EPS HEP – Vienna, July 22-29th 2015

Expected b-tagging performance for ATLAS in LHC Run 2

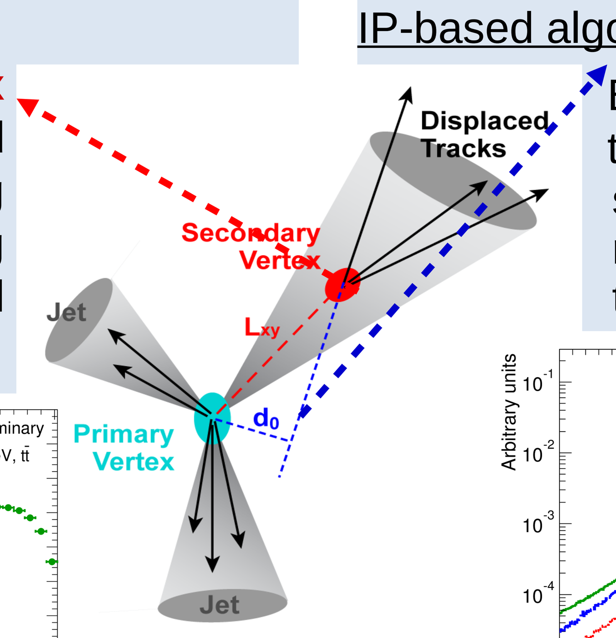
Introduction

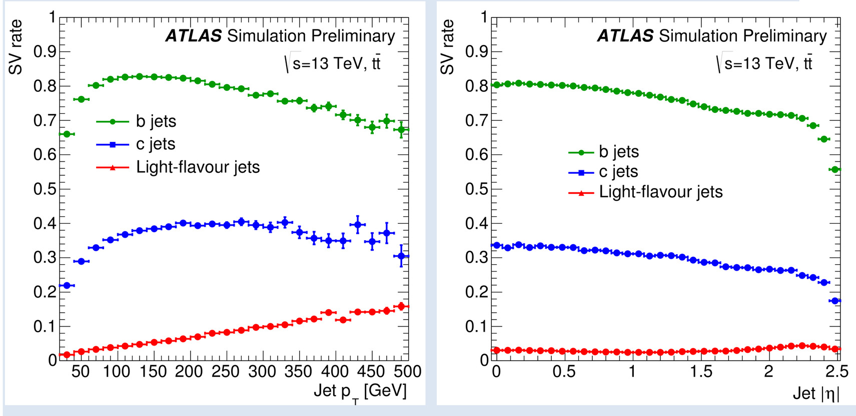
The identification of jets originating from b-quarks is an important aspect of many analyses in ATLAS (e.g. top physics, H->bb, new physics searches). The discrimination of b and light-jets (udsg) is mainly possible thanks to the long lifetime of b-flavored hadrons, leading to significantly displaced secondary vertices (SV), and reconstructed charged particle tracks with large impact parameters (IP). Thanks to the addition of a pixel detector layer, the Insertable-B-Layer (IBL), which is closer to the beam pipe (~3.3cm instead of ~5cm), and has smaller pixels (50μm x 250μm instead of 50μm x 400μm), the impact parameter resolution is improved in Run 2. Together with improvements to the b-tagging algorithms [1], the b-tagging performance is expected to be significantly enhanced in LHC Run 2. All the plots shown here are made using a ttbar simulated sample, and use a loose jet selection: $p_T > 20$GeV and $|\eta| < 2.5$. The basic b-tagging algorithms all use tracks spatially matched to the jets as input.

Basic b-tagging algorithms

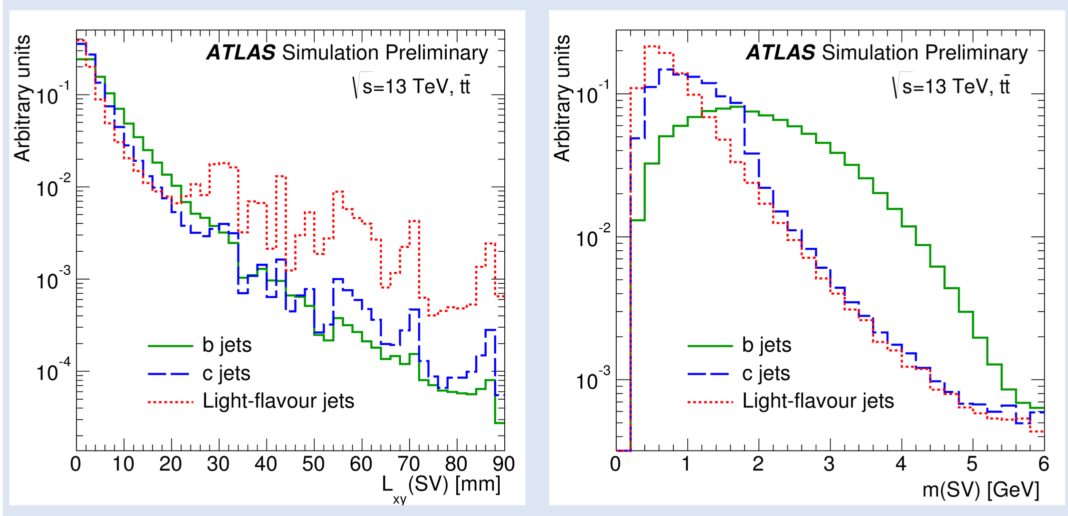
SV-based algorithms: SVF

Reconstructs one secondary vertex inclusively per jet, using displaced tracks that are not identified as coming from long-living particles, and removing vertices consistent with material interactions.





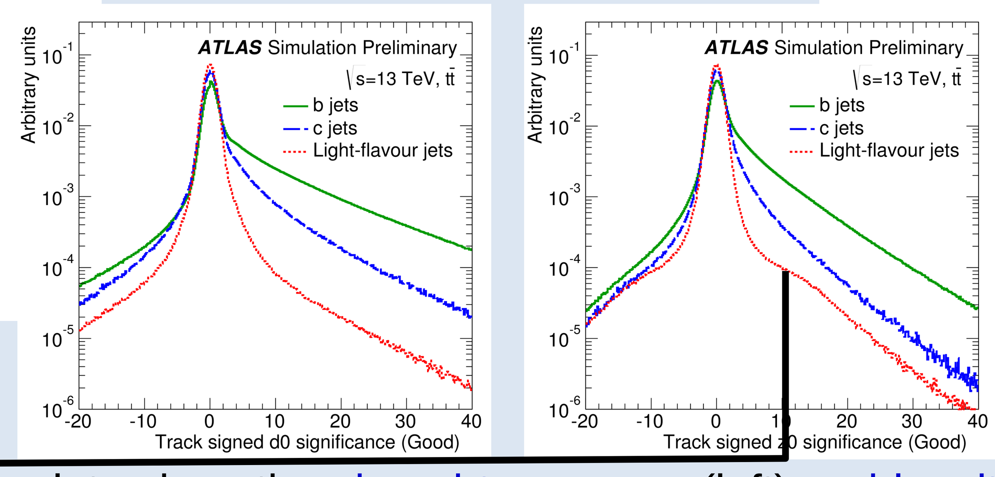
With the Run 2 configuration, SVF typically manages to reconstruct a SV in 70-80% of b-jets.



Two properties of the reconstructed vertex are shown: the transverse decay length (left), and the mass of the secondary vertex (right).

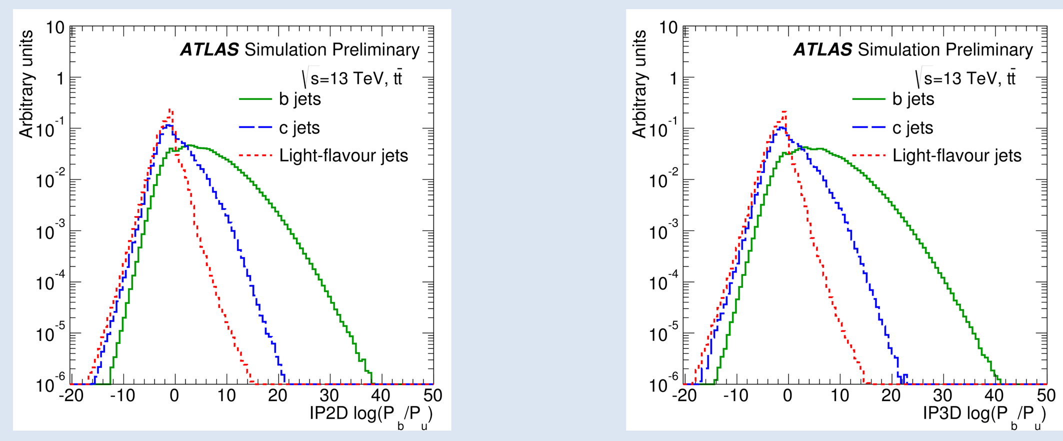
IP-based algorithms: IP2D and IP3D

Based upon the signed IP significance of the tracks, which is defined as positive if it is in the same direction as the jet momentum direction relative to the primary vertex.



These plots show the signed transverse (left) and longitudinal (right) IP significances.

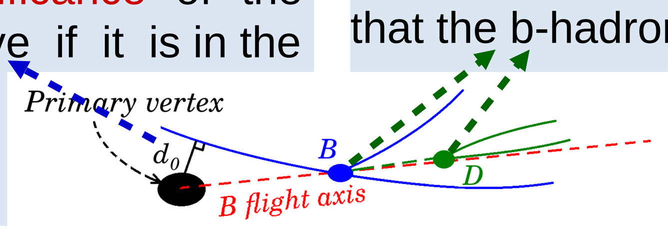
An additional component comes from pile-up tracks. Since IP2D does not use longitudinal informations, it is expected to be more robust against pile-up than IP3D.

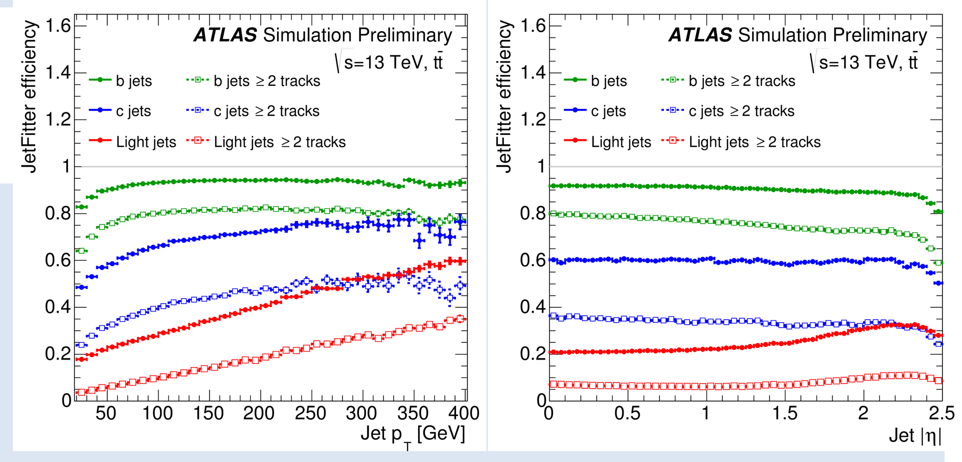


Probability density functions of the signed IP significance are used to define ratios of the b- and light-jet hypotheses which are then combined into a single log likelihood ratio discriminant (LLR), which are shown above for IP2D (left) and IP3D (right)

Multi-vertex fit algorithm: JetFitter

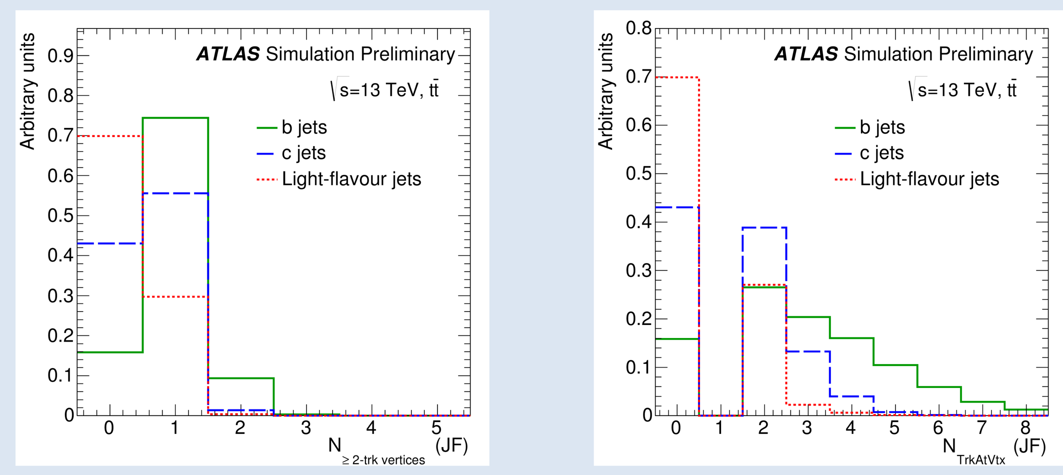
JetFitter reconstruct the full PV→B→C decay chain assuming that the b-hadron decays along the jet axis.





The two plots above show the vertices reconstruction efficiencies against jet p_T (left), and jet |η| (right).

Although the efficiency to reconstruct a decay chain with at least one two track vertex is lower, they have a much higher purity compared to the one-track vertex decay chains.



The two plots above show the number of vertices associated to at least two tracks (left), and the number of tracks from vertices with at least two tracks (right).

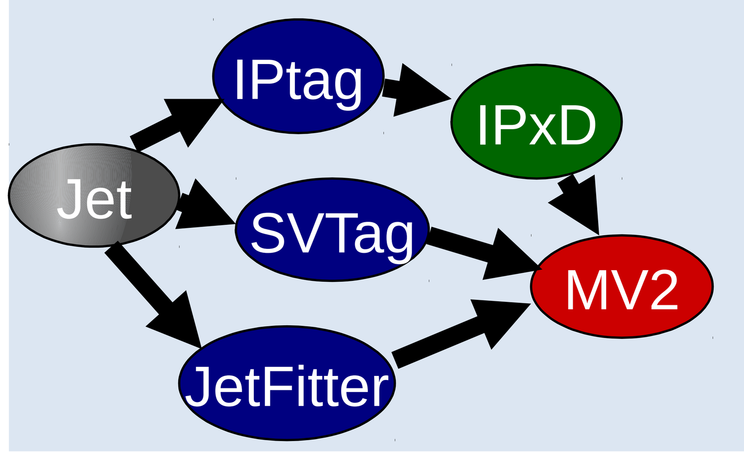
MVx tagging algorithms : from Run 1 to Run 2

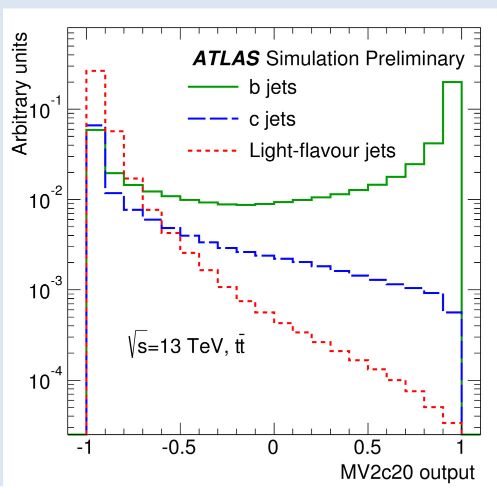
MVx: a combined tagging algorithm

- Combine variables from the three basic algorithms

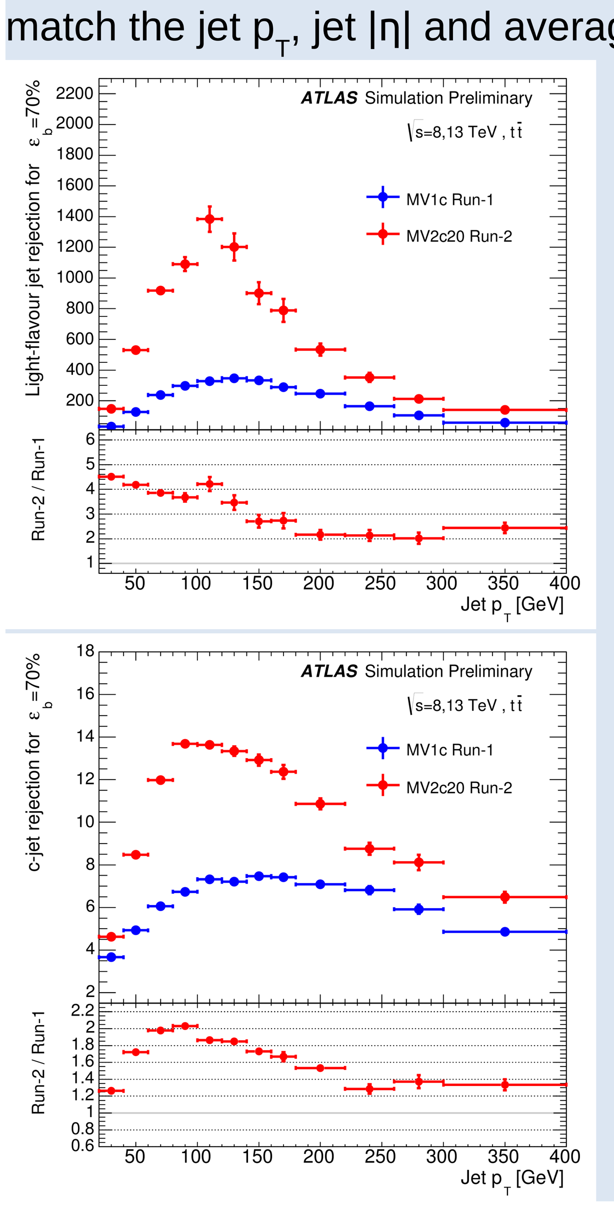
Two Neural Network multivariate tools were used in Run 1, MV1 (trained purely against light-jets) and MV1c (trained against a mixture of light and c-jets), based on the inputs from intermediate MVA tools.

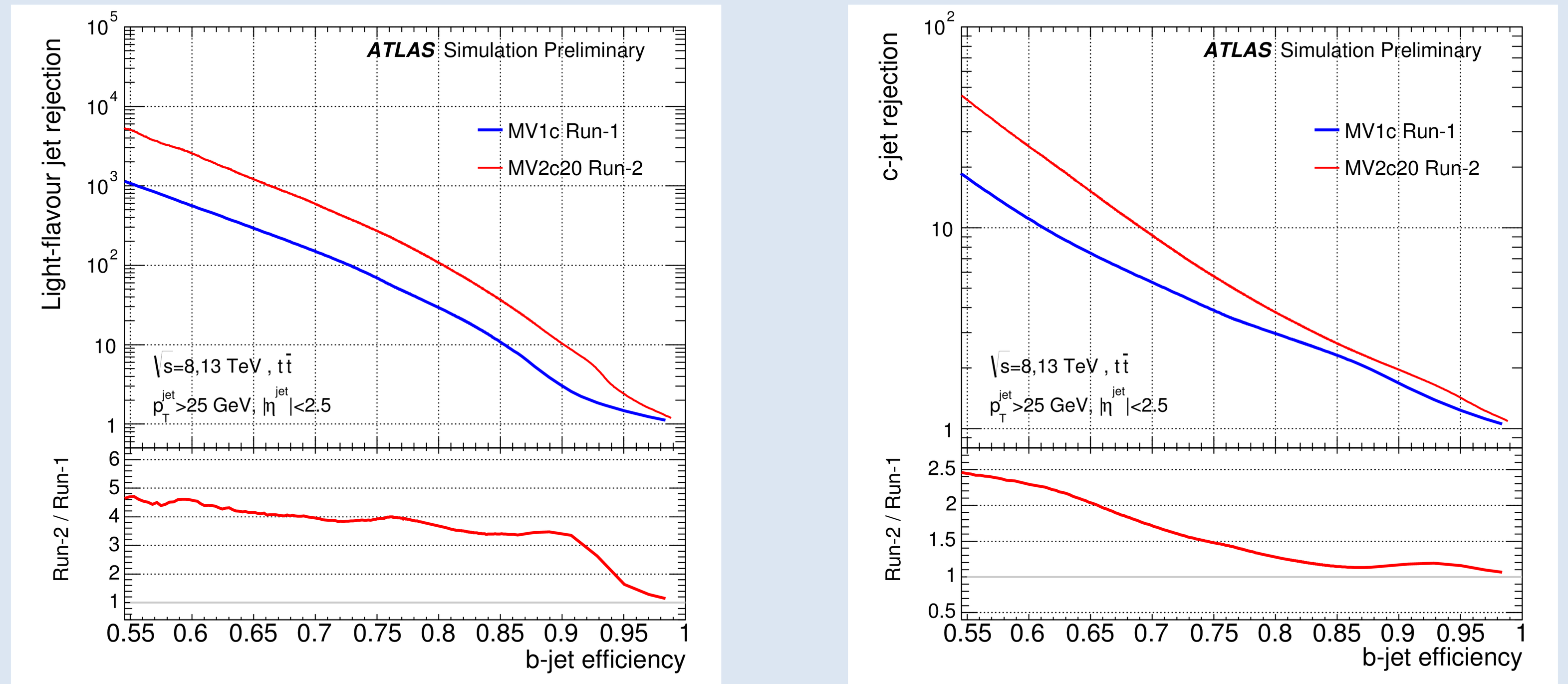
- Several improvements have been made for Run 2:
  - ➔MV2c20 (trained with 80% light and 20% c-jets) is the default Run 2 b-tagging algorithm.
  - ➔Uses Boosted Decision Tree (BDT).
  - ➔Uses inputs directly from the basic algorithms, which allows us to better exploit correlations between the input variables.
  - ➔Simplifies the algorithm by omitting additional intermediate multivariate tools.





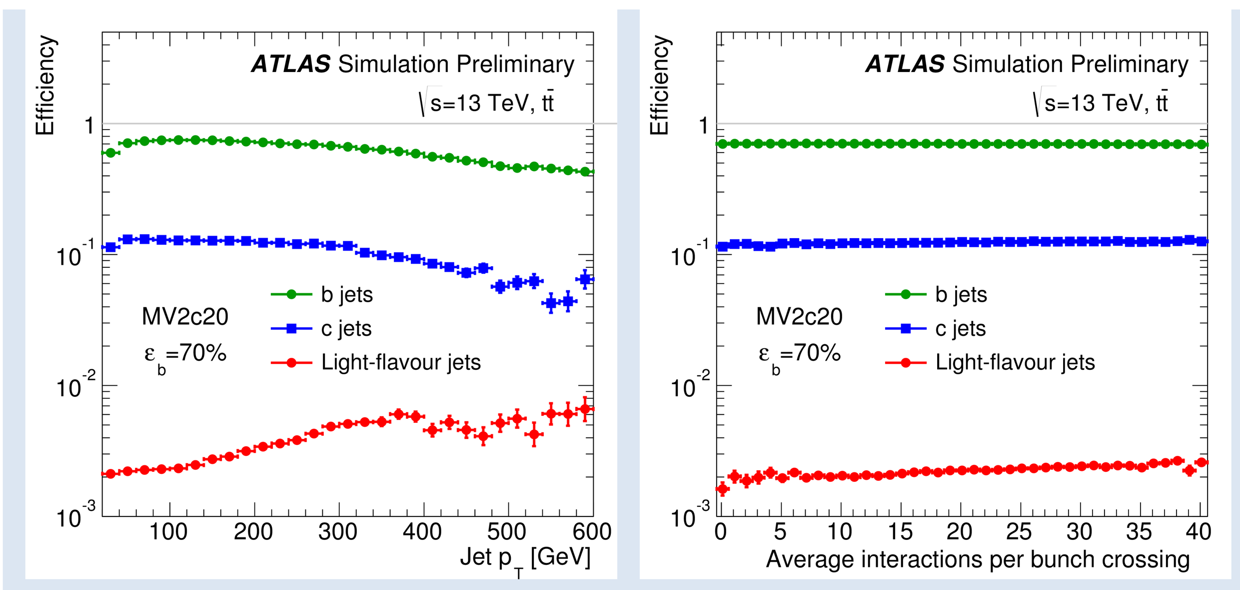
Here we compare MV1c with the Run 1 detector and reconstruction software, to MV2c20 with the Run 2 detector and reconstruction software, on a ttbar simulated sample at 8TeV for MV1, and 13 TeV for MV2. The 13 TeV sample is re-weighted to match the jet p_T, jet |η| and average number of interactions, μ, in the 8TeV sample to allow for an unbiased comparison.





- The Run 2 performance is significantly improved in terms of both light and c-jet rejection.
  - ➔For a 70% b-tagging efficiency the light-jet rejection is increased by a factor of about 4.
  - ➔Or, fixing the light-jet rejection at the value achieved for the 70% b-jet efficiency in Run 1, a relative gain of 10% in b-tagging efficiency is expected. In an analysis with four b-quarks in the final state (e.g. ttH(bb)), this increase represents a gain of 40-50% in signal acceptance.
- The improvement at high p_T is mostly due to updated b-tagging and tracking [2] algorithms, whilst the improvement at low and medium p_T is mostly due to the addition of the IBL.

Differential b-tagging performance



b-jet, c-jet, and light-jet efficiencies versus p_T (left), and pile-up (right), for a global cut on the MV2c20 output, which corresponds to a 70% b-jet efficiency on the ttbar sample.

Conclusions and outlook

Run 2 b-tagging improvements:

- New pixel layer inserted closer to the beam pipe.
- Improvements to both basic and multivariate b-tagging algorithms implemented for Run 2.
- Expect significant improvements in b-tagging performance for Run2.

We are now taking Data. To ensure we achieve these expected improvements, it is vital that the detector, tracking, and b-tagging algorithms are fully commissioned with the new data. Such efforts are underway, and are presented elsewhere [3].

References

[1] Expected performance of the ATLAS b-tagging algorithms in Run2:
https://twiki.cern.ch/twiki/bin/view/AtlasPublic/FlavorTaggingPublicResults

[2] The optimization of ATLAS Track Reconstruction in Dense Environments:
ATL-PHYS-PUB-2015-006

[3] Data commissioning:
https://twiki.cern.ch/twiki/bin/view/AtlasPublic/InDetTrackingPerformanceApprovedPlots

Thomas CALVET (CPPM, Aix-Marseille Université), for the ATLAS Collaboration