1. ATLAS b-tagging

The identification of b-quark jets ("b-tagging") plays a vital role for the ATLAS experiment. MV2c10 is the ATLAS Run2 default tagger. It is a boosted decision tree (BDT) algorithm combining one impact parameter and two secondary vertex based algorithms. [1]

2. Event Selection

Event-level selection:
- Exactly 2 leptons: $\ell \ell$ $\mu \mu$
- Exactly 2 jets, $\geq$ 1 tagged

Probe jets definition:
- 1 tagged events $\rightarrow$ the other jet is a probe
- 2 tagged events $\rightarrow$ both jets are probes

After the selection more than 90% of the candidates originate from $t \bar{t}$ events.

3. Efficiency Measurement

The b-tagging efficiency for b-quark-initiated jets is measured from the number of probe jets in data before ($N_{\text{data}}$) and after ($N_{\text{data}}^{\text{pass}}$) the b-tagging requirement. Two MC-based corrections are applied to account for light and charm-jets in the selected probe samples, and for non-$t \bar{t}$ events:
- The number of probes from non-$t \bar{t}$ events predicted by the simulation is subtracted from the data probe yields, thus obtaining the so-called "uncorrected efficiency":
  \[
  \varepsilon_{\text{data}}^{\text{uncorr}} = \frac{N_{\text{pass, non-}\bar{t}}^{\text{data}} - N_{\text{data}}^{\text{pass, non-}\bar{t}}}{N_{\text{data}}^{\text{pass, non-}\bar{t}}}
  \]
- The bias on the uncorrected efficiency due to non-b jets in the selected $t \bar{t}$ events is removed using a $t \bar{t}$ MC simulation. The correct b-tagging efficiency $\varepsilon_{\text{data}}$ is obtained from:
  \[
  \varepsilon_{\text{data}} = \varepsilon_{\text{data}}^{\text{uncorr}} \times f_{t \bar{t}}^{\text{MC}} + (1 - f_{t \bar{t}}^{\text{MC}}) \times f_{\text{non-b}}^{t \bar{t}}
  \]

Here $f_{t \bar{t}}^{\text{MC}}$ is the fraction of the probe jets in the $t \bar{t}$ simulation that originate from a b-quark, and $f_{\text{non-b}}^{t \bar{t}}$ is the b-tagging efficiency of non-b-jets predicted by the $t \bar{t}$ simulation.

The measurement is performed in bins of jet kinematic variables such as $p_T$.

4. A dedicated BDT : improve b-jet purity

A multivariate classifier to distinguish $b \bar{b}$ jet pair candidates in $t \bar{t}$ events from other type of jet pairs is applied to improve the b-jet purity in the selected sample. The classifier is a BDT trained on simulated samples using as signal $b \bar{b}$ pairs in $t \bar{t}$ events and as background all other jet pairs selected in $t \bar{t}$ and single top events. Eleven kinematic variables computed from the jet and lepton momenta and directions and the global properties of the event (the missing transverse momentum and the number of the forward jets) are used as input to the BDT. The BDT improves the b-jet purity from 49% to 67% for jet $p_T \in [20,30]$ GeV.

5. Results

Calibration results are expressed as scale factors - the ratio of the efficiencies measured in data to those measured in simulation. The scale factors are generally consistent with unity for Anti-$k_T$, R=0.4 calorimeter jets. No significant dependence of the scale factors on pile-up has been observed. [2]

References


Changqiao Li, (Univ. of Sci. and Tech. of China & Laboratoire de Physique Nucléaire et de Hautes Energies - Paris) for the ATLAS Collaboration