ATLAS NOTE

ATL-PHYS-PUB-2015-022

July 24, 2015

Expected performance of the ATLAS *b*-tagging algorithms in Run-2

The ATLAS Collaboration

Abstract

The expected performance of the ATLAS *b*-tagging algorithms for Run-2 is presented. Significant improvements with respect to Run-1 are achieved due to the addition of the Insertable B-Layer, which results in an extra pixel layer closer to the beampipe, and due to several enhancements to the tracking and *b*-tagging algorithms. The new default multivariate *b*-tagging algorithm for Run-2 is described and a detailed set of performance plots are presented.

1 Introduction

The identification of b-quark jets, typically referred to as b-tagging, will play a vital role for the ATLAS experiment [1] during the LHC Run-2. It is important for both precise Standard Model measurements, including the Higgs sector, and for exploring New Physics scenarios, which have a significantly extended reach thanks to the higher centre of mass energy ($\sqrt{s}$ = 13 TeV) of the proton-proton collisions delivered by the LHC.

The most vital input required for b-tagging are the charged particle tracks reconstructed in the Inner Detector [2], which has an acceptance up to a pseudorapidity[1], η, of 2.5. The Inner Detector is immersed in an axial 2 T magnetic field and consists of four layers of silicon pixel sensors (Pixel), four layers of silicon micro-strip sensors (SCT), each equipped with two layers of modules, and a straw tube tracker (TRT), which provides approximately 36 measurements for track reconstruction.

Between Run-1 and 2, a fourth pixel layer, the Insertable B-Layer (IBL) [3], was installed in the ATLAS detector between a new beam pipe with a smaller radius and the previously existing pixel detector. Due to the significantly improved impact parameter resolution, the IBL is expected to have a major impact on the b-tagging performance. In addition, the tracking and b-tagging algorithms have been revisited. In the track reconstruction domain one of the main changes is an improved handling of pixel hits shared between multiple tracks in the core of high transverse momentum, p_T, jets [4], based on a Neural Network pixel hit clustering [5]. The improvements to the b-tagging algorithms will be briefly described in this note.

Both the hardware upgrade from the additional pixel layer and the algorithmic enhancements in the tracking and b-tagging reconstruction software lead to a significant improvement in the b-tagging performance, which is characterized in detail in this note. A comparison is also presented to the performance achieved in Run-1.

This note is organized as follows. In Section 2 the samples and object selection are described. The basic b-tagging algorithms are briefly presented in Section 3. This includes a description of the algorithmic changes made since the end of Run-1 and a visualization of the results in terms of the main output variables. These variables are then used as input to a multivariate algorithm, which is described in Section 4, along with the operating points recommended for use in physics analyses. The expected b-tagging performance is then characterized in detail in Section 5, while Section 6 presents a comparison to the performance achieved in Run-1. Finally conclusions are presented in Section 7.

2 Samples and Physics Object Selection

All the following performance plots are produced with $t\bar{t}$ events corresponding to 13 TeV proton-proton collisions simulated with Powheg+Pythia6 [6, 7] and CT10 [8] parton distribution functions. EvtGen [9] is used to model the decays of b and c-hadrons. Only $t\bar{t}$ decays with at least one lepton from a subsequent W decay are included. Minimum bias interactions consistent with the expected 2015 run conditions are generated with Pythia8 [10] and are overlaid on the $t\bar{t}$ events. The propagation of particles through the detector and the detector response are modeled using GEANT4 [11].

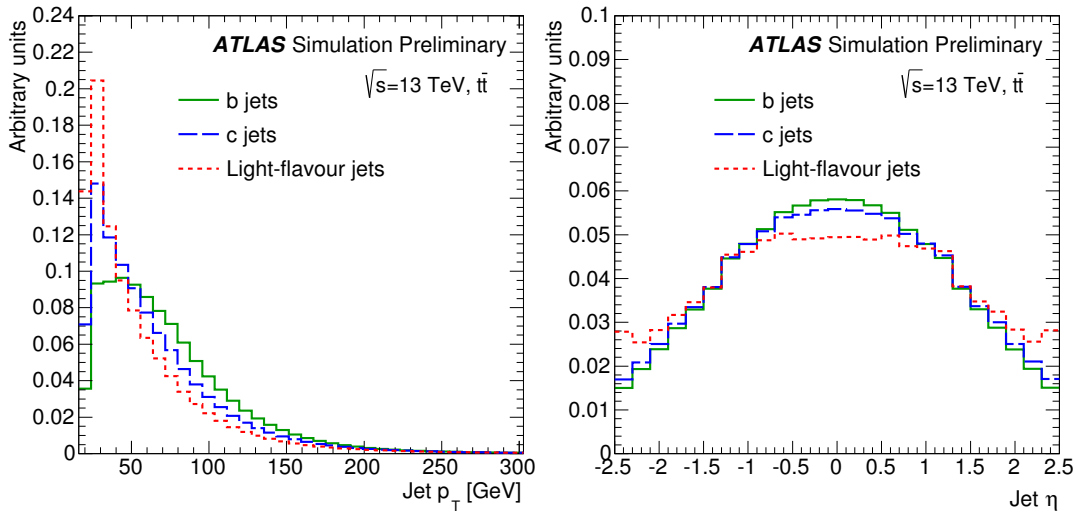
The same ATLAS reconstruction software is run on simulated samples and experimental data. An event-level selection is then applied. It is required that a primary vertex is reconstructed in the event and in the case of several candidate vertices, the primary vertex is defined as the vertex with the largest

[1] ATLAS uses a coordinate system with its origin at the nominal interaction point (IP) in the center of the detector and the z-axis along the beam pipe. Cylindrical coordinates (r, ϕ) are used in the transverse plane, ϕ being the azimuthal angle around the beam pipe. The pseudorapidity is defined in terms of the polar angle θ as $\eta = -\ln\tan(\theta/2)$, while $\Delta R \equiv \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

sum of squared transverse momenta of the associated tracks. Jets are reconstructed by clustering energy deposits in the calorimeter with the anti-k_t algorithm [12] and 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. A preliminary version of the Run-2 jet energy scale calibration is applied to the jets [13]. In this note, only jets with p_{T} above 20 GeV and $|\eta| < 2.5$ are considered. A new algorithm, referred to as the jet vertex tagger (JVT), is used to reject jets from pileup. The JVT algorithm combines two track-based variables, which are sensitive to the vertex origin of jet, in a likelihood discriminant. Jets with $p_{\mathrm{T}} < 50$ GeV and $|\eta| < 2.4$ are rejected if they have a JVT output of less than 0.641 [14]. This corresponds to an expected efficiency of about 92% for jets from the hard-scatter and a 2% efficiency for pile-up jets. The JVT selection is close to 100% efficient for b-tagged b and c-jets. To minimise any bias when making a comparison between the b-tagging performance in Run-1 and Run-2, it was necessary to make some adjustments to the event selection and samples, this is discussed in more detail in Section 6.

For simulated jets, a flavour label is assigned by matching jets to the truth-level weakly decaying b and c hadrons with $p_{\mathrm{T}} > 5$ GeV, in a ΔR cone of less than 0.3. The flavour labelling is exclusive, with the hadron matched to the closest jet in the ΔR phase-space. If a b-hadron is found within the cone the jet is labeled as a b-jet. If no b-hadron is found, the search is repeated for c-hadrons, then for τ leptons. If no match is found for b, c, or τ, the jet is labeled as a light-flavour jet.

The p_T and η distributions of the b, c, and light-flavour jets are shown in Fig. 1.



(a) (b)

Figure 1: The p_T (a) and η (b) distributions for b (solid green), c (dashed blue) and light-flavour (dotted red) jets in $t\bar{t}$ events.

3 Basic Algorithms

The basic b-tagging algorithms use charged particle tracks to produce a set of variables which discriminate between different jet flavour. Tracks are first associated to a jet and are then required to pass a quality selection, as described in Section 3.1. ATLAS uses three distinct basic b-tagging algorithms,

which provide complementary information:

- Impact parameter based algorithm (Section 3.2).
- Inclusive secondary vertex reconstruction algorithm (Section 3.3).
- Decay chain multi-vertex reconstruction algorithm (Section 3.4).

The output of these b-tagging algorithms are later combined in a multivariate discriminant (Section 4), which provides the best separation between the different jet flavours. This note primarily focusses on the major changes made to the b-tagging algorithms in Run-2 compared to Run-1: the updated track categories for the impact parameter tool and the new multivariate tool.

3.1 Track Selection

Tracks are associated to calorimeter jets based on their angular separation $\Delta R(\text{track}, \text{jet})$. The ΔR association requirement varies as a function of the jet p_T, resulting in a narrower cone for jets at high p_T which are more collimated. A given track is associated with only one jet; if it satisfies the association criterion with respect to more than one jet, the jet with the smallest ΔR is chosen.

The track selection depends on each specific b-tagging algorithm. For the impact parameter based algorithm, a tight selection is applied. The most important requirements include a requirement on the track p_T above 1 GeV, the transverse and longitudinal impact parameters to be limited to $|d_0| <$ 1 mm and $|z_0 \times \sin\theta|$<1.5 mm, and at least two hits in the pixel detector. For the secondary vertex based algorithms a looser selection is used, relying on the secondary vertex reconstruction to provide additional purity. This includes requiring track p_T to be above 700-800 MeV and significantly looser requirements in terms of impact parameter and track quality.

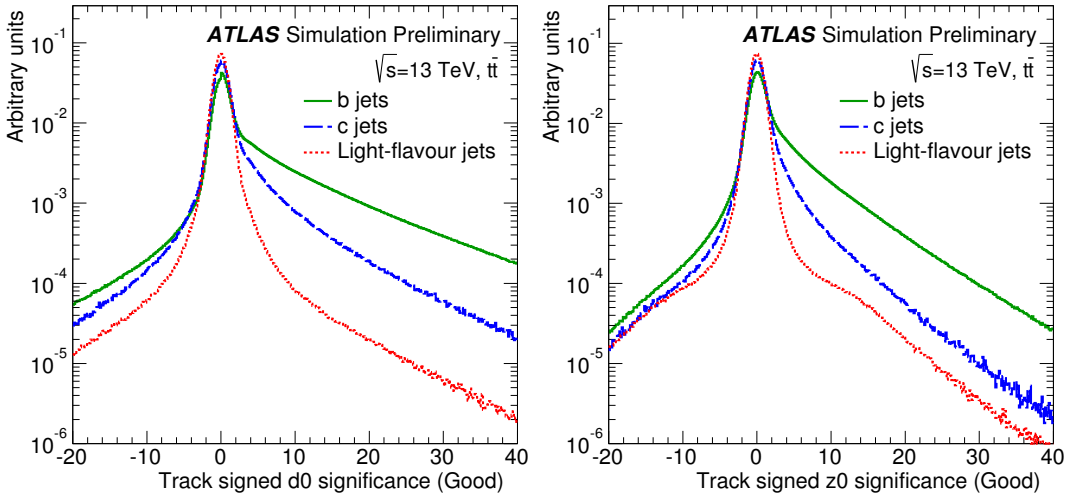
3.2 Impact Parameter based Algorithms: IP2D, IP3D

The IP2D and IP3D algorithms [15], make use of the signed impact parameter significance of the tracks matched to the jet. The sign is defined 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 probability density functions (PDFs) for the signed impact parameter significance of these tracks are used to define ratios of the b- and light-flavour jet hypotheses, and these are then combined in a single log likelihood ratio discriminant (LLR). The LLR formalism also allows for the use of different PDF sets for different track categories. In Run-2, the categorisation of tracks has been significantly refined with respect to the version used in Run-1. Table 1 describes the different track categories and shows the rate at which tracks from b-, c- and light-flavour jets populate them. Alternative LLR discriminants can be constructed based on ratios of the b- and c-jet, or c- and light-flavor jet hypotheses.

IP3D uses both the transverse and longitudinal impact parameters taking into account their correlations, while IP2D only uses the transverse impact parameters. Compared to IP3D, IP2D is more robust against the effects of pile-up, as it does not take account of the longitudinal impact parameter significance, which will typically be large for tracks from pileup jets. Figure 2 shows the transverse and longitudinal impact parameter distributions for tracks from b-, c- and light-flavour jets. In the distribution of transverse impact parameter significances for light-flavour jets a clear exponential tail at high positive values is seen, corresponding to tracks from K_S or Λ decays, from photon conversions and interactions in the detector material. In the case of the longitudinal impact parameter significance, an additional component is seen in the tail, symmetric around zero, corresponding to tracks from pileup. The final discriminants for both the IP2D and IP3D algorithms are illustrated in Fig. 3.

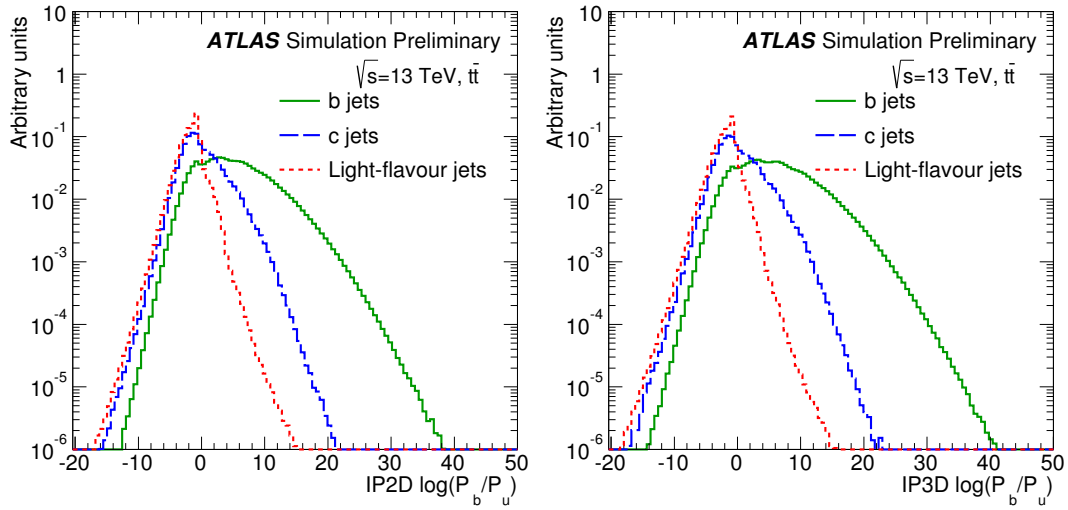
#	Description	Rate [%]		
		b jets	c jets	light jets
0	No hits in first two layers; exp. hit in L0 and L1	1.5	1.6	1.6
1	No hits in first two layers; exp. hit in L0 and no exp. hit in L1	0.1	0.1	0.1
2	No hits in first two layers; no exp. hit in L0 and exp. hit in L1	0.03	0.03	0.03
3	No hits in first two layers; no exp. hit in L0 and L1	0.03	0.03	0.02
4	No hit in L0; exp. hit in L0	2.4	2.3	2.1
5	No hit in L0; no exp. hit in L0	0.9	0.9	0.9
6	No hit in L1; exp. hit in L1	0.5	0.5	0.5
7	No hit in L1; no exp. hit in L1	2.4	2.4	2.3
8	Shared hit in both L0 and L1	0.01	0.01	0.04
9	Shared pixel hits	2.1	1.6	1.8
10	Two or more shared SCT hits	2.4	2.2	2.2
11	Split hits in both L0 and L1	1.2	1.1	0.8
12	Split pixel hit	2.1	1.6	1.1
13	Good: a track not in any of the above categories	84.3	85.5	86.6

Table 1: Description of the track categories used by IP2D and IP3D along with the fraction of tracks in each category for the $t\bar{t}$ sample. The categories are constructed with respect to the track quality, which is defined by the clusters (hits), from the silicon layers of the Inner Detector, used in the track reconstruction. The clusters in the innermost (L0) and next-to-innermost (L1) layers of the pixel detectors are of particular importance, as is the knowledge of whether a cluster was expected (exp.) or not, based on the detector coverage and dead module maps. Shared hits are clusters which are shared among more than one track, degrading track quality, while split hits are clusters which have been identified as originating from overlapping tracks and have therefore be split into sub-clusters.



(a) (b)

Figure 2: The transverse (a) and longitudinal (b) signed impact parameter significance of tracks in $t\bar{t}$ events associated with b (solid green), c (dashed blue) and light-flavour (dotted red) jets for the "Good" category defined in Table 1.

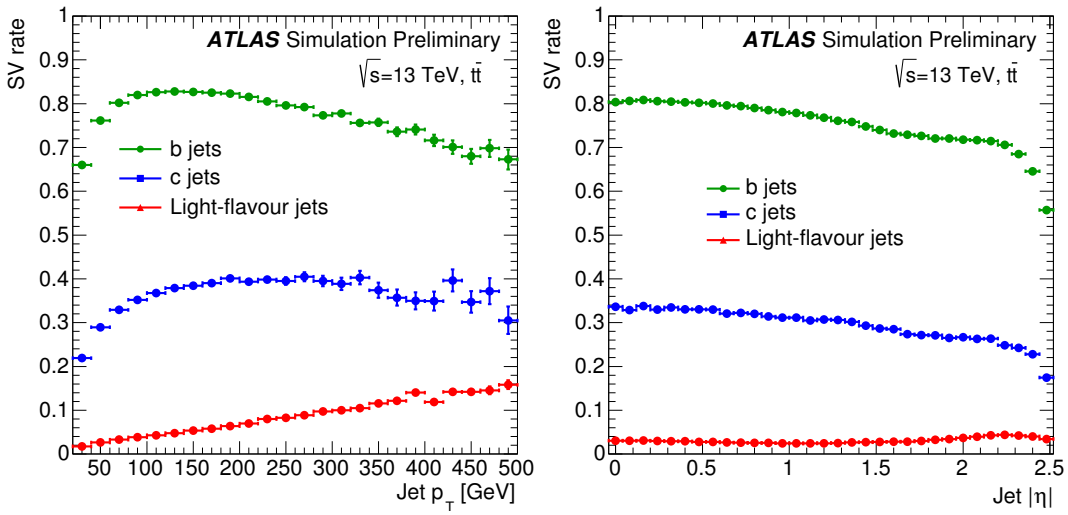


(a) (b)

Figure 3: The log likelihood ratio for the IP2D (a) and IP3D (b) b-tagging algorithm for b (solid green), c (dashed blue) and light-flavour (dotted red) jets in $t\bar{t}$ events. If no tracks are found in the jet, a large negative value is assigned as the algorithm output. This happens for less than 0.5% of b and c-jets, and for about 2% of light-flavour jets.

3.3 Secondary Vertex Finding Algorithm: SV

The secondary vertex based algorithm [15] aims to explicitly reconstruct an inclusive displaced secondary vertex within the jet. The first step consists of reconstructing two-track vertices using the candidate tracks. Tracks are rejected if they form a secondary vertex which can be identified as likely originating from the decay of a long-lived particle (e.g. K_S or Λ), photon conversions or hadronic interactions with the detector material. A single vertex is then reconstructed using the tracks that survive this preselection, with outlier tracks iteratively removed. Figure 4 shows the secondary vertex reconstruction efficiency as function of jet p_T and η for b-, c- and light-flavour jets. Figure 5 shows the distribution of some of the properties of the reconstructed secondary vertex comparing vertices from b, c and light-flavour jets.



(a) (b)

Figure 4: Secondary vertex reconstruction rate as function of jet p_T (a) and η (b) for b (green), c (blue) and light-flavour (red) jets in $t\bar{t}$ events.

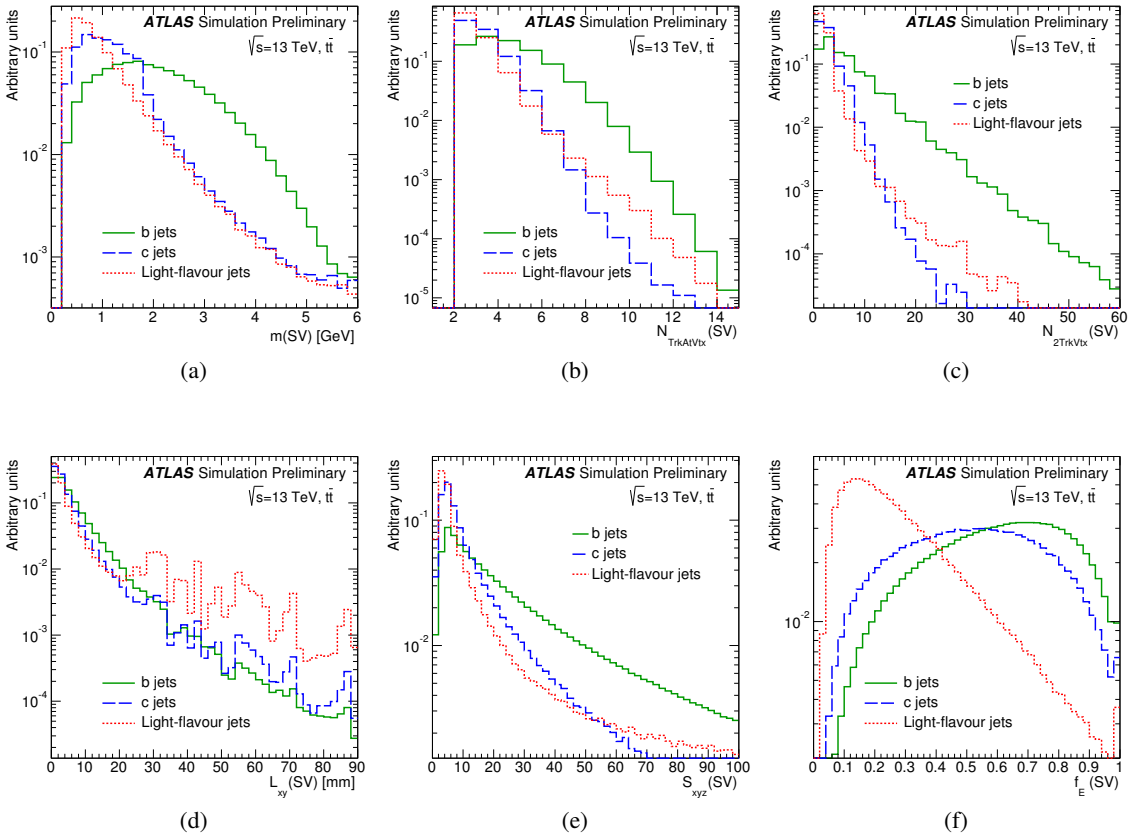
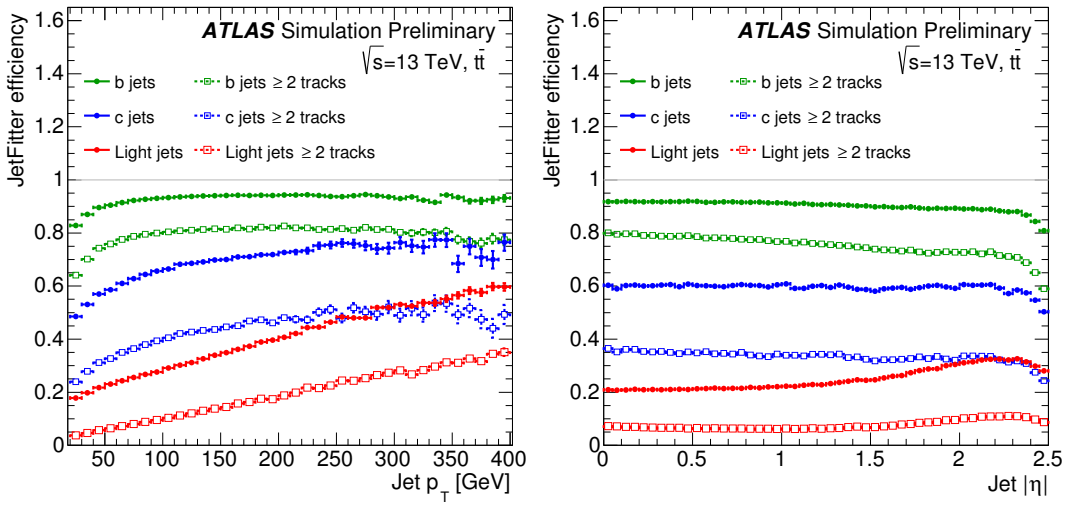


Figure 5: Properties of the secondary vertices reconstructed by the SV algorithm for b- (solid green), c- (dashed blue) and light-flavour (dotted red) jets in $t\bar{t}$ events: the invariant mass (a), the number of tracks (b), the number of two-track vertices reconstructed within the jet (c), the transverse decay length (d), 3D decay length significance (e) and the energy fraction, defined as the energy from the tracks in the displaced vertex relative to all tracks reconstructed within the jet (f). The increased rate of light-flavour jets at at high transverse decay length values is due to material interactions and long-lived particles (K_s^0 and Λ^0), with the jagged structure due to the vertex cleaning cuts.

3.4 Decay Chain Multi-Vertex Algorithm: JetFitter

The decay chain multi-vertex reconstruction algorithm, JetFitter [16], exploits the topological structure of weak b- and c-hadron decays inside the jet and tries to reconstruct the full PV $\rightarrow$ b $\rightarrow$ c-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 each of them.

The efficiency to reconstruct a vertex with at least one or two tracks is illustrated in Fig. 6. As the plots clearly show the efficiency to have at least a single-track vertex is significantly higher than the efficiency to have a vertex with at least two tracks. This comes however at the expense of a higher rate of reconstructed vertices in light-jets. Distributions of some of the most important output variables are shown in Fig. 7, based on either the decay topology or the properties of the reconstructed vertices.



(a) (b)

Figure 6: JetFitter vertex reconstruction efficiency as function of jet p_T (a) and $|\eta|$ (b) for b (green), c (blue) and light-flavour (red) jets in $t\bar{t}$ events. The solid lines with closed markers represent the efficiency to reconstruct any JetFitter decay chain, the dashed line with open markers requires that at least one vertex has two or more tracks.

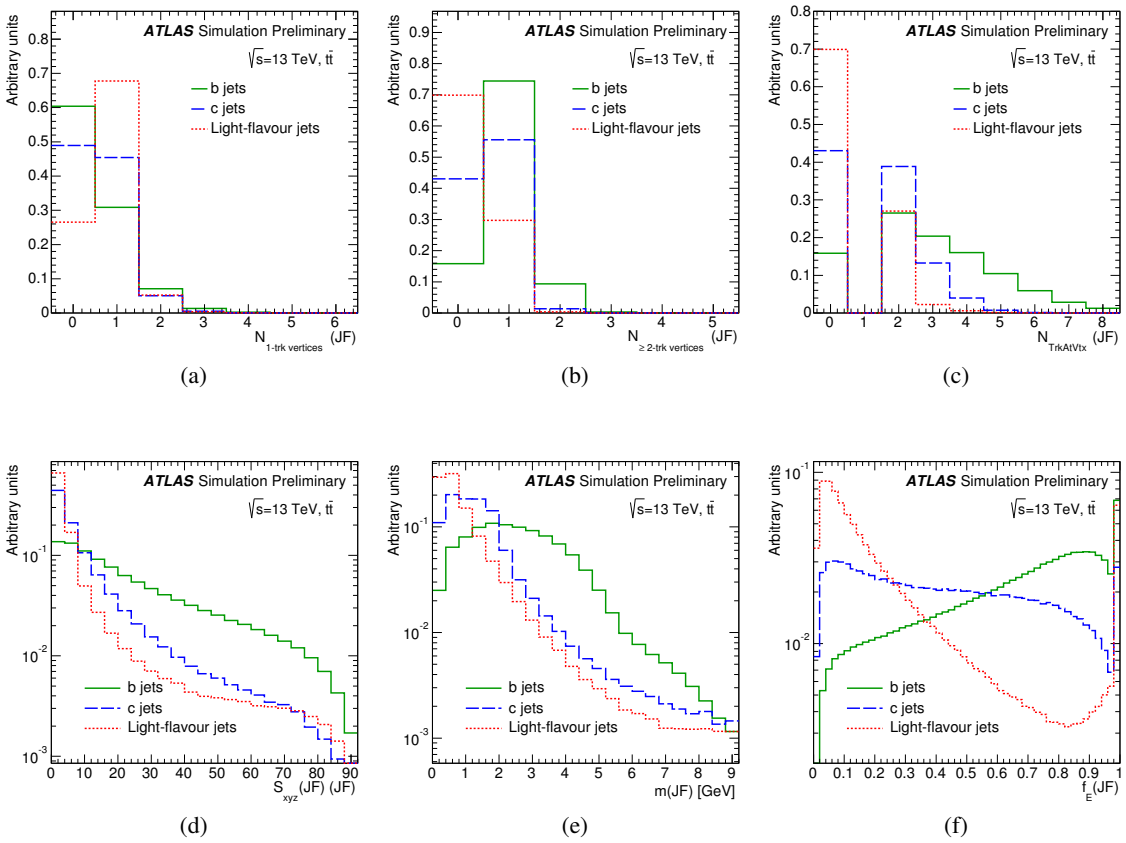


Figure 7: Properties of the decay topology and secondary vertices reconstructed by JetFitter for b (solid green), c (dashed blue) and light-flavour (dotted red) jets in $t\bar{t}$ events: the number of 1-track vertices (a), the number of vertices with at least two tracks (b), the number of tracks from vertices with at least two tracks (c), the average flight length significance of the reconstructed vertices (d), the invariant mass of tracks fitted to one or more displaced vertices (e) and the energy fraction, defined as the energy from the tracks in the displaced vertex relative to all tracks reconstructed within the jet (f).

4 Multivariate Algorithm: MV2

The input variables obtained from the three basic algorithms are combined using a boosted decision tree (BDT) algorithm to discriminate b-jets from light (u,d,s-quark or gluon jets) and c-jets. The training is performed on a set of approximately 5 million $t\bar{t}$ events. The MV2c20 algorithm is defined as the output of such a BDT with the training performed assigning b-jets as signal and a mixture of 80% light-flavour jets and 20% c-jets as background. The list of input variables used for the training is summarized in Table 2. The kinematic properties (p_T and η) of the jets are included in the training to take advantage of correlations with the other input variables. In order to avoid any difference in the kinematic spectra between signal and background jets being interpreted as discriminant by the training, the signal jets are reweighted to match the spectrum of the background jets. The MV2c20 output distribution is shown in Fig. 8 for b, c and light-flavour jets.

The performance for several background mixtures of c- and light-flavour jets in the training has been compared: the mixture adopted in MV2c20 gave a good trade-off between light- and c-jet rejection from an analysis standpoint. In a second variant, denoted MV2c00, the training is performed with only light jets as background. As shown in Fig. 9, a comparison between the MV2c20 and MV2c00 algorithms shows that a small admixture of c-jets only slightly degrades the light-flavour jet rejection, but significantly improves the c-jet rejection.

The MV2 algorithm constitutes a significant revision of the main b-tagging algorithm used during Run-1, denoted MV1 [17], which combined inputs from intermediate multivariate tools trained for each of the basic b-tagging algorithms. MV1 was based on a neural network approach rather than a BDT. The new approach not only improves the performance, as will be shown in Section 6, but also significantly simplifies the algorithm by directly using the variables from the basic algorithms, omitting the additional intermediate multivariate tools.

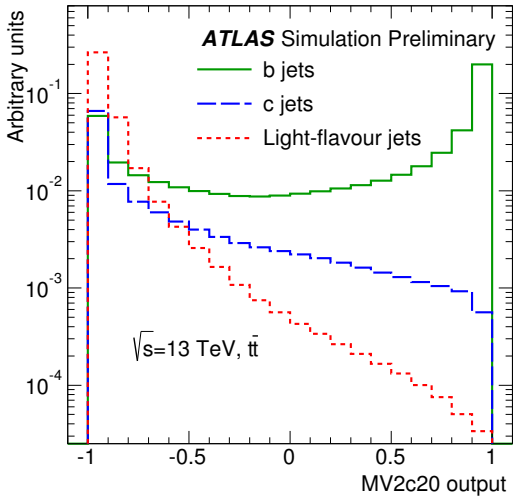


Figure 8: The MV2c20 output for b- (solid green), c- (dashed blue) and light-flavour (dotted red) jets in $t\bar{t}$ events.

Input	Variable	Description
Kinematics	$\mathrm{p}_T(jet)$	Jet transverse momentum
	$\eta(jet)$	Jet pseudo-rapidity
IP2D, IP3D	$\log(P_b/P_{\mathrm{light}})$	Likelihood ratio between the b- and light jet hypotheses
	$\log(P_b/P_c)$	Likelihood ratio between the b- and c-jet hypotheses
	$\log(P_c/P_{\mathrm{light}})$	Likelihood ratio between the c- and light jet hypotheses
SV	$m(\mathrm{SV})$	Invariant mass of tracks at the secondary vertex assuming pion masses
	$f_{\mathrm{E}}(\mathrm{SV})$	Fraction of the charged jet energy in the secondary vertex
	$N_{\mathrm{TrkAtVtx}}(\mathrm{SV})$	Number of tracks used in the secondary vertex
	$N_{\mathrm{2TrkVtx}}(\mathrm{SV})$	Number of two track vertex candidates
	$L_{xy}(\mathrm{SV})$	Transverse distance between the primary and secondary vertices
	$L_{xyz}(\mathrm{SV})$	Distance between the primary and secondary vertices
	$S_{xyz}(\mathrm{SV})$	Distance between the primary and secondary vertices divided by its uncertainty
	$\Delta R(\mathrm{jet},\mathrm{SV})$	ΔR between the jet axis and the direction of the secondary vertex relative to the primary vertex
Jet Fitter	$N_{\mathrm{2TrkVtx}}(\mathrm{JF})$	Number of 2-track vertex candidates (prior to decay chain fit)
	$m(\mathrm{JF})$	Invariant mass of tracks from displaced vertices assuming pion masses
	$S_{\mathrm{xyz}}(\mathrm{JF})$	Significance of the average distance between the primary and displaced vertices
	$f_{\mathrm{E}}(\mathrm{JF})$	Fraction of the charged jet energy in the secondary vertices
	$N_{\text{1-trk vertices}}(\mathrm{JF})$	Number of displaced vertices with one track
	$N_{\geq\text{2-trk vertices}}(\mathrm{JF})$	Number of displaced vertices with more than one track
	$N_{\mathrm{TrkAtVtx}}(\mathrm{JF})$	Number of tracks from displaced vertices with at least two tracks
	$\Delta R(\vec{p}_{\mathrm{jet}},\vec{p}_{\mathrm{vtx}})$	ΔR between the jet axis and the vectorial sum of the momenta of all tracks attached to displaced vertices

Table 2: The 24 input variables used by the MV2c00 and MV2c20 b-tagging algorithm.

4.1 Operating Points

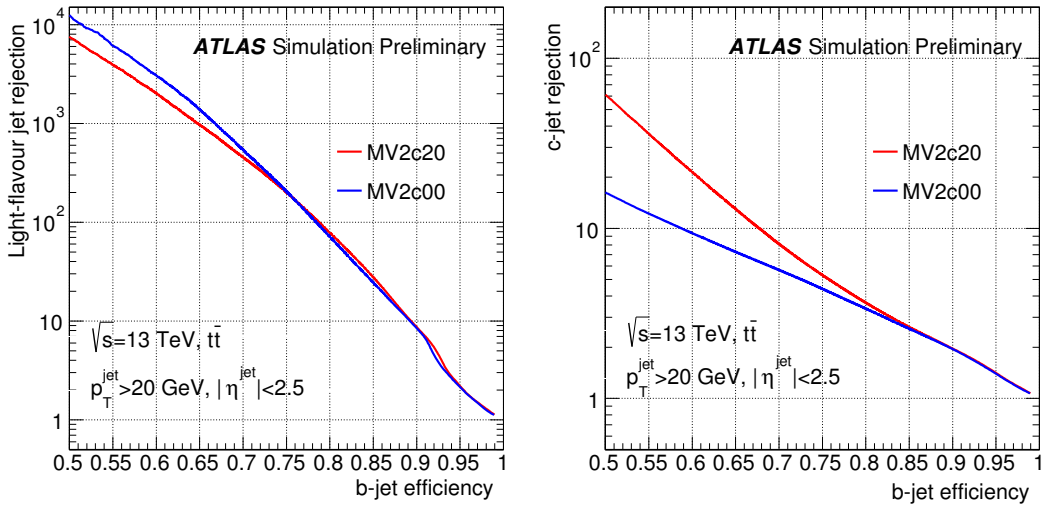
Operating points are defined by a single cut value on the MV2 output distribution and are chosen to provide a specific b-jet efficiency on a $t\bar{t}$ sample. Table 3 shows the four operating points defined for the MV2c20 b-tagging algorithm, with benchmark performance numbers calculated when intergrating over all jets in the $t\bar{t}$ sample.

Cut Value	b-jet Efficiency [%]	c-jet Rejection	τ-jet Rejection	Light-jet Rejection
0.4496	60	21	93	1900
-0.0436	70	8.1	26	440
-0.4434	77	4.5	10	140
-0.7887	85	2.6	3.8	28

Table 3: Operating points for the MV2c20 b-tagging algorithm, including benchmark numbers for the efficiency and rejections rates. The statistical uncertainties on the rejection and efficiency estimates are negligible and so are not shown.

5 Algorithm Performance

The performance of the MV2c00 and MV2c20 b-tagging algorithms are shown Fig. 9 for the b-jet efficiency versus both the light and c-jet rejection. MV2c20 has a substantially larger rejection for c-jets at the expense of a slightly lower light-flavour jet rejection.

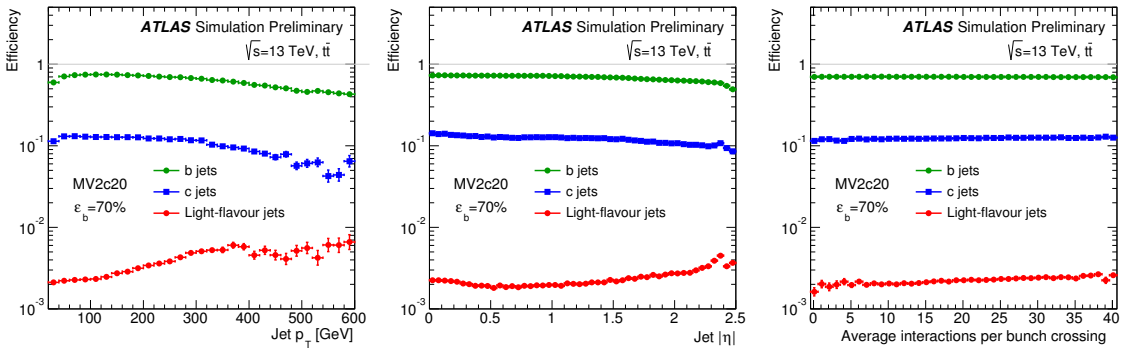


(a) (b)

Figure 9: The light-flavour (a) and c-jet (b) rejection versus b-jet efficiency for the MV2c20 (red) and MV2c00 (blue) b-tagging algorithms in $t\bar{t}$ events.

5.1 Differential Performance

The tagging performance for different flavours are shown for the 70% operating point in Fig. 10 as a function of the jet p_T, $|\eta|$ and average pile-up, when integrating over all jets in the event. The performance of the three other operating points defined in Table 3 are shown in Appendix A.1.



(a) (b) (c)

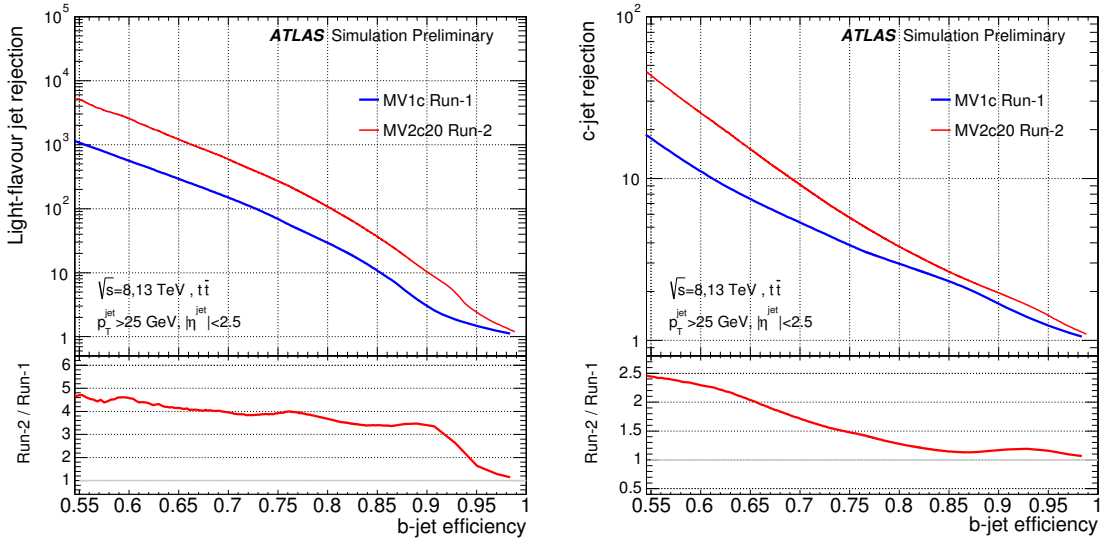
Figure 10: The efficiency to tag b (green), c (blue) and light-flavour (red) jets for the MV2c20 tagging algorithm with the 70% operating point. Efficiencies are shown as a function of the jet p_T (a), $|\eta|$ (b) and the average number of interaction per bunch crossing (c).

6 Performance Enhancement from Detector and Algorithmic Improvements

It is of interest to determine the total improvement in the b-tagging performance achieved between Run-1 and Run-2 due to the addition of the IBL and the algorithmic updates. This requires a comparison to be made between samples produced using the Run-1 simulation and reconstruction software to those using the Run-2 setup. It is therefore necessary to slightly alter the selection and simulation setup to ensure a meaningful comparison. As opposed to the JVT requirement, the jets are required to be matched to a truth hard-scatter jet. The minimum jet p_T requirement is raised from 20 to 25 GeV. Additionally the same $t\bar{t}$ sample, but without EvtGen was used, since EvtGen has only been widely used in ATLAS since the start of Run-2. Not using EvtGen results in a positive shift of about 2 to 3% in the b-jet efficiency, when compared to Pythia6 with EvtGen. The Run-1 and Run-2 samples have been produced with different centre of mass energies (8 and 13 TeV) and different pile-up conditions. In order to get an unbiased comparison, the distributions of jet p_T, jet $|\eta|$ and of the average number of interactions μ for the Run-2 sample are re-weighted in such a way to reproduce the corresponding Run-1 distributions.

Figure 11 shows such a comparison between the performance of the default b-tagging algorithm with improved c-jet rejection for Run-1, MV1c [18], with the Run-1 detector and reconstruction software, compared to the default Run-2 b-tagging algorithm, MV2c20, with the Run-2 detector and reconstruction software. Comparing the two algorithms, the light-flavour jet rejection is improved by a factor of about 4 and the c-jet rejection by a factor of between 1.5-2 for a 70% b-jet efficiency. Alternatively for the same c and light-flavour jet rejection obtained in Run-1 for the 70% working point, a relative

10% improvement in the b-jet efficiency is achieved. Figures 12 and 13 show the same comparison for the light-flavour and c-jet rejection respectively, but as a function of jet p_T and jet $|\eta|$ given a fixed b-jet efficiency of 70% in each bin. The IBL is expected to improve the impact parameter resolution of tracks mostly at p_T up to 5-10 GeV [3], and thus the improvement due to the addition of IBL is concentrated in the low to medium jet p_T region, while at high jet p_T most of the improvement comes from the new algorithms. Appendix A.4 has the same comparison for the default Run-1 b-tagging algorithm, MV1, and the equivalent Run-2 b-tagging algorithm, MV2c00.

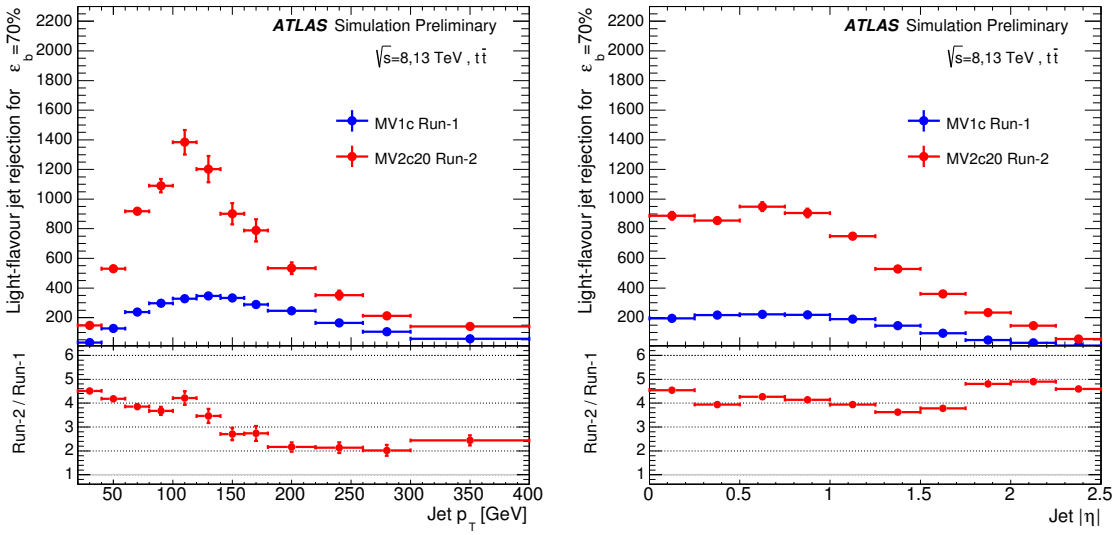


(a) (b)

Figure 11: The light (a) and c-jet rejection (b) versus b-jet efficiency for the MV1c b-tagging algorithm using the Run-1 detector and reconstruction software (blue) compared to the MV2c20 b-tagging algorithm using the Run-2 setup (red).

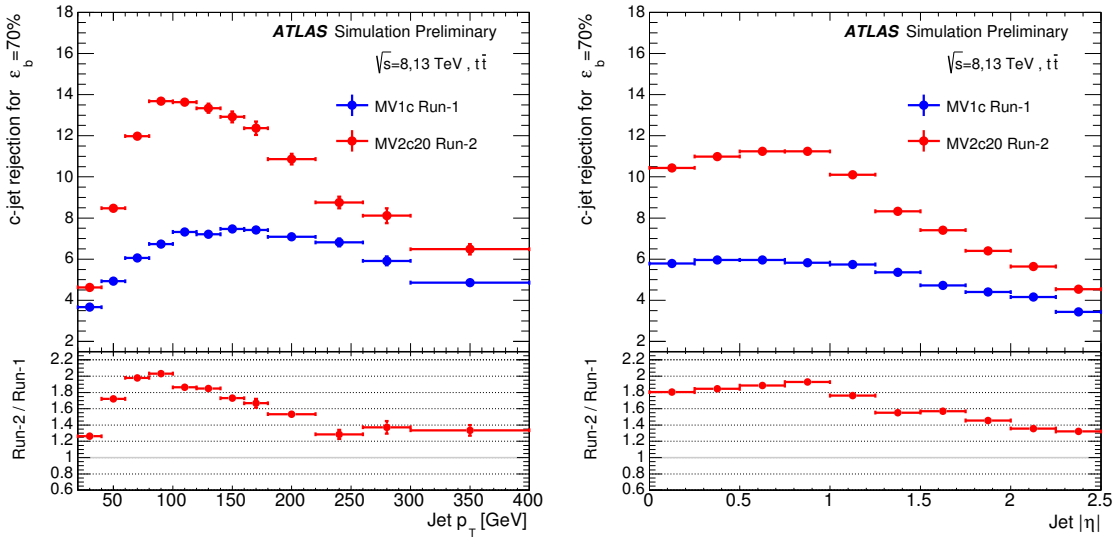
7 Conclusions

The expected performance of the ATLAS b-tagging algorithms for Run-2 has been presented, including a description of the new default multivariate tool and a detailed set of performance plots. Significant improvements are expected due to the addition of the Insertable B-Layer (IBL), which results in an extra pixel layer in the detector, and from several enhancements to the tracking and b-tagging algorithms.



(a) (b)

Figure 12: The light-flavour jet rejection in bins of jet p_T (a) and $|\eta|$ (b) for the MV1c b-tagging algorithm using the Run-1 detector and reconstruction software (blue) compared to the MV2c20 b-tagging algorithm using the Run-2 setup (red). In each p_T or $|\eta|$ bin the b-tagging cut value has been chosen in such a way to yield a constant b-jet efficiency of 70%.



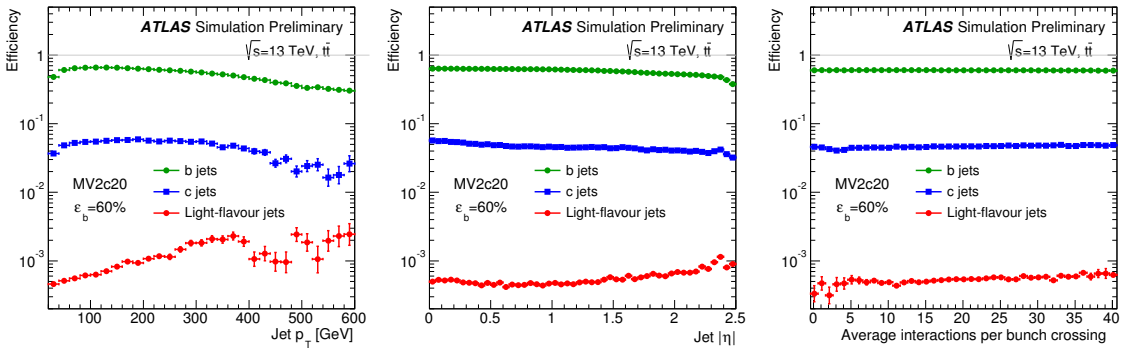
(a) (b)

Figure 13: The c-jet rejection in bins of jet p_T (a) and $|\eta|$ (b) for the MV1c b-tagging algorithm using the Run-1 detector and reconstruction software (blue) compared to the MV2c20 b-tagging algorithm using the Run-2 setup (red). In each p_T or $|\eta|$ bin the b-tagging cut value has been chosen in such a way to yield a constant b-jet efficiency of 70%.

A Operating Point Comparisons

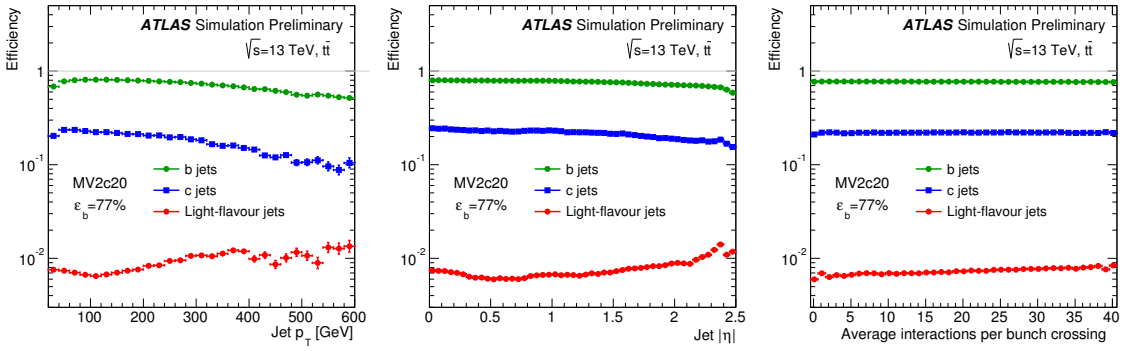
A.1 Differential Performance: 60, 77 and 85% Operating Points

The performance of the four operating points defined in Table 3, except for the 70% operating point which is shown in Section 5.1, are shown in Figs. 14–16 as a function of the jet p_T, $|\eta|$ and average pile-up, when integrating over all jets in the event.



(a) (b) (c)

Figure 14: The efficiency to tag b (green), c (blue) and light-flavour (red) jets for the MV2c20 tagger with the 60% operating point. Efficiencies are shown as a function of the jet p_T (a), $|\eta|$ (b) and the average number of interaction per bunch crossing (c).

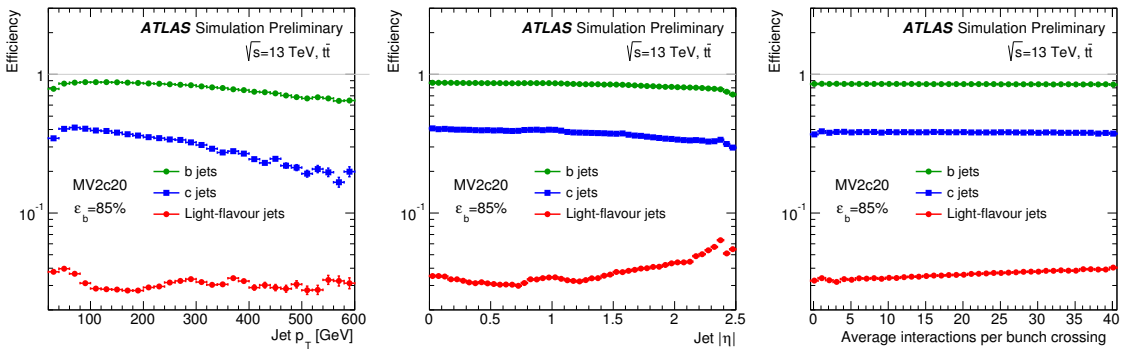


(a) (b) (c)

Figure 15: The efficiency to tag b (green), c (blue) and light-flavour (red) jets for the MV2c20 tagger with the 77% operating point. Efficiencies are shown as a function of the jet p_T (a), $|\eta|$ (b) and the average number of interaction per bunch crossing (c).

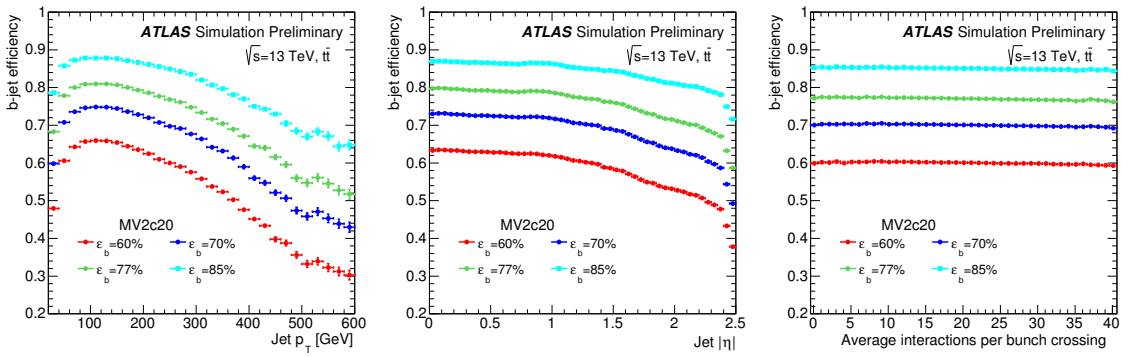
A.2 Differential Operating Point Comparisons

Figures 17-19 show comparisons of the b-jet efficiency, c-jet rejection and light-flavour jet rejection respectively for the four operating points defined in Section 4.1.



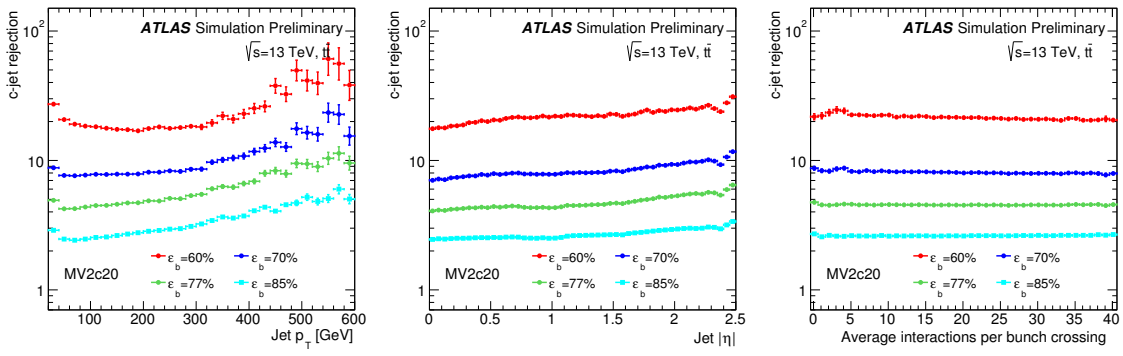
(a) (b) (c)

Figure 16: The efficiency to tag b (green), c (blue) and light-flavour (red) jets for the MV2c20 tagger with the 85% operating point. Efficiencies are shown as a function of the jet p_T (a), $|\eta|$ (b) and the average number of interaction per bunch crossing (c).



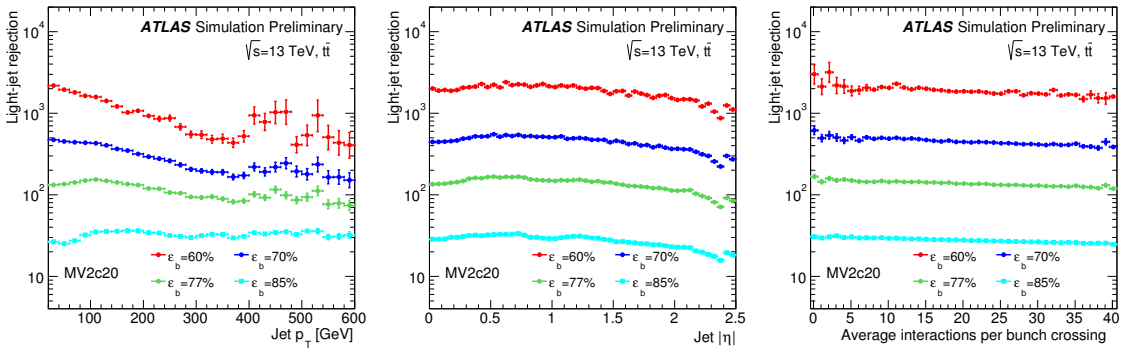
(a) (b) (c)

Figure 17: The b-jet efficiency for the four operating points of the MV2c20 tagger: 60% (red), 70% (blue), 77% (green) and 85% (light blue). Efficiencies are shown as a function of the jet p_T (a), $|\eta|$ (b) and the average number of interaction per bunch crossing (c).



(a) (b) (c)

Figure 18: The c-jet rejection for the four operating points of the MV2c20 tagger: 60% (red), 70% (blue), 77% (green) and 85% (light blue). Efficiencies are shown as a function of the jet p_T (a), $|\eta|$ (b) and the average number of interaction per bunch crossing (c).

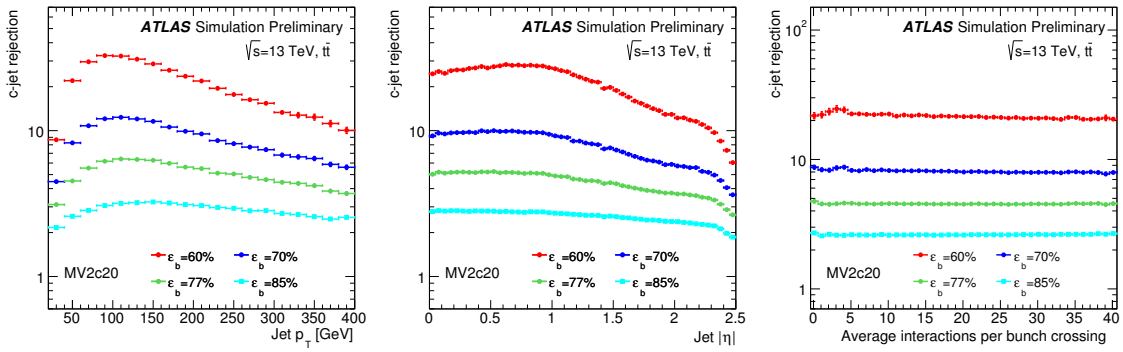


(a) (b) (c)

Figure 19: The light-flavour jet rejection for the four operating points of the MV2c20 tagger: 60% (red), 70% (blue), 77% (green) and 85% (light blue). Efficiencies are shown as a function of the jet p_T (a), $|\eta|$ (b) and the average number of interaction per bunch crossing (c).

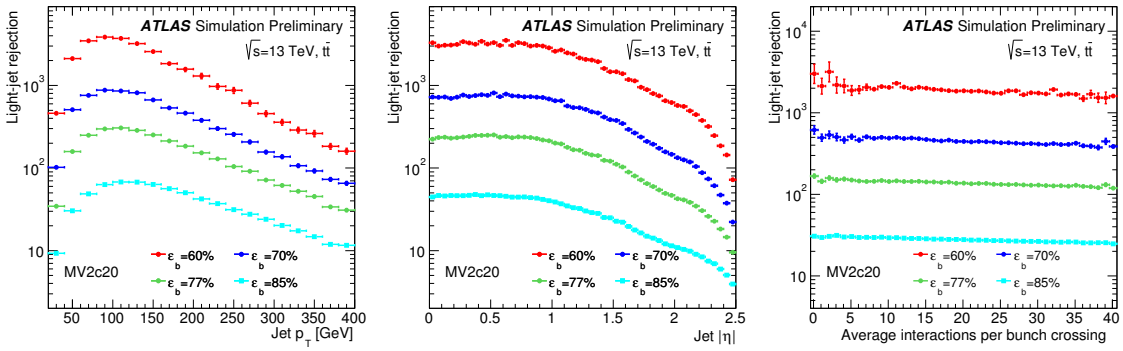
A.3 Differential Operating Point Comparison for a Fixed Efficiency

Figures 20 and 21 show a comparison of the c and light-flavour jet rejections respectively for the four operating points defined in Section 4.1. In constrast to the previous set of plots which relate to the a fixed cut on the MV2c20 output, the cut value is varied in each bin to produce a fixed b-jet efficiency.



(a) (b) (c)

Figure 20: The c-jet rejection for a fixed b-jet efficiency of 60% (red), 70% (blue), 77% (green) and 85% (light blue) in each bin for the MV2c20 b-tagging algorithm. Rejection is shown as a function of the jet p_T (a), $|\eta|$ (b) and the average number of interaction per bunch crossing (c).

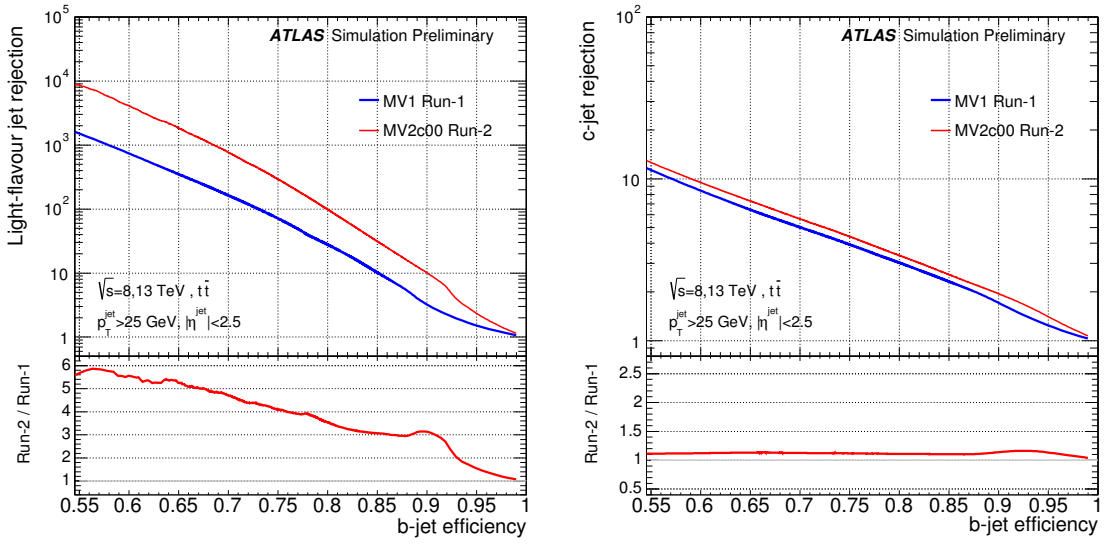


(a) (b) (c)

Figure 21: The light-flavour jet rejection for a fixed MV2c20 b-jet efficiency of 60% (red), 70% (blue), 77% (green) and 85% (light blue) in each bin for the MV2c20 b-tagging algorithm. Rejection is shown as a function of the jet p_T (a), $|\eta|$ (b) and the average number of interaction per bunch crossing (c).

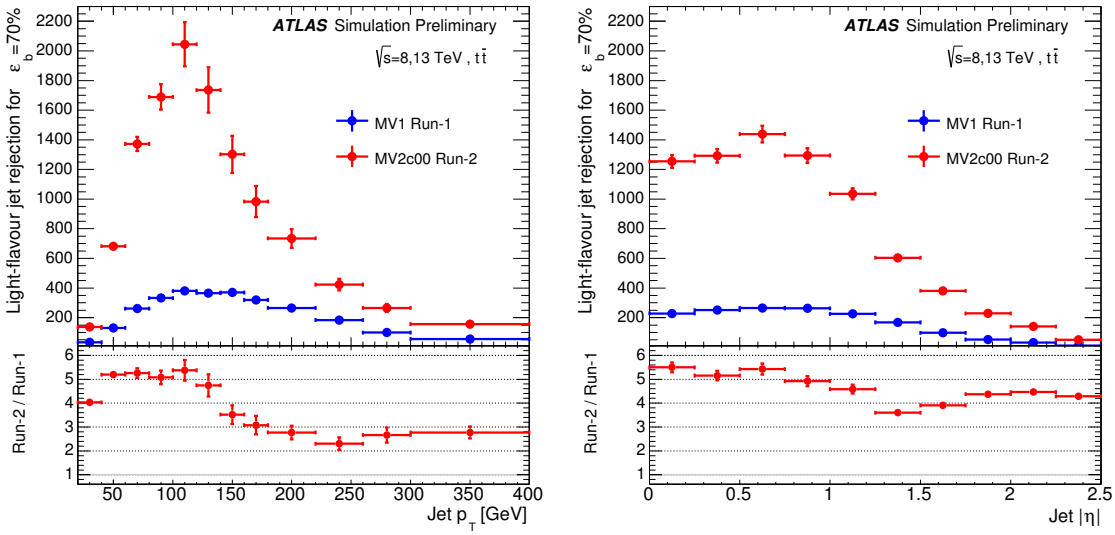
A.4 Comparison of Run-1 MV1 and Run-2 MV2c00

Figure 22 shows such a comparison between the performance of the default b-tagging algorithm for Run-1, MV1, with the Run-1 detector and reconstruction algorithms, compared to the updated Run-2 b-tagging algorithm, MV2c00, with the Run-2 detector and reconstruction software. As outlined in Section 6, to have a meaningful comparison the object selection and sample have been slightly altered compared to those used elsewhere in this note. MV1 and the new MV2c00 multivariate algorithms are both trained with only light-flavour jets as background. Comparing the two algorithms, the light-flavour jet rejection is improved by a factor of 4-5 and the c-jet rejection by about a factor 1.1 for a 70% b-jet efficiency. Alternatively for the same light-flavour jet rejection obtained in Run-1 for the 70% working point, a relative 10% improvement in the b-jet efficiency is achieved. Figures 23 and 24 show the same comparison, for light-flavour and c-jet rejection respectively, but as a function of jet p_T and jet $|\eta|$ given a fixed b-jet efficiency of 70% in each bin.



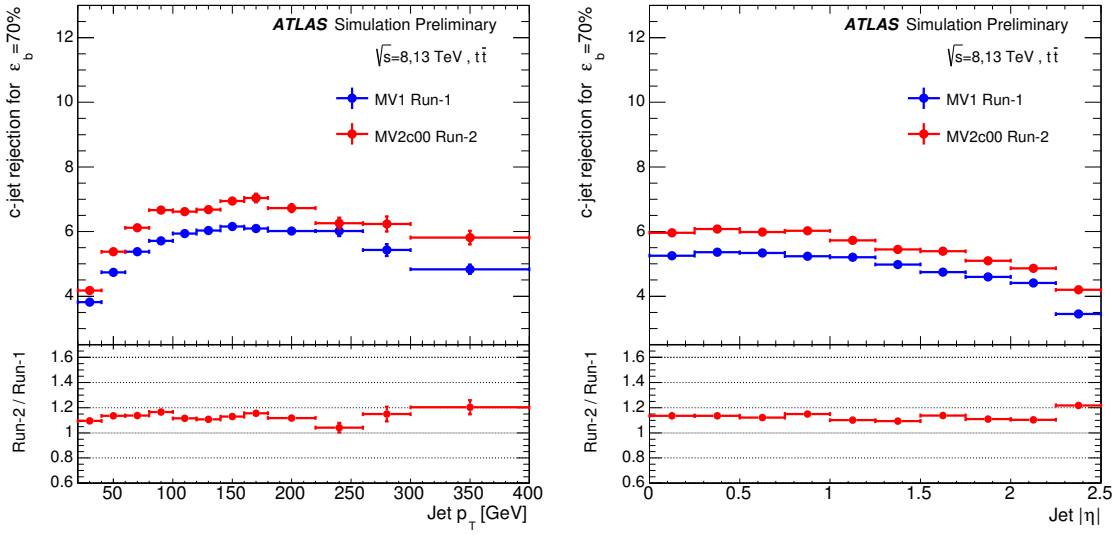
(a) (b)

Figure 22: The light (a) and c-jet rejection (b) versus b-jet efficiency for the MV1 b-tagging algorithm using the Run-1 detector and reconstruction software (blue) compared to the MV2c00 b-tagging algorithm using the Run-2 setup (red).



(a) (b)

Figure 23: The light-flavour jet rejection in bins of jet p_T (a) and $|\eta|$ (b) for the MV1 b-tagging algorithm using the Run-1 detector and reconstruction software (blue) compared to the MV2c00 b-tagging algorithm using the Run-2 setup (red). In each p_T or $|\eta|$ bin the b-tagging cut value has been chosen in such a way to yield a constant b-jet efficiency of 70%.



(a) (b)

Figure 24: The c-jet rejection in bins of jet p_T (a) and $|\eta|$ (b) for the MV1 b-tagging algorithm using the Run-1 detector and reconstruction software (blue) compared to the MV2c00 b-tagging algorithm using the Run-2 setup (red). In each p_T or $|\eta|$ bin the b-tagging cut value has been chosen in such a way to yield a constant b-jet efficiency of 70%.

References

[1] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003. http://iopscience.iop.org/1748-0221/3/08/S08003.

[2] ATLAS Collaboration, *The ATLAS Inner Detector commissioning and calibration*, Eur. Phys. J. **C70** (2010) 787–821, arXiv:1004.5293 [physics.ins-det]. http://arxiv.org/abs/1004.5293.

[3] ATLAS Collaboration, *ATLAS Insertable B-Layer Technical Design Report*, CERN-LHCC-2010-013. ATLAS-TDR-19. https://cds.cern.ch/record/1291633.

[4] ATLAS Collaboration, *The Optimization of ATLAS Track Reconstruction in Dense Environments*, ATL-PHYS-PUB-2015-006. http://cds.cern.ch/record/2002609.

[5] ATLAS Collaboration, *A neural network clustering algorithm for the ATLAS silicon pixel detector*, JINST **9** (2014) P09009, arXiv:1406.7690 [hep-ex]. http://arxiv.org/abs/1406.7690.

[6] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **11** (2004) 040, arXiv:0409146 [hep-ph]. http://arxiv.org/abs/hep-ph/0409146.

[7] T. Sjöstrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **05** (2006) 026. http://arxiv.org/abs/hep-ph/0603175.

[8] H. L. Lai et al., *New parton distributions for collider physics*, Phys.Rev. **D 82** (2010), arXiv:1007.2241 [hep-ph]. http://arxiv.org/abs/1007.2241.

[9] D. Lange, *The EvtGen particle decay simulation package*, Nucl.Instrum.Meth. **A462** (2001) 152–155. http://www.sciencedirect.com/science/article/pii/S0168900201004533.

[10] T. Sjöstrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput.Phys.Commun. **178** (2008) 852–867, arXiv:0710.3820 [hep-ph]. http://arxiv.org/abs/0710.3820.

[11] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4: A simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250–303. http://geant4.web.cern.ch/geant4/.

[12] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP **04** (2008) 063, arXiv:0802.1189 [hep-ph]. http://arxiv.org/abs/0802.1189.

[13] ATLAS Collaboration, *Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at $\sqrt{s}$=13 TeV*, in preparation. http://cds.cern.ch/record/TBD.

[14] ATLAS Collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*, ATLAS-CONF-2014-018. http://cds.cern.ch/record/1700870.

[15] ATLAS Collaboration, *Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data*, ATLAS-CONF-2011-102. http://cds.cern.ch/record/1369219.

[16] ATLAS Collaboration, "Expected Performance of the ATLAS Experiment – Detector, Trigger and Physics." CERN-OPEN-2008-020, December, 2008. http://arxiv.org/abs/0901.0512.

[17] ATLAS Collaboration, *Calibration of the performance of b-tagging for c and light-flavour jets in the 2012 ATLAS data*, ATLAS-CONF-2014-046. http://cds.cern.ch/record/1741020.

[18] ATLAS Collaboration, G. Aad et al., *Search for the $b\bar{b}$ decay of the Standard Model Higgs boson in associated $(W/Z)H$ production with the ATLAS detector*, JHEP **01** (2015) 069, arXiv:1409.6212 [hep-ex].