

Measurement of c -jet tagging efficiency in ATLAS with $W+c$ -jet events

Alessandro Lapertosa*[†]

Universita e INFN Genova (IT)

E-mail: alessandro.lapertosa@cern.ch

The correct identification of the jet flavour (Flavour Tagging) is of fundamental importance for hadron collider experiments such as ATLAS and CMS at the Large Hadron Collider. In particular, it is an important tool for many physics analyses with heavy flavour jets in the final state (Higgs and top quark physics and SUSY searches). The new b -tagging algorithms developed for Run 2 improve the c -jet rejection with respect to Run 1 performance. A method to measure the efficiency to mistakenly tag a c -jet as a b -jet is presented. It uses a sample of c -jets produced in association with a W boson, with 2011 data collected by the ATLAS detector. The $W+c$ -jet sample, with the W boson reconstructed via its decay into an electron and a neutrino and the c -jet identified via a soft muon stemming from a semileptonic c -hadron decay, allows to select a high-purity c -jet sample, exploiting the charge correlation between the electron and the muon.

*38th International Conference on High Energy Physics
3-10 August 2016
Chicago, USA*

*Speaker.

[†]on behalf of the ATLAS collaboration.

1. Introduction

ATLAS is a multi-purpose experiment located at the CERN Large Hadron Collider. The correct identification of jets containing b -hadrons (hence b -tagging) is of capital importance for the ATLAS physics programme. Measurements that benefit from b -tagging include Standard Model analyses, like Higgs and top-quark physics, as well as searches for new physics phenomena beyond the Standard Model.

Exploiting some typical properties of the jets, it is possible to separate jets coming from a beauty quark (b -jets), from a charm quark (c -jets) or from a gluon or a up, down or strange quark (light jets). The ATLAS experiment developed b -tagging algorithms in order to separate b -jets from c -jets and light jets: the most recent algorithms are introduced in Sec. 2.

The algorithm's performance is first measured on Monte Carlo simulated samples (MC) containing top-quark pairs, then the efficiency to identify a b -jet is separately measured on simulated samples and on collision data, thus extracting the calibration in terms of data-to-MC scale factors for different b -tagging working points. The b -tagging efficiency is measured with different methods based on selected samples of b -jets, c -jets and light jets: the measurement of the b -tagging efficiency on a sample of c -jets produced in association with W bosons is described in Sec. 3.

2. ATLAS b -tagging algorithms

The identification of b -jets is performed with several algorithms, exploiting the long lifetime, high mass and high decay multiplicity of b -hadrons. The b -hadron decay inside the jet produces a secondary vertex, whose tracks have large impact parameters with respect to the tracks stemming from the primary vertex of the collision. Due to similar properties, c -jets are easily misidentified as b -jets.

During Run 1, the ATLAS experiment developed different algorithms able to identify b -jets. The three basic b -tagging algorithms are: SV1, based on the secondary vertex reconstruction; IP3D, based on the large impact parameters of the tracks stemming from the secondary vertex and JetFitter, based on the reconstruction of the topological structure of the b -hadron and c -hadron decays inside the jet. The output variables of these basic algorithms were combined into a Neural Network to obtain a final discriminant variable: MV1 [1]. Adding a c -jet component into the training sample as background hypothesis along with light jets produced an improvement of the c -jet rejection performance. The resulting enhanced variable is referred to as MV1c.

Also Run 2 b -tagging discriminant variables (MV2) are produced combining a set of sensitive variables with multivariate techniques. In particular, MV2 discriminants are produced processing 24 input variables through a Boosted Decision Tree. The performance for several background mixtures of c -jets and light jets in the training sample has been evaluated. During 2015, the reference b -tagging discriminant variable was MV2c20 [2], trained with b -jets as signal and a mixture of about 80% light jets and 20% c -jets as background.

The MV2c20 performance in terms of c -jet and light jet rejection is significantly improved with respect to the Run 1 discriminant variable, thanks to the addition of an extra pixel layer in the detector (the Insertable B-Layer [3]) and thanks to several enhancements to the tracking and b -tagging algorithms. The IBL grants a better performance for jets in the low p_T region, because

of the improved impact parameter resolution of tracks up to 10 GeV, while the enhancements in tracking and b -tagging algorithms result in a better performance for jets in the medium and high p_T regions.

Many physics analyses were found to be more limited by the low c -jet rejection than by the light jet rejection performance. For this reason, a new variable (MV2c10 [4]) has been developed for 2016, with the optimization of the multivariate parameters. Thanks to these enhancements, the b -tagging performance of MV2c10 (obtained with about 93% light jets and 7% c -jets as background training sample) is expected to improve, granting an improvement in the c -jet rejection (up to 40%) for a wide range of jet transverse momentum at the 77% b -tagging efficiency working point with respect to MV2c20 performance.

3. Measurement of the b -tagging efficiency on c -jets

To perform the measurement of the b -tagging efficiency on c -jets, a high purity c -jet sample is needed. During Run 1, the b -tagging efficiency was measured on a sample of c -jets containing charged D^* mesons [5, 6]. The main disadvantage of this method is the high uncertainty due to the generalization of the result to an inclusive sample of c -jets, since D^* mesons occur in only a limited subset of c -jets.

A different method was also developed, based on a sample of c -jets produced in association with a W boson [7]. This measurement was performed with 7 TeV data collected by the ATLAS detector during 2011. The W boson is reconstructed via its decay into an electron and a neutrino, and the c -jet is identified via a soft muon stemming from the semi-leptonic c -hadron decay. The c -jet sample is extracted exploiting the electron-muon charge correlation, separating events with opposite sign (OS) and same sign (SS) charged leptons. Due to the associated production, the $W+c$ signal has mainly OS leptons, while backgrounds almost evenly populate the OS and SS lepton samples. The main background contribution comes from multi-jet production and from the production of a W boson in association with a light jet, both estimated with data-driven methods. After the OS-SS subtraction, the remaining sample is mainly composed of $W+c$ -jet events (90%), with a small contamination of W +light and multi-jet events.

The efficiency, defined as the fraction of $W+c$ signal events passing the b -tagging requirement, is measured separately for data and MC. The data efficiency is measured taking into account the possibility to tag background events, while the MC efficiency is measured directly on $W+c$ signal events. The data-to-MC scale factors are calculated as the ratio of data to MC efficiencies, for each of the b -tagging variable working points.

This measurement is based on a particular sample of c -jets, where the c -hadron inside the c -jet decays semi-leptonically. Since the selection of semi-leptonically decaying c -hadrons has an effect on the probability to select different types of c -hadrons, the composition of this sample differs slightly with respect to an inclusive c -jet sample. Extrapolation coefficients are determined for data and MC, in order to extrapolate the results to be valid for a general sample of c -jets. The efficiency measurement performed with $W+c$ events has the benefit of applying to all c -jets containing c -hadrons decaying semi-leptonically, which requires significantly less extrapolation to inclusive c -jets than the D^* analysis.

The b -tagging efficiency of the MV2c10 variable will be measured on c -jets with $W+c$ events, using 13 TeV data collected by the ATLAS detector during 2015 and 2016, to confirm the expected improvements in the c -jet rejection of MV2c10 with respect to MV1.

4. Conclusion

Beauty quarks are present in the final states of many interesting physics processes: Higgs boson decay, top-quark decay and searches of particles beyond the Standard Model. The identification of jets containing b -hadrons is of fundamental importance for many measurements performed at high energy physics experiments. The ATLAS experiment developed specific algorithms able to discriminate between b -jets, c -jets and light jets, with the purpose of identifying interesting b -jets and reject c -jets and light jets as background. The most recent algorithms developed for Run 2 (MV2) are expected to have a better performance with respect to the one used during Run 1 (MV1). The improvements are mainly due to the addition of an extra pixel layer and improvements of the tracking and b -tagging algorithms. The b -tagging efficiency of MV1 algorithm on c -jets was measured with $W+c$ events, selecting a sample of c -jets produced in association with a W boson. The measurement was performed with 7 TeV data collected during 2011: the same method will be used to measure the MV2c10 efficiency on c -jets with 13 TeV data collected during 2015 and 2016 by the ATLAS detector.

References

- [1] The ATLAS Collaboration, Performance of b -jet identification in the ATLAS Experiment, JINST11 2016 P04008, CERN-PH-EP-2015-216, arXiv:1512.01094 [hep-ex].
- [2] The ATLAS Collaboration, Expected performance of the ATLAS b -tagging algorithms in Run II, ATL-PHYS-PUB-2015-022 (2015).
- [3] The ATLAS Collaboration, ATLAS Insertable B-Layer Technical Design Report, CERN-LHCC-2010-013, ATLAS-TDR-19.
- [4] The ATLAS Collaboration, Optimisation of the ATLAS b -tagging performance for the 2016 LHC Run, ATL-PHYS-PUB-2016-012 (2016).
- [5] The ATLAS Collaboration, b -jet tagging calibration on c -jets containing D^{*+} mesons, ATLAS-CONF-2012-039 (2012).
- [6] The ATLAS Collaboration, Calibration of the performance of b -tagging for c and light-flavour jets in the 2012 ATLAS data, ATLAS-CONF-2014-046 (2014).
- [7] The ATLAS Collaboration, Calibration of the b -tagging efficiency for c jets with the ATLAS detector using events with a W boson produced in association with a single c quark, ATLAS-CONF-2013-109 (2013).