PROCEEDINGS OF SCIENCE

First performance of the real-time reconstruction at LHCb

Daniel Cervenkov[a,1]

[a] *University of Oxford*

E-mail: daniel.cervenkov@physics.ox.ac.uk

We present an overview of the real-time reconstruction at LHCb. A new software trigger provides "offline-quality" reconstructed objects at 30 MHz and allows for a sophisticated trigger selection. This allows LHCb to deal with an increased luminosity and pile-up and to improve the hadronic trigger efficiency. We also present a new particle identification calibration approach and tools, which allow a more accurate efficiency estimation.

41st International Conference on High Energy Physics - ICHEP2022
6-13 July 2022
Bologna, Italy

---

[1] On behalf of the LHCb collaboration

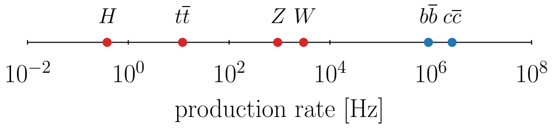


Figure 1: Production rates of typical signals at ATLAS, CMS (red), and LHCb (blue).

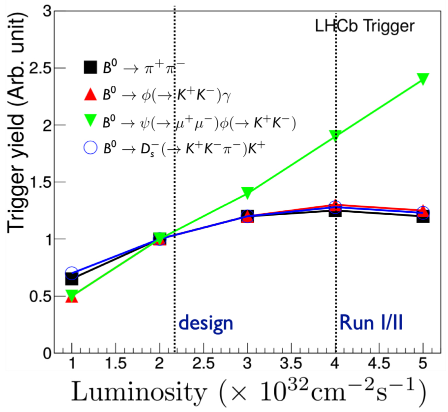


Figure 2: LHCb trigger yields for several channels around Run 1–2 luminosity [1].

1. Introduction

In order to understand the requirements for the real-time reconstruction at LHCb, it is instructive to look at the data rates at some of the LHC experiments. The typical luminosity of ATLAS and CMS during Run 2 was an order of magnitude above LHCb's $\mathcal{L} = 2 \times 10^{33}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$; however, the production rates of their typical signals were much lower; see Fig. 1. Furthermore, the $b\bar{b}$ and $c\bar{c}$ events look a lot like the underlying background events from proton-proton collisions, so triggering on these events is challenging. This is especially true on the fast hardware trigger level. In fact, the trigger was already saturated for many hadronic channels at Run 1–2 luminosity; see Fig. 2.

The solution to this issue was to move to a fully software trigger. The new trigger does an “offline-quality” online reconstruction at 30 MHz. All this happens almost in real-time, the caveat being the existence of a 11 PB buffer, which gives us about two weeks to catch up when necessary. Most events (TURBO) use “selective persistence”, meaning not all the information from the event is saved, but only the potentially interesting parts. This selectivity saves us a lot of bandwidth as the TURBO events are up to 5 times smaller than the full events. The complete information is kept for a part of the events, which are then used for calibration and specialized analyses. The software trigger allows us to use a more sophisticated selection than would be possible using a hardware trigger, reducing the background data rate. The approach increases the hadronic trigger efficiency by a factor of 2–4 compared to Run 2. The LHCb dataflow schematic is shown in Fig. 3.

2. LHCb Upgrade

The LHCb underwent a major upgrade for Run 3. The innermost tracker system, VELO, was upgraded to a pixel vertex detector with 55 µm × 55 µm pixels. The change will result in lower

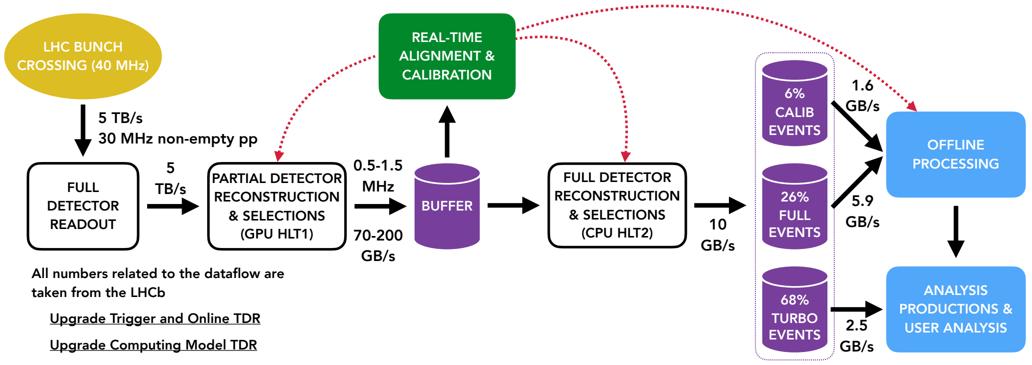


Figure 3: LHCb Run 3 dataflow [2].

occupancy and up to 40% interaction point resolution improvement. There is a new fine-granularity scintillating fiber detector (SciFi) and silicon-strip tracker (UT), with a 250 µm fiber diameter and 190/95 µm strip pitch, respectively. Finally, LHCb has a new MaPMT-based high-resolution hadron particle identification Cherenkov detector (RICH), with a 0.45–0.78 mrad Cherenkov angle resolution.

All of the readout electronics and DAQ were replaced to deal with the new conditions. In particular, the luminosity will increase by a factor of 5 to $\mathcal{L} = 2 \times 10^{33}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, and we will have to deal with pile-up, which will increase from 1.1 to 5–6. The plan is to record $15\,\mathrm{fb}^{-1}$ in Run 3 and $50\,\mathrm{fb}^{-1}$ in Run 3+4; compared to $6\,\mathrm{fb}^{-1}$ in Run 2.

3. LHCb Real-time Tracking

There are several types of tracks in LHCb that differ in the tracking procedure used. An overview can be seen in Fig. 4. “VELO” tracks are reconstructed by pattern recognition of straight tracks solely in the VELO detector. The efficiency is around 99% across the whole momentum range and can be seen in Fig. 5a. “Upstream tracks” pass through the VELO and the UT, but no other detectors; their HLT1 efficiency can be seen in Fig. 5b. “Long tracks” pass through the VELO, the UT, and the SciFi. The HLT2 “forward tracking” technique utilizes the fringe magnetic field in the UT to measure momentum and extrapolate the tracks to the SciFi. The extrapolation makes the tracking algorithm much more efficient by limiting the SciFi search to a small window consistent with the measured momentum and the extrapolated track position. The efficiency can be seen in Fig. 5c. SciFi seeds are tracks reconstructed only in the SciFi layers. Seeds from B decays for non-$e^{\pm}$ $p_{\mathrm{T}} > 1\,\mathrm{GeV}$ tracks are found with 95% efficiency; see Fig. 5d. Finally, long tracks can also be reconstructed via track matching — using a neural network to match VELO tracks with SciFi seeds. This approach is more efficient than forward tracking, so it is tried first. Residual unmatched VELO tracks are then used to find long tracks via forward tracking. The combined long track efficiency can be seen in Fig. 5e. Ghost (fake) tracks are defined as tracks with less than 70% of hits from the same tracks. The ghost fraction is shown in Fig. 5f.

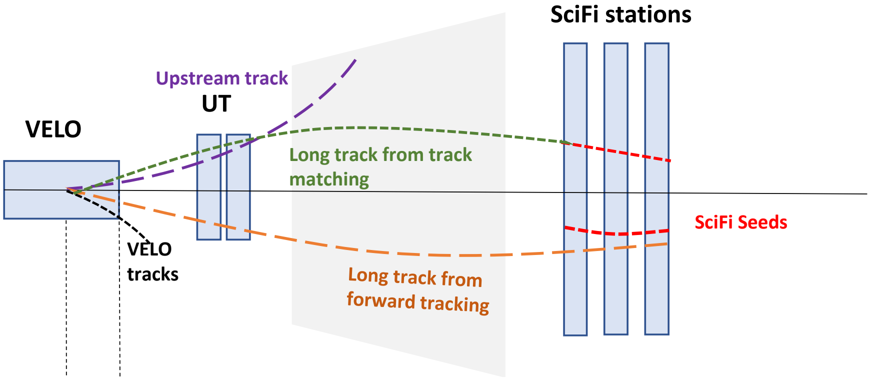


Figure 4: Schematic showing the different types of tracks in LHCb [3].

4. Tracking without the UT

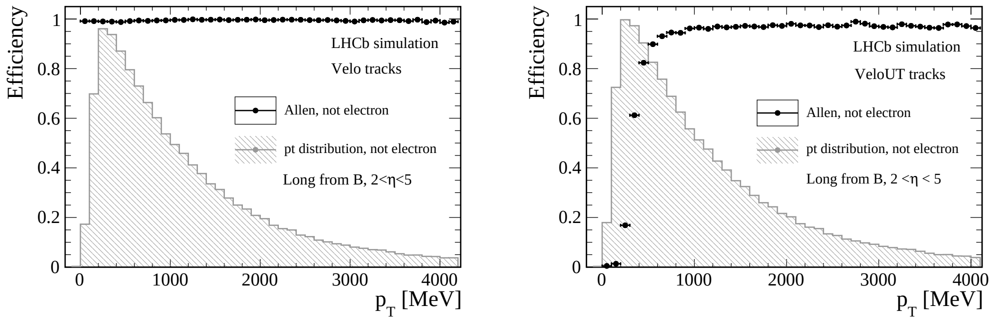
Unfortunately, the UT will not be installed until the end of the year. Therefore, the tracking procedure must be adapted to work without momentum and charge information. The nominal forward tracking technique is not possible without the UT, so a new method was developed. It extrapolates each VELO track as a straight line and makes two windows, one assuming a positive charge and the other assuming a negative charge. This trick allowed the tracking efficiency and throughput to be maintained at the baseline design level. The tradeoff is an increase in the ghost rate. Both the efficiency and ghost rate are shown in Fig. 6.

5. PID Calibration

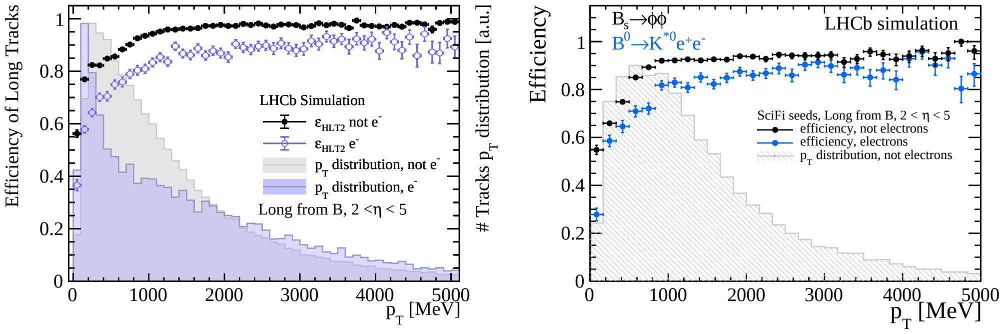
Another essential prerequisite of the LHCb physics program is particle identification (PID). The PID algorithms combine information from the RICH, calorimeters, and the muon system to determine the particle type. PID requirements are used extensively in the trigger. The importance of the PID at LHCb is illustrated by the plots in Fig. 7. Simulation of the PID response is notoriously difficult, so LHCb instead relies on data-driven techniques. The procedure exploits self-tagging channels, such as $D^{*+} \to [K^-\pi^+]_{D^0}\pi^+_{\text{tag}}$ and $D^{*-} \to [K^+\pi^-]_{\overline{D}^0}\pi^-_{\text{tag}}$, where the charge of π_{tag} is used to determine the particle type of the positive and negative probe tracks.

The procedure starts with selecting events without PID information using dedicated trigger lines. The selected events serve as a proxy for the studied signal. Subsequently, an *sPlot* technique is used to obtain pure samples to which PID cuts tailored for a particular analysis are applied. The events are then binned in p_{T}, η, and n_{tracks} — the PID response is non-uniform in these variables. A PID-cut efficiency is then assigned to each track using a look-up in the calibration histogram, effectively reweighting the proxy and signal datasets to match.

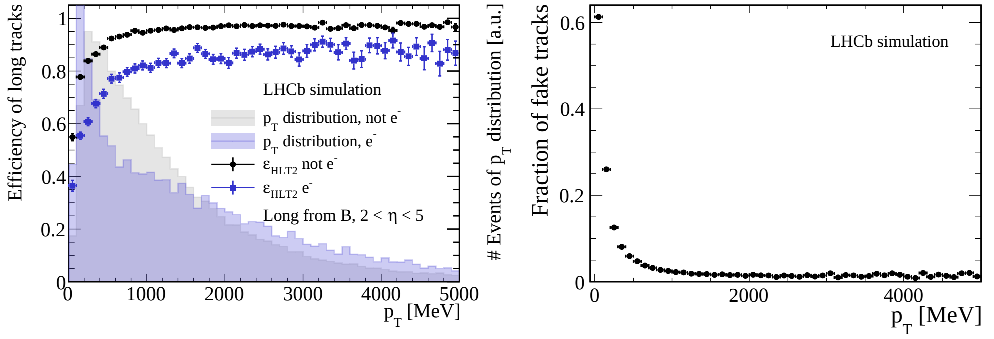
In Run 2, we used a global *sPlot* fit to obtain the pure signal. However, the mass resolution in those fits slightly depends on the decay kinematics. We want to increase the precision of our approach for Run 3, so we conduct the *sPlot* fits in kinematic bins. To this end, we developed a new Python-based fitting framework built on JAX. It includes an automatic pipeline that creates tuples from the output of the HLT2 PID calibration trigger lines, runs the *sPlot* fits, produces the



(a) VELO tracks **(b)** Upstream tracks



(c) Forward long tracks **(d)** SciFi seeds



(e) Combined long tracks **(f)** Combined ghost fraction

Figure 5: Track reconstruction efficiency (a)–(f) and ghost fraction (f) versus transverse momentum [4–7].

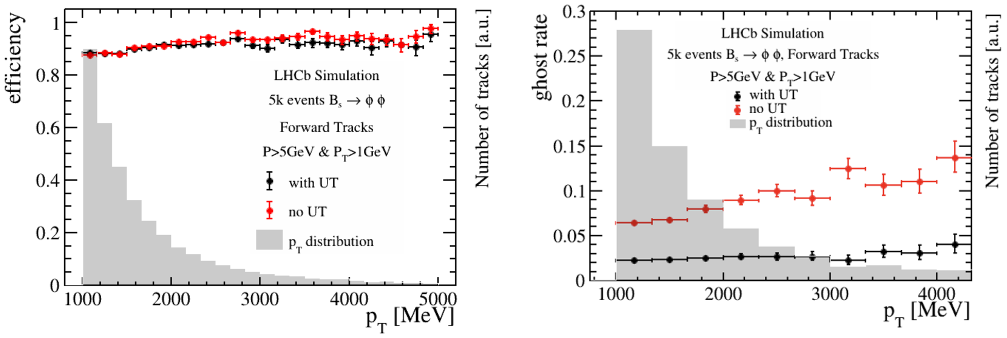


Figure 6: HLT1 tracking efficiency (left) and ghost rate (right) with and without the UT [8].

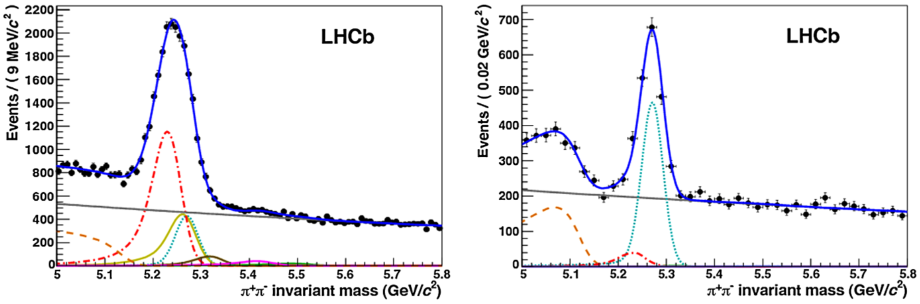


Figure 7: Invariant mass distribution of $B^0 \to \pi^+\pi^-$ candidates with (right) and without (left) PID requirements [9]. The solid turquoise line corresponds to the signal, while the dash-dotted red line shows the misidentified $B^0 \to K^+\pi^-$ decays.

calibration tables, and takes care of the bookkeeping. We also have new tools that automatically process and assign PID efficiency to analysts' samples using chosen PID cuts.

Ackwnowledgements

The authors would like to thank the LHCb computing and simulation teams for their support and for producing the simulated LHCb samples used in the paper.

References

[1] A. Piucci, *The LHCb Upgrade*, *J. Phys. Conf. Ser.* **878** (2017) 012012.

[2] LHCb collaboration, *RTA and DPA dataflow diagrams for Run 1, Run 2, and the upgraded LHCb detector* , `https://cds.cern.ch/record/2730181`.

[3] P. Li, E. Rodrigues and S. Stahl, *Tracking Definitions and Conventions for Run 3 and Beyond*, `https://cds.cern.ch/record/2752971`.

[4] LHCb collaboration, *Performance of the GPU HLT1 (Allen)*, `https://cds.cern.ch/record/2722327`.

[5] LHCb collaboration, *HLT2 reconstruction throughput and Forward Tracking performance for Run 3 of LHCb*, `https://cds.cern.ch/record/2810226`.

[6] LHCb collaboration, *Standalone track reconstruction and matching algorithms for the GPU-based High Level Trigger at LHCb*, `https://cds.cern.ch/record/2811214`.

[7] LHCb collaboration, *Selected HLT2 reconstruction performance for the LHCb upgrade*, `https://cds.cern.ch/record/2773174`.

[8] LHCb collaboration, *HLT1 forward tracking performance*, `https://cds.cern.ch/record/2810803`.

[9] LHCb RICH collaboration, *Performance of the LHCb RICH detector at the LHC*, *The European Physical Journal C* **73** (2013) 2431.