

# Real-time data analysis model at LHC and connections to other experiments and fields

Arantza Oyanguren<sup>1,2,\*</sup>

<sup>1</sup>Instituto de Física Corpuscular (IFIC), Carrer del Catedratic José Beltrán Martínez, 2, 46980 Paterna, Valencia, Spain.

<sup>2</sup>Copyright 2020 CERN for the benefit of the ATLAS, CMS, and LHCb Collaborations. CC-BY-4.0 license.

**Abstract.** With the upcoming increase of proton-proton collision rates at the Large Hadron Collider (LHC) experiments, and the corresponding increase of data volumes, real-time analysis becomes a key ingredient to be able to analyse and select the interesting data within the available computing resources. In this talk I will review the main features of the techniques followed by the ATLAS, CMS and LHCb experiments. Similar challenges have to be faced in other fields, such as astronomy and cosmology, and I will comment about them.

## 1 Introduction

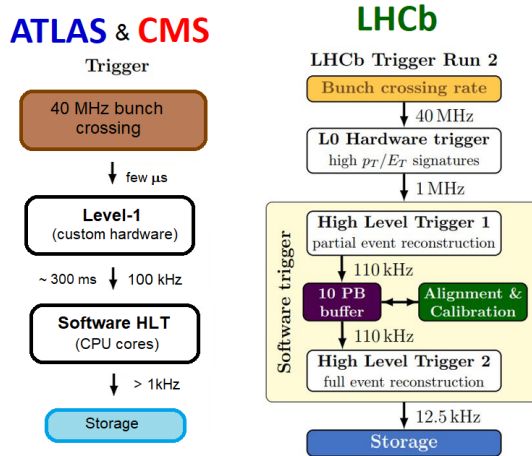
The Large Hadron Collider LHC is a proton-proton circular accelerator placed at CERN<sup>1</sup> in the French-Swiss border, and at present, operating at a maximum energy of 13 TeV. Three of the four main experiments (ATLAS [1], CMS [2], LHCb [3] and ALICE [4]) aim to discover new particles that explain the unknowns of our well-established Standard Model of particle physics, and resolve questions such as what dark matter is, what the particle mass is or where the antimatter is in our Universe. The LHC collides two beams of protons, each carrying 2808 nominal bunches with  $10^{11}$  particles. Bunches are spaced in time by 25 ns so that the bunch collision frequency is 40 MHz. This corresponds to 40 millions collisions per second. Considering that the readout of detectors produces about 1 MB raw data per collision, we have to manage an amount of 40 TB/s and search among them very rare processes such as the  $B_s \rightarrow \mu^+ \mu^-$  decay [5, 6] or decays of the recently discovered Higgs boson [7–10].

With the existing computing resources we can not keep and store all data produced in the LHC collisions, but we have to select the events of interest. For this we define the *trigger* systems, a procedure using simple criteria to rapidly decide which events are recorded for a posterior analysis. Traditional trigger systems [11] have a first stage where an initial decision is made by custom hardware processors (typically FPGAs [Field Programmable Gate Arrays]) using information of the calorimeters and muon stations. Typical selection rates are from 40 MHz to around 100 kHz at ATLAS and CMS, and to 1 MHz at LHCb, within a few  $\mu$ s latency. The ALICE experiment has several specific hardware triggers with different latencies and rates ranging from 500 Hz to 2 KHz. After this process one or more high-level triggers (HLT), which consist of a farm of processors, make fast object reconstruction using

\*e-mail: Arantza.Oyanguren@ific.uv.es

<sup>1</sup>European Organization for Nuclear Research.

mainly the information of tracking detectors. Figure 1 shows the trigger systems for ATLAS [12], CMS [13] and LHCb [14] in the last data taking period (2015-2018). In practice, hundreds of *trigger paths* are coded at the HLT stage, each selecting events for a particular physics signature, based on simple selection criteria such as track momenta, quality of track vertex fits or invariant masses. Trigger output rates after HLT are about 1.5 kHz at ATLAS and CMS and 12.5 kHz at LHCb.



**Figure 1.** Trigger systems at ATLAS, CMS and LHCb experiments during the last data taking period (2015-2018).

The amount of computational resources needed for the storage of the events is determined by the trigger output bandwidth, defined as:

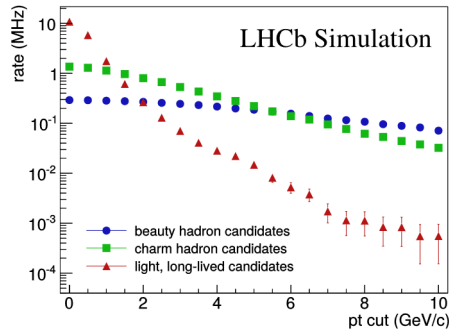
$$\text{bandwidth [GB/s]} \sim \text{trigger rate [kHz]} \times \text{event size [MB]}.$$

With raw data event sizes of 1 MB at ATLAS and CMS, and 0.1 MB at LHCb, that means a trigger bandwidth larger than 1 GB/s, which will be even larger in future data taking periods, since the number of events per bunch crossing will increase. New data processing models have to be developed if we want to maximize the trigger output bandwidth.

## 2 Real-time analysis

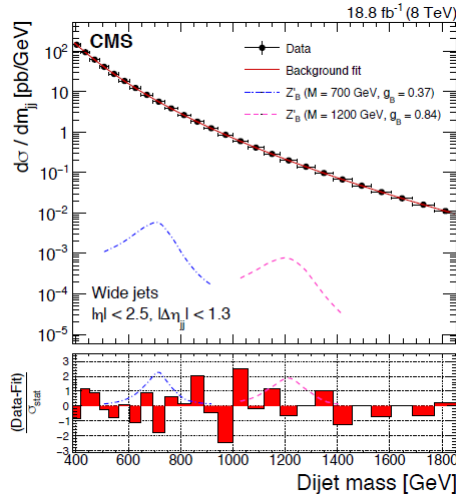
One of the major issues concerning the trigger bandwidth is that the output rates saturate and we cannot reduce it without losing our signals of interest. This can be seen in Figure 2, where high trigger rates for beauty and charm signal candidates are shown at LHCb even if hard selection criteria against background events are applied. The only way to decrease the bandwidth would be to reduce the event size: instead of keeping raw data, store only the relevant information of an event. For this we need to reconstruct and analyse the events to select them in real time.

The CMS experiment was a pioneer in using this technique for the search of narrow resonances using the dijet mass spectrum [15]. Online reconstructed events from HLT with a jet transverse momentum larger than 350 GeV or a dijet mass greater than 400 GeV, were



**Figure 2.** HLT partially reconstructed signal rates as a function of transverse momentum cuts at LHCb [LHCb-PUB-2014-027].

recorded and saved in a reduced format. This *data scouting* concept allowed to take data in the dijet mass region below 1 TeV, that otherwise would have been rejected by the usual trigger filters. The idea is that only some physics objects reconstructed during the HLT processing, such as particle flow (PF) jets<sup>2</sup> are stored for each event. Figure 3 shows the measured differential cross-section and the fit residuals as a function of dijet mass using the *scouting* technique [16] for 18.8 fb<sup>-1</sup>. An updated analysis with more statistics is also available [17].



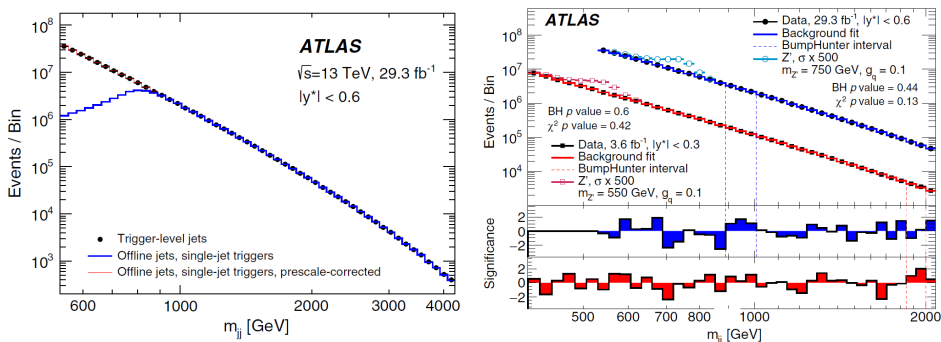
**Figure 3.** Dijet mass spectrum from HLT wide jets (points) compared to a smooth fit (solid) and to predictions of several resonances.

In practice, several *scouting* streams containing a number of HLT trigger paths are implemented. The *Calo-scouting* stream reconstructs jets from calorimeter deposits. It allows one to reduce the cuts on the transverse momenta of the jet ( $H_T$ ) to 250 GeV. The calorimeter

<sup>2</sup>Algorithm based on the concept of global event reconstruction: it performs a combination of the basic elements (i.e. tracks and clusters) obtained from all subdetector systems, in order to identify all particles in the event and measure their properties.

jet, the missing transverse momentum (MET), and a measure of the average energy density in the event,  $\rho$ , are stored. The size of the event content in this stream is about 100 times smaller than the standard raw data format in physics trigger paths. The *PF-scouting* stream uses the online version of the PF sequence to reconstruct and select the events. The  $H_T$  cut is set to 450 GeV, and the event content includes reconstructed PF jets, MET, PF candidates,  $\rho$ , primary vertices, and electron, muon and photon objects. The event size is about 10 KB. Thanks to the *scouting* technique the trigger output rate can be increased by a factor of 10, with a bandwidth 10 times lower than the standard trigger paths, largely improving the search range for dijet resonances. This is a key for the understanding of the Standard Model and for the search of new physics candidates such as the ones which could explain the dark matter in the Universe.

A similar strategy to the CMS *scouting* is applied by the ATLAS experiment, called trigger-object level analysis (TLA) [18, 19]. This approach allows jet events to be recorded at a peak rate of up to twice the total rate of events using the standard triggers, and uses less than 1% of the trigger bandwidth. Jets are reconstructed at the HLT level using information of the energy deposit in the calorimeter cells (topological clusters), and the stored information includes the four-momentum of each jet (sum over the four-momentum of topological clusters), and a set of calorimeter variables related to the jet quality and structure. Figure 4 (right) shows how the number of selected dijets events are increased by using this technique for dijet masses below 0.9 TeV.



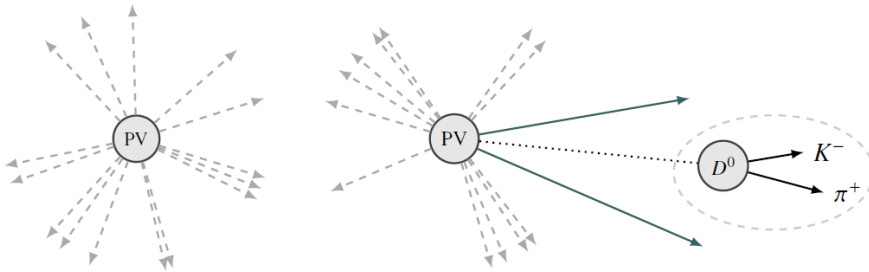
**Figure 4.** Left: comparison of the number of dijet events vs mass selected by normal trigger lines (blue line) and dedicated TLA (black points). Right: dijet mass distribution (black points) with the prediction of several dark-matter  $Z'$  resonances (empty points). Horizontal lines show the largest discrepancy with the QCD background hypothesis (solid lines)

The ATLAS experiment has performed searches of low-mass dijet resonances in the 0.45-1 TeV range using the TLA technique, shown in Figure 4 (right), providing strong limits for dark-matter particles and other new physics models.

The LHCb experiment has also developed a real-time analysis strategy to widen its physics program within the available computing resources. The Turbo<sup>3</sup> model [20] exploits the event topology reconstructed at HLT and stores only a subset of the objects which are relevant for a posterior analysis. As compared to raw data events, the event size for a typical decay channel is reduced by a factor of 20. Several models of persistence are available: Turbo, *selective reconstruction* and *complete reconstruction*. Signal tracks, neutral objects,

<sup>3</sup>Originally it was called Tesla.

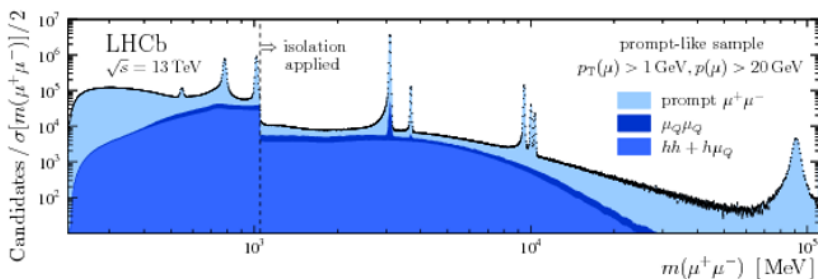
decay vertices and tracking detector clusters associated to the candidate are saved in the Turbo stream, in addition to all of reconstructed primary vertices (PV). In the trigger lines with selective reconstruction persistence additional objects linked to the candidate are also stored, such as new tracks compatible with a decay vertex. All reconstructed objects in the event are stored for the trigger lines with complete reconstruction persistence, allowing for more inclusive triggers, but still keeping rates lower than raw events since the detector hit information is not saved. Figure 5 shows an example of the selective reconstruction persistence in the Turbo model for the  $D^0 \rightarrow K^- \pi^+$  decay channel. The Turbo model is crucial for the charm physics program at LHCb due to the large cross section of  $pp \rightarrow c\bar{c}X$  events. The observation of CP



**Figure 5.** A reconstructed event using the selective reconstruction persistence level of the Turbo model. The trigger line is fired by the  $D^0 \rightarrow K^- \pi^+$  signal decay. In addition to the candidate, primary vertices and other possible tracks associated to it are recorded.

Violation in charm decays [21] and the observation of the doubly charmed baryon  $\Xi_{cc}^{++}$  [22] have been possible thanks to the Turbo model. Among others, one of the analyses which has largely benefited by this trigger strategy is the search for dark photons which can be linked to dark matter particles [23]. For example a dark photon can decay into two muons. A wide range of the dimuon mass spectrum has been analyzed thanks to the possibility of storing only the information of reconstructed objects. Stringent exclusion limits are set by LHCb on the existence of these new particles. A similar analysis has been also performed at CMS with the *scouting* [24].

Figure 6 shows the dimuon mass spectrum reconstructed for prompt-like muons (clear blue) and background from misreconstruction, recorded using the Turbo model.

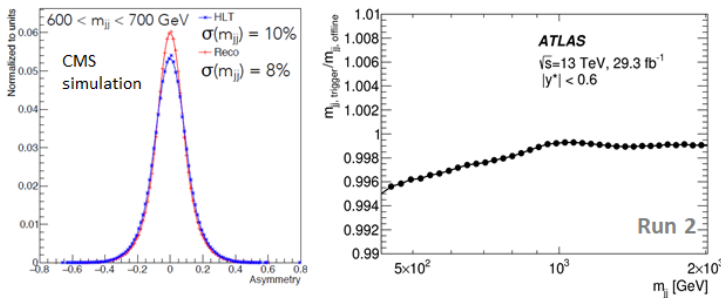


**Figure 6.** Dimuon mass spectra recorded using the Turbo model at LHCb.

During the next data taking period of the LHCb experiment (Run 3, from 2021 onward) the number of events will increase by a factor of five. A purely software based trigger has been developed and will process 30 MHz of bunch crossings on average. This has been achieved thanks to the development of improved algorithms and data structures, in parallel with the use of new processors and vectorisation. In this conditions the trigger output rate will be of 1 MHz, with a raw event size of about 150 kB. Because storage resources are limited to a bandwidth of the order of 10 GB/s, the usage of the LHCb real-time analysis model has to be extended to a more than 70% of the physics program [25].

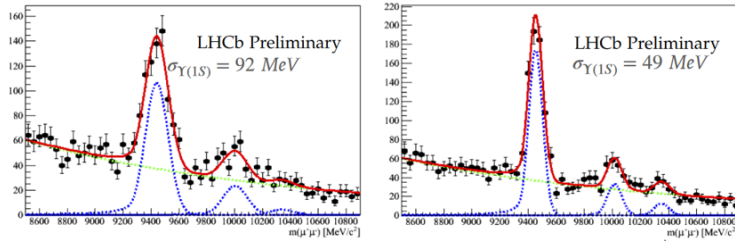
### 3 Alignment and calibration

In order to perform physics analyses at the trigger level precise detector alignment and calibration are needed. Alignment provides the best known position and orientation of the detector elements. Calibration applies corrections to the detector element response to properly reproduce the information provided by the data. Tracking and calorimeter object reconstruction is usually less sophisticated than the "offline" processing, where larger computing time is available. Less iterations in tracking reconstruction and smaller detector regions for hit searches are used in the trigger algorithms. For jet reconstruction at the trigger level, corrections are applied using simulation and calibrated objects from data, such as  $Z \rightarrow \ell^+ \ell^-$ , photons and multijets. As shown in Figure 7, physics performance for trigger-level-reconstructed objects is not very different from offline reconstruction.



**Figure 7.** Left: comparison of the dijet mass resolution for trigger level and offline reconstruction from CMS simulation. Right: ratio of trigger and offline reconstruction efficiencies vs the dijet mass in ATLAS.

At LHCb, all detectors are aligned and calibrated online using a selected set of events, stored in a 10 PB buffer, before data go to the following stage of the high level trigger. The procedure runs at the start of each data fill, and depending on the subdetector involved it can take from a few minutes to several hours. This is necessary because the detector encounters mechanical movements, changes in temperature and pressure, etc. The buffer has a contingency of about two weeks, providing accurate alignment and calibration constants that make the trigger level reconstruction as good as the offline one. No data reprocessing is needed, with the consequent gain in timing and effective CPU resources. Figure 8 shows how the physics performance of the dimuon mass spectra of Upsilon resonances improves after the alignment of the trackers. This has been a key in physics analysis such as the search of new spin-0 bosons decaying into two muons [26].

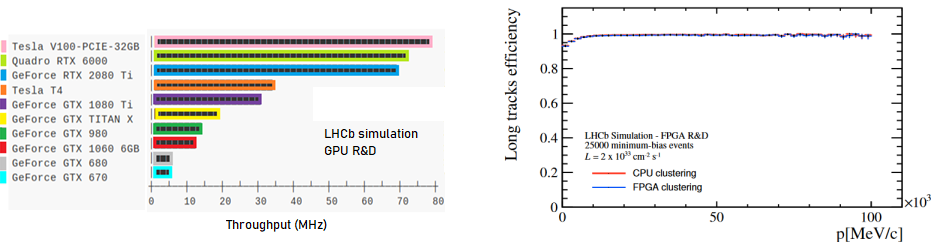


**Figure 8.** Comparison of the dimuon mass resolution before (left) and after (right) the alignment of the trackers of the LHCb experiment, using a Kalman filter in a track subset. The three peaks correspond to, from left to right,  $\Upsilon(1S) \rightarrow \mu^+\mu^-$ ,  $\Upsilon(2S) \rightarrow \mu^+\mu^-$  and  $\Upsilon(2S) \rightarrow \mu^+\mu^-$ .

## 4 Using accelerators

Hardware accelerator platforms such as GPUs (Graphic Processing Units) or FPGAs are promising architectures for speeding up object reconstruction and will be implemented in the trigger system. The Allen project [27] allows LHCb to execute the full first level trigger stage on 500 GPUs at about 4 TB/s. The trigger algorithms include reconstructing proton-proton collision points and the trajectories of the charged particles, identifying the type of particles (hadrons or muons), and finding displaced decay vertices of long-lived particles. Figure 9 (left) shows the high throughput of Allen on various models of GPUs. The achieved physics performance is identical to conventional processors.

R&D projects devoted to accelerate several tasks of the reconstruction chain by means of FPGA cards are being developed. They could help to expedite the trigger decisions. At LHCb the hit clustering of the first tracker is being implemented on FPGAs for the next period of data taking, using an artificial retina [28]. Performing pattern recognition is also being investigated with the same technique [29]. Figure 9 (right) shows the track reconstruction efficiency as a function of track momentum for CPU and FPGA-based clustering algorithms. An FPGA based track finder for the first level of the trigger decision at high luminosity rates is also investigated at the CMS experiment [30].



**Figure 9.** Left: throughput of the Allen project at LHCb on several GPUs. Right: track efficiency of clustering algorithms using FPGAs at LHCb [LHCb-FIGURE-2019-011].

## 5 Connections to other experiments and fields

Real-time analysis is not only important for high energy physics experiments but also in many other fields such as astronomy and astrophysics. Large sky surveys aim to discover transient objects that change brightness over time-scales of seconds to months. Cosmic explosions (supernovae, gamma-ray bursts...), relativistic phenomena such as black hole formation and jets, or potentially hazardous asteroids are some examples of the kind of events that need to be traced in our sky. The Large Synoptic Survey Telescope (LSST), currently under construction in Chile, will take 800 images (30 TB data) and will see millions of transient objects per night, allowing to alert other telescopes and instruments in real time. For this, object classification before taking any decision is crucial [31]. Telescopes such as the Square Kilometer Array (SKA), a radio telescope to be built in south Africa and Australia, can also detect very interesting objects such as fast radio bursts [32], coming from distant astronomical sources and lasting less than a millisecond. From 2026 onward SKA will receive 1 EB/s of raw data, and only a tiny fraction of the data collected by the antennas will be used for scientific analysis. The increasing data volume of upcoming telescopes poses indeed new computing challenges for image reconstruction and selection.

## 6 Conclusions

Data volumes and complexity are exponentially increasing in high energy physics experiments and in other fields. In these conditions it is impossible to record all raw data within the available computing resources. Real-time analyses are crucial to analyse the input data and take a quick decision to keep the important one, extending the phase space for upcoming discoveries.

## References

- [1] G. Aad *et al.* [ATLAS Collaboration], *JINST* **3**, S08003 (2008). doi:10.1088/1748-0221/3/08/S08003
- [2] S. Chatrchyan *et al.* [CMS Collaboration], *JINST* **3**, S08004 (2008). doi:10.1088/1748-0221/3/08/S08004
- [3] A. A. Alves, Jr. *et al.* [LHCb Collaboration], *JINST* **3**, S08005 (2008). doi:10.1088/1748-0221/3/08/S08005
- [4] K. Aamodt *et al.* [ALICE Collaboration], *JINST* **3**, S08002 (2008). doi:10.1088/1748-0221/3/08/S08002
- [5] V. Khachatryan *et al.* [CMS and LHCb Collaborations], *Nature* **522**, 68 (2015) doi:10.1038/nature14474 [arXiv:1411.4413 [hep-ex]].
- [6] R. Aaij *et al.* [LHCb Collaboration], *Phys. Rev. Lett.* **118**, no. 19, 191801 (2017) doi:10.1103/PhysRevLett.118.191801 [arXiv:1703.05747 [hep-ex]].
- [7] G. Aad *et al.* [ATLAS Collaboration], *Phys. Lett. B* **716**, 1 (2012) doi:10.1016/j.physletb.2012.08.020 [arXiv:1207.7214 [hep-ex]].
- [8] S. Chatrchyan *et al.* [CMS Collaboration], *Phys. Lett. B* **716**, 30 (2012) doi:10.1016/j.physletb.2012.08.021 [arXiv:1207.7235 [hep-ex]].
- [9] G. Aad *et al.* [ATLAS Collaboration], *Phys. Rev. D* **101**, no. 1, 012002 (2020) doi:10.1103/PhysRevD.101.012002 [arXiv:1909.02845 [hep-ex]].
- [10] A. M. Sirunyan *et al.* [CMS Collaboration], *Eur. Phys. J. C* **79**, no. 5, 421 (2019) doi:10.1140/epjc/s10052-019-6909-y [arXiv:1809.10733 [hep-ex]].



- [11] ATLAS Collaboration, CERN-LHCC-98-014; CMS Collaboration, CERN-LHCC-2000-038, CMS-TDR-6-1; LHCb Collaboration, CERN-LHCC-2003-031, LHCb-TDR-10; LHCb Collaboration, CERN-LHCC-2014-016, LHCb-TDR-016,
- [12] M. Aaboud *et al.* [ATLAS Collaboration], *Eur. Phys. J. C* **77**, no. 5, 317 (2017) doi:10.1140/epjc/s10052-017-4852-3 [arXiv:1611.09661 [hep-ex]].
- [13] V. Khachatryan *et al.* [CMS Collaboration], *JINST* **12**, no. 01, P01020 (2017) doi:10.1088/1748-0221/12/01/P01020 [arXiv:1609.02366 [physics.ins-det]].
- [14] R. Aaij *et al.* [LHCb Collaboration], *JINST* **14**, no. 04, P04013 (2019) doi:10.1088/1748-0221/14/04/P04013 [arXiv:1812.10790 [hep-ex]].
- [15] S. Chatrchyan *et al.* [CMS Collaboration], *JHEP* **1301**, 013 (2013) doi:10.1007/JHEP01(2013)013 [arXiv:1210.2387 [hep-ex]].
- [16] V. Khachatryan *et al.* [CMS Collaboration], *Phys. Rev. Lett.* **117**, no. 3, 031802 (2016) doi:10.1103/PhysRevLett.117.031802 [arXiv:1604.08907 [hep-ex]].
- [17] A. M. Sirunyan *et al.* [CMS Collaboration], *JHEP* **1808**, 130 (2018) doi:10.1007/JHEP08(2018)130 [arXiv:1806.00843 [hep-ex]].
- [18] The ATLAS Collaboration, ATL-DAQ-PUB-2017-003 (1017).
- [19] M. Aaboud *et al.* [ATLAS Collaboration], *Phys. Rev. Lett.* **121**, no. 8, 081801 (2018) doi:10.1103/PhysRevLett.121.081801 [arXiv:1804.03496 [hep-ex]].
- [20] R. Aaij *et al.*, *Comput. Phys. Commun.* **208** (2016) 35 doi:10.1016/j.cpc.2016.07.022 [arXiv:1604.05596 [physics.ins-det]].
- [21] R. Aaij *et al.* [LHCb Collaboration], *Phys. Rev. Lett.* **122**, no. 21, 211803 (2019) doi:10.1103/PhysRevLett.122.211803 [arXiv:1903.08726 [hep-ex]].
- [22] R. Aaij *et al.* [LHCb Collaboration], *Phys. Rev. Lett.* **119**, no. 11, 112001 (2017) doi:10.1103/PhysRevLett.119.112001 [arXiv:1707.01621 [hep-ex]].
- [23] R. Aaij *et al.* [LHCb Collaboration], *Phys. Rev. Lett.* **120**, no. 6, 061801 (2018) doi:10.1103/PhysRevLett.120.061801 [arXiv:1710.02867 [hep-ex]].
- [24] CMS Collaboration, CMS-PAS-EXO-19-018.
- [25] LHCb Collaboration, CERN-LHCC-2018-014.
- [26] R. Aaij *et al.* [LHCb Collaboration], *JHEP* **1809**, 147 (2018) doi:10.1007/JHEP09(2018)147 [arXiv:1805.09820 [hep-ex]].
- [27] R. Aaij *et al.*, *Comput. Softw. Big Sci.* **4**, 7 (2020) [arXiv:1912.09161 [physics.ins-det]].
- [28] L. Ristori, *Nucl. Instrum. Methods Phys. Res.* **A453** 425 (2000).
- [29] G. Punzi *et al.*, *The 28th International Workshop on Vertex Detectors*, 047 (2019). <https://pos.sissa.it/373/047/pdf>.
- [30] R. Aggleton *et al.*, *JINST* **12**, 12019 (2017).
- [31] S.G. Djorgovski *et al.*, DOI: 10.1109/eScience.2014.7.
- [32] K. W. Bannister *et al.*, *Science* **365**, 565 (2019).