b-tagging Performance in ATLAS

Grant Gorfine
(University of Wuppertal)

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

International Conference on Particle Physics

Istanbul, October 27-31, 2008

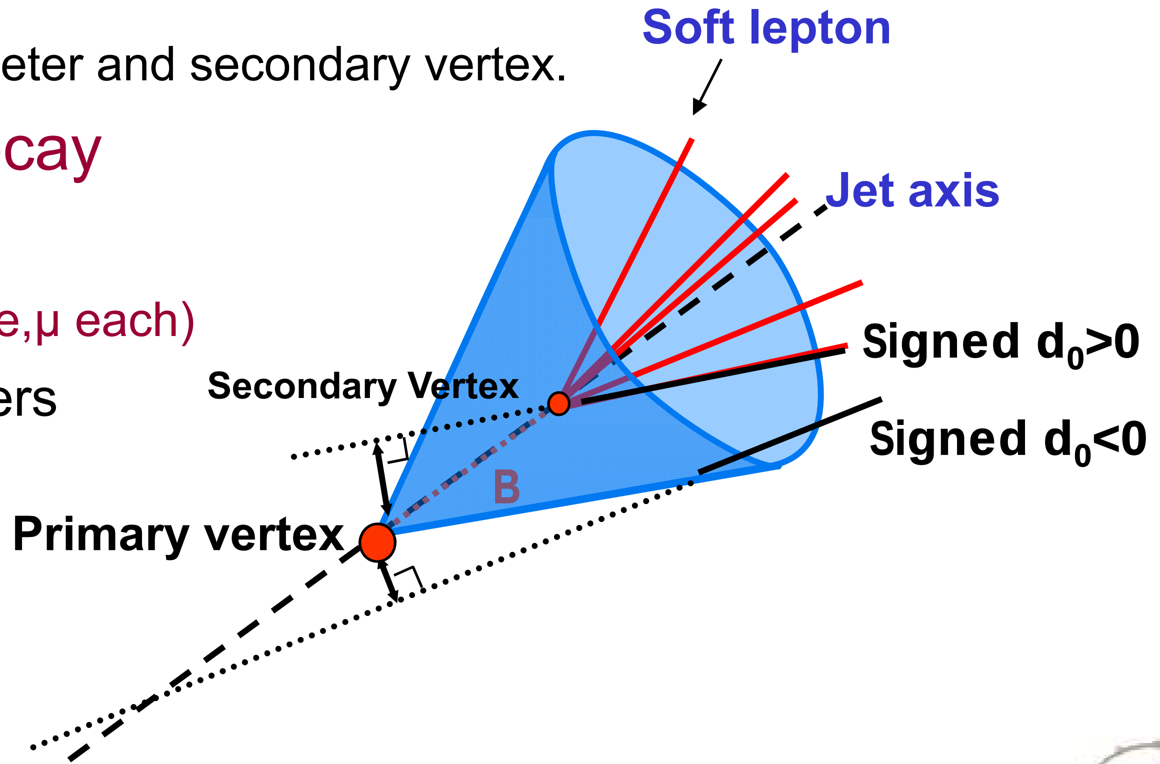
In Memoriam Engin Arik and Her Colleagues

Physics Motivation

- b-tagging vital for high p_T physics program at LHC.
- Top physics.
  - Large cross section, so moderate ε_b (>50%) OK.
    - Helps to reduce combinatoric and W+jets background
    - Important for highest precision measurements.
    - S/B 2x (4x) better requiring one (two) jets to be b-tagged
  - ttbar - Dominant background to many channels
- Higgs
  - H→bb largest decay mode for m_H < 135 GeV
  - ttH, (H→bb). Distinct signature, 4 b jets!
    - Comparatively low cross sections. Require high ε_b ~70%.
- New Physics
  - Higgs in SUSY (eg H^+→tb, bbH/A)
  - SUSY decay chains, Heavy Gauge bosons (eg Z′→bb ~TeV jets)

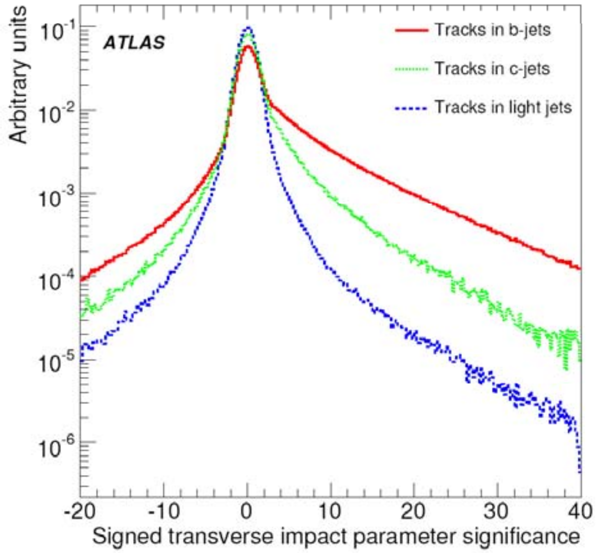
b-tagging overview

- b-jet properties
  - B hadron cτ~430 μm, ie travels order few mm with typical boost (eg ~3 mm @ p_T=50GeV, d_0~500 μm)
    - Spatial taggers
      - Impact parameter and secondary vertex.
  - Semi-leptonic decay
    (*b→lνX, b→c→lνX,*
    branching ratio ~20% e,μ each)
    - Soft lepton taggers
      - Use p_T, rel



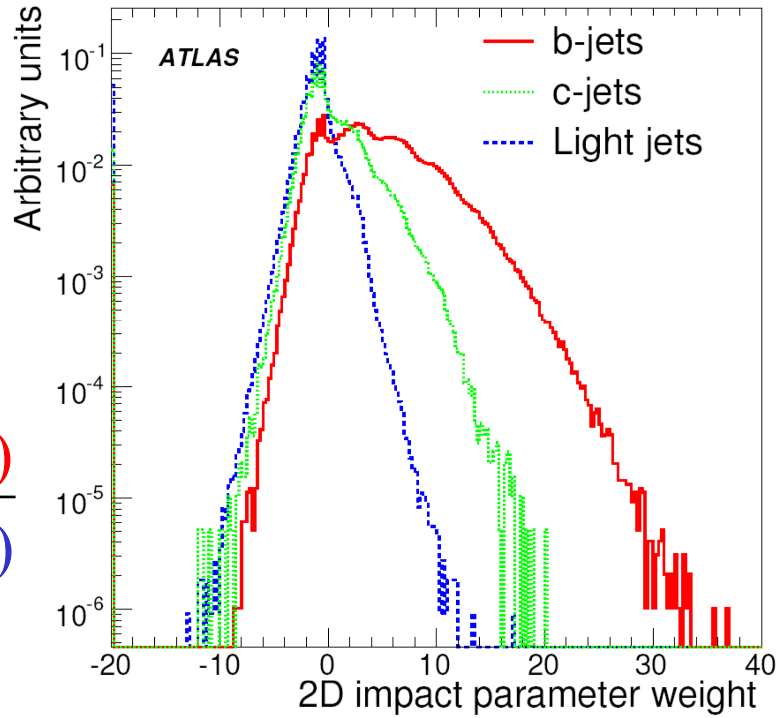
Impact parameter tagging

Impact parameter significance:
$S = d_0 / \sigma$ (Signed with jet axis)



Normalized distribution → pdfs for b-jets *b(S)* and light jets *u(S)*

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



Simplest tagger: Count tracks with large d_0 or d_0 / σ

JetProb (à la ALEPH). Signed d_0 compared with Resolution function of prompt tracks (obtained from data - negative signed d_0)

Rej ($1/\varepsilon_{light}$) ~30 @ ε_b = 60%

IP2D (d_0):

Rej ($1/\varepsilon_{light}$) = 46 @ ε_b = 60%

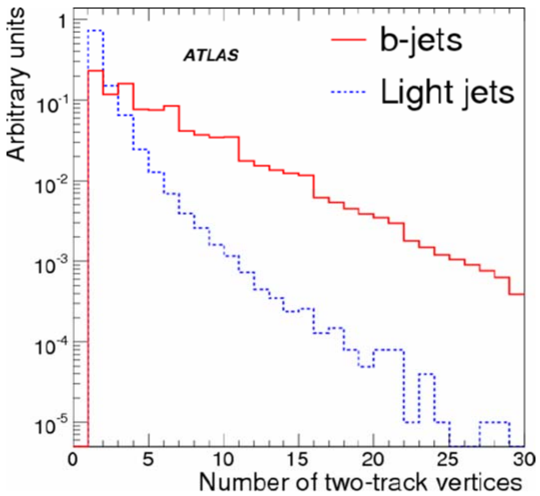
IP3D (d_0+z_0 , 2D pdfs):

Rej ($1/\varepsilon_{light}$) = 67 @ ε_b = 60%

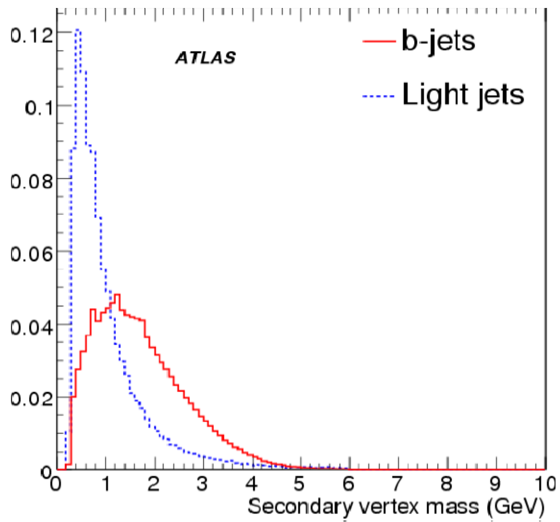
Secondary Vertex Tagger

- Inclusive secondary vertex reconstruction
  - Removal of V^0s, conversions, material interactions
- Use 3 variables from vertex:

Number of two track vertices

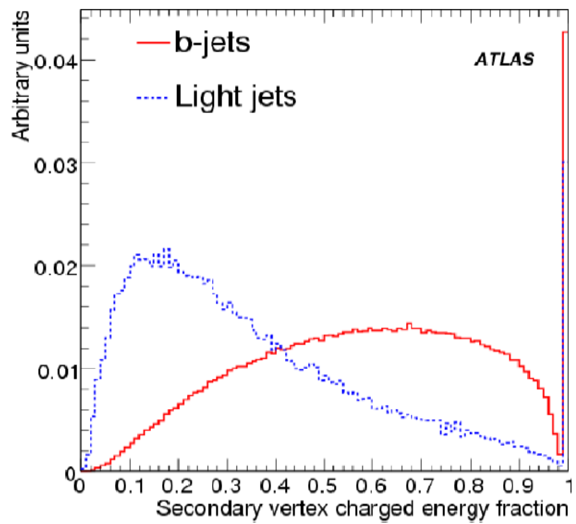


Vertex mass



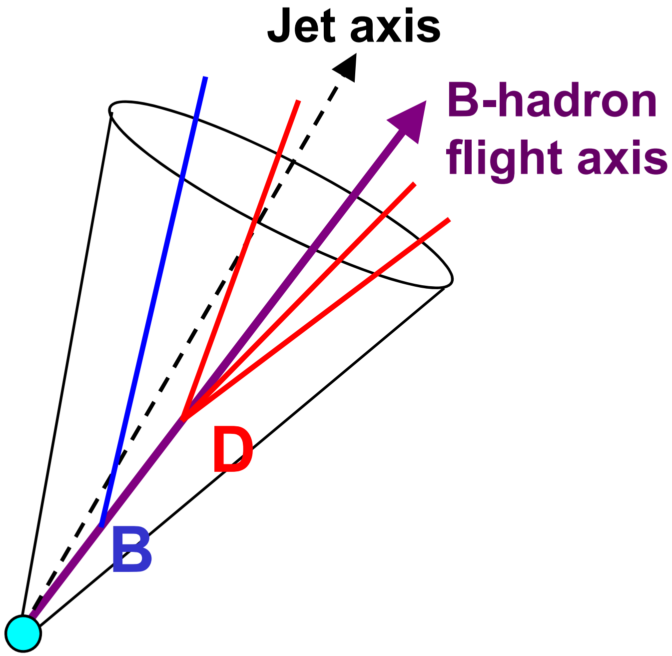
Energy fraction

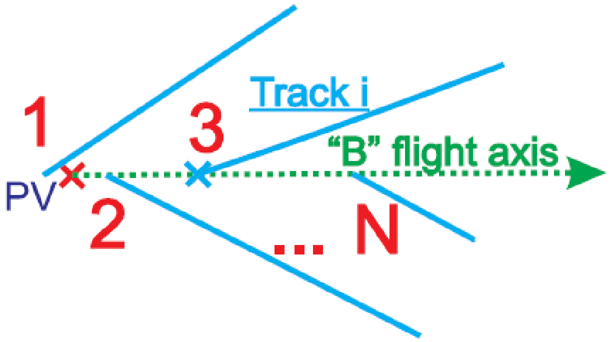
$$F = \frac{E(\text{tracks in vertex})}{E(\text{tracks in jet})}$$



- Likelihood ratio method (+ fold in efficiency for vertex)
- Combined with IP3D: "IP3D+SV1": **Rej 154 @ ε_b = 60%**

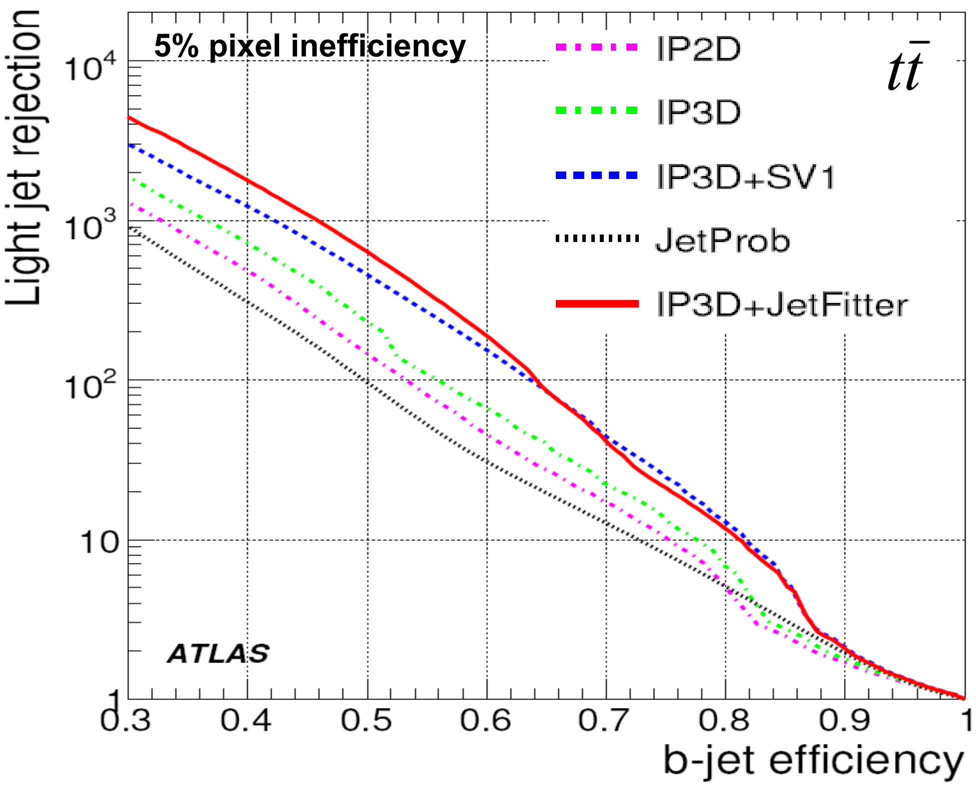
Topological Secondary Vertex ("JetFitter")

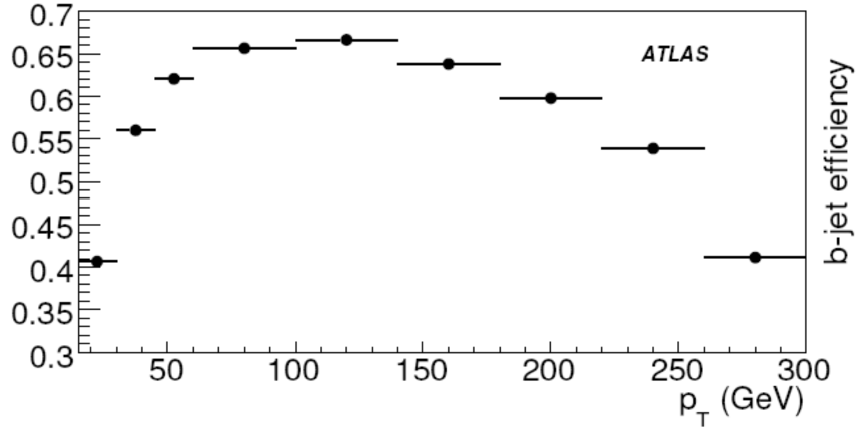


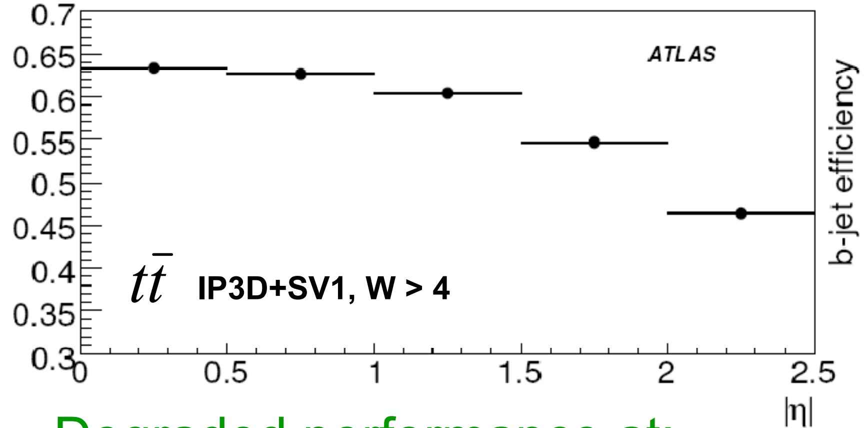


- Reconstruct topology of decay chain B→D.
- Makes use of property that D and B vertex approximately along B hadron flight direction
  - Single track vertices possible.
- Vertex variables used in Likelihood:
  - Vertex mass
  - Energy fraction
  - Flight length significance
- Split into different topology categories, eg:
  - 1 vertex
  - 1 vertex + single track
  - 2 vertices
  - …
- ~20% improvement in light jet rejection
- Promising for c/b separation

b-tagging performance







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

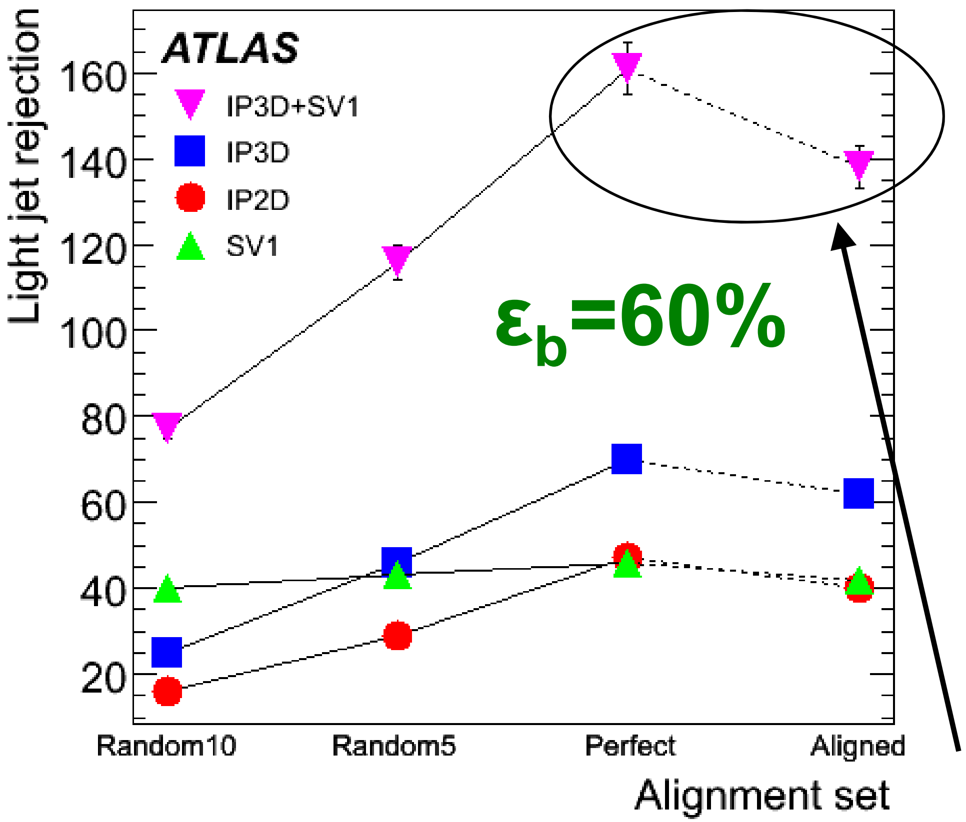
Degraded performance at:

Low p_T (material)

High p_T (high track density)

High η (large lever arm, material)

Effects of Misalignment



Random10: Shifts/rotations in Pixel layers, disks, and modules: 10 μm in $r\varphi$, 30 μm in η

Random5: ~ half as big

Perfect: Perfectly aligned detector

- **Aligned:** Misalignments put in simulation typical of expected assembly:
  - ~30-100μm shifts of modules, layer, disks
  - ~mm shifts of sub-system
  - clocking effects, rotations, etc.

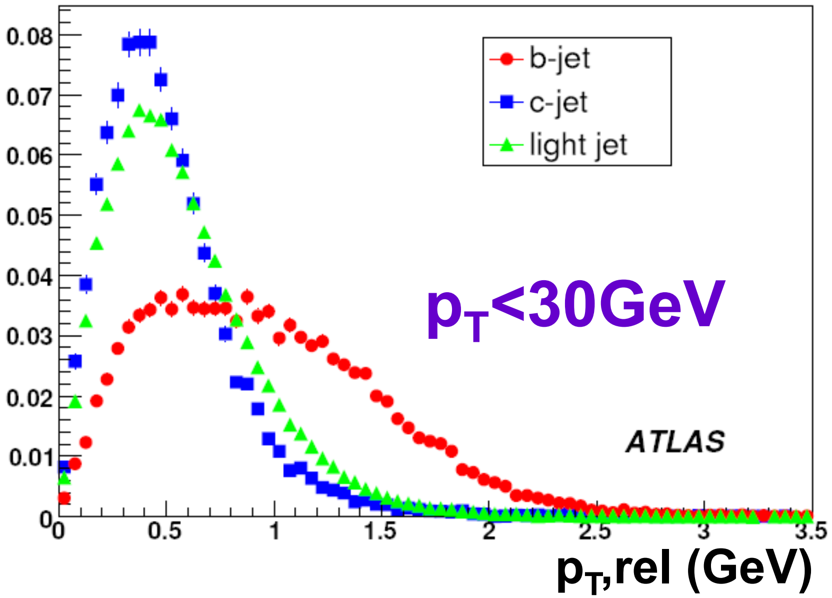
  Then aligned with actual ATLAS alignment procedures.

  → Includes any systematic deformations introduced by alignment procedure itself
- Cluster errors tuned.
- Moderate (~15%) degradation in b-tagging performance wrt to Perfect alignment
- Encouraging but not all systematic deformations (eg pixel stave bow, twists, …)

b-tagging calibration with di-jet events

- Main idea: use QCD di-jets events
  - b-jet enriched by Muon+Jet signature (dedicated Trigger)

Method 1: p_T,rel
- Templates for b,c (from MC) and light jets (from data), Likelihood fit to data, extract b/light fraction. Before/after tag → ε_b



Method 2: Non-linear system (à la DØ)
- Less MC dependence
- Two samples with different b fraction
- Two uncorrelated taggers (soft muon & spatial tagger)
- 8 equations, 8 unknowns (including ε_b) Solve analytically

- Both methods work well for jet p_T < 80 GeV
- Precision on b-tag efficiency: 6% @ $50pb^{-1}$. Dominated by systematics.

b-tagging calibration using ttbar events

Tag counting method

- Well defined flavour content in ttbar. Typically 2 b-jets
  - Count number of 1,2,3 b-tags
  - If exactly 2 b-jets and 100% reconstruction efficiency

$$N(1\text{ b-tag}) = 2 N_{tot} \varepsilon_b (1-\varepsilon_b)$$

$$N(2\text{ b-tags}) = N_{tot} \varepsilon_b^2$$

  - But need to take into account reconstruction efficiency and other heavy flavour content, charm and light jet rejection.
    - Use MC to get fraction of expected b, charm and light jets
  - Use likelihood to extract ε_b, ε_c and cross-section
- Method requires low statistics
  - For 100 pb^{-1} at $\varepsilon_b = 0.6$,
    - Semileptonic $\sigma_{stat} = 2.7\%$ syst 3.4%
    - Dilepton $\sigma_{stat} = 4.2\%$ syst 3.5%
- Efficiency averaged over p_T spectrum of jets

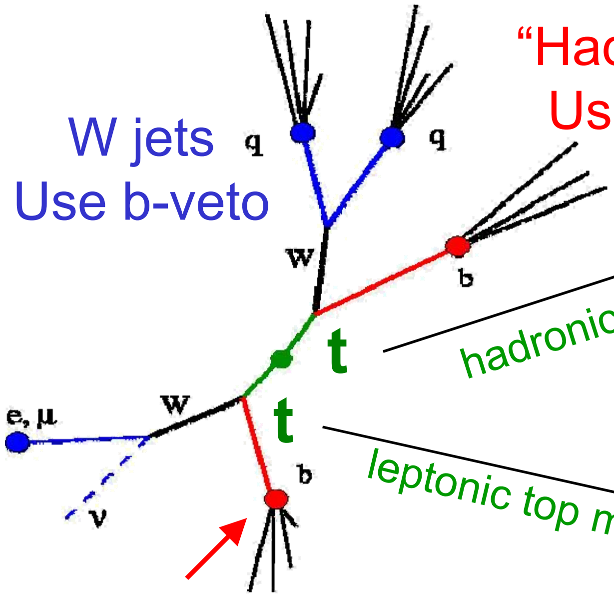
b-tagging calibration using ttbar events

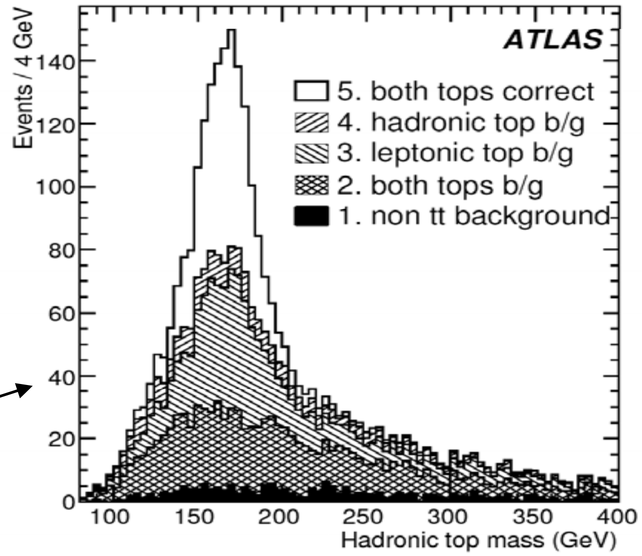
Topological Method:

Select semi-leptonic ttbar events and identify b-jet sample:

W jets
Use b-veto

"Hadronic b"
Use b-tag



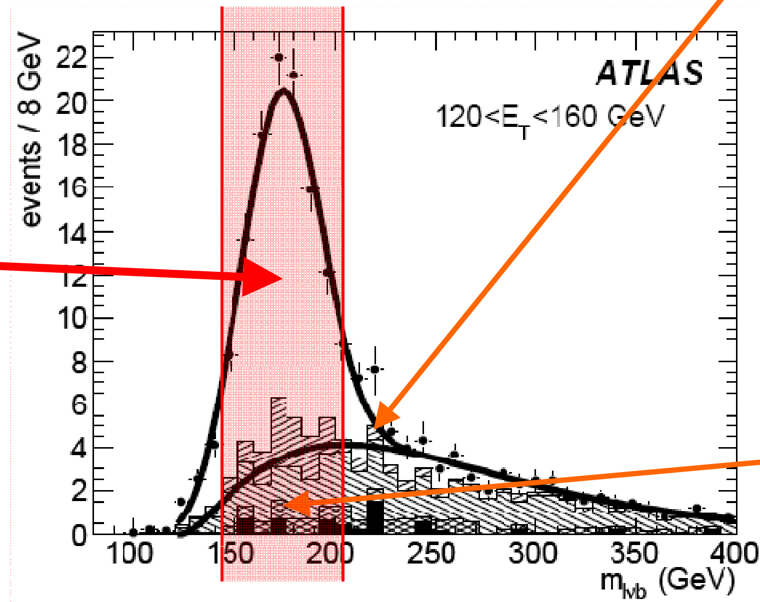


Events in hadronic mass sideband + b-veto on leptonic side
→provides good background shape estimate.
→Used in **fit (signal+bkg)**

"Leptonic b" as b-jet sample

No b-tagging - Unbiased

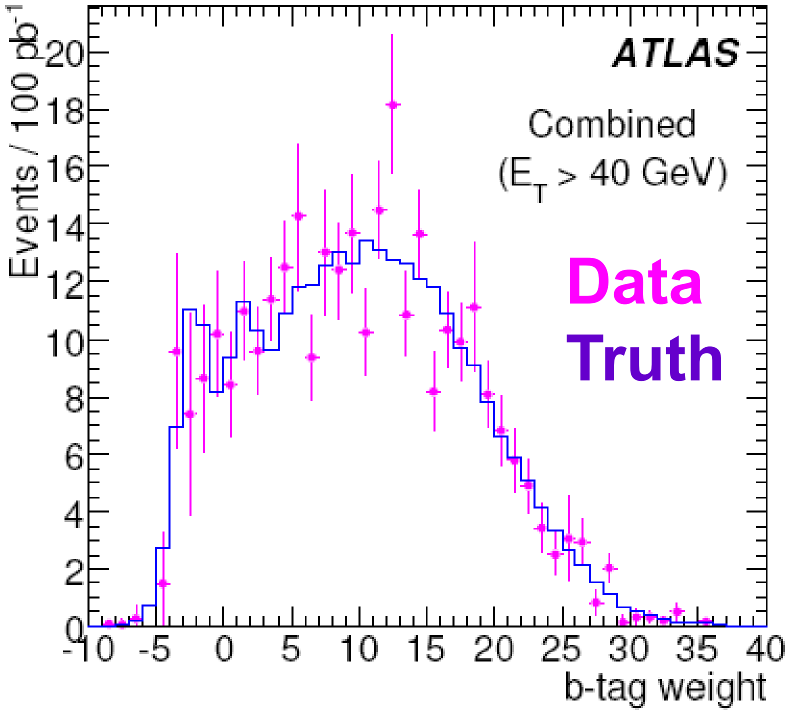
60-80% purity (p_T dependent), requires background subtraction



Leptonic sideband sample describes well flavour mixture in background

b-tagging calibration using ttbar events

- "b-jet sample" = jets in signal region – background: obtain any distribution on statistical basis.
- eg. "IP3D+SV1" b-jet weight → b-jet efficiency.
- For E_T>40GeV, error on ε_b: ±6%(stat) ± 3%(syst.) @ 200pb^{-1}
- ε_b as function of p_T.
- Also other distributions. eg. pdfs



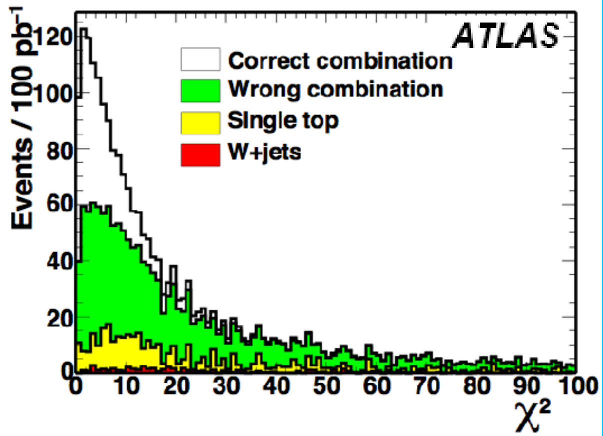
Two other similar methods:

Likelihood method

- Calculate discriminant based on templates of correct/incorrect assignments

Kinematic fit method

- Use χ^2 from kinematic fit
- Background subtraction: High χ^2 region analogous to mass side band in topological method



Conclusion

- Suite of b-tagging algorithms
  - Simpler taggers for commissioning
    - Track counting, JetProb: light jet rejection 30 @ ε_b 60%,
  - More sophisticated taggers for ultimate performance
    - Combined impact parameter and secondary vertex taggers
      - Light jet rejection 150 @ ε_b 60%
- Many studies to prepare for reality
  - Impact of misalignments ~15% degradation with real alignment procedures.
- Techniques developed to obtain b-tagging efficiency from data.
  - 6% accuracy achievable in first few 100 pb^{-1}
  - Can obtain any b-jet distribution
- Looking forward to LHC collisions and tagging b's!