Real-time analysis from the trigger candidates and novel calibration strategy at the LHCb experiment

Giulio Gazzoni[1] on behalf of the LHCb collaboration

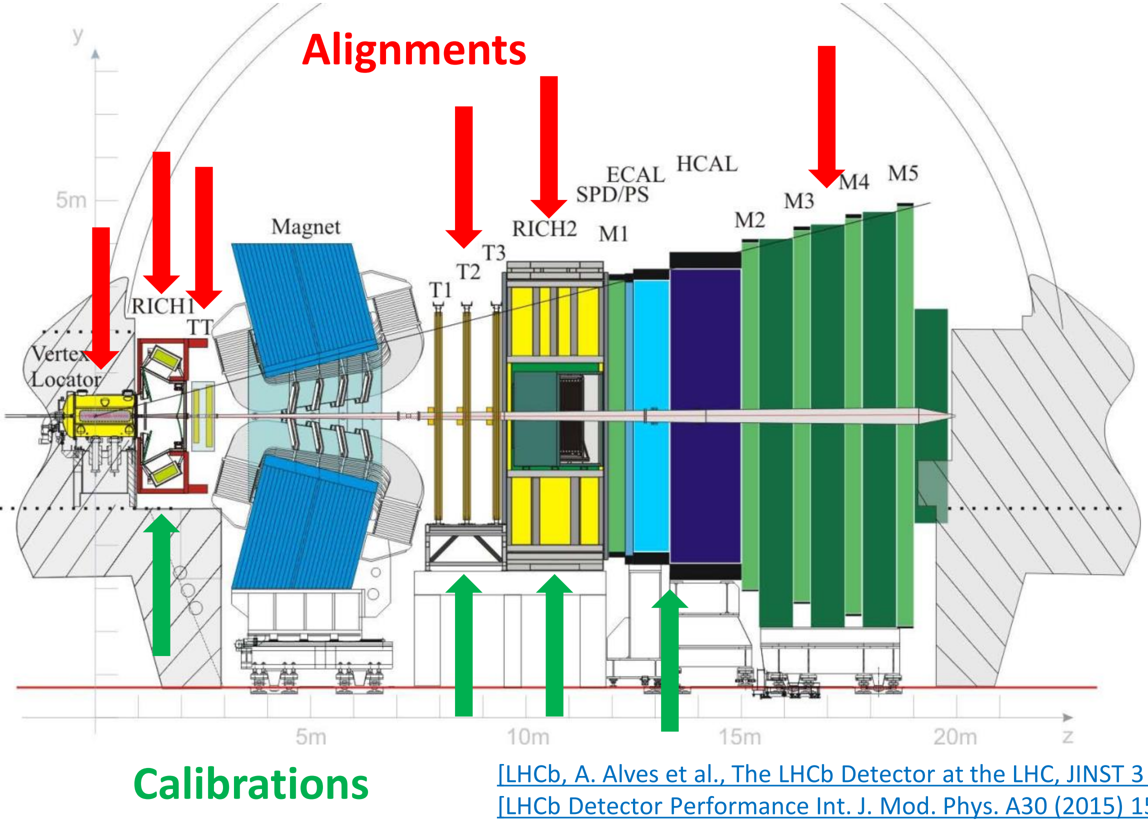
[1]Université Clermont-Auvergne, LPC Clermont

EPS-HEP 2017

Venice, 6 July 2017

The LHCb detector

- The LHCb detector is a single-arm spectrometer with a forward geometry dedicated to beauty and charm physics and designed to cover a unique pseudorapidity range, i.e. $2 < \eta < 5$.



- **Tracking System:** VELO, Tracker Turicensis (TT), Tracking Stations (T1-3) composed by Inner and Outer tracker.

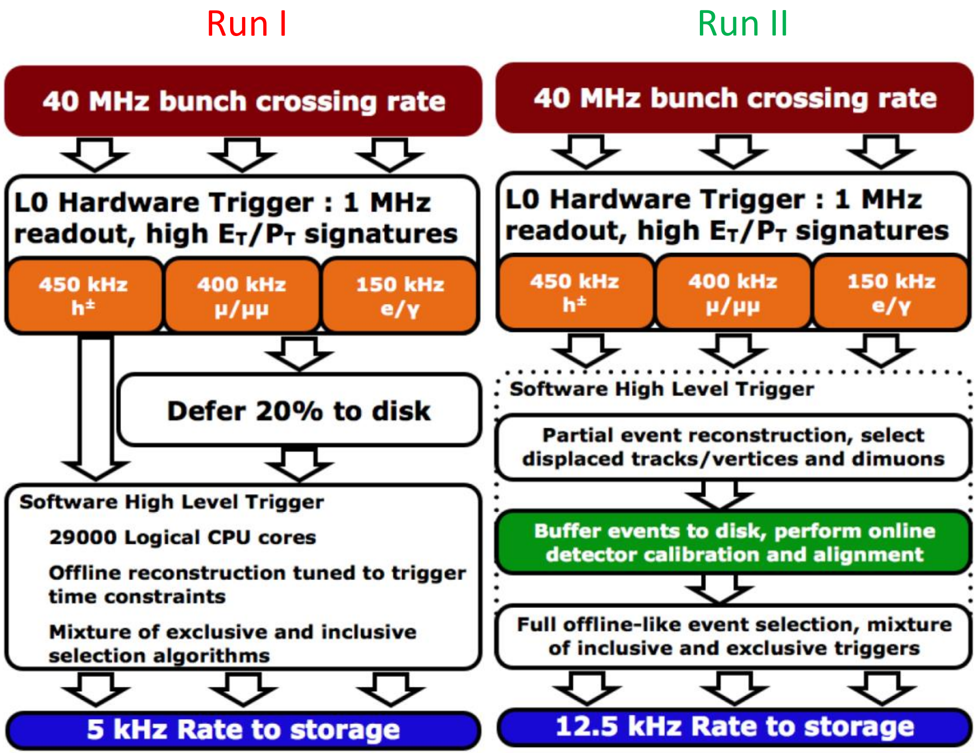
- **Particle Identification System:** RICH1, RICH2, SPD/PS, ECAL, HCAL and Muon Stations (M1-M5).

- LHCb has excellent Tracking and PID performances.

[LHCb, A. Alves et al., The LHCb Detector at the LHC, JINST 3 (2008) S08005]
[LHCb Detector Performance Int. J. Mod. Phys. A30 (2015) 1530022]

Evolution of the LHCb trigger

- During the first long shutdown the LHCb High Level Trigger (HLT) farm was almost doubled and the HLT software was improved. This allowed to deeply revisit the trigger strategy for Run 2.



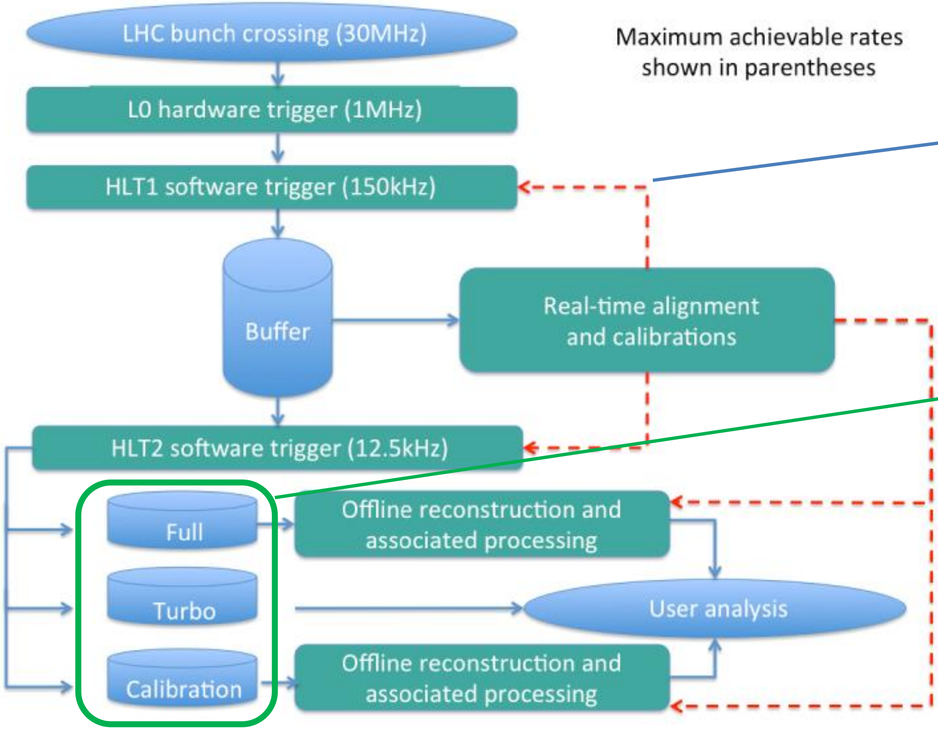
Run I:

- After the L0 trigger the HLT software is split in two stages.
- HLT1: partial event reconstruction.
- HLT2: full simplified event reconstruction and preliminary detector alignment and calibration.
- Marginal use of PID information and partial buffer (2012) after L0 to allow processing between LHC fills.

Run II:

- L0 trigger: the same as in Run I.
- All events passing the HLT1 are buffered to disk to perform a **full real-time alignment and calibration** of PID detectors and tracking system.
- Events out the HLT2 stage have the same quality of the full offline reconstructed events.

Data processing model in Run II



[arXiv: 1604.05596v1]

- Dedicated samples from the HLT1 are taken as input to align and calibrate the detector.
- The **procedure is fully automated** and performed at regular interval depending on needs.
- The calibrations are implemented before HLT2 to guarantee the best quality event selection.

The LHCb workflows:

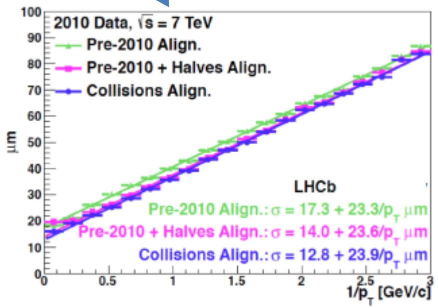
- **Full stream**: full event reconstruction and analysis pre-selections.
- **Turbo stream**: new approach that saves only candidate information.
- **Calibration stream**: dedicated to reference channels used for calibration purposes.

Alignment and calibration importance

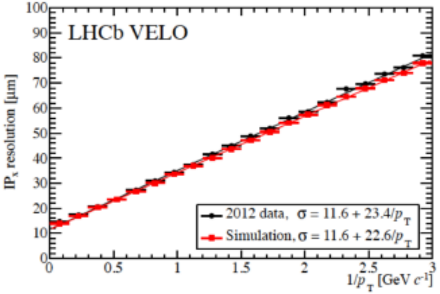
Physics performances crucially rely on spatial alignment and detector calibration.

- VELO alignment: better PV discrimination, IP and decay-time resolution.
- Tracking alignment: better $\delta p/p$ and mass resolution
- RICH calibration: PID allows more exclusive selections and clean signal peaks

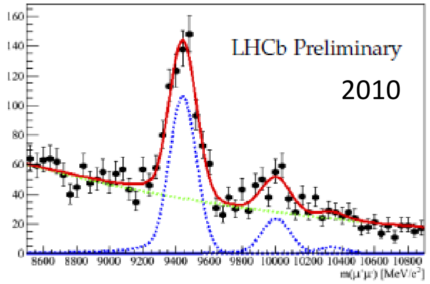
[2014 JINST 9 P09007]



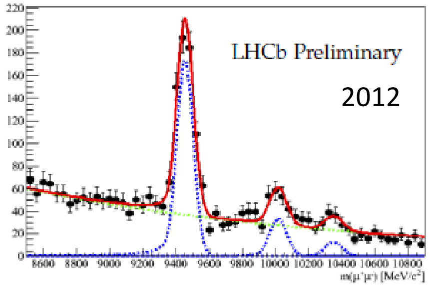
First: σ_{IP} (high p_T) = 14.0 μm



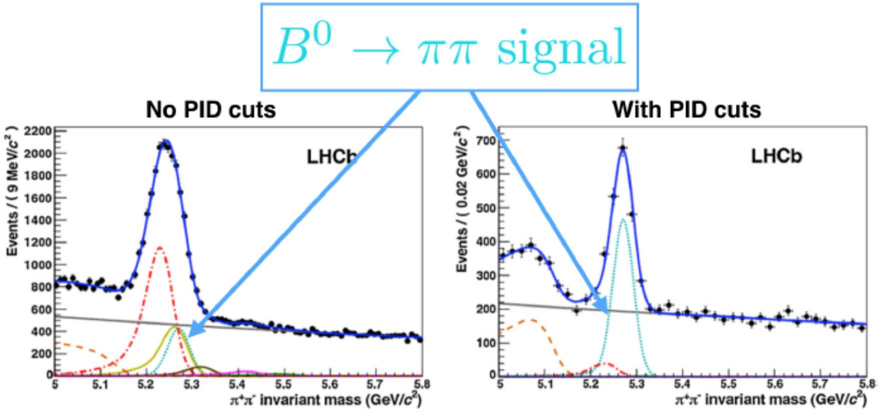
Latest: σ_{IP} (high p_T) = 11.6 μm



First alignment: σ_Υ = 92 MeV/c^2



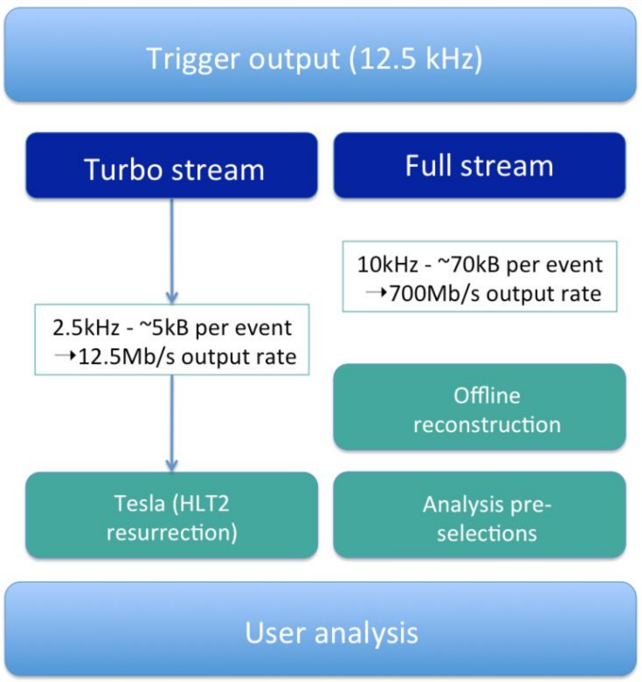
Latest alignment σ_Υ = 49 MeV/c^2



Invariant mass distribution for $B^0 \to \pi\pi$ decay
($B^0 \to \pi\pi$, $B^0 \to K\pi$, $B^0 \to$ 3-bodies, $B_s \to K\pi$, $\Lambda_b \to pK$, $\Lambda_b \to p\pi$)

[Eur. Phys. J. C (2013) 73: 2431]

The LHCb Turbo Stream



[The LHCb Turbo Stream, 2015 J. Phys.: Conf. Ser. 664 082004]

- Part of the physics program needs billions of recorded candidates (e.g. charm analyses) but with no need for the rest of the event.
- A novel approach that exploits the LHCb real-time alignment capabilities and 20% of the HLT2 output bandwidth (2.5 kHz) was implemented for Run 2 data taking.

TURBO STREAM

- Candidates are saved directly after the trigger without any further reconstruction and raw events are removed.
- 10kHz of the HLT2 output undergo an additional offline reconstruction using the sub-detectors raw data banks.

- The turbo stream has many advantages:

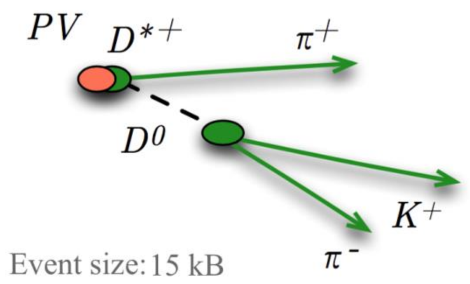
Great disk space saving: ~5kB vs ~70kB

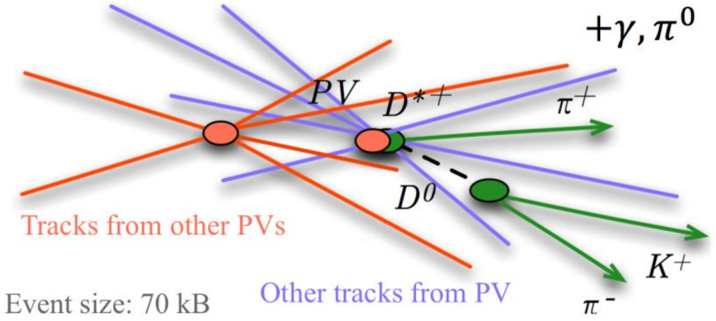
Much higher speed: <1h vs ~30h

Fast data availability for physics analyses

Turbo Stream in 2015 and 2016

2015 - Turbo





2016 – Turbo++

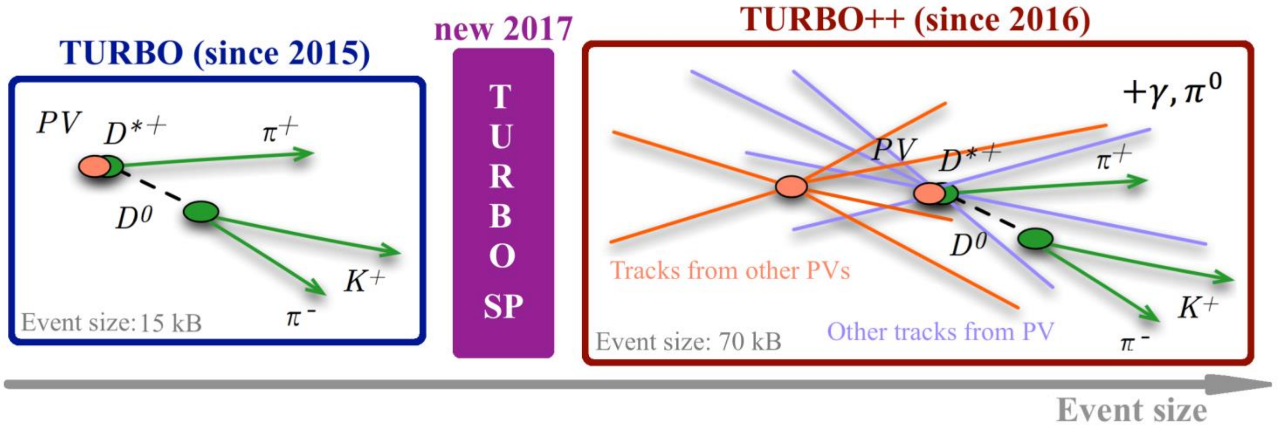
- Saves only objects selected by the trigger
- Output limited to a reduced set of variables

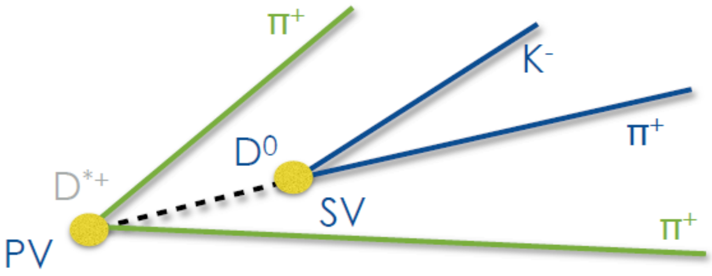
- Can save trigger candidates plus any reconstructed objects
- Allows to save arbitrary variables like isolation, hits in a cone region

- The goal of 2016 implementation is, according to the physics channel and desired measurement, to choose how much (and which variables) of the event has to be saved. By this way more full analyses can be performed using the trigger output (e.g. charm spectroscopy, flavour tagging).

Out of the 420 HLT2 lines in the 2016 physics program, 150 choose Turbo, ~60 more lines compared to 2015.

Turbo Stream in 2017 - Turbo SP





- New intermediate solution for Turbo in 2017: Turbo SP (Selective Persistence) that allows to save candidates and a subset of the reconstruction.
- Allows to choose particles nearby the PV for further analysis.
- Utility to help selecting interesting particles.

Calibration samples

- In order to perform a complete physics analysis we need to evaluate PID and Tracking efficiencies with a data-driven approach using dedicated high statistics calibration samples. This step is needed since the simulation does not reproduce these variables with sufficient accuracy.

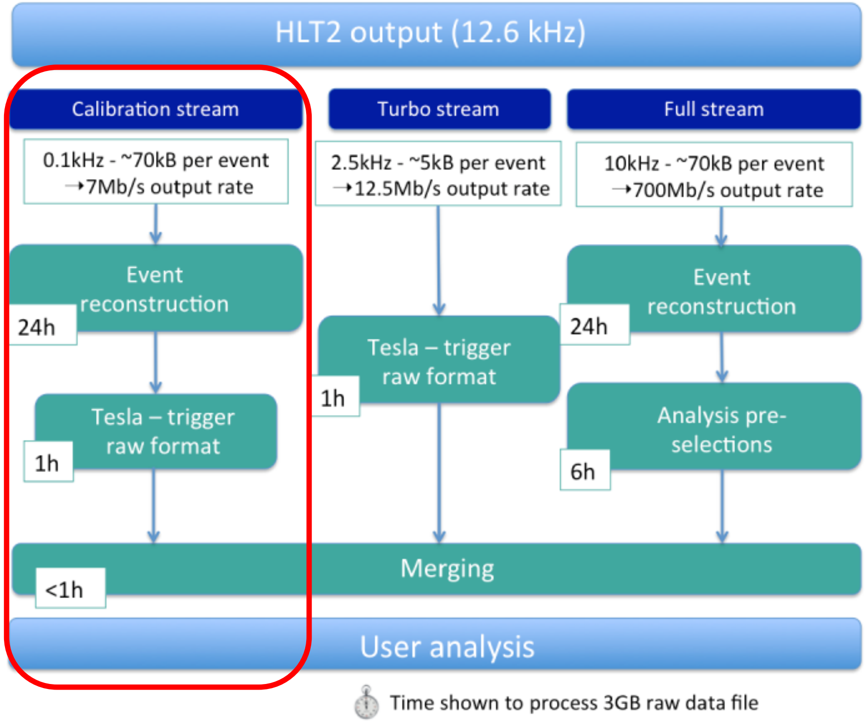
- During the Run 1 the calibration samples were produced with offline selections, giving rise in some cases to lack of statistics and poor coverage in some phase-space regions.

- It is crucial to have larger statistics calibration samples in order to:
  - have smaller statistical uncertainty
  - perform extensive systematic studies, including those with detector low-level information.

- Larger samples could also be used for development of new algorithms in view of the LHCb upgrade.

- In Run 2 the production strategy radically changed exploiting the Turbo Stream implementation.

TurboCalib - the PID paradigm



[LHCb-PUB-2016-020]

- Thanks to the novel real-time alignment and calibration procedure we can exploit the high quality selected candidates as out of the HLT2.

- Calibration samples selected in this way are saved in a dedicated stream called **TurboCalib**, allowing the calculation of PID related variables for both online (Turbo) and offline (Full) streams.

- The stream is then processed centrally to perform the matching of online and offline candidates.

- With a further processing step the matched candidates are used to produce background subtracted samples that will be exploited by the final user in order to extract the real efficiency for a given PID cut. The efficiencies can be evaluated using official LHCb software (PIDCalib package).

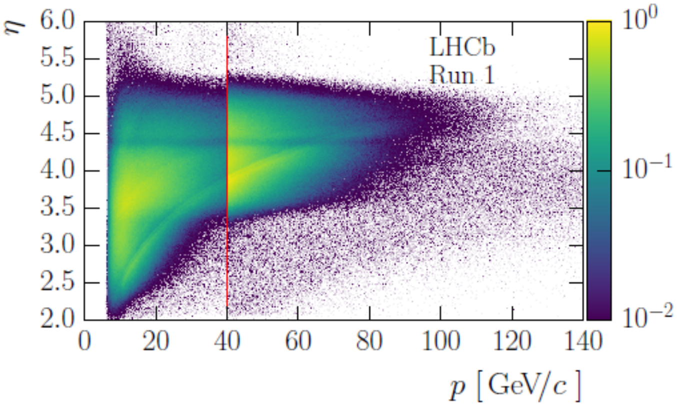
PID calibration samples

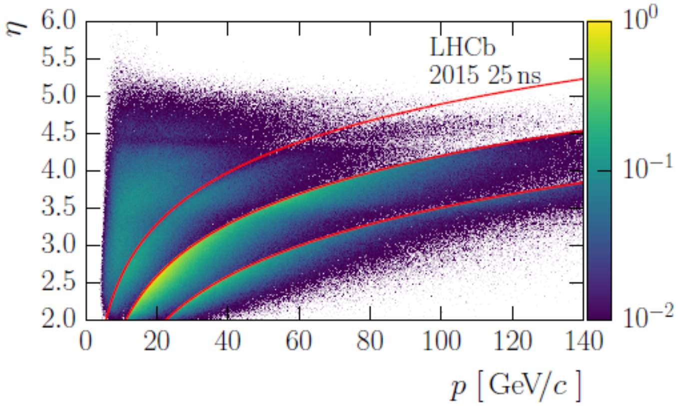
- Run 2 calibration samples contain hundreds of millions of events of high purity charged-ID candidates.

Species	Soft p_T	Hard p_T
$e^\pm$	-	$B^+ \to [e^+e^-]_{J/\psi}K^+$
$\mu^\pm$	$D_s^+ \to [\mu^+\mu^-]_\phi \pi^+$	$J/\psi \to \mu^+\mu^-, B^+ \to [\mu^+\mu^-]_{J/\psi}K^+$
$\pi^\pm$	$K_s \to \pi^+\pi^-$	$D^{*+} \to [K^-\pi^+]_{D^0}\pi^+$
$K^\pm$	$D_s^+ \to [K^+K^-]_\phi \pi^+$	$D^{*+} \to [K^-\pi^+]_{D^0}\pi^+$
$p^\pm$	$\Lambda^0 \to p^+\pi^-$	$\Lambda^0 \to p^+\pi^-$, $\Lambda_c^+ \to p^+K^-\pi^+$, $\Lambda_b^0 \to [p^+K^-\pi^+]_{\Lambda_c^+}\mu^-\nu$
π^0	-	$D^{*+} \to [K^-\pi^+\pi^0]_{D^0}\pi^+$
γ	$D_s^{*+} \to [K^+K^-\pi^+]_{D_s^+}\gamma$, $D_{(s)}^+ \to [\rho^0\gamma]_{\eta'}\pi^+$	$B^0 \to [K^+\pi^-]_{K^*}\gamma$, $B_s^0 \to [K^+K^-]_\phi\gamma$

[LHCb-PUB-2016-005]

Proton sample in Run 1



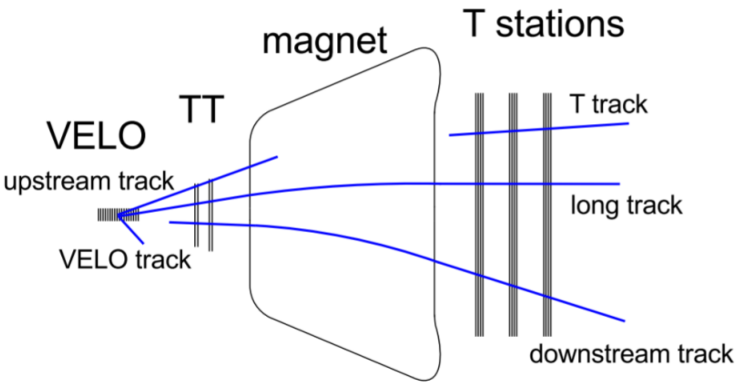


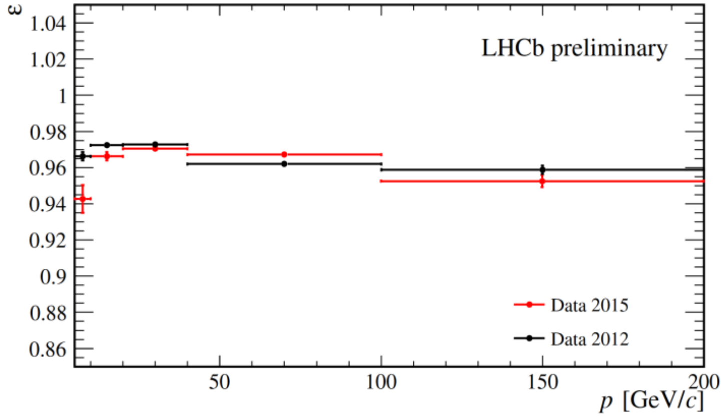
Proton sample in Run 2

- In 2017 the same production strategy has also been extended to neutral objects.

Track reconstruction efficiency

- Data-driven method to measure the tracking reconstruction efficiency since the MC simulation is not accurate enough.
- The efficiency is measured using the Tag-and-Probe method with the $J/\psi \rightarrow \mu^+ \mu^-$ decays. Long track + probe track which is only partially reconstructed, not using information from at least one sub-detector.
- The TurboCalib stream developed for the PID calibration samples, has also been used to produce the Tracking calibration samples.
- A dedicated LHCb software analogous to the PIDCalib package is available to determine the real tracking efficiency (TrackCalib).
- See Agnieszka's talk for more details.





Conclusions

- The LHCb computing model significantly evolved during the first long shutdown and the high level trigger underwent important updates. This allowed to deeply revisit the LHCb trigger strategy for Run II and to implement a fully automated real-time alignment and calibration procedure.
- A novel approach that exploits such operation was implemented. This new concept, called Turbo stream, allows to save candidates directly after the HLT2 stage without any further reconstruction and with offline quality level.
- In 2017 a more flexible variation of this paradigm has been developed. This extension, the Turbo SP (Selective Persistence), allows to save candidates and a subset of the reconstruction with user-selectable granularity.
- The production strategy of PID and Tracking calibration samples changed thanks to the TurboCalib stream implementation. This allowed to speed up the process and to guarantee higher statistics samples.
- In 2017 the same approach has been extended to provide analogous calibration samples also for neutral objects, *i.e.* π^0 and γ.

Backup slides

Data-MC discrepancies for tracking efficiency in Run II

The charge collected in the LHCb VELO sensors is affected by radiation damage. One such effect, which is more pronounced in the outer regions of downstream sensors, arises from charge induction on second metal layer routing lines [2]. Prior to the start of Run 2, modifications were made to the digitization step in the LHCb simulation framework to model this effect. An error was made in the parametric implementation resulting in a reduction of the track reconstruction efficiency in simulation compared to data for tracks with low pseudorapidity. The tracking efficiency calibration procedure that was applied in this paper to the data and simulation [3] was unable to correct the mismodelling.