Measurement of the inclusive $b$ production cross section in pp collisions at $\sqrt{s} = 7$ TeV

Lea Caminada
(ETH|PSI)
on behalf of the CMS collaboration

July 22, 2010
Heavy Quark Production

- Heavy Quark production is an important process for the study of QCD
- Previous measurements at other colliders (Tevatron, HERA, LEP, …)
  - reasonable agreement with NLL/NLO QCD predictions
  - sizeable uncertainties
- Large $b\bar{b}$ production cross section in pp collisions at $\sqrt{s} = 7$ TeV at LHC
  - provides access to new regions in phase space
  - $b$ events present important background to many searches
- $b$ production at LHC
  - LO: Flavor creation (FCR)
  - Large NLO contributions: Flavor excitation (FEX) and Gluon splitting (GSP)

b Identification at CMS

- Use of distinct properties of b quarks
  - long lifetime, large mass, hard fragmentation
- Semi-leptonic and hadronic decays
- CMS very well suited for b physics due to excellent tracking and muon detectors
  - Pixel detector for precise reconstruction of secondary vertices
  - Muon system with ability to trigger on low $p_T$ muons ($p_T > 3$ GeV)

Double b-jet candidate
CMS DPS -2010/015

Muon event
CMS PAS MUO-10-002
CMS Inclusive b Production Results

- Based on LHC data collected by the CMS experiment between March and July 2010

- CMS PAS BPH-10-007: Open beauty production cross section with muons in pp collisions at $\sqrt{s} = 7$ TeV

- CMS PAS BPH-10-009: Measurement of the inclusive b-jet production in pp collisions at $\sqrt{s} = 7$ TeV
  - Poster by S. Honc:
    "Inclusive b-jet production measurement on early CMS data"

- Two independent measurements with their own systematic uncertainties and covering different regions in phase space
Open b Production with Muons

- Semi-leptonic b decays into muons
  - Single muon trigger ($p_T > 3$ GeV)
  - Muon $p_T > 6$ GeV, $|\eta| < 2.1$
    (Efficiency: trigger ~82%, reconstruction ~97%)
  - Signal events discriminated from background events based on muon $p_T^{\text{rel}}$ (on average harder in b events due to larger b mass)
  - b direction reconstructed from tracks only
    - Tracks clustered by anti-$k_T$ (R=0.5) algorithm
    - Very good angular resolution (2-8%)
    - Efficiency of 74% to almost 100% depending on muon $p_T$

- Measurement of total cross section and differential cross section as a function of muon $p_T$ and pseudo-rapidity with an integrated luminosity of $L = 8.1$ nb$^{-1}$
Cross Section Measurement

- Binned maximum likelihood fit to measured $p_T^{\text{rel}}$ distribution
  - $b$ and $c$ templates from MC (signal validated in $b$-enriched data)
  - Data-driven template for muons from light quarks and gluons (measurement of in-flight decays)
  - Background combined in fit
  - Different templates for each bin in muon $p_T$ and $\eta$

$$N_{b}^{\text{data}} = f_{b}^{\text{fit}} N_{b}^{\text{data}}$$
$$\sigma \equiv \sigma(pp \rightarrow b\bar{b} + X \rightarrow \mu + X', p_T^{\mu} > 6 \text{ GeV}, |\eta^{\mu}| < 2.1) = \frac{N_{b}^{\text{data}}}{\mathcal{L} \varepsilon}$$

- $N_{b}^{\text{data}}$ : number of $b$ events in data
- $\varepsilon$ : trigger and reconstruction efficiency
- $\mathcal{L}$ : integrated luminosity
b Cross Section with Muons at $\sqrt{s} = 7$ TeV

- Visible b cross section

\[ \sigma = (1.48 \pm 0.04_{\text{stat}} \pm 0.22_{\text{syst}} \pm 0.16_{\text{lumi}}) \mu b \]

\[ \sigma_{\text{MC@NLO}} = [0.84^{+0.36}_{-0.19}(\text{scale}) \pm 0.08(m_b) \pm 0.04(\text{pdf})] \mu b \]

\[ (\mu_F = \mu_R = p_T) \]

- Differential b cross section

- Measurement in agreement with MC@NLO for muon $p_T > 12$ GeV, while data is above the prediction in the central region at low $p_T$
Systematic Uncertainties

- Systematic uncertainty dominated by the description of the light quark background template and the underlying event as well as the luminosity uncertainty
- Modelling of $b$ production and decay are better understood and have less impact

<table>
<thead>
<tr>
<th>source</th>
<th>uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trigger</td>
<td>3–5%</td>
</tr>
<tr>
<td>Muon reconstruction</td>
<td>3%</td>
</tr>
<tr>
<td>Tracking efficiency</td>
<td>2%</td>
</tr>
<tr>
<td>Background template shape uncertainty</td>
<td>1–10%</td>
</tr>
<tr>
<td>Background composition</td>
<td>3–6%</td>
</tr>
<tr>
<td>Production mechanism</td>
<td>2–5%</td>
</tr>
<tr>
<td>Fragmentation</td>
<td>1–4%</td>
</tr>
<tr>
<td>Decay</td>
<td>3%</td>
</tr>
<tr>
<td>MC statistics</td>
<td>1–4%</td>
</tr>
<tr>
<td>Underlying Event</td>
<td>10%</td>
</tr>
<tr>
<td>Luminosity</td>
<td>11%</td>
</tr>
<tr>
<td>total</td>
<td>16–20%</td>
</tr>
</tbody>
</table>
Inclusive b-jet Production

- Measurement of the inclusive b-jet cross section and ratio to the inclusive jet production with $L = 60\text{ nb}^{-1}$
- Events collected with a combination of minimum bias and jet triggers
- Jets ($18 < p_T < 300\text{ GeV}, |y| < 2$) reconstructed by anti-$k_T$ algorithm ($R=0.5$) using tracker and calorimeter information (Particle Flow) [CMS PAS PFT-10-002]
- b-tagging based on secondary vertex reconstruction
  - Vertex with at least 3 tracks and large flight length significance
  - b-tagging efficiency from MC, verified in subsample by measurement of scale factor using $p_T^{rel}$
  \[ \frac{\epsilon_b^{Data}}{\epsilon_b^{MC}} = 0.98 \pm 0.08 \pm 0.18 \]
  - Mistag rate from MC, constrained by data-driven study using negative-tag discriminator

\[ \sqrt{s} = 7\text{ TeV} \]

\[ p_T \text{ (GeV)} \]
b-tagged Sample Purity

- Estimated using two complementary approaches
  1) Data-based: Fit to secondary vertex mass
  2) MC-based: \( f_b = \frac{F_b \varepsilon_b}{F_b \varepsilon_b + F_c \varepsilon_c + F_l \varepsilon_l} \) (\( F \): flavor fraction)

- Good agreement between data and MC: \( \frac{\text{Data}}{\text{MC}} = 0.976 \pm 0.022 \)
- Central values taken from MC for proper treatment of \( p_T \) and \( y \) dependence
b-Jet Cross Section at $\sqrt{s} = 7$ TeV

$$\frac{d^2\sigma_{b-jets}}{dp_Tdy} = \frac{N_{\text{tagged}} f_b C_{\text{smear}}}{\epsilon_{\text{jet}} \epsilon_b \Delta p_T \Delta y \mathcal{L}}$$

with $C_{\text{smear}}$: unfolding correction
$\epsilon_{\text{jet}}$: jet reconstruction efficiency
$\epsilon_b$: b-tagging efficiency
Ratio to Inclusive Jet Cross Section

• Inclusive jet cross section measurement
• Measurement of ratio reduces experimental uncertainty from jet energy reconstruction and luminosity
• Fit of measured ratio of data and PYTHIA for $30 < p_T < 150$ GeV and $|y| < 2$ yields scale factor of $0.99 \pm 0.02$ (stat) $\pm 0.21$ (syst)

**Talk by M. Voutilainen**

[CMS PAS QCD-10-011]
Conclusions

• Presented the first measurements of the inclusive b production at √s = 7 TeV

• Open b production with muons:
  - Measurement for muon p_T = 6-30 GeV, |η| < 2.1 with statistical error of 5-20% and systematic uncertainty of 16-20%
  - Good agreement with MC@NLO at muon p_T > 12 GeV, while data are above the prediction in the central region at low p_T

• Inclusive b-jet production:
  - Measurement for jet p_T = 18-300 GeV, |y| < 2
  - Overall good agreement with PYTHIA within ~2% statistical and 21% systematic uncertainty
  - Reasonable agreement with MC@NLO for overall cross section, but shape differences in p_T and y
Backup
CMS Detector
CMS Muon System

$\eta = 0.8$

RPC

ME 1

ME 2

ME 3

ME 4

CSC
Muon Trigger Efficiency

Muons with $|\eta| < 1.2$
- Data, 84 nb$^{-1}$
- Simulation

CMS Preliminary, $\sqrt{s} = 7$ TeV

Muons with $1.2 < |\eta| < 2.1$
- Data, 84 nb$^{-1}$
- Simulation

CMS Preliminary, $\sqrt{s} = 7$ TeV
Fraction of pions, kaons and protons mis-identified as muons
Jet Trigger

- Minimum bias and single jet triggers $p_T > 6, 15, 30$ GeV
- Combined exclusively at ~99% turn-on
Unfolding

• Ansatz method to correct jet $p_T$ back to particle level
• Phenomenological power law motivated by parton model (Feynman, Field, Fox), extended at the Tevatron and updated at CMS for low $p_T$ and b-jets

$f(p_T) = N_0 p_T^{-\alpha} \left(1 - \frac{2p_T \cosh(y_{\text{min}})}{\sqrt{s}} \right)_{\text{high } p_T}^{\beta} \exp(-\gamma/p_T)_{\text{low } p_T \text{ and b-jets}}$

\[
F(p_T) = \int_0^\infty f(p'_T)R(p'_T - p_T; \sigma)dp'_T
\]

$R(p'_T - p_T; \sigma)$: smearing function

$C_{\text{res}} = f(p_T)/F(p_T)$

\[\text{CMS preliminary, 60 nb}^{-1} \quad \sqrt{s} = 7 \text{ TeV}\]
B tagging: Secondary Vertex Properties and Discriminator
B tagging Efficiency

- 12% systematic uncertainty derived from study of jet $p_T$ and $\eta$ modelling (4-8%), muon selection (2-8%), jet flavor assignment (2%), pile-up (3%), shape of light quark background (3-5%)
- Additional systematic uncertainty of 15% to effects not yet studied ($p_T^{\text{rel}}$ shape for $b$ and non-$b$ jets, fragmentation, effect of trigger, jet energy scale uncertainty)

<table>
<thead>
<tr>
<th>Tagger+Operating Point</th>
<th>$\epsilon_b^{\text{data}}$</th>
<th>$\epsilon_b^{\text{MC}}$</th>
<th>$SF_b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSVHPT</td>
<td>0.203 ± 0.015</td>
<td>0.207± 0.002</td>
<td>0.98 ± 0.08± 0.18</td>
</tr>
</tbody>
</table>