PROCEEDINGS OF SCIENCE

Development of a new Soft Muon Tagger for the identification of b-jets in ATLAS

Andrea Sciandra[*†]
On behalf of the ATLAS Collaboration
Physikalisches Institut, University of Bonn, Germany
E-mail: andrea.sciandra@cern.ch

B-tagging plays a fundamental role at LHC, as it helps in the identification of heavy particles that decay to bottom quarks, as the top quark and the Higgs boson or heavy exotic particles. The Soft Muon Tagger (SMT) allows to identify jets from b-quarks taking advantage of the presence of a muon coming from semileptonic decays of b-hadrons. The development of this new b-tagger is described, showing that, despite the low efficiency of the jet-muon association based on the angular distance, the discriminating power of the associated muon variables is remarkable to reject light jets. An enhanced performance has been reached for all light jet rejection working points by adding the SMT output to the best performing multivariate b-tagger in ATLAS (MV2). A good modeling of input and output variables is shown, comparing simulation with Run 2 data.

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

[*] Speaker.
[†] Supported by European Research Council grant ERC-CoG-617185.

PoS(EPS-HEP2017)768

1. Introduction

The Soft Muon Tagger (SMT) is based on the reconstruction of muons coming from semileptonic decays of heavy-flavour hadrons. An earlier version of the SMT was based on a simple cut on a χ^2 distribution [1]. These muons usually have a sizable transverse momentum (though smaller than the typical p_{T} of leptons from electroweak bosons decays, hence the label "Soft"), as well as a large transverse momentum relative to the jet axis, $p_{\mathrm{T}}^{\mathrm{rel}}$. The presence of a muon is enhanced in b-jets with respect to c- and light-flavour jets due to the significant semileptonic decay branching ratio of b hadrons ($BR(b \to \mu \nu X) \approx 11\%$) and c-hadrons produced by the b-hadron decay (sequential semileptonic decay, $BR(b \to c \to \mu \nu X) \approx 10\%$) [2].

2. Selection and composition

The new SMT is based on calorimetric jets and combined muons, i.e. muon candidates with matching tracks reconstructed in ATLAS Inner Detector (ID) and Muon Spectrometer (MS) systems [3]. Muons are associated to the closest selected jet by requiring an angular separation between the muon and the jet-axis of $\Delta R < 0.4$. Muons are required to have $p_{\mathrm{T}} > 5$ GeV, $d_0 < 4$ mm and $|\eta| < 2.5$. Since minimum ionising particles lose on average ~ 3 GeV in the ATLAS calorimeter system, the efficiency for reconstructing muons with p_{T} below this value is low. The fraction of b-jets in the $t\bar{t}$ simulated sample described in Ref. [4] with a reconstructed muon passing the requirements listed above is $\sim$12%. There are three main background sources in light jets that give rise to muon candidates passing these requirements and they are evaluated in the $t\bar{t}$ simulated sample: prompt muons from the nearby W boson randomly associated to light jets ($\sim 1\%$ contamination), muons coming from the decay in flight of light hadrons, mostly pions and kaons ($\sim 1\%$ contamination) and energetic hadrons (“punch-through”) that travel through the calorimeter system and reach the MS ($< 0.1\%$ contamination).

3. Input variables

Three kinematic variables separating muons in light jets from those from b- or c- hadron decays are used:

- ΔR: angular distance between the muon and the associated jet;
- $p_{\mathrm{T}}^{\mathrm{rel}}$: orthogonal projection of the muon p_{T} onto the jet axis;
- d_0: muon impact parameter measured with respect to the interaction primary vertex.

The $p_{\mathrm{T}}^{\mathrm{rel}}$ value of decay products is correlated to the parent particle mass. Therefore, muons from direct b decays tend to have a larger $p_{\mathrm{T}}^{\mathrm{rel}}$ value than those from background. Muons from the cascade $b \to c$ decay are more difficult to identify since they tend to be softer in the $p_{\mathrm{T}}^{\mathrm{rel}}$ spectrum. Muons from π and K decays in flight have tracks that do not in general extrapolate close to the primary vertex. Nevertheless, if the decay occurs at a small angle, the decaying hadron and the muon can be reconstructed as a single track and the muon candidate can pass the track selection.

In addition to ΔR, d_0 and $p_{\mathrm{T}}^{\mathrm{rel}}$, the SMT algorithm makes use of three observables defining the quality of the muon track. The scattering neighbour significance ($\mathcal{S}$) is computed by considering pairs of adjacent hits along the track and evaluating the significance of the angular difference $\Delta\phi$ between the two half tracks ending/starting at each of them, summing significances up over the whole track times the particle charge (q):

$$\mathcal{S} = q \times \sum_{\mathrm{i}} \frac{\Delta\phi^{i}_{\mathrm{scat}}}{\sigma_{\Delta\phi^{i}_{\mathrm{scat}}}}.$$

A higher significance is more likely to correspond to a π or K decay in flight. The momentum imbalance significance ($\mathcal{M}$) is defined as follows:

$$\mathcal{M} = \frac{p_{\mathrm{ID}} - p^{\mathrm{extr}}_{\mathrm{MS}}}{\sigma_{E_{\mathrm{loss}}}},$$

where p_{ID} is the muon momentum measured by the Inner Detector, $p^{\mathrm{extr}}_{\mathrm{MS}}$ is the momentum measured by the MS and extrapolated to the vertex and $\sigma_{E_{\mathrm{loss}}}$ is the uncertainty on the energy loss measured by the calorimeters. Another quantity sensitive to muons originating from these decays through the p_{ID} to p_{MS} comparison is:

$$\mathcal{R} = \frac{(q/p)_{\mathrm{ID}}}{(q/p)_{\mathrm{MS}}},$$

where $(q/p)_{\mathrm{ID}}$ is the charge-to-momentum-ratio, i.e. track curvature, measured by the Inner Detector and $(q/p)_{\mathrm{MS}}$ is the same measured by the MS. Figure 1 shows SMT discriminating variables for b-, c- and light-flavour jets in simulated $t\bar{t}$ events.

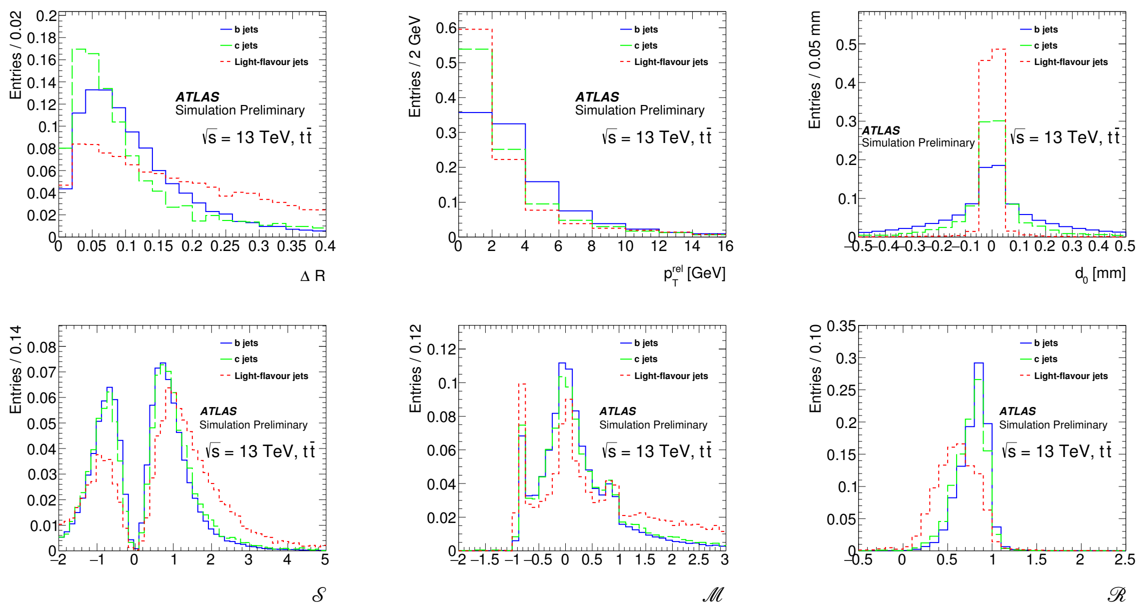

Figure 1: Normalised distributions of the six discriminating variables for reconstructed muons associated to b-jets (blue), c-jets (green) and light-flavour jets (red) [4].

4. Performance of SMT algorithm

In the SMT tagging the full set of six variables is used as input in a dedicated MVA [5], using a gradient-boosted BDT. Figure 2 shows the new discriminant output for *b*-, *c*- and light-flavour jets containing a muon candidate passing the SMT selection. By cutting on the SMT discriminant (BDT>-0.15), the efficiency for accepting a jet having a candidate SMT muon is 85% for *b*-jets and 15% for light-flavoured jets in $t\bar{t}$ events; for an overall *b*-jet tagging efficiency of 10%, the mistag rate is 0.02%. Figure 3 shows the event display of a $t\bar{t}$ dilepton candidate event where a soft muon is associated to a b-jet candidate tagged by the new BDT-based Soft Muon Tagger algorithm.

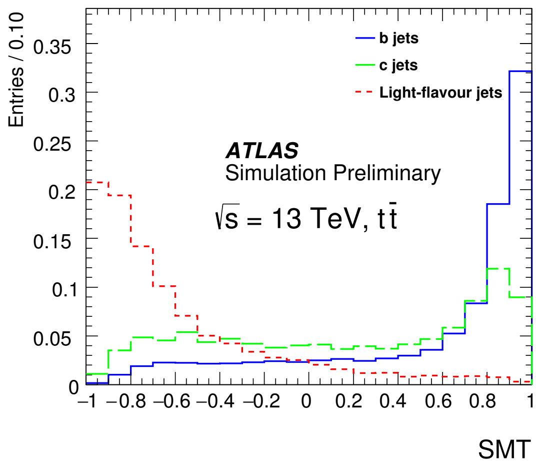

Figure 2: Normalised BDT response in simulated $t\bar{t}$ events of SMT for reconstructed muons associated to *b*-jets (blue), *c*-jets (green) and light-flavour jets (red) [4].

5. Modelling and implementation in MV2

Figure 4 shows data/MC comparison for the BDT output on $t\bar{t}$ events; good agreement between data and MC simulation is observed in the SMT input variables as well as in the final BDT discriminant. Details on the selection adopted for such data/MC comparison can be found in Ref. [4].

The MV2 *b*-tagging algorithm combines 24 input variables based on properties of impact parameter, secondary vertex and weak decay topology algorithms into a BDT. Among the variants of the MV2 taggers a new option has been developed, including in addition the SMT output ("MV2Mu"), and added for the 2017 data-taking campaign. The implementation of the SMT output itself as an additional input variable, as shown in Figure 5, leads to a $20-25\%$ improvement in light-jet rejection in the $70-85\%$ *b*-jet efficiency range (relevant for most of the physics analyses).

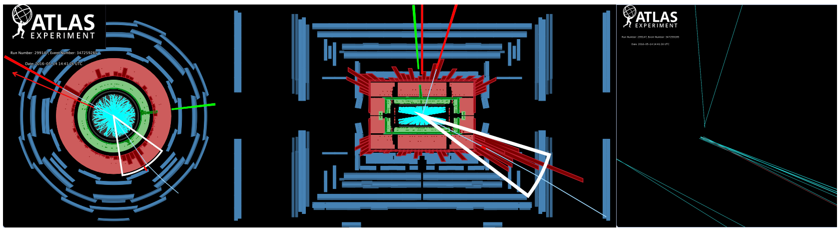

Figure 3: Event display of a $t\bar{t}$ dilepton candidate event from proton-proton collisions recorded by ATLAS with LHC stable beams at a collision energy of 13 TeV. In addition to two prompt leptons (an electron of 107 GeV and a muon of 268 GeV) a soft muon (7 GeV) is associated to a b-jet candidate (in $\Delta R < 0.4$) tagged by the new BDT-based Soft Muon Tagger algorithm. On the right a zoom on the reconstructed primary vertex (PV, two tracks) and secondary vertex (SV, six tracks) in the $z-\rho$ plane. The track of the muon within the jet tagged by new BDT-based SMT b-tagging algorithm is showed in red. The distance between PV and SV is about 2 mm [6].

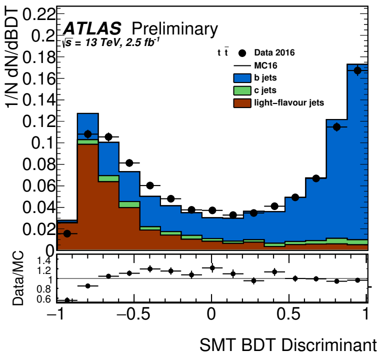

Figure 4: Data/MC comparison of the SMT BDT output for a sample dominated by $t\bar{t}$ dilepton events [4].

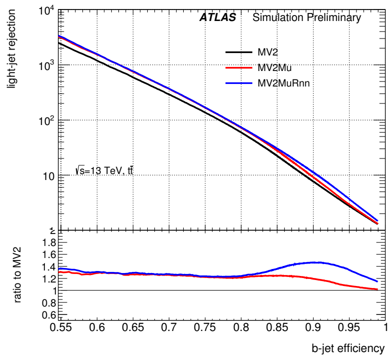

Figure 5: Light-flavour and c-jet rejection as a function of b-jet efficiency for MV2 (black line), MV2Mu (red line), MV2MuRnn (blue line). The latter is a variant of MV2 including both SMT and RNN inputs [7]. The algorithm evaluation is performed on $t\bar{t}$ events. The ratio reported on the bottom of the figure is calculated for two MV2 variants with respect to MV2. A $20-25\%$ improvement in light-jet rejection is expected in the $70-85\%$ b-jet efficiency range [4].

References

[1] ATLAS Collaboration, Performance of b-Jet Identification in the ATLAS Experiment, JINST 11 (2016) P04008.

[2] M. Acciarri et al., Measurement of the branching ratios $b \to e\nu X,\ \mu\nu X,\ \tau\nu X$ and νX, Zeitschrift für Physik C: Particles and Fields, Volume 71, Issue 1, pp. 379-389 (1996).

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

[4] ATLAS Collaboration, Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run, ATL-PHYS-PUB-2017-013.

[5] A. Hocker et al., TMVA - Toolkit for Multivariate Data Analysis, PoS ACAT (2007) 040, arXiv: physics/0703039 [physics.data-an].

[6] ATLAS Collaboration, Event Displays from Run 2 physics analyses, `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/EventDisplayRun2Physics`.

[7] ATLAS Collaboration, Identification of jets containing b-hadrons with Recurrent Neural Network at the ATLAS experiment, ATLAS-PHYS-PUB-2017-003.