Real-time $b$-jet identification in ATLAS

A. Buzatu on behalf of the ATLAS Collaboration
Research Associate, University of Glasgow, Glasgow, G12 8QQ, UK
E-mail: adrian.buzatu@glasgow.ac.uk

Abstract. There are interesting physics processes studied at the ATLAS detector at the LHC, whose signatures contain quark jets, but no charged lepton: top quark pair production where both quarks decay hadronically, Higgs boson produced in vector boson fusion and decaying to a bottom quark pair, or supersymmetric signatures with no charged lepton. Data for these processes can only be collected using jet-based triggers. As LHC instantaneous luminosity increases, jet transverse-energy thresholds need to be raised, lowering the signal acceptance for a wide range of signatures. However, the need to raise jet thresholds can be made less pressing if jets originating from bottom quarks are identified as such ($b$-tagging) already at the trigger stage if at least one jet originates from a bottom quark. ATLAS started using $b$-jet real-time identification in 2011, increased their performance in 2012 and plans to use even more advanced taggers in 2015.

1. Introduction
The Large Hadron Collider (LHC) at CERN has had a successful Run 1 from 2010 to 2012 in providing head-on proton-proton collisions at ever increasing centre-of-mass ($\sqrt{s}$) energies, instantaneous luminosities, and number of interactions per bunch crossing. The ATLAS detector [1] is one of the two general purpose detectors that record LHC data. In order to record events whose signatures do not contain a charged lepton, ATLAS uses triggers based on streams of collimated particles produced by the hadronization and fragmentation of a quark or gluon, collectively called jets. As the instantaneous luminosity increases, an option to maintain trigger rates constant is to increase the jet transverse-energy thresholds. For processes such as top quark pair production where both quarks decay hadronically [2], Higgs boson produced in vector boson fusion and decaying to a bottom quark pair, or supersymmetric signatures with no charged lepton, at least one jet contains a $B$ hadron. For such signatures, the need to raise jet thresholds can be made less pressing if such jets are identified ($b$-tagged) in real time by the ATLAS trigger system. The need to increase thresholds is reduced by the fact that most multi-jet candidates originate in $u$, $d$, $s$ quarks or gluons (have a light-jet content). Applying the $b$-tagging requirement rejects light jets and thus produces an extra rejection factor that can substitute the one possibly coming from increased thresholds. These triggers are collectively called $b$-jet triggers and improved during the data taking to cope with the changing conditions. In the 2011 (2012) run at $\sqrt{s}$=7 TeV ($\sqrt{s}$=8 TeV), the maximum instantaneous luminosity was of $3.65 \times 10^{33}$ cm$^{-2}$s$^{-1}$, producing a maximum of about 19 (33) interactions per bunch crossing ($\mu$). For the Run 2 data taking starting in 2015, LHC will have a $\sqrt{s}$ of about 13-14 TeV, as well as a higher instantaneous luminosity of about $2 \times 10^{34}$ cm$^{-2}$s$^{-1}$, which translates into an average $\mu$ of around 60. ATLAS started using $b$-jet real-time identification in 2011,
increased its performance in 2012 and plans to use even more advanced taggers in 2015. This paper discusses the evolution of these triggers during the LHC Run 1 and the plans for upgrade for the LHC Run 2 data taking, as well as the data-driven calibrations that are necessary in order for data analyses to be able to use these triggers.

2. b-tagging Algorithms

b-tagging algorithms are developed taking into account several distinguishing features of a \( B \) hadron \[3\]. The \( B \) hadron lifetime is about 1.5 ps, allowing it to travel a measurable distance before decaying. Also, its mass is about 5 GeV. The trajectories of its daughter charged particles (tracks) create a secondary vertex displaced from the primary interaction vertex. Reconstructed tracks is therefore a key to b-tagging, as b-tagging algorithms take as inputs high-level-objects reconstructed from tracks, such as the reconstructed primary vertex, secondary vertex and the distribution of track impact parameters, as seen in Figure 1.

![Figure 1. Schematic view of tracks emerging from a primary and secondary vertex in a b-jet.](image)

3. The ATLAS Detector

As a general-purpose detector, ATLAS has a cylindrical and backward-forward symmetry, which allows it to detect particles heading in all directions from the most probable collision spot, in the centre of the detector. ATLAS is formed of four highly-granular and hermetic subdetectors arranged in concentric layers around the direction of the beams of colliding protons. The subdetector that reconstructs tracks of charged particles by their ionization of matter, being therefore crucial for jet b-tagging, is the closest to the beam line and is suggestively called the Inner Detector (ID).

The ID is made up of three independent subsystems. Starting at only 5 cm from the beam line, the silicon pixel detector \[5\] is made of about 80 million silicon pixels arranged in three layers. The size of one pixel is 50 x 400 \( \mu \)m\(^2\). Next is the silicon micro-strip detector consisting of four cylindrical layers that have a stereoscopic geometry, which allows the measurement of the position of particle ionisation energy deposits (hits) in three dimensions. The last from the beam line is the transition radiation tracker, which is made of straw tubes of 4 mm diameter, which provide \( r-\phi \) position hit measurement up to a radius of about 111 cm from the beam line. The ID measures the position of origin of tracks with a precision of 10 microns in the transverse plane, thus allowing for vertex reconstruction, mostly due to the first subsystem. Furthermore, being immersed in a solenoid-generated 2 T magnetic field, tracks are curved, which allows for precise transverse momentum measurement.

Besides the ID, there are the subdetectors of electromagnetic calorimeter, which measures the energy of electrons, positrons and photons, the hadronic calorimeter, which measures the energy of hadrons, and the muon spectrometer, which measures the momentum of muons and contributes to pattern recognition.
4. **b-jet Trigger Menu**

The ATLAS trigger system and its performance is described in detail in [4]. It consists of three levels and selects every second about 700 events to be recorded from 20 million events. Each trigger level is more complex than the previous one. The first level (L1) is based on hardware that uses coarse information from the electromagnetic and hadronic calorimeters, as well as from the dedicated trigger chambers from the muon spectrometer. The second level (L2) and the third level, the Event Filter (EF), are based on almost-commercial computing farms. Together they are called the High Level Trigger (HLT). The L1 trigger identifies regions of interest (RoIs) in the detector around electron, photon, jet and muon candidates. The L2 is the first that has access to the ID detector and reconstructs tracks in the RoIs identified at L1. Then the EF has access to the full detector and uses software that is similar to the offline reconstruction software, albeit a bit simpler. The L1 reduces the event rate to about 70 kHz and has a time latency of 2.5 µs. The L2 reduced the rate further to about 5 kHz and has a latency of about 75 ms. The EF reduced the rate further to about 700 Hz and has a latency on the order of 1 s. In 2012, the totality of the typical b-jet triggers selected every second was 5,000/900/45 events at L1/L2/EF.

5. **b-jet Triggers in 2011**

The b-jet triggers were already implemented in the 2010 data run, using for b-tagging the JetProb [6] algorithm. They were used to commission the real-time tracking reconstruction and b-tagging, but were not actively selecting events [4]. In the 2011 run the b-jet triggers were used to collect data. The algorithm takes as input the distribution of impact parameter of tracks inside a jet and computes the probability for all the tracks to originate from the primary vertex. For a b-jet, the value is close to zero.

We denote $d_0$ the transverse impact parameter, defined as the distance of closest approach between a particle track and the primary interaction vertex. We denote the measured uncertainty on $d_0$ as $\sigma(d_0)$ and the significance of $d_0$ as $S(d_0) = d_0/\sigma(d_0)$. The sign of $d_0$ is positive (negative) if the track crosses the jet axis in front of (behind) the primary vertex. The sign of $d_0$ is positive for most tracks associated with b-jets due to the long lifetime of the B hadron. However, tracks from jets that originate from light quarks or from gluons may appear also positive due the finite impact parameter resolution, leading to light-jets being indentified incorrectly as b-jets (mistag or fake b-tagging). Figure 2 a) shows the comparison of the impact parameter significance for tracks associated with b-jets, from c-jets and from light-jets in a simulated sample of di-jet QCD to those measured in 2011 data. The trigger does not rely on simulation, but on the knowledge of the distribution of the tracks from light-jets, which can be easily derived from data.

![Graph](image_url)

**Figure 2.** a) Signed transverse impact parameter significance of reconstructed tracks at the Event Filter level [7]. b) Resolution for the primary vertex z position estimate as a function of the number of online tracks at L2 and EF. [7]. c) The typical transverse distribution of primary vertices in the 2011 data taking. [8].
The correct reconstruction of the primary vertex is crucial for all the $b$-tagging algorithms, including the JetProb. Its $z$ position is measured for each bunch crossing using a sliding window algorithm that counts the number of tracks around each $z$ value. In 2011 the vertexing was done in every RoI using all the tracks with $p_T > 1$ GeV reconstructed in the RoI. The resolution for the primary vertex $z$ position estimate as a function of the number of online tracks at L2 and EF is seen in Figure 2 b). To evaluate the $d_0$ of tracks precisely, one needs to measure also the $x$ and $y$ coordinates of the primary vertex. This is not done for each bunch crossing, but rather an average for many bunch crossings that share the same running conditions. The result is called the beam spot and is seen in Figure 2 c).

There were three main $b$-jet triggers in 2011: selecting four jets (two of which are $b$-tagged), selecting four jets (one of which is $b$-tagged), and selecting two jets (both jets $b$-tagged). The typical rates for the former trigger, which was used by ATLAS for a measurement of the top quark production where one top quark decays hadronically and the other decays to a hadronically-decaying tau lepton [2], are shown in Figure 3 a).

6. $b$-jet Triggers in 2012

In the 2012 run more advanced $b$-tagging algorithms replaced the relatively simple JetProb in order to cope with the higher instantaneous luminosity. Two algorithms were combined. The first algorithm used both the transverse and longitudinal impact parameter distributions. All per-track likelihoods to originate from the primary vertex are then combined into a per-jet quantity. The second algorithm is a likelihood of the secondary vertex reconstruction based on the mass, the number of two-track vertices and the fraction of the energy of the jet in the secondary vertex, as seen in Figure 3 b). The primary vertex reconstruction was very similar to the one in 2011. However, in 2012, all the tracks from all the RoIs were used at the same time for estimating a $z$ primary vertex, thus achieving a better resolution and robustness against pile-up. Both algorithms use prior knowledge from simulated top quark events of how these distributions look like for $b$-jets and light-jets (log likelihood method).

Figure 3. a) Rate for a trigger with four jets, two of which are $b$-tagged [7]. b) Ratio of energy sum of quality tracks associated with the prescaled Event Filter jets’ secondary vertex and the energy sum of all quality tracks in the jet for Event Filter jets. [7]

7. Plans for $b$-jet Triggers in 2015

For the Run 2 data taking, ATLAS plans to insert a new pixel detector layer at only 3.2 cm from the beamline, whereas the current closest layer is at 5 cm. This will allow better tracking reconstruction, which will translate into a better primary vertex reconstruction, which determines better $b$-tagging. ATLAS is investigating the online 3D reconstruction of a primary vertex, which already is performed offline as well as using online the $b$-tagging algorithms.
currently used in the offline analyses. Such an algorithm is for example the multivariate decision based on the jet information already used in the 2011 and 2012 algorithms, but adding also more information, such as the reconstruction of a third vertex from the decay of a c quark produced in the decay of the b-quark. Though in the LHC Run 1 the b-jet triggers were used only for signatures that contained no charged lepton, in the LHC Run 2 they may be used for signatures that contained charged leptons as well, for example by asking for at least one charged lepton and one b-jet, a typical signature for a SUSY process.

8. Calibrations
For data collected with the b-jet triggers to be used in offline analyses, trigger selection efficiencies need to be measured [1]. Furthermore, such inefficiencies differ in data and simulated sample and introduce a bias with respect to the offline event selection, as seen in Figure 4.a) for different operating points. Since ATLAS adopts, among others, methods for b-tagging calibration using semileptonic bottom-quark decays, triggers that select events with jets with a low transverse-momentum muon inside were used to record data enhanced in b-jets since 2010 in such a way as the rate is constant for a wide range of transverse momenta, as seen in Figure 4.b).

![Figure 4.](image)

**Figure 4.** a) The measured offline JetProb distribution for jets b-tagged at the HLT at three different operating points (yellow-QCD, black-data, red/blue/green-loose/medium tight trigger cut). b) Distribution of triggers with jets that have loose muon inside as a function of the jet transverse momentum in offline muon-in-jets [7].

9. Conclusions and Outlook
The b-jet triggers have been active in ATLAS during the Run 1 data taking in 2011 and 2012, using ever more advanced algorithms. Trigger efficiency calibrations have been performed, allowing these triggers to be used by analyses that have no charged lepton in their signature. In preparation for the conditions for the Run 2 data taking in 2015, the tracking is improved with a new pixel-detector layer, and advanced offline b-tagging algorithms will be ported to the online trigger HLT level.

References
[1] ATLAS Collaboration, The ATLAS Experiment at the CERN Large Hadron Collider, JINST 3 S08003 (2008)
[2] ATLAS Collaboration, Measurement of the t\(\bar{t}\) production cross section in the tau+jets channel using the ATLAS detector, Eur.Phys.J.C 73 (2012)
[3] J. Beringer et al. (Particle Data Group), Phys. Rev. D86, 010001 (2012)
[4] ATLAS Collaboration, Performance of the ATLAS Trigger System in 2010, Eur. Phys. J. C 72 (2012)
[5] ATLAS Collaboration, ATLAS pixel detector electronics and sensors, JINST 3 P07007 (2008)
[6] ALEPH Collaboration, A Precise measurement of \(\Gamma(Z \rightarrow b\overline{b})/\Gamma(Z \rightarrow \text{hadrons})\), Phys. Lett. B, 313, 535 (1993)
[7] https://twiki.cern.ch/twiki/bin/view/AtlasPublic/BJetTriggerPublicResults
[8] https://twiki.cern.ch/twiki/bin/view/AtlasPublic/BeamSpotPublicResults