BTAGGING AT ATLAS

G. Watts (U Washington – Seattle/CPPM – Marseille)
Anatomy of a B-Hadron Decay

A B is Long Lived
(decays via weak force)

- Top, Higgs contain b-quark jets
- Most backgrounds do not
- Jets look like any light quark jet
- Other than contain a B meson
- Has finite life time
- Travels some distance from the vertex before decaying
  ~ 1mm
- With charm cascade decay, about 4.2 charged tracks

Several algorithms under development

Impact Parameter (d)
Decay Length (L_{xy})

Impact Parameter Resolution
Decay Length Resolution

d/\sigma(d)
L_{xy}/\sigma(L_{xy})
Tagging Algorithms

- **Impact Parameter**
  - Displaced vertices generate displaced tracks
    - Large impact parameters
  - Rely on impact parameter significance ($IP/\sigma_{IP}$).
  - Likelihood fit or straight out cut

- **Secondary Vertex**
  - Attempt to reconstruct the bottom-quark’s decay vertex.
  - Cuts on track quality, $L_{xy}$, $L_{xy}/\sigma_{L_{xy}}$, etc.

- **Soft Lepton Tagger**
  - Look for muon/electron from $B$ or $B \rightarrow D$ cascade decay
  - Rate is low. Very difficult to do electron tagger
Combined Algorithms

- Use the results of the previously mentioned algorithms
  - Combined with other input variables
  - Use MVA techniques
  - Can see gains of 20% or so
  - Calibration will be very difficult!!
Variables We Care About To Tag

- **Track Parameters**
  - In particular the impact parameter significance!
  - 2D and 3D parameters
- **SV Parameters**
  - Decay length in particular
- **Soft electron, muon**
- **Sum of track pT’s**
- **Jet specific variables**
  - Total energy, $E_T$, etc.
- **Count of 2-track vertices in jet**
- **# of tracks in the jet**
- **Vertex Mass**

These are all per-jet variables! Are there Wb/Wbb variables that could be used to enhance this?

G. Watts (UW/CPPM)
Rejection

Efficiency

Exactly what you would expect:

\[
\frac{\text{# of } b\text{-jets tagged}}{\text{# of } b\text{-jets}}
\]

Rejection

\[
\frac{1}{\frac{\text{# of non-}\text{b/c -jets tagged}}{\text{# of non-}\text{b/c-jets}}}
\]

Easy to get fooled by large changes in the rejection – which are small changes in the actual tag rate!
Just like any other section cut there is a turn-on.
We are less sensitive to low $p_T$ b’s than higher $p_T$ b’s.
We are less sensitive to very high $p_T$ b’s!

**Calibration**

Tells us what this effect actually is
And allows us to “deconvolute” it from our actual trigger efficiencies
Why?

Multiple Scattering!

Collimated jets – much denser core, so increased tracking errors, more fragmentation tracks... and... out of range!!!

<table>
<thead>
<tr>
<th></th>
<th>$R_B &gt; 2.9$ cm</th>
<th>$R_B &gt; 5.1$ cm</th>
</tr>
</thead>
<tbody>
<tr>
<td>all $E_T$</td>
<td>9.0%</td>
<td>2.8%</td>
</tr>
<tr>
<td>$E_T &gt; 100$ GeV</td>
<td>12.2%</td>
<td>3.9%</td>
</tr>
<tr>
<td>$E_T &gt; 200$ GeV</td>
<td>21.1%</td>
<td>7.9%</td>
</tr>
</tbody>
</table>
Side Note About Semileptonic Decays

If you miss the jet you get an isolated lepton and missing $E_T$.

Theorist should ignore: these are detector effects and we have a very hard time (or impossible time) simulating them.
Gluon Decay’s

When we calibrate we do it be jet $p_T$ and eta

Implicit in this is the assumption that the B hadron $p_T$ is strongly correlated with the jet $p_T$!

Correlation clearly does not work when we have gluon splitting in the same jet!

On average a B hadron from gluon splitting in a jet will have about half the $p_T$ of a b from the same $p_T$ jet if it was from the hard scatter.

Tag Rate is Lower
There are two of them
Tag Rate is higher

Tevatron: about 15% lower than expected

So we are going to mis-calibrate those unless we account for them!

G. Watts (UW/CPPM)
Based on the ghost-track algorithm first implemented in SLD

Regular fit hypothesizes a single vertex in the jet.

Try two vertices, with a ghost track from the B to the D vertex.

Currently the best performing single tagger in ATLAS’ arsenal.

A similar technique could be used for gluon splitting

ATLAS (and others) need to know how important this is.

G. Watts (UW/CPPM)
Schedule – What will be ready when?

- The Tagging Algorithms
  - Most are fully coded and working well on MC now
  - Some require likelihood calibrations to work at their best (derived from data)

- Calibration
  - This is the weakest part – not much is production quality yet

- Infrastructure
  - Good at the 80% level
Schedule – What will be ready when?

- The Tagging Algorithms
  - Most are fully coded and working well on MC now
  - Some require likelihood calibrations to work at their best (derived from data)
- Calibration
  - This is the weakest part, not much is production quality yet
- Infrastructure
  - Good at the 80% level

Good To Go for First Data

G. Watts (UW/CPPM)

Should be ready – but perhaps best not to make it a show-stopper!
Early Data

- Performance likely much worse than what I’ve shown.

- But… we can re-run on early data and improve the performance as we improve our understanding of the detector

  - If b-tagging misses the first results, it will be there for the second results — even if no further data is taken.
You can find out almost everything we know in ATLAS from the CSC exercise public document

http://arxiv.org/abs/0901.0512