PROCEEDINGS OF SCIENCE

Real-time flavour tagging selection in ATLAS

Carlo Varni, on behalf of the ATLAS collaboration[a,*]

[a] *INFN and Università degli Studi di Genova*

E-mail: carlo.varni@cern.ch

In high-energy physics experiments, online selection is crucial to select interesting collisions from the large data volume. ATLAS b-jet triggers are designed to identify heavy-flavour content in real-time and provide the only option to efficiently record events with fully hadronic final states containing b-jets. In doing so, two different, but related, challenges are faced. The physics goal is to optimise as far as possible the rejection of *light* jets, while retaining a high efficiency on selecting b-jets and maintaining affordable trigger rates without raising jet energy thresholds. This maps into a challenging computing task, as tracks and their corresponding vertices must be reconstructed and analysed for each jet above the desired threshold, regardless of the increasingly harsh pile-up conditions. We present an overview of the ATLAS strategy for online b-jet selection for the LHC Run 2, including the use of novel methods and sophisticated algorithms designed to face the above mentioned challenges. The evolution of the performance of b-jet triggers in Run 2 data is presented, including the use of novel triggers designed to select events containing heavy-flavour jets in heavy-ion collisions.

40th International Conference on High Energy physics - ICHEP2020
July 28 - August 6, 2020
Prague, Czech Republic (virtual meeting)

[*] Speaker

https://pos.sissa.it/

PoS(ICHEP2020)792

1. Introduction

Processes containing b-jets are of particular interest in the ATLAS [1] physics program. The ability to select events containing heavy-flavour-initiated jets represents a pivotal feature for these processes, especially in the absence of any additional distinguishing characteristics, such as high-p_T light leptons (electrons or muons), or missing transverse momentum. In particular, for processes such as $HH \rightarrow bb\bar{b}\bar{b}$, all hadronic $ttH(H \rightarrow b\bar{b})$, VBF $H(H \rightarrow b\bar{b})$ and $bA(A \rightarrow b\bar{b})$. The ATLAS trigger system [2] provides such an ability with b-jet triggers, which use flavour tagging as a tool for correctly identifying and separating jets stemming from light quarks or gluons (*light*-jets), and heavy quarks (c- or b-jets).

In this contribution, the performance of b-jet triggers in Run 2 data (2015-2018) is presented, including the use of novel triggers designed to select events containing heavy flavour jets in heavy-ion collisions.

2. The evolution of *b*-jet triggers

b-jet triggers collect information from several domains (i.e. tracking, jets and flavour tagging) and assemble them in a coherent way. As a consequence, updates in these areas have a direct impact on b-jet triggers. Between Run 1 (2009 to early 2013) and Run 2 these domains underwent major revisions of their strategies and during Run 2 additional improvements were applied to all these areas, together resulting in improved b-jet trigger performance.

2.1 Tracking strategy

During Run 1 the track reconstruction strategy consisted of running the full tracking individually in the full z-width Regions of Interest (RoIs) centred on jets. Albeit such an approach is feasible in a low event multiplicity environment, it is not advisable once the number of collisions occurring in parallel grows and RoIs start overlapping. In this scenario the same reconstruction algorithm would run multiple times in the same regions of the detector, thus leading to a waste of CPU resources and a potential bias on primary vertex finding, by double-counting tracks in overlapping regions.

For Run 2, the tracking strategy was revised to cope with the increased beam energy and interaction multiplicity. A multistage strategy was adopted. The first stage performs a fast track reconstruction on an ensemble of RoIs constructed around calorimetric jets. This is the so-called *Super-RoI*, whose constituents are narrow in η and ϕ ($|\Delta\eta| < 0.1$ and $|\Delta\phi| < 0.1$ with respect to the jet axis) but without constraints on the z-coordinate. Running tracking algorithms on this composite RoI provides tracks that are then used by specialised algorithms to identify the primary vertex position. In order to reduce the CPU consumption, only track candidates with $p_T > 5$ GeV are considered at this stage. The second stage performs tracking in large RoIs around the calorimetric jets ($|\Delta\eta| < 0.4$ and $|\Delta\phi| < 0.4$ with respect to the jet axis) and with an additional constraint on the z-coordinates provided by the primary vertex ($|\Delta z| < 10$ mm). At first, *fast tracking* is run considering all track candidates with $p_T > 1$ GeV, which are then used to seed *precision tracking*: offline tracking algorithms [3, 4] that resolve any ambiguity with respect to duplicated hits or hits wrongly attributed to tracks.

This newly-adopted strategy, illustrated in Figure 1, resulted in a more robust and reliable track reconstruction. Also, this guarantees the possibility of applying an additional filter between the two stages: RoIs with no tracks pointing to the primary vertex are not taken into account. Track efficiencies are shown in Figure 1.

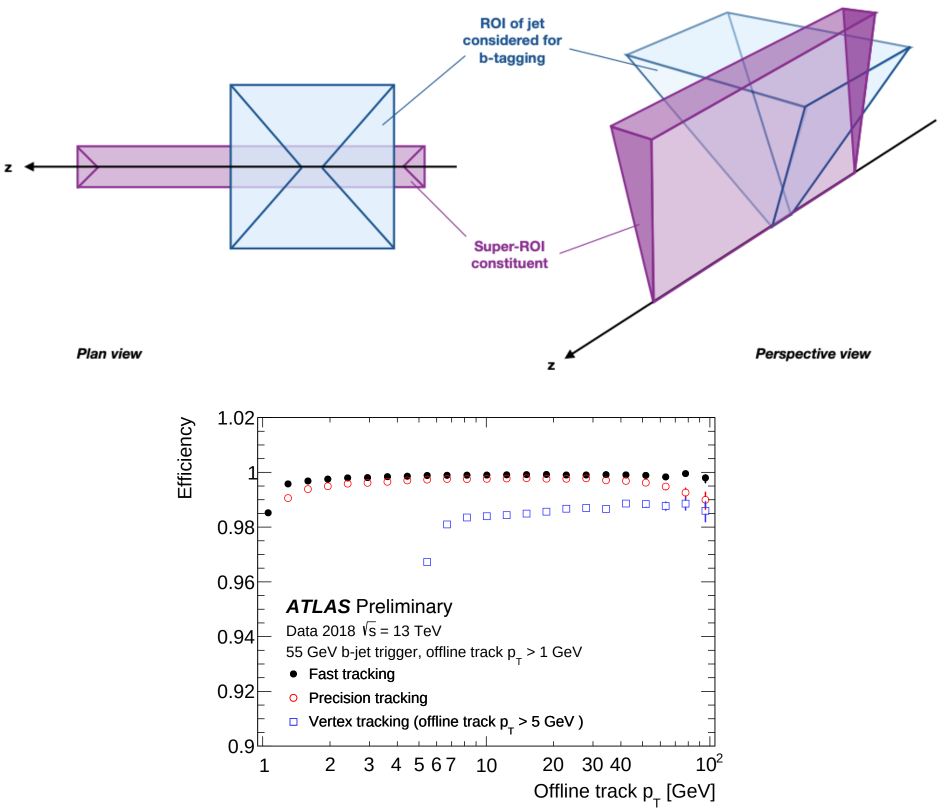


Figure 1: (upper plot) The RoIs used in the multi-stage approach used by b-jet triggers during Run 2. The purple Super-RoI is used to identify track candidates for the primary vertex determination, while the blue RoI is used for reconstructing tracks for the flavour tagging computation. (lower plot) The tracking efficiencies of the different tracking stages: vertex, fast and precision tracking [5].

2.2 Jet calibration

The Global Sequential Calibration (GSC) [6] was introduced in 2017 for reducing fluctuations in the jet energy measurement. It is a sequential jet calibration that uses the transversal and longitudinal properties of the jet structure (longitudinal shower shapes of jets, and characteristics of associated tracks) in order to correct the energy scale of jets and provide a better energy resolution. Thus it allows to correct for the dead material and non-compensation effects of the ATLAS calorimeters. Due to the high CPU consumption of tracking, two different GSC versions are available: the first one using only the information from calorimeters; the other one using also tracks associated to jets. The latter provides improved performance with respect to the former, with sharper p_T turn-ons, as shown in Figure 2. In Run 2, b-jet triggers were the only signature that could afford to use this improved GSC correction since tracks were always reconstructed for the flavour tagging computation.

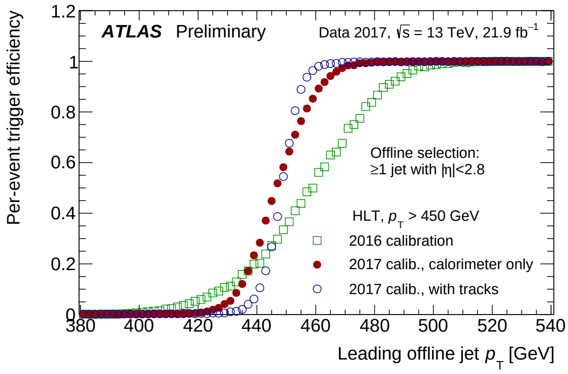


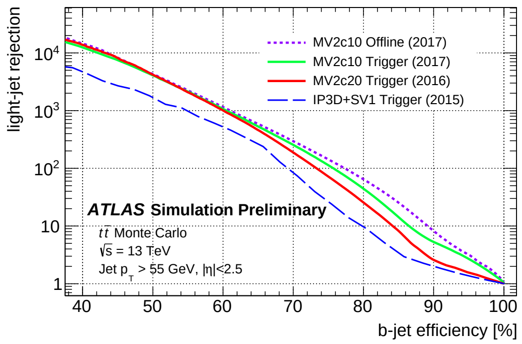
Figure 2: Efficiencies are shown for a single-jet trigger with three different calibrations applied to jets in the ATLAS high-level trigger. In green the calibration applied in 2016 data, in red the updated calibration applied in 2017, utilising only calorimeter information, and in blue this updated calibration with track information included [7].

2.3 Flavour tagging algorithms

The differences between b- and c- or *light*-jets is reflected in several physical quantities bringing different and complementary – albeit partially correlated – information that must be combined. Several specialised algorithms have been developed to analyse these properties: algorithms that analyse track d_0 and z_0 significances and their correlations; algorithms that reconstruct a single, inclusive secondary vertex; and algorithms that take into account the reconstruction of the complete b-hadron decay chain. The outputs from these complementary algorithms can then be combined with multivariate techniques, resulting in more versatile and powerful tools.

In Run 2 the same flavour tagging algorithms used at the offline level were adopted also at the trigger level. Such an approach is highly desirable since it centralises the developments of the algorithms and it maximises the correlation between online and offline flavour tagging. In particular, a Boosted Decision Tree based algorithm called MV2 [8] was adopted as the recommended tagger at the trigger level during Run 2.

The training and testing samples of the flavour tagging algorithms were initially based on simulated $t\bar{t}$ events, and therefore limited in statistics at high p_T. As a consequence, the b-tagging algorithm does not efficiently learn to discriminate b-jets from *light*- or c-jets in high-p_T kinematic regions. The introduction of a new training based on a mixed sample of $t\bar{t}$ and Z' (mass of 1000 GeV) simulated events – the so-called *hybrid tuning* – resulted in enhanced performance of b-jet triggers at high p_T [9], with negligible effects at lower p_T, where the training sample is dominated by the $t\bar{t}$ sample. This is shown in Figure 3. b-jet triggers that use this hybrid tuning were commissioned in 2018.



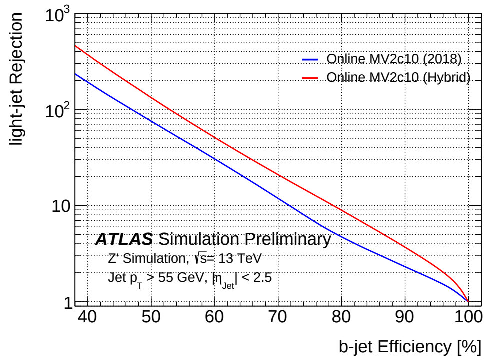


Figure 3: Performance of online b-tagging: (left plot) the evolution of performance during Run 2, from 2015 to 2017 compared to offline b-tagging in 2017; (right plot) the comparison between the hybrid and the $t\bar{t}$-based tuning evaluated on $Z \rightarrow t\bar{t}$ simulated events [10].

3. Performance of b-jet triggers

The performance of b-tagging algorithms with respect to the true flavour of jets is expressed in terms of *light*-jet and c-jet rejection as a function of b-jet efficiency. All the above-mentioned upgrades resulted in enhanced performance of b-jet triggers. The evolution of performance is illustrated in Figure 3. In the course of time the online *light*-jet rejection has increasingly gotten closer to the offline performance, resulting in better b-jet identification at trigger level.

Flavour tagging performance at both the online and offline levels must be understood and calibrated before being used in any physics analysis. Trigger scale factors – derived from both data and simulated $t\bar{t}$ events – are applied to Monte Carlo simulations to correct for any mismodelling of the b-jet trigger performance, and must be applied in addition to the offline b-tagging scale factors. The calibration is based on a likelihood-based method, as described in Refs. [11, 12].

4. Triggers in heavy ion collision data

Heavy ion collisions are extremely dense environments. The use of traditional b-jet triggers in such an environment is not feasible due to the complexity of the computation and the resulting high trigger rate caused by the large number of jets and high track multiplicity. However, approximately 20% of b-hadrons undergo semi-leptonic decays, with at least one electron or muon produced at large angles relative to the jet axis (although overlapped with the jet itself: $\Delta R \leq 0.5$). Dedicated triggers have been designed to target low-p_T muons that are geometrically matched to a jet. They represent one of the rare ways to select events containing b-jets during heavy ion collisions, since the presence of muons provides an early acceptance criterion to run tracking on all jets. The matching requirement improves the purity of the selected sample and increases the rejection power against *light*-jet backgrounds. As a result, these semi-leptonic b-jet triggers can reach lower jet p_T ranges compared to the standard b-jet triggers. They are also used to provide a sample of b-jet-enriched data used to calibrate the b-tagging algorithms in proton-proton collisions.

5. Conclusions

A brief description of b-jet triggers in ATLAS, which allow real-time flavour tagging selection, has been presented. The upgrades that these triggers have undergone between Run 1 and Run 2 have been illustrated: these include the adoption of novel tracking strategies, as well as new jet calibration procedures and the adoption of new flavour tagging algorithms. These upgrades resulted in enhanced performance of b-jet triggers. Moreover, calibration of flavour tagging and the use of triggers designed to select events containing heavy-flavour jets in heavy-ion collisions have been introduced.

References

[1] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[2] ATLAS Collaboration, *Performance of the ATLAS trigger system in 2015*, Eur. Phys. J. C **77** (2017) 317.

[3] T. Cornelissen et al., *The new ATLAS track reconstruction (NEWT)*, J. Phys. Conf. Ser. **119** (2008) 032014.

[4] ATLAS Collaboration, *Performance of the ATLAS Silicon Pattern Recognition Algorithm in Data and Simulation at $\sqrt{s} = 7$ TeV*, ATLAS-CONF-2010-072.

[5] ATLAS Collaboration, *HLT Tracking Public Results*, `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/HLTTrackingPublicResults`.

[6] ATLAS Collaboration, *Determination of jet calibration and energy resolution in proton-proton collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector*, CERN-EP-2019-057.

[7] ATLAS Collaboration, *Public Jet Trigger Plots for Collision Data*, `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/JetTriggerPublicResults`.

[8] ATLAS Collaboration, *Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run*, ATL-PHYS-PUB-2016-012.

[9] ATLAS Collaboration, *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*, ATL-PHYS-PUB-2017-013.

[10] ATLAS Collaboration, *Public b-Jet Trigger Plots for Collision Data*, `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/BJetTriggerPublicResults`.

[11] ATLAS Collaboration, *ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **79** (2019) 970.

[12] ATLAS Collaboration, *Measurements of b-jet tagging efficiency with the ATLAS detector using $t\bar{t}$ events at $\sqrt{s} = 13$ TeV*, JHEP **08** (2018) 89.