Measurement of $t\bar{t}$ Production Cross Section Using b-tagged Semileptonic Events

Sadia Khalil
for the CMS Collaboration
Young Scientific Forum(4)
Rencontres de Moriond (EWK), La Thuile, AO, Italy
3-10 Mar, 2012
**Introduction**

- **Pair production**
  - $\sigma \approx 7.5$ pb (Tevatron 1.96 TeV ppbar)
  - $\sigma \approx 165$ pb (LHC 7 TeV)

- **Top Pair final state decays**
  - Golden channels: lepton + jets and dilepton
  - Challenging channels: full hadronic, tau + mu

- **Top Pair Branching Fractions**
  - "alljets" 46%
  - $\tau+\text{jets}$ 15%
  - $\mu+\text{jets}$ 15%
  - "dileptons" 15%
  - "lepton+jets"

- $\sim 90\%$(Tevatron)
- $\sim 15\%$(LHC)
- $\sim 10\%$(Tevatron)
- $\sim 85\%$(LHC)
Cross-section Measurement

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

- CMS 2011 combination
  TOP-11-024 (L=0.8-1.1/fb)
  $166 \pm 2 \pm 11 \pm 8$
  (val. ± stat. ± syst. ± lumi.)

- CMS e/µ+jets+btag
  TOP-11-003 (L=0.8-1.1/fb)
  $164 \pm 3 \pm 12 \pm 7$
  (val. ± stat. ± syst. ± lumi.)

- CMS dilepton (ee,µµ, eµ)
  TOP-11-005 (L=1.1/fb)
  $170 \pm 4 \pm 16 \pm 8$
  (val. ± stat. ± syst. ± lum)

- CMS all-hadronic
  TOP-11-007 (L=1.1/fb)
  $136 \pm 20 \pm 40 \pm 8$
  (val. ± stat. ± syst. ± lumi.)

- CMS dilepton (µτ)
  TOP-11-006 (L=1.1/fb)
  $149 \pm 24 \pm 26 \pm 9$
  (val. ± stat. ± syst. ± lumi.)

https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsTOP

Friday, March 9, 2012
Event Selection

- **Triggers**
  - Single Electron/Muon

- **Electrons**
  - Electron ID with 70% (95%) reconstruction efficiency (second electron for dilepton veto)
  - Relative isolation < 0.01 and conversion rejection
  - $|\eta| < 2.5$, $P_T > 45$ GeV

- **Muons**
  - Global and tracker muon, relative isolation < 0.125
  - $|\eta| < 2.1$, $P_T > 35$ GeV

- **Jets**
  - anti-$kT$ jet clustering algorithm with $R=0.5$
  - $P_T > 30$ GeV and $|\eta|< 2.4$, JetID
  - $\geq 1$ btag jet using standard high efficiency tagger

- **MET**
  - Negative sum of transverse momenta of all reconstructed particles
  - MET > 20.0 GeV

Friday, March 9, 2012

http://cdsweb.cern.ch/record/1386709?ln=en
Simultaneous Heavy Flavor and Top

- **Discriminating variables**
  - N-Jet distribution - Distinguish Top from Wbb
  - “Flavor separator” - Distinguish Wbb from W+charm and W+LF
    - Secondary vertex (secvtx) mass - invariant mass of tracks originating at identified secvtx
    - Good separation of bottom from charm and light flavor jets
- **Perform a simultaneous 3D($N_{Jets}$, $N_{Tags}$, secvtx mass) Maximum Profile likelihood Fit**
Systematic Uncertainties (1)

- How do we put systematic uncertainties into fitter?
  - Create and compare the nominal and shifted MC templates
- All functional forms are parameterized as polynomials
  - Generally 2\textsuperscript{nd} order and are applied as multiplicative factors.
- For example: Able to extract btag efficiency in-situ
  - Use b-tagging efficiency and mistag SF constrained to 1.0 ±0.1

![CMS Simulation](image)

\(\mu\)-channel, \(\geq 1\) tag

- Measured: 97± 1 %

Btag group measurement of 95 ± 1 % is covered in our 1σ interval

Friday, March 9, 2012
### Systematic Uncertainties (2)

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Constraint (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b$-tag Efficiency Scale Factor</td>
<td>$100 \pm 10$</td>
</tr>
<tr>
<td>$b$-tag Mistag Scale Factor</td>
<td>$100 \pm 10$</td>
</tr>
<tr>
<td>Jet energy scale relative to nominal</td>
<td>$100 \pm 3$ ($\eta, p_T$ dependent)</td>
</tr>
<tr>
<td>$W$+jets renormalization/factorization scales</td>
<td>$100_{-50}^{+100}$</td>
</tr>
<tr>
<td>$W$+jets background normalization</td>
<td>unconstrained</td>
</tr>
<tr>
<td>QCD background normalization</td>
<td>$100 \pm 100$</td>
</tr>
<tr>
<td>Single-top background normalization</td>
<td>$100 \pm 30$</td>
</tr>
<tr>
<td>$Z$+jets background normalization</td>
<td>$100 \pm 30$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>Muon Analysis</th>
<th>Electron Analysis</th>
<th>Combined Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantity</td>
<td>Uncertainty (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lepton ID/reco/trigger</td>
<td>3.4</td>
<td>3</td>
<td>3.4</td>
</tr>
<tr>
<td>$E_T$ resolution due to unclustered energy</td>
<td>&lt; 1</td>
<td>&lt; 1</td>
<td>&lt; 1</td>
</tr>
<tr>
<td>$t\bar{t}$+jets $Q^2$ scale</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>ISR/FSR</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>ME to PS matching</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Pile-up</td>
<td>2.5</td>
<td>2.6</td>
<td>2.6</td>
</tr>
<tr>
<td>PDF</td>
<td>3.4</td>
<td>3.4</td>
<td>3.4</td>
</tr>
<tr>
<td>Profile Likelihood Parameter</td>
<td>Uncertainty (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jet energy scale and resolution</td>
<td>4.2</td>
<td>4.2</td>
<td>3.1</td>
</tr>
<tr>
<td>$b$-tag efficiency</td>
<td>3.3</td>
<td>3.4</td>
<td>2.4</td>
</tr>
<tr>
<td>$W$+jets $Q^2$ scale</td>
<td>0.9</td>
<td>0.8</td>
<td>0.7</td>
</tr>
<tr>
<td>Combined</td>
<td>7.8</td>
<td>7.8</td>
<td>7.3</td>
</tr>
</tbody>
</table>

**Important!** The full result is not the sum of the squares. Have to account for correlations (fitter does this automatically)

**systematic uncertainties not covered in fit**

**systematic uncertainties covered in fit**
Results

CMS Preliminary $\sqrt{s} = 7$ TeV $\int L dt = 1.1$ fb$^{-1}$, Muons

- # of events / 0.5 GeV
- Secondary Vertex Mass (GeV)

CMS Preliminary $\sqrt{s} = 7$ TeV $\int L dt = 0.8$ fb$^{-1}$, Electrons

- # of events / 0.5 GeV
- Secondary Vertex Mass (GeV)

Friday, March 9, 2012
Results

<table>
<thead>
<tr>
<th>Channel</th>
<th>luminosity</th>
<th>Measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>μ+jets</td>
<td>1.1/ fb</td>
<td>163.2 ± 3.4 (stat.) ± 12.7 (syst.) ± 7.3 (lum.) pb</td>
</tr>
<tr>
<td>e+jets</td>
<td>0.8/ fb</td>
<td>163.0 ± 4.4 (stat.) ± 12.7 (syst.) ± 7.3 (lumi.) pb</td>
</tr>
<tr>
<td>combine</td>
<td>0.8-1.1/ fb</td>
<td>164.4 ± 2.8 (stat.) ± 11.9 (sys.) ± 7.4 (lumi.) pb</td>
</tr>
</tbody>
</table>

- **Btag SF**
  - 96.7 ± 1 %
- **Jet Energy Scale:**
  - 99.2 ± 1.2 % of the nominal in agreement with 100%
- **W+Jets Q^2:**
  - 9 ± 10 % in agreement with the nominal value
- **W+b-jets:**
  - K=1.2 ± 0.3 * SM expectation
- **W+c-jets:**
  - K=1.7 ± 0.1 * SM expectation
Extra material
### Results

#### Fit factors

<table>
<thead>
<tr>
<th></th>
<th>Uncorrected ( \sigma(Top) ) in pb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top</td>
<td>164.2(\pm^{6.1}_{-5.9})</td>
</tr>
<tr>
<td>SingleTop</td>
<td>1.44(\pm^{0.27}_{-0.27})</td>
</tr>
<tr>
<td>Wbx</td>
<td>1.21(\pm^{0.27}_{-0.27})</td>
</tr>
<tr>
<td>Wcx</td>
<td>1.66(\pm^{0.06}_{-0.06})</td>
</tr>
<tr>
<td>Wqq</td>
<td>0.58(\pm^{0.08}_{-0.07})</td>
</tr>
<tr>
<td>ZJets</td>
<td>1.34(\pm^{0.29}_{-0.29})</td>
</tr>
</tbody>
</table>

#### Fit Shifts

<table>
<thead>
<tr>
<th>Shifts</th>
<th>Uncorrected ( \sigma )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q2</td>
<td>0.09(\pm^{0.10}_{-0.10})</td>
</tr>
<tr>
<td>btag</td>
<td>-0.33(\pm^{0.11}_{-0.11})</td>
</tr>
<tr>
<td>jes</td>
<td>-0.28(\pm^{0.39}_{-0.40})</td>
</tr>
<tr>
<td>iftag</td>
<td>0.11(\pm^{1.00}_{-1.00})</td>
</tr>
</tbody>
</table>

- Uncorrected \( \sigma(Top) \) in pb
- Fraction of SM
- Shift from nominal value \( [\sigma] \)
- Constrained to \( 0 \pm 1 \)
- 100\% uncertainty
- Constrained to \( 0 \pm 1 \)
### Results (Correlation Matrix)

<table>
<thead>
<tr>
<th></th>
<th>Top</th>
<th>SingleTop</th>
<th>Wbx</th>
<th>Wcx</th>
<th>Wqq</th>
<th>Zjets</th>
<th>Q2</th>
<th>btag</th>
<th>jes</th>
<th>lftag</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top</td>
<td>1.000</td>
<td>-0.285</td>
<td>-0.180</td>
<td>0.288</td>
<td>0.032</td>
<td>0.074</td>
<td>-0.135</td>
<td>-0.627</td>
<td>-0.835</td>
<td>0.002</td>
</tr>
<tr>
<td>SingleTop</td>
<td>-0.285</td>
<td>1.000</td>
<td>-0.731</td>
<td>0.049</td>
<td>0.047</td>
<td>-0.041</td>
<td>0.069</td>
<td>-0.104</td>
<td>0.134</td>
<td>-0.006</td>
</tr>
<tr>
<td>Wbx</td>
<td>-0.180</td>
<td>-0.731</td>
<td>1.000</td>
<td>0.068</td>
<td>0.123</td>
<td>-0.145</td>
<td>0.295</td>
<td>0.195</td>
<td>0.269</td>
<td>-0.002</td>
</tr>
<tr>
<td>Wcx</td>
<td>0.288</td>
<td>0.049</td>
<td>0.068</td>
<td>1.000</td>
<td>0.053</td>
<td>0.034</td>
<td>0.673</td>
<td>-0.428</td>
<td>-0.204</td>
<td>-0.011</td>
</tr>
<tr>
<td>Wqq</td>
<td>0.032</td>
<td>0.047</td>
<td>0.123</td>
<td>0.053</td>
<td>1.000</td>
<td>-0.139</td>
<td>0.311</td>
<td>-0.058</td>
<td>-0.048</td>
<td>-0.763</td>
</tr>
<tr>
<td>ZJets</td>
<td>0.074</td>
<td>-0.041</td>
<td>-0.145</td>
<td>0.034</td>
<td>-0.139</td>
<td>1.000</td>
<td>0.129</td>
<td>0.000</td>
<td>-0.100</td>
<td>0.002</td>
</tr>
<tr>
<td>Q2</td>
<td>-0.135</td>
<td>0.069</td>
<td>0.295</td>
<td>0.673</td>
<td>0.311</td>
<td>0.129</td>
<td>1.000</td>
<td>-0.022</td>
<td>0.231</td>
<td>-0.016</td>
</tr>
<tr>
<td>btag</td>
<td>-0.627</td>
<td>-0.104</td>
<td>0.195</td>
<td>-0.428</td>
<td>-0.058</td>
<td>0.000</td>
<td>-0.022</td>
<td>1.000</td>
<td>0.460</td>
<td>-0.011</td>
</tr>
<tr>
<td>jes</td>
<td>-0.835</td>
<td>0.134</td>
<td>0.269</td>
<td>-0.204</td>
<td>-0.048</td>
<td>-0.100</td>
<td>0.231</td>
<td>0.460</td>
<td>1.000</td>
<td>0.003</td>
</tr>
<tr>
<td>lftag</td>
<td>0.002</td>
<td>-0.006</td>
<td>-0.002</td>
<td>-0.011</td>
<td>-0.763</td>
<td>0.002</td>
<td>-0.016</td>
<td>-0.011</td>
<td>0.003</td>
<td>1.000</td>
</tr>
</tbody>
</table>
Simple Fit For Top Cross Section

- Take all events (electrons and muons) with
  - $\geq 4$ jets
  - $\geq 1$ b-tag
- Put secondary vertex mass of leading b jet into fit.
- Constrain
  - $Wbb (200\% \pm 200\% SM)$
  - Single top (100\% $\pm$ 30\% SM)
  - QCD (to MET fits $\pm 100\%$)
- Result:
  - $\sigma_{tt} = 150 \pm 17$ (stat + $Wbb$)
  - $\pm 11$ (Btag) $\pm 11$ (JES)
  - $\pm 9$ (Syst) pb
  = $150 \pm 22$ pb

- Most of this is due to uncertainties on $Wbb$!

Theory value $\sigma_{top\; pair} = 157$ pb.

How Can We Reduce Uncertainties due to $Wbb$, B-tagging, and JES?
Simultaneous Fit

- Fit separately lepton flavor (electron, muon), number of jets (1, 2, 3, 4, ≥ 5), and number of tags (1, ≥ 2).
- Fitter maximizes Poisson likelihood $P(N_{\text{pred}}|\text{Data})$ for each bin.
  > It needs to understand how much of each sample (e.g., top) is there for each lepton flavor, jet, and tag region.
- For example, top:

$$\mathcal{N}_{\text{pred}}(\text{lepton, jet, tag}) = \sigma_{t\bar{t}} \cdot MC_{\text{pred}}(\text{lepton, jet, tag}) \cdot P_{B\text{tag}}^N(\text{lepton, jet, tag}|\mathcal{R}_{B\text{tag}}) \cdot P_{\text{JES}}^N(\text{lepton, jet, tag}|\mathcal{R}_{\text{JES}}) \cdot ...$$

- Top cross section (floats in fit)
- MC prediction of events for a given lepton, jet, tags region
- Number of predicted top pair events for a given lepton, jet, tag region
- Polynomial function that describes fractional change to prediction as a function of $R_{b\text{tag}}$. Always define to be exactly 1 when $R_{b\text{tag}} = 0$. 
- Nuisance parameter for the B-tagging
• Weight the tag jets in simulation

\[ \mathcal{W} = \begin{cases} 
S_{b\text{tag}} & \text{heavy flavor jet tagged during simulation} \\
0 & \text{heavy flavor jet not tagged during simulation} \\
S_{\text{mis}} & \text{light flavor jet}
\end{cases} \]

**Efficiency in Data**

---

**Figure 3**: Efficiency of the secondary vertex tagger expected in data for 3.8 \( B \)-Tagging

---

**Figure 5**: Trigger efficiency for electron triggers vs. number of probe passing HLT::pass value for all triggers used.

---

**Figure 6**: Data to MC scale factor is calculated in the following way. We get ID efficiency as a function of transverse momentum for different electron triggers. We are using \( p_T \) cut of 2.5 GeV. The choice of \( B \)-tagging algorithm is the simple secondary vertex high efficiency tagger at the medium operating point. In the analyses that follow the scale factor is taken into account by weighting the actual tags in the simulation regardless of whether the SF is larger or smaller than unity. Each jet gets assigned a weight according to this formula: 

\[ \mathcal{W} = \begin{cases} 
S_{b\text{tag}} & \text{heavy flavor jet tagged during simulation} \\
0 & \text{heavy flavor jet not tagged during simulation} \\
S_{\text{mis}} & \text{light flavor jet}
\end{cases} \]

---

**Figure 10**: Trigger efficiencies shown on the Figure 5. One can see a rise of efficiency with respect to transverse momentum for different electron triggers. We are using \( p_T \) cut of 2.5 GeV to have single efficiency value for all triggers used.
Coverage

- How do we know that the fitter doesn’t get lost somewhere?
- We run pseudo-experiments to determine 15%, 50% and 84% points for various pull distributions
- We correct for these small under coverage/bias

Fitted values | PE Corr. [$\sigma$] | Corr. values
--- | --- | ---
| + Err | - Err | Bias | +Err | -Err | +Err | -Err |
| Electron | 162.4 | 8.9 | -8.3 | -0.06 | 1.04 | 1.01 | 163.0 | 9.3 | -8.4 |
| Muon | 162.9 | 7.9 | -7.7 | -0.03 | 1.01 | 1.03 | 163.2 | 8.0 | -7.9 |
| e+mu | 164.2 | 6.1 | -5.9 | -0.03 | 1.01 | 1.03 | 164.4 | 6.2 | -6.1 |

Top pull distribution for the combine Mu+Ele channel generated with Q2 systematic uncertainty
PDF Uncertainty

- Examine the uncertainties in the PDF set CTEQ6M (default) which are defined in PDF set CTEQ61
- The systematic uncertainty on the acceptance is estimated \((1.004, 0.997)\)

\[
\Delta = \frac{1}{N_{\text{events}}} \sum_{\text{events}} \left[ \sum_{i=1}^{22} \left( \frac{\text{PDF}_i}{\text{PDF}_0} \right)^2 + \sum_{j=1}^{22} \left( \frac{\text{PDF}_j}{\text{PDF}_0} \right)^2 \right] \]

PDF uncertainty eigenvector weights

- The additional uncertainty due to Njets shape is estimated by examining the ratio between each of the Njet distribution corresponding to 44 eigenvectors in CTEQ61 and the nominal
- A flat 3.4% uncertainty for the 68% C.L. brackets the variation seen conservatively
These plots are for events containing at least 3 jets and at least 1 tag.

Figure 2: The muon + jets channel's kinematic distributions of the missing transverse energy (\(E_T\)).

- **Higgs boson search**: \(L_{dt} = 1.1 \, fb\)
- **QCD**: \(W+\text{Jets}\), \(\text{SingleTop}\), \(Z+\text{jets}\), \(t\bar{t}\)
- **Data**: \(\text{Muons}\)

- **Missing Transverse Energy (GeV)**
- **Muon p_T (GeV)**
- **W Transverse Mass (GeV)**
- **Secondary Vertex Mass (GeV)**

**CMS Preliminary, \(\sqrt{s} = 7 \, \text{TeV}\), \(L_{dt} = 1.1 \, fb^{-1}\), Muons**
momentum. These plots are for events containing at least 3 jets and at least 1 tag.

Figure 3: The electron + jets channel's kinematic distributions of the missing transverse energy.

### 4.4 Simultaneous Muon and Electron Channel Analysis

<table>
<thead>
<tr>
<th>CMS Preliminary, √s = 7 TeV, ∫Ldt = 0.8 fb⁻¹, Electrons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td># of events / 10 GeV</td>
</tr>
<tr>
<td># of events / 5 GeV</td>
</tr>
<tr>
<td># of events / 1.25 GeV</td>
</tr>
<tr>
<td># of events / 0.1 GeV</td>
</tr>
</tbody>
</table>

CMS Preliminary, √s = 7 TeV, ∫Ldt = 0.8 fb⁻¹, Electrons

- Transverse mass of the secondary vertex, \( M_{T} \)
- Electron transverse momentum, \( p_{T}^{e} \)
- Missing transverse energy, \( E_{T}^{miss} \)
- Data
- T
- Z+jets
- SingleTop
- W+Jets
- QCD

Friday, March 9, 2012