Neural density estimators

in search for binary systems in our Galaxy

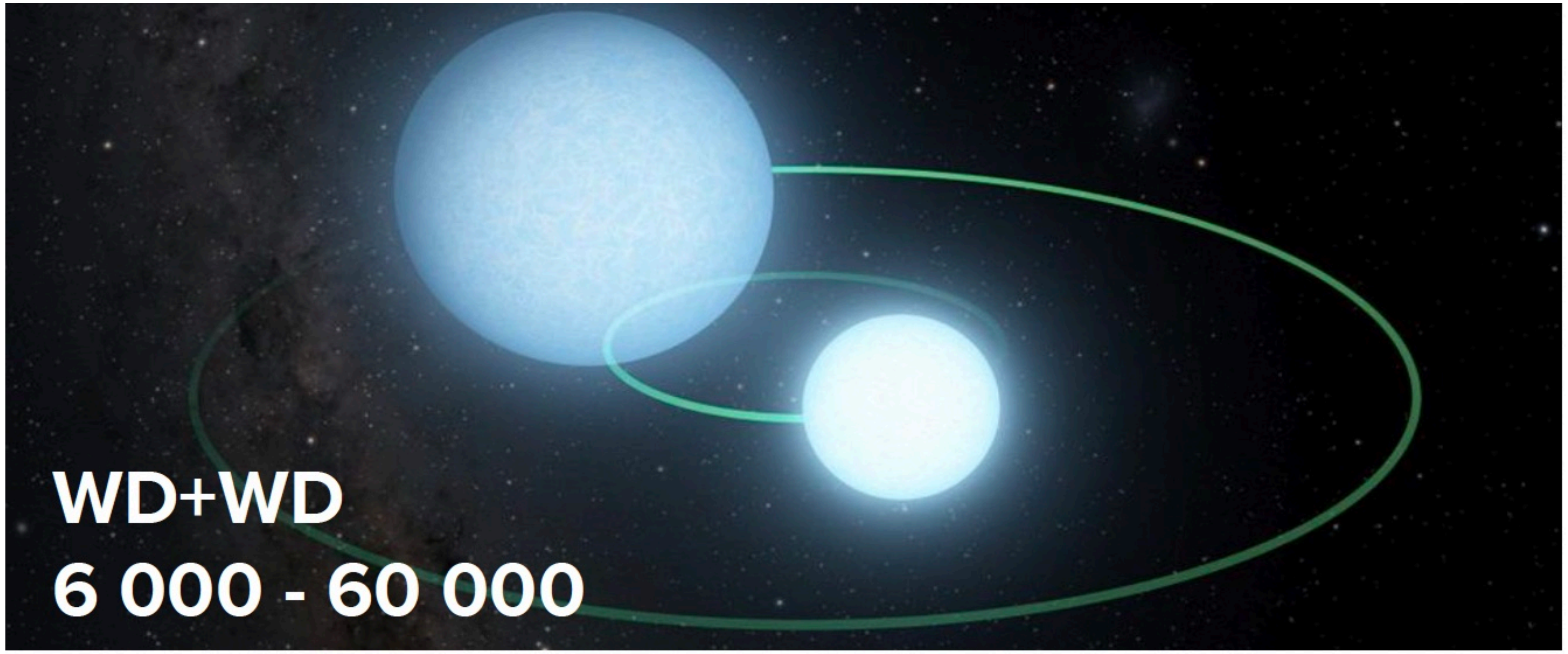
Natalia Korsakova

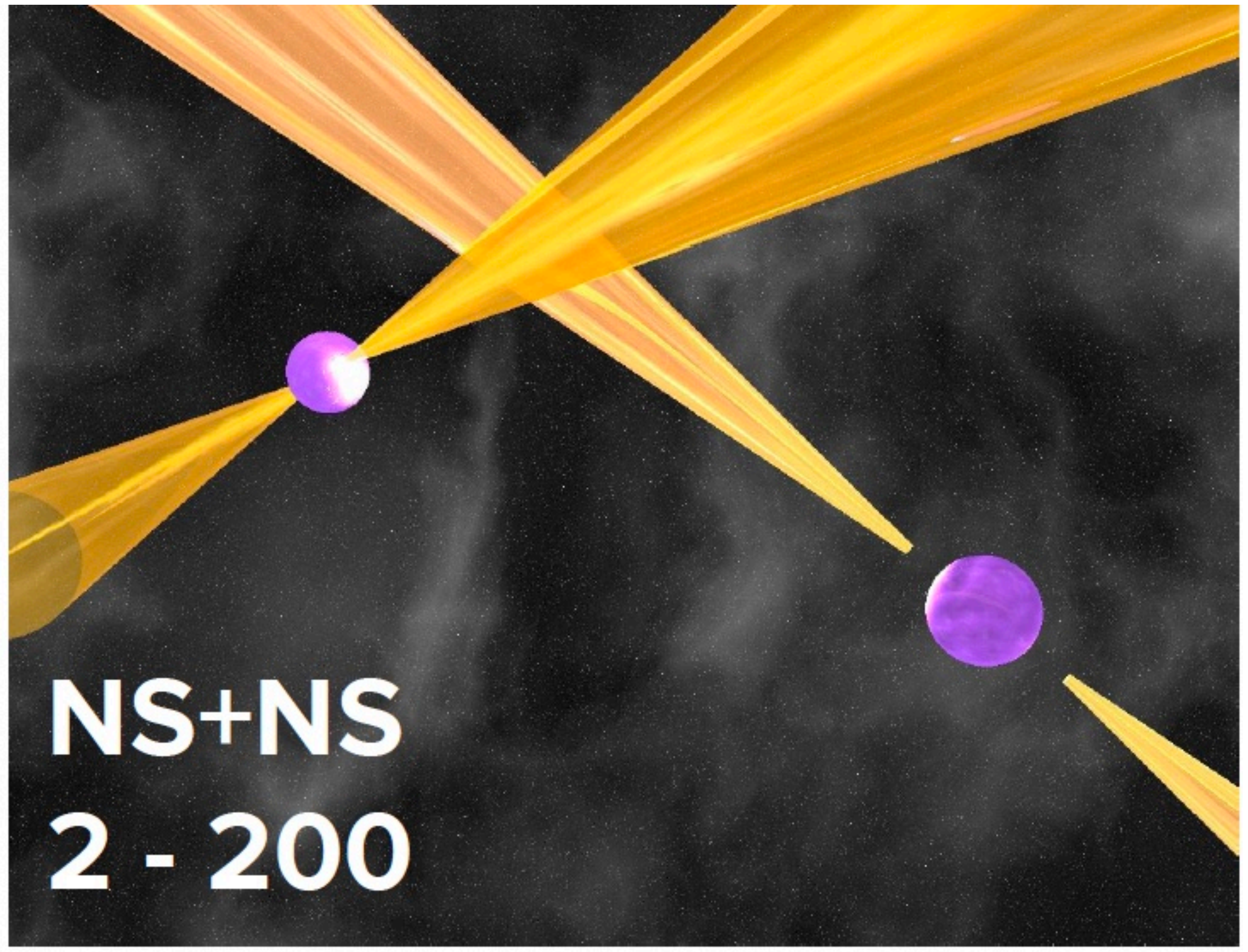
APC, CNRS, FRANCE

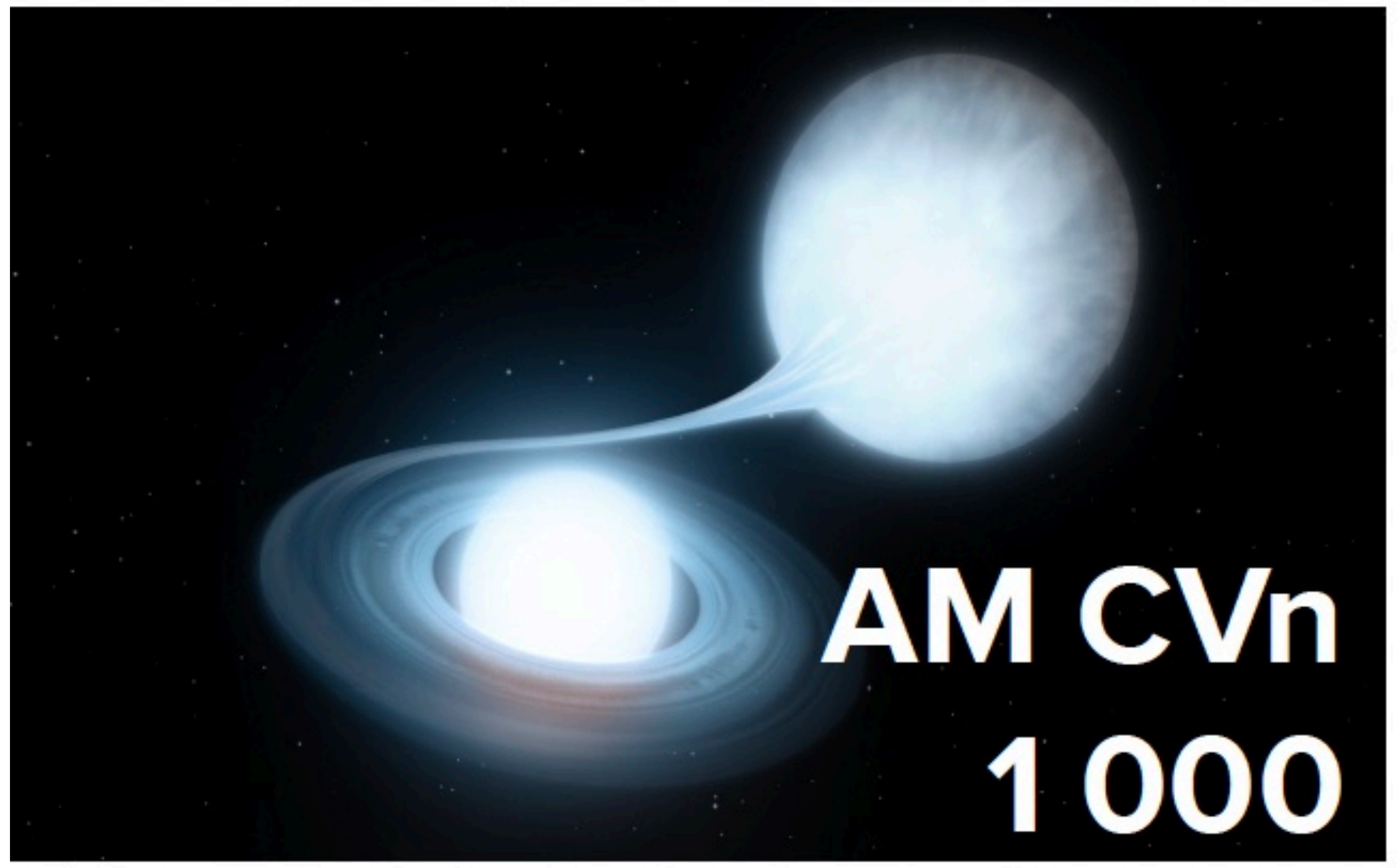
1 October 2024 Toulouse

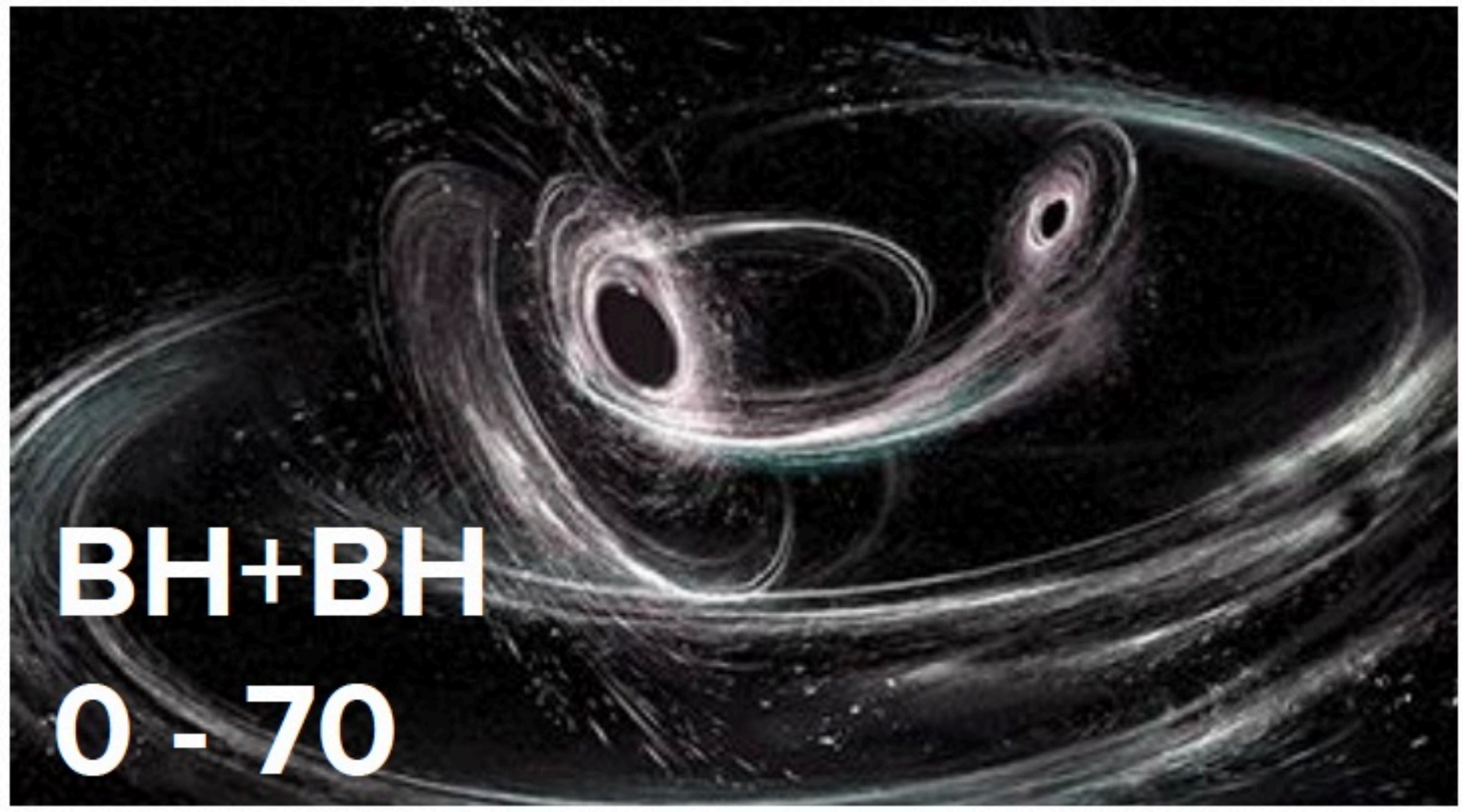
Galactic Binaries

Image credit: Valeriya Korol



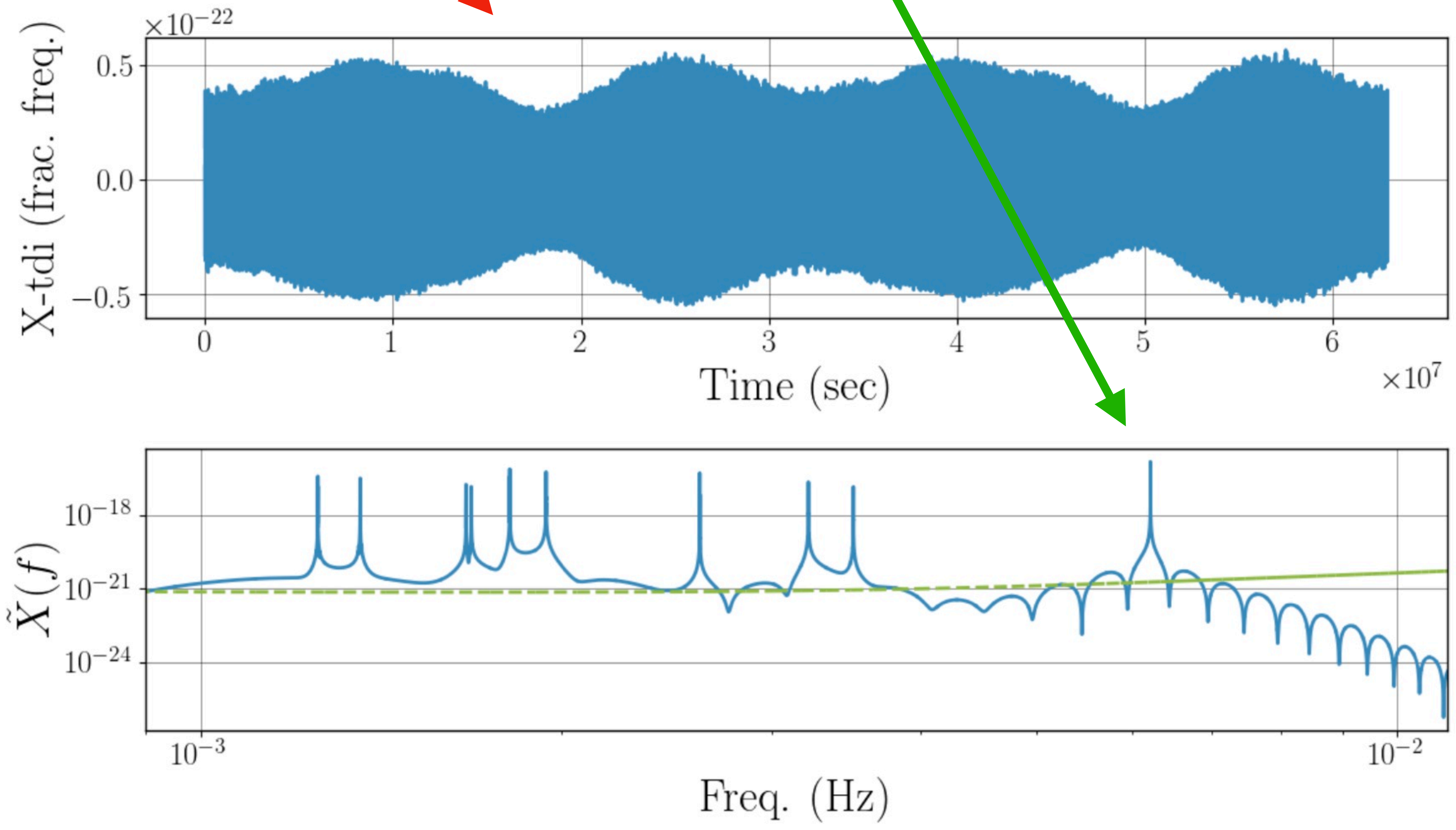






Signal

in time and frequency domain



Parameters that we can extract:

- frequency
- frequency derivative
- sky localisation
- distance
- position of the orbit relative to observer

Galactic Binaries

in Milky Way

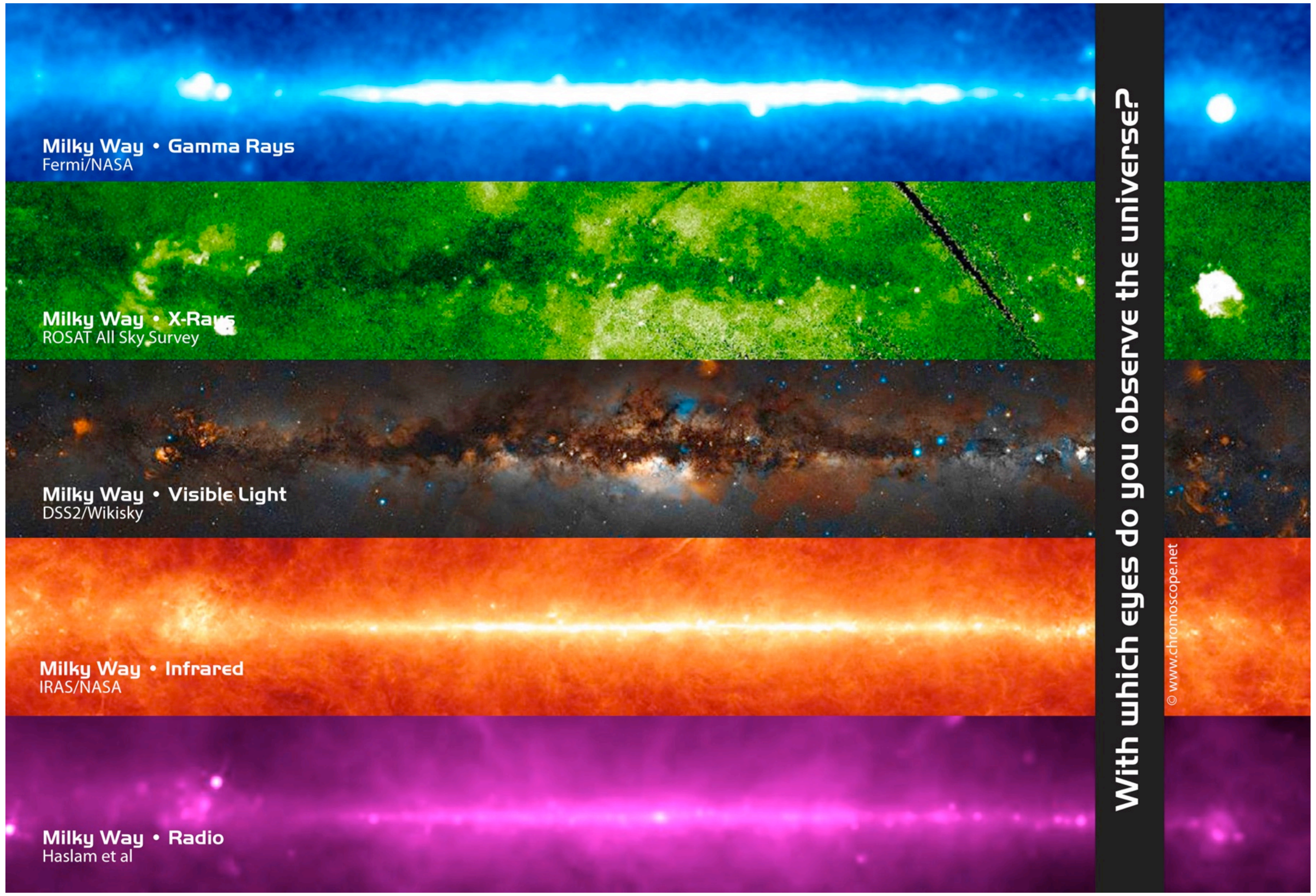
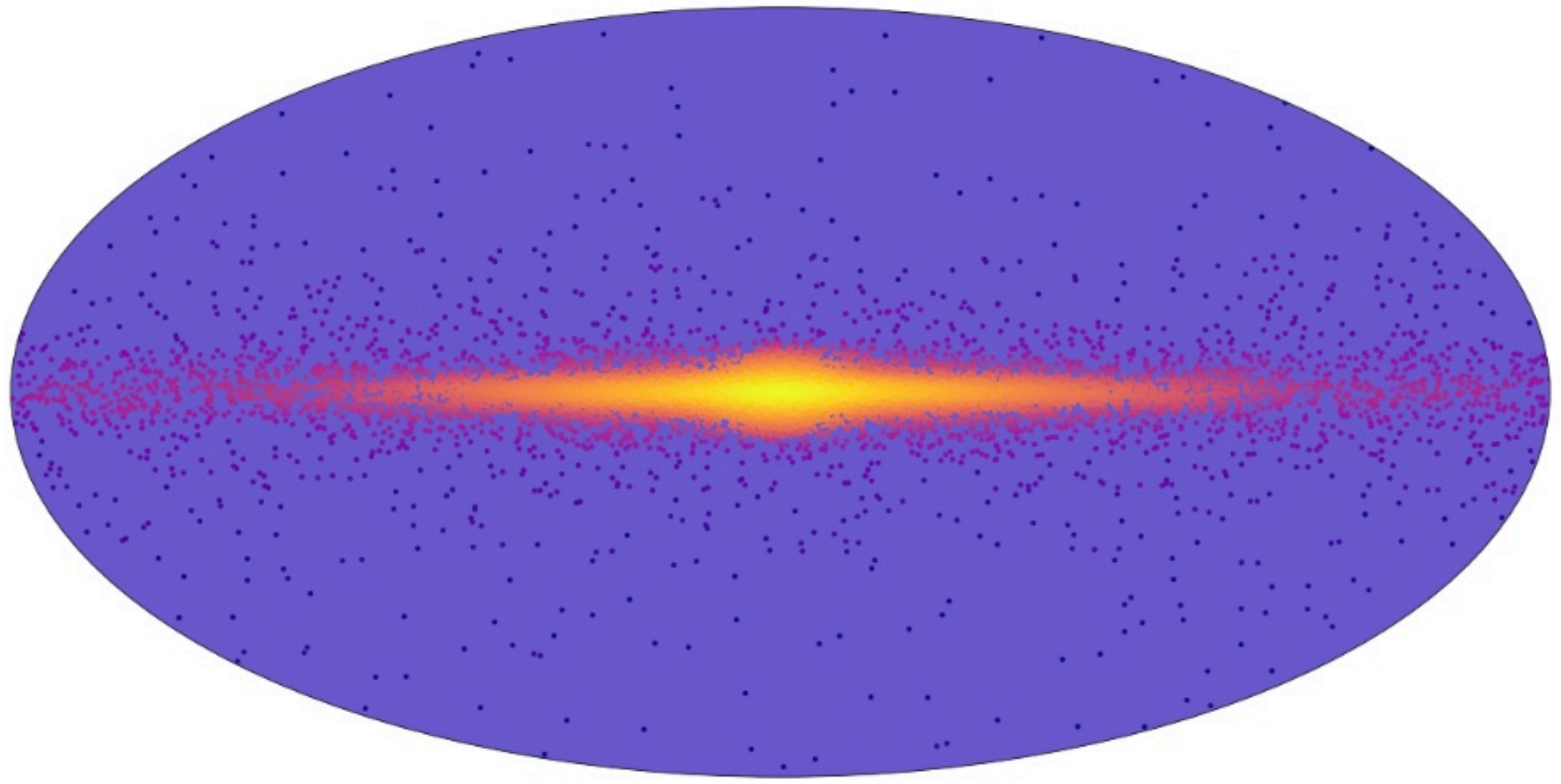


Image credit: ESO

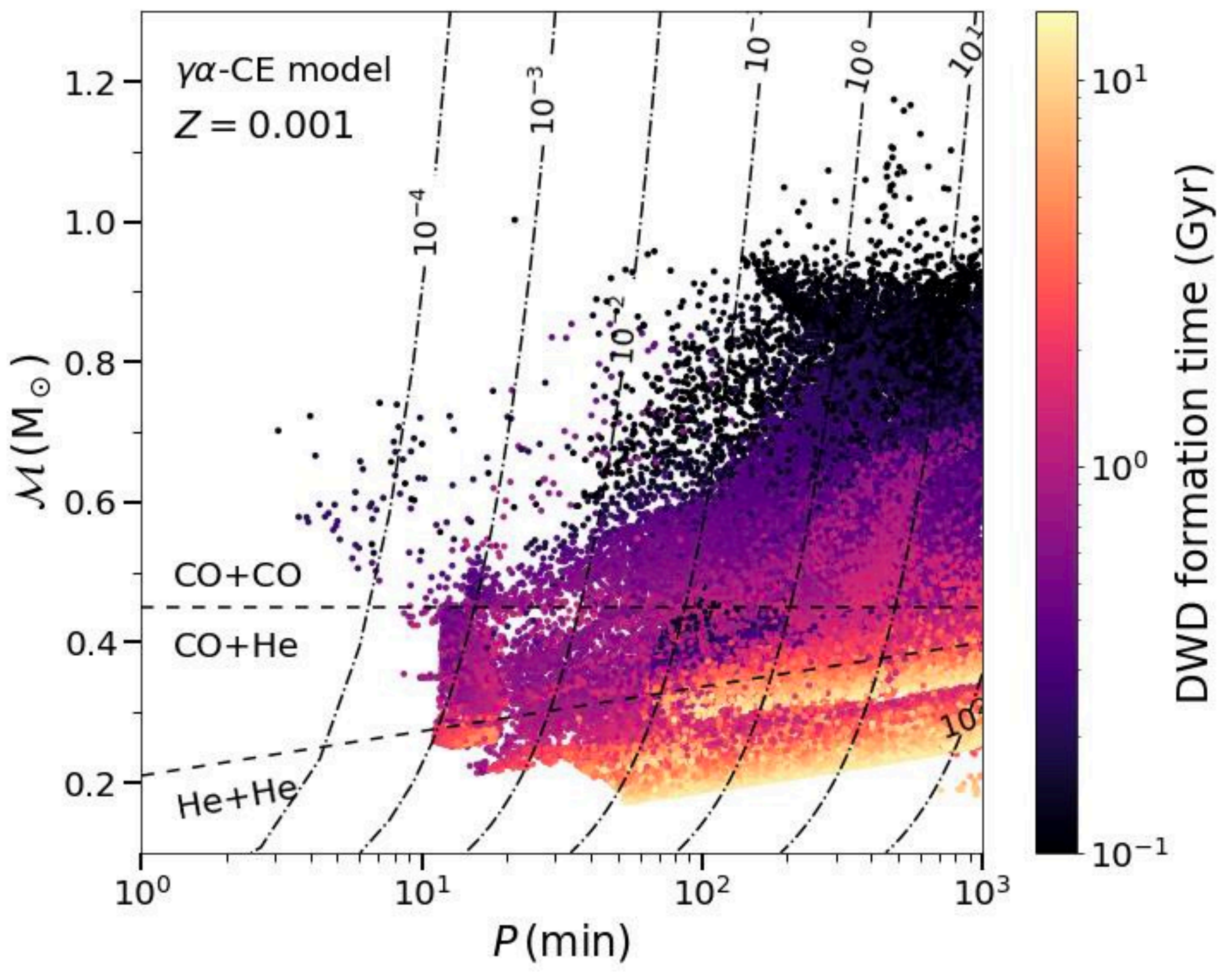


Milky Way seen with Galactic Binaries

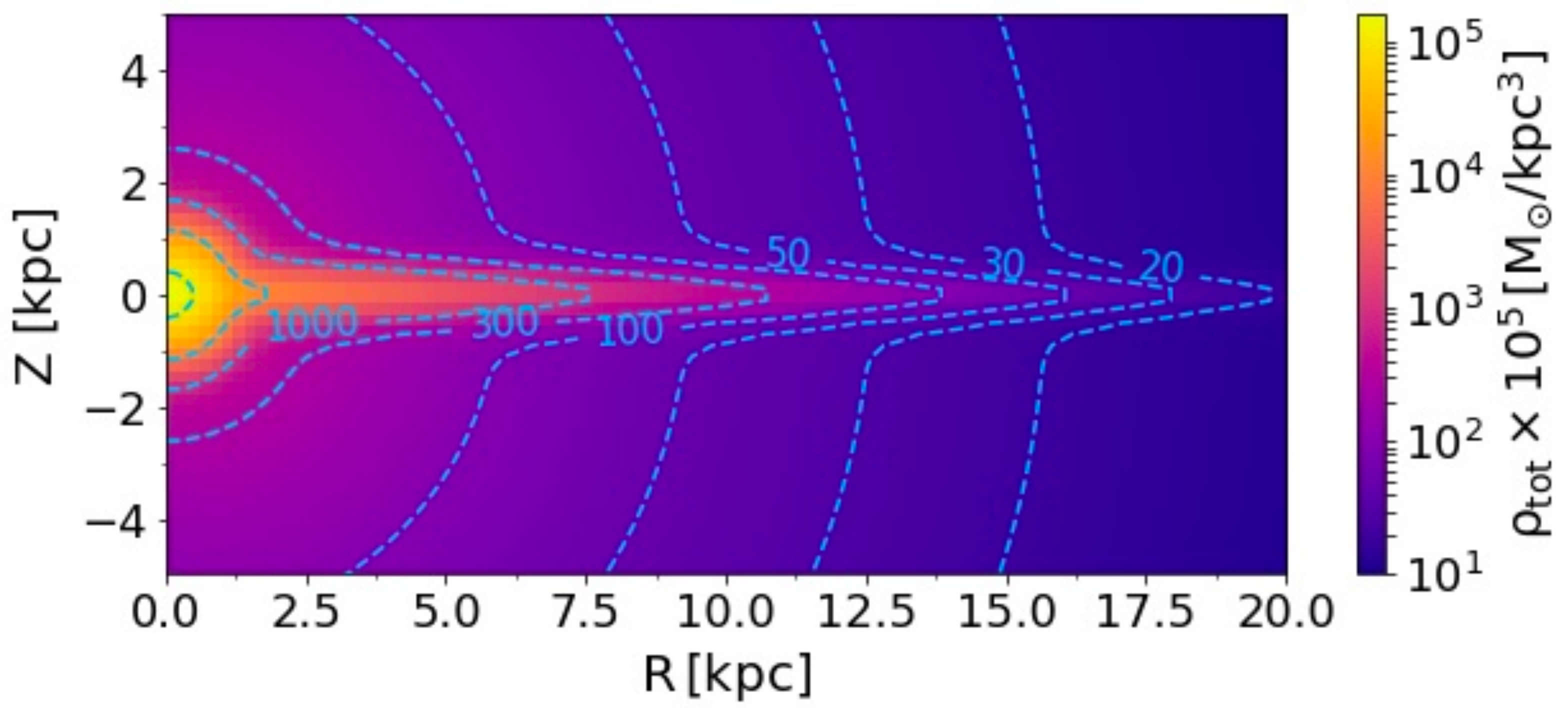
Image credit: Valeriya Korol

Galactic Binaries

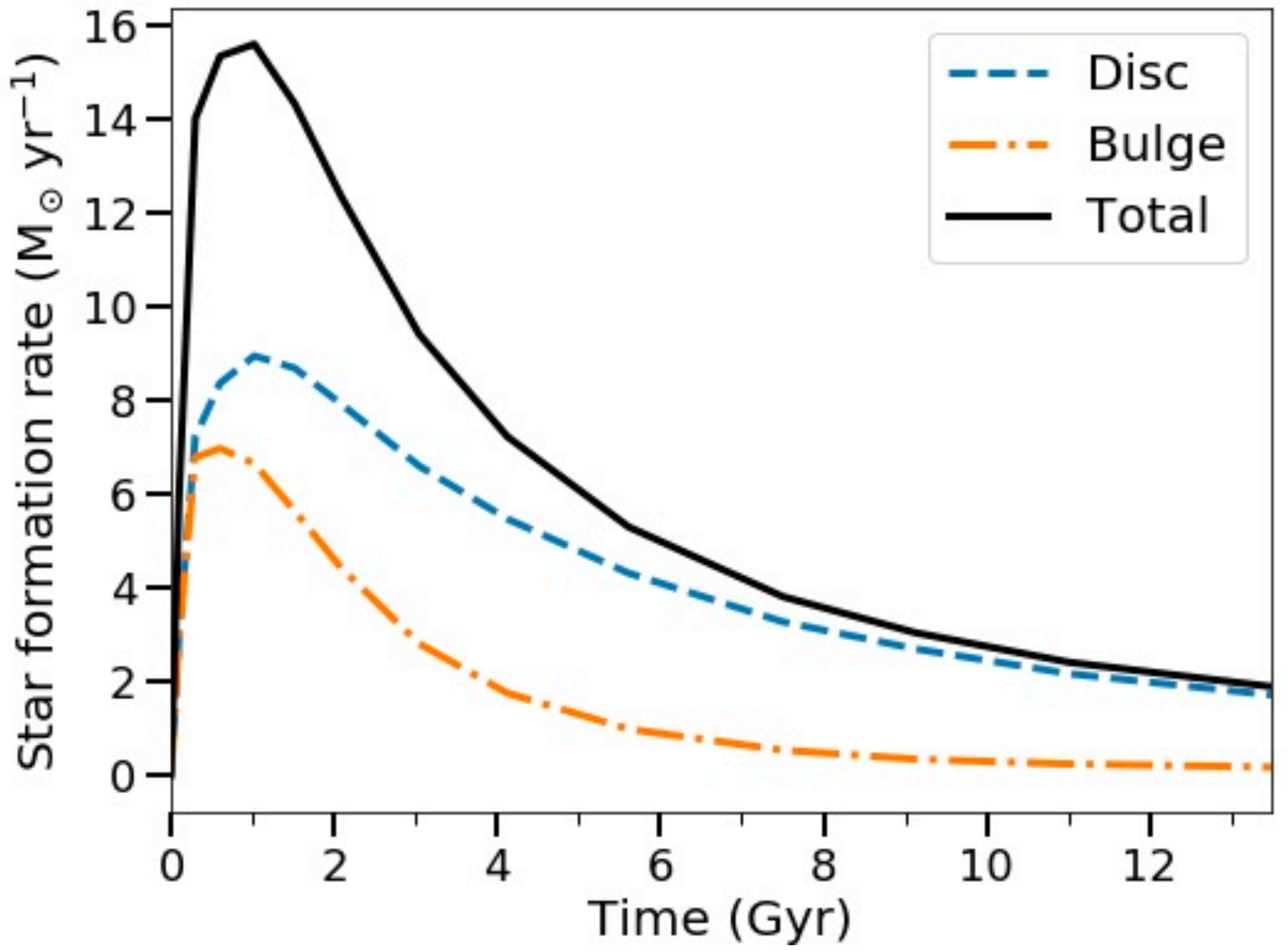
modelling the population



+



+



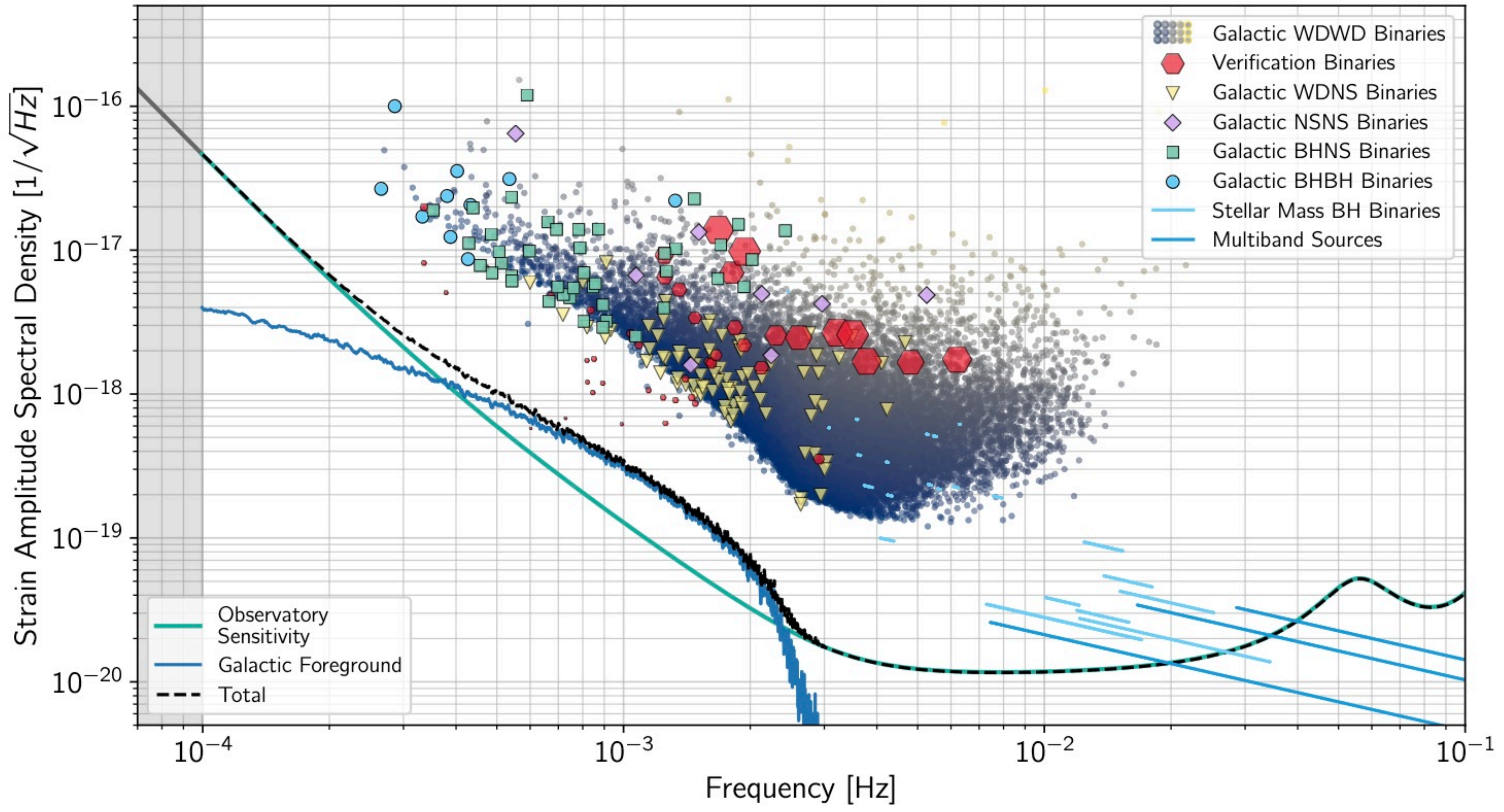
Synthetic population of GBs

Milky Way potential

Star formation rate

Image credit: Valeriya Korol

LISA sensitivity and Galactic Binaries



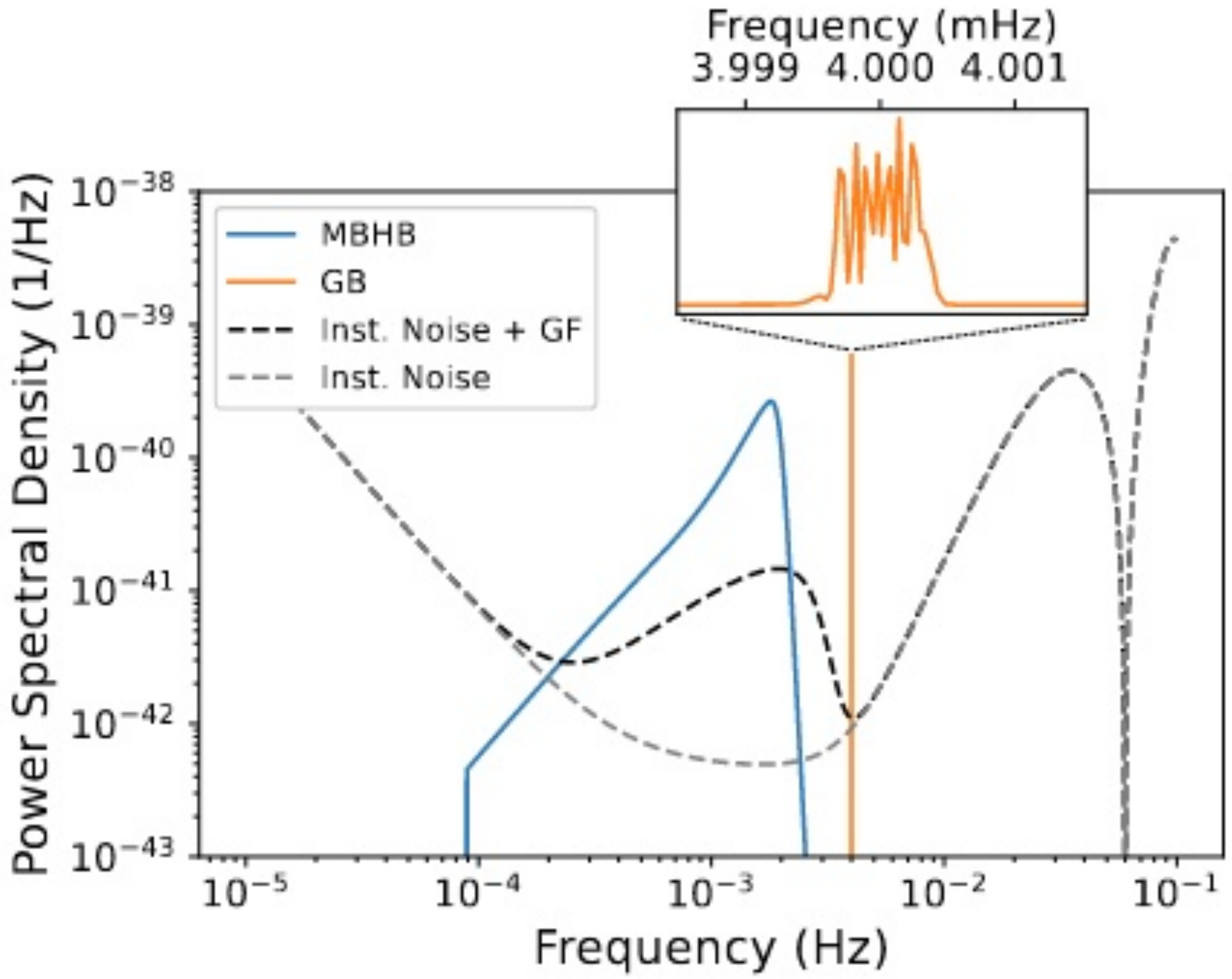
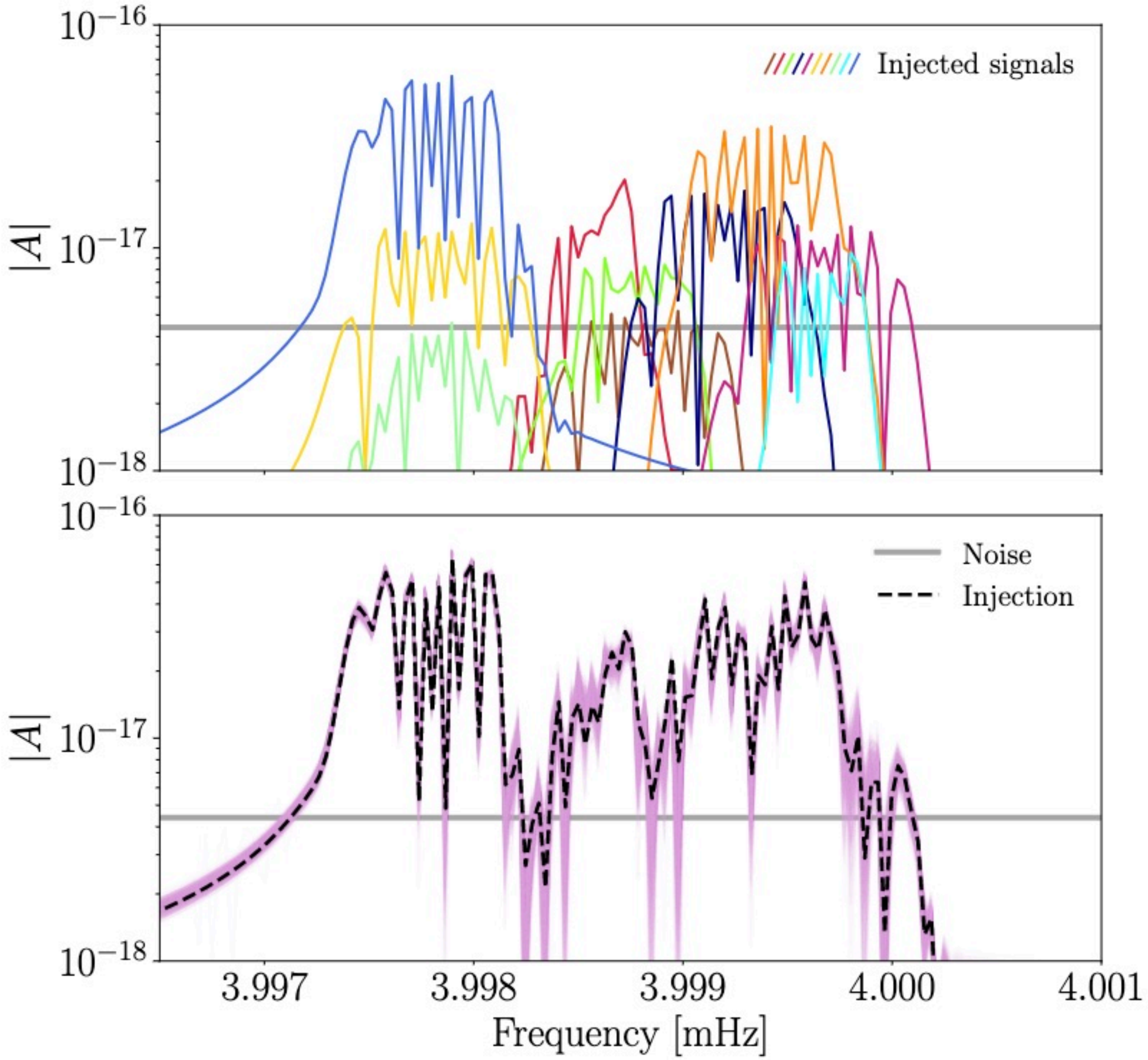


Image credit: Red book, arXiv:2402.07571

Image credit: arXiv:2405.04690

Why they are a problem for data analysis

Unknown number of signal, unknown noise



Main problems:

- at some point signals mix and become confusion noise
- we do not know the number of the signals in each band
- we have to estimate the noise at the same time as we estimate the parameter of the signals

Image credit: arXiv:2303.02164

Parameter estimation

Bayes equation

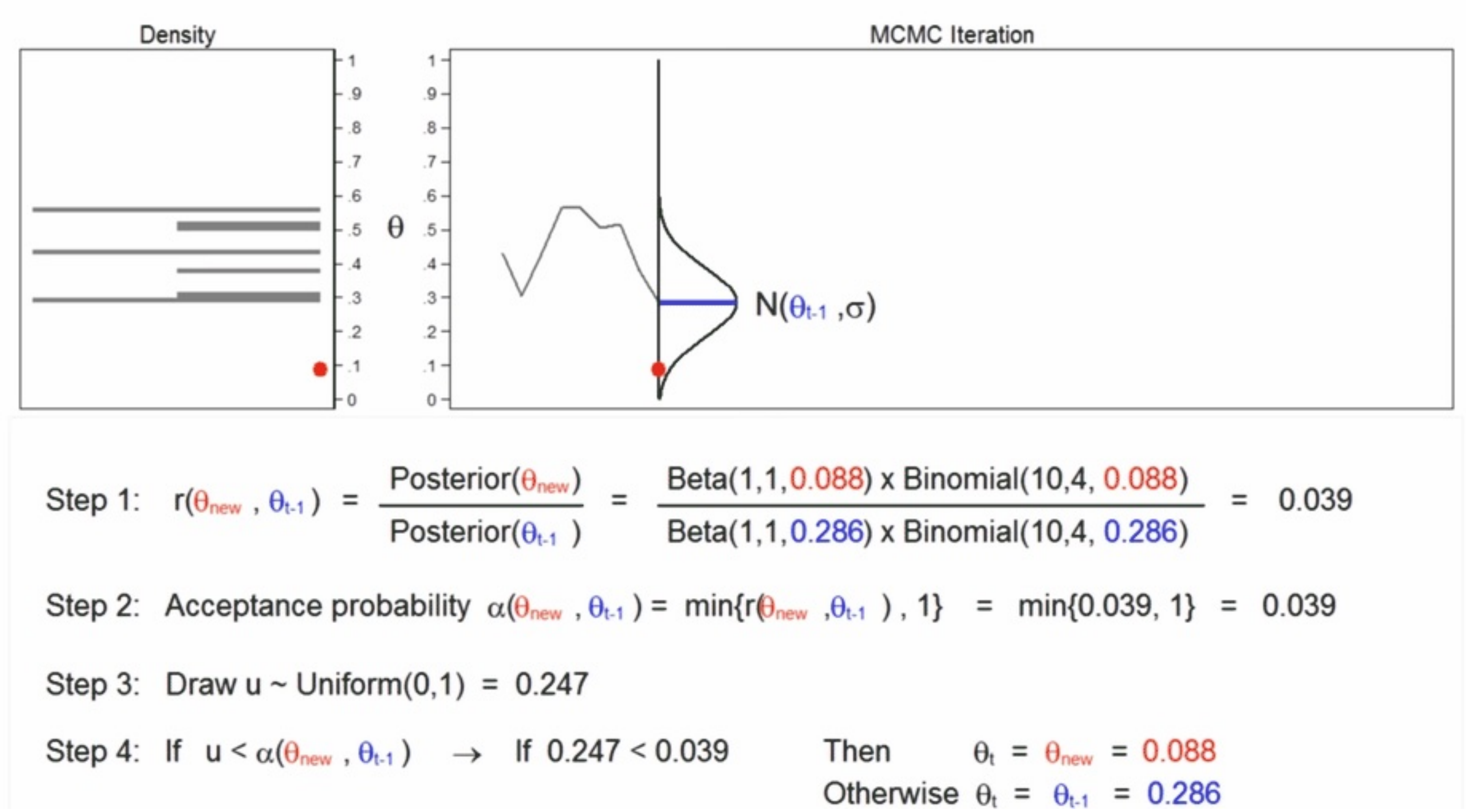
$$p(\theta \mid x) = \frac{p(x \mid \theta)p(\theta)}{p(x)}$$

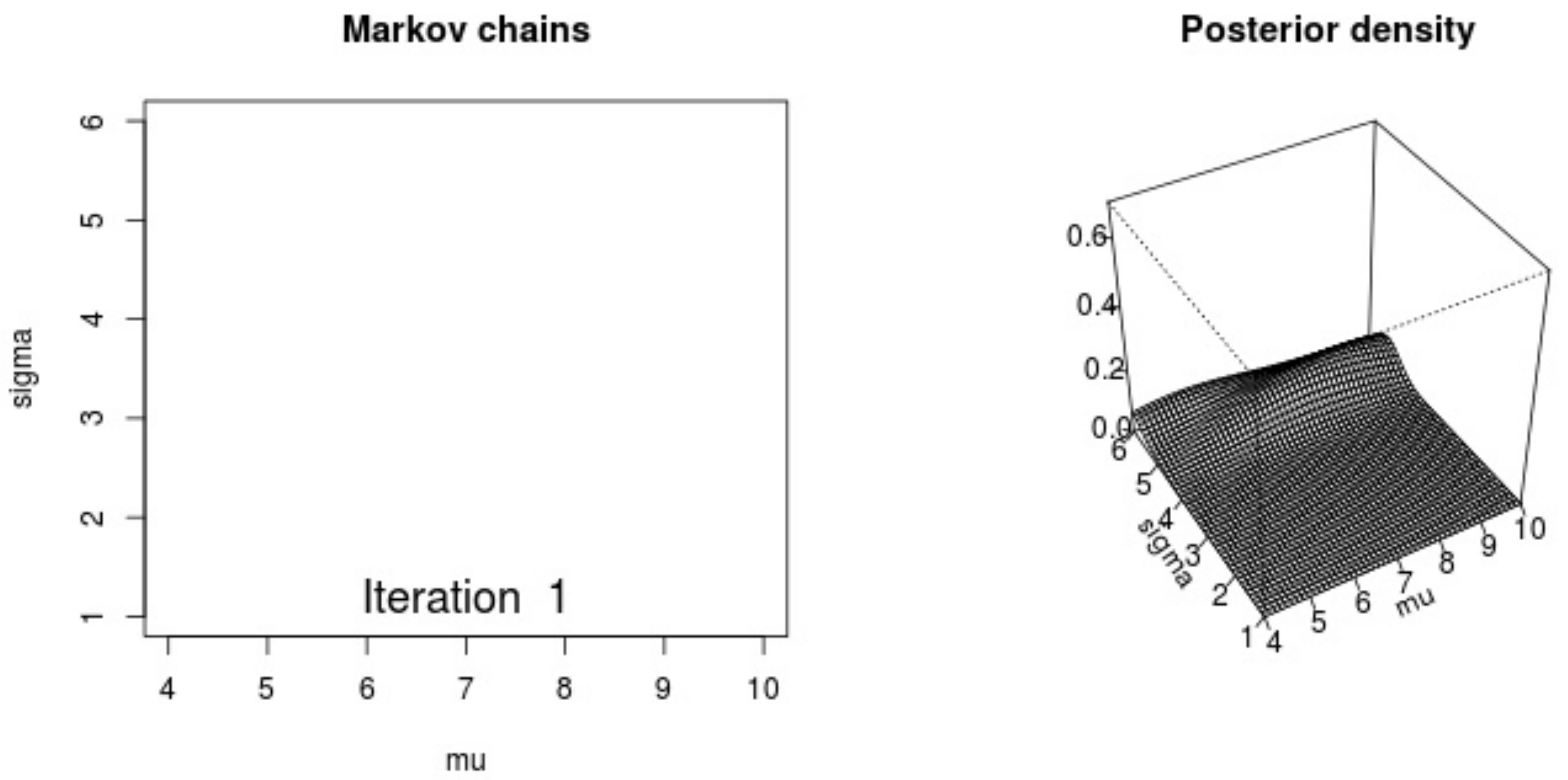
- approximate inference:
  - MCMC/Nested sampling
    requires likelihood evaluation

- simplification to the model:
  - Invertible models

Sampling to solve Bayes equation

Markov Chain Monte Carlo: fixed dimensionality





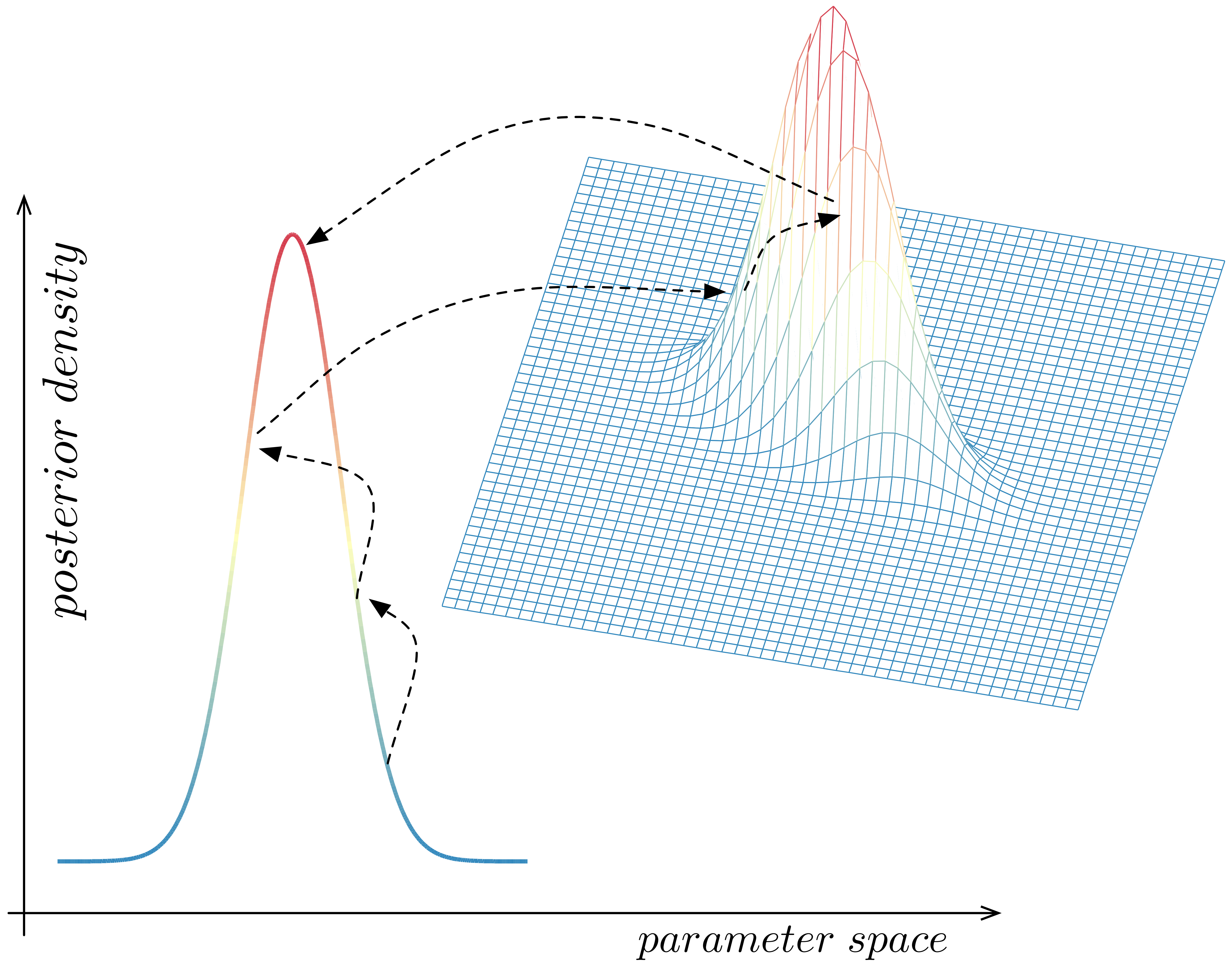
https://blog.stata.com/
https://blog.revolutionanalytics.com/

- Start from theta_0
- Propose a new point from proposal distribution q
- Accept, or reject with a probability

$$\alpha = \min\left[1, \frac{p(\vec{\theta_1}|y)q(\vec{\theta_0},\vec{\theta_1})}{p(\vec{\theta_0}|y)q(\vec{\theta_1},\vec{\theta_0})}\right]$$

Image credit: Nikos Karnesis

Unknown number of dimensions



- Same procedure, now generalized for ***k***-order of model. It is organized in two steps.
- Before all, we begin with $\boldsymbol{\theta}_\mathbf{k}$ for model ***k***.

1. In-Model Step: The usual MH step, for model ***k***.
2. Outer-Model Step:

- Propose new $\boldsymbol{\theta}_\mathbf{m}$ for model ***m*** from a given proposal distribution *q*.
- Essentially propose the "birth" or "death" of dimensions at each iteration.
- Accept, or reject with a probability:

$$\alpha = \min\left[1, \frac{p(y|\vec{\theta}_k)p(\vec{\theta}_k)q(\{k,\vec{\theta}_k\},\{m,\theta_m\})}{p(y,\vec{\theta}_m)p(\vec{\theta}_m)q(\{m,\vec{\theta}_m\},\{k,\theta_k\})}\right]$$

Image credit: Nikos Karnesis

Unknown number of dimensions

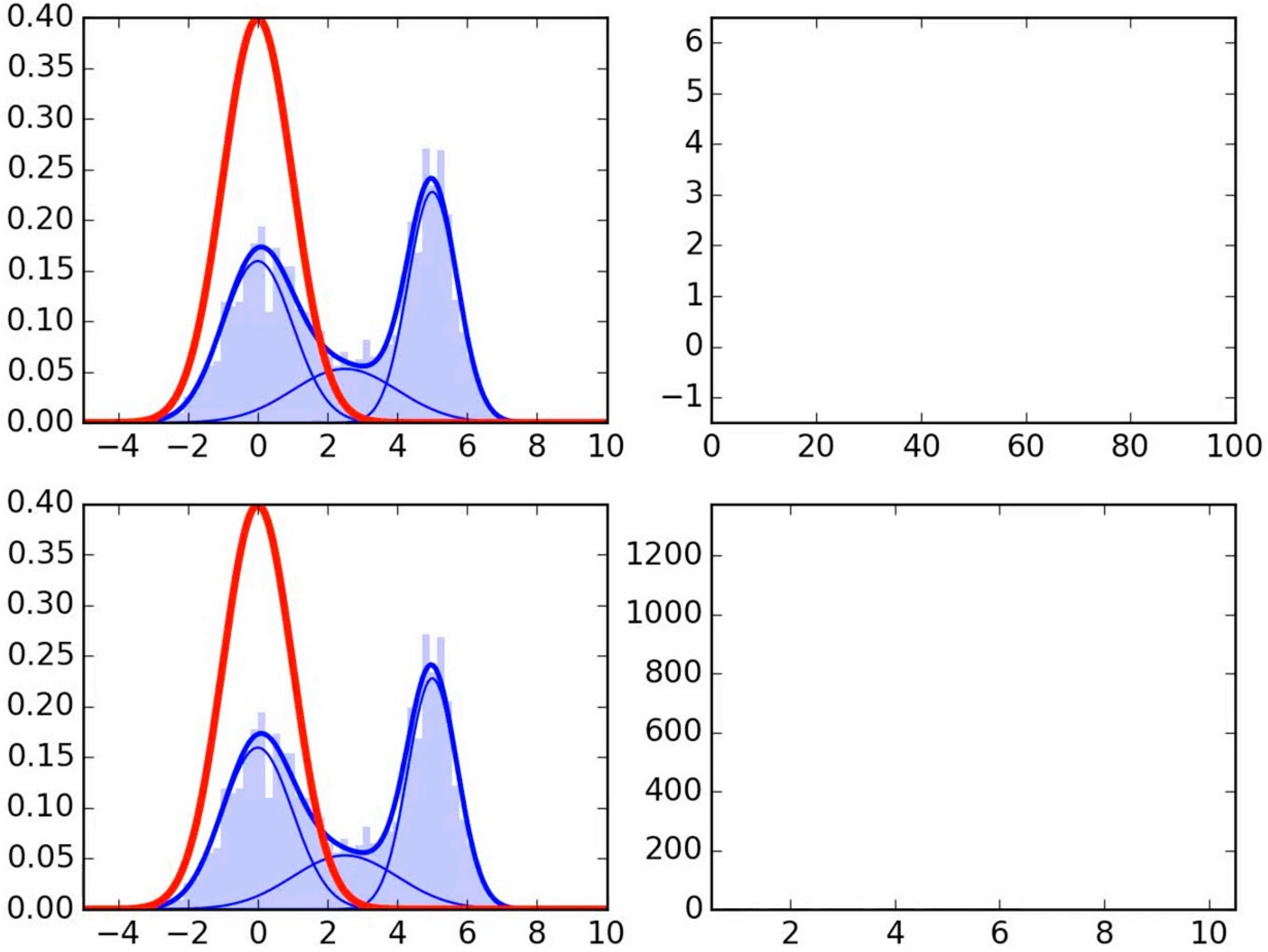
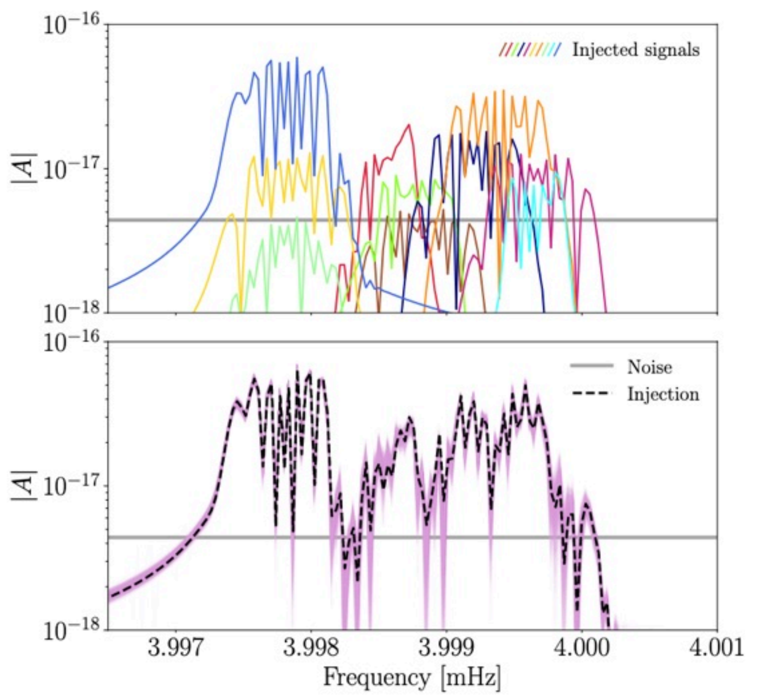


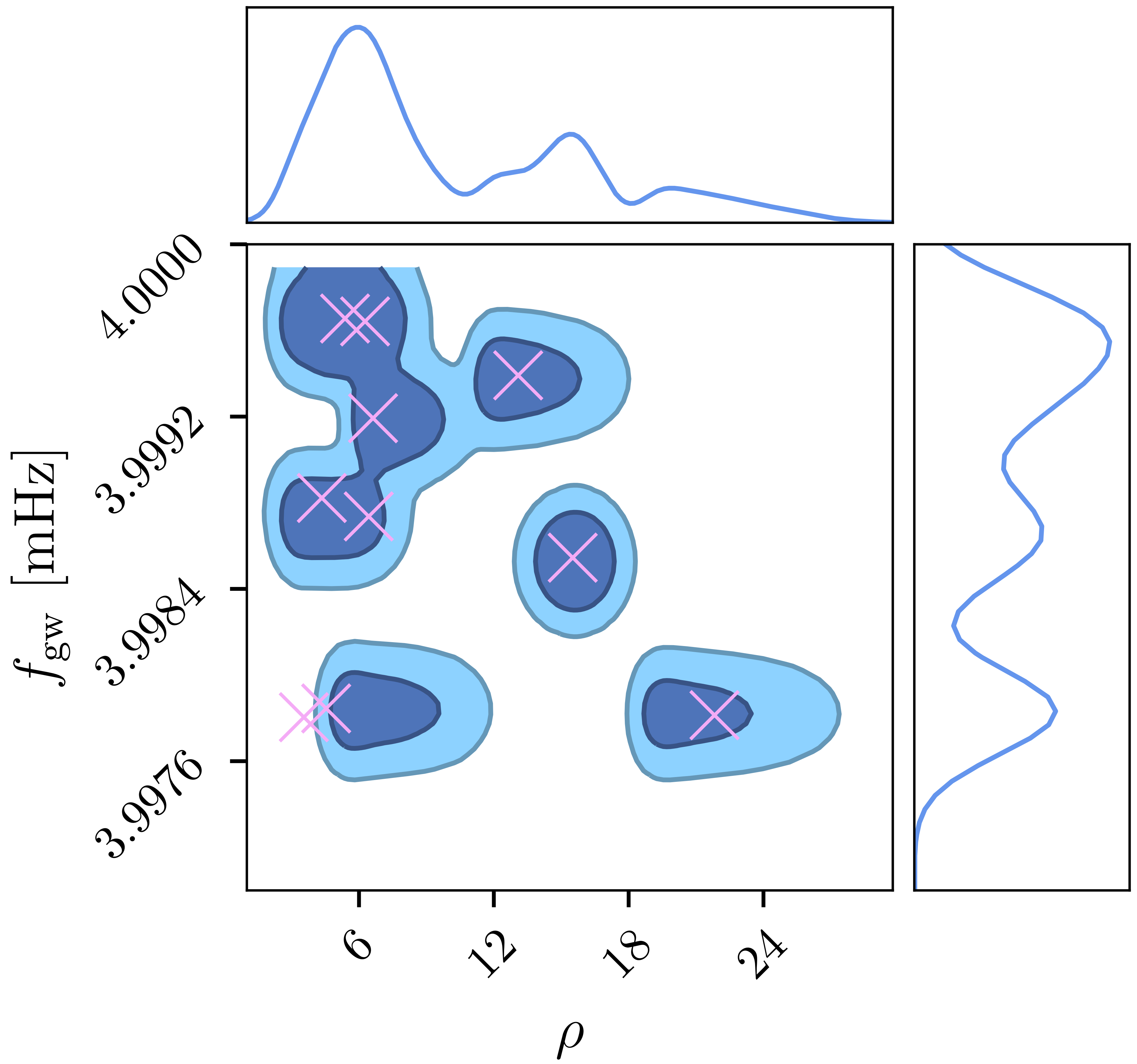
Image credit: Nikos Karnesis

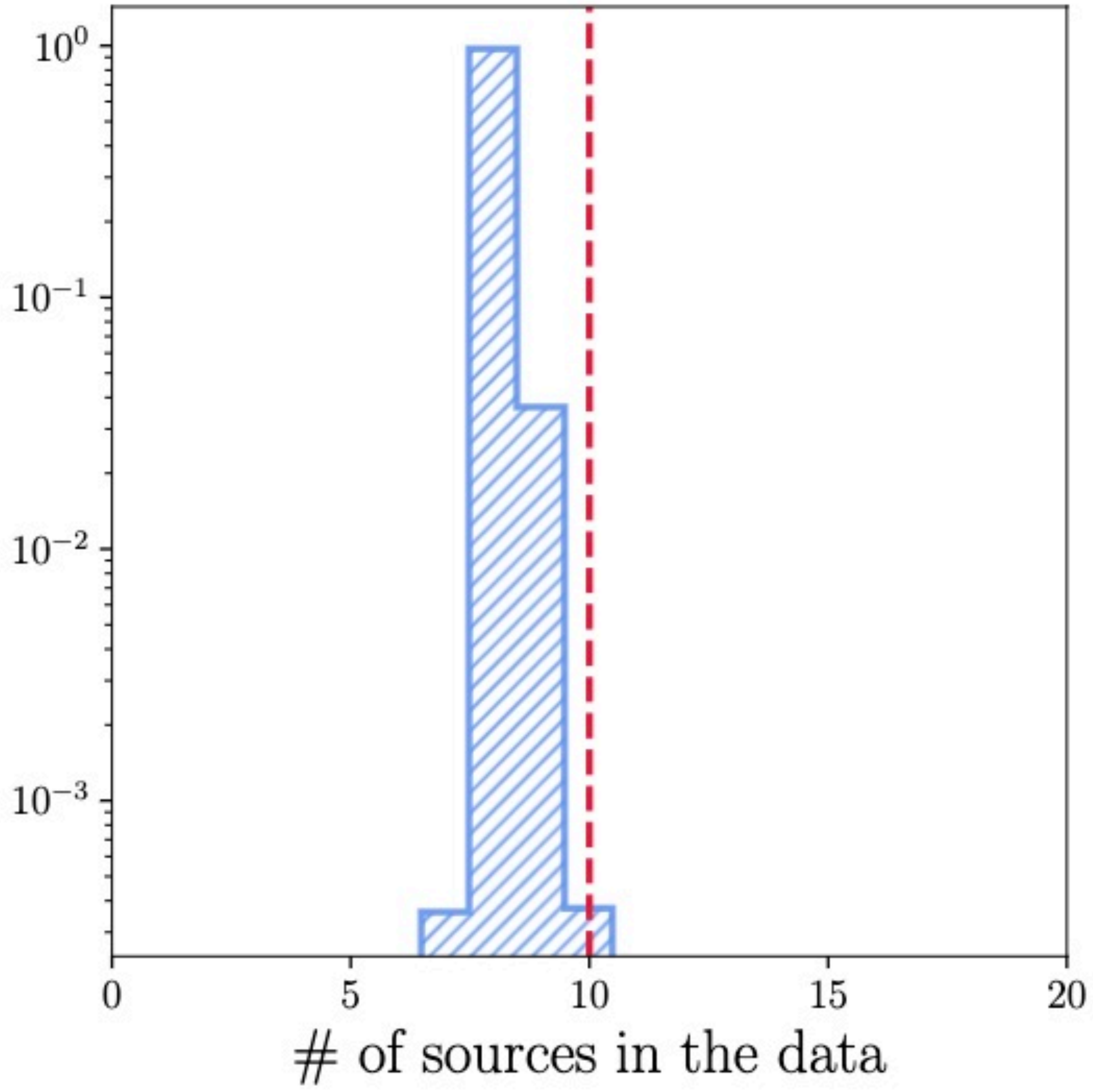
Video source: https://www.youtube.com/watch?v=wBTGoA_dIIo

Unknown number of dimensions

Galactic Binaries, single band example







Proposals

Efficiency of proposals

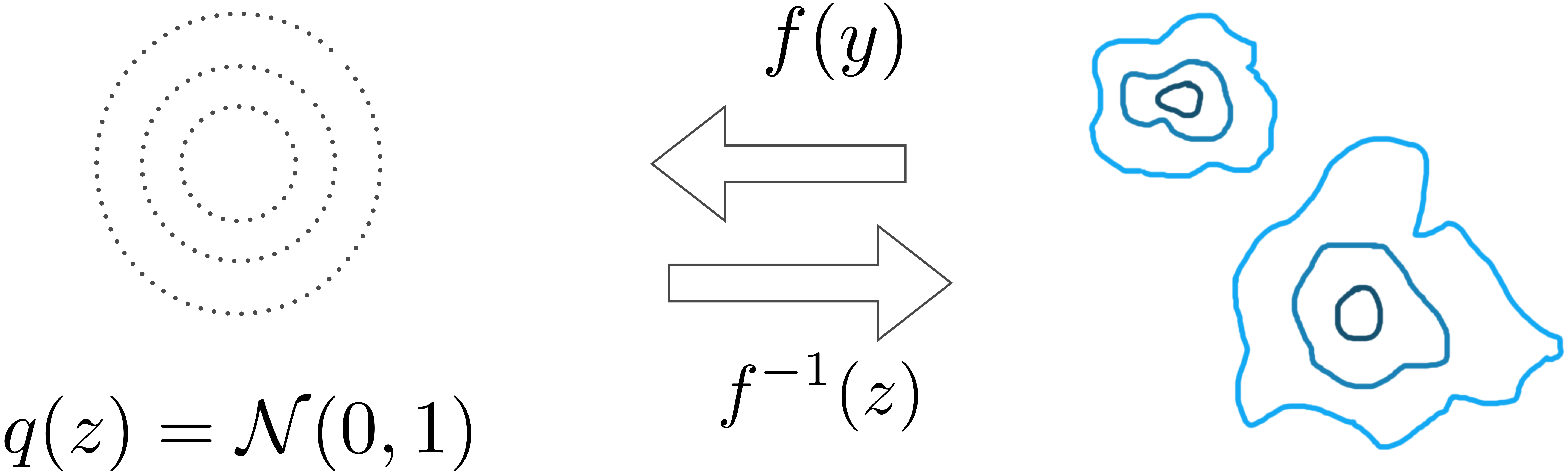
We rely on two criteria to evaluate the performance of proposals:

- rate of accumulation of effective samples
  - execution time
  - autocorrelation length
- faithfulness of the posterior

Neural density estimators for proposals

Estimating the densities

1. We have simple random generator
2. We want to sample from a more complex distribution
3. We can estimate a bijective transformation which will allow us to do that



Neural density estimators for proposals

Estimating the densities

$$p(y) = \boxed{q(f^{-1}(y))}\,\boxed{\left|\det\left(J_{f^{-1}}(y)\right)\right|}$$

- has to be a bijection
- and f^{-1} have to be differentiable
- Jacobian determinant has to be tractably invertable

Neural density estimators for proposals

Estimating the densities

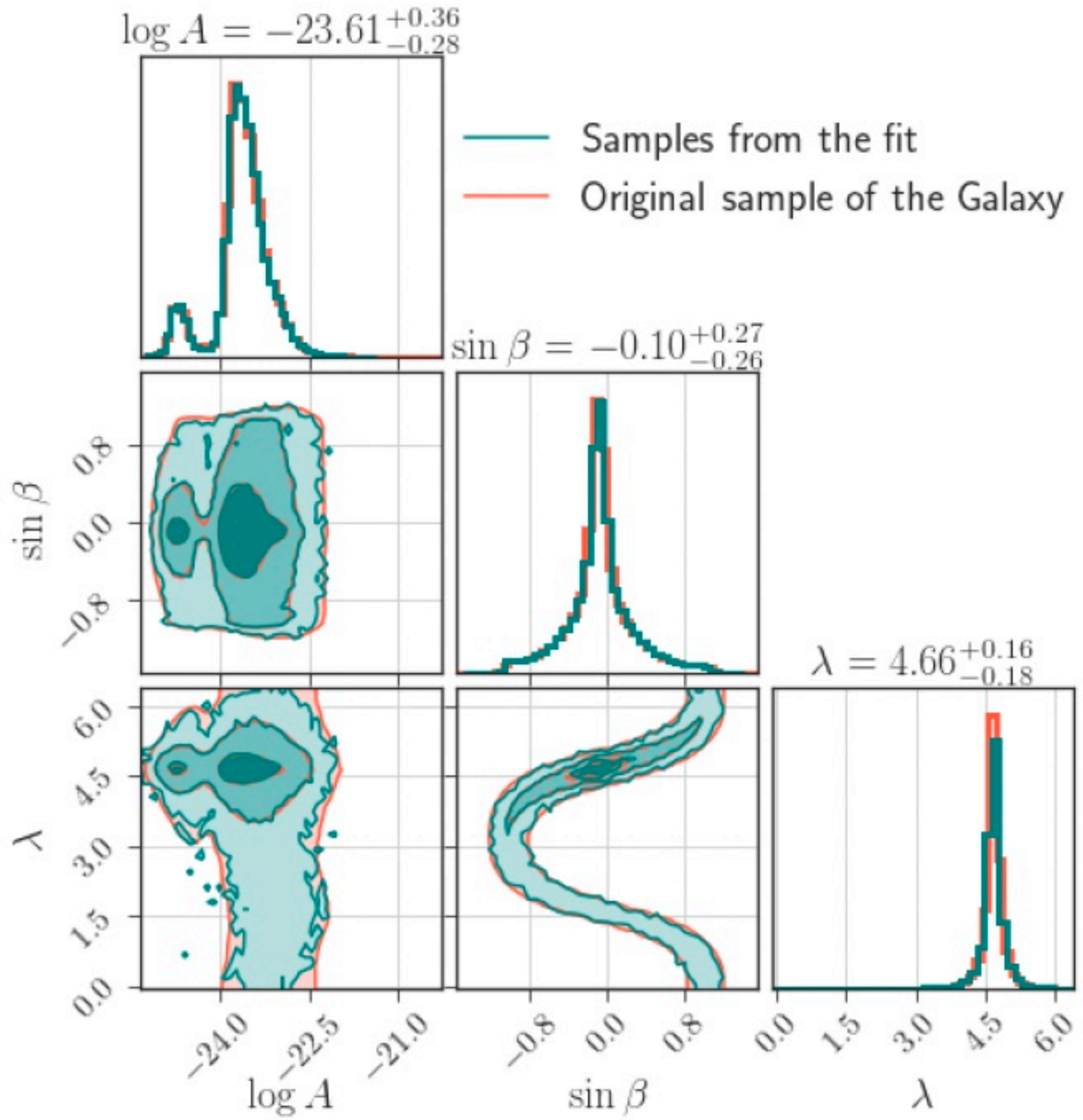
- Fit probability distribution function from the samples.
- Use Normalising Flows as a density estimator.
- Train network by optimising Kullback–Leibler divergence between samples and transformed base distribution.

$$KL(p||q) = \sum_x p(x) \log \left[\frac{p(x)}{q(x)} \right]$$

- Use estimated distribution for proposals.

Neural density estimators for proposals

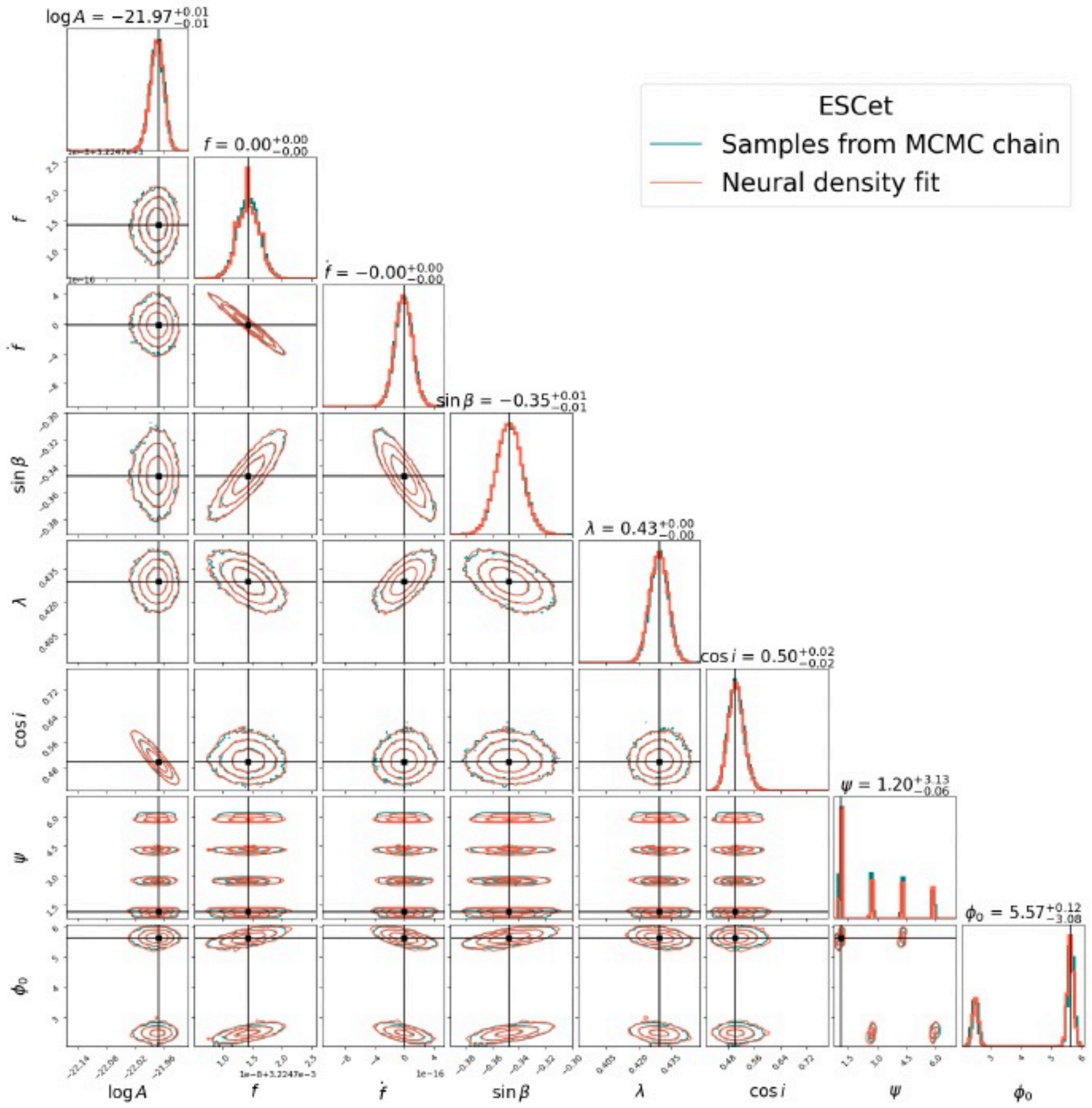
Priors and proposal for the Galaxy



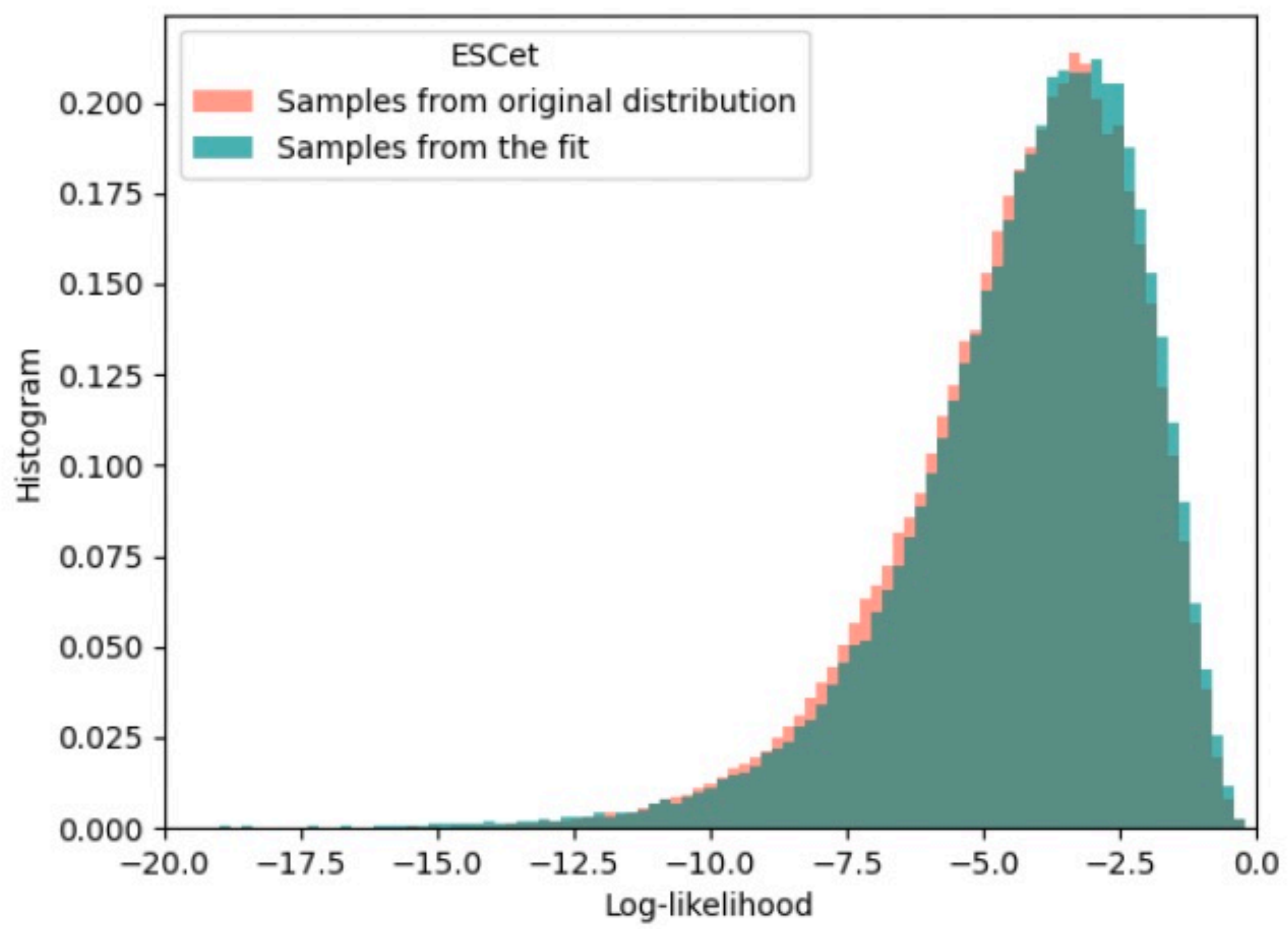
- The knowledge on the Galaxy distribution can be used either and a prior or as a proposal.

Neural density estimators for proposals

Priors and proposals from the previously estimated sources

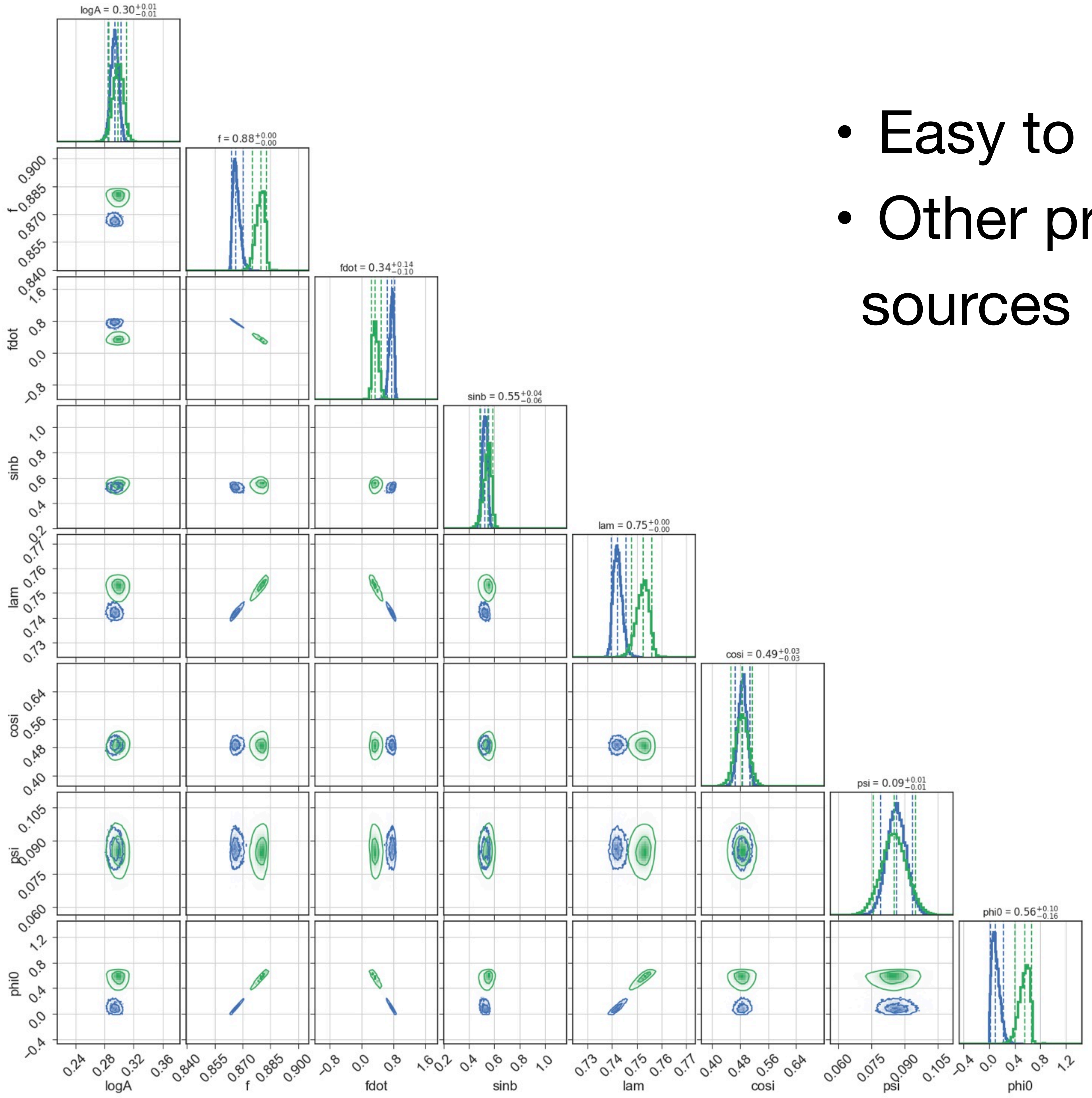


- Search for GBs from high to low SNRs
- Overtime we accumulate more data, so we need to update our estimated for the parameters
- We can use proposals based on the probabilities for the density fits to the already acquired posteriors



Neural density estimators for proposals

Case of overlapping signals

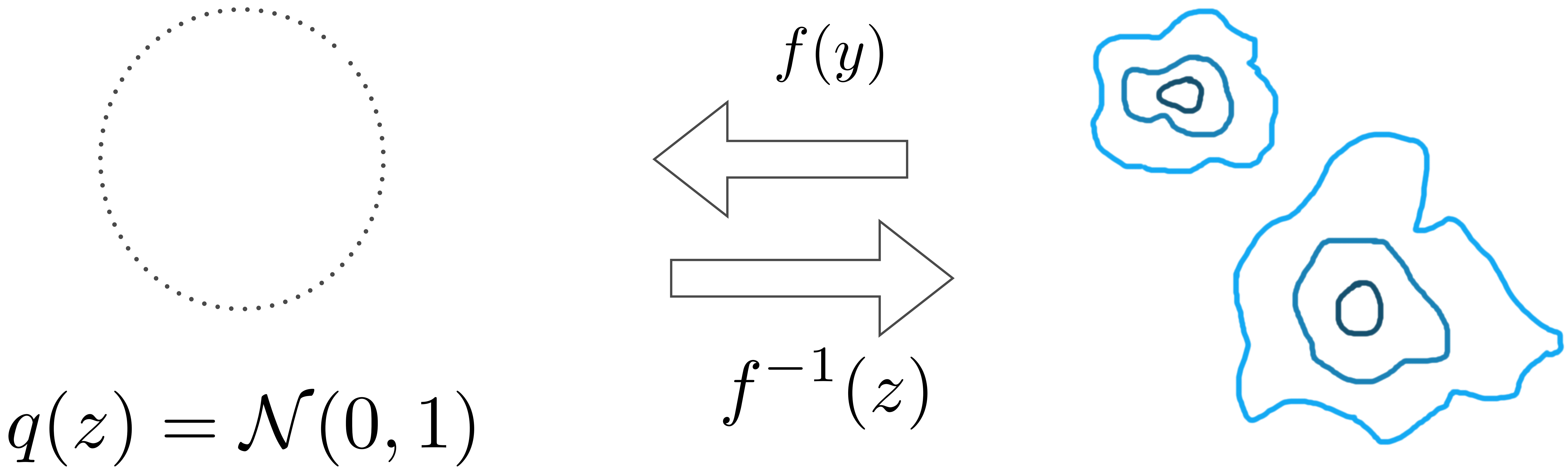


- Easy to extend to high dimensional data
- Other proposals will fail on the low SNR overlapping sources

Normalising flows for parameter estimation

Conditional density estimation

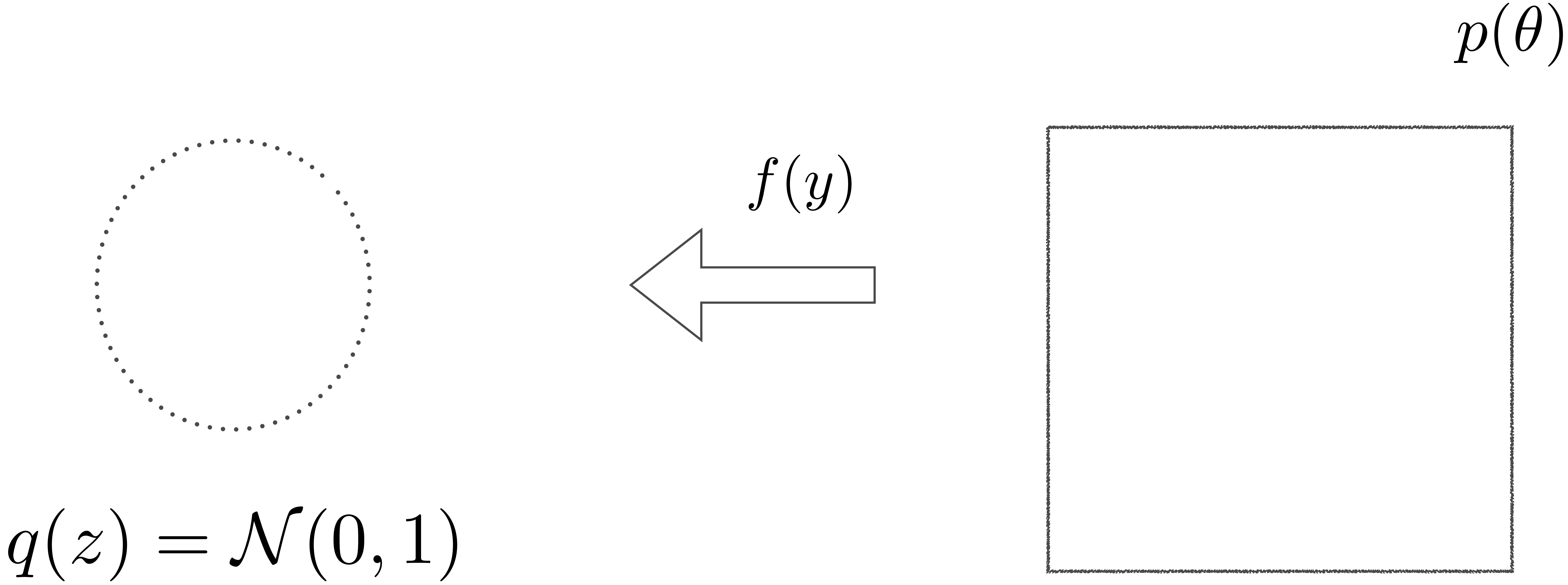
- Do not have access to samples from posterior



Normalising flows for parameter estimation

Conditional density estimation

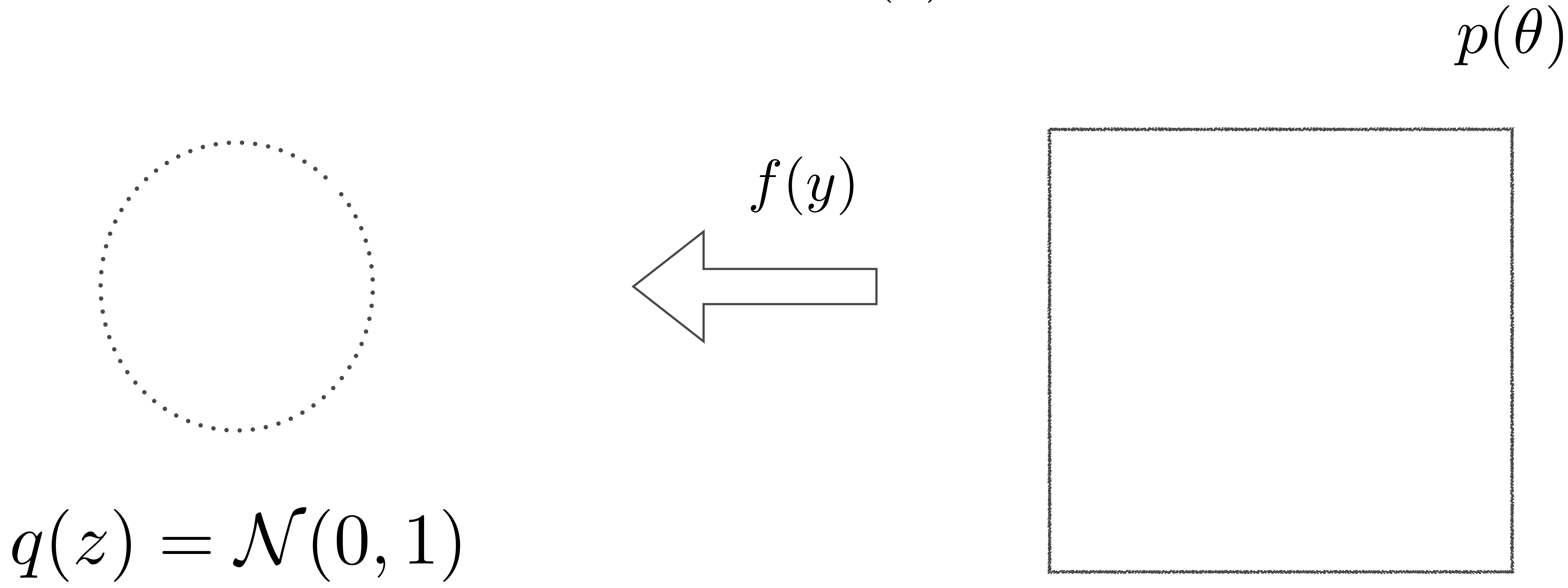
- Do not have access to samples from posterior
- Have access to samples from prior +



Normalising flows for parameter estimation

Conditional density estimation

- Do not have access to samples from posterior
- Have access to samples from prior +
- Can generated simulated data $x = h(\theta) + n$

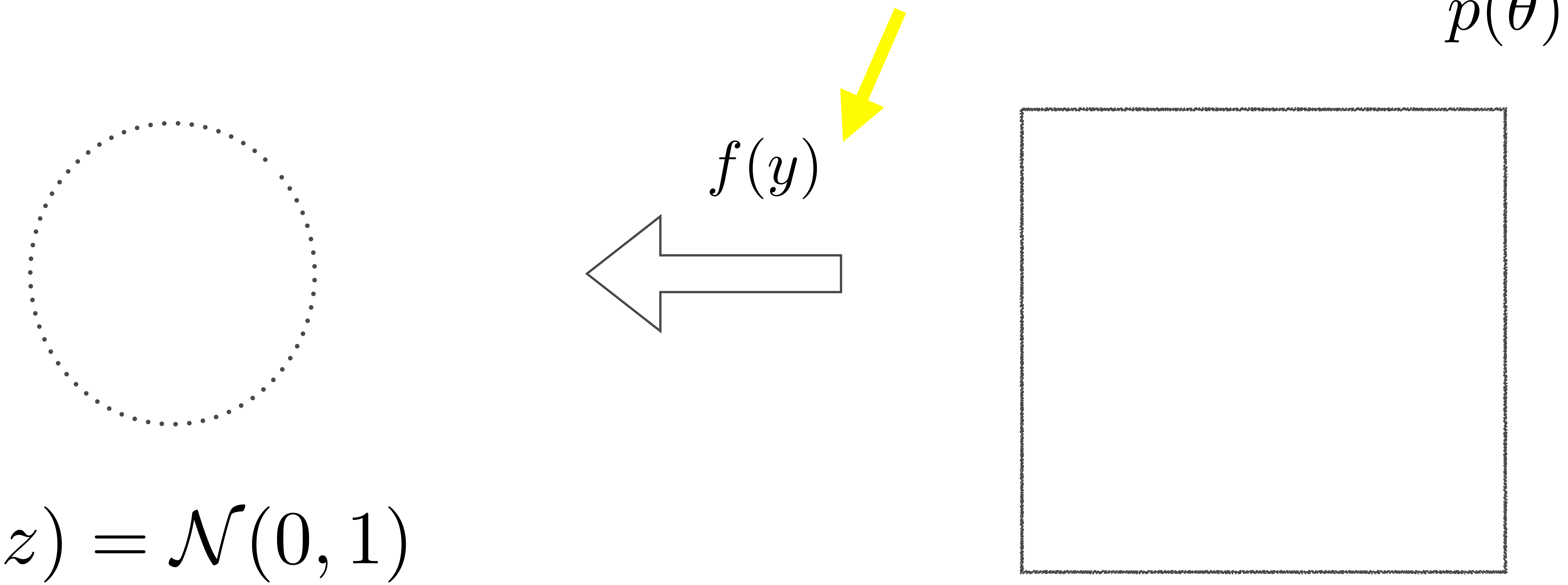


Normalising flows for parameter estimation

Conditional density estimation

Condition inverted map on real data

- Do not have access to samples from posterior
- Have access to samples from prior +
- Can generated simulated data $x = h(\theta) + n$



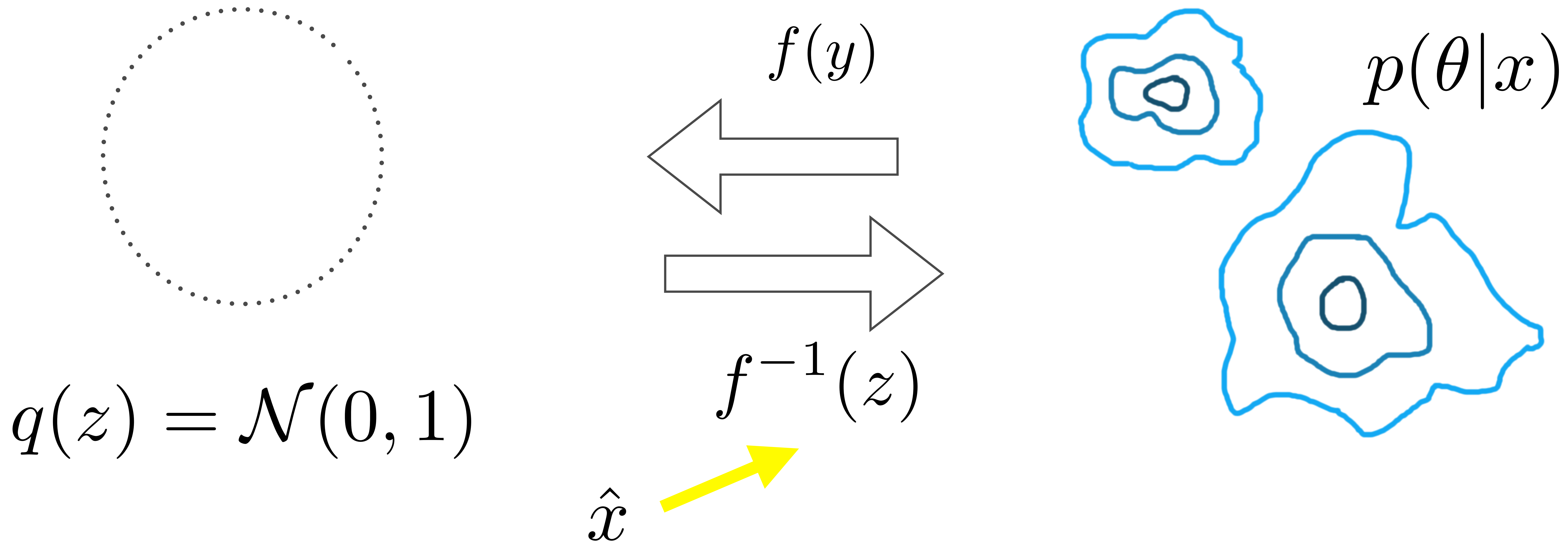
Therefore have access to the joint sample $p(x, \theta) = p(x \mid \theta)p(\theta)$

Normalising flows for parameter estimation

Conditional density estimation

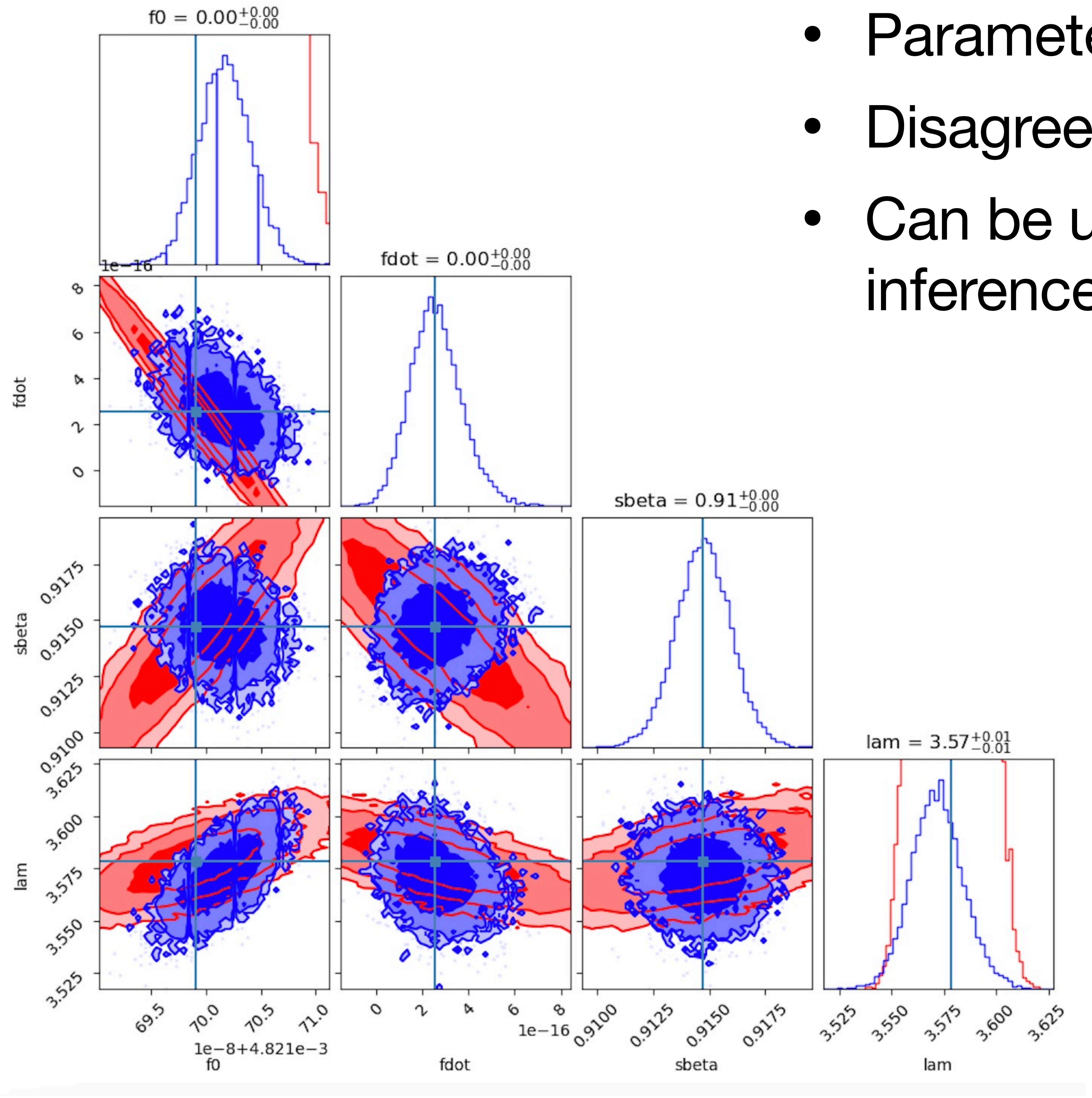
- Do not have access to samples from posterior
- Have access to samples from prior +
- Can generated simulated data $x = h(\theta) + n$

Condition inverted map on real data



Normalising flows for parameter estimation

Preliminary Results



- Parameter estimation results for single Galactic Binary
- Disagreement in posteriors can be solved by Importance Sampling
- Can be used as initial proposal or as a separate way to perform inference

Conclusions and outlook in the future

- Improves considerably sampling efficiency
- Still have to be properly incorporated to the Global analysis
- Combining sampling with the flow on the deeper level
- Hierarchical inference for the population of Galactic Binaries directly from the results of Global fit