Software and Computing activities at UB

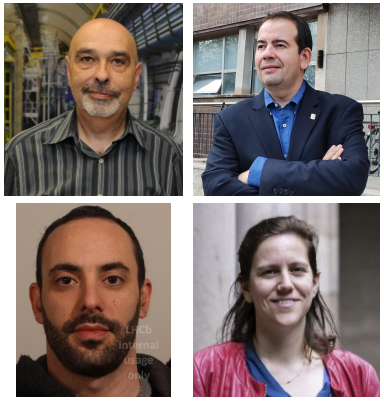
Carla Marin
Experimental High Energy Physics group

UNIVERSITAT DE BARCELONA

Experimental High Energy Physics group @UB

Seniors:

- Dr. Lluís Garrido
- Dr. Eugeni Graugés
- Dr. Ricardo Vázquez
- Dr. Carla Marin



PhD students:

- Dr. Aniol Lobo (→ LAPP, France)
- Albert Lopez
- Paloma Laguarta
- Pol Vidrier
- Alejandro Rodriguez
- Ernest Olivart

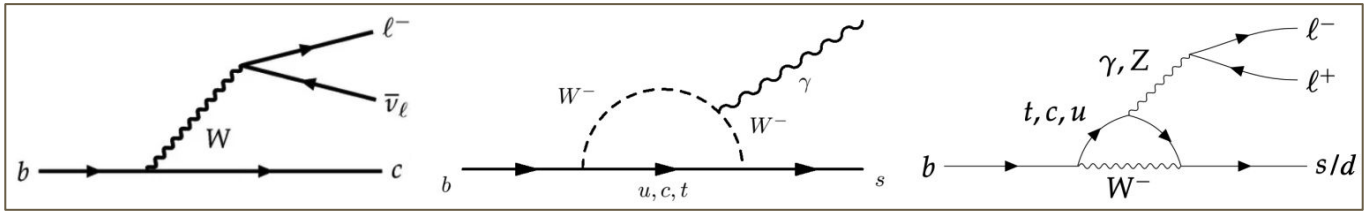
Postdoctoral researchers:

- Dr. Alessandra Gioventu (→ USC)
- Dr. Lukas Calefice
- Dr. Felipe Souza

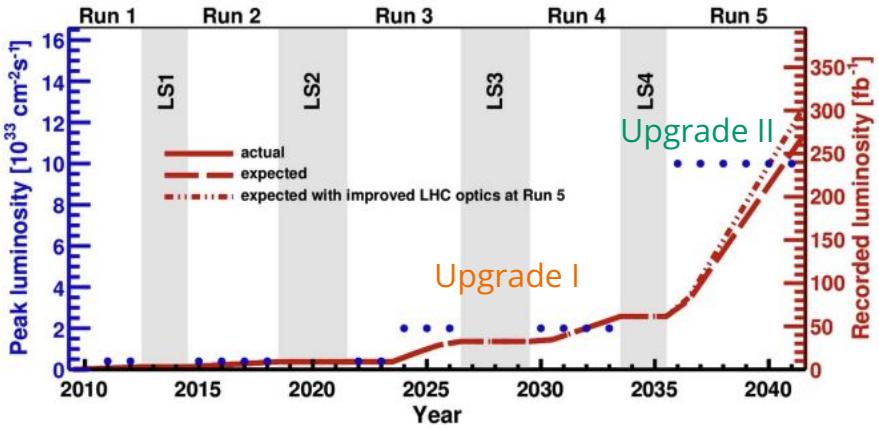
Experimental High Energy Physics group @UB

Long-standing trajectory, member of LHCb since 1998

- strong expertise in calorimetry:
  - design, construction and operation of SPD detector in Run 1 & 2
  - design of calorimeter ASIC (ICECAL), operation and monitoring for LHCb Upgrade (Run 3)
  - design of calorimeter ASIC for LHCb Upgrade II (Run 5)
- strong expertise in Real Time Analysis:
  - γ, $e^{\pm}$ and $\mu^{\pm}$ reconstruction, identification (PID) and calibration in Run 2 & 3
  - development of ML-based trigger selections for Run 2 & 3
- strong involvement in rare and semileptonic b-hadron decays



LHCb upgrades



Upgrade I: $L_{inst} = 2 \times 10^{33}\ cm^{-2}s^{-1}$ (x5)

- Higher granularity tracking detectors
  - ECAL: same hardware → much higher occupancy than in Run 2
- Triggerless readout + fully-software based trigger system
  - HLT1: Real-Time reconstruction at 30 MHz
  - HLT2: offline-quality reconstruction, after real-time detector & calibration

Upgrade II: $L_{inst} = 10 \times 10^{33}\ cm^{-2}s^{-1}$ (x5 again!)

- Higher granularity detectors, (most) including timing
  - ECAL partly upgraded in LS3
- Fully-software based trigger exploiting 4D reconstruction (including timing)

Software & computing activities @UB

Run 1 & 2: muon misidentification (misID) calibration

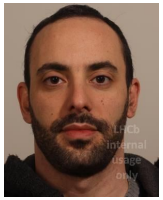
Run 3:

- calorimeter and trigger operation, and data quality monitoring
- photon and electron reconstruction
- electron particle identification (PID) and its calibration
- Machine Learning (ML)-based HLT2 selections:
  - inclusive radiative trigger
  - autoencoder for $\Lambda_b \rightarrow p\pi^-\mu^+\mu^-$ decays - see Paloma's talk

Run 4 & 5:

- LHCb software and computing strategy
- calorimeter reconstruction using GNN's - see Felipe's talk

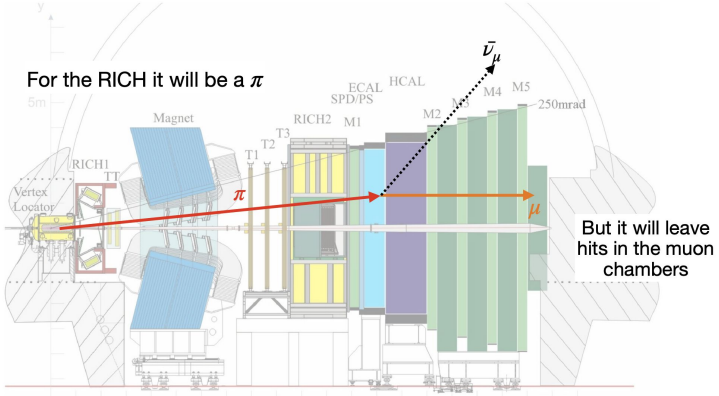
See Alejandro's talk at COMCHA workshop

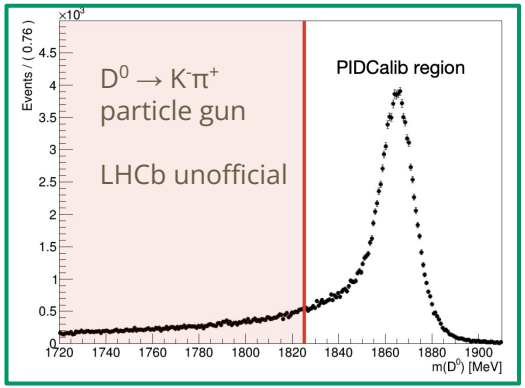



Muon misidentification calibration

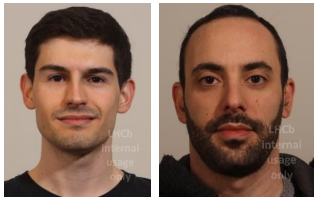
π → μν decays in flight: significant background for μ identification

1. tail of calibration sample out of range → correction from simulation





See Alejandro's talk at COMCHA workshop

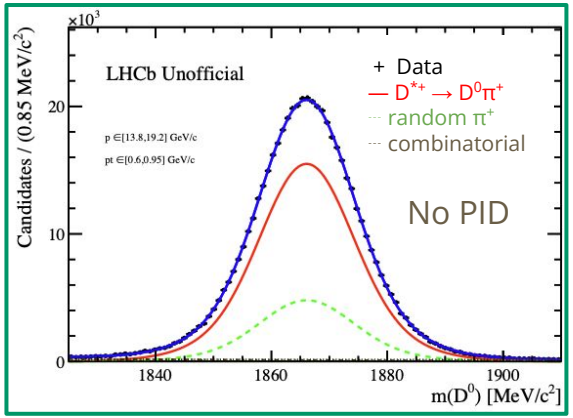


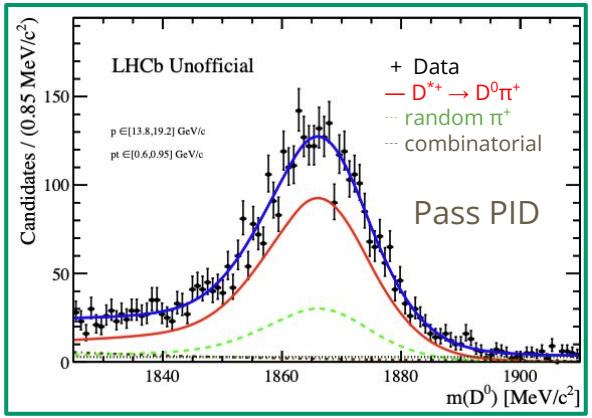
Muon misidentification calibration

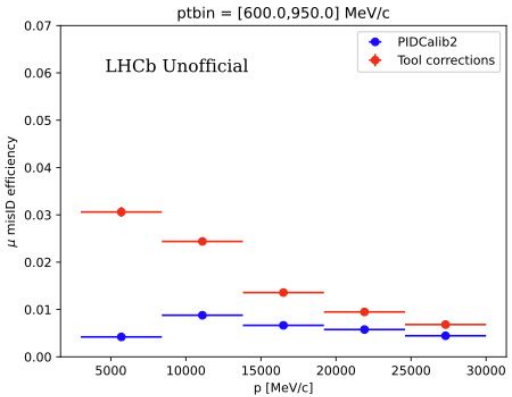
π → μν decays in flight: significant background for μ identification

1. tail of calibration sample out of range → correction from simulation
2. distortion of mass shape → mass model including tail

calibration sample: $D^{*+} \to D^0 (\to K^- \pi^+) \pi^+$







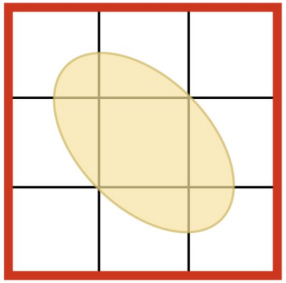
See Alessandra's talk at COMCHA workshop

γ reconstruction in Run 3

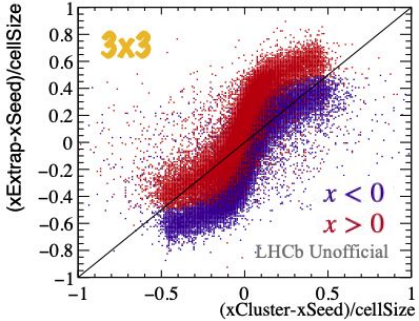
ECAL cluster reco: energy & position corrections critical for resolution

1. first estimate: weighted-average barycenter

$$E = \sum_i E_i \qquad x_b = \frac{1}{E}\sum_i x_i E_i \qquad y_b = \frac{1}{E}\sum_i y_i E_i$$



first estimate



See Alessandra's talk at COMCHA workshop

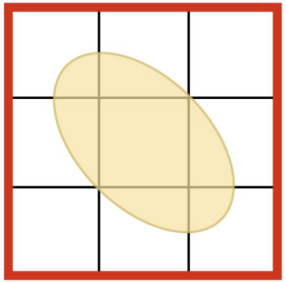
γ reconstruction in Run 3

ECAL cluster reco: energy & position corrections critical for resolution

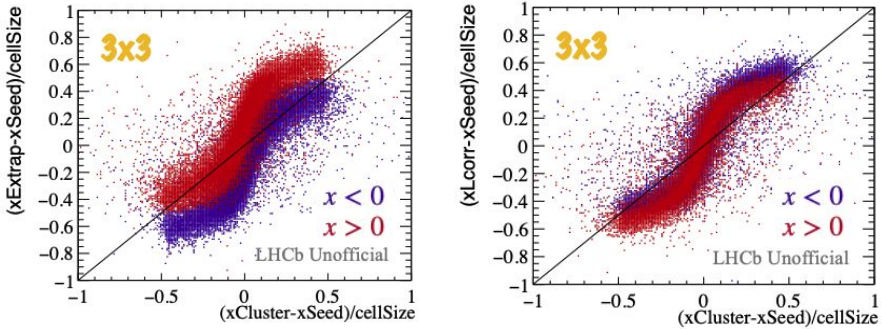
1. first estimate: weighted-average barycenter
2. L correction: redefine (x,y) at z of shower maximum

$$E = \sum_i E_i \qquad x_b = \frac{1}{E}\sum_i x_i E_i \qquad y_b = \frac{1}{E}\sum_i y_i E_i$$

$$z^L = \gamma_0 \ln(E) + \delta_0$$



first estimate → after L correction



See Alessandra's talk at COMCHA workshop

γ reconstruction in Run 3

ECAL cluster reco: energy & position corrections critical for resolution

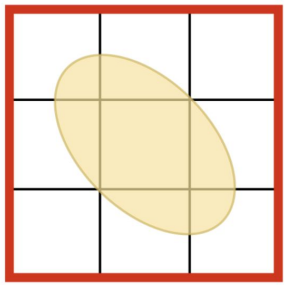
1. first estimate: weighted-average barycenter
2. L correction: redefine (x,y) at z of shower maximum
3. S correction: correct for non-linear energy distribution among cluster cells
   a. angular and residual corrections

$$E = \sum_i E_i \qquad x_b = \frac{1}{E}\sum_i x_i E_i \qquad y_b = \frac{1}{E}\sum_i y_i E_i$$

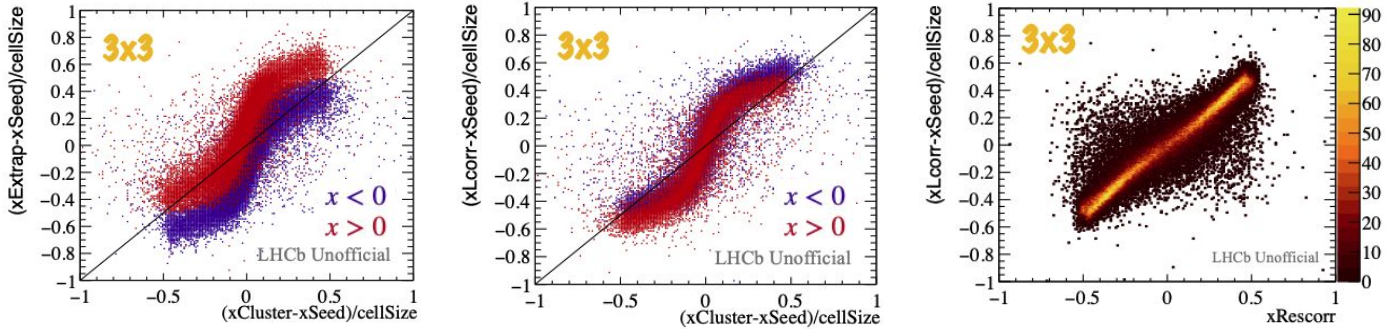
$$z^L = \gamma_0 \ln(E) + \delta_0$$

$$(x_i)^S_{\text{reco}} = b \sinh^{-1}\left(\frac{(x_i)_{\text{Cluster}} - (x_i)_{\text{seed}}}{\Delta}\cosh\frac{\Delta}{b}\right)$$

WIP: test on full MC



first estimate → after L correction → after all corrections

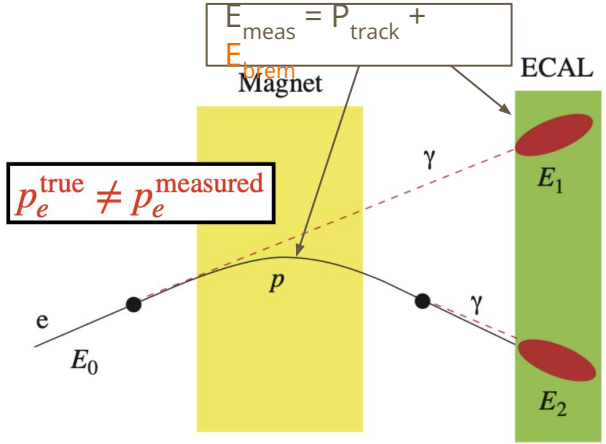


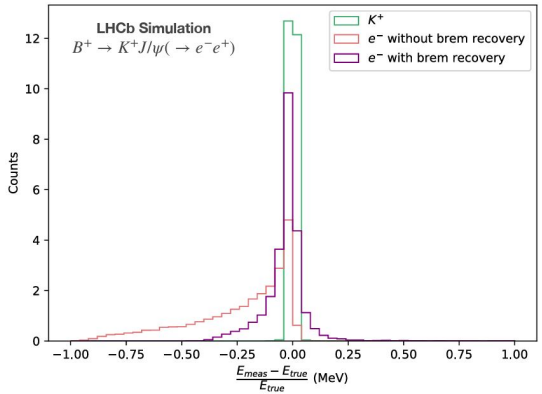
See Paloma's talk at COMCHA workshop & Lukas' at last CPAN

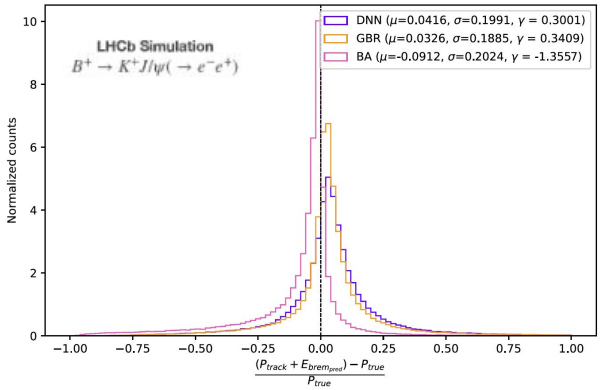
$e^{\pm}$ reconstruction in Run 3

Significant energy (E) loss due to Bremsstrahlung → degraded momentum (p) resolution → Idea: E_{brem} recovery with ML inference (DNN and GBR)

- calorimeter + tracking information to infer total brem energy → $E_{brem,pred}$
- new approach (WIP): correct existing brem algorithm instead → $E_{brem} = E_{BremAdder} * \alpha_{corr,pred}$
- challenge: impact on low-mass background shape







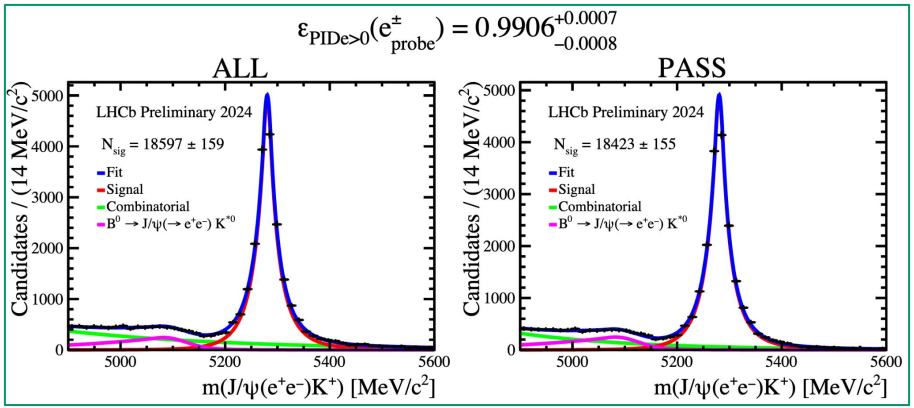
Electron PID calibration in Run 3

See Pol's talk at COMCHA workshop

Accurate calibration of $e^{\pm}$ ID and misID critical for analyses. Method:

- $B^+ \to J/\psi\,(\to e^+e^-)\,K^+$ with tag-and-probe selection
  - π^+_{probe} from $D^{*+} \to (D^0 \to K^-\pi^+_{\text{probe}})\pi^+$ to estimate $\pi^+ \to e^{\pm}$ misID
- apply PID to $e_{\text{probe}} \to$ obtain selection efficiency from data, using mass fits
- fit-and-count in kinematic bins to account for resolution variations

LHCb-FIGURE-2024-038



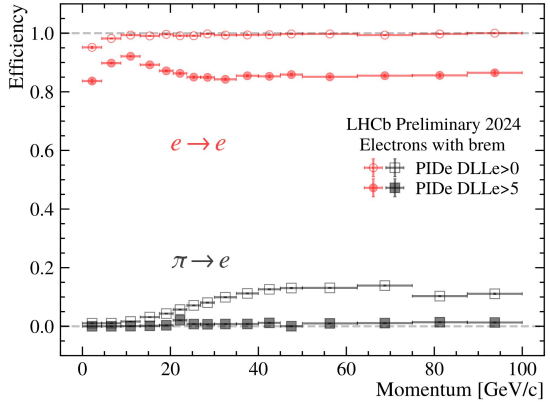
Electron PID calibration in Run 3

See Pol's talk at COMCHA workshop

Results:

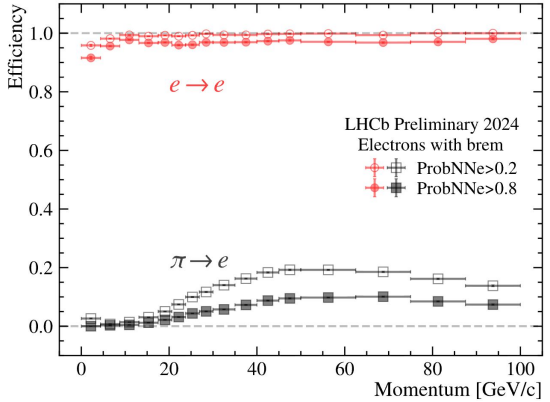
- Performance comparable to Run 2
- Efficiency and misID tables available for all Run 3 analyses (through PIDCalib2)

Log-likelihood method



LHCb-FIGURE-2024-038

NN method



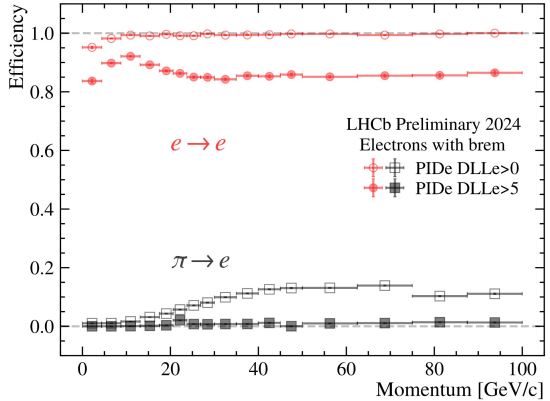
Electron PID calibration in Run 3

See Pol's talk at COMCHA workshop

Results:

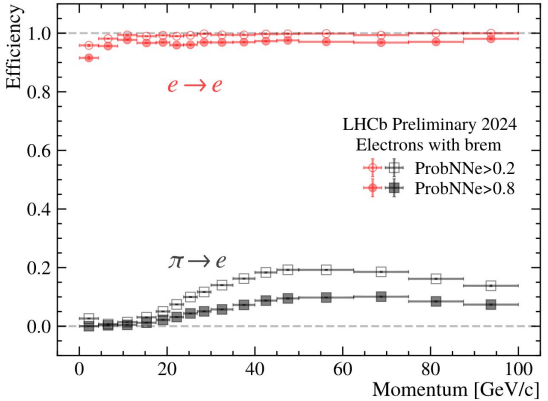
- Performance comparable to Run 2
- Efficiency and misID tables available for all Run 3 analyses (through PIDCalib2)
- Room for improvement in NN-based classifier → see Ernest's poster

Log-likelihood method

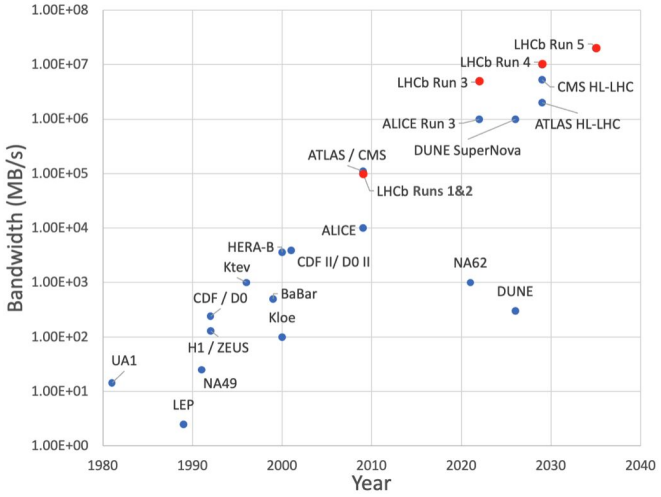


LHCb-FIGURE-2024-038

NN method



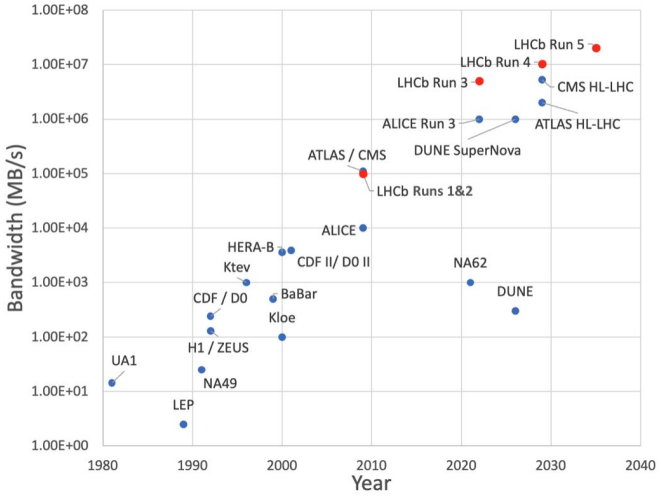
LHCb Upgrade II



Biggest bandwidth challenge in HEP: > 10^7 MB/s

- heterogeneous system:
  - FPGA-based DAQ cards
  - high-speed dedicated network cards
  - GPUs for partial (HLT1) and full-quality reco (HLT2), both exploiting timing info
- large simulation needs → flash simulation and parallelisation (CPUs and GPUs)
- huge samples for offline analysis → optimised storage formats and access models
- see LHCb-PUB-2025-004 submitted to ESPPU

LHCb Upgrade II



Biggest bandwidth challenge in HEP: > 10^7 MB/s

- heterogeneous system:
  - FPGA-based DAQ cards
  - high-speed dedicated network cards
  - GPUs for partial (HLT1) and full-quality reco (HLT2), both exploiting timing info
- large simulation needs → flash simulation and parallelisation (CPUs and GPUs)
- huge samples for offline analysis → optimised storage formats and access models
- see LHCb-PUB-2025-004 submitted to ESPPU

ECAL reconstruction also very challenging due to high occupancy

- baseline reco: Run 3 approach (Graph Clustering + cluster corrections) + timing
- alternative: Graph-Neural-Network (GNN) inference → see Felipe's + Uzzi's talks

Summary

High Energy Physics group @UB strongly involved in LHCb software & computing, leveraging expertise in calorimeter.

- Past years focused on Run 3 developments: trigger, reco, calibration
- Shifting to Run 4 (partial ECAL upgrade) and Run 5 (full upgrade, x5 L_{inst})

ML approaches proving crucial in most tasks.

A lot of work is WIP, happy to get input and discuss!

BACK-UP