Tools and techniques in high-precision particle physics: a case study of the LHCb experiment

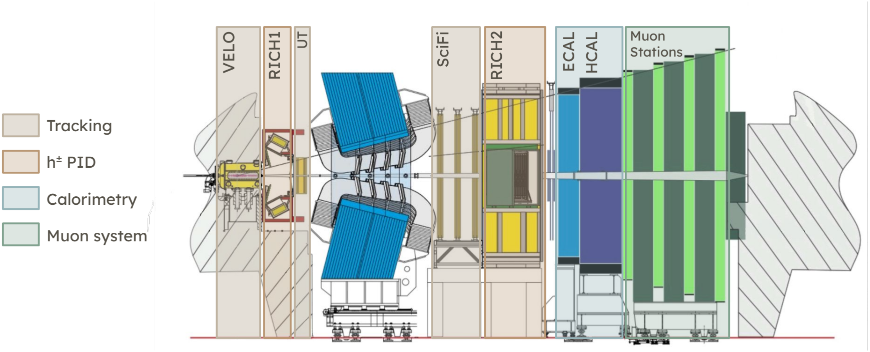
Marion Lehuraux, from the University of Warwick,
on behalf of the LHCb collaboration

EDSU Tool 2024

June 3, 2024

LHCb in one slide: WHY and HOW ?

- Dedicated heavy flavour experiment @LHC → forward spectrometer
  - Study CPV in the beauty sector and rare heavy hadrons decays
- … but also a general purpose detector in the forward region
  - QCD, heavy ions, electroweak, exotic spectroscopy etc.
- Difficulty: targets branching ratios from $\sim 10^{-9}$ to $\sim 10^{-1}$

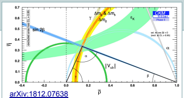


LHCb Upgrade I

Why do we need cutting-edge tools & techniques on LHCb ?

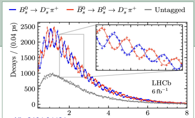
- The rich physics program of LHCb for Run 3 and post-Upgrade II

Precise ($\mathcal{O}$ 0.35°) measurement of CKM angle γ



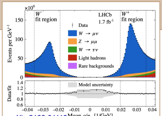
arXiv:1812.07638

Search for CPV in B^0 and B^0_s mixing

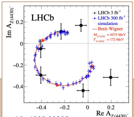


arXiv:2104.04421

Precise ($\mathcal{O}$ MeV, stat) measurement of W mass

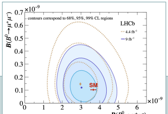


arXiv:2109.01113



arXiv:1808.08865

Exotic hadron spectroscopy

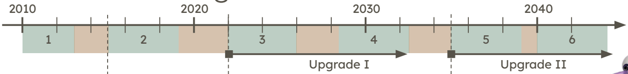


arXiv:2108.09284

$\mathcal{B}(B^0\to\mu\mu)/\mathcal{B}(B^0_s\to\mu\mu)$

and so much more...

- Cope with the experimental challenges

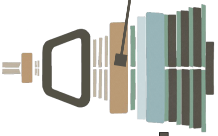


	Run 1	Run 2	Upgrade I	Upgrade II
Nominal instantaneous luminosity	~4×10^{32} cm^{2}s^{-1}	~4×10^{32} cm^{2}s^{-1}	~2.5×10^{33} cm^{2}s^{-1}	~1.5×10^{34} cm^{2}s^{-1}
Recorded luminosity	~3 fb^{-1}	~8 fb^{-1}	~50 fb^{-1}	~300 fb^{-1}
Pile Up	~1	~1	~5	~42
Trigger input	1 MHz, partial	1 MHz, partial	30 MHz, full detector readout	-
Maximum trigger output rate (1st stage)	5 kHZ	12.5 kHZ	2 MHz	-
Bandwidth to storage	0.3 GB/s	0.6 GB/s	10 GB/s	-
Number of events to simulate	–	~215×10^9 events	~3500×10^9 events	~8000×10^9 events

Tools & techniques on the LHCb experiment

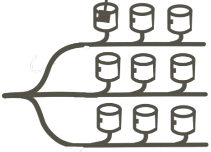
Detector

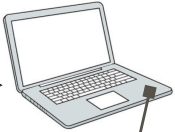
- State of the art detectors & electronics
  (Not detailed in this presentation)



Online Analysis

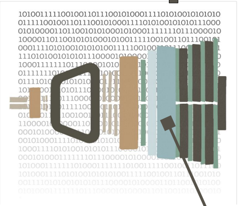
- Full detector readout & event builder at 30 MHz
- ML-based tracking algorithms
- Fake tracks rejection
- Global multivariate classifiers for PID
- π^0 - γ separation
- ...





Offline Analysis

- Gradient Boosted Reweighting for MC/data corrections
- BDTs/NNs for selections
- Amplitude models
- ...

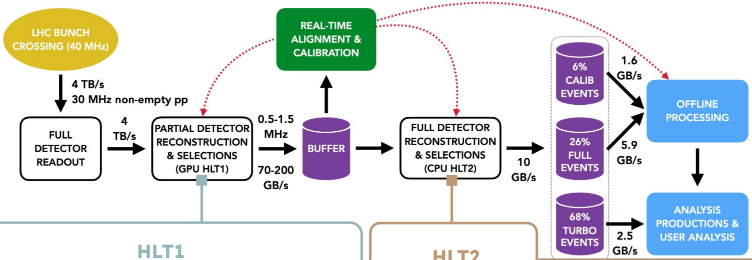


- R&D: Using generators on GPUs
- R&D: Hadronization using ML
- R&D: ML to speed up simulations
- ...

Simulation

LHCb software trigger [*Parallel architecture, Algorithms*]

- Aims to reduce the data volume from 4 TB/s to about 10 GB/s
- Too many b and c hadrons to select events based on $p_T \rightarrow$ *Real time analysis approach*



HLT1

Inclusive high-level trigger based on charged particles reconstruction

Looking Forward, Search by triplet, CompassUT, HybridSeeding, VELO-SciFi matching...

HLT2

Full offline-quality reconstruction on all events using real-time alignment & calibration

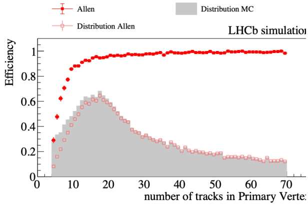
Support $\mathcal{O}(1000)$ independent selection “lines”

Pattern recognition + Kalman fit → track
Global multivariate classifiers → PID

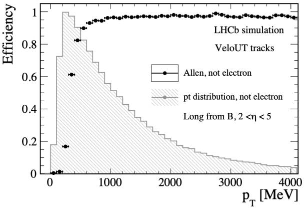
HLT1 @30 MHz [*Parallel architecture, Algorithms*]

- Copy full event ($\sim$ 100 kB) to GPU
- Run HLT1 @30 MHz on $\sim$ 500 GPUs:
  - Decode RAW data in tracking sub-detectors
  - Clustering and track reconstruction in tracking sub-detectors
  - Track matching & primary vertices reconstruction
  - 2-track secondary vertices reconstruction
  - Apply selections
- Selections copied back to CPU

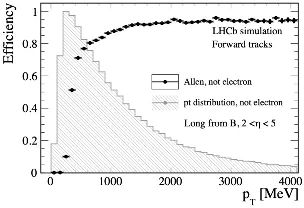
PV Reconstruction efficiency



VELO-UT tracking efficiency

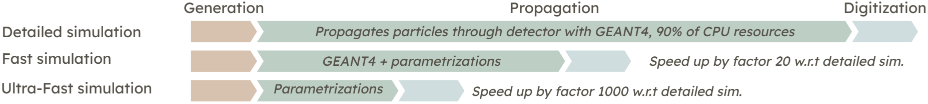


Sci-Fi tracking efficiency



(Ultra) Fast simulation in LHCb [*ML, AI*]

- Fast simulation is crucial to be able to analyze the larger datasets to be collected in Run 3 *(and later)* with the available computing resources



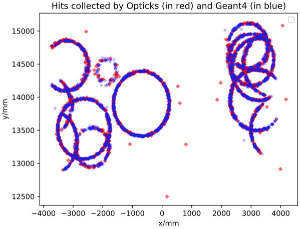
- Several R&D projects to address the question:
  - LAMARR, *ultra-fast* simulation using ML-based parametrizations
  - Propagating optical photons using NVIDIA OptiX ray tracing [Paper]
  - Fast simulation of the ECAL response using VAEs and GANs
  - ...

VAEs = Variational Auto-Encoders, GANs = Generative Adversial Networks

(Ultra) Fast simulation in LHCb: Promising results [*ML, AI*]

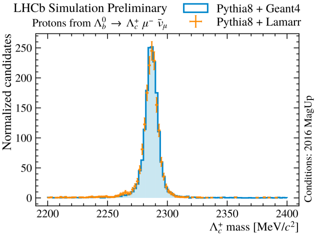
Propagating optical photons using NVIDIA OptiX ray tracing [Paper]

- GPU-accelerated optical photon interface
- Tested using RICH1 geometry
- Factor 5-10 gain in propagation time
- Need for more physics processes
- Challenging integration in LHCb simulation



LAMARR, *ultra-fast* simulation using ML-based parametrizations

- Uses Deep Generative models to parametrize detector response and reconstruction algorithms
- Experimental errors and uncertainties introduced in the detection and reconstruction phases encoded in NNs
- Factor 1000 w.r.t detailed simulation



LHCb would be nothing without its people ! [*Soft skills*]

"The Early Career, Gender and Diversity (ECGD) office oversees the well-being and working environment of all LHCb members."

1716 members, 100 institutes, 22 countries

(As of 24th May 2024)

- Plenary meetings at each LHCb week
- LHCb note: Effects of Covid on LHCb scientists
- Conducted a work-stress related survey
- Laura Bassi initiative [Anja Beck, Janina Nicolini]
  - Tackle under-representation of women in HEP
  - Promote discussion and ideas on related issues
  - Provide an informal network to colleagues
- LHC-wide Soft-skills workshops
  - Effective teamwork in large collaborations
  - Networking


William Barter
Sneha Malde
Lorenzo Paolucci
Eliot Jane Walton

Conclusions

- Use of cutting edge tools & techniques allowed LHCb to achieve great performances and publish important results over the years
- Upgrade of the LHCb detector motivated by a rich physics program
- Coping with such increasing luminosity implies unprecedented challenges for particle physics experiment
- LHCb is rising to the occasion:
  - Achieving 30 MHz full event readout ✓
  - Work in progress to speed up simulations ✓
  - Gain up to a factor 1000 in simulation
  - Some limitations inherent to data-driven simulations to tackle
- And it doesn't stop there...

Towards Upgrade II

- From Run 5, LHCb will enter the Upgrade II phase
- Tougher challenges that motivate further developments
- **Objective:** add an extra-dimension to the data to recover Upgrade I complexity

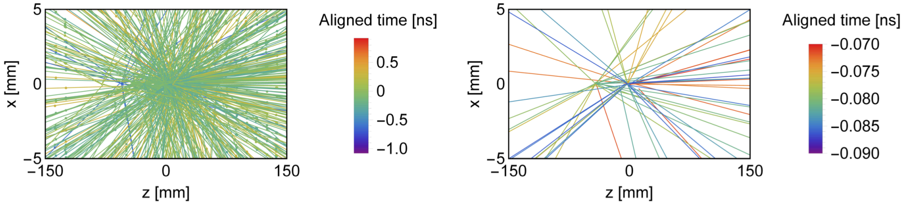


Illustration of the track density inside the VELO generated by 42 collisions spread over a bunch crossing. **Left:** whole bunch crossing time period ($\sim$ 1 ns). **Right:** Time window of 20 ps.

Thank you for your attention

References

- Links to arXiv papers listed on Slide 3:
  - arXiv:1812.07638
  - arXiv:1808.08865
  - arXiv:2104.04421
  - arXiv:2108.09284
  - arXiv:2109.01113
- Sources for numbers on Slide 3:
  - U2 Workshop May 2021, C. Bozzi
  - *Operation and Performance of the LHCb experiment*
  - *An LHCb Vertex Locator (VELO) for 2030s*
  - *Looking Forward (Track Following Algorithm)*
- Search by triplet: An efficient local track reconstruction algorithm for parallel architectures
- LHCb ECGD Office web page

References

- Hadronization using ML
- Search by triplet
- Standalone track reconstruction and matching algorithms for GPU-based High level trigger at LHCb
- LHCb Upgrade II FTDR