



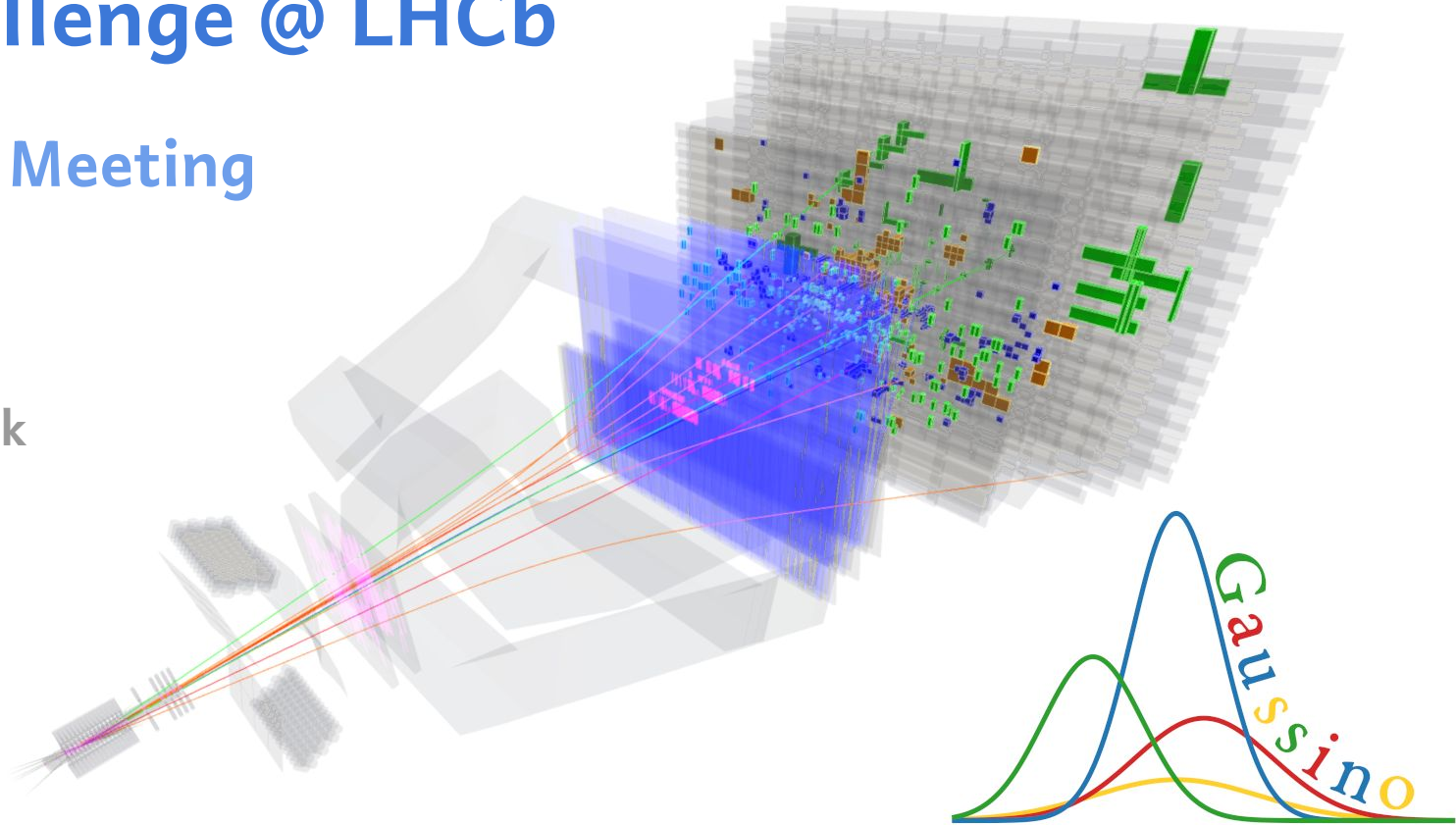
NATIONAL
CENTRE
FOR NUCLEAR
RESEARCH
ŚWIERK



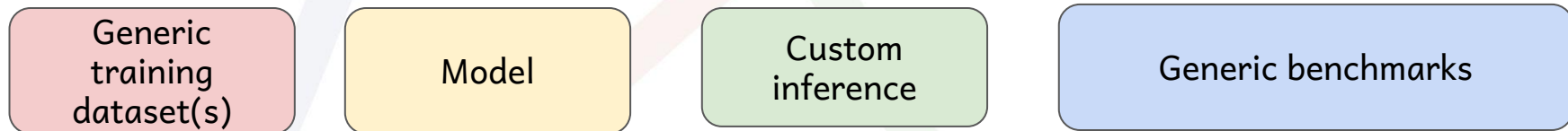
CaloChallenge @ LHCb

HSF & IML Meeting

Michał Mazurek
on behalf of the
Simulation Project

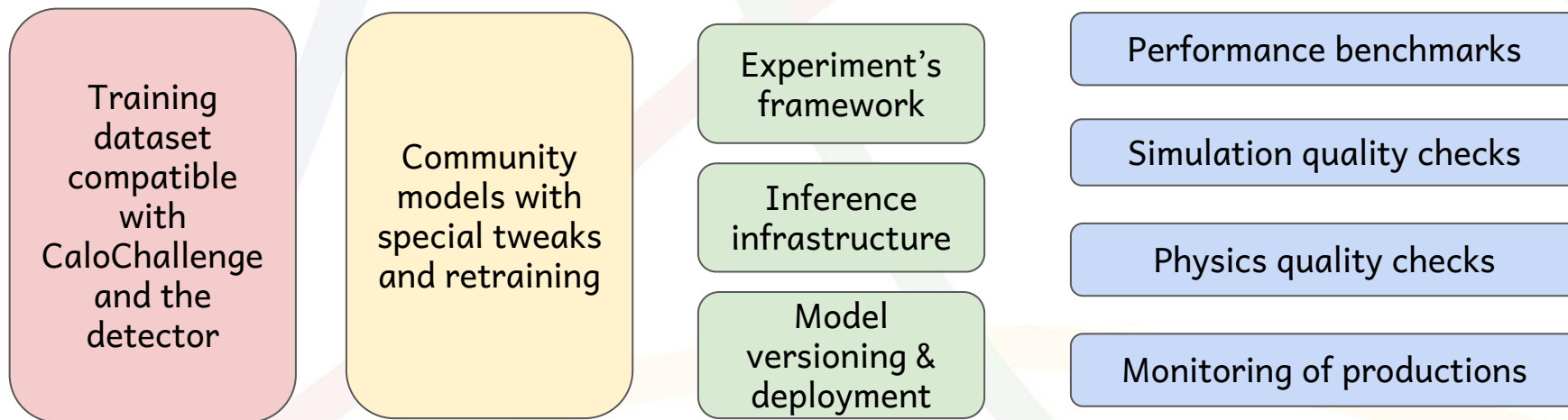


CaloChallenge workflow



This is already a lot of work!!

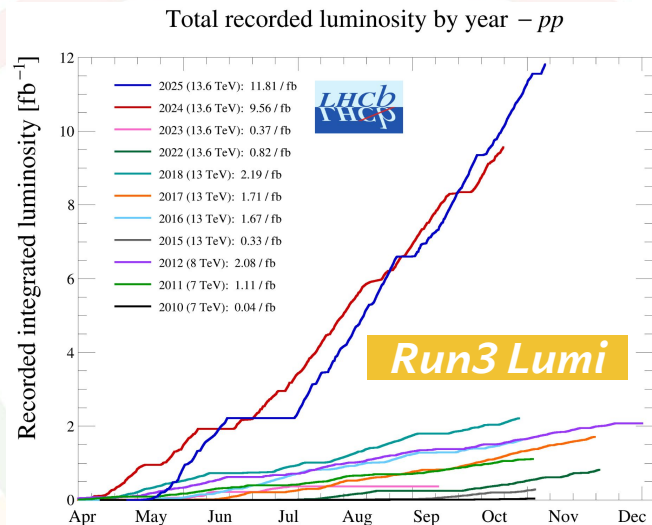
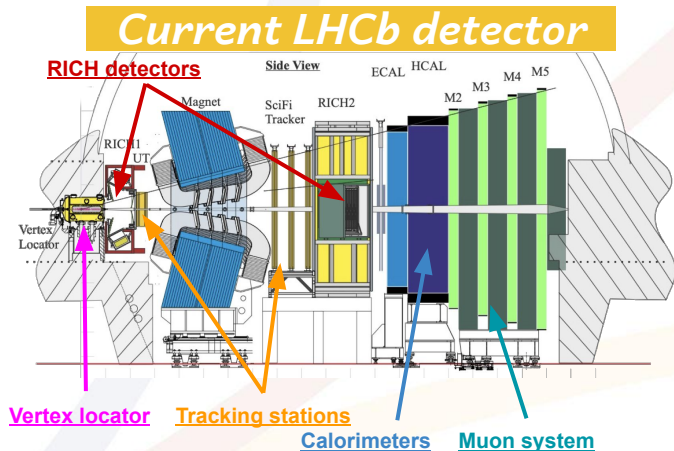
What experiments have to do extra



Let's take a look at the LHCb's case...

LHCb experiment in Run 3

- 👉 single-arm forward spectrometer
- 👉 designed to study decays involving b and c quarks
- 👉 taking data now with its upgraded detector!



- 👉 large increase in luminosity for Run3 and beyond very challenging for computing more events = more simulated events!
- 👉

Simulations in LHCb

Facts

- 👉 key to the **design** of the detectors and **interpretation** of the physics results
- 👉 must provide simulated samples for all physics runs & detector versions
- 👉 detailed simulations are **CPU intensive**

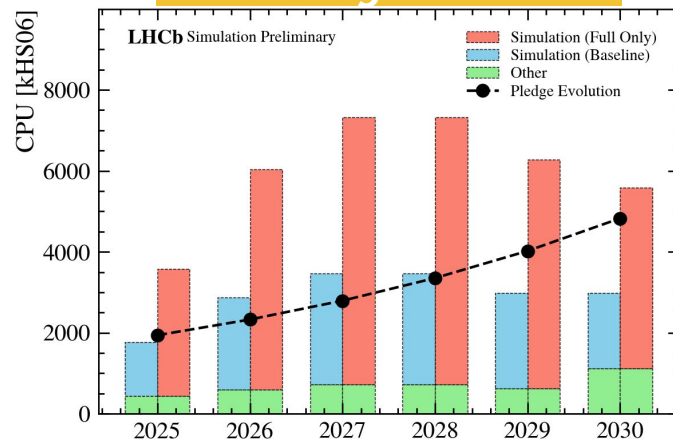
Landscape of Run 3 and future runs

- 👉 simulation **takes up to 90%** of the experiment offline computing resources
- 👉 still have to produce more samples
 - **efficient simulation framework**
 - complete palette of **fast simulations**



Faster simulations = Time vs. Quality

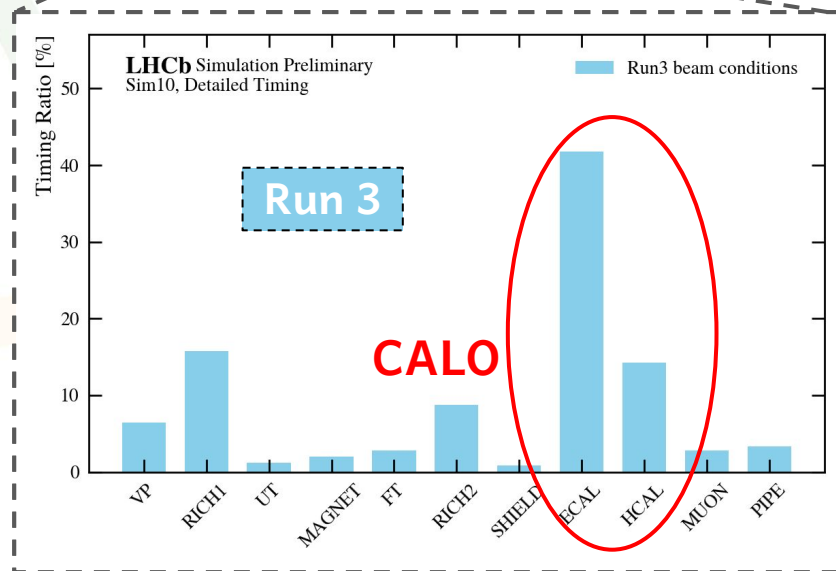
CPU usage forecast



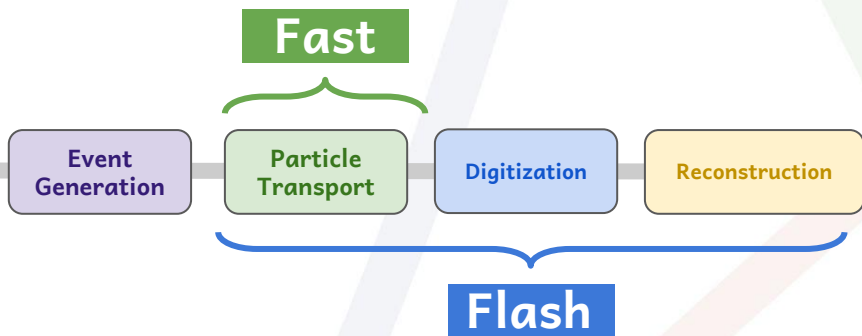
What is expensive in the LHCb Simulation?



- 👉 **particle transport** remains the main bottleneck in MC productions
- 👉 **calorimeter shower propagation** takes around **50 %** time of all the processes in the detector
- 👉 **Run 3 beam conditions** bring a noticeable increase in the simulation time



Fast & flash simulations in LHCb



CaloML

[CHEP 2024 talk](#)

*fast simulation option where the calorimeter response from Geant4 is replaced with **Generative AI***

Model	Event Generation	Particle Transport	Reconstruction
ReDecay	✓	✓	✗
ParticleGun	✓	✓	✗
SplitSim	✓	✓	✗
RICHless	✗	✓	✗
TrackerOnly	✗	✓	✗
Lamarr	✗	✓	✓
Point library	✗	✓	✗
Generative AI	✗	✓	✗

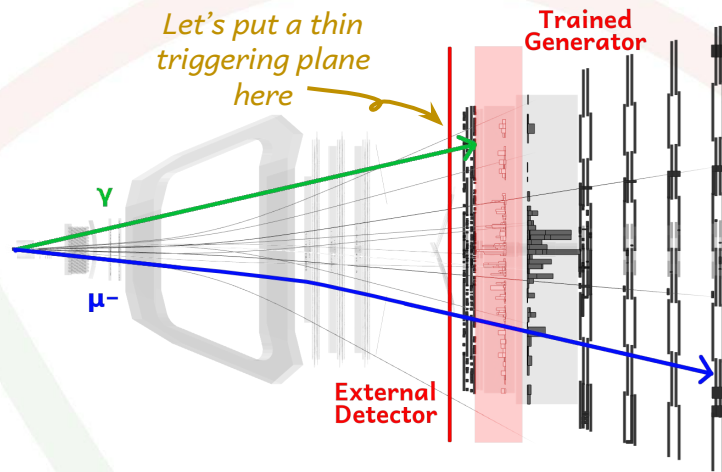
Lamarr

[CHEP 2023 talk](#)

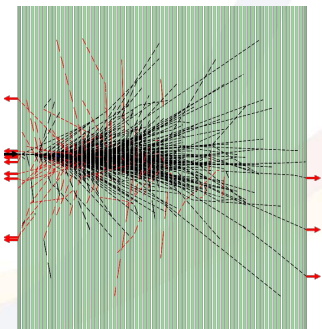
***ultra-fast simulation** option where not only the detector response, but also the **reconstruction** is parametrized*

ML fast simulations in calorimeter

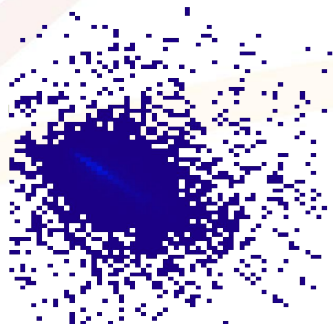
- 👉 stop detailed simulation in a particular region of the detector
- 👉 use a **parametric model** to produce **similar output**



What happens in Geant4?



What is actually stored?

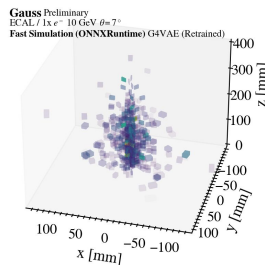
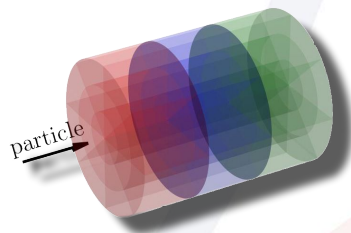


- 👉 **train an AI model** on the output
- 👉 produce hits by running **inference** on the **generator** part

Generic showers in LHCb with CaloChallenge

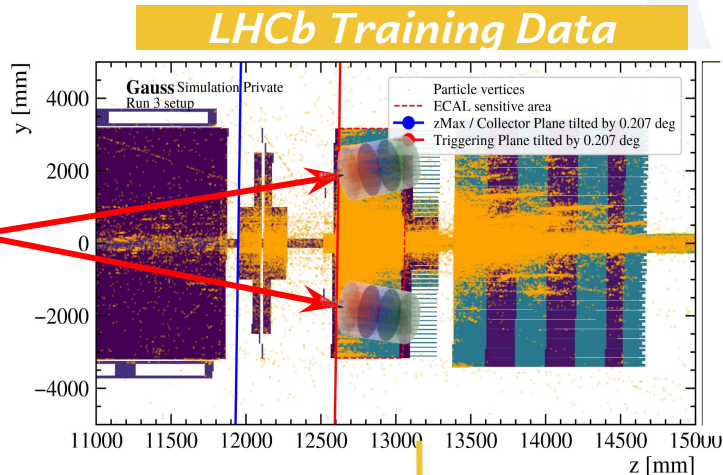
arXiv:2410.21611
MetaHEP, A. Zaborowska et al.

Build a *cylinder of virtual hits*
around a particle shower



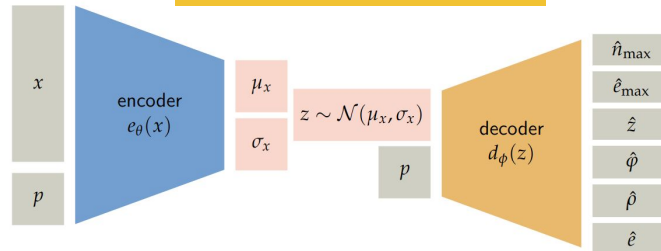
making things generic = minimizing the work later

- compare models **objectively** across geometries
- adapt easily by **retraining** on LHCb's calo data



Training

ML Model



Inference workflow

Gaussino

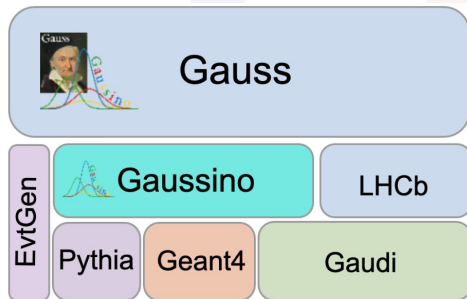
- 👉 core simulation framework for LHCb
- 👉 interface to fast simulations to a given detector
→ via Geant4
- 👉 interface to ML libraries [see CHEP 2023 talk](#)
→ pyTorch and ONNXRuntime

Gauss

- 👉 introduces the LHCb geometry & conditions

PROTOTYPE

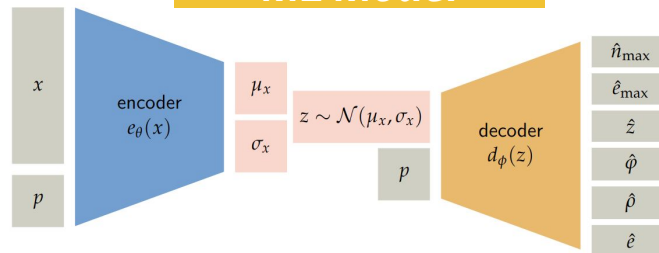
COMMISSIONING PRODUCTION



we're here!

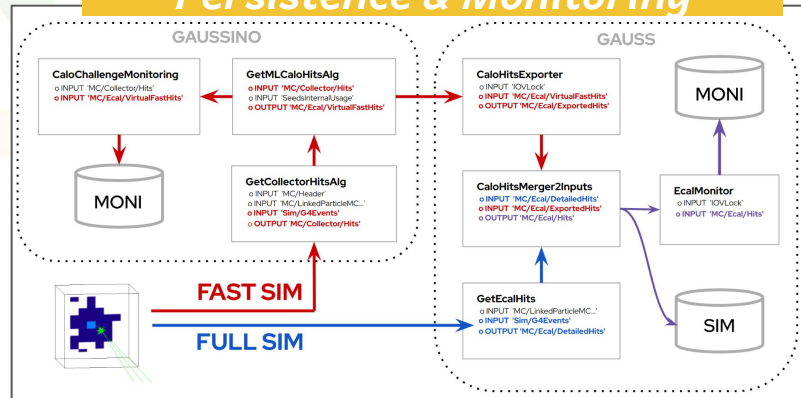
First official productions planned this year 🎉

ML Model



Inference

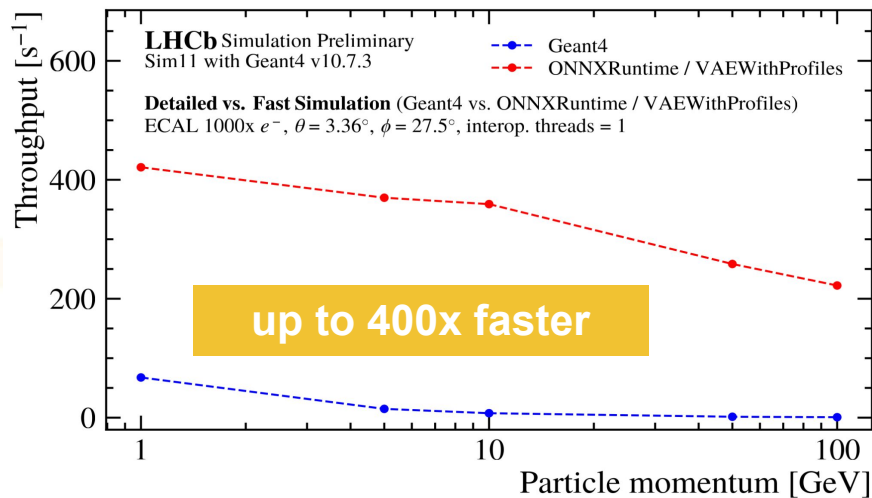
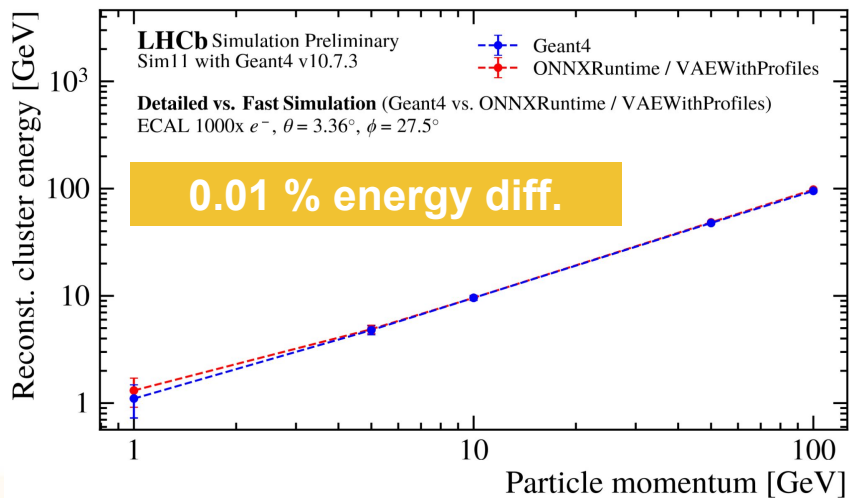
Persistence & Monitoring



Performance

- 👉 one simple model (VAE) for all PIDs (electrons, photons)
- 👉 around **0.01% energy difference** on reconstructed objects
- 👉 up to **400x times faster** for electrons and photons in ECAL
- 👉 hadrons not included yet **WIP**

NOW
NOW
NOW

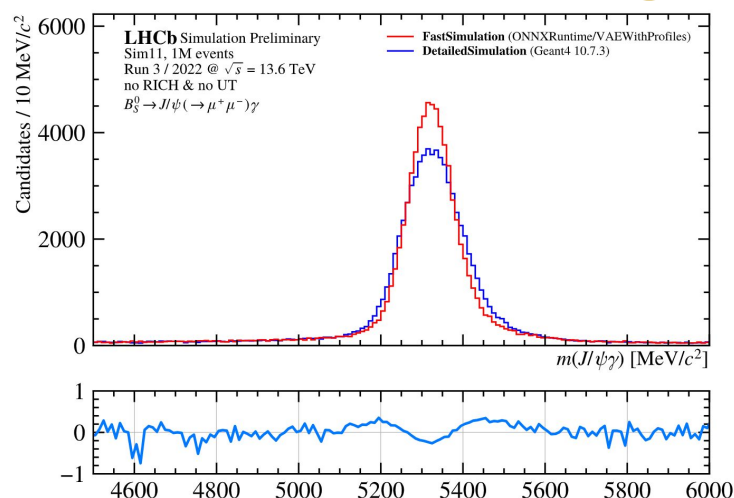
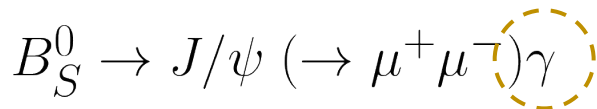
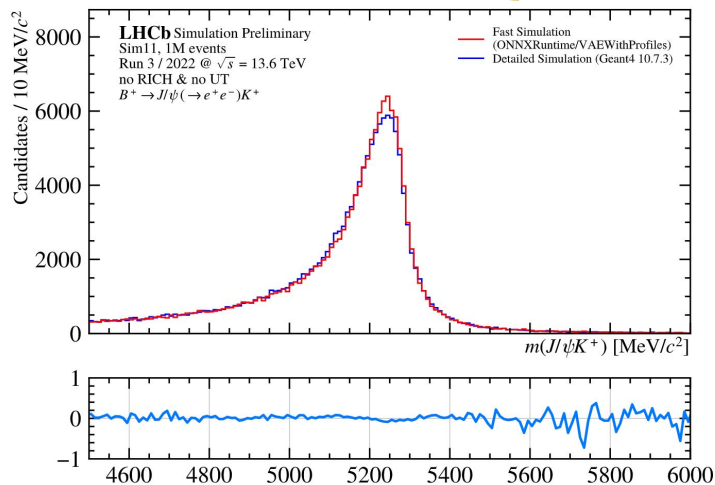
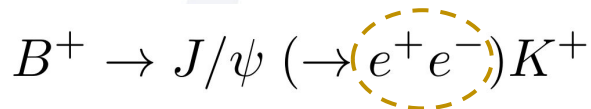


Validation with representative LHCb decay channels

- 👉 very good agreement with reconstructed physics observables
- 👉 small differences to be solved with more advanced models

NOW

WIP



Conclusions and future outlook

- 👉 The **increased luminosity in Run 3 and beyond** of the LHCb experiment places high demands on computational resources → **more simulated data is needed**,
- 👉 ML-based **fast and flash simulations** are essential to meet this challenge:
 - **CaloML = CaloChallenge at LHCb**: provides a fast simulation of calorimeter responses using generative AI, achieving up to 400x speedup in ECAL while maintaining high accuracy and very good agreement with physics observables
- 👉 **Plans 2025/2026:**
 - **New architectures,**
 - **Optimisation of the setup,**
 - **Run 3 & Upgrade 2 studies,**
 - **Hadron and punch-through studies.**

Thank you! Questions?