Calibration of light-flavour jet b-tagging rates on ATLAS data at $\sqrt{s} = 13$ TeV

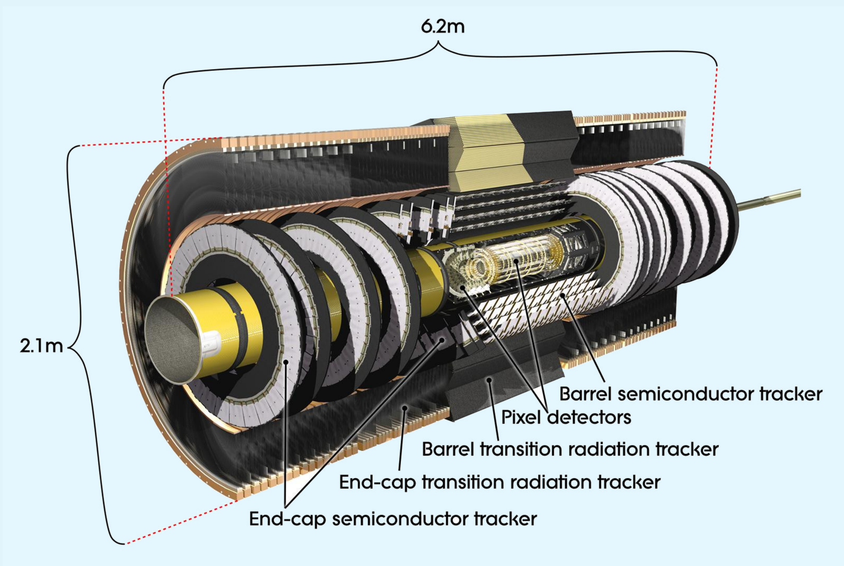
Matthias Saimpert[1],
on behalf of the ATLAS collaboration

[1] Deutsches Elektronen-Synchrotron (DESY)

EPS-HEP 2017
Venice, Italy 5-12 July 2017

Prelude: the ATLAS experiment at the LHC

- **The ATLAS detector** [1] is a general-purpose experiment recording high energy hadronic collisions occuring at the Large Hadron Collider.



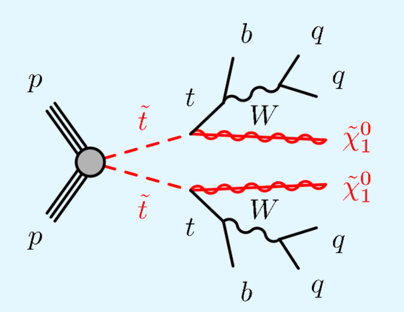
- Most relevant subdetector for this work: **the inner detector**, providing vertex and tracking capabilities for $|\eta| < 2.5$.

  - → composed of 4 layers of silicon pixel sensors, followed by 8 (18) silicon microstrip sensors arranged in cylinders (disks) in its barrel (endcap) region,
  - → followed by a straw tube tracker, providing on average 36 additional hits.
  - → Access to charged particle transverse momentum (axial 2 T magnetic field).

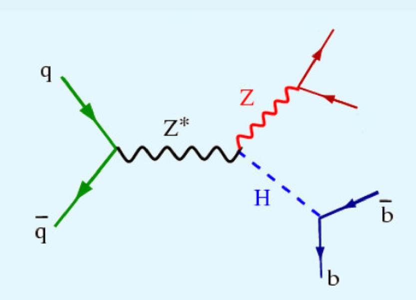
Introduction

- Identifying the jets originating from b-quarks (b-tagging) is essential to many ATLAS physics analysis:

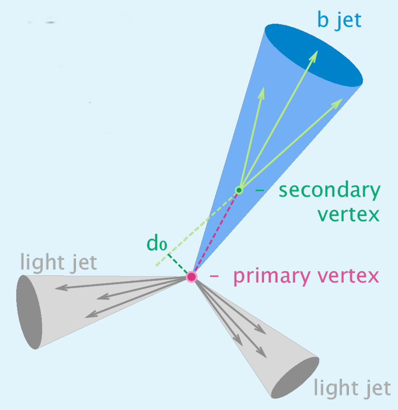
Top Physics / New Phenomena



Higgs Physics



b- and light jet topology



- ATLAS uses a boosted decision tree (BDT) algorithm to identify the b-jets, **MV2c10** [2], based on:
  - → the presence of tracks with non-zero impact parameters inside the jet
  - → the presence of secondary vertices inside the jet
  - → the decay chain topology of the tracks/vertices inside the jet
- b-tagging efficiency for light-flavour (LF) jets, a.k.a. **mistag rate,** is a crucial ingredient for background estimation.
- Mistag originates from track resolution, material interactions and long-lived particles within LF jets → **needs calibration.**

b-tagging efficiency working point definition

- b-tagging working points (WP) [3] are used to define b-jets in physics analysis.
- A WP is defined as a fixed cut on the MV2c10 BDT score output, e.g. MV2c10 Output $> X$, MV2c10 Output $\in$ [-1, 1].
- The cut values are chosen to provide a specific b-jet efficiency on jets in a $t\bar{t}$ simulated sample:

WP	Cut value X	ε_b^{MC}	C Rejection	τ Rejection	LF Rejection
85%	0.18	85%	3	8	34
77%	0.65	77%	6	22	134

Measurement of the mistag rate in ATLAS data

- Selection of well-measured di-jet events in 2015 + 2016 data:
  - → Use of a set of prescaled and unprescaled jet triggers (L = 0.02 to 36100 fb^{-1})
  - → At least 2 jets with $p_T^{jet} > 20$ GeV (lower bound of jet energy calibration), $|\eta^{jet}| < 2.5$ (tracker geometrical acceptance)
  - → Selection of the two highest p_T^{jet} jets in the event (pileup jets excluded)
  - → Good separation between the two selected jets in the transverse plane ($\Delta\phi_{jj} > 2$ radians) to reduce gluon splitting and non-collision background
- The b-tagging rate of the selected jets in the data sample is given by:

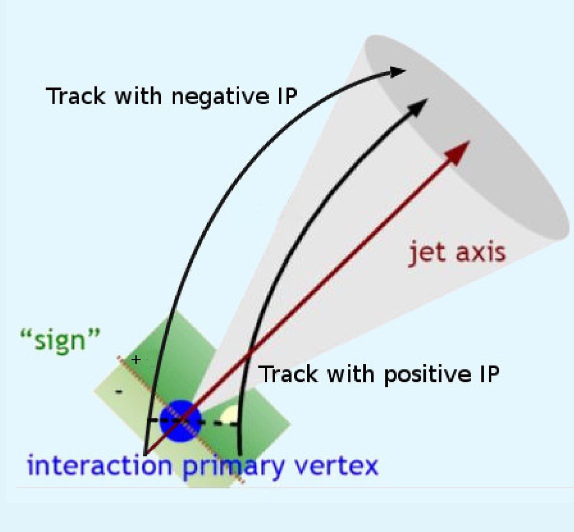
$$(N_{jets,btag}/N_{jets,all})^{data} = f_c \cdot \epsilon_c + f_b \cdot \epsilon_b + (1 - f_c - f_b) \cdot \epsilon_l$$

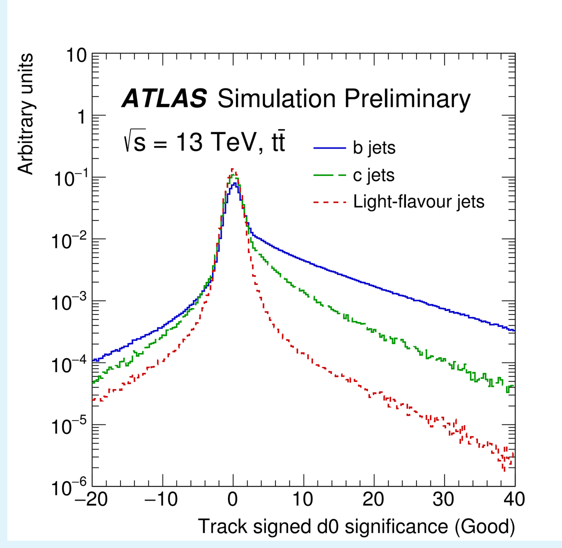
f_c, f_b: fraction of true b- and c-jets. **4-8% and 1-4% according to simulation.**
ϵ_l, ϵ_c, ϵ_b: b-tagging efficiency for true LF, c- and b-jets, **typically** $\epsilon_c, \epsilon_b > 10 \times \epsilon_l$.

Too high contamination by true b- and c-jets (HF jets). An alternative method is needed → use of the negative tag method [4]

The negative tag method

- **Assumption:** LF jets misidentified due to track resolution effects
- Example: consequence on the signed track impact parameter significance





- High tails for b- and c-jets due to the long lifetime of b and c hadrons, used as discriminants in MV2c10.
- Distribution symmetric for LF jets → mistag rate driven by track resolution effects

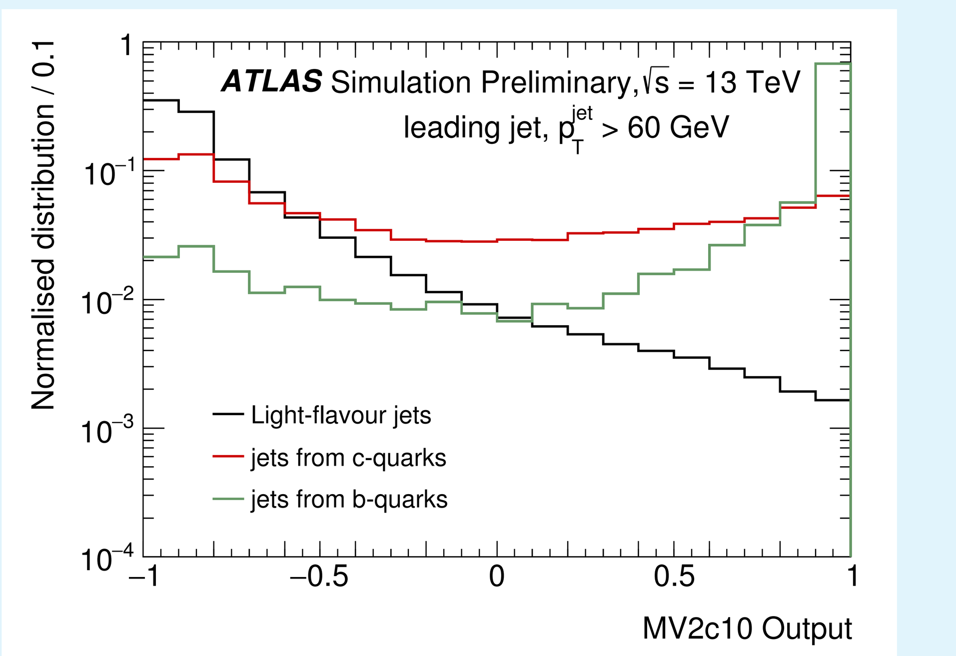
The mistag rate can be accessed by tagging the jets with negative attributes. Definition of a new "flipped" algorithm to do so → MV2c10Flip

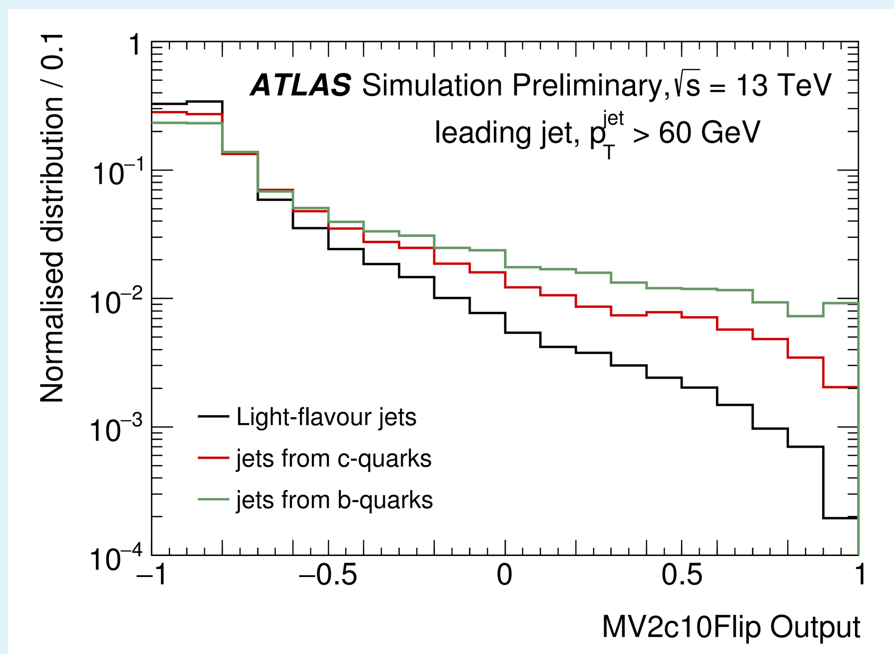
- Additional transfer factors extracted from simulation are used to correct for:
  - → HF contamination (K_{HF}^{MC}): true c- and b-jets do not have exactly the same "flipped" efficiency as LF jets.
  - → LF jets with true secondary vertices (K_{LL}^{MC}): nominal LF b-tagging efficiency > "flipped" b-tagging efficiency due to material interactions and long-lived mesons (K_s and Λ^0 mainly)

Final mistag rate measurement: $\epsilon_l^{corr} = \epsilon_{neg}^{data} \cdot K_{HF}^{MC} \cdot K_{LL}^{MC}$

K_{HF}^{MC} ranges from 0.9 to 0.3, K_{LL}^{MC} ranges from 1.3 to 5.

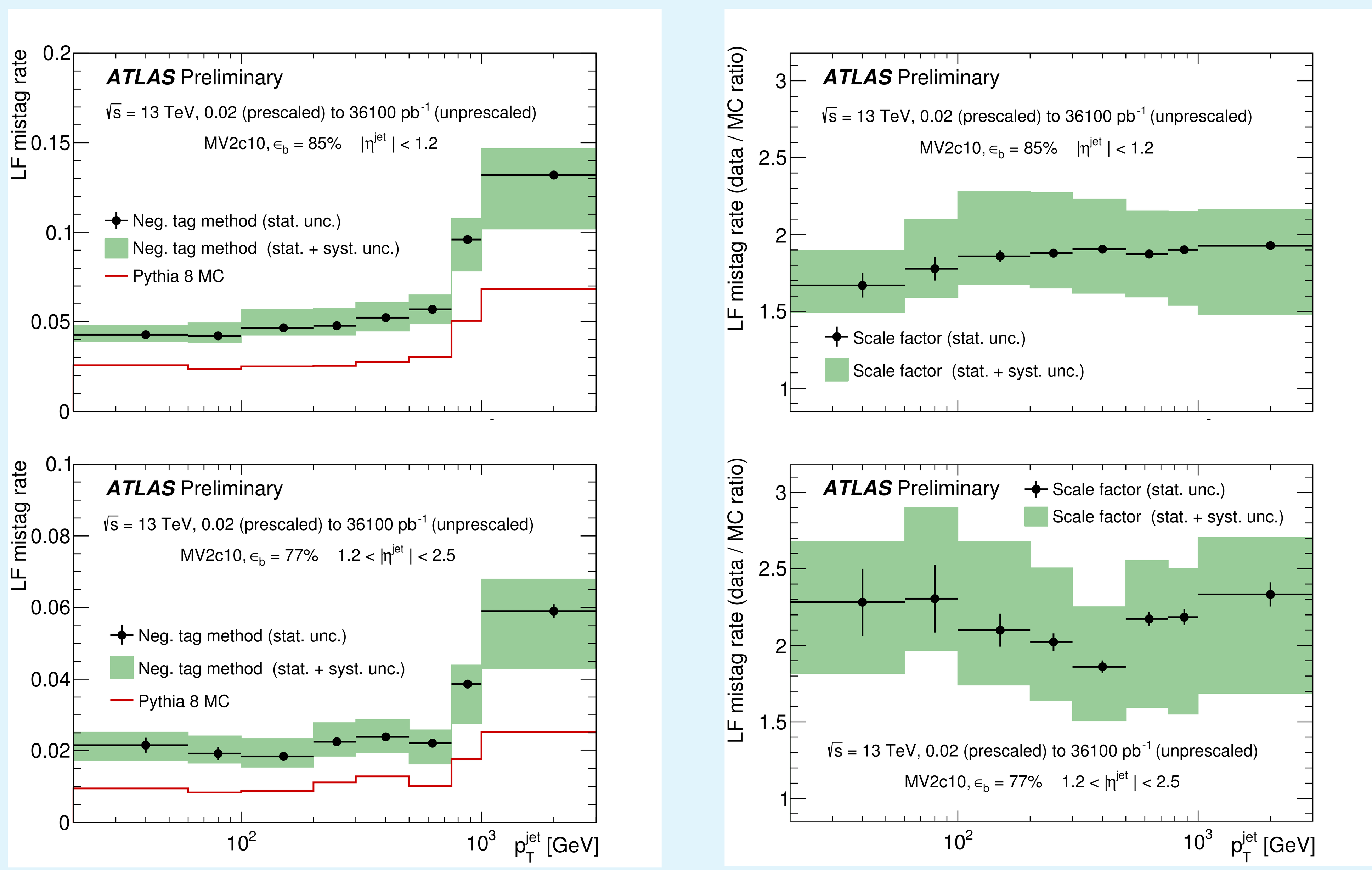
Expected performance of the "flipped" algorithm





- MV2c10Flip discriminates much less between true LF, c and b-jets with respect to MV2c10.
- MV2c10/MV2c10Flip similar for LF jets (at first order).

Mistag rate measurements for the 85% and 77% WPs in $p_T^{jet}/|\eta^{jet}|$ bins



- **Partial cancelation of uncertainties** shared between LF/HF jets and MV2c10/MV2c10Flip modeling (use of transfer factors K_{LL}^{MC} and K_{HF}^{MC}).
- **Precision (±10-35%)** limited by the modeling of the track impact parameter resolution (detector simulation) and HF contamination in the data sample (MC generator).
- **Data/MC ratios** used to correct simulated events in ATLAS (per jet correction).

References

[1] **ATLAS Collaboration**, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST 3 (2008) S08003.
[2] **ATLAS Collaboration**, *Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run*, ATL-PHYS-PUB-2016-012, 2016.
[3] **ATLAS Collaboration**, *Performance of b-Jet Identification in the ATLAS Experiment*, JINST 11 (2016), P04008.
[4] **ATLAS Collaboration**, *Calibration of the performance of b-tagging for c and light-flavour jets in the 2012 ATLAS data*, ATLAS-CONF-2014-046, 2014.