

Precision-Machine Learning for the Matrix Element Method

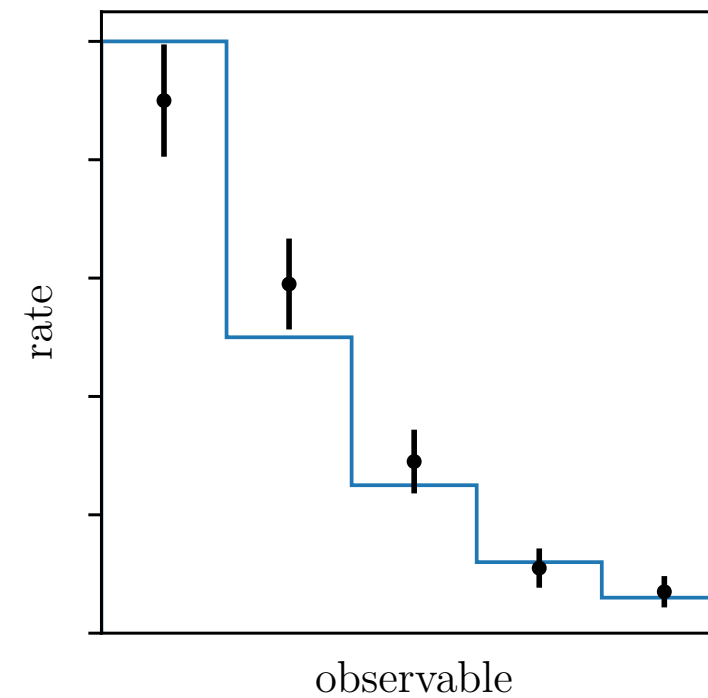
Theo Heimel
April 2024

Institut für theoretische Physik
Universität Heidelberg

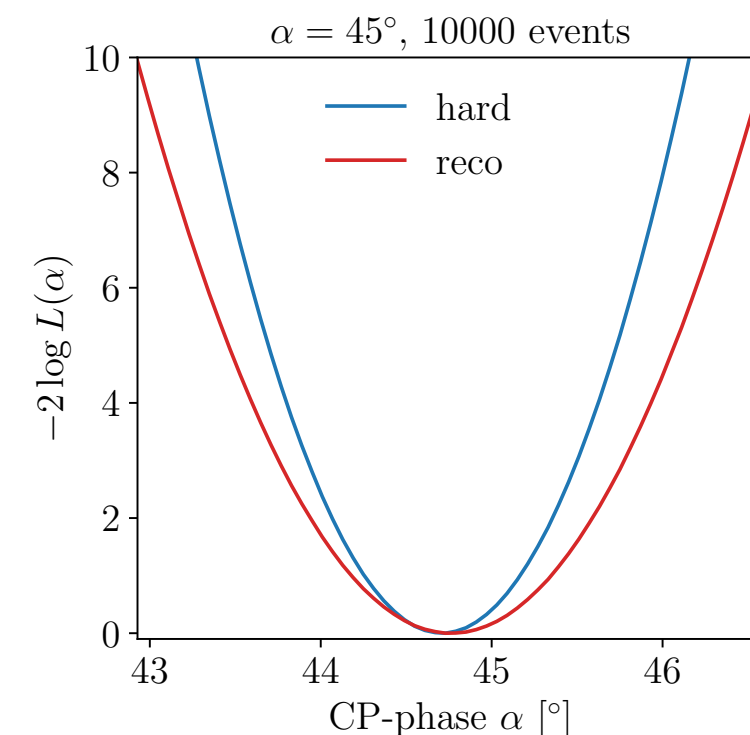


UNIVERSITÄT
HEIDELBERG
ZUKUNFT
SEIT 1386

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



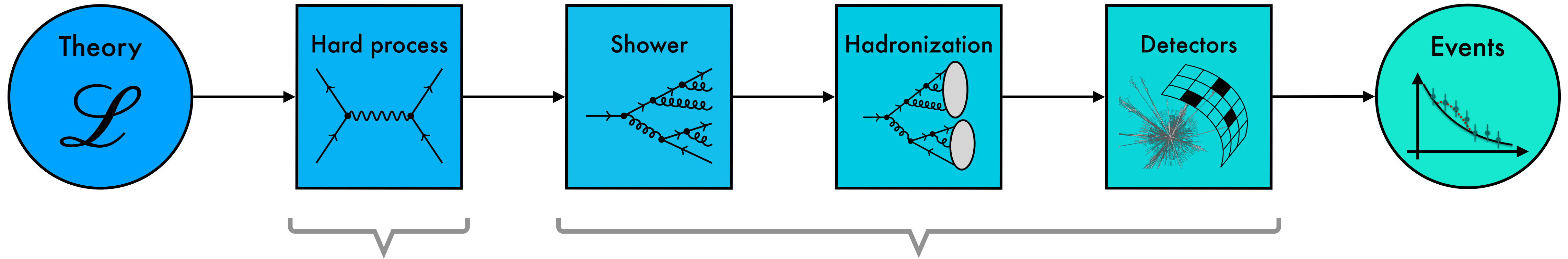
Binned, low-dimensional data
→ **loss of information**



Use theory knowledge to
extract likelihood

→ **matrix element method**

Precision-Machine Learning for the Matrix Element Method



known from theory

likelihood intractable
→ use machine learning

Precision-Machine Learning for the Matrix Element Method
 Theo Heinzel, Nathan Huetsch, Ramon Winterhalder, Tilman Plehn, Anja Butter

Classical analysis

- hand-crafted observables
- limited 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}

$$p(x_{\text{reco}} | \alpha) = \int d\mathcal{M}_{\text{hard}} p(\mathcal{M}_{\text{hard}} | \alpha) p(x_{\text{reco}} | \mathcal{M}_{\text{hard}}) \epsilon(\mathcal{M}_{\text{hard}})$$

Efficient MC integration
importance sampling with normalizing flow
 $p_{\text{hard}} = p(\mathcal{M}_{\text{hard}} | \alpha)$

Theory knowledge
diff. cross-section $\frac{1}{\text{dvol}} \frac{d\sigma}{d\text{vol}}$

Transfer function
density estimation; normalizing flow and transformer

Acceptance function
learn with simple classifier network

Learning the transfer function

LHC example
Single top and Higgs production with anomalous CP-phase α
Hadronic decay of top → $b\bar{b}$: $t\bar{t} \rightarrow (b\bar{b})(\gamma\gamma) + \text{OCD jets}$

- low total cross section (few events)
- low relation of rate
- kinematic observables still sensitive
- ideal use case for MEM

Results

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

Come to my poster to see how this can be done with transformers, normalizing flows, classifiers and neural importance sampling!