PROCEEDINGS OF SCIENCE

Calibration of the ATLAS b-tagging algorithm in $t\bar{t}$ events with high multiplicity of jets

Francesco La Ruffa[*†]
Universitá della Calabria e INFN (IT)
E-mail: francesco.la.ruffa@cern.ch

The calibration of the ATLAS b-tagging algorithm in environments characterised by high multiplicity of jets is presented. The calibration uses reconstructed $t\bar{t}$ candidate events collected by the ATLAS detector in proton-proton collisions at LHC with a centre-of-mass energy $\sqrt{s}$ of 13 TeV, with a final state containing one charged lepton, missing transverse momentum and at least four jets. The b-tagging efficiencies are measured not only as a function of the most relevant kinematic quantities, such as the transverse momentum or the presudo-rapidity of the jets, but also as a function of quantities that are sensitive to close-by jet activity. The results extend the regions where data-to-simulation b-tagging scale factors are derived when using dilepton $t\bar{t}$ events.

ATL-PHYS-PROC-2017-148
21 September 2017

PoS(EPS-HEP2017)739

The European Physical Society Conference on High Energy Physics
5-12 July, 2017
Venice

[*] Speaker.
[†] on behalf of the ATLAS collaboration

1. Introduction

The correct identification of jets containing b-hadrons, the b-tagging, has a crucial role in the physics programme of the ATLAS experiment [1], a multi-purpose particle detector that records collision events produced by the Large Hadron Collider [2] at CERN. The b-tagging is important for various final states, so analyses relying on it benefit from a precise efficiency calibration of jets coming from a beauty quark (b-jets).

In order to distinguish between a b-jet and a jet coming from a charm quark (c-quark) or from a gluon or lighter quarks (light jets), b-tagging algorithms exploit peculiar properties of the jets. In particular, the b-tagging algorithms exploit the long lifetime of hadrons, high mass and high decay multiplicity of hadrons containing a b-quark and the presence of reconstructed vertices with tracks that have larger impact parameters compared to the tracks originated by the primary vertex of the collision.

The b-tagging algorithm used in ATLAS during Run 2 is MV2c10 [3], a Boosted Decision Tree (BDT) algorithm that combines the input of the algorithms developed during Run 1: IP3D, based on the track impact parameters, SV1 and JetFitter, based on the properties of a displaced vertex reconstructed inside the jet. The c-jet fraction of the training for MV2c10 is 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.

2. b-tagging calibrations

In general, the b-tagging calibrations in ATLAS are based on $t\bar{t}$ events, as the decay of a $t\bar{t}$ pair has a clear signature and the top quark decays into a W-boson and a b-quark, providing a large and pure b-jet sample. In particular, the standard b-tagging calibrations are based on $t\bar{t}$ di-leptonic events, in which both the W-bosons decay leptonically, since dilepton channel is the purest among $t\bar{t}$ decay and allows to have smaller systematic uncertainties in the measurements with respect to the others.

Two different methods are used on $t\bar{t}$ dileptonic events to calibrate the ATLAS b-tagging algorithms, the combinatorial likelihood method (dilepton $t\bar{t}$ PDF) [4] and the tag-and-probe method (dilepton $t\bar{t}$ T&P) [5] and they have been proven to provide precise calibrations.

During Run 1, a $t\bar{t}$ ℓ+jets based b-tagging calibration has been performed, using tag-and-probe method on single lepton $t\bar{t}$ events ($t\bar{t}$ SL T&P) [6]. Exploiting a semileptonic $t\bar{t}$ decay, it is possible to extract data-to-simulation scale factors at higher transverse momentum p_{T} of the jet and reduces the statistical uncertainties on the calibration, especially for tighter b-tagging working points. In addition, it allows to explore an environment with high multiplicity of jets and gives the possibility to study the effect of the close-by jet activity on the b-tagging performance.

3. Measurement of the b-tagging efficiency in ℓ+jets events

The calibration of the MV2c10 algorithm has been performed using $t\bar{t}$ ℓ+jets candidate events collected by the ATLAS detector during 2015 and 2016 [7]. In order to improve the purity of the sample and reduce the background contamination, a selection has been applied to the events, requiring a final state with one charged lepton, missing transverse momentum and at least four reconstructed jets. Then, a kinematic fitter is used to reconstruct the $t\bar{t}$ decay and accordingly assign the jets to the leptonic b-jet, the W-hadronic jets and the hadronic b-jet. The hadronic b-jet is considered as the probe jet in the application of the tag-and-probe method on the $t\bar{t}$ ℓ+jets candidate events. The hadronic side of the $t\bar{t}$ decay is choosen since it provides higher jet multiplicity near the probe jet, so it is more suitable to study the close-by jet activity on the b-tagging performance. The b-tagging efficiency in data is calculated taking into account the estimated fractions of b-, c- and light flavour jets in the simulation, the fraction of jets tagged by MV2c10 and the fraction of non-prompt leptons background measured in data, as well as the mistag efficiencies for c-, light flavour jets and the non-prompt lepton background.

4. Results

The b-tagging efficiencies and the relative data-to-simulation scale factors are measured as a function of the transverse momentum p_T and the absolute value of the pseudo-rapidity $|\eta|$ of the probe jet, but also as a function of the angular separation between the probe jet and its nearest neighbouring jet ΔR^{min}. This quantity does not need any b-tagging requirement to be calculated and provides information on the close-by jet activity near the probe jet. The b-tagging efficiency are shown in Figure 1, while the corresponding scale factors are presented in Figure 2. In addition, Figure 3 shows a comparison between the results obtained with the tag-and-probe method on $t\bar{t}$ single lepton events and the standard $t\bar{t}$ dilepton-based calibrations as a function of the transverse momentum of the jet. All the results are shown for a representative 77% working point. The calibration scale factors are compatible with unity within their combined statistical and systematic uncertainties.

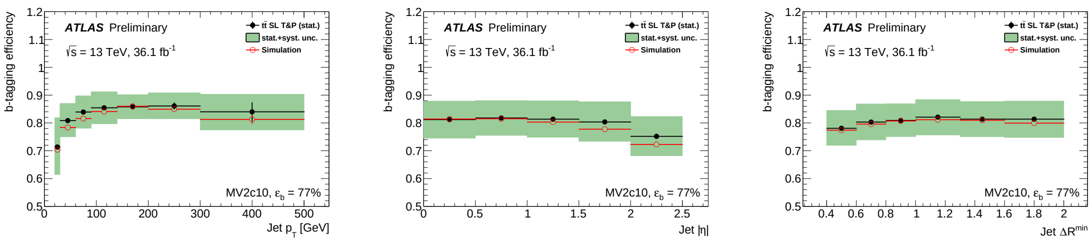

Figure 1: b-tagging efficiencies for the MV2c10 algorithm at the 77% working point [7] as a function of the transverse momentum (left), the absolute value of the pseudo-rapidity (middle) of the probe jet and the angular separation between the probe jet and its nearest neighbouring jet (right). The vertical error bars represent the statistical uncertainty.

5. Conclusions

A measurement of the b-tagging efficiencies and corresponding calibration scale factors of the MV2c10 algorithm has been performed on $t\bar{t}$ single lepton candidate events, providing an environment characterized by relatively high jet multiplicity that is suitable to test the close-by jet

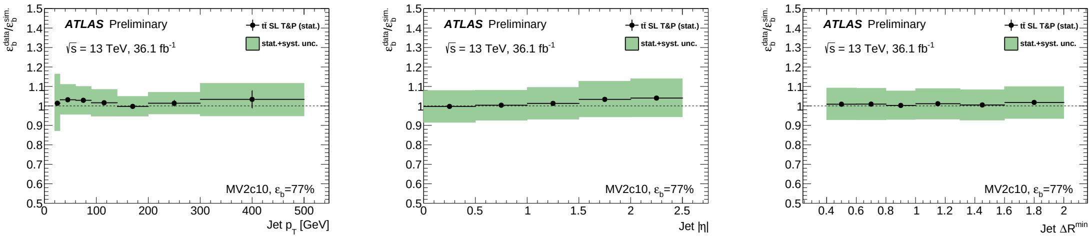

Figure 2: Data-to-simulation scale factors for the MV2c10 algorithm at the 77% working point [7] as a function of the transverse momentum (left column), the absolute value of the pseudo-rapidity (middle column) of the probe jet and the angular separation between the probe jet and its nearest neighbouring jet (right column). The vertical error bars represent the statistical uncertainty.

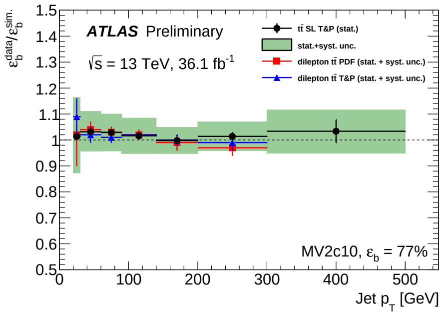

Figure 3: Comparison between the data-to-simulation scale factors for the MV2c10 algorithm obtained using the tag and probe method (T&P) applied to $t\bar{t}$ single lepton (SL) candidate events and those obtained using $t\bar{t}$ dilepton events [7, 8, 9]. The vertical error bars represent the statistical uncertainty for T&P SL, while indicate the total statistical and systematic uncertainties for the $t\bar{t}$ dilepton calibrations.

activity on the *b*-tagging performance. The results of the $t\bar{t}$ calibration using ℓ+jets events are compatible with the results obtained by using dilepton $t\bar{t}$ events with PDF and T&P methods. This technique allows the calibration to be extracted for jets of up to 500 GeV, beyond the reach of the dilepton $t\bar{t}$ calibrations.

References

[1] ATLAS Collaboration, JINST 3 (2008) S08003.

[2] L. Evans and P. Bryant (editors) JINST 3 (2008) S08001.

[3] ATLAS Collaboration, ATL-PHYS-PUB-2016-012, https://cds.cern.ch/record/2160731.

[4] ATLAS Collaboration, ATLAS-CONF-2014-004, https://cds.cern.ch/record/1664335.

[5] ATLAS Collaboration, ATLAS-CONF-2012-097, https://cds.cern.ch/record/1460443.

[6] ATLAS Collaboration, ATLAS-CONF-2016-001, https://cds.cern.ch/record/2127958.

[7] ATLAS Collaboration, FTAG-2017-003, http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/FTAG-2017-003/.

[8] ATLAS Collaboration, FTAG-2016-003, http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/FTAG-2016-003/.

[9] ATLAS Collaboration, FTAG-2017-001,
http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/FTAG-2017-001/.