Algorithmic improvements and calibration measurements for flavour tagging at the ATLAS experiment

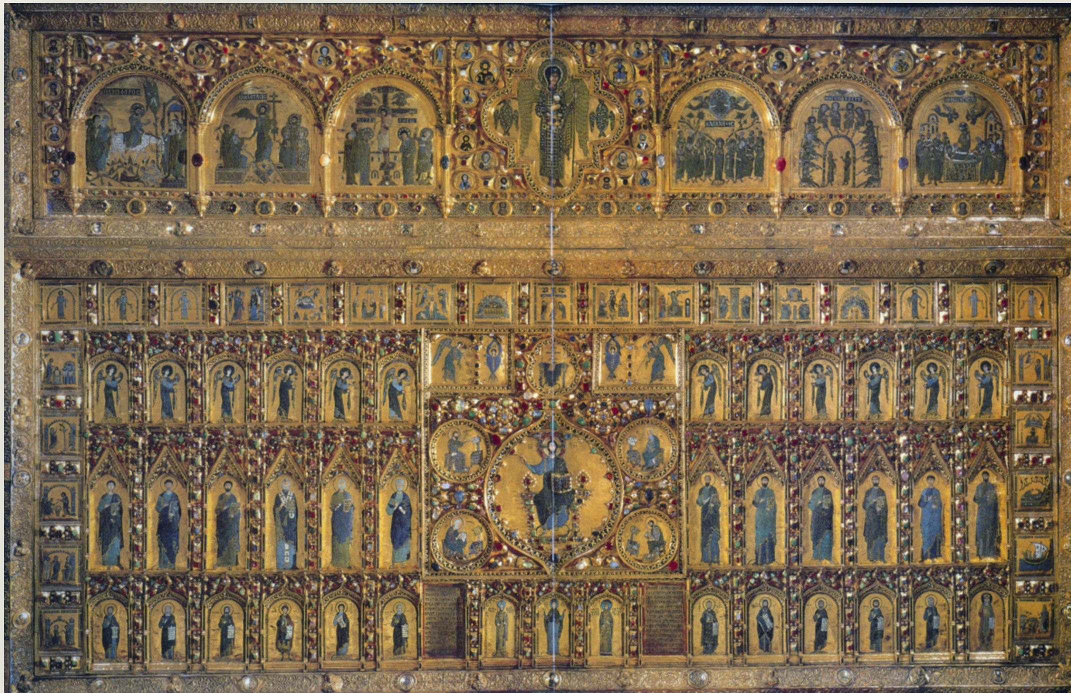
Marco Battaglia
on behalf of the ATLAS Collaboration

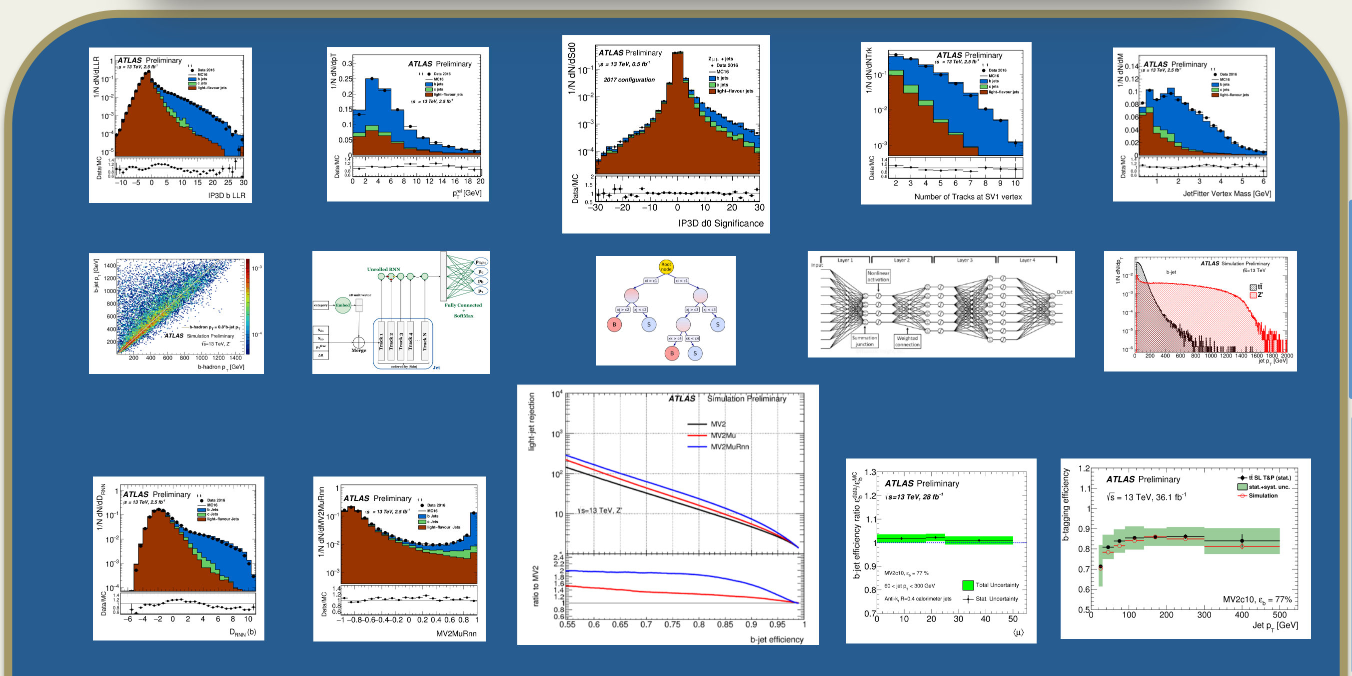
Venezia, July 6th 2017

EPS Conference on High Energy Physics
Venice, Italy 5-12 July 2017

Jet Flavour Tagging, ATLAS, 2017

Pala d'Oro, S. Marco, Venezia (XI cent. - 1345)



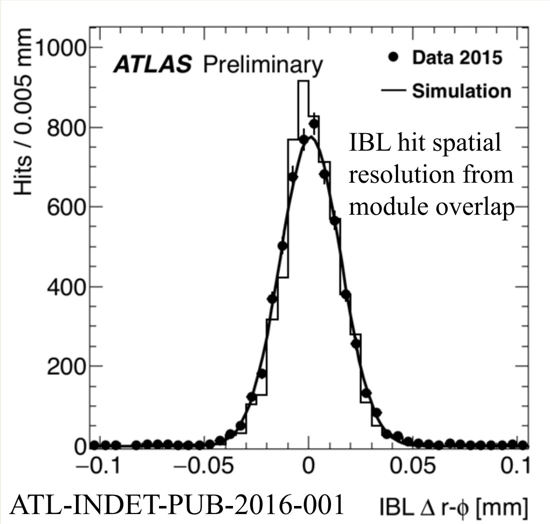


Physics Taggers

High-level Discriminants & Training

Performance, Data/MC Response & Calibration

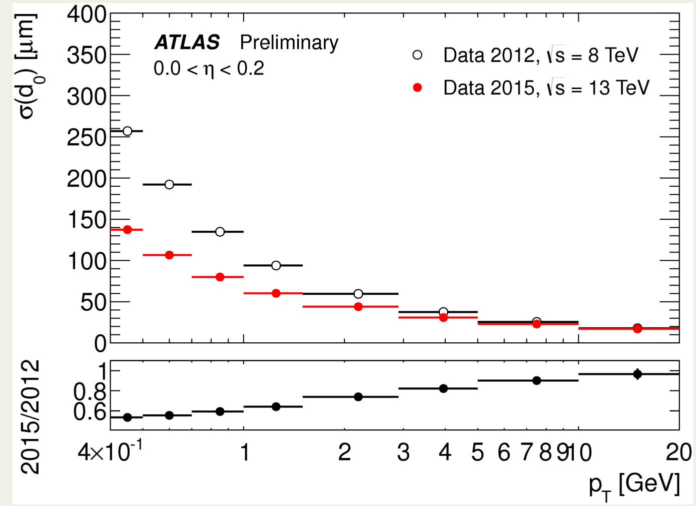
Jet Flavour Tagging in ATLAS: from particle tracks to jet tagging



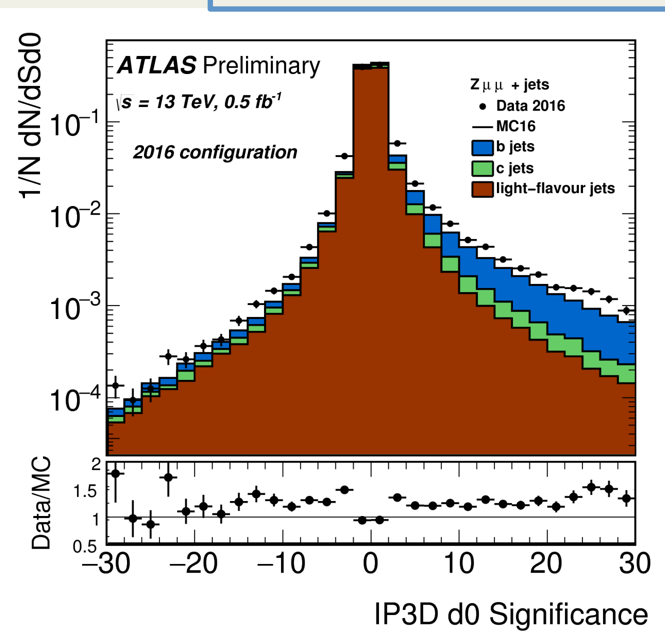
ATL-INDET-PUB-2016-001

Flavour tagging based on particle tracks and their extrapolation to colliding beam envelope;

Introduction of IBL pixel layer located at R~3.3cm with 10μm hit resolution in Run 2 improved track extrapolation resolution up to 40%:

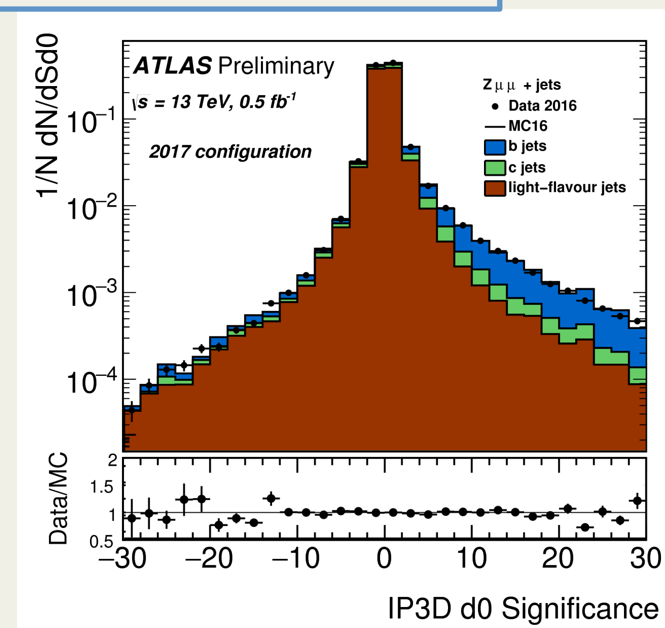


Introduction of Bichsel model of energy deposition in pixel Si
and more realistic ID material modelling in 2017 software configuration:



Improved data/MC agreement of d_0 resolution

Improved data/MC agreement of track-based taggers



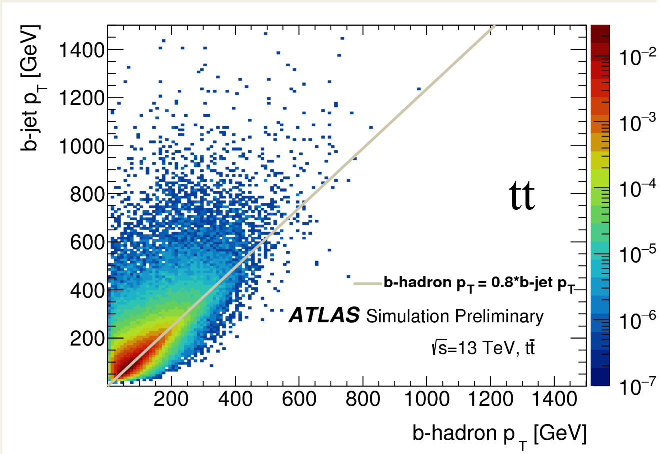
ATL-PHYS-PUB-2017-013

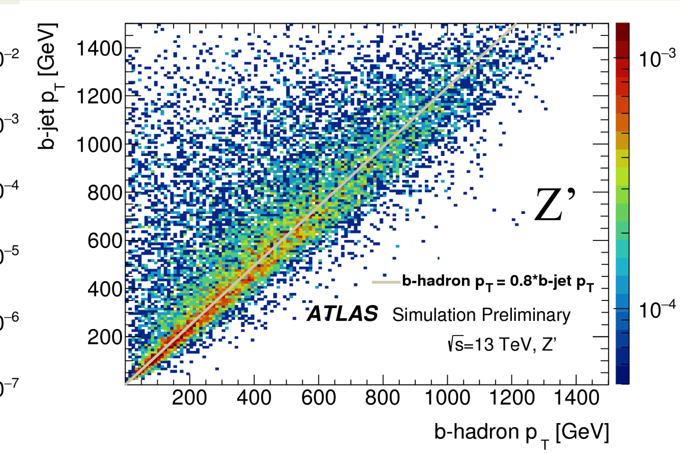
Jet Flavour Tagging in ATLAS: from heavy flavour decays to jet tagging

Jet pT and B hadron pT

Relation between jet p_T and energy to those of original *b*-quark and hadron is key to understanding flavour tagging response as a function of p_T for *b*-jets emitted by process at mass scale *m*.

Correlation between jet p_T and heavy hadron p_T is observed only up $p_T \sim O(m)$, higher p_T jets determined by nearby hadronic activity unrelated to *b* hadron decay.

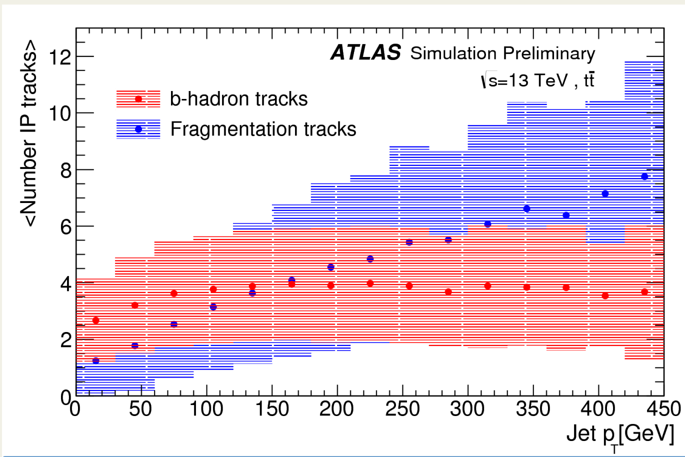




ATL-PHYS-PUB-2017-013

Jet tracks and B decay products

Flavour tagging performance depends on <u>number of charged decay products</u> from *b*- and *c*-hadrons, their <u>separation from event PV</u> and <u>fraction of jet energy</u> that they carry. Discrimination of charged decay products improves with increasing heavy hadron energy while their jet energy fraction decreases with increasing jet energy.

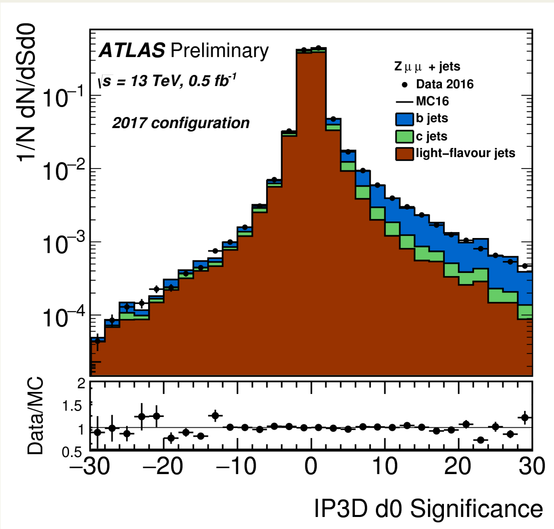


ATL-PHYS-PUB-2016-012

Adopt hybrid (tt + Z') sample for training of high-level discriminants

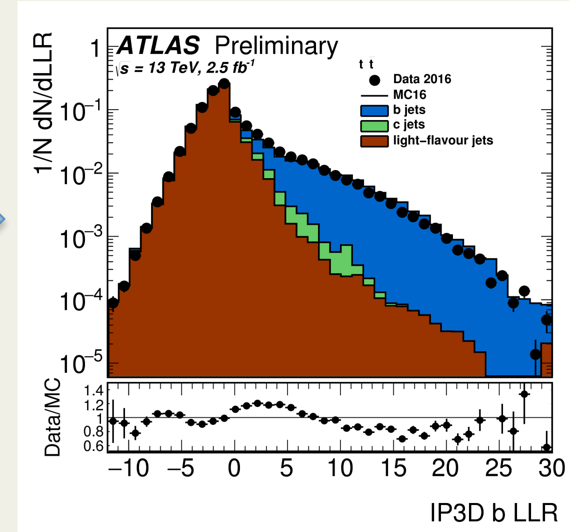
IP significance-ordered tracks in Physics Tagger

Physics Taggers: Track-based Taggers

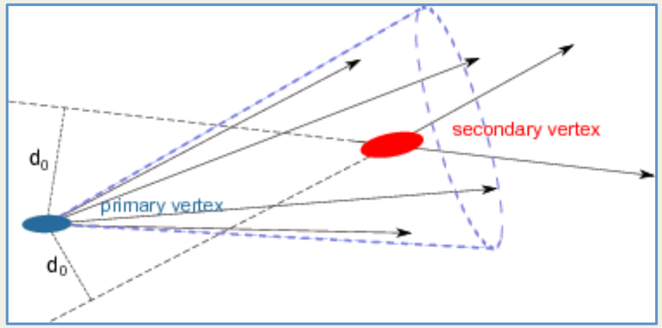


IPTag LLR from IP/σ templates

$$w_{\mathrm{trk}} = p_b/p_q$$
$$w_{\mathrm{jet}} = \sum \log w_{\mathrm{trk}}$$

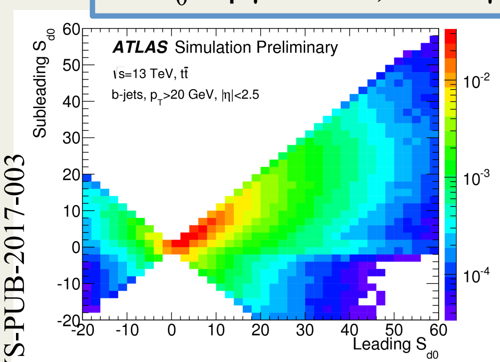


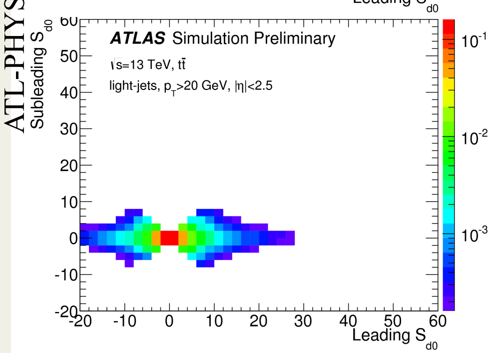
ATL-PHYS-PUB-2017-013



Track Impact Parameter (IP)
$d_0 = \beta\gamma c\tau \sin\theta;\ \theta \sim 1/\gamma$

IP3D & RNNIP Discriminant

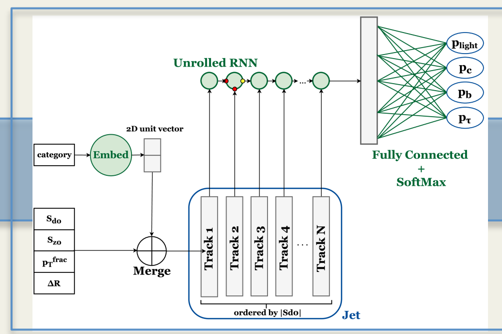




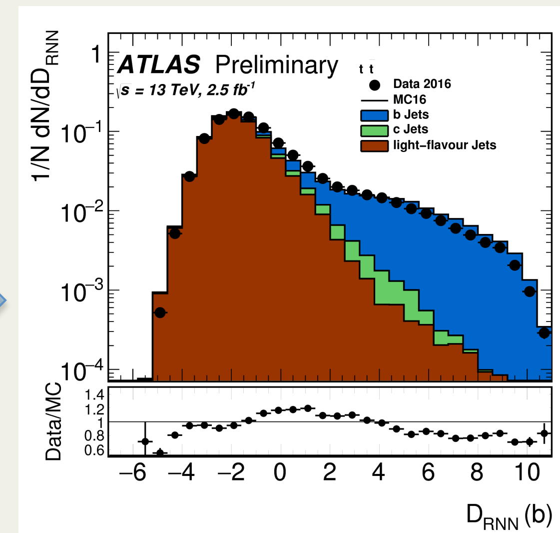
ATL-PHYS-PUB-2017-003

Track Variable
S_{d_0}
S_{z_0}
Category [24]
p_T^{frac}
ΔR(track, jet)

RNNIP Recurrent Neural Net



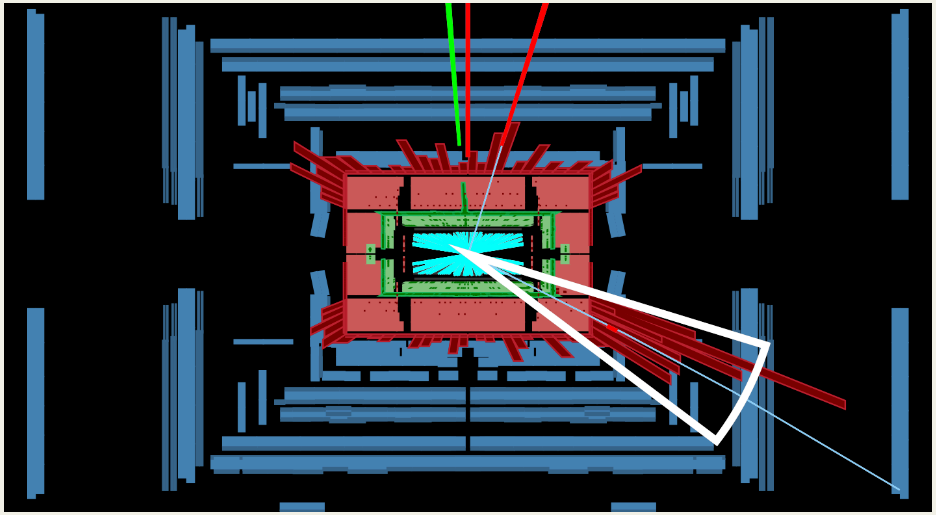
RNN algorithm processes tracks exploiting spatial and kinematic correlations w/o reliance on secondary vertex finding.



x2.5 gain in light jet rejection w.r.t. IP3D at 70% b-jet Eff.

Physics Taggers: Particle-based Taggers

Soft Muon Tagging



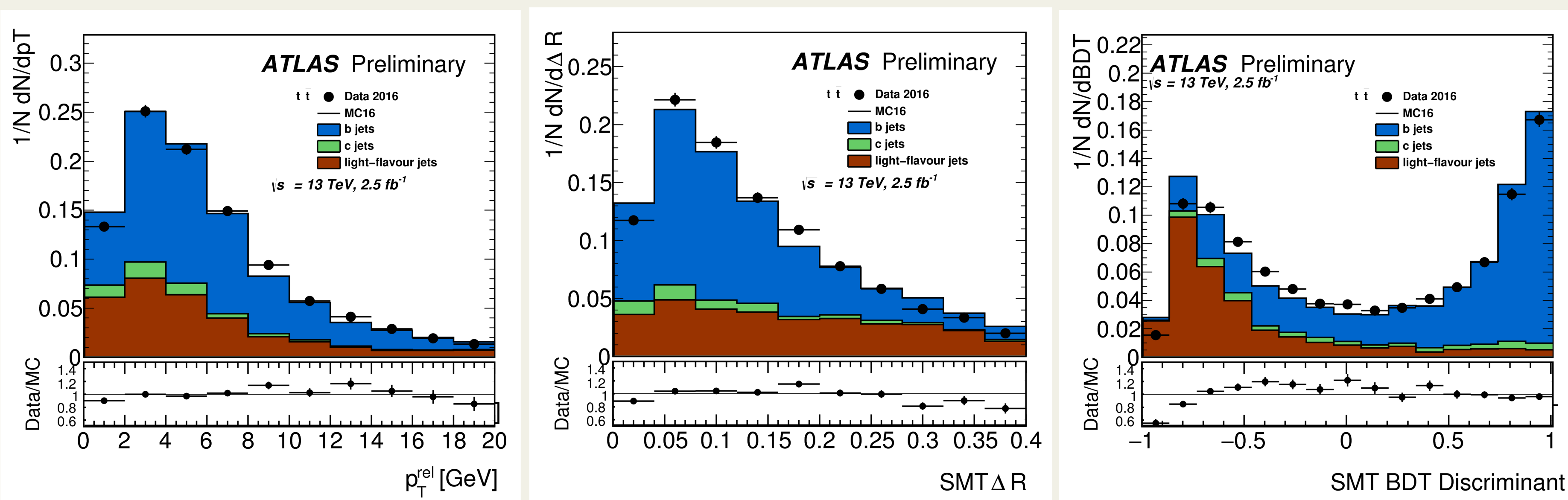
Identify muons from s.l. heavy hadron decay in jet:

- identify genuine prompt muons;
- discriminate muons from b decay chain;
- 3+3 variables fed to SMT BDT;

	Eff.	Misid. Prob
μ in Jet	65%	1.8%

b-jet Tag Eff.	light jet Tag Prob.
10%	0.2%

See also contribution in Poster Session (A. Sciandra)



ATL-PHYS-PUB-2017-013

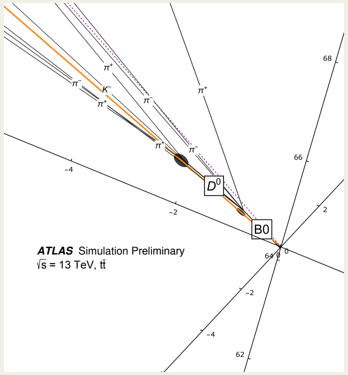
Physics Taggers: Vertex-based Taggers

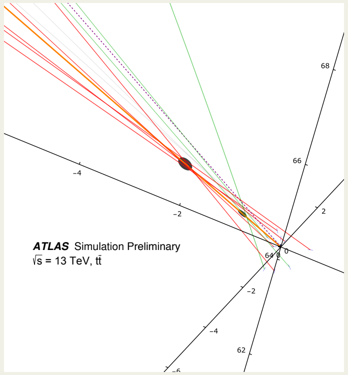
Inclusive Secondary Vertex Finding (SV1)

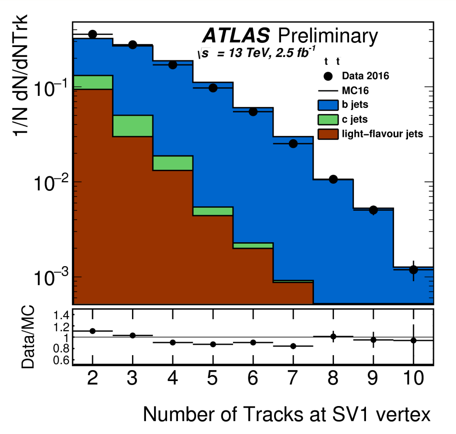
ATL-PHYS-PUB-2017-011

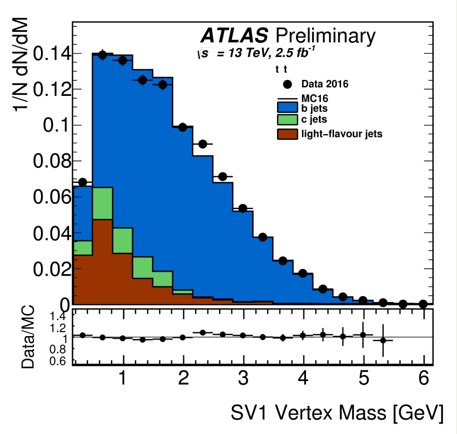
- build two-track vertices using tracks fulfilling strict quality criteria;
- drop those consistent with V^0 decays and interactions;
- merge selected vertices into single inclusive secondary vertex.

ATLAS-FTAG-2017-004









ATL-PHYS-PUB-2017-013

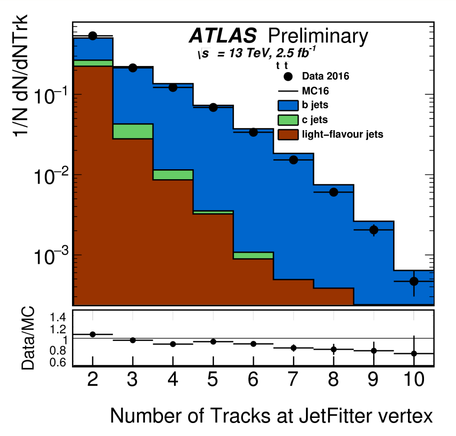
Data/MC Comparison of Rate of Reco Vtx.

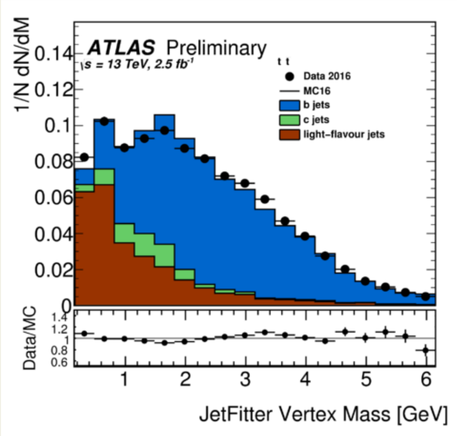
Tagger ($Z \rightarrow \mu^+\mu^-$+jets)	Rate in data	Rate in MC
SV1	0.109±0.001	0.108±0.001
JF secondary vertex	0.145±0.001	0.147±0.001
JF tertiary vertex	0.0247±0.0005	0.0249±0.0005

Tagger ($t\bar{t}$)	Rate in data	Rate in MC
SV1	0.189±0.001	0.187±0.001
JF secondary vertex	0.229±0.001	0.217±0.001
JF tertiary vertex	0.0467±0.0006	0.0504±0.0005

Topological Decay Reconstruction (JetFitter)

- reconstruct decay chain from crossings of selected tracks with jet axis;
- apply modified KF formalism where distance from crossing with jet axis replaces vtx position;
- provides secondary and tertiary vertices (including single prong)





ATL-PHYS-PUB-2017-013

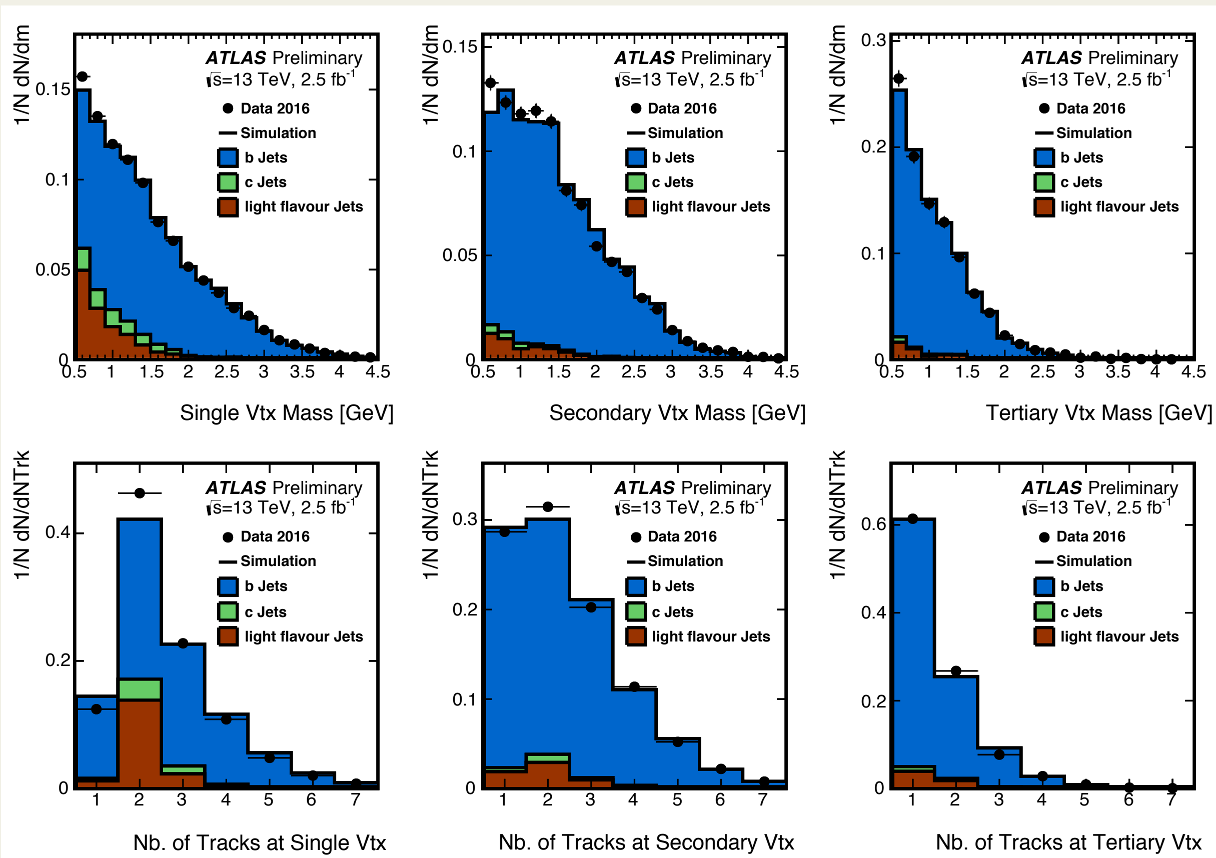
Physics Taggers: Vertex-based Taggers

Topological reconstruction of B-hadron decays important to efficiently and cleanly separate b- from c- and light-jets

Single Vertex

Secondary Vertex

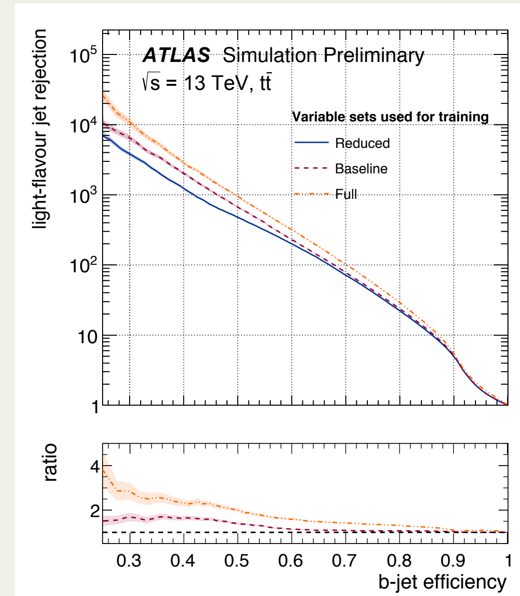
Tertiary Vertex

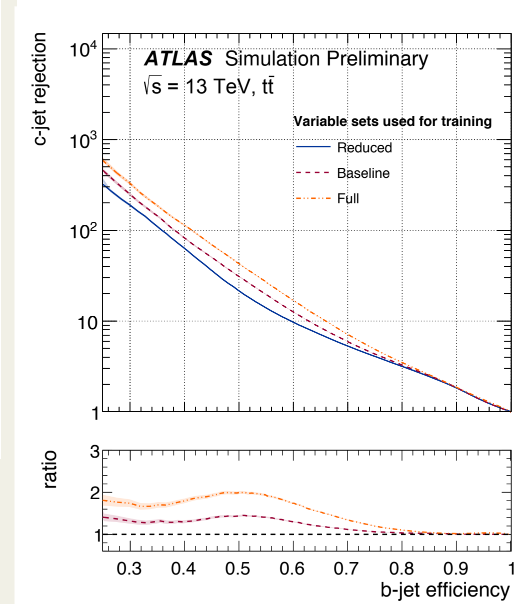


ATLAS-FTAG-2017-004

Detailed information on reconstructed topology and kinematics of detached vertices shows clear advantages in rejection of non-b jets:

b-Tagging with Topological Variables





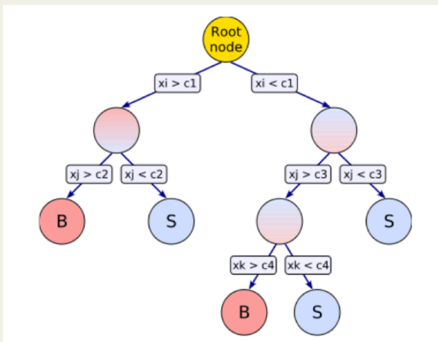
See also contribution in Poster Session (G. Gilles)

ATLAS-FTAG-2017-004

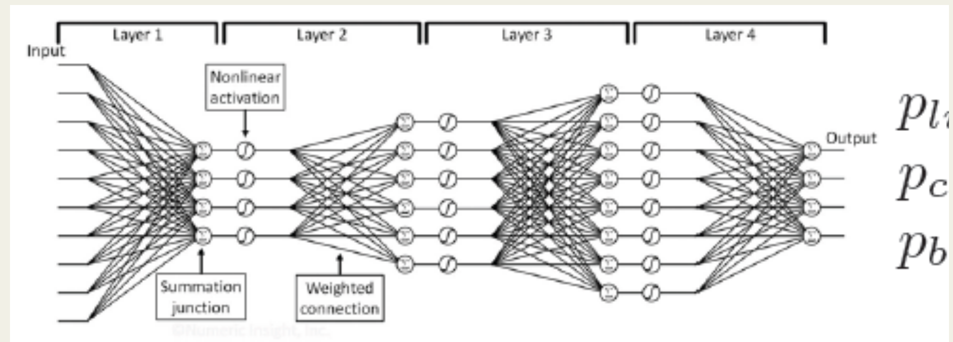
Multivariate Discriminants: High-level Taggers and Training

High level taggers combine inputs from track-, particle and vertex-based physics taggers using multivariate classifier to maximise the *b*-tagging performance.
2017 ATLAS configuration adopts two high-level tagger variants:

Boosted Decision Tree (BDT) MV2
evolution of tagger used in 2016



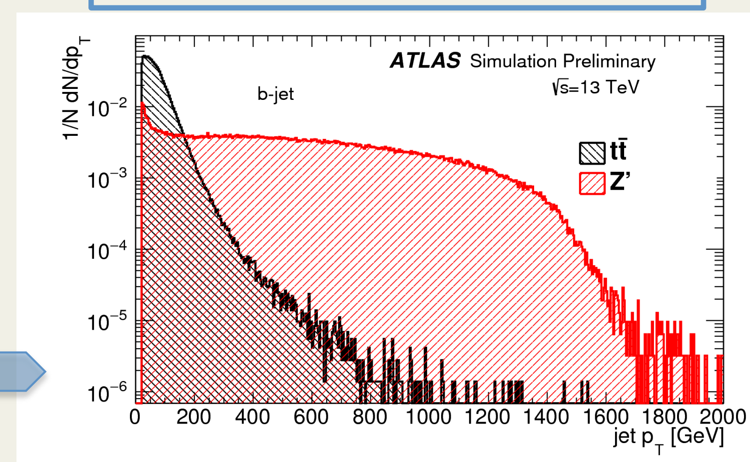
Deep Learning Neural Network
new DL1 tagger



New training strategy uses sample made of *tt* events, to characterise the low p_T region, and Z' events with flat p_T spectrum, to probe high p_T regime.

New *hybrid* sample obtained by including *b*-jets from $t\bar{t}$ for *b*-hadron p_T <250 GeV and from Z' for *b*-hadron p_T >250 GeV.

Jet p_T of Hybrid Training Sample

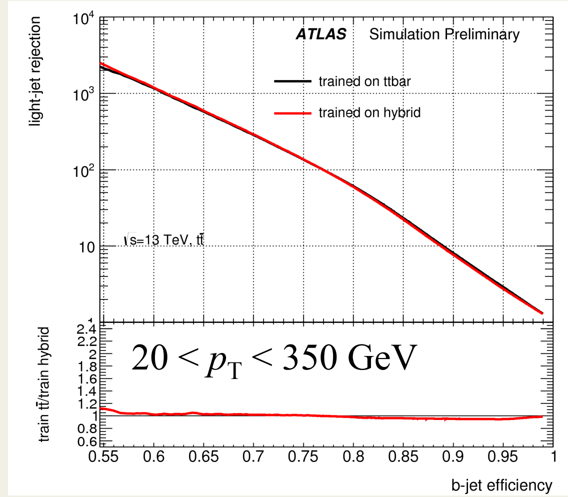


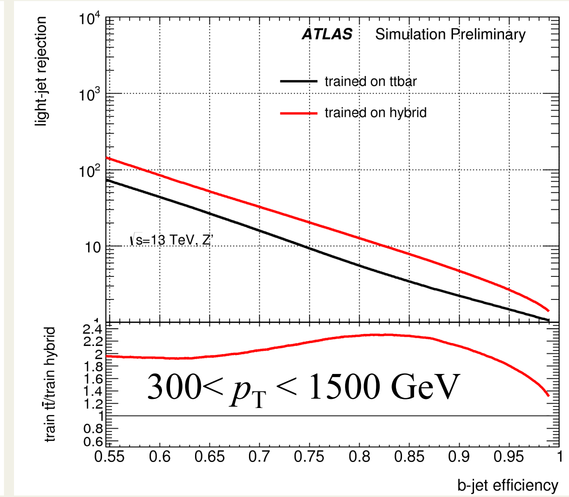
See also contribution in Poster Session (F. Di Bello and M. Lanfermann)

ATL-PHYS-PUB-2017-013

Multivariate Discriminants: Performance

Validation of hybrid training: light-jet rejection vs. b-jet eff. for tt and hybrid training

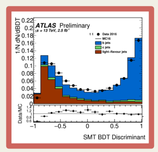




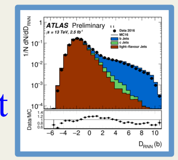
See also contribution in Poster Session (F. Di Bello)

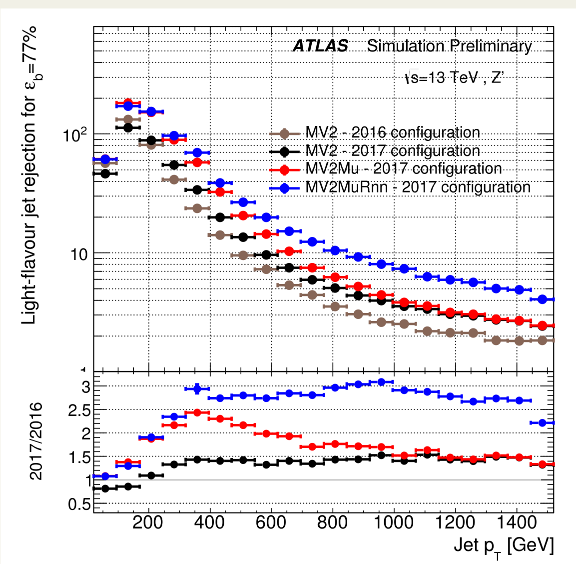
MV2 Rejection for constant 77% b-jet efficiency vs jet p_T

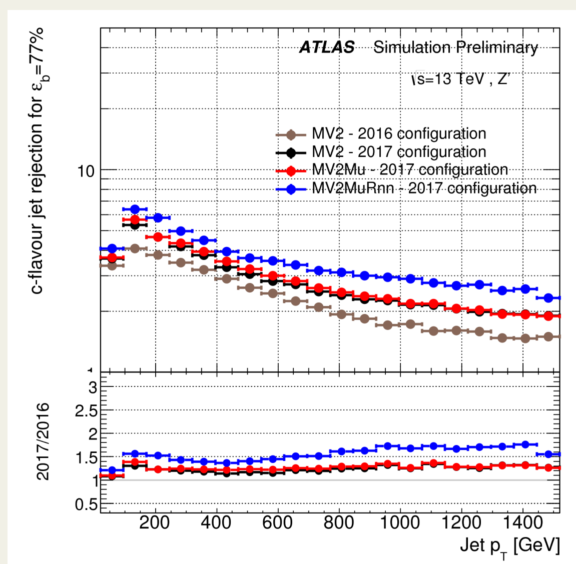
Track IPxD Variables + Vertexing Variables + Particle SMT Discriminant



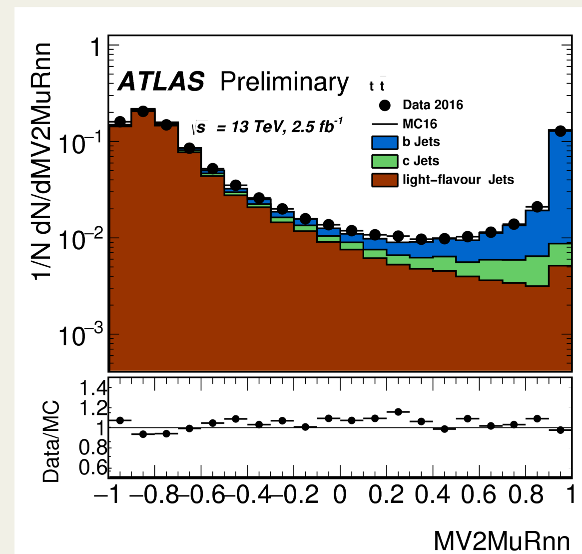
+ RNNIP Discriminant







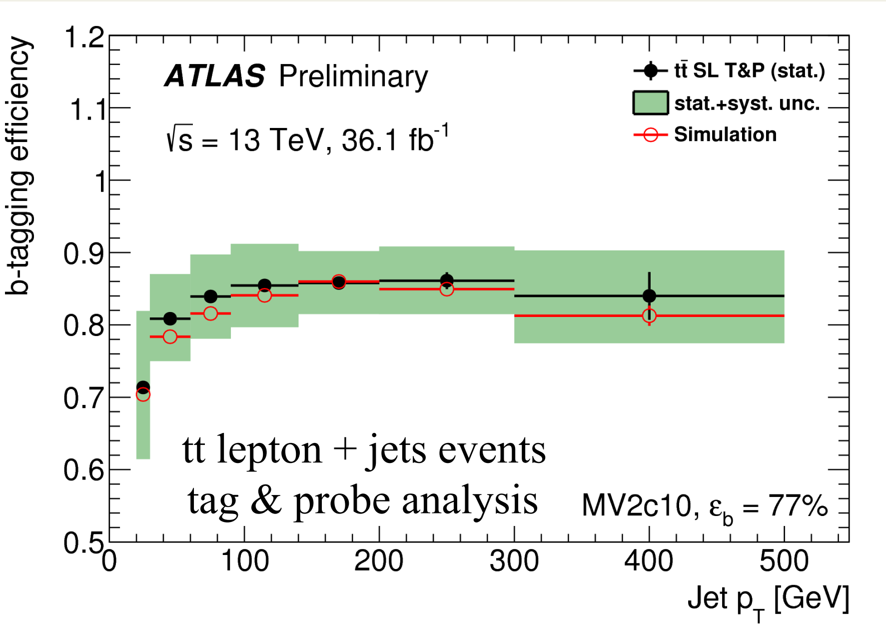
MV2 Data/MC (eμ+jets events)

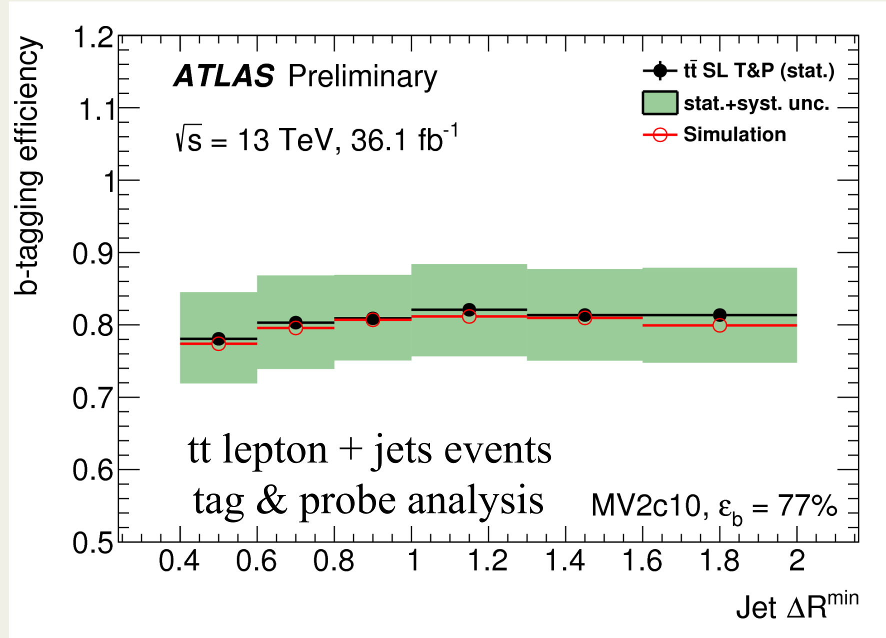


ATL-PHYS-PUB-2017-013

Multivariate Discriminants: Performance in Physics Channels and Universality

Tagging performance depends mainly on: jet p_T, jet Rapidity and jet distance to closest jet / track density :





ATLAS-FTAG-2017-003

See also contribution in Poster Session

Test universality of tagger performance on different physics processes:

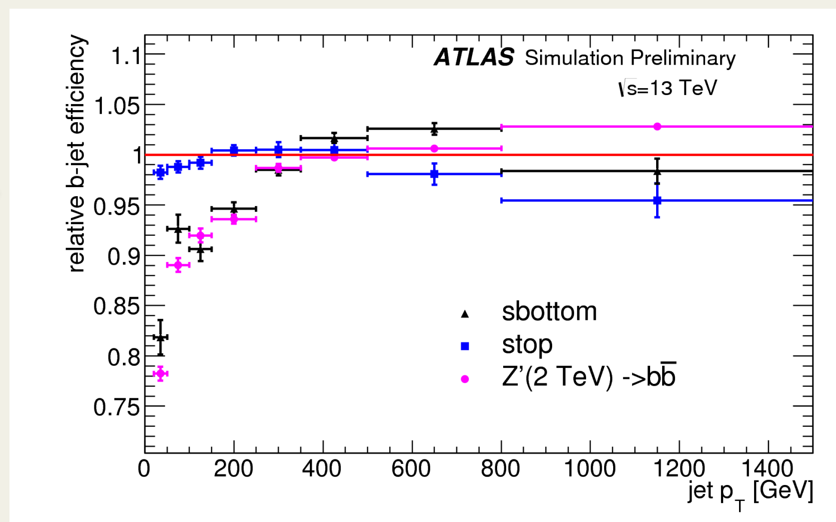
$$pp \to \tilde{t}\tilde{t} \to t\chi_1^0 t\chi_1^0,\ m(\tilde{t}) = 600\ \text{GeV}$$

$$pp \to \tilde{b}\tilde{b} \to b\chi_1^0 b\chi_1^0,\ m(\tilde{b}) = 600\ \text{GeV}$$

$$pp \to Z' \to bb,\ m(Z') = 2\ \text{TeV}$$

$$pp \to VH,\ H \to b\bar{b}$$

b-jet Eff. VH/hybrid	1.03±0.02
c-jet misid. Prob. V+jets/hybrid	0.87±0.08
light-jet misid. Prob. V+jets/hybrid	0.88±0.11



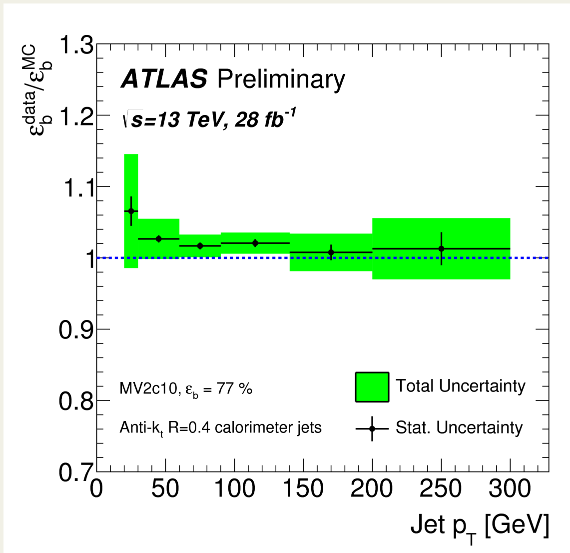
ATL-PHYS-PUB-2017-013

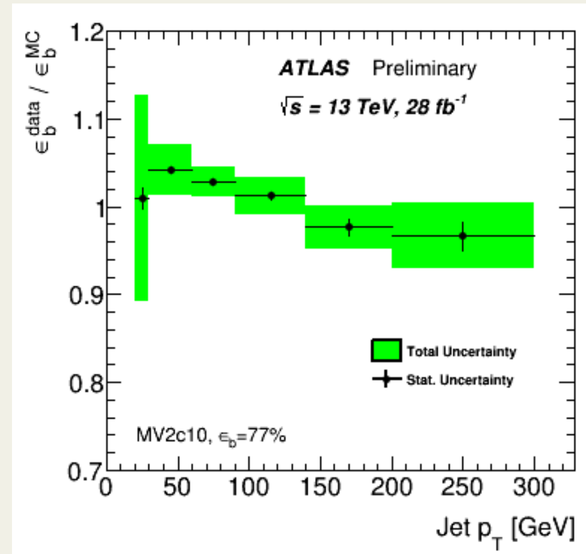
Multivariate Discriminants: Calibration for b-jet Efficiency

Calibrate b-jet Tagging efficiency (2016 configuration) using ttbar dilepton events:

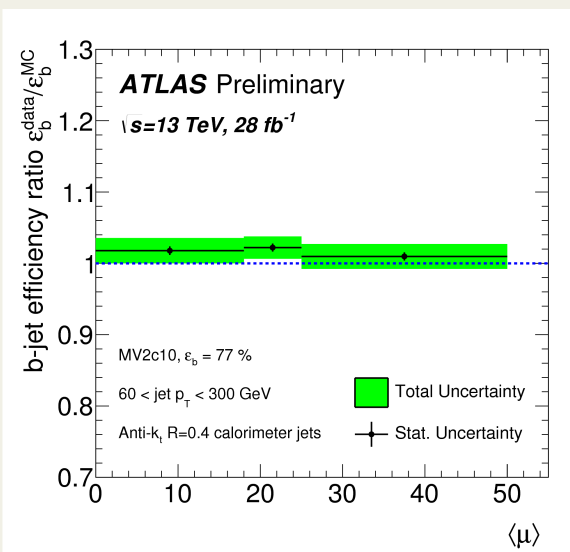
Tag and probe in tt eμ + jets events	Combinatorial likelihood in tt 2l + jets events

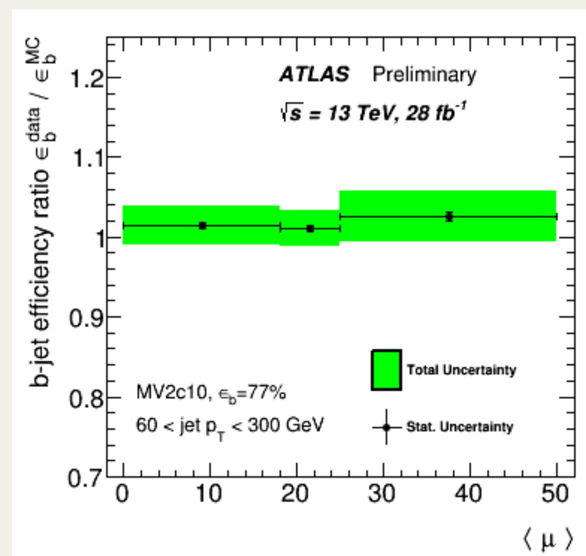
b-jet Eff. Data/MC Scale Factors vs jet p_T





b-jet Eff. Data/MC Scale Factors vs Pile-Up





ATLAS-FTAG-2017-001

ATLAS-FTAG-2016-003
Updated from ATLAS-CONF-2014-004

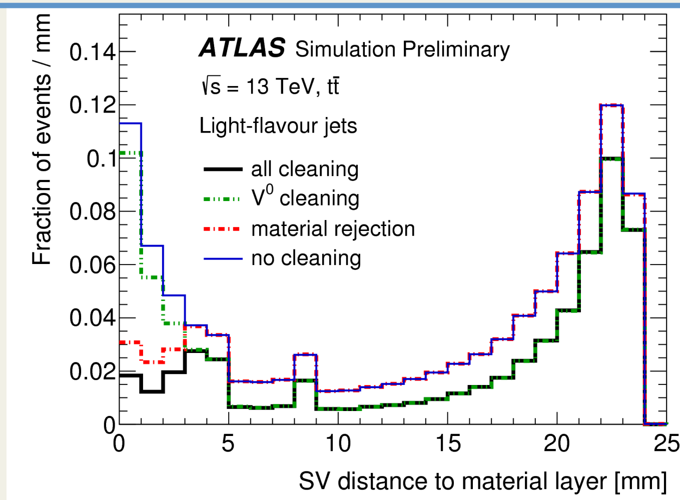
Multivariate Discriminants: Calibration for light-jet Misidentification

Mis-identification of light jets due to several sources including:

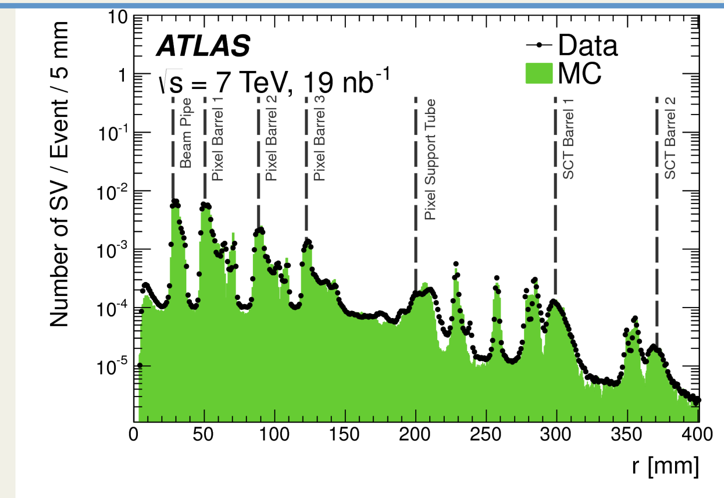
- resolution effects
- material interactions
- long-lived particles

Rejection of Mis-id Sources in Secondary Vtx Reco

ATL-PHYS-PUB-2017-011



Material Mapping with Hadronic Interaction Vertices

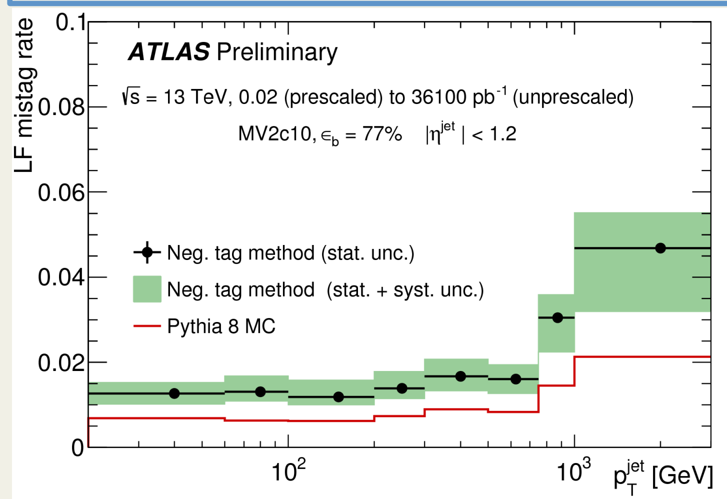


JINST 11 (2016) P11020

Track extrapolation resolution effects can be extracted by impact parameter resolution adjusted MC and taggers fed with tracks with negative lifetime-signed impact parameters:

Light-flavour jet Mis-identification Rate (2016 configuration) from di-jet events in data and MC using <u>negative tag method</u>:

Light-flavour jet Mis-identification Rate



ATLAS-FTAG-2017-002

See also contribution in Poster Session

Conclusions and Outlook

Improvements and innovations in physics taggers combined with new approaches to MVA and training samples resulted in more performant flavour-tagging algorithms available for analysis of 2017-2018 ATLAS data.

New RNNIP tagger exploits information from track-by-track correlations within a jet, providing superior and complementary tagging to i.p.-based IPxD taggers.
Soft muon tagging, offers additional discriminating power for *b* jets with s.l. decays.

New hybrid training sample including tt events and hadronic decays of Z' with flat p_{T} boosts tagger response at high jet p_{T}. Deep Learning NN-based algorithm developed in parallel with BDT-based MV2 tagging algorithm offer higher degree of integration of low- and high-level taggers and new training opportunities to mitigate impact of modelling systematics.

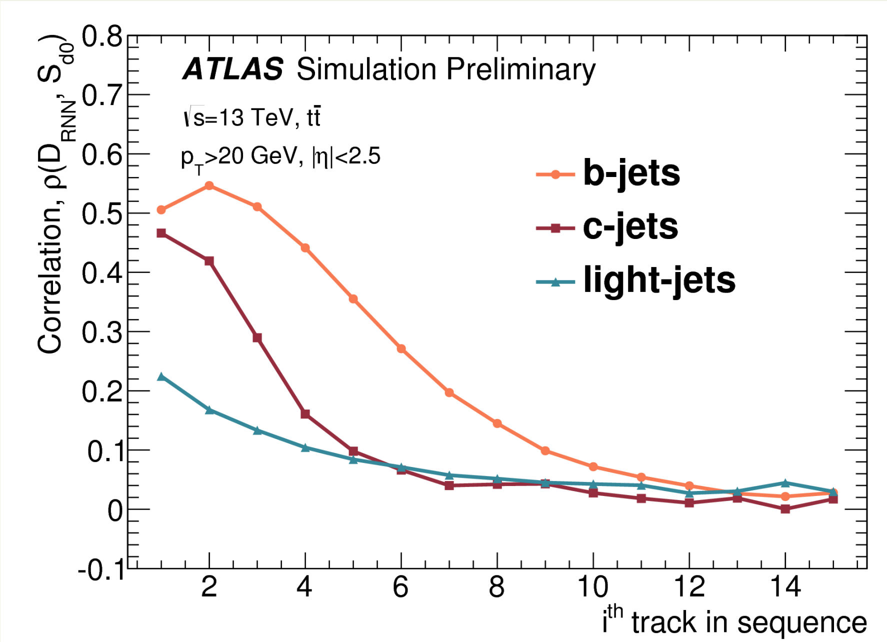
Optimisation of flavour tagging response included 2016 collision data and MC comparisons, from track variables to physics tagger observables and high-level tagger response for event samples enriched and depleted in *b* jets. Changes in tracking simulation in 2017 configuration improved data/MC agreement in *b*-tagging inputs and output distributions.

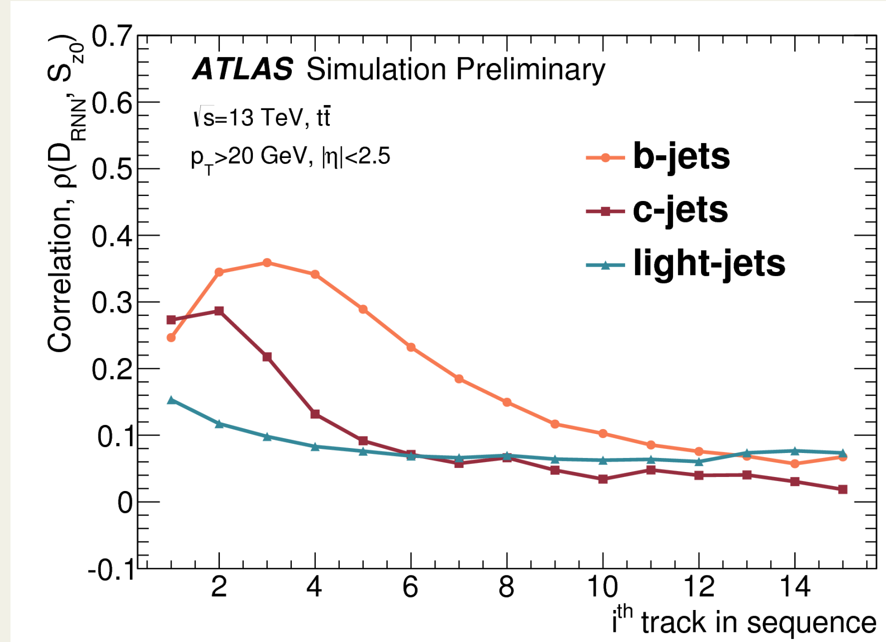
Baseline MV2 tagger efficiency verified on physics samples of different kinematics:
$\tilde{t}$ and $\tilde{b}$ pair production, Z' and VH, $H \rightarrow bb$.
Sensitivity to pile-up conditions and modelling of *b* production and decay characterised.

Several posters describe ATLAS strategies for physics and high-level taggers in detail.

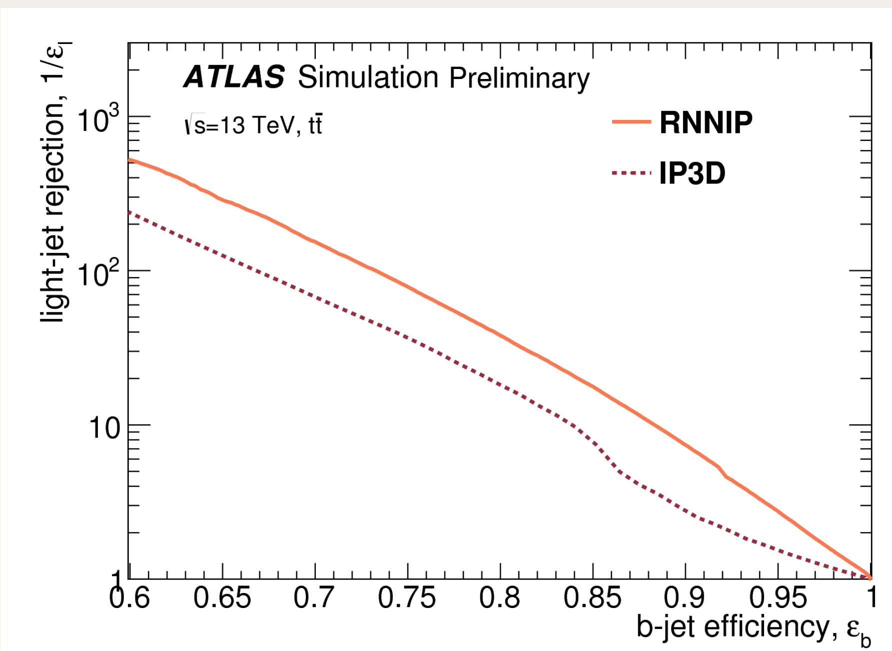

Back-up Information

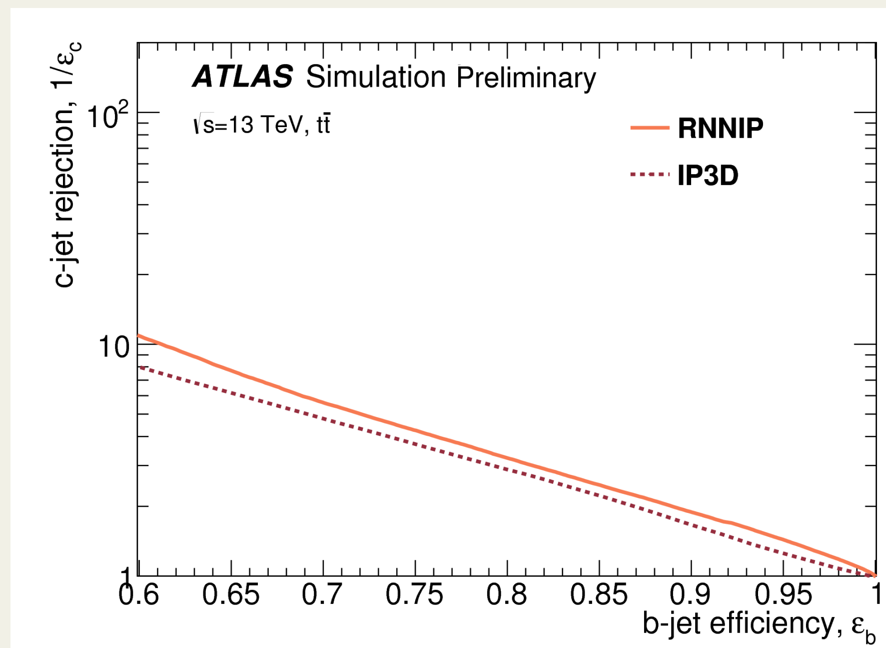
Physics Taggers: Track-based RNNIP Tagger





ATL-PHYS-PUB-2017-003

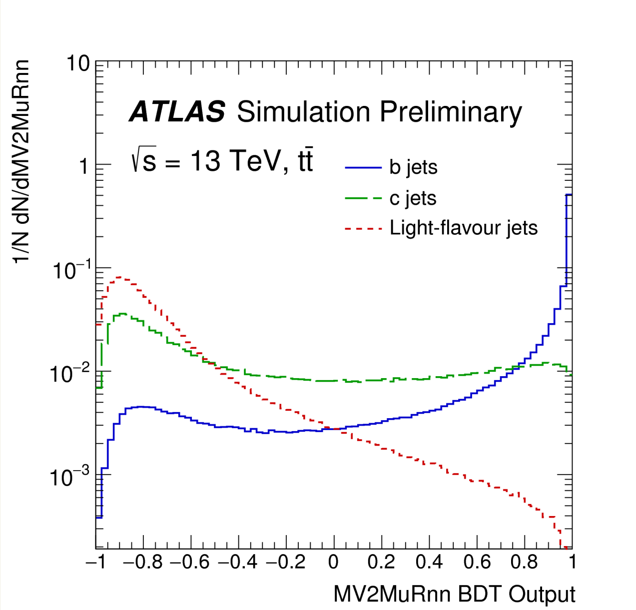


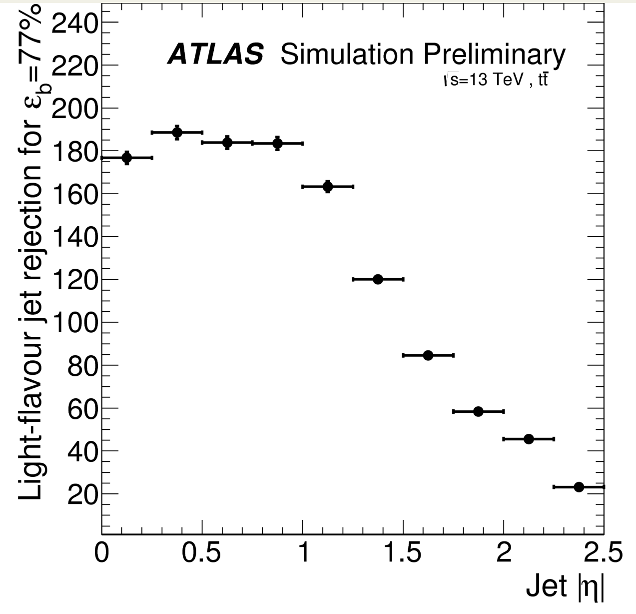


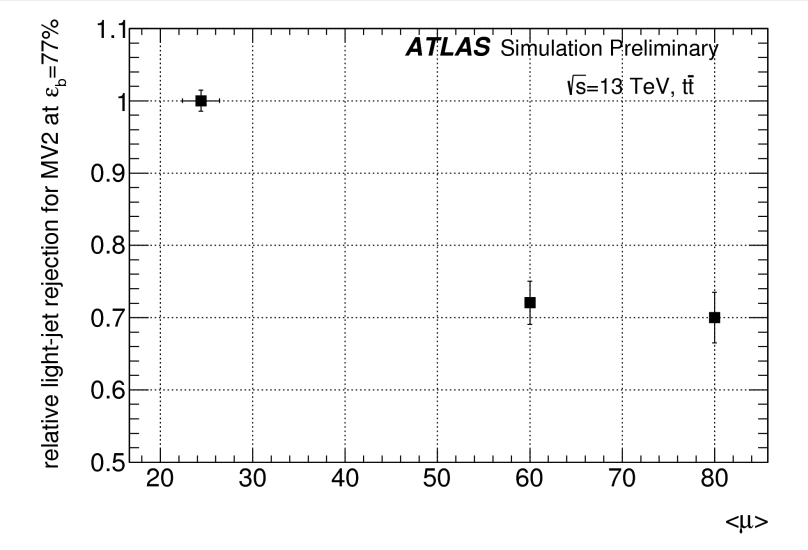
ATL-PHYS-PUB-2017-013

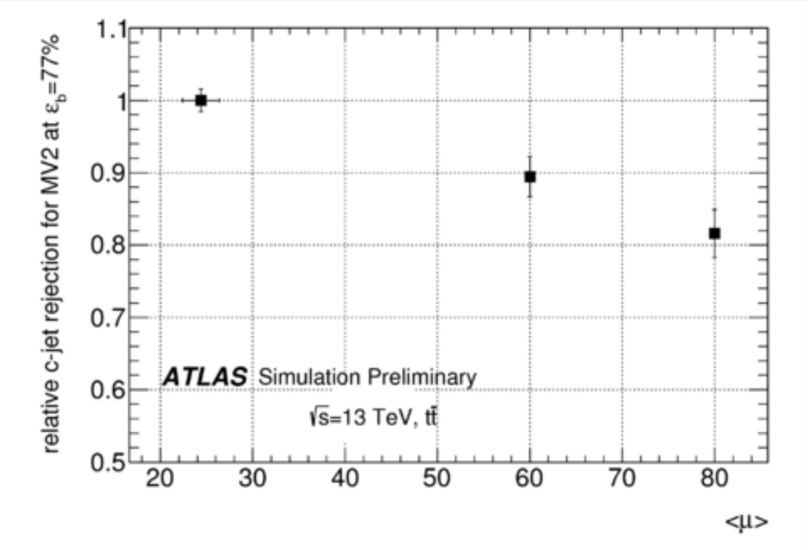
Multivariate Discriminants: MV2 Performance

ATL-PHYS-PUB-2017-013









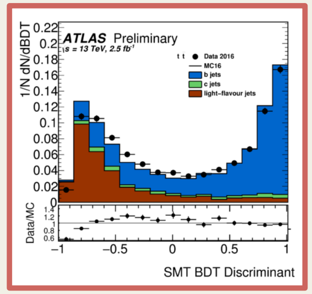
Multivariate Discriminants: MV2 Performance

ATL-PHYS-PUB-2017-013

Track IPxD Variables
+ Vertexing Variables

+

Particle SMT Discriminant



+

RNNIP Discriminant