Neutral pion identification at Future Circular Collider

Matej Haviernik
Institute for Particle and Nuclear Physics

Supervisor: Mgr. Jana Faltová, Ph.D.

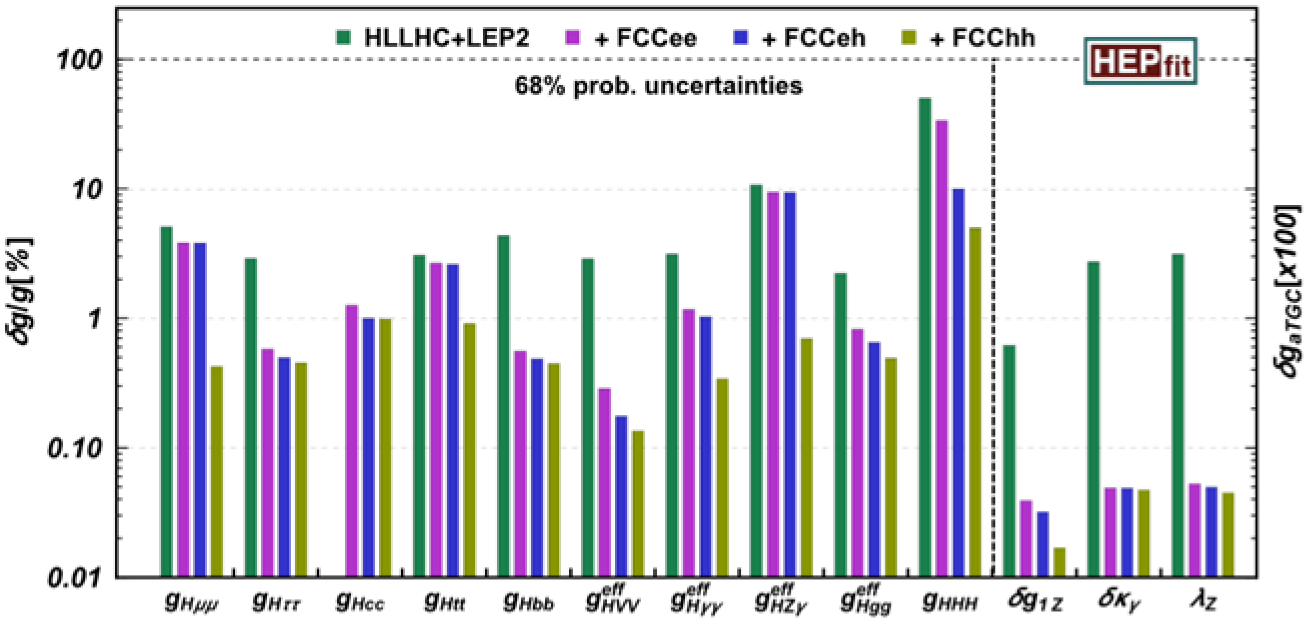
FCC

100 km collider at CERN following up on HL-LHC
2 stages:
FCC-ee: e^+e^- collider (2040)

- 88-365 GeV, multiple runs
- EW sector + Higgs & t
  - Vector boson mass & decay width, A_{FB}, $\alpha_S(M_Z)$, couplings

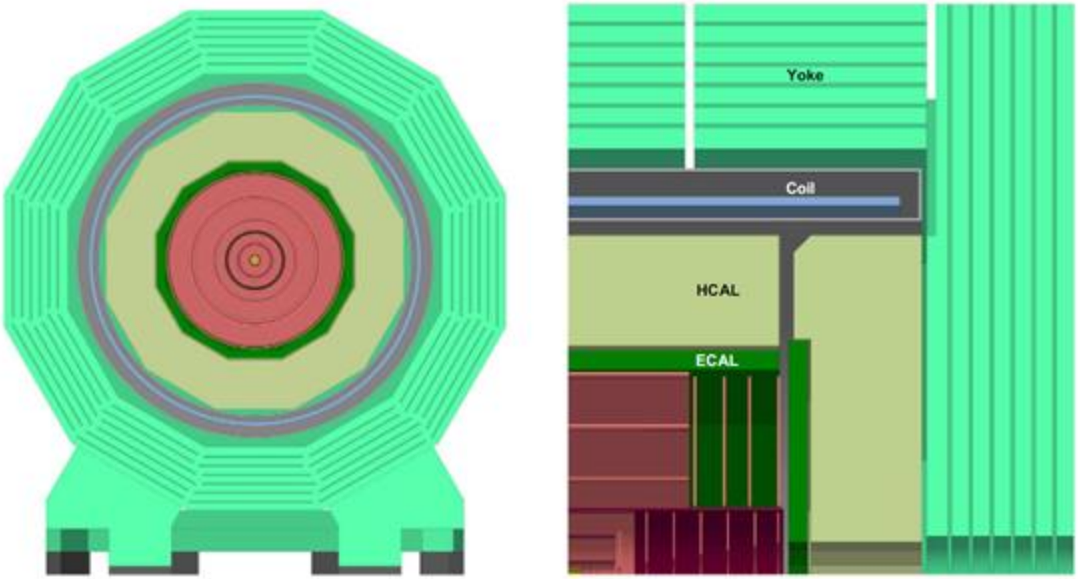
Working point	Tot. lum./year [ab^{-1}/year]	Run [years]	Goal [ab^{-1}]
Z pole (88-94 GeV)	24	2	150
Z pole (88-94 GeV)	48	2	
WW threshold ($\sim$ 161 GeV)	6	1-2	10
ZH threshold ($\sim$ 240 GeV)	1.7	3	5
$t\bar{t}$ threshold (340-350 GeV)	0.20	1	0.2
$t\bar{t}$ threshold ($\sim$ 365 GeV)	0.34	4	1.5

FCC-hh: pp collider with max. CMS energy of 100 TeV



Abada, A., Abbrescia, M., AbdusSalam, S.S. *et al.* FCC Physics Opportunities. *Eur. Phys. J. C* **79**, 474 (2019)

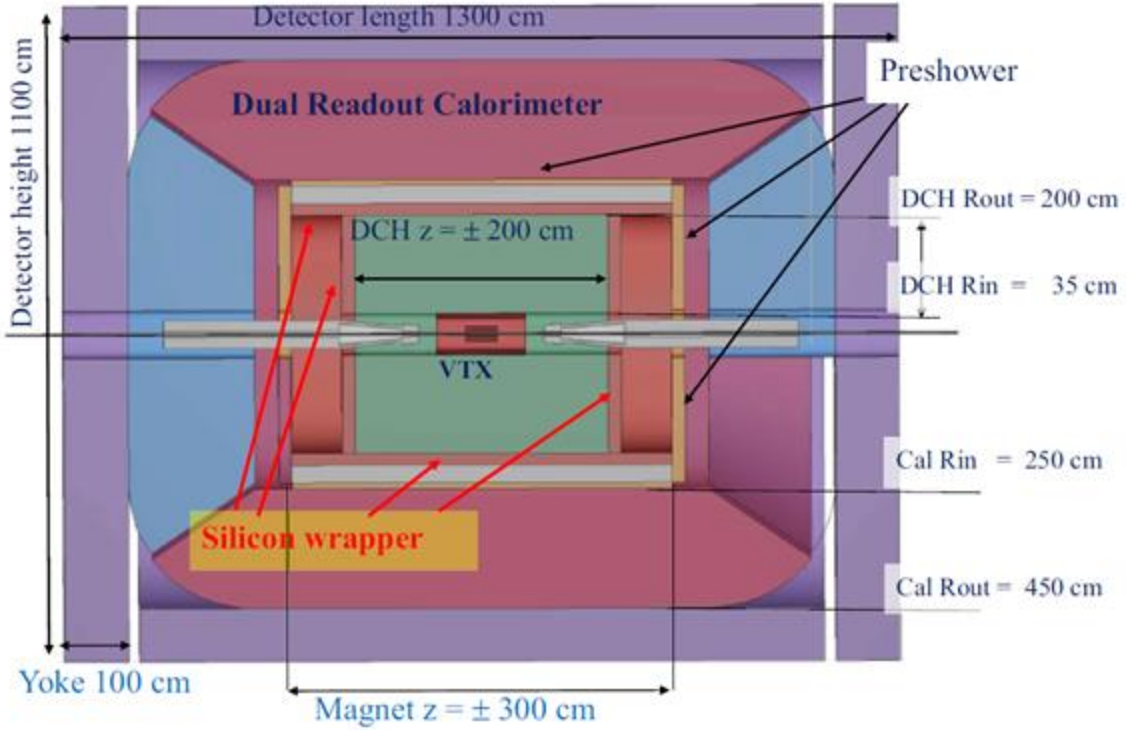
Proposed detectors for FCC-ee



A. Abada et. al. FCC-ee:The Lepton Collider. The European Physical Journal Special Topics, 228:261–623, 201

CLD – CLIC-Like Detector

- Pixel vertex detector + Si tracker
- ECAL: Si-W sampling calorimeter
- HCAL: Polystyrene scintillator



A. Abada et. al. FCC-ee:The Lepton Collider. The European Physical Journal Special Topics, 228:261–623, 201

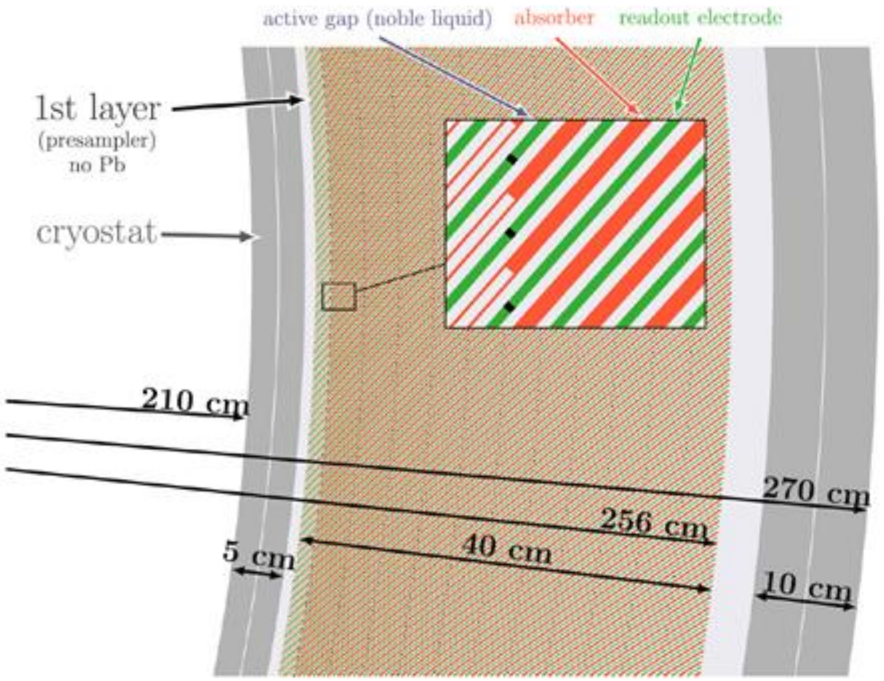
IDEA – Innovative Detector for Electron-Positron Accelerators

- Pixel vertex detector + wire chamber
- MPGD & lead layers
- CALO: Dual-readout design with Cherenkov & scintillators

FCC-ee noble liquid calorimetry

Modified IDEA design - LAr+lead sampling ECAL calorimeter & HCAL calorimeter with scintillating tiles

- Advantages: good energy resolution, uniform response
- LAr calorimeter projected also for FCC-hh due to radiation hardness



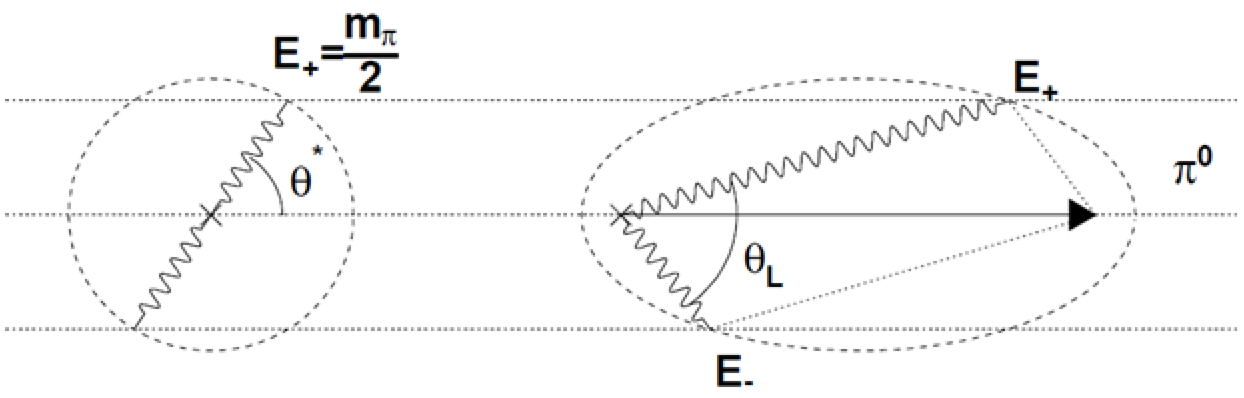
N. Morange. Noble Liquid Calorimetry for FCC-ee. Instruments, 6(4), 2022.

ATLAS TileCal-like sampling calorimeter with scintillating tiles & steel absorbers

LAr sampling calorimeter with lead absorbers

M. Aleksa. Noble Liquid Calorimetry: Input proposals in Track 2. Presented at: ECFA Detector RD Roadmap Task Force 6: 2nd Calorimetry Community Meeting, [online]: https://indico.cern.ch/event/1246381

Neutral pion decay



$\pi^0 \rightarrow \gamma\gamma$ with BR = 0.98823 ± 0.00034

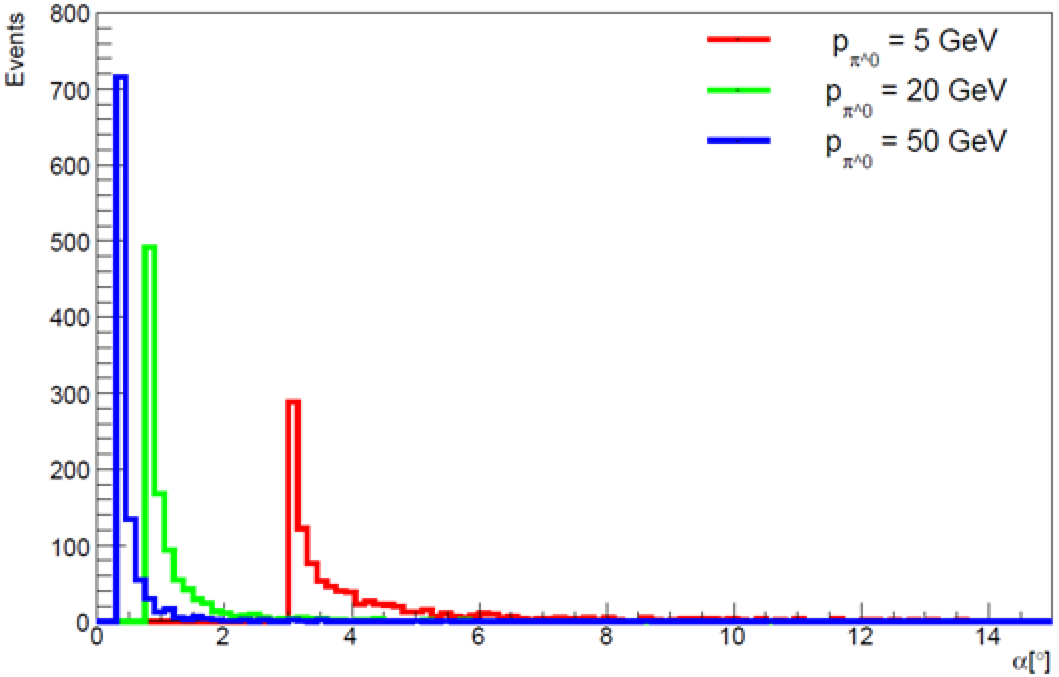
Important decay product

- Decays of τ, Z & hadrons

γγ decay angle very small in LAB frame & most decays clustered around α_{min}

Problem: easy to misidentify as a single photon in a detector

$$\sin\frac{\alpha_{min}}{2} = \frac{m_\pi}{E_\pi}$$



Data & simulated environment

Geant4 - 100 000 π^0 and γ

- E = [0; 100] GeV
- Uniform φ, ϑ and E distribution
- Threshold for shower particles = 0.05 GeV

Calorimetric environment – Barrel LAr sampling calorimeter in DD4HEP

- Granularity - $\Delta\vartheta \times \Delta\varphi = 0.57^\circ \times 0.47^\circ$ vs. $\Delta\vartheta \times \Delta\varphi = 0.14^\circ \times 0.47^\circ$ (strip layer motivation)
- 12 radial layers

Final state – clustering information & cells (position, index, deposited E)

Task: identify neutral pions against photon background & evaluate the impact of high hranularity on identification efficiency

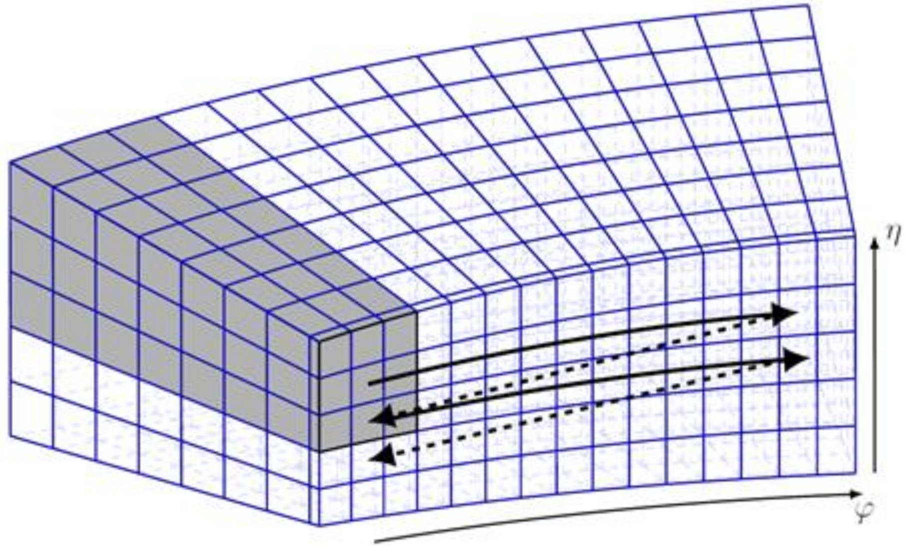
Clustering

Sliding window

- $N_\eta \times N_\varphi$ = 9x17 cluster (e- optimalization) for $E > E_{cut}$
- Low number of clusters per event

Topological clustering – cluster growing from seed cells with significance above a cut $\xi_{cell} = \left| \frac{E_{cell}}{\sigma_{cell}^{noise}} \right| \geq S$

- Large number of clusters per event with irregular shape
- S = 4



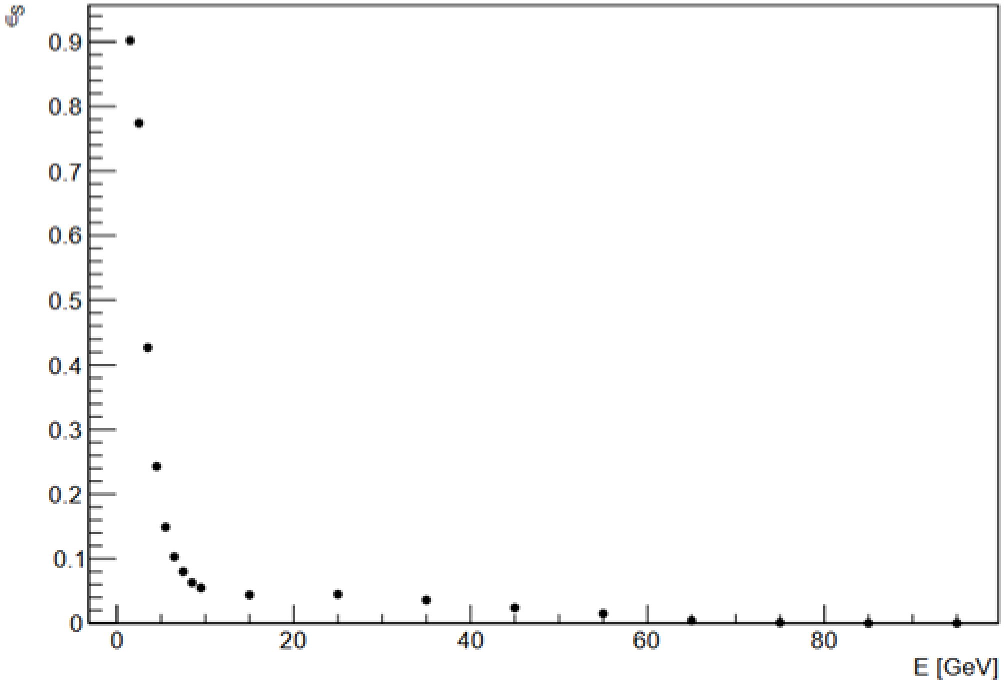
M. Aleksa et. al. Calorimeters for the FCC-hh, 2019. [online]: https:// arxiv.org/abs/1912.09962.

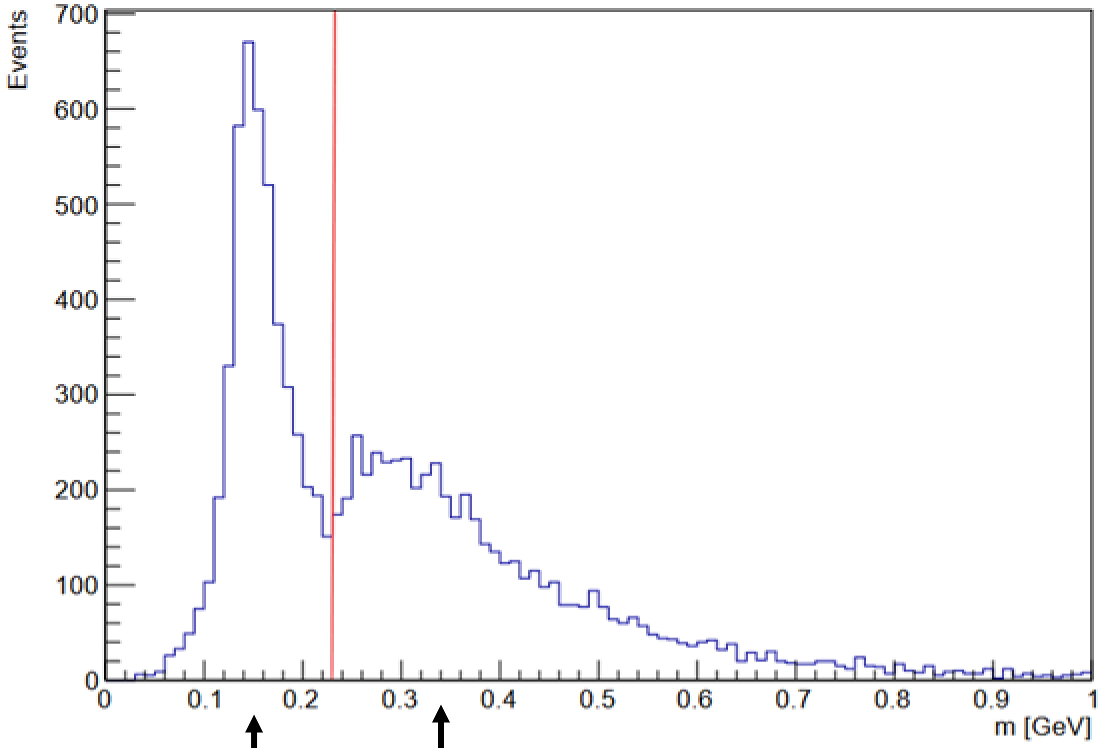
Pion identification using clusters?

Ineffective (sliding window algorithm unoptimized for π^0)

Sliding window – identify two-cluster events with $m = m_\pi$

- Rapid drop-off beyond E = 10 GeV
- Cluster merging due to overlapping electromagnetic showers





True neutral pion peak

Contributions from high energy π^0 decays – spurious 2nd clusters

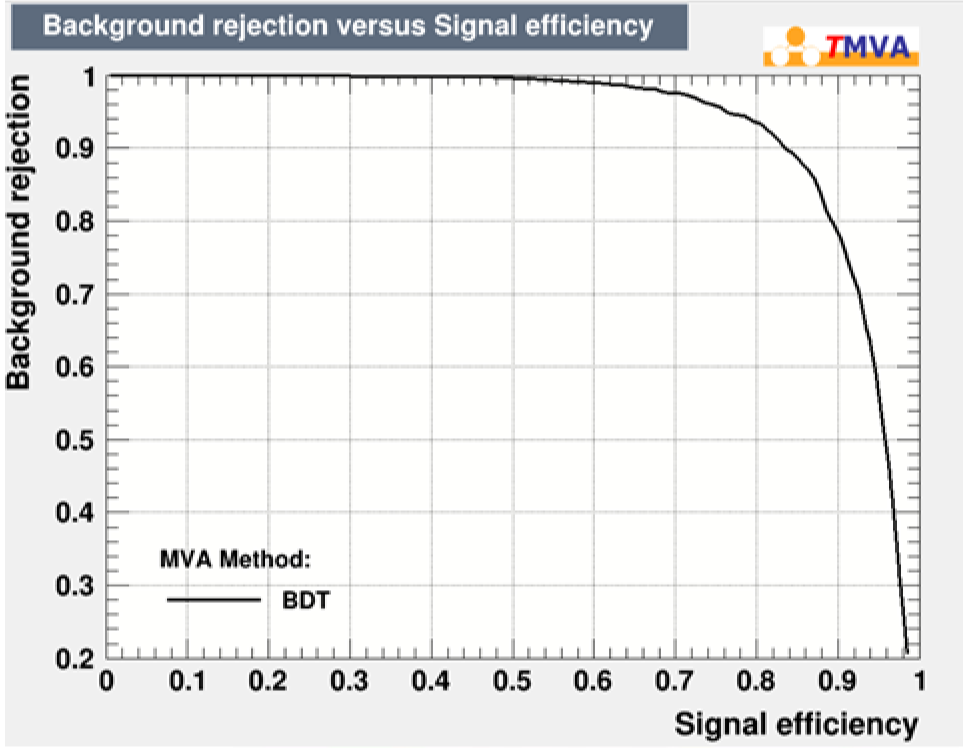
Multivariate analysis

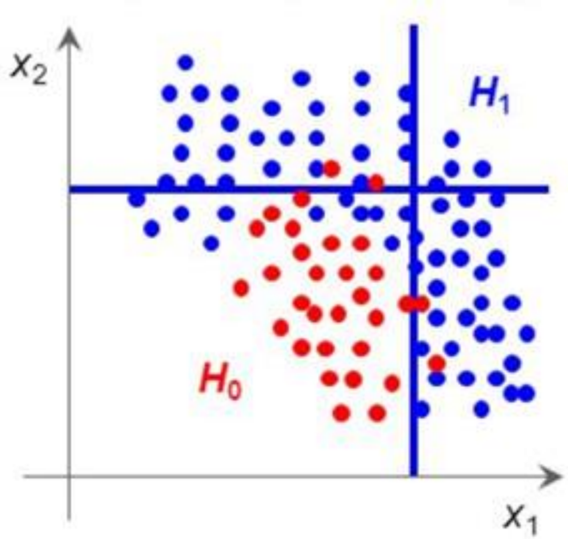
Characterize events by a set of discriminating variables

Classify events into signal (S) or background (B) with ROC curve as a function of r_B against ε_S

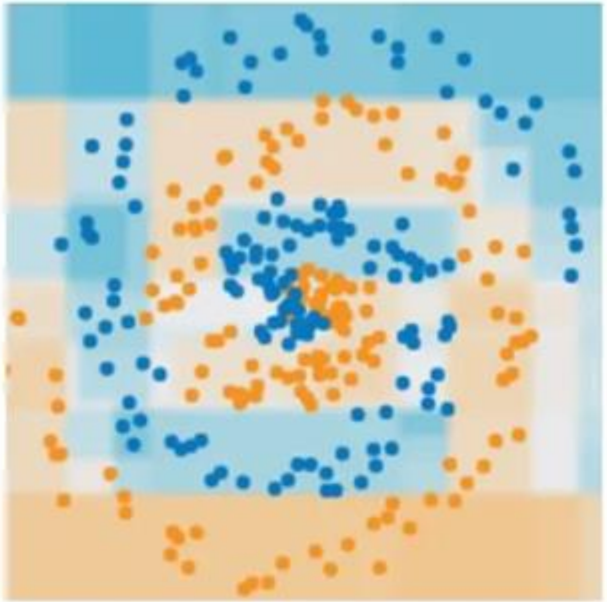
TMVA ROOT, methods used:

- Rectangular cuts – transparent, resistant to overtraining
- Boosted Decision Trees – better performance





R. Quagliani. Study of Double Charm B Decays with the LHCb Experiment at CERN and Track Reconstruction for the LHCb Upgrade. First edition. Springer Cham, 2018.

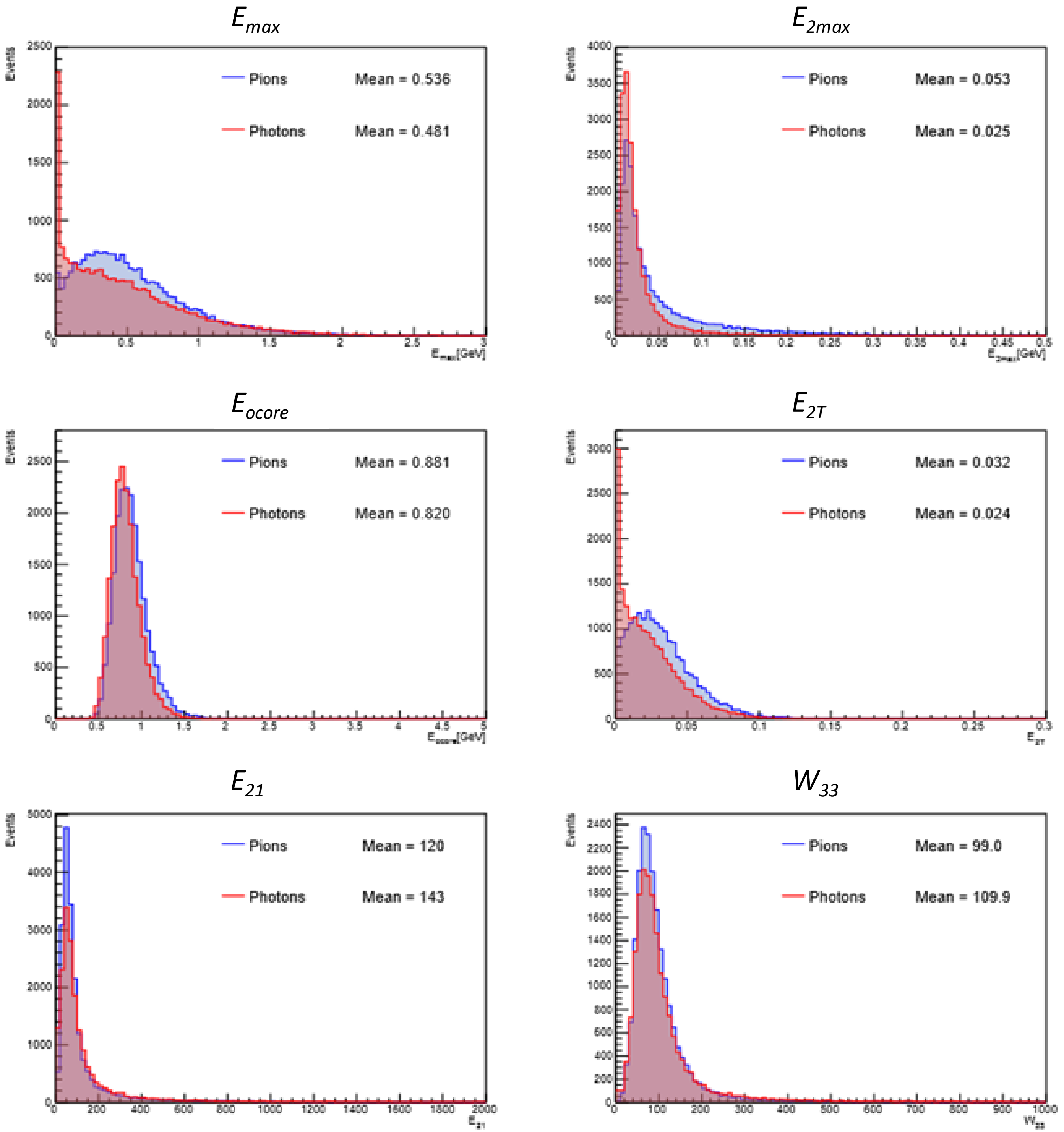


[online]: https://amva4newphysics.wordpress.com

Discriminating variables

Obtained from energy deposition in cells

- Longitudinal profile – E_{max}, E_{2max}, E_{iT}, E_{i1}…
- Transversal profile – E_{ocore}, W_{33}



Cuts

Genetic Algorithm sampling of solutions

Parameters

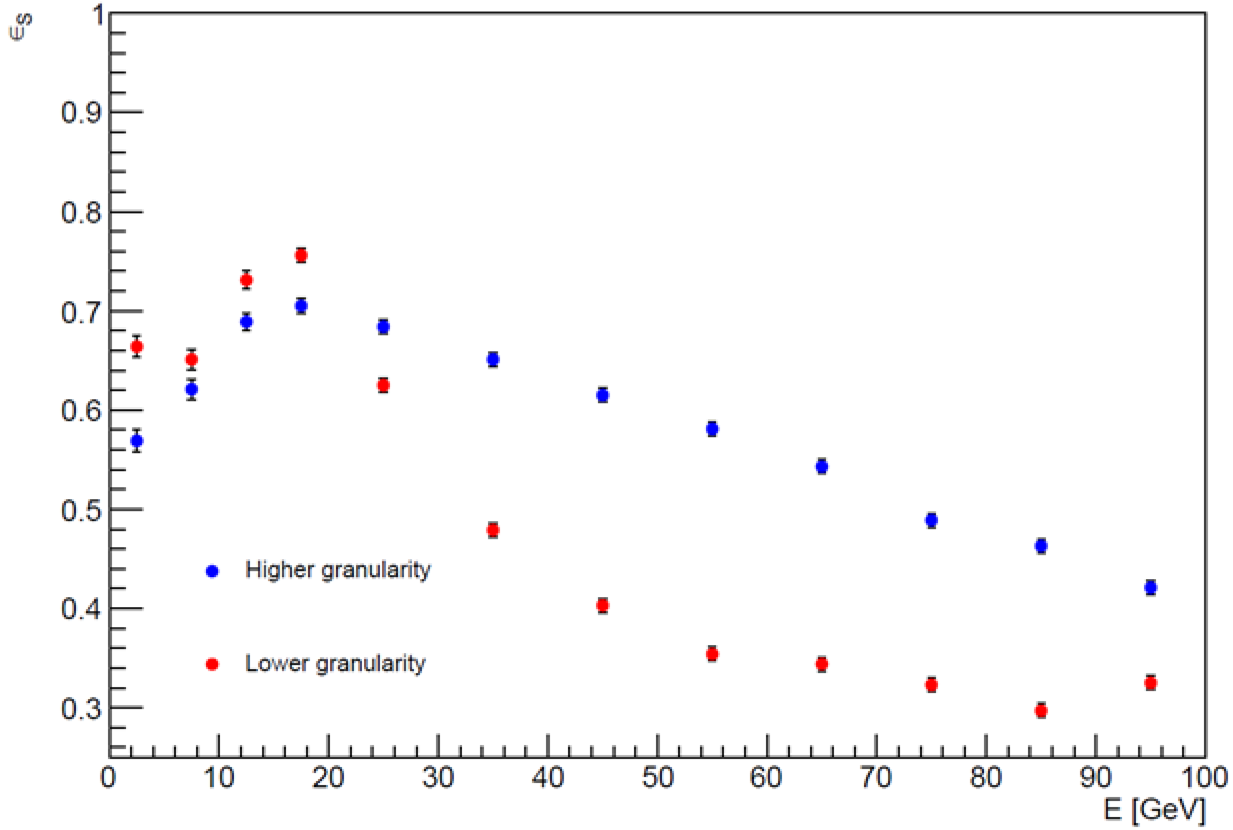
PopSize	800
Steps	40
Cycles	5

Efficiency calculated for $r_B = 0.8$

Maximum at $E = 20$ GeV

- Drop in efficiency below 20 GeV
- Visibly better performance for HG

Eliminate the (counterintuitive) dropping below 20 GeV & raise performance – move on to BDT

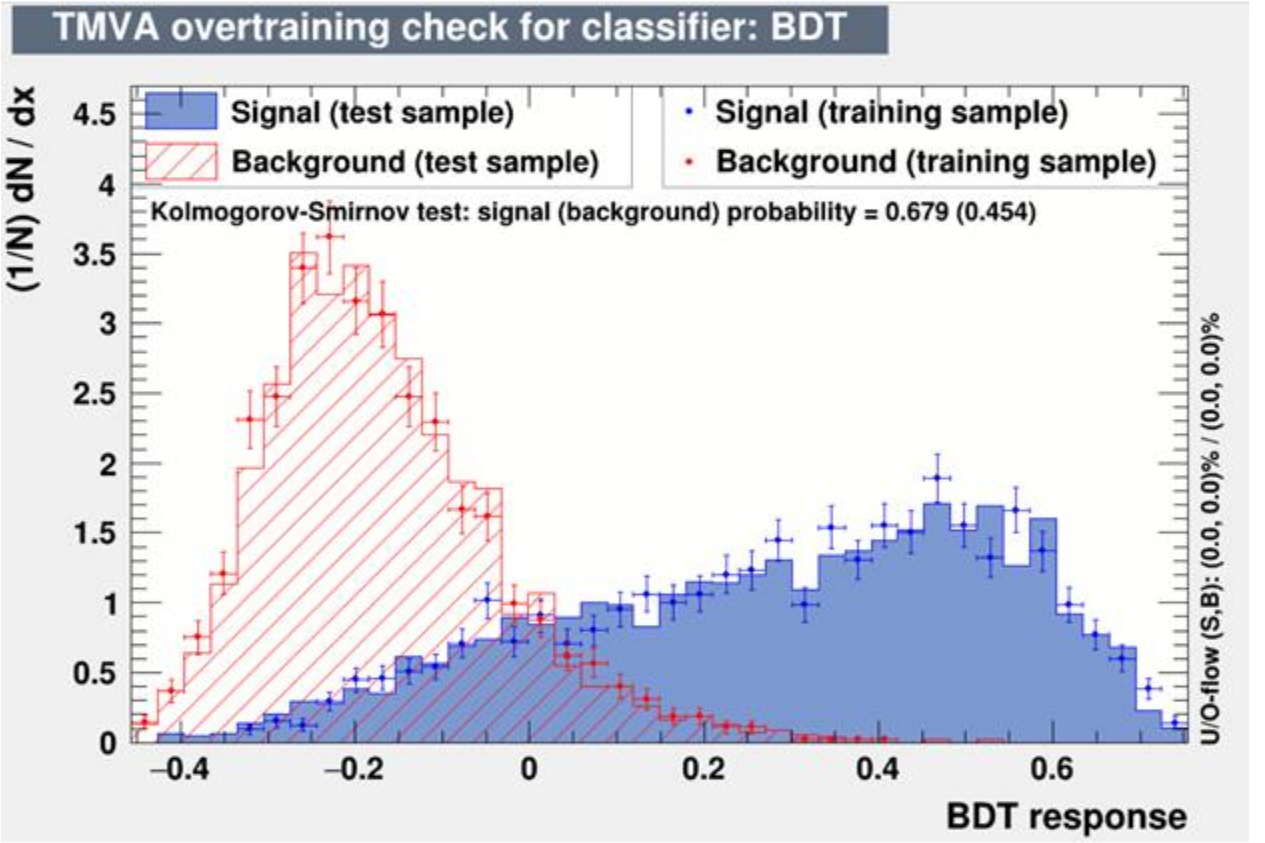


BDT parameter optimization

Maximize AUROC + check Kolmogoroff-Smirnoff test

Default values found for high granularity - due to overtraining, values for low granularity were lowered

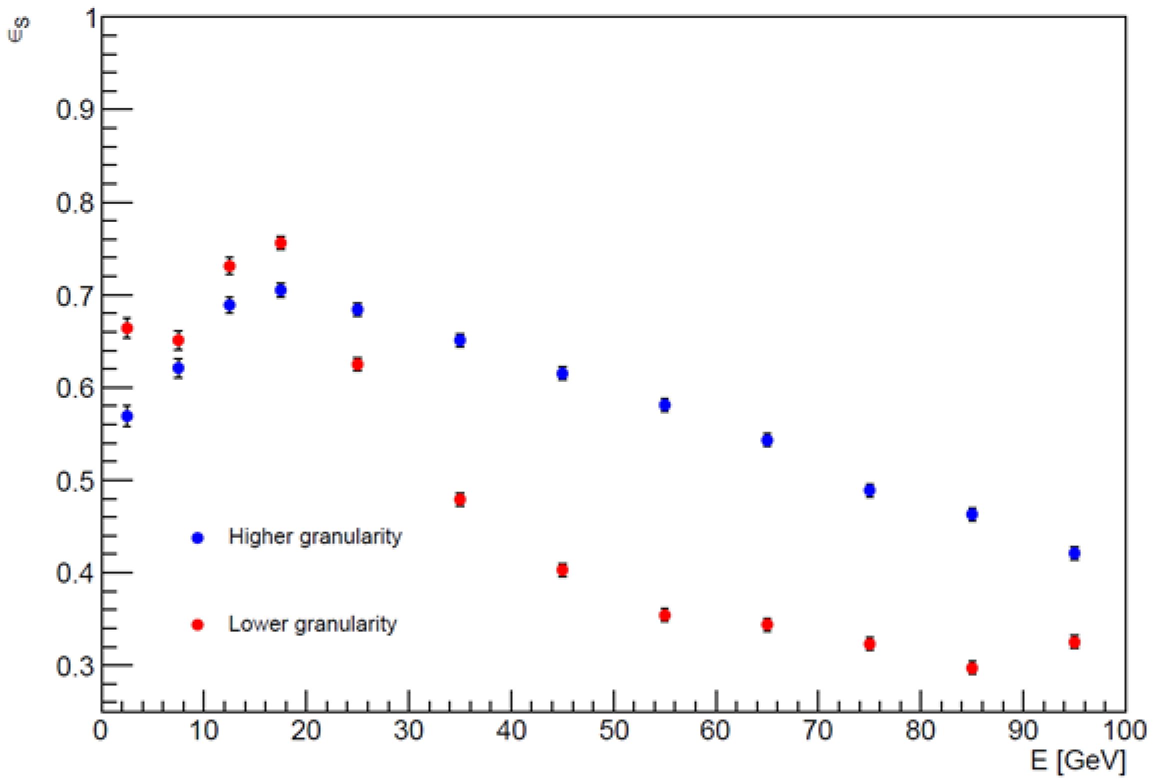
	Higher granularity	Lower granularity
NTrees	250	200
AdaBoostBeta	0.5	0.5
MinNodeSize	1.5 %	1.5 %
MaxDepth	3	2



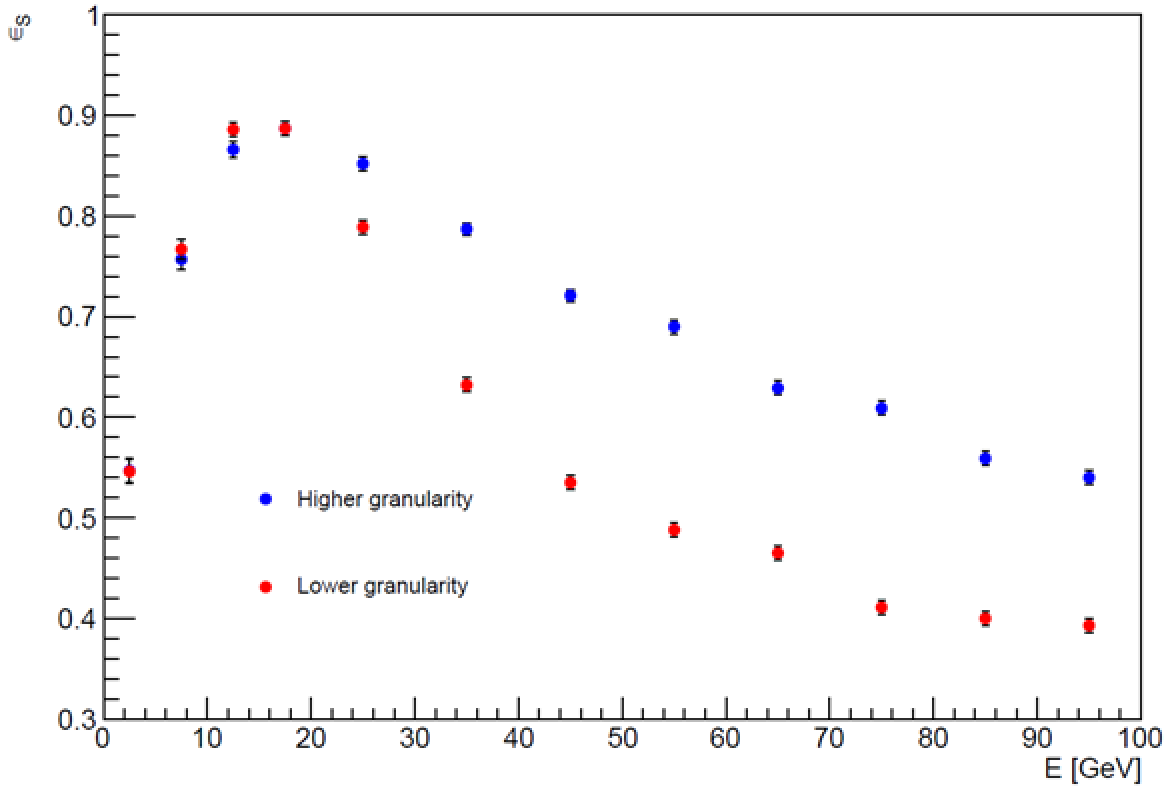
BDT results

- Higher efficiency – success, but still rapid descend below 20 GeV

Results obtained using Rectangular cuts



Results obtained using Boosted Decision Trees

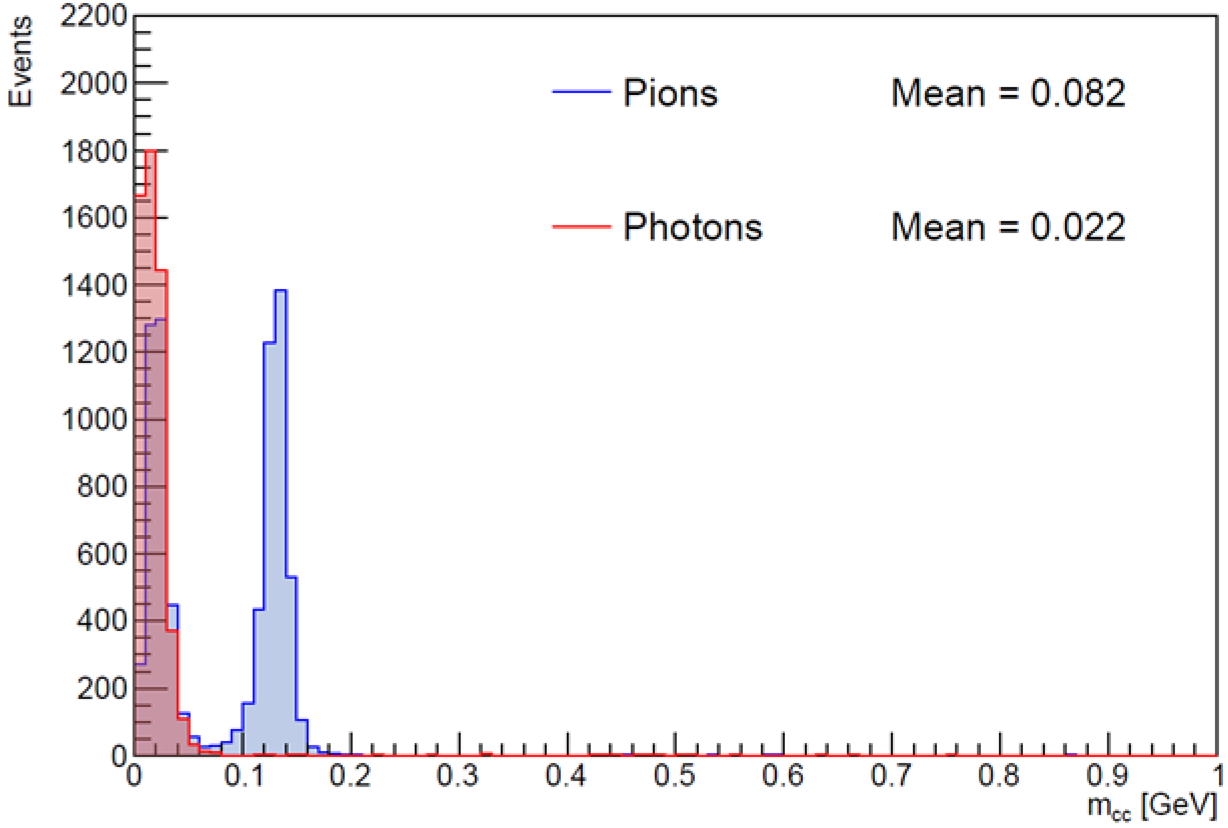


Low-energy efficiency correction

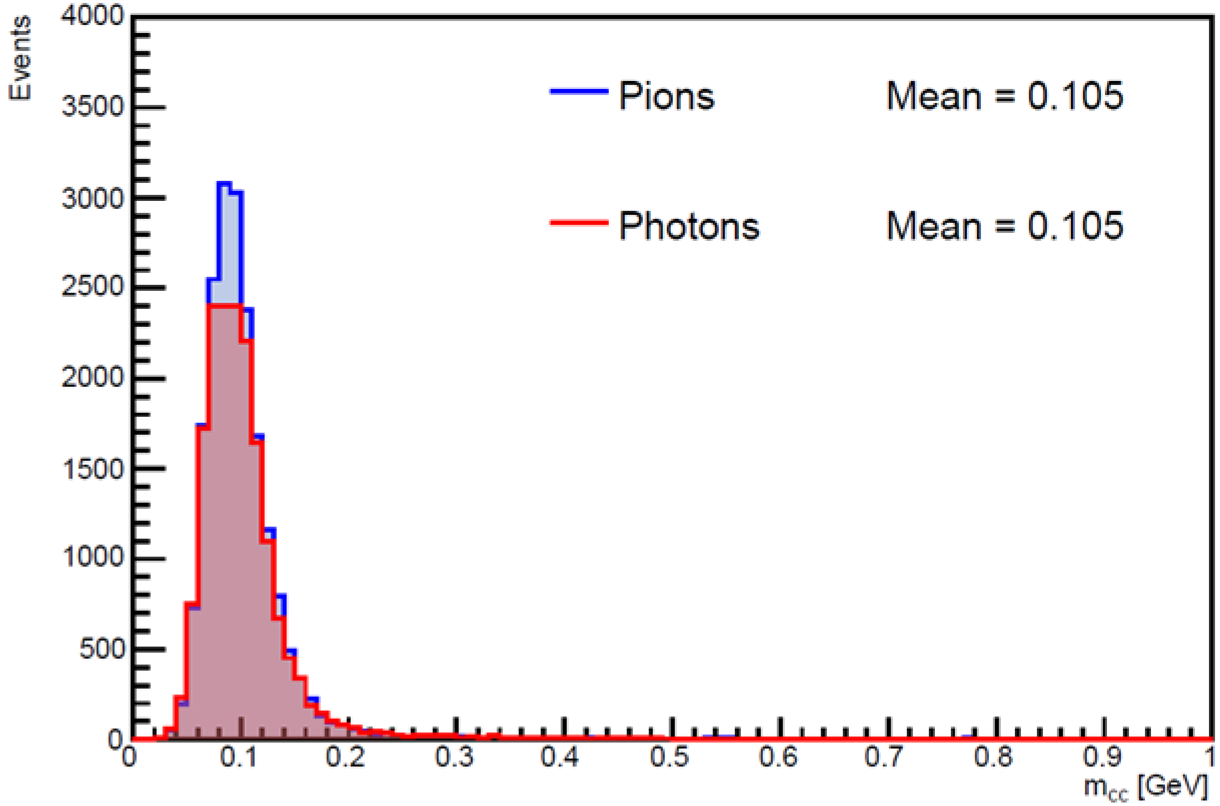
m_{cc} variable – invariant mass of the two topoclusters with highest E deposit

- Displayed very good separation at low E

E = [0; 10] GeV



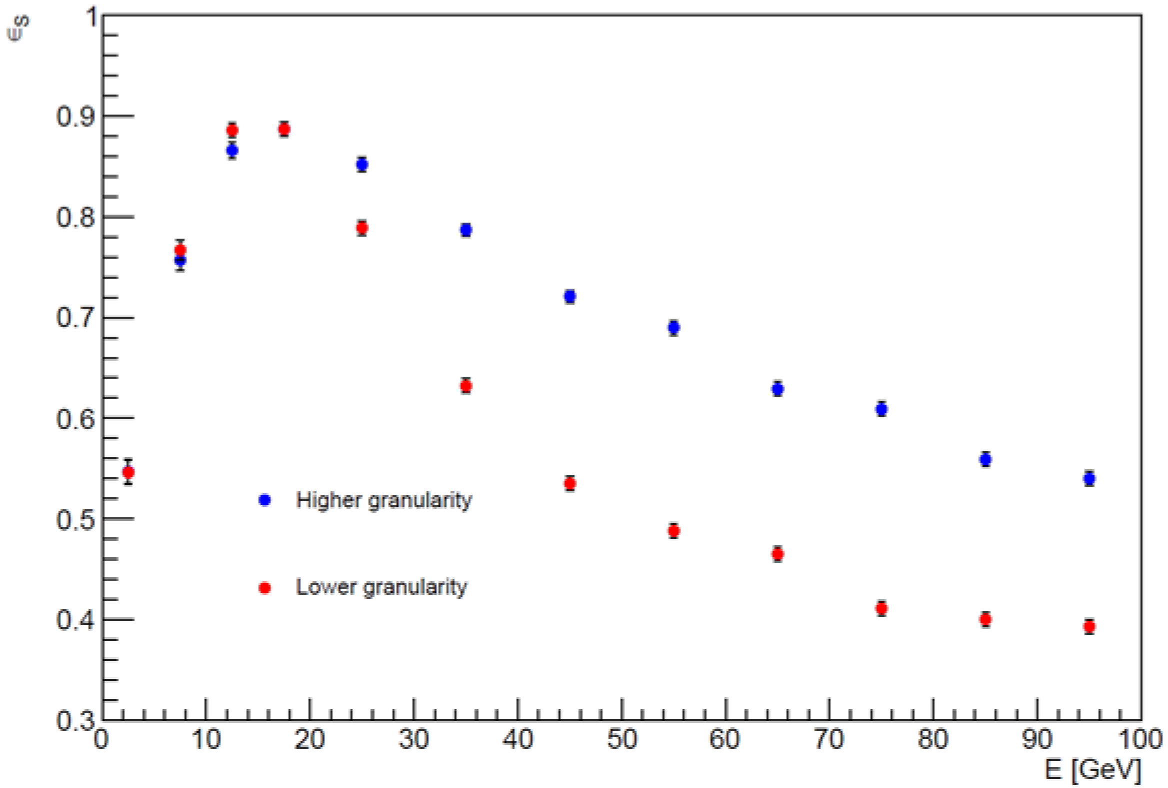
E = [50; 70] GeV



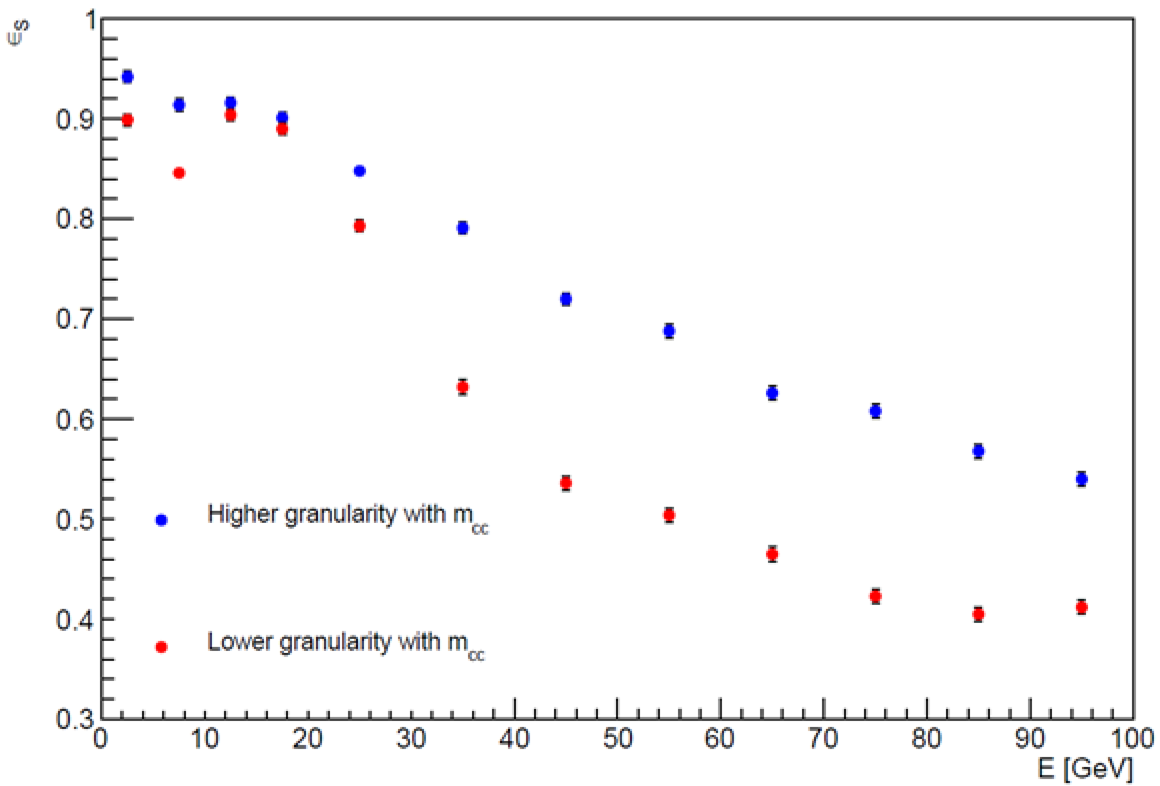
Results comparison

The efficiency at low energies remains high at approx. 0.9, drop-off eliminated

Results obtained using Boosted Decision Trees



Results obtained using Boosted Decision Trees with m_{cc}



In closing

Using MVA methods we obtained good π^0 discrimination against single γ background

By combining cell information with topological clustering we eliminated the drop in ε_S at low E

Demonstrated the importance of high granularity strip layer – instrumental for resolution of highly collimated decay products

How to follow up?

- Cluster merging & optimization
- More realistic simulated environment (varying granularity across layers)
- More advanced MVA methods (CNN)
- Improving the high energy regime

Thank you for your attention

Discriminating variable definition

- $\boldsymbol{E_{max}}$ - The energy contained in a cell with the largest energy deposit in the second layer of the calorimeter
- $\boldsymbol{E_{2max}}$ - The second largest energy deposit in the second layer of the calorimeter
- $\boldsymbol{E_{ocore}}$ - Energy deposited in cells surrounding the shower centre defined as

$$E_{ocore} = \frac{E(3) - E(1)}{E(1)}$$

where $E(n)$ is the total energy deposited in $\pm n$ cells surrounding the cell with the highest energy deposit

- $\boldsymbol{E_n}$ **-** The sum of energy contained in the first n layers of the calorimeter
- $\boldsymbol{E_{i1}}$ - Energy deposited in the *ith* layer of the calorimeter divided by energy deposited in the first layer
- $\boldsymbol{E_{iT}}$ - Energy deposited in the *ith* layer of the calorimeter divided by the total energy deposited in the calorimeter
- $\boldsymbol{W_{nl}}$ - Variable determining shower width in a calorimeter layer l, defined as a normalized sum of energy over $\pm n$ cells in the η coordinate and ± 1 cells in ϕ, weighted by the distance in the $\eta \times \phi$ space

$$W_{nl} = \frac{\sum_{i=1}^{n} E_i \times \Delta R^2}{\sum_{i=1}^{n} E_i}$$

where ΔR is defined as $\Delta R^2 = (\eta - \eta_{max})^2 + (\phi - \phi_{max})^2$

Table 1 Summary table of the expected energy resolution for the different technologies. The values are measurements where available, otherwise obtained from simulation. Those values marked with "?" are estimates since neither measurement nor simulation exists

Detector technology (ECAL & HCAL)	E.m. energy res. stochastic term	E.m. energy res. constant term	ECAL & HCAL had. energy resolution (stoch. term for single had.)	ECAL & HCAL had. energy resolution (for 50 GeV jets)	Ultimate hadronic energy res. incl. PFlow (for 50 GeV jets)
Highly granular Si/W based ECAL & Scintillator based HCAL	15 – 17 % [12, 20]	1 % [12, 20]	45 – 50 % [20, 45]	≈ 6 % ?	4 % [20]
Highly granular Noble liquid based ECAL & Scintillator based HCAL	8–10 % [24, 27, 46]	< 1 % [24, 27, 47]	≈ 40 % [27, 28]	≈ 6 % ?	3–4 % ?
Dual-readout Fibre calorimeter	11 % [48]	< 1 % [48]	≈ 30 % [48]	4–5 % [49]	3–4 % ?
Hybrid crystal and Dual-readout calorimeter	3 % [30]	< 1 % [30]	≈ 26 % [30]	5–6 % [30, 50]	3–4 % [50]