Beauty Production and Identification at CMS

Alexander Schmidt

Universität Zürich
Physik-Institut

presented at:

BEACH 2008
The 8th International Conference on Hyperons, Charm, and Beauty Hadrons
22 - 28 June 2008 UNIVERSITY OF SOUTH CAROLINA
Outline

- the CMS detector

- $b$-production in hadron collisions

- applications of $b$-identification methods

- some technical details about $b$-tagging at CMS and expected performance

- prospects for inclusive $b$-production measurement

- further methods of measuring flavour fractions in jets
**magnet:**
4 Tesla superconducting solenoid successfully tested

**calorimeters:**
ECAL: 76000 crystals of PbWO₄
HCAL: brass absorber + scintillator sandwich

**muon system:**
barrel: drift tubes
decap: cathode strip chambers
barrel+endcap: resistive plate chambers

**trigger:**
40 MHz bunch crossing rate reduced to 100Hz by Level-1 and High Level Triggers
Silicon Tracker

of special importance for b-physics:

- is the largest ever built:
  - 200m² of active silicon
  - 1440 pixel modules
  - 15148 strip detector modules
  - hit rate density of 1MHz/mm² (at radius of 4cm)

- pixel cell size of $100 \times 150\mu m^2$
- spatial resolution of $15 - 20\mu m$

- 66 million pixels = readout channels
the $b$ quark production cross-section was first measured by UA1 at CERN in 1988 at $\sqrt{s} = 630\text{GeV}$

$$\sigma(p\bar{p} \rightarrow b\bar{b} + X) = 10.2 \pm 3.3\mu\text{b}$$


later measurements at CDF and D0 showed significant deviations from calculations (initially factor >2) even when running at UA1 energy:


meanwhile, updated fragmentation schemes reduced the discrepancy

=> looking forward to LHC results at much higher transverse energies
Further Sources of b Quarks

b quarks play a crucial role in the search for new physics at LHC:

**Standard Model Higgs Boson:**
- decays into b¯b for low Higgs masses (B.R. > 50%)
- candidate channel: association with top
- suffers from t¯t + jets backgrounds

**SUSY decay chains lead to final states with b¯b:**

example h^0 at LM5:

an accurate knowledge of the b¯b production rate is critical for an understanding of backgrounds in many BSM searches

**Higgs signal**

background from mis-identified jets

=> more work needs to be done in this case
**Observables used for b-Tagging**

**goal:** identify b-decay in jet as efficient as possible

- closest approach of track to primary vertex
  - positive (negative) if decay occurs upstream (downstream) the jet direction.

- other observables based on **secondary vertex**:
  - flight distance (separation from primary vertex)
  - invariant mass at vertex (large b mass compared to c)
  - number of tracks at vertex (on average 5 for b)
  - ....

- include all available information into combined tagging algorithms based on Likelihood methods or neural networks

**=> using significance of second track**
(ordered by the significance itself)

- because first track is likely to be mismeasured.
- second track => higher efficiency
- third track => higher purity
Detector Alignment

the achievable performance strongly depends on detector alignment and calibration.

scenarios under study:
- **startup**: installation precision + survey + cosmics (e.g. 20µm for pixel barrel sensors)
- **first collisions**: 10pb⁻¹: using low mass resonances like J/psi and Upsilon improvement by a factor of ~5 for pixel detector
- 100pb⁻¹: using high mass resonances like W and Z
- 1000pb⁻¹: tracker can be considered to be aligned

Impact Parameter Significance of second track

=> broader for startup due to large fake track rate (IP value larger)

=> more narrow for startup due to large measurement errors (IP error larger)
example: **impact parameter based b-tagging**
(using second track, ordered by IP)
- very simple, does not require training of any kind (likelihood, NN)
  => application with first data is straightforward
- still reaches high purities
- but it is quite sensitive to misalignment
Performance of b-Tagging (II)

secondary vertex based b-tagging, two examples:

A. secondary vertex only, using flight distance significance:
- also very simple, no training
- still reaches high purities
- turns out to be robust against misalignment
- limited to SV finding efficiency

=> 2% mis-tags at 65% b-efficiency
(startup: 3% mis-tags at 35% b-efficiency)

decrease of performance with misalignment for various algorithms:

B. combined algorithm
- combining all available information
- requires training
- strongly affected by misalignment

light flavour misidentification for combined algorithm:
Performance of b-Tagging (III)

Performance depends on transverse momentum and pseudo-rapidity

Misidentification rates at fixed b-efficiency of 50%

Optimal at ~80 GeV
- degraded track quality at lower $p_T$ (multiple scattering)
- at higher $p_T$ increased track multiplicity and more difficult tracking in dense jets

Optimal in the barrel
- higher material budget and worse detector resolution in forward region

CMS Note 2006/019

Light flavours

Charm
study of differential cross sections of incl. b-production in a full detector simulation:

- generated event samples for the simulation:
  - QCD jets with PYTHIA containing 6% b-jets
  - $\tau\tau$ production as background \( \sigma = 500 \text{pb} \)

- Trigger: using a muon+b-jet trigger giving a total event rate of 6.1 Hz at $L = 2 \times 10^{33} \text{cm}^{-2} \text{s}^{-1}$

### Table

<table>
<thead>
<tr>
<th>$\hat{p}_T$, GeV/c</th>
<th>$\sigma^\text{QCD}$, $\mu$b</th>
</tr>
</thead>
<tbody>
<tr>
<td>50 – 80</td>
<td>20.9</td>
</tr>
<tr>
<td>80 – 120</td>
<td>3.0</td>
</tr>
<tr>
<td>120 – 170</td>
<td>0.5</td>
</tr>
<tr>
<td>170 – 230</td>
<td>0.1</td>
</tr>
<tr>
<td>230 – 300</td>
<td>$2.4 \times 10^{-2}$</td>
</tr>
<tr>
<td>300 – 380</td>
<td>$6.4 \times 10^{-3}$</td>
</tr>
<tr>
<td>380 – 470</td>
<td>$1.9 \times 10^{-3}$</td>
</tr>
<tr>
<td>470 – 600</td>
<td>$6.9 \times 10^{-4}$</td>
</tr>
<tr>
<td>600 – 800</td>
<td>$2.0 \times 10^{-4}$</td>
</tr>
<tr>
<td>800 – 1000</td>
<td>$3.6 \times 10^{-5}$</td>
</tr>
<tr>
<td>1000 – 1400</td>
<td>$1.1 \times 10^{-5}$</td>
</tr>
</tbody>
</table>

High Level Trigger can also use lifetime tagging algorithms:
(track counting algorithm)
The cross-section was measured as a differential in $p_T$ and $\eta$.

$\Rightarrow$ Transverse energy and pseudo-rapidity of jets need to be well reconstructed.

The $p_T$ resolution was found to be about 13% on average.
to measure the cross section, one needs:

1. number of selected events
2. integrated luminosity
3. signal selection efficiency
4. signal purity (flavour fraction)

=> measure 1. and 2.
for 3. and 4. rely as few as possible on simulation.

tagging efficiencies and flavour fractions can also be measured, e.g. by fitting the relative transverse momentum of the muon to the jet

=> can be done with a precision of about 10% for the b-content in 10 fb⁻¹
=> contamination from \( \bar{t}t \) events is 1% for \( p_T<500 \text{GeV} \) and 2.4% for \( p_T > 500 \text{ GeV} \)
systematic uncertainties taken into account:
- jet energy scale: 12% (depending on energy)
- model of detector response and lepton identification: 6%
- b-tagging: 5%
- luminosity: 5%
- trigger: < 3%
- fragmentation modeling

=> systematic uncertainty of 10% dominant below 1 TeV

=> measurement limited to 1.5 TeV by statistical uncertainty
Measurement of b-Tagging Efficiencies

- **top quarks:** select high purity b-jet sample based on top quark observables: mass resonances, leptons, angles, ...

- **muons in jets:** fitting relative $p_T$ of muon

- **System8 method:**
  three uncorrelated identification criteria combined:
  1. working point of (lifetime) algorithm
  2. cut on muon $P_{T_{rel}}$
  3. presence of second b-tagged-jet
      (b-Quark pair production)

=> get system of 8 linear equations with 8 unknowns (flavour content)

- **negative tags:**
  using the negative tail in the IP distribution to estimate mistagging rates
  light flavour jets show symmetric distribution in the ideal case

=> negative tags allow measurement of mistagging rates with data

$$\varepsilon_{data}^{mistag} = \varepsilon_{data}^- \cdot R_{light}$$

can be done with 7% relative uncertainty for a mistag rate of 1%

(CMS PAS BTV 07 002)
- a large number of physics studies depend on efficient and clean identification of b-quark jets:
  Higgs physics, BSM searches, top physics, ...

- knowledge of $b\bar{b}$ production is crucial for these searches

- CMS will be able to measure the differential $b\bar{b}$ production cross section up to 1.5 TeV

- the tools to do this all seem to be in place

- detector commissioning is making good progress

=> waiting for data...