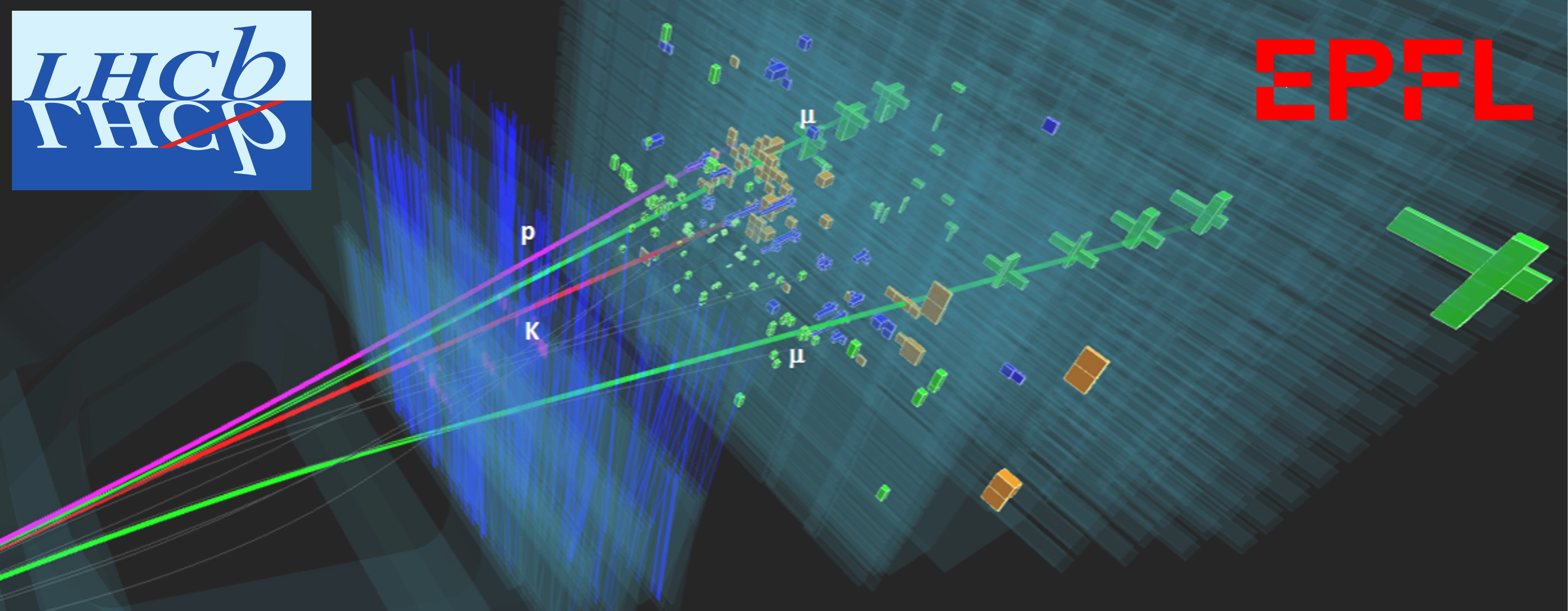


Anomaly detection at the trigger level for LLPs

Luca Hartman, Louis Henry, Lesya Shchutska, Juliette Beaubis

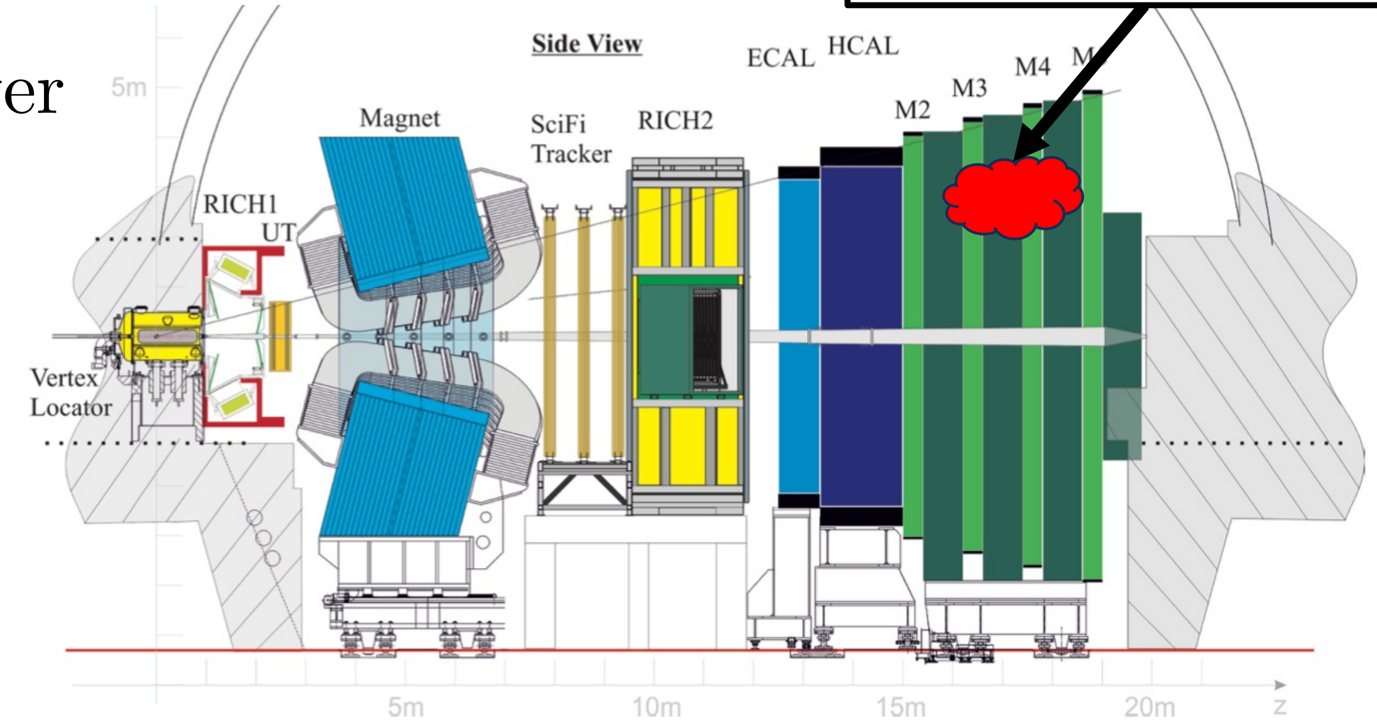
Unique detector signatures

- **Feebly interacting particles** appear in many BSM scenarios
  - This work focuses on:
    - Heavy neutral leptons (HNL)
    - Axion like particles (ALP)
- Long lifetimes lead to **unique challenges and opportunities**
  - Particles could decay beyond the last tracking station
- We suggest to use the **muon detector** as a **sampling calorimeter**
  - Very rare signature in the SM
  - Similar searches done by ATLAS [1] and CMS [2, 3]
  - LHCb could contribute in a short timescale [4]

LHCb muon detector

- Four multiwire proportional chambers (M2-M5)
- Three iron layers of each $4.8\lambda_I$ (80 cm of iron)
- Large decay volume
- But **not** designed for shower detection
  - No energy deposit measurements
  - Rough timing only

Unexpected shower

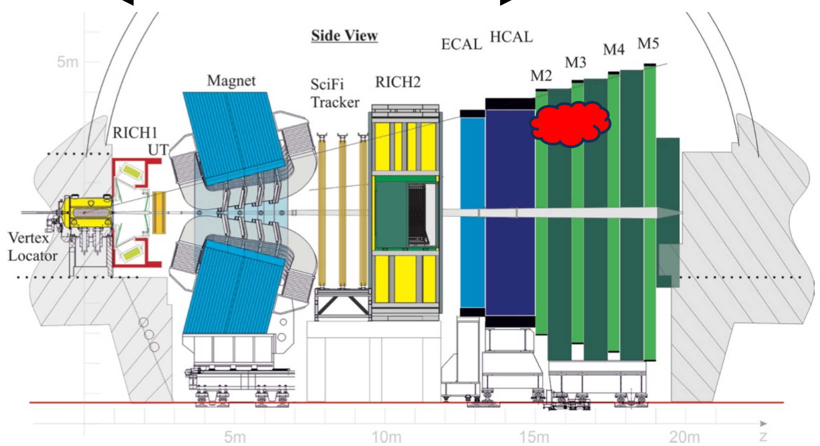


[LHCb reoptimized detector design and performance : Technical Design Report, LHCb Collaboration, 2003](#)

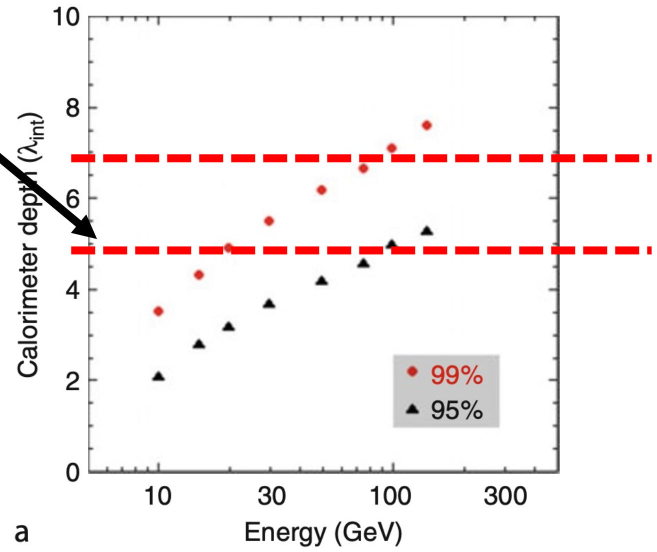
LHCb muon detector

- Four multiwire proportional chambers
- Three iron layers of each $4.8\lambda_{\mathrm{I}}$ (80 cm of iron)
- Large decay volume
- But **not** designed for shower detection
  - No energy deposit measurements
- Very **clean** environment
  - First plane (M2) after $6.7\lambda_{\mathrm{I}}$ of material

$\sim 7\lambda_{\mathrm{I}}$



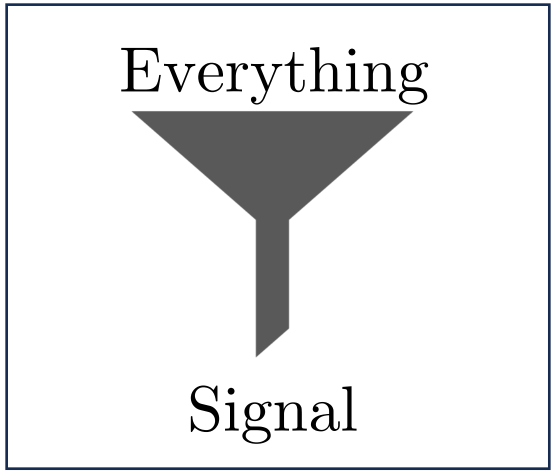
LHCb reoptimized detector design and performance : Technical Design Report, LHCb Collaboration, 2003

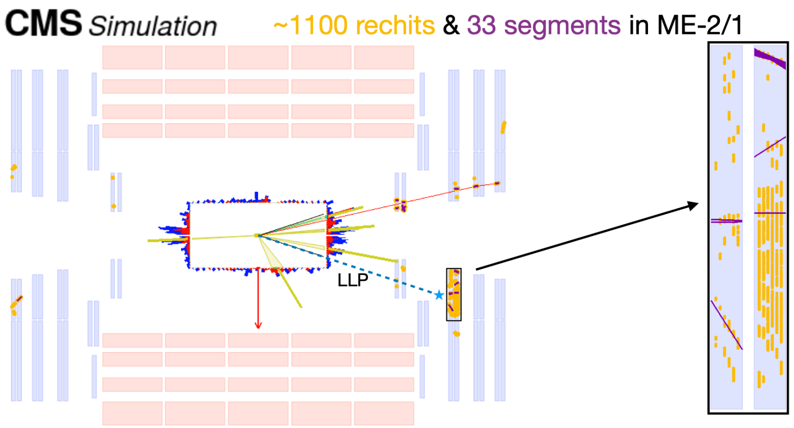


Handbook of Particle Detection and Imaging, R. Wigmans, pp 497–517

How to trigger on such signatures?

- Full-software HLT1 and HLT2
  - HLT1: 40 TB/s to 2 TB/s → GPUs
  - HLT2: 2 TB/s to a 80 GB/s → CPUs
- **No existing trigger line** for such events
  - One HLT2 line since April 2024
- Information at our disposal:
  - Number of hits in (parts of) the muon system
  - Hit coordinates (x, y, z) (and time)
- Similar trigger in CMS [HNLs at CMS, L. Lunerti, LHCP 2024]
  - Based on hit counting only
  - 10^7 background rejection when combined with vetos





[CSC High Multiplicity Trigger in Run 3, CMS Collaboration, 2022]

Anomaly detection elevator pitch

- Need an inclusive trigger for HLT1
- Usual triggers identify the **rare signal events**
  - Based on the **signal properties**
  - Set selection threshold to reduce the rate of "boring" events
- What if we tried to recognise the **boring events**?
- Anomaly detection:

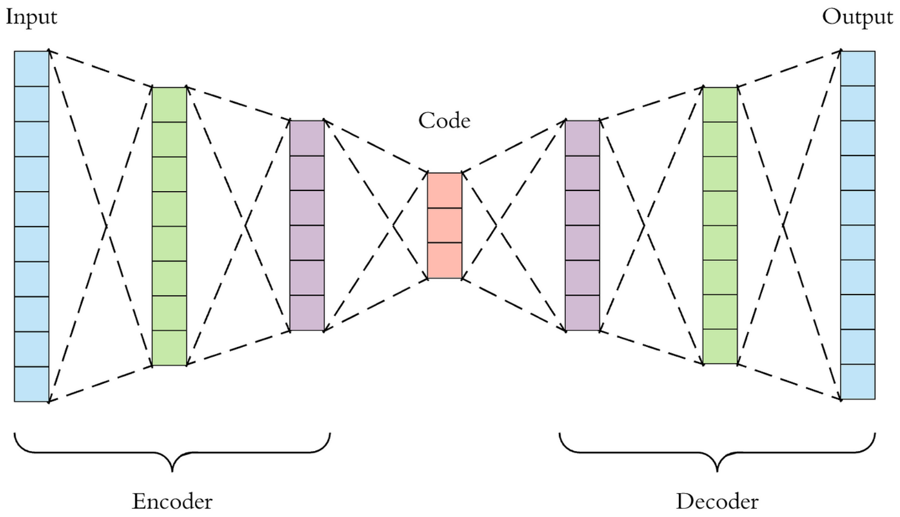
What we don't recognise as boring = interesting events

L. Hartman

How can we do that?

Autoencoder Under Normalization Constraints, S. Yoon *et al*, 2021, arXiv:2105.05735

- Use an normalised **autoencoder** (AE)
  - Encoder neural network → Bottleneck → Decoder neural network
- Punish the model if it has too low error on non-MinBias events
  - i.e., reconstruct well the minimum bias events, **and only them**
  - Measure this by MCMC sampling of fake events
- Measure error between input and reconstruction (MSE)
- Train to on **minimum bias events**
- Events with large error are **anomalous** → **Signal**
- Other models tried and benchmarked
- Model info in backup



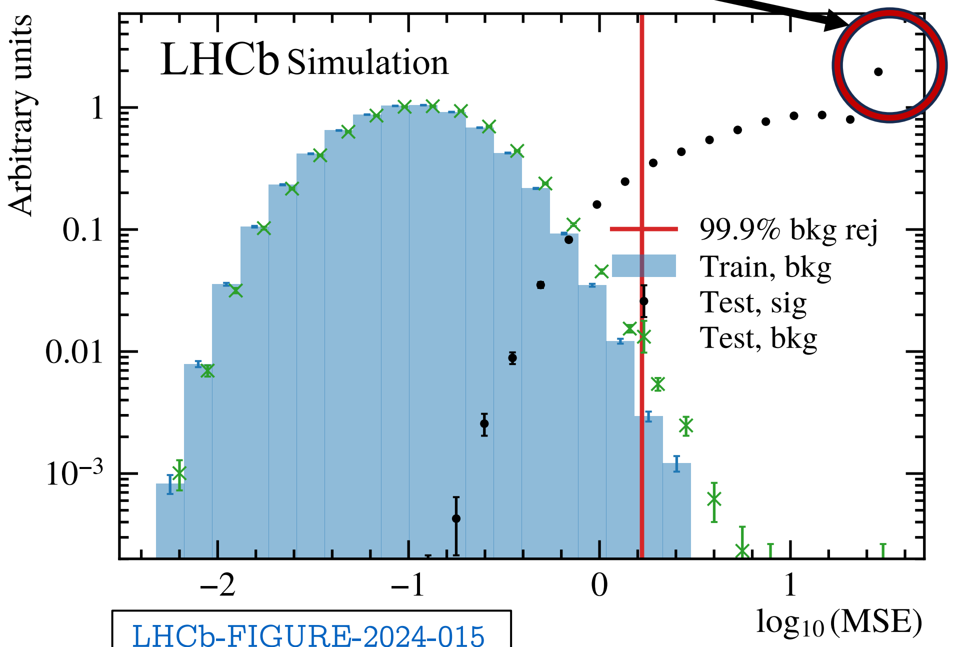
Applied Deep Learning - Part 3: Autoencoders, Arden Dertat

How well does it work?

- Excellent efficiency
  - Outperforms BDT thanks to larger background data set
  - Increased model independence for the HNLs
- Z and B_s samples still effectively rejected
- Could be much faster than BDT
  - Thanks to HLT1 being on GPUs

Sample	ϵ_S@ 99.9% [%]	ϵ_S@ 99.99% [%]
Axion, 10 GeV	93.7 ± 0.1	83.7 ± 0.1
HNL, e, 4 GeV	33.0 ± 0.4	15.9 ± 0.3
HNL, e, 1.6 GeV	26.2 ± 0.4	10.0 ± 0.3
$B_s \to \phi\phi$	0.5 ± 0.1	0.3 ± 0.1
$Z \to \mu\mu$	1.6 ± 0.1	0.3 ± 0.1

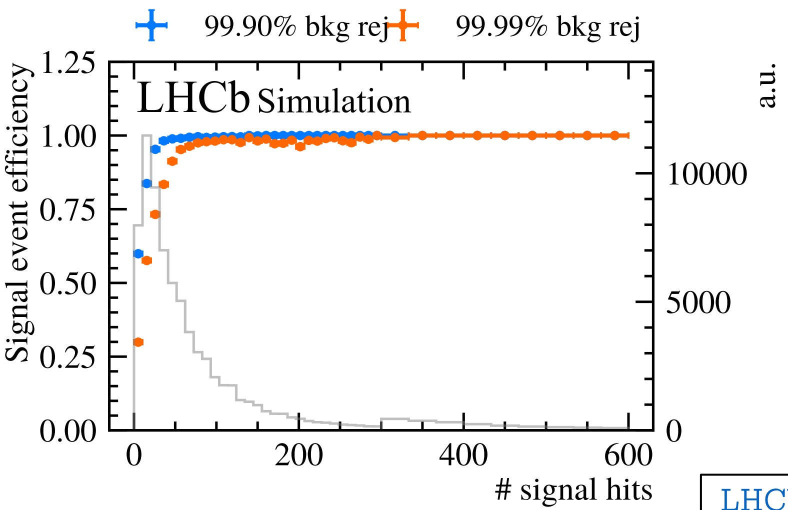
Last signal bin contains overflow

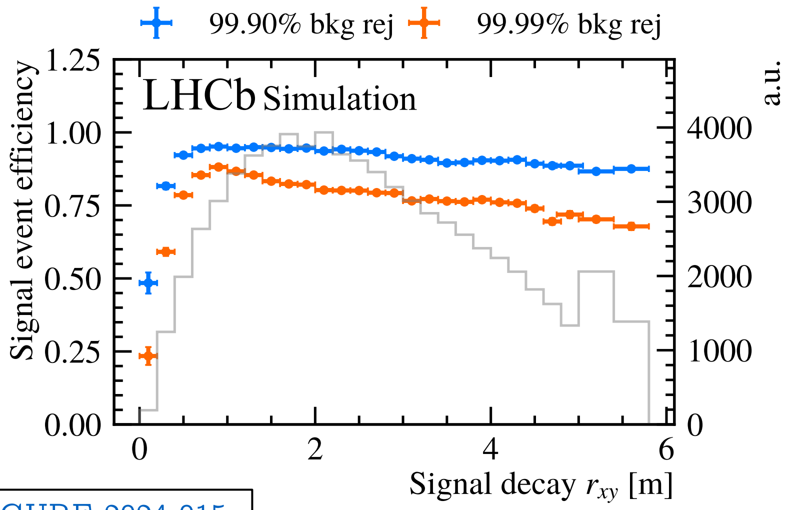


LHCb-FIGURE-2024-015

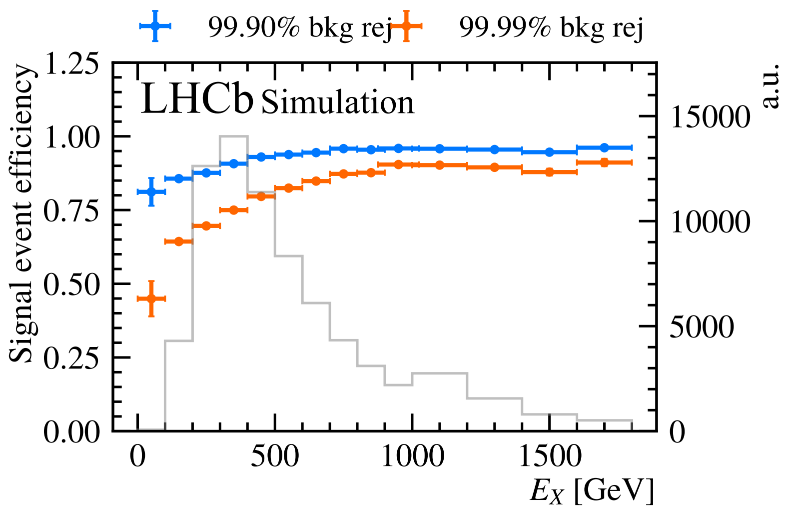
Efficiency curves

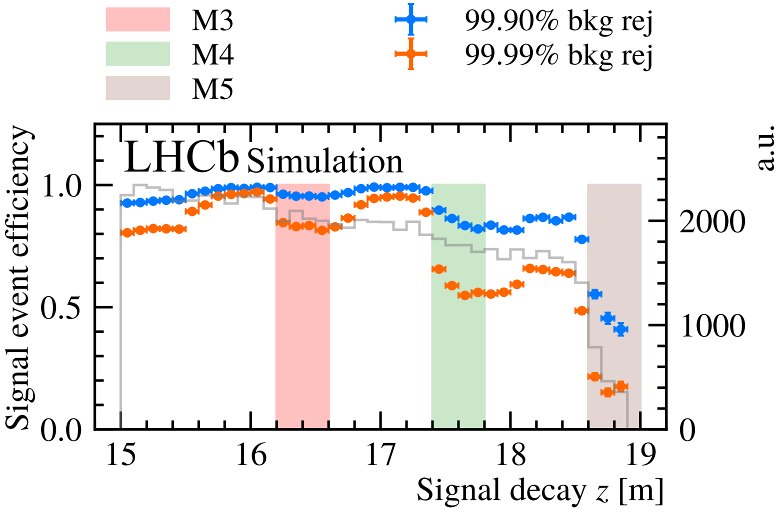
- Can pick up much smaller showers
  - Keeps high efficiency for large showers
- Much flatter efficiency as a function of the distance to the beam pipe
- Better efficiency at low energy





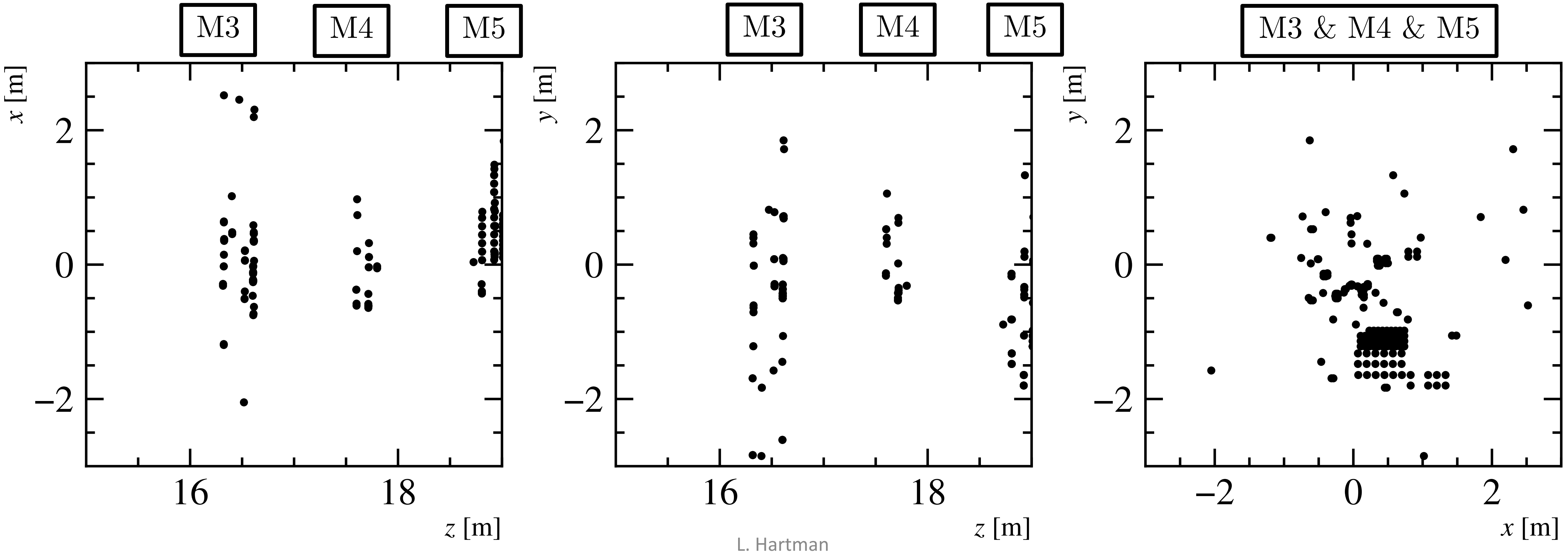
LHCb-FIGURE-2024-015





How to study such events?

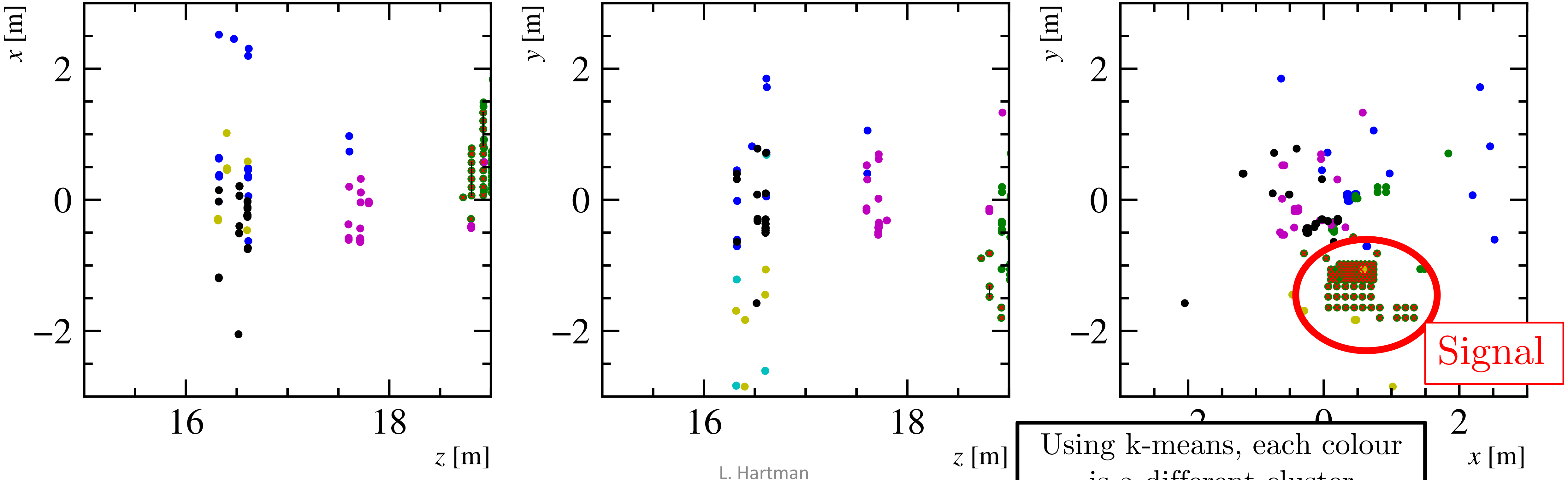
- Can we find the signal shower among the hits? → clustering
  - Active layers clearly visible along z
  - And **one big group** of hits → the signal shower



Can we make clusters?

Algorithms & parameters in backup

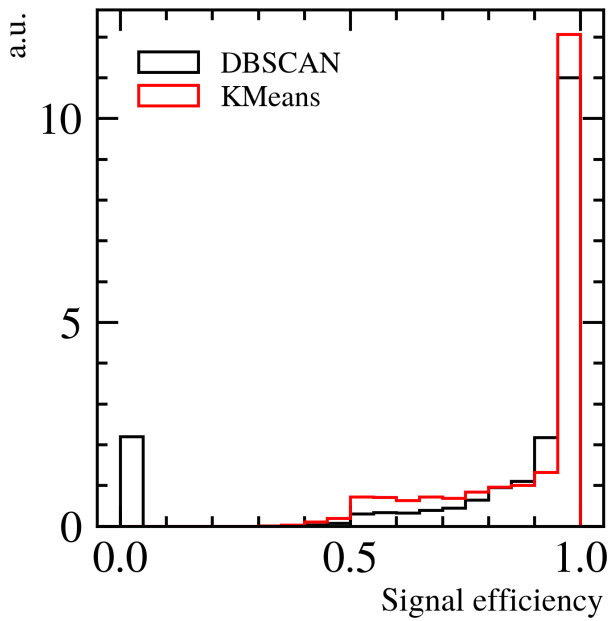
- **Clusters** using the distances between the points
  - Using k-means (6 clusters) or DBSCAN ($\epsilon = 1$, $\mathrm{n_{min}} = 10$) samples)
- The cluster containing the most signal is the **Best Cluster** (BC)

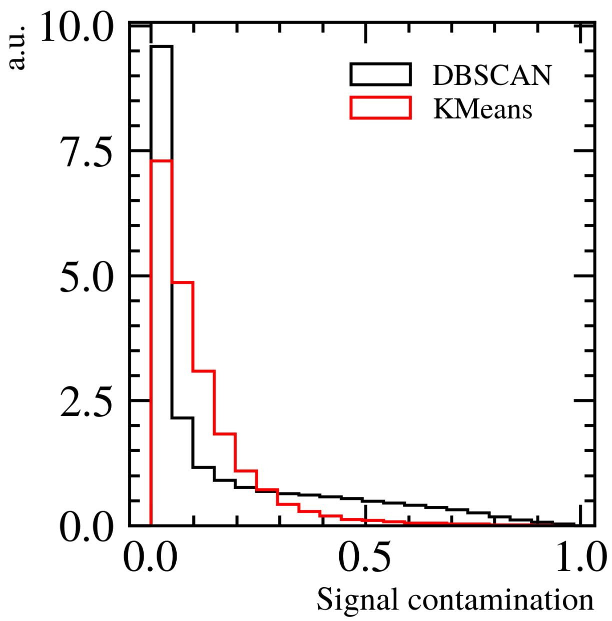


Using k-means, each colour is a different cluster

Is it clean enough?

- We want **all of the signal** to be all within **one cluster**
  - Signal efficiency = (signal hits in BC) / (all signal hits)
- We want the cluster to **only contain signal**
  - Contamination = (non-signal hits in BC) / (all hits in best cluster)
- Most of the background within the BC is **unassociated hits**
  - Caused by electronic noise or low energy particles discarded by the simulation
- Extraction of kinematic information under work
  - See backup for preliminary results





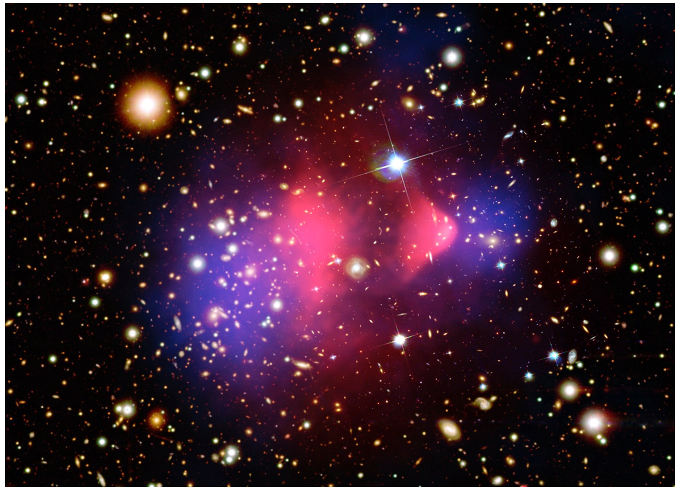
Conclusions

- **Exciting** trigger lines for LLP decays
  - On GPUs: developed, to be included soon
  - On CPUs: taking data since April
- Anomaly detection will be used for the **first time** in a LHCb trigger
  - Model independent
  - High efficiencies
- Shower can be cleanly selected
  - Develop kinematic information extraction

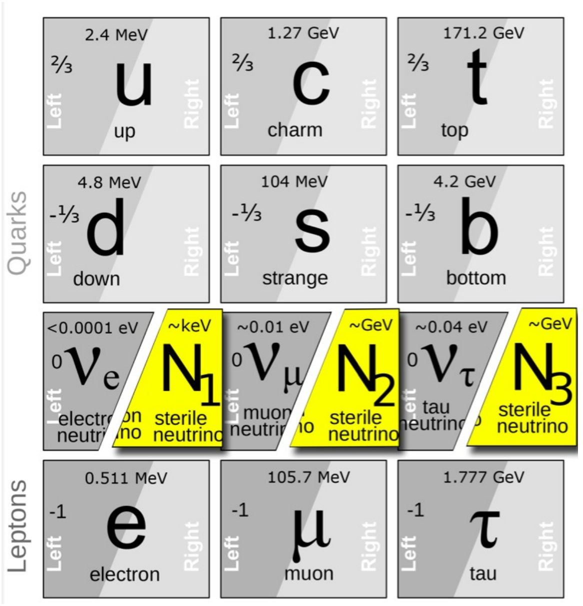
BACKUP

Feebly interacting particles

- SM: unparalleled success, but known to be **incomplete**
- Many suggested solutions
- **Feebly interacting particles** have become the focus of many searches
- **Heavy Neutral Leptons (HNLs)** are sterile right-handed neutrinos
  → Neutrino mass, Dark matter candidate, Matter asymmetry
- **Axion-like particles (ALPs)** come from any anomalous spontaneously broken $U(1)$ symmetry
  - E.g. to explain lack of CP violation in QCD



The Bullet Cluster, ESA, 2007



Heavy neutral leptons - minimal and testable explanation for Beyond Standard Model phenomena, K. Bondarenko, 2021

The samples

- HNLs N
  - $B_u \to N\mu,\ N \to eh$ where h is anything hadronic
  - N masses $1.6,\ 4$ GeV and lifetime 1 ns
- Axions A (main one used for development)
  - $H \to AA,\ A \to \tau\tau,\ \tau \to \pi^{\pm}\pi^{\pm}\ \pi^{\pm}\ \nu$, where H is the SM 125 GeV Higgs boson
  - Mass $m_A = 10$ GeV and lifetime 1 ns
- Signal requirements:
  - Decay in muon detector ($15\ \mathrm{m} < z < 19\ \mathrm{m}$)
  - At least 5 hits from the signal shower
- Minimum bias events
- Two SM physics processes to benchmark
  - $Z \to \mu\mu$
  - $B_s \to \phi\phi$

How to measure the performance?

- *All signal events
  - Contain ≥ 5 hits in layers M2-M3-M4-M5
  - Decay position $15 \text{ m} < z < 19 \text{ m}$
- Minimum bias rejection:

$$1 - \frac{\text{events passing selection}}{\text{all events in acceptance}}$$

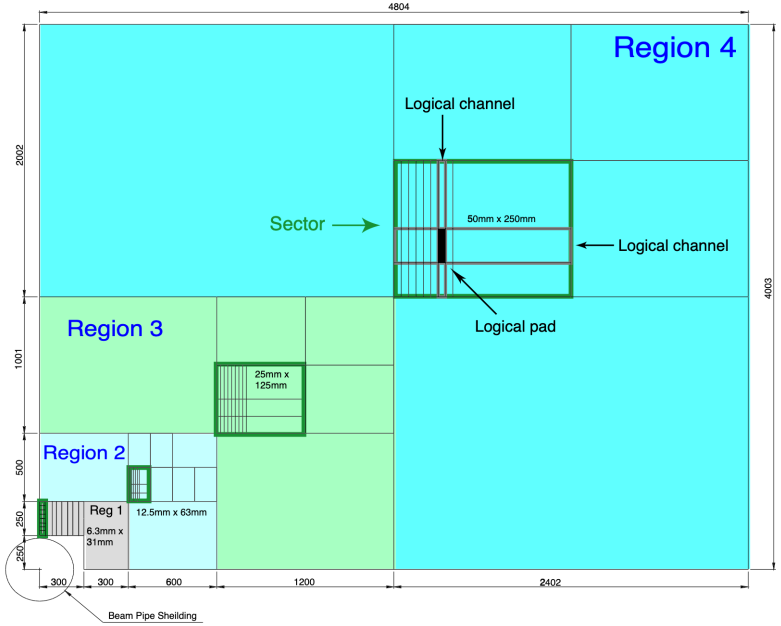
- Signal efficiency:

$$\frac{\text{signal events passing selection}}{\text{all signal events with decays in muon system}^*}$$

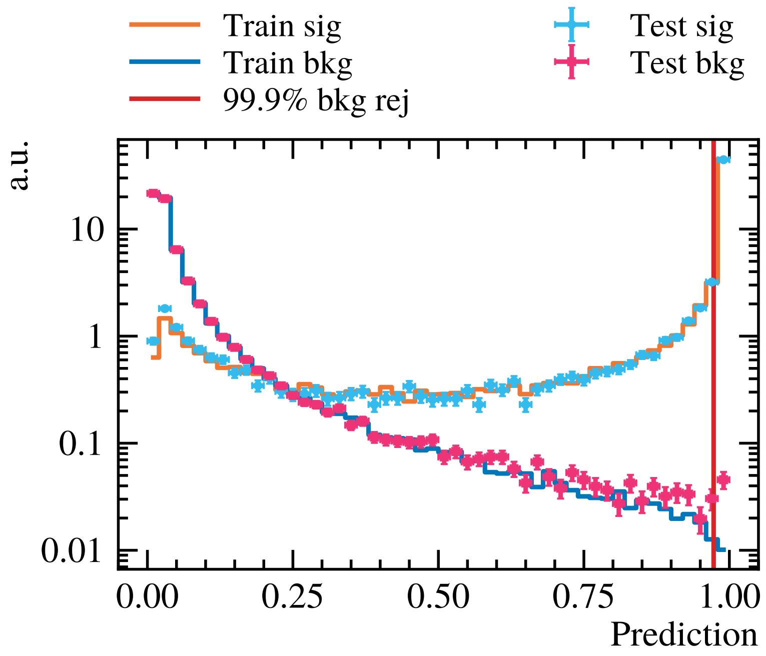
Hit counting

- Simple cuts are not enough
- Use a BDT
  - Based on xGBoost
  - More info in backup
- 12 Features are the number of hits
  - Per station (M3, M4, M5) and per region
- Currently implemented in HLT2

Sample	ϵ_S@ 99.9% [%]	ϵ_S @ 99.99% [%]
Axion, 10 GeV	63.1 ± 0.4	44.5 ± 0.4
HNL, 4 GeV	16.7 ± 0.3	6.4 ± 0.2
HNL, 1.6 GeV	13.1 ± 0.3	4.8 ± 0.2
$B_s \to \phi\phi$	1.8 ± 0.2	0.3 ± 0.1
$Z \to \mu\mu$	3.5 ± 0.1	0.6 ± 0.1

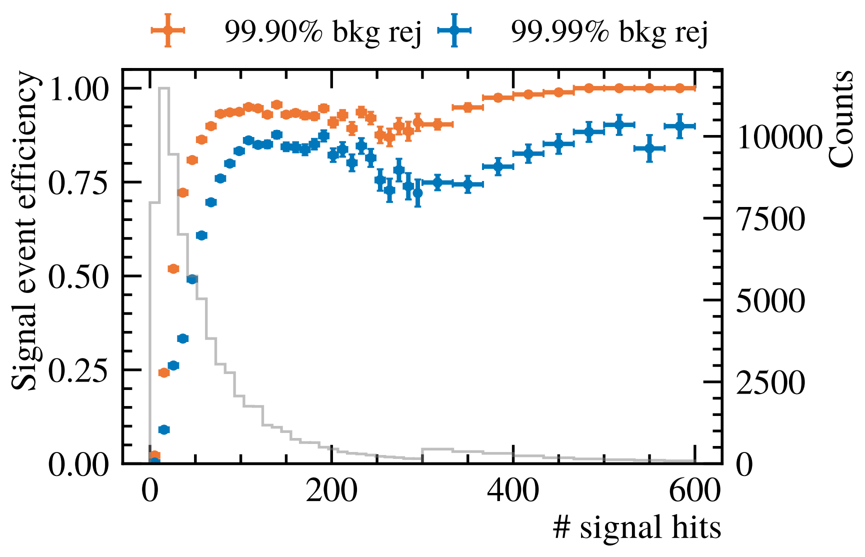


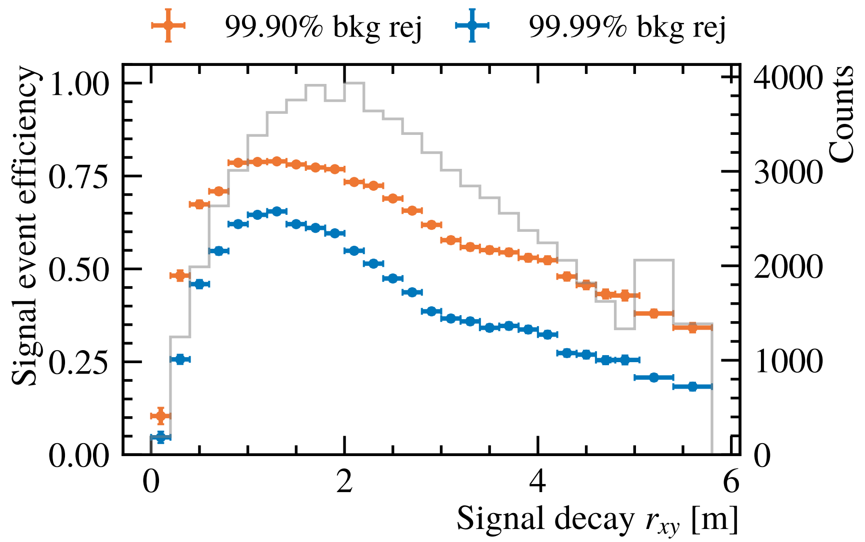
LHCb muon system : Technical Design Report, 2001

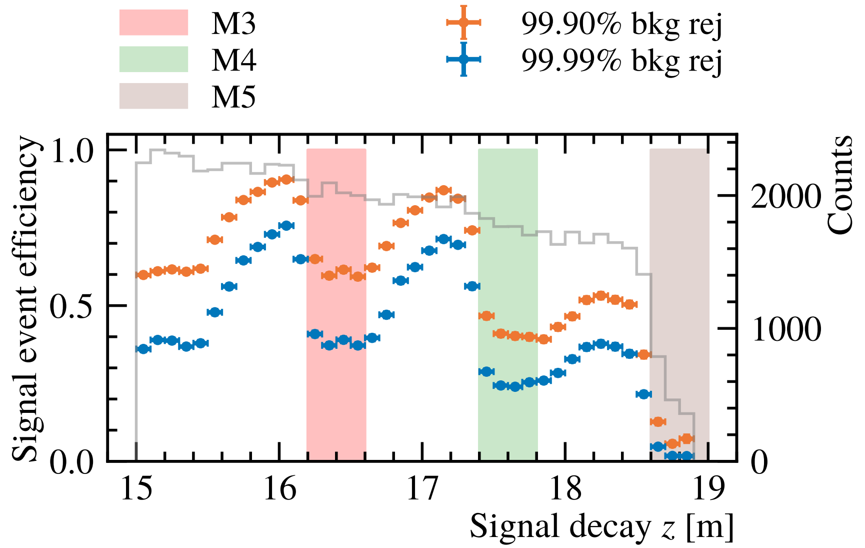


Efficiency curves

- Turn on effect with larger showers
- Better efficiencies when decays happen right before the active layers
- Large loss of efficiency at further away from the beam pipe

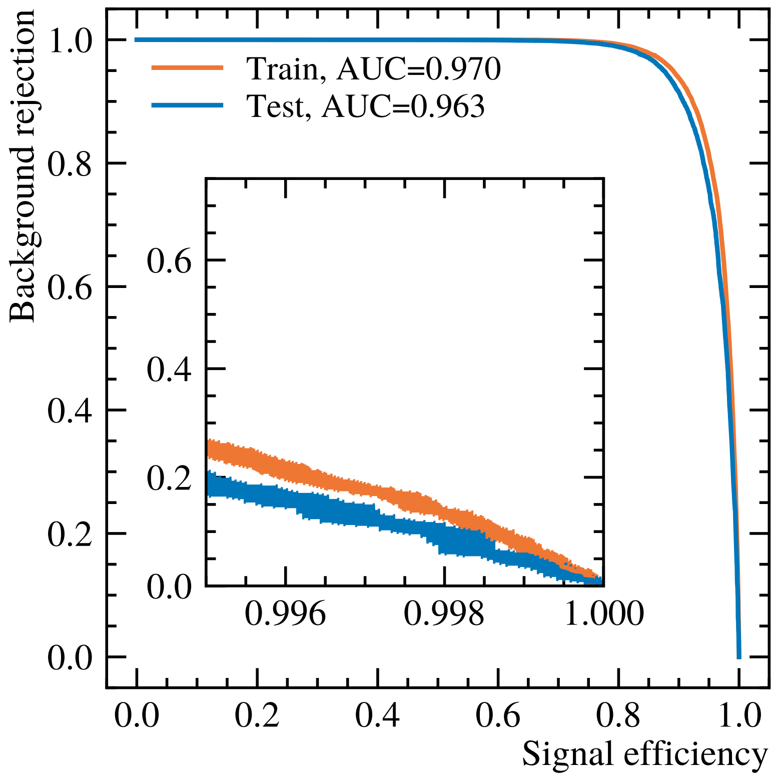


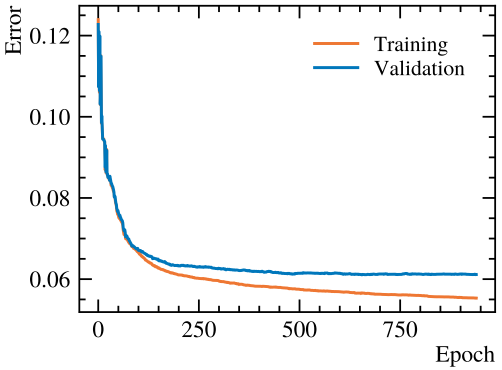


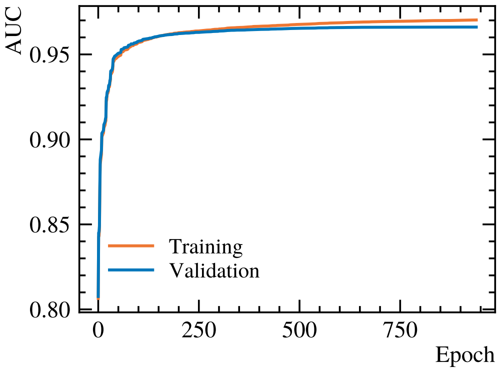


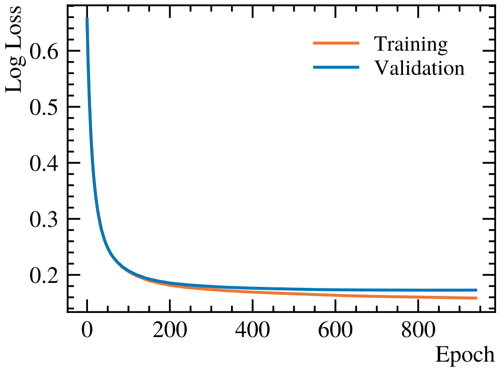
Hit counting BDT

- Based on XGBoost
- Hyperparameters are optimised using optuna
  - Optimised for maximal AUC under the ROC curve
- Hyperparameters:
  - Number of trees: 940
  - Maximum depth: 4
  - Learning rate: 0.06499









Hit clustering algorithm: k-means

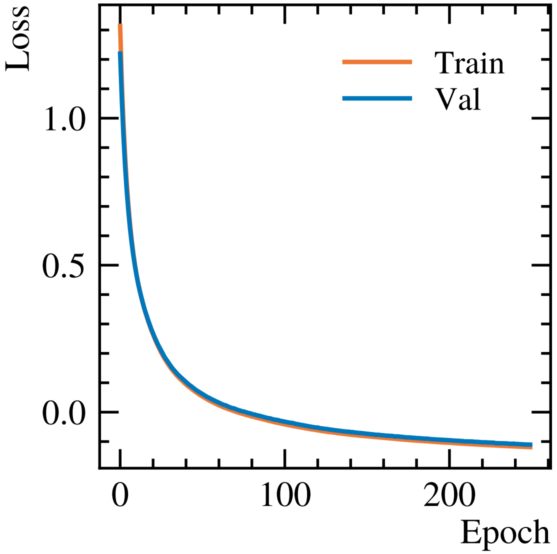
- Based on the k-means algorithm
  - Minimises the distances within each cluster
- Using the scikit-learn implementation
- Depends on only two parameters:
  - Number of clusters: 6
  - Number of initialisations: 10
- The parameters were optimised to
  - Reduce the impact of the random initialisation
  - Increase the signal efficiency at 99.9% minimum bias rejection

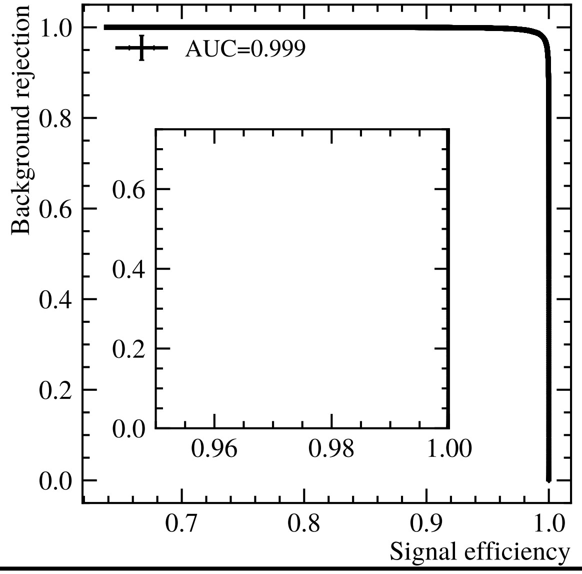
Hit clustering algorithm: DBSCAN

- Based on the DBSCAN algorithm
  - Connects hits if within distance ϵ
  - Consider it to be a cluster if at least $n_{\min}$ hits are sufficiently close
- Advantages:
  - Computationally faster
  - More stable (independent on initialisation)
  - Rejects noise
- Using the dbscan-python pip package
- Depends on only two parameters (more physical than k-means)
  - Distance for a cluster: $\epsilon = 1$ m
  - Number of hits for a cluster: $n_{\min} = 10$
- The parameters were optimised to maximise signal significance
  - Significance = efficiency / contamination

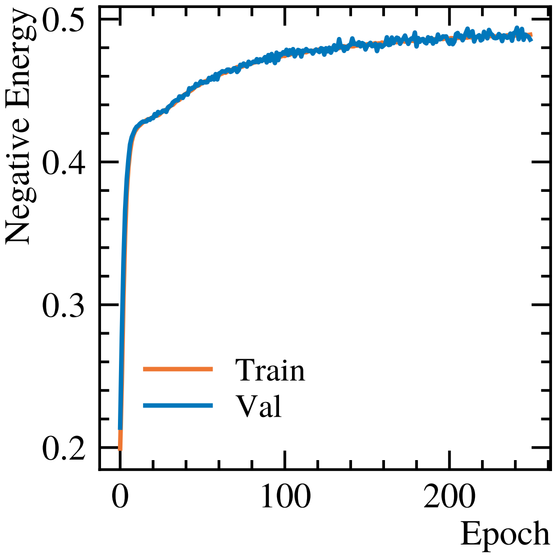
Autoencoder model

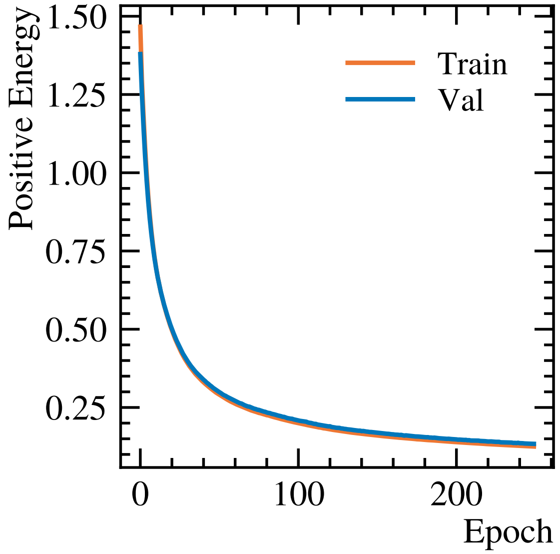
- Implemented using [pyTorch]
- Symmetric encoder and decoder
  - Input: 72 features
  - 2 convolutional layers
  - 3 Dense layers
  - Bottleneck: 40 variables
  - Total: 4.8×10^{6} paramètres
- Hyper-parameters:
  - Batch size $=$ 256
  - Learning rate $= 10^{-6}$
  - Maximum 250 epochs





Negative energy $=$ normalisation constant
Positive energy $=$ reconstruction error





Particle reconstruction: proof of concept

- Reconstruct the position (r_{xy}, z) of the decay in a using the centre of the best cluster
- For the particle momentum, we train a DNN regressor
- r_{xy} is good, z is biased, p needs further work