B-Tagging at LHC

Andrea Rizzi (CMS), SNS and INFN Pisa

MCWS Frascati, 28/02/06

Atlas contact: Tommaso Lari
Outline

- Introduction to b-tagging @ LHC
- Efficiency and rejection
  - vs Energy and Eta
  - Low/High luminosity
- Trigger
- Detector effects
- Calibration
b-Tagging introduction

Goal: identify jets originating from b-quarks

• lifetime based algorithms

Those algorithms exploit the fact that B hadrons decays far (~mm) from the primary interaction vertex.

• soft lepton algorithm

Those algorithms are based on the presence of a “soft lepton” (muon/electron) in the jet.

b-tagging is crucial in many channels (top studies, ttH, susy, ... ) to tag signal or reject bkgrnd.
Life time based algorithms

In b-jets it is likely to find:

- tracks with high impact parameter wrt P.V.
- Displaced secondary vertex

Good inner tracking needed

**Different algorithms exist based on:**

- transverse / longitudinal / 3D impact parameter (or combination of those)
- secondary vertex reconstruction and properties: decay length, mass at vertex, fraction of charged tracks,

**Different methods are then used:**

- Significance of the i.p. / d.l.
- Probability density functions
- Likelihood methods
- Neural networks
Life time based algorithms

3D Impact parameter distribution for tracks of b-jets vs tracks of uds and c jets

distribution of significance

\[ S = \frac{ip}{\sigma_{ip}} \]
Two algorithms described...

**CMS Track counting**
- Order track in decreasing impact parameter significance $S$
- Use the significance of $n^{th}$ track as “discriminator”
- Parameter $n$ can be tuned according to analysis needs

**Atlas 2D algorithm**
- Compute $S$
- Probability of a track to come from “b” or “uds” $S$-distribution is computed
- Ratio between $b$ and uds probability computed
- Ratios of all tracks combined to give a “jet weight”

Algorithms based on secondary vertex reconstruction are more powerful and are implemented by both experiments
The typical output of a b-tagging algorithm is a float variable indicating the "b-likeness" of a jet.
Identification efficiency

Atlas and CMS obtain similar b-tagging performances: rejection $\sim 300$ (=mis-tagging $3 \times 10^{-3}$) at efficiencies $\sim 50\%$
The rejection of uds, g and c jets is better in central eta region.
b-tagging reconstruction is limited by tracking to the region
-2.5 < eta < 2.5

CMS @ 50% b-eff
Atlas @60% b-eff
mis-identification vs Pt

At too low energies tracks have a low Pt and the multiple scattering rise the I.P. uncertainty

At high energy:
- track reconstruction is harder
- more tracks -> easier to find high I.P. tracks in light quark jets
Soft lepton tagging

- limited by BR of B to leptons (~10% per lepton)
- can be used without vertex information
- exploit different information wrt lifetime base algorithms
b-tag trigger

Trigger based on b-tagging can be used at High Level Triggers (>L1) when tracker data is available

- regional track reconstruction is needed
- simple algorithms
- reconstruction applied to 1\textsuperscript{st} and 2\textsuperscript{nd} most energetic jets

useful e.g. in pure hadronic ttbar
Ideal vs actual performances

Atlas studies on performance loss:
- initial detector (2 pixel layers)
- pile up events
- pixel inefficiencies

CMS studies on misalignment:
- Perfect detector
- Initial alignment
- Track based “realignment”
Two type of “calibrations” on data:

- **studies of track impact parameter distributions** (used by several b-tagging algorithms)
- **direct study of efficiency on ttbar sample**
  - select pure top sample without b-tagging (or tagging only one b)
  - using only kinematic constraints (W mass, top mass) identify q-jets and b-jets
  - measure algorithm efficiency

*semileptonic ttbar is studied by both CMS and Atlas for this purpose*
Generators

Generator used for b-tag studies:

- **Pythia** (CMS and Atlas)
- **MC@NLO** (top Atlas)

Possible generator effects:

- **Number of charged tracks in light jets**
- **B-decay** (number of tracks, angular distribution)

*no specific study on that (as far as I know)*
Italian activities

Italian groups are deeply involved in b-tag both in Atlas and CMS

• CMS
  – Pisa: Impact Parameter based algorithms, HL Trigger
  – Firenze: Soft Muons algorithm

• ATLAS
  – Genova: Algorithms, trigger and calibration
  – Milano: Susy b-tag
  – Udine: top studies
Backup

Backup slides
Identification efficiency

DISCLAIMER:
Different sample/Jet energy/Eta range

CMS: CombinedBTag

Atlas
More on mis-alignment

Misalignment effects on secondary vertex based algorithm

Pixel misalignment increased by a factor 3
Conclusion

- Different type of algorithms are implemented
- Atlas and CMS experiments reach similar performances of efficiency vs rejection
- Expected “detector effects” (misalignment, inefficiency, staging in initial detectors) can change rejection up to a factor 2
- $t\bar{t}b$bar is expected to be best channel for calibration