Precision-Machine Learning for the Matrix Element Method

Theo Heimel, Nathan Huetsch, Ramon Winterhalder, Tilman Plehn, Anja Butter

UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

Classical analysis

- hand-crafted observables
- binned data

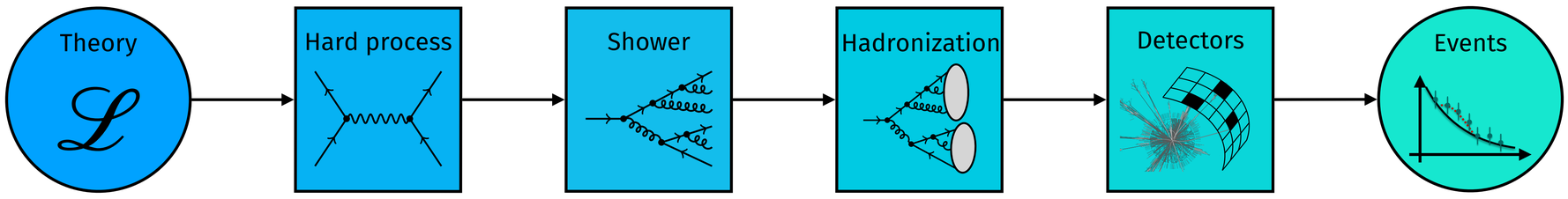
→ loss of information

How can we extract all the available information from LHC data?

Matrix Element Method (MEM)

- based on first principles
- estimates uncertainties reliably
- optimal use of information

→ perfect for processes with few events



Theory parameter α

known from theory

likelihood intractable → learn with neural network

Reconstructed momenta x_{reco}

$$p(x_{\text{reco}} \,|\, \alpha) = \int \mathrm{d}x_{\text{hard}}\; p(x_{\text{hard}} \,|\, \alpha)\; p(x_{\text{reco}} \,|\, x_{\text{hard}})\; \epsilon(x_{\text{hard}})$$

Efficient MC integration

importance sampling with normalizing flow

$x_{\text{hard}} \sim p(x_{\text{hard}} \,|\, x_{\text{reco}}, \alpha)$

Theory knowledge

diff. cross-section

$$\frac{1}{\sigma(\alpha)} \frac{\mathrm{d}\sigma(\alpha)}{\mathrm{d}x_{\text{hard}}}$$

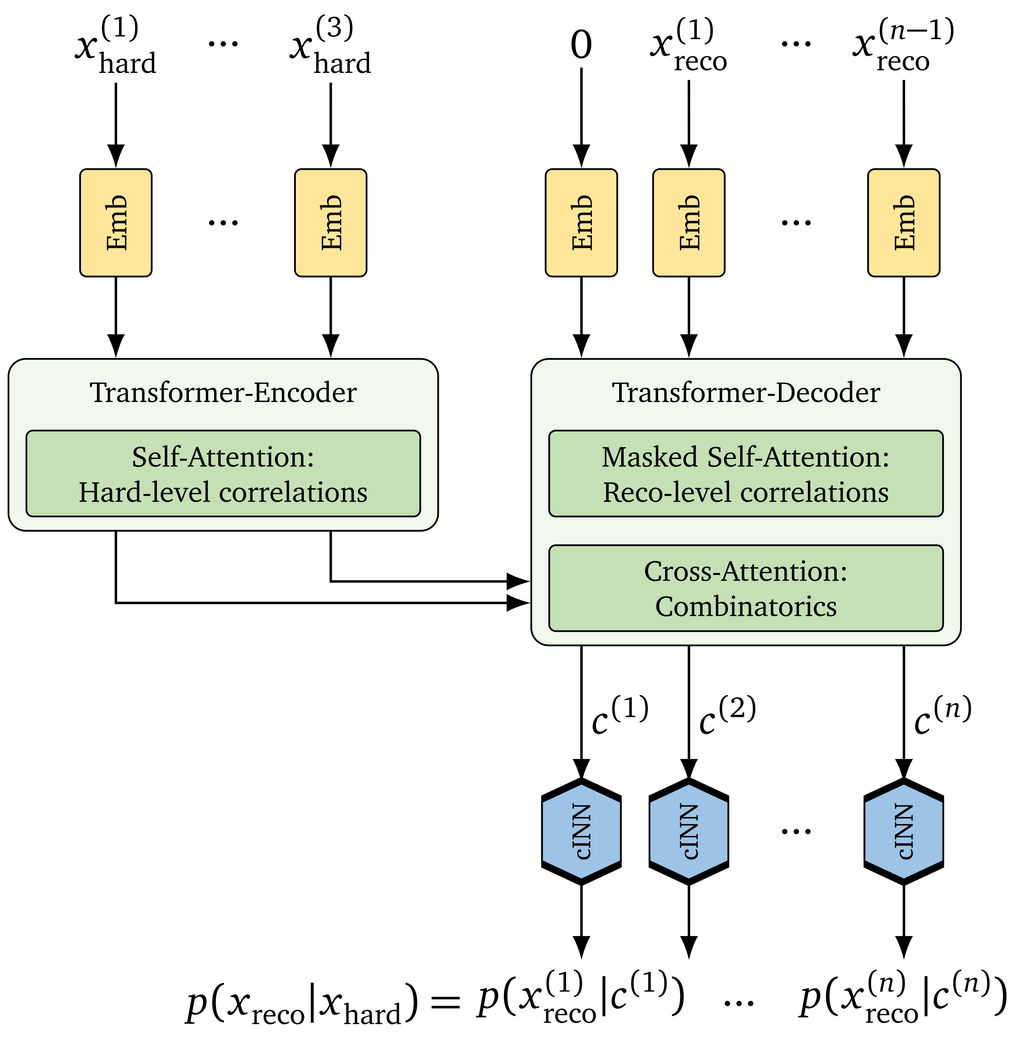
Transfer function

density estimation: normalizing flow and transformer

Acceptance function

learn with simple classifier network

Learning the transfer function



- **transformer**
  correlations between momenta, combinatorics
- **normalizing flow**
  likelihood for individual momenta
- **Bayesian networks**
  estimate training uncertainties

LHC example

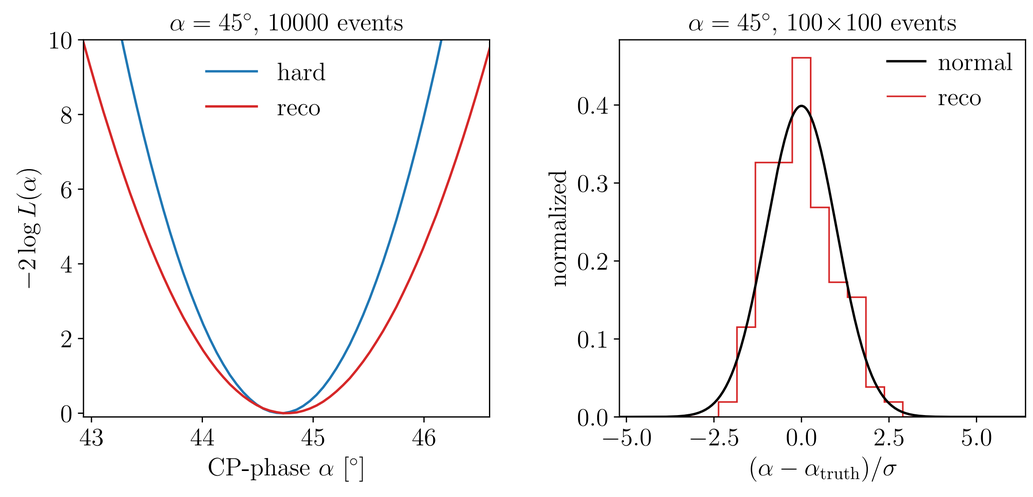
Single top and Higgs production with anomalous CP-phase α

Hadronic decay of top + ISR: $\mathrm{tHq} \to (\mathrm{bjj})\ (\gamma\gamma)\ \mathrm{j} + \text{QCD jets}$

- low total cross section (few events)
- low variation of rate
- kinematic observables still sensitive

→ ideal use case for MEM

Results



- smooth and well-calibrated likelihoods, both for low and high event counts
- close to optimal information
- Uncertainty bands: MC integration error & systematic error from limited training statistics (BNN)