PROCEEDINGS OF SCIENCE

Calibration of flavour tagging algorithms in ATLAS on $t\bar{t}$ and Z+jets final states

Laura Pereira Sánchez[a,*] **on behalf of the ATLAS Collaboration**

[a] *Stockholm University and Oscar Klein Center*

E-mail: laura.pereira.sanchez@cern.ch

The correct identification of b-initiated jets against c-initiated, hadronic tau-initiated and u-, d-, s-initiated jets is primordial for ATLAS Collaboration searches and precision measurements e.g. $t\bar{t}$ and single-top measurements, supersymmetry, Higgs boson. In consequence, the ATLAS Collaboration has developed several so-called b-tagging algorithms to discriminate b-jets against other flavour jets. Expected performance is evaluated on Monte Carlo events, but due to some inaccuracies in the simulation, it might differ from the performance in real data events. These algorithms are calibrated to account for differences between data and simulated events in the b-tagging efficiency for b-jets and mis-tagging efficiency in c- and light-jets. The b-jet efficiency calibration is performed on two-lepton $t\bar{t}$ events in order to exploit the large purity of this final state in b-jets. The c-jet mis-tag rate is measured on one-lepton $t\bar{t}$ events, which allow the presence of an hadronically decaying W-jet enriched in c-jets. Both calibrations adopt a combinatorial likelihood approach to retrieve the efficiency of the jet for a given flavour. Finally, due to the large light-jet rejection of b-tagging algorithms, the calibration of light-jets mis-tag rate is performed on a Z+jet sample since $t\bar{t}$ doesn't allow for a pure enough light-jet sample. This proceeding presents an overview of the DL1r b-tagging algorithm and the calibration techniques explained above. Results in the form of data - Monte Carlo scale-factors are obtained for the b-jet efficiency and c- and light-jet mis-tag rate with 140 fb^{-1} of data recorded by the ATLAS detector from 2015 to 2018.

**** The European Physical Society Conference on High Energy Physics (EPS-HEP2021), ****
**** 26-30 July 2021 ****
**** Online conference, jointly organized by Universität Hamburg and the research center DESY ****

[*] Speaker

https://pos.sissa.it/

PoS(EPS-HEP2021)434

ATL-PHYS-PROC-2021-079
15 October 2021

1. Introduction

The correct identification of jets containing b-hadrons against other jet flavours, referred as b-tagging, is crucial in many searches and precision measurements performed with the ATLAS detector [1]. A jet containing a b-hadron is referred to as a b-jet, a jet containing a c-hadron but no b-hadron is referred to as a c-jet and a jet containing neither a b-hadron nor a c-hadron is referred to as a light jet. Due to the relatively long lifetime of b-hadrons, reconstructed b-jets are characterised by having a secondary vertex (SV) that is displaced from the primary vertex (PV) or collision point and tracks with high impact parameter (IP), defined as the distance of closest approach from the track trajectory to the PV. These jets are also characterised by having larger invariant mass and track multiplicity than c- or light-jets. The DL1 algorithm [2] is based on an artificial deep neural network that combines information from low level algorithms that exploit the individual features of b-hadrons, to maximise the tagging of b-jets while minimising the mis-tagging of other jet flavours. This proceeding focuses on the calibration of the DL1r algorithm, a variation of the DL1 that uses additional information from the RNNIP algorithm [3] which is based on a recurrent neural network. The performance of the b-tagging algorithms [4] is estimated with the flavour tagging efficiency (ϵ_i), defined by Eq. (1), where $i = b$, c or light.

$$\epsilon_i = \frac{\#\ \text{tagged } i\text{-jets}}{\#\ i\text{-jets}} \tag{1}$$

A jet is considered b-tagged if its DL1r score is above a certain threshold, also known as working point (WP), that results in a specific b-jet efficiency (ϵ_b). Figure 1 shows the light jet rejection $= 1/\epsilon_{\text{light}}$ as a function of ϵ_b for three variations of the DL1 algorithm. The tighter WP tags only 60% of the b-jets but has a very high light-jet rejection whereas the looser WP tags 85% of the b-jets but mis-tags more light-jets.

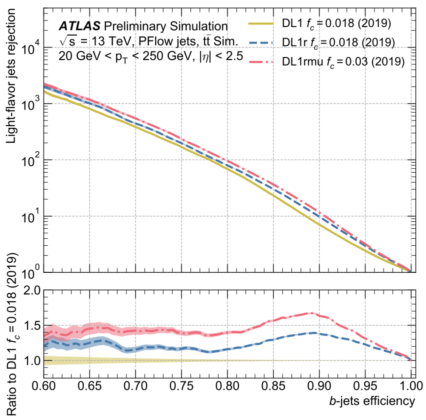


Figure 1: Comparison of ROC curves for three variations of DL1 algorithms used in 2019 [5].

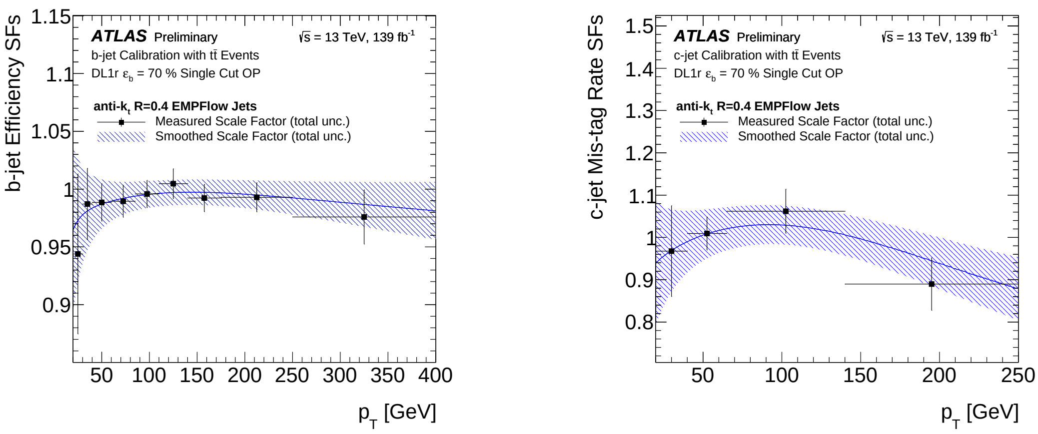
Calibrations of the DL1r algorithm are performed to account for differences in algorithm behaviour between Monte Carlo (MC) simulation and real data. Scale factors, defined in Eq. (2), are derived for each b-tag WP in bins of jet p_{T} to correct the ϵ_i^{MC} for each flavour i. A smoothing [6] is performed to avoid unrealistically sharp variations of the calibrations across jet p_{T}. The final SFs, which are assumed to be independent of the physics process [7], are applied to simulated events:

$$SF_i = \frac{\epsilon_i^{\mathrm{data}}}{\epsilon_i^{\mathrm{MC}}}. \tag{2}$$

2. Calibrations with $t\bar{t}$ events

The calibration of b- and c-jets is performed with $t\bar{t}$ events because they are rich in b-jets due to the large branching ratio of the $t \to Wb$ decay. The dileptonic $t\bar{t}$ decay, where both W bosons decay to a charged lepton and a neutrino, ensures a low contamination of non b-jets and is used to measure SF_b [8]. The semileptonic $t\bar{t}$ decay, where one W-boson decays to leptons and the other decays to cs is used to measure SF_c [9].

A likelihood fit to data is performed to estimate respectively $\epsilon_b^{\mathrm{data}}$ and $\epsilon_c^{\mathrm{data}}$ in bins of jet p_{T} for the different b-tag WPs. The scale factors from the b- and c-jet calibrations are shown respectively in Figure 2a and 2b for the 70% WP.

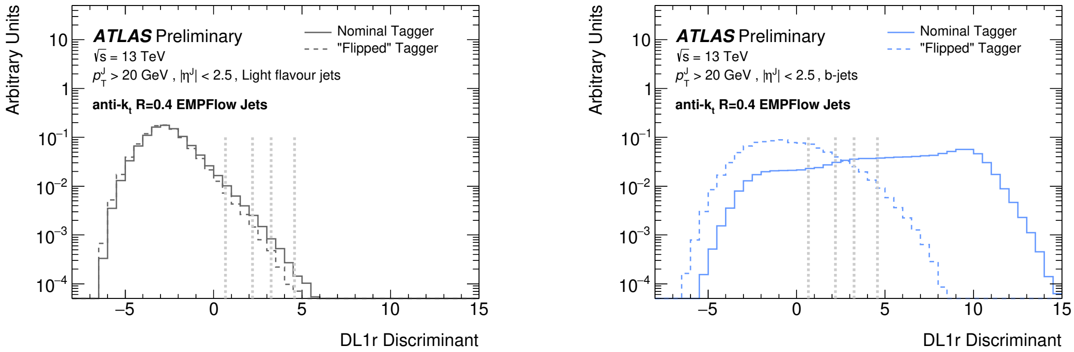


(a) SF_b **(b)** SF_c

Figure 2: Calibration of the b-jet efficiency and c-jet mis-tag rate with $t\bar{t}$ events [10].

3. Calibration with Z+jets events

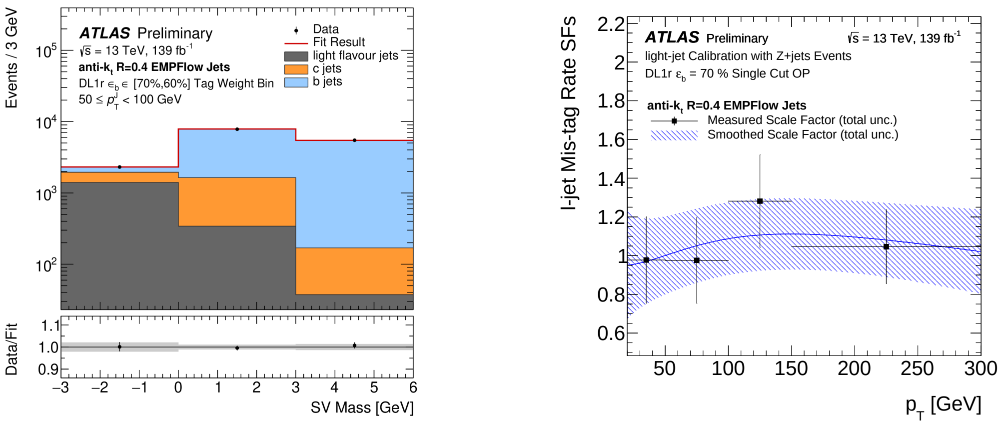
The mis-tag rate ($\epsilon_{\mathrm{light}}$) calibration is performed with Z+jets events in a final state with two oppositely charged leptons and at least one jet. This sample is rich in light-jets. But due to the large light-jet rejection of the b-tagging algorithms, very few light-jets pass the DL1r cut, making a direct fit to data impossible. To obtain a tagged sample with a larger light-jet contribution, a modified version of the DL1r tagger, is used. The modified tagger, referred to as the flipped tagger (DL1rFlip), is based on tracks and SVs for which the sign of the IPs and decay lengths have been inverted. Figure 3 shows the DL1rFlip and DL1 discriminant distributions for *b*- and light-jets. The DL1rFlip tagger has a considerably lower *b*- (and also *c*-) jet tagging efficiency than the DL1 tagger but it has similar mis-tag rate resulting in a tagged sample enriched in light-jets.



(a) light-jets **(b)** *b*-jets

Figure 3: Discriminant of the DL1 and DL1rFlip taggers for different flavour jets. The vertical lines represent the thresholds for each of the 4 calibrated working points [11].

The mis-tag calibration [12] is performed on the DL1rFlip tagger, where a fit to data is possible. The calibration exploits a simultaneous fit to the DL1rFlip discriminant and the SV mass distribution, shown in Figure 4a. This allows to simultaneously extract SF_{light} and fit ϵ_b, minimising the effect of a potential mis-modelling of the *b*-jet efficiency on the mis-tag rate. To account for any discrepancies between the SFs for the DL1r and DL1rFlip taggers, MC is used to derive an extrapolation uncertainty. The final scale factors from the mis-tag rate calibration are shown in Figure 4b.



(a) SV mass distribution [11]. **(b)** SF_{light} [10].

Figure 4: Calibration of the light-jet mis-tag rate with *Z*+jets events.

References

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

[2] ATLAS Collaboration, *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*, Tech. Rep. ATL-PHYS-PUB-2017-013, CERN, Geneva (Jul, 2017).

[3] ATLAS Collaboration, *Identification of Jets Containing b-Hadrons with Recurrent Neural Networks at the ATLAS Experiment*, Tech. Rep. ATL-PHYS-PUB-2017-003, CERN, Geneva (Mar, 2017).

[4] ATLAS Collaboration, *Performance of b-jet identification in the atlas experiment*, *Journal of Instrumentation* **11** (2016) P04008–P04008.

[5] ATLAS Collaboration, "Expected performance of the 2019 ATLAS *b*-taggers." `http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/FTAG-2019-005/`, September, 2019.

[6] ATLAS Collaboration, *Optimisation of the smoothing of b-jet identification efficiency and mistag rate simulation-to-data scale factors in ATLAS*, Tech. Rep. ATL-PHYS-PUB-2020-004, CERN, Geneva (2020).

[7] ATLAS Collaboration, *Monte Carlo to Monte Carlo scale factors for flavour tagging efficiency calibration*, Tech. Rep. ATL-PHYS-PUB-2020-009, CERN, Geneva (May, 2020).

[8] ATLAS Collaboration, *ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at* $\sqrt{s}$ = 13 *TeV*, *Eur. Phys. J. C* **79** (2019) 970 [`1907.05120`].

[9] ATLAS Collaboration, *Calibration of the b-tagging efficiency on charm jets using a sample of W+c events with $\sqrt{s}$ = 13 TeV ATLAS data*, Tech. Rep. ATLAS-CONF-2018-055, CERN, Geneva (Dec, 2018).

[10] ATLAS Collaboration, "Flavour-tagging efficiency corrections for the 2019 ATLAS PFlow jet and VR track jets *b*-taggers with the full LHC Run II dataset." `http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/FTAG-2021-001/`, February, 2021.

[11] ATLAS Collaboration, "Light flavour mistag calibration of the 2019 ATLAS particle-flow jets *b*-taggers using the Z+jets Negative Tag method." `http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/FTAG-2021-002/`, March, 2021.

[12] ATLAS Collaboration, *Calibration of light-flavour b-jet mistagging rates using ATLAS proton-proton collision data at* $\sqrt{s}$ = 13 *TeV*, Tech. Rep. ATLAS-CONF-2018-006, CERN, Geneva (Apr, 2018).