B-Tagging in ATLAS: expected performance and its calibration in data

Giacinto Piacquadio (University of Freiburg) on behalf of the ATLAS Collaboration

Charged Higgs 2008 Conference
(Uppsala: 15-19 September 2008)
Outline

- Why b-Tagging?
- Review of “spatial” algorithms
- Soft-lepton Taggers
- Misalignment studies
- Calibration on data
- Conclusion
- Outlook
Why b-Tagging?

Vital ingredient for the high pT physics program at the LHC, e.g.:

- to obtain a very pure top quark sample
- for Higgs analyses: e.g. \( ttH \rightarrow ttbb \) (4 b-jets to tag!)
- SUSY Higgses, like in charged Higgs (2-4 b-Jets) or bbH/A or inclusive searches
- Many exotic scenarios

Two available signatures:

- spatial (lifetime of b-hadrons: \( c\tau \sim 450\mu m \rightarrow \beta\gamma\tau \sim 3-15 \text{ mm} \))
  1. Impact Parameter based algorithms
     relies on the (in)compatibility of single tracks with the primary vertex
  2. Secondary vertex based algorithms
     explicit determination of the weak B hadron decay vertex → use its production and decay properties (mass, fragmentation function, track multiplicity)
- lepton based
  3. Lepton-ID based algorithms
     identify muon or electron from semileptonic B or \( B \rightarrow D \) decay (e and \( \mu \) ~20% each)
Jets

- Jets typically from calorimeter information are considered as b-jet candidates (need direction!)

Tracks

- Tracks are assigned to the jet if $\Delta R(\text{Track},\text{Jet})<0.4$
- Tracks must satisfy quality criteria ($p_T>1\text{GeV}$, loose IP cuts, b-Layer hit requirement,...)

Impact parameter resolution essential for “spatial” tagging

- Resolution of the (innermost) 3 barrel pixel layers is around $10\ \mu\text{m}$ in $r\phi$ and $115\ \mu\text{m}$ in $z$.
- Transverse resolution of tracks goes from $\sim100\ \mu\text{m}$ ($p_T=1\ \text{GeV}$) to $\sim10\ \mu\text{m}$ ($p_T=100\ \text{GeV}$)

Displacement is computed wrt. Primary Vertex (PV)

- Transverse plane: PV well constrained by beam spot ($\sim15\ \mu\text{m}$)
- PV reconstruction essential to get PV $z$ coordinate ($\sigma \sim 50\mu\text{m}$)
**Impact Parameter based b-Tagging algorithm**

Determine Lifetime Sign of Impact Parameters: in 3d

\[
\text{sign}(IP) = \text{sign}\left( (\hat{p}_\text{JE}\times\hat{p}_\text{TR}) \cdot (\hat{p}_\text{TR} \times (\hat{r}_\text{PV} - \hat{r}_\text{PCA})) \right)
\]

Consider signed IP significances

\[
S_i = \frac{d^*}{\sigma(d^*)}
\]

Define PDFs:

\[
P_b(S_i) \quad \text{and} \quad P_\text{light}(S_i)
\]

Define Jet weight:

\[
W_{JET}^{IP} = \sum_{\text{Tracks}} \log \left( \frac{P_b(S_i)}{P_\text{light}(S_i)} \right)
\]

- **IP2D**: only transverse IPs
- **IP3D**: also longitudinal IPs

(2-dim PDFs)

- Likelihood ratio formalism adopted for both IP based algorithms
- Simpler algorithms based only on background hypothesis (JetProb, à la Aleph) or on counting high IP tracks also available → important for commissioning!
Inclusive secondary vertex reconstruction in Jet (I)

Find all displaced two-track vertices → remove $V^0$ decays, material interactions

Jet Axis

- Fit surviving tracks from two-track vertices into one inclusive geometrical vertex
- Remove iteratively most incompatible ones

ATLAS

$K_s^0$

$\Lambda^0$

Beam-pipe, pixel layers
Define templates based on:

- Invariant mass at vertex
- Energy of charged particles at vertex
- Energy of charged particles in jet
- Number of good two-track vertices

\[ P(x_1, x_2, x_3) = \]

Add probability to find vertex \( \epsilon \):

\[ PDF = \begin{cases} \epsilon \cdot P(x_1, x_2, x_3) & \text{[vertex found]} \\ 1 - \epsilon & \text{[no vertex]} \end{cases} \]

Define Jet weight based on likelihood ratio:

\[ W^{SV}_{JET} = \log \left( \frac{PDF_b}{PDF_{light}} \right) \]

Combine IP and SV based weights:

\[ W^C_{JET} = W^{IP}_{JET} + W^{SV}_{JET} \]
Topological reconstruction of the $PV \rightarrow B \rightarrow D$ decay chain (I)

The “JetFitter” algorithm tries to disentangle the weakly decaying B and D vertices.

- b and c vertices approx. on same line of flight
- intersect b-hadron flight direction with tracks

Principle used by SLD in “ghost track” algorithm

[SLAC-PUB-8225 (1999)]

JetFitter is based on an original extension of the Kalman Filter formalism commonly used for vertexing

[J. Phys.: Conf. Ser. 119 032032]

Finding strategy

- Initialization of:
  1) Primary Vertex
  2) “B” flight axis (from calorimeter jet direction)
- First fit under the hypothesis that each track represents a single vertex along the “B” flight axis
- optimal $(\phi_{AXIS}, \theta_{AXIS}, d_1, d_2, ..., d_N)$
Topological reconstruction of the $PV \rightarrow B \rightarrow D$ decay chain (II)

For all combinations of two vertices (including the Primary Vertex) the probability of having a common vertex is evaluated.

1. Merge pair of vertices with highest probability
2. Perform a new “full fit” and repeat from 1
3. Stop when no pair of vertices needs to be merged anymore ($P_{xy} < \text{cut value}$)

Likelihood function:

$$L_{b,l,c}^x(x) = \sum_{\text{cat}} \text{coeff}(\text{cat}) \cdot \text{PDF}_{\text{cat}}(\text{mass}) \cdot \text{PDF}_{\text{cat}}(\text{energyFraction}) \cdot \text{PDF}_{\text{cat}} \left( \frac{d}{\sigma(d)} \right)$$

Variables used for B-Tagging:

- Decay topology (number of vertices, tracks at vertices, additional single tracks on flight axis)
- Invariant mass of charged particles of decay chain
- Fractional charged tracks energy
- Decay length significance $d/\sigma(d)$

Jet weight is again defined according to likelihood ratio. Analogously combined with IP3D.
Performance of spatial algorithms

- B-Tagging performance tested on a sample of >1M of fully simulated pp → tt and pp→ttjj events.
- Tagging efficiency: \( \epsilon_q = \frac{\text{Number of jets of flavour } q \text{ tagged as } b}{\text{Number of jets of flavour } q} \)
- Rejection: \( r_u = 1/\epsilon_u, \quad r_c = 1/\epsilon_c \)

### Light jet rejections

<table>
<thead>
<tr>
<th>( \epsilon(b\text{-jet}) )</th>
<th>JetProb</th>
<th>IP2D</th>
<th>IP3D</th>
<th>IP3D+SV1</th>
<th>IP3D+JetFitter</th>
</tr>
</thead>
<tbody>
<tr>
<td>50%</td>
<td>83±1</td>
<td>116±2</td>
<td>190±3</td>
<td>458±13</td>
<td>555±17</td>
</tr>
<tr>
<td>60%</td>
<td>30±0</td>
<td>42±0</td>
<td>59±1</td>
<td>117±2</td>
<td>134±2</td>
</tr>
</tbody>
</table>

### Charm jet rejections

<table>
<thead>
<tr>
<th>( \epsilon(b\text{-jet}) )</th>
<th>JetProb</th>
<th>IP2D</th>
<th>IP3D</th>
<th>IP3D+SV1</th>
<th>IP3D+JetFitter</th>
</tr>
</thead>
<tbody>
<tr>
<td>50%</td>
<td>8.4±0</td>
<td>9.5±0</td>
<td>10.6±0</td>
<td>12.4±0.1</td>
<td>12.3±0.1</td>
</tr>
<tr>
<td>60%</td>
<td>5.1±0</td>
<td>5.8±0</td>
<td>6.5±0</td>
<td>7.4±0</td>
<td>7.4±0</td>
</tr>
</tbody>
</table>

- Ideal geometry and 5 % pixel inefficiency (3 % pixel inefficiency expected at the end of 2008) assumed in the simulation
- No specific charm rejection implemented
Soft Lepton based Tagging algorithms

- Efficiency is a-priori limited by semi-leptonic branching ratios:
  - $\text{BR}(b \rightarrow l X) \sim 11\%$, $\text{BR}(b \rightarrow c \rightarrow l X) \sim 10\%$ ($l=e,\mu$)
- Correlation with “spatial” algorithms is very low:
  - perfect for obtaining b-Tagging efficiency from data
- Both algorithms make use of the relative pT ($p_T^{\text{rel}}$) of the lepton with respect to the Jet axis

Soft Muon Tagging algorithm
- Background given by fake muons (e.g. punch-throughs) and from decay of light hadrons
- IP significance of lepton not used (avoid correlations with spatial)
- Rejection: $\sim 300$ for 10% b-Tagging efficiency

Soft Electron Tagging algorithm
- Low pT Electron-ID in dense Jet environment very challenging: use dedicated likelihood discriminator which at 80% electron efficiency gives:
  - rejection of $\sim 200$ against charged pions
  - rejection of $\sim 2-3$ against conv./$\pi^0$ Dalitz decays
- IP significance of lepton used in addition
- Rejection: $\sim 100$ for 7% b-Tagging efficiency
Effect of detector misalignment

- Detector misalignment affects tracking efficiency and IP resolutions, thus B-Tagging.
- Dedicated study:
  - Simulation with randomly misaligned detector (~10-100 μm, including some global deformations)

Reconstruction with 2 alignment sets:
- **Perfectly aligned**: equivalent to no misalignment
- **Aligned**: residual misalignment after realistic track based alignment procedure

- Error scaling procedure on the track hits used to deal with residual misalignment
- On real data, after alignment, a degradation of less than 25 % in light-Jet rejection with respect to the ideal case seems feasible.

60% b-Tagging Efficiency
IP and SV Tagging algorithms
Calibration on data

The b-Tagging efficiency can be extracted using:

- Two uncorrelated b-Tagging algorithms ("spatial" and Soft Muon)
  - use QCD dijet Events
  - b-flavour enriched by Muon+Jet signature (dedicated Trigger)

- Extraction of light-Jet mistagging rates still under study...

Kinematic reconstruction of $pp \rightarrow t\bar{t}$ events

- Select a pure top sample
- Subtract the residual background

$\sim 100k$ events in 30 hours at $10^{31}$ cm$^{-2}$s$^{-1}$!
Calibration on data using QCD dijet events (System 8 Method)

- Based on:
  - 2 samples with different flavour composition
    - 1. Jet+Muon [n]
    - 2. Jet+Muon + additional back to back b-Tagged Jet [p]
  - 2 uncorrelated Taggers (muon, "spatial")
    → 4 combinations [no tag], [μ tag], [“spatial tag”], [both tags] to be applied on 2 samples
  
  → 8 equations with 8 unknowns

- Solve equation: obtain flavour composition of samples and b-tagging efficiency

\[
\begin{align*}
  n &= n_b + n_{cl} \\
  p &= p_b + p_{cl} \\
  n_\mu &= \varepsilon^\mu n_b + r^\mu n_{cl} \\
  p_\mu &= \varepsilon^\mu p_b + r^\mu p_{cl} \\
  n_{Tr} &= \varepsilon^{Tr} n_b + r^{Tr} n_{cl} \\
  p_{Tr} &= \beta \varepsilon^{Tr} p_b + \alpha r^{Tr} p_{cl} \\
  n_{all} &= k_b \varepsilon^\mu \varepsilon^{Tr} n_b + k_{cl} r^\mu r^{Tr} n_{cl} \\
  p_{all} &= k_b \beta \varepsilon^\mu \varepsilon^{Tr} p_b + k_{cl} \alpha r^\mu r^{Tr} p_{cl}
\end{align*}
\]

- Method is dominated by systematic uncertainties with more than 50 pb\(^{-1}\) of data

- A pT and \(\eta\) dependent measurement of the b-Tagging efficiency with a precision of 6 % up to 150 GeV Jet pT seems feasible.
Calibration on data using ttbar events

Topological selection (I)

- Basic preselection of one hadronic top and one leptonic top
- Use b-Tagging on the hadronic side
  - leptonic top left unbiased
- Solve equation for missing $p_z$ of the neutrino, take smallest solution
- More combinations? Take the one with largest $\sum p_T$ of the two tops

W jets:
- $E_T > 40 \text{ GeV}$
- $E_T > 20 \text{ GeV}$
- $+b$-veto

Lepton: $E_T > 20 \text{ GeV}$

Hadronic side b-jet:
- $E_T > 40 \text{ GeV}$
- b-tag

E$_{T\text{miss}} > 20 \text{ GeV}$

Leptonic side b-jet: $E_T > 20 \text{ GeV}$
- No tag requirement
Calibration on data using ttbar events

Topological selection (II)

- Reachable b-Jet purity: 54-86 % depending on pT
- Shape for background obtained from a signal depleted control sample
- Simultaneous fit on selection + control sample
- b-Tagging weight distribution from signal region after background subtraction (gives efficiency as a function of discriminator cut)

With 200pb\(^{-1}\) and \(E_T > 40\)GeV get a relative precision on the b-Tag. efficiency of \(\pm 7.7\%(\text{stat}) \pm 3.2\%(\text{syst.})\)

Permits to get any distribution/property of the b-Jets on a statistical basis! (~1 fb\(^{-1}\) needed)
Outlook

Many algorithmic improvements to increase B-Tagging performance still on the way.

E.g. one interesting development:

- Use JetFitter's different decay chain topologies to improve the charm-Quark rejection

Possible decay chain topologies:

- Neural Network to discriminate b-Jets against light and charm-Jets
- A consistent increase in charm-Quark rejection is possible at the cost of a lower light-Quark rejection
- The method is being applied to the recent analysis of:
  \[ pp \rightarrow W(\rightarrow \mu \nu)H(\rightarrow bb) \text{ with } p_T(H)>200\text{GeV} \]

where the bottom/charm-Jets from the top are a severe background for the Higgs→bb
Conclusions

- The LHC has started!

- Performance achievable by b-Tagging algorithms in ATLAS at 60 % b-Tagging efficiency, in order of expected commissioning:
  - JetProb → light Jet rejection of ~30 (only input: resolution function for prompt tracks in data)
  - IP3D → light Jet rejection of ~60
  - Sec Vtx based algorithms → light-Jet rejection of ~120-140

- The effect of residual misalignment is expected to degrade these rejections by less than 25 %

- A 8-15 % discrepancy in the detector material description in the Monte Carlo simulation would impact these rejections by ~10%

- Methods established to measure the b-Tagging efficiency on data to 5 % accuracy with 100 pb$^{-1}$ of data

- Mistagging rate determination under study: 10 % precision expected from Tevatron experience.

- Many improvements still on the way!