

b-TAGGING CALIBRATION USING $t\bar{t}$ EVENTS WITH THE ATLAS EXPERIMENT

A. LEYKO
on behalf of the ATLAS Collaboration

*Physikalisches Institut, Rheinische Friedrich-Wilhelms-Universität Bonn,
Nussallee 12, 53115 Bonn, Germany*

Results of b -tagging calibrations measured in 35 pb^{-1} of data collected by the ATLAS detector from the Large Hadron Collider pp collisions at $\sqrt{s} = 7 \text{ TeV}$ in 2010 are presented. Two different approaches, using jets containing muons and jets originating from top quark pair event candidates are discussed. The measurement on the dijet sample was performed using the *pTrel* and *system8* methods, while in the sample of top quark pairs a *kinematic selection* and a *tag counting* method were used. The results are provided in the form of data-to-simulation scale factors.

1 Introduction

Many physics analyses at the LHC expect to have jets originating from b -quarks and the algorithms that allow to identify those jets are of great importance¹. It is thus crucial to understand their performance. In the final state of top quark pair decays at least two b -jets are present. This b -enriched sample provides a perfect environment for calibration of b -tagging algorithms for analyses with large multiplicity of high p_T jets, for example Higgs or SUSY searches. This approach takes advantage of the large cross-section of the top quark pair production at the LHC and a good understanding of this process after the initial phase of data taking with the ATLAS detector. To measure the b -tagging efficiency in the single-lepton channel a modified tag-and-probe method is applied to top quark pair events selected from data. Alternatively, in both single-lepton and dilepton channels, the number of b -tagged jets per event can be counted. The latter also provides an estimation of the top quark pairs production cross-section.

2 b -tagging algorithms

b -tagging algorithms are designed to identify reconstructed jets originating from b -quarks. A b -quark hadronizes to a B -hadron, which decays and forms a secondary vertex that can be reconstructed in the tracker of the ATLAS detector. The SV0 b -tagging algorithm², which uses the presence of an inclusively reconstructed secondary vertex to separate b -jets from light-flavour jets, will be used to demonstrate the performance of various calibration methods in these proceedings. However, there is a variety of algorithms calibrated in the ATLAS experiment, such as IP3D³ which relies on the impact parameter of the tracks associated to jets. There are also more advanced taggers that are combination of other tagging algorithms, IP3D+JetFitter³ and IP3D+SV1³.

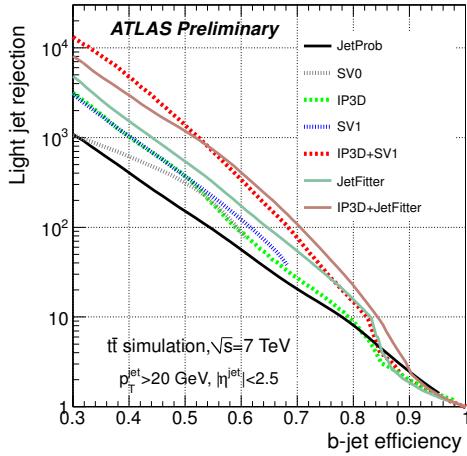


Figure 1: Expected rejection rate as a function of the b -jet tagging efficiency for various b -tagging algorithms³.

The performance of a b -tagging algorithm can be assessed by studying the rate with which the algorithm rejects light-flavour jets for a given efficiency to tag a jet originating from a b -quark. The rejection rate is defined as the ratio of all light-flavour jets to those tagged by the b -tagging algorithm. Figure 1 shows rejections rate as a function of b -tagging efficiency for various b -tagging algorithms.

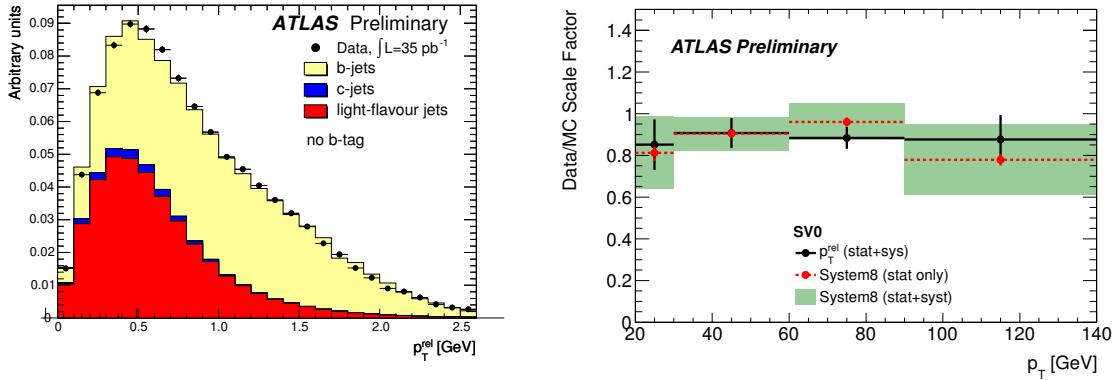


Figure 2: Left: Example of a template fit to the p_T^{rel} distribution in data before b -tagging. All jets with $p_T > 20$ GeV are considered in the fit and uncertainties shown are for data statistics only⁴. Right: Data-to-simulation b -jet tagging efficiency scale factors as a function of jet p_T , obtained with *system8* and $pTrel$ methods, for the SV0 tagging algorithm, measured in 35 pb^{-1} of data collected in 2010⁵.

3 Calibration with jets containing muons

In approximately 20% of B -hadron decays there will be a muon present. As shown in Figure 2 (left) muons originating from b -quarks have in average higher p_T^{rel} , where p_T^{rel} denotes the muon transverse momentum p_T with respect to the $\mu + \text{jet}$ direction. In the $pTrel$ method⁴ the b -tagging efficiency is extracted by fitting templates of p_T^{rel} for b -, c - and light-flavour jets to data before and after tagging with the algorithm under test. The templates are obtained from Monte-Carlo simulations.

The *system8* method⁵ uses the same sample of jets with reconstructed muons, however is applying additional selection criteria, which divide it into three subsamples with different b -purity. This

allows for a construction of a system of eight equations from which the efficiency to tag a b -jet can be extracted.

The results of both calibration methods in form of data-to-simulation scale factors are presented in Figure 2 (right). Both approaches rely on low p_T jet triggers to enable the measurement for the low p_T jets, which has some statistical limitation. Moreover for high p_T jets the p_T^{rel} variable is not anymore a good discriminant between b -, c - and light-jets. And thus the methods using jets containing muons cannot be applied to jets with very high p_T .

4 Calibration with $t\bar{t}$ events

The signature of a top quark pair decay is characterized by at least two b -jets and at least one lepton (e or μ). These events are easy to select and the production rate of $t\bar{t}$ events at LHC is sufficiently large due to the high center-of-mass energy of the pp collisions. These factors make this sample a perfect environment for b -tagging efficiency studies.

The measurement was performed in dilepton channel where both W -bosons from $t \rightarrow Wb$ decay leptonically (exactly two high p_T isolated leptons required) and single-lepton channel where one of the W -bosons decays hadronically (exactly one high p_T isolated lepton required). These two samples are statistically uncorrelated.

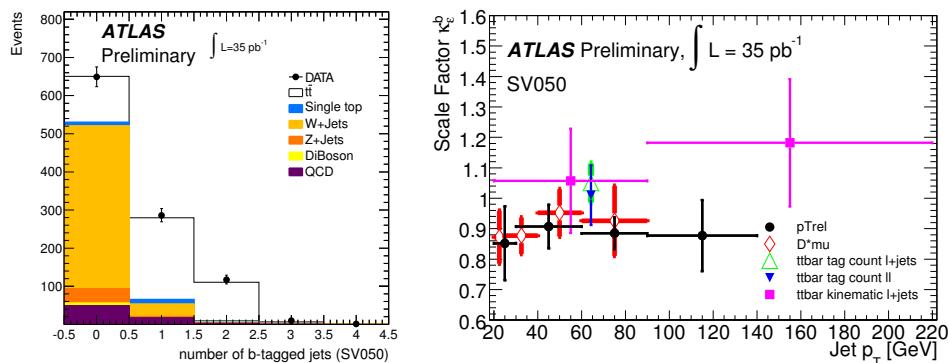


Figure 3: Left: Number of observed and expected events in the single-lepton channel with n jets tagged by the SV0 tagger in the *tag counting* measurement⁴. The $t\bar{t}$ signal is normalized to the NNLO cross-section. Right: The measured data-to-simulation b -tagging efficiency scale factors for the $pT\text{rel}$ and $t\bar{t}$ methods for the SV0 algorithm⁴. The data points are placed in the middle of the jet p_T bin, except for the *tag counting* methods that do not provide the measurement as a function of p_T . For these the scale factor points are instead placed at the average jet p_T in the $t\bar{t}$ sample.

With the *tag counting* method the cross-section of top quark pair production and b -tagging efficiency can be simultaneously measured. It relies on the fractions F_{ijk} of events with i b -jets, j c -jets and k light-jets which are derived from Monte-Carlo simulations separately for all contributing processes¹. After combining those factors the expected number of events with n tagged jets can be estimated and compared with the value measured in data which allows to calculate the tagging efficiency⁴. The number n of tagged jets for the single lepton channel is shown in Figure 3.

The *kinematic selection* method directly uses the knowledge about the jet flavour composition in the $t\bar{t}$ signal and the background sample. This information is used to calculate the jet flavour fractions and tagging efficiencies of non- b -jets in the analyzed sample. The b -tag efficiency can then be extracted from the number of jets before and after tagging in data⁴. The *kinematic selection* method is applied in the single-lepton channel only. Here, additionally to the standard $t\bar{t}$ selection⁴, at least one jet is required to be tagged with the SV0 algorithm. This is done in

Table 1: Total relative systematic and statistical uncertainties on the b -jet tagging efficiency for the pT_{rel} , *system8, kinematic selection, tag counting* method in the single-lepton and dilepton channels.

method	pTrel	system8	kin. sel.	tag count. single-lepton	tag count. dilepton
Stat.	3-10%	1-5%	10-12%	7%	9%
Syst.	4-11%	4-12%	11-12%	5-9%	4%

order to increase the signal-to-background ratio and thus increase the b -purity of the sample. However the information about which jet was tagged in this preselection is not used in the further analysis. In the measurement of b -tagging efficiency only four jets with the highest p_T are taken into account, assuming they are coming from the top quark pair decay⁴.

The results of both $t\bar{t}$ based methods are presented in Figure 3 (right) in the form of data-to-simulation scale factors. For the *kinematic selection*, the scale factors are obtained as a function of jet p_T , whereas for the *tag counting* method only a global value for all jets from selected events is obtained.

5 Conclusions and outlook

The biggest difficulty to properly estimate the b -tagging efficiency with both approaches is that it strongly depends on the flavor composition of the analyzed sample. The flavor fractions are measured in the simulated events, which introduces a systematical uncertainty related to modeling of physics processes in the Monte-Carlo simulations.

The relative statistical and total systematic uncertainties for all calibration methods discussed in this study are listed in Table 1. The dominant uncertainties in the $t\bar{t}$ -based methods are the choice of simulation generator and modeling of initial and final state radiation. Another important contribution is the uncertainty on the background estimation, which is often obtained directly from data. Finally, there is an uncertainty on the simulation of the detector response, where the leading one is the jet energy scale and the resolution. More details about systematic uncertainties can be found in references^{4,5}.

With more data collected with the ATLAS detector and better understanding of the $t\bar{t}$ sample, the $t\bar{t}$ based b -tagging calibration methods are expected to achieve much lower statistical and systematic uncertainties as well as to reach much higher jet p_T range than is possible with the conventional methods based on jets containing muons.

References

1. ATLAS Collaboration, *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics*, <http://arxiv.org/abs/0901.0512>
2. ATLAS Collaboration, *Calibrating the b -Tag and Mistag Efficiencies of the SV0 b -Tagging Algorithm in 3 pb^{-1} of Data with the ATLAS Detector*, <http://cdsweb.cern.ch/record/1312145>
3. ATLAS Collaboration, *Commissioning of the ATLAS high-performance b -tagging algorithms in the 7 TeV collision data*, <http://cdsweb.cern.ch/record/1369219>
4. ATLAS Collaboration, *Calibrating the b -Tag Efficiency and Mistag Rate in 35 pb^{-1} of Data with the ATLAS Detector*, <http://cdsweb.cern.ch/record/1356198>
5. ATLAS Collaboration, *b -Jet Tagging Efficiency Calibration using the System8 Method*, <http://cdsweb.cern.ch/record/1386703>