

Performance and calibration of b -tagging with the ATLAS experiment at LHC Run-2

Ian Connelly^{1,a}, on behalf of the ATLAS Collaboration

¹Royal Holloway, University of London

Abstract. The identification of b -flavoured jets is key to many physics analyses at the LHC, including measurements involving Higgs bosons or top quarks, and searches for physics beyond the Standard Model. The capacity of ATLAS to efficiently tag b -jets has been enhanced for Run-2 with the addition of the Insertable B Layer (IBL), and improvements in the tracking and b -tagging algorithms. In the algorithm optimisation special emphasis has been placed in improving the performance for reconstructing high- p_T b -jets, addressing the challenges posed by track and vertex reconstruction in such an environment. The efficiency and rejection power of these algorithms have been calibrated on data taken in 2015, in particular by exploiting the copious production of b -jets in top quark decays, complemented by studies in multi-jet events.

1 Introduction

The Large Hadron Collider (LHC) has provided proton-proton collisions with centre-of-mass energies of 7 and 8 TeV during Run-1 and at 13 TeV during Run-2. Using these datasets, the ATLAS experiment has recorded and released over 500 publications which analyse the collisions with immense precision. Some of these analyses require the ability to identify whether a collimated group of particles, collectively known as a hadronic jet, originated from the decay of a b -hadron.

A hadronic jet which originated from the production of a b -hadron is commonly referred to as a b -jet. Their identification is a crucial factor in a range of physics analyses; measurements of top-quark properties (where a top-quark will decay to a b -quark with a branching fraction close to unity), searches for supersymmetric particles (where the production of the lightest supersymmetric particle in a model will often result in a b -jet) and searches and measurements of the Higgs boson (where the favoured decay is to a pair of b -quarks).

2 The ATLAS Detector

The ATLAS detector¹ is a general purpose detector providing an inclusive picture of proton collisions at the LHC. The ATLAS detector has been recording data at $\sqrt{s} = 13$ TeV since June 2015.

^ae-mail: ian.connelly@cern.ch

¹ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the centre of the detector and the z -axis along the beam pipe. The x -axis points from the IP to the centre of the LHC ring, and the y -axis points upward. Cylindrical coordinates (r, ϕ) are used in the transverse plane, ϕ being the azimuthal angle around the z -axis. The pseudorapidity is defined in terms of the polar angle θ as $\eta = -\ln(\theta/2)$.

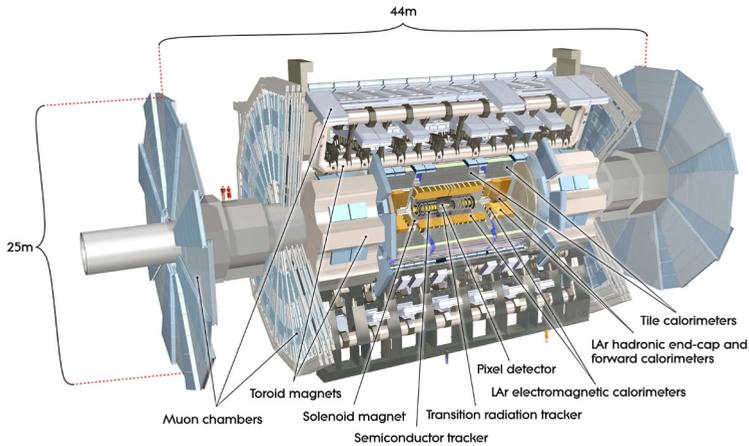


Figure 1. Cross-section of the ATLAS detector, highlighting the different detector components [1].

The ATLAS detector can be broadly categorised into three detector components: inner tracker, calorimeters and muon spectrometers. For the purposes of b -tagging, the performance of the inner tracker and the charged track reconstruction efficiency are the key considerations.

The inner detector consists of four silicon pixel layers, four silicon strip layers and a transition radiation tracker. The inner detector has an acceptance of $|\eta| < 2.5$ and sits in a 2 T axial magnetic field. The closest pixel layer is known as the insertable b-layer (IBL) which was installed during the long shutdown between Run-1 and Run-2. The IBL was installed 33.25 mm radially away from the beam line and has a smaller pixel size compared to the other pixel layers ($50 \times 250 \mu\text{m}^2$ compared to $50 \times 400 \mu\text{m}^2$ in the R-z plane). The addition of a new pixel layer added increased redundancy to the existing detector, providing an additional space-time point to reconstruct charged particle tracks close to the beam line. This closer proximity also improves the robustness of the tracking algorithms as the impact of multiple scattering is much reduced when extrapolating tracks back to the interaction region.

When reconstructing a track and extrapolating back towards to the interaction point, the point of closest approach is known as the impact parameter. This value is measured in the transverse plane (d_0) and in the longitudinal direction (z_0). The resolution of these variables feeds directly into the ability to identify b -jets as typically tracks originating from a b -hadron decay will have larger impact parameters as they originate from a different production vertex. The transverse and longitudinal impact parameter resolution is shown in Figure 2a and 2b respectively. The comparison is shown between Run-1 data collected in 2012 at 8 TeV and Run-2 data collected in 2015 at 13 TeV. The improvement at low transverse momentum comes from the inclusion of the IBL which reduced the impact of multiple scattering. The improvement in the longitudinal impact parameter across the entire transverse momentum range is due to the improved granularity in the z-direction. One additional change from Run-1 to Run-2 has been an improvement in the track reconstruction algorithms. Neural network pixel hit clustering has improved the reconstruction of tracks that share the same hit in the pixel layers. This has primarily improved the reconstruction of tracks in the core of high- p_T jets [2].

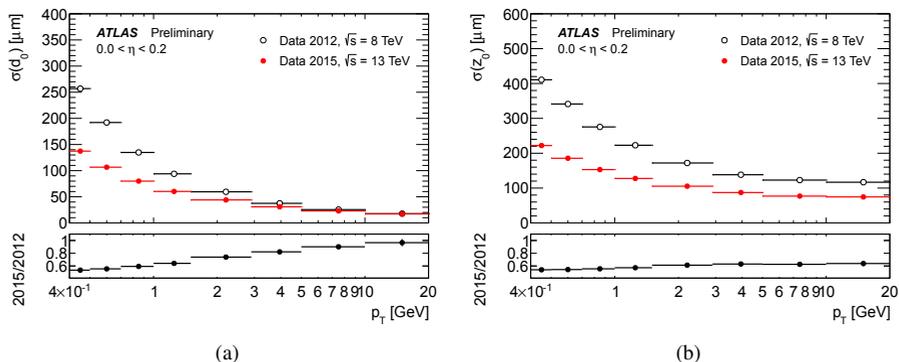


Figure 2. Comparison of impact parameter resolutions ($\sigma(d_0)$, $\sigma(z_0)$) between Run-1 and Run-2 [3].

3 b -Tagging in ATLAS

b -tagging is defined as the identification of a hadronic jet which originated from the fragmentation and hadronisation of a b -hadron, with the resulting jet being known as a b -jet. b -tagging allows for precise Standard Model measurements, plays a key role in a number of Higgs searches, and opens the door to exploring New Physics scenarios with heavy flavours.

The rate at which a true b -jet is identified defines the b -tagging efficiency for a particular b -tagging algorithm. The true flavour of a jet is defined in simulated data using a spatial ΔR matching between stable hadrons and reconstructed jets. A hierarchical matching is performed, first checking whether a b -hadron can be matched, followed by a charm hadron and followed by a τ -lepton. This matching procedure results in a jet being classified respectively as either a b -jet, c -jet, τ -jet or a light-jet (if no match is found).

b -hadrons have a collection of unique properties which can be exploited in order to be identified in data. An illustration of the production of a b -jet is shown in Figure 3. The decay of a b -hadron is suppressed by a CKM factor which results in a longer flight path before decaying in the inner detector compared to charm and light hadrons. In the lab frame, this means that a b -hadron will travel on the order of a few millimetres before subsequently decaying². This gives rise to a displaced secondary vertex which is a key signature of a b -hadron decay. In addition, the b -quark is the heaviest quark which can form hadrons which results in a heavy hadronic decay signature. Finally, the decay of a b -hadron will on average have a higher charged track multiplicity in the decay than other hadron decays.

3.1 b -Tagging Algorithms

There are three main categories of b -tagging algorithms commonly used in ATLAS. There are impact parameter based algorithms (IP2D, IP3D); inclusive secondary vertex reconstruction algorithms (SV); and decay chain reconstruction algorithms (JetFitter). These algorithms contribute complementary information and can be combined using a multivariate function to create a single b -tagging discriminant (MV). Whilst b -tagging is a status used to describe a hadronic jet, the b -tagging algorithms use information relating to tracks which have been associated to a jet. This association is performed using a p_T

²A B^+ hadron has $c\tau$ of approximately 0.5 mm. If this hadron was produced with p_T of 100 GeV, the resultant distance travelled radially into the detector is given by $\langle l \rangle = \beta\gamma c\tau \approx 10$ mm.

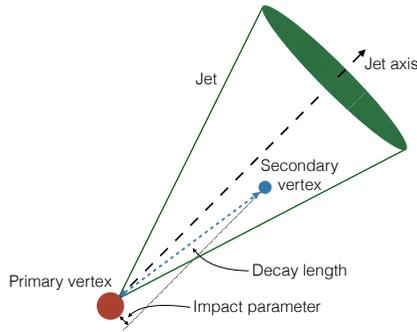


Figure 3. Illustration of the production of a b -jet.

dependent spatial ΔR matching between tracks and jets. The tracks are then subjected to an algorithm-dependent selection in order to determine whether they are used by the b -tagging algorithms.

The impact parameter algorithms use the signed impact parameter significance,

$$\text{sig}(x) = \frac{x}{\sigma_x} \quad (1)$$

of the tracks which are matched to a jet. The sign is defined as positive (negative) if the point of closest approach of the track to the primary vertex is in front (behind) the primary vertex with respect to the jet direction. The IP2D algorithm uses templates of the signed impact parameter significance in the transverse direction for different jet flavours. The IP3D algorithm uses templates of the signed impact parameter significance in both the transverse and longitudinal directions. The distributions for the d_0 and z_0 signed impact parameter significances are shown in Figure 4a and Figure 4b respectively. These distributions provide probability distribution functions (PDFs) for tracks coming from different jet flavour hypotheses. A ratio can be defined between the b -jet hypothesis and light jet hypothesis to combine this information into a single likelihood-ratio discriminant. The final IP2D and IP3D discriminants can be seen in Figure 5a and Figure 5b for three different flavour hypotheses. It can be seen that one could choose to cut on this variable to select a sample in data with a high b -jet purity.

The secondary vertex algorithms attempt to reconstruct an inclusive displaced secondary vertex which is formed by selected tracks within a jet. All possible two-track vertices are identified from the track candidates. Vertices which are likely from long-lived particles (K_s , Λ), photon conversions or hadronic interactions with the detector material are removed. A single inclusive vertex is reconstructed by iteratively removing outlier tracks until a good vertex candidate is identified. Once reconstructed, the kinematic properties of the tracks can be used to offer discrimination between the b -jet and light-jet hypotheses. The invariant mass of the reconstructed secondary vertex is shown in Figure 6 where a b -jet typically has a higher value compared to non- b -jets.

The multi-vertex decay chain algorithm attempts to reconstruct the weak decay chain of a b -hadron produced at the primary vertex which subsequently decays to a charm hadron (producing a tertiary vertex). A Kalman filter is used to identify the b -hadron flight path using the assumption that the tertiary vertex lies along the same flight path. The tracks which are associated to the reconstructed vertices can also provide discriminating information, in a similar way to the SV algorithm. The invariant mass of tracks fitted to one or more of the displaced vertices is shown in Figure 7a and the fraction of energy carried by the fitted tracks compared to the total energy of the jet is shown in Figure 7b.

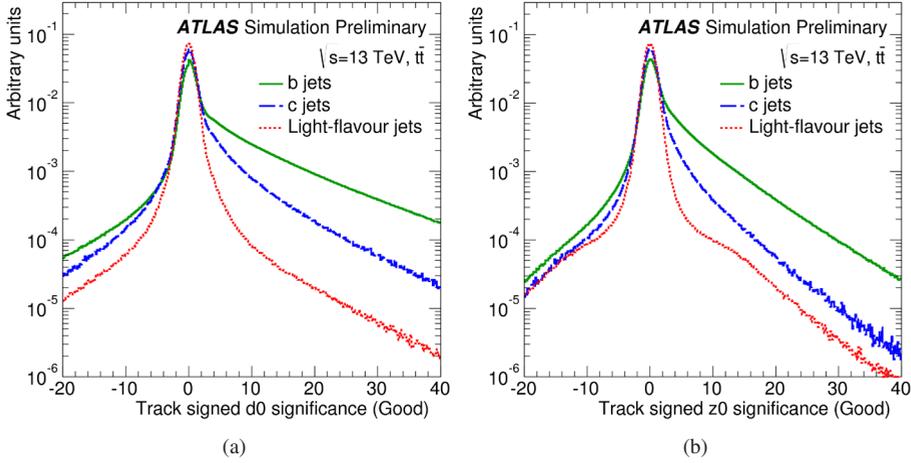


Figure 4. Impact parameter significances (d_0 , z_0) for b , c and light-jets [4].

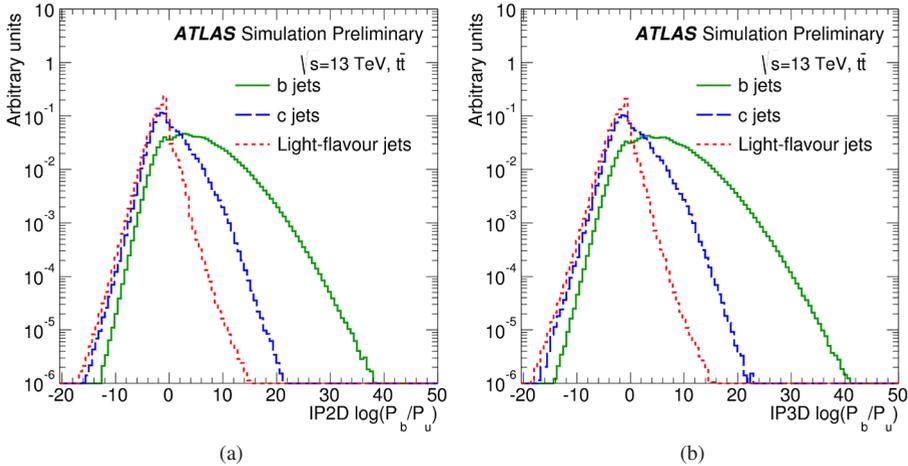


Figure 5. Impact parameter b -tagging algorithms (IP2D, IP3D) for b , c and light-jets [4].

3.2 Performance in 2016

The three types of algorithms provide input to a multivariate classifier (MV). In Run-2 this classifier is a boosted decision tree (BDT) trained to discriminate b -jets from light-jets. The background samples of light-jets can contain an admixture of charm and light-jets in order to improve the charm-jet rejection. The discrimination performance is shown in Figure 8a for the MV2 algorithm with a 10% charm-jet admixture (MV2c10). The performance between the Run-1 MV algorithm (MV1) and the Run-2 MV algorithm (MV2) has seen an improvement of 10%, primarily driven by the inclusion of the IBL. Between 2015 and 2016, retraining of the b -tagging classifier has further improved the

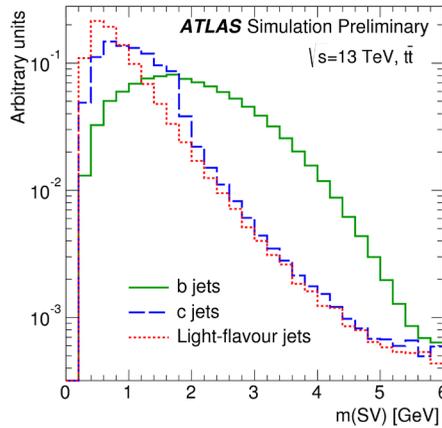


Figure 6. Invariant mass of the displaced vertex identified with the SV algorithm [5].

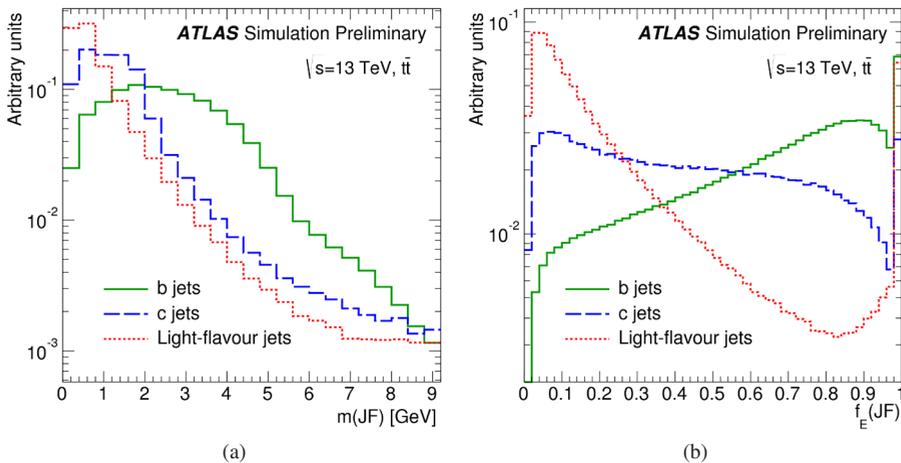


Figure 7. Distributions showing the invariant mass of all tracks associated with the identified displaced vertices using the JetFitter algorithm and the energy fraction of tracks identified with the JetFitter algorithm compared to the jet [4].

charm-jet rejection by around 40% at the 77% working point. The charm-jet rejection rate as a function of b -tagging efficiency is shown in Figure 8b. Jets which have been selected using a top-pair event selection are shown in Figure 8c comparing the performance of the b -tagging weight (MV2c10) in simulation and data, where a good agreement can be observed.

4 Calibration Techniques

Calibrations of b -tagging algorithms derive a correction factor applied to simulated data to match the b -tagging efficiency measured in data. This correction factor is dependent on the true flavour

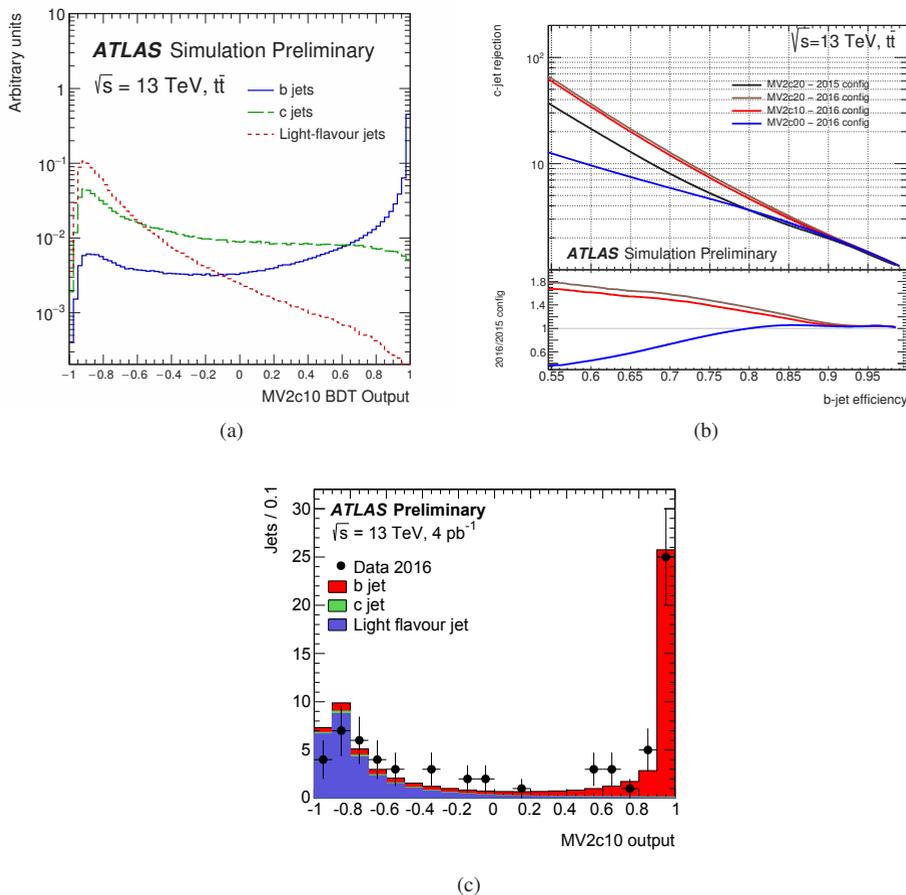


Figure 8. The performance of the MV2c10 algorithm is shown for different jet flavours [5], a comparison between the MV2c algorithms trained in 2015 and trained in 2016 [5] and the performance of the MV2c10 algorithm compared to 2016 data [7].

of the simulated jet and the jet kinematics. A correction factor is derived both for true *b*-jet which are correctly identified as well as for light and charm-jets which are incorrectly *b*-tagged. In order to derive these scale factors, the performance of the *b*-tagging algorithm needs to be measured using events in data which are enriched in a particular jet flavour whilst correctly modelling the backgrounds.

To measure the *b*-tagging efficiency, events are measured using dileptonically decaying top-pair events as these events are pure in *b*-jets using the Run-1 combinatorial likelihood method [6]. Events with one electron, one muon and two or three jets are used to measure the efficiency. Events with two same-flavour leptons are used to derive data-driven corrections to the *Z*+jets and top-pair samples. A combinatorial likelihood method is applied to extract the *b*-tagging efficiency in data by applying a two-dimensional maximum likelihood fit as a function of leading and sub-leading jet p_T , which allows the method to include correlations between the two jets. The current *b*-tagging efficiency scale factors calculated with this method as a function of jet p_T is shown in Figure 9 where the scale factors agree

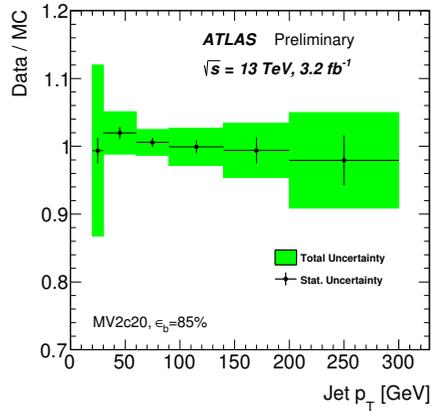


Figure 9. Scale factors derived from the preliminary Run-2 b -tagging efficiency measurement on dilepton top-pair events in 2015 data [9].

with unity within errors. The intermediate jet range has a total uncertainty of 2-4% and is dominated by the effect of top-pair modelling.

Charm and light-jet calibrations are underway using the Run-2 dataset. The techniques used in Run-1 are being applied to the Run-2 dataset [8]. The charm-jet calibration uses the D^* method, where a sample of events enriched in charm-jets are selected by reconstructing the decay chain of $D^* \rightarrow D_0(\kappa\pi)\pi_s$. The yields before and after applying a b -tagging requirement comes from some fraction of true b -jets correctly selected with an efficiency and some fraction of charm-jets incorrectly selected with an efficiency. By fitting the proper lifetime of the events, the contamination from b -hadron decays can be estimated and used to extract the charm-tagging efficiency. The light tagging efficiency has become harder to estimate since the performance of the b -tagging algorithms has improved in Run-2 compared to Run-1. The negative-tag method was used in Run-1 where parameters such as the signed impact parameter significance and the signed decay length significance have their sign requirements inverted. This enhances the fraction of light-jets selected by the IP and SV algorithms. The inclusive distribution of the negatively tagged light-jets has a similar shape to the mis-tagged light-jets, and once the contamination from heavy flavour and long-lived particles is accounted for, can be used to extract the light-jet tagging efficiency.

5 Concluding Remarks

The b -tagging performance in Run-2 shows a marked improvement compared to Run-1. The inclusion of the IBL has improved the track reconstruction efficiency and accounts for the majority of the 10% improvement in b -tagging efficiency. The optimisation of tracking algorithms and b -tagging algorithms has improved the performance at high p_T and the latest 2016 optimisation has improved the charm-jet rejection factor by around 40%. Efforts are currently underway to calibrate the b -tagging algorithms with the 13 TeV data collected in 2015 and 2016. Preliminary b -tagging efficiencies have been successfully extracted from 13 TeV data and charm-tagging efficiency measurements are being finalised.

References

- [1] ATLAS Collaboration, The ATLAS experiment at the CERN Large Hadron Collider, JINST 08003 (2008)
- [2] ATLAS Collaboration, The Optimization of ATLAS Track Reconstruction in Dense Environments, ATL-PHYS-PUB-2015-006 (2015)
- [3] ATLAS Collaboration, Impact parameter resolution Run-1 vs. Run-2, ITDR-2015-007 (2015)
- [4] ATLAS Collaboration, Expected performance of the ATLAS b -tagging algorithms in Run-2, ATL-PHYS-PUB-2015-022 (2015)
- [5] ATLAS Collaboration, Optimisation of the ATLAS b -tagging performance for the 2016 LHC Run, ATL-PHYS-PUB-2016-012 (2016)
- [6] ATLAS Collaboration, Calibration of b -tagging using dileptonic top pair events in a combinatorial likelihood approach with the ATLAS experiment, ATLAS-CONF-2014-004 (2014)
- [7] ATLAS Collaboration, b -tagging performance plots in a $t\bar{t}$ -dominated sample from early 2016 ATLAS data, FTAG-2016-001 (2016)
- [8] ATLAS Collaboration, Calibration of the performance of b -tagging for c and light-flavour jets in the 2012 ATLAS data, ATLAS-CONF-2014-046 (2014)
- [9] ATLAS Collaboration, b -tagging calibration plots using dileptonic $t\bar{t}$ events produced in pp collisions at $\sqrt{s} = 13$ TeV and a combinatorial likelihood approach, FTAG-2016-002 (2016)