PROCEEDINGS OF SCIENCE

Real-time physics, alignment, and reconstruction in the LHCb trigger

Alex Pearce[*], **on behalf of the LHCb collaboration**
CERN, Switzerland
E-mail: alex.pearce@cern.ch

Since 2015, the LHCb experiment has employed an exclusively-real-time analysis strategy for a large fraction of its physics programme. Full physics analyses are performed directly on the objects reconstructed in the final stage of the software trigger, negating the need for subsequent offline reconstruction and reducing the output event size, without a loss of performance. In mid-2017, an extension of the associated persistency model was made to allow a completely flexible set of physics objects to be saved for subsequent study, greatly increasing the potential for speculative analysis and data mining. This model and its recent extension are motivated and described, as are the real-time alignment and calibration techniques that permit the strategy to provide offline-equivalent performance.

Sixth Annual Conference on Large Hadron Collider Physics (LHCP2018)
4–9 June 2018
Bologna, Italy

PoS(LHCP2018)226

---

[*] Speaker.

© Copyright owned by the author(s) under the terms of the Creative Commons
Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0).

1. Introduction

The LHCb experiment [1] is nominally dedicated to the study of production and decay properties of heavy flavour hadrons, those containing charm or beauty quarks. However, owing to the excellent performance of the LHCb detector [2] and the flexibility of the two-stage software trigger [3], the collaboration has expanded its physics programme to be truly general, ranging from Higgs and electroweak measurements [4, 5], to studies of central exclusive production [6] and heavy ion collisions [7].

In a typical high-energy physics (HEP) experiment, the trigger system decides whether to accept or reject a given event for permanent storage offline. This must be done in 'real time', as computational resources in the online system only allow for a handful of consecutive events to be buffered. In a multi-stage system, each trigger stage filters events such that subsequent stages have more available processing time, however compromises must still be made due to timing constraints. For example:

1. The reconstruction of physics information is simplified in some way or omitted, to allow for faster event evaluation; and

2. The alignment and calibration constants available in the trigger may not reflect the current state of the detector, as there is not sufficient time to re-evaluate them.

Because the physics performance in a traditional trigger is not optimal, the full raw detector information is persisted for a triggered event, with the trigger reconstruction being discarded. An improved reconstruction is then run offline, and may be re-run several times as new alignment and calibration constants become available. It is then highly desirable to define selections of physics objects in the trigger which are *looser* than those applied offline, otherwise the analyst must deal with the convolution of online and offline resolution effects, complicating the analysis and often increasing systematic uncertainties. The costs of looser selections are lower signal purities and higher trigger rates, and hence a higher trigger output bandwidth requiring more offline resources.

Since 2015, LHCb has employed a data processing model that allows for full physics analyses to be performed in the final stage of the software trigger without a loss of physics performance. This has been facilitated by both a real-time alignment and calibration of the detector [8] and the porting of the offline reconstruction into the trigger, such that the online reconstruction is identical to that performed offline. In turn, this enables the storage of a reduced event format in which only the subset of the event relevant for offline analysis is kept, discarding the raw information, which reduces the trigger output bandwidth by up to a factor 20 for a given trigger rate [9].

This contribution continues in Section 2 by motivating the need for an offline-quality real-time analysis strategy in LHCb in the context of Run 2 and Run 3. The implementation of the online alignment and calibration is described in Section 3, followed by a description of the reduced event format in Section 4, which saw a substantial evolution in 2017 that allowed for additional bandwidth savings. Prospects for the future are outlined in Section 4.1, and a summary is given in Section 5.

2. Real-time analysis

A coarser trigger reconstruction with respect to that offline is ideally balanced by looser trigger selections. This comes at the cost of increased trigger rate, and this is problematic for high-rate signal processes, such as charm and beauty production, as well as for suppressed signal processes as tight selections are required to reduce the trigger rate to manageable levels.

During Run 2 of the Large Hadron Collider (LHC), in which the proton-proton centre-of-mass energy is $\sqrt{s} = 13\,\text{TeV}$, the instantaneous luminosity delivered to the LHCb experiment is $4 \times 10^{32}\,\text{cm}^{-2}\,\text{s}^{-1}$. Using the measured $pp \to c\bar{c}X$ and $pp \to b\bar{b}X$ cross-sections [10, 11] and an average raw event size of 60 kB, the total charm and beauty 'bandwidth' within the detector acceptance is $60\,\text{GB}\,\text{s}^{-1}$ and $2.5\,\text{GB}\,\text{s}^{-1}$, respectively. Clearly the full heavy flavour rate cannot be selected with 100 % efficiency, however it also cannot be selected with 100 % purity, and so it is still the case that the bandwidth is overwhelming, given that offline computational resources only permit an output bandwidth of $700\,\text{MB}\,\text{s}^{-1}$.

For Run 3 of the LHC, due to commence in 2021, the instantaneous luminosity will increase by a factor 5. The LHCb detector as well as the data acquisition and online systems will undergo a near-complete upgrade to cope with the resulting increase in particle densities and data rates [12]. Naively scaling the heavy flavour bandwidths previously discussed by a factor 5, neglecting an increased raw event size and any possible improvements in the trigger efficiency, the beauty bandwidth alone is then sufficient to saturate the available resources.

Several complementary solutions exist to reduce trigger output bandwidth: tighten the trigger selections significantly, resulting in a large loss of signal efficiency; reduce the scope of the physics programme; or save a reduced event format. The latter option can be implemented by storing high-level reconstructed information computed in trigger, discarding the raw event, and comes with its own costs: the trigger reconstruction is of poorer quality than that performed offline, reducing the resolution of physics observables; and the exact subset of the reconstruction to be persisted must be defined upfront, and this is often not known at the time of the trigger decision (such as for inclusive triggers, which typically constitute the majority of an experiment's trigger rate).

Both the ATLAS and CMS collaboration currently employ a reduced event format for the output of triggers for low-mass dijets [13, 14], which are produced at a very high rate yet serve as important input for searches for dijet resonances. However, these approaches do not make use of the standard offline reconstruction at the trigger level and require special tools to analyse.

To overcome the usual downsides of persisting trigger objects exclusively, the LHCb experiment invested resources that have allowed an offline-equivalent reconstruction to be run in their final software trigger stage. A key ingredient to achieving this remarkable result is in the implementation of a real-time alignment and calibration of the full detector before the final trigger is executed, discussed in the following Section.

3. Online alignment and calibration

The creation of objects in software or hardware that represent physical particles present in an event is the process of reconstruction. In order to correctly infer the properties of reconstructed objects, the geometry and responses of the various sub-detectors are required. This allows for hits

PoS(LHCP2018)226

to be accurately related in space for tracking, and for electrical signals to be converted into physical values. Offsets can be applied to correct for deviations in a set of base position and response models to what the detector actually looked like at the moment the data were taking. The *alignment* and *calibration* of the detector are the processes by which these offsets are computed.

Alignment and calibration are most commonly performed offline using well-known control processes, and this may be repeated as knowledge of the detector and of calibration techniques improves. New constants require a reprocessing of the raw data such that a new reconstruction can use them. Analysts must then move to using the new data. The constants used online, in the trigger system, usually represent some past state of the detector, and hence can be sub-optimal.

3.1 The disk buffer

During LHC Run 1, the commodity disks included in the HLT farm servers were used to provide a buffer between L0 and HLT1, allowing an average increase in HLT1 input of 20 %, as the trigger could be run during LHC downtime. In between LHC Run 1 and Run 2 (2013–2015), the disk capacity was increased to 10 PB to be used in between the first and second stages of the LHCb software trigger, giving the dataflow presented in Figure 1. This permits around 11 days of HLT1 output to be buffered and increases the effective CPU resources available in HLT2 as it can run whilst the LHC is not delivering stable beams.

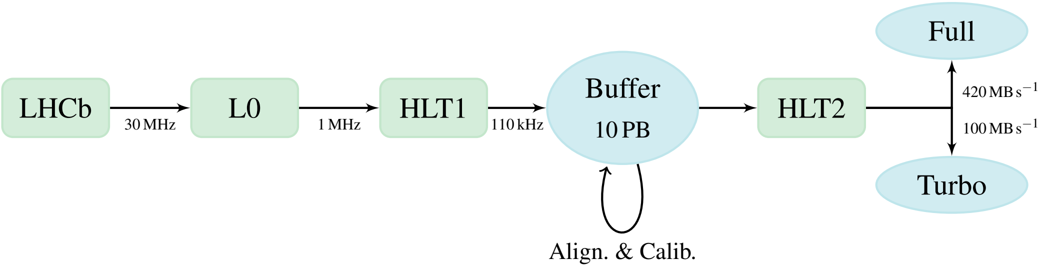

Figure 1: The flow of events in the LHCb online system in Run 2. Events reaching terminal nodes are sent to permanent offline storage.

The addition of the disk buffer buys enough time for the alignment and calibration procedures to be run on the data that have just been taken before they must be processed by HLT2. Along with the experience of calibrating and aligning the detector gained in Run 1, this has allowed an offline-quality alignment and calibration to be implemented in the online system, such that HLT2 runs with as accurate and precise constants as could be achieved offline. The alignment and calibration algorithms run across the whole trigger computing farm using dedicated data streams from HLT1 as input, with each algorithm being executed as soon as a sufficiently large dataset has been collected in the current LHC fill. The constants are updated if the latest ones differ significantly. The additional computing time available in HLT2 permitted the porting of the offline reconstruction in to the final trigger stage, such that the HLT2 reconstruction is now identical to that run offline.

3.2 Sub-detector workflows

The technical implementation of the alignment and calibration tasks within the online system

PoS(LHCP2018)226

are detailed in Ref. [8]. Two example tasks are motivated and described here.

The LHCb vertex locator (VELO) is a silicon-strip detector surrounding the proton-proton interaction region, with the closest active elements around 8 mm from the beam spot. As the detector is so sensitive, it is comprised of two independent halves that must be retracted outside of stable beams. The halves are centred around the beam spot at the beginning of each fill, and so the VELO alignment can in principle change with the same frequency. The alignment is performed by minimising the residuals of tracks created with a Kalman fit that uses primary vertex positions as input constraints. The stability of this procedure over six-week period is shown in Figure 2.

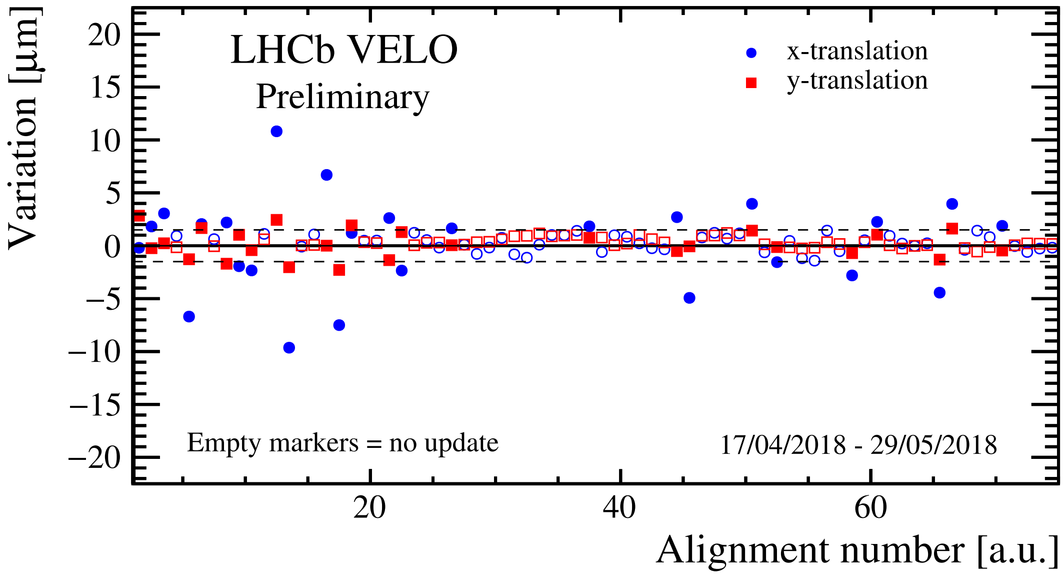

Figure 2: The deviation of two VELO alignment parameters from their previous values as a function of time. Filled points represent significant deviations that resulted in updates of the constants used in HLT2.

Two ring-imaging Cherenkov detectors (RICH1 and RICH2) provide particle identification (PID) information for charged tracks that allows discrimination primarily between pion, kaon, and proton hypotheses. Good performance of these sub-detectors is critical to disentangle decays which share topologies but have different final state particles. Mass hypotheses are assigned to tracks through a maximum likelihood procedure. This determines the set of hypotheses for all tracks simultaneously by comparing the expected photon yield for that hypothesis set with what is observed in the RICHs. The refractive index of the gases in the sub-detectors is a function of their temperature and pressure and plays an important role in determining the photon yield. These quantities are monitored using hardware sensors, but this does not provide a precise enough estimate of the refractive indices for physics analysis. Instead, the deviation from expected Cherenkov angle and those measured in the RICHs is measured and used to compute calibration constants that scale the measured indices.

4. Turbo: a flexible reduced event format

With a full, offline-quality reconstruction in HLT2, and an up-to-date alignment and calibration database, trigger objects can now be persisted and analysed directly offline without a loss of

PoS(LHCP2018)226

physics performance [9]. These objects correspond to the same C++ instances created in the offline reconstruction, and so can be analysed using standard tools.

Each trigger line that fires in HLT2 is assigned to one or more output streams which are sent offline. An event in a given output stream contains a subset of all available raw *banks*, each corresponding to the information created by different sub-detectors and by the trigger itself. Events fired by trigger lines assigned to the *full stream* are persisted with the complete set of raw banks. The trigger reconstruction is discarded, and the events are reconstructed again offline; this corresponds to the traditional processing model employed at most experiments. Trigger lines assigned to the *Turbo stream* persist the reconstructed objects used to make the trigger decision. These objects are serialised into a specific raw bank and only this bank is sent offline, along with trigger banks that are used to analyse how the event passed through the trigger system.

PoS(LHCP2018)226

For a typical exclusive decay reconstructed by a Turbo trigger line, such as the $D^0 \to K^-\pi^+$ charm decay, saving only the reconstructed objects rather than the whole raw event reduces the average event size by a factor of 20, allowing for much higher trigger rates within the available computing resources. In addition to persisting the triggering object and the set of all primary vertices, Turbo lines may each request the persistence of additional reconstructed objects in the event. This is available with two levels of granularity, illustrated in Figure 3:

Selective persistence The trigger line can select any number of additional reconstructed objects in a completely flexible manner, such as the set of all tracks which form a good-quality vertex with the trigger candidate, or the set of all reconstructed neutral objects.

Complete persistence The entirety of the HLT2 reconstruction is persisted. This is typically slightly smaller than saving the full raw information.

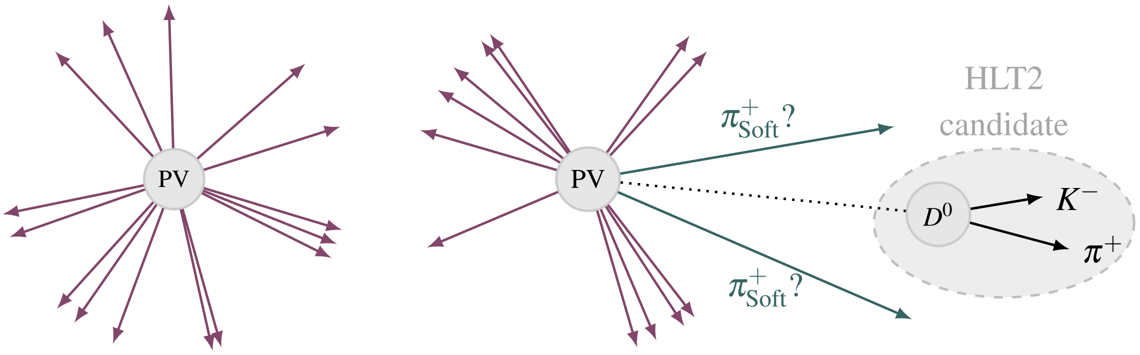

Figure 3: A reconstructed event, in which a Turbo trigger line has fired by selecting a candidate $D^0 \to K^-\pi^+$ decay. The candidate decay and primary vertices are persisted, and then additional objects are persisted depending on the configuration of the trigger line. Objects related to the trigger candidate (in green) may be selectively persisted, or the full reconstruction (in purple) may be chosen instead.

Between the two extremes of candidate-only and complete persistence, this Turbo model can accommodate analyses that rely on exclusive or inclusive triggers. As an example, selective persistence allows for a single exclusive trigger for a ground-state charm decay to persist all objects which might be used for subsequent excited charm spectroscopy offline, without imposing knowledge on exactly which excited states one may want to analyse.

Today, Turbo triggers account for around half the rate of all triggers used for physics analysis, but their output bandwidth is only around a quarter that of the full stream. Physics measurements made possible thanks to the Turbo model include charm and J/ψ cross-sections [15, 16], the discovery of new ground-state and excited charm baryons [17, 18], searches for dark photons [19], the characterisation of charmonium within jets [20]. The dimuon spectrum used in the dark photon search is shown in Figure 4. Without the possibility of using a reduced event format, the triggers used in these analyses would together add too much bandwidth to fit within available computing resources. This widespread exploitation of the real-time analysis model is unique to LHCb.

PoS(LHCP2018)226

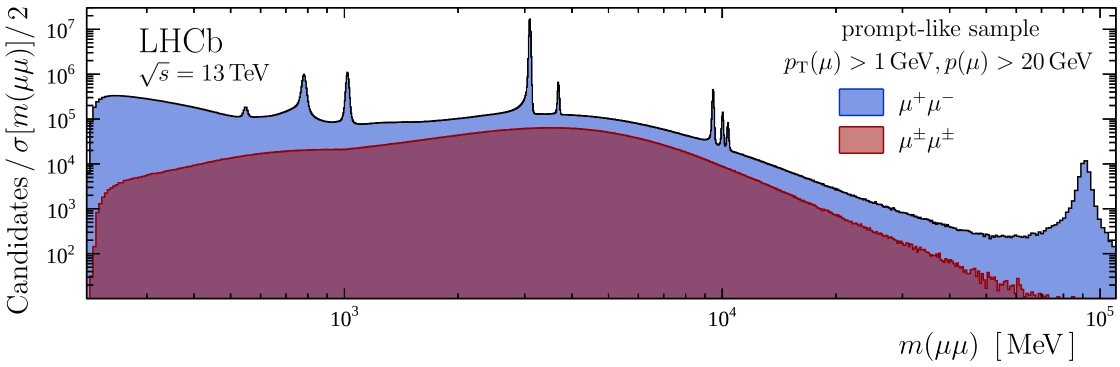

Figure 4: Prompt like-sign (red) and opposite-sign (blue) dimuon mass spectra record by a single Turbo trigger line [19].

4.1 Prospects for Run 3

A major challenge for the LHCb experiment in Run 3 is the factor-five increase in instantaneous luminosity. With today's trigger selections, such an increase would cause the beauty programme to saturate the output bandwidth permissible within Run 3 resources.

The beauty programme is captured primarily by a set of inclusive triggers which select displaced, high-p_{T} vertices characteristic of heavy-flavour decays [21, 22]. By definition, the objects entering the trigger decision do not necessarily constitute the full decay which is required for analysis offline, and so saving only the trigger candidate is insufficient. Today, the full raw event is saved for theses triggers. With selective persistence, such inclusive lines can be moved to the Turbo model, given that a sufficiently broad set of additional objects can be defined. This could be, for example, the set of tracks that have some maximum distance of closest approach to the displaced trigger vertex, and the set of tracks which are associated to the same primary vertex as the trigger candidate.

5. Summary

Thanks to the addition of a large disk buffer installed between the first and second stages of the LHCb software trigger, a broad real-time analysis strategy has been implemented without a loss of physics performance. This includes a real-time alignment and calibration of the data and an offline-quality reconstruction in the final trigger stage. By persisting the set of objects associated to positive trigger decisions, the trigger output bandwidth has been significantly reduced as the raw event can be discarded, allowing for new physics lines to be included. Arbitrary sets of additional persisted objects may be specified on a per-line basis, such that inclusive triggers can be accommodated in the model. In Run 2, these achievements have allowed for the continuation of the charm physics programme, which would otherwise be too high-rate to fit within available computing resources offline. In Run 3, this approach is critical in assuring the success of the beauty programme, which will become similarly bandwidth-constrained.

References

[1] LHCb collaboration, A. A. Alves Jr. et al., *The LHCb detector at the LHC*, *JINST* **3** (2008) S08005.

[2] LHCb collaboration, R. Aaij et al., *LHCb detector performance*, *Int. J. Mod. Phys.* **A30** (2015) 1530022, [`1412.6352`].

[3] R. Aaij et al., *The LHCb trigger and its performance in 2011*, *JINST* **8** (2013) P04022, [`1211.3055`].

[4] LHCb collaboration, R. Aaij et al., *Limits on neutral Higgs boson production in the forward region in pp collisions at* $\sqrt{s} = 7$ *TeV*, *JHEP* **05** (2013) 132, [`1304.2591`].

[5] LHCb collaboration, R. Aaij et al., *Measurement of* $Z \to \tau^+\tau^-$ *production in proton-proton collisions at* $\sqrt{s} = 8$ *TeV*, *Submitted to: JHEP* (2018) , [`1806.05008`].

[6] LHCb collaboration, R. Aaij et al., *Central exclusive production of* J/ψ *and* $\psi(2S)$ *mesons in pp collisions at* $\sqrt{s} = 13$ *TeV*, *Submitted to: JHEP* (2018) , [`1806.04079`].

[7] LHCb collaboration, R. Aaij et al., *Prompt* Λ_c^+ *production in pPb collisions at* $\sqrt{s_{NN}} = 5.02$ *TeV*, `1809.01404`.

[8] G. Dujany and B. Storaci, *Real-time alignment and calibration of the LHCb Detector in Run II*, *J. Phys. Conf. Ser.* **664** (2015) 082010.

[9] R. Aaij et al., *Tesla: an application for real-time data analysis in High Energy Physics*, *Comput. Phys. Commun.* **208** (2016) 35–42, [`1604.05596`].

[10] LHCb collaboration, R. Aaij et al., *Measurements of prompt charm production cross-sections in pp collisions at* $\sqrt{s} = 13$ *TeV*, *JHEP* **03** (2016) 159, [`1510.01707`].

[11] LHCb collaboration, R. Aaij et al., *Measurement of the b-quark production cross-section in 7 and 13 TeV pp collisions*, *Phys. Rev. Lett.* **118** (2017) 052002, [`1612.05140`].

[12] *LHCb Trigger and Online Upgrade Technical Design Report*, Tech. Rep. CERN-LHCC-2014-016. LHCB-TDR-016, May, 2014.

[13] CMS collaboration, *Data Parking and Data Scouting at the CMS Experiment*, Tech. Rep. CMS-DP-2012-022, Sep, 2012.

[14] ATLAS collaboration, *Trigger-object Level Analysis with the ATLAS detector at the Large Hadron Collider: summary and perspectives*, Tech. Rep. ATL-DAQ-PUB-2017-003, CERN, Geneva, Dec, 2017.

PoS(LHCP2018)226

[15] LHCb collaboration, R. Aaij et al., *Measurement of forward J/ψ production cross-sections in pp collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **10** (2015) 172 LHCb-PAPER-2015-037, CERN-PH-EP-2015-222, [`1509.00771`].

[16] LHCb collaboration, R. Aaij et al., *Measurements of prompt charm production cross-sections in pp collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **03** (2016) 159 LHCb-PAPER-2015-041, CERN-PH-EP-2015-272, [`1510.01707`].

[17] LHCb collaboration, R. Aaij et al., *Observation of five new narrow Ω_c states decaying to $\Xi_c^+ K^-$*, *Phys. Rev. Lett.* **118** (2017) 182001 LHCb-PAPER-2017-002, CERN-EP-2017-037, [`1703.04639`].

[18] LHCb collaboration, R. Aaij et al., *Observation of the doubly charmed baryon Ξ_{cc}^{++}*, *Phys. Rev. Lett.* **119** (2017) 112001 LHCb-PAPER-2017-018, CERN-EP-2017-156, [`1707.01621`].

[19] LHCb collaboration, R. Aaij et al., *Search for dark photons produced in 13 TeV pp collisions*, *Phys. Rev. Lett.* **120** (2018) 061801 LHCb-PAPER-2017-038, CERN-EP-2017-248, [`1710.02867`].

[20] LHCb collaboration, R. Aaij et al., *Study of J/ψ production in jets*, *Phys. Rev. Lett.* **118** (2017) 192001 LHCb-PAPER-2016-064, CERN-EP-2017-006, [`1701.05116`].

[21] V. V. Gligorov and M. Williams, *Efficient, reliable and fast high-level triggering using a bonsai boosted decision tree*, *JINST* **8** (2013) P02013, [`1210.6861`].

[22] T. Likhomanenko et al., *LHCb topological trigger reoptimization*, *J. Phys. Conf. Ser.* **664** (Oct, 2015) 082025.

PoS(LHCP2018)226