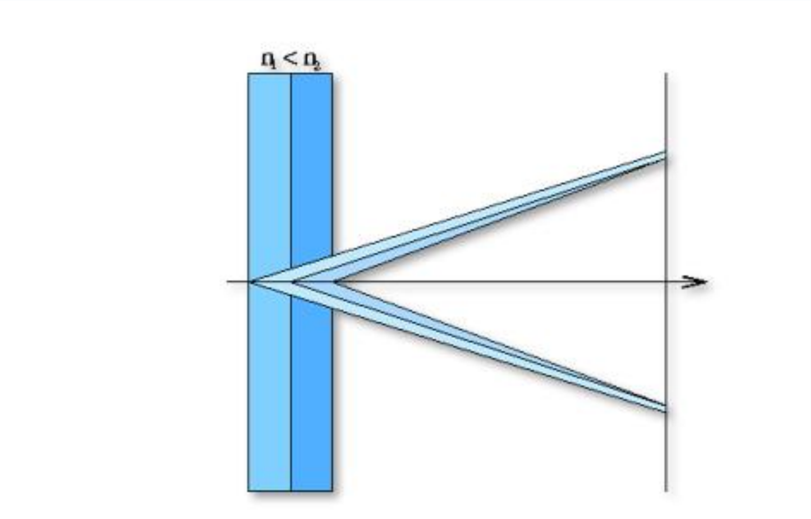
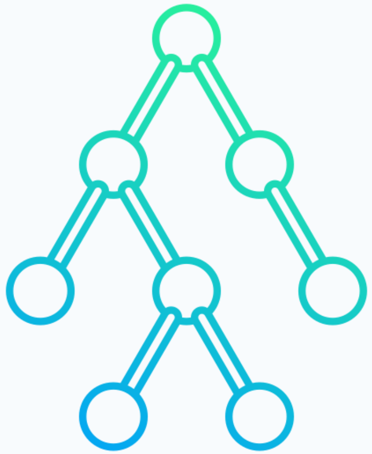
24

Machine-Learning-Driven Particle Identification with the FARICH Detector under realistic operating conditions





P. Rogozhin, M. Chadeeva, T. Uglov

HSE University · P. N. Lebedev Physical Institute

The 10th International Conference in Deep Learning in Computational Physics

FARICH turns particle velocity into a photon ring

This is the physical observable behind the ML features.

A charged particle crosses a multilayer aerogel radiator.

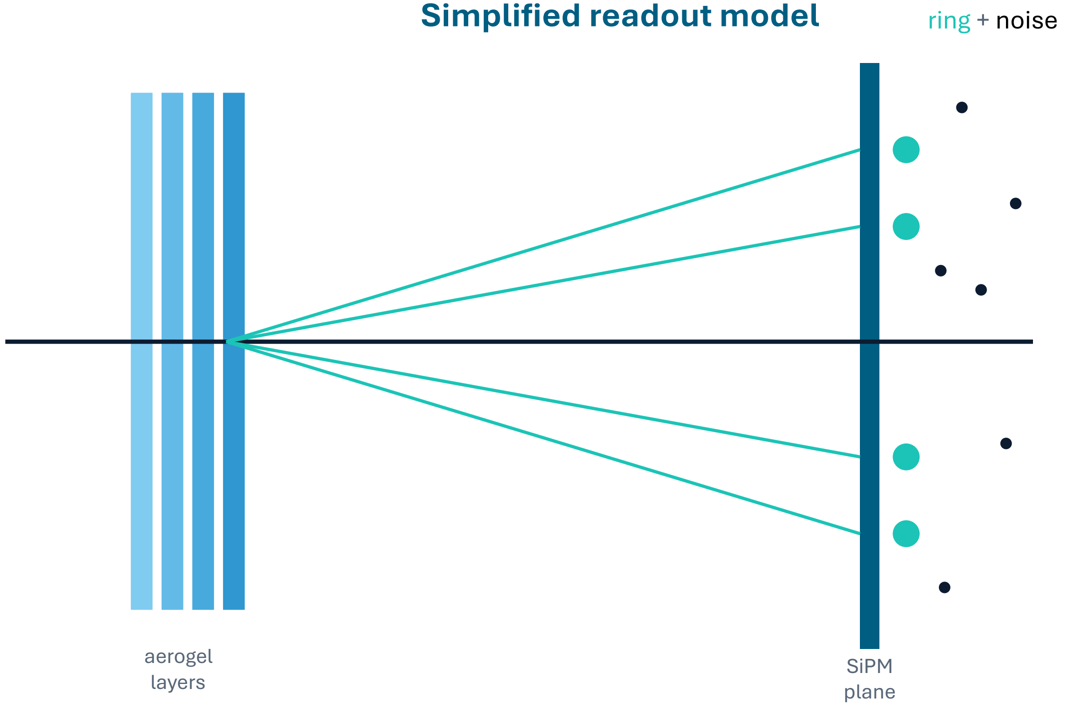
Cherenkov angle follows $\cos\theta_c = \frac{1}{n\beta}$

Layers with different refractive indices focus photons onto a compact ring.

The SiPM plane records sparse signal hits plus dark-count noise.

With track momentum and incident angle, the reconstructed ring radius becomes species-sensitive.

Simplified readout model



R_{reco} ring radius	p, θ_{in} tracker input	Z_{DCR} noise proxy

ML role: learn species boundaries from physically meaningful, low-dimensional observables rather than from detector pictures.

A detector-PID problem cast as a supervised ML task

Under realistic geometry, timing response, and dark-count noise, simple ring-based cuts become fragile; physics-informed ML provides tunable PID scores.

Input
SiPM hits + track geometry
→
Feature engineering:
ring radius, hit density, DCR proxy
→
BDT classifier:
P(μ), P(π), P(K)
→
Analysis operating point

≤0.5%
target residual background from π/μ mis-ID

1 MHz/mm^2
stress-test dark count rate

μ/π/K
three-class supervised PID

100×
background suppression at practical signal efficiency

Talk logic: from detector response to analysis impact

01 FARICH principle and physics objective

02 Simulation and digitisation as labelled data factory

03 Reconstruction as physics-informed feature engineering

04 BDT model, baselines, robustness checks

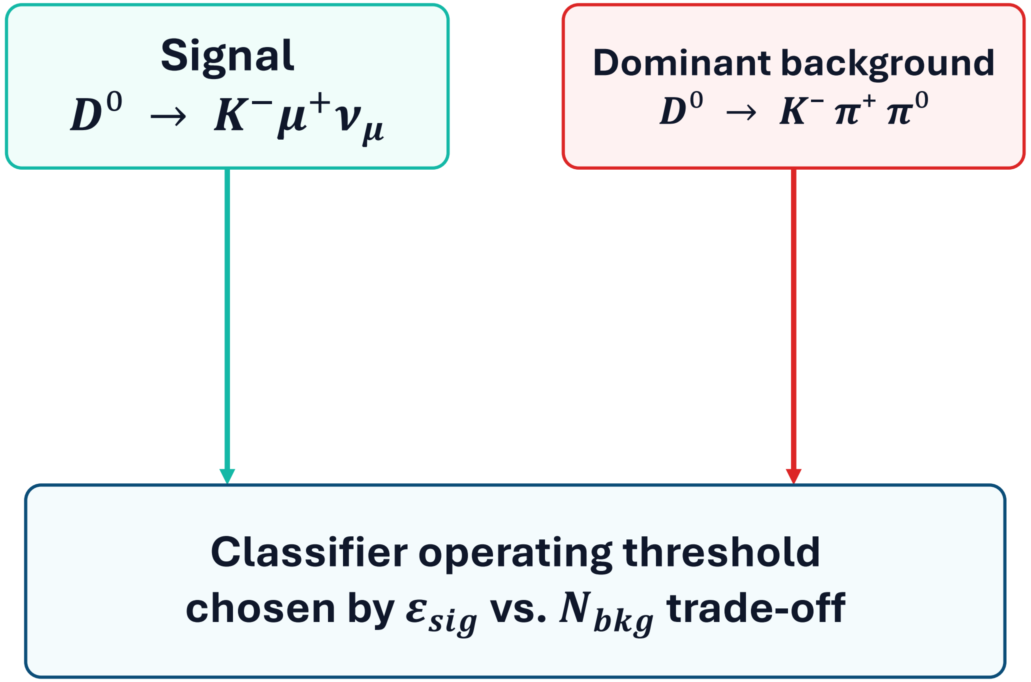
05 Transfer validation in $D^0 \rightarrow K^-\mu^+\nu_\mu$

Detector principle first; then ML pipeline, transfer-validation strategy, and robustness.

FARICH PID controls the dominant charm-analysis background

A key validation channel is semileptonic charm decay where π/μ separation directly controls background leakage.

- Process: $e^+e^- \to \psi(3770) \to D^0\overline{D^0}$ @ $\sqrt{s} = 3.77$ GeV ($\overline{D^0}$ forced to invis.decay)
  - two D^0 decay channels:
    - signal: $D^0 \to K^-\mu^+\nu_\mu$
    - dominating background: $D^0 \to K^-\pi^+\pi^0$
  - number of simulated events: 10^6
  - branching ratio: $\mathfrak{B}_{sig}/\mathfrak{B}_{bgd} = 0.034/0.144$
- Detector model in the Aurora framework
  - tracker+FARICH of SCTF
  - particle identification with FARICH only
- Event selection accounting for:
  - π^0 reconstruction efficiency
  - detector acceptance and reconstr. eff.
- After cuts and before PID: $\mathrm{Nsig_{sel}}$, $\mathrm{Nbgd_{sel}}$



Validation question: can a classifier trained on controlled single-particle samples remain reliable in realistic event kinematics?

Detector simulation turns response into labelled ML data

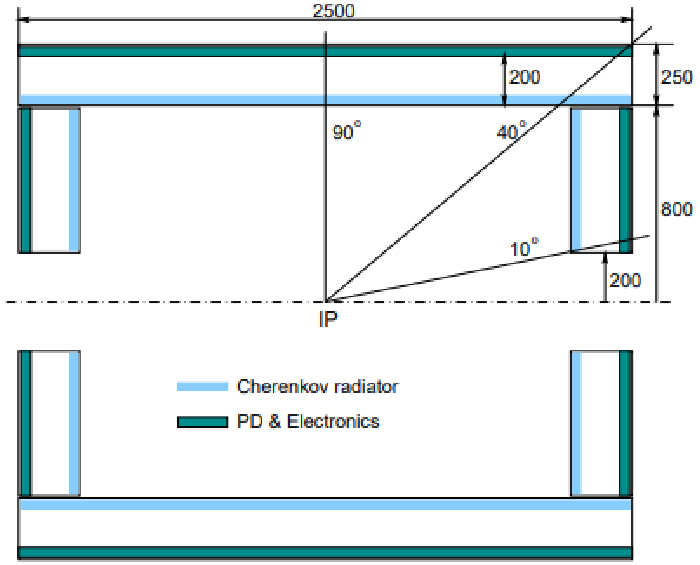
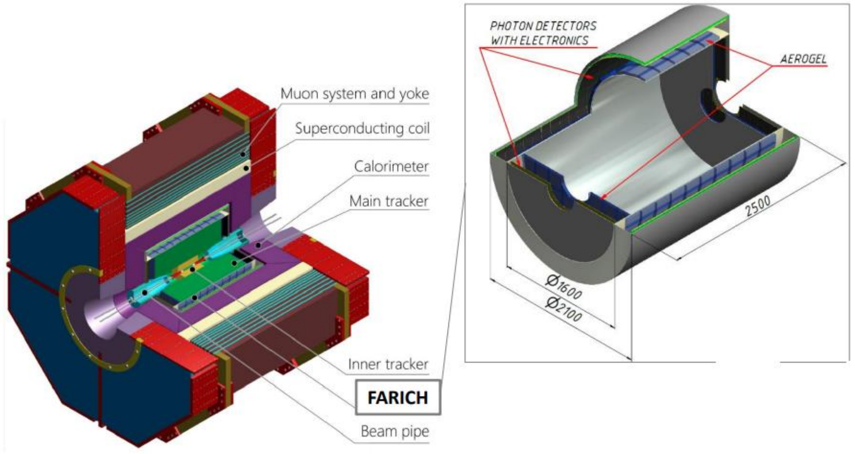
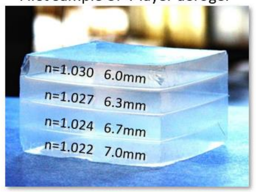


Figure 2.13. FARICH PID system for the SCTF.

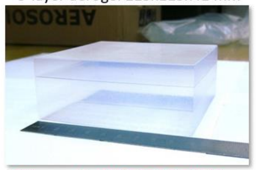


- Aurora + Geant4 optical-photon transport for FARICH
- Barrel-only acceptance: approximately 40°–130° w.r.t. beam
- Four aerogel layers, n = 1.04–1.05, 200 mm gap to photodetector plane
- SiPM digitisation includes PDE, geometrical efficiency, dark-count noise

First sample of 4-layer aerogel



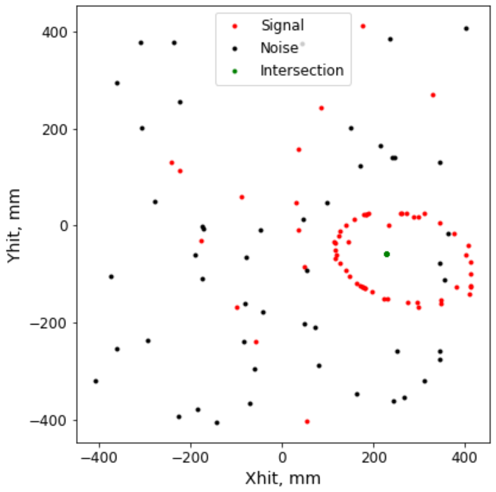
3-layer aerogel 115x115x41 mm^3



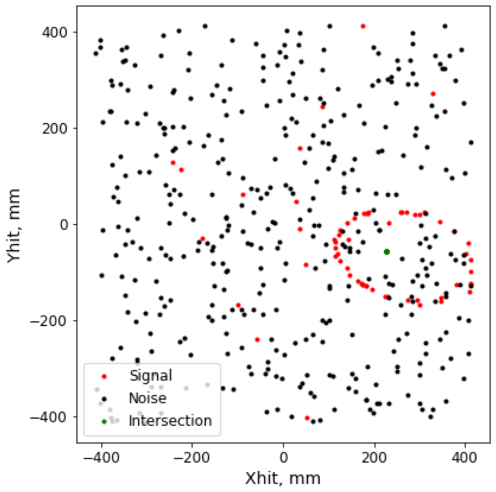
T.Iijima et al., NIM A548 (2005) 383
A.Yu.Barnyakov et al., NIM A553 (2005) 70

FARICH radiator and SiPM-photodetector option

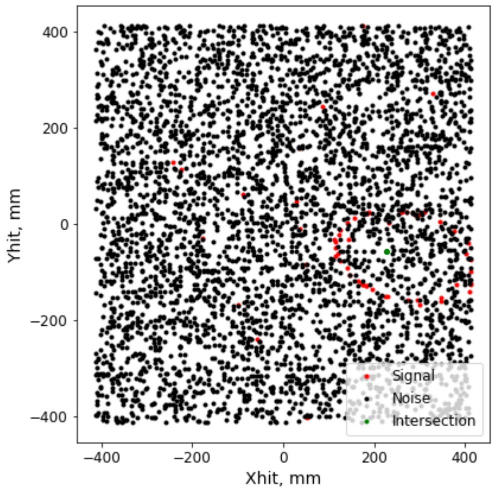
Realistic dark-count noise defines the robustness target



Signal hits + low-noise scenario



No cooling: 100 kHz/mm^2



Damaged sensor stress test: 1 MHz/mm^2

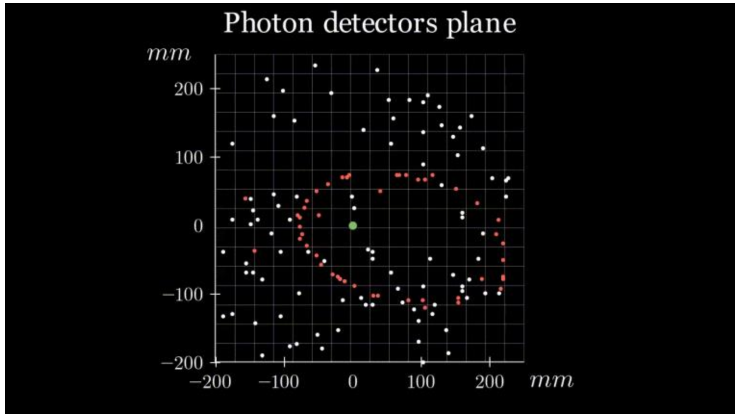
The ML task is intentionally trained and tested against conditions where naïve ring interpretation becomes fragile.

10 kHz/mm^2
cooled sensor reference

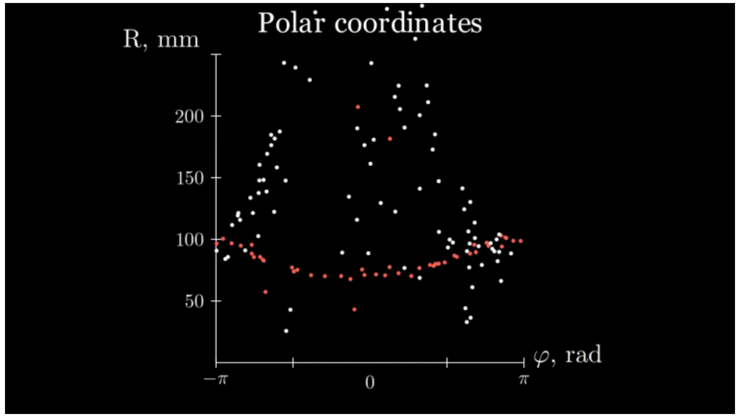
100 kHz/mm^2
realistic no-cooling scenario

1 MHz/mm^2
conservative radiation-damage scenario

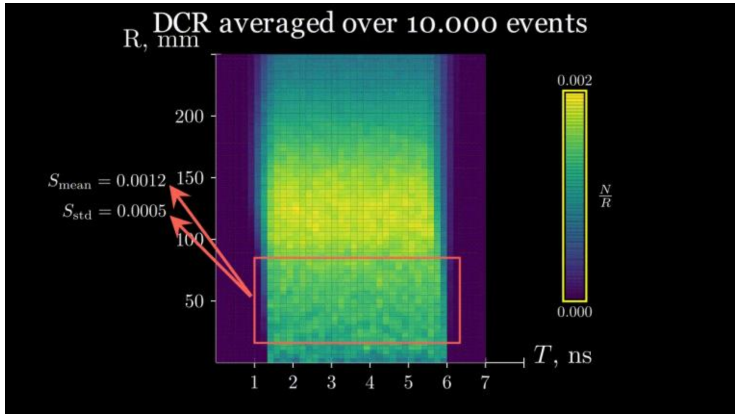
Ring reconstruction becomes ML feature engineering



Hit plane



Polar / RT coordinates



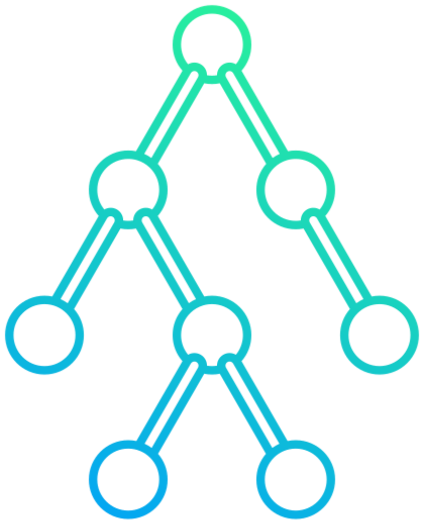
DCR-corrected observable

Track impact point and hit timing anchor the transformation → DCR density map gives a noise-aware local baseline → Output features feed supervised classifier

Main message: the classifier is not trained on raw hits; it uses physically motivated observables designed to be stable under noise.

ML feature set: compact, physical, and deployment-oriented

Feature	Physical meaning	ML role
p_in	primary-particle momentum	from MC truth in this study; tracker value in deployment
θ_in	incident polar angle	accounts for geometry and path length
R_reco	reconstructed Cherenkov-ring radius	main species-separation observable
Z_DCR	noise proxy / Z-score	event-level DCR measure for robustness



XGBoost BDT

version 2.1.2, three classes: μ, π, K

Single-particle labels

supervised training on known species

Caveat: MC-truth p and θ are used here to isolate FARICH response; deployment uses reconstructed tracker quantities.

Why a BDT is a good first ML model here

The goal is a robust physics-informed PID score, not an end-to-end image model.

Tabular features

p_in, θ_in, R_reco and Z_DCR are heterogeneous, low-dimensional, and physically calibrated.

Tunable score

BDT probabilities naturally define analysis operating points: efficiency vs mis-ID.

Stable baseline

Gradient-boosted trees are strong and interpretable for tabular detector observables.

Non-ML / pre-ML reference points

- Prototype ring/β reconstruction was validated on 1.2×10^{6} pions in the 0.6–1.5 GeV range.
- β-resolution and Gaussian μ/π separation are used as physics sanity checks.
- The classifier extends this into a 3-class score and directly optimizes analysis-level mis-ID.

Why not start with deep learning?

- The reconstruction already compresses detector response into physically meaningful features.
- A BDT gives a transparent baseline under realistic noise before moving to raw-hit, graph, or CNN models.
- DL use is more natural for fast simulation or trigger studies performed by another group.

Training strategy separates model learning from physics validation

Training sample

μ, π, K particle gun
uniform p = 300–2500 MeV/c
isotropic angles

→

Model

BDT classifier
learns species boundaries
under chosen DCR

→

Physics validation

$D^0 \to K^- \mu^+ \nu_\mu$
vs $D^0 \to K^- \pi^+ \pi^0$
realistic kinematics

3 × 150k

training events for μ, π, K

3 × 50k

held-out test events

10^6

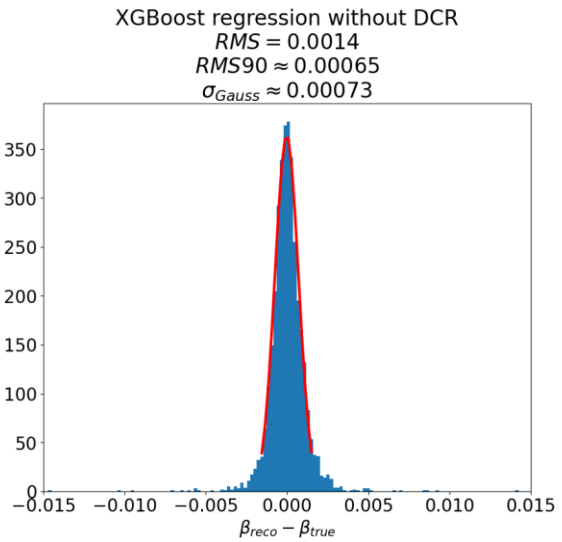
simulated D^0 decay events for validation

2 DCR modes

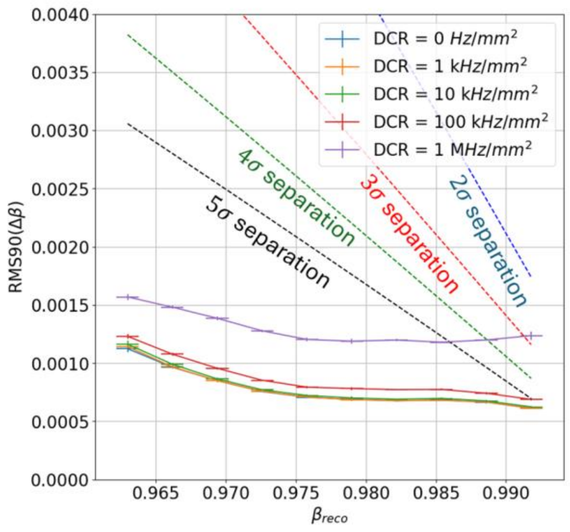
100 kHz/mm^2 and 1 MHz/mm^2 classifiers

Key ML validation: transfer from uniform single-particle training to realistic decay kinematics, rather than training directly on the analysis distribution.

Regression cross-check: reconstructed β remains physically meaningful



β residual distribution



Expected μ/π separation vs β

Dedicated XGBoost regressor used as a cross-check, not as the final PID model

Central 90% RMS of Δβ validates stability of reconstructed ring radius

5σ μ/π separation up to β ≈ 0.995 for DCR ≤ 100 kHz/mm^2; up to β ≈ 0.985 in the 1 MHz/mm^2 stress test

The classification stage is built on observables that already pass a physics-resolution sanity check.

Track-angle uncertainty and noise proxy are explicit caveats

These checks reduce the risk that the ML score is only simulation-specific.

Track angle

Simulation uses the IP angle; in operation the tracker provides the angle after the outer tracker wall.

Δθ study

Multiple scattering estimate: $\Delta\theta_{tracker} \approx 5 \times 10^{-3}$ vs $\Delta\theta_{IP} \approx 10^{-3}$, raising $\Delta\beta$ by $\approx 2 \times 10^{-4}$.

Still usable

The β-resolution remains within the relevant separation region across the considered DCR scenarios.

Noise-aware feature motivation

Pure-DCR fake rings typically stay within $\approx 8\sigma_{DCR}$ of the mean noise level.

Real signal peaks are much farther from the DCR baseline.

This motivates Z_{DCR} as a compact robustness feature.

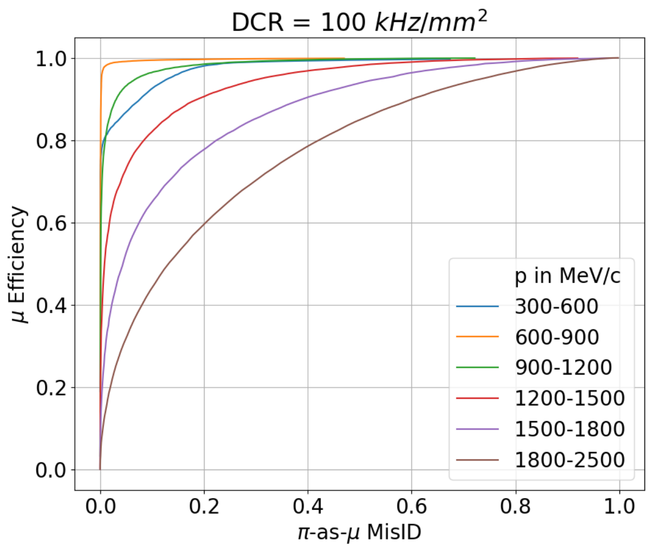
Deployment caveat

Current classifier inputs use MC-truth p and θ to isolate FARICH performance.

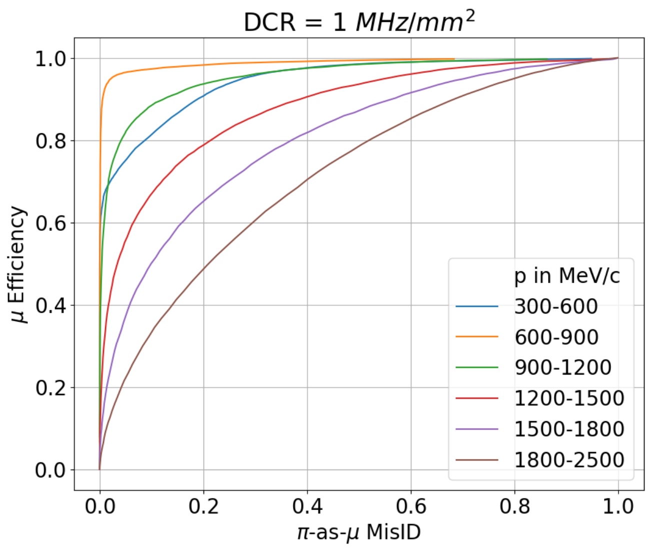
Next validation: replace them with reconstructed tracker quantities and calibrate the DCR proxy from dedicated runs.

Main message: the classifier is robust because the feature engineering already accounts for geometry, timing and dark-count structure.

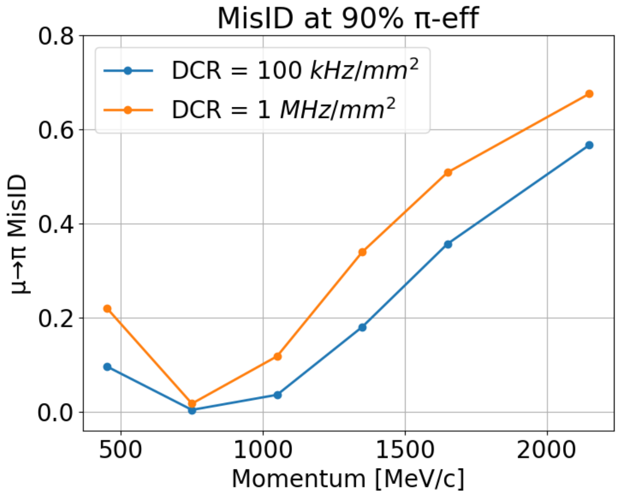
Single-particle validation: μ/π separation is learned where it matters



ROC: 100 kHz/mm²



ROC: 1 MHz/mm²



μ as π mis-ID at 90% efficiency

600–900 MeV/c

best separation window

<0.2

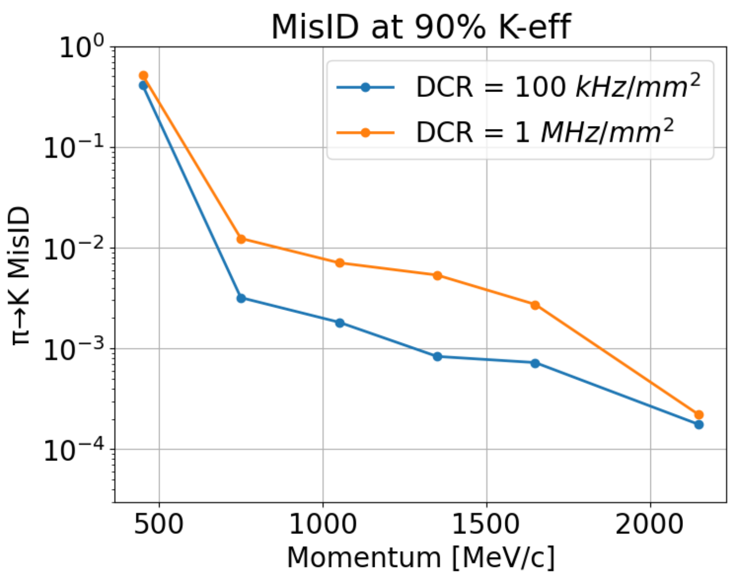
π→μ and μ→π mis-ID up to 1200 MeV/c at 90% efficiency

<0.1

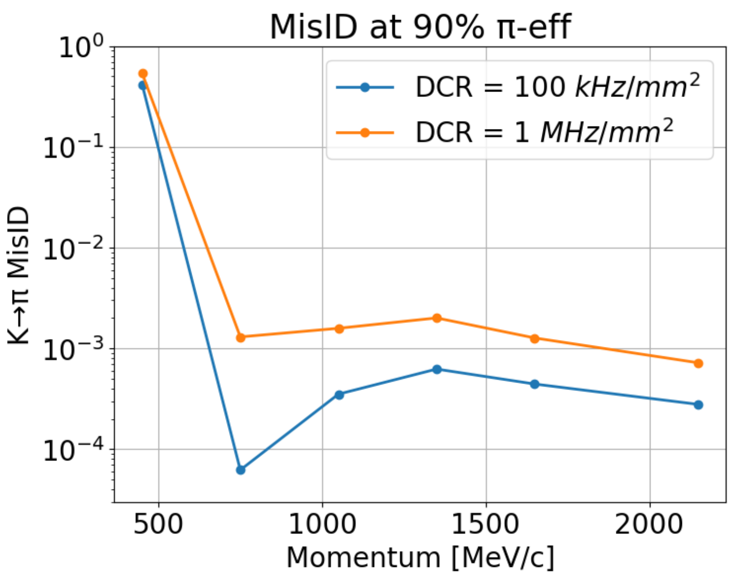
mis-ID up to 1200 MeV/c at 80% efficiency

Low momentum is limited by missing Cherenkov signal; high momentum by radius differences approaching the geometrical resolution limit.

Single-particle validation: π/K separation is robust above threshold



π as K mis-ID at 90% efficiency



K as π mis-ID at 90% efficiency

<0.01

π→K mis-ID above 900 MeV/c for all DCR values

$<3\times10^{-3}$

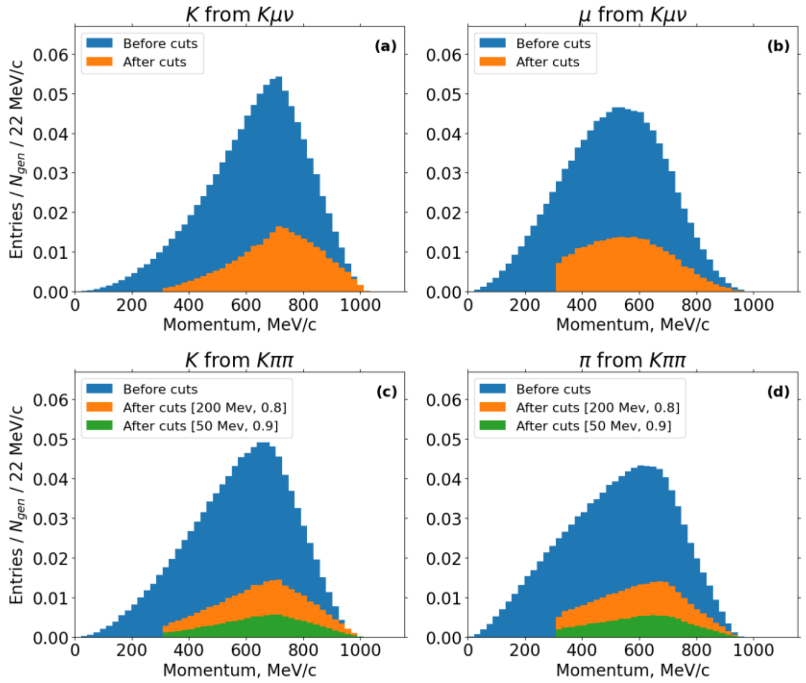
K→π mis-ID above 900 MeV/c for all DCR values

90%

fixed signal-efficiency operating point for comparison

The classifier maintains strong π/K rejection even under the conservative 1 MHz/mm² dark-noise scenario.

From uniform training samples to realistic kinematics



Momentum spectra after event selection

- Classifier trained on single particles (uniformly distr. momenta)
- Performance metrics:
  - signal efficiency:
    $\varepsilon_{PID} = \mathrm{Nsig}_{\mathrm{PID}}(K\mu) \,/\, \mathrm{Nsig}_{\mathrm{sel}}$
  - background misID:
    $\tau_{PID} = \mathrm{Nbgd}_{\mathrm{PID}}\,(K\mu) \,/\, \mathrm{Nbgd}_{\mathrm{sel}}$
- **100**-fold background suppression at **75%** efficiency for worst-case noise and **85%** for realistic noise

Analysis metrics and classifier transfer

$D^0 \rightarrow K^- \mu^+ \nu_\mu$

signal channel

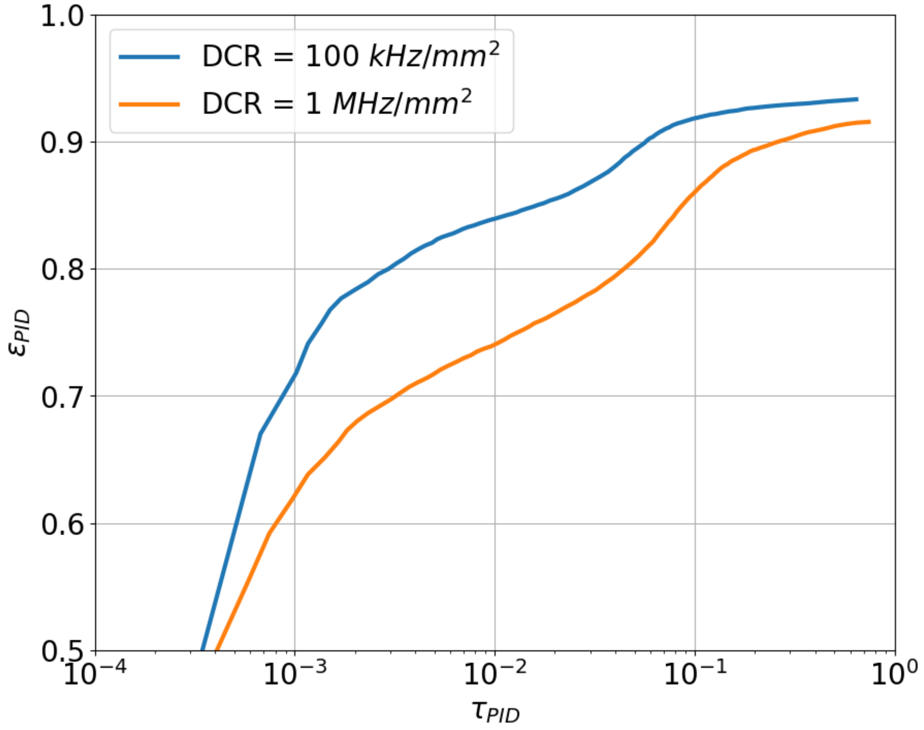
$D^0 \rightarrow K^- \pi^+ \pi^0$

dominant background

FARICH only

PID performance isolated from global PID

Classifier score enables an analysis-tunable operating point



Signal efficiency vs. background mis-ID threshold

Operating point examples

100×

background suppression

75%

signal efficiency at 1 MHz/mm^2

85%

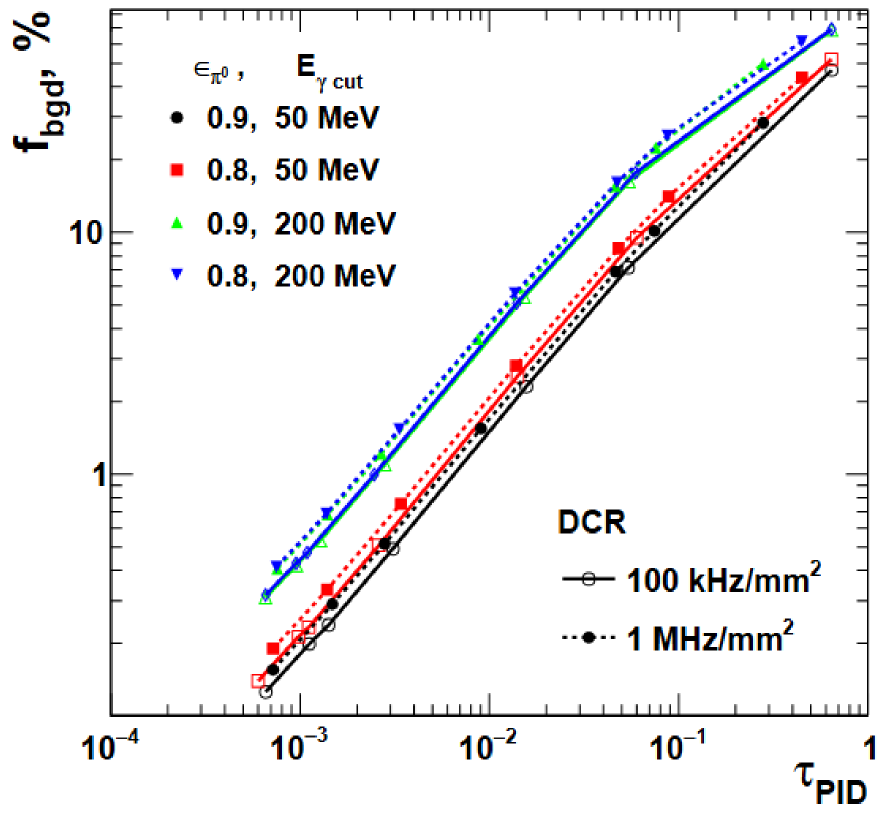
signal efficiency at 100 kHz/mm^2

Threshold can be chosen per analysis objective

Noise stress test remains within useful signal-efficiency range

The same score provides a natural input to future global PID combination

Analysis impact: π/μ mis-ID background contribution below 0.5%



Residual background fraction after classifier

Why this matters

- Directly reduces the background component from $D^0 \to K^-\pi^+\pi^0$
- Limits PID-induced background contamination at the sub-percent level
- Confirms transfer from single-particle training to physics-analysis kinematics

<0.5%

background contribution reachable in the analysis study

0.7–0.8

typical signal-efficiency range at this background level

Take-away: physics-informed ML makes FARICH PID robust, tunable, and analysis-ready

1 **Feature engineering** Dedicated reconstruction converts noisy FARICH hit patterns into stable physics observables.

2 **Supervised PID** XGBoost BDT separates μ, π, K using $p_{in}, \theta_{in}, R_{reco}$, and a DCR proxy.

3 **Noise robustness** Performance remains useful up to a conservative 1 MHz/mm^2 SiPM dark-count rate.

4 **Physics validation** $D^0 \rightarrow K^-\mu^+\nu_\mu$ study confirms analysis-level background suppression.

Next steps: replace MC-truth p and θ by reconstructed tracker quantities; integrate FARICH scores into global PID and fast simulation.

Backup slides

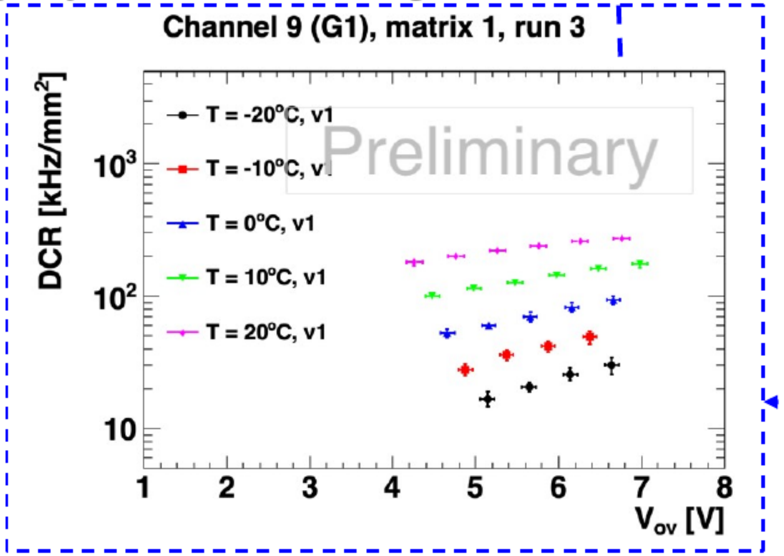
Original full data-processing workflow

Data processing workflow

Simulation Aurora, Geant4 → Digitisation PDE DCR → Ring reconstruction from track impact point, hit time, hit density → Classification BDT from R_{reco}, p_{in}, θ_{in}, DCR measure

Dedicated procedure for FARICH described in our paper

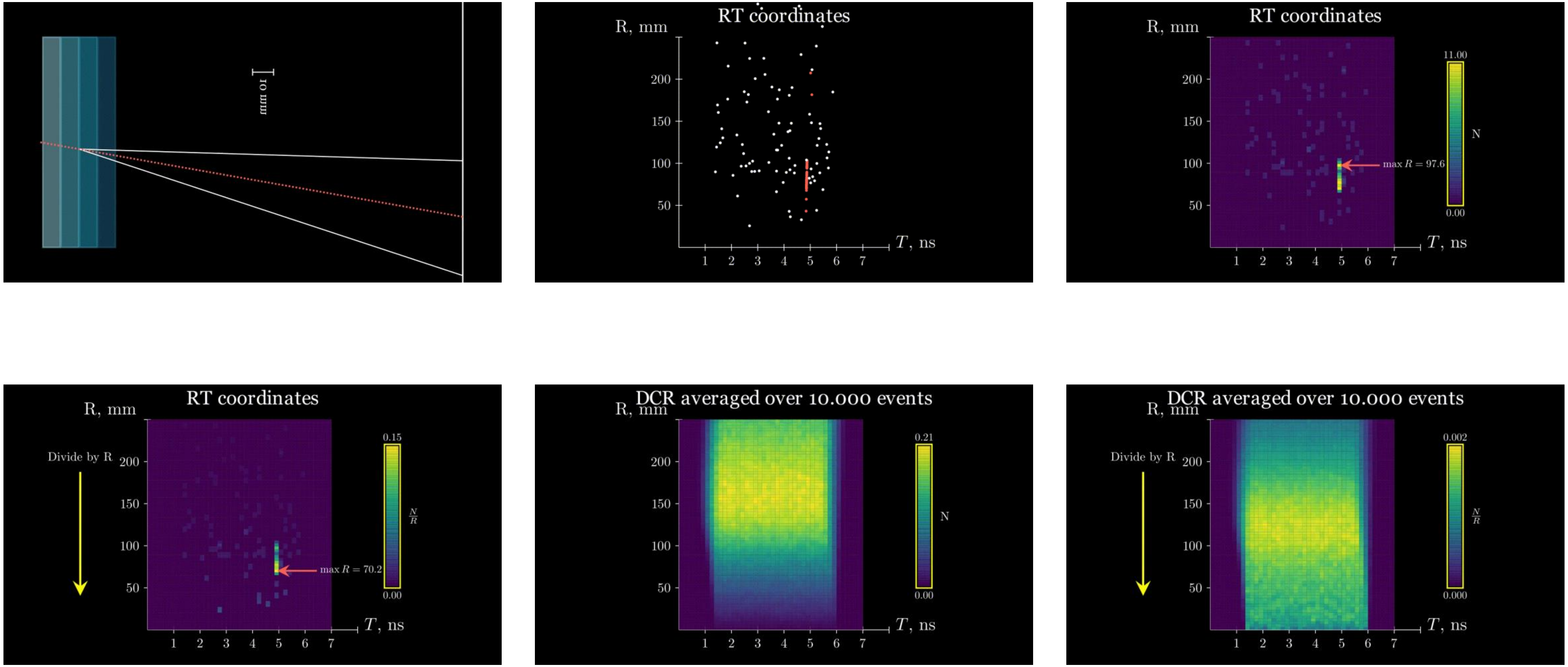
Calibration and Reconstruction Algorithm for FARICH System of the Detector at Super Charm-Tau Factory / A. Yu. Barnyakov, V. S. Bobrovnikov, S. A. Kononov, P. D. Rogozhin& M. V. Chadeeva// Bulletin of the Lebedev Physics Institute – 2023 – Volume 50 – pages 534 – 539



- KETEK SiPM matrix 8x8 (3x3 mm²)
- DCR vs. overvoltage and temperature measured at LPI
- preliminary

27.11.2025 Platon Rogozhin 5

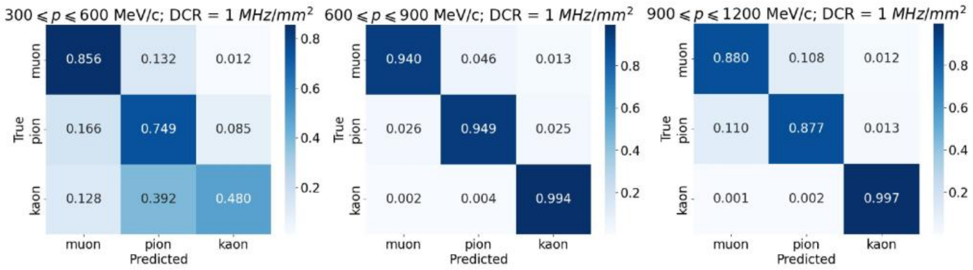
Ring-reconstruction animation sequence



Classification quality: mixing matrices

Classification Quality: mixing matrices

In momentum bins

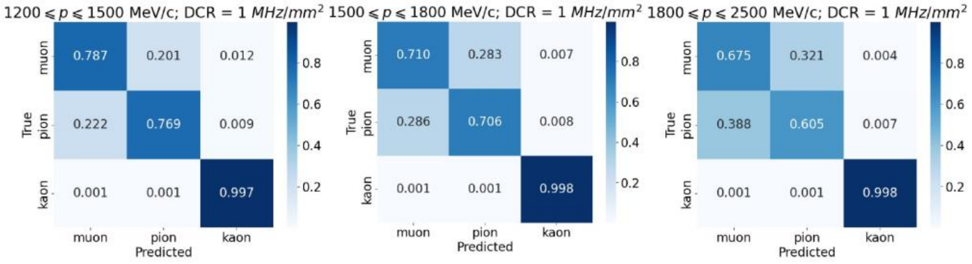


- Large π/K mixing in low momentum range due to lack of signal hits from both particles
- Almost no mixing with K in medium momentum range due to lack of signal hits only from K

27.11.2025 Platon Rogozhin 31

Classification Quality: mixing matrices

In momentum bins



- Increase in π/μ mixing due to rings getting closer with higher momentum
- Low mixing with K due to rings being of significantly different sizes

27.11.2025 Platon Rogozhin 32

Event selection and PID performance estimate

PID performance estimation
Event selection

$\mathcal{B}_{signal}$	ϵ_{tr}^{signal}	ϵ_{F}^{signal}	N_{sel}^{signal}	$\mathcal{B}_{bgd}$	ϵ_{tr}^{bgd}	ϵ_{F}^{bgd}	$P_{lost\pi^0}$ $[\epsilon_{\pi^0}, E_{\gamma cut}]$	N_{sel}^{bgd}
0.034	0.58	0.84	47415	0.144	0.58	0.78	0.32 [0.9, 50 MeV]	65293
							0.40 [0.8, 50 MeV]	80352
							0.81 [0.9, 200 MeV]	164219
							0.83 [0.8, 200 MeV]	168285

- Tracker reconstruction efficiency
- Efficiency of hitting barrel

$\Big\} \ \varepsilon_{tr} \approx 0.58$

- Not reconstructing π^0: $E_\gamma < E_{cut} \in \{50, 200\}$ MeV
- Calorimeter efficiency: $\varepsilon_\pi \in \{0.8, 0.9\}$
- Minimum required momentum to reach FARICH (ϵ_F): $p \geq p_{min} = 310$ MeV/c
  - $N_{sel}^{signal} = N_{gen}^{signal} \cdot \epsilon_{tr}^{signal} \cdot \epsilon_{F}^{signal}$
  - $N_{sel}^{bgd} = N_{gen}^{bgd} \cdot \epsilon_{tr}^{bgd} \cdot \epsilon_{F}^{bgd} \cdot P_{lost\pi^0}$

PID performance estimation
Fraction of background events

- The main contribution to the improvement of the signal-to-background ratio comes from the particle identification system.
- It is possible to achieve a background contamination below **0.5%**.
- Even at the lowest π⁰ reconstruction efficiency, the signal-to-background ratio before the application of the classifier is **47,415 / 168,285**, while after applying the classifier, a background contribution of **0.5%** can be reached with a signal efficiency of **0.7–0.8**.