Real-time analysis in HLT1 at LHCb in Run 3

Kate Richardson on behalf of the LHCb collaboration

US LUA Annual Meeting, SLAC

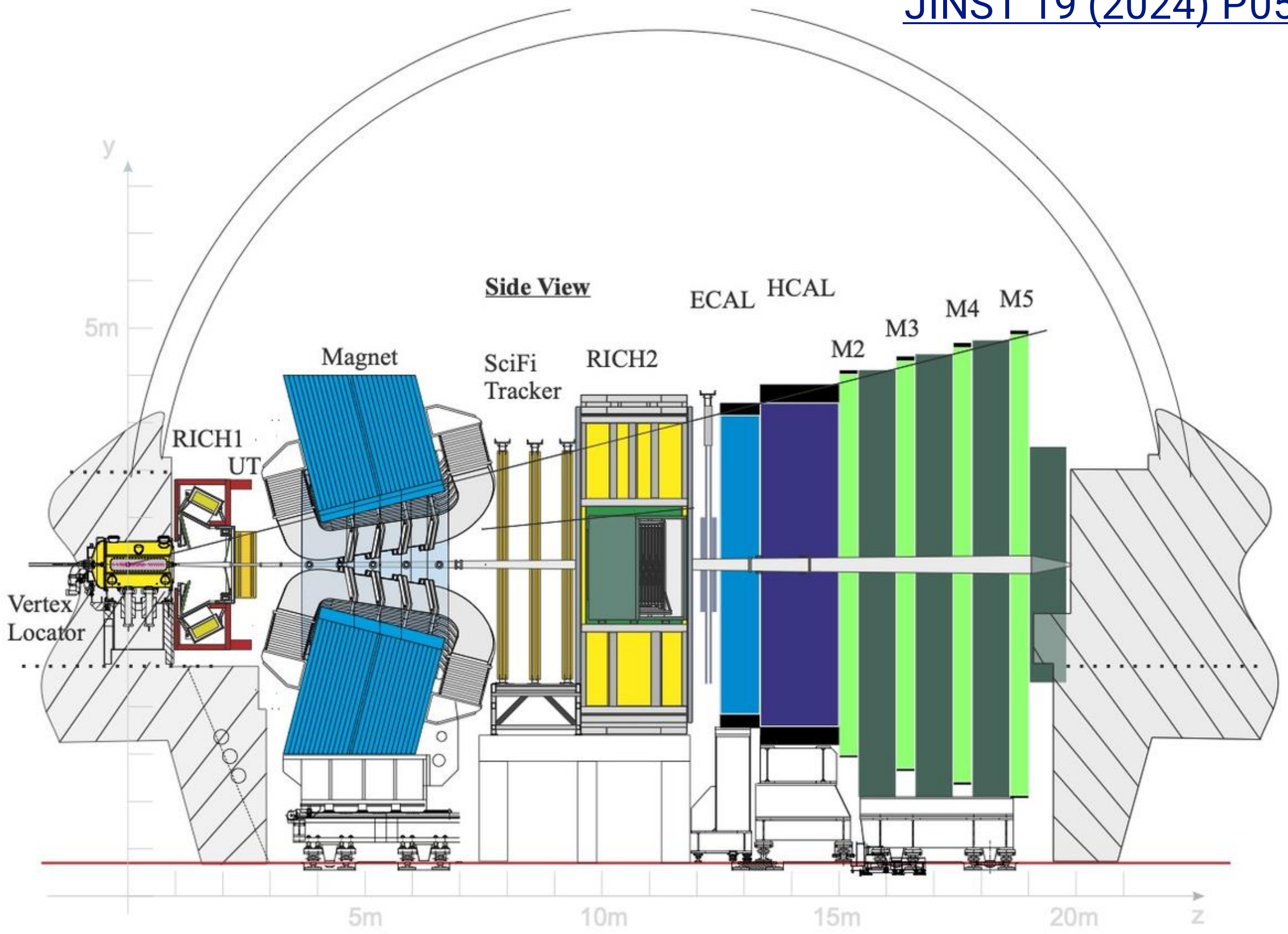
17 December 2024

Massachusetts Institute of Technology

The upgraded LHCb detector

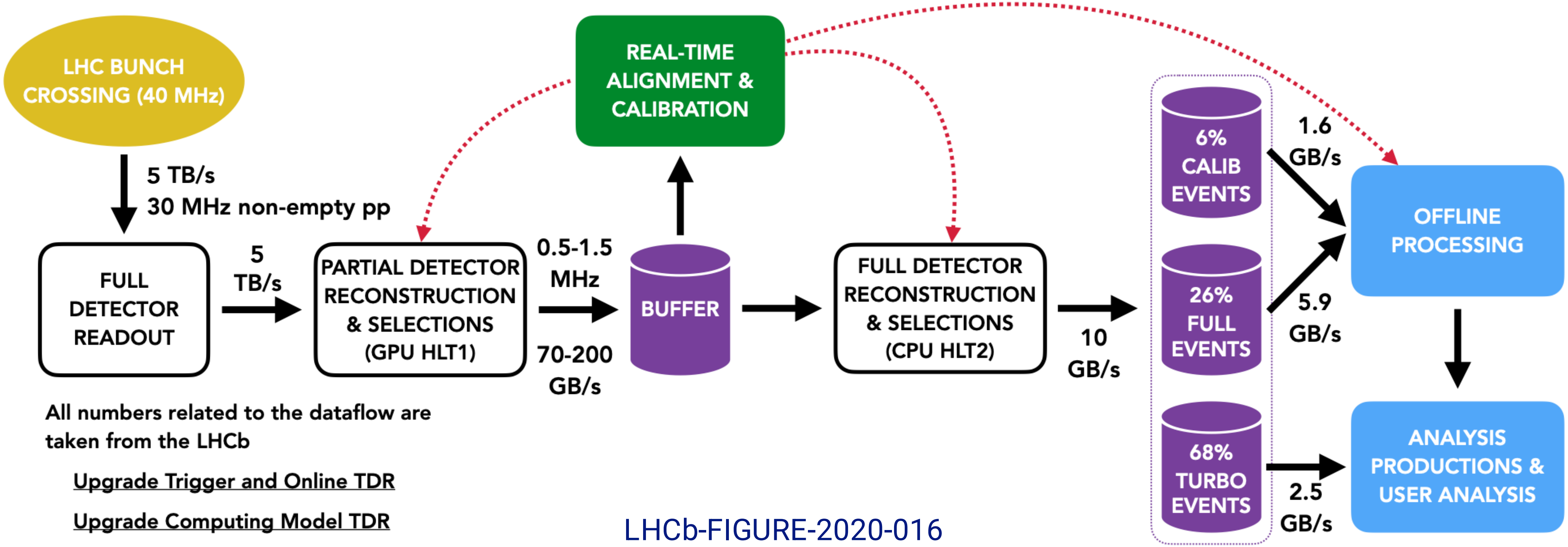
JINST 19 (2024) P05065

- Increased instantaneous luminosity $5 \times$ from Run 2 to 2×10^{33}cm^{-2}s^{-1}
- New tracking detectors
- Replaced readout electronics to meet rate requirements
- **Removed hardware trigger**



Dataflow in LHCb Run 3

- Software-only trigger allows more flexible selections



All numbers related to the dataflow are taken from the LHCb

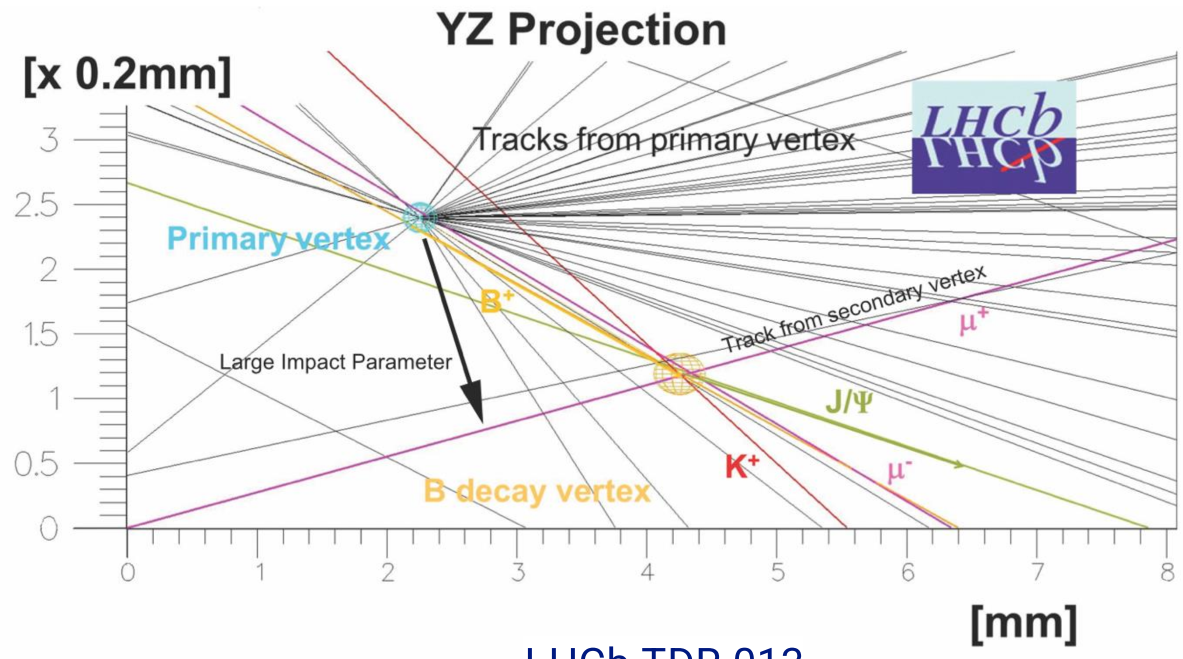
Upgrade Trigger and Online TDR

Upgrade Computing Model TDR

LHCb-FIGURE-2020-016

HLT1 requirements

- To perform selections at HLT1 level, we need:
  - Subdetector reconstruction for VELO, UT, SciFi, ECAL, and MUON
  - Primary and secondary vertex reconstruction
  - Track fitting
  - Electron and muon PID
- All at a 30 MHz rate!



LHCb-TDR-013

Why GPUs?

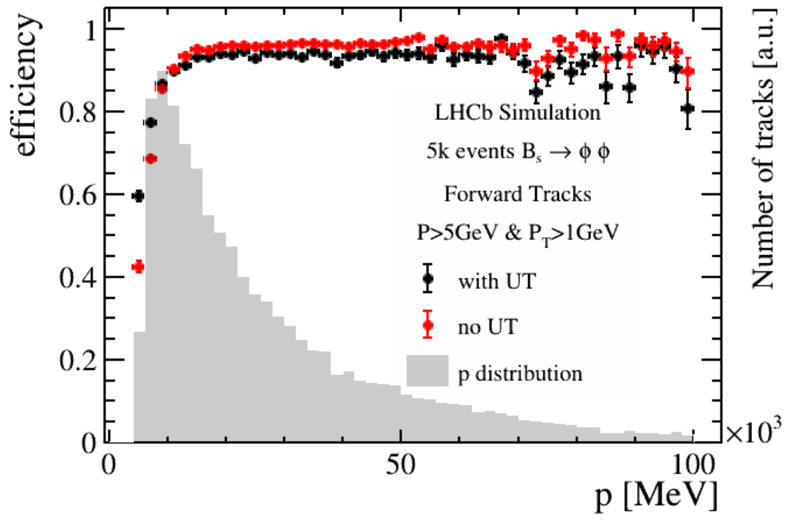
- HLT1 is an inherently parallelizable task, at multiple levels
  - We run multiple streams on each GPU, each with a slice of events
  - Within algorithms, threads are used to parallelize over objects (vertices, tracks, etc.)
- Limited I/O bandwidth is acceptable because small raw event data $\mathcal{O}$(100 kB/event) means thousands of events still fit in $\mathcal{O}$(10 GB) memory
- Cheaper and more scalable than CPU alternative
- Fit well into LHCb DAQ architecture
- Run with 489 Nvidia RTX A5000 GPUs

Comput. Softw. Big Sci. 4 (2020) 7

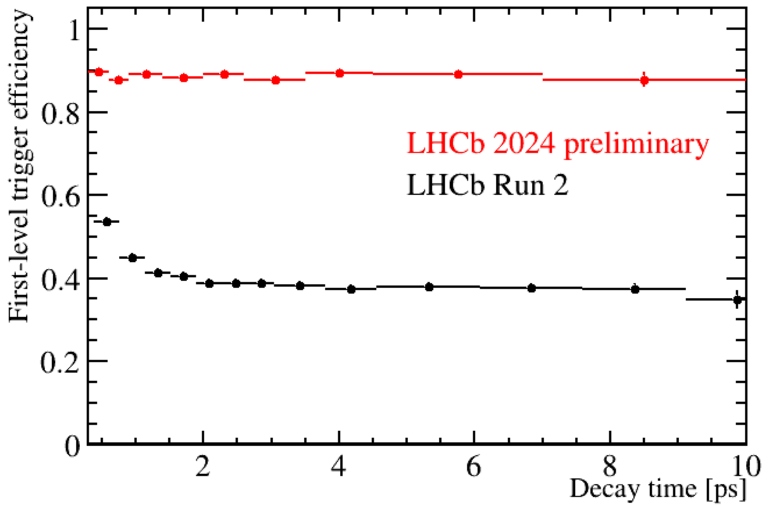
HLT1 performance

- Excellent track reconstruction efficiency
- Efficiencies equal or better compared to Run 2

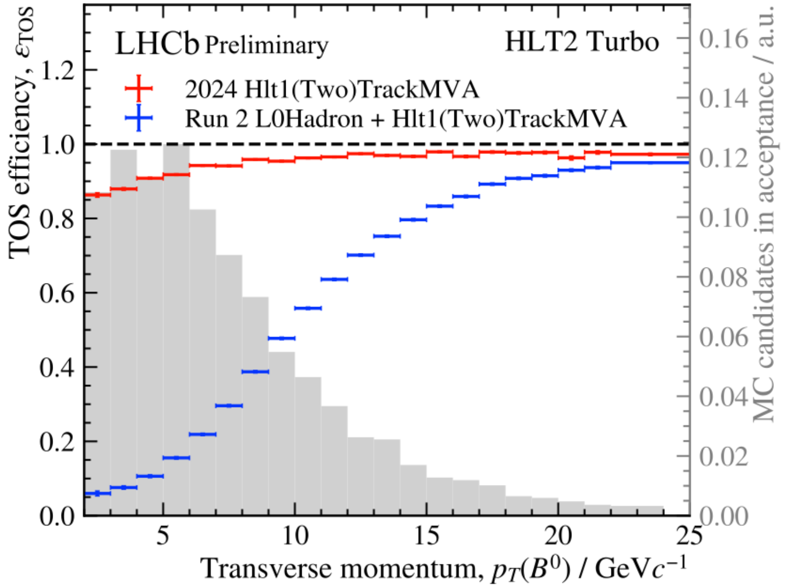
MVA lines



LHCb-FIGURE-2022-007



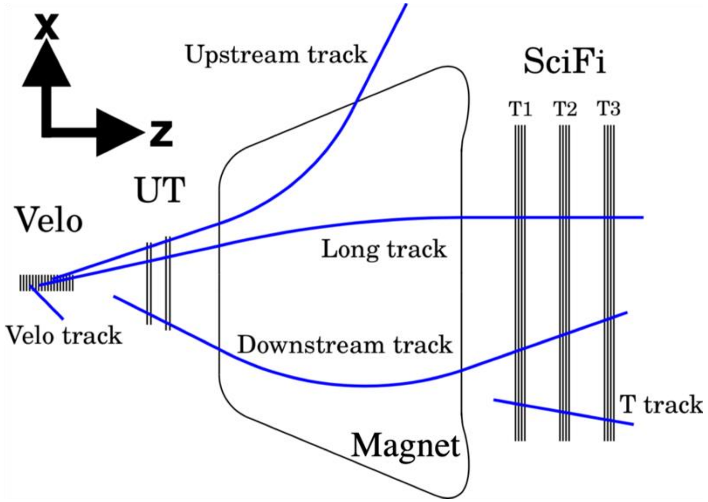
LHCb-FIGURE-2024-014



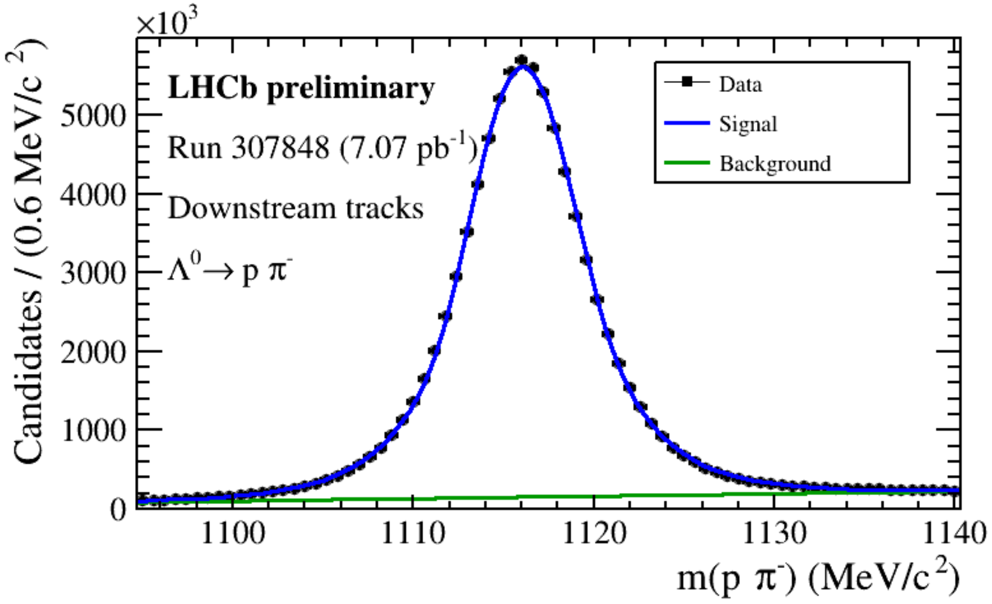
LHCb-FIGURE-2024-030

Going beyond the design requirements

- Original goal was only to make selections using long tracks in HLT1 but now also reconstructing downstream tracks
- Jet reconstruction using an iterative cone algorithm is performed and used for selection
- And more!



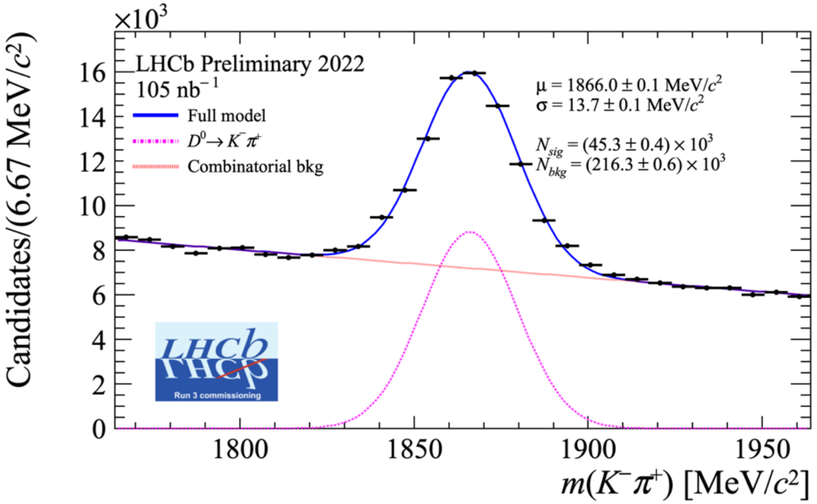
LHCb-DP-2022-002



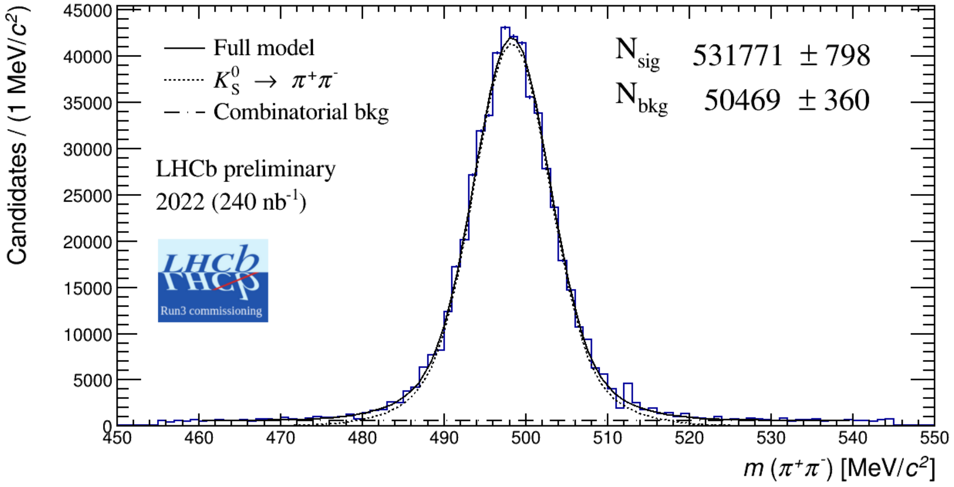
LHCb-FIGURE-2024-035

HLT1 online monitoring

- Monitoring is necessary to allow for real-time supervision of reconstruction and selections in HLT1 to find issues quickly
- LHCb has a monitoring infrastructure for aggregation and display but HLT1 monitoring buffers need to be periodically transferred to host
- Monitor all events at full 30 MHz rate → access to events discarded by HLT1



LHCb-FIGURE-2023-009

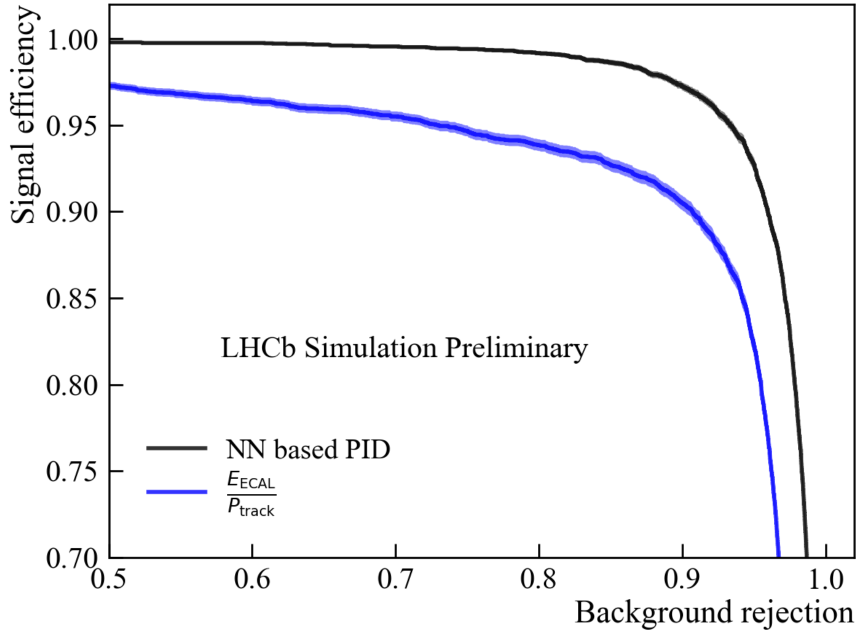


LHCb-FIGURE-2023-005

Machine learning in HLT1

- Requirements of running in the HLT1 environment are that any model needs to be small and fast
- For physics we want our models to be robust and possibly monotonic
- Use Lipschitz neural networks to achieve this
- Currently in use for PID, selections, and more

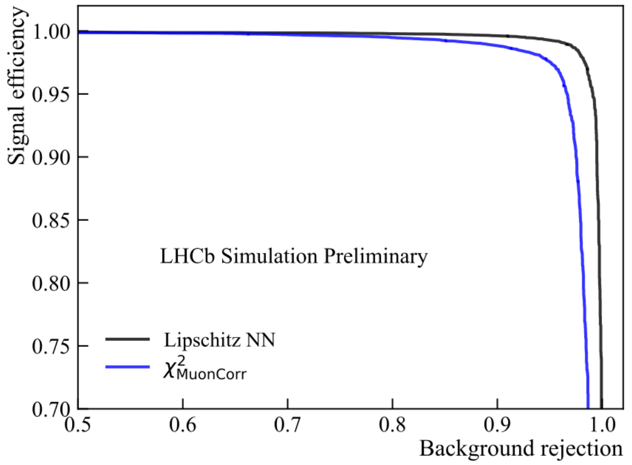
Mach.Learn.Sci.Tech. 4 (2023) 3, 035020

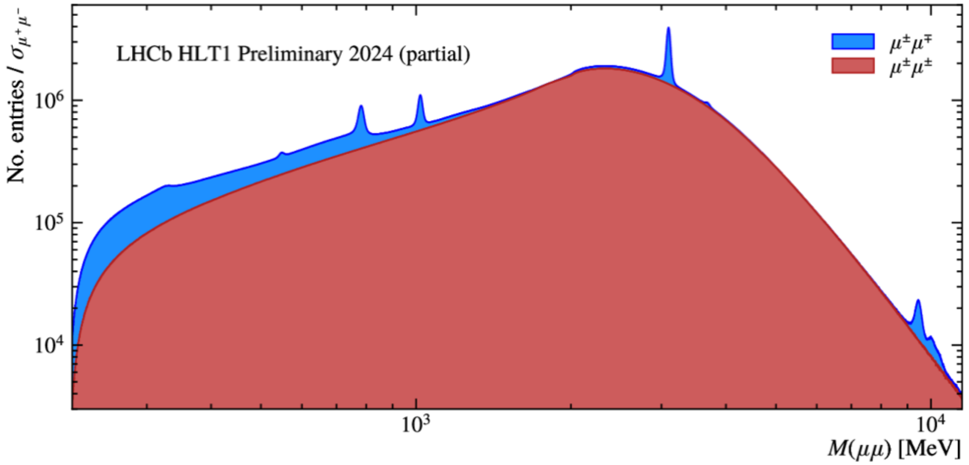


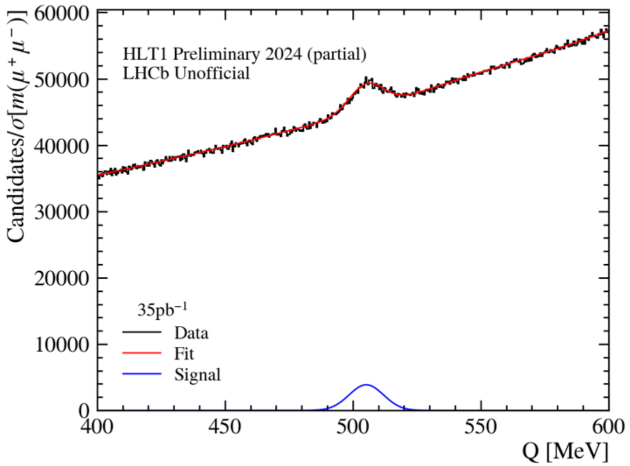
LHCb-FIGURE-2024-003

Case study: $A' \to \mu^+\mu^-$

- All prompt candidates included in monitoring histogram
- Muon ID NN used for selection to reduce background
- Approximately **20 times increase** in $\eta \to \mu^+\mu^-$ yield per fb^{-1} over Run 2







LHCb-FIGURE-2024-029

$$\mathcal{Q} \equiv \sqrt{m(\mu\mu)^2 - 4m(\mu)^2}$$

Summary

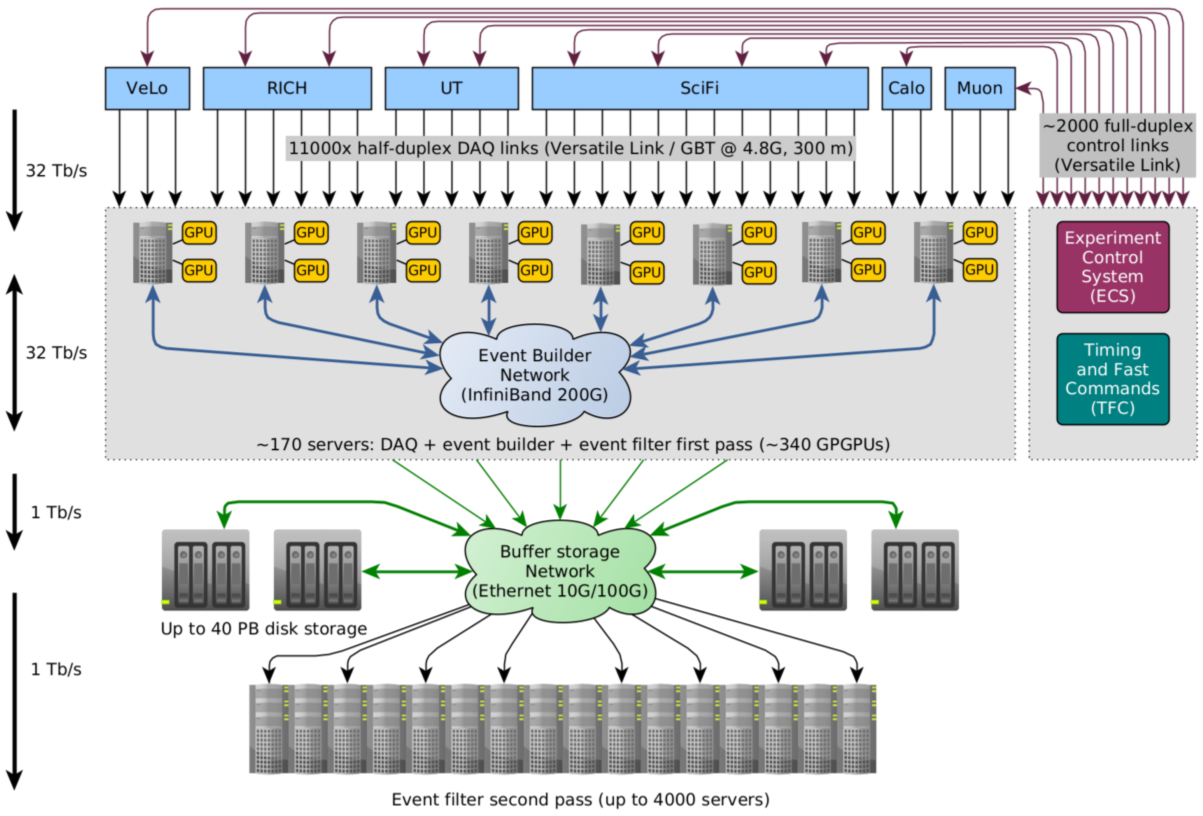
- LHCb is running a completely software-based trigger at 30 MHz (5 TB/s)
- HLT1 completes a partial reconstruction and selection at the full event rate by utilizing parallelization on GPUs
- The HLT1 implementation includes many algorithms beyond the original design scope, opening doors for additional interesting physics
- Large improvements in trigger efficiencies compared to the Run 2 system

Thank you for listening!

Backup

DAQ architecture

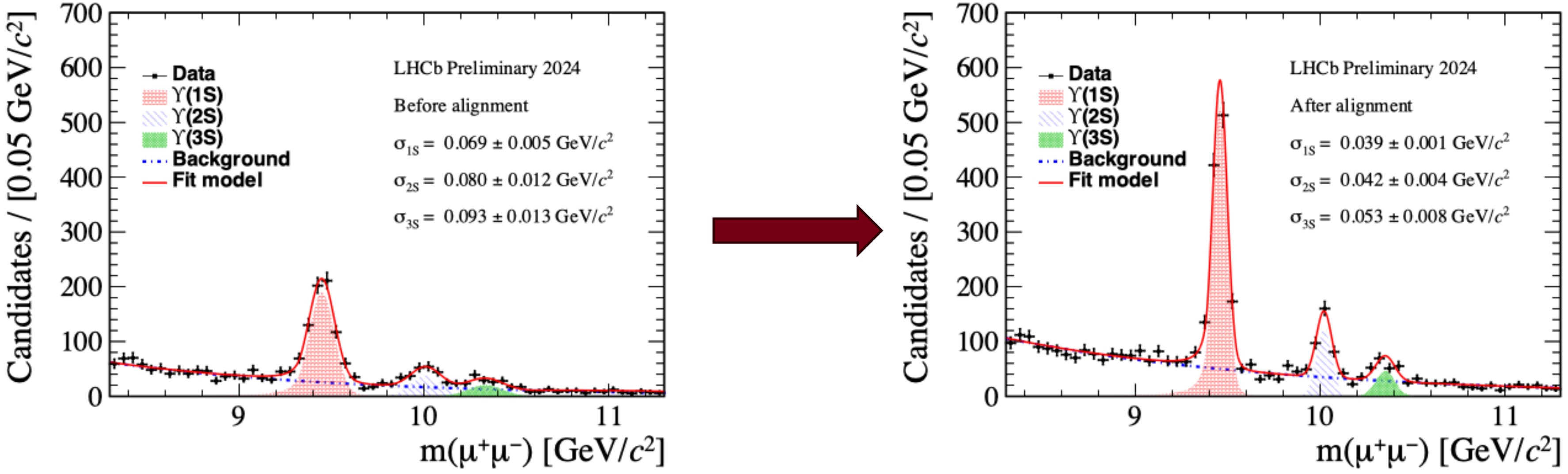
- $\mathcal{O}(500)$ FPGA readout boards receive data from subdetectors at 30 MHz
- Event builder units reorder raw data from front-end boards into event packets to be processed by HLT1
- Throughput of 5 TB/s
- Off-the-shelf components reduce cost



JINST 19 (2024) P05065

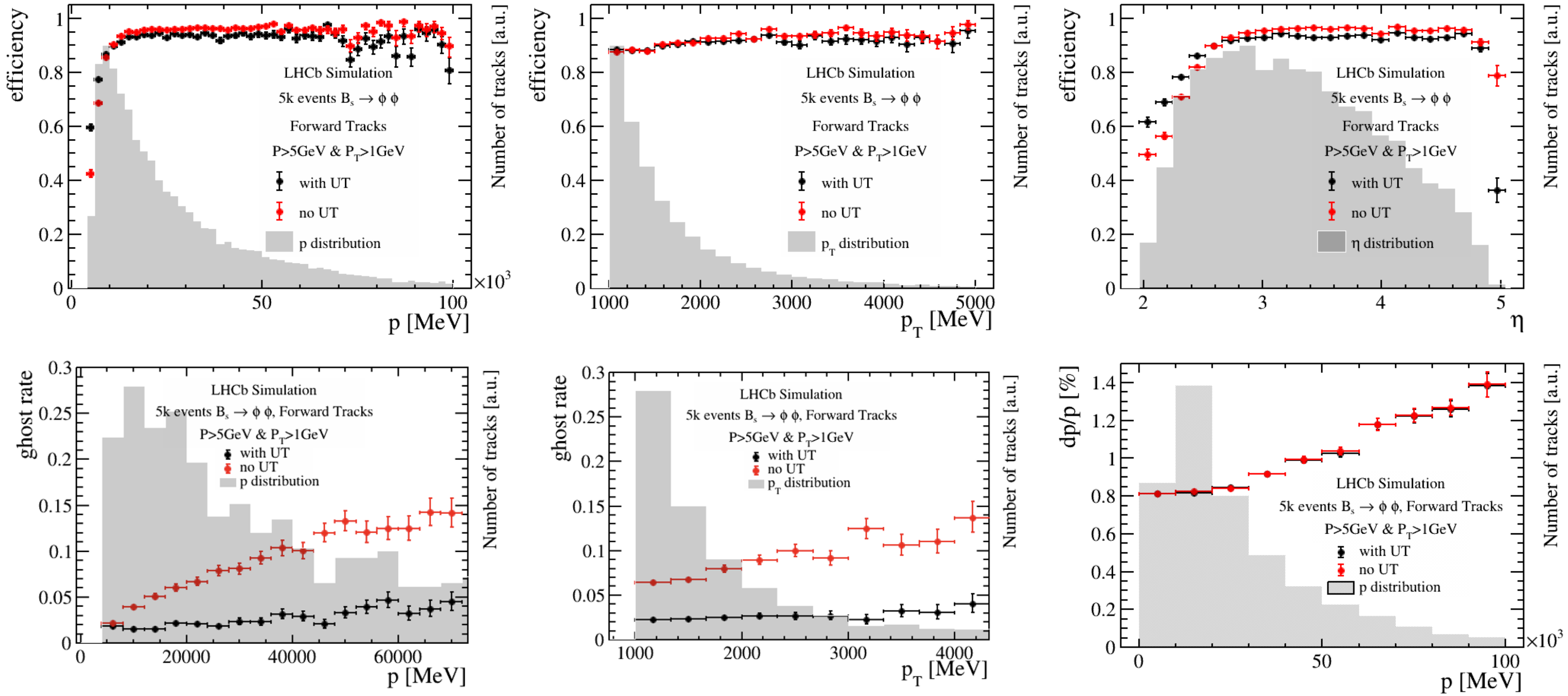
Alignment and calibration

- Alignment for VELO, RICH mirrors, UT, SciFi, and MUON and calibration for RICH, ECAL, and HCAL run automatically every fill
- Alignment process based on analyzer and iterator to obtain convergence
- Constants calculated in this process are periodically updated in HLT1



LHCb-FIGURE-2024-037

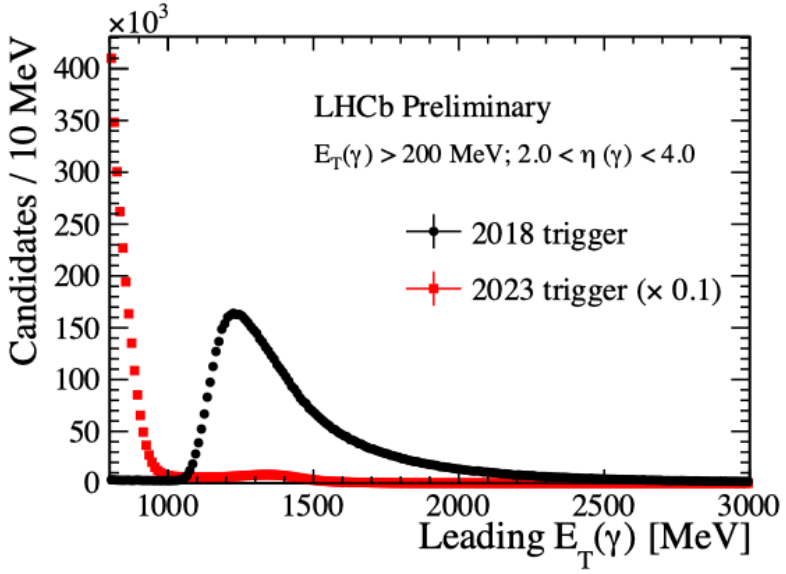
HLT1 tracking efficiencies



LHCb-FIGURE-2022-007

More beyond the design requirements!

- Able to lower the E_T to 200 MeV for diphoton PbPb (compared to 1 GeV in the hardware trigger for Run 2)
- First ever K_S^0 selection in a first-level trigger enabling new $A^{CP}(D^0 \rightarrow K_S^0 K_S^0)$ measurement from only 3 months of Run 3 data due to increased efficiencies



LHCb-FIGURE-2024-012

Stay tuned!