SCHOOL OF DATA ANALYSIS

Charged Particle Identification in LHCb

Denis Derkach

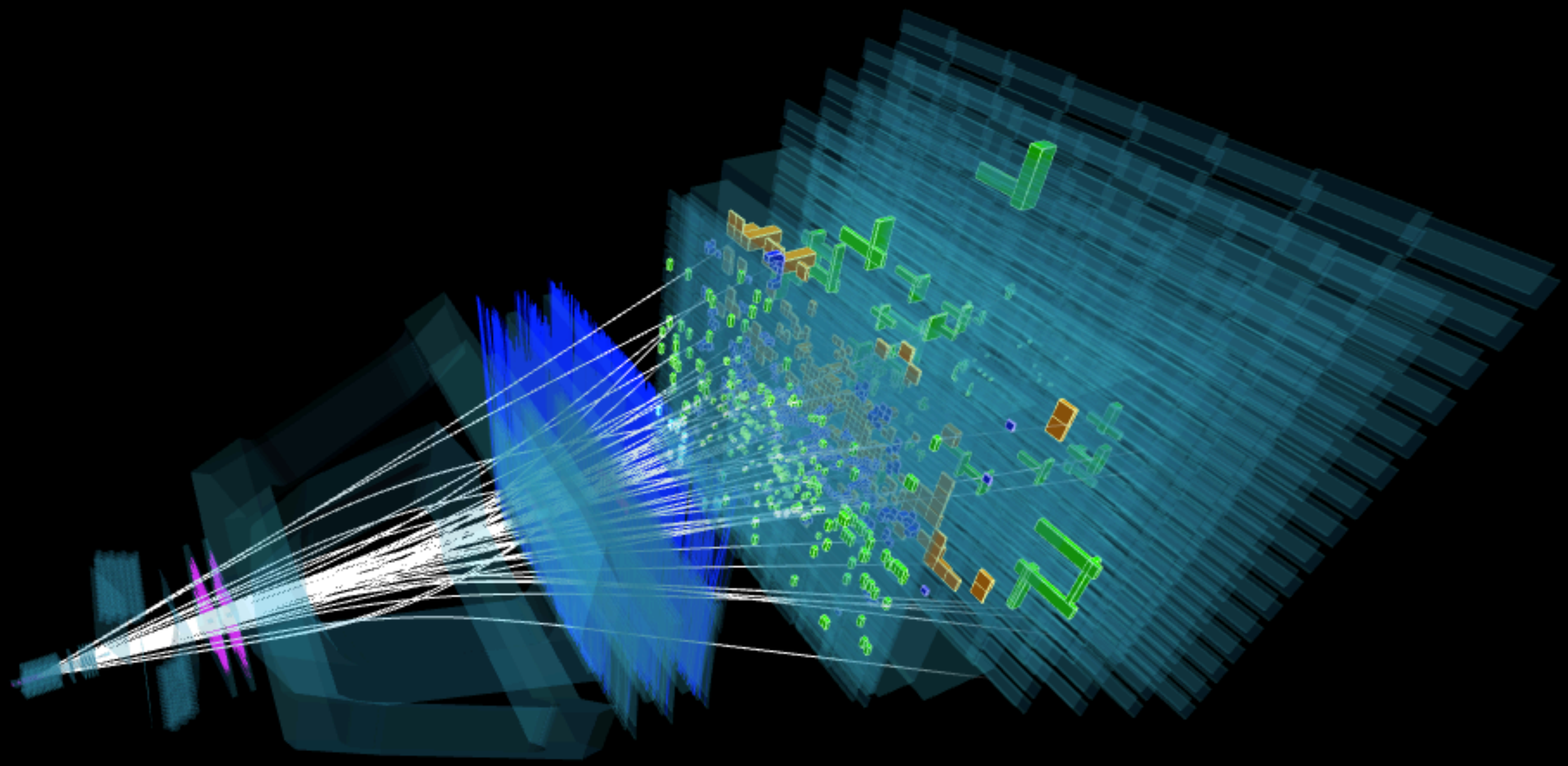
Yandex School of Data Analysis and Higher School of Economics, Moscow, Russia

Lambda Seminar

02 October 2017

NO PID

Event 2717213713
Run 199643
Thu, 28 Sep 2017 10:39:17



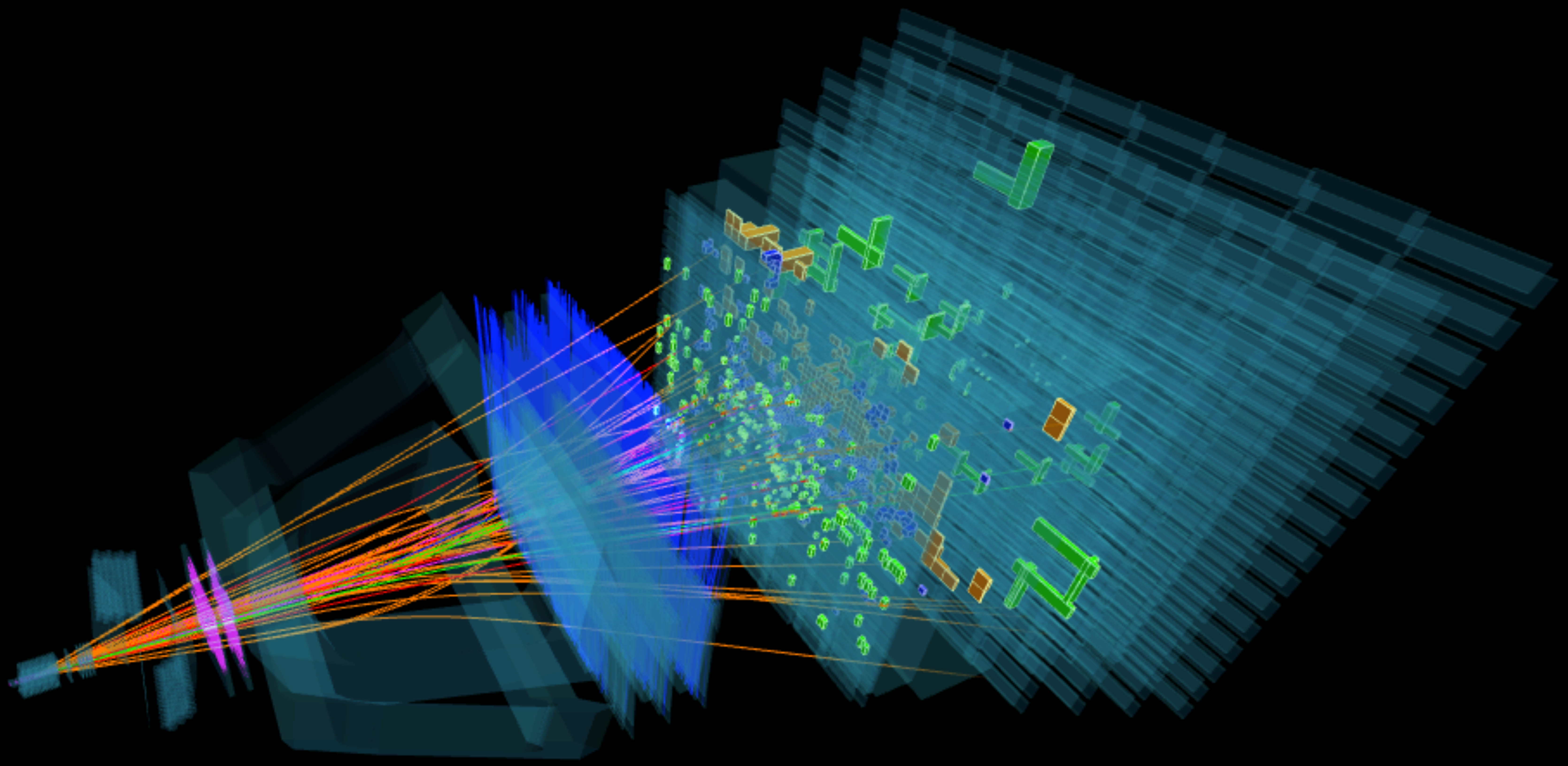
Open Controls

LHCb

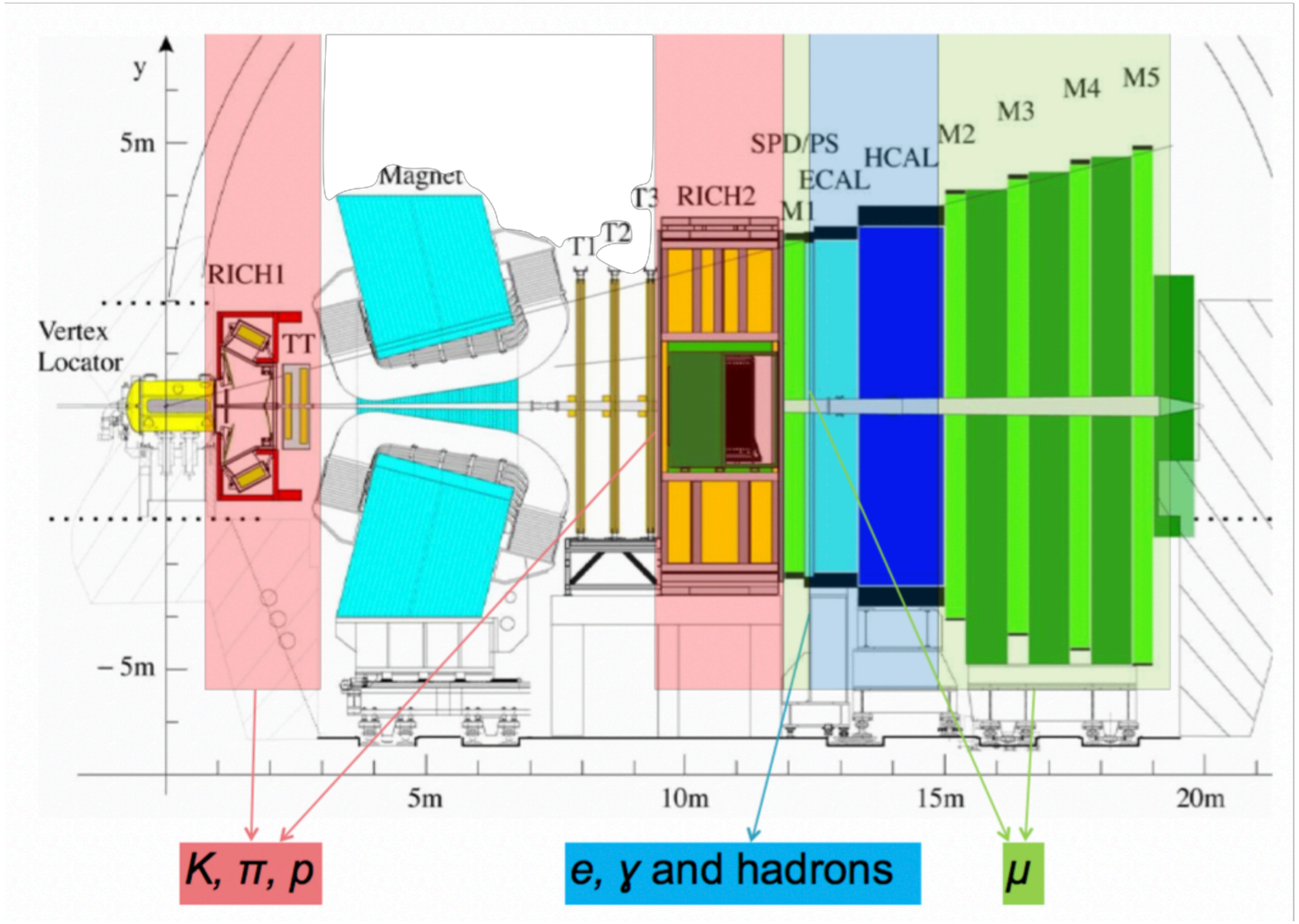
Event 2717213713

Run 199643

Thu, 28 Sep 2017 10:39:17



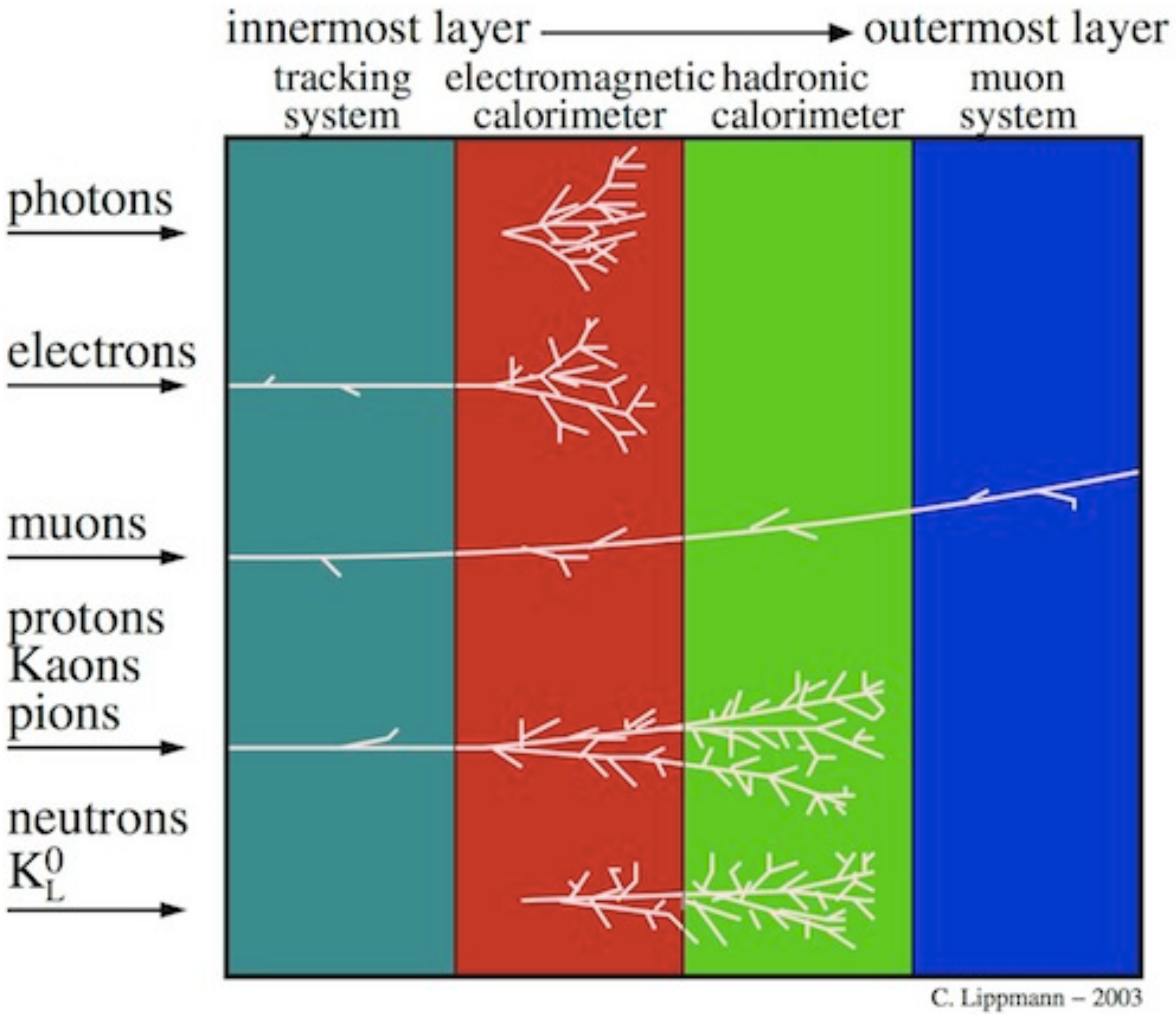
Event Characterisation



- Identify charged particle associated with a track
- Use information, collected by subdetectors:
  - trackers;
  - Cherenkov detectors;
  - calorimeters;
  - muon chamber.

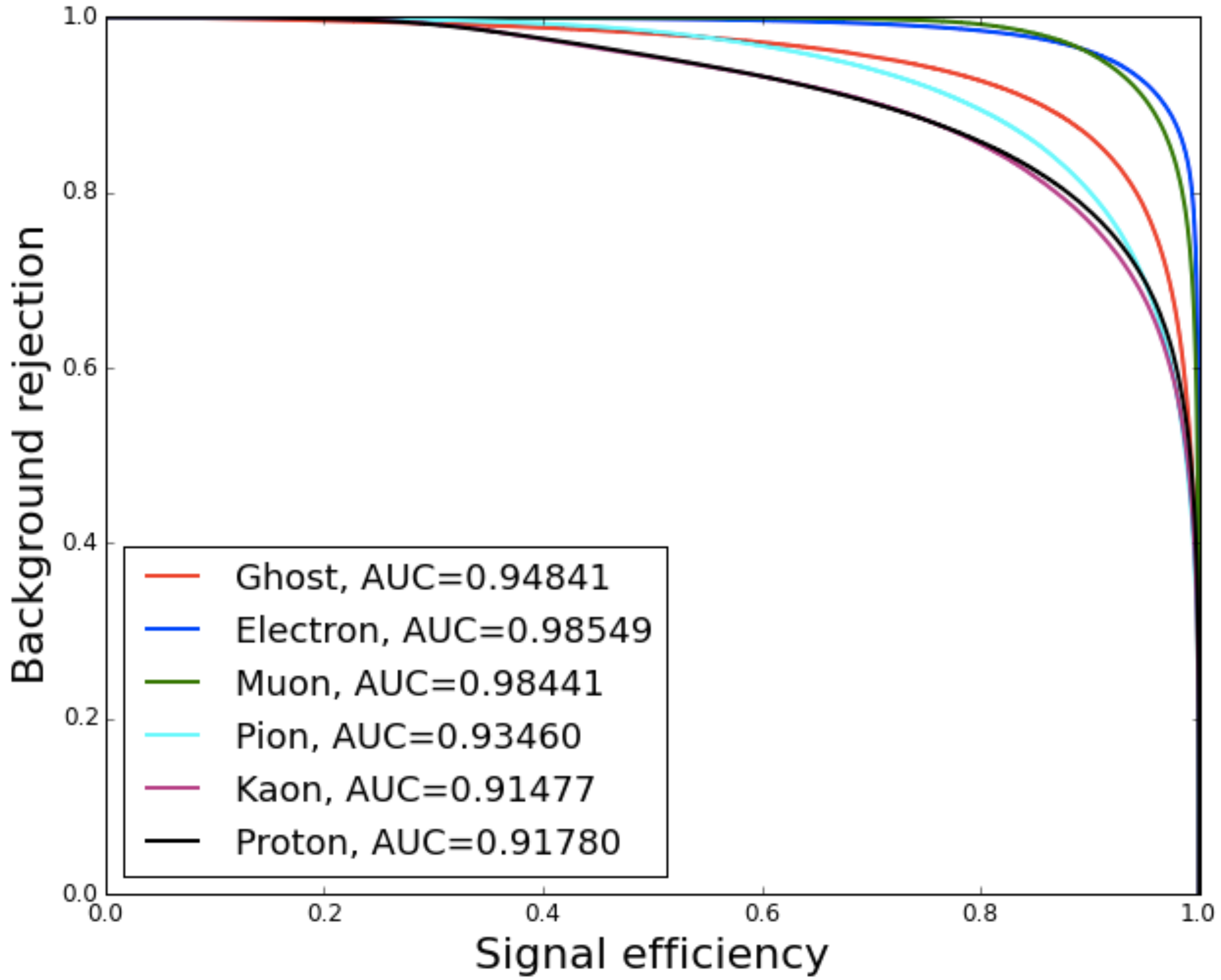
Problem

- particle types: electron, muon, pion, kaon, proton.
  - extra type: ghost track (reconstruction error).
- Aim: efficiently combine information from subdetectors.
- Sample: fully simulated events are used for training (12 Million events)



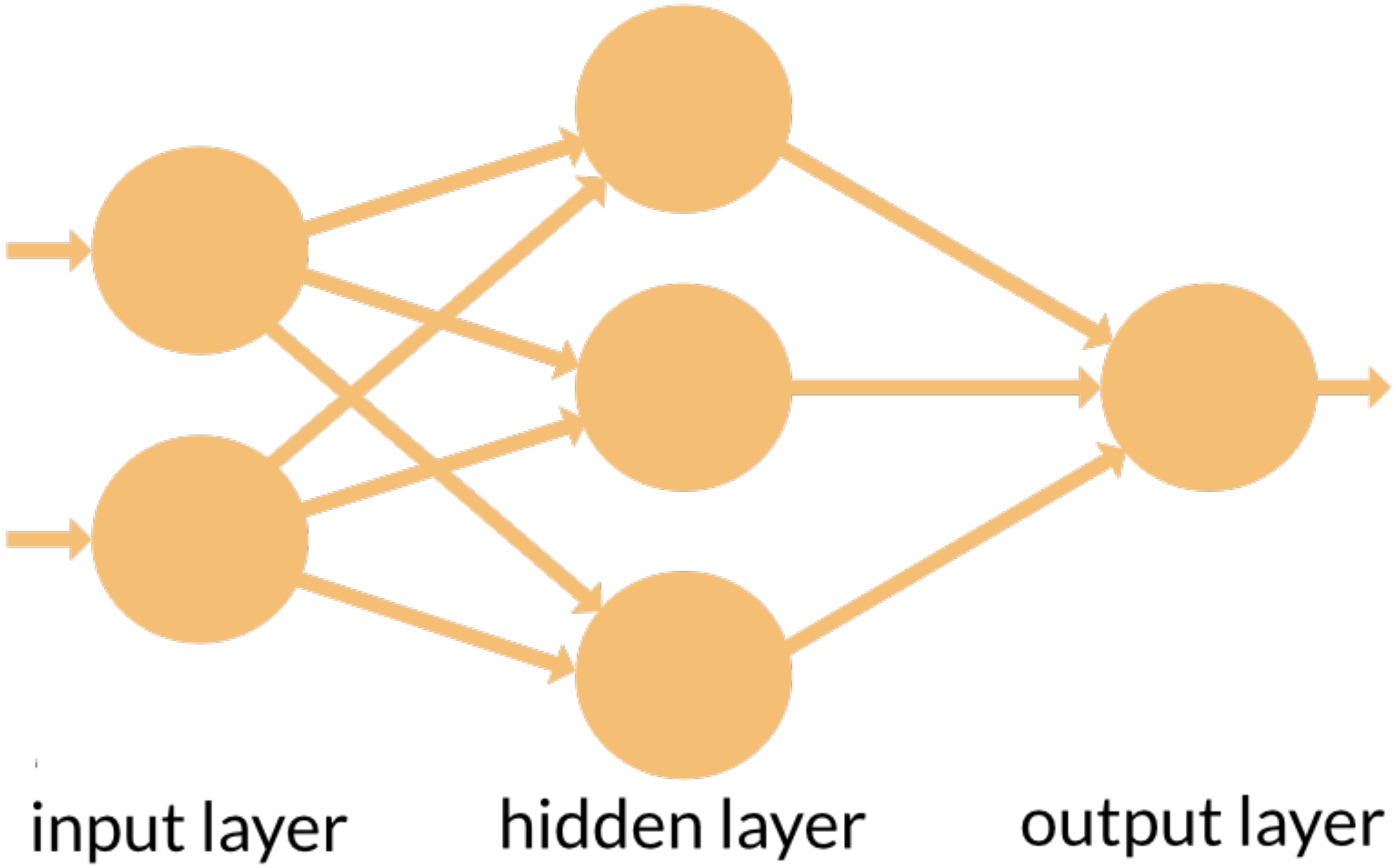
Total Model Quality

- In analyses, we are mostly interested in selecting one type (e.g., muon).
  - We use one-vs-rest ROC curves and area under the curves (AUC) to measure quality of the classification.



Baseline Model

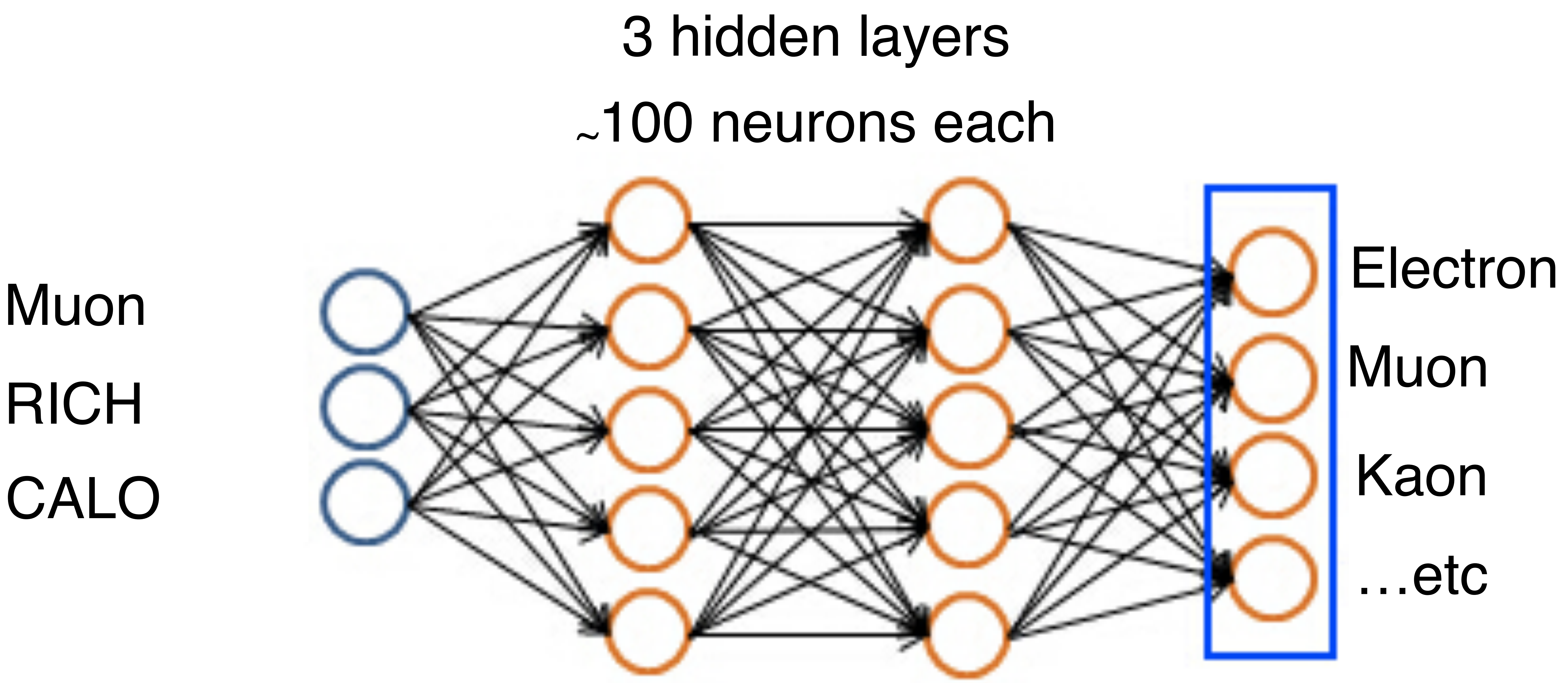
- Simple neural network with one hidden layer
- Consists of 6 binary classification models: one versus rest



Multiclass classification for Neural Networks

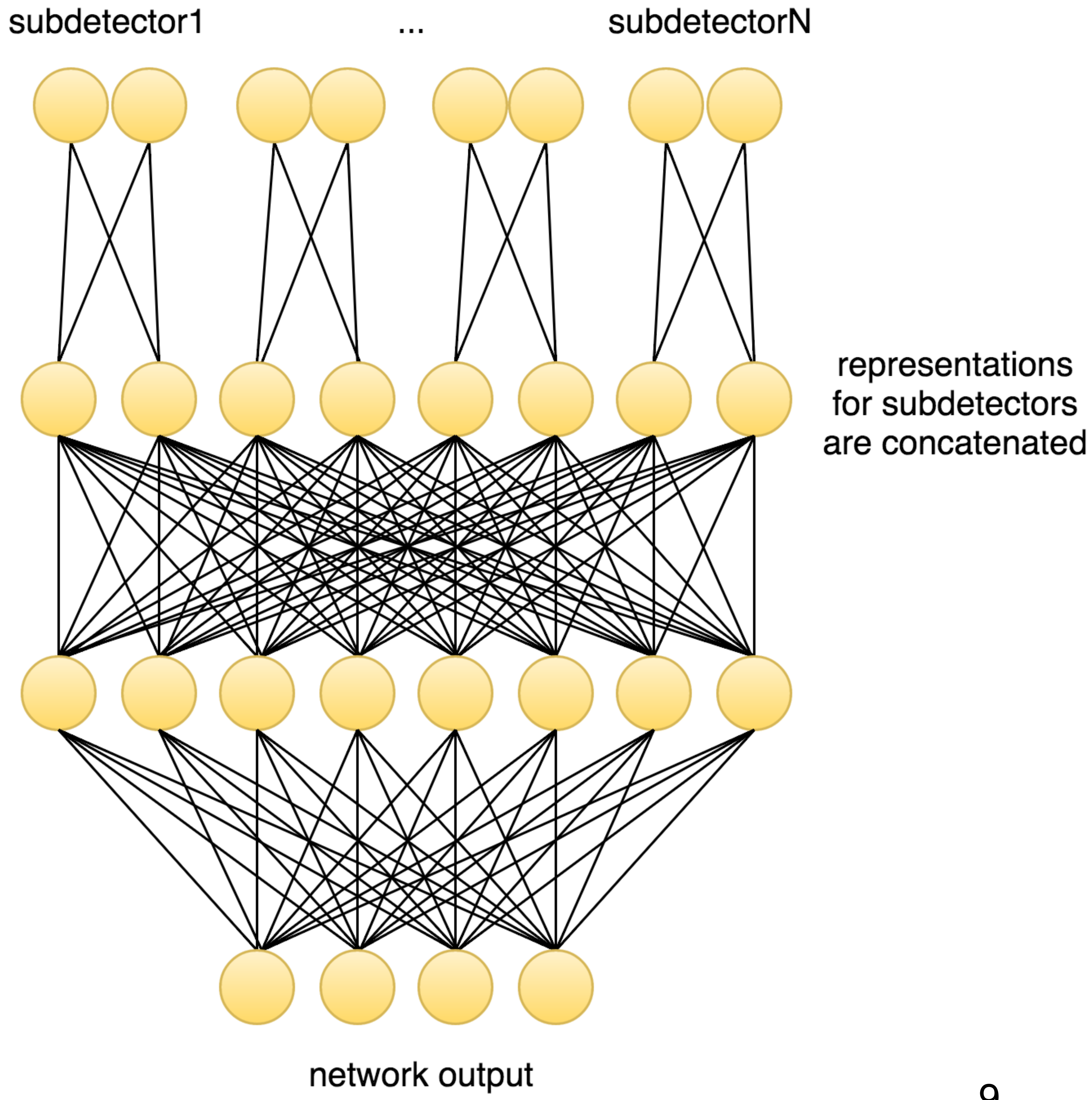
- The number of weights (parameters) in Neural Network is similar for binary classification and multiclass classification

- Computationally multiclass classification has (almost) the same complexity as binary classification (unless there are hundreds of classes and more)



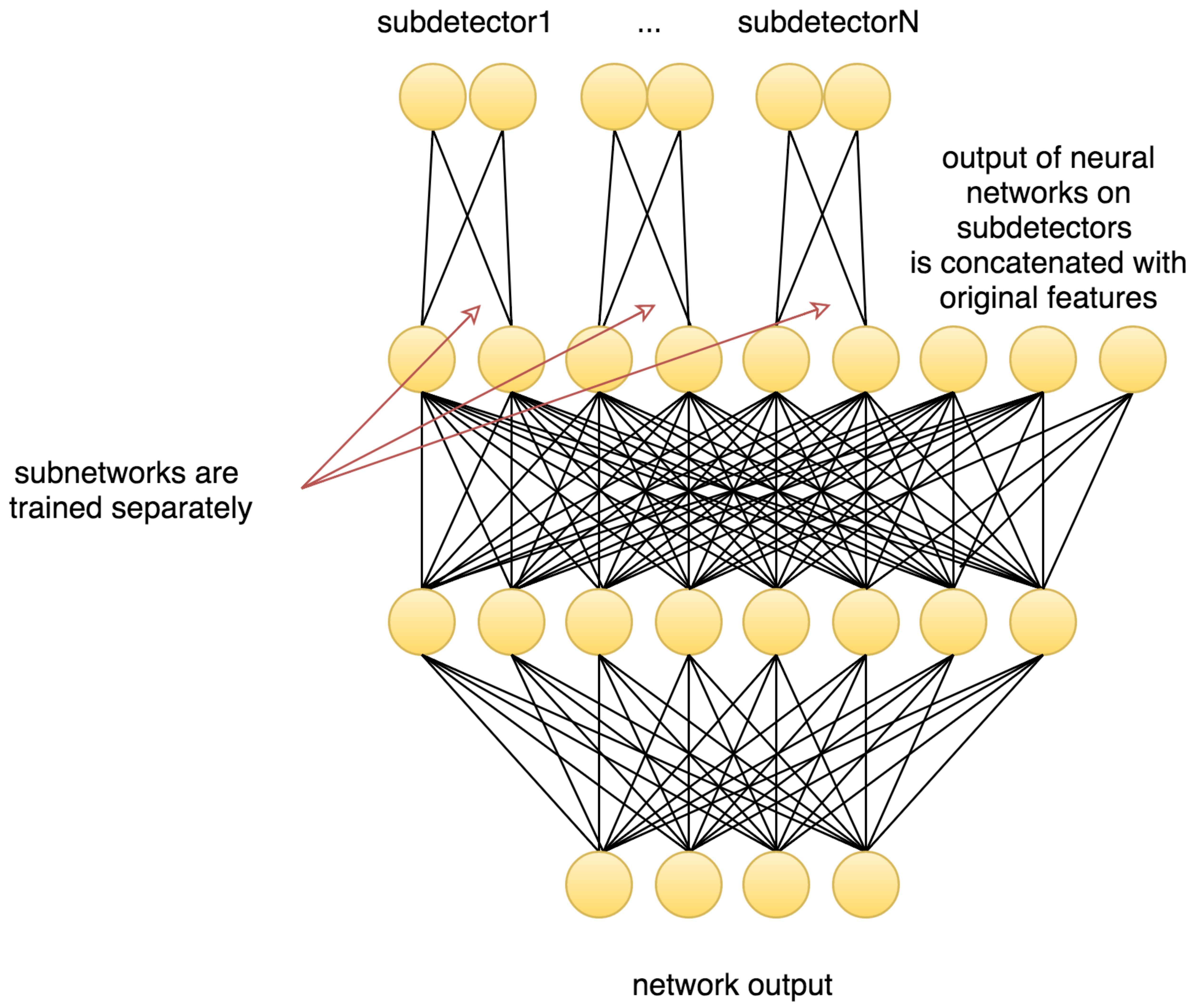
Neural Networks: Special Structure

- Linear combination of features for a subdetector seems to be informative.
- Initial layers combine information separately for each subdetector.
- These initial layers are optimized simultaneously with the rest of network.



Neural Networks: Stacking

- Linear combination of features for a subdetector seems to be informative.
- Separately train NN on each subdetector features.
- Later use this NNs as additional features for another NN (stacking approach)



Gradient Boosting on Decision Trees

The winning solutions were obtained using XGBoost and CatBoost.

Add linear combinations of initial features which can help in trees construction (NNs can reconstruct those itself)

Best-efficiency Models: ROC AUCs (one-vs-rest)

The higher the better; 1 means ideal separation.

Baseline → ProbNN

LHCb soft → CatBoost

LHCb soft → deep NN

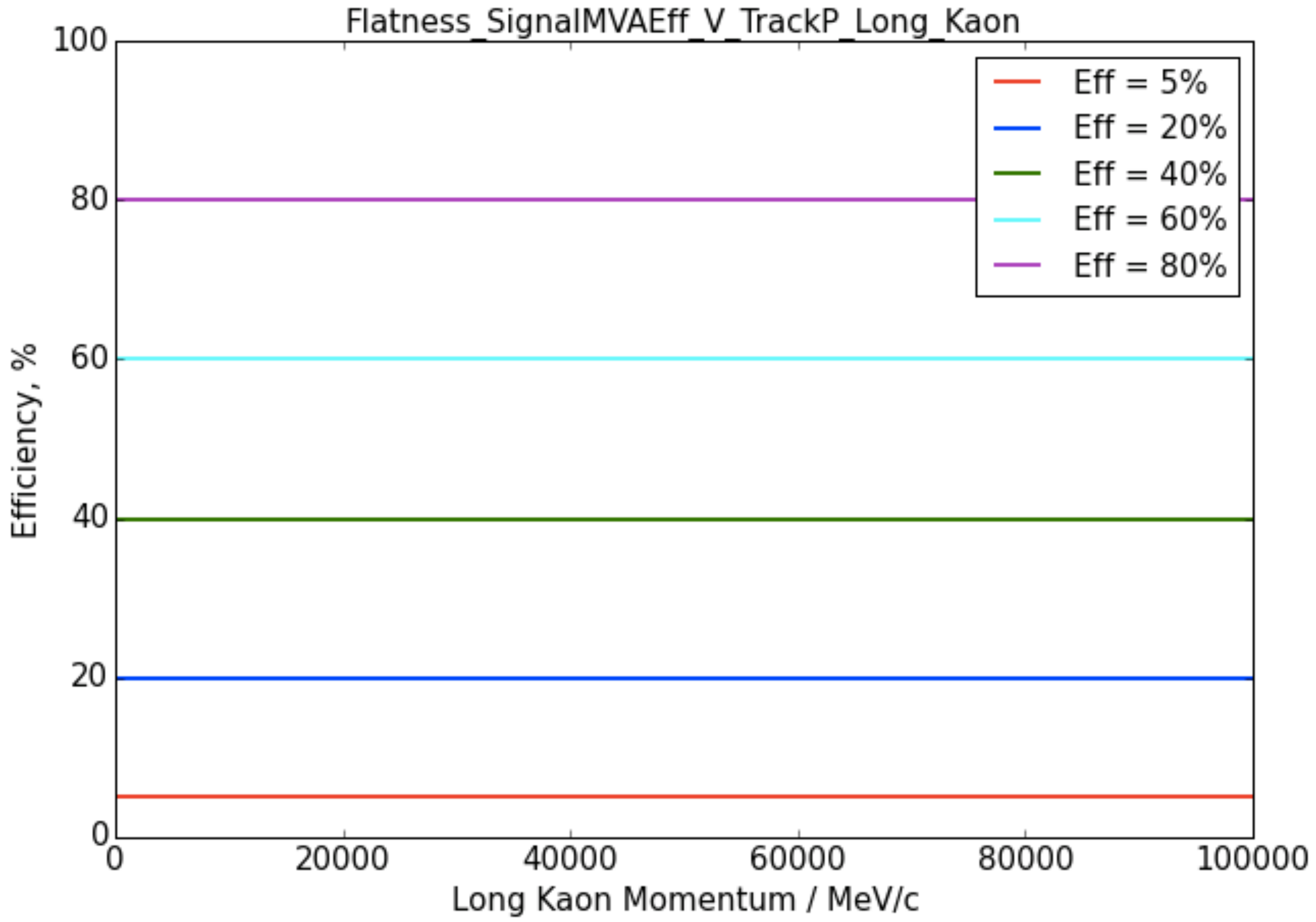
Particle type	Ghost	Electron	Muon	Pion	Kaon	Proton
ProbNN	0.9484 $\pm$0.0002	0.9855 $\pm$0.0001	0.9844 $\pm$0.0001	0.9346 $\pm$0.0003	0.9148 $\pm$0.0002	0.9178 $\pm$0.0001
special GB	0.9637 $\pm$0.0002	0.9914 $\pm$0.0001	0.9927 $\pm$0.0001	0.9577 $\pm$0.0003	0.9309 $\pm$0.0002	0.9310 $\pm$0.0001
XGBoost	0.9609 $\pm$0.0002	0.9908 $\pm$0.0001	0.9922 $\pm$0.0001	0.9568 $\pm$0.0003	0.9303 $\pm$0.0002	0.9302 $\pm$0.0001
CatBoost	**0.9641** $\pm$0.0002	**0.9917** $\pm$0.0001	**0.9929** $\pm$0.0001	0.9586 $\pm$0.0002	**0.9322** $\pm$0.0002	**0.9323** $\pm$0.0003
simple NN	0.9615 $\pm$0.0002	0.9910 $\pm$0.0001	0.9922 $\pm$0.0001	0.9574 $\pm$0.0002	0.9305 $\pm$0.0002	0.9304 $\pm$0.0001
deep NN	0.9632 $\pm$0.0002	0.9915 $\pm$0.0001	0.9925 $\pm$0.0001	**0.9587** $\pm$0.0002	0.9320 $\pm$0.0002	0.9319 $\pm$0.0001
stack NN	0.9623 $\pm$0.0002	0.9911 $\pm$0.0001	0.9924 $\pm$0.0001	0.9578 $\pm$0.0002	0.9315 $\pm$0.0002	0.9312 $\pm$0.0001
special NN	0.9621 $\pm$0.0002	0.9910 $\pm$0.0001	0.9923 $\pm$0.0001	0.9576 $\pm$0.0002	0.9308 $\pm$0.0001	0.9301 $\pm$0.0001

- BDT has similar quality to DL

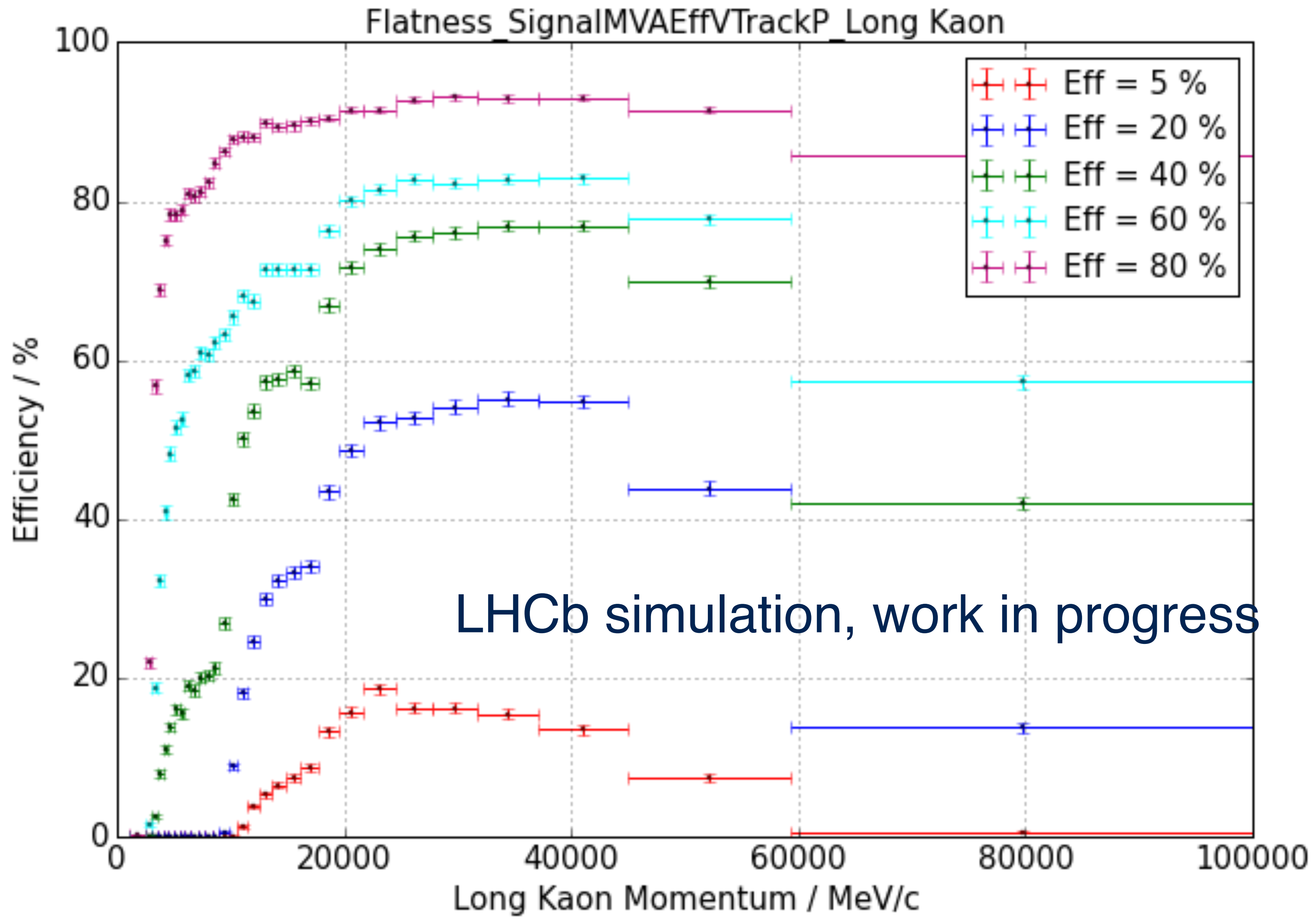
Improving PID with flat models

- The whole PID information strongly depends on particle momentum, that leads to strong dependency between PID efficiency and momentum.

- In some analyses we need to have flat PID along signal particle momentum to avoid additional systematics.



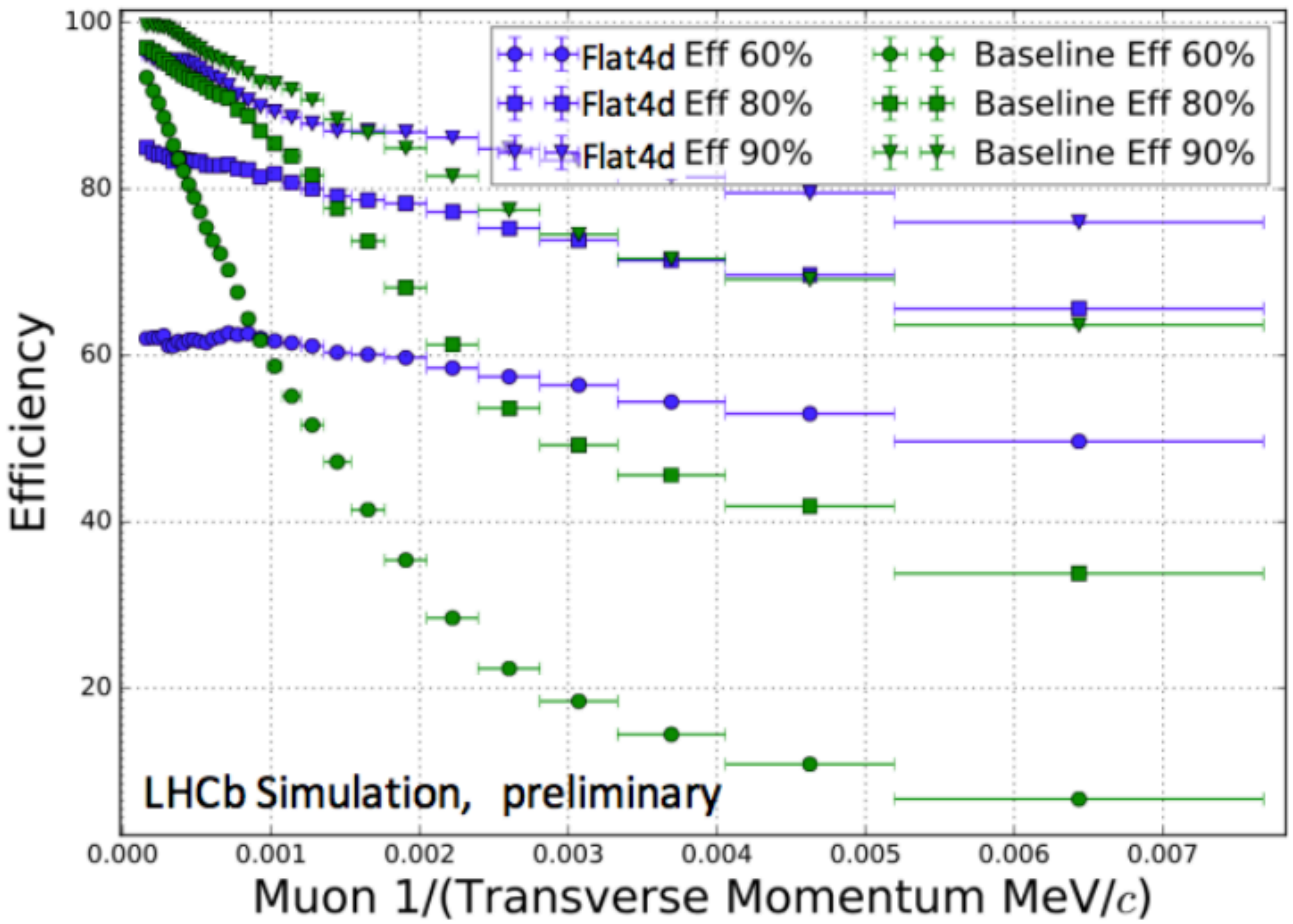
Ideal world



Real world

Flat Model vs Baseline

- There is a strong dependency on momentum for conventional models
(see baseline efficiencies on the plot)
- Uniform boosting approach suppresses this dependency.



Flat Models: ROC AUC

- Flatness is a very strong restriction, holding this restriction leads to quality decreasing.
- Still, flat model is not worse than baseline in ROC AUCs

	Ghost	Electron	Muon	Pion	Kaon	Proton
baseline	0.9484	0.9854	0.9844	0.9345	0.9147	0.9178
P + Pt flatness	0.9605	0.9883	0.9886	0.9514	0.9146	0.9139
2d(P, Pt) flatness	0.9594	0.9868	0.9865	0.9494	0.9049	0.8993
4d(P, Pt, eta, nTracks) flatness	0.9593	0.9861	0.9864	0.9474	0.9062	0.8976
P + Pt + eta + nTracks flatness	0.9600	0.9874	0.9884	0.9503	0.9130	0.9129

Conclusion

- Application of modern machine learning solutions brings significant improvements into the workflow of LHCb experiments.
- Advanced techniques may allow to tackle the source of systematics uncertainties with only a small deterioration in quality.
- Next steps will be a global optimisation of charged particle identification based on the deep subdetector feature combination.