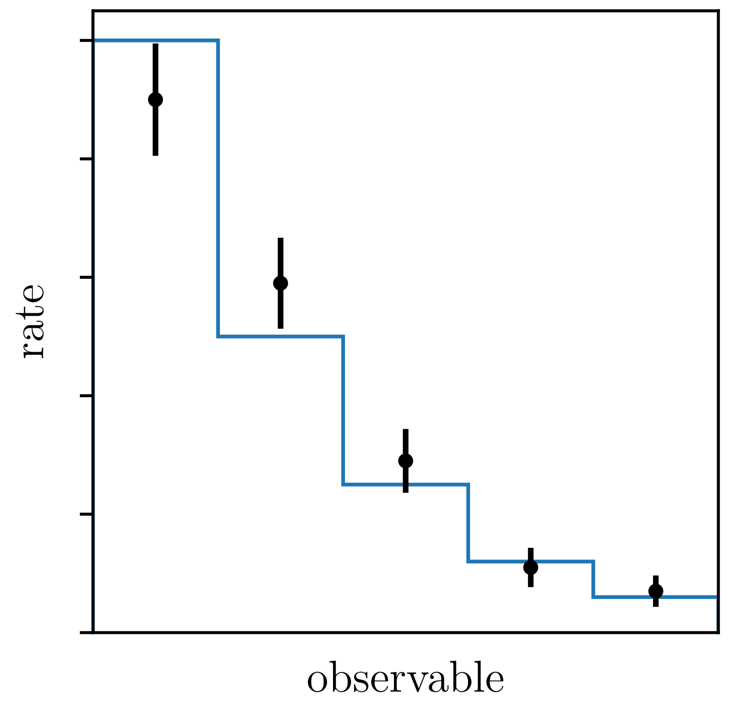
Precision-Machine Learning for the Matrix Element Method

Theo Heimel
April 2024

Institut für theoretische Physik
Universität Heidelberg

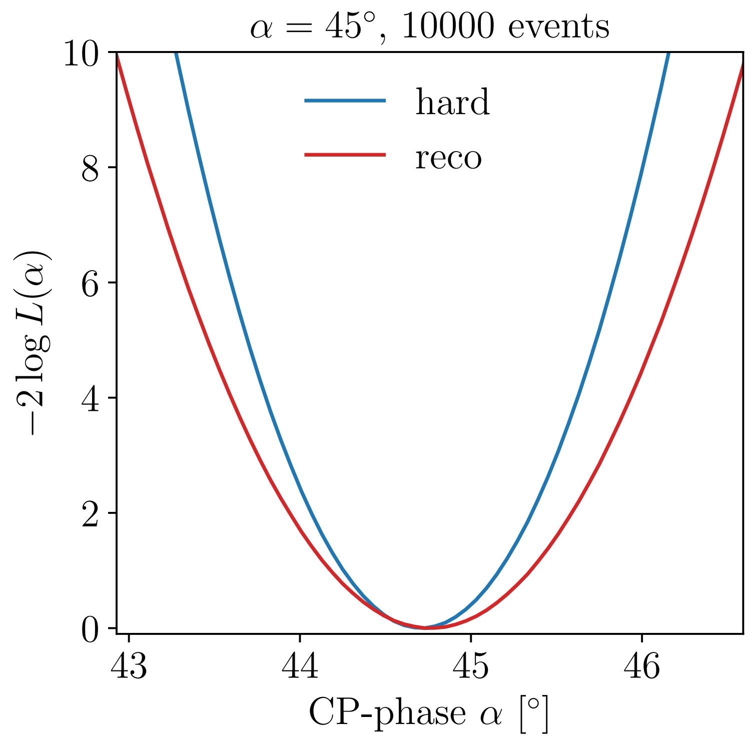
UNIVERSITÄT
HEIDELBERG
ZUKUNFT
SEIT 1386

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



Binned, low-dimensional data

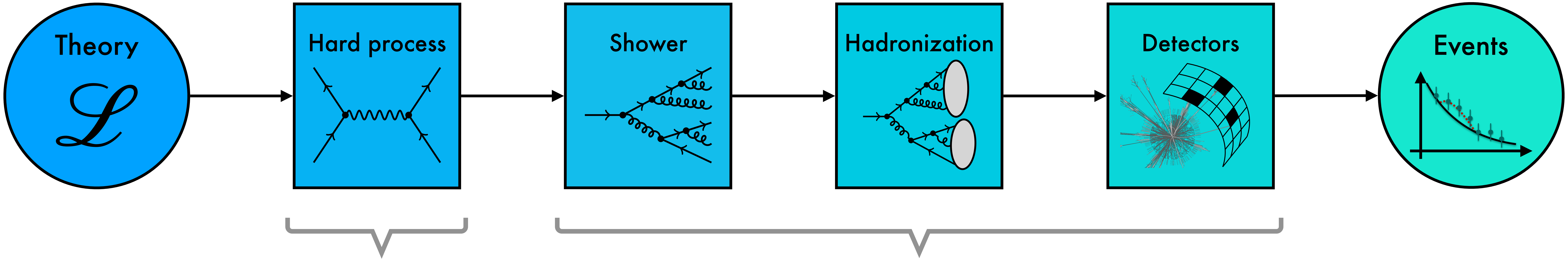
→ loss of information



Use theory knowledge to extract likelihood

→ matrix element method

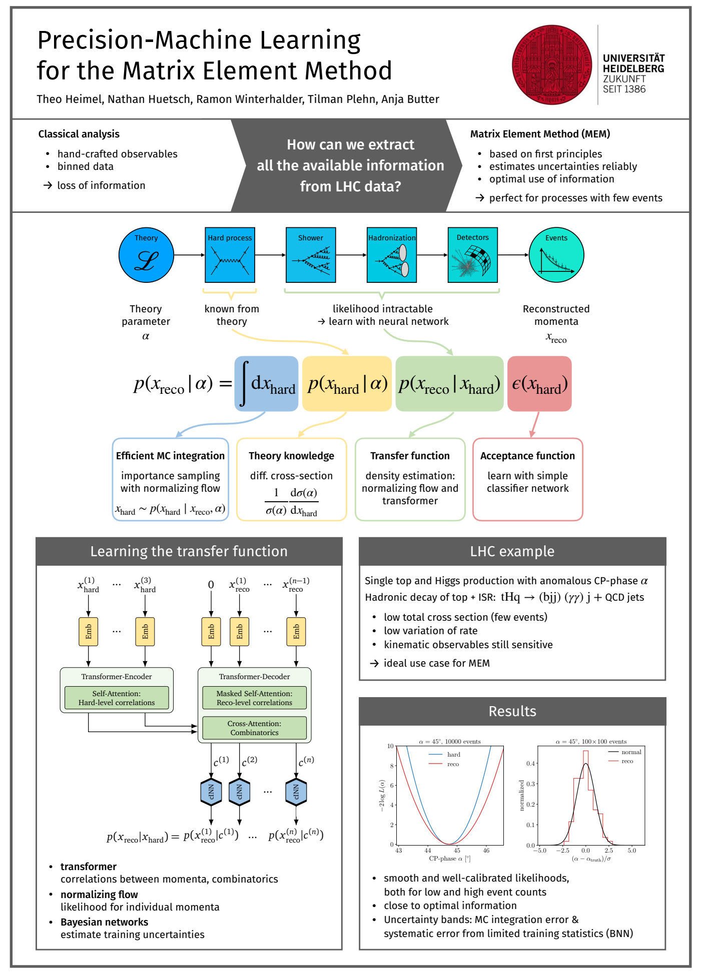
Precision-Machine Learning for the Matrix Element Method



known from theory

likelihood intractable

→ **use machine learning**



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

PDF uncertainties in the presence of inconsistent training data

Mark N. Costantini

EuCAIF Conference 2024, Amsterdam

POSTER: #100

Funded by the European Union

European Research Council
Established by the European Commission

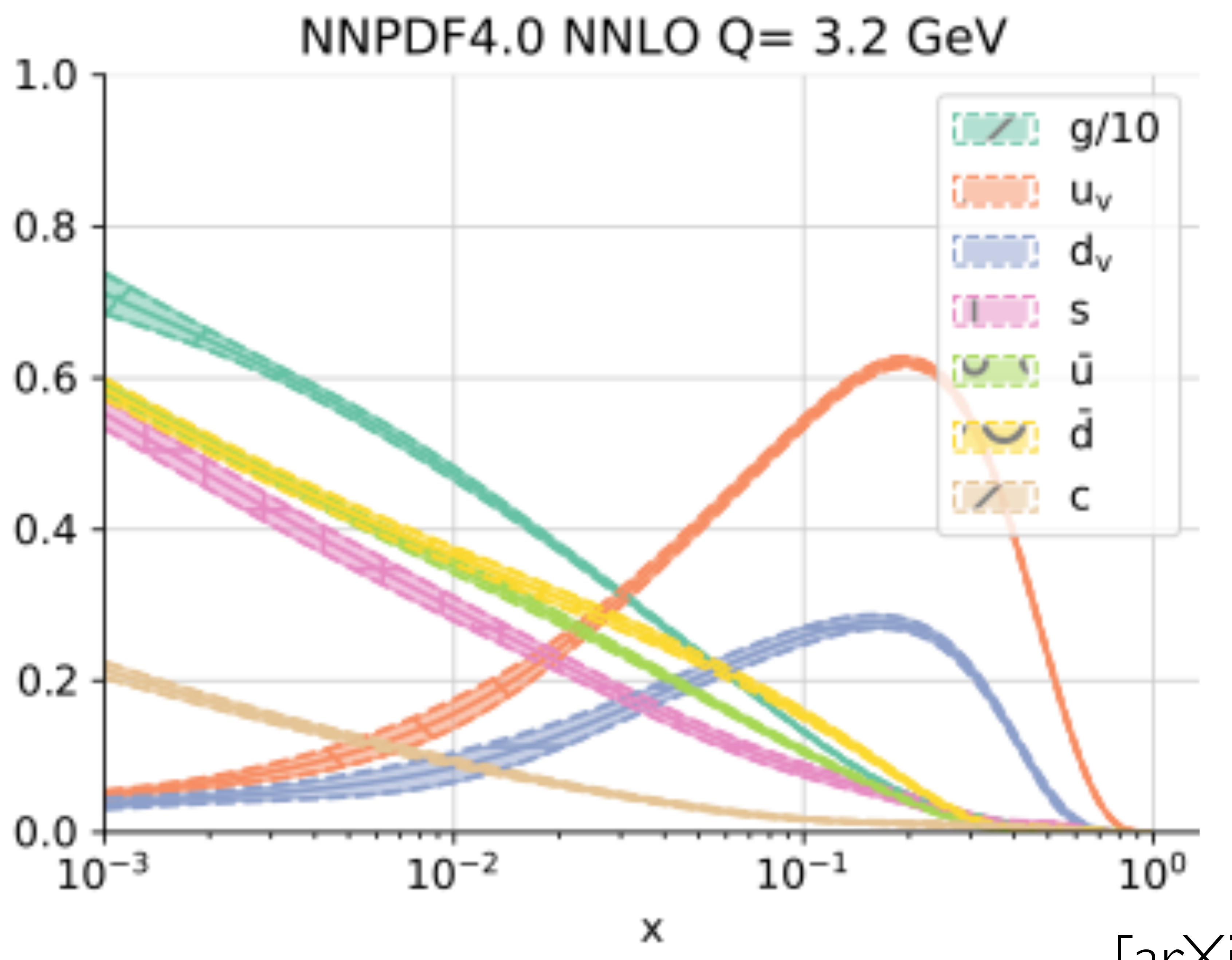
UNIVERSITY OF CAMBRIDGE

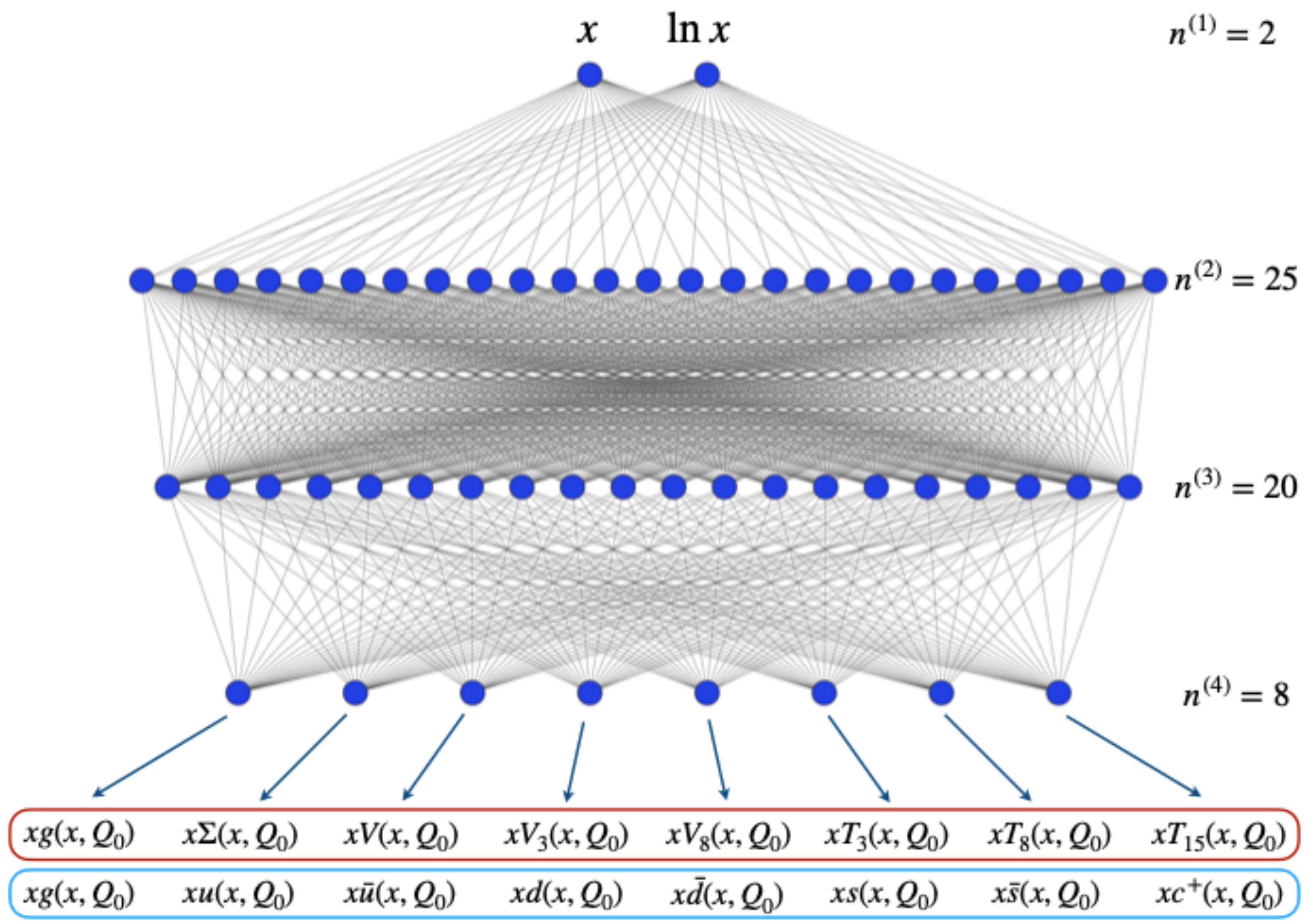
NNPDF

PBSP

Parton Distribution Functions and their uncertainties

- PDFs parameterise the structure of the proton in terms of its subnuclear constituents
- NNPDF: PDF model is parameterised by a dense Neural Network and uncertainties are estimated trough the Monte Carlo replica method [arXiV: 2404.10056]
- Fitted from data from several experimental sources → several sources of inconsistencies can affect fit



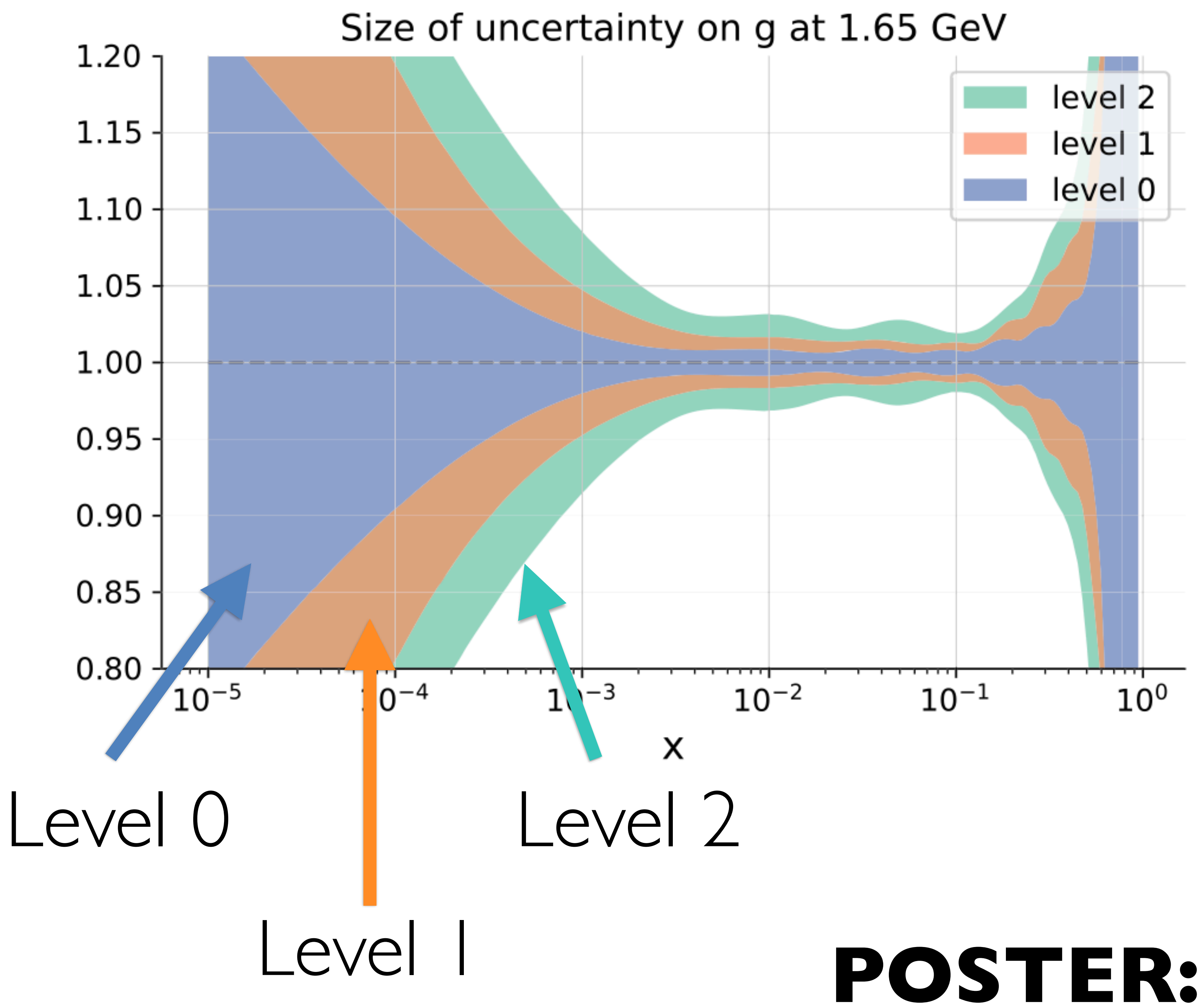


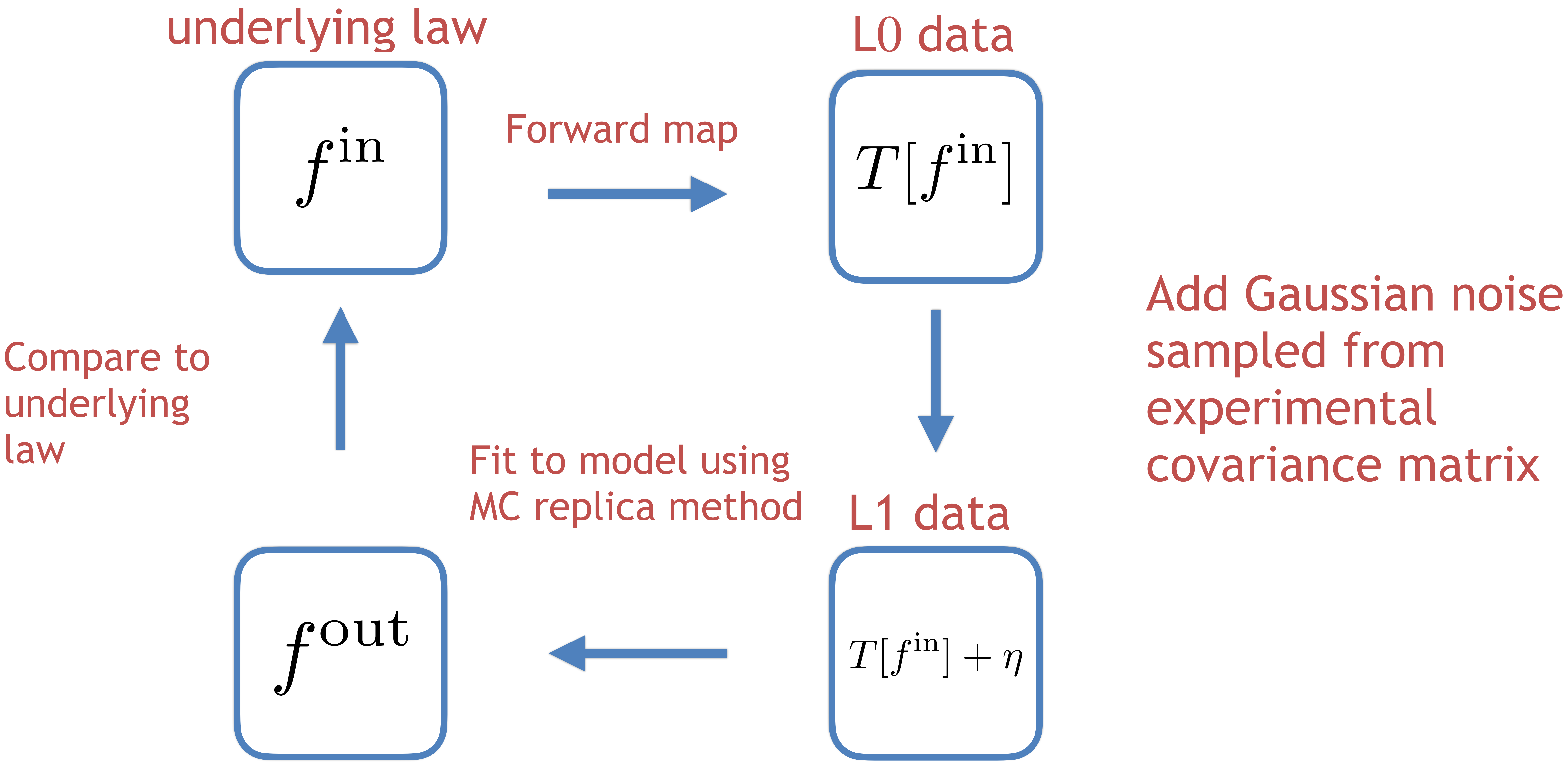
POSTER: #100

[arXiv: 2109.02653]

Inconsistent training data

- Inconsistency is defined as a missing or underestimated experimental uncertainty
- Closure test: powerful framework to test fitting methodology [arXiV: 2111.05787]





POSTER: #100

Evaluating Generative Models with non-parametric two-sample tests

Samuele Grossi

Amsterdam, EuCAIFCon, 30th April 2024

Based on work in collaboration with Riccardo Torre

Two-sample test in high energy physics

Two-Sample Test: understand if two independent data samples are drawn from the same probability density function (PDF)

- **PARAMETRIC**: Some assumptions on the underlying distributions of the samples are needed to perform the test
- **NON-PARAMETRIC**: Only the data are used to perform the test, without any assumption on the underlying distributions

In high energy physics: two-sample test to compare data provided by two different generators.
Example: Powheg/MadGraph vs Neural Networks

Objectives and procedure

Purpose of the work: provide a systematic analysis of non-parametric two-sample test using different evaluation metrics

- Use reference distributions to :
  - Extract toy samples to build null hypothesis distributions
  - Perform the likelihood-ratio test

- Carry out non-parametric tests using different metrics. Compare performances to the likelihood-ratio

Particle physics → high dimensional datasets

We focused on *univariate integral probability measure* metrics:
easy to implement, fast results, scale well with the dimensions

Estimation of ML model uncertainty in particle physics event classifiers

Miguel Cárdenas Montes[1] miguel.cardenas@ciemat.es

José M. Hernández[1] jose.hernandez@ciemat.es **Julia Vázquez Escobar**[1] julia.vazquez@ciemat.es

[1]Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (**CIEMAT**)

EuCAIFCon 2024

May 30, 2024

GOBIERNO DE ESPAÑA MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES

Ciemat Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas

cfp CIEMAT física de partículas

Julia Vázquez-Escobar, J.M. Hernández, Miguel Cárdenas-Montes, Computer Physics Communications, 268, (2021)

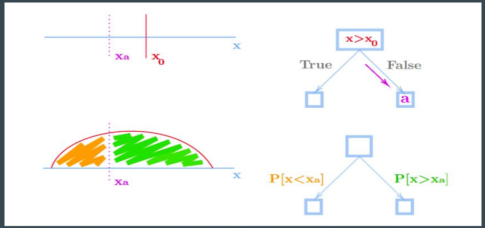
Methods

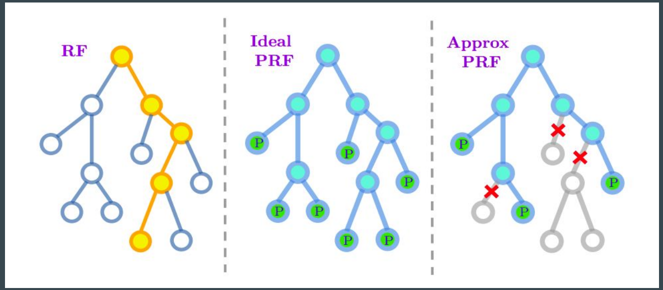
Bayesian approx.

$$\widetilde{\mathbb{E}}\left[\mathbf{y}^*\right] := \frac{1}{T}\sum_{t=1}^{T}\mathbf{f}^{\hat{\omega}_t}\left(\mathbf{x}^*\right) \underset{T\to\infty}{\longrightarrow} \mathbb{E}_{q^*_\theta(\mathbf{y}^*|\mathbf{x}^*)}\left[\mathbf{y}^*\right]$$

$$\widetilde{\mathrm{Var}}\left[\mathbf{y}^*\right] := \tau^{-1}\mathbf{I} + \frac{1}{T}\sum_{t=1}^{T}\mathbf{f}^{\hat{\omega}_t}\left(\mathbf{x}^*\right)^T\mathbf{f}^{\hat{\omega}_t}\left(\mathbf{x}^*\right) - \widetilde{\mathbb{E}}\left[\mathbf{y}^*\right]^T\widetilde{\mathbb{E}}\left[\mathbf{y}^*\right] \underset{T\to\infty}{\longrightarrow} \mathrm{Var}_{q^*_\theta(\mathbf{y}^*|\mathbf{x}^*)}\left[\mathbf{y}^*\right]$$

Probabilistic RF





Local ensembles

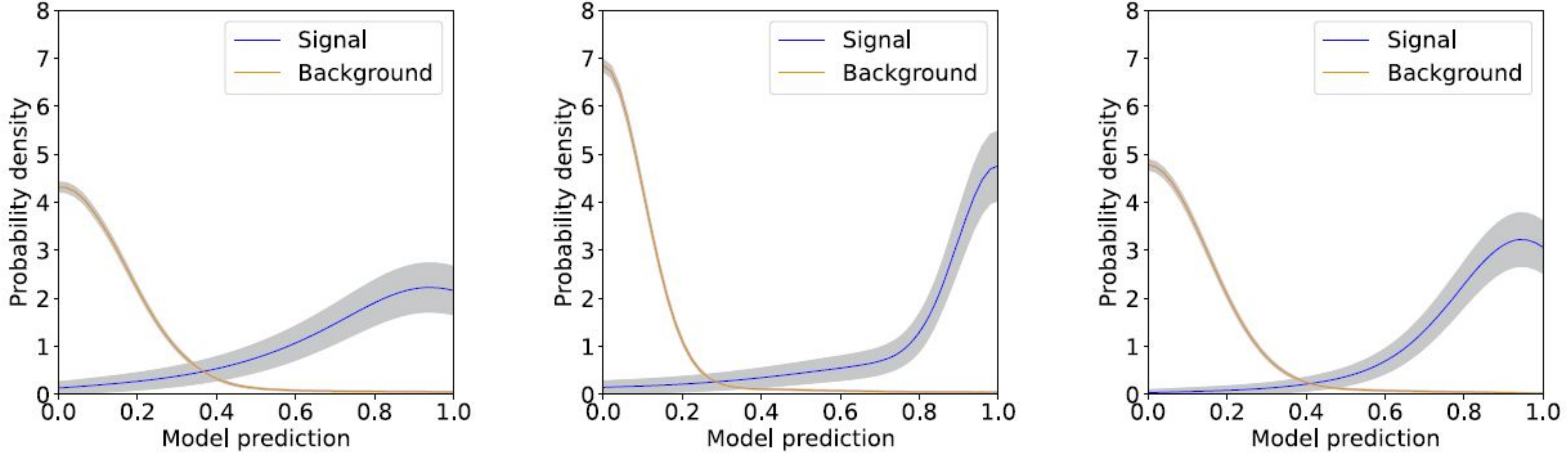
Proposition 1.1 *Let Δ_θ be the projection of a random perturbation with mean zero and covariance proportional to the identity $\epsilon \cdot I$ into the ensemble subspace spanned by $\{\xi_{(j)} : j > m\}$. Let P_Δ be the linearized change in prediction induced by the perturbation*

$$P_\Delta(x') := g_{\theta*}(x')^\top\Delta_\theta \approx \hat{y}(x', \theta * + \Delta_\theta) - \hat{y}(x', \theta*)$$

then $\mathcal{E}_m(x') = \varepsilon^{-1/2} \cdot SD(P_\Delta(x'))$.

Results

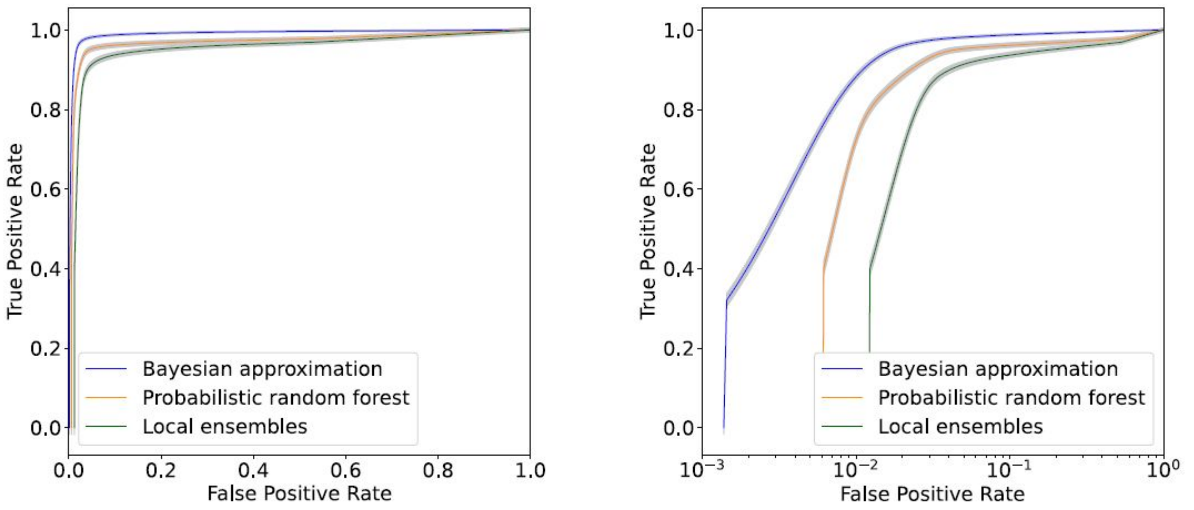
The probability density functions of the classification parameter for true signal and background events are shown.



(a) Local Ensembles (b) Probabilistic Random Forest (c) Bayesian approximation

Results

Model	AUC
Probabilistic Random Forest	0.969 ± 0.005
Local Ensembles	0.951 ± 0.006
Bayesian Approximation	0.990 ± 0.001



(a) ROC curve.

(b) ROC curve with x-axis in logarithmic scale.