



ATLAS NOTE

ATL-PHYS-PUB-2016-012

13th June 2016



Optimisation of the ATLAS b -tagging performance for the 2016 LHC Run

The ATLAS Collaboration

Abstract

The performance of the current ATLAS b -tagging algorithms has been optimised ahead of the 2016 Run-2 data taking. The algorithms used to identify jets containing b -hadrons are briefly described and the improvements relative to the 2015 LHC run are highlighted. Particular emphasis is given to the optimisation and performance of the multivariate b -tagging algorithm, MV2. Overall for a b -tagging efficiency of 77%, the light-quark rejection factor has been held constant while increasing the charm rejection factor by 40%. The performance at high jet p_T has also been improved.

Note: Updated 28/2/17 to correct LLR discriminant formula on page 4.



1 Introduction

The identification of jets containing b -hadrons (b -tagging) is an important ingredient in the physics program of the ATLAS experiment [1]. This note is intended to provide an update on the main b -tagging algorithms with respect to those used for the 2015 LHC, which are documented in Ref. [2]. The main modifications for the 2016 run are briefly described in this note and a comparison is presented between the performance of the new and the 2015 version of the algorithms.

The inputs to the b -tagging algorithms are the trajectories of charged particles (tracks) reconstructed in the Inner Detector (ID) [3]. The ATLAS inner detector system, consisting of a silicon pixel detector, a silicon micro-strip detector (SCT) and the straw tubes of the transition radiation tracker (TRT), is used to measure the trajectories and momenta of charged particles in the region¹ $|\eta| < 2.5$. The ID surrounds the beam pipe, and is located inside a solenoid magnet that provides a 2 T axial magnetic field. A series of upgrades were made to the ATLAS detector between Run 1 and Run 2, including improvements to the ID. In order to have better track and vertex reconstruction performance at the higher luminosities expected during Run 2, a fourth pixel layer, the insertable B -layer (IBL), has been added to ATLAS [4]. This inclusion leads to significantly improved b -tagging performance, mostly in the low and medium jet p_T region, compared to that achieved at Run 1 [2, 5].

This note is structured as follows. Section 2 gives an overview of the samples employed for this study, the object definitions and the track selection used. Section 3 briefly summarizes the current basic b -tagging algorithms used by the ATLAS collaboration. The outputs resulting from these basic tagging algorithms are used as inputs to a multivariate-based algorithm (MV2) described in Section 4. Finally, conclusions are reported in Section 5.

2 Event Samples and Selections

The results presented in this note are obtained using a Monte Carlo sample of $t\bar{t}$ events produced in proton-proton collisions generated at a centre-of-mass energy (\sqrt{s}) of 13 TeV and simulated with Powheg [6] interfaced with Pythia6 [7] for parton shower and hadronization. The CT10 [8] parton distribution function set is used in the simulation. The EvtGen [9] program is used to model the decays of b - and c -hadrons. In addition, minimum bias interactions are produced with Pythia8 [10] and additional inelastic pp collisions accompanying the hard-scatter interactions (pileup) are included in the simulation. The simulation is generated with an average of 20 pileup interactions per event (μ). Only $t\bar{t}$ decays with at least one lepton from a subsequent W -boson decay are included in the simulation. In Section 3.1, $Z' \rightarrow t\bar{t}$ events simulated with Pythia8 [10] at a Z' resonance mass of 4 TeV are also used. The NNPDF2.3 [11] parton distribution function set is used in the simulation and, as for the $t\bar{t}$ case, the EvtGen program models the decays of b - and c -hadrons. The propagation of particles through the detector and the detector response are modelled using GEANT4 [12, 13].

The ATLAS reconstruction software is run on the simulated samples. A common selection used as

¹ ATLAS 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 upwards. Cylindrical coordinates (r, ϕ) are used in the transverse plane, being ϕ the azimuthal angle around the z -axis. The pseudo-rapidity is defined in terms of the polar angle θ as $\eta = -\ln \tan(\theta/2)$ while $\Delta R^2 = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

a basis for all performance studies is then applied. The hard scatter event vertex is chosen as the one with the largest sum of squared transverse momenta (p_T) of the associated tracks. Jets are reconstructed by clustering energy deposits in the calorimeter with the anti- k_t algorithm [14] with a radius parameter of 0.4, where clusters are calibrated at the EM-scale and the hadronic scale is obtained through a p_T and η dependent correction factor that brings the energy scale in Monte Carlo close to that measured in data [15].

In this note, only jets with p_T above 20 GeV and $|\eta| < 2.5$ are selected. Furthermore, in order to reject jets originating from pileup interactions, a requirement on the output of the Jet Vertex Tagger algorithm (JVT) [16] is used. This algorithm is based on track-based variables combined in a multivariate discriminant and is designed to suppress pileup jets in such a way that the resulting efficiency for jets from the hard scattering collision is stable as a function of the number of primary vertices in the event. Jets with $p_T < 60$ GeV and jet $|\eta| < 2.4$ are discarded if the JVT score is less than 0.59. This choice results in 92% efficiency for jets from the hard-scatter interaction and 2% residual observed rate from pileup jets.

Finally, in order to label jets as being b , c - or light-flavour jets in the simulation for training and performance studies, the jets are matched to the b and c -hadrons within a cone of radius $\Delta R=0.3$ around the jet axis. The labelling procedure is performed by first searching for b -hadrons within the cone centered on the axis of the jet. If no b -hadron is found, the algorithm is repeated for c -hadrons, and for τ leptons. In addition, the transverse momenta of b - and c - hadrons are required to be larger than 5 GeV. Ultimately the remaining jets which do not fall in the previous categories are considered as light-flavour jets. Furthermore, a requirement on the possible overlap between jets and leptons is applied: jets are removed if they overlap with a truth electron or muon from W/Z boson decays in a cone of $\Delta R < 0.3$.

The tracks used in the b -tagging algorithms are associated to jets reconstructed in the calorimeter by using the angular separation ΔR between the track and the jet axis. The ΔR requirement varies as a function of jet p_T , being wide for low p_T jets and narrower for high p_T jets which tend to be more collimated. For instance, at 20 GeV, it is 0.45 while for more energetic jets with a p_T of 150 GeV the ΔR cut is 0.26 [5]. Further track selection depends on the specific algorithm and will be described in Sections 3.1, 3.2 and 3.3.

3 Basic Algorithms

The identification of b -quark jets in ATLAS is based on distinct strategies encoded in three basic b -tagging algorithms: impact parameter-based algorithms described in Section 3.1, an inclusive secondary vertex reconstruction algorithm described in Section 3.2 and a decay chain multi-vertex reconstruction algorithm described in Section 3.3. The output of these b -tagging algorithms are combined in a multivariate discriminant (MV2) that is the output of the default algorithm used by ATLAS, described in Section 4, which provides the best separation between the different jet flavours. These sections also describe the algorithmic changes and improvements made since the release of the previous note [2].

3.1 IP2D and IP3D: The Impact Parameter based Algorithms

The long lifetime of hadrons containing a b -quark (~ 1.5 ps, $c\tau \sim 450$ μm) can be exploited to build lifetime-based tagging algorithms. The typical b -hadron topology is characterized by at least one vertex

displaced from the point where the hard-scatter collision takes place. The transverse impact parameter (d_0) is defined as the distance of closest approach in the r - ϕ plane of the track to the primary vertex while the longitudinal impact parameter ($z_0 \sin \theta$) is defined as the distance of the track to the primary vertex in the longitudinal plane at the point of closest approach in r - ϕ . Due to the long lifetime of b -hadrons, tracks generated from b -hadron decay products tend to have large impact parameters enabling their contribution to be separated from the contribution of tracks from the primary vertex. It is possible to assign a sign to the impact parameter by determining, relatively to the jet direction, if the primary vertex is in front of the secondary vertex or behind. Secondary vertices behind the primary are normally due to background and are used to help to calibrate the light quark tag rate [5]. The impact parameter taggers, IP2D and IP3D, generally referred to as IP in what follows, are based on this sign convention.

The track selection applied for the impact-parameter algorithm is based on the following requirements:

- track p_T above 1 GeV;
- $|d_0| < 1$ mm and $|z_0 \sin \theta| < 1.5$ mm;
- seven or more silicon hits, with at most two silicon holes, at most one of which is in the pixel detector, where a hole is defined as a hit expected to be associated with the track but not present [17].

The IP2D tagger makes use of the transverse impact parameter significance, d_0/σ_{d_0} , as discriminating variable whereas IP3D uses both the transverse and the longitudinal impact parameter significance, $z_0 \sin \theta/\sigma_{z_0 \sin \theta}$, in a two-dimensional template to account for their correlation. Probability density functions (PDF) obtained from reference histograms for the transverse and longitudinal impact parameter significances are derived from Monte Carlo simulation. They are separated into exclusive categories that depend on the hit pattern of a given track to increase the discriminating power and used to calculate ratios of the b - and light flavour jet probabilities. This computation is performed on a per-track basis. Different PDFs are used for the various track categories and for each jet-flavour hypothesis (b , c , light). The track categorisation exploited in the algorithm can be found in Table 1 and the categories are determined in a $t\bar{t}$ sample. By knowing whether the hit is expected or not in the innermost (IBL) and in the next-to-innermost (b-layer) layers, a quality criterion is defined. The expectations for which hits should be present are derived from the knowledge of the detector geometry and inactive module maps. A hit is considered *split* if, during the ambiguity solver stage of the track reconstruction at the pattern recognition level [17], it was identified as being created by multiple charged particles (the decision is made by looking at its typical cluster shape [18, 19]). If a Pixel or SCT hit is associated to more than one track, and not already marked as *split*, it is regarded as *shared*.

Afterwards, a log-likelihood ratio (LLR) discriminant is computed as the sum of the per-track contributions, $\sum_{i=1}^N \log \left(\frac{p_b}{p_u} \right)$, where N is the number of tracks of a given jet and p_b , p_u are the template probability density functions (PDF) for the b - and light-flavour jet flavour hypotheses, respectively assuming no correlation among the various tracks contributing to the sum. In addition to the LLR separating b - and light-flavour jets, LLR functions are also computed to separate b - from c -jets, or c -jets from light-flavour jets. The distributions for the transverse and longitudinal impact parameter significances are shown in Figure 1 for tracks from light, c - and b -jets. It can be seen that the distribution of d_0/σ_{d_0} for light-flavour jet features a small exponential tail at high values that is formed by tracks coming from Λ , K_S decays, hadronic interactions or from photon conversions, while the tail at high d_0/σ_{d_0} for b -jets is an effect of the long b -hadrons lifetime. In addition, z_0/σ_{z_0} shows a large tail that is typically generated by tracks from pileup jets (very close-by pileup vertices).

#	Category	Fractional contribution [%]		
		<i>b</i> -jets	<i>c</i> -jets	light-jets
0	No hits in first two layers; expected hit in IBL and b-layer	1.9	2.0	1.9
1	No hits in first two layers; expected hit in IBL and no expected hit in b-layer	0.1	0.1	0.1
2	No hits in first two layers; no expected hit in IBL and expected hit in b-layer	0.04	0.04	0.04
3	No hits in first two layers; no expected hit in IBL and b-layer	0.03	0.03	0.03
4	No hit in IBL; expected hit in IBL	2.4	2.3	2.1
5	No hit in IBL; no expected hit in IBL	1.0	1.0	0.9
6	No hit in b-layer; expected hit in b-layer	0.5	0.5	0.5
7	No hit in b-layer; no expected hit in b-layer	2.4	2.4	2.2
8	<i>Shared</i> hit in both IBL and b-layer	0.01	0.01	0.03
9	At least one <i>shared</i> pixel hits	2.0	1.7	1.5
10	Two or more <i>shared</i> SCT hits	3.2	3.0	2.7
11	<i>Split</i> hits in both IBL and b-layer	1.0	0.87	0.6
12	<i>Split</i> pixel hit	1.8	1.4	0.9
13	<i>Good</i>	83.6	84.8	86.4

Table 1: Description of the track categories used by IP2D and IP3D together with the fraction of tracks in each category for jets in $t\bar{t}$ events. The order of the layers is explained in the text. The categories further down in the list can be more inclusive than the first ones because, when a category is not fulfilled, the next one is evaluated.

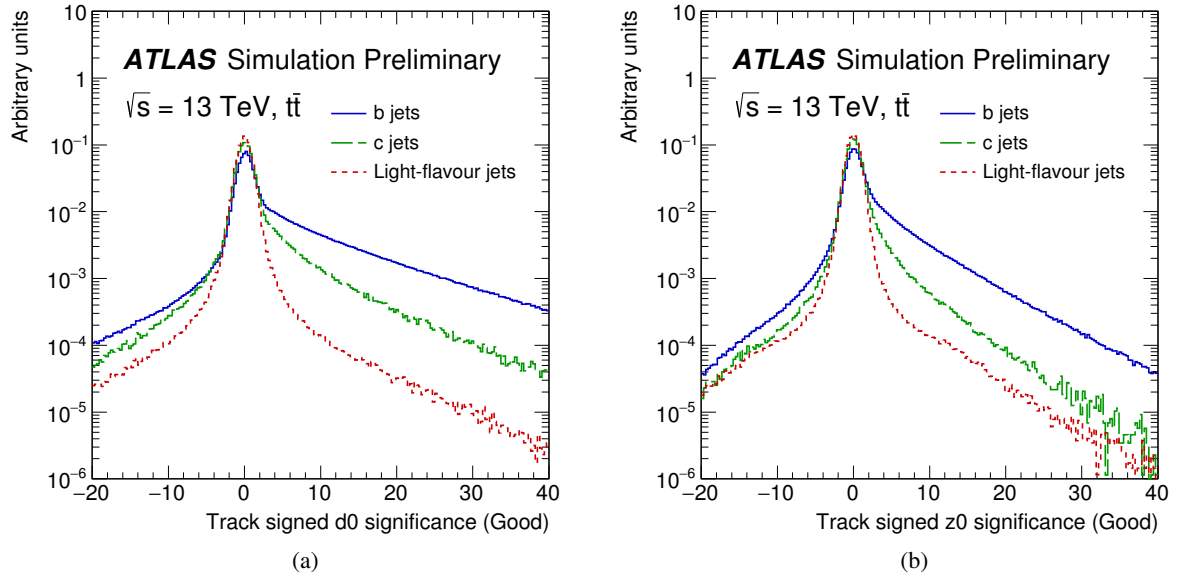


Figure 1: The transverse (a) and longitudinal (b) signed impact parameter significance of tracks in $t\bar{t}$ events for *b* (solid blue), *c* (dashed green) and light flavour (dotted red) jets for the *Good* category as defined in Table 1.

The distributions of the final LLR discriminant for IP2D and IP3D are shown in Figures 2 (a) and 2 (b) for the b - versus light-flavour separation.

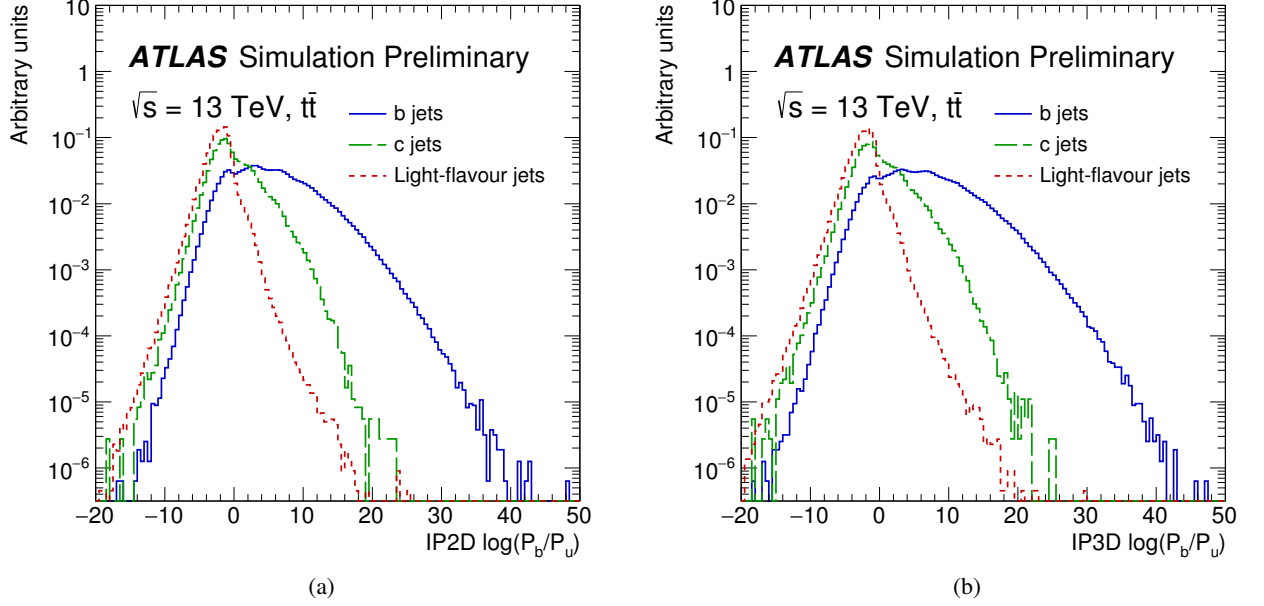


Figure 2: The log-likelihood ratio for the IP2D (a) and IP3D (b) b -tagging algorithm for b - (solid blue), c - (dashed green) and light-flavour (dotted red) jets in $t\bar{t}$ events. The log-likelihood ratio shown here is computed as ratio of the b - and light-flavour jet hypotheses. If no tracks are found in the jet, a large negative value that is not indicated in the plot is assigned as the algorithm output.

Correctly measuring the d_0/σ_{d_0} and z_0/σ_{z_0} shapes for even classes with a small track population can significantly improve the rejection because even a small class of tracks can significantly contribute to the already very small fake rate. This is shown in Figure 3 that illustrates the distributions of the transverse and longitudinal impact parameter significances for the fourth category in Table 1 (no hits in IBL while expected). The comparison with the templates for the *Good* category shown in Figure 1 illustrates the difference of the distributions.

Several refinements in the algorithm have been introduced for the new version of the IP tagger compared to the version described in [2]. They are listed below.

- The requirement on the number of pixel hits is relaxed from at least two to at least one in the pixel detector. The previous requirement induced some inefficiency in the high b -jet p_T region because a significant fraction of high p_T b -hadrons decay after the IBL ($R=3.3$ cm) and the b-layer ($R=5.5$ cm). The new requirement increases the performance at high b -jet p_T , while not significantly affecting the low-medium p_T regime.
- It is found that by ignoring tracks originating from conversions, Λ and K_S decays, material interactions as identified by the SV algorithm described in Section 3.2, a sizeable gain of 15% in light-flavour jet rejection for a b -jet efficiency working point of 77% is achieved.

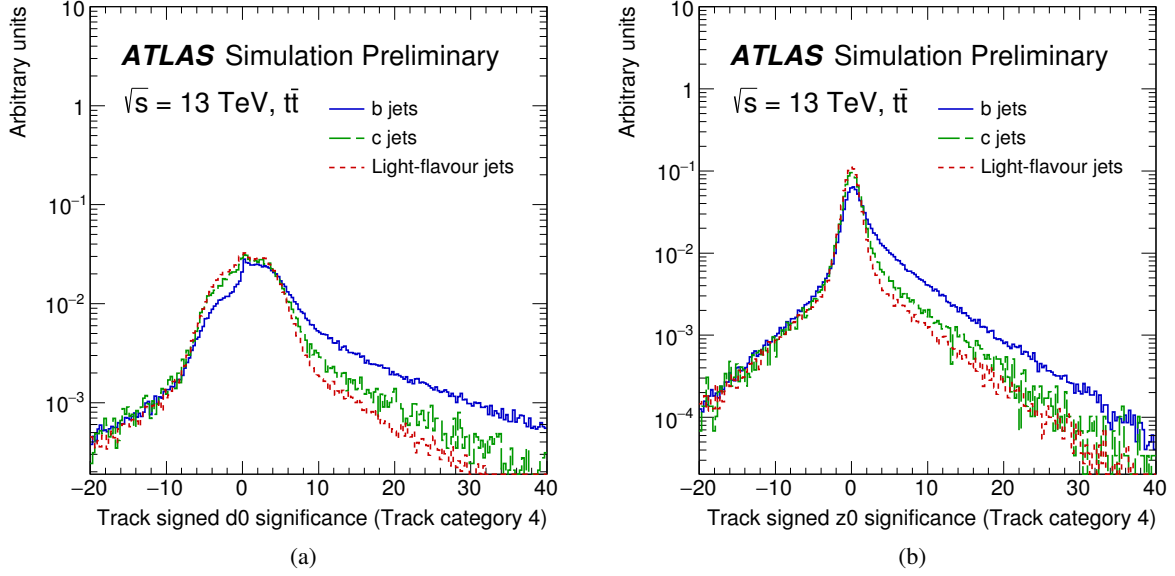


Figure 3: The transverse (a) and longitudinal (b) signed impact parameter significance of tracks in $t\bar{t}$ events with b (solid blue), c (dashed green) and light flavour (dotted red) jets for track category 4 (no hit in IBL while expected) as defined in Table 1.

- Reference histograms are produced with a mixture of $t\bar{t}$ and $Z' \rightarrow t\bar{t}$ (50%/50%) for the two categories with no hits in the first two layers despite them being expected (categories 0 and 1 in Table 1). Tracks from b -jets in $t\bar{t}$ events (low-medium p_T range) mostly populate categories where hits in the first two innermost layers are present. However, b -jets (or the b -hadrons) from $Z' \rightarrow t\bar{t}$ events are characterised by a harder p_T spectrum, hence their associated tracks tend to start outside the IBL/ b -layer since the corresponding b -hadron decay occurs beyond these layers and they populate the categories where no hits in the inner layers are present. This allows us to overcome the limited statistics available in these categories in the $t\bar{t}$ -only sample. Figure 4 shows the variation of the fraction of tracks in each of the 14 categories as a function of jet- p_T . By using a sample with higher p_T b -jets, it is therefore possible to significantly enhance the number of tracks in the aforementioned categories.

Figure 5 shows the distributions of the average multiplicity of tracks from b -hadron decays including tracks from subsequent c -hadron decays and tracks stemming from the fragmentation process (the charged multiplicity at generator level is around 5). The distribution in Figure shows the average multiplicity of tracks directly from b -hadron and from subsequent c -hadron decays selected for the IP algorithm as a function of the jet p_T . The contribution of tracks from fragmentation inside b -jets is observed to be dominant in the medium-high jet p_T regime (> 200 GeV) compared to tracks from b -hadron decays and therefore represents an important source of background inside the jets. Figure 6 shows the light (a) and c -jet rejection (b) as a function of the b -jet efficiency for the IP3D tagger when comparing the 2016 configuration to the 2015 version. The improvement in light-flavour jet rejection is mostly present in the medium-low b -jet efficiency region (below 75%). Gains of approximately 7% in light-flavour jet rejection and 3% in c -jet rejection are observed at the 77% b -jet efficiency working point. The small wiggles

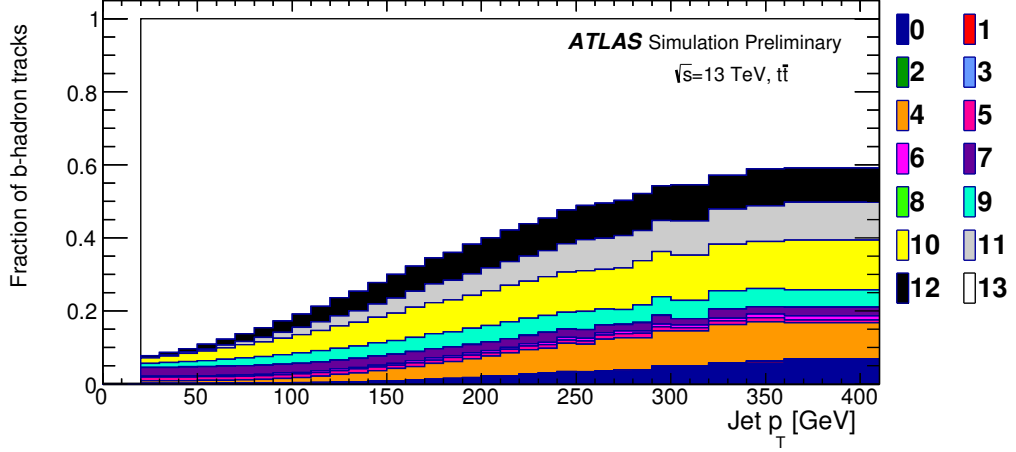


Figure 4: Fraction of tracks from b -hadron decays as a function of the jet p_T for $t\bar{t}$ events for IP categories reported in Table 1.

characterising the light-flavour jet rejection curve are due to jets with no tracks surviving the IP selection such that they carry a default value for the likelihood discriminant.

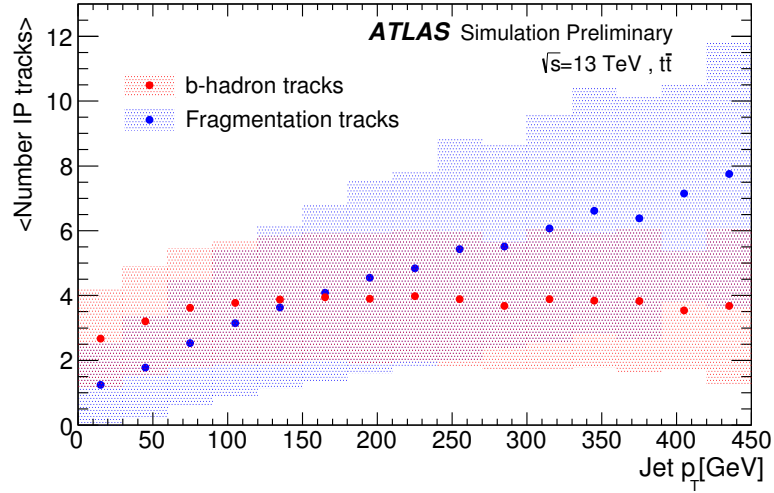


Figure 5: Average number of b -hadron and jet fragmentation tracks selected for the IP algorithm as a function of the jet p_T . The shaded band around the two contributions represents the RMS for each p_T bin.

3.2 Secondary Vertex Finding Algorithm: SV

The secondary vertex finding algorithm [20] explicitly reconstructs an inclusive displaced secondary vertex within the jet. All track pairs within a jet are tested for a two-track vertex hypothesis. Any found two track vertices are rejected if they were likely to originate from the decay of a long-lived particle (e.g.

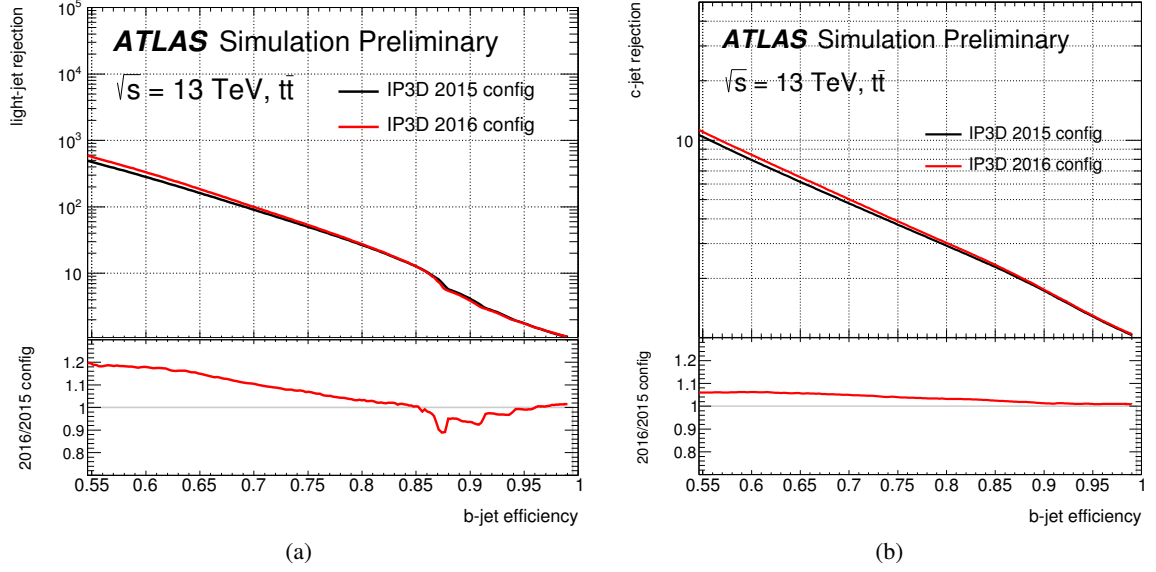


Figure 6: Light-flavour(a) and c -jet (b) rejection as a function of the b -jet efficiency for the IP3D algorithm comparing the current configuration (2016 config) with the previous one (2015 config).

K_s or Λ), photon conversions or hadronic interactions with the detector material. If a two-track vertex remains, a new vertex is then fitted with all tracks from the accepted two-track vertices, where outliers are iteratively removed from this set of tracks. Several refinements in the selections of tracks and two-track vertices have been made in order to improve the performance of the algorithm, especially against pileup contamination and for high- p_T jets.

Each selected track is required to have at least seven silicon hits (sum of pixel and SCT hits) and at most one shared hit in these two detectors. The χ^2 per degree of freedom of the track fit is required to be less than 3 [17]. In addition to these track quality selections, extra track requirements are used to improve the performance of the algorithm for the 2016 LHC run:

- Highly-energetic jets (with a $p_T > 300$ GeV), as illustrated in Figure 5, suffer from high multiplicity of tracks produced in jet fragmentation which increases the probability of reconstructing fake vertices. In order to reduce the number of fake vertices, tracks are ordered according to their p_T and at most 25 tracks with largest p_T are used in the secondary vertex reconstruction. With an average charged particle multiplicity from b -hadron decays of about 5, this selection is inclusive enough to retain all reconstructed tracks from any b -hadron decay, including the case of jets containing a gluon splitting to $b\bar{b}$.
- In addition, the reconstruction of tracks associated to jets in the high pseudorapidity region ($|\eta| > 1.5$) suffers from an increasing amount of detector material leading to worse track parameter resolution and an increased rate of hadronic interactions. Additional track cleaning is applied for these jets to mitigate the negative influence on the secondary vertex finding efficiency. To increase the quality of the selected tracks, the minimal number of required hits in the silicon detectors is increased by one for tracks with $|\eta| > 1.5$.

- In order to reduce the impact of pileup, tracks with low d_0/σ_{d_0} (<2) and high z_0/σ_{z_0} (>6) are removed, again reducing the number of fake vertices.

As previously mentioned, the set of tracks passing the track selections described above are used to form all possible two-track vertices. The two-track vertex candidates are required to be significantly displaced from the primary interaction vertex. Each track of the pair should exceed a significance of the distance between the track and the primary vertex of 2, and the sum of the two IP significances of the tracks in the two-track vertex should be higher than 2. Finally the χ^2 of the fitted tracks should be less than 4.5. In order to additionally reduce fake vertices and material interactions, extra cleaning selections on two-track vertices are considered for the 2016 LHC run:

- since secondary tracks cannot have associated detector hits with radii smaller than the production vertex radius, the track pixel hit pattern is required to correspond to the reconstructed vertex radius for both tracks;
- two-track vertices with an invariant mass greater than 6 GeV are removed as those are not likely to be from b or c hadron decays.
- Some of the reconstructed 2-track vertices come from K_s , Λ decays, photon conversions and hadronic interactions with the detector material; the tracks coming from K_s and Λ decays can be rejected by using the mass of the vertex. The invariant mass spectra of $\pi^+\pi^-$ and $p\pi$ are used to identify the peaks due to K_s and Λ decays, charged tracks coming from such vertices are rejected by the algorithm.

Figure 7 shows the distribution of some of the properties of the reconstructed secondary vertex comparing vertices from b , c and light-flavour jets. Figure 8 displays the fraction of jets which have a reconstructed secondary vertex, as a function of jet- p_T (a) and η (b) for b -jets, c -jets, and vertices in light-flavour jets. The increase of tracks from fragmentation in the high jet p_T region is the main reason for the performance degradation. The number of fake vertices is increasing with jet p_T in light jets, while the secondary vertex reconstruction efficiency for b - and c - jets slightly decreases with jet p_T . Figures 8 (c) and 8 (d) show the secondary vertex reconstruction rate for light-flavour jets when employing the previous track selection documented in [2] and the updated requirements. Notably, the new track selection results in significantly smaller fake rate of SV reconstruction for light-flavour jets both as function of jet p_T and η .

3.3 Decay Chain Multi-Vortex Algorithm: JetFitter

The decay chain multi-vertex reconstruction algorithm, JetFitter [21], exploits the topological structure of weak b - and c -hadron decays inside the jet and tries to reconstruct the full b -hadron decay chain. A Kalman filter is used to find a common line on which the primary vertex and the bottom and charm vertices lie, approximating the b -hadron flight path, as well as their positions. With this approach, the b - and c -hadron vertices, whenever resolution allows, can be resolved, even when only a single track is attached to any of them. No modification of this tagger has been made compared to the previous results reported in [2].

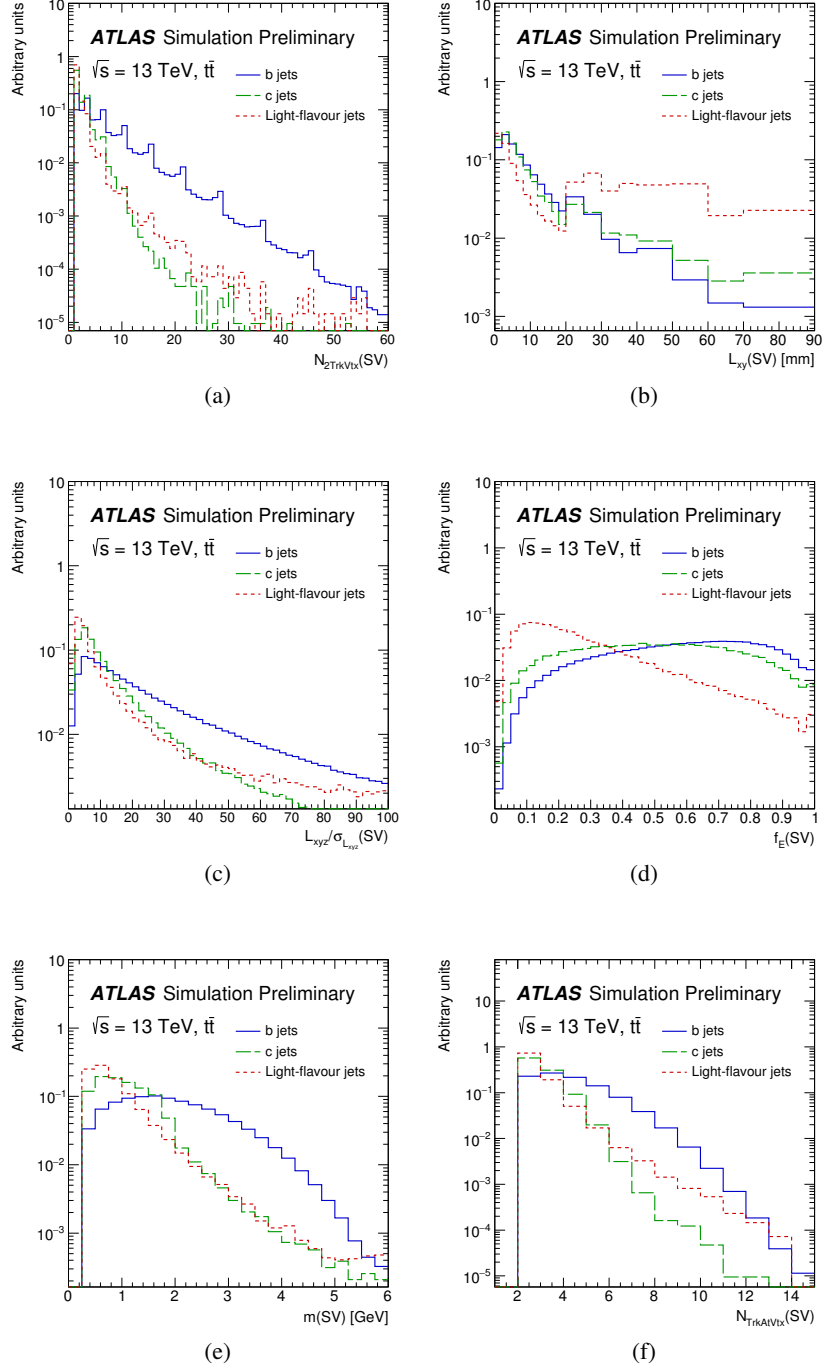
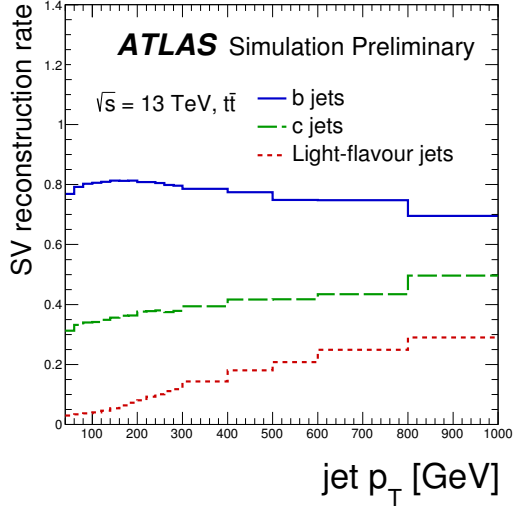
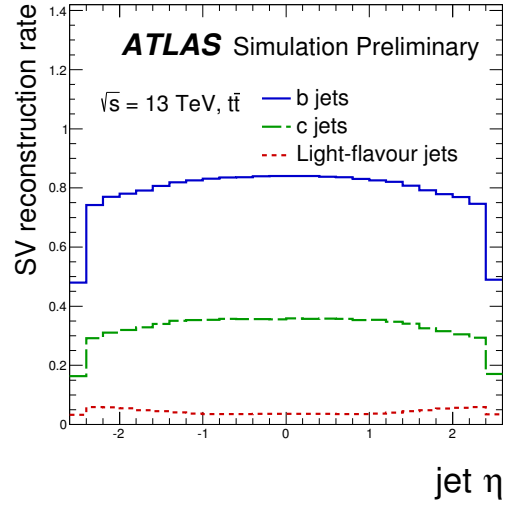


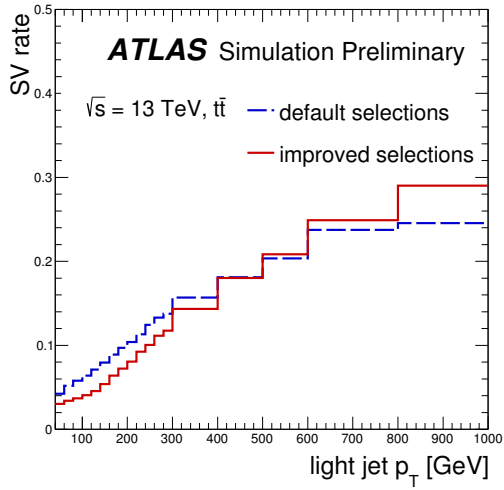
Figure 7: Properties of secondary vertices reconstructed by the SV algorithm for b - (solid blue), c - (dashed green) and light-flavour (dotted red) jets in $t\bar{t}$ events: the number of two-track vertices reconstructed within the jet (a), the transverse decay length (b), the 3D decay length significance (c), the energy fraction (d), defined as the energy of the tracks in the displaced vertex relative to the energy of all tracks reconstructed within the jet, the invariant mass (e) and the number of tracks (f) associated with the vertex. The increased rate of light-flavour jets at high transverse decay length values is due to residual material interactions.



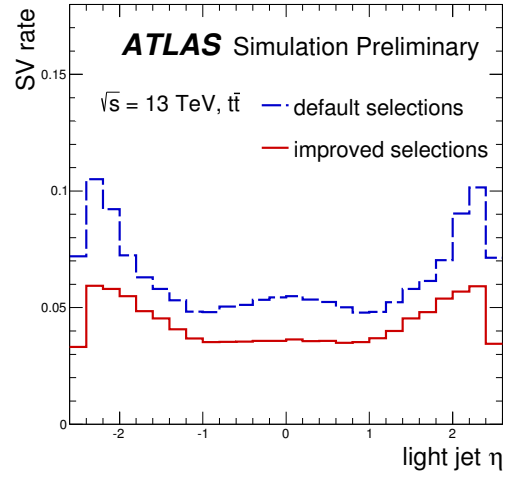
(a)



(b)



(c)



(d)

Figure 8: Secondary vertex reconstruction rate as a function of jet p_T (a) and η (b) for b -jets (blue), c -jets (dashed green), and light-flavour jets (dotted red). Fake rate of secondary vertices for light-flavour jets as a function of jet p_T (c) and η (d) when comparing the previous track selection and the current requirements described in Section 3.2.

4 Multivariate Algorithm: MV2

To achieve a better discrimination than any of the basic algorithms can exploit individually, a Boosted Decisions Tree (BDT) algorithm, using the ROOT Toolkit for Multivariate Data Analysis (TMVA) [22], is employed. It combines the output of the basic taggers described in Section 3.

4.1 Description of the algorithm and training procedure

The list of input variables that are found to be discriminating among the different flavour hypotheses have not been modified with respect to the previous version of the MV2 algorithm documented in Ref. [2].

The training of the multivariate classifier is performed on jets from $t\bar{t}$ events with b -jets (1 million) being considered as signal, and c - (0.5 million) and light-flavour jets (1 million) being considered as background. The performance is tested on an independent sample of 5 million $t\bar{t}$ events. The kinematic properties of the jets, namely p_T and $|\eta|$, illustrated in Figure 9, are included in the training in order to take advantage of the correlations with the other input variables. However, to avoid differences in the kinematic distributions of signal and background (light-flavour jets) being interpreted as discriminating by the training, the signal jets are reweighted in p_T and $|\eta|$ to match the spectrum of the light-flavour background jets. No kinematic reweighting is applied at the evaluation stage of the multivariate classifier.

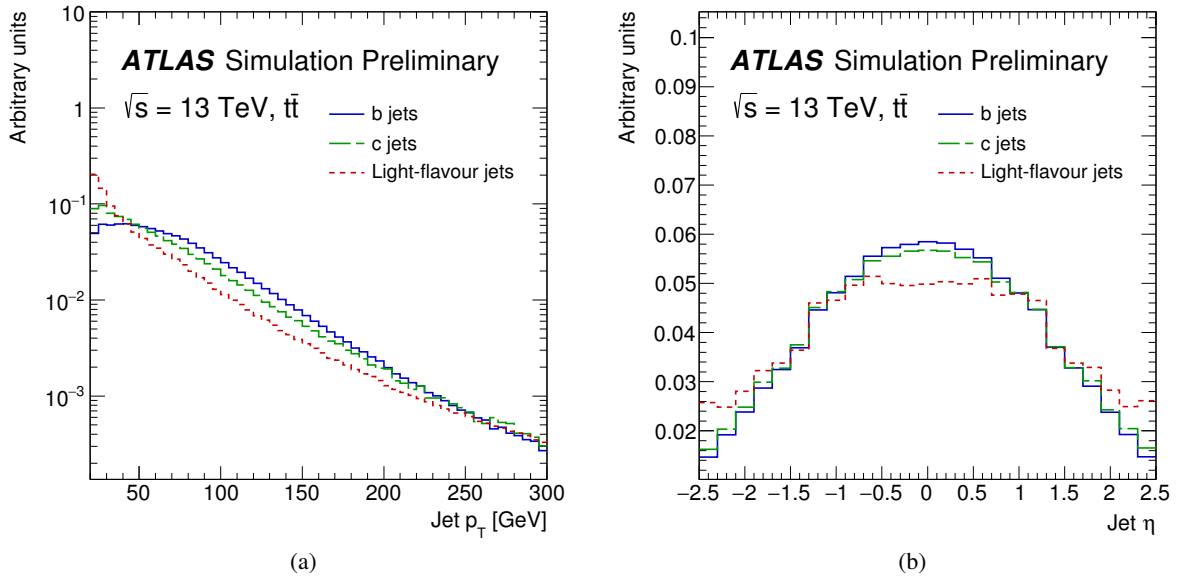


Figure 9: Kinematic distributions of p_T (a) and $|\eta|$ (b) for b (solid blue), c (dashed green) and light-flavour (dotted red) jets in $t\bar{t}$ events.

The new version of the MV2 algorithm constitutes a significant refinement of that documented in Ref. [2]. Each tagging algorithm can fail to produce a result for a jet (e.g. if it does not reconstruct a secondary vertex in the case of SV or JetFitter). In the previous version of the algorithm, jets failing to produce results in any of the algorithms were given a penalty weight (10^{-6}) in the training procedure (downgrading

procedure). This approach was found to be sub-optimal since such events are not properly taken into account in the training. In addition, the probability of reconstructing secondary vertices for different flavour hypotheses (b , c and light) is very useful information for the multivariate classifier to exploit in order to enhance the overall signal-to-background separation. In the current downgrading procedure, the penalty factor (10^{-6}) is applied if and only if all the three underlying tagger algorithms are found to be invalid. The gain in performance originating from this change is gauged by the $u(c)$ -jet rejection versus b -jet efficiency at the 77% b -jet efficiency working point: 40% gain in light-flavour jet rejection and 20% in c -jet rejection. The fraction of jets affected by this procedure, namely the ones not having valid output for all taggers is 0.6% in the case of light-flavour jets and 0.1% for c - and b -jets.

In the previous version of the algorithm, three MV2 variants were released, MV2c00, MV2c10 and MV2c20, where the names of the taggers indicate the c -jet fraction in the training, e.g. in MV2c20, the background sample is composed of 20% (80%) c - (light-flavour) jets. It is possible to modify the light versus c -jet rejection performance by changing the fraction of c -jets in the training. Given that the majority of physics analyses are presently more limited by the c - rather than the light-flavour jet rejection, the c -jet background fraction in the training has been chosen in such a way to enhance the charm rejection by keeping a similar light-flavour jet rejection compared to the previous approach. The c -jet fraction of the training for MV2c10 is therefore set at 7% such that the training is performed assigning b -jets as signal and a mixture of 93% light-flavour jets and 7% c -jets as background. In a second variant, denoted MV2c20, the c -jet fraction used is 15%; lastly, no c -jet contribution is present in the training is used for MV2c00.²

The training parameters for the BDT have also been optimized. The minimum node size (MNS) parameter defining the minimal fraction of jets in a leaf over the total number of jets in the training is reduced from 0.5% to 0.05% as the new downgrading procedure results in a larger training sample. The gain in performance for the light-flavour jet rejection as a function of the number of trees in the BDT forest or as a function of the typical depth of the trees (30) is also checked for fixed values of MNS. An additional test is performed on the number of trees, and although the number of trees and depth are intrinsically correlated, it is found that a non-negligible extra gain is brought by extending the number of trees from 400 to 1000. Lastly, a scan over the number of cuts in the training, namely the number of grid points over the variable range used in finding an optimal cut in node splitting, is performed and results are observed to be relatively insensitive to the specific value used for this variable. The final improvements in light- (c) -jet rejection at 77% b -jet efficiency working point brought by the optimization of the parameters of the BDT for the training is approximately 30% (20%).

4.2 Studies on the MV2 performance and training validations

The performance of the optimized MV2c00, MV2c10 and MV2c20 b -tagging algorithms is shown in Figure 10 for the light and c -jet rejection as a function of the b -jet efficiency in comparison to the 2015 MV2c20 configuration. It is noted that the current MV2c10 (2016 configuration) discriminant provides a similar light-flavour jet rejection (improvement of approximately 4% at 77% b -jet efficiency) to the 2015 MV2c20 configuration, but a significantly better c -jet rejection of +40% at the 77% b -jet efficiency working point. The 2016 MV2c20 setup provides even better charm rejection, but at the expense of a reduced light-flavour jet rejection, which may be detrimental to some analyses. The MV2c10 tagger

² It was decided to keep the same nomenclature MV2XX where X=00,10,20 for the three variants of the MV2 algorithm even though the value of XX no longer reflects the c -jet fraction used in the training sample.

algorithm has therefore been chosen as the standard b -tagging discriminant for 2016 analyses. Figure 11 shows the MV2c10 BDT output for the signal and background components. The Monte Carlo simulation has also been adjusted to better represent the data conditions expected in 2016 (pileup profile for 2016 data-taking), and this is found to have a small effect on the performance when re-training the MV2 algorithm (5-7% improved light-flavour jet rejection and 2-3% gain in c -jet rejection at 77% b -jet efficiency).

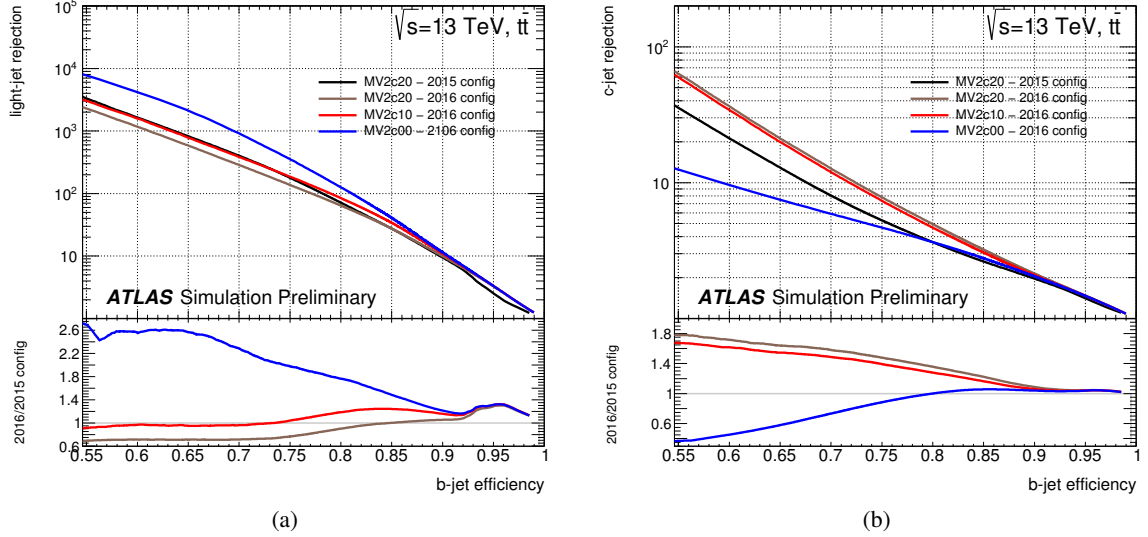


Figure 10: Light-flavour jet (a) and c -jet (b) rejection versus b -jet efficiency for the previous (2015 config) and the current configuration (2016 config) of the MV2 b -tagging algorithm evaluated on $t\bar{t}$ events. As explained in Section 4.2, MV2c00 denotes the MV2 algorithm where no c -jet contribution was present in the training. MV2c10 (MV2c20) denote the MV2 outputs where a 7% (15%) c -jet fractions was present in the background sample (for the 2016 configuration).

Operating points are defined by a single cut value on the MV2 output and are chosen to provide a specific b -jet efficiency on a $t\bar{t}$ sample. Table 2 shows the operating points defined for the recommended tagger b -tagging algorithm, MV2c10, with benchmark performance values.

BDT Cut Value	b -jet Efficiency [%]	c -jet Rejection	Light-jet Rejection	τ Rejection
0.9349	60	34	1538	184
0.8244	70	12	381	55
0.6459	77	6	134	22
0.1758	85	3.1	33	8.2

Table 2: Operating points for the MV2c10 b -tagging algorithm, including benchmark numbers for the efficiency and rejections rates. These values have been extracted from $t\bar{t}$ events, the main requirement being jet p_T above 20 GeV.

Figure 12 displays the comparison of the b -jet efficiency as a function of jet p_T between the current tagger MV2c10 and the 2015 MV2c20 algorithm. Figure 13 shows the comparison between the baseline 2015 configuration (MV2c20) and the current 2016 approach (MV2c10) for the light-flavour and c -jet rejection as a function of jet p_T and jet η . The distributions have been produced with a fixed cut at the 77% b -jet efficiency operating point. The light-jet rejection is slightly decreased in the high b -jet p_T region (>100

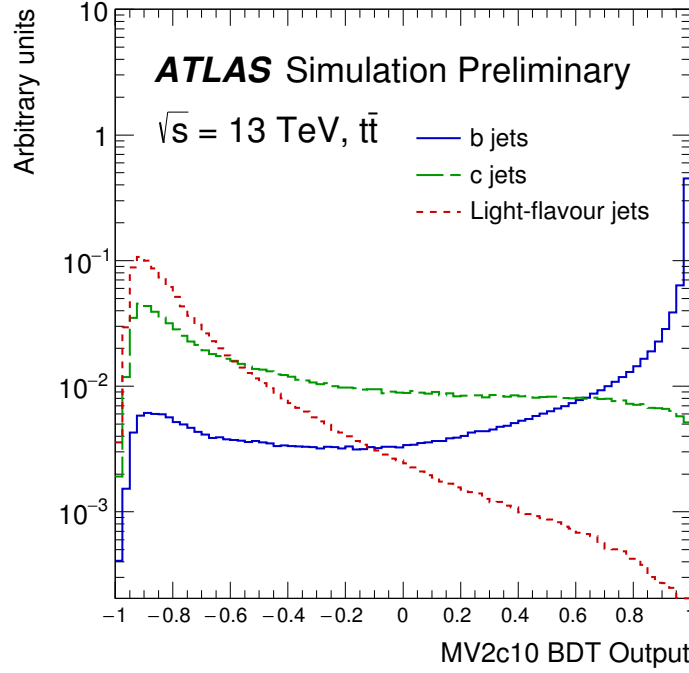


Figure 11: MV2c10 BDT output for b - (solid blue), c - (dashed green) and light-flavour (dotted red) jets evaluated with $t\bar{t}$ events.

GeV), but, at the same time, the b -jet efficiency is also increased in the same kinematic region. The reason for this is that the new MV2 training with the updated downgrading procedure results in a different effective tagging requirement as a function of jet p_T : at high jet p_T , the 77% working point is slightly worse as compared to the 2015 configuration. The b -jet efficiency distribution as a function of jet η has also been inspected: no major differences in efficiency are found when comparing the results in 2015 and 2016 MV2 trainings.

Figure 14 displays the comparison between the baseline 2015 configuration (MV2c20) and the current 2016 approach (MV2c10) for the light-flavour and c -jet rejection. In each bin of the p_T distribution, the b -tagging cut value has been chosen in such a way to yield a constant b -jet efficiency of 77%. For fixed b -jet efficiency, the new MV2 training shows a sizeable improvement in all kinematic regions.

Similarly to the rejection v s efficiency curves reported in Figure 10, the enhancement in c -jet rejection brought by the optimization of the MV2 multivariate classifier is also clearly visible over the full p_T and η range. Furthermore, as a consequence of the choice of the c -jet fraction in the training for MV2c10, the τ -rejection has increased by approximately a factor 2 with respect the 2015 algorithm.

Additional cross-validation checks on the final BDT configuration were performed in order to assess the impact of the usage of different training samples on the performance and to verify the effect of the finite Monte Carlo statistics for the training of the classifier. Cross-training tests have been performed in order to make sure that the performance is consistent when independent sub-samples are used as input to the multivariate classifier. It is observed that the tagging performance for c - and light-flavour jet rejection

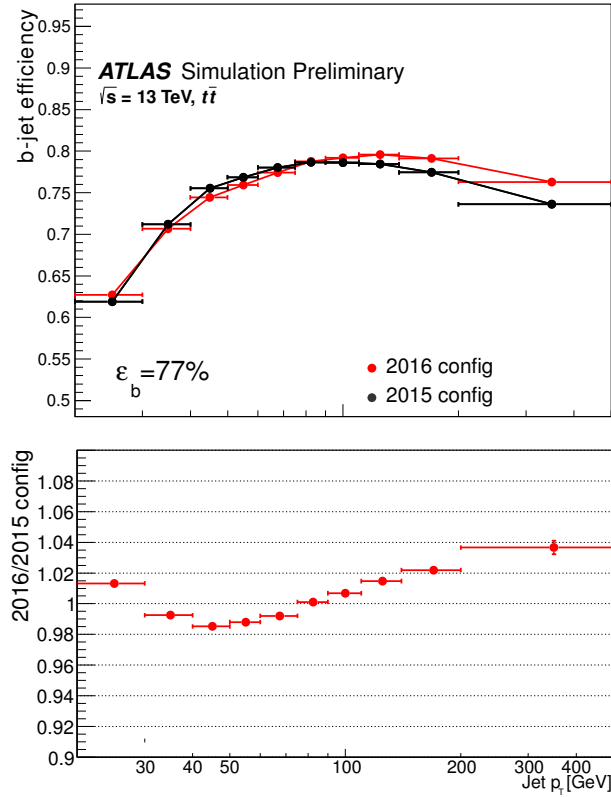


Figure 12: b -jet efficiency for the fixed cut working point with a b -jet efficiency of 77% as a function of the jet p_T for the comparison between the MV2c10 b -tagging algorithm employed for the 2016 analyses (denoted as 2016 config) and the previous version of the tagger, MV2c20 (2015 config).

is stable at the level of few % while the statistical uncertainty on the rejection values evaluated on the test sample is negligible.

5 Conclusions

The current performance of the ATLAS flavour-tagging algorithms for the 2016 LHC run has been described in this note including a review of the current algorithmic strategy, the new features of the main basic taggers, IP and the Secondary Vertex Finder algorithms, and a complete overview of the new training procedure for MV2. Particular emphasis is given to the recent developments of MV2 in light of the improved c -jet rejection resulting in a gain of approximately 40% in c -jet rejection at 77% b -jet efficiency and approximately the same light-flavour jet rejection as achieved by the previous algorithm used for 2015 analyses. In addition, the rejection power at high b -jet p_T is also significantly improved.

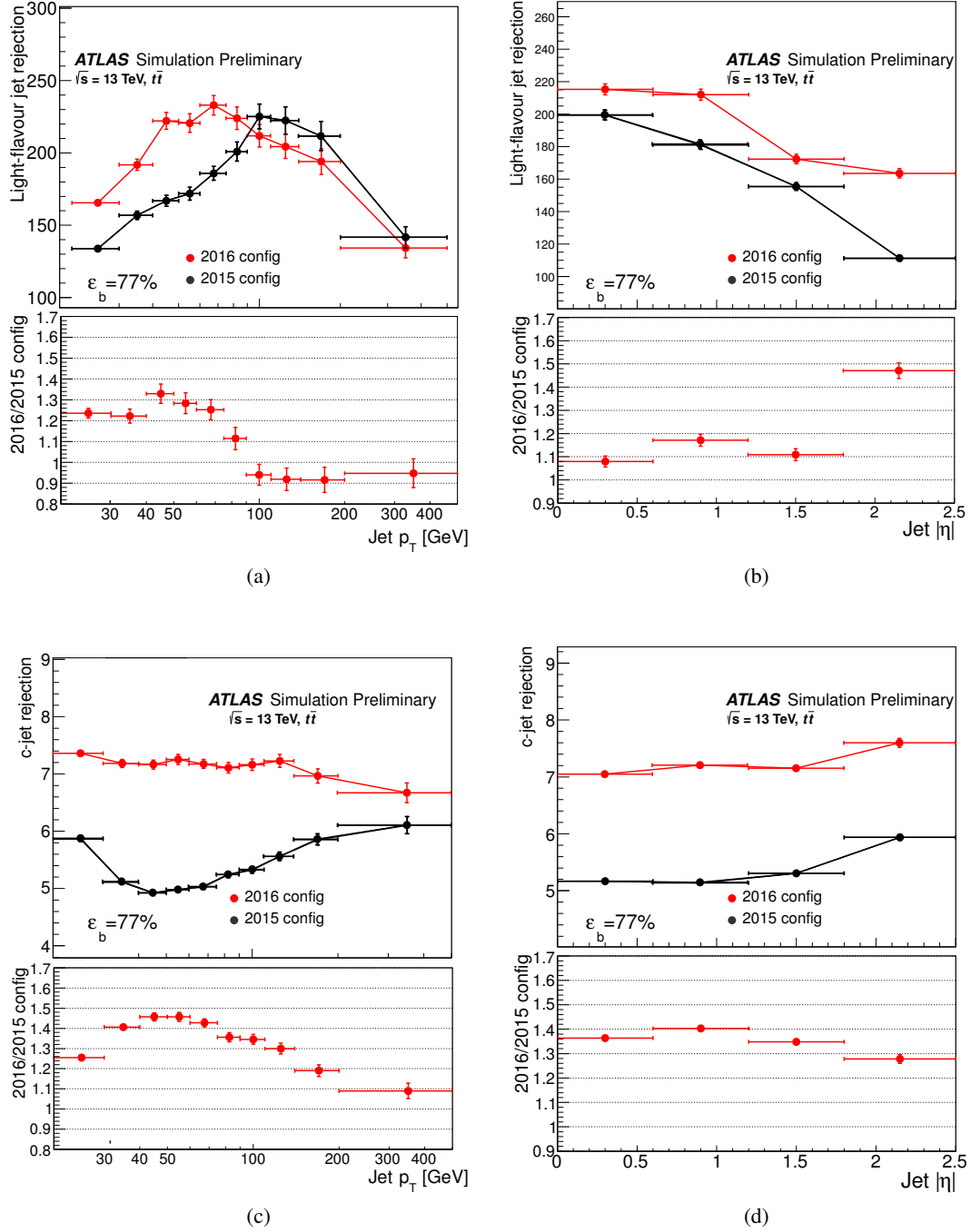


Figure 13: Light-flavour jet and c -jet rejection as a function of jet p_T , (a) and (c), and jet η , (b) and (d), for the previous (2015 config) MV2c20 and the current MV2c10 configuration (2016 config). A fixed cut at 77% b -jet efficiency operating point is employed.

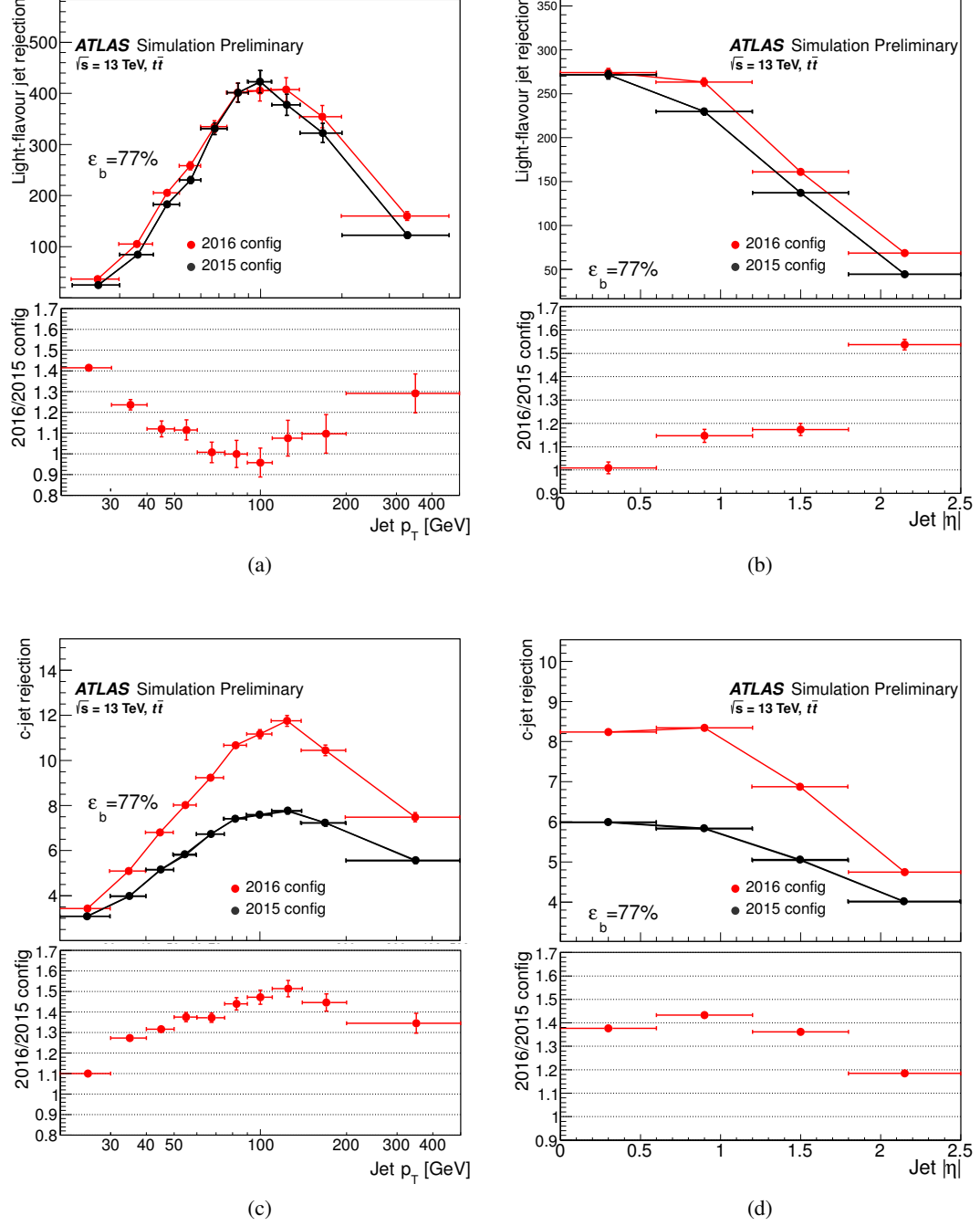


Figure 14: Light-flavour jet and c -jet rejection as a function of jet p_T , (a) and (c), and jet η , (b) and (d), for the previous (2015 config) MV2c20 and the current MV2c10 configuration (2016 config). In each bin of the p_T distribution, the b -tagging cut value has been set in such a way to yield a constant b -jet efficiency at 77%.

References

- [1] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, **JINST** **3** (2008) S08003.
- [2] ATLAS Collaboration, *Expected performance of the ATLAS b-tagging algorithms in Run-2*, ATL-PHYS-PUB-2015-022, URL: <https://cds.cern.ch/record/2037697>.
- [3] ATLAS Collaboration, *The ATLAS Inner Detector commissioning and calibration*, **Eur. Phys. J. C** **70** (2010) 787.
- [4] ATLAS Collaboration, *ATLAS Insertable B-Layer Technical Design Report*, CERN-LHCC-2010-013. ATLAS-TDR-19, URL: <https://cds.cern.ch/record/1291633>.
- [5] ATLAS Collaboration, *Performance of b-Jet Identification in the ATLAS Experiment*, **JINST** **11** (2016) P04008.
- [6] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, **JHEP** **11** (2004) 040.
- [7] T. Sjöstrand, S. Mrenna and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, **JHEP** **05** (2006) 026.
- [8] H.-L. Lai et al., *New parton distributions for collider physics*, **Phys. Rev. D** **82** (2010) 074024.
- [9] D. J. Lange, *The EvtGen particle decay simulation package*, **Nucl. Instrum. Meth. A** **462** (2001) 152.
- [10] T. Sjöstrand, S. Mrenna and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, **Comput. Phys. Commun.** **178** (2008) 852.
- [11] R. D. Ball et al., *Parton distributions with LHC data*, **Nucl. Phys. B** **867** (2013) 244.
- [12] S. Agostinelli et al., *GEANT4: A simulation toolkit*, **Nucl. Instrum. Meth. A** **506** (2003) 250.
- [13] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, **Eur. Phys. J. C** **70** (2010) 823.
- [14] M. Cacciari, G. P. Salam and G. Soyez, *The anti- k_t jet clustering algorithm*, **JHEP** **04** (2008) 063.
- [15] ATLAS Collaboration, *Jet energy measurement and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, **Eur. Phys. J. C** **75** (2015) 17.
- [16] ATLAS Collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*, ATL-CONF-2014-018, URL: <https://cds.cern.ch/record/1700870>.
- [17] ATLAS Collaboration, *Track Reconstruction Performance of the ATLAS Inner Detector at $\sqrt{s} = 13$ TeV*, ATL-PHYS-PUB-2015-018, URL: <https://cds.cern.ch/record/2037683>.
- [18] ATLAS Collaboration, *The Optimization of ATLAS Track Reconstruction in Dense Environments*, ATL-PHYS-PUB-2015-006, URL: <http://cds.cern.ch/record/2002609>.
- [19] ATLAS Collaboration, *A neural network clustering algorithm for the ATLAS silicon pixel detector*, **JINST** **9** (2014) P09009.
- [20] ATLAS Collaboration, *Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data*, ATL-CONF-2011-102, URL: <https://cds.cern.ch/record/1369219>.

- [21] G. Piacquadio and C. Weiser,
A new inclusive secondary vertex algorithm for b-jet tagging in ATLAS,
[J. Phys. Conf. Ser. **119** \(2008\) 032032](#).
- [22] A. Höcker et al., *TMVA: Toolkit for multivariate data analysis*, PoS **ACAT** (2007) 040,
arXiv: [physics/0703039 \[physics\]](#).