b-tagging with ATLAS

Laurent Vacavant

on behalf of the ATLAS Collaboration

All material to be available in: ***Expected Performance of the ATLAS Experiment, Detector, Trigger and Physics*****, CERN-OPEN-2008-20**

Physics at LHC 2008 – Split, Croatia – October 2 2008

View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by CERN Document Server

Identification of b-jets is a key ingredient for the whole high-p_T physics program of ATLAS

- Top physics: (talks from F.P.Shiling and W.Verkerke, B. Clement, P. Cavalleri next)
  - favorable situation: S/B 10x higher than Tevatron
  - b-tagging still useful to kill light jet background:
    - example for lepton+jets channel, 100 pb^{-1}:
      - S/B 2x (4x) better requiring one (two) jets to be b-tagged (no remaining light jet background expected in the latter case)

modest ε OK

- Standard Model Higgs (searches & properties): (talk by L. Fayard)
  - Natural interest for low mass region ($\Gamma_{ff} \sim m_f^2$)
  - revived interest for WH(bb) (w/ high-p_T jets) (Butterworth et al., PRL 100,242001,2008)
  - ttH(bb) channel + other ttH channels (talk by I. Ludwig)
    - signal ~0.4 pb, 4 b-jets in signal, huge background
  - background suppression (top) for other channels

$\varepsilon \gtrsim 70\%$ needed !

- SUSY:
  - mostly Higgs sector (coupling to b enhanced by sinα/cosβ for 2HDM models like MSSM) (talk by L. Fayard)
  - also inclusive searches: +20% in reach (Kadala et al.,arXiv:0803.0001, see posters)

higher jet p_T or soft (bbH), taus

- Exotics: Z', little Higgs, etc

TeV-ish jets !

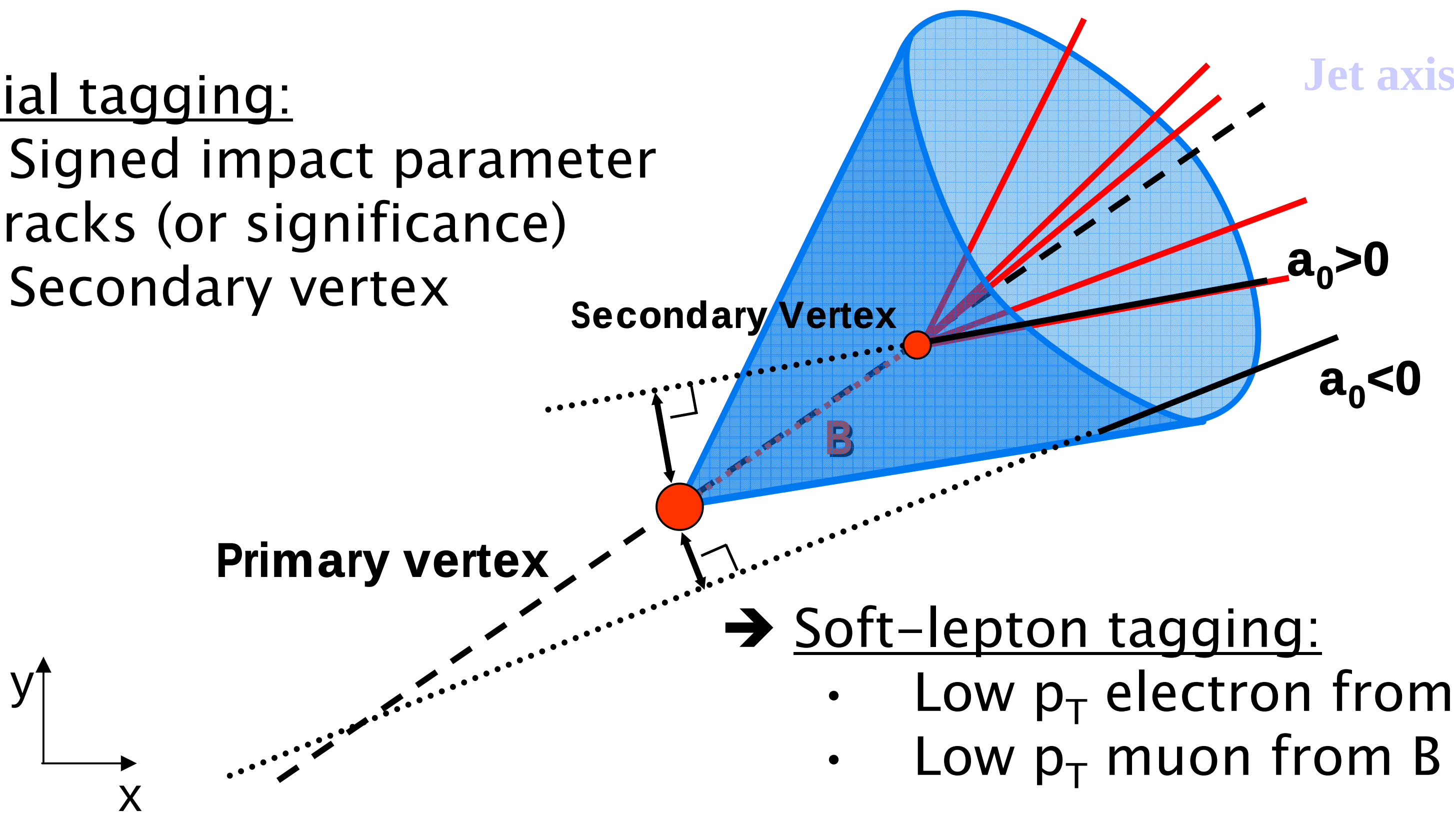
b-tagging basics

Hard fragmentation of b quarks $x_B \sim 70\%$
High mass $m_B \sim 5$ GeV
Lifetime of B hadrons:

$c\tau \sim 470\ \mu m$ (mixture $B^+/B^0/B_s$) , $\sim 390\ \mu m$ (Λ_b)
for E(B) ~ 50 GeV, flight length ~ 5 mm, $d_0 \sim 500\ \mu m$

➔ Spatial tagging:

- Signed impact parameter of tracks (or significance)
- Secondary vertex



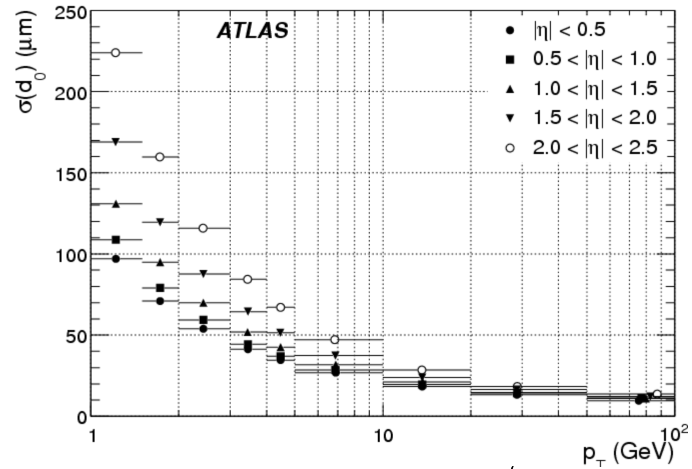
➔ Soft-lepton tagging:

- Low p_T electron from B (D)
- Low p_T muon from B (D)

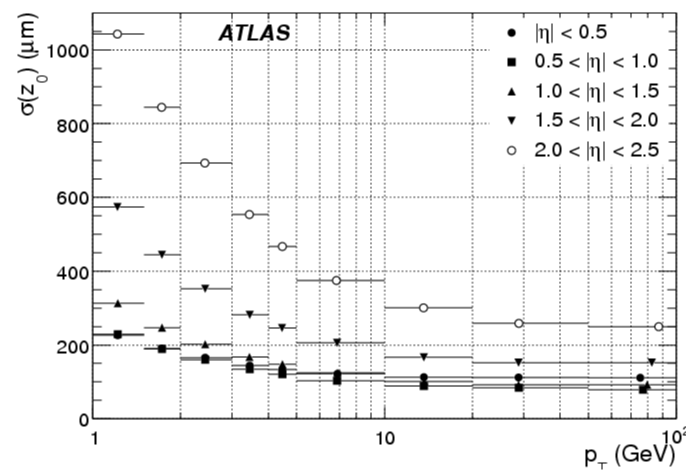
(limited by Br: around 20% each)

b-tagging ingredients

- Jets (direction):
  - to assign tracks ($\Delta R<0.4$)
  - to sign impact parameters
- Tracks:
  - impact parameter resolution (35μm for a central 5 GeV track)
  - tracking in (dense) jets
  - specific quality requirements
    - N(hit B-layer) > 0
    - shared hits, etc
- Primary vertex:
  - mostly needed along Z ($\sigma\sim50\mu m$)
  - @ high-luminosity
- Lepton ID (soft leptons)



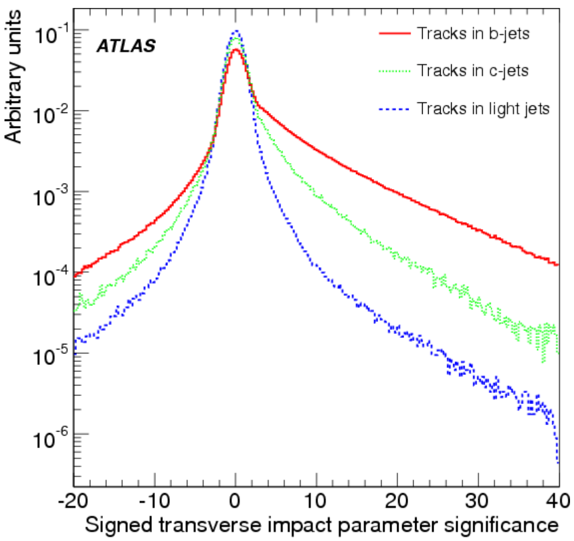
$\sigma(d0) \sim 10 \oplus 150/p_T\sqrt{\sin\theta}\ \mu m$



Impact parameter-based b-tagging

Impact parameter:

- signed w.r.t. jet direction
- significances (d_0/σ)



Simplest algorithm:

- counting tracks with large d_0 or d_0/σ

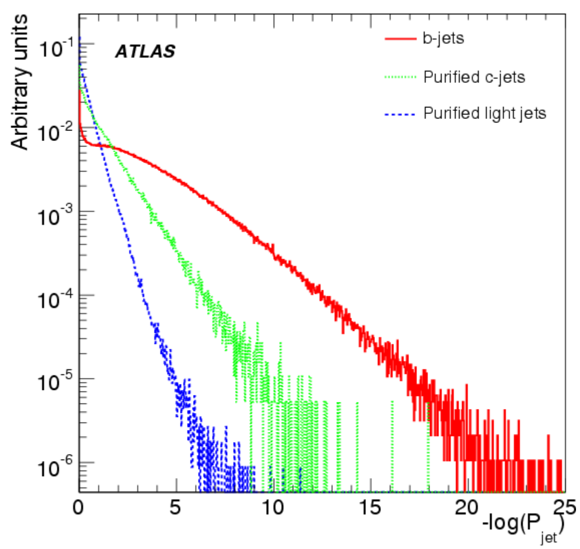
JetProb algorithm: (à la ALEPH)

- compatibility of tracks with primary vertex
- resol. function from data

$$\mathscr{P}_i = \int_{-\infty}^{-|d_0^i/\sigma_{d_0}^i|} \mathscr{R}(x)dx$$

$$\mathscr{P}_0 = \prod_{i=1}^{N} \mathscr{P}'_i$$

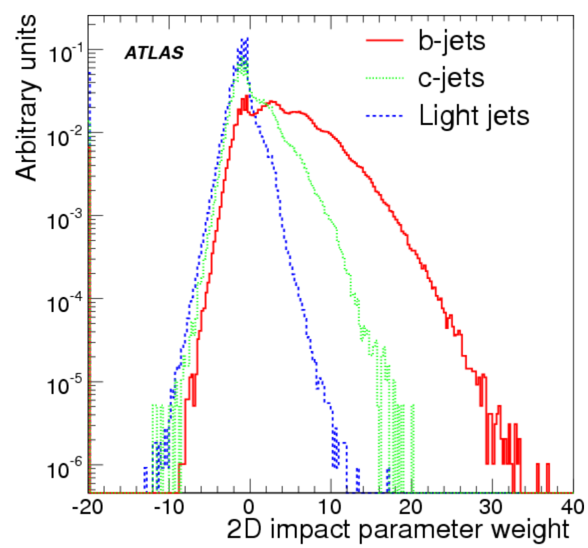
$$\mathscr{P}_{jet} = \mathscr{P}_0 \sum_{j=0} \frac{(-ln\mathscr{P}_0)^j}{j!}$$



IP likelihood ratios:

- smoothed & normalized significances for the two hypotheses: b(S), u(S)
- for each track: ratio b(S)/u(S)
- 2D PDF for (d_0,z_0)

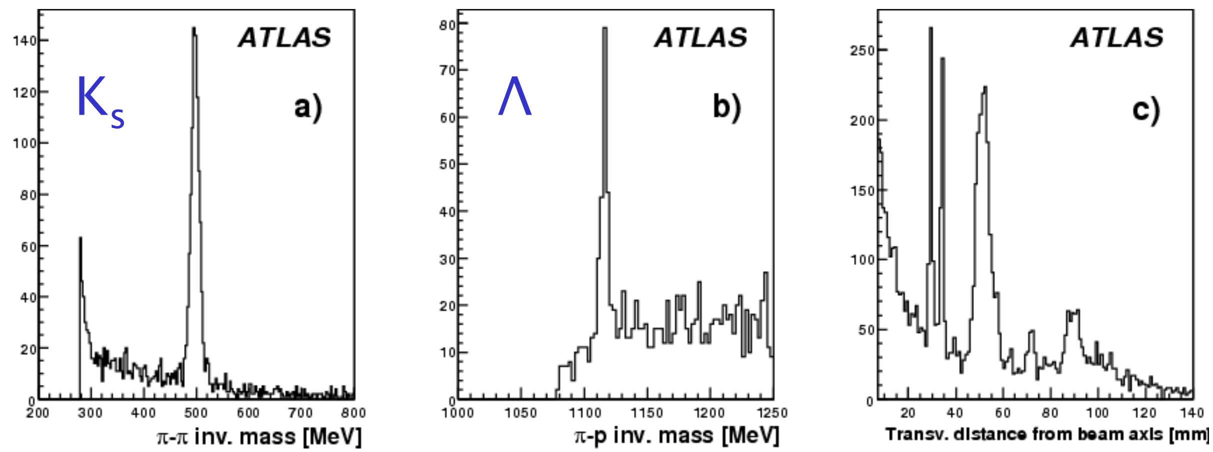
$$W_{jet} = \sum_{i=1}^{N_{tr}} \ln \frac{b(S_i)}{u(S_i)}$$



Secondary vertex-based b-tagging

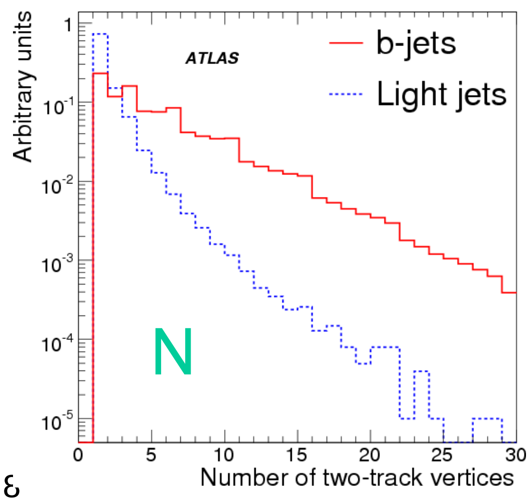
Inclusive vertex:

1. Find all two-track vertices (tracks away from primary vertex: $L_{3D}/\sigma > 2$)
2. Remove vertices compatibles w/ γ conv., K_s, Λ, material inter.



3. Fit all remaining tracks into single inclusive vertex
4. Remove iteratively worst tracks
5. Use distance to PV or properties of the fitted vertex for discrimination:
   - number of two-track vertices
   - mass of the vertex
   - energy fraction

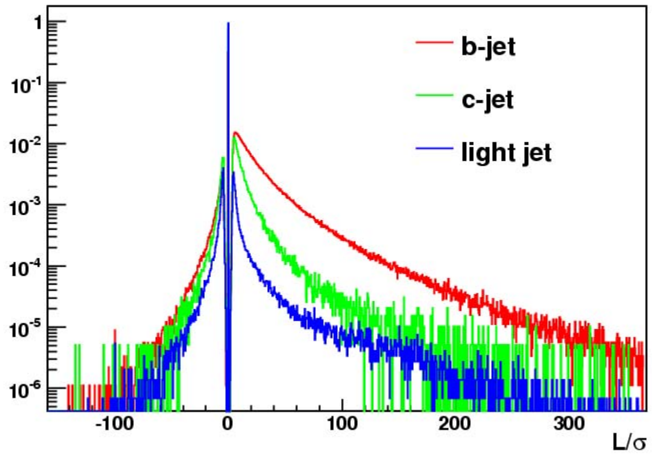
(NB: L_{xy} not used, correlated with IP)



Secondary vertex-based b-tagging

<u>Basic algorithm:</u>

- SV0 tagger: returns L_{3D}/σ

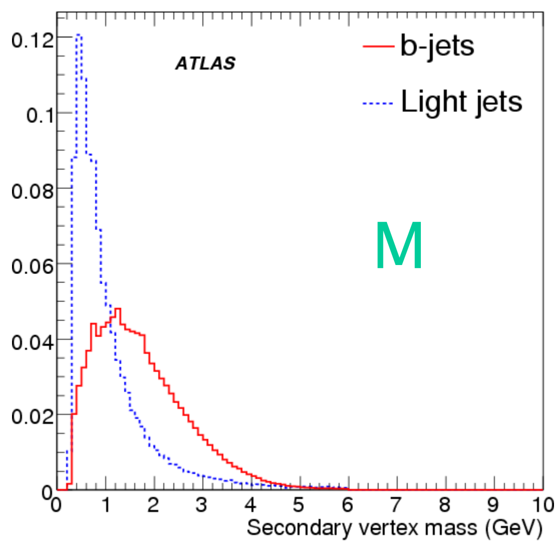


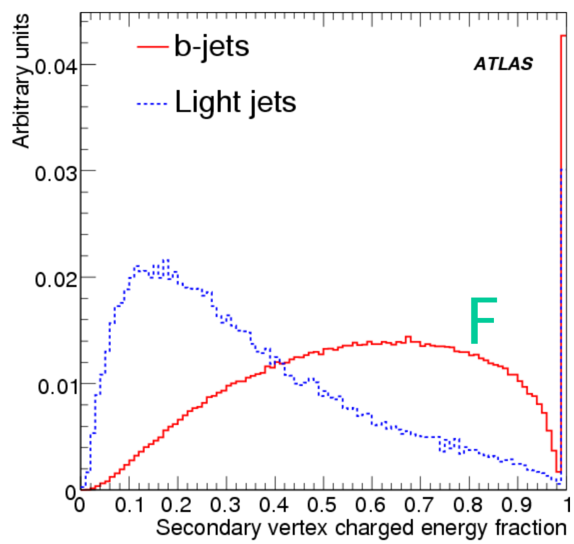
<u>LR approach:</u> templates (PDFs) for b and light hypotheses:

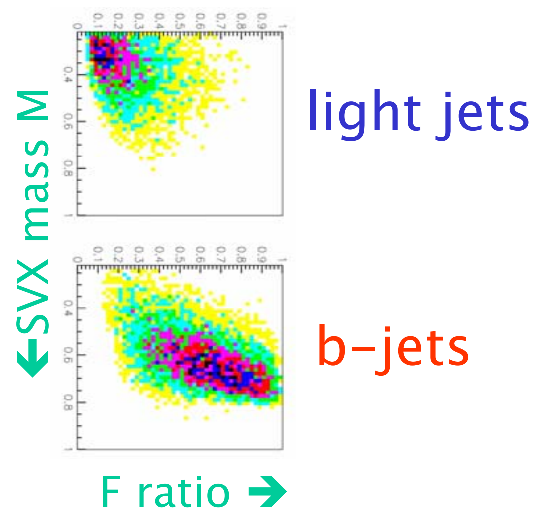
- SV1 tagger: 1D for N', 2D for (M', F')
- SV2 tagger: 3D for (N', M', F')
- Fold in also probability ε to find vertex

$$W_{jet} = W_{tracks} + W_{vertex}$$

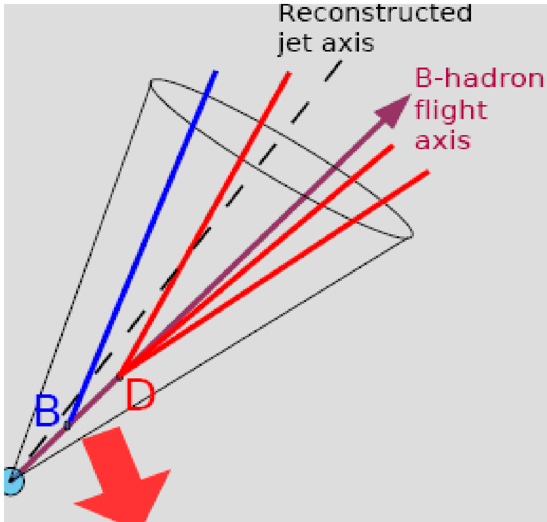
$$= \sum_{i=1}^{N_{track}} \log \frac{b(d_{0i}, z_{0i})}{u(d_{0i}, z_{0i})} + \log \frac{b(M,F,N)}{u(M,F,N)}$$

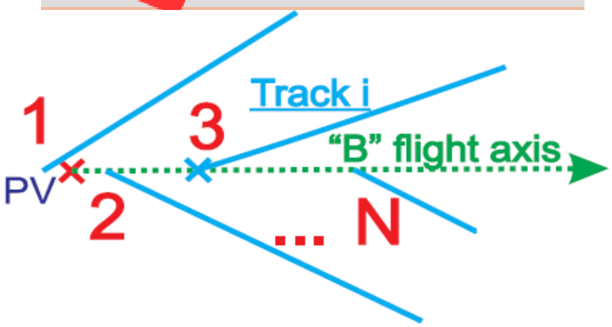






Topological reconstruction of B/D decay





Main idea: try to constrain all tracks from the B/D hadron decays to intersect the same flight axis (instead of the common vertex approach)

JetFitter: a new dedicated Kalman filter fitting:

$$\{x_v, y_v, z_v, \phi, \theta, dist_1, dist_2, dist_3, \dots, dist_N\}$$

Treatment of incomplete/1-prong topologies:

- on average 1.9 (1.7) OK tracks from D (B) decays

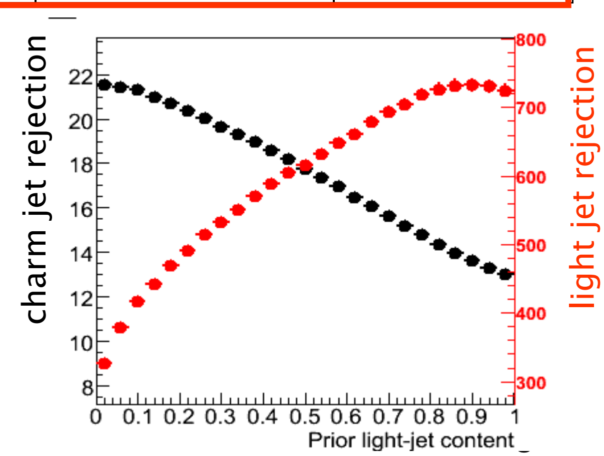
Algorithm	Topology	Track efficiency	Track purity
BTagVrtSec	1 inclusive B/D vertex	69 %	92 %
JetFitter	1 vertex	74 %	91 %
	1 vertex + 1 track	80 %	85 %
	2 vertices	85 %	89 %

Based on likelihood:

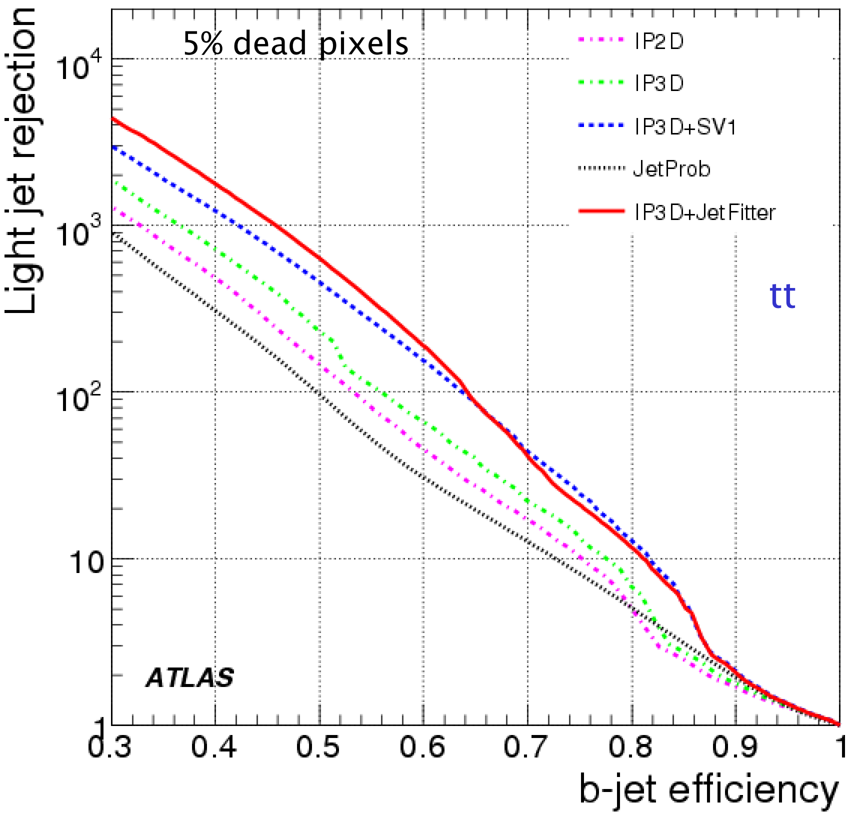
- categories for topology
- vertex mass
- energy fraction
- significances of $\{dist_i\}$

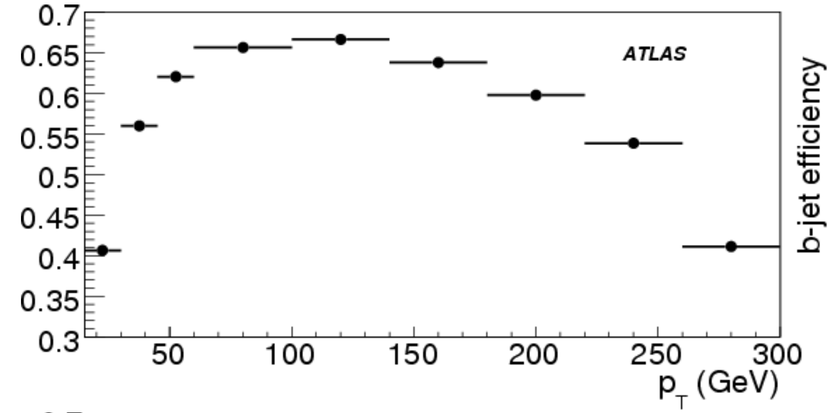
➔~20% improvement in light jet rej.

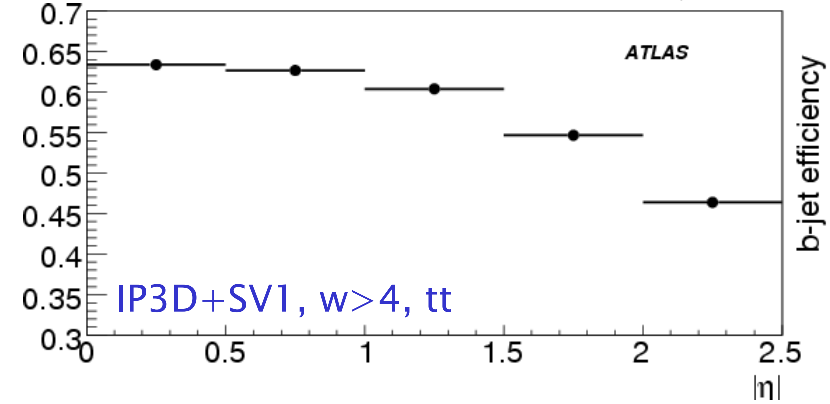
➔promising for b/c separation



Performance of b-tagging algorithms







- Track counting: R~30 @ 60%
- Soft muon: R~300 @ 10% (i.e. 80% w/ BR)
- Soft electron: R~100 @ 8%
- HLT: R~20 @ 60%
- Charm rejection: 5 to 7 @ 60%, up to 20 with JetFitter

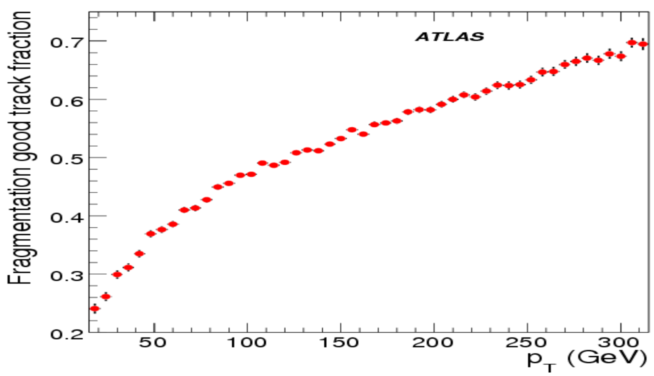
Factorization (≠ channels):
dependency on jet pT, η and env. (ΔR)

Degraded performance:

- low p_T: MS, secondaries
- high η: lever-arm (z_0),secondaries
- high p_T: next slide

Tagging high and very high-p_T jets

Challenges:
- fixed ΔR for track/jet assoc → dilution for jets of high p_T

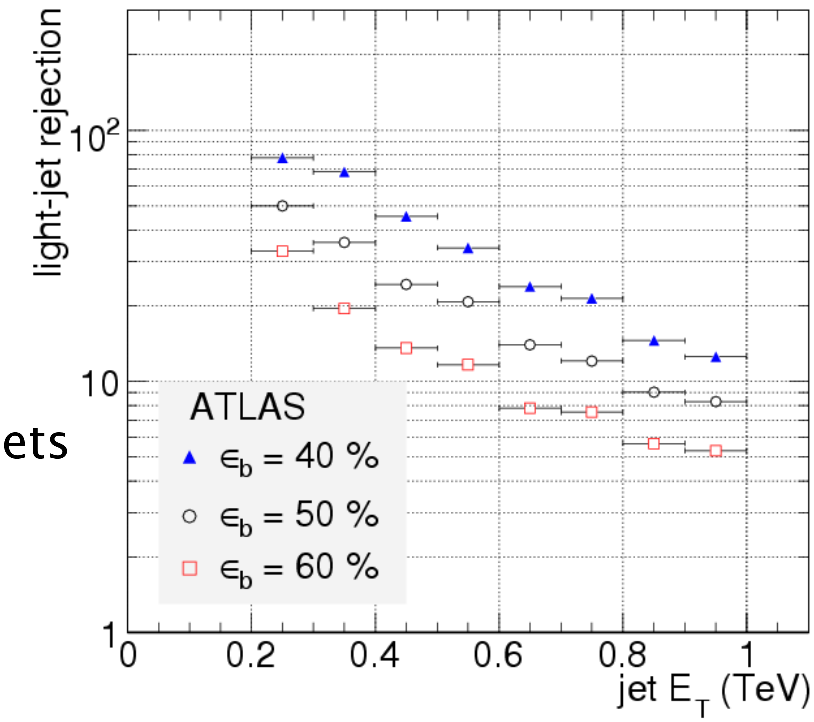


- gains (x2,x3) possible for non-isolated jets
- pattern-recognition issues
- 'late' B-decays

Fraction of jets (WH400):

	$R_B > 2.9$ cm	$R_B > 5.1$ cm
all E_T	9.0%	2.8%
$E_T > 100$ GeV	12.2%	3.9%
$E_T > 200$ GeV	21.1%	7.9%

Example: Z'(2 TeV) → bb, uu
after retuning of IP3D+SV1:
(x3 worse otherwise)



→ Require dedicated treatment for clustering, pattern-recognition

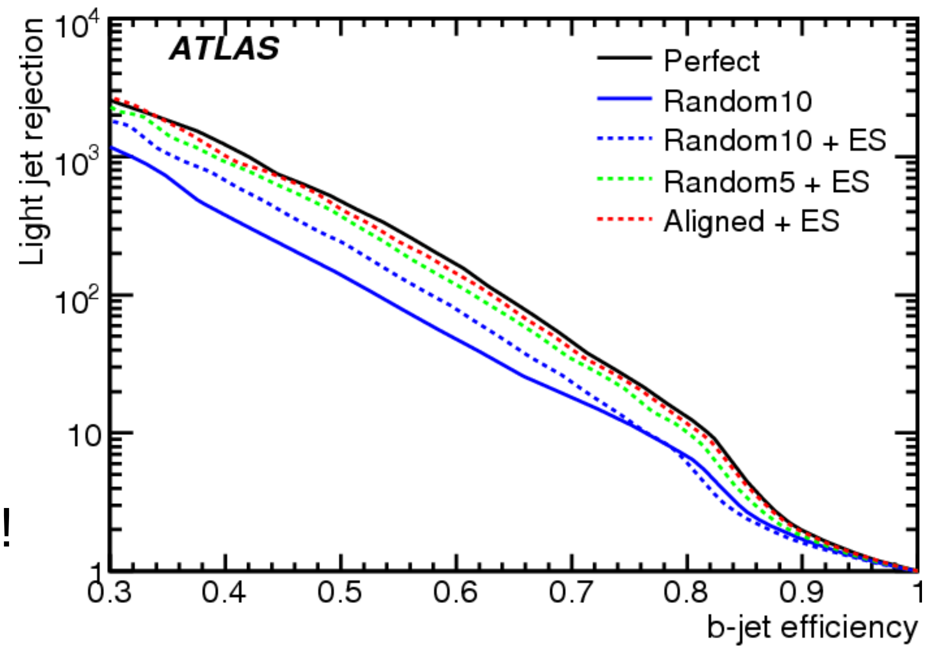
Impact of residual misalignments

Studies:

MC geometry includes misplacements as expected in detector (incl. surveys)
10–100 μm for modules, mm for elements, some systematic deformations (global shifts, clocking effects,…)

1. all displacements known in reco
2. 1 + additional pixel random misalignments introduced in reco only
3. actual ATLAS alignment procedure ! (cf. talk by R. Moles)

Impact on b-tagging:



Both 1 and 2 require a specific procedure to increase the hit errors → fully recovers tracking efficiency and fake levels

Alignment procedure suggests residual misalignments of ~ 3 μm in rφ (pixels)

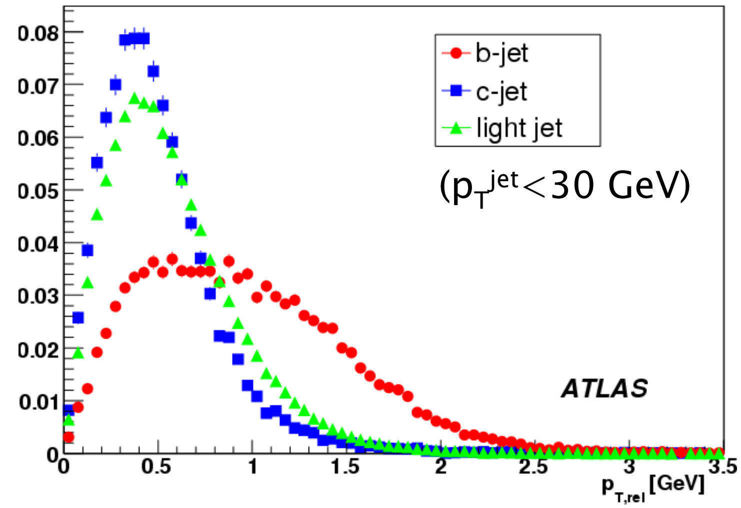
➔ after alignment, relative loss in rejection between 11–18%

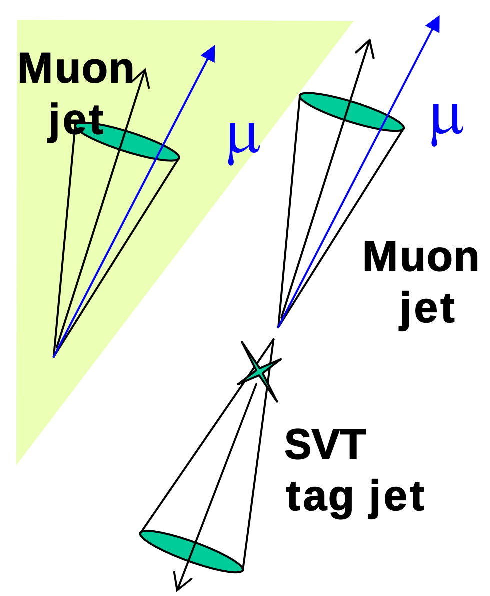
➔encouraging, even though probably not all systematic deformations were accounted for

Calibrating b-tagging efficiency with di-jet events

- Key-ingredient: soft muons
- Dedicated μ-jet trigger:
  - staged jet E_T thresholds
  - 1 Hz → 100k in 30h (1 pb^{-1}@10^{31} $cm^{-2}s^{-1}$)
- Method 1: p_T^{rel} templates
  - b & c templates from MC
  - fit in p_T, η bins
- Method 2: non-linear system (à la D0)
  - 2 samples
  - 2 different b-jet/light jet fractions
  - 2 non-correlated taggers
  - ➔ system can be solved analytically
- Both methods work well for p_T^{jet}<80 GeV
- Systematic uncertainties dominate for > 50 pb^{-1}

➔ Current precision on b-tag efficiency: 6%





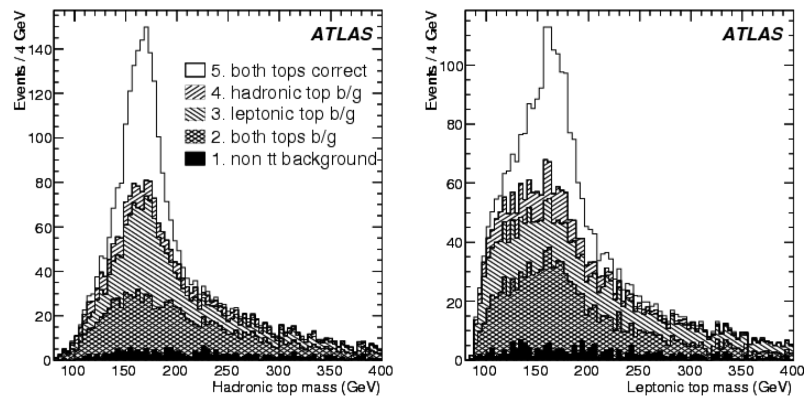
Calibrating b-tagging efficiency with top events

Tag counting method:

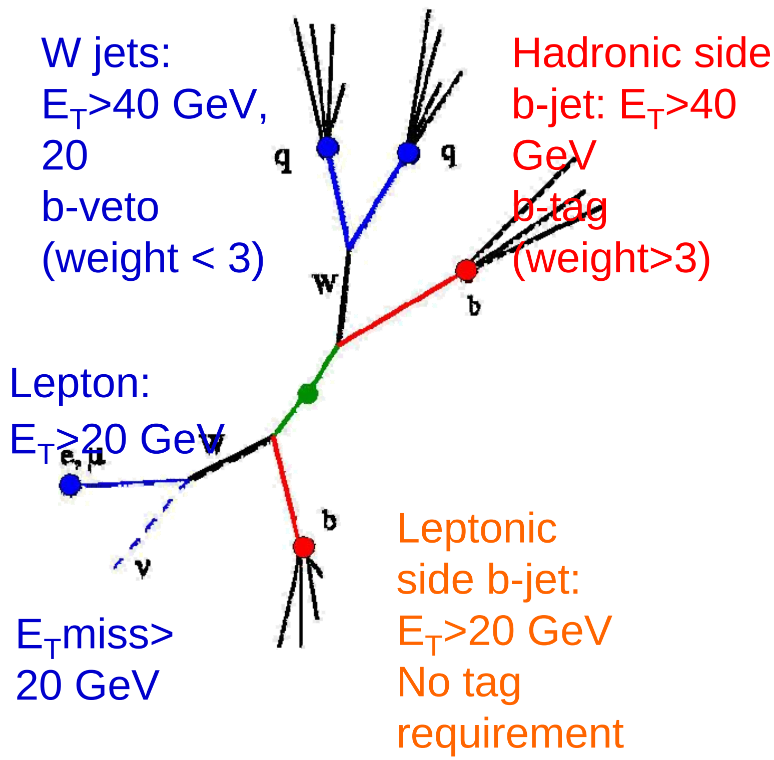
- count 1,2,3 b-tags → likelihood
- best accuracy on ε_b: for 100 pb^{-1} $\pm 2.7\%(stat) \pm 3.4\%(syst)$
- integrated efficiency only

Several selection methods:
- topological
- likelihood
- kinematic fit

→ signal purity 60-80%



Selecting unbiased b-jets (leptonic side)



b-tagging efficiency determination:

	Topo.	Like.	Kine.
Stat. (200 pb^{-1})	6.4%	4.4%	5.5%
Syst. error	3.4%	14.2%	6.2%

➔ can also be used to extract b's p.d.fs

Conclusion

- Expected b-tagging performance of ATLAS studied in details
- For simple taggers, light jet rejection of 30 (for ε_b=60%)
- Secondary vertex taggers can achieve a rejection of 150
- Most advanced taggers can reach a rejection of 50 for ε_b=70%, very important for multi b-jet signatures (more is needed: NN, BDT)
- Many critical aspects studied: misalignment, material, state of detector, Monte Carlo modelling, etc – More to do: misalignment, pile-up

- Methods to measure the b-tagging efficiency in data have been established, allowing a ~7% accuracy with 200 pb^{-1} in top events
- Work is on-going for the measurement of mistag rates

- New developments, e.g. b/c separation, can broaden the physics reach of ATLAS

- Commissioning of the detector progressing well, our work will now focus on the commissioning of the b-tagging to transform our expectations into reality !