Neural Empirical Bayes: Source Distribution Estimation and its Applications to Simulation-Based Inference

Maxime Vandegar, Michael Kagan, Antoine Wehenkel & Gilles Louppe

AISTATS

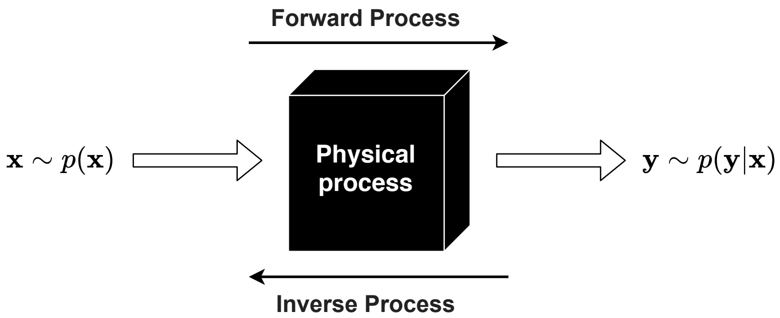
SLAC NATIONAL ACCELERATOR LABORATORY

Stanford University

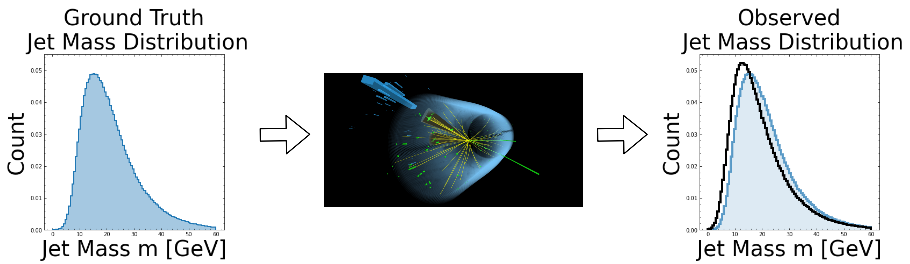
LIÈGE université

Motivation

- Aim to learn an unseen source distribution $p(x)$ that has generated observations $\{y_i\}_{i=1}^{N}$.



- Useful for learning:
  - Data-informed priors.
  - Latent variables with a physical meaning.
  - Densities from noise-corrupted samples (unfolding).



Method

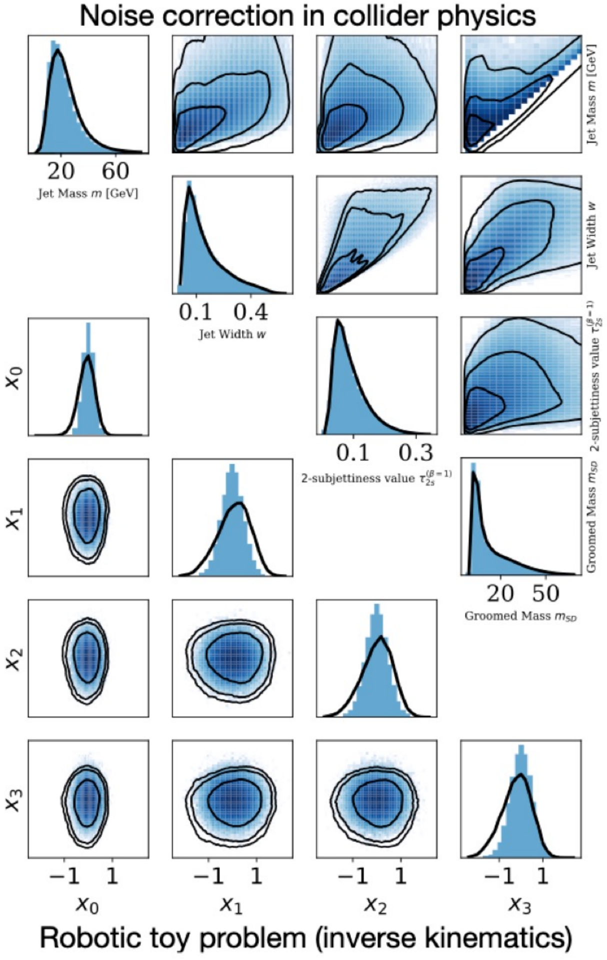
- Build an estimator $\mathcal{L}_K(\theta)$ of the log marginal likelihood & use it to learn θ, the parameters of a model of p(x).

$$\mathscr{L}_K(\theta) = \log \frac{1}{K} \sum_{k=1}^{K} p(y \mid x_k), \quad x_k = G_\theta(\epsilon_k), \quad \epsilon_k \sim p(\epsilon)\,.$$

- Biased but consistent.
  - $\mathbb{E}\left[\mathscr{L}_{K+1}(\theta)\right] \geq \mathbb{E}\left[\mathscr{L}_K(\theta)\right]$.
  - $\lim_{K\to\infty} \mathscr{L}_K(\theta) = \log p(y)$.
- Can be unbiased with the Russian roulette estimator.

- Generative model that approximates p(x).
- If it enables density estimation, variational methods can be used for learning θ.

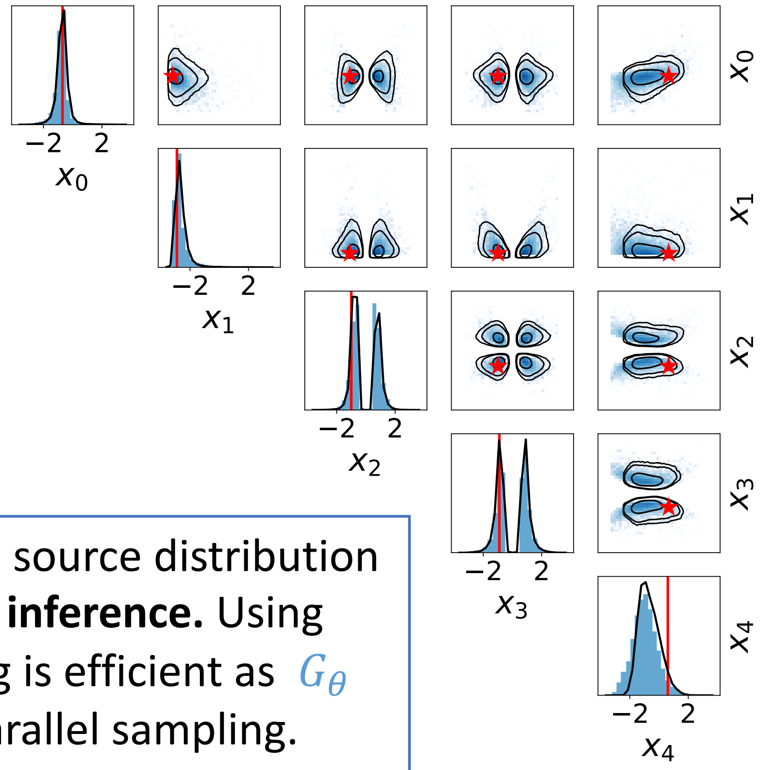
Results (biased estimator)



Source distributions $p(x)$ in blue against the estimated source distributions in black.

The **source distribution** learned with NEB can approximate well $p(x)$ in toy and collider physics experiments.

The exact posterior distribution in blue against its approximation in black obtained with rejection sampling. The generating source data are shown in red.



Once learned, the source distribution enables **posterior inference.** Using rejection sampling is efficient as G_θ allows efficient parallel sampling.

For more details, come and meet us during our poster session!