ATLAS b-tagging performance during LHC Run-2 with the Insertable B-layer

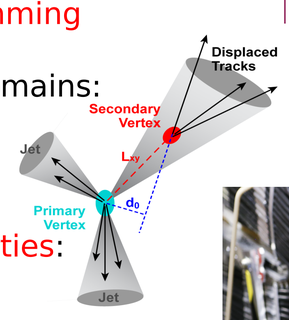
LHCP – St. Petersburg, August 31 - September 5, 2015

Yulia Rodina (CPPM, Aix-Marseille Université & Universitat Autònoma de Barcelona)
on behalf of the ATLAS collaboration

Introduction

Goal of b-tagging: identify jets stemming from b-quarks

B-tagging used in many analysis domains:
- Top physics
- Higgs physics
- Beyond Standard Model physics



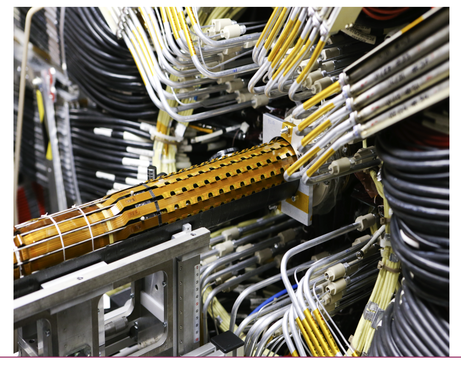
Algorithms rely on b-hadron properties:
- High mass (~5 GeV)
- Relatively long lifetime (~1.5 ps)

B-tagging algorithms:
- Spatial tagging
  - Impact parameter based: IP2D, IP3D
  - Secondary vertex based: SV, JetFitter
  - Multivariate: MV1, MV2
- Soft lepton tagging

Insertable B-layer (IBL)

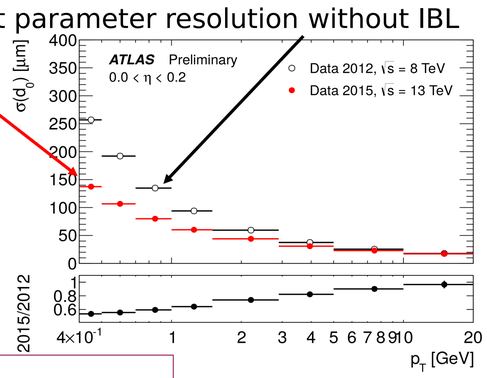
Major inner detector upgrade for Run-2: addition of Insertable B-layer (IBL) as fourth innermost pixel layer
- Radius (IBL) ~ 3.3cm
- Radius (former Layer 0) ~ 5cm

Advantages of IBL:
- Proximity to the interaction region
- Higher granularity (50μm x 250μm instead of 50μm x 400μm the case of former Layer 0)
- Increase pixel measurements of a track from 3 to 4

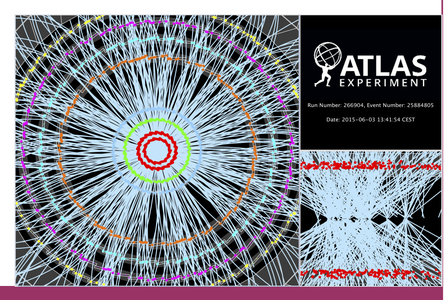


Transverse impact parameter resolution without IBL and with IBL [3]

IBL significantly improves track reconstruction: d0 resolution for low-p_T tracks is improved up to factor of 2!



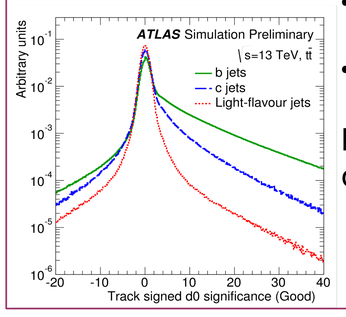
Collision event collected by full ATLAS tracker (June 2015):



Baseline b-tagging algorithms

Impact parameter based — IP2D and IP3D

- Use lifetime signed impact parameter (IP) significance of tracks matched to a jet
- IP - distance of closest approach to the primary vertex (PV)

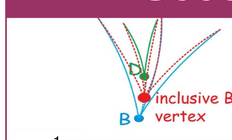


- IP2D: only use transverse IP in 1D likelihood
- IP3D: combine with longitudinal IP in 2D likelihood

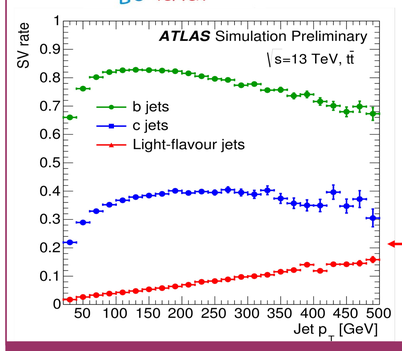
Information from multiple tracks combined into a jet weight:

$$w_{track} = \frac{p_b}{p_{light}};\ w_{jet} = \sum_{tracks} \log\left(w_{track}\right).$$

Secondary vertex finding — SV



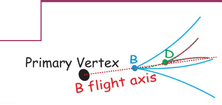
- Removes tracks compatible with long-lived particles' decays or material interactions
- Fits an inclusive secondary vertex



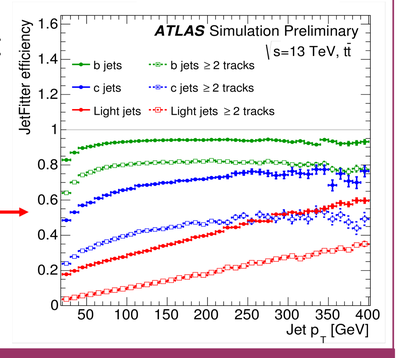
Secondary vertex reconstruction rate as function of jet p_T [1]

Multi-vertex fit — JetFitter

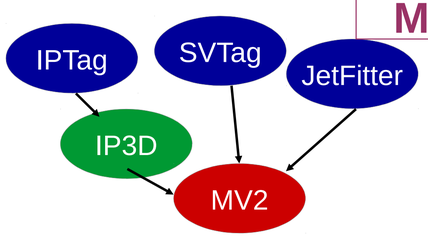
- Reconstructs more than one displaced vertex
- Attempts to reconstruct full PV to b- to c-hadron decay chain



Rate to reconstruct a vertex with at least one or two tracks as function of jet p_T [1]



Multivariate algorithm — MV2



Combines discriminant observables from IP3D, SV and JetFitter into a boosted decision tree (BDT) algorithm

The default algorithm for Run-2, MV2c20, is a BDT which is trained using b-jets as signal and a mixture of 80% light-flavour jets and 20% c-jets as background.
- MV2 is a revision of MV1
- MV1: the main b-tagging algorithm used during Run-1
- MV1: based on a Neural Network that combined inputs from intermediate tools
- MV1c: version of MV1 trained using same background mixture as MV2c20

Advantages of MV2:
- Better performance
- Easier retraining and software maintenance

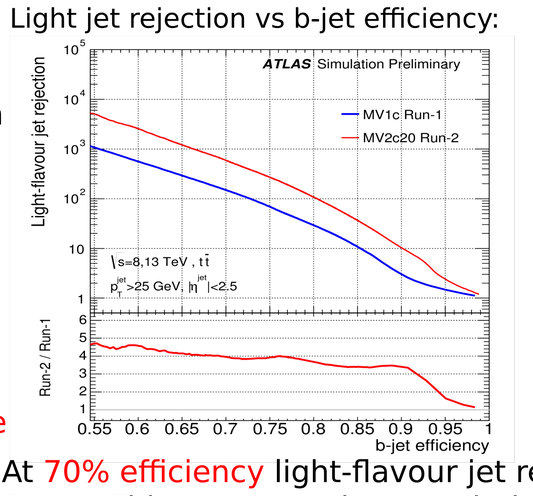
Expected b-tagging performance

Run-2 vs Run-1

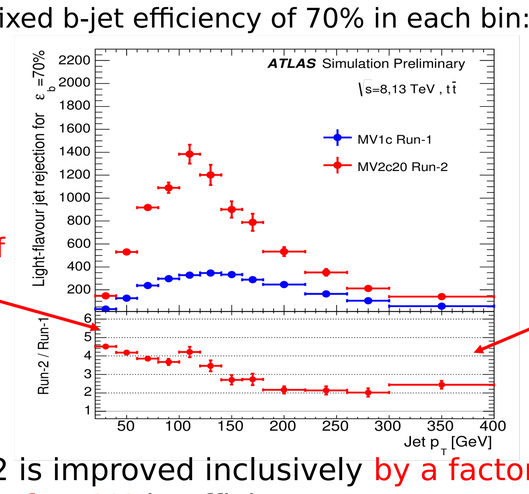
A b-tagging performance improvement is expected to be achieved in Run-2 due to:
- Addition of the IBL
- Many algorithmic updates in track reconstruction [4] and b-tagging, both in the basic taggers and final MVA algorithm

The plots below [1] show a comparison of the default Run-2 b-tagging algorithm MV2c20 and the equivalent Run-1 b-tagging algorithm MV1c.

Light jet rejection vs b-jet efficiency:



Light jet rejection as a function of jet p_T given a fixed b-jet efficiency of 70% in each bin:



Improvement at low p_T is mostly due to addition of the IBL

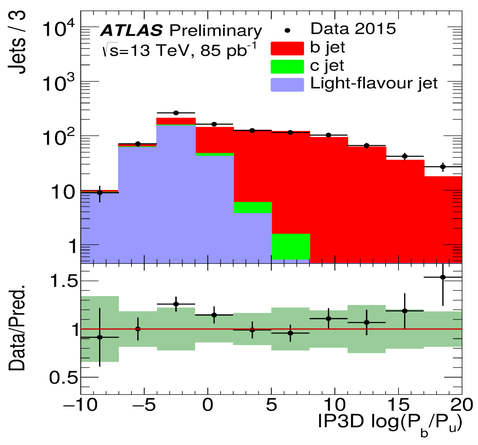
Improvement at high p_T is due to algorithm improvements

At 70% efficiency light-flavour jet rejection in Run-2 is improved inclusively by a factor of 4 compared to Run-1. This corresponds to a relative improvement of ~10% in efficiency at a constant light-jet rejection!

Run-2 data to Monte Carlo (MC) comparison

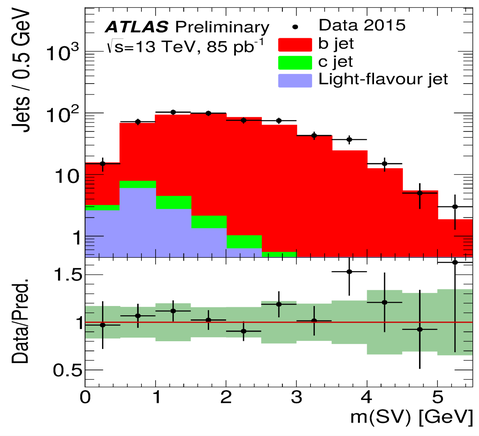
- To confirm the MC performance, simulation needs to be compared with data. The early Run-2 data commissioning studies are a first step in this direction.
- First study of the b-tagging modeling performed in 13 TeV on high purity b-jet sample of e+μ di-leptonic $t\bar{t}$ events [2]. The leading two jets with p_T> 20 GeV and |η|<2.5 are considered.

IP3D weight:

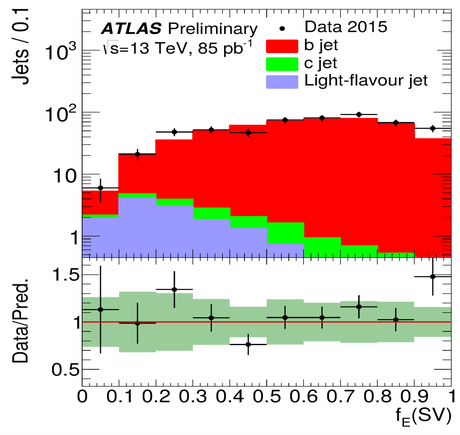


Inclusive secondary vertex properties:

(a) Invariant mass

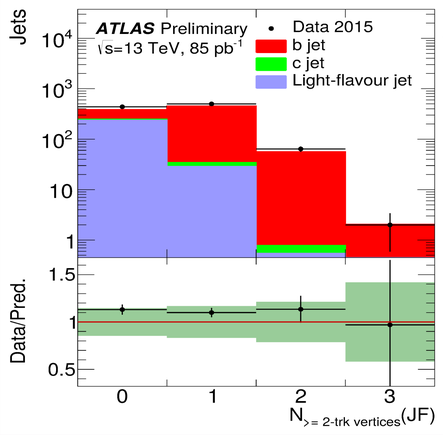


(b) Energy fraction of tracks at vertex w.r.t. all tracks in the jet

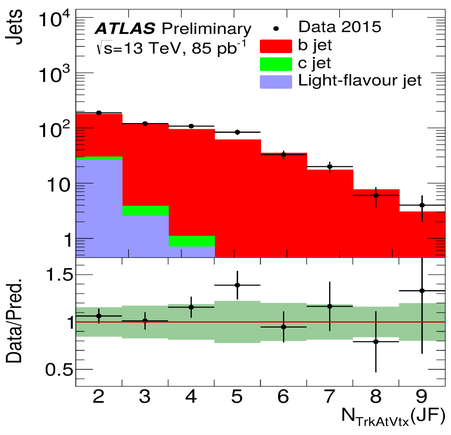


Vertex decay chain fit properties:

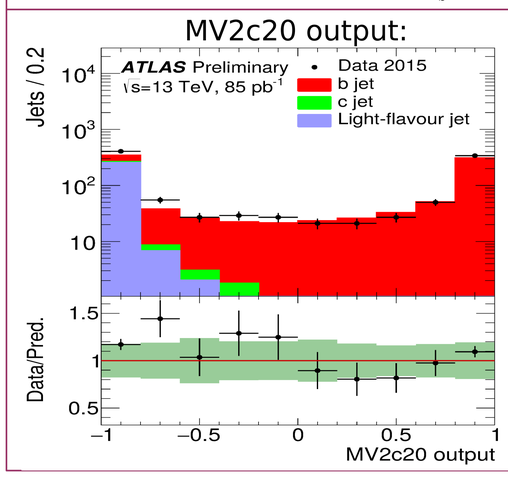
(a) Number of vertices with ≥2 tracks



(b) Number of tracks at vertices with ≥2 tracks



MV2c20 output:



Conclusions

The b-tagging performance in ATLAS is expected to be significantly improved in Run-2 thanks to the addition of the IBL and to updates in the tracking and b-tagging algorithms. MC studies showed significant improvement in light-flavour jet rejection in Run-2 compared to Run-1. First commissioning studies on Run-2 data show promising agreement of data with MC.

References

[1] Expected performance of the ATLAS b-tagging algorithms in Run-2 (ATL-PHYS-PUB-2015-022): https://cds.cern.ch/record/2037697
[2] Commissioning of the ATLAS b-tagging algorithms using $t\bar{t}$ events in early Run-2 data (ATL-PHYS-PUB-2015-039)
[3] https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/IDTR-2015-007/
[4] The Optimization of ATLAS Track Reconstruction in Dense Environments (ATL-PHYS-PUB-2015-006): https://cds.cern.ch/record/2002609