Inclusive $b$ Production Measurements at CMS

Lea Caminada (LBNL)

LBNL RPM Seminar
September 30, 2010
Outline

• Introduction

• The CMS Detector and its Performance

• Open Beauty production with Muons

• Inclusive b-jet production

• Conclusions
Introduction

- Heavy Quark production is an important process for the study of QCD

- Many physics processes at the LHC produce b-jets in the final state
  - Pure QCD (access to $b$ PDFs)
  - Decays of various heavy particles (top, W, Z, H, SUSY particles, ...)
  - Associated production (with W, Z, H, ...)

- 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 and rates
  - $b$ events provide major background to many searches
Introduction

- Previous measurements at other colliders (Tevatron, HERA, LEP, ...)
  - reasonable agreement with NLL/NLO QCD predictions
  - sizeable uncertainties

- Great interest to verify the results at higher center-of-mass energy provided by the LHC
Heavy Quark Production

• LO
  - Flavor creation (FCR): $gg$ fusion (dominant) and $q\bar{q}$ annihilation

• Large NLO contributions
  - Flavor excitation (FEX): $b\bar{b}$ from the sea, only one $b$ participates in hard scattering
  - Gluon splitting (GSP): $g \rightarrow b\bar{b}$ in initial or final state

• Production mechanism not separated in analyses presented here

Herwig 6.5


1

0.5

0.1

b-jet pt [GeV]

GSP

FEX

100

1000

b-jet pt [GeV]
The CMS Detector

- **Magnet**
  - 3.8 T

- **Tracking**
  - 200 m² silicon
  - pixels and strip

- **Calorimeter**
  - ECAL: 76 000 PbWO$_4$ crystals
  - HCAL: brass absorbers and scintillators

- **Muon System**
  - Drift Tube Chambers
  - Cathode Strip Chambers
  - Resistive Plate Chambers

- **Trigger System**
  - L1: hardware (40 MHz → 100 kHz)
  - HLT: software (100 kHz → 100 Hz)
b Identification at CMS

- Use of distinct properties of b quarks
  - long lifetime, large mass, hard fragmentation
- Semi-leptonic and hadronic decays
- Tracking and muon detectors are main subdetectors for early heavy flavor physics
  - Pixel detector for precise reconstruction of secondary vertices
  - Muon system with ability to trigger on low $p_T$ muons ($p_T > 3$ GeV)
CMS Tracker

10 M strips

60 M pixels

Material budget estimate from $\gamma$ conversion

=> Data and simulation agree within 10%
Tracker Performance

- CMS Tracker is well understood and performing as expected from simulation
  - Momentum scale measurement using $K_S$ mass
  - Primary vertex resolution
  - Track impact parameter resolution
• Three different algorithms for muon reconstruction:
  - Standalone muons
  - Tracker muons
  - Global muons (combined fit of hits in tracker and muon system)

used in the analyses presented here
Performance of Global Muon Reconstruction

- Muon system well commissioned in cosmic runs 2008/09 and early collision data taking in 2009/10

Kinematic distributions in minimum bias events

Global muon reconstruction efficiency (measured from data using $J/\Psi$ resonance)
Measurements of Inclusive b Production at CMS

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

• Presented at ICHEP 2010

• Open beauty production cross section with muons in pp collisions at $\sqrt{s} = 7$ TeV ($L = 8$ nb$^{-1}$)
  CMS PAS BPH-10-007

• Measurement of the inclusive b-jet production in pp collisions at $\sqrt{s} = 7$ TeV ($L = 60$ nb$^{-1}$)
  CMS PAS BPH-10-009

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

• Semi-leptonic $b$ decays into muons
  - Direct ($b \rightarrow \mu X$) and cascade ($b \rightarrow c \rightarrow \mu X$) decays
  - Kinematic selection: muon $p_T > 6$ GeV, $|\eta| < 2.1$
  ⇒ Acceptance $\approx 1\%$

• Background
  - Charm decays to muons
  - Fake muons from $\pi/K$ in-flight decays and hadronic punch through
Methodology

• Signal events discriminated from background based on muon $p_T^{\text{rel}}$
  - harder in $b$-events than in background events due to larger mass of $b$-quark

• Binned maximum likelihood fit to measured $p_T^{\text{rel}}$ distribution based on simulated template distributions

• Measurement of total cross section and differential cross section as a function of muon $p_T$ and pseudo-rapidity
Event Selection

- Data collected in April/May 2010 (L = 8 nb⁻¹)
- Single muon trigger (p_T > 3 GeV)
- Primary vertex with >= 3 tracks
- Muon p_T > 6 GeV, |η| < 2.1
  - Efficiency: trigger ~82%, reconstruction ~97%
- b direction reconstructed from tracks only
  - Tracks clustered by anti-k_T (D=0.5) algorithm (→ TrackJets)
  - Muon momentum subtracted from TrackJet momentum
  - p_T^{rel} between muon and closest TrackJet
  - TrackJet provide very good angular resolution (2-8%)
  - Efficiency of 74% to almost 100% depending on muon p_T
Muon Trigger Efficiency

- Derived from data using two independent methods
  - Tag & probe on di-muons from $J/\Psi$
  - Single muons in minimum bias events
- Efficiency turn-on curve well described by simulation, plateau few percent lower in data
- Efficiency as measured from data is used in analysis
Binned Maximum Likelihood Fit

- Binned maximum likelihood fit to measured $p_T^{rel}$ 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
  - $c$-to-udsg background composition treated as systematic uncertainty
  - Different templates for each bin in muon $p_T$ and $\eta$
Validation of signal $p_{T}^{rel}$ templates

- Systematic uncertainty of $p_{T}^{rel}$ shape in b-events
  - due to modelling of $b$ production mechanism, $b$-quark fragmentation and decay
  - ~7% (determined by comparing different MC tunes and generators)

- Data-driven validation in $b$-enriched sample obtained by cut on muon impact parameter significance
  - $b$ purity of 86% for $d_{o}/\sigma_{d} > 20$
  - Data and MC agree within limited statistics
Determination of background $p_T^{\text{rel}}$ templates

- Muon fake probability measured using low mass resonances $K_S^0 \rightarrow \pi^+\pi^-$, $\Lambda \rightarrow p\pi^-$ and $\phi \rightarrow K^+K^-$.  
- Data-driven udsg template obtained by re-weighting the hadronic track spectrum in minimum bias events by the muon fake probability and measuring the $p_T^{\text{rel}}$ between any track and the closest TrackJet.
- $p_T^{\text{rel}}$ distribution in data significantly harder (covered by systematics)
Cross Section Measurement

• Visible cross section defined by muon kinematic range

\[ \sigma \equiv \sigma(pp \to b\bar{b} + X \to \mu + X', p_T^\mu > 6 \text{ GeV}, |\eta^\mu| < 2.1) = \frac{N_b^{\text{data}}}{\mathcal{L}\varepsilon} \]

- \( N_b^{\text{data}} \): number of b events in data determined by the fit
- \( \varepsilon \): trigger and reconstruction efficiency
- \( \mathcal{L} \): integrated luminosity

• Result

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

• Compared to

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

---

**MC@NLO+HERWIG**

CTEQ6M PDF

\( m_b = 4.75 \text{ GeV} \)

\( \mu_F = \mu_R = p_T \)
Differential b Cross Section at $\sqrt{s} = 7$ TeV

- 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>
Recent Data/Theory Comparison
(Cacciari and Nason)

- CMS data compared to FONLL calculation
Inclusive b-Jet Production
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 ($D=0.5$) using tracker and calorimeter information (Particle Flow) to extend measurement to low $p_T$
- B-tagging based on secondary vertex reconstruction
- Data-driven techniques to control b-tagging efficiency and purity
- Unfolding technique to correct $p_T$ bin migration
Minimum Bias and Jet Trigger

- Minimum bias and single jet triggers $p_T > 6, 15, 30$ GeV
- Combined exclusively at $\sim 99\%$ turn-on
- Low $p_T$ results limited to run periods with negligible pile-up ($10$ nb$^{-1}$)
Particle Flow

- Particle Flow algorithm combines information from all subdetectors to create a unique list of reconstructed particles
- This list is then used as input to the jet clustering algorithm
- Significant improvement of jet response and resolution at low $p_T$

![Particle Flow Diagram]

Lea Caminada  LBNL RPM Seminar  28
Jet Energy Corrections

- JEC currently derived from MC, cross-checked by in-situ jet calibration studies using di-jet and $\gamma$+jet events
- Relative $\eta$
  - Residual corrections determined from di-jet balance (with reference jet in central region)
- Absolute $p_T$
  - 5% JEC uncertainty for tracking-based jets
  - Supported by missing $E_T$ projection fraction method in $\gamma$+ jet events
Jet $p_T$ Resolution

- Jet resolution estimated from di-jet asymmetry:
  \[ A = \frac{p_T^{\text{jet1}} - p_T^{\text{jet2}}}{p_T^{\text{jet1}} + p_T^{\text{jet2}}} \]

- For approximately equal values of jet $p_T$:
  \[ \frac{\sigma(p_T)}{p_T} = \sqrt{2\sigma_A} \]

- Resolution in data and MC agree within 10%
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

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

$f(p_T)$: Ansatz function to parametrize true jet $p_T$ spectrum

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

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

$$C_{\text{res}} = f(p_T)/F(p_T)$$
Inclusive Jet Cross Section

- Inclusive jet $p_T$ spectrum in good agreement with NLO theory
- Main systematic uncertainties from jet energy scale (5%), jet resolution (10%) and luminosity (11%)
B-tagging Performance

- Based on reconstruction of secondary vertices (SV)
- SV with at least 3 tracks and large flight length significance
- SV tagging commissioned with first lumi
  => MC simulation well reproduces the measured b-tagging observables
B-tagging Efficiency

- b-tagging efficiency as a function of $p_T$ and $y$ is taken from MC
- Verified in subsample by measurement of data/MC scale factors based on $p_T^{\text{rel}}$:

$$\epsilon_b^{\text{data}} = \frac{f_b^{\text{tag}} \cdot N_{\text{data}}^{\text{tag}}}{f_b^{\text{tag}} \cdot N_{\text{data}}^{\text{tag}} + f_b^{\text{untag}} \cdot N_{\text{data}}^{\text{untag}}}$$

$\Rightarrow$ Measured scale factors are compatible with 1 within the systematic uncertainty (20%)
B-tagging Mistag Rate

- Mistag rate constrained by data-driven study using negative tag discriminators

\[
\varepsilon_{\text{mistag}} = \varepsilon_{\text{data}} \cdot R_{\text{light}},
\]

where

\[
R_{\text{light}} = \frac{\varepsilon_{\text{mistag}}}{\varepsilon_{\text{MC}}}
\]

SSVHEM tagger

\[|\eta(\text{jet})| < 2.4\]
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: \( \text{Data/MC} = 0.976 \pm 0.022 \)
- Central values taken from MC for proper treatment of \( p_T \) and \( y \) dependence
\[ \frac{d^2 \sigma_{b\text{-}jets}}{dp_T dy} = \frac{N_{\text{tagged}} f_b C_{\text{smear}}}{\epsilon_{\text{jet}} \epsilon_b \Delta p_T \Delta y \mathcal{L}} \]

with \( C_{\text{smear}} \): unfolding correction
\( \epsilon_{\text{jet}} \): jet reconstruction efficiency
\( \epsilon_b \): b-tagging efficiency

**b-Jet Cross Section at \( \sqrt{s} = 7 \text{ TeV} \)**

- MC@NLO
- Pythia
- Exp. uncertainty (centered on ansatz)

**Data / NLO theory**

- \( |y| < 0.5 \times 125 \)
- \( 0.5 \leq |y| < 1 \times 25 \)
- \( 1 \leq |y| < 1.5 \times 5 \)
- \( 1.5 \leq |y| < 2 \)

**Uncertainty on b-jet production (%)**

- Total uncertainty
- b-tag efficiency (20%)
- Jet energy scale (5%)
- Charm mistag (20%)
- Light mistag (50%)
Ratio to Inclusive Jet Cross Section

- 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)
Conclusions

• CMS detector is working very well which made the first measurements of the inclusive $b$ production possible

• Open $b$ production with muons:
  - Measurement for muon $p_T = 6$-$30$ GeV, $|\eta| < 2.1$ 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 $p_T$
  - Recent comparison to FONLL

• 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
Differential b Cross Section

\[ \frac{d\sigma}{dp_T} \rightarrow b + X \rightarrow \mu + X' [\text{nb/GeV}] \]

CMS data
MC@NLO (CTEQ6M, \(m_\mu = 4.75 \text{ GeV}\))
MC@NLO scale variation (0.5-2)
PYTHIA (MSEL 1, CTEQ6L1)
HERWIG (1500, CTEQ6L1)
HERWIG (1705, CTEQ6L1)
CASCADE

\( \sqrt{s} = 7 \text{ TeV} \)
\(L = 8.1 \text{ nb}^{-1}\)
\( |\eta^\mu| < 2.1 \)
Fraction of pions, kaons and protons mis-identified as muons
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 ($p_T^{rel}$ shape for b and non-b jets, fragmentation, effect of trigger, jet energy scale uncertainty)

\[
\varepsilon_b^{data} = \frac{f_b^{tag} \cdot N_{data}^{tag}}{f_b^{tag} \cdot N_{data}^{tag} + f_b^{untag} \cdot N_{data}^{untag}}
\]

<table>
<thead>
<tr>
<th>Tagger+Operating Point</th>
<th>$\varepsilon_b^{data}$</th>
<th>$\varepsilon_b^{MC}$</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>
Inclusive Jet Cross Section: Systematic

CMS preliminary, 60 nb$^{-1}$, $\sqrt{s} = 7$ TeV

- Total uncertainty
- Absolute $p_T$ ($\pm 5\%$)
- Relative $p_T$ ($\pm 1\%$)
- $p_T$ resolution ($\pm 10\%$)

Theory uncertainty:
- PDF (CTEQ6.6)
- NP (Pythia-Herwig++)
- Scale ($\mu/2 \rightarrow 2\mu$)

Anti-$k_T$, $R=0.5$ PF

Anti-$k_T$, $R=0.5$ Jets
Jet Relative Response

\[ R(\eta^{\text{probe}}, p_T^{\text{dijet}}) = \frac{2+ < B >}{2- < B >} \]

where \[ B = \frac{p_T^{\text{probe}} - p_T^{\text{barrel}}}{p_T^{\text{dijet}}} \]

- For \( \sqrt{s} = 7 \text{ TeV} \), anti-\( k_t \) \( R = 0.5 \)
  - Raw PF Jets
  - 40 < dijet \( p_T < 63 \text{ GeV} \)
    - CMS Preliminary
  - Data (20 nb\(^{-1}\))
  - CMS Simulation

- For \( \sqrt{s} = 7 \text{ TeV} \), anti-\( k_t \) \( R = 0.5 \)
  - Raw PF Jets
  - 63 < dijet \( p_T < 80 \text{ GeV} \)
    - CMS Preliminary
  - Data (71 nb\(^{-1}\))
  - CMS Simulation

- For \( \sqrt{s} = 7 \text{ TeV} \), anti-\( k_t \) \( R = 0.5 \)
  - Raw PF Jets
  - 80 < dijet \( p_T < 120 \text{ GeV} \)
    - CMS Preliminary
  - Data (71 nb\(^{-1}\))
  - CMS Simulation

- For \( \sqrt{s} = 7 \text{ TeV} \), anti-\( k_t \) \( R = 0.5 \)
  - Raw PF Jets
  - 120 < dijet \( p_T < 150 \text{ GeV} \)
    - CMS Preliminary
  - Data (71 nb\(^{-1}\))
  - CMS Simulation
Jet Response

\[ p_T^\gamma + p_T^{\text{recoil}} = 0 \]

\[ R_{\gamma} p_T^\gamma + R_{\text{recoil}} p_T^{\text{recoil}} = -E_T^{\text{miss}} \]

\[ R_{\gamma} = 1 \]

\[ R_{\text{recoil}} = 1 + \frac{E_T^{\text{miss}} \cdot p_T^{\gamma \text{ gamma}}}{(p_T^{\gamma})^2} \equiv R_{\text{MPF}} \]
Photon Pair Invariant Mass Distribution

ECAL Barrel

ECAL Endcap
Jet $p_T$ Resolution

![Graphs showing $d(p_T^{\text{eff}})/dp_T$ for different jets.](image)
Jet $p_T$ Resolution: Asymmetry Method

$$A = \frac{p_T^{\text{jet1}} - p_T^{\text{jet2}}}{p_T^{\text{jet1}} + p_T^{\text{jet2}}}$$
Jet Position Resolution ($\eta$)
Jet Position Resolution ($\phi$)
Tracking Efficiency for Pions

$D^0 \rightarrow K^- \pi^+ \pi^- \pi^+ \text{ ("K3\pi")}$

$D^0 \rightarrow K^- \pi^+ \text{ ("K\pi")}$

$$R = \frac{N_{K3\pi}}{N_{K\pi}} \cdot \frac{\epsilon_{K\pi}}{\epsilon_{K3\pi}}$$

$R\text{ (PDG)} = 2.08 \pm 0.05$

$$\frac{\epsilon\text{ (data)}}{\epsilon\text{ (MC)}} = \sqrt{\frac{R}{R\text{ (PDG)}}}$$
Measurement of $\sigma(pp \rightarrow b\bar{b}X)$ at LHCb
CMS Measurement of non-prompt $J/\Psi$ Production