Anomaly Detection on BESIII EMC Using Machine Learning

Mingrun Li, Chunxiu Liu, Xiaobin Ji

mrli@ihep.ac.cn

22 Oct. 2024

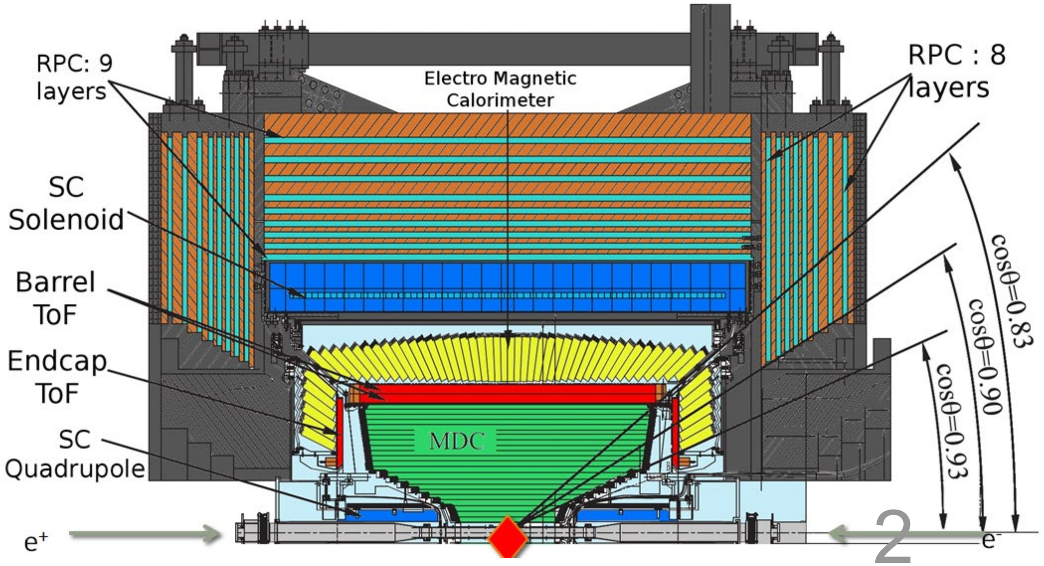
Introduction of BESIII and BEPCII

Beijing Electron Positron Collider (BEPCII)

- ➢ Peak luminosity: $10^{33}\ \mathrm{cm}^{-2}\mathrm{s}^{-1}$
- ➢ E_{CM} : $2.0 \sim 4.95$ GeV (5.8 GeV in the future)

Beijing Spectrometer (BESIII)

- ➢ Dataset: $10\mathrm{B}\ J/\psi$, $2.7\mathrm{B}\ \psi(3686)$, $20\mathrm{fb}^{-1}\ \psi(3770)$, $29\mathrm{fb}^{-1}$ $4.0\sim4.95$ GeV data
- ➢ Plays an important role in charmed flavor and hadron physics
- ➢ Composed of
  - Multilayer drift chamber (MDC)
  - Time-of-flight system (TOF)
  - Electro-Magnetic Calorimeter (EMC)
  - Muon Identifier (MUC)



Electro-Magnetic Calorimeter (EMC) on BESIII

- Consists of **6240** CsI(Tl) crystals, divided into 1 barrel and 2 endcaps
- Energy resolution: $\sigma_E/E \sim 2.5\% \ (1\ \mathrm{GeV})$
- Position resolution: $0.6\ \mathrm{cm}\ (1\ \mathrm{GeV})$

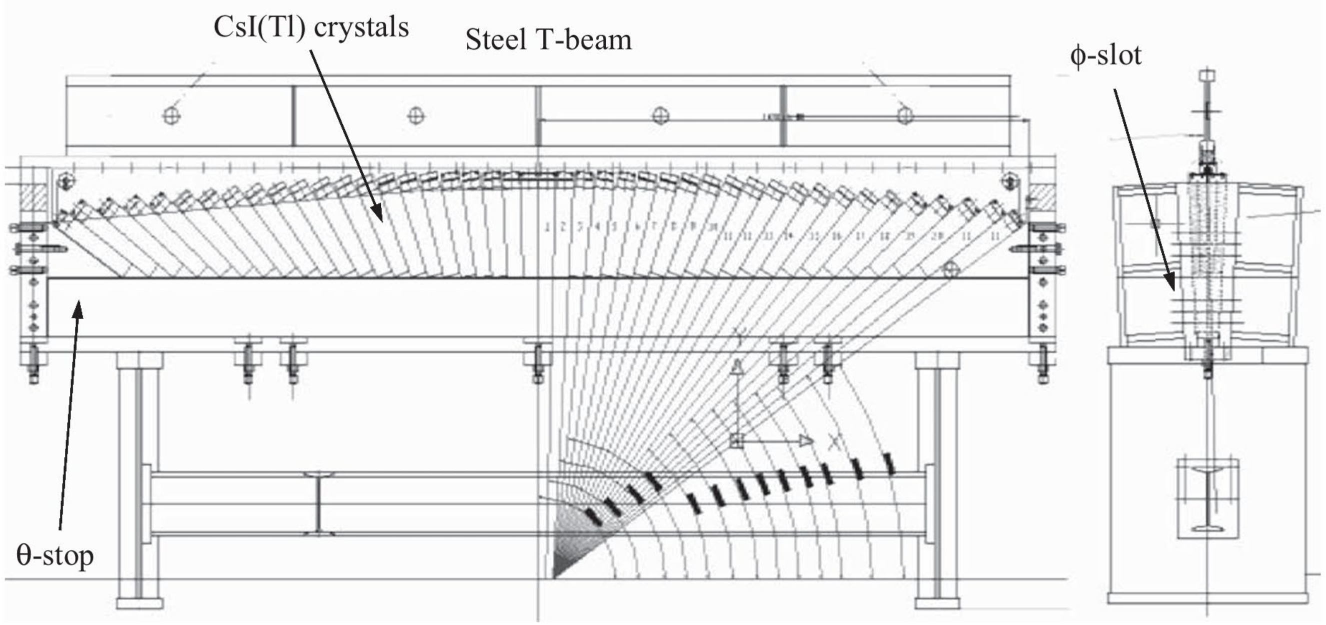
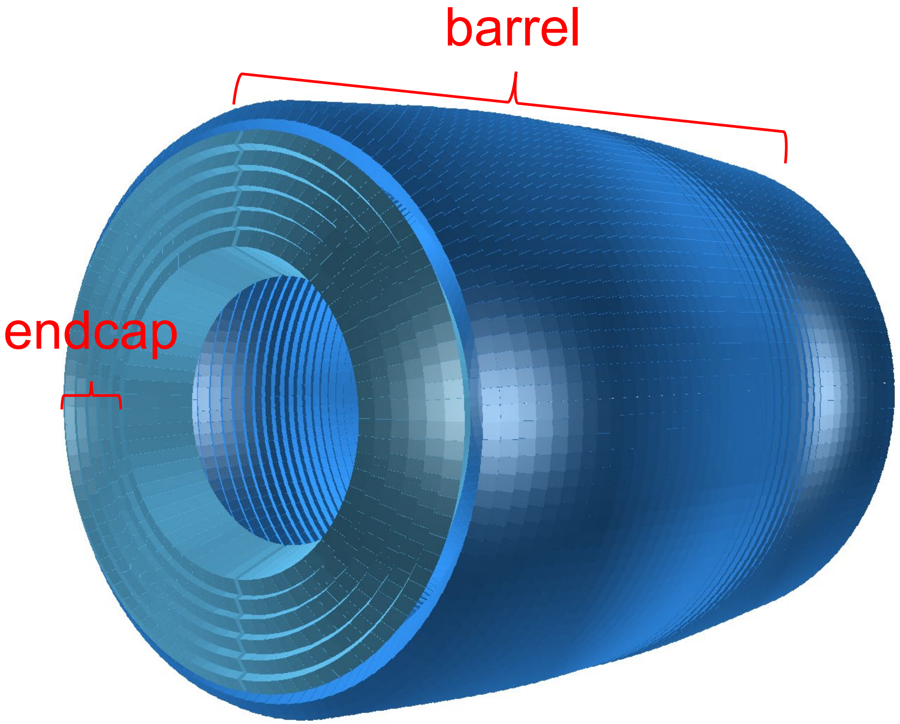
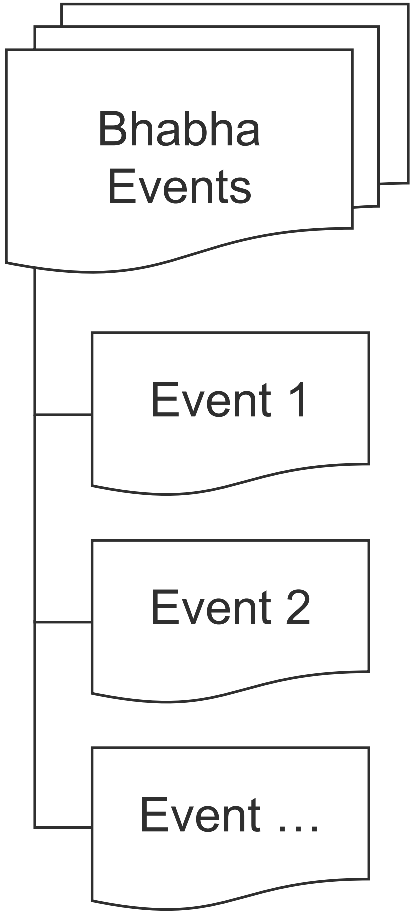


Fig. 45. The side and cross-sectional views of the barrel super module assembly jig.



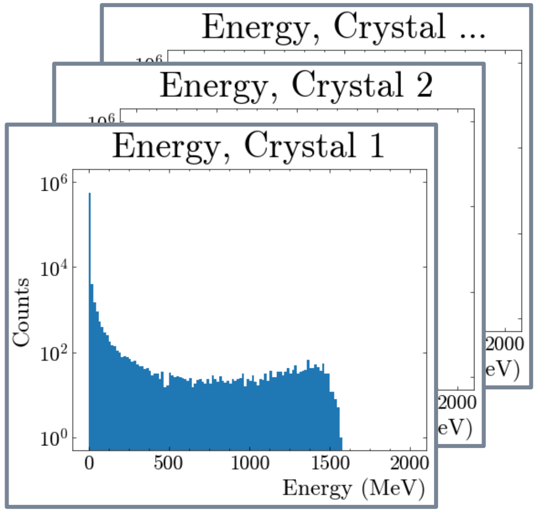
Data Preparation



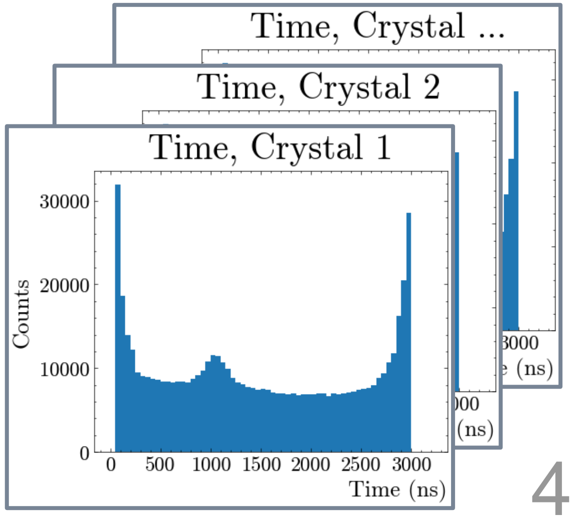
EMC Hits Format

Crystal ID	Energy	Time
ID 1	Energy 1	Time 1
ID 3	Energy 2	Time 2
ID 2	Energy 3	Time 3
ID 1	Energy 4	Time 4
ID 2	Energy 5	Time 5
…	…	…

6240 Energy Histograms



6240 Time Histograms



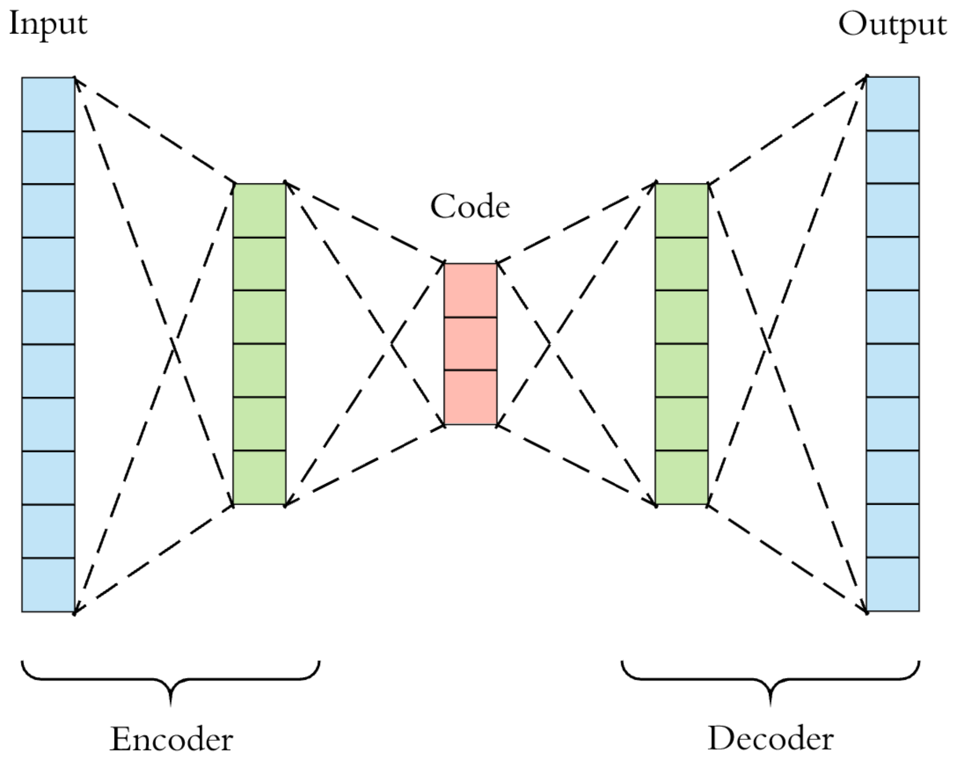
➢ Loop over EMC hits, fill histograms by Crystal ID

Detect Anomaly with Autoencoder

Autoencoder (AE): A special neural network
- Input shape = Output shape;
- Tighter in the middle layers.

Training
- Minimize loss between input and output;
  - Equivalent to ask model to copy what it receives.

To minimize total loss, AE will learn encode/decode rule for **majority** of data, which is **not suitable** for the remaining **minority** data.



Model Structure

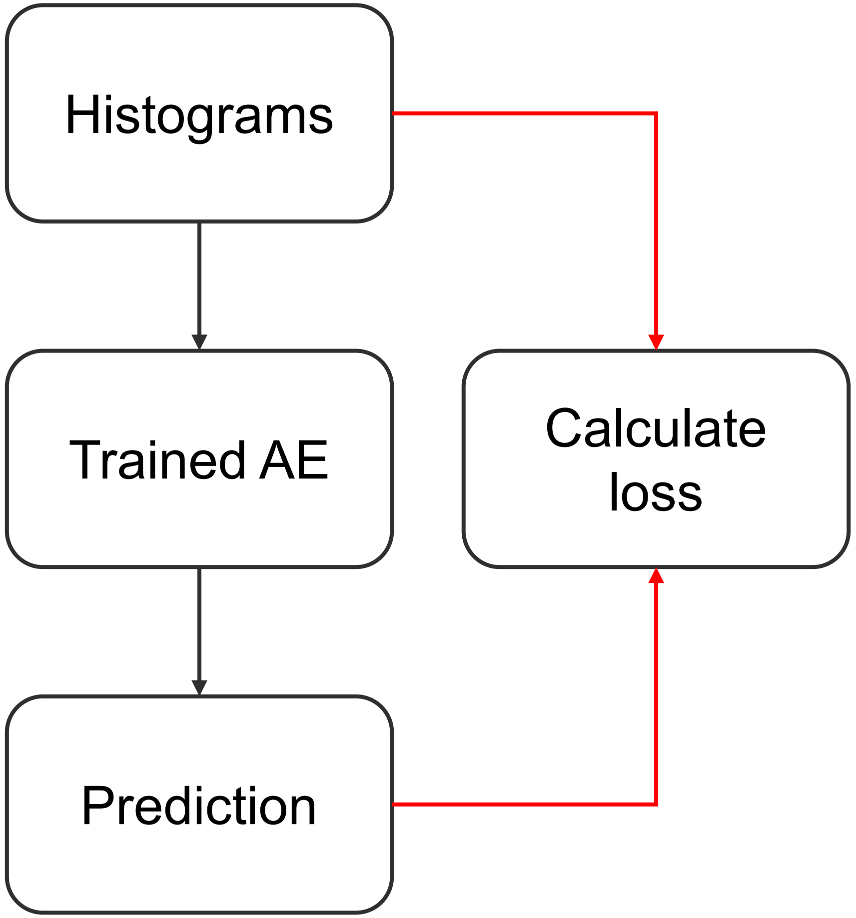
Detect Anomaly with Autoencoder

1. Assumptions:
   - "Majority" is "Normal";
   - "Minority" is "Anomaly".

2. Train autoencoder with histograms;
   (AE will perform badly for minority data)

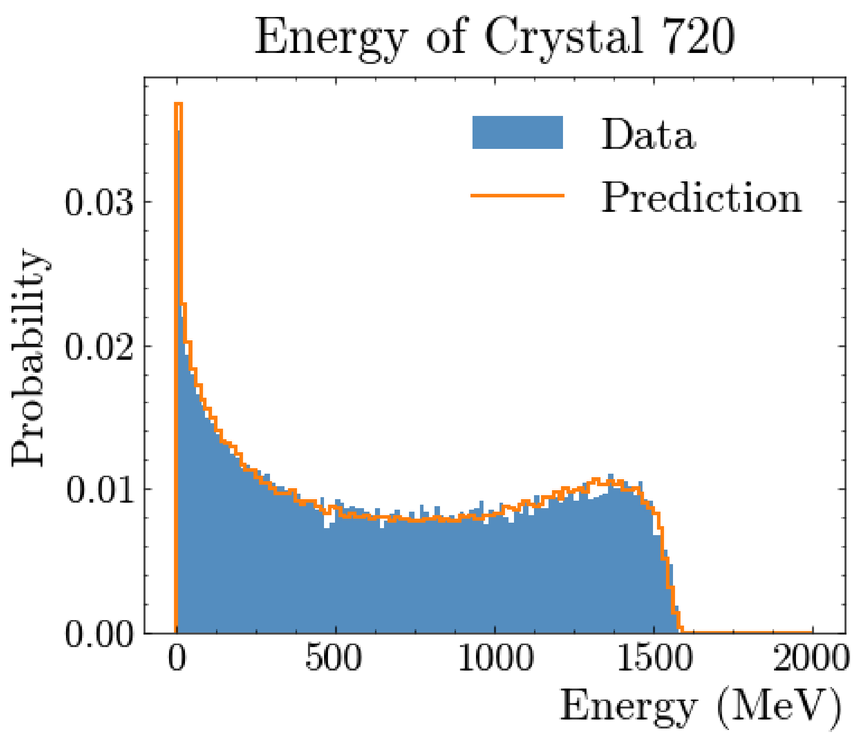
3. Obtain model's predictions for each histogram.

4. Calculate $\mathrm{Loss}(\mathrm{histogram}, \mathrm{prediction})$, consider histograms with high loss as anomalies.

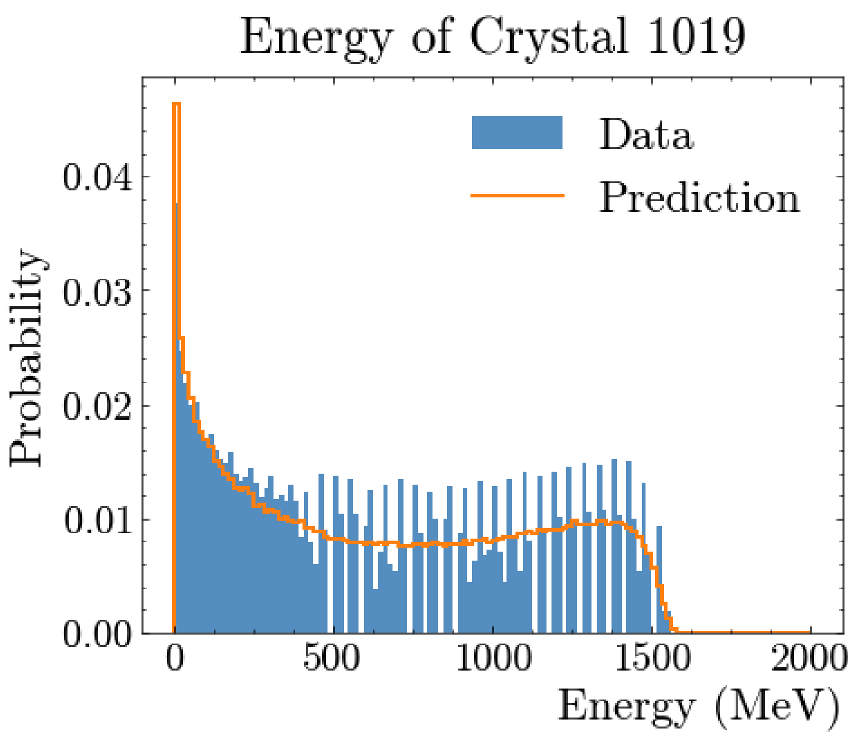


Detect Anomaly with Autoencoder

Model predictions for normal (720) and abnormal (1019) crystal

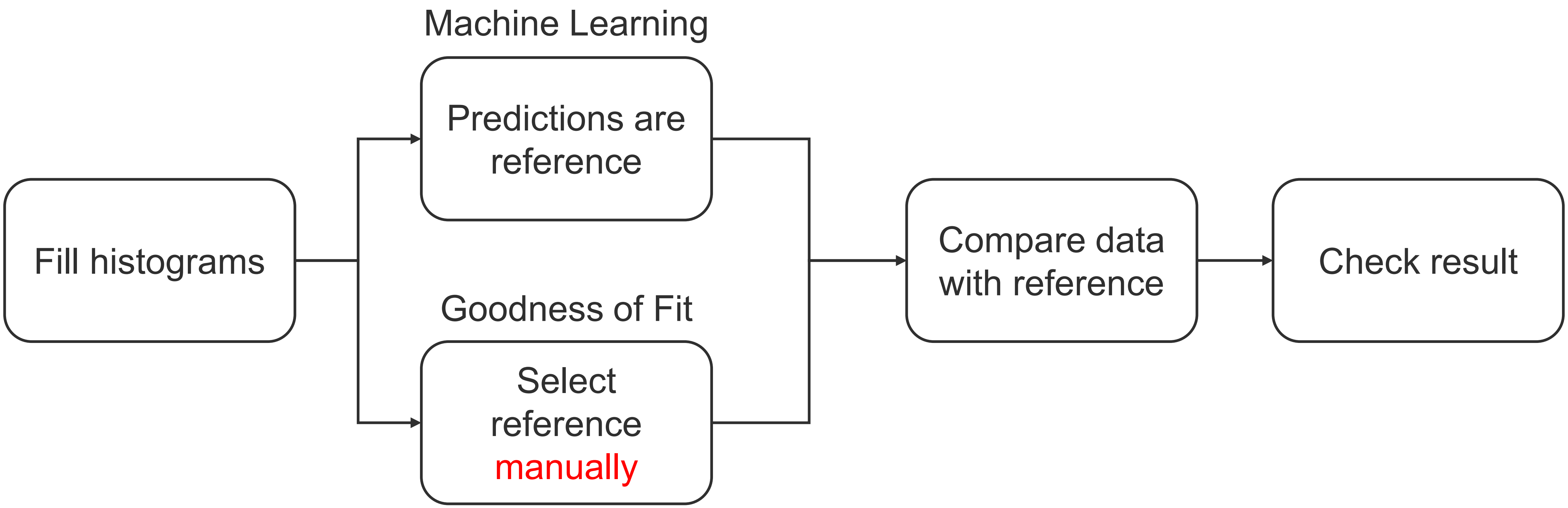


Low loss



High loss

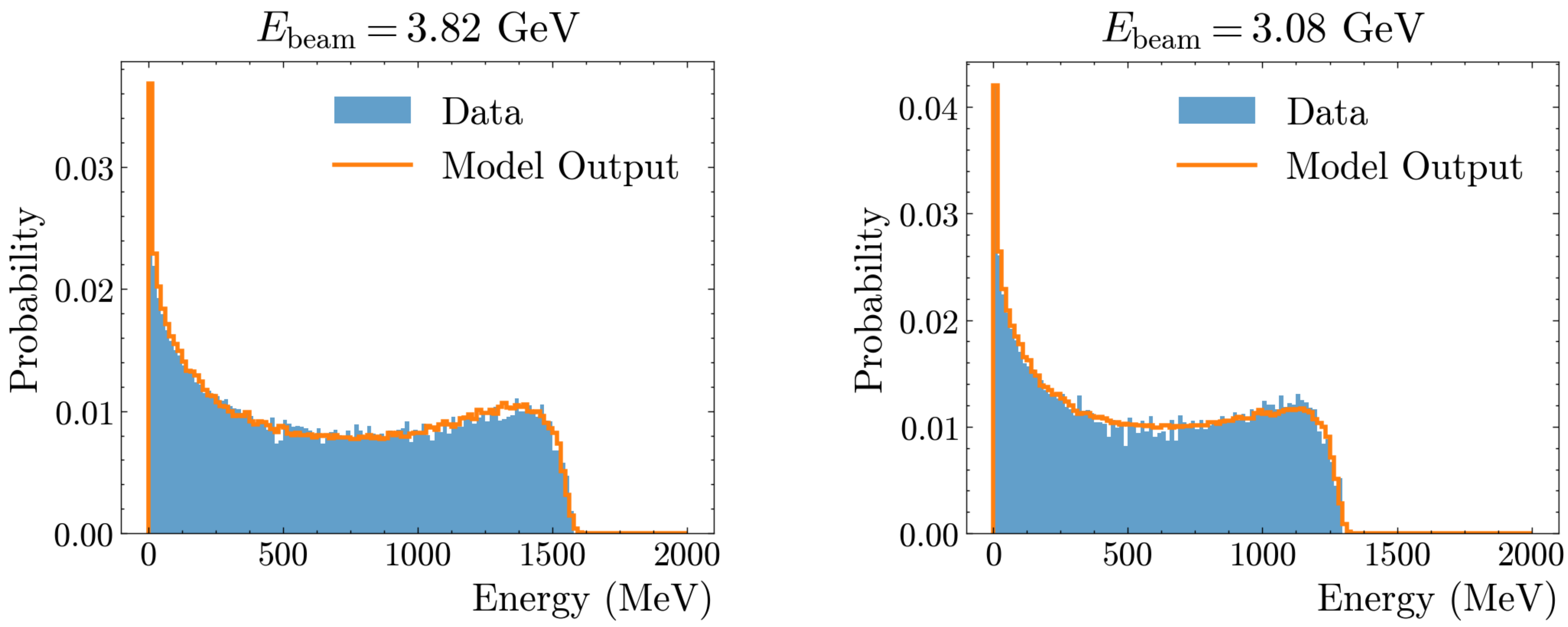
Advantages of Machine Learning Method



For Goodness of Fit method, we need to select new reference manually once outer condition (such as beam energy) changes.

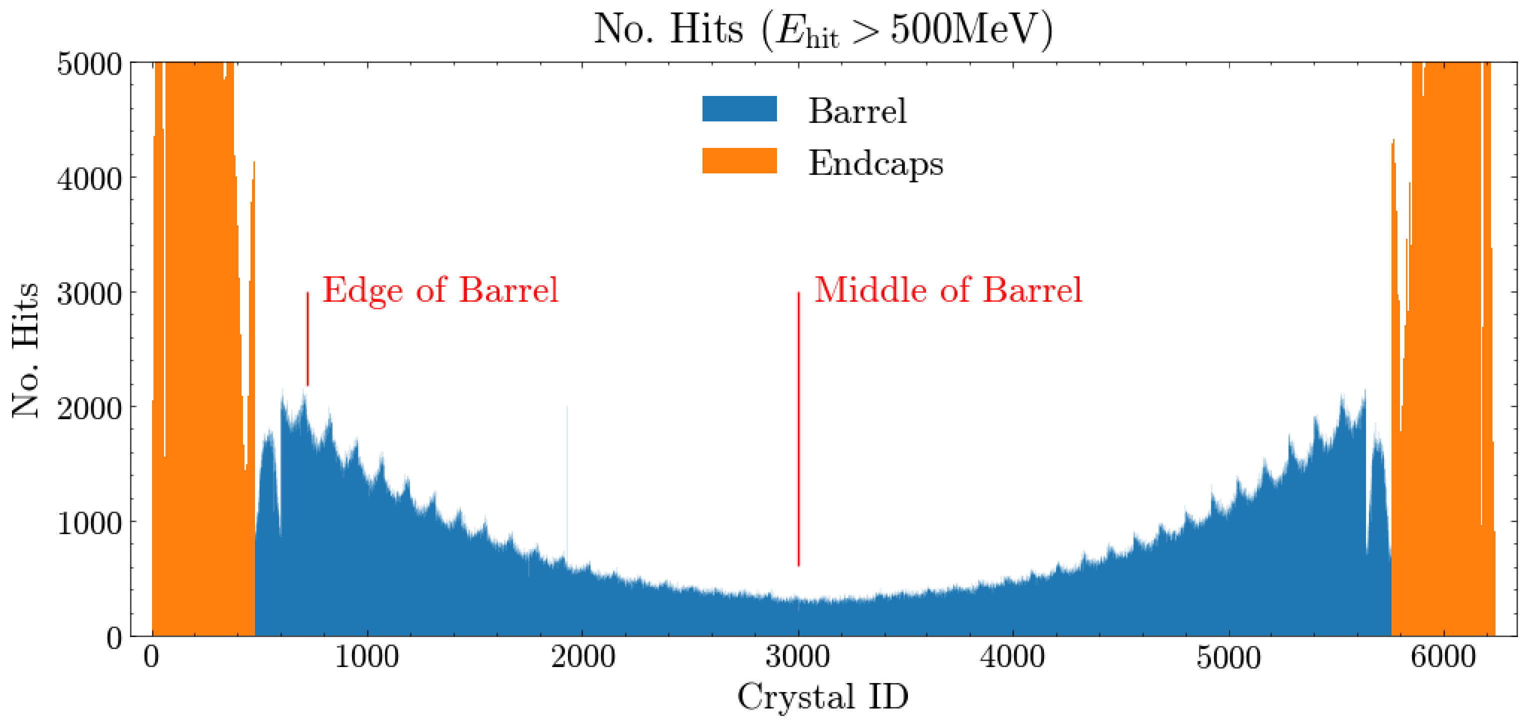
Advantages of Machine Learning Method

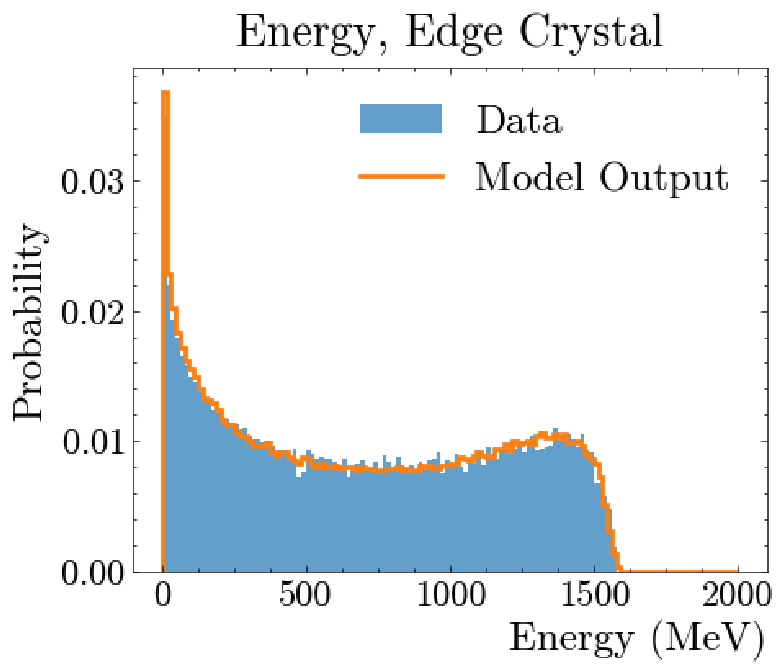
For machine learning method, models trained with same parameters can well adapt outer condition changing.

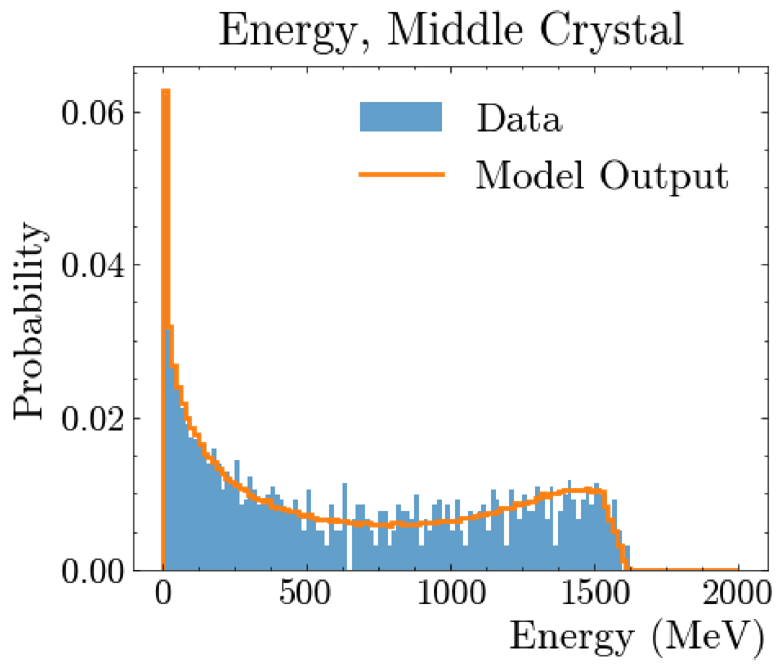


Advantages of Machine Learning Method

Machine learning method can also well handle position-depend difference:

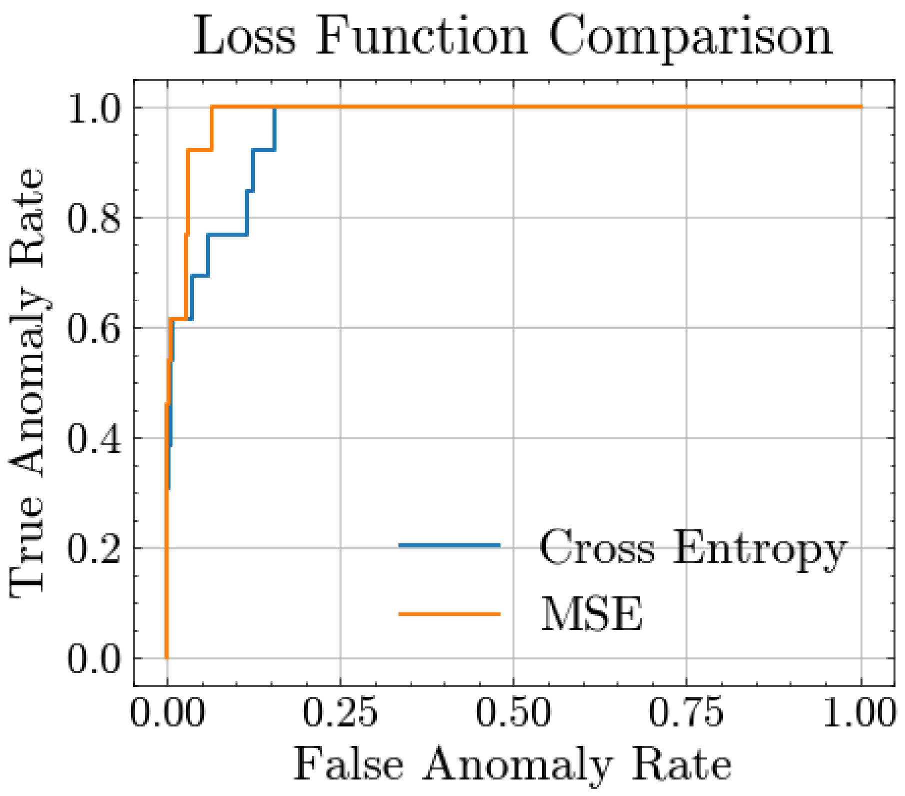


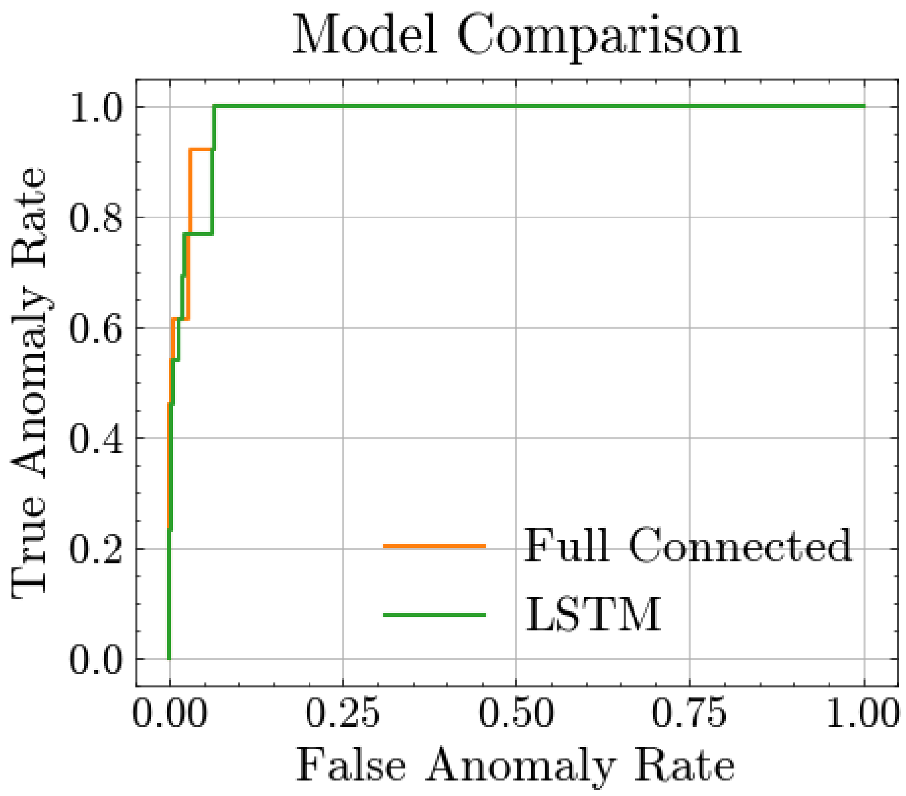




Anomaly Detection in Practice on EMC

① Select Mean Squared Error (MSE) as loss

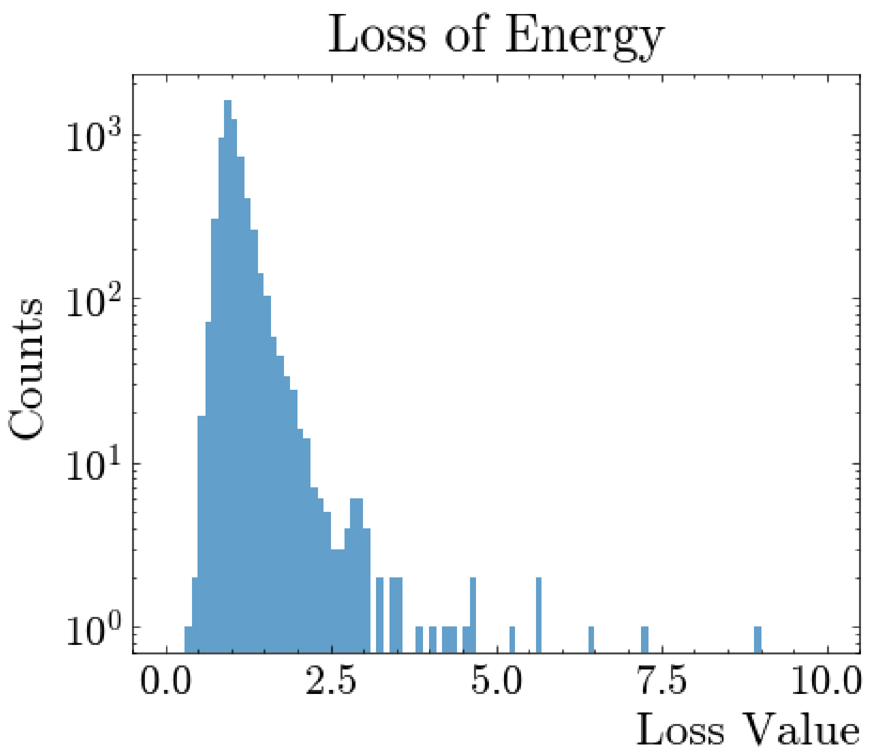


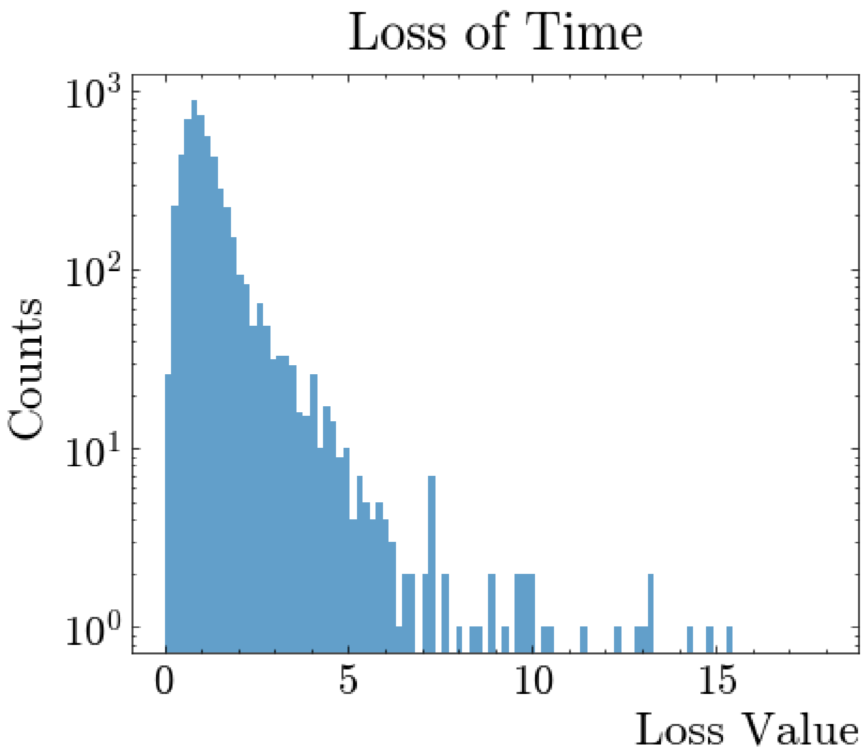


Anomaly Detection in Practice on EMC

② Train model, obtain predictions, calculate losses for all crystals;

(Separately apply on energy and time histograms)

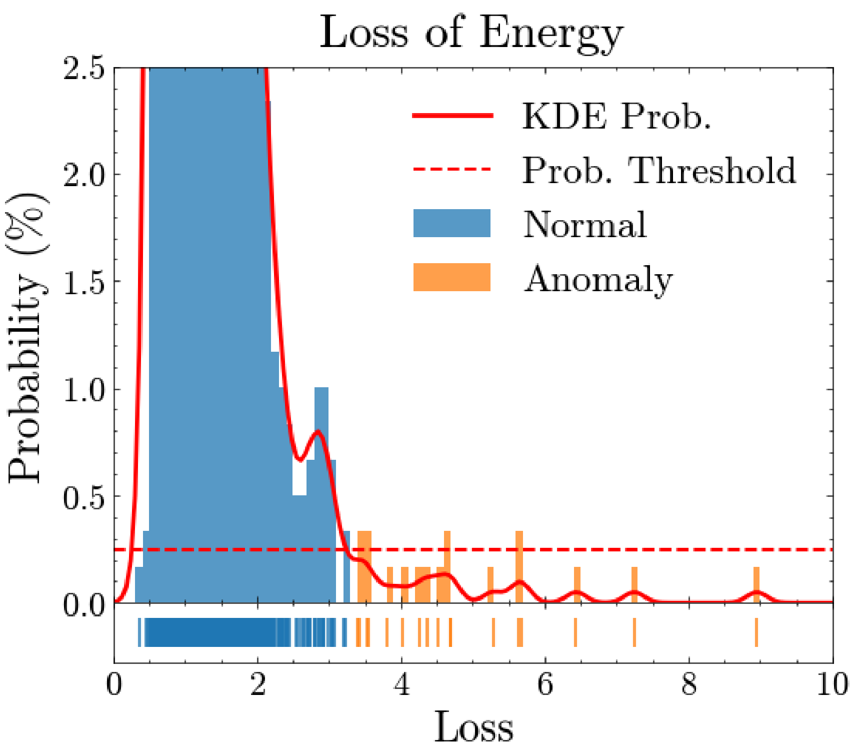


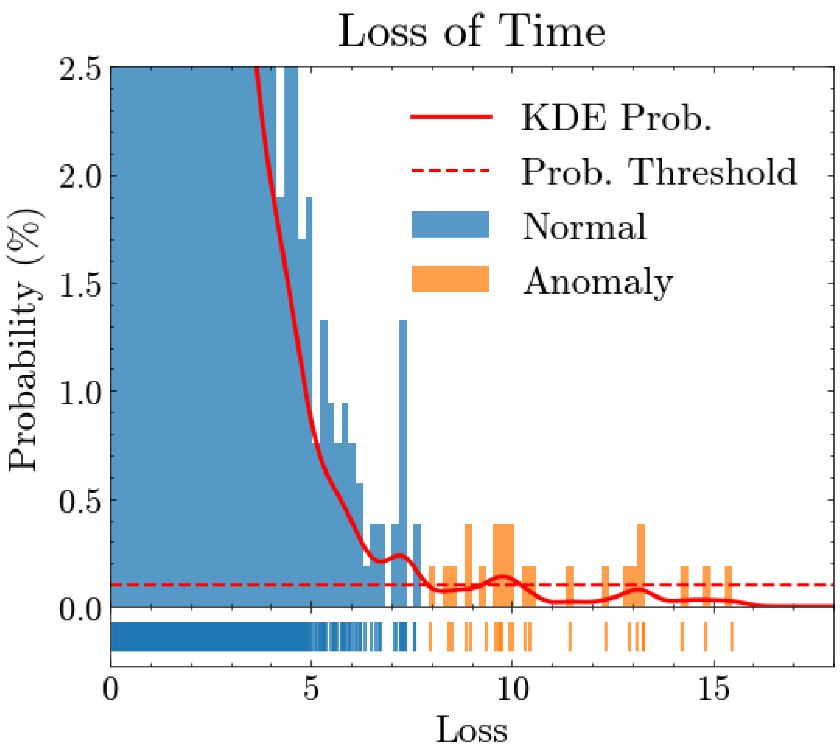


Anomaly Detection in Practice on EMC

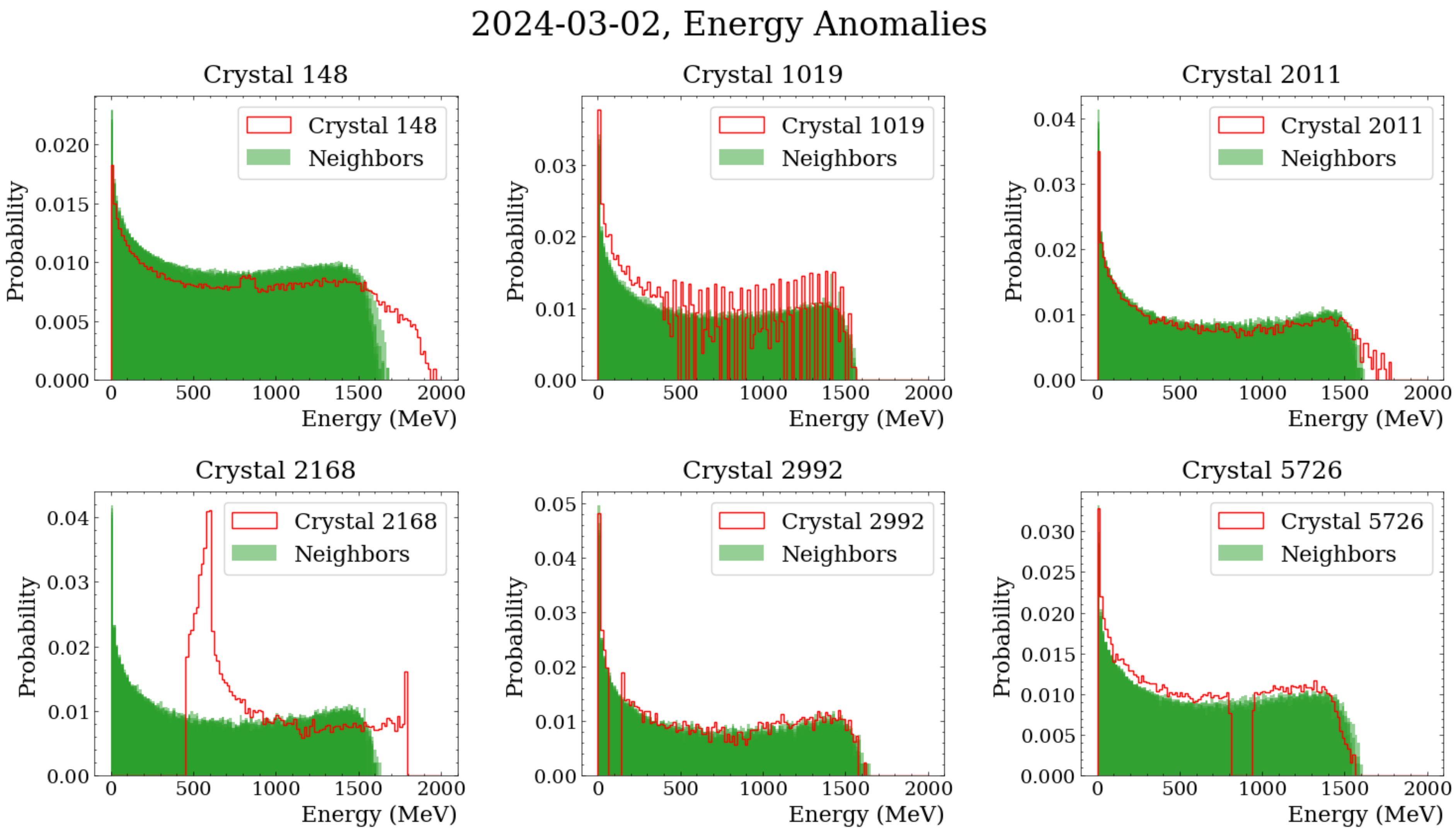
③ Estimate loss probability density by Kernel-Density Estimation (KDE);

④ Anomalies: (1) Energy: $\text{Prob.} < 0.25\%$ (2) Time: $\text{Prob.} < 0.1\%$



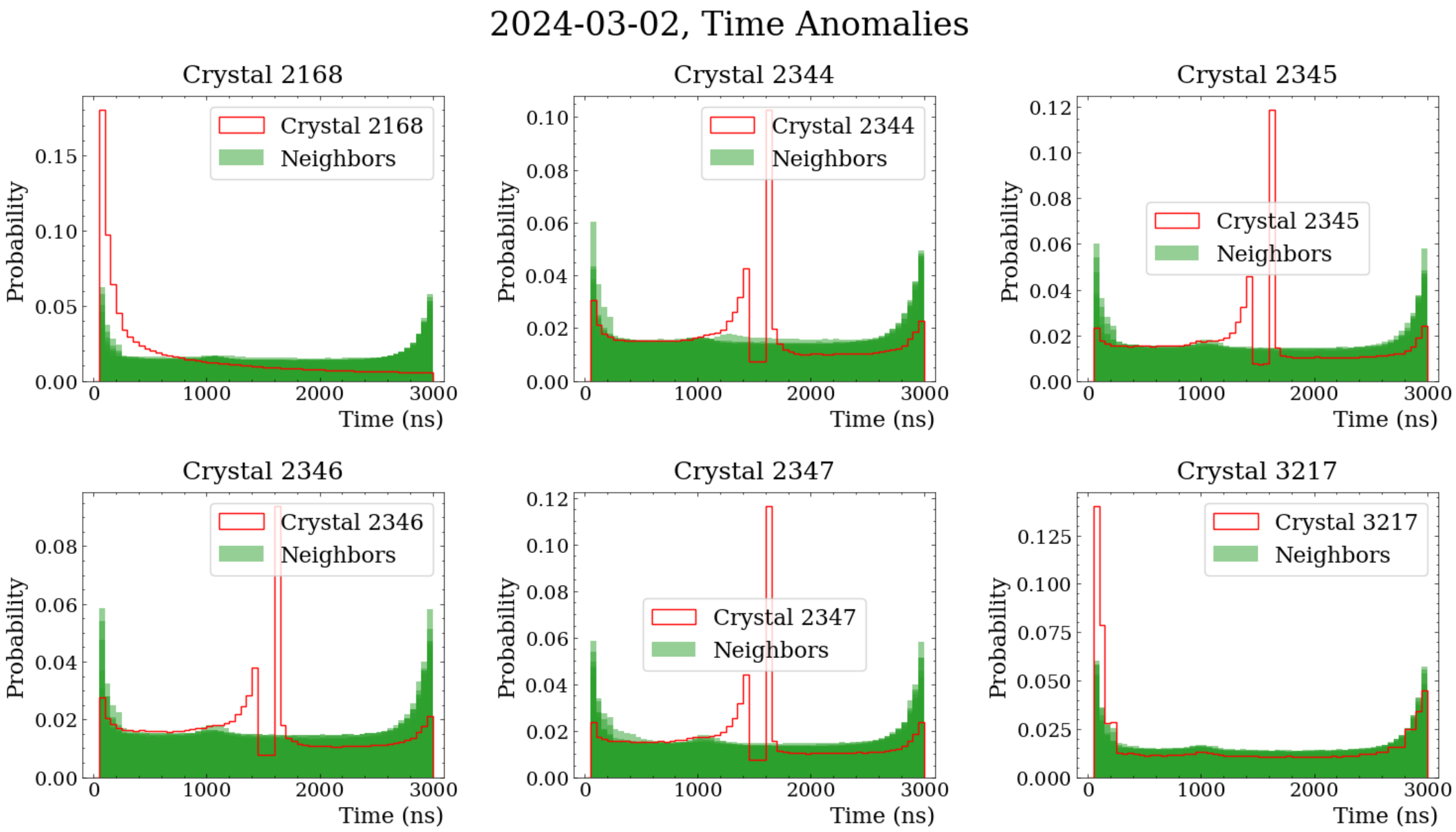


Anomaly Detection in Practice on EMC



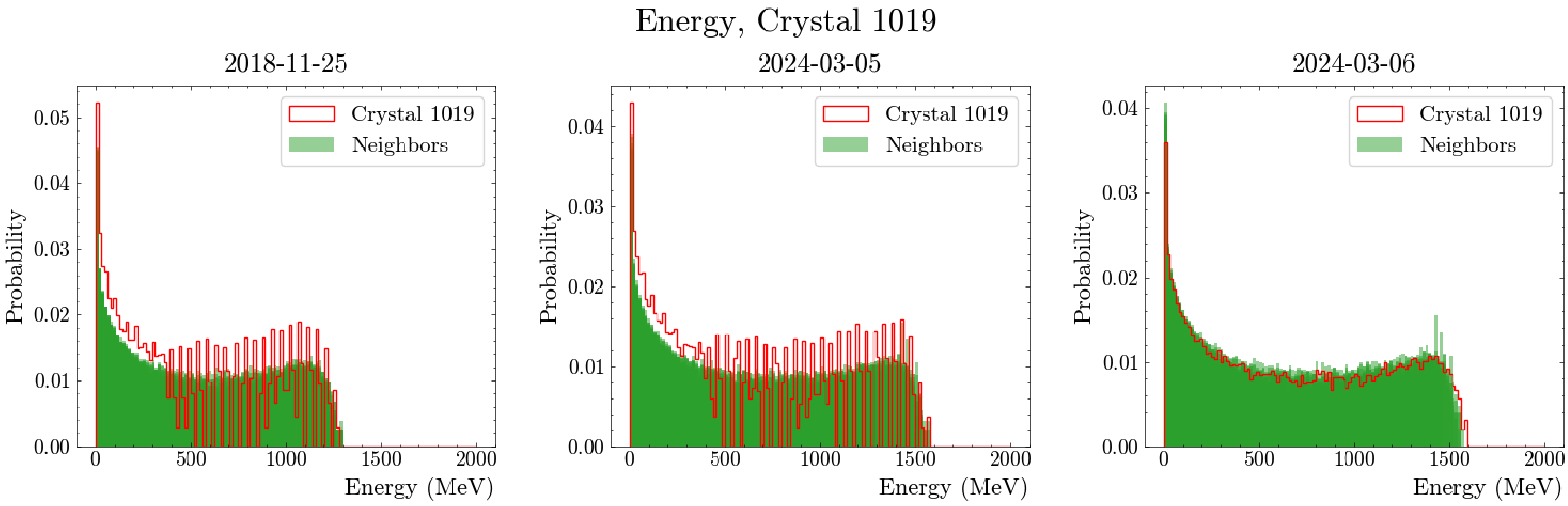
"Neighbors": Normal crystals around the anomaly.

Anomaly Detection in Practice on EMC



“Neighbors”: Normal crystals around the anomaly.

Anomaly Detection in Practice on EMC



Crystal 1019 has been abnormal for years, and got repaired after we found it.

Summary

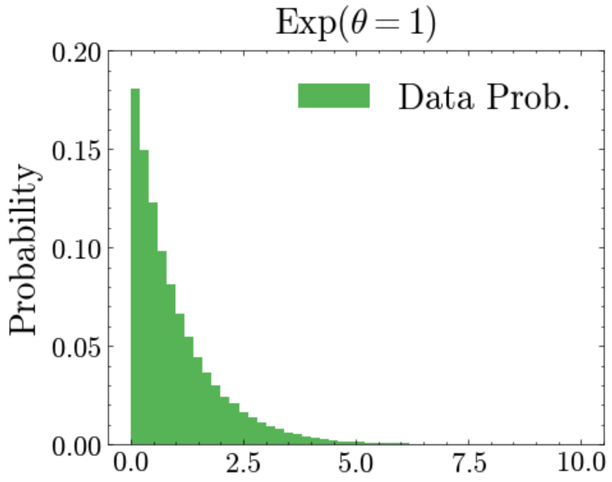
- An effective Machine-Learning based EMC monitoring method is developed;
- Some hidden anomalies on EMC crystals were found by the new method;

Next Step

- Try this method on other subdetectors;
- Develop more checking algorithms;
- Deploy this method for daily check;

Back Up

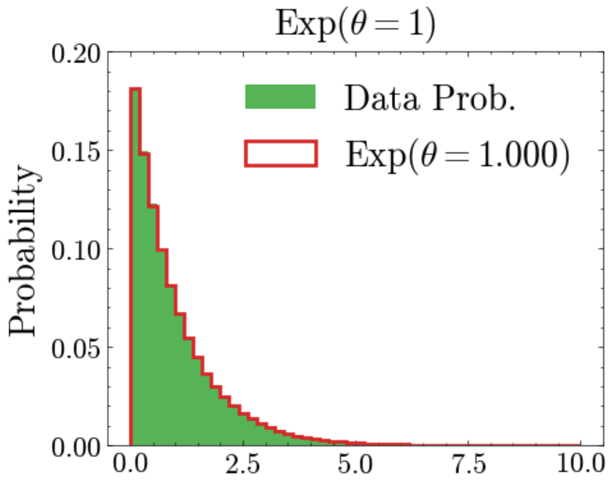
3. Detect Anomaly by Autoencoder



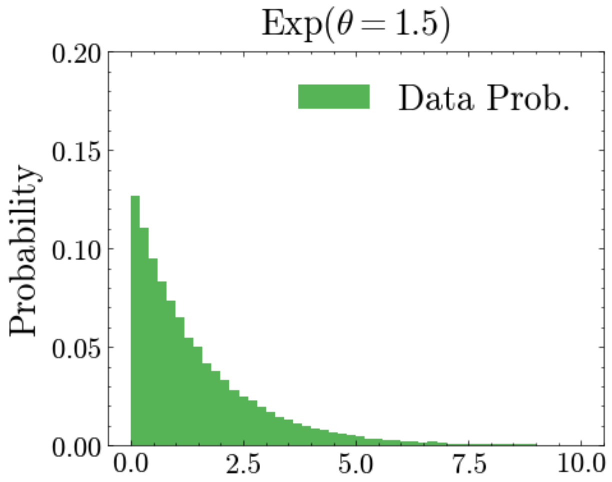
Encode

$\mu = 1.000$

Decode



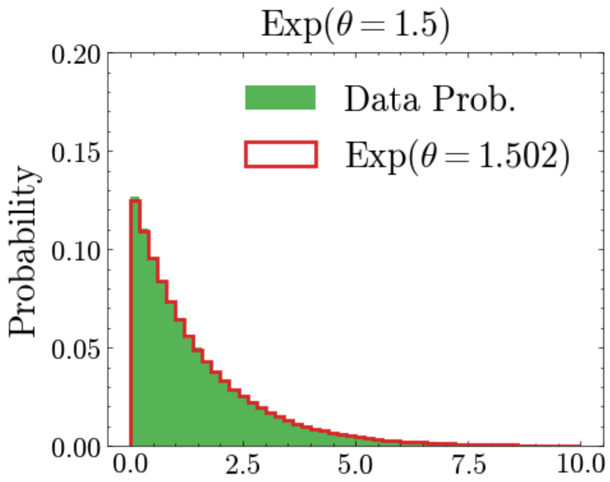
Low Loss(in, out)



Encode

$\mu = 1.502$

Decode



Low Loss(in, out)

If histogram always follows exponential distribution, we can use mean value to “encode” and “decode” its contents without suffering high loss.