Low-Latency Online Particle-Hypothesis Studies of Combined Detectors with ML

2nd GSI/FAIR AI Workshop

A. Belias (GSI),

D. Veretennikov (DESY), G. Schnell (Basque U., Bilbao, IKERBASQUE)

Low-Latency Online Particle-Hypothesis Studies of Combined Detectors with ML

2nd GSI/FAIR AI Workshop

A. Belias (GSI),

D. Veretennikov (DESY), G. Schnell (Basque U., Bilbao, IKERBASQUE)

SHARP

Introduction

First studies

Future work

cost EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

Funded by the European Union

Intro – Towards “Edge Computing”

High beam intensities and interaction rates at FAIR, at CERN (HL-LHC, FCC), EIC, will provide unique experimental conditions, yet pose formidable challenges to detectors, readout electronics/DAQ and physics event selections:

→ **Large number of simultaneous collisions**

→ **Large increase in continuous data rates**

→ **Radiation harsh regions**

Faced with such great challenges, the future experiments are increasingly considering to use computational methods on FPGAs, close to the detectors.

Such ”**edge computing**” strategies combined with ML based models allow fast, low-latency intelligent filtering close to the detector front-end readout.

Focus in SHARP

Many studies of ML methods on FPGAs examine specific detectors
e.g. (a) the EMC “PicoCal” at LHCb, (b) SiPM-based RICH in CBM/FAIR.

Our studies focus on ML based models in FPGAs combining data from different detector types for on-line physics/background discrimination.

We investigate two aspects:

(a) examine ML models across detectors (off-line)

(b) potential to implement ML models on FPGAs (on-line)

SHARP - Structure and Spectroscopy of Hadrons Research Project

Work Groups include: Novel Methods in Data Analysis, Detectors and DAQ

EU funded COST Action, 2025-2029, (https://www.cost.eu/actions/CA24159/)

Particle hypothesis study with straw tubes and ToF using multilayer perceptron (MLP)

D. Veretennikov (DESY),
A. Belias (GSI), G. Schnell (Basque U., Bilbao, IKERBASQUE)

A configurable toy Monte Carlo defines the first use case

ROOT-based event generation with adjustable detector geometry and response.

EVENT MODEL

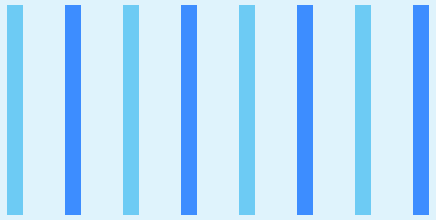
1–5 particles/event, μ± π± K± p / $\bar{p}$, p = 2–15 GeV/c

Straight tracks, No energy loss

Interaction point

→

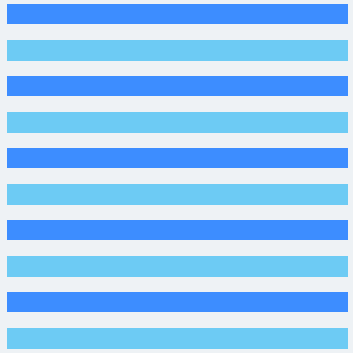
STRAW STATION



z = 5 m • 2 × 2 m
5 mm straw diameter

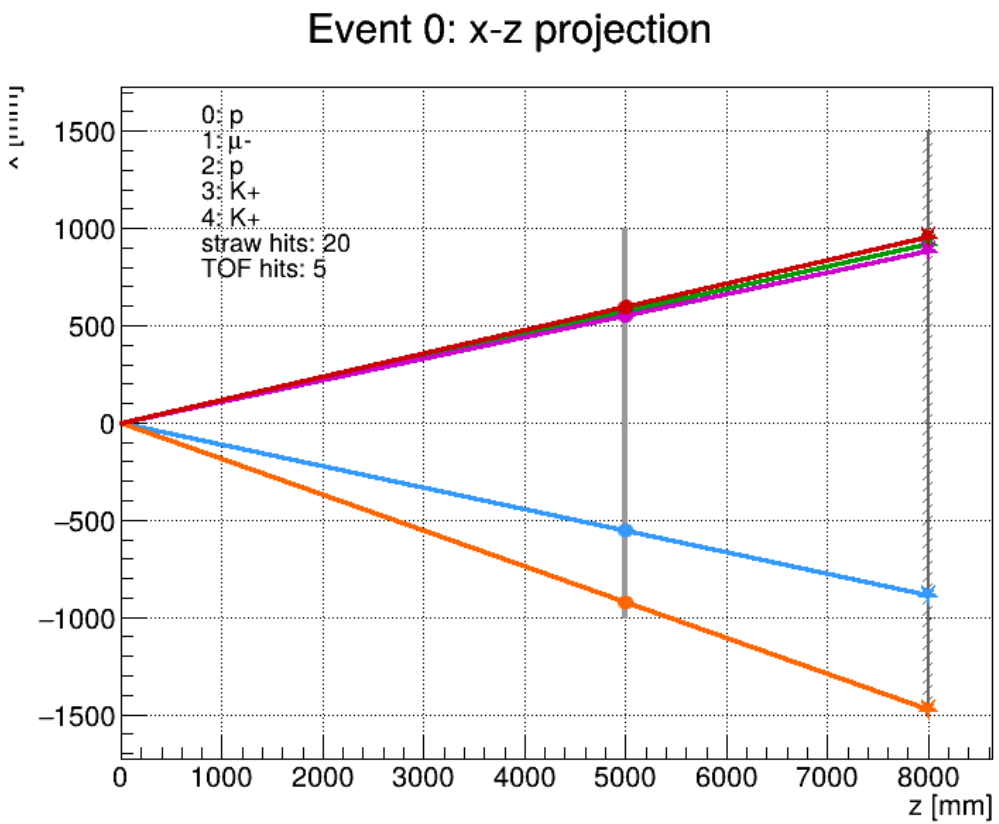
→

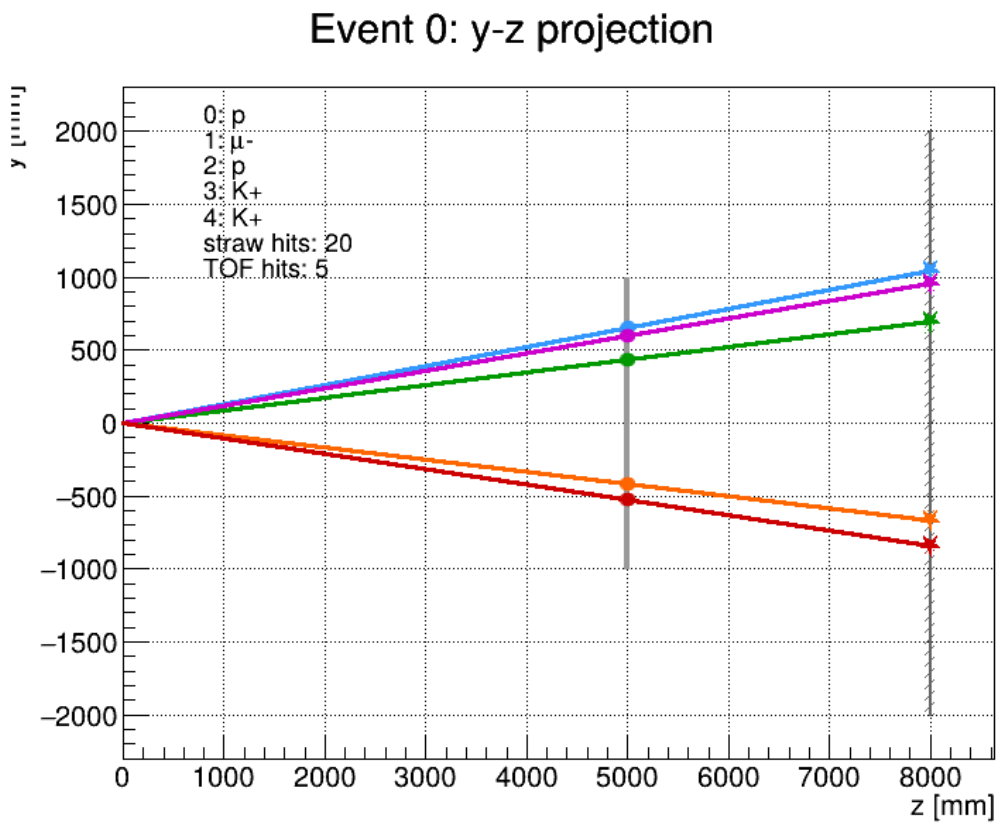
TOF WALL



z = 8 m • 3 × 4 m
3 m × 10 cm bars

Output:
truth particles, straw hits and TOF hits in linked ROOT trees


Event 0: x-z projection
0: p
1: μ-
2: p
3: K+
4: K+
straw hits: 20
TOF hits: 5
x [mm]
z [mm]
1500
1000
500
0
−500
−1000
−1500
0
1000
2000
3000
4000
5000
6000
7000
8000


Event 0: y-z projection
0: p
1: μ-
2: p
3: K+
4: K+
straw hits: 20
TOF hits: 5
y [mm]
z [mm]
2000
1500
1000
500
0
−500
−1000
−1500
−2000
0
1000
2000
3000
4000
5000
6000
7000
8000

Classical χ^2 provides a transparent physics baseline

Each associated straw–TOF candidate is tested against the muon hypothesis.

$$\chi^2_\mu = \chi^2_{staw} + \chi^2_{ToF}$$

TRACK CONSISTENCY

$$\chi^2_{staw} = \sum_i \left(\frac{r_{i,meas} - r_{i,track}}{\sigma_r} \right)^2$$

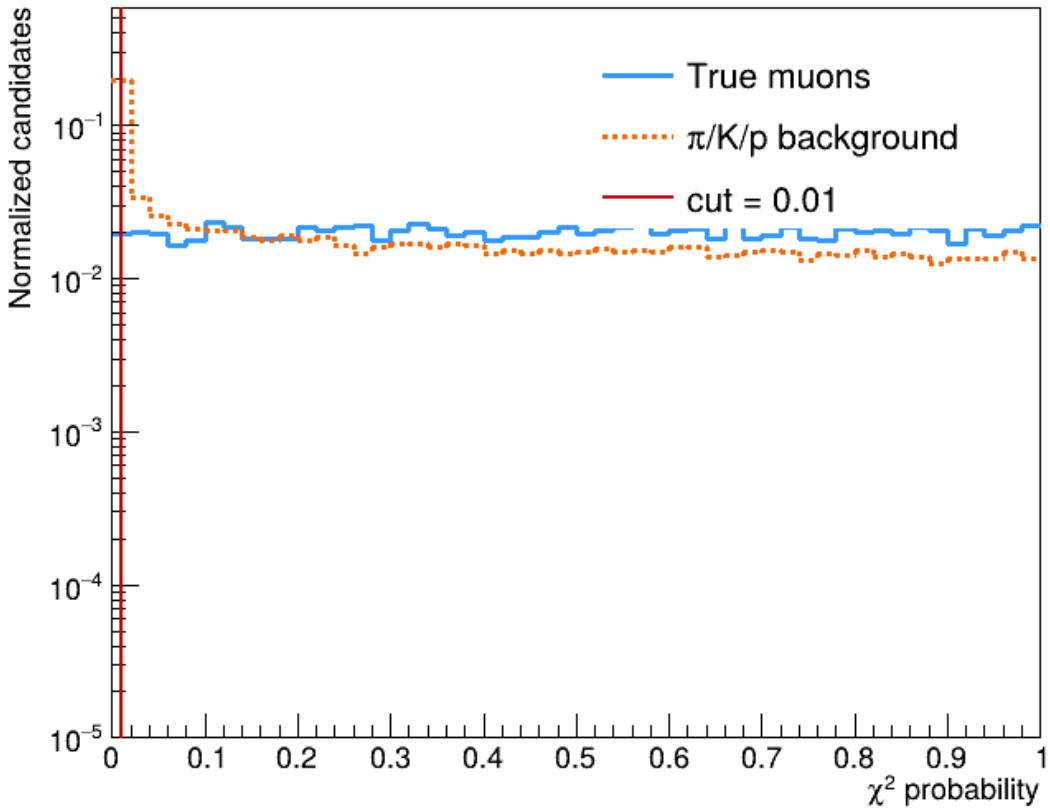
- straight-line fit through the known vertex
- drift time converted to drift radius
- residuals evaluated layer by layer

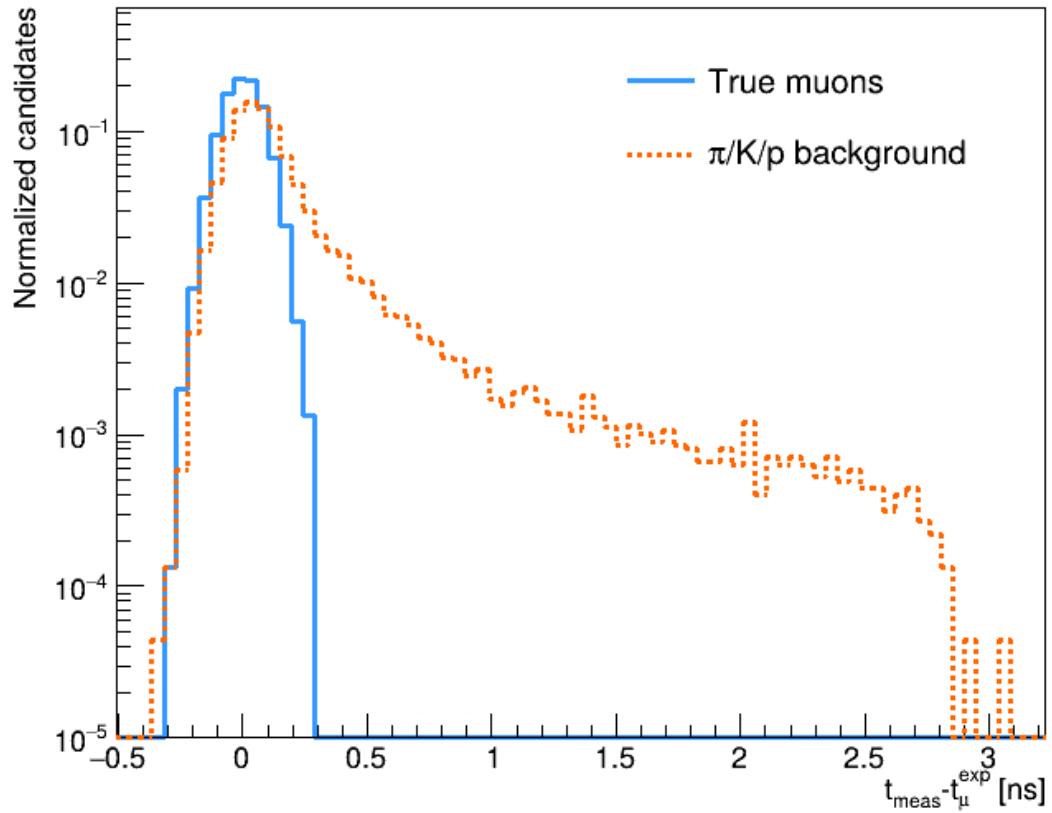
MUON TIME-OF-FLIGHT

$$\chi^2_{ToF} = \left(\frac{\Delta_{t,meas} - L/\beta\mu c}{\sigma_{ToF}} \right)^2$$

βμ is calculated from the measured momentum and muon mass. The χ^2 probability provides an adjustable compatibility threshold.

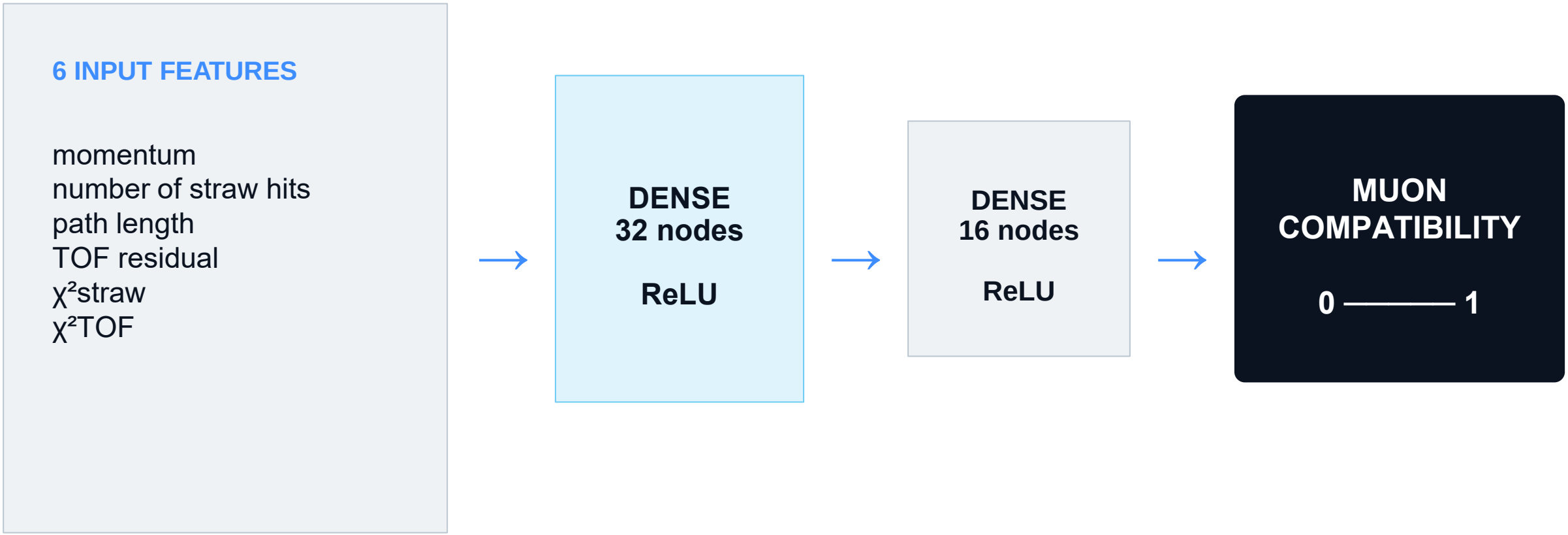
Physics baseline: ideal hit association and generator momentum used as an external measurement


Normalized candidates
True muons
π/K/p background
cut = 0.01
χ² probability


Normalized candidates
True muons
π/K/p background
t_meas-t_μ^exp [ns]

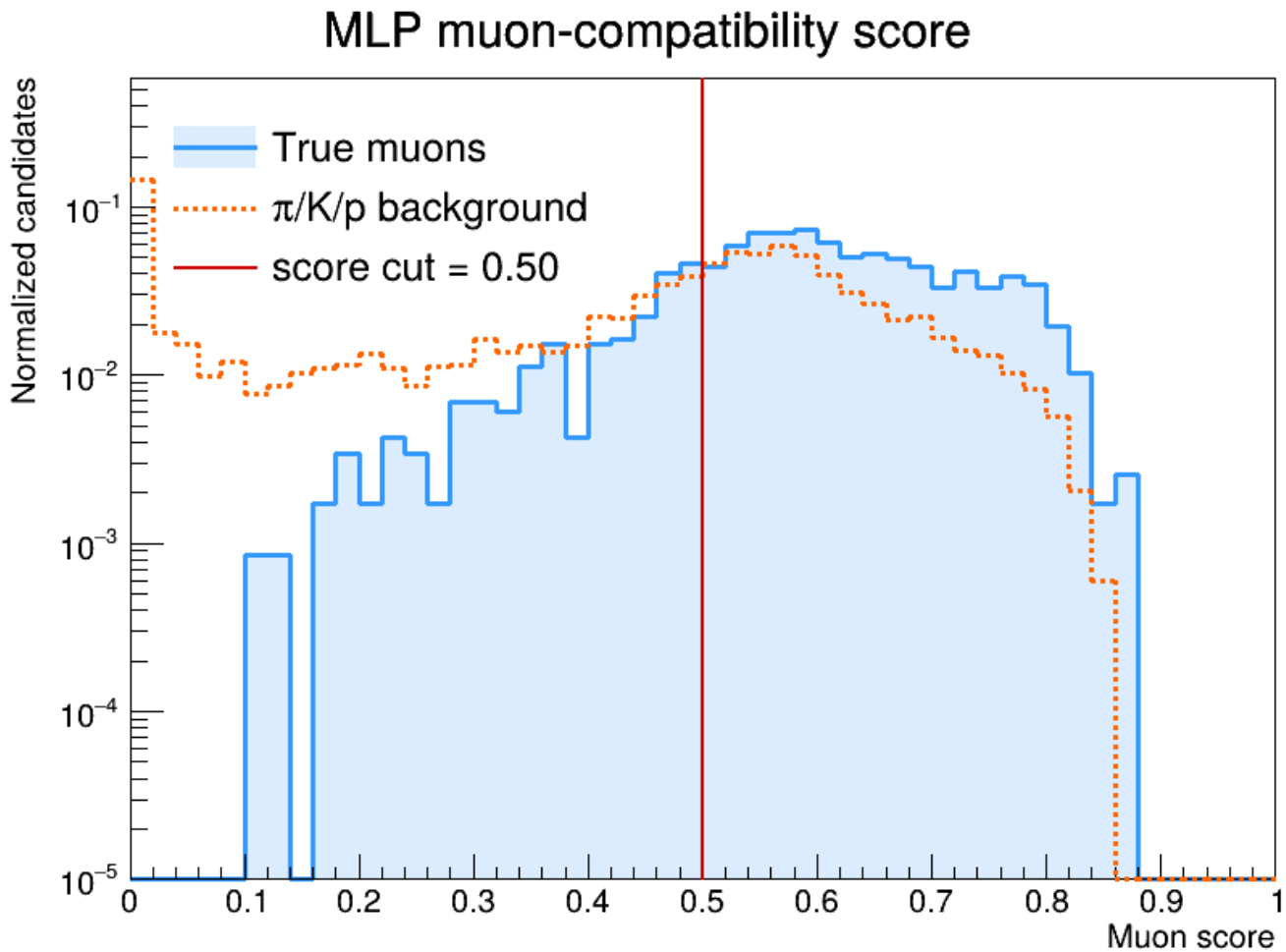
A compact MLP learns a nonlinear compatibility score

PyTorch model trained only on reconstructed quantities; truth is used solely as the label.



BCEWithLogitsLoss • class weighting • event-wise train/test split • early stopping

Export: TorchScript model with normalization embedded for later inference



MLP muon score distributions for true muons and background particles, red line showing the chosen classification cut at 0.5

The two approaches serve complementary roles

Quantitative performance must be established on independent samples and reported versus momentum.

CLASSICAL χ^2

+ physically interpretable
+ no training required
+ stable reference baseline

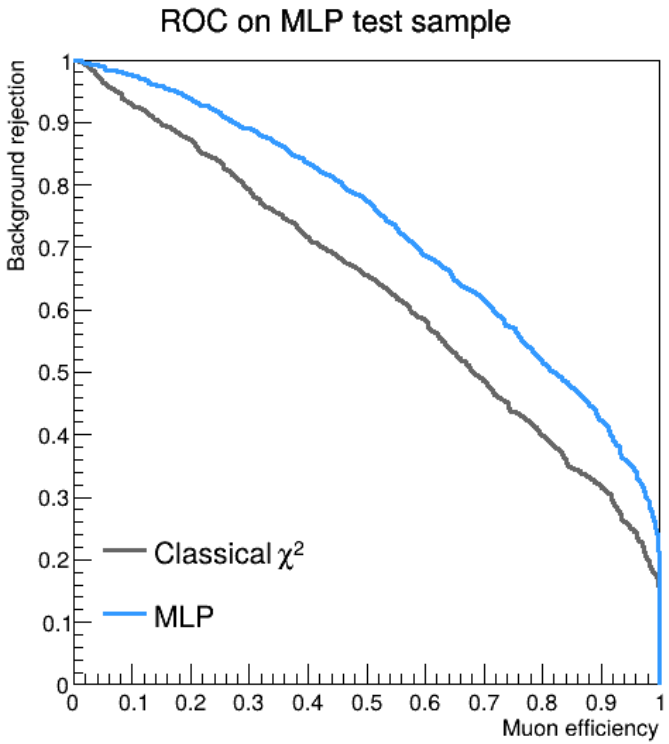
− limited nonlinear correlations
− assumes reliable reconstruction

PYTORCH MLP

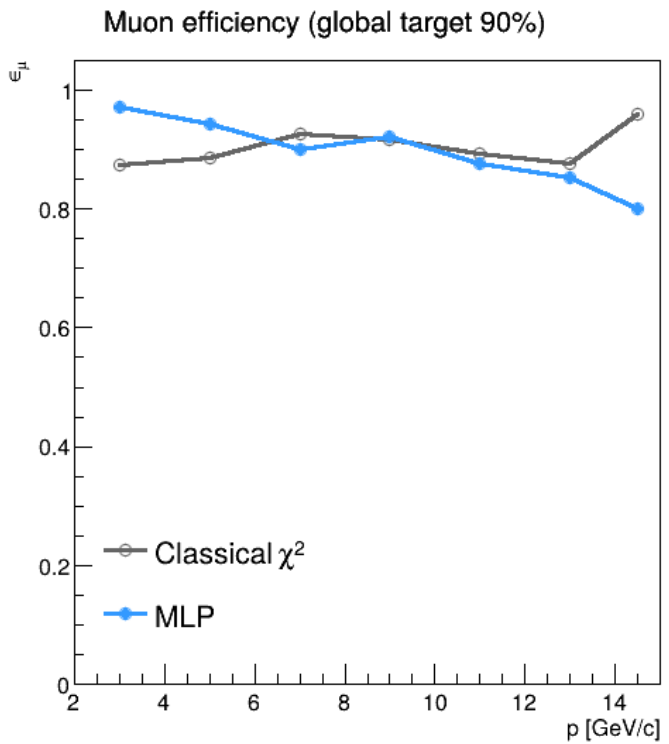
+ combines correlated observables
+ compact and quantization-ready
+ adaptable to realistic backgrounds

− depends on simulation quality
− requires validation and calibration

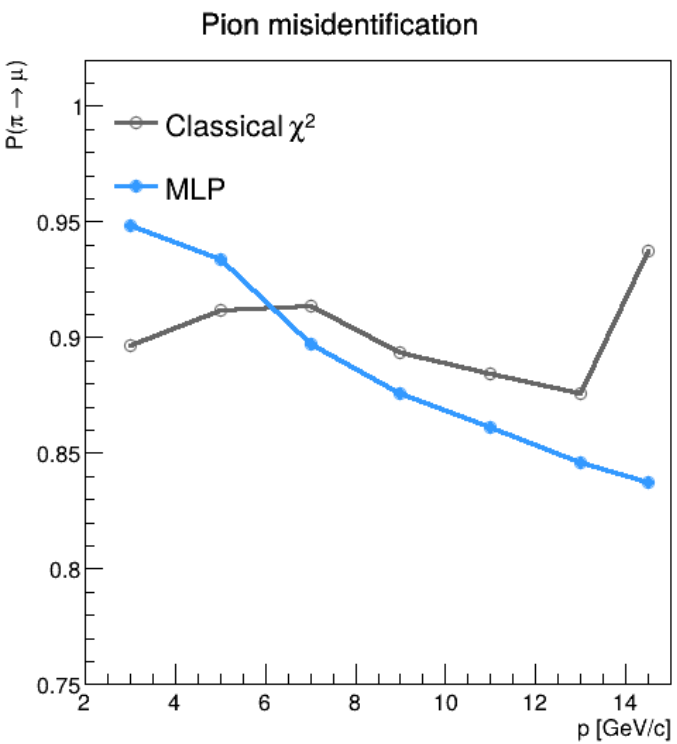
Conclusion: use χ^2 as the physics benchmark, then quantify whether the MLP improves background rejection at fixed muon efficiency—especially as a function of momentum.



MLP has better background rejection than classical χ^2 at the same muon efficiency



Muon efficiency versus momentum for the selected global 90% target



Pion misidentification versus momentum, showing how often pions pass the muon selection

Future Work

In our studies we focus on ML based models in FPGAs combining data from different detector types for on-line physics/background selection.

We investigate two aspects:

(a) examine ML models across detectors (off-line) → Just started

Next steps: realistic MC • add noise and ambiguities • extend use cases

(b) potential to implement ML models on FPGAs (on-line) → Due to start

First steps: quantization • implementation (hls4ml, i.e. Zynq SoC)

Interested??

Join SHARP - Structure and Spectroscopy of Hadrons Research Project

EU funded COST Action (CA24159), (https://www.cost.eu/actions/CA24159/)