PROCEEDINGS OF SCIENCE

Efficiency calibration for ATLAS b-jet identification algorithms

Angela Burger[a,*] **on behalf of the ATLAS collaboration**

[a] *Oklahoma State University,*
Department of Physics, 223 Physical Sciences Bldg, Stillwater, OK 74078, United States

E-mail: angela.maria.burger@cern.ch

Many analyses in ATLAS rely on the identification of jets containing b-hadrons (b-jets) at high efficiency while rejecting more than 99% of non-b-jets. Identification algorithms, called b-taggers, exploit b-hadron properties such as their long lifetime, their high mass, and high decay multiplicity to achieve this. Recently developed ATLAS b-taggers using neural networks are expected to outperform previous b-taggers by a factor of two in terms of non-b-jet rejection for the same b-tagging efficiency. The performance of these taggers is measured in data, and simulated LHC collision events are corrected to reflect the data performance through calibration scale factors. Due to recent improvements in measurement procedures, the data efficiency precision is at the level of a few percent for b-jet identification efficiency and at the level of 10–20% for light- and charm-jet mistag rates. The methods to calibrate the b-jet identification and the charm- and light-jet mistag efficiencies of the recent b-taggers and the calibration results will be presented in this poster.

40th International Conference on High Energy physics - ICHEP2020
July 28 - August 6, 2020
Prague, Czech Republic (virtual meeting)

*Speaker

https://pos.sissa.it/

PoS(ICHEP2020)799

ATL-PHYS-PROC-2020-101
09 November 2020

1. The ATLAS b-tagging algorithms

Many analyses in the ATLAS [1] experiment at the LHC, for example measurements of top quark or Higgs boson properties and searches for new physics, rely on algorithms identifying jets from b-hadrons (b-jets) with high efficiency while rejecting 99% of jets from charm and light hadrons. These algorithms exploit the distinct signatures of b-jets, like their long lifetime, high mass and high decay multiplicity [2].

Impact-parameter based low-level taggers use the impact parameter significances of the tracks in the jets to construct variables which are discriminating between the jet flavours, either using a template-based approach (`IP2D`, `IP3D`) or a recurrent neural network (`RNNIP`). The latter has been added recently and improves the light jet rejection by taking into account correlations between the tracks in the jet. Secondary vertex based low-level taggers reconstruct either a single displaced vertex (`SV1`) or the hadron decay chain (`JetFitter`). The output of these low-level taggers are fed into multivariate algorithms. A discriminant is constructed from the output probabilities for a given jet to originate from a light, charm or b-hadron. The current taggers for particle-flow jets are based on deep neural networks, either with the RNNIP low-level tagger included (`DL1r`) or without (`DL1`).

Several changes with respect to past b-tagger versions improve the performance: The light jet rejection has been improved by a factor of 2 and the charm jet rejection by a factor of 1.5 at a fixed average b-tagging efficiency of 70%. A new dedicated training on particle-flow jets, an extension of the jet p_{T} spectrum of the training sample, the use of the RNNIP low-level tagger and of deep neural networks are responsible for these improvements. Even further improvement of the performance can be achieved by replacing the RNNIP with an algorithm based on deep sets (DIPS) [3].

2. Calibration of the b-tagging algorithms

Mismodeling of tagger training input variables can cause differences in tagging efficiencies between data and simulation. b-jet efficiencies and charm and light mistag rates are measured in 139 fb^{-1} of ATLAS data and Monte-Carlo (MC) simulations and correction factors ("scale factors") for simulations are derived.

The b-efficiency calibration selects $t\bar{t}$ events decaying into two leptons, neutrinos, and exactly two jets. A combined template fit approach in signal and control regions with different flavour compositions of the jet pairs reduces the uncertainty to percent level [4]. The charm mistag rate measurement selects $t\bar{t}$ events decaying into one lepton and exactly four jets. A kinematic likelihood fitter assigns jets in the event to the $t\bar{t}$ decay products and the charm mistag calibration is extracted using a combined likelihood fit to the b-tagging multiplicity of the jets assigned to the $W \rightarrow jj$ decay [5]. The light mistag rate calibration is performed on the leading jet in $Z(\rightarrow \ell\ell)$+jets events. A 2-dimensional template fit is used to extract the calibration and to constrain the contamination by other jet flavours. The high rejection of light jets of the b-tagging algorithms makes a direct calibration impossible and a modified tagger with reduced b-jet efficiency and largely unchanged light jet response is calibrated and an additional uncertainty is applied.

3. Results and post-processing

Scale factors are defined as the ratio between data and MC efficiencies. The results in Figure 1 show that data and MC efficiencies are compatible. Uncertainties are at percent level (1–2%) for the b-tagging efficiency, at the level of a few percent for the charm mistag, and of the order of 10–20% for the light jet mistag calibration. Data-driven corrections in the case of b and light jet calibrations and an improved inner detector simulation in case of the light jet calibration reduce the uncertainties.

The scale factors and uncertainties are smoothed as a function of the jet p_T using a non-parametric regression technique to remove discontinuities at bin boundaries. The smoothing parameters are optimized using a statistical approach [6]. The measured calibration is extrapolated beyond a jet p_T of 400 GeV. Additional uncertainties due to physics and detector modeling are added to the measured uncertainties at the smoothed central value of p_T = 400 GeV.

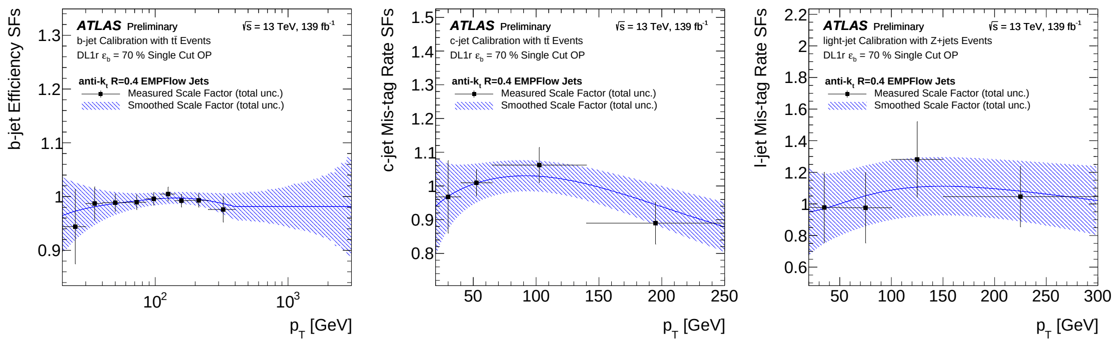

Figure 1: Scale factors for b-jet (left), charm (center) and light jet mistag (right) efficiencies for a fixed average b-tagging efficiency of 70%. Measured scale factors are represented by black dots and post-processed results by the blue shaded area. The results have been obtained with particle-flow jets. From Ref. [7].

4. Conclusion and outlook

ATLAS analyses profit from an improved b-tagging performance. The b-tagging efficiencies and light and charm mistag rates are measured, the results in data are compatible with those in MC simulation. Calibration factors are calculated and their uncertainties are at the percent level for the b-efficiency calibration and of the order of 10–20% for the light jet mistag calibration. There is constant effort in the ATLAS experiment at the LHC to improve the taggers and the calibrations, to extend the calibration to new phase spaces, and to explore new algorithms and methods.

References

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

[2] ATLAS Collaboration, ATL-PHYS-PUB-2017-013.

[3] ATLAS Collaboration, ATL-PHYS-PUB-2020-014.

[4] ATLAS Collaboration, *Eur. Phys. J. C* **79** (2019) 970

[5] ATLAS Collaboration, ATLAS-CONF-2018-001.

[6] ATLAS Collaboration, ATL-PHYS-PUB-2020-004.

[7] ATLAS Collaboration, FTAG-2020-001 http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/FTAG-2020-001/.