Improving Particle Identification in the Belle II TOP detector through Machine Learning

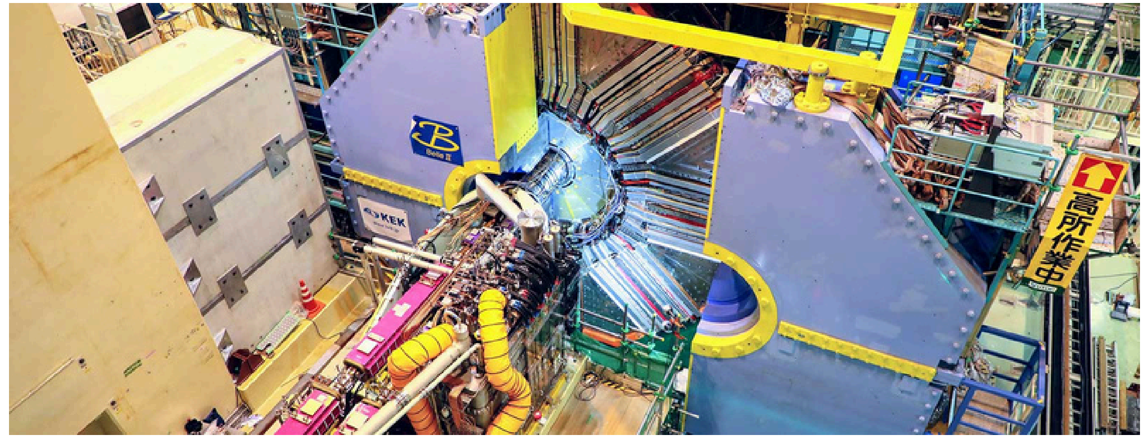
Cecilia Antonioli
University and INFN of Padova

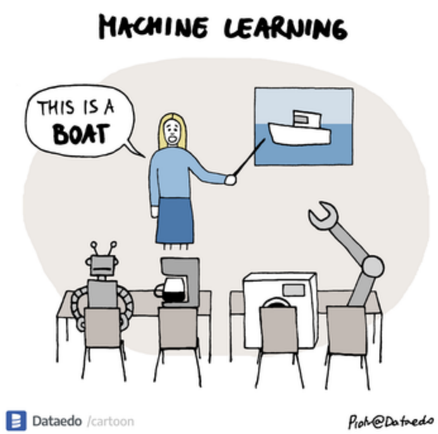
61st International Winter Meeting on Nuclear Physics

January 31, 2025, Bormio, Italy

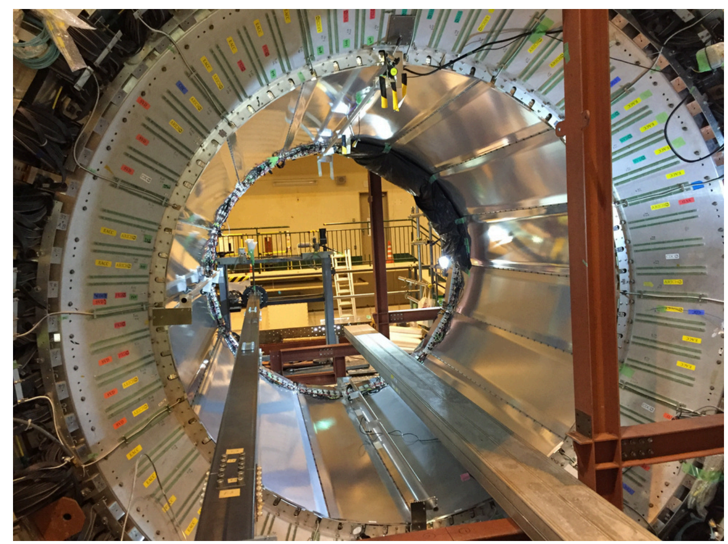
Outline

- Introduction
- Belle II experiment
- TOP counter
- TOP working principle
- TOP PID Likelihood
- TOP PID with Machine Learning





https://dataedo.com



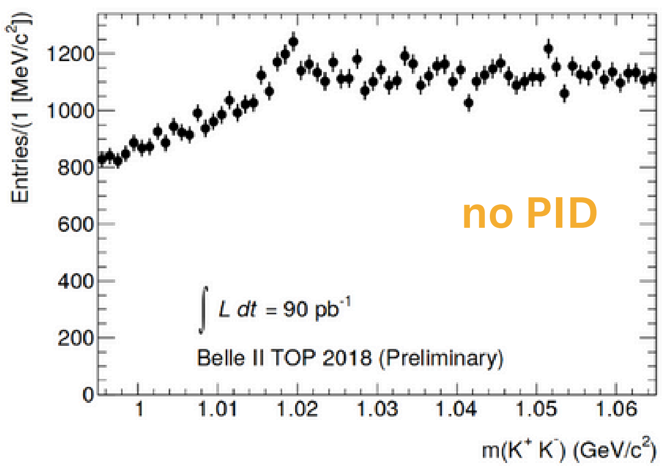
Introduction

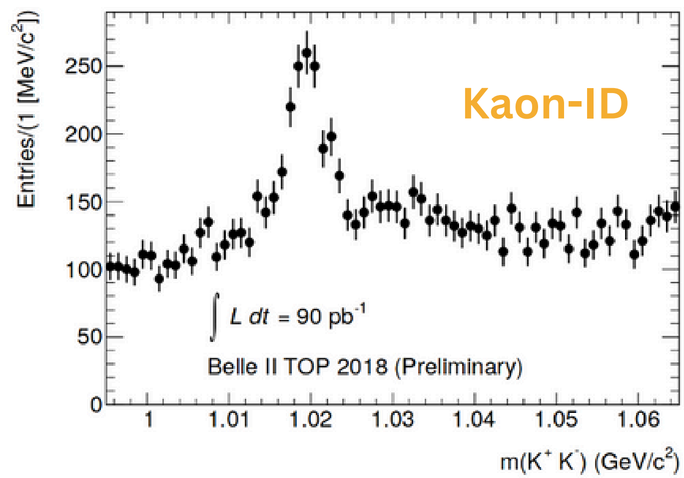
Reliable particle identification (PID) is essential for a flavor physics experiment

At *B*-factories, PID is required for

- the tagging of the *B*-meson flavor
- precision measurements of rare *B*/*D* decays
- ...

An example of PID effect on *Φ → K+ K-* reconstruction





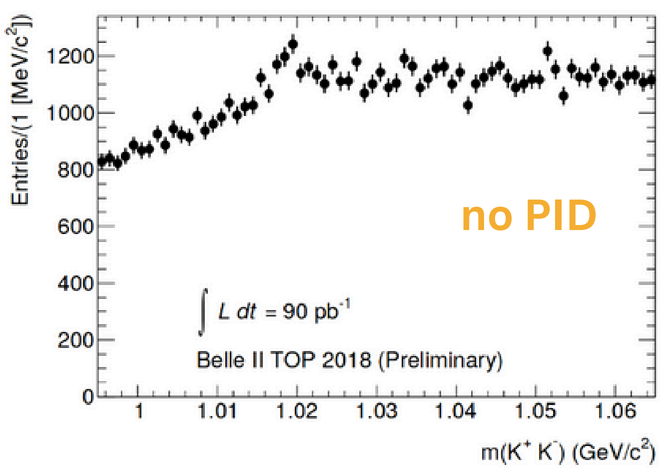
Introduction

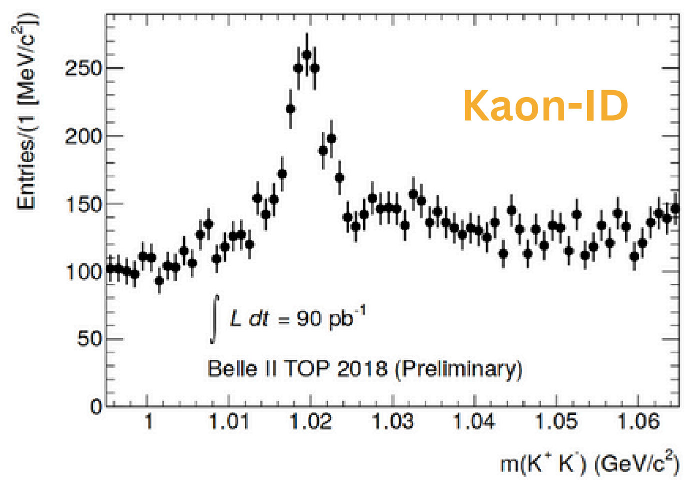
Reliable particle identification (PID) is essential for a flavor physics experiment

At *B*-factories, PID is required for

- the tagging of the *B*-meson flavor
- precision measurements of rare *B*/*D* decays
- ...

An example of PID effect on $\Phi \to K+\ K-$ reconstruction

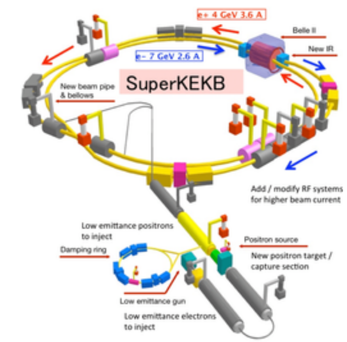




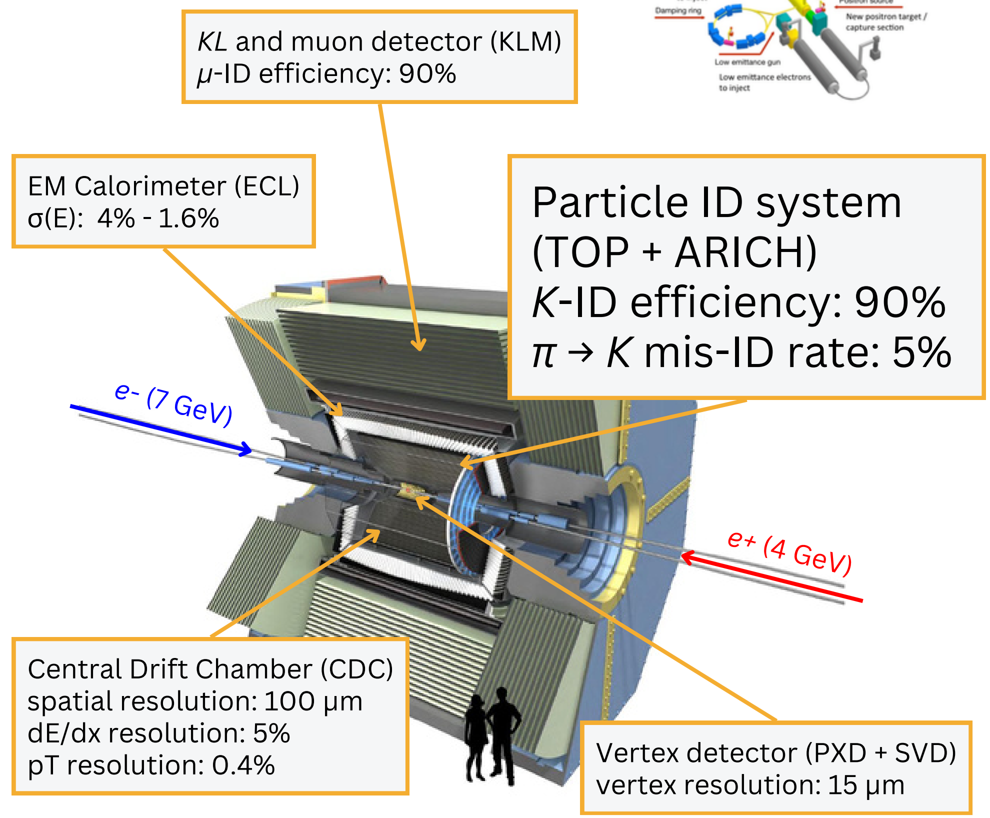
Belle PID performance

- *K*-ID efficiency: 90%
- $\pi \to K$ mis-ID rate: 10%

Belle II experiment at SuperKEKB



- Belle II is a luminosity frontier experiment to search for Physics Beyond the Standard Model
- It is located at SuperKEKB, the asymmetric *e+e-* collider at KEK in Tsukuba, Japan
- Collisions are (mostly) at the *Y(4S)* resonance, $\sqrt{s}$ = 10.58 GeV
- Peak luminosity record (Dec. 2024) $L = 5.1 \times 10^{34}$ cm^{-2} s^{-1}
- Accumulated 575 fb^{-1}

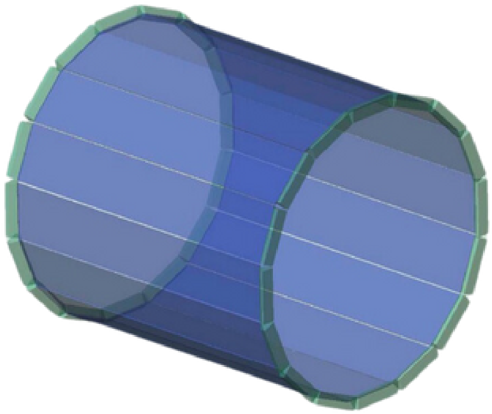


PID at Belle II

- Two dedicated PID devices:
  - **TOP in the barrel**
  - ARICH in the forward endcap
- All Belle II sub-detectors (except PXD) contribute to PID
- Six types of "stable" charged particles: *e*, *μ*, *π*, *K*, *p*, *d*

TOP counter

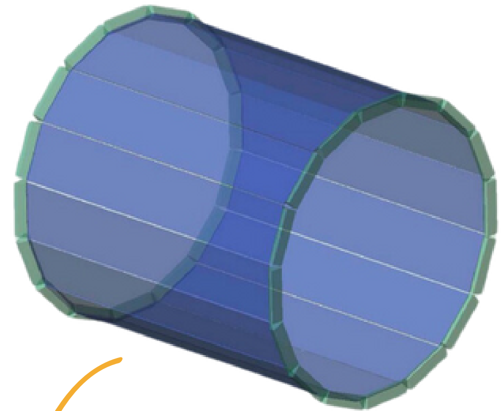
TOP stands for **Time-Of-Propagation**



It consists of 16 modules arranged around the Central Drift Chamber (CDC) in the barrel region

TOP counter

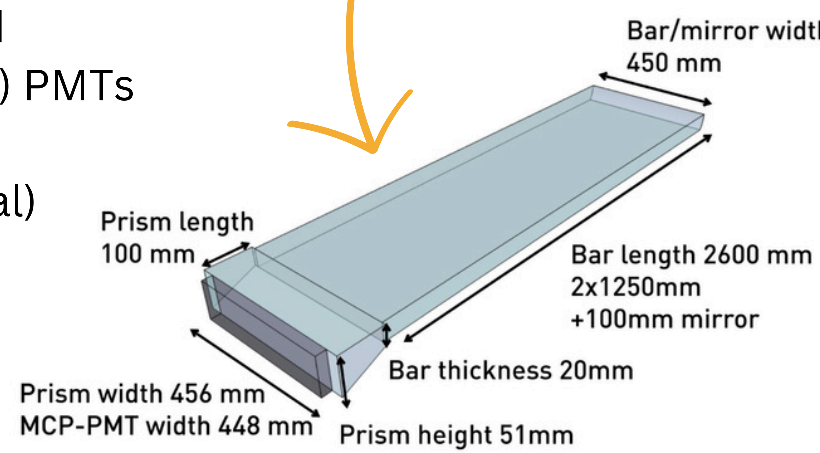
TOP stands for **Time-Of-Propagation**



It consists of 16 modules arranged around the Central Drift Chamber (CDC) in the barrel region

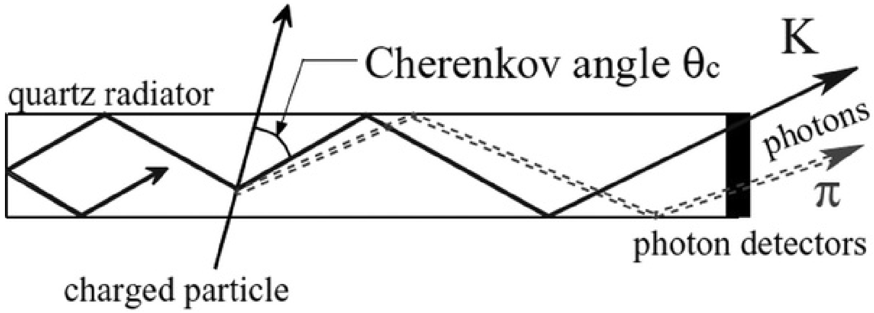
Each module consists of

- 2 quartz (n = 1.47) bars glued together
- a focusing mirror at one end
- a small expansion prism at the other end
- an array of 32 Micro Channel Plate (MCP) PMTs
  - arranged in 2 rows
  - 4 × 4 channels each (512 pixels in total)
  - transit time spread < 50 ps



TOP working principle

TOP is a ring-imaging Cherenkov device



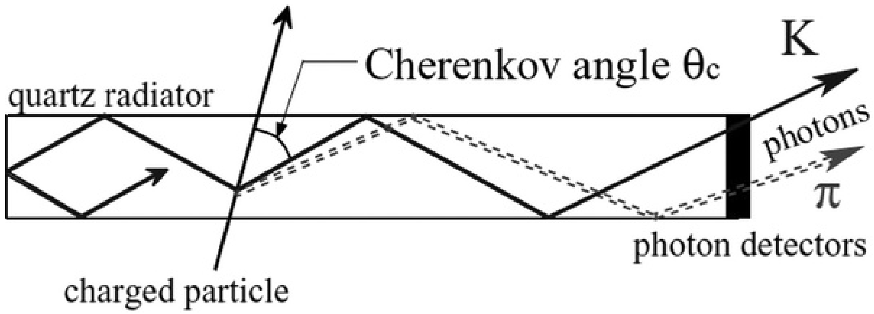
Dependence on particle mass and momentum

$$\cos\theta_C = \frac{1}{n\beta}$$

- Photons propagate within the quartz bar by internal reflection
- The 2D information of a Cherenkov ring-image is given by the **arrival time** and the **hit position** of the Cherenkov photons at the photon detector plane

TOP working principle

TOP is a ring-imaging Cherenkov device



Dependence on particle mass and momentum

$$\cos\theta_C = \frac{1}{n\beta}$$

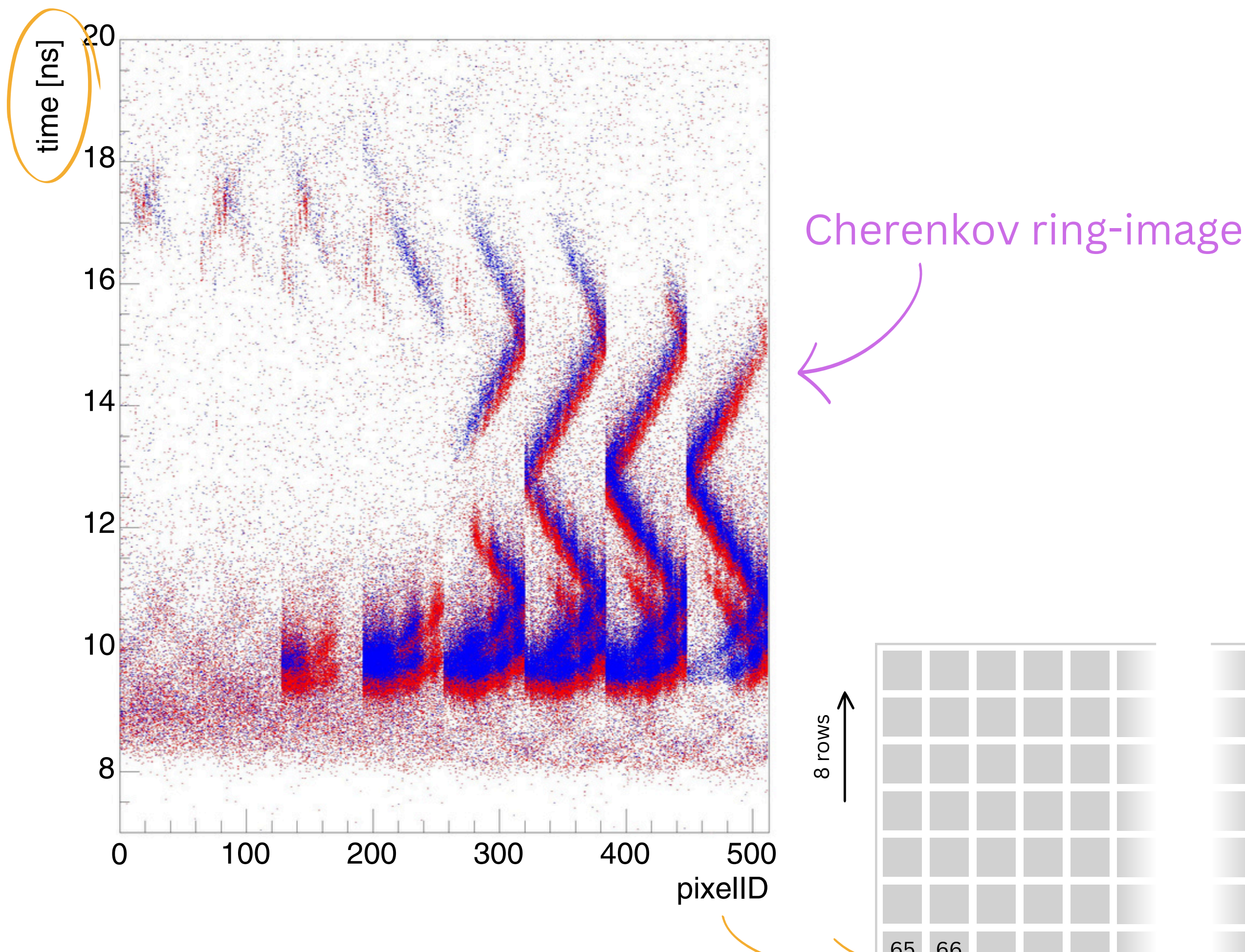
- Photons propagate within the quartz bar by internal reflection
- The 2D information of a Cherenkov ring-image is given by the **arrival time** and the **hit position** of the Cherenkov photons at the photon detector plane

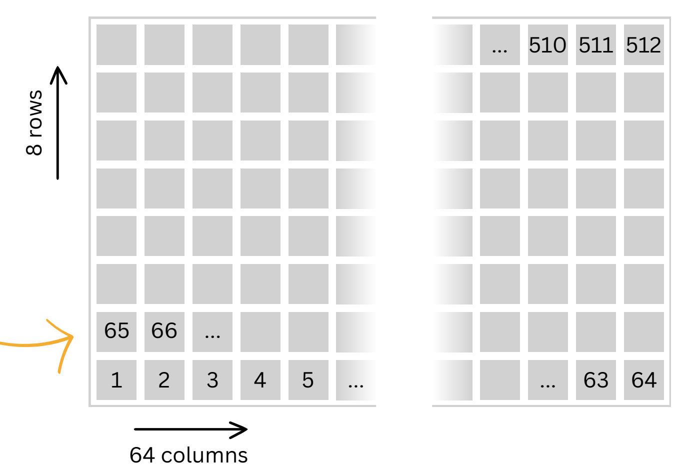
K* and *π (with same momentum, impact position and impact angle on the bar) emit Cherenkov photons at **different θc** → different path length → **different time of propagation and hit position on the photon detector plane**

- Time resolution < 100 ps for good PID performance

Photon pattern

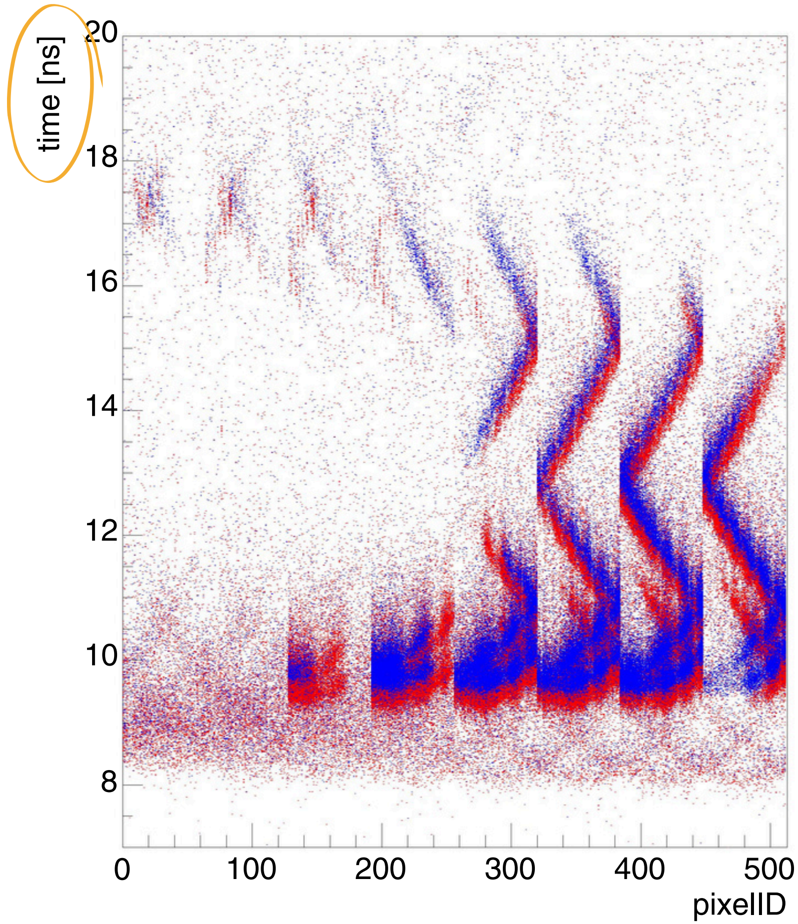
Example of measured pattern for thousands of photons, for simulated **pions** and **kaons** with p = 2.0 GeV/c, θ = 90 deg, φ = 41.5 deg



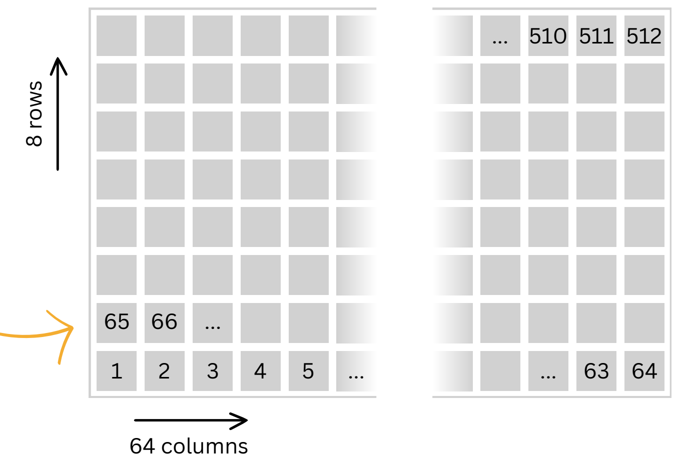


Photon pattern

Example of measured pattern for thousands of photons, for simulated **pions** and **kaons** with p = 2.0 GeV/c, θ = 90 deg, φ = 41.5 deg



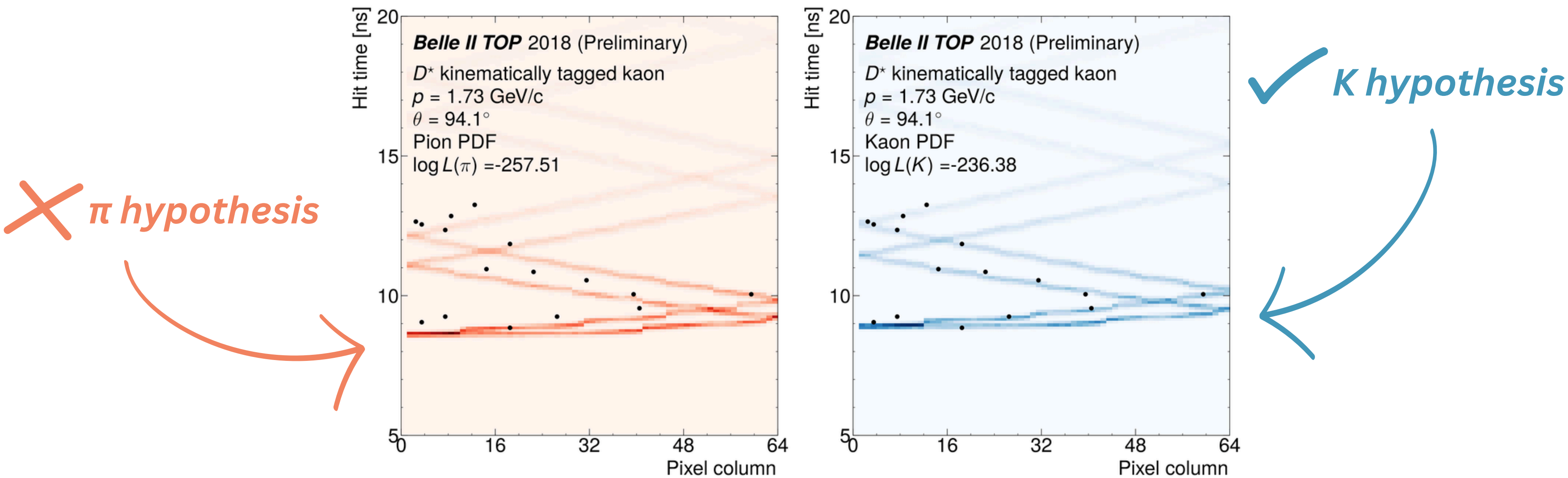
Cherenkov ring-image



Note: 20-40 photons are typically detected per event!

TOP PID Likelihood

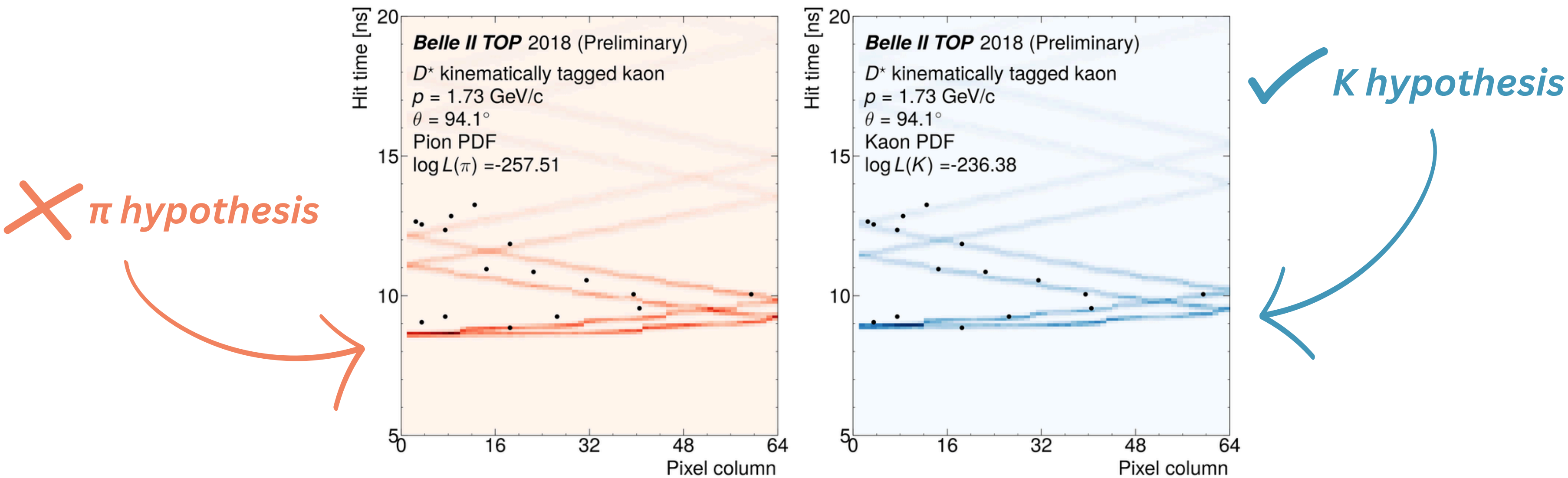
We compare the **measured 2D distribution *x* - *t*** (*x* = pixel column) (black dots) of the photon hits with the **expected pattern analytically calculated** (colored lines) for *K* and *π* hypotheses → *K* and *π* TOP likelihoods



Note: the tracking system measures particle momentum, impact position and impact angle on the TOP bar

TOP PID Likelihood

We compare the **measured 2D distribution *x* - *t*** (*x* = pixel column) (black dots) of the photon hits with the **expected pattern analytically calculated** (colored lines) for *K* and *π* hypotheses → *K* and *π* TOP likelihoods



Note: the tracking system measures particle momentum, impact position and impact angle on the TOP bar

The ratios of the TOP likelihoods are used to assign TOP PID probabilities

$$P_K = \frac{\mathcal{L}_K}{\mathcal{L}_K + \mathcal{L}_\pi} \qquad P_\pi = 1 - P_K$$ (binary PID)

TOP only binary PID gives 85% of *K*-ID efficiency with 10% *π* → *K* mis-ID rate

TOP PID with Machine Learning (ML)

Why ML?

- The calculation of analytical TOP likelihoods requires a highly detailed simulation of all components of the TOP
- The simulation cannot perfectly reproduce the geometry of the TOP modules and the differences between them

→ With ML we can train using the data measured by the TOP (and the tracking system), so we do not need to simulate the TOP counter very precisely

TOP PID with Machine Learning (ML)

Why ML?

- The calculation of analytical TOP likelihoods requires a highly detailed simulation of all components of the TOP
- The simulation cannot perfectly reproduce the geometry of the TOP modules and the differences between them

→ With ML we can train using the data measured by the TOP (and the tracking system), so we do not need to simulate the TOP counter very precisely

Training flow

Training on particle gun data → Training on generic MC data → Training on experimental data

Past results

→ Boosted Decision Tree (BDT) for K and π separation
  - Improved PID performance for particle gun events
  - Performance much decreased for generic MC events

(Very) Preliminary results with CNN

Proposal

→ Use a Convolutional Neural Network (CNN) for K and π separation
- CNNs are highly effective for pattern recognition

(Very) Preliminary results with CNN

Proposal

Use a Convolutional Neural Network (CNN) for *K* and *π* separation

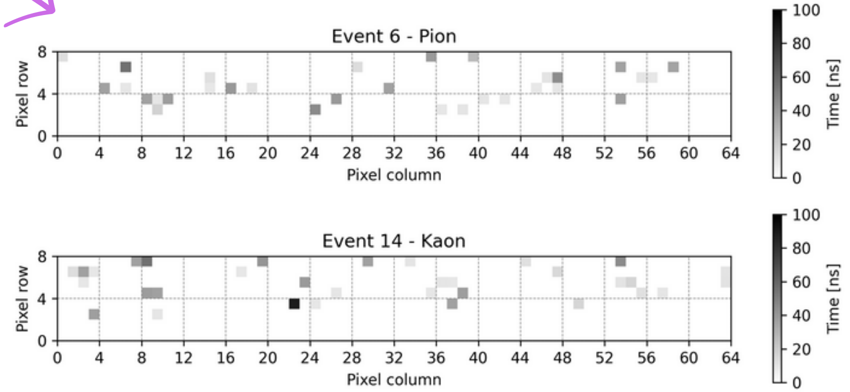
- CNNs are highly effective for pattern recognition

Simulated particle gun events

- ~200k particles (*K+* and *π+*) from the interaction point
- *p* = 2.0 GeV/c, *θ* = 90 deg, *φ* = 41.5 deg

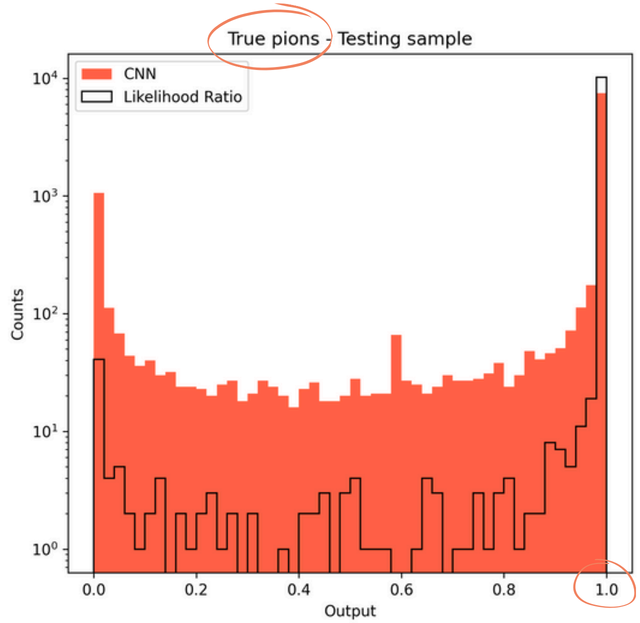
CNN model

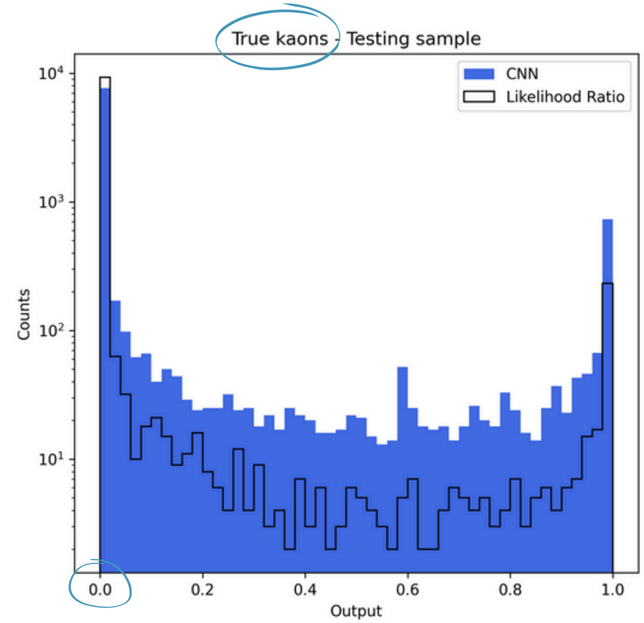
- Supervised learning
- Input:
  - images (8 x 64 x 1) (row x col x time)
- Network:
  - convolutional layers + dense layers
- Output:
  - binary classification: 1 for *π*, 0 for *K*
- Training sample (90%) + Testing sample (10%)
- Training for 1000 epochs



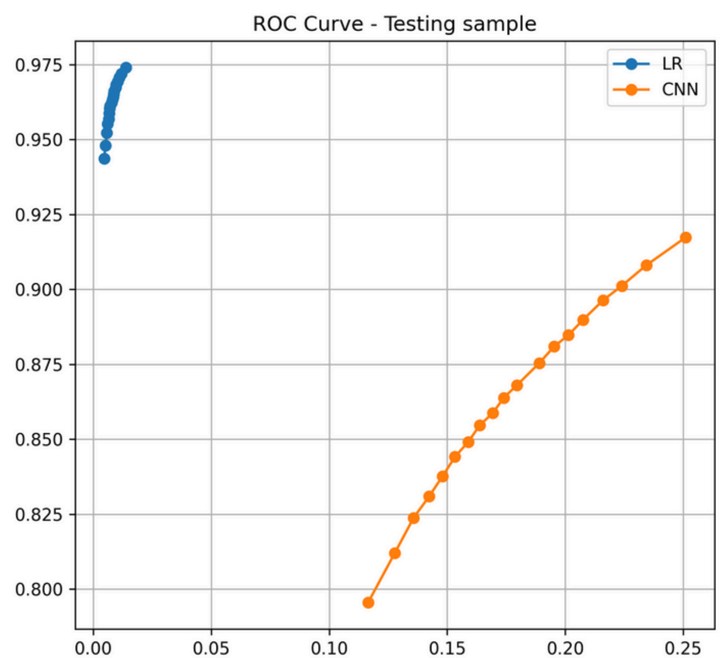
(Very) Preliminary results with CNN

Output:
- 1 for π
- 0 for K





K-ID efficiency



$\pi \to K$ mis-ID rate

The network is able to separate pions and kaons, but there is no improvement in PID performance (for now!)

Conclusion

- Belle II PID performance is better than Belle one
- TOP is a new concept Cherenkov detector type for PID that relies on precise timing of individual photons
- TOP only binary PID gives 85% of K-ID efficiency with 10% $\pi \rightarrow K$ mis-ID rate

Can we do better?

- We are exploring ML techniques for the TOP counter with the aim of improving PID performance
- Improved PID performance with BDT for particle gun events
- **Work in progress with CNN for particle gun events**
  - First tentative with images (row x col x time) as input
  - It is planned to use images (col x time) as input
- When we move to real data, we will have to take into account the dependence of PID performance on background conditions

Conclusion

- Belle II PID performance is better than Belle one
- TOP is a new concept Cherenkov detector type for PID that relies on precise timing of individual photons
- TOP only binary PID gives 85% of K-ID efficiency with 10% $\pi \rightarrow K$ mis-ID rate

Can we do better?

- We are exploring ML techniques for the TOP counter with the aim of improving PID performance
- Improved PID performance with BDT for particle gun events
- **Work in progress with CNN for particle gun events**
  - First tentative with images (row x col x time) as input
  - It is planned to use images (col x time) as input
- When we move to real data, we will have to take into account the dependence of PID performance on background conditions