Machine Learning for FARICH Reconstruction at NICA SPD

F.Shipilov[1], A.Barnyakov[2,3], A.Ivanov[4], F.Ratnikov[1]

(1) HSE University (2) Budker Institute of Nuclear Physics (3) Novosibirsk State Technical University (4) JINR

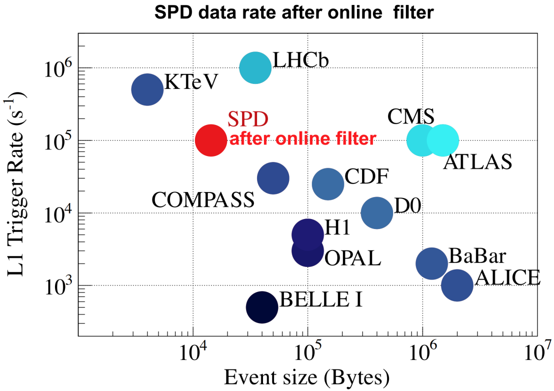
foma@shipilov.ru

LAMBDA • HSE

May 15, 2025

Introduction: Spin Physics Detector

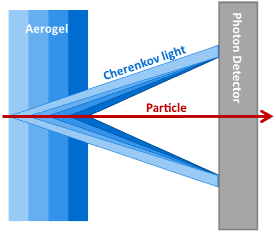
- A universal detector proposed by the SPD collaboration at NICA
- Study the Drell–Yan (DY) processes, J/Ψ production processes, elastic reactions, spin effects in one and two hadron production processes, polarization effects in heavy ion collisions, and more
- The SPD is a medium energy experiment, offering unique possibilities of beam operation and bridging the gap between the low-energy measurements, e.g. ANKE-COSY and the high-energy measurements, such as RHIC
- High luminosity up to 10^{32} cm^{-2}s^{-1} and free-flowing (triggerless) running mode

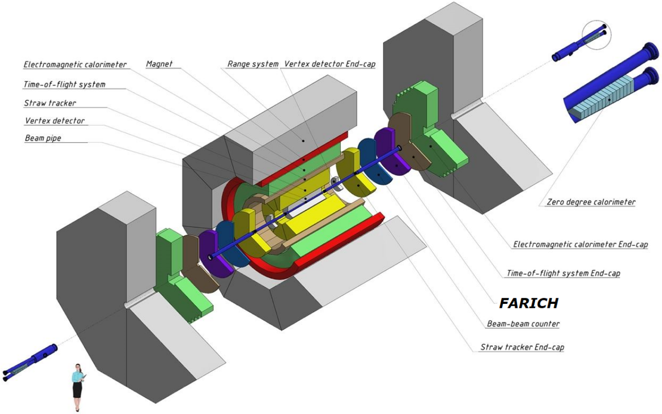


Focusing Aerogel Ring Imaging Cherenkov

Compact particle ID detector

- Trade off between resolution and number of Cherenkov photons
- Improve resolution by using layers of aerogel with varying refractive index





https://nica.jinr.ru/projects/spd.php

Aim and Objectives

- Particle identification (PID): the goal is to identify the most probable particle type
- Ambiguous with FARICH only, momentum p from straw tracker is needed

$$\beta = \frac{1}{n \cos \theta_c}, \quad m_0 = \frac{p\sqrt{1-\beta^2}}{\beta} \tag{1}$$

- In terms of optimization, we seek a solution to a classification problem

$$\mathbb{E}_{(X,y)\sim q(X,y)} \left[Q\big(f_W(X), y\big) \right] \to \max_W$$

- End-to-end neural networks can combine both steps from (1)
- Potentially better performance by optimizing a more appropriate objective

Objectives: develop and implement a variety of both β-outputting and end-to-end models, evaluate on a synthetic dataset, analyze and compare with the baseline

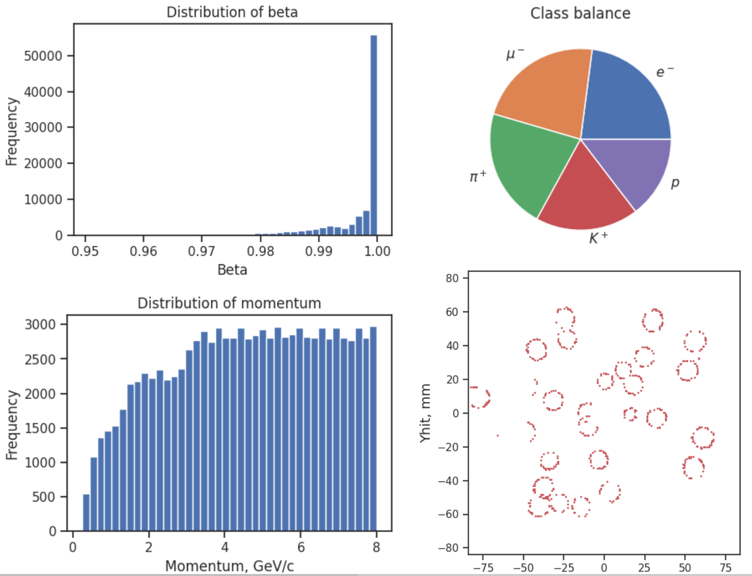
Balanced Dataset

Provided input:

- Track parameters: x_p, y_p, z_p – coordinates of the particle when entering aerogel detector, θ_p, ϕ_p – polar angle and azimuth of the direction of travel, momentum p
- Photon hits: x_c, y_c, z_c – coordinates of triggered pixels in the photosensitive matrix

Expected output:

- PID statistics
- Particle type



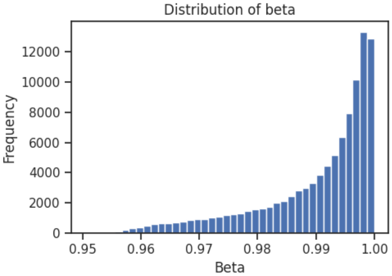
Production Dataset

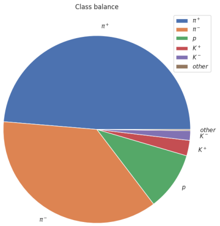
Provided input:

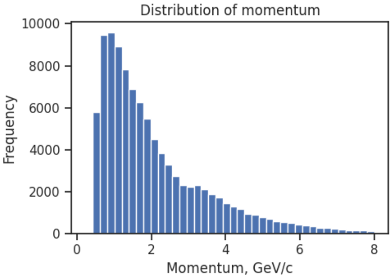
- Track parameters: x_p, y_p, z_p – coordinates of the particle when entering aerogel detector, θ_p, ϕ_p – polar angle and azimuth of the direction of travel, momentum p
- Photon hits: x_c, y_c, z_c – coordinates of triggered pixels in the photosensitive matrix

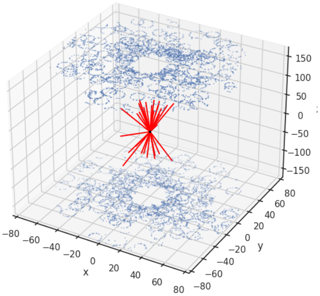
Expected output:

- PID statistics
- Particle type



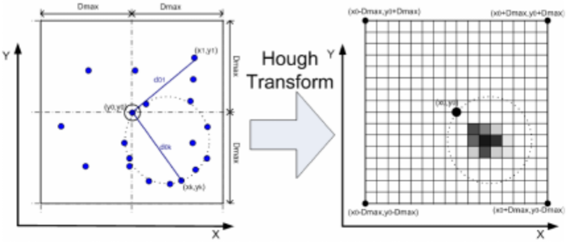






Balanced Baseline: Hough Transform

- Based on the RICH reconstruction from the CBM experiment at FAIR
- Utilizes Hough transform for ellipses, more precisely the Taubin method
- Outputs estimated parameters of the ellipse, such as a, b – semi-axes, x, y – center coordinates, and ϕ – angle of rotation
- Does not account for refraction, does not use the information from the straw tracker, such as θ_p and ϕ_p



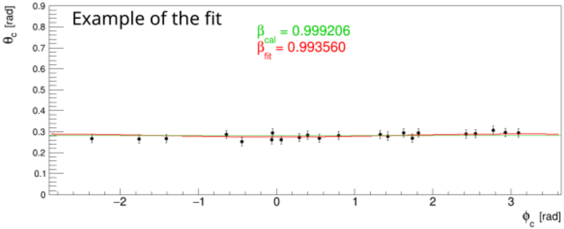
[1]

Production Baseline: θ_c by ϕ_c Fit

- Based on fitting

$$\theta_c(\phi_c|\beta, n, \theta_t) = \arccos\left(\frac{1}{n\beta}\right) + \arccos\left(n(1-(\mathbf{n}_0\mathbf{n}_\gamma)^2)\right) + (\mathbf{n}_0\mathbf{n}_\gamma)\sqrt{1-n^2(1-(\mathbf{n}_0\mathbf{n}_\gamma)^2)}$$

- Outputs estimated β – particle velocity
- Accounts for refraction, uses track information
- Is incomplete / broken (we have no control here)



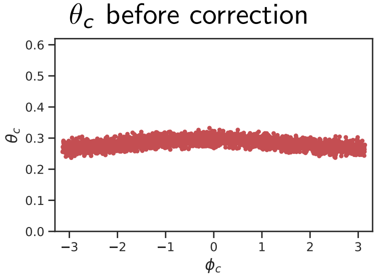
Refraction Correction

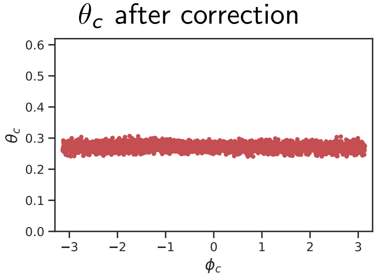
First order approximation

$$\alpha \approx \beta - (n-1)\tan\beta \tag{2}$$

Or fixed-point iteration

$$\sin\alpha_{k+1} = \left(n\sqrt{1+\left(\frac{d}{r-h\tan\alpha_k}\right)^2}\right)^{-1} \tag{3}$$



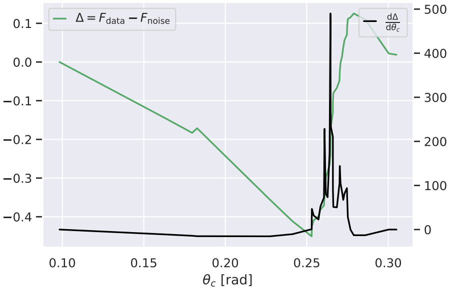


Median:

- Compute θ_c distribution
- Take a median value $\hat{\theta}_c$

MLE:

- Compute θ_c distribution
- Drop unphysical velocities $\beta = v/c \geqslant 1$
- Construct eCDF F_{data}
- Take a numerical derivative and find its peak $\hat{\theta}_c = \arccos(1/n\hat{\beta})$

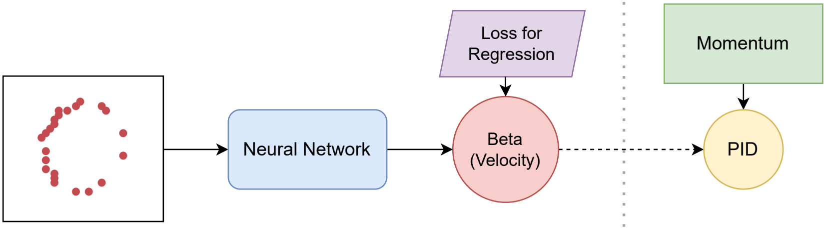


MLE example

Machine Learning Models: Beta NN

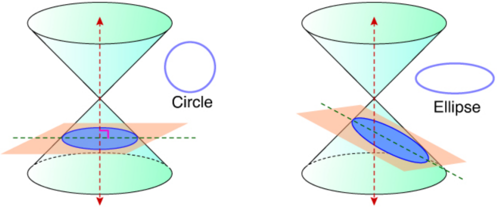
- Algorithmic models output $\hat{\theta}_c$ that can be converted to the rest mass of the particle using momentum p from the straw tracker
- The underlying task is regression
- A simple approach is to train a neural network to predict particle velocity β and then convert it into mass

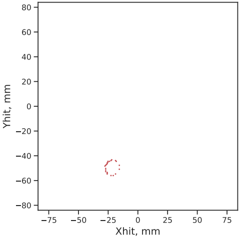
We implement a neural regressor that learns β by optimizing MSE objective.

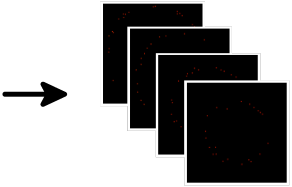


Data Transform

Simple reproject + Fourier features



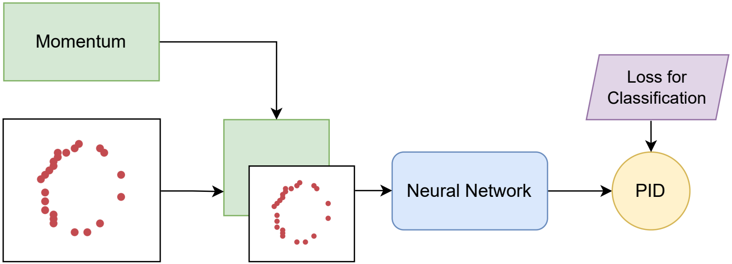




Machine Learning Models: Extra Channel NN

- The objective of RICH detector is to separate different particle types, in particular π/K separation
- ML end-to-end models are able to skip the intermediate step of computing $\hat{\theta}_c$
- Moreover, the regression task is not well suited for an ML model because of the significant value imbalance

A neural classifier can receive **momentum** p as an extra channel in the input tensor along with **hit coordinates** x_c, y_c, optimizing cross entropy objective directly.

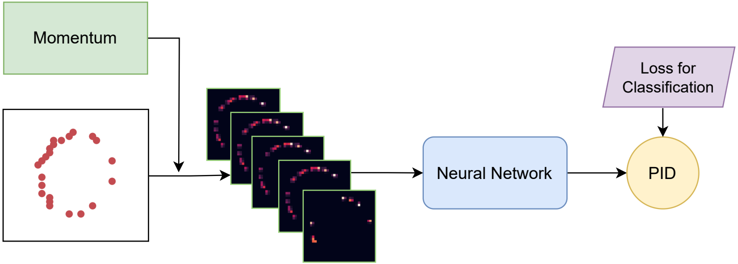


Machine Learning Models: Renormalize NN

The non-linear dependence of the ellipse dimensions on p can be disentangled:

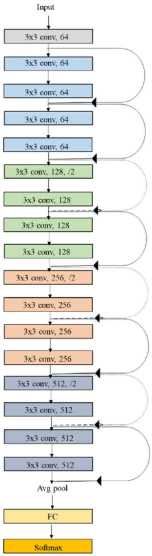
$$\cos\theta_c = \frac{1}{n}\sqrt{m_0^2/p^2 + 1} \tag{4}$$

- Use (4) with rest masses m_0 of all N particle types in the data to transform θ_c into N rings, diameter of the ring corresponding to true particle type (class) is constant and known

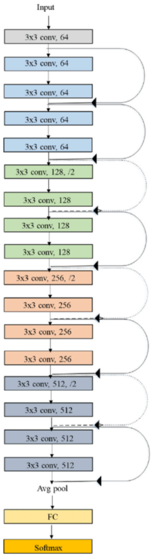


Training Details (Balanced)

- 3 configurations
  - Beta NN
  - Extra Channel NN
  - Renormalize NN
- ResNet-18 CNN architecture
  - Changed input conv, max pooling and classifier head to accommodate input and output data formats
- Input format: $N \times 32 \times 32$ tensor
- Output format: 5 class classification
- Trained for 1700000 samples with Adam optimizer, cosine annealing scheduler, learning rate $3 \cdot 10^{-4}$ and batch size 128

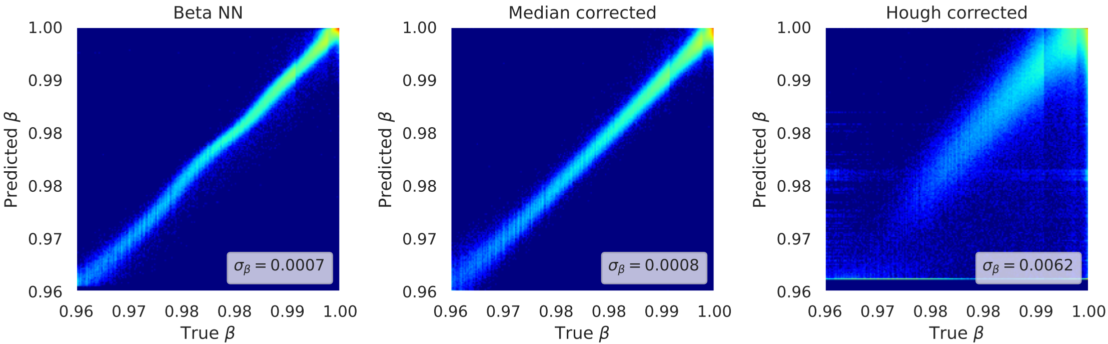


Training Details (Production)

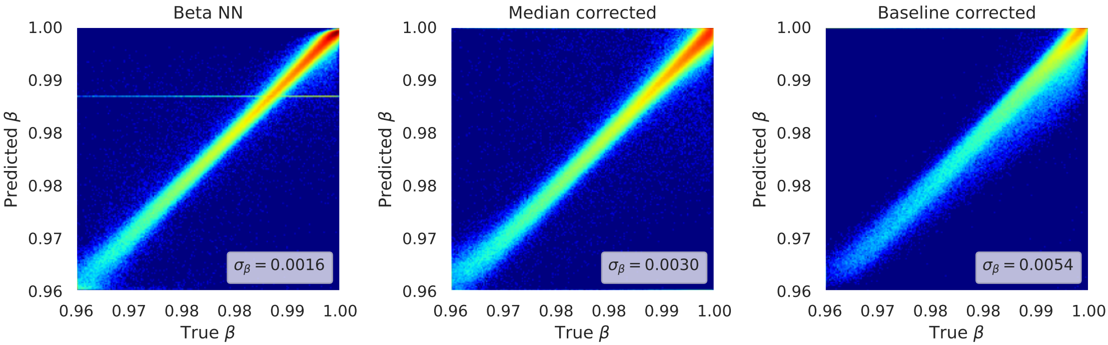


- Output format: binary classification (π/K)
- Countering class imbalance 20/1
  - Positive class weight
  - Downsampling
  - No adjustment (control)
- For β-outputting NN MSE can be weighted according to class
- The best performing training procedure (AUCROC on validation) evaluated on the test set

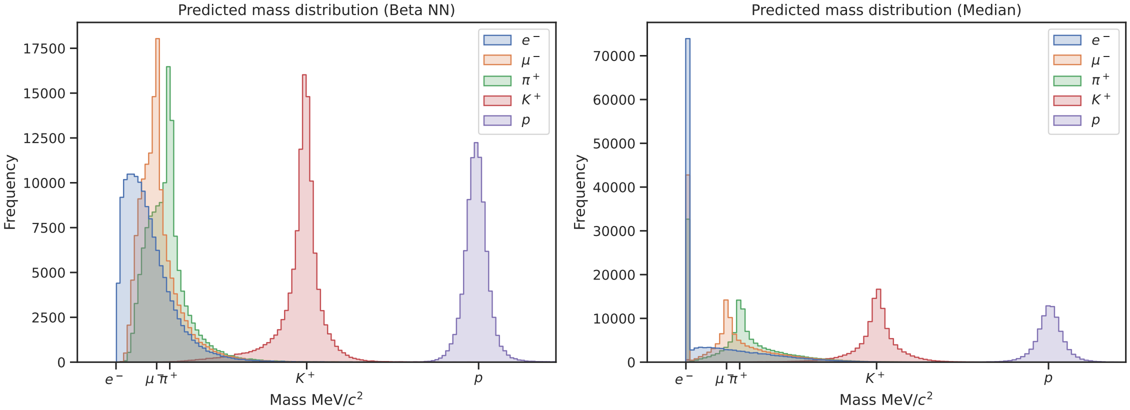
Results (Balanced)



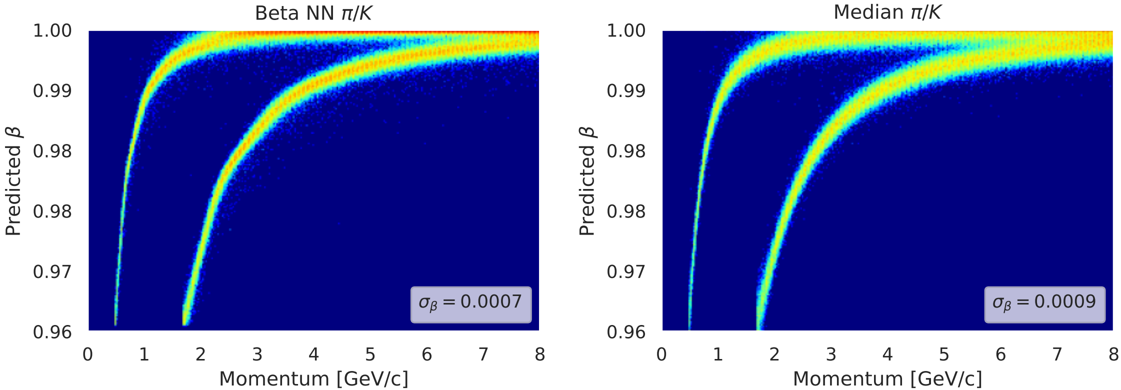
Results (Production)



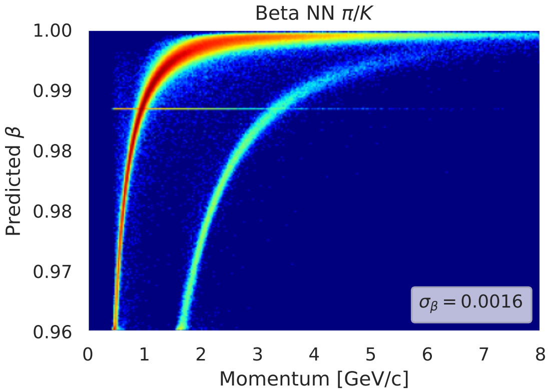
Results (Balanced)

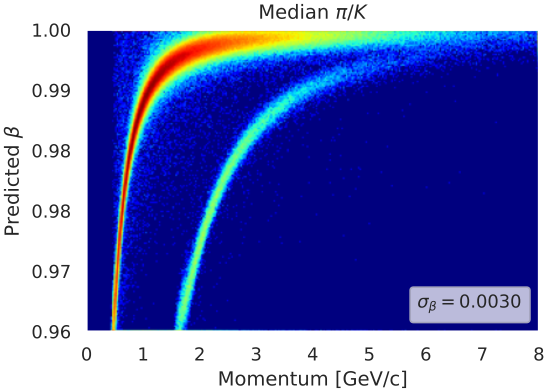


Results (Balanced)

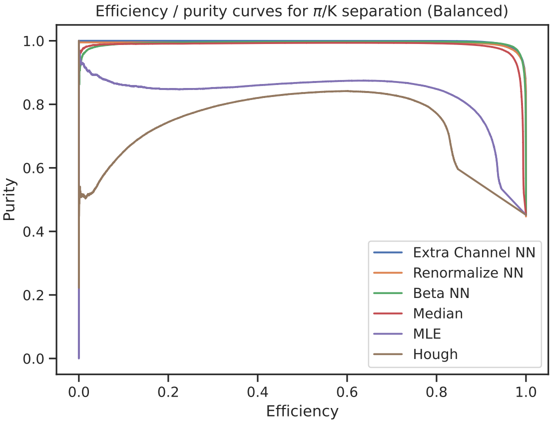


Results (Production)

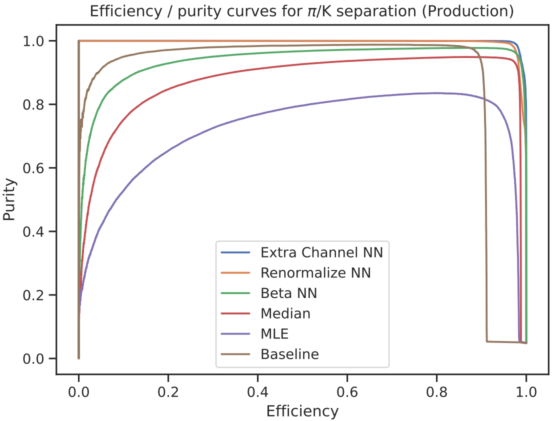




Results

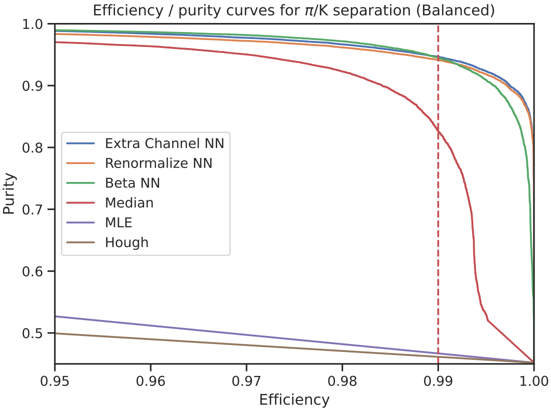


Efficiency / purity curves for π/K separation (Balanced)
Purity
Efficiency
1.0
0.8
0.6
0.4
0.2
0.0
Extra Channel NN
Renormalize NN
Beta NN
Median
MLE
Hough

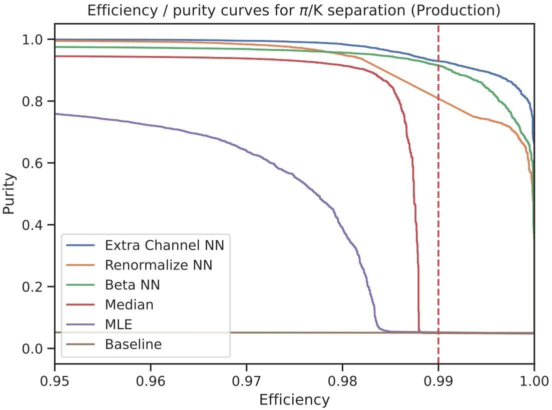


Efficiency / purity curves for π/K separation (Production)
Purity
Efficiency
1.0
0.8
0.6
0.4
0.2
0.0
Extra Channel NN
Renormalize NN
Beta NN
Median
MLE
Baseline

Results



Efficiency / purity curves for π/K separation (Balanced)
1.0
0.9
0.8
0.7
0.6
0.5
Purity
0.95
0.96
0.97
0.98
0.99
1.00
Efficiency
Extra Channel NN
Renormalize NN
Beta NN
Median
MLE
Hough



Efficiency / purity curves for π/K separation (Production)
1.0
0.8
0.6
0.4
0.2
0.0
Purity
0.95
0.96
0.97
0.98
0.99
1.00
Efficiency
Extra Channel NN
Renormalize NN
Beta NN
Median
MLE
Baseline

Results

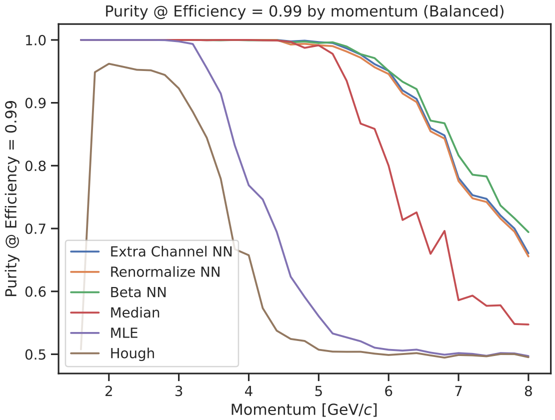
Method	π, K Purity@Efficiency $= 0.99$	π, K AP	5 class Accuracy	σ_β
Ex Channel NN	**0.95**	**0.998**	**0.68**	N/A
Renormalize NN	0.95	0.998	0.67	N/A
Beta NN	0.92	0.993	0.64	0.0007
Median	0.83	0.985	0.65	0.0008
MLE	0.47	0.836	0.53	0.0018
Hough	0.46	0.721	0.50	0.0062

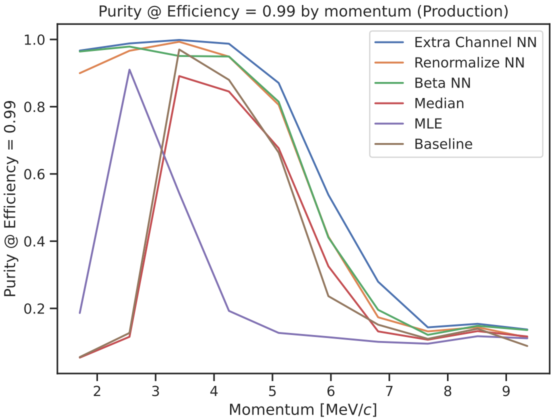
Table 1. Balanced dataset

Method	Purity@Efficiency $= 0.99$	AUROC	AP	σ_β
Ex Channel NN	**0.91**	**0.9998**	**0.997**	N/A
Renormalize NN	0.77	0.9996	0.993	N/A
Beta NN	0.90	0.999	0.919	0.0018
Median	0.75	0.989	0.870	0.0030
MLE	0.05	0.978	0.733	0.0037
Baseline	0.05	0.923	0.905	0.0054

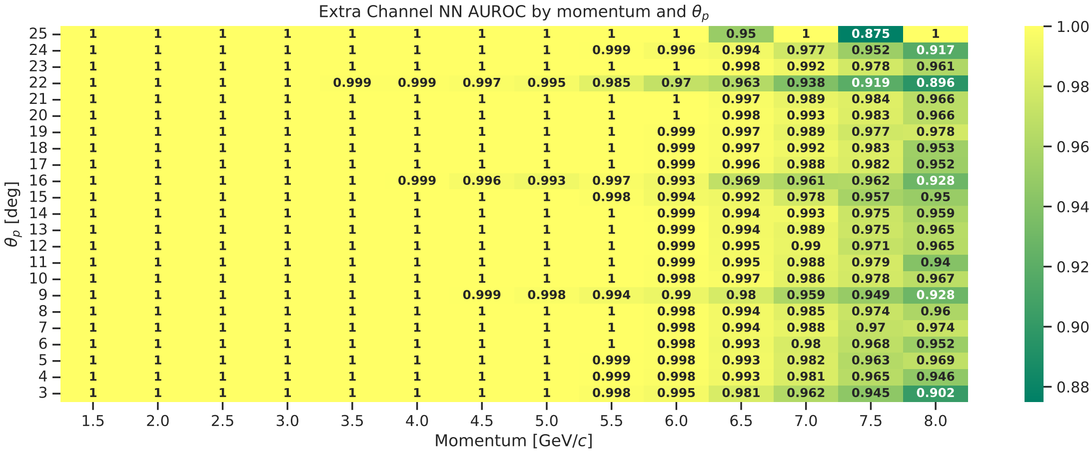
Table 2. Production dataset. Only binary π, K classification is considered

Detailed Results

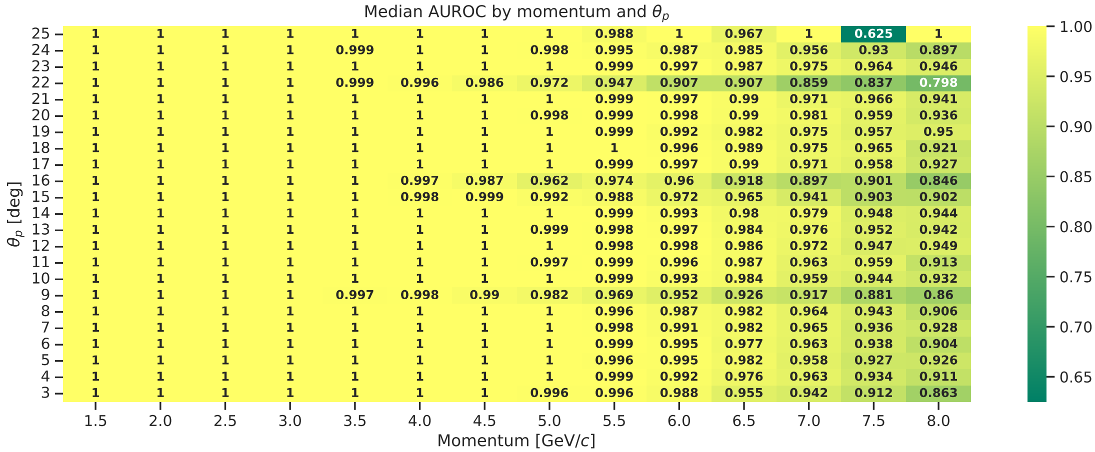




Detailed Results (Balanced)



Detailed Results (Balanced)

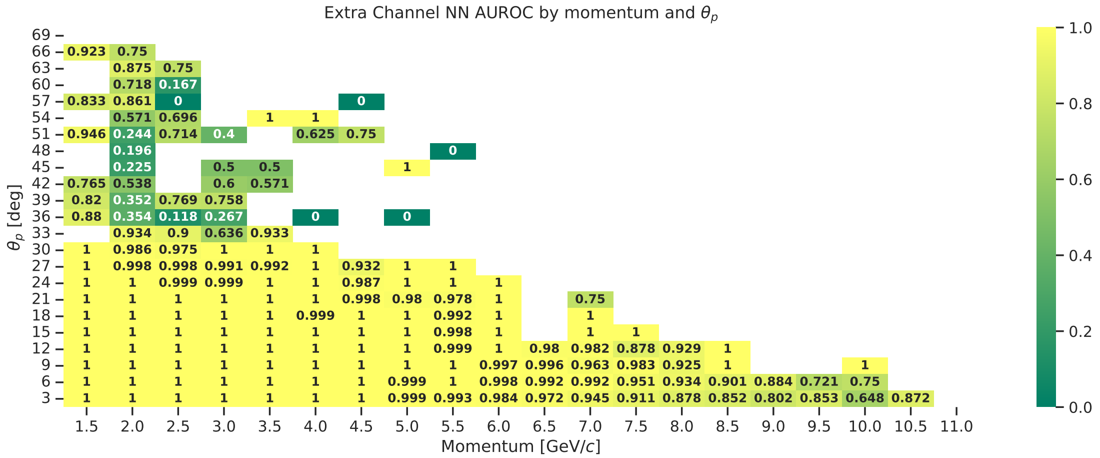


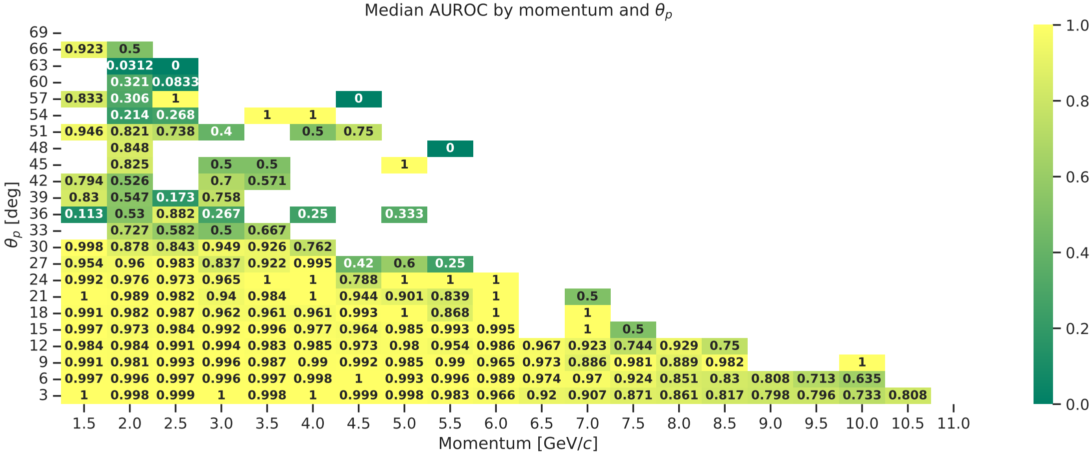
Detailed Results (Balanced)

Hough AUROC by momentum and θ_p

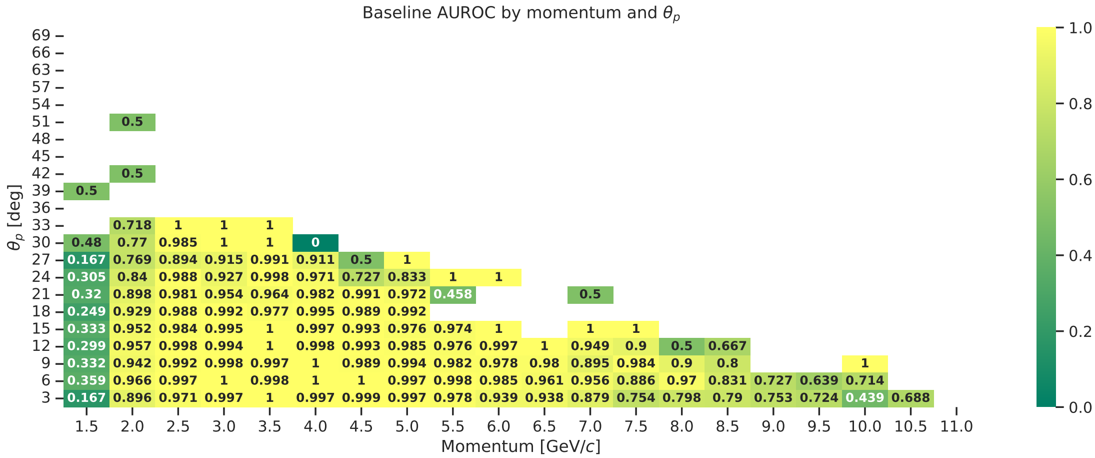
θ_p [deg] \ Momentum [GeV/c]	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0
24	0.994	0.987	0.978	0.935	0.902	0.812	0.775	0.709	0.698	0.635	0.64	0.616	0.587	0.588
23	0.998	0.995	0.993	0.999	0.976	0.954	0.886	0.827	0.774	0.705	0.69	0.615	0.592	0.635
22	0.935	0.862	0.774	0.714	0.717	0.686	0.657	0.619	0.597	0.598	0.574	0.567	0.543	0.542
21	0.999	0.995	0.991	0.992	0.981	0.952	0.893	0.844	0.788	0.752	0.702	0.639	0.649	0.593
20	0.999	0.995	0.997	0.99	0.981	0.955	0.878	0.826	0.768	0.731	0.695	0.687	0.645	0.632
19	0.999	0.995	0.993	0.988	0.979	0.952	0.884	0.827	0.818	0.749	0.686	0.66	0.648	0.598
18	0.996	0.999	0.993	0.998	0.97	0.945	0.904	0.838	0.787	0.715	0.696	0.671	0.642	0.642
17	0.998	0.992	0.996	0.995	0.98	0.943	0.893	0.86	0.801	0.72	0.72	0.664	0.63	0.63
16	0.925	0.87	0.782	0.788	0.738	0.689	0.697	0.626	0.662	0.626	0.595	0.553	0.615	0.576
15	0.99	0.965	0.917	0.915	0.864	0.871	0.808	0.77	0.725	0.677	0.649	0.639	0.633	0.593
14	0.999	0.997	0.994	0.999	0.99	0.979	0.912	0.852	0.781	0.746	0.705	0.68	0.64	0.615
13	0.997	0.996	0.995	0.994	0.99	0.979	0.911	0.874	0.799	0.765	0.694	0.674	0.641	0.591
12	0.998	0.998	0.998	0.991	0.997	0.979	0.946	0.859	0.778	0.78	0.7	0.673	0.684	0.654
11	0.998	0.995	0.99	0.994	0.992	0.976	0.943	0.892	0.832	0.772	0.704	0.675	0.708	0.613
10	0.999	0.999	0.997	0.993	0.996	0.982	0.957	0.893	0.831	0.81	0.73	0.715	0.645	0.667
9	0.929	0.92	0.821	0.791	0.767	0.79	0.732	0.735	0.685	0.626	0.619	0.627	0.606	0.622
8	0.999	0.998	0.992	0.995	0.985	0.974	0.958	0.926	0.864	0.816	0.761	0.714	0.7	0.61
7	1	0.998	0.998	0.992	0.992	0.979	0.961	0.925	0.874	0.816	0.794	0.749	0.719	0.737
6	0.998	0.994	0.997	0.993	0.983	0.975	0.945	0.929	0.899	0.831	0.799	0.749	0.725	0.682
5	0.998	0.996	0.994	0.99	0.988	0.972	0.965	0.926	0.879	0.817	0.792	0.754	0.721	0.728
4	0.998	0.993	0.988	0.964	0.984	0.964	0.915	0.879	0.838	0.795	0.741	0.733	0.704	0.584
3	0.964	0.978	0.948	0.93	0.902	0.857	0.823	0.756	0.726	0.692	0.672	0.646	0.616	0.609

Detailed Results (Production)




Median AUROC by momentum and θ_p
θ_p [deg]
69 66 63 60 57 54 51 48 45 42 39 36 33 30 27 24 21 18 15 12 9 6 3
0.923 0.5
0.0312 0
0.321 0.0833
0.833 0.306 1 0
0.214 0.268 1 1
0.946 0.821 0.738 0.4 0.5 0.75
0.848 0
0.825 0.5 0.5 1
0.794 0.526 0.7 0.571
0.83 0.547 0.173 0.758
0.113 0.53 0.882 0.267 0.25 0.333
0.727 0.582 0.5 0.667
0.998 0.878 0.843 0.949 0.926 0.762
0.954 0.96 0.983 0.837 0.922 0.995 0.42 0.6 0.25
0.992 0.976 0.973 0.965 1 1 0.788 1 1 1
1 0.989 0.982 0.94 0.984 1 0.944 0.901 0.839 1 0.5
0.991 0.982 0.987 0.962 0.961 0.961 0.993 1 0.868 1 1
0.997 0.973 0.984 0.992 0.996 0.977 0.964 0.985 0.993 0.995 1 0.5
0.984 0.984 0.991 0.994 0.983 0.985 0.973 0.98 0.954 0.986 0.967 0.923 0.744 0.929 0.75
0.991 0.981 0.993 0.996 0.987 0.99 0.992 0.985 0.99 0.965 0.973 0.886 0.981 0.889 0.982 1
0.997 0.996 0.997 0.996 0.997 0.998 1 0.993 0.996 0.989 0.974 0.97 0.924 0.851 0.83 0.808 0.713 0.635
1 0.998 0.999 1 0.998 1 0.999 0.998 0.983 0.966 0.92 0.907 0.871 0.861 0.817 0.798 0.796 0.733 0.808
1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0 5.5 6.0 6.5 7.0 7.5 8.0 8.5 9.0 9.5 10.0 10.5 11.0
Momentum [GeV/c]
1.0 0.8 0.6 0.4 0.2 0.0

Detailed Results (Production)



Discussion of Results

- Beta NN and Median – nearly identical performance
- Indicates the limit of β-based approaches
- End-to-end models are also close to each other; qualitative improvement over β-based models
- Almost perfect π/K separation on production data
- Machine learning-based approaches demonstrate flexibility and achieve better quality

Discussion of Results

Problems/Difficulties

- ResNet-18 is a large computationally slow model
- Prior data transforms (reprojection, Fourier features extraction, renormalization) are even more computationally expensive, training and inference are CPU-bound
  - We implemented data parallelism and achieved 7x speed up

Future Plans

- Architecture optimization
  - knowledge distillation, pruning and quantization
- Multi-ring PID, automatic track-ring matching using ML
- Integration into the SPD data processing pipeline; change point & calibration
- FARICH fast simulation

References

[1] Lebedev, S. A. and Ososkov, G. A. Fast algorithms for ring recognition and electron identification in the CBM RICH detector. Phys. Part. Nucl. Lett. 6, 161–176