PROCEEDINGS OF SCIENCE

Boosted $\mathrm{H} \rightarrow \mathrm{b}\bar{\mathrm{b}}$ Tagger in Run II

Merve Sahinsoy* **on behalf of the ATLAS Collaboration**
Kirchhoff Institute for Physics, Heidelberg University
E-mail: merve.sahinsoy@cern.ch

Many searches for Higgs bosons decaying to b quark pairs benefit from the increased Run II centre-of-mass energy by exploiting the boosted kinematic regime at large transverse momenta of the Higgs boson, where the two b-jets are merged into one large radius (R) jet. ATLAS [1] uses a boosted $H \rightarrow b\bar{b}$ tagger algorithm to separate Higgs signal from background processes (QCD, W and Z bosons, top quarks). The tagger takes as input a large $R = 1.0$ jet with calibrated pseudorapidity, energy and mass scale. It employs b-tagging, Higgs candidate mass, and substructure information. The performance of several operating points in Higgs boson signal, QCD, and $t\bar{t}$ all-hadronic backgrounds are presented. Systematic uncertainties are evaluated so that this tagger can be used in analyses.

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

*Speaker.

© Copyright owned by the author(s) under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0).

PoS(ICHEP2016)1129

1. Introduction

After the discovery of the Higgs boson in Run I, searches including the Higgs boson gained great prominence. Many searches are ongoing in Run II based on the new physics models which predict new resonances or Higgs partners [2]. Having the largest branching ratio of 57%, the $b\bar{b}$ decay channel provides a great opportunity to study processes involving Higgs bosons. Decay products of the Higgs boson are collimated proportionally to the transverse momentum (p_T) of the boosted Higgs. Thus, a boosted H boson can be identified as a large-radius (R) jet which contains a $b\bar{b}$ pair. The tagger selects large-R jets based on their p_T, mass and topological substructure, and it applies various b-tagging criteria to the small-R track jets associated to the large-R jet area. Several benchmarks are defined using different cuts on jet mass and substructure variables. In this paper, the baseline boosted $H \to b\bar{b}$ identification strategy for Run 2 data is described, and the performance of the tagger is presented [2].

2. Large-R Jet Reconstruction

Jets are reconstructed from topological clusters of calorimeter cell energy deposits. For large-R jet reconstruction, the anti-k$_t$ algorithm with a radius parameter $R = 1.0$ is used. Applying trimming with f$_{cut}$=0.05 and R_{subjet}=0.2 parameters [3], pile-up and underlying event contaminations are removed from the reconstructed large-R jet area. Muons from the semi-leptonic decays of b-hadrons are identified and the momentum of the large-R jet is corrected for their presence [2].

One of the most important observables to identify $H \to b\bar{b}$ jets is the large-R jet mass, which by default is calculated using the calorimeter information only, m^{calo}. An alternative jet mass definition, *track assisted jet mass* m^{TA} is also studied for the Higgs tagged jets [3], which gets benefits from the tracker. First, the total mass of the tracks which are associated to the large-R jet area is calculated, and then missing neutral components are corrected by the p_T ratio of the calorimeter jet to the sum of tracks. A correction for muons overlapping with the jet is also applied for m^{TA} by considering the correction in calorimeter jet p_T, while for m^{calo} the entire 4-momentum of the jet is corrected. In Figure 1, the effect of the muon correction is shown for Higgs jets mass.

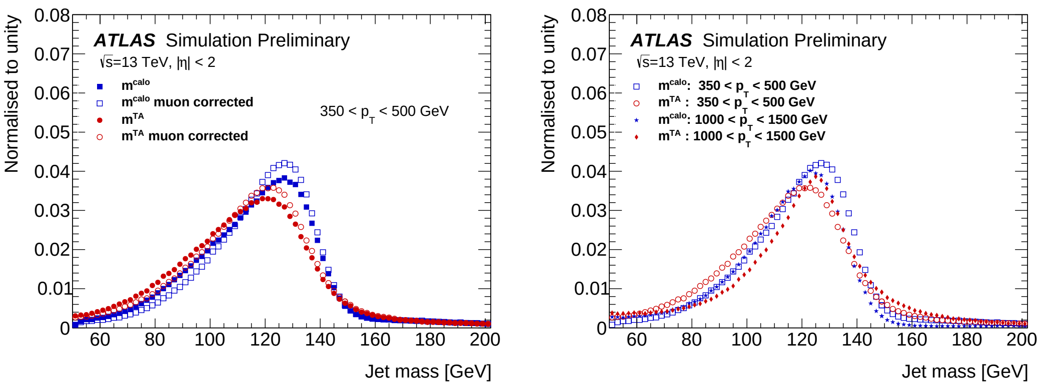

Figure 1: m^{calo} and m^{TA} distributions for Higgs jets before and after the muon corrections (left). m^{calo} and m^{TA} are compared in different p_T ranges after muon correction (right) [2].

PoS(ICHEP2016)1129

The Higgs tagger uses m^{calo} as default jet mass definition and two different mass window requirements are applied on it: 93-134 GeV, *tight mass window*, and 76-146 GeV *loose mass window*, which correspond to 68% and 90% of the groomed Higgs-jet mass distributions.

3. Track Jet b-tagging

For boosted H bosons, it is hard to identify b-jets using regular calorimeter jets with $R = 0.4$ because of the collimation of the $H \rightarrow b\bar{b}$. Therefore, track jets with $R = 0.2$ are reconstructed with the anti-k$_t$ algorithm and used for b-tagging. One or two track jets are required to be associated to the large-R jet area. Using several working points (WP) of the MV2c10 algorithm [4], track jets are identified as b-jets. The performance of the different track jet b-tagging selections are shown for the Higgs jet efficiency versus multi-jet and hadronic top background rejections in Figure 2.

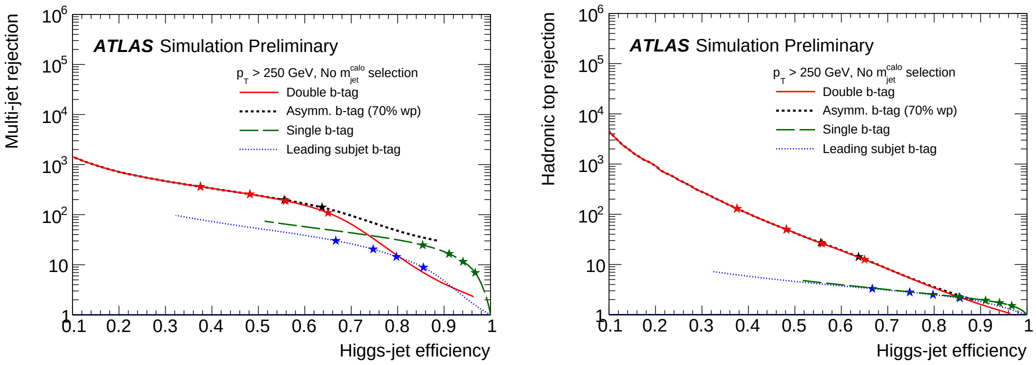

Figure 2: The rejection of inclusive multi-jets (hadronic top) versus Higgs-jet identification efficiency using large-R jets with $p_T > 250$ GeV on the left (right), for various b-tagging requirements. The stars correspond to the 60%, 70%, 77% and 85% b-tagging WPs from left to right [2].

4. Jet Substructure

The internal structure of a jet is a useful observable to discriminate Higgs jets from the multi-jets and hadronic top decays. The $D_2^{(\beta=1)}$ substructure variable [2], defined as a ratio of the two and three point energy correlation functions, is used in this analysis.

5. Performance and Validation of the Tagger

Several benchmarks are defined and provided together with the uncertainties. The performance of those benchmarks can be seen in Figure 3. In order to ensure an accurate modelling of the Higgs tagger in simulation, it is possible to cross-check the performance of b-tagging and evaluate the modelling of jet substructure variables in data by MC, using the $g \rightarrow b\bar{b}$ splitting process in QCD multi-jet events. Applying similar selections as for the Higgs tagger [2], kinematic distributions in relevant observables is studied. As a result of those studies, it is found that data and MC agree within systematical uncertainty. In Figure 4, the double b-tagging rate is shown for data and MC.

PoS(ICHEP2016)1129

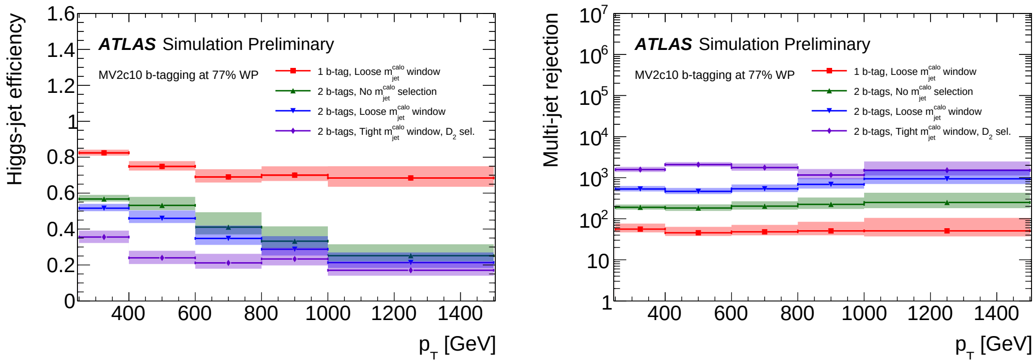

Figure 3: Higgs-jet signal efficiency (left) and multi-jet background rejection (right) as a function of the p_T of large-R jets. Statistical and systematic uncertainties are given by shaded bands [2].

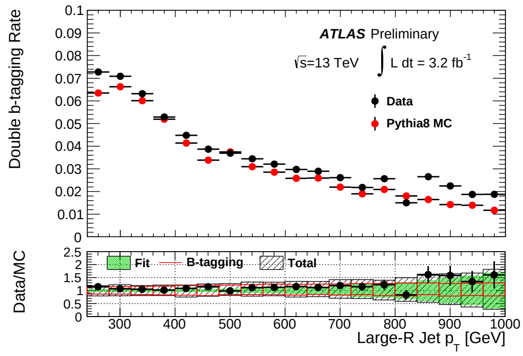

Figure 4: Double b-tagging rate for both data and MC as a function of the large-R jet p_T [2].

6. Conclusion

Several Higgs-tagger selections are studied and five working points are provided for physics analyses. Performance studies for jet mass and track jet b-tagging are presented. Data/MC comparisons are done and it is found that important variables for the Higgs tagger are reasonably well modelled in simulations for $\sqrt{s}$=13 TeV ATLAS Run 2 Data.

References

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

[2] ATLAS Collaboration, Boosted Higgs → $b\bar{b}$ Boson Identification with the ATLAS Detector at $\sqrt{s}$ =13 TeV, ATLAS-CONF-2016-039.

[3] ATLAS Collaboration, Jet Mass Reconstruction with the ATLAS Detector in Early Run 2 Data, ATLAS-CONF-2016-035.

[4] ATLAS Collaboration, Performance of b-Jet Identification in the ATLAS Experiment, JINST 11 (2016) P04008, arXiv: 1512.01094 [hep-ex].

PoS(ICHEP2016)1129