Slow pion identification at the Belle II PXD with machine learning

J. Bilk, S. Käs, S. Lange, E. Prencipe, T. Schellhaas
II. Physikalisches Institut, Justus-Liebig-Universität Giessen

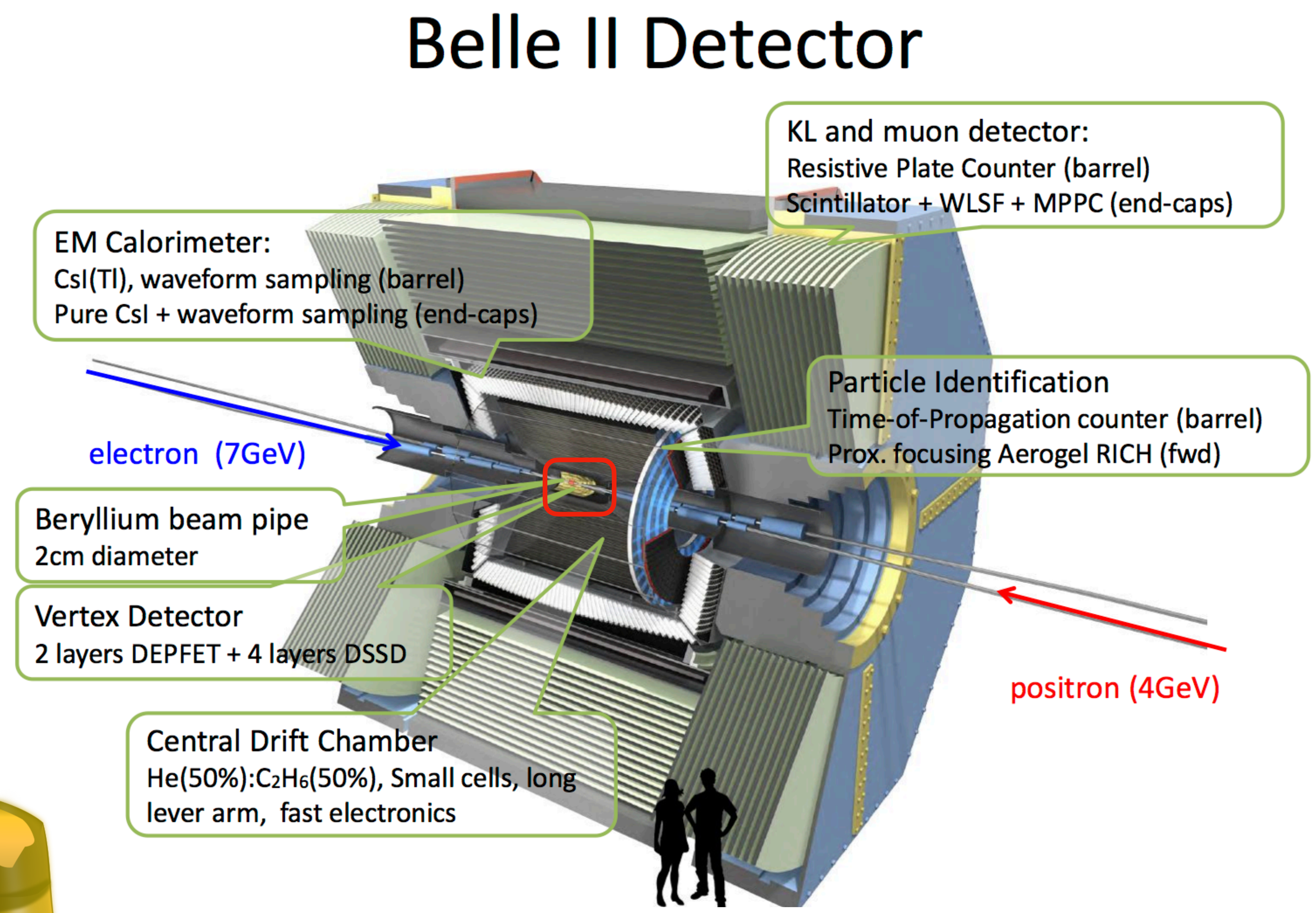
DPG Talk, 8th March 2024

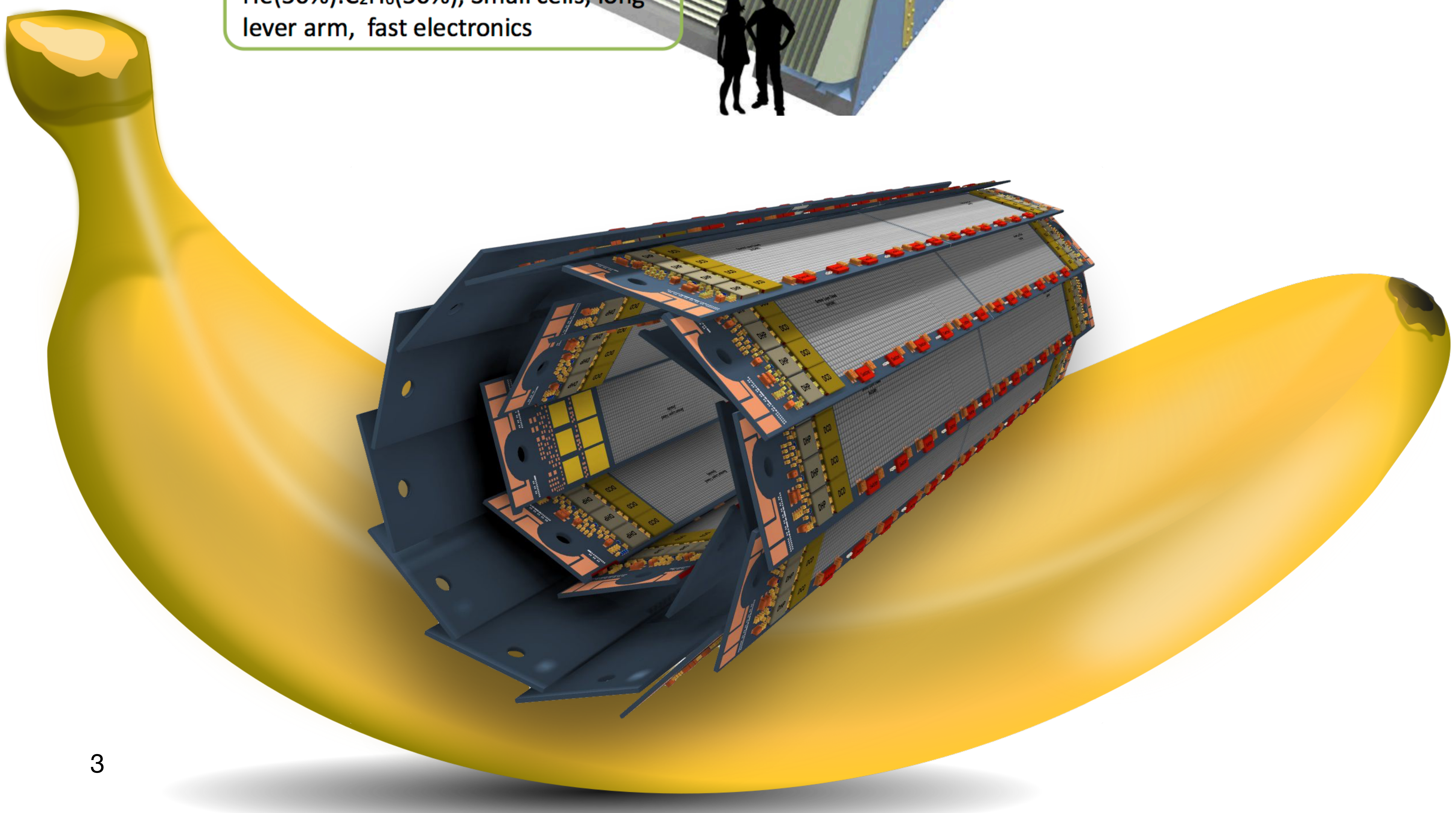
Supported by BMBF under (05H21RGKB1)

Introduction

Belle II Detector

- B-meson factory at Tsukuba, Japan
- e^+e^- collisions at 10.58 GeV
- our group is focused on the data acquisition of the pixel detector
- the pixel detector (PXD) consists of two layers of DEPFET silicon pixels
- they have a radius of 14mm and 22mm and are made up of 8 and 12 modules respectively
- each module has 256 × 768 pixels (in total 8 mil. pixels)





Slow Pions

- signal:
    - semi-leptonic decays of B mesons to charged D* mesons, which are sensitive to new physics
    - charged D* mesons decay to slow pions by $D^{*+(-)} \to D^0 \ pi^{+(-)}$
    - due to the low momentum often do not reach drift chamber, and therefore do not generate a reconstructed track

idea: identify slow pions based on cluster properties in the PXD

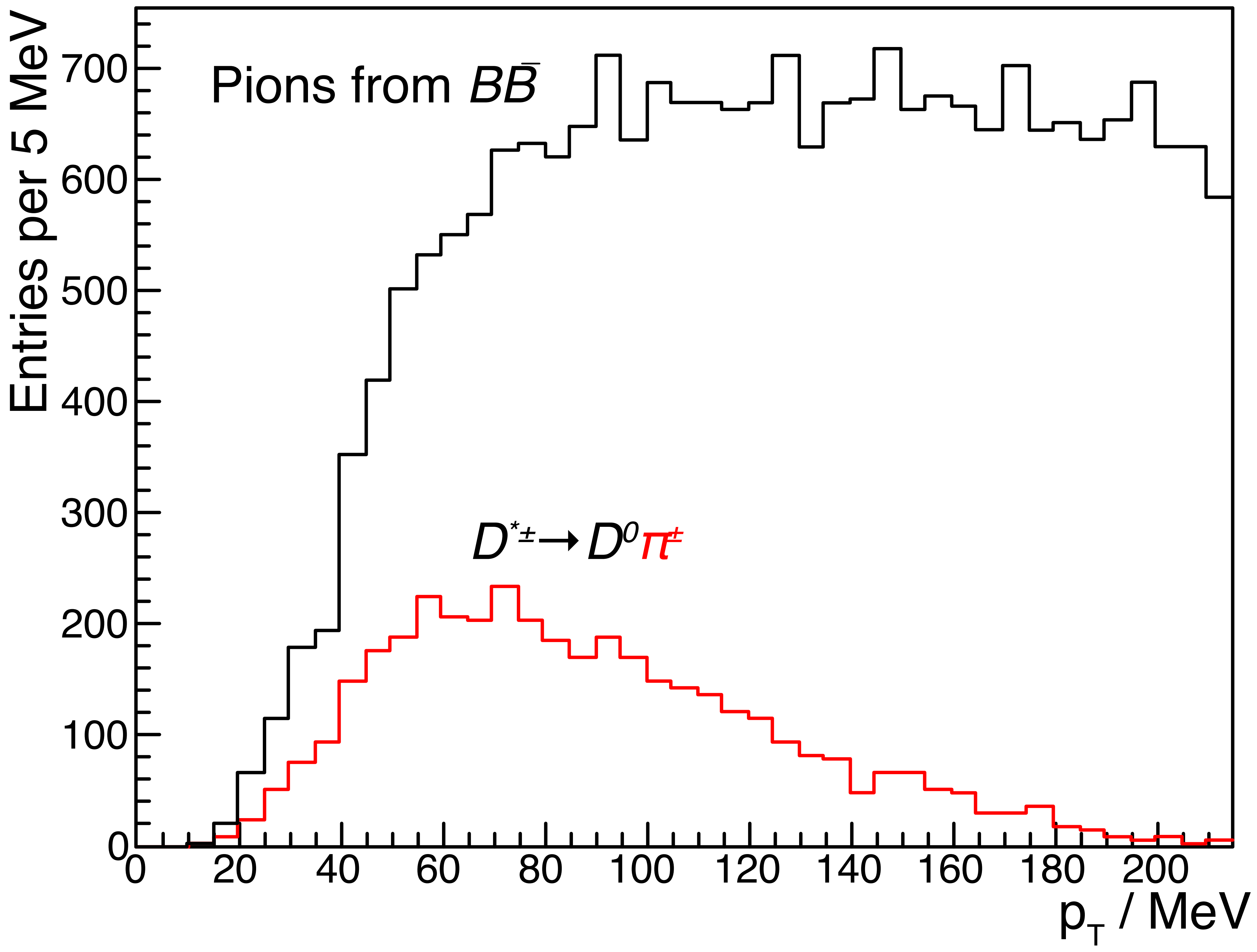
- background:
    - electrons from beam background and QED processes such as $e^+e^- \to e^+e^-e^+e^-$

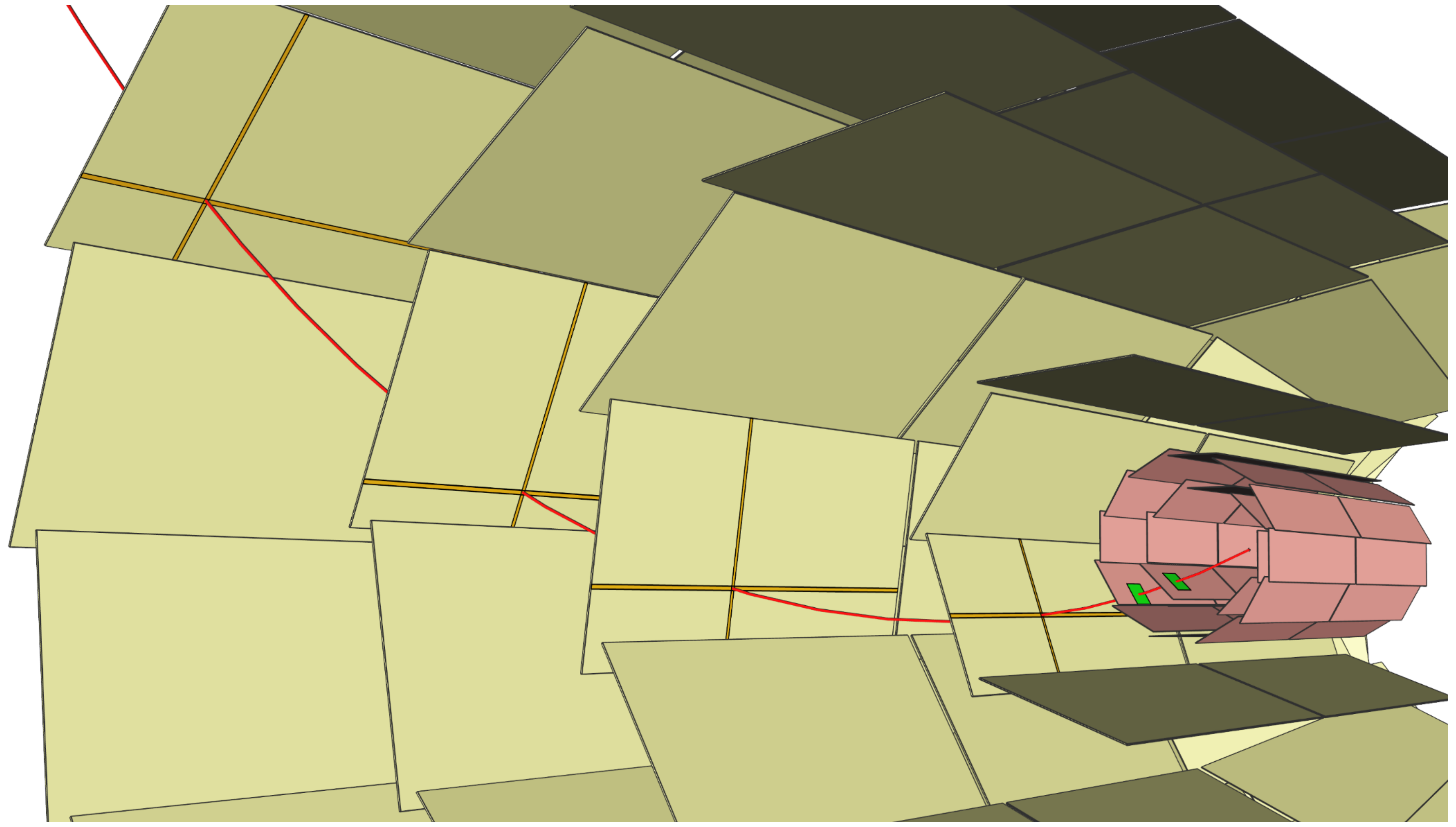
$$B^0 \to D^{*-}(2010) l^+ \nu_l$$
$$\hookrightarrow \bar{D}^0 \pi^-_{slow}$$

$$\bar{B}^0 \to D^{*+}(2010) l^- \bar{\nu}_l$$
$$\hookrightarrow D^0 \pi^+_{slow}$$

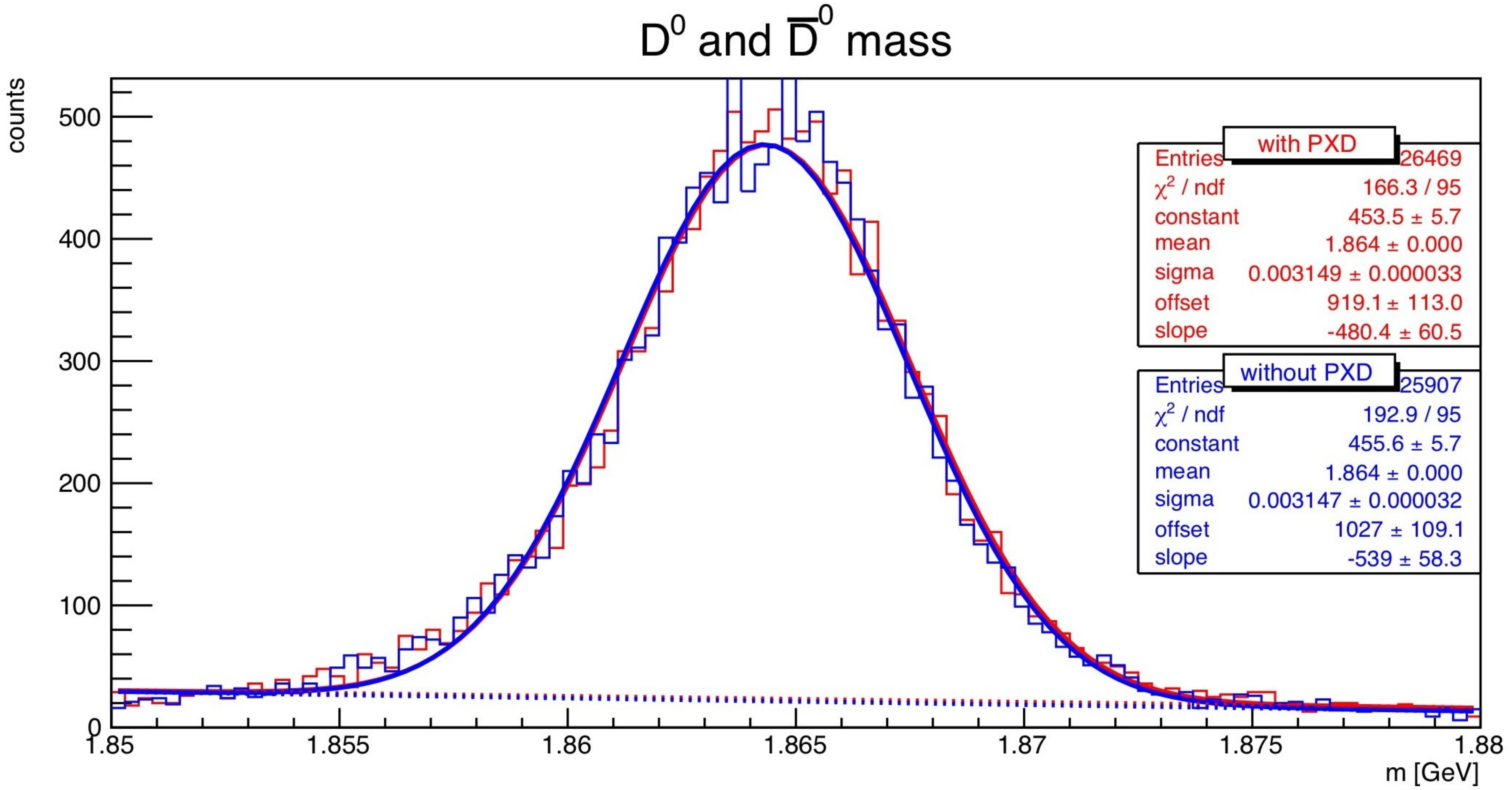
Slow Pions

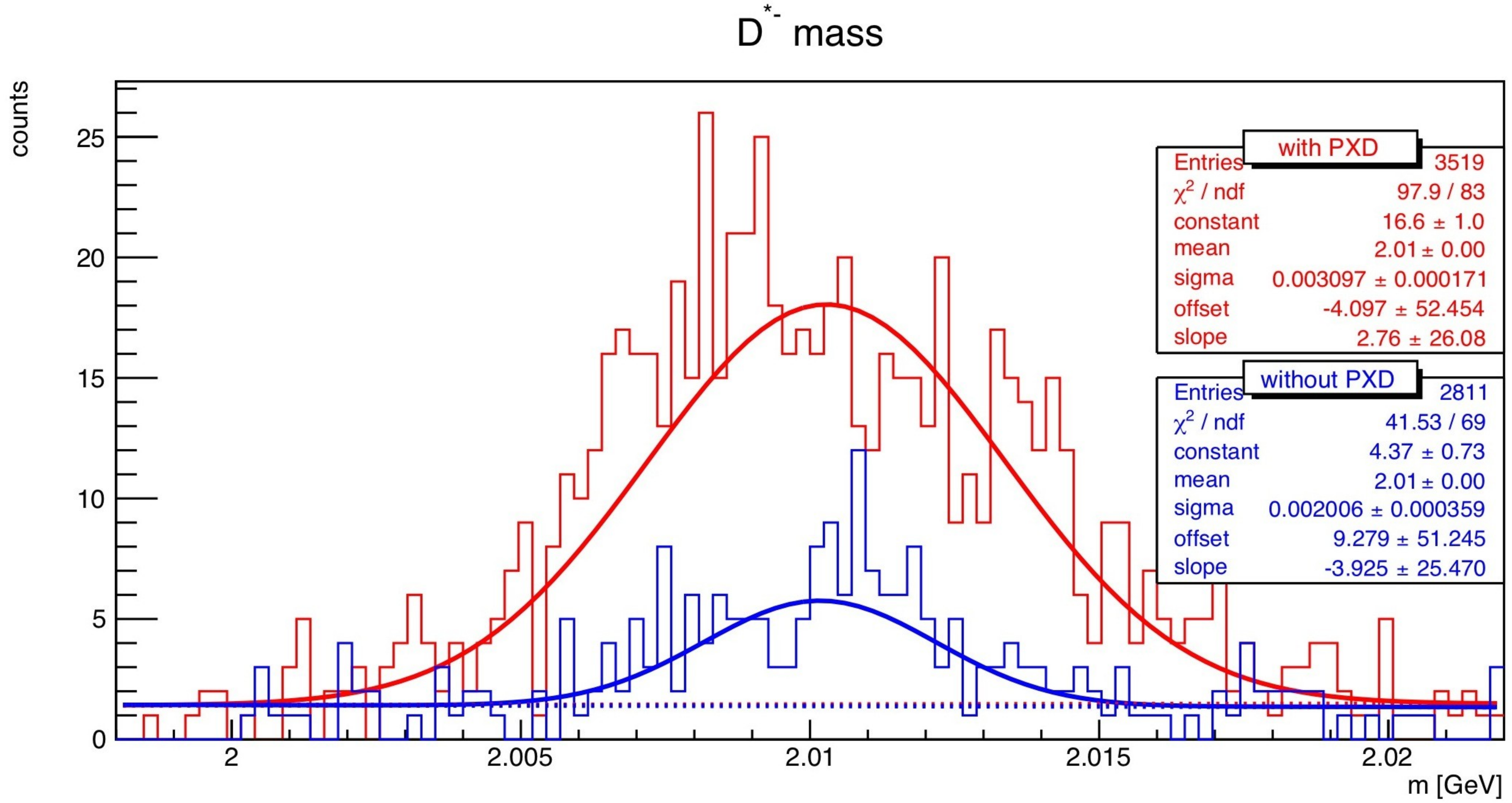
- due to the low momentum often do not reach drift chamber
  - no reconstructed track on trigger system
  - no region-of-interest generated on trigger system
  - PXD data are deleted on ONSEN system, before reaching data storage
- long term project: ML decision online on FPGA (cluster rescue)





D0 mass & D*- mass





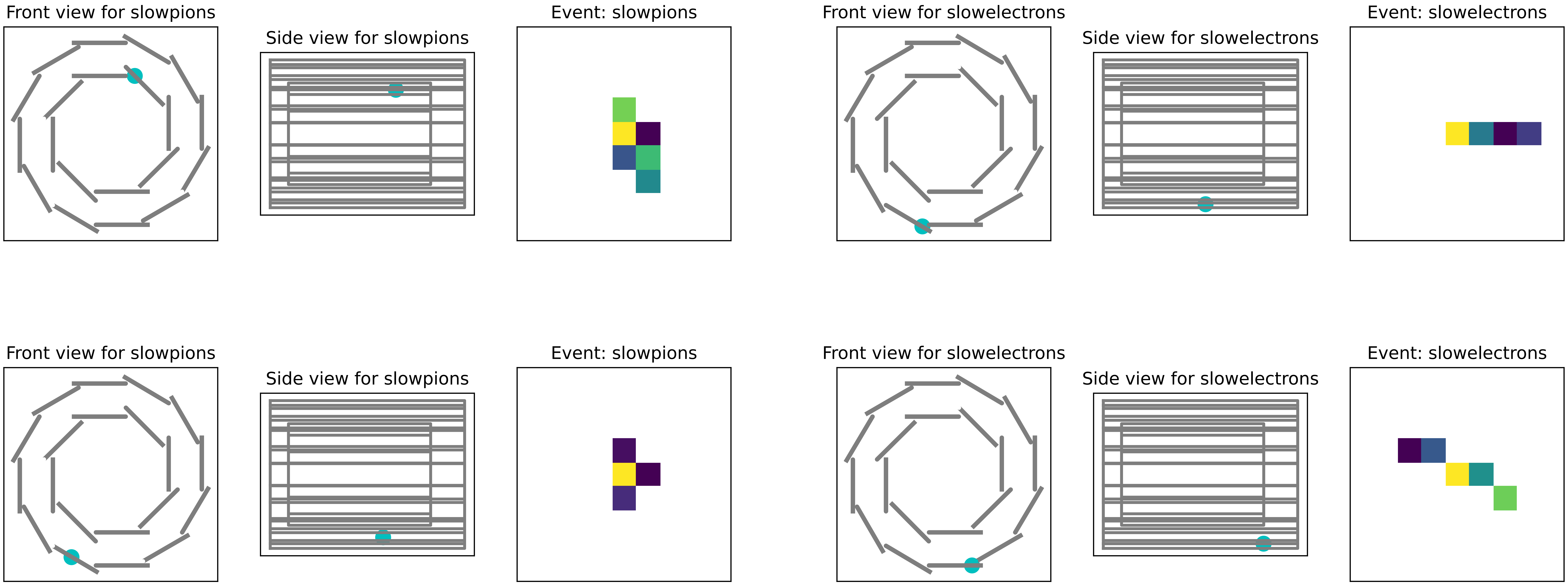
10k simulated events
- Each event contains 1 D0 and 1 D0 mesons from B0
- Additional D0 or D0 from B0 possible
- With PXD: 11932 events, without PXD: 11980 events → no significant difference found!

With 3σ-cut on D0-mass

	Mean:	Yields:
• with PXD:	2.01 GeV (fixed)	537±25
• without PXD:	2.01 GeV (fixed)	92±10

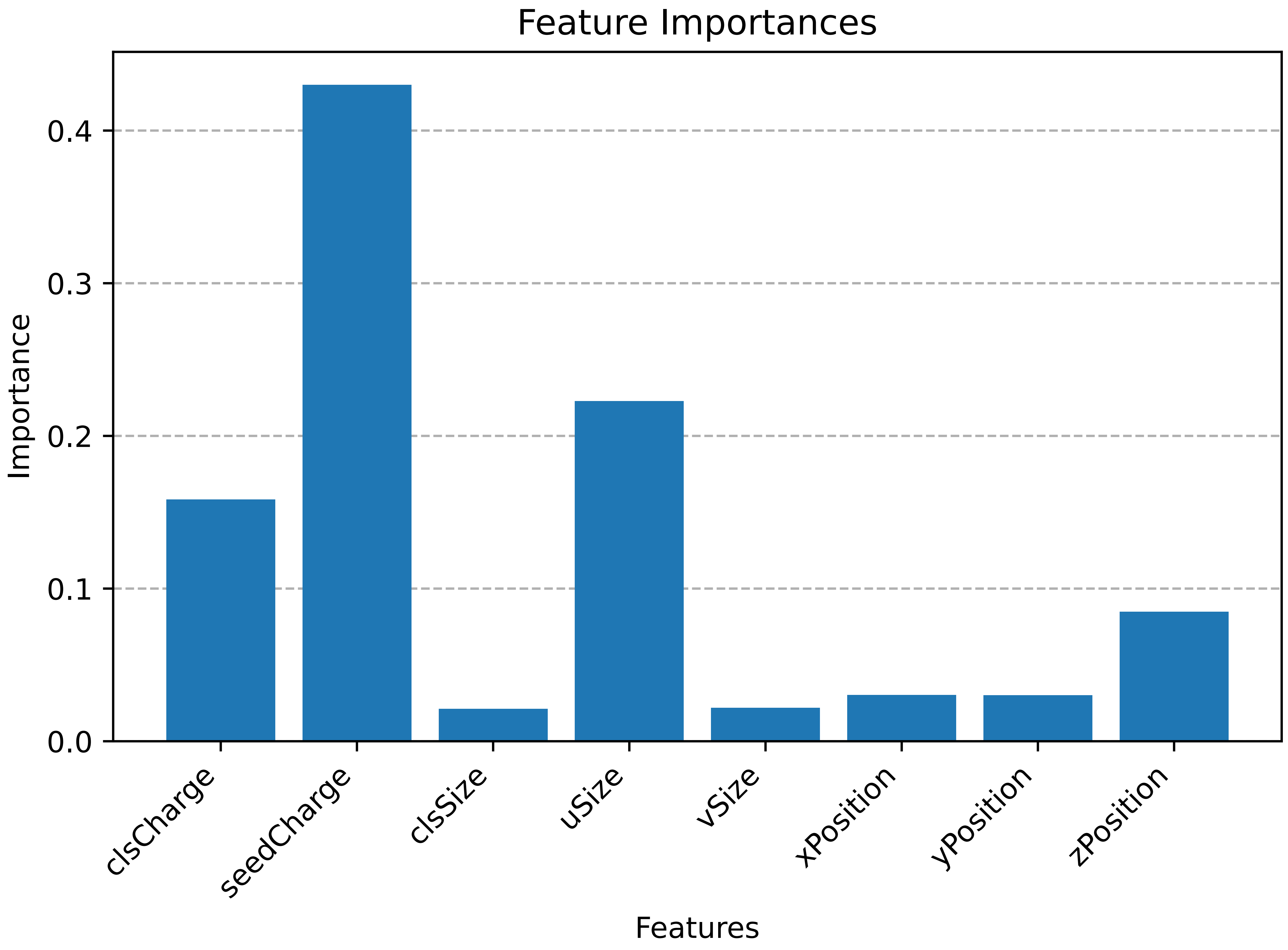
A Look at the Events

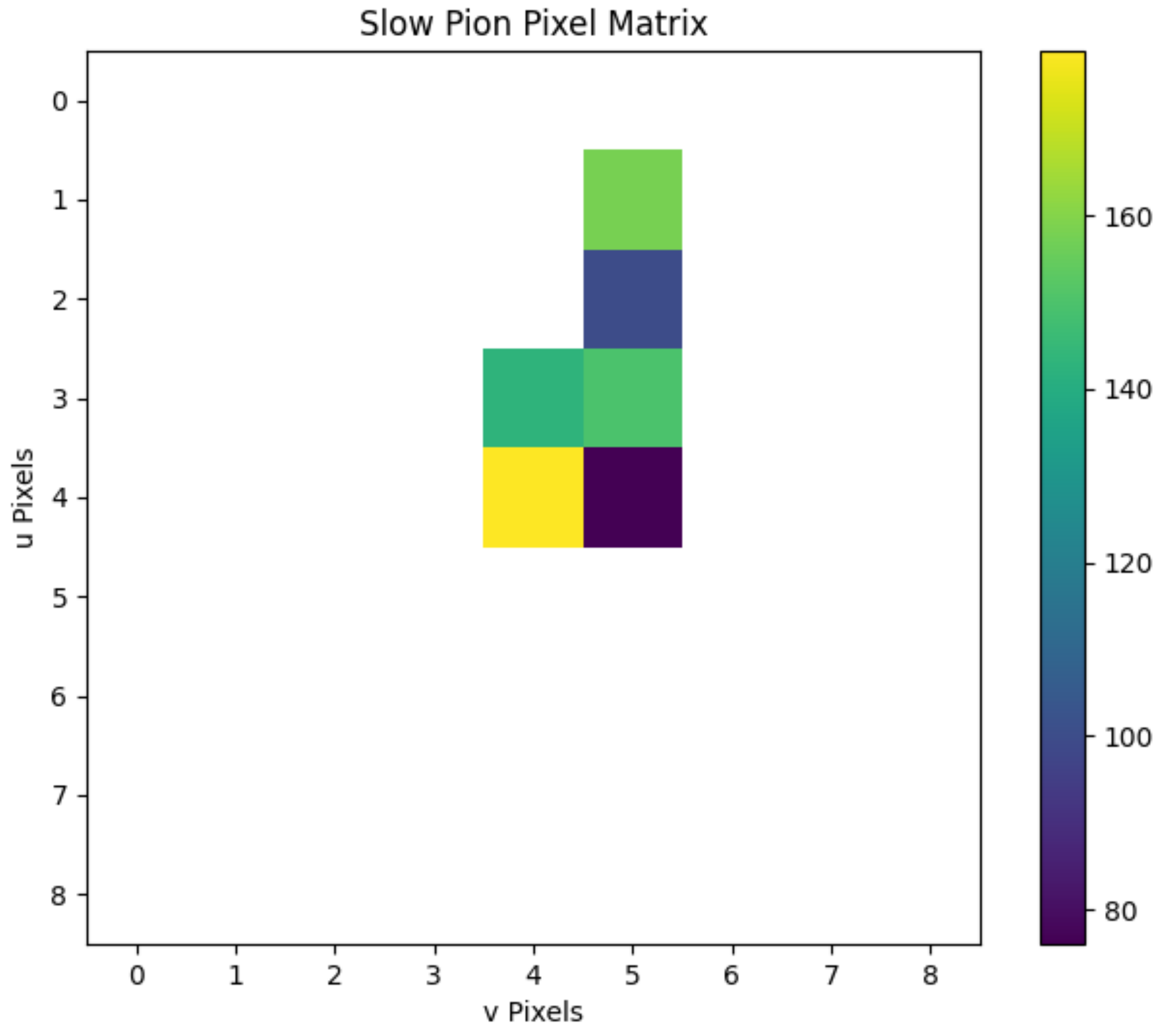
PXD Cluster Examples



Slow Pions vs Electrons

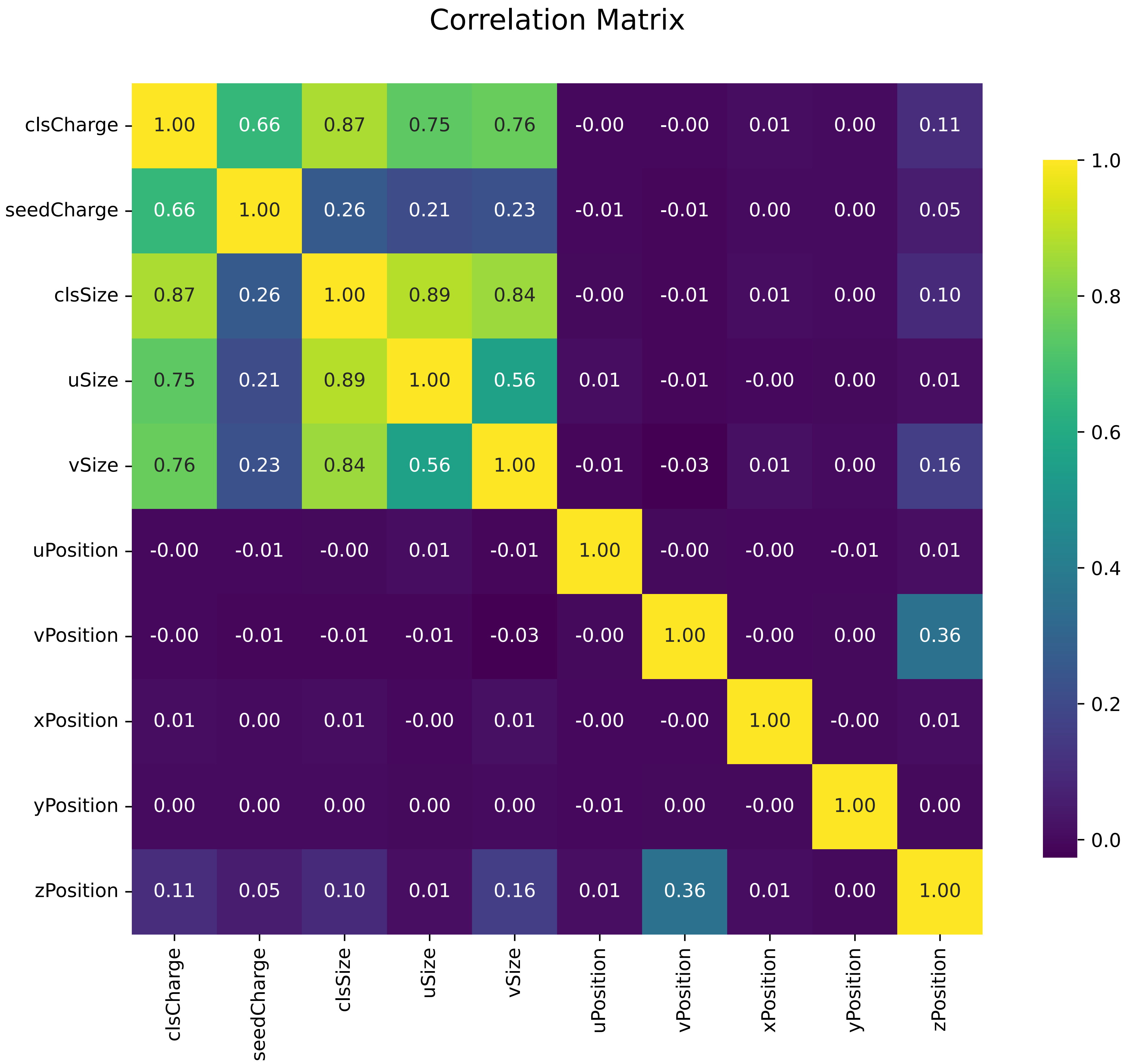
Feature Importance

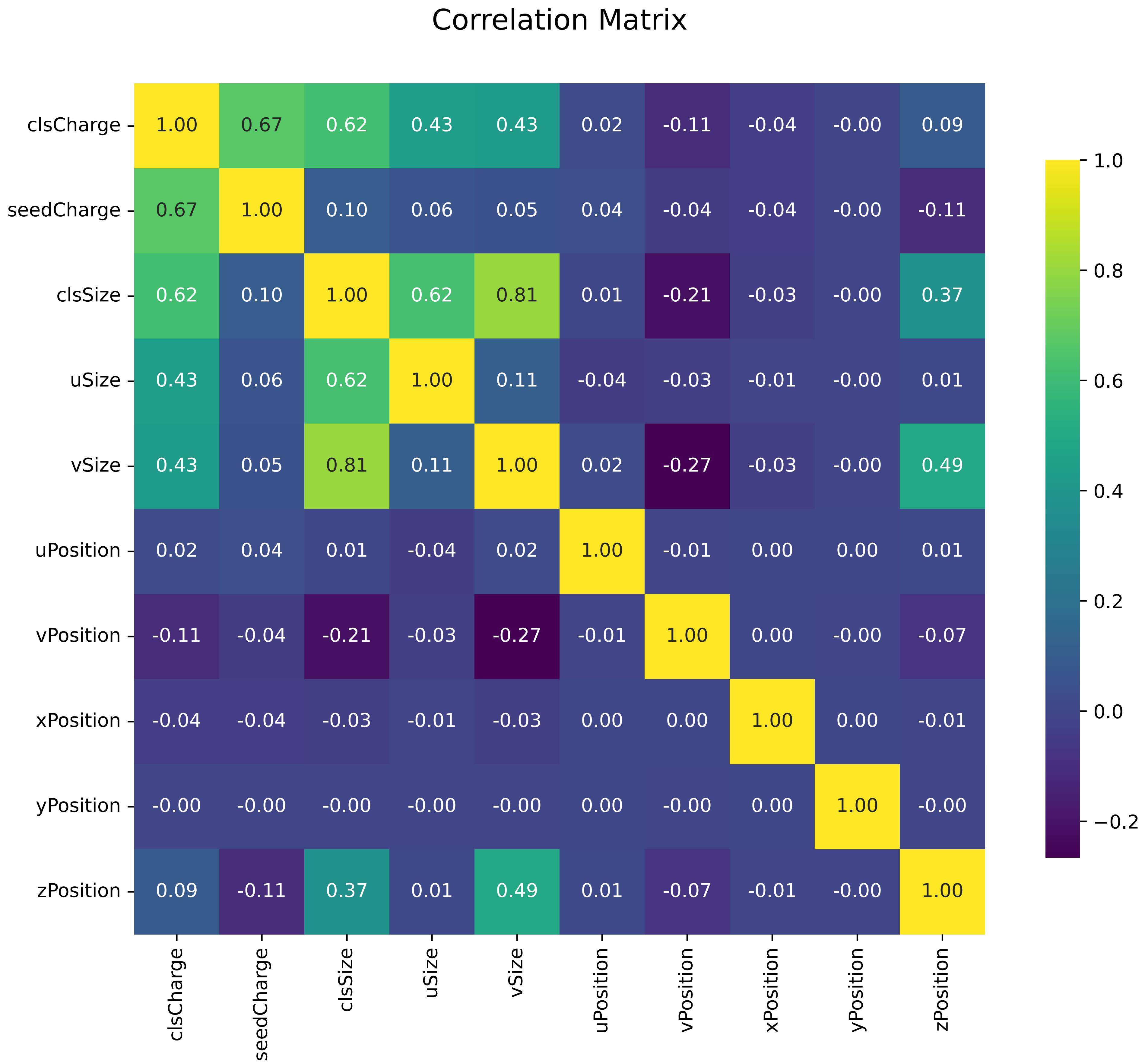




Slow Pions vs Electrons

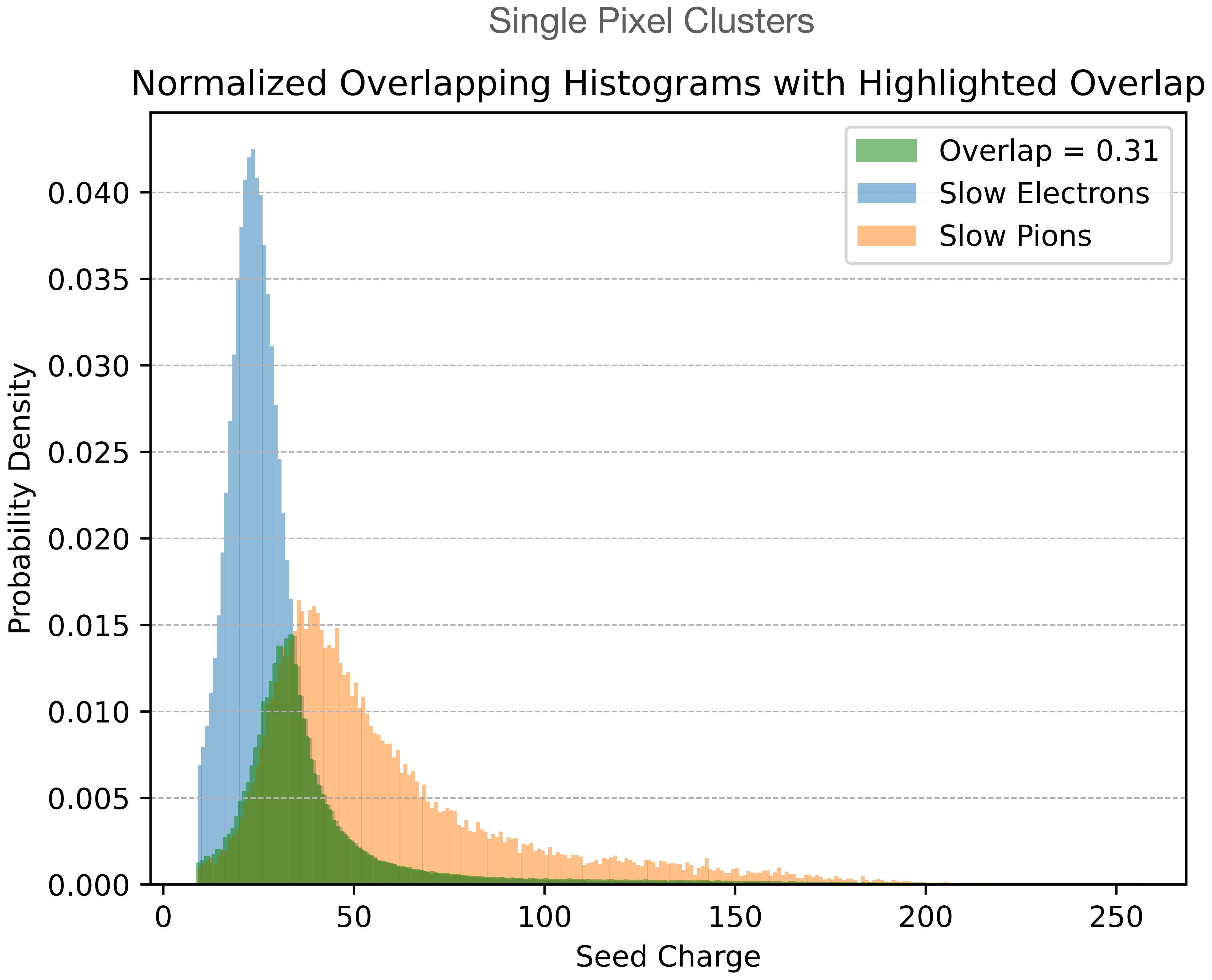
Correlation Matrices

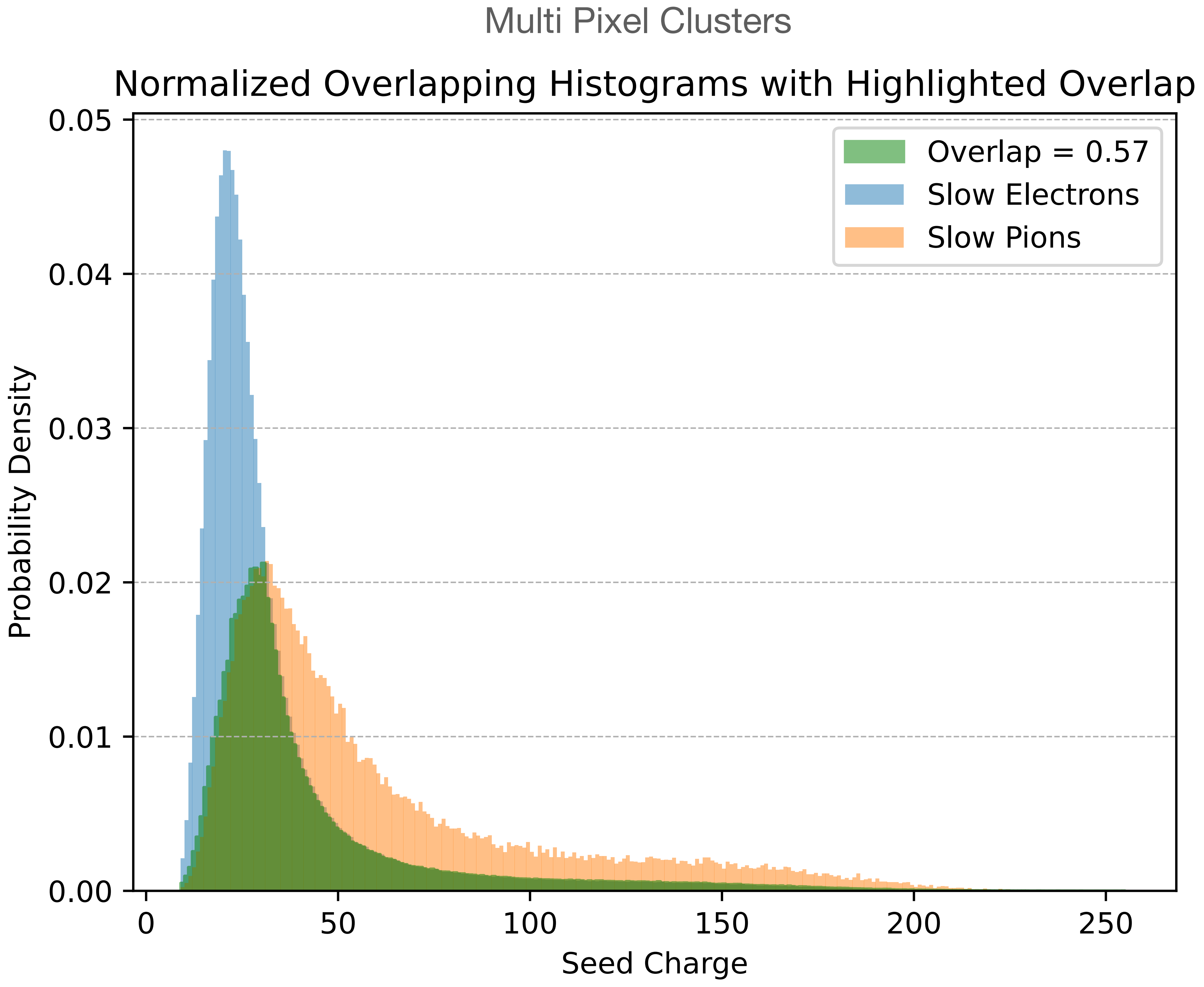




Slow Pions vs Electrons

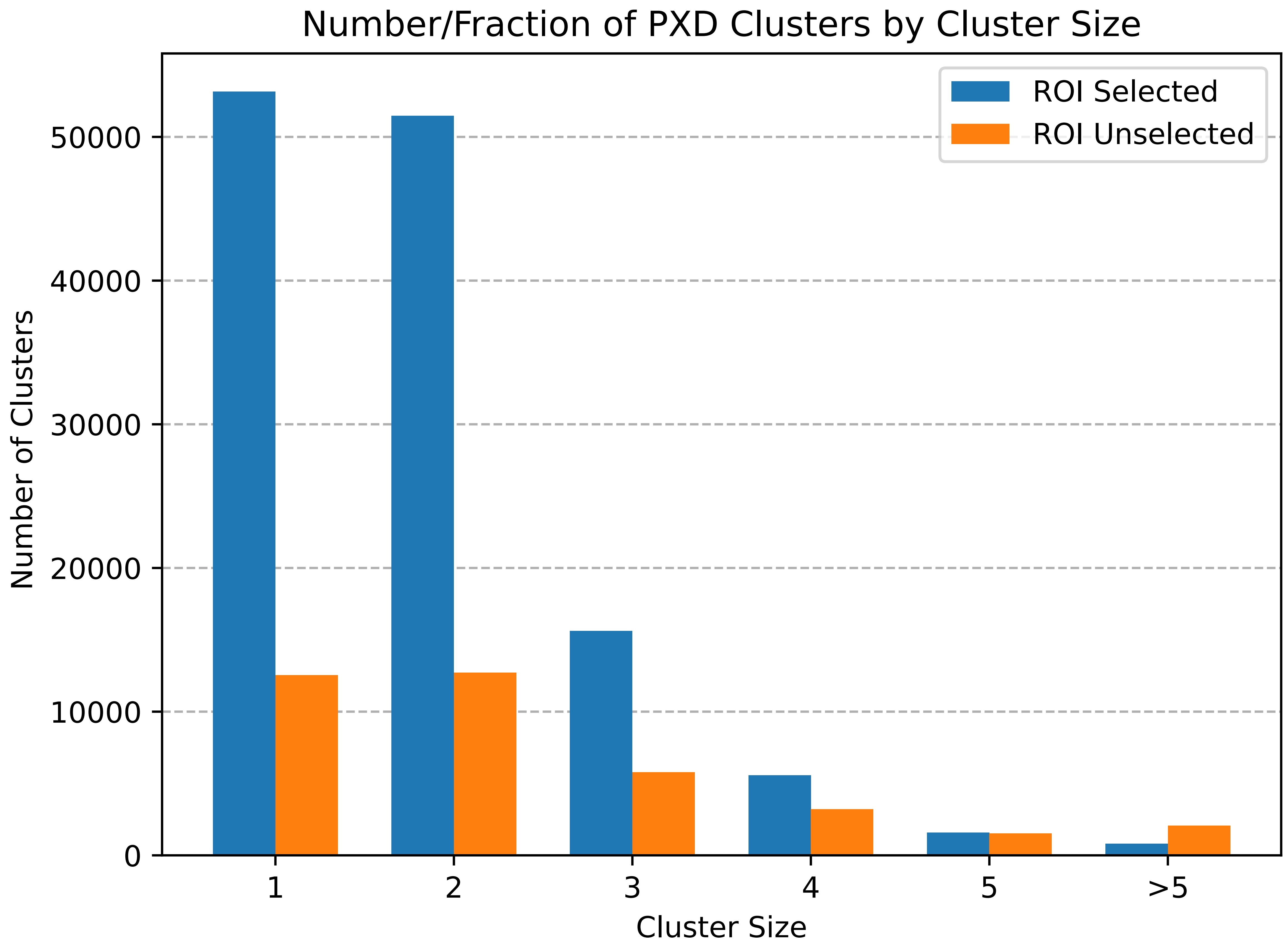
Seed Charge Histogram

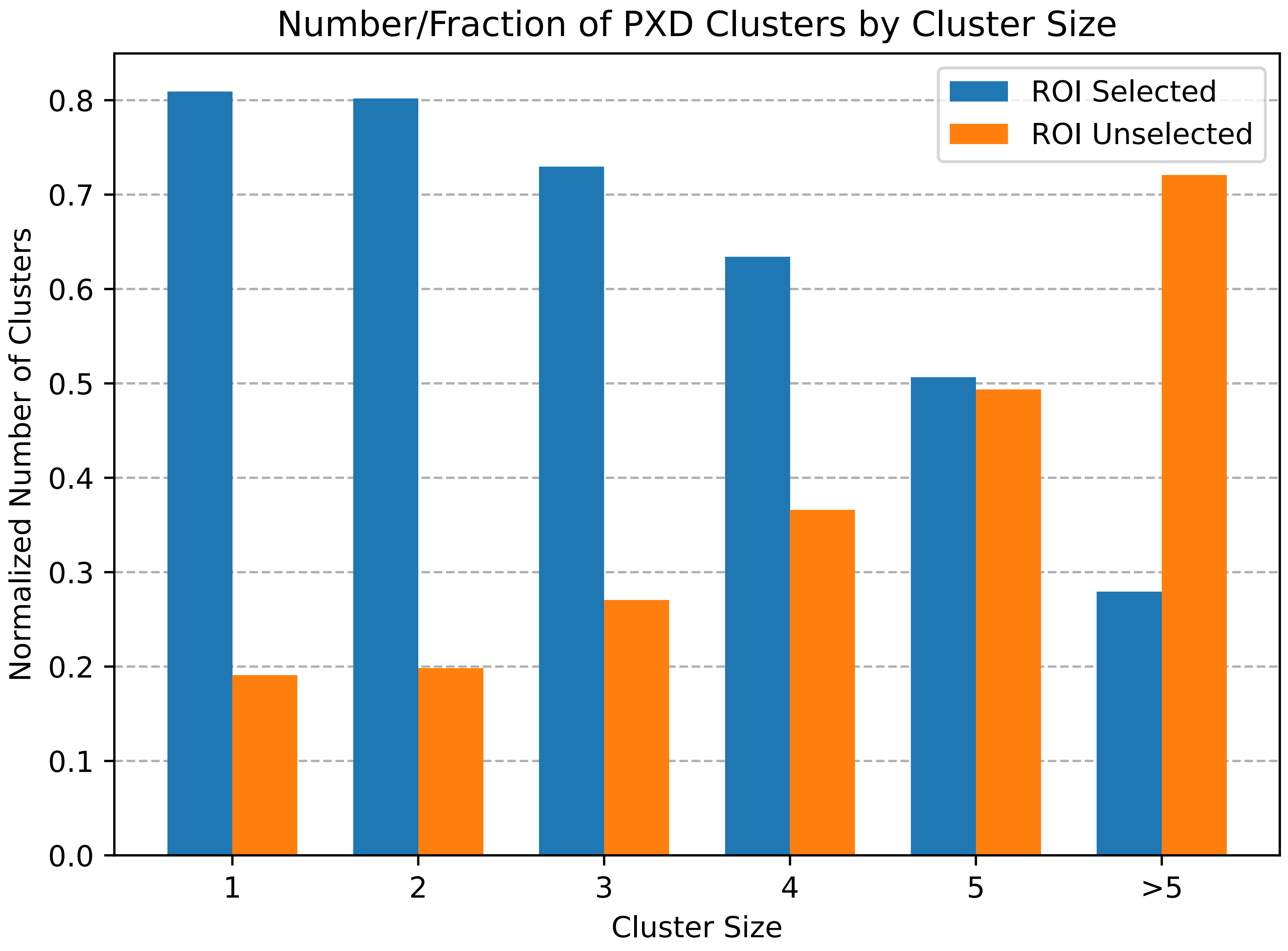




Cluster size impact on ROI selection

Cluster Size



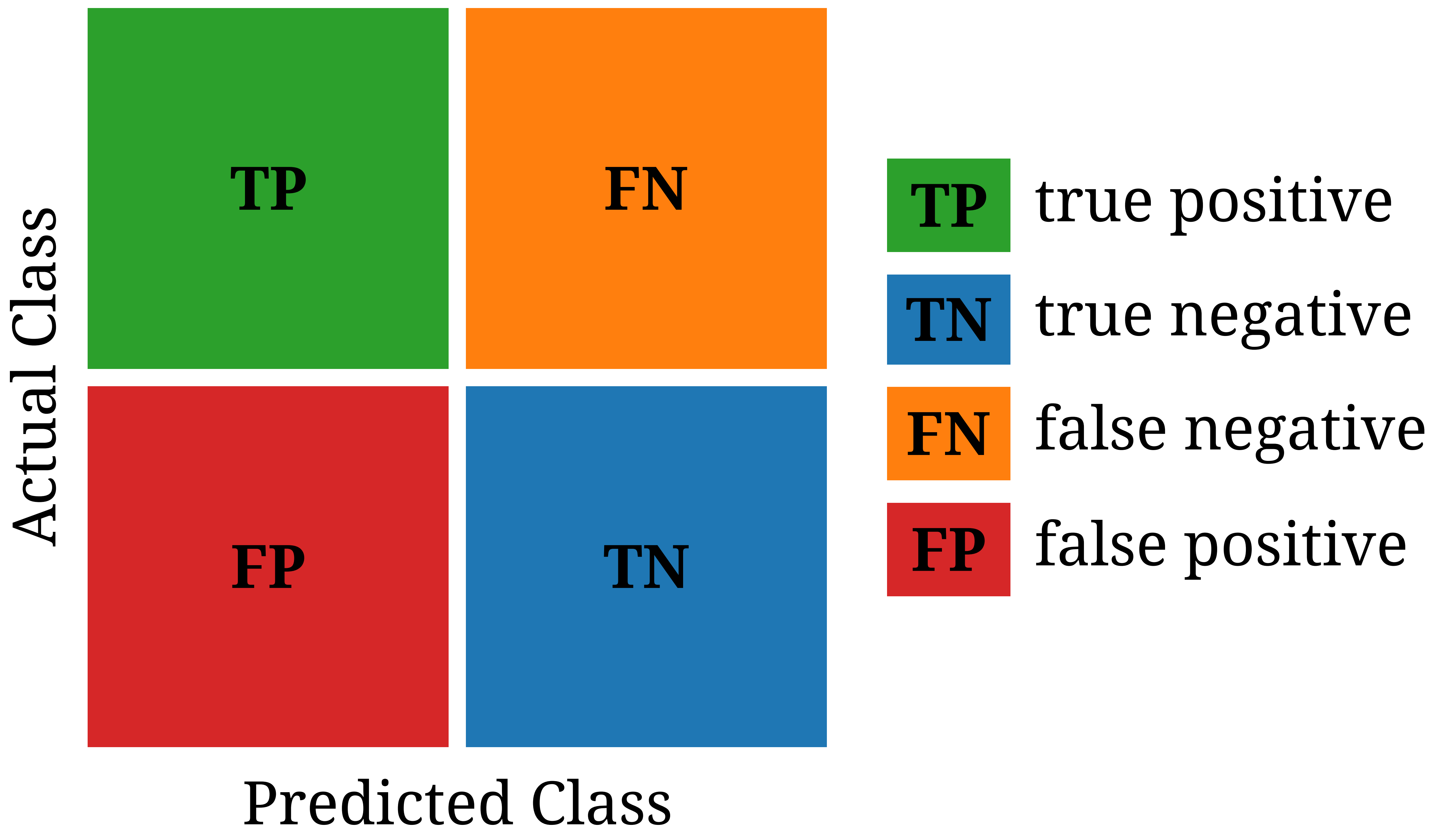


Results

The Confusion Matrix

Binary Case

- we have two classes
- TP, TN, FN and FP can be read directly from the matrix



Evaluating Performance Machine Learning Algorithm

- Sensitivity

  Number of events correctly identified by the NN as a signal (TP), divided by all real signal events (TP + FN) is also called "efficiency" in particle physics

- Precision

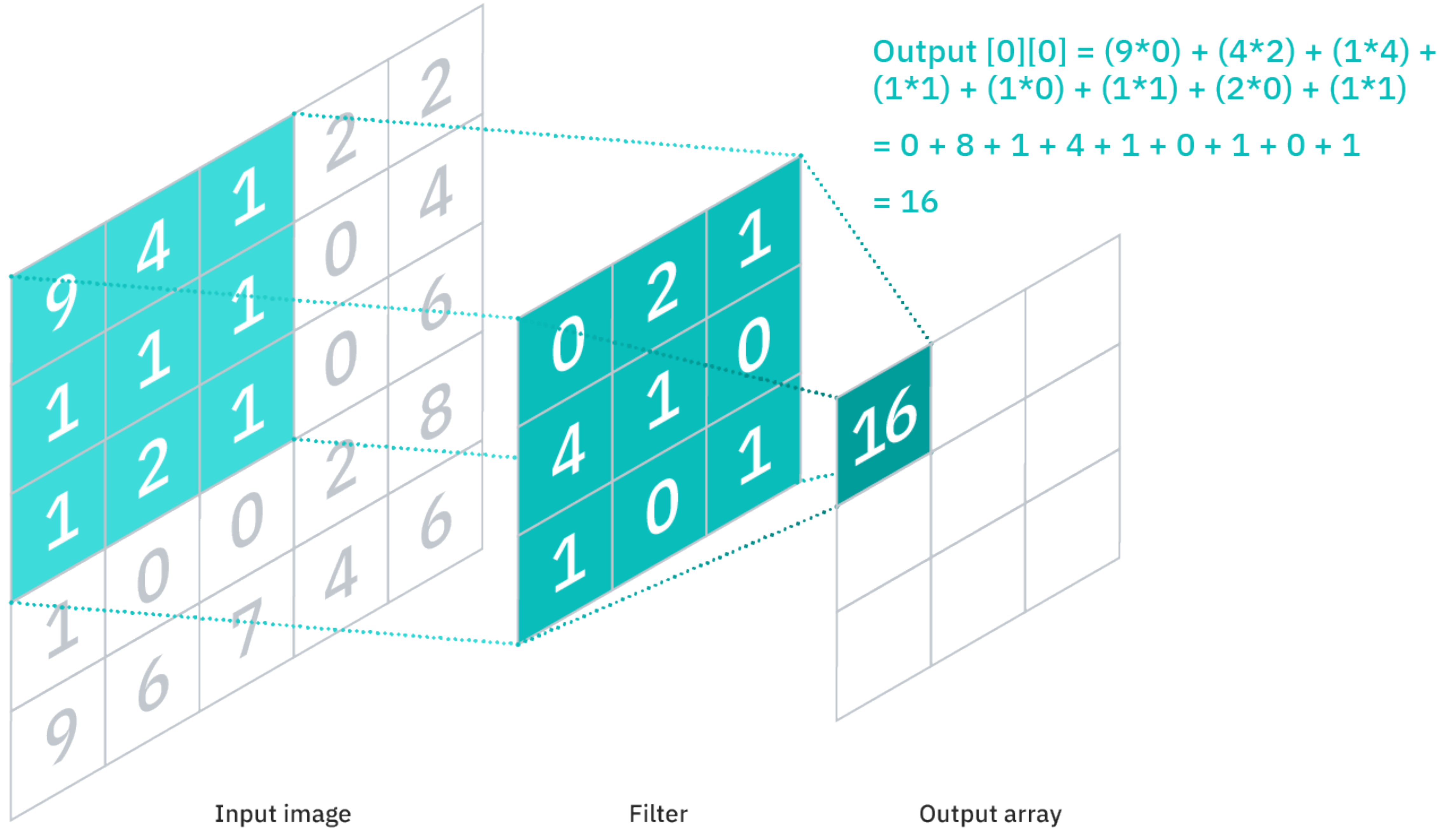
  Number of events correctly identified by the NN as a signal (TP), divided by all events identified by the NN as a signal (TP + FP) is also called "purity" in particle physics

- Specificity or Selectivity

  Number of events correctly identified by the NN as background (TN), divided by all real background events (TN + FP) is called "background rejection" in particle physics

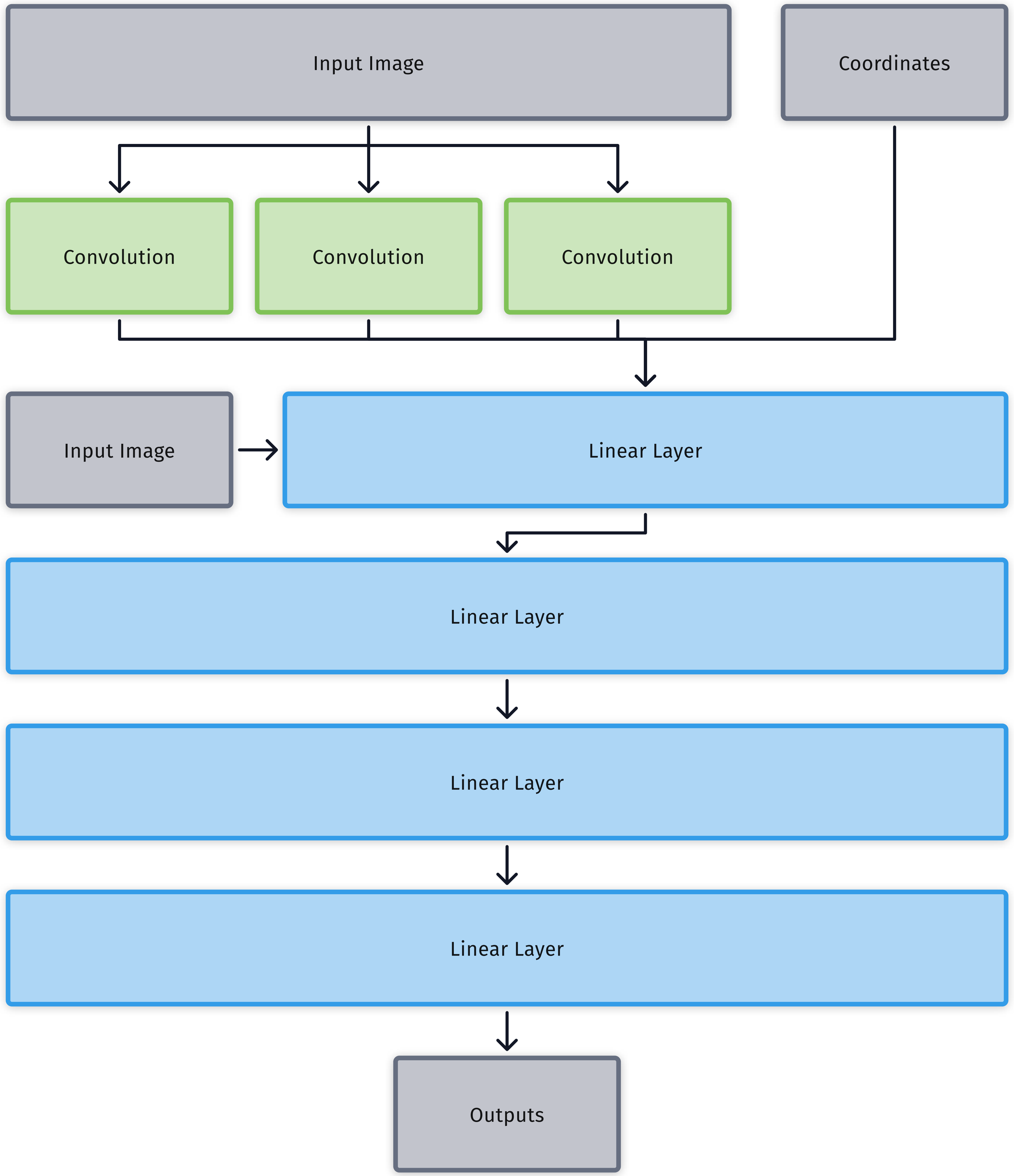
CNN

- Master thesis Johannes Bilk
  Precision (Efficiency) 82 %
- Sensitivity (Purity) 81%
- excluding 1-pixel cluster reduces precision by 3%
- multiclass vs. binary (one NN vs. many NN, all particle species) reduces precision by factor 2



CNN

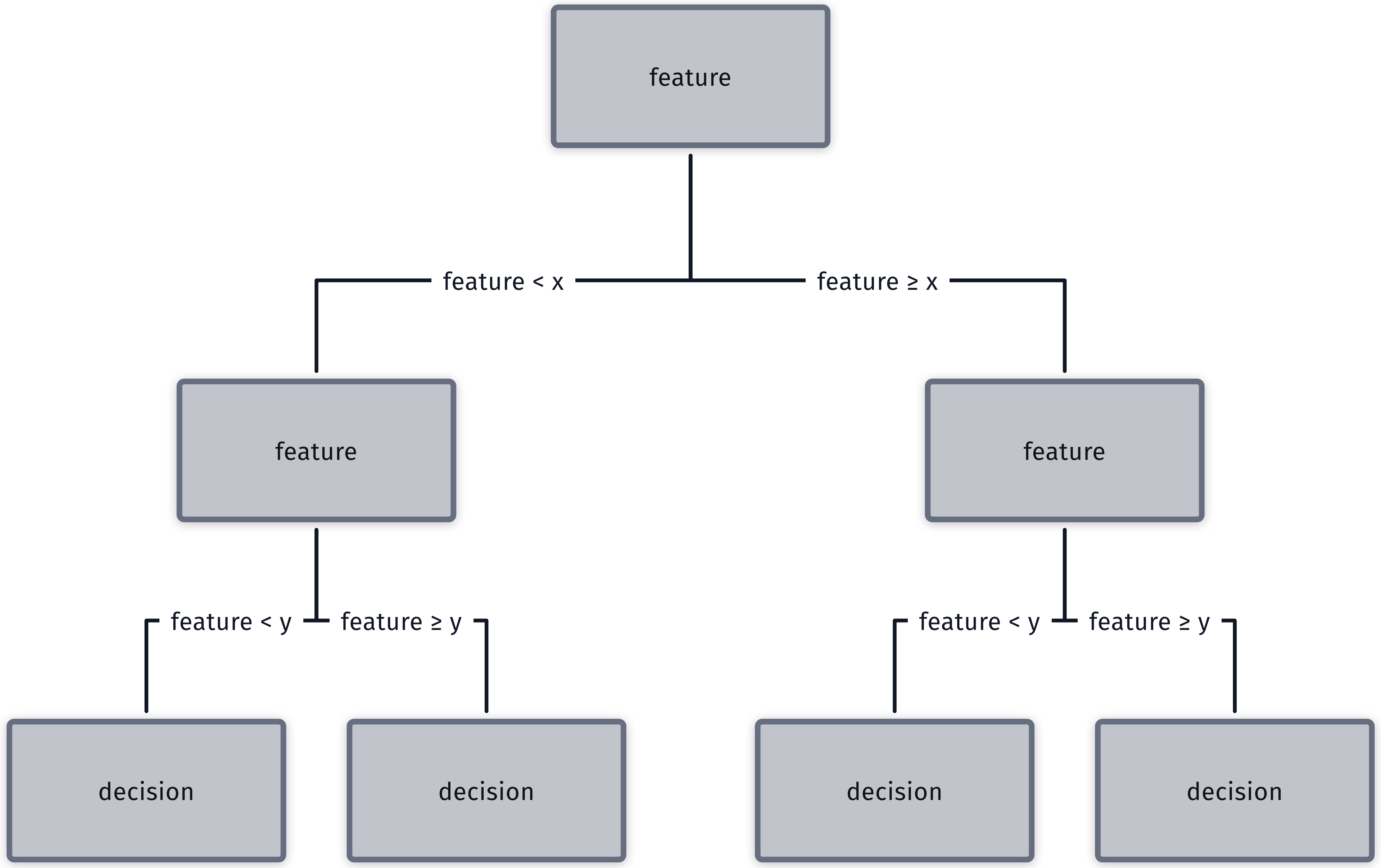
- Master thesis Johannes Bilk
  Precision (Efficiency) 82 %
- Sensitivity (Purity) 81%
- excluding 1-pixel cluster reduces precision by 3%
- multiclass vs. binary (one NN vs. many NN, all particle species) reduces precision by factor 2



Decision Trees

- Master thesis Stephanie Käs Precision (Efficiency)
- see table next page Sensitivity (Purity)
- Small Decision Tree with only 3 layers reaches 81% accuracy

1. cluster charge (97.6 % information content)
    - (reminder: no cluster reco, total charge is just the 9x9 ADC sum)
2. ADC value of pixel 41
    - (center of matrix, 1.1% information content)
3. z coordinate (1.3% information content).

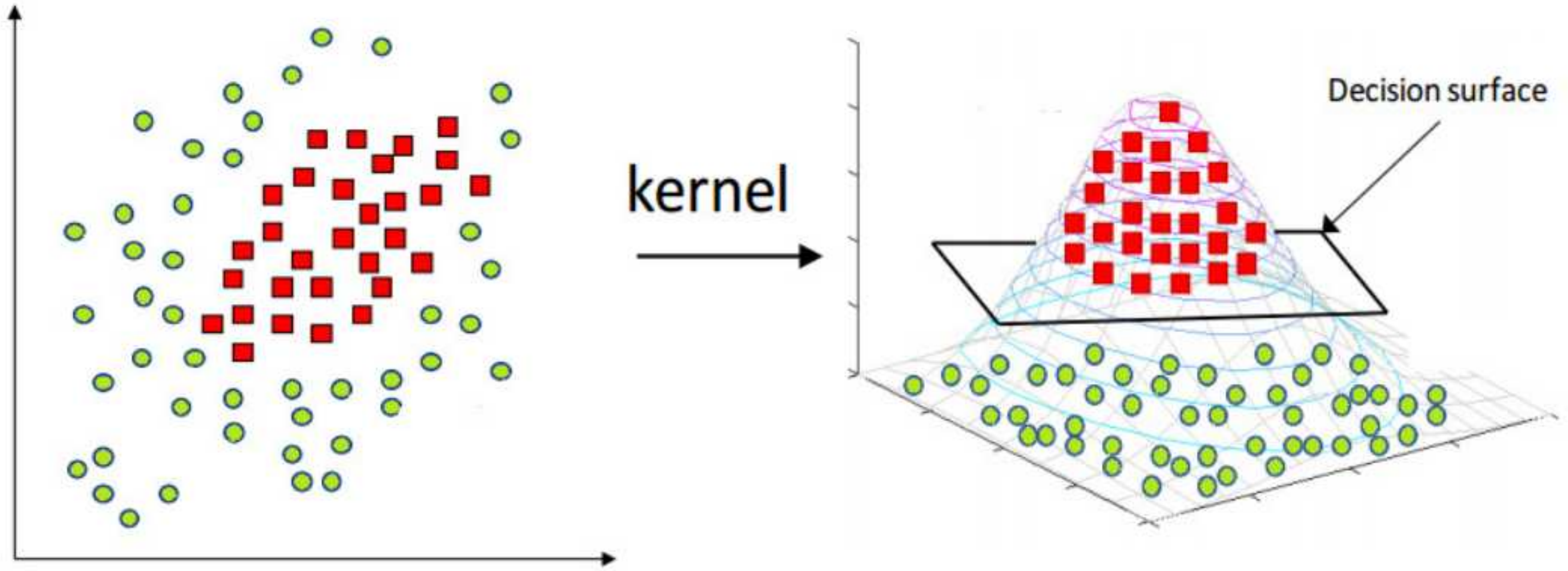


Decision Trees

Typ	max. Tiefe	Präzision [%]	Sensitivität [%]	Genauigkeit [%]
Std. Tree	2	82	80	81
	3	82	80	81
	5	82	80	81
	10	80	84	81
	15	78	82	79
	20	77	80	78
AdaBoost	-	82	80	81
RandomForest	2	77	68	74
	3	80	65	74
	10	81	80	81
	15	82	80	81
	50	80	79	80
	100	80	78	79
XGBoost	2	80	83	81
	3	80	84	81
	10	80	84	82
	15	80	83	81
	50	79	82	80
	100	79	82	80

Support Vector Machine

- Bachelor thesis Timo Schellhaas
  Precision (Efficiency) 82.7 %
- Sensitivity (Purity) 68.5 %
- can be improved ?
- dimension n+1 is generated with radial basis functions
  - maybe wrong function for our problem



Source: Simple Tutorial on SVM and Parameter Tuning in Python and R | HackerEarth Blog

Summary

- PXD essential for slow pion tracking
- slow pion classification with machine learning was tested:
    - CNN, Decision trees, SVM (shown today)
    - MLP, SOM, Auto-encoder, Hopfield Network and GNN (not shown today)
    - Efficiency > 80 % and Purity > 80 % with almost all methods
- Flat decision tree with only 3 layers maybe best candidate algorithm for FPGA implementation

Thank You for Listening

Back Up