FERMILAB-SLIDES-22-051-SCD

DeepQuasar: Simulation-based inference for lensed quasar modeling

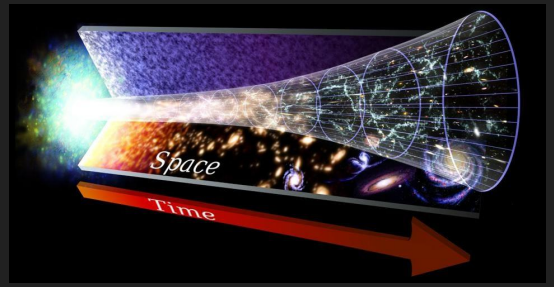
E. Danilov[1,*], A. Ciprijanovic[2] and B. Nord[2,3,4]

1. Ecole Polytechnique Federale de Lausanne (EPFL), Switzerland
2. Fermi National Accelerator Laboratory, USA
3. Department of Astronomy and Astrophysics, University of Chicago, USA
4. Kavli Institute for Cosmological Physics, University of Chicago, USA

Research supported by Deep Skies Lab
Funding by Fermilab and U.S. Department of Energy
Contact email: egorssed@gmail.com

Hubble constant H_0 in LCDM

Expansion of the Universe

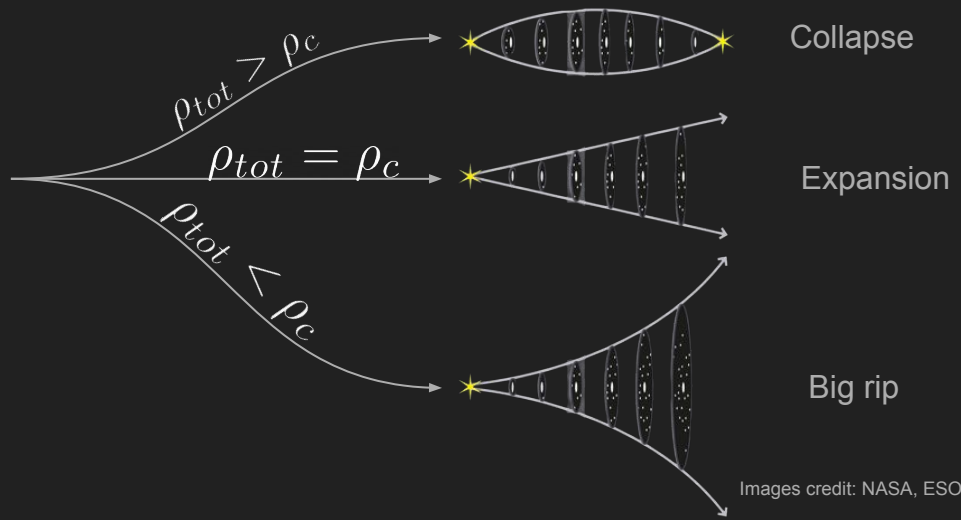


Expansion of the Universe $a(t) = a_0 e^{Ht} + \ldots$

Insight into dark energy $\Lambda = 3H_0^2 + \ldots$

Fate of the Universe

Critical density of the Universe $\rho_c = \dfrac{3H_0^2}{8\pi G}$



Images credit: NASA, ESO Supernova

Measurement discrepancies

Hubble rate H_0

Direct methods → $H_0 = 74.03 \pm 1.42\ km\ s^{-1} Mpc^{-1}$

Indirect methods → $H_0 = 67.4 \pm 0.5\ km\ s^{-1} Mpc^{-1}$

>3σ discrepancy

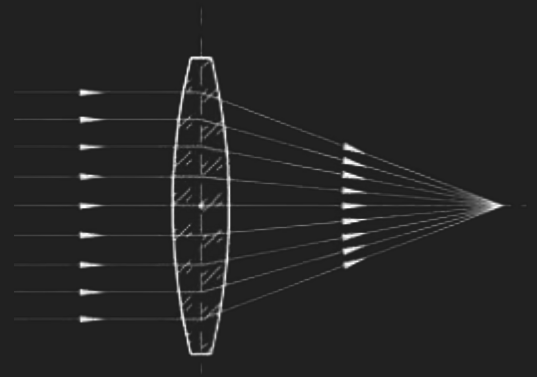
→ Mistakes in detectors

→ New physics

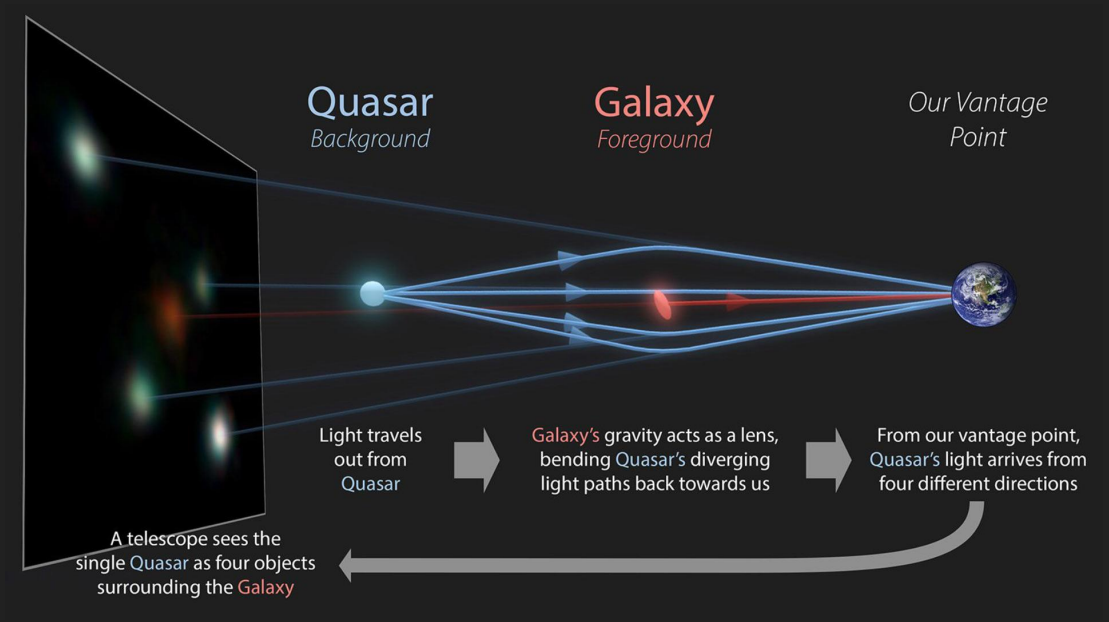
Solution: Time delay cosmology

Gravitational lensing

Ordinary lensing



Gravitational lensing

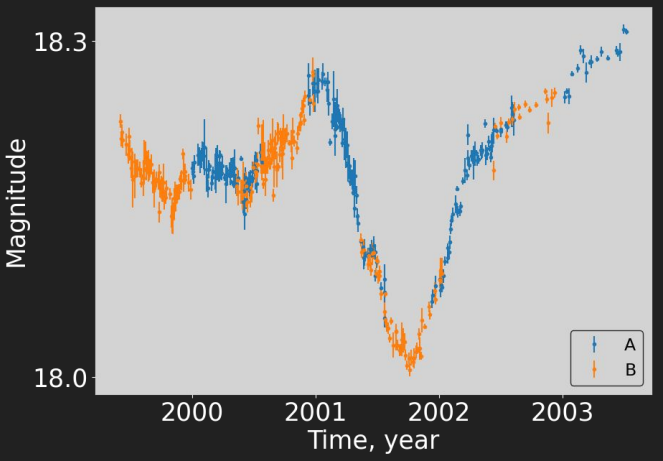


Time delay: $t(\vec{\theta}) = \frac{D_{\Delta t}}{c}\phi(\vec{\theta}) \propto \frac{1}{H_0}\phi(\vec{\theta})$

Image credit: R. Hurt, IPAC & Caltech / Gaia Gravitational Lenses group

Gravitationally Lensed Quasars (LQSO)

Quasar light curves

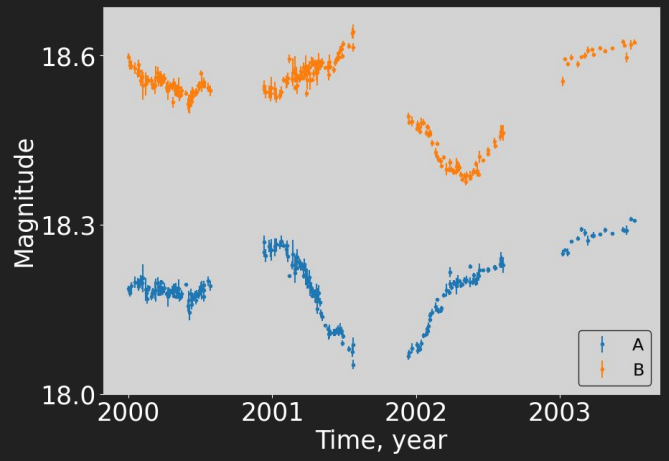


Lensing

Shifts: time and magnitude

$$\Delta t \propto \frac{1}{H_0}$$

Lensed quasar light curves



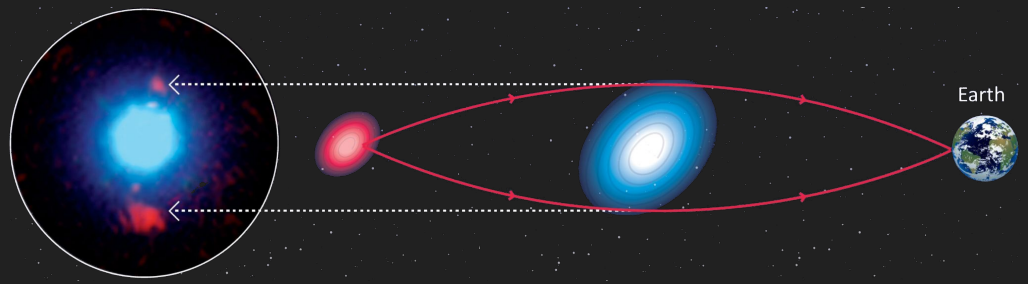


Image credit: fftech.net

Outline of the presentation

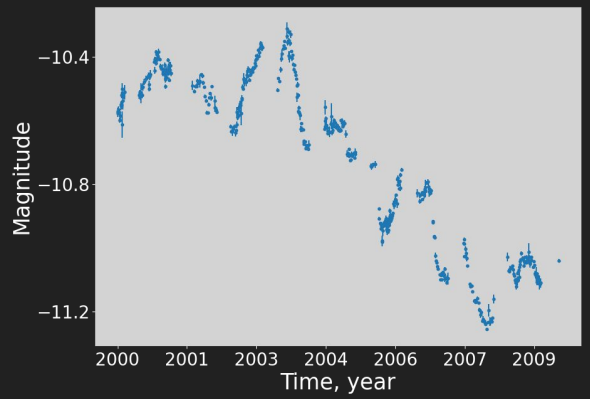
1. Challenges of inference
2. Neural networks-based method
3. Performance of neural networks

Challenges of inference

Quasar model - accretion + reverberation

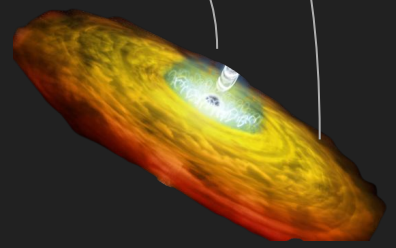
$$\underbrace{I(t)} \propto \int_0^\infty \underbrace{f_{acc}(t - t_\lambda \rho)}\underbrace{P_{disk}(\rho)} d\rho \sim \underbrace{\mathcal{N}(\mu, \Sigma[\hat{\sigma}^2, \tau, t_\lambda])}$$

Quasar light curve

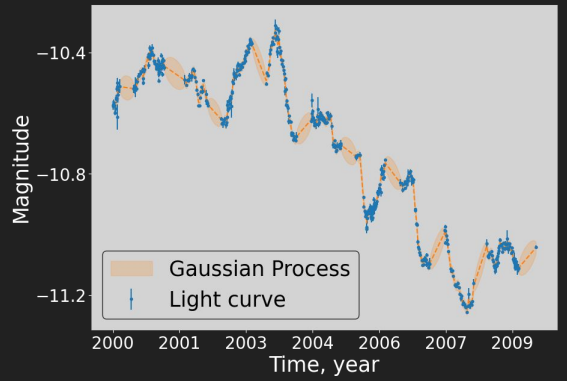


Accretion onto black hole

Reverberation of the disk



Gaussian stochastic process



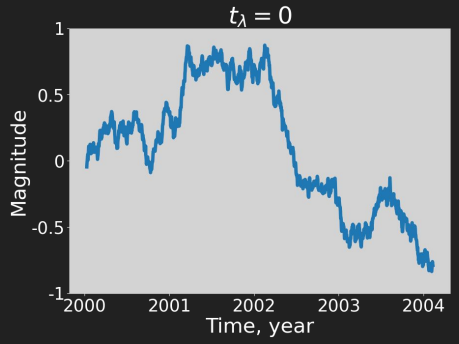
Images credit: NASA

Quasar model - Reverberating Damped Random Walk

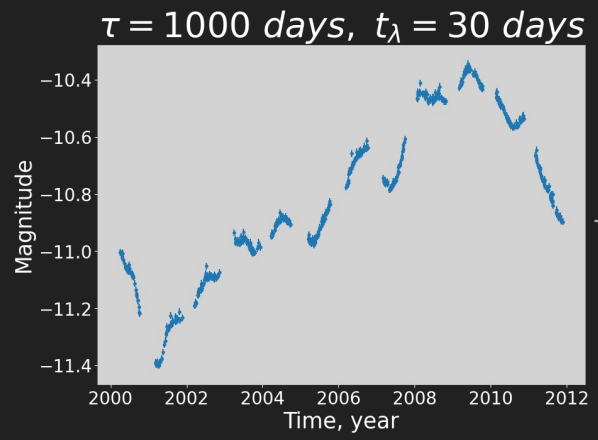
Autocorrelation: $$\int_{-\infty}^{\infty} f(t'+\Delta t) f(t') dt' \propto \tau \cdot exp\left(-\frac{|\Delta t|}{\tau}\right) - 4t_\lambda \cdot exp\left(-\frac{|\Delta t|}{4t_\lambda}\right)$$

Correlation time τ
global evolution (years)

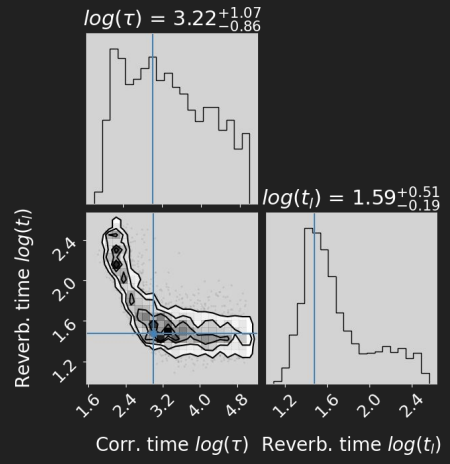
Reverberation time t_λ
local evolution (days)



Challenge of inference



Gaussian likelihood MCMC

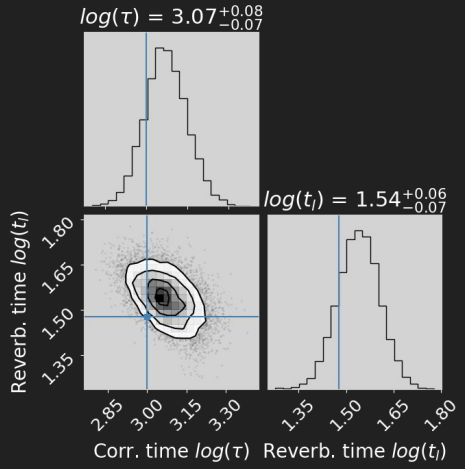


1σ constraints:

$$200 < \tau < 20000\ days$$

$$25 < t_\lambda < 125\ days$$

Neural Network

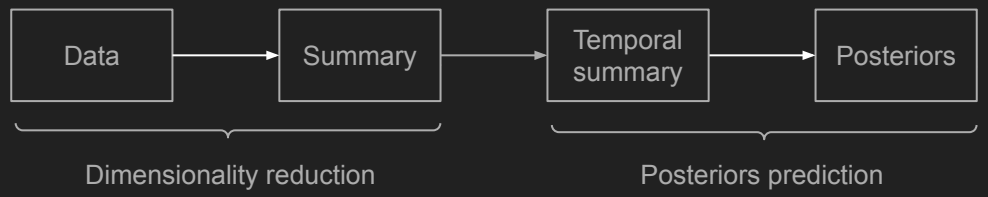


1σ constraints:

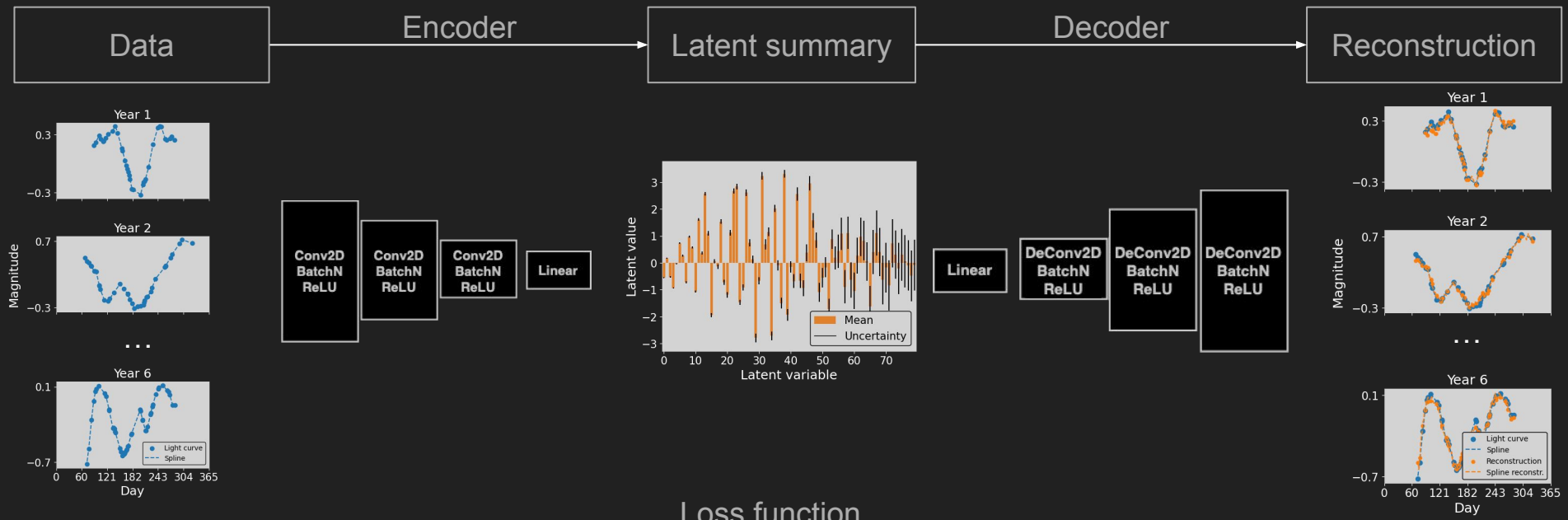
$$990 < \tau < 1410\ days$$

$$29 < t_\lambda < 40\ days$$

Neural networks-based method



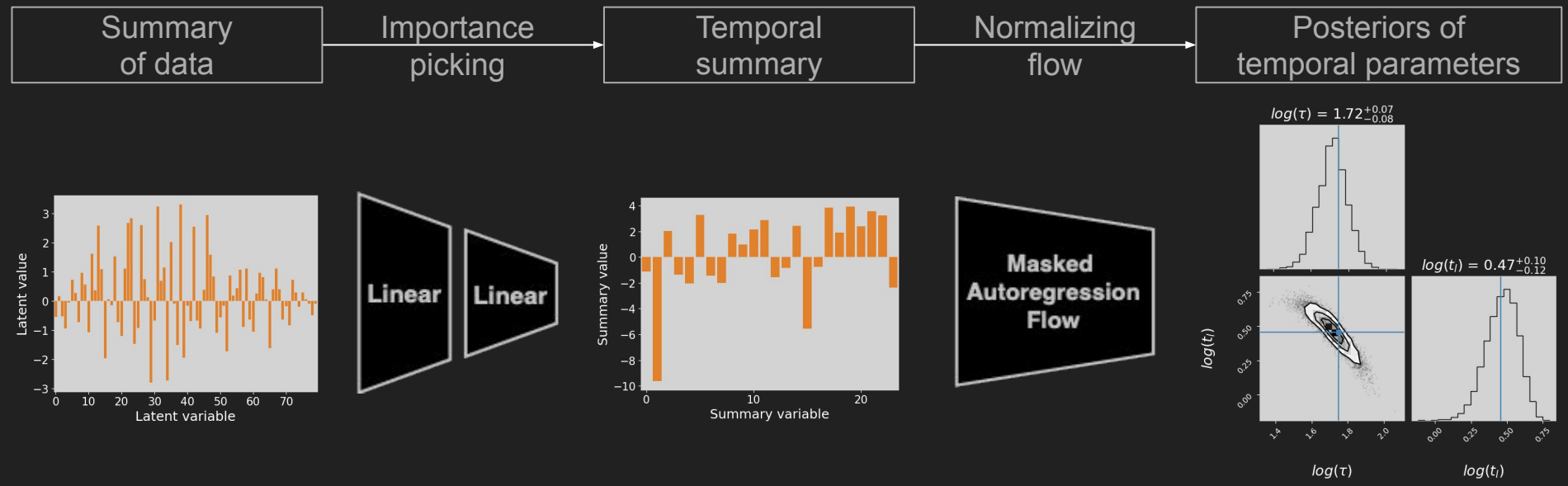
Dimensionality reduction with VAE



Loss function

$$\mathcal{L} = \underbrace{MSE(x, x')}_{\text{Quality of reconstruction}} + \beta \cdot \underbrace{D_{KL}(p(z|x)||p(z))}_{\text{Amount of information}}$$

Posteriors from Normalizing flow

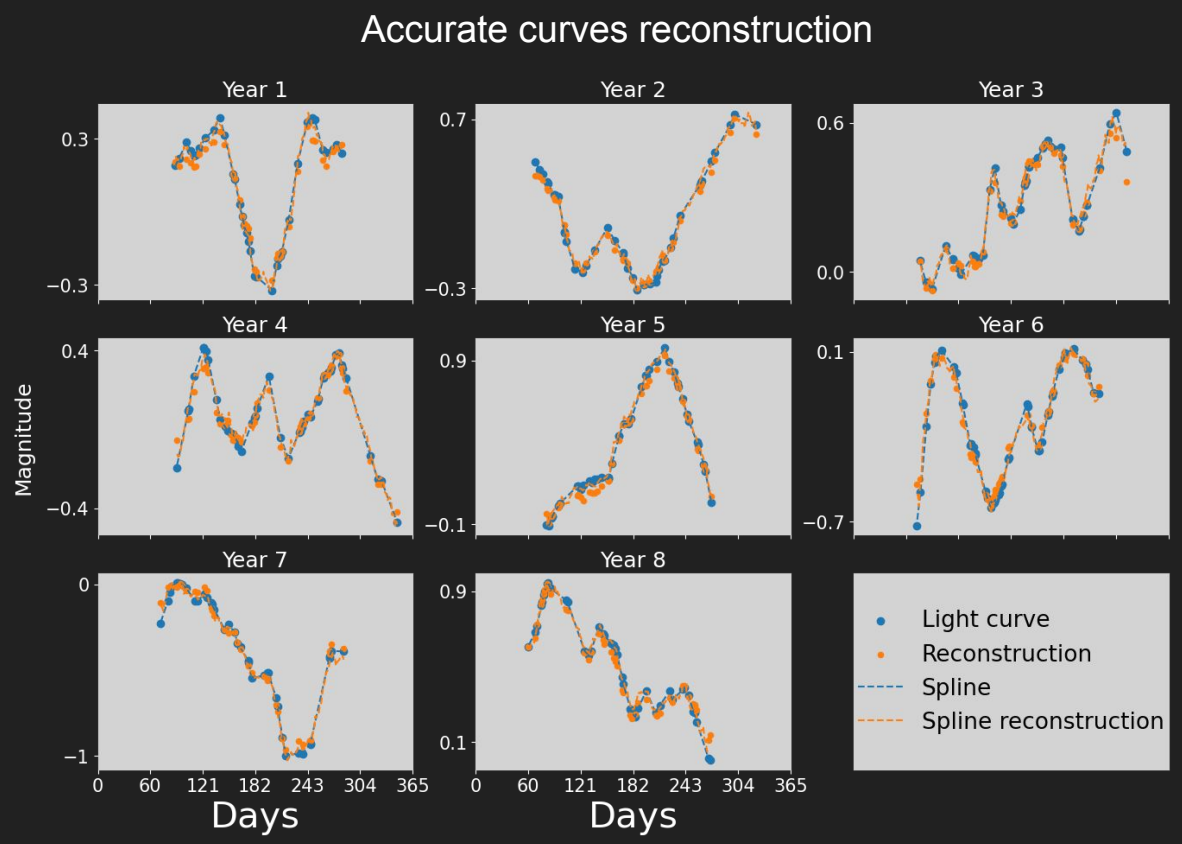


Result performance

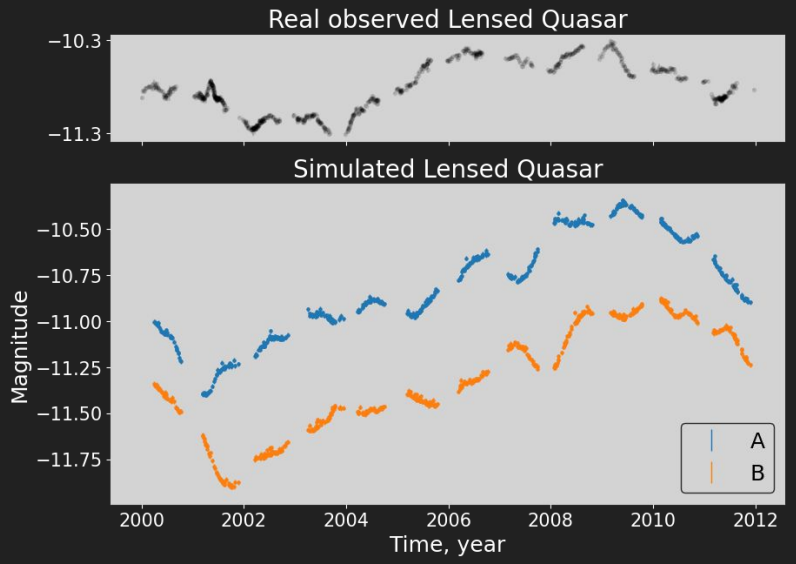
Variational autoencoder's reconstruction

Relative error: 4%

Error: RMSE=0.08



Quasar parameters inference



Truth:

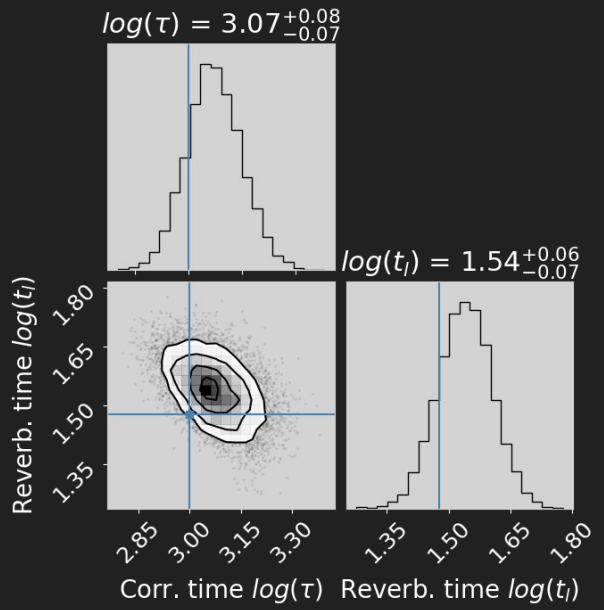
$\tau = 1000\ days$

$t_\lambda = 30\ days$

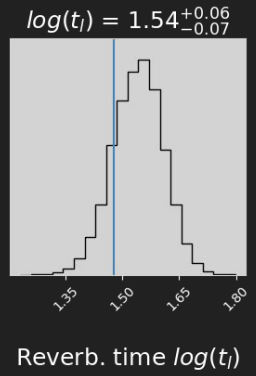
1σ constraints:

$990 < \tau < 1410$

$29 < t_\lambda < 40$



Errors of prediction



Prediction:

$$x = x_{pred} \pm \sigma_x$$

Absolute error:

$$x_{pred} - x_{g.t.}$$

Should be:

$$\Delta x \to 0$$

Normalized error (Chi):

$$\chi(x) = \frac{x_{pred} - x_{g.t.}}{\sigma_x}$$

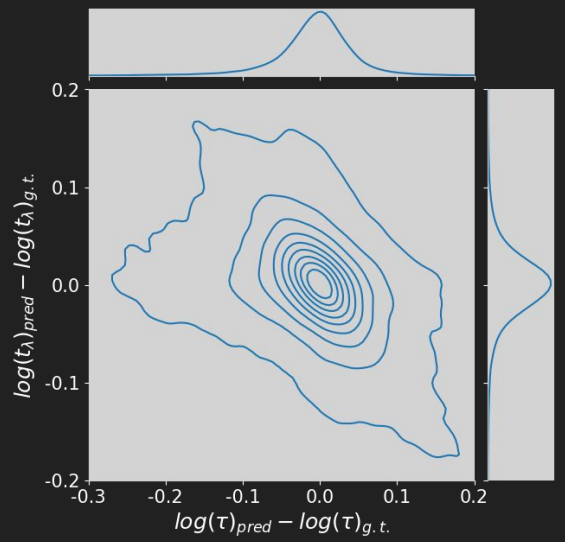
Should be:

$$\chi(x) \to \mathcal{N}(0,1)$$

Statistical distribution of errors

Distribution of absolute errors

$$x_{pred} - x_{g.t.}$$

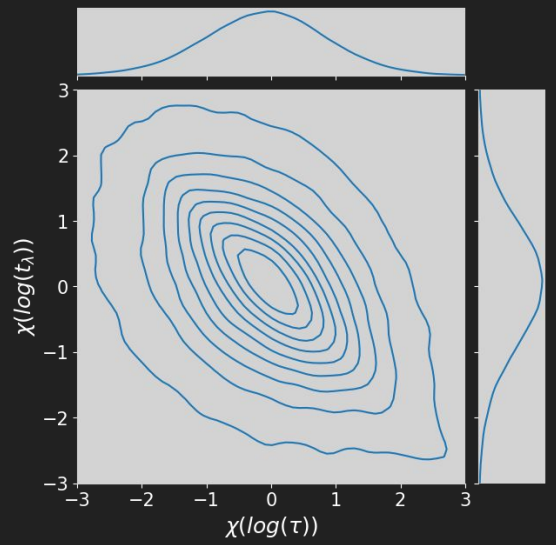


Conclusion: accurate prediction

$$0.5 < \frac{x_{pred}}{x_{g.t.}} < 2$$

Distribution of normalized errors

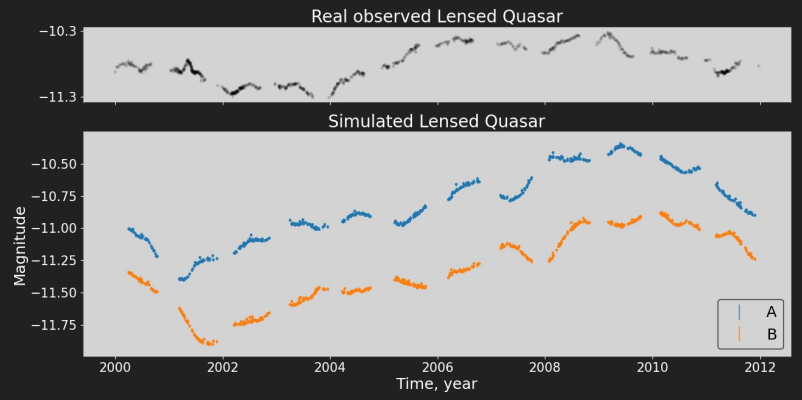
$$\chi(x) = \frac{x_{pred} - x_{g.t.}}{\sigma_x}$$



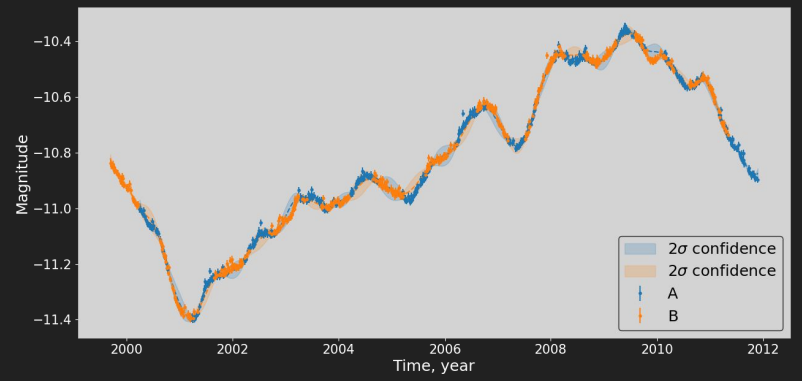
Conclusion: reliable uncertainties

$$\sigma_{pred} \approx 0.93 \cdot \sigma_{truth}$$

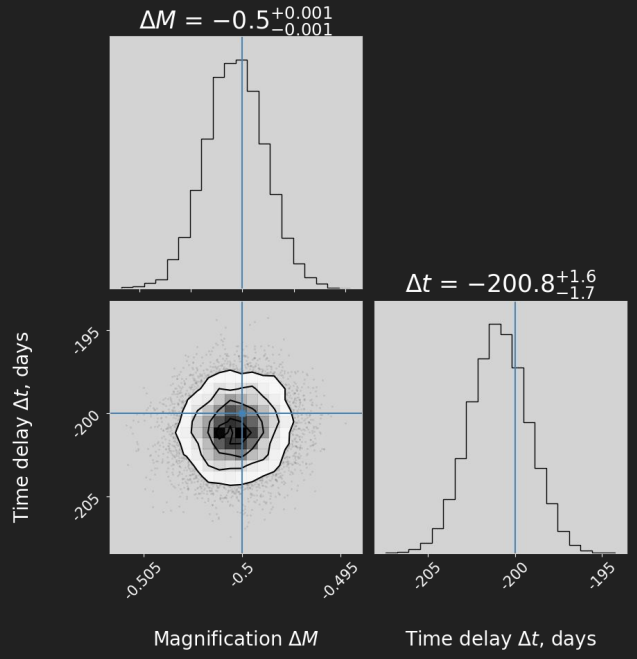
Lensing inference



Removed lensing effect



Posteriors of lensing parameters



- Negligible magnification error
- Accurate time delay prediction

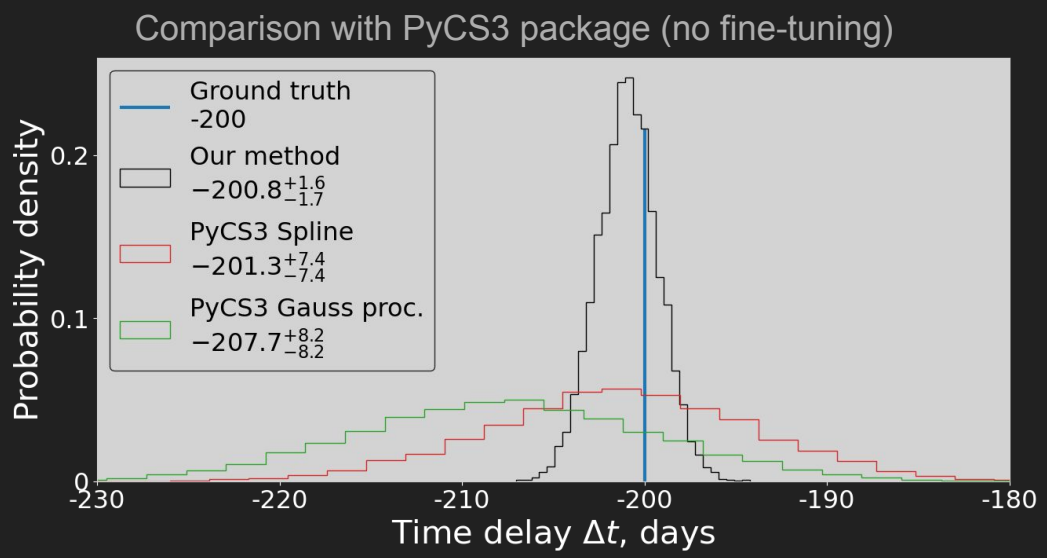
Comparison with standard software

PyCS3 setup:

- Adjusted necessary hyperparameters
- The rest left to default

PyCS3 warnings:

- does not guarantee realistic time-delay uncertainty estimates
- Please try to avoid publishing time delays with overly optimistic uncertainty estimates



Method	Inference time	Uncertainty
Our method	**20 sec**	**±1.6 days**
PyCS3 Splines	13 min	±7.4 days
PyCS3 Gauss proc.	9 min	±8.2 days

Package credit: Millon et al. 2020, Tewes et al. 2013

Improvement of Hubble rate H_0

Uncertainty from time delay:

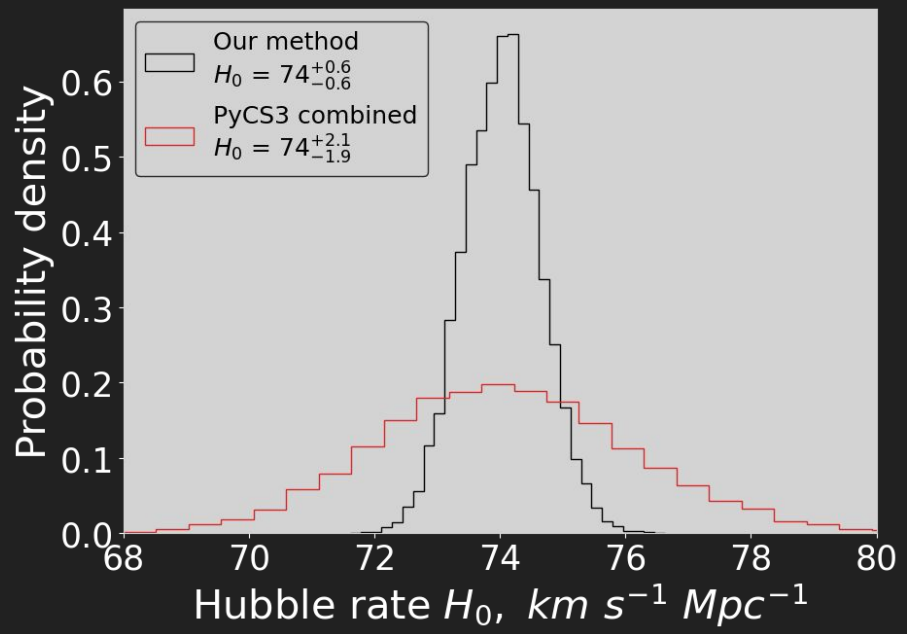
$$H_0 \pm \sigma_H \propto \frac{\Delta\phi_{g.t.}}{\Delta t \pm \sigma_t}$$

PyCS3: $\sigma_H = 2.1\ km\ s^{-1}\ Mpc^{-1}$

Our method: $\sigma_H = 0.6\ km\ s^{-1}\ Mpc^{-1}$

3.5 times less

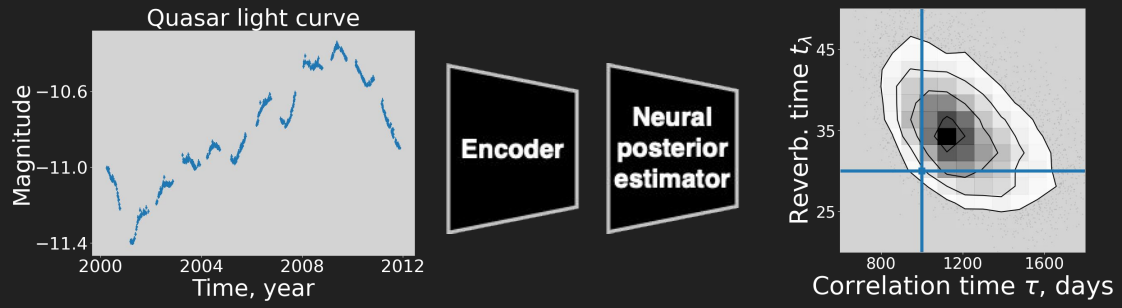
Comparison with PyCS3 package (no fine-tuning)



Summary

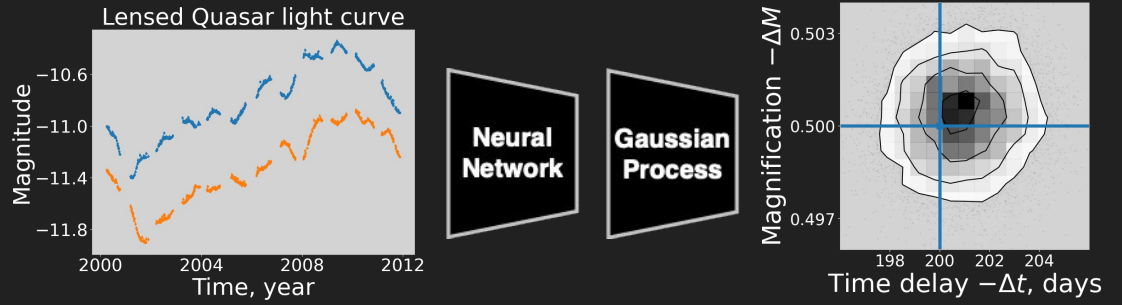
Quasar light curves

- Relative error: $0.5 < \frac{x_{pred}}{x_{g.t.}} < 2$
- Uncertainty $\sigma_{pred} \approx 0.93 \cdot \sigma_{truth}$



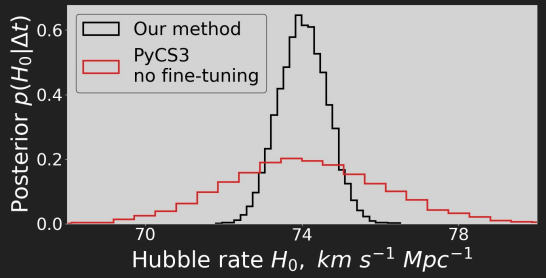
Lensing inference

- Magnification uncertainty 0.001
- Time delay uncertainty ±1.6 days



Hubble rate improvement

- Uncertainty from time delay 0.6
- Uncertainty improved in 3.5 times



Contact email:
egorssed@gmail.com

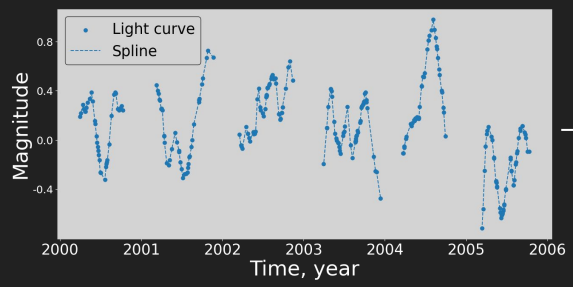
Thank you for your attention

Contact email: egorssed@gmail.com

Additional materials

Data preprocessing

Target parameter	Time scale	Image axis
Correlation time $\mathcal{T}$	Years	Y axis
Reverberation time t_λ	Days	X axis



Annual split
Spline interpolation

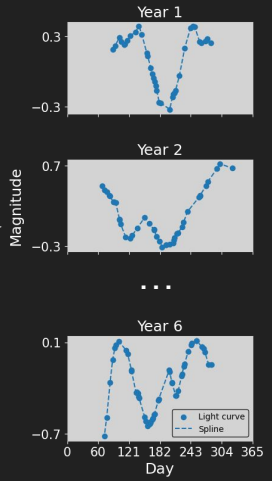
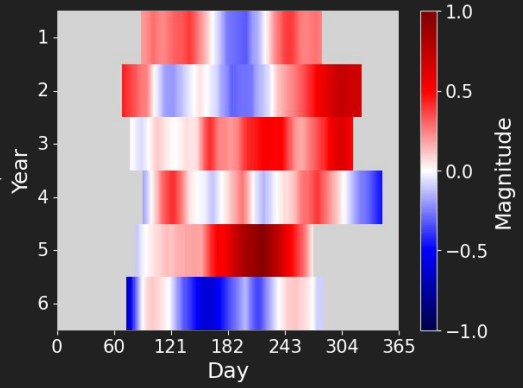
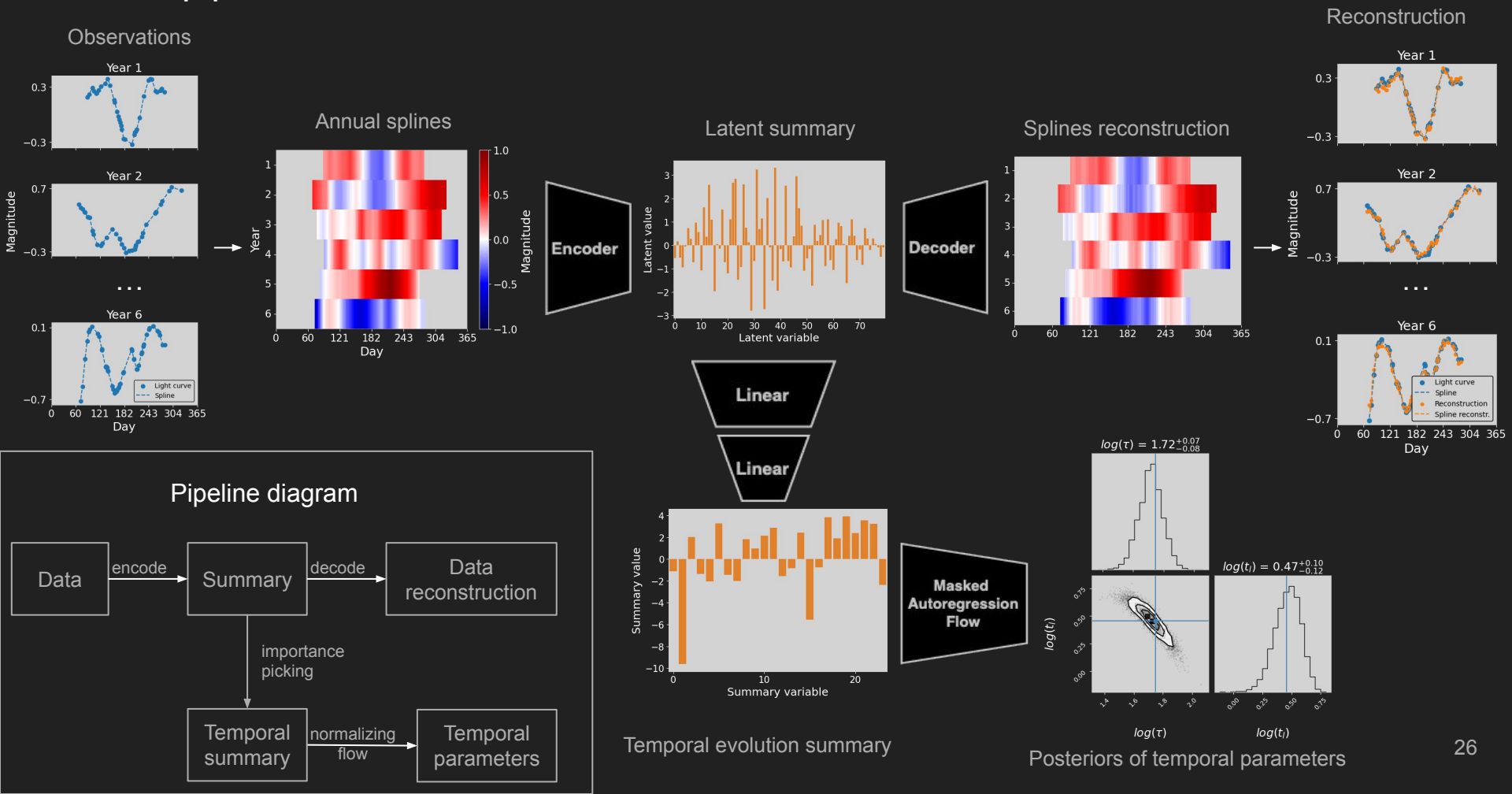


Image-like
representation



Inference pipeline architecture



Variational autoencoder's reconstruction

Dataset:

- 12 years-long mock light curves
- 800k / 100k / 100k split

Architecture:

- Latent space size 128
- Temporal summary size 24

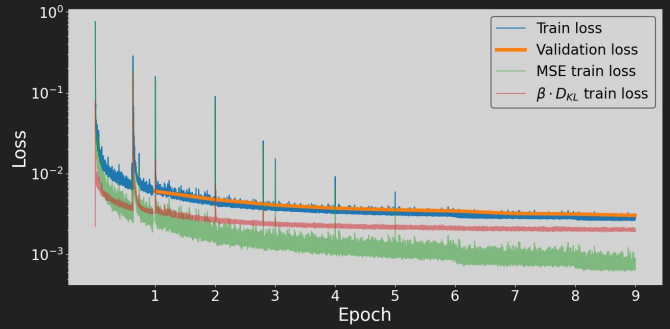
Training:

- 9 epochs, batch size = 128
- Regularization β = 0.05
- Loss:

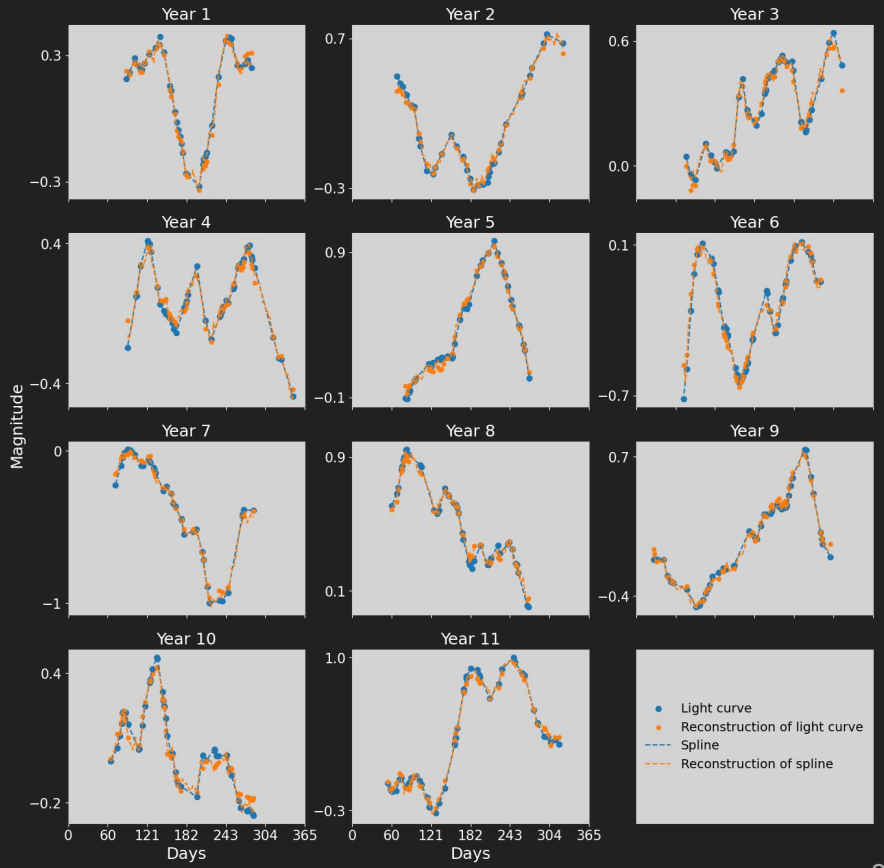
$$\mathcal{L} = MSE(x, x') + \beta \cdot D_{KL}(p(z|x)||p(z))$$

Result:

- RMSE = 0.04 for data in [0,1]
- Learning curve

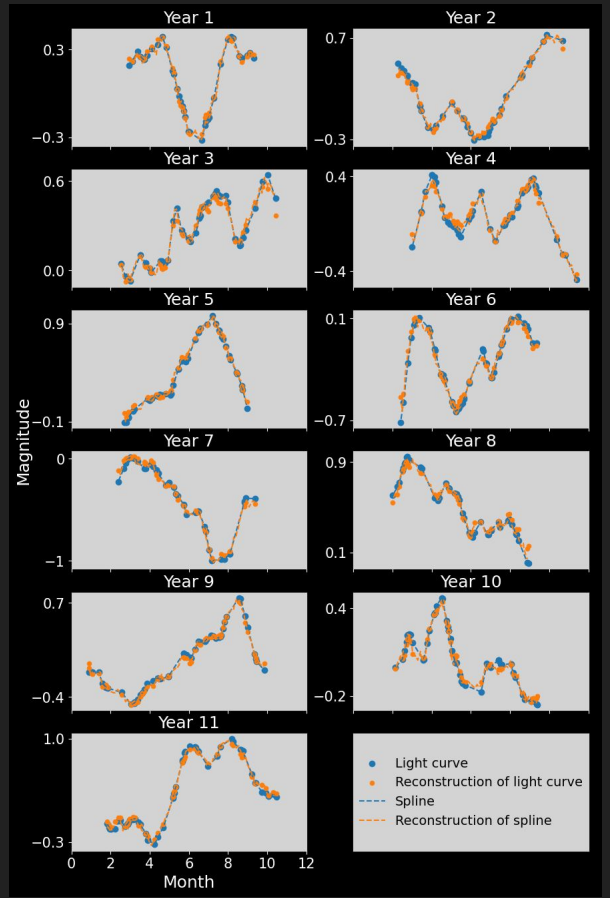


Accurate curves reconstruction

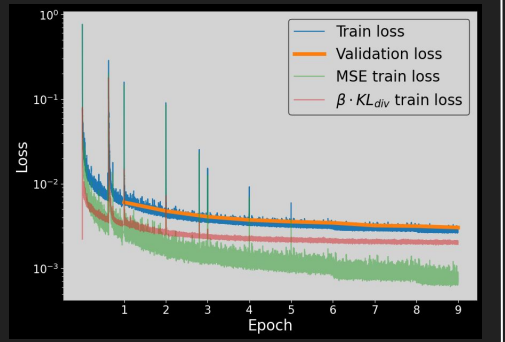


Variational autoencoder

Reconstruction performance



Learning curve



Dataset:

- 12 years long light curves
- 800k / 100k / 100k split

$log(\tau)$	$\sim$	$\mathcal{N}_{tr}(\mu = 2.8, \sigma = 0.52)$
$log(t_\lambda)$	$\sim$	$\mathcal{N}_{tr}(\mu = 0.64, \sigma = 0.43)$
$Cadence$	$\sim$	$\exp(\frac{1}{\lambda} = 4.9)$

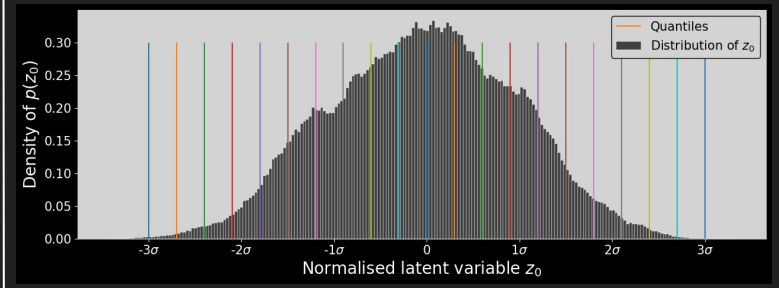
Training:

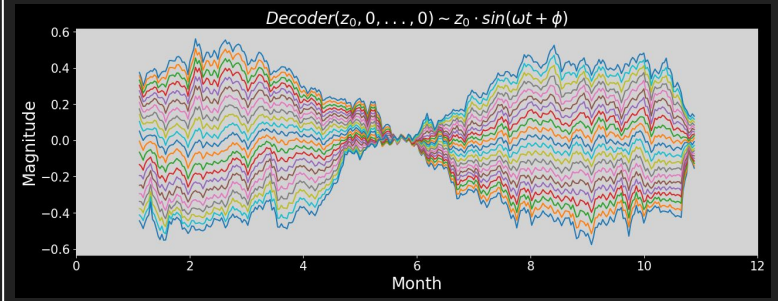
- 9 epochs, batch size = 128
- Loss - MSE and KL divergence
- Regularization β = 0.05

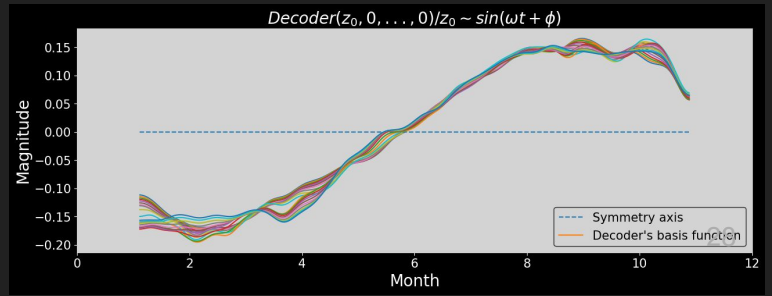
Result:

- RMSE = 0.04 for data in [0,1]

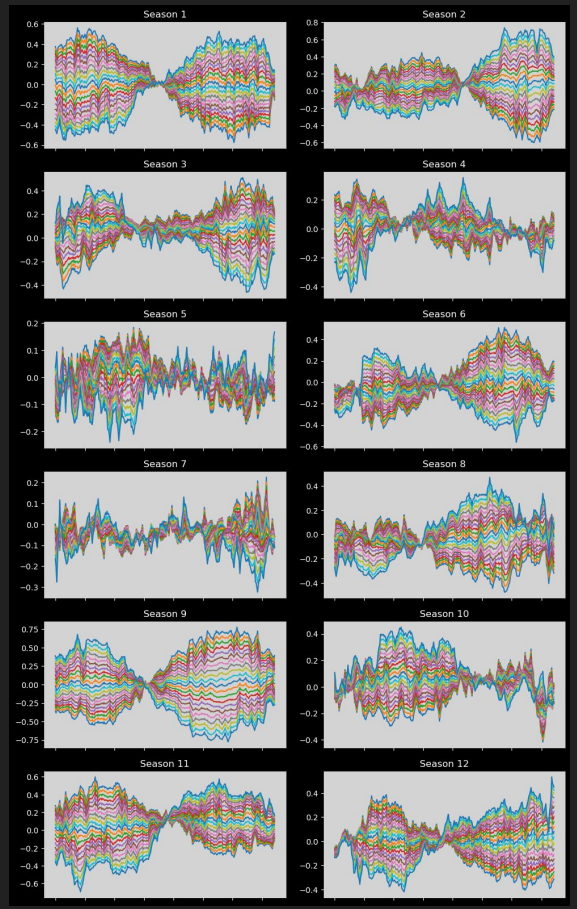
Vary one latent variable

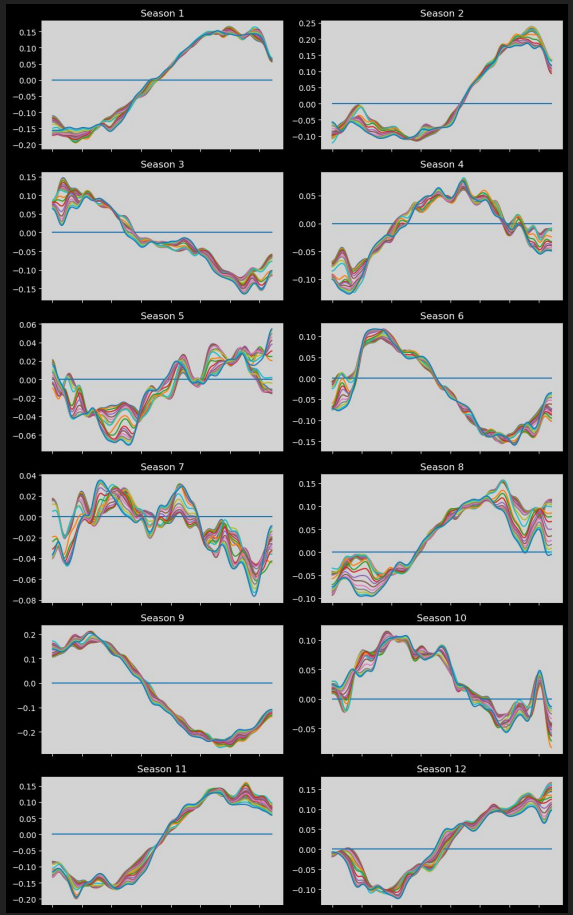






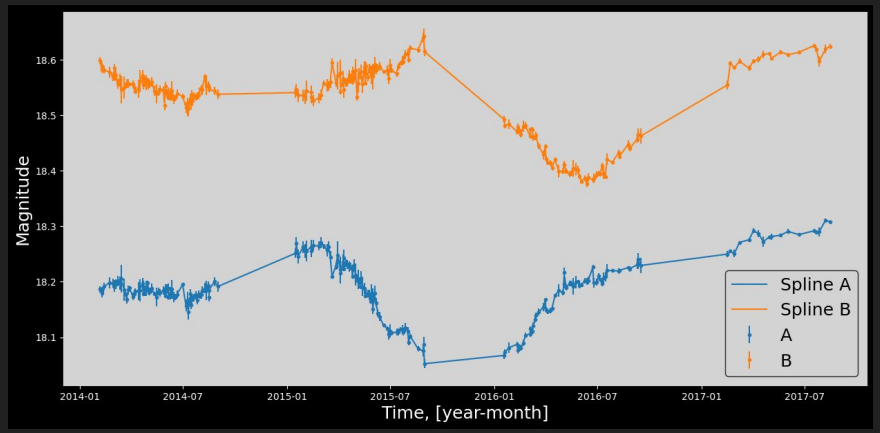
Variational autoencoder latent effects





Models of a quasar light curve

Spline



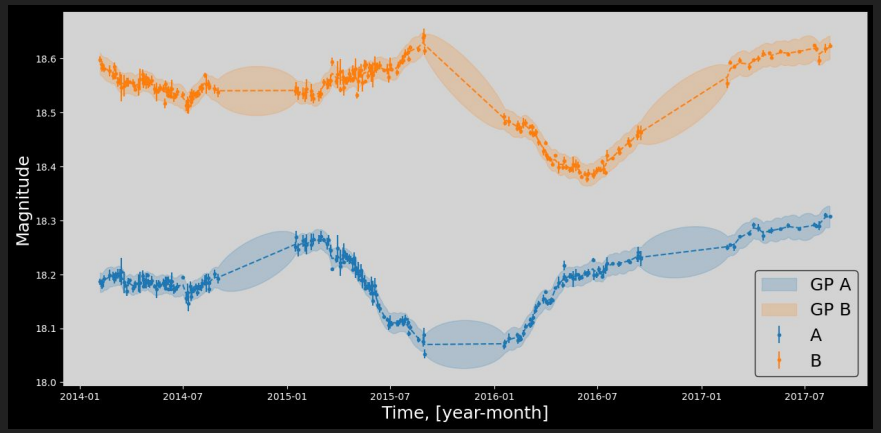
Advantages:

- No hyperparameters except for degree

Disadvantages:

- Not interpretable
- No uncertainties

Gaussian process (GP)



Advantages:

- Physically interpretable parameters
- Get uncertainties for free

Disadvantages:

- Degeneracies between parameters
- Difficult optimization for disbalanced gradients