# Innovative analysis techniques

Alessio Gianelle[1], Donatella Lucchesi[1,2], Davide Nicotra[1], Lorenzo Sestini[1], Davide Zuliani[1,2]
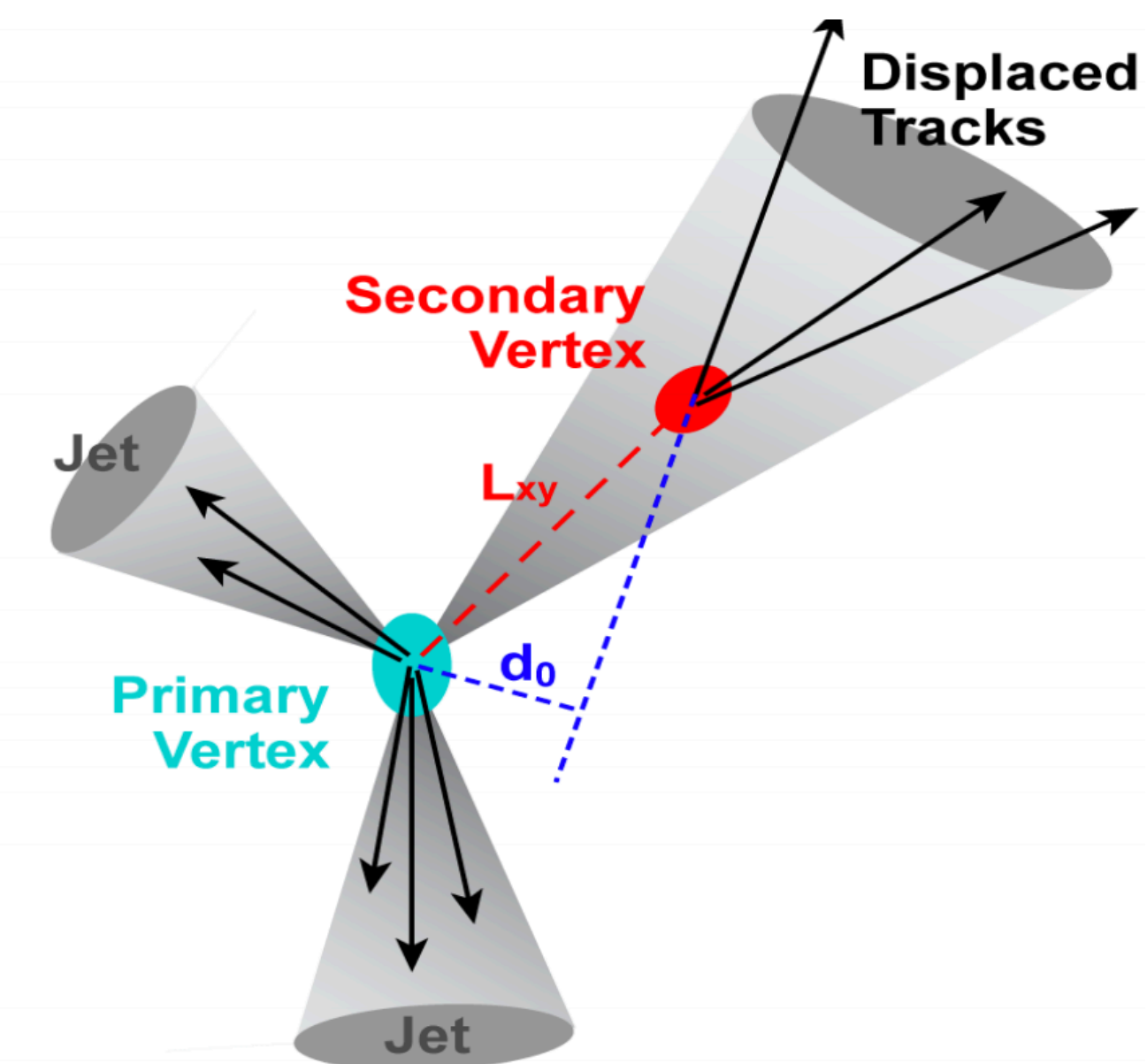
Riunione LHCb Padova — 01/12/2020

Tagging

- The task is to identify jet generated by heavy flavour quarks (b and c quark)

- Up to now LHCb relies on *Secondary Vertex tagging* (SV tagging) algorithm

- Good performances for b-tagging, room for improvement for c-tagging
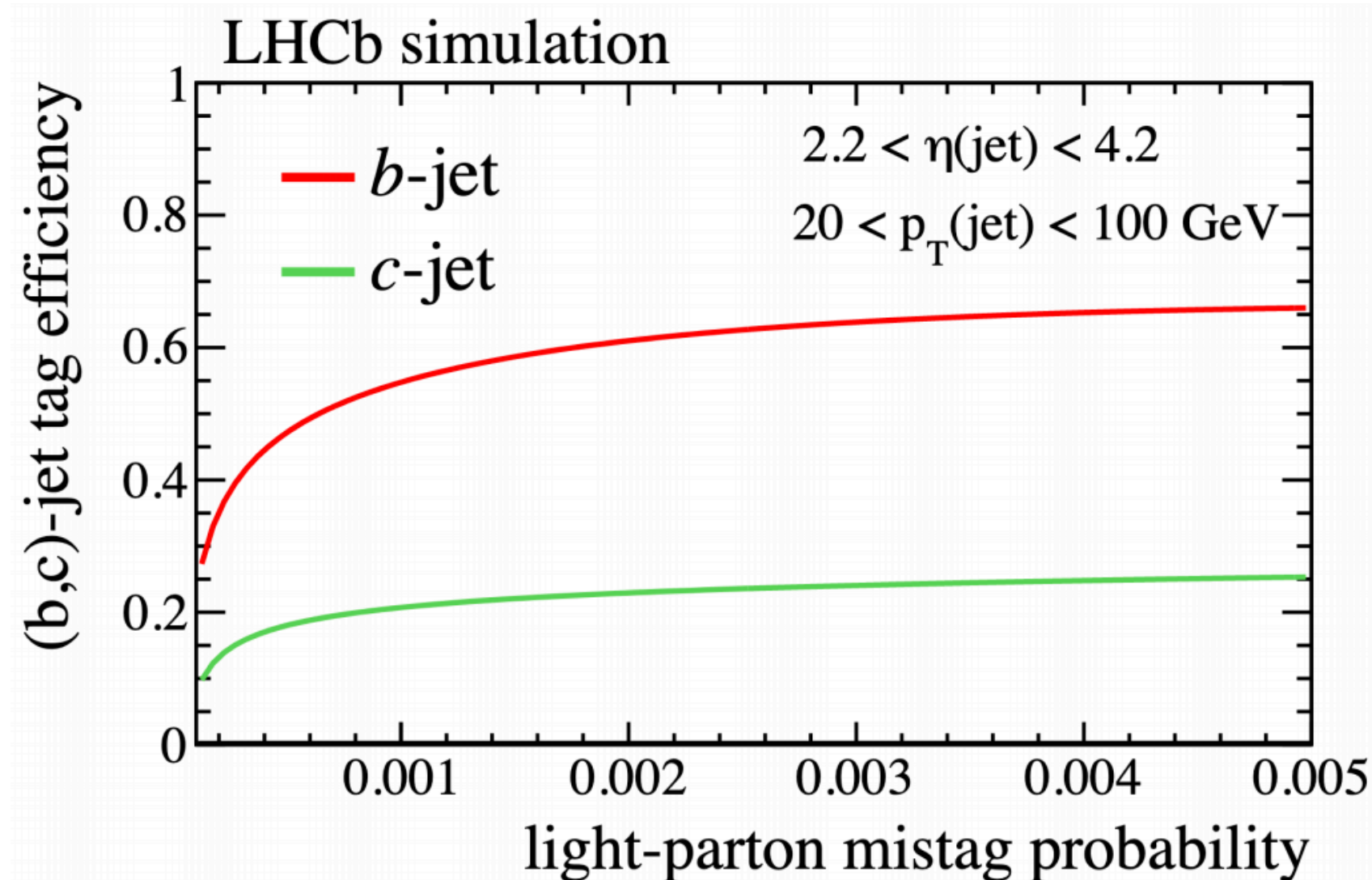
- **The jet tagging system takes advantage of LHCb features → precise vertex reconstruction!**

- A jet is identified to be generated from a **b** or **c** quark (**b-jet** or **c-jet**) if a **Secondary Vertex** is reconstructed within the jet cone (**ΔR <0.5**).

- Single tracks used to build the **Secondary Vertex** are **not required** to have **ΔR <0.5** with respect to the jet axis.

- Two **Boosted Decision Trees** are used to identify b and c jets.

BDT(bc|udsg)
To separate **heavy flavour** jets from **light** jets

BDT(b|c)
To separate **b-jets** from **c-jets**

New algorithms to be studied!

b vs. $\bar{b}$ tagging with TTN

T. Felser, A. Gianelle, D.Lucchesi, S. Montangero, M. Trenti, L. Sestini, D.Z.

- At LHCb we are interested in jets flavour tagging
- Up to now classical ML techniques are used (*Boosted Decision Trees*)
- We developed a quantum-inspired Tree Tensor Network and measured its performances
  - Comparison with state-of-the-art DNN
  - Timing performances (possible real time application)
  - Understanding the physics behind jets
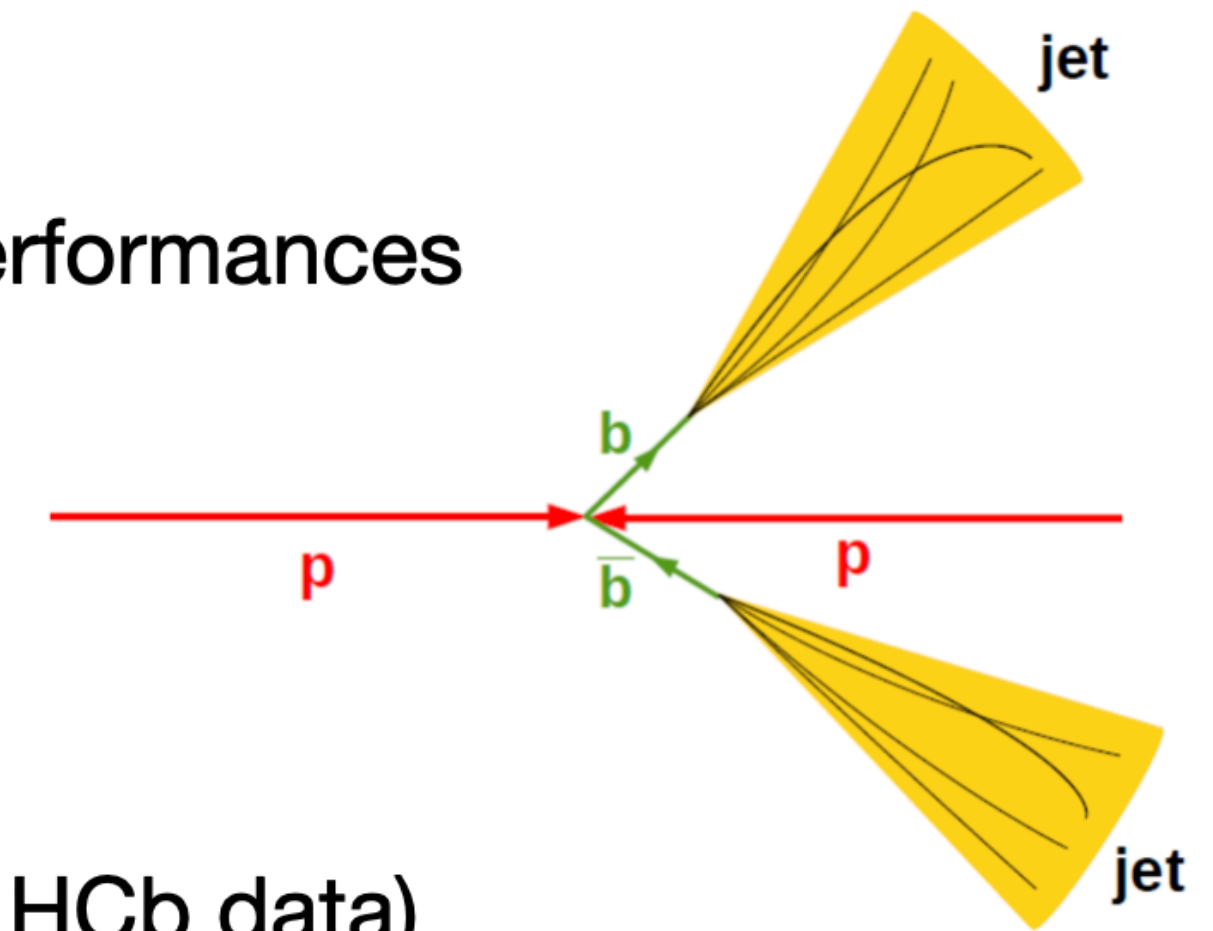
- In our study case we tried to tag $b-$ and $\bar{b}-$jets
- The main idea is to use this new techniques to measure the asymmetry (on LHCb data)
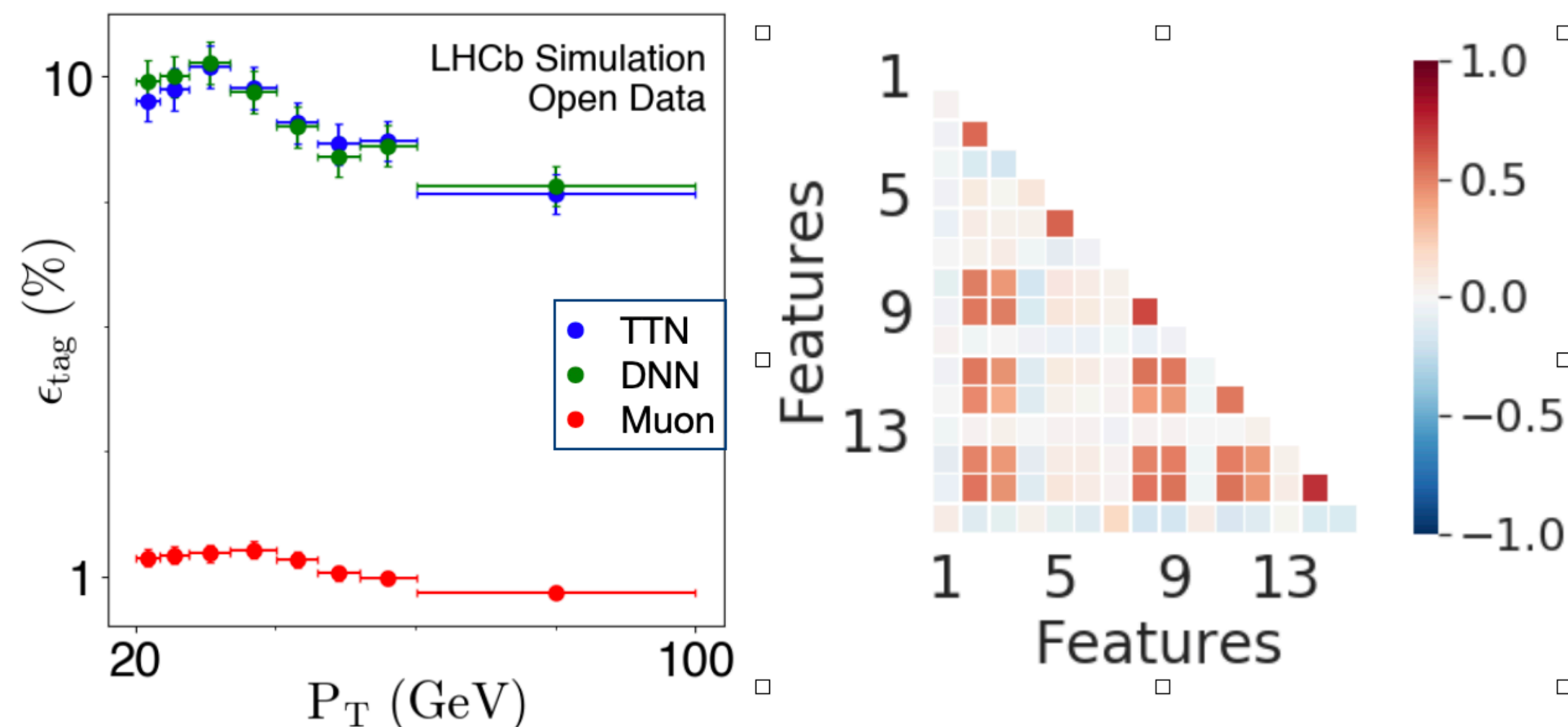
- Comparable performances with DNN
- Possibility to measure correlations and entropy between variables
- Lower prediction times (possible to reach MHz rates)

- A pre-print is available (arXiv:2004.13747)
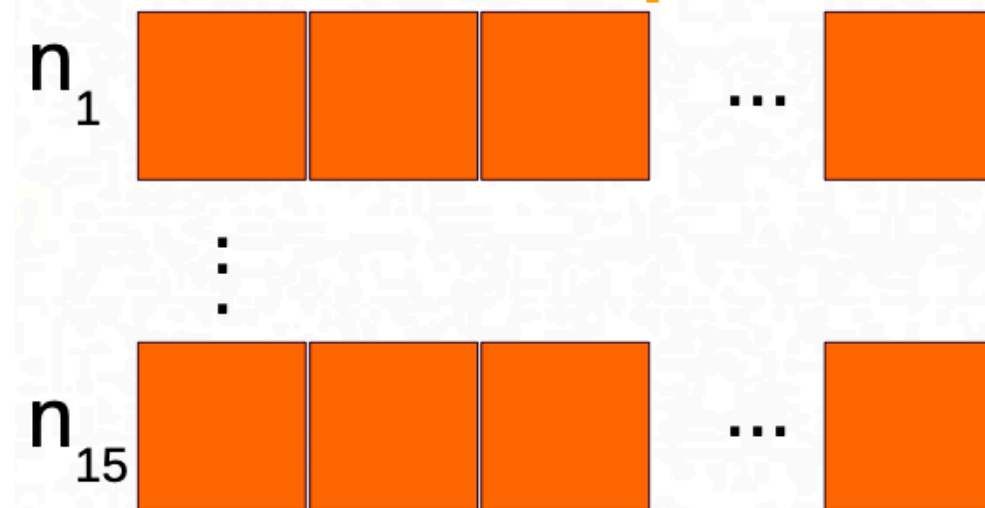- The paper has been submitted to journal

- **Matrix with jets constituents as input (inspired by CMS Deep Tagging)**

Charged particles: 16 features

c_1 [ ][ ][ ] ... [ ][ ][ ][ ][ ]

c_{10} [ ][ ][ ] ... [ ][ ][ ][ ]
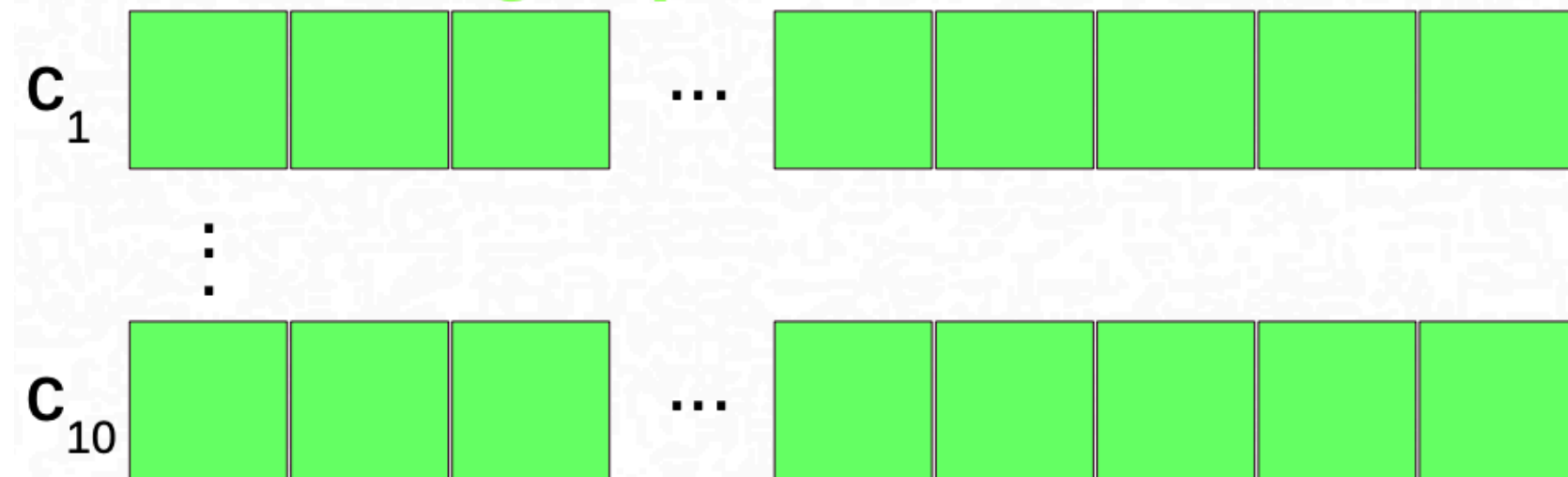
Neutral particles: 12 features

n_1 [ ][ ][ ] ... [ ]

n_{15} [ ][ ][ ] ... [ ]

Secondary vertex: 14 features

[ ][ ][ ] ... [ ][ ][ ][ ]

**Global observables: 7 features**

[ ][ ][ ][ ][ ][ ][ ]
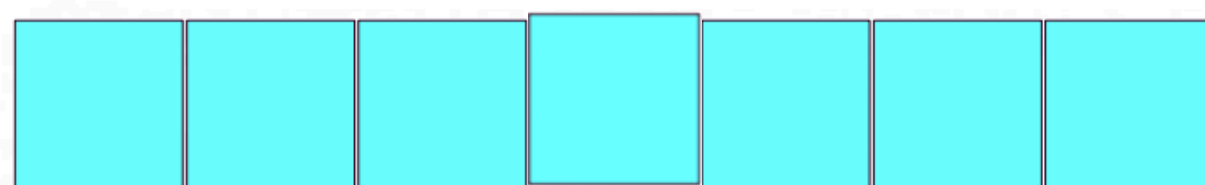
- Tracks, calorimeter clusters, and SV. **Particle ID informations included**

- Fixed number of particles: when a charged, neutral or SV is not present all its variables are set to 0.

- LSTM tecnique is used to exploit correlations among particles.

- Output of LSTM is given to a DNN.

- Ternary classification with 3 probabilities p_b, p_c, p_q.

- Keras and Tensorflow are used.

- **First results are promising, network optimization ongoing.**

Receiver operating characteristic

efficiency (TPR) vs mistag (FPR)

- ROC b Vs c curve (area = 0.8355)
- ROC b Vs q curve (area = 0.9513)
- ROC c Vs q curve (area = 0.8226)
- ROC bc Vs q curve (area = 0.9207)

probability distribution for $b\bar{b}$ di-jet events
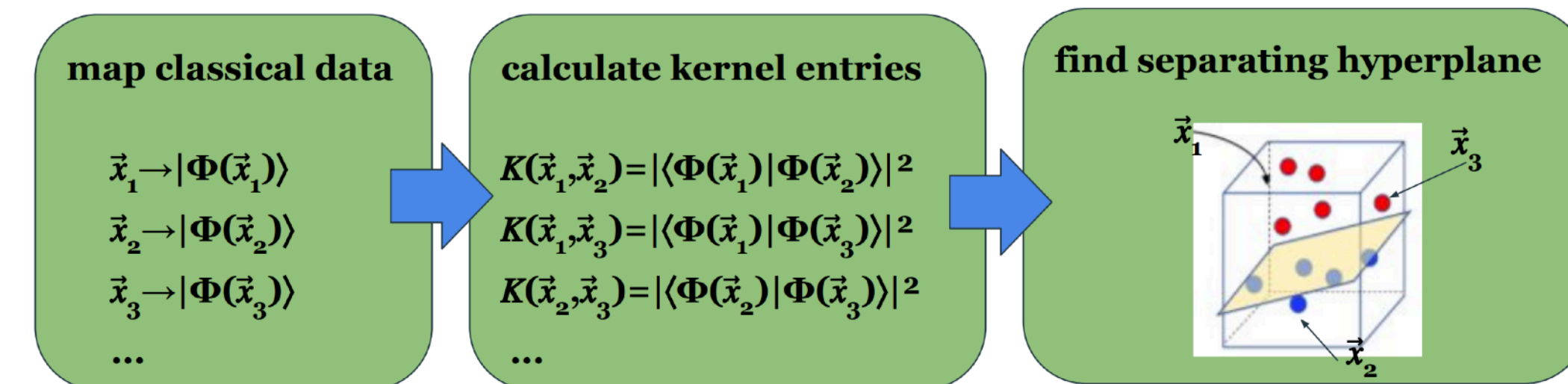
bb
cc
qq

Probability
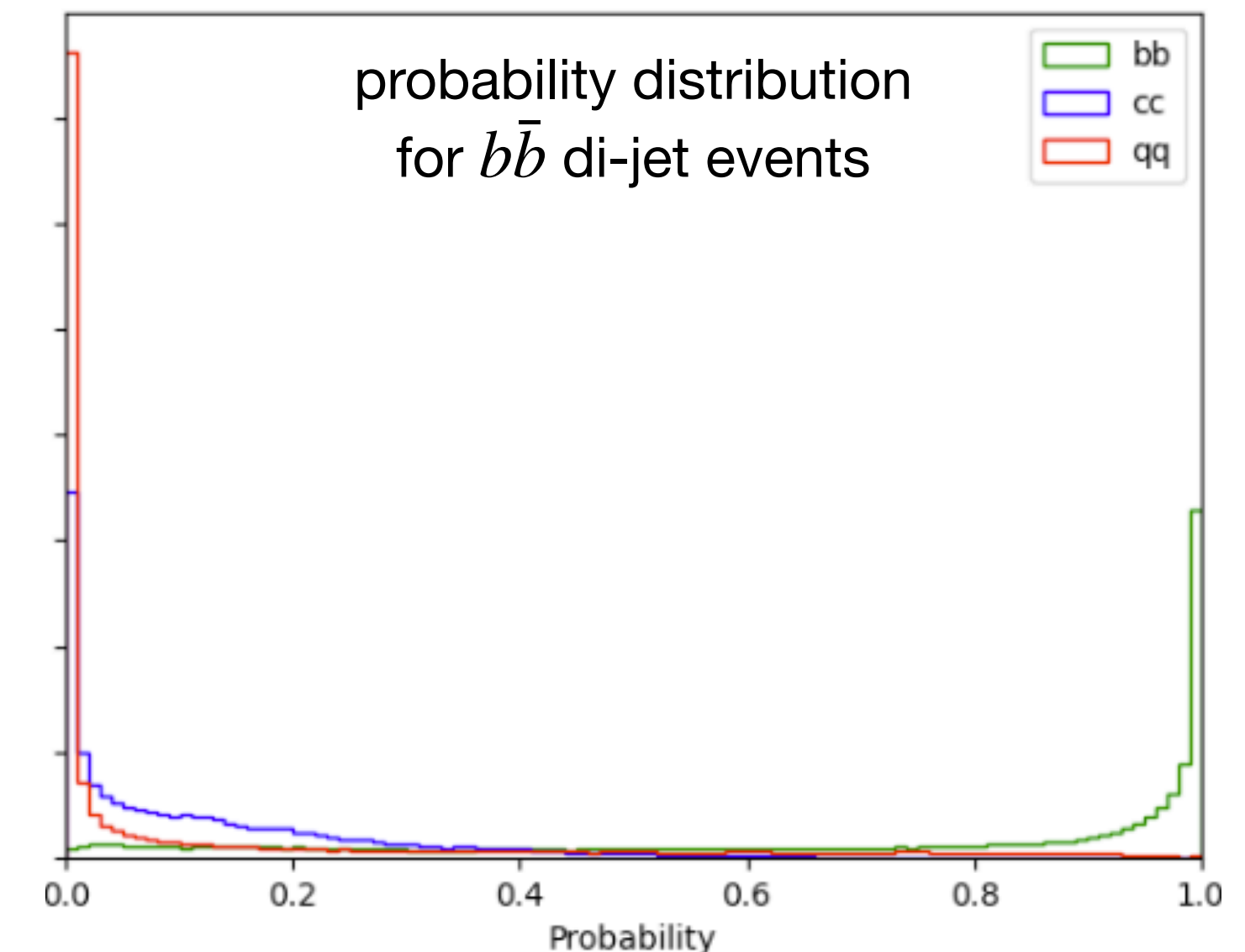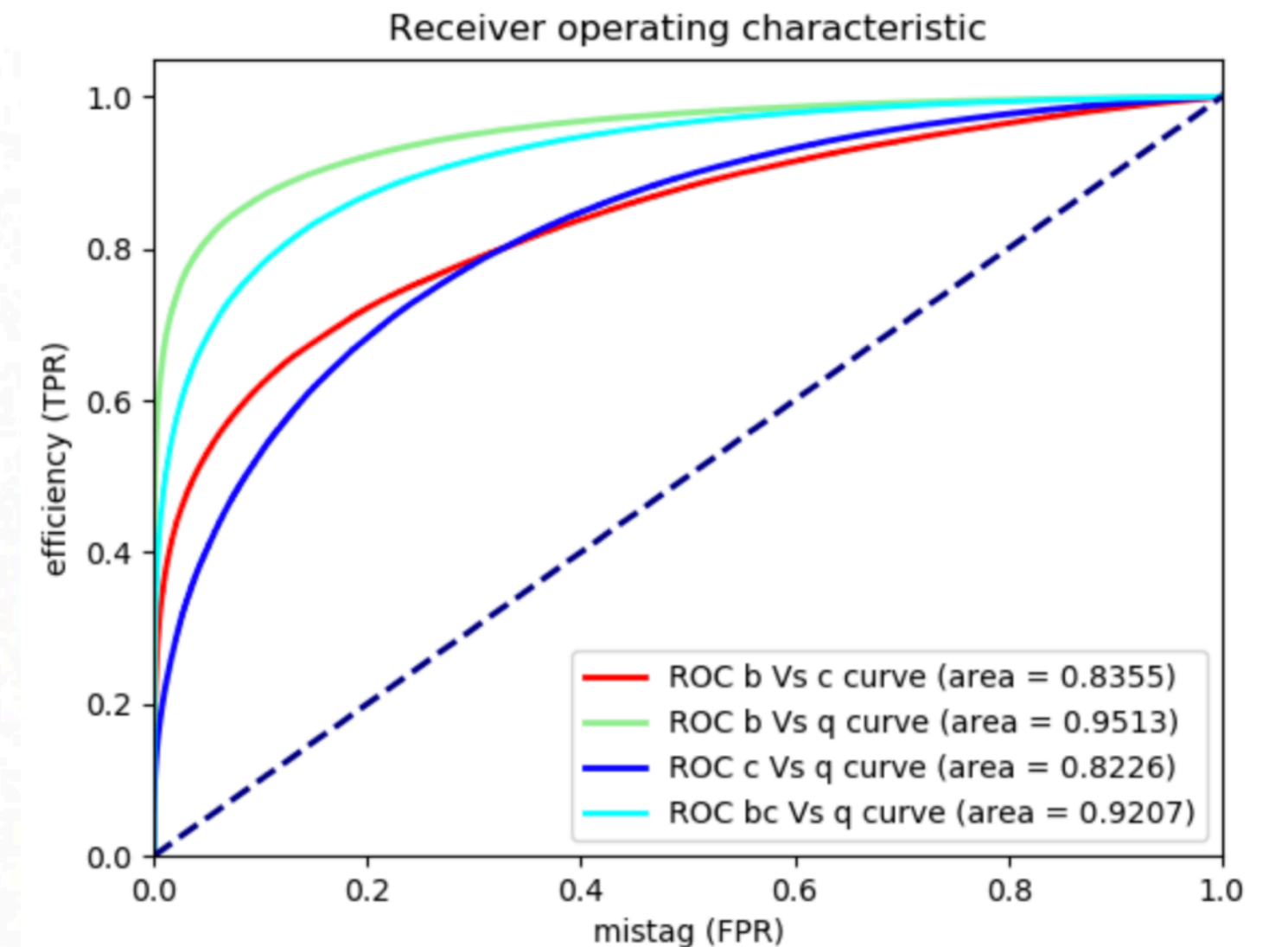
- The aim is to study tagging algorithms that can run on quantum simulations and (possibly) devices
- First time tagging is considered for quantum algorithms
- b vs. $\bar{b}$ is considered as *use case*
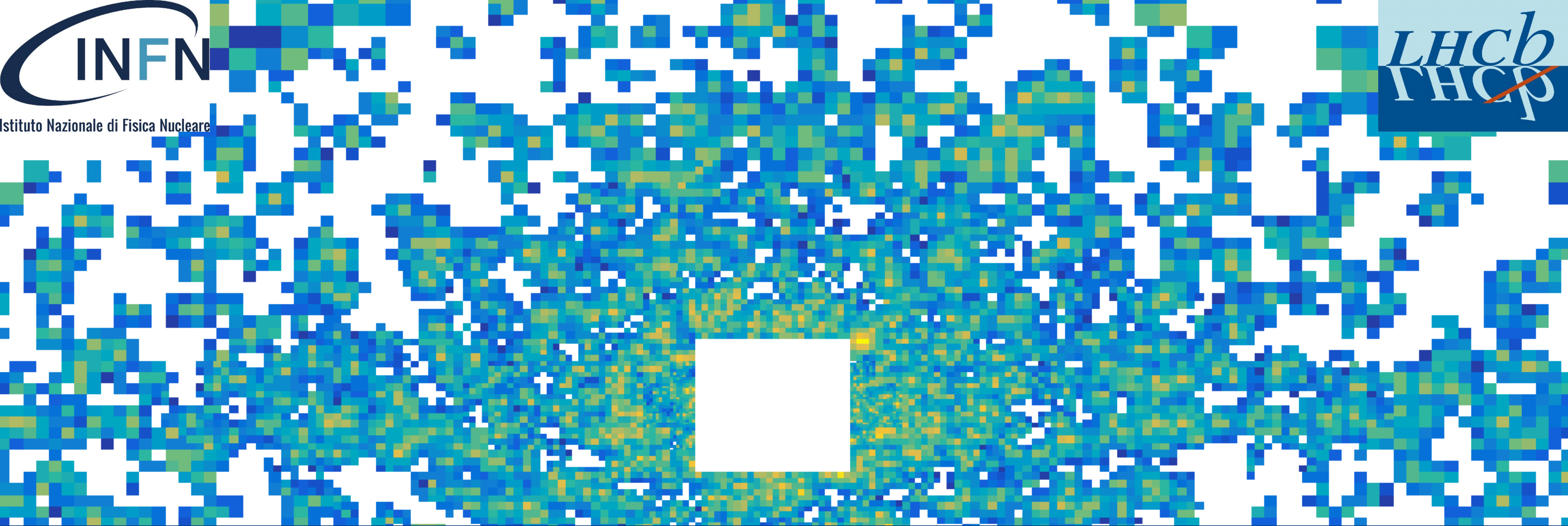
Variational Quantum Classifier (VQC)
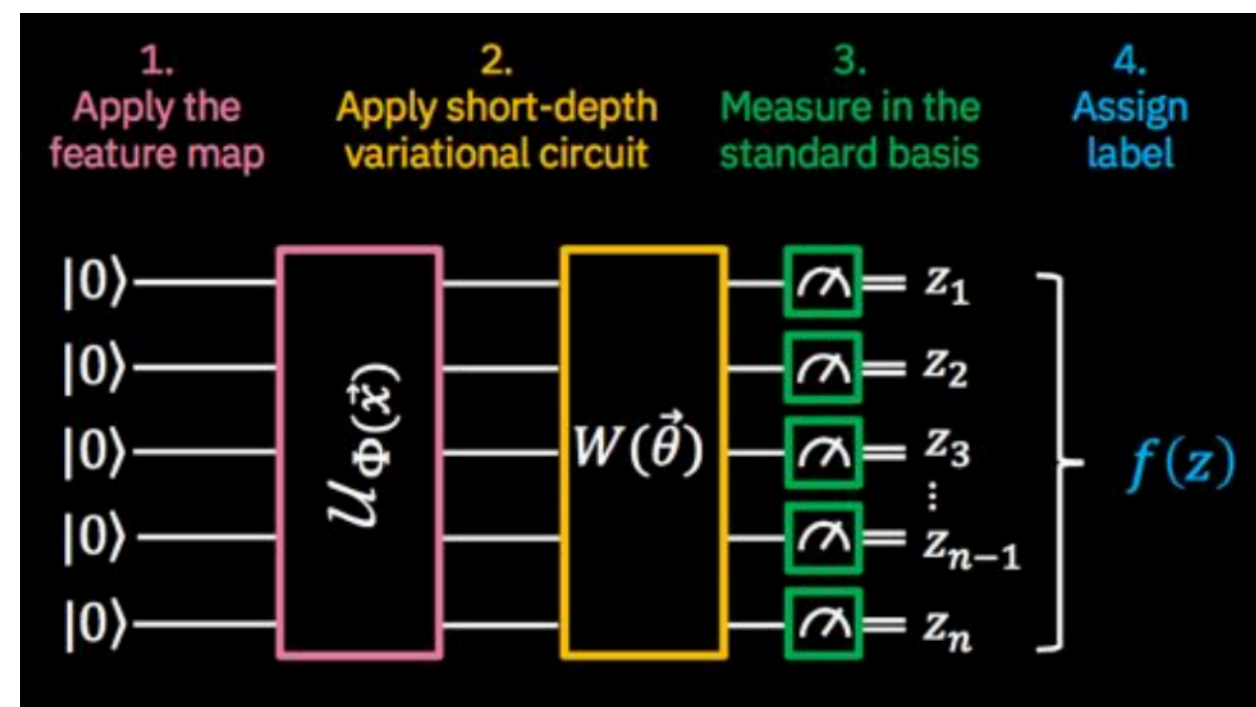
Quantum Support Vector Machine (QSVM)

QSVM Kernel method

- Davide Nicotra (master student) will work on this topic for his master thesis
- This activity is also managed in the LHCb DPA WP4 (directed by Donatella)
- Strict correlation with Quantum Technologies Initiative by CERN OpenLab
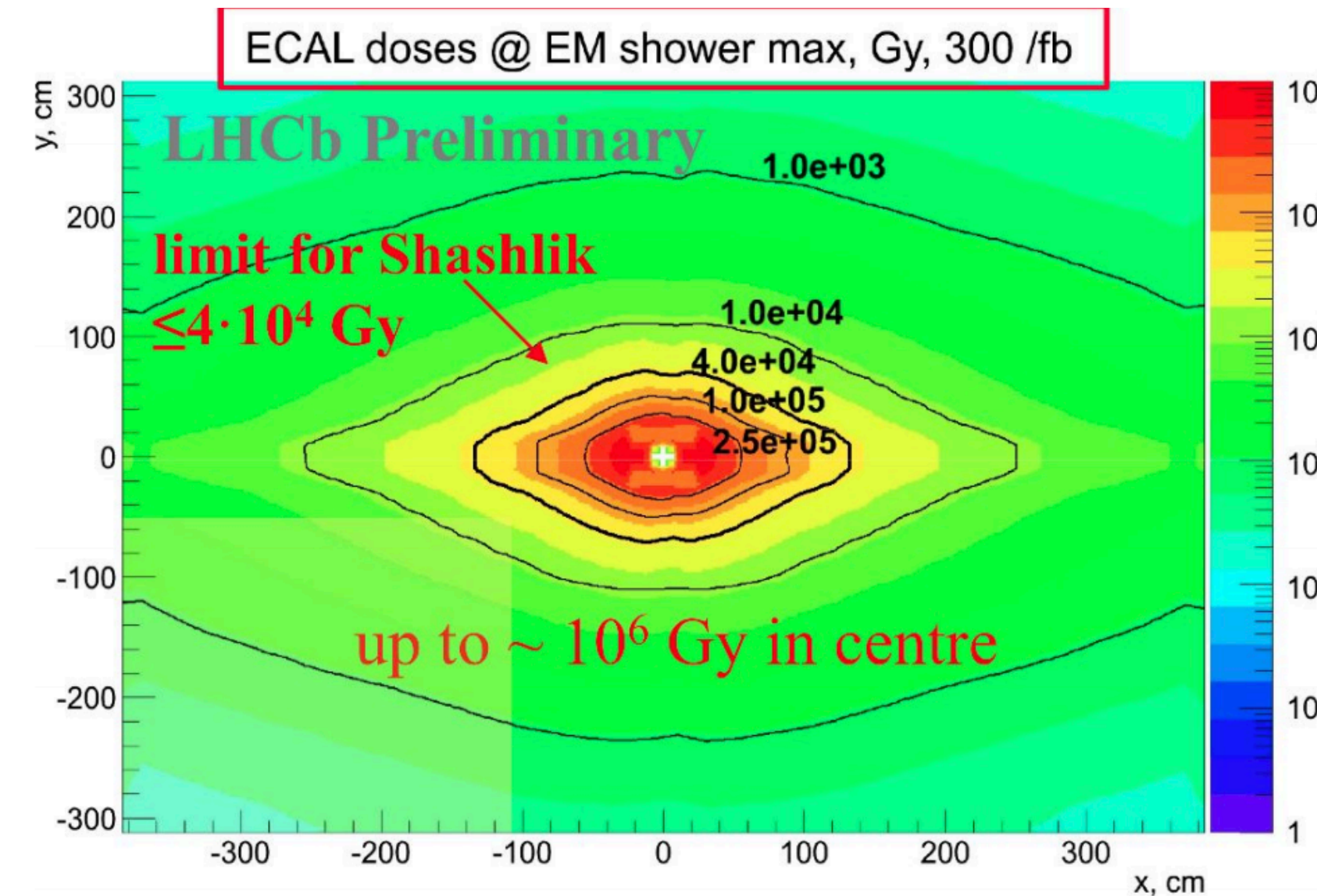
Calorimeter Upgrade

Alessio Gianelle[1], Donatella Lucchesi[1,2], Lorenzo Sestini[1], Davide Zuliani[1,2]

Riunione LHCb Padova — 01/12/2020

[1] INFN, [2] University of Padova

ECAL upgrade for Upgrade 2 ([WIP](#))

- During Run 5 the ECAL will be redesigned
  - High luminosity
  - High radiation on sensors

ECAL doses @ EM shower max, Gy, 300 /fb

LHCb Preliminary

limit for Shashlik
≤4·10⁴ Gy

up to ~ 10⁶ Gy in centre

- Therefore the following requirements are needed:
  - New geometry (SPACAL, Shashlik)
  - New cells dimension
  - New materials

- Our main focus is to study high p_T physics to find suitable ECAL parameters
  - $Z \to e^+ e^-$ is used as main channel
  - Possible figures of merit: $m_{inv}(e^+ e^-)$, σ_E, σ_{p_T}
  - Test different energy and time resolutions
  - Study Bremsstrahlung recovery
  - Study ADC saturation

Shashlik

Spaghetti

- Studies are ongoing
- Test-beam @ DESY currently ongoing
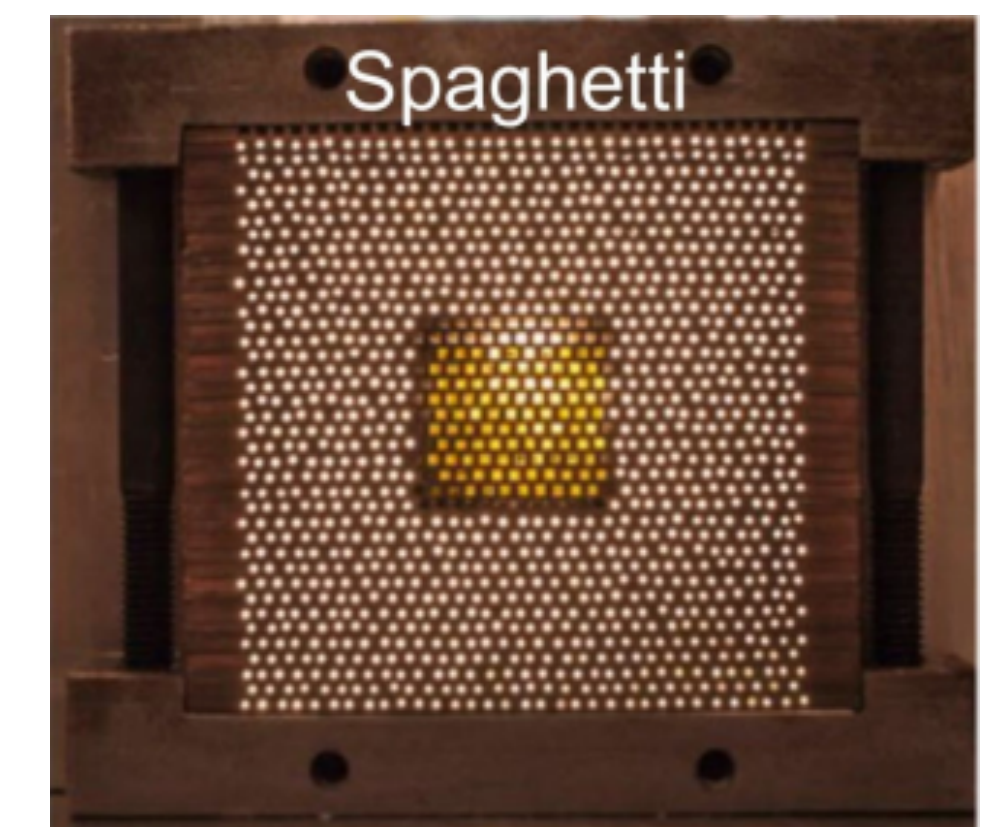
ECAL upgrade for Upgrade 2 ([WIP](#))

- Du

Current calorimeter:
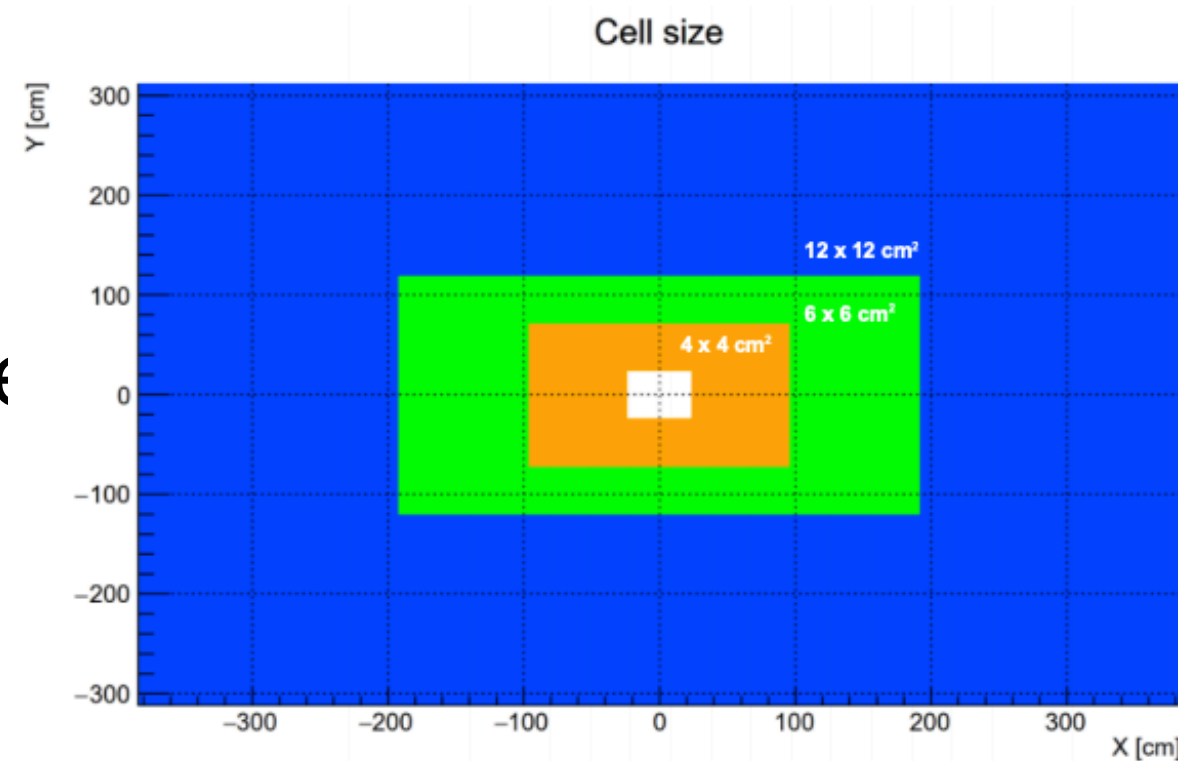
Option 1:
(32 new modules)

Option 2:
(32+144 new modules)

- The

- Our main focus is to study high p_T physics to find suitable ECAL parameters

  - $Z \to e^+e^-$ is used as main channel

  - Possible figures of merit: $m_{inv}(e^+e^-)$, σ_E, σ_{p_T}

  - Test different energy and time resolutions
  - Study Bremsstrahlung recovery
  - Study ADC saturation

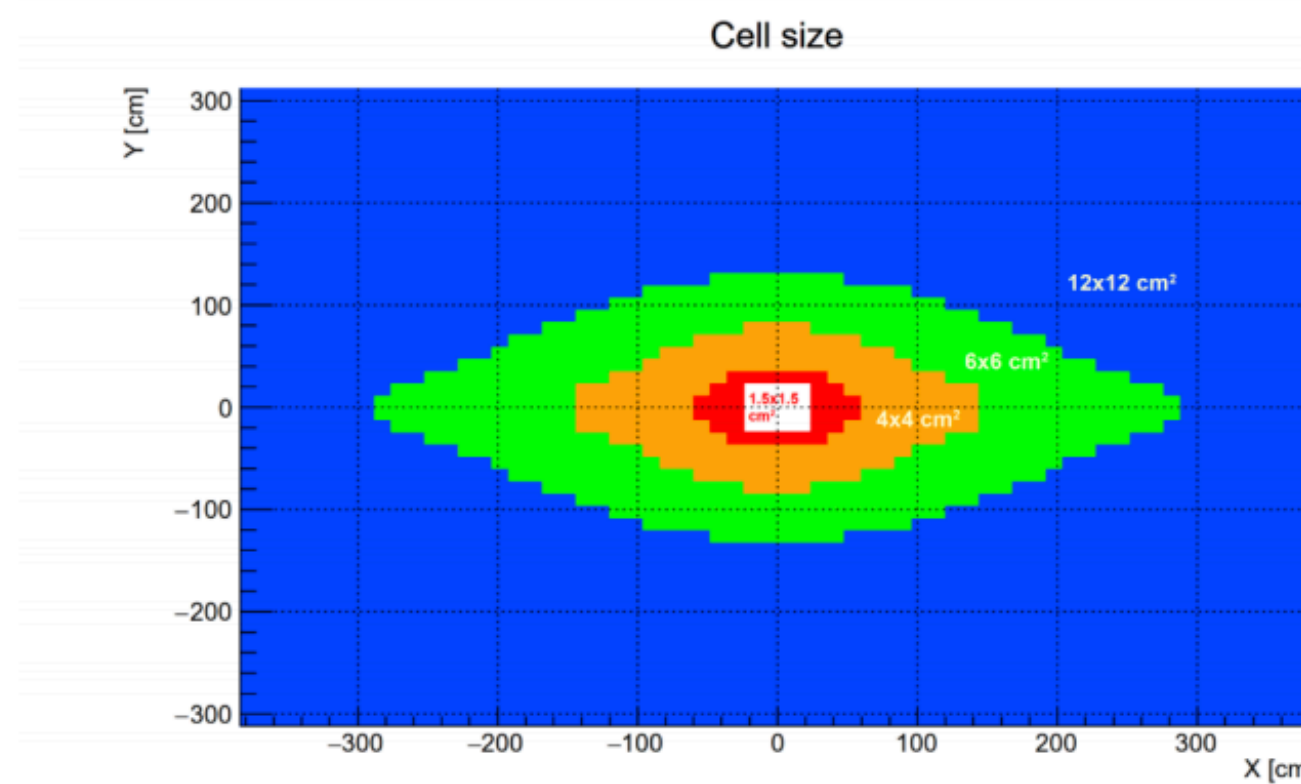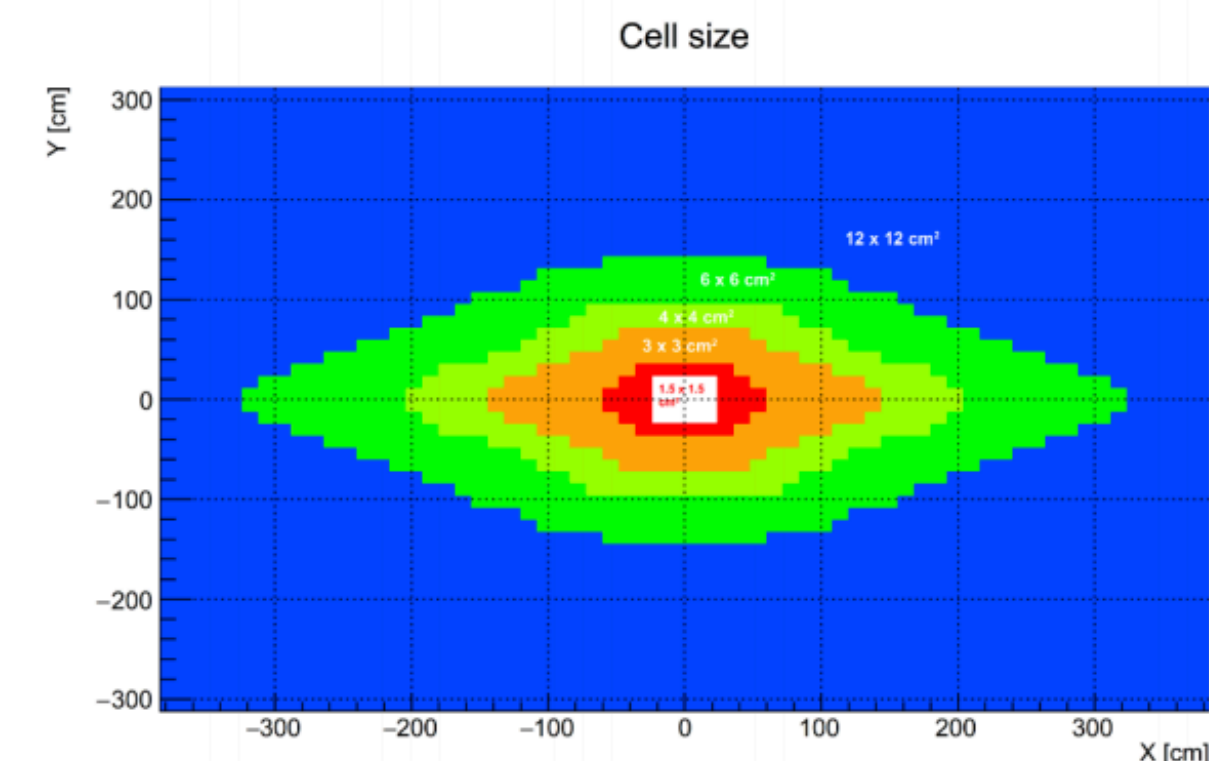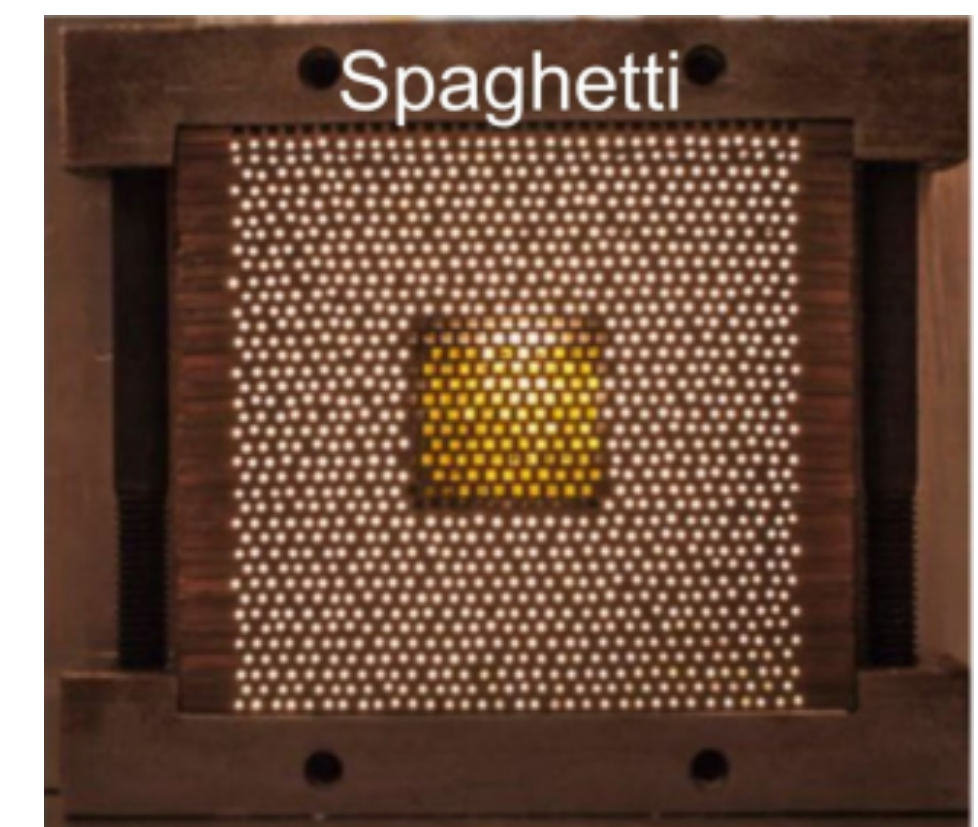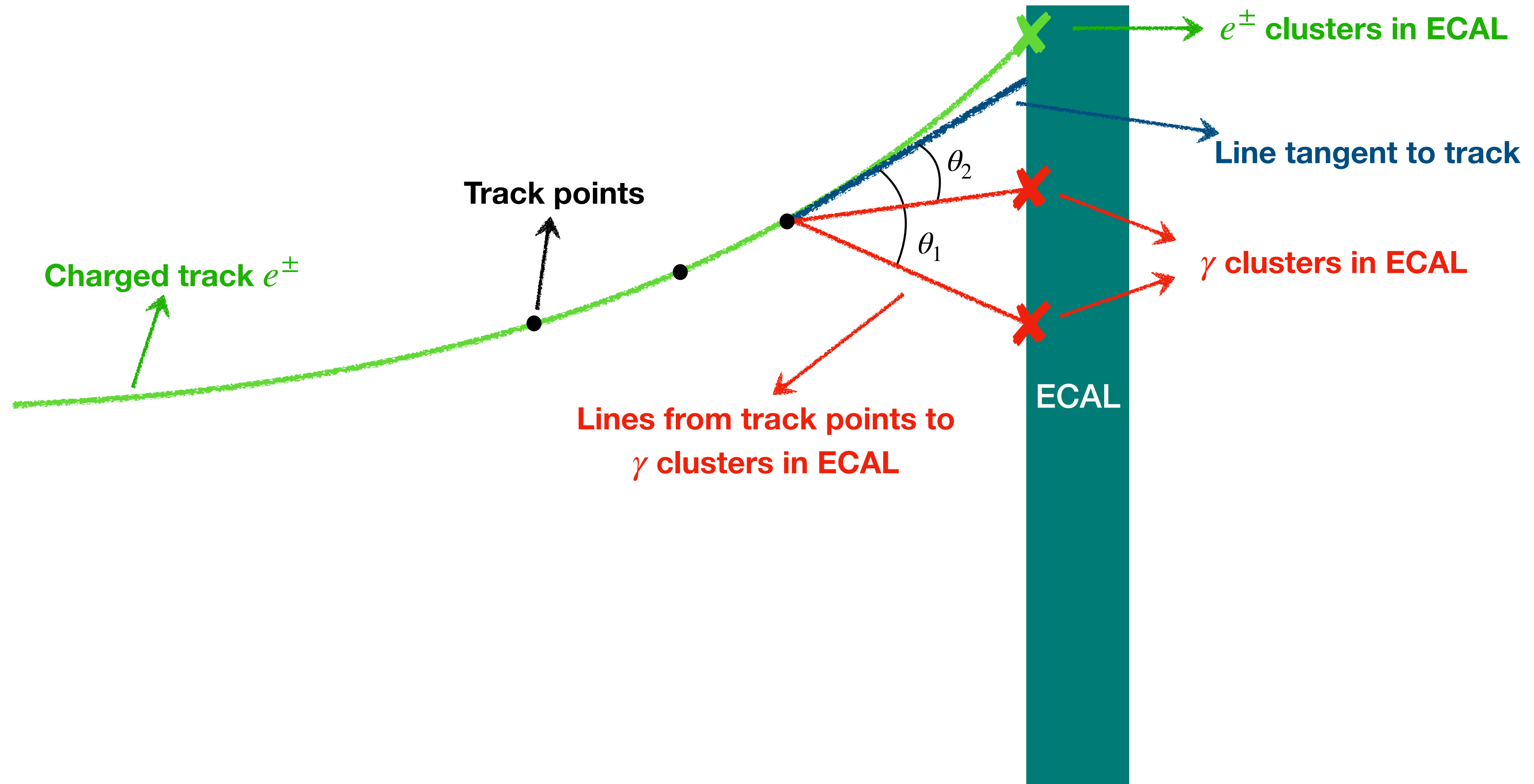Shashlik

Spaghetti

- Studies are ongoing
- Test-beam @ DESY currently ongoing

$e^{\pm}$ **clusters in ECAL**

Line tangent to track

θ_2

Track points

θ_1

Charged track $e^{\pm}$

γ **clusters in ECAL**

ECAL

Lines from track points to
γ **clusters in ECAL**
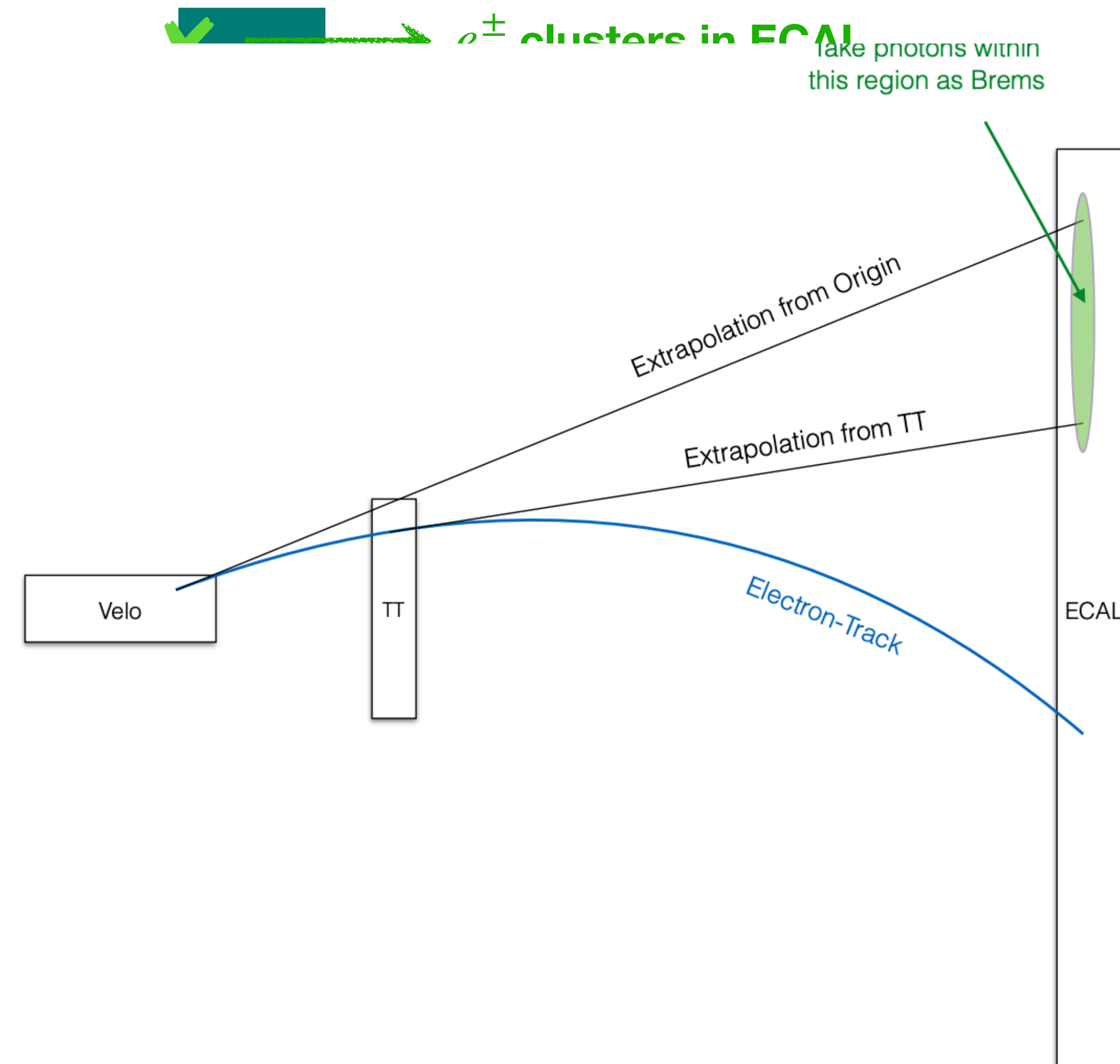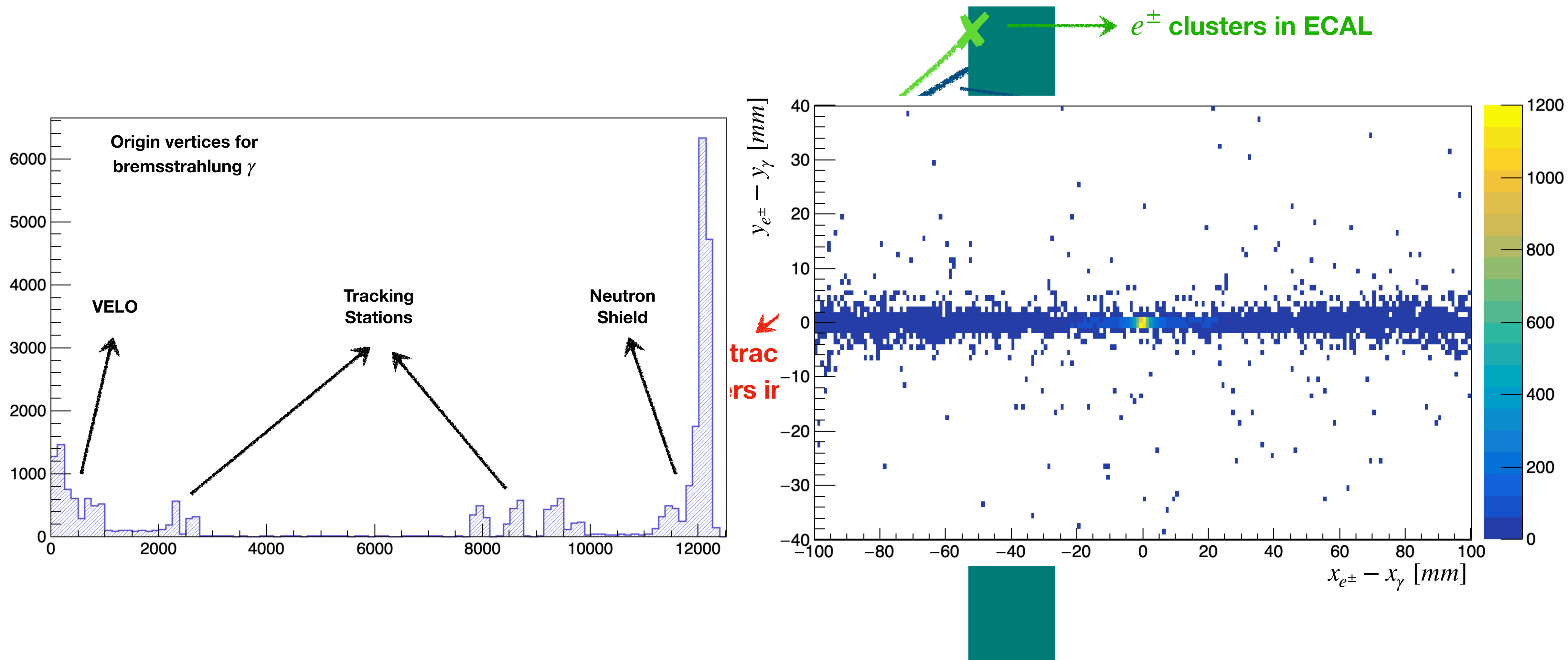
BASELINE LHCb PROCEDURE

LHCb-INT-2019-027

74 For this, a tool called `BremAdder` is used. It adds the 4-momenta of reconstructed photons
75 in the ECAL which match a reconstructed electron. The definition of "matching" is
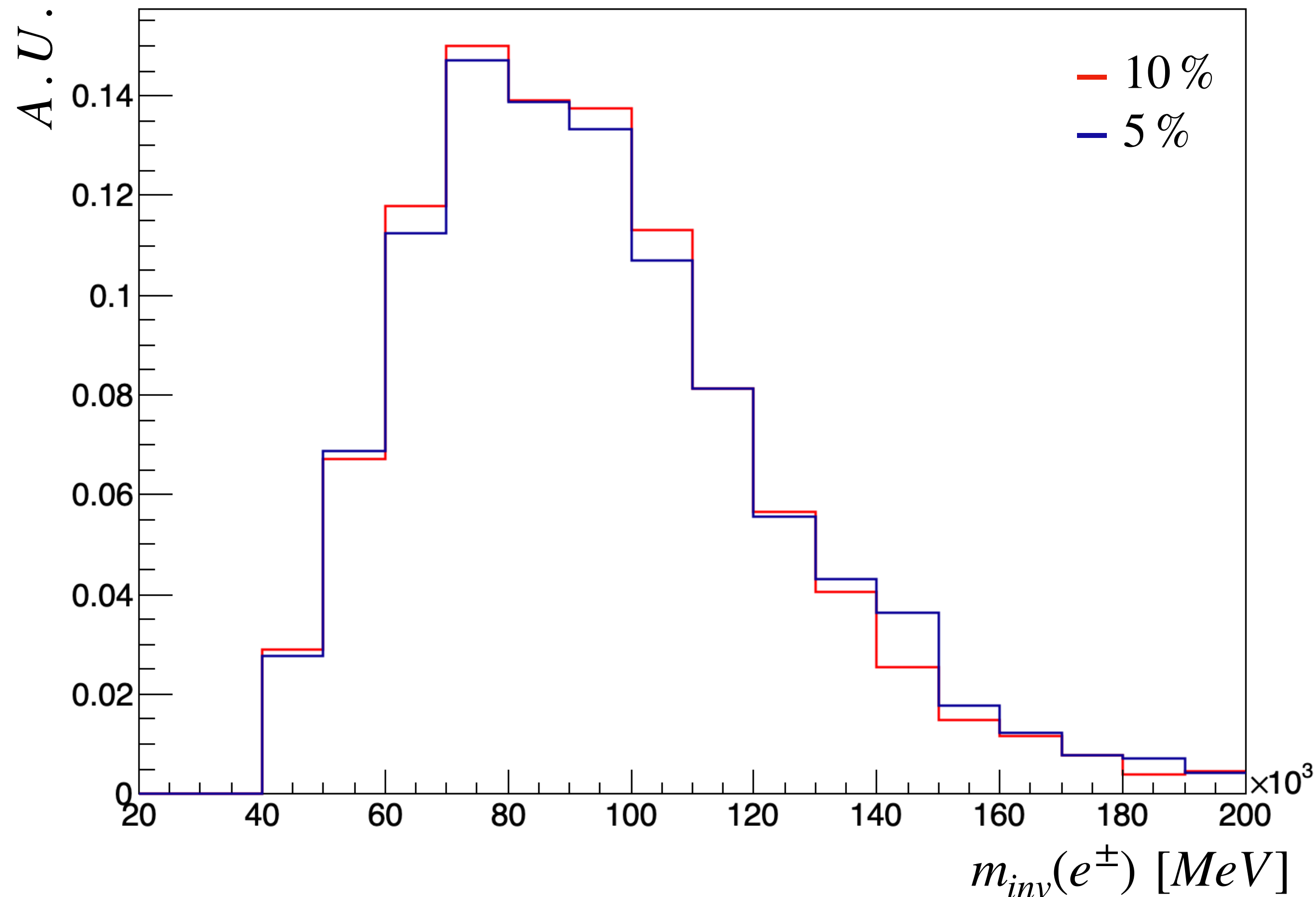76 illustrated in figure 4 and described in the following paragraph:

77 The fitted electron track is linearly extrapolated to the ECAL from two points: the starting
78 point (origin vertex) of the track and the intersection of the track with the end of the
79 "Tracker Turicensis" (TT). 4-momenta of photons that are reconstructed in the ECAL
80 around the connection line between the endpoints of the two extrapolated lines (green
81 area) are added to the 4-momentum of the electron. The width of this area is defined by
82 the 2σ interval of the combined error of the extrapolation and the position-estimation
83 of a given photon in the calorimeter. Thus, the exact width can slightly vary between
84 different photons that are currently considered as being bremsstrahlung. Since only energy
85 depositions are measured within the ECAL, the momentum direction is reconstructed
86 by assuming the photon comes from the primary vertex and its momentum direction is
87 pointing to the 3D barycenter of the shower in the calorimeter [7].

$e^{\pm}$ **clusters in ECAL**

Take photons within
this region as Brems

Extrapolation from Origin

Extrapolation from TT

Electron-Track

Velo

TT

ECAL

Bremsstrahlung recovery ([WIP](#))

$e^{\pm}$ **clusters in ECAL**

Origin vertices for bremsstrahlung γ

VELO

Tracking Stations

Neutron Shield

trac
rs ir

$y_{e^{\pm}} - y_{\gamma}$ [mm]

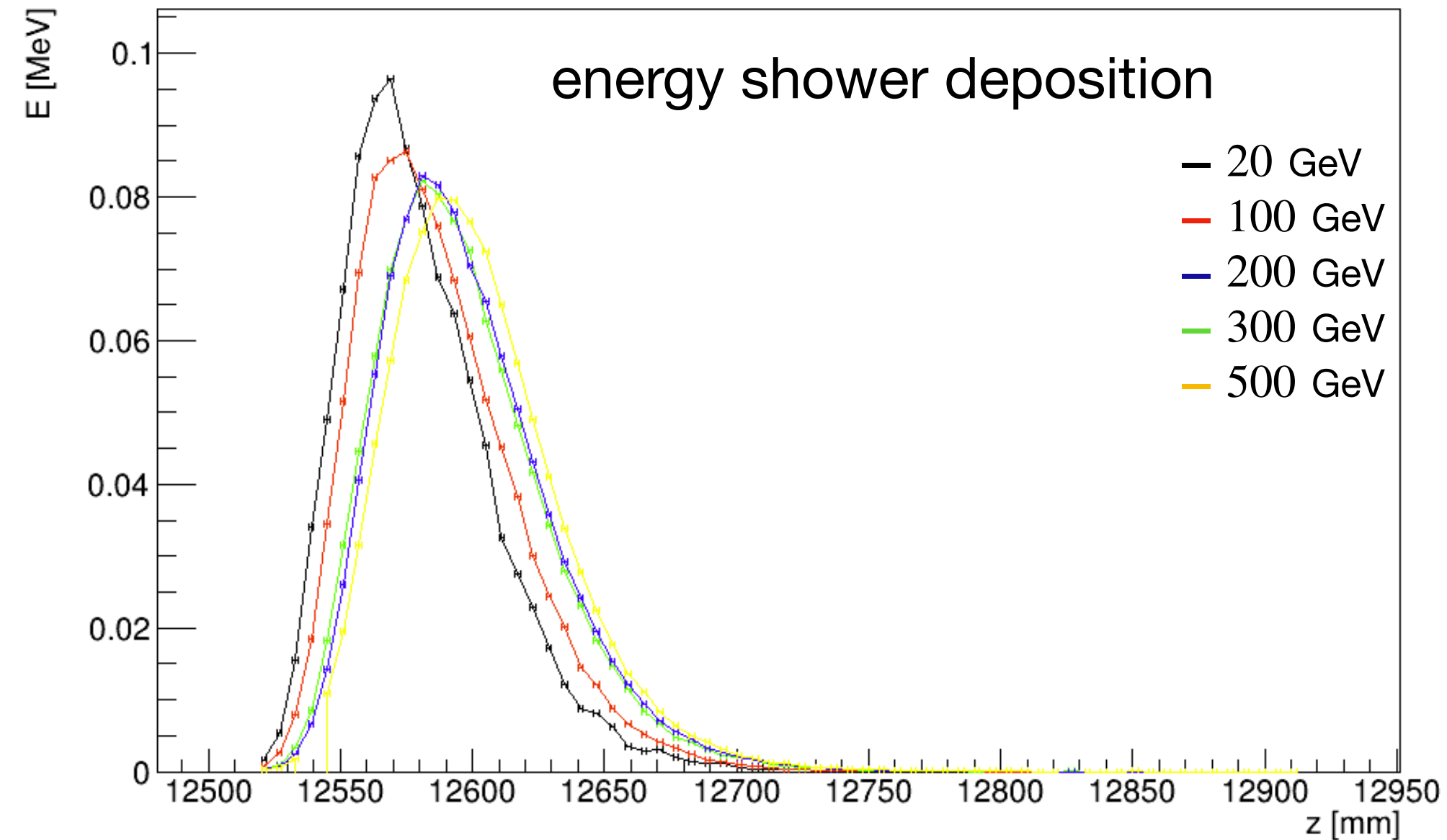$x_{e^{\pm}} - x_{\gamma}$ [mm]
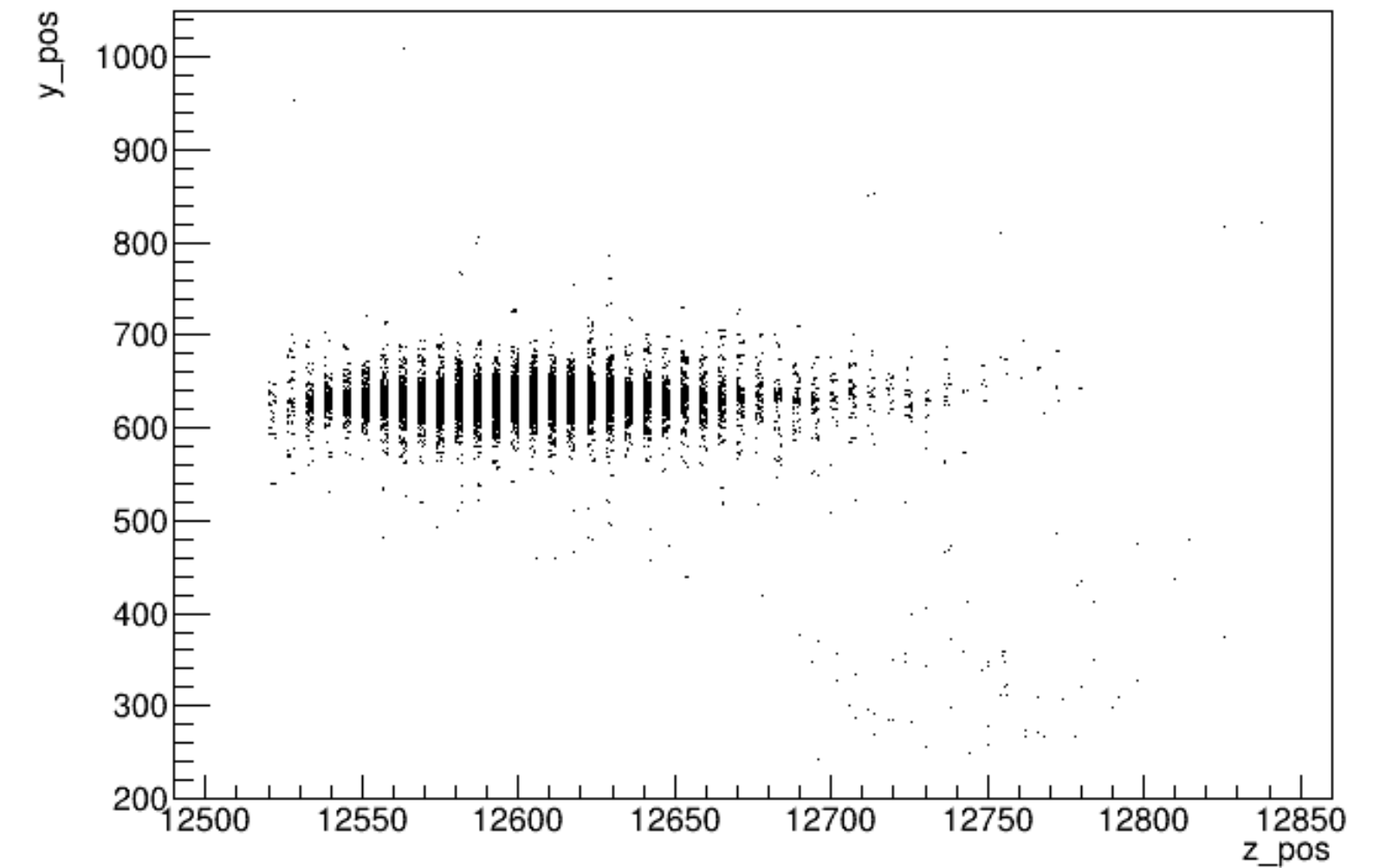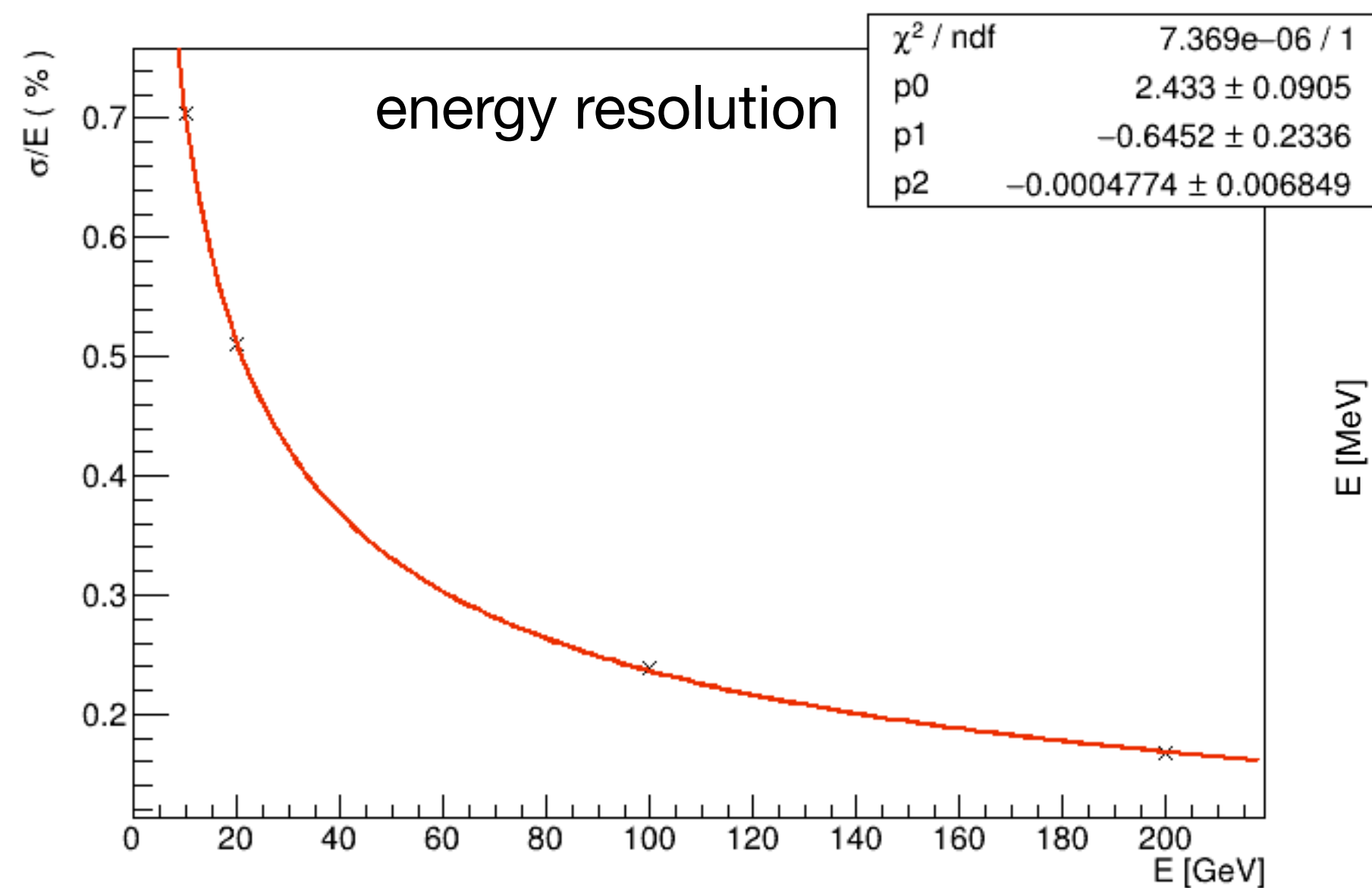
Mass reconstruction ([WIP](#))

- Up to now mass reconstruction has been performed just considering calorimeter clusters

- We are developing a tracks+clusters reconstruction approach

- ADC saturation will impact on mass reconstruction (up to now is $E_T = 20$ GeV)
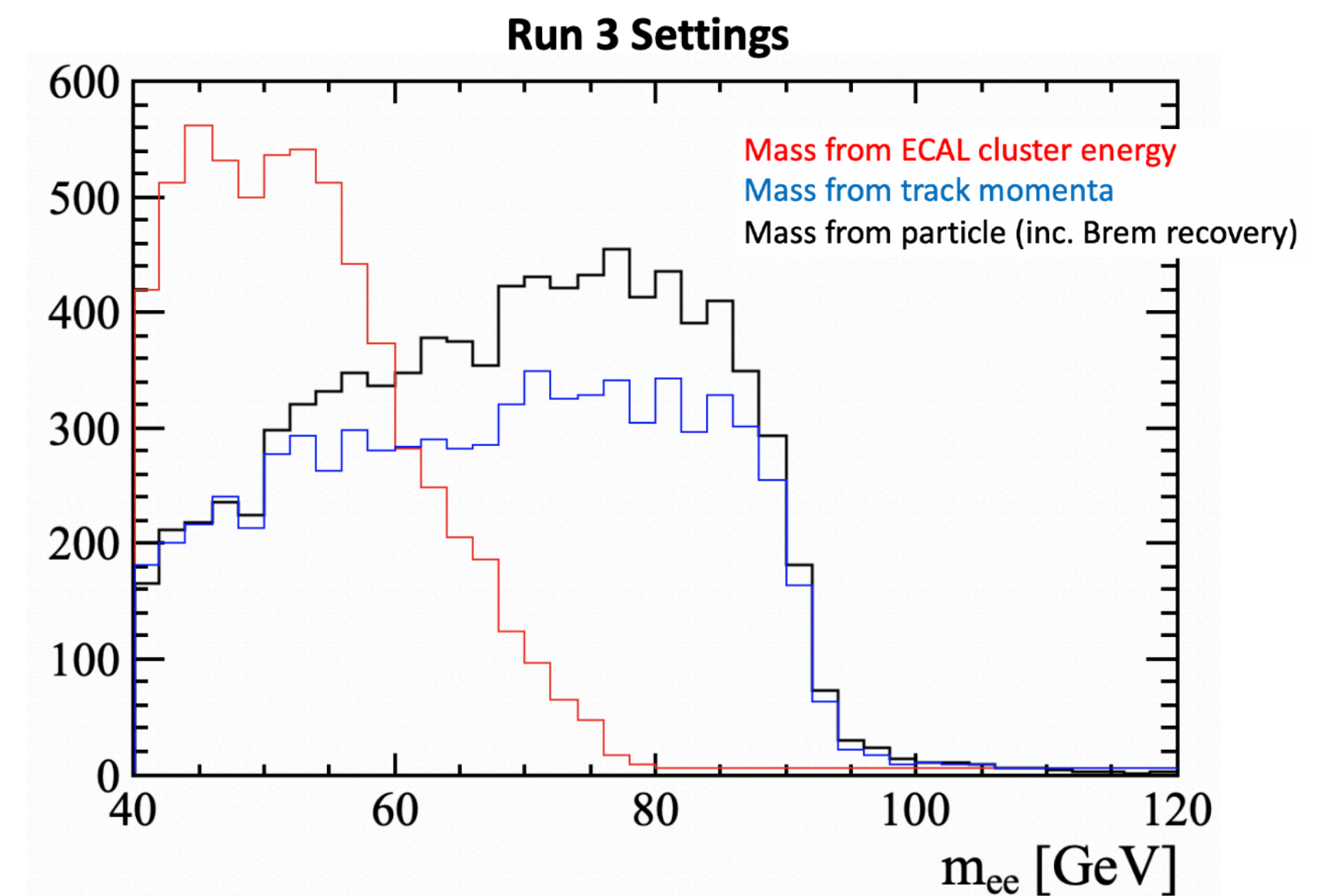
- We are developing some studies to consider z-segmentation for ECAL
- Parametric simulation obtained using GFlash package from Geant 4
- Up to now just some plots showing:

energy resolution

χ^2 / ndf	7.369e−06 / 1
p0	2.433 ± 0.0905
p1	−0.6452 ± 0.2336
p2	−0.0004774 ± 0.006849

energy shower deposition

— 20 GeV
— 100 GeV
— 200 GeV
— 300 GeV
— 500 GeV

Grazie per l'attenzione