b-jet identification algorithms and **performance** in the **ATLAS** experiment

Kazuya MOCHIZUKI(CPPM, Aix-Marseille Université)
on behalf of the ATLAS collaboration

LHC & ATLAS

- Physics Goals
  - Higgs boson, New Physics (SuSy, Extra dimension, etc.)
- Proton-proton collision
- $\sqrt{s} = 8\ (7)$ TeV in 2012 (2011)
- $\int L\,dt \sim 20\ (5)$ fb^{-1}

ATLAS detector
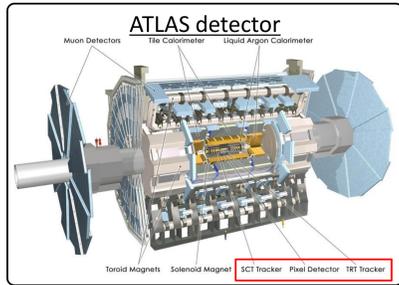
Inner detector (tracker) ∋ Pixel, SCT, TRT
Accurate tracking and vertexing vital for *b*-tagging.
$\sigma(d_0) \sim 35\ \mu m$ for typical track $p_T = 5$ GeV

Calorimeter ∋ Liquid Argon, Tile
Calorimeter jets define the axis around which tracks are clustered for *b*-tagging.

Muon Detectors
Muon tracks are combined with inner detector tracks to form muon candidates used in semi-muonic tagging.

b-tagging

Goal: identify *b*-jets against both light and *c*-jets
A crucial tool for many key analyses
Higgs physics ($H \to bb$, ttH, $H \to \tau\tau$, WW <for rejection>), SuSy, top physics, etc.
Key differences between the jet flavors are exploited:

$L_{xy} \sim 5$ mm
when
$p_T^{jet} \sim 50$ GeV
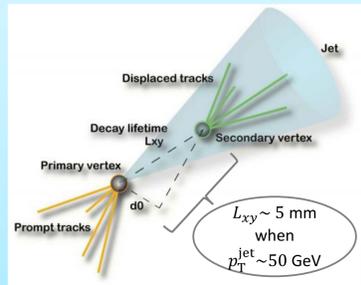
b-jet
- Relatively high track multiplicity
- Tracks displaced from primary vertex:
  Tracks with large impact parameter (IP)
- Higher secondary vertex (SV) invariant mass ← *B* hadron mass:
  m_B up to ~5 GeV
- SV information of *B* hadron decay:
  - Mass
  - Track multiplicity
  - Direction
  - SV decay length from primary vertex coming from lifetime: $c\tau_B \sim 490\ \mu m$

Light jet
Jet originating from light quarks (u, d, s) and gluons
- Relatively low track multiplicity
- Low SV mass

c-jet
- Typically very similar but weaker features compared to *b*-jet:
  $c\tau_{D^+} \sim 310\ \mu m$
- SV typically with mass of *C* hadron: m_C up to ~2 GeV

b-tagging at ATLAS: algorithms

IP3D
Track based algorithm. Uses 2D probability density functions of transverse and longitudinal IP of the tracks to form a jet 'weight' based on the log-likelihood ration method (LLR).

SV1
Secondary vertex fitting algorithm. Uses SV information of mass, direction, track multiplicity and number of vertices. LLR is taken as output weight.

JetFitter
Fit vertices along *b* flight direction which is assumed the same as the jet axis. Uses SV information. Outputs LLR weight.

MV1
- Neural network (NN) combination of 3 taggers' information
- Best performing *b*-tagging algorithm in ATLAS
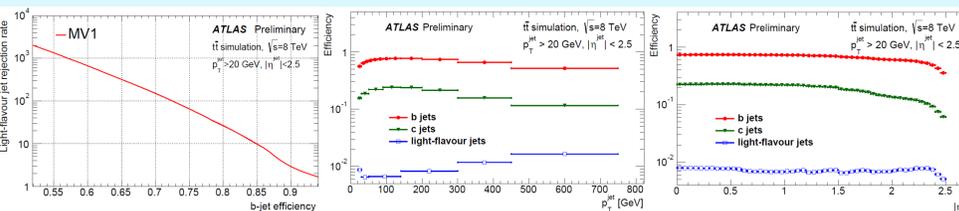- Recommended tagger in the Run 1 analyses

MV1c
- A successor of MV1, trained also against *c*-jet background
- Improved *c*-jet rejection for modest reduction in light jet rejection compared to MV1 tagger
- Now used in H->bb analyses
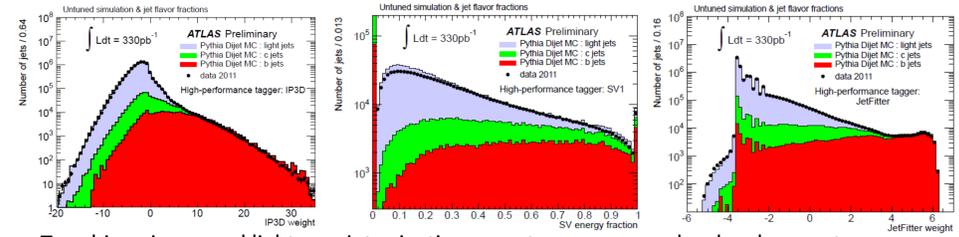
To achieve improved light or *c*-jet rejection, new taggers are under development.

b-tagging performance & calibration

Light jet rejection as a function of *b*-jet efficiency.

b-, *c*- and light jet efficiency as a function of η^{jet}.

*b*-, *c*- and light jet efficiency as a function of p_T^{jet}.

Differences between data/Monte Carlo performance are corrected for using scale factors derived in several dedicated calibration methods for *b*-, *c*- and light jets.

D^* meson calibration: *c*-tagging efficiency

- The *c*-jet tagging efficiency is measured using a sample enriched with D^{*+} mesons.
- Uses template fit before and after the *b*-tagging to extract *b*-tagging efficiency of the D^{*+} candidates.
- Using the measured D^{*+} meson efficiency, the *c*-jet efficiency is extracted by:

$$\epsilon_c = \frac{\epsilon_{D^{*+}} - f_b \epsilon_b}{1 - f_b}$$

ATLAS-CONF-2014-046
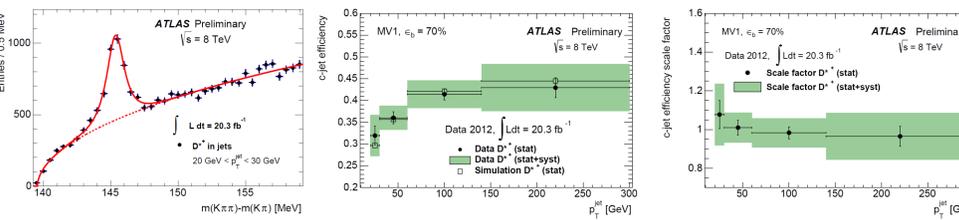
$Wc(c \to \mu X)$ calibration: *c*-tagging efficiency

ATLAS-CONF-2013-109

- Dominant production mechanism: $gs \to Wc$
- Event selection:
  - 1 isolated electron and a jet associated to exactly 1 muon and E_T^{miss}
  - Opposite sign (OS) for electric charges of selected electron and muon
  - Same sign (SS) events are used as part of the background estimation
- Charge correlation of 2 leptons in $W(\to e\nu) + c(\to \mu X)$ signal process allows to obtain very high *c*-jet purity(~90%).
- *Wc* yield is obtained from OS-SS subtraction, exploiting the fact that most backgrounds are symmetric in the OS/SS regions whereas the signal is predominately in OS.

$$\epsilon_{c(\mu)}^{data} = \frac{N_{Wc}^{b-tag}}{N_{Wc}}$$
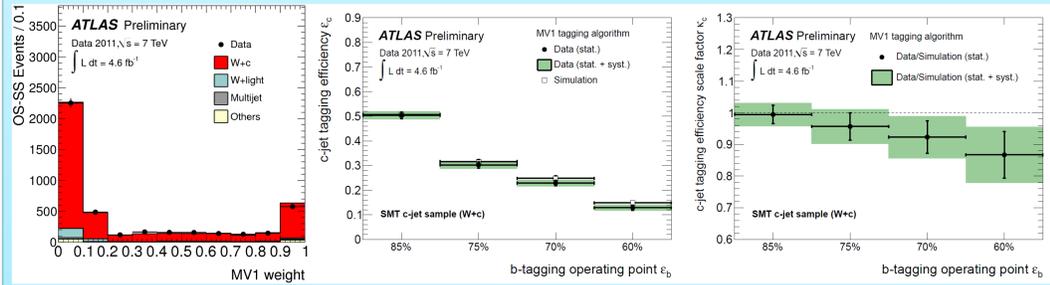
- Semi-muonic *c*-jet scale factors vary between 0.87 and 0.99 with total relative uncertainties of 10% to 4%.
- Inclusive *c*-jet scale factors are obtained by simulation-based extrapolation, scale factors vary between 0.75 and 0.92 with total relative uncertainties of 13% to 5%.

## Combinatorial likelihood calibration: *b*-tagging efficiency

- $t\bar{t}$ dilepton events can provide very pure *b*-jet sample, using $e\mu$, $ee + \mu\mu$ channels in 2 and 3 jet bins.
- Event level correlation between the jet flavors can be exploited to achieve greater precision.

$$f_{2\ tags} = f_{bb}\epsilon_b^2 + f_{bj}\epsilon_j\epsilon_b + (1 - f_{bb} - f_{bj})\epsilon_j^2$$
$$f_{1\ tag} = 2f_{bb}\epsilon_b(1-\epsilon_b) + f_{bj}[\epsilon_j(1-\epsilon_b) + (1-\epsilon_j)\epsilon_b] + (1 - f_{bb} - f_{bj})2\epsilon_j(1-\epsilon_j)$$

- $\epsilon_b(\epsilon_j)$ is the *b*-jet (non *b*-jet) efficiency.
- $f_{bb}(f_{bj})$ is the fraction of events with a true *bb* (*bj*) jet pair.

- *b*-tagging efficiency can be obtained by maximizing a likelihood function, taking both $f_{1\ tag}$ and $f_{2\ tags}$ from data, with f_{bb}, f_{bj} and ϵ_i from simulation.
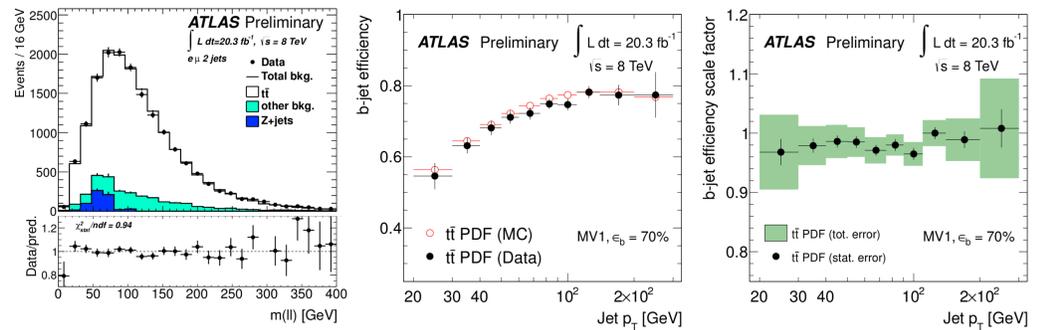

The total uncertainty on the scale factors is ~2% for jets with p_T around 100 GeV for MV1 70% operating point.

ATLAS-CONF-2014-004

Mis-tag rate calibration

- Fake SVs reconstructed in light jets are expected to have approximately symmetric decays around the primary vertex while those from true SVs are not. The same is expected for IP of tracks.
- Mis-tag rate is estimated by using negatively signed tracks and vertex decay lengths, correcting for the presence of heavy flavor jets and long-lived particles/material interactions in light-jets.
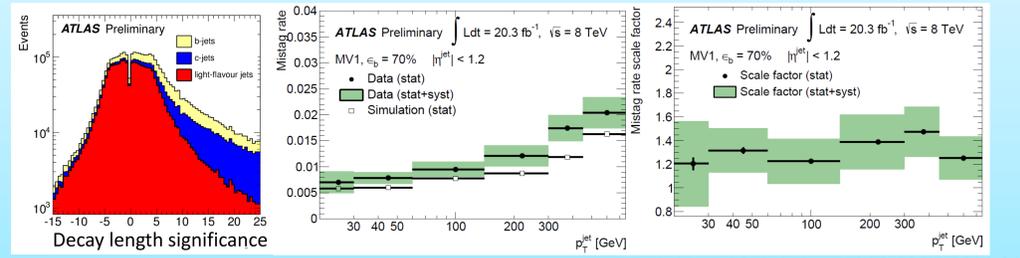
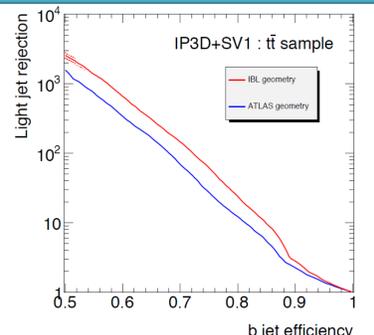**ATLAS-CONF-2012-040** **ATLAS-CONF-2014-046**

Summary & outlook

Conclusion

- Calibrations with improved precision:
  - Updated calibration with full 8 TeV data
  - New developments:
    - *Wc* calibration
    - Combinatorial likelihood calibration

Outlook for Run 2

- Insertable B-Layer:
  A new innermost layer of pixel detectors improves *b*-tagging performance:
  ~ two times better rejection for the same *b*-tagging efficiency.

- Development of new taggers
  - Multi-output NN for charm jet tagging.
  - Improved version of MV1 using a boosted decision tree and re-optimized list of input variables.
  - Dedicated tagger for high p_T and merged jet from top decay.

Thank you for your attention!