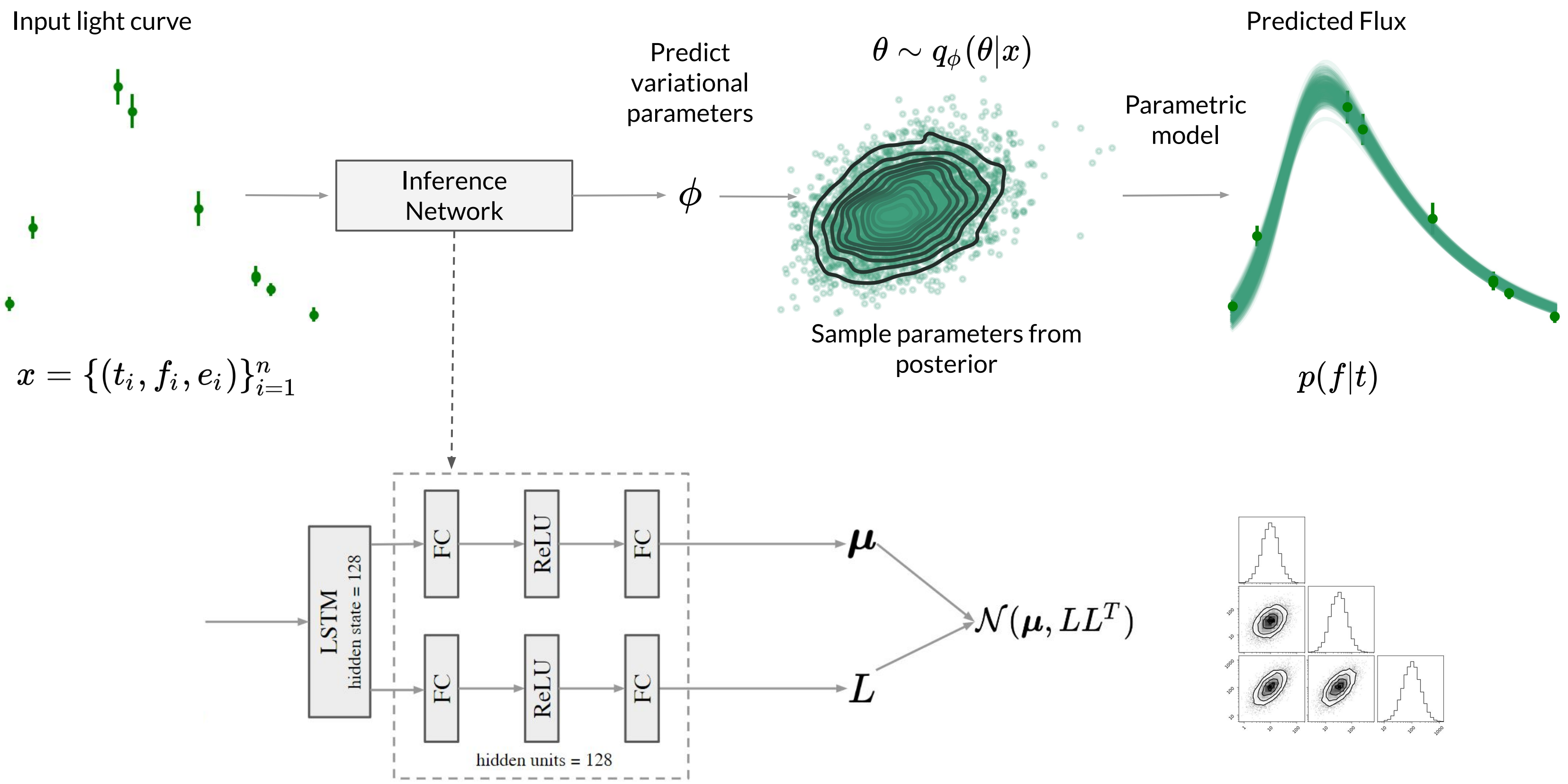
We can consider other approximate methods, such as Variational Inference (VI):

- Define an approximate posterior distribution and find the parameters that optimize the ELBO:

$$q_\phi(\theta)$$

$$\mathbb{E}_{q_\phi}\left[\log p(D|\theta)\right] - D_{KL}(q_\phi(\theta)||p(\theta))$$

Getting fast posteriors from light curves

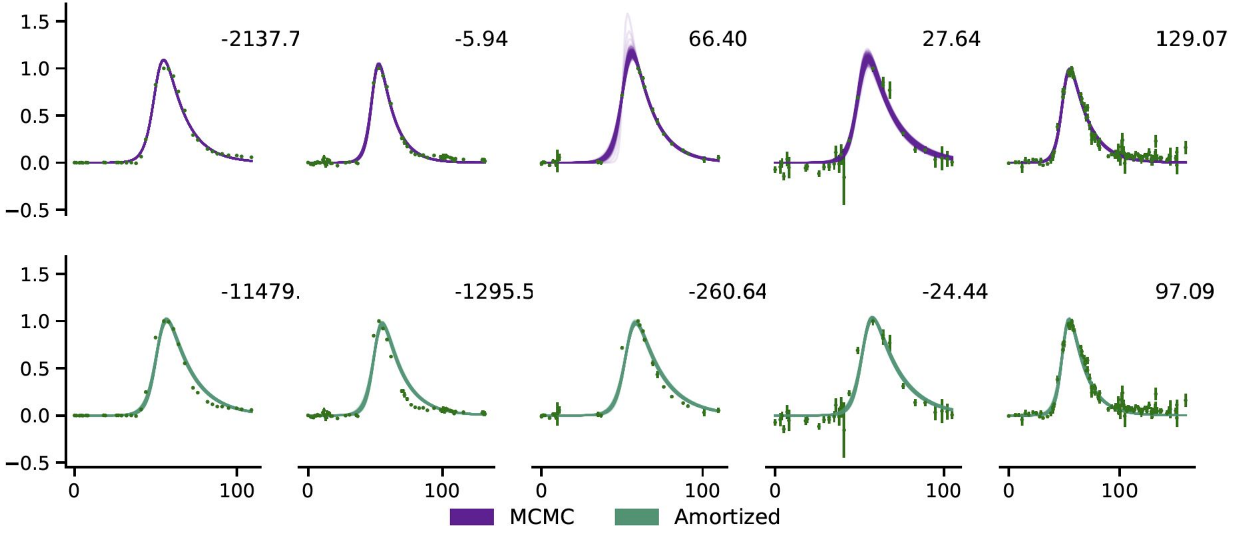


Results

We compare our results with the No-U-Turn Sampler (NUTS) and AVI on real and simulated (from prior) data:

- Fits
- Execution Time
- Difference in parameter estimation

Results



Fits on real type Ia supernovae light curves (ZTF g-band)

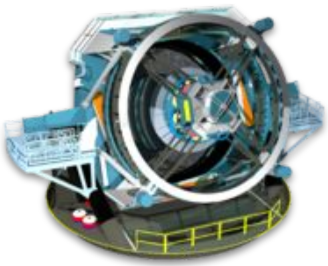
Results

Timing comparison

10,000 LCs:

Method	Training	Inference
MCMC	–	$1,468.27 \pm 16.01$
AVI (CPU)	$5,737.96 \pm 870.76$	133.44 ± 3.53
AVI (GPU)	539.04 ± 151.27	4.57 ± 0.05

LSST alert rate:
~10,000 LCs / 30 s.



Results

	Synthetic data		Real data	
Percentile	MCMC	Amortized	MCMC	Amortized
5%	−382.60	−764.96	−2112.87	−11227.21
50%	38.26	3.38	31.37	−261.57
95%	122.63	90.66	291.32	96.09

Table 1. Percentiles of the average log-likelihood distribution.

Parameter	MCMC	Amortized
A	0.063	0.066
t_0	0.556	0.549
τ_{fall}	2.204	2.577
τ_{rise}	0.214	0.277

Table 3. Median absolute deviation between the median marginal posterior and the parameter used for the simulation.

Results

Difference between MCMC and AVI posteriors

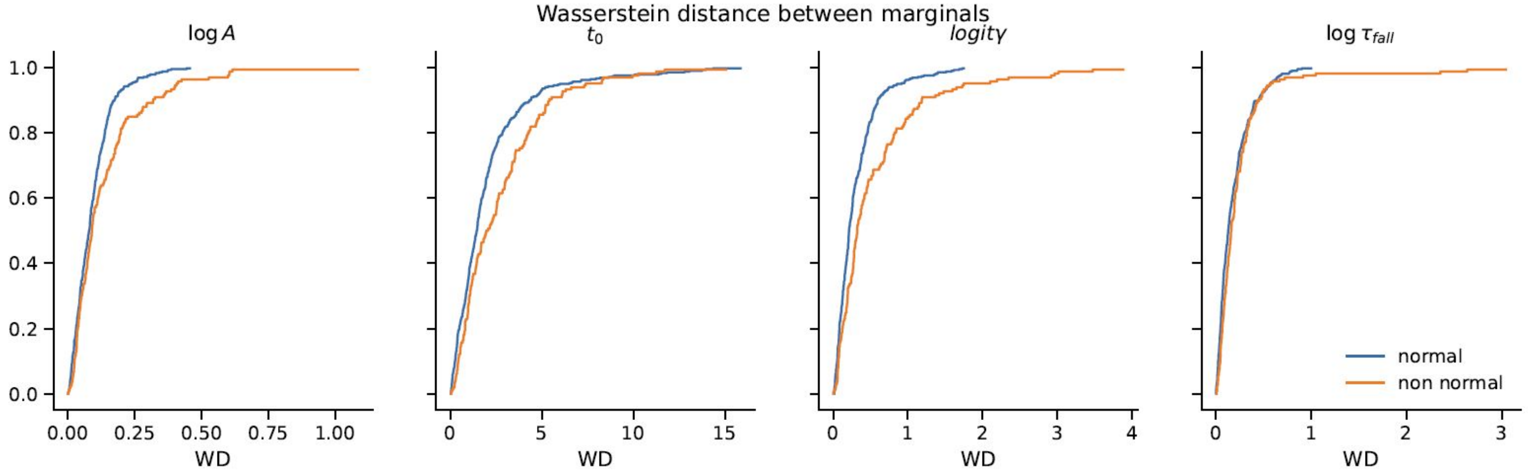


Figure 7. Cumulative distribution function of Wasserstein distances between the MCMC and amortized marginal posteriors separating those cases where the MCMC posterior is normal or not. We see that the only case where the distribution of distances is significantly larger for non-normal posteriors is in the logit γ marginal posteriors.

Results

Main difference is in the parameter that models rise of the light curve, MCMC is able to model lower values.

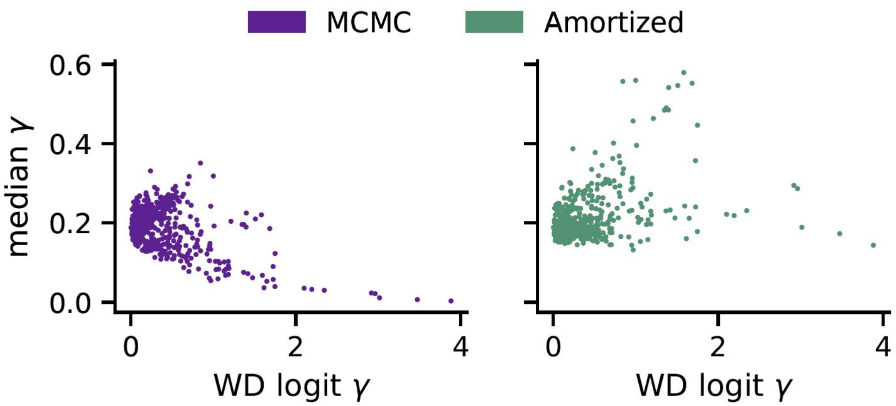


Figure 8. Wasserstein distance between logit γ and the median of the marginal posterior of γ.

Summary

Parameter inference is useful for studying supernovae.

We can obtain **approximate posteriors** with **amortized variational inference (AVI)**.

AVI is much **faster** than MCMC, allowing online **real time inference** for e.g. LSST.

Our approach can be **generalized** to other phenomena by using the appropriate parametric model (e.g., Zhang et al. 2021 for microlensing events).

Thank you!