Evidence for single top production at the Tevatron

Supriya Jain
University of Oklahoma
on behalf of D0 and CDF collaborations

Rencontres de Moriond QCD 2008
11 March 2008
Production of top quarks at the Tevatron

- **Strong interaction**
  - Distinct event signature from the decay of a massive object

- **Electroweak interaction**
  - Smaller cross sections
  - One less massive object hence difficult to identify

\[\sigma_{\text{ttbar}} = 6.77 \pm 0.42 \text{ pb}\]

\[\sigma_{s\text{-channel}} = 0.88 \pm 0.14 \text{ pb}\]

\[\sigma_{t\text{-channel}} = 1.98 \pm 0.30 \text{ pb}\]

Signal : Background \(\sim 3 : 1\)

Signal : Background \(\sim 1 : 15\)

Quoted cross sections at \(M_{\text{top}} = 175 \text{ GeV}\)

--- N. Kidonakis, R. Vogt, Phys. Rev. D 68, 114014 (2003);
Single top production: Goals

- Claim discovery of single top production
- Measure production cross sections
  - $\sigma_s$, $\sigma_t$, $\sigma_{s+t}$
- Perform direct measurement of CKM matrix element $|V_{tb}|$
- Study top quark polarization
- Establish techniques useful for searches for small signals, like the Higgs search
- Probe new physics effects
New physics searches in single top production

- Recent results
  - Limits on W' from CDF
    - \( M(W') > 800 \text{ GeV} \) to \( 825 \text{ GeV} \), depending on couplings and decays
      
    \[ \text{[D0: PLB 641:423-431 (2006)]} \]
  - FCNC gluon coupling limits from D0
    
    \[ \text{[PRL 99:191802 (2007)]} \]
    - \( \kappa_c/\Lambda < 0.15 \text{ TeV}^{-1} \) and \( \kappa_u/\Lambda < 0.038 \text{ TeV}^{-1} \)
      
    \( (\Lambda \text{ is the new physics scale}) \)

Supriya Jain, University of Oklahoma
Evidence for single top production at Tevatron

**D0**
- First evidence at $3.4 \sigma$ using 0.9 fb$^{-1}$ data [PRL 98, 181802 (2007)]
- Improved two of its three analyses (Bayesian NN, and Matrix Element) [arXiv.org:0803.0739, submitted to PRD]

**CDF:**
- Also saw evidence at $3.1 \sigma$ (using 1.5 fb$^{-1}$ data)
- Latest results using 2.2 fb$^{-1}$

**Graph:**
Integrated Luminosity 3586.37 (1/pb)
3.5 fb$^{-1}$ delivered
General analysis strategy

Data → Event Selections → Signal/Background Monte Carlo → Multivariate Analysis
- Likelihood (CDF)
- Neural network (CDF)
- Matrix element (CDF, D0)
- Decision tree (D0)
- Bayesian NN (D0)

Statistical Analysis
- Cross section
- Discriminant
One high-$E_T$ lepton (**electron** or **muon**)

Missing transverse energy **MET** (neutrino)

$\geq 2$ jets

- **s-channel**: 2 b-quark jets
- **t-channel**: 2 b-quark jets, and 1 light quark jet, $q'$

**Event sample composition**

- Top quark pairs
- multijet
- W+jets
- Single top
D0 Results

[arXiv.org:0803.0739]
### Event yields

<table>
<thead>
<tr>
<th>Source</th>
<th>2 jets</th>
<th>3 jets</th>
<th>4 jets</th>
</tr>
</thead>
<tbody>
<tr>
<td>$tb$</td>
<td>16 ± 3</td>
<td>8 ± 2</td>
<td>2 ± 1</td>
</tr>
<tr>
<td>$tqb$</td>
<td>20 ± 4</td>
<td>12 ± 3</td>
<td>4 ± 1</td>
</tr>
<tr>
<td>$t\tilde{t} \rightarrow ll$</td>
<td>39 ± 9</td>
<td>32 ± 7</td>
<td>11 ± 3</td>
</tr>
<tr>
<td>$t\tilde{t} \rightarrow l+\text{jets}$</td>
<td>20 ± 5</td>
<td>103 ± 25</td>
<td>143 ± 33</td>
</tr>
<tr>
<td>$W+b\bar{b}$</td>
<td>261 ± 55</td>
<td>120 ± 24</td>
<td>35 ± 7</td>
</tr>
<tr>
<td>$W+c\bar{c}$</td>
<td>151 ± 31</td>
<td>85 ± 17</td>
<td>23 ± 5</td>
</tr>
<tr>
<td>$W+jj$</td>
<td>119 ± 25</td>
<td>43 ± 9</td>
<td>12 ± 2</td>
</tr>
<tr>
<td>Multijets</td>
<td>95 ± 19</td>
<td>77 ± 15</td>
<td>29 ± 6</td>
</tr>
</tbody>
</table>

Total background: 686 ± 41
Data: 697

- Combined all channels as a product of Poisson likelihoods

### Percentage of single top $tb+tqb$ selected events and S:B ratio

(white squares = no plans to analyze)

<table>
<thead>
<tr>
<th>Electron + Muon</th>
<th>1 jet</th>
<th>2 jets</th>
<th>3 jets</th>
<th>4 jets</th>
<th>≥ 5 jets</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 tags</td>
<td>10%</td>
<td>25%</td>
<td>12%</td>
<td>3%</td>
<td>1%</td>
</tr>
<tr>
<td>1 : 3,200</td>
<td>1 : 390</td>
<td>1 : 300</td>
<td>1 : 270</td>
<td>1 : 230</td>
<td></td>
</tr>
<tr>
<td>1 : 100</td>
<td>1 : 20</td>
<td>1 : 25</td>
<td>1 : 40</td>
<td>1 : 53</td>
<td></td>
</tr>
<tr>
<td>1 : 11</td>
<td>1 : 15</td>
<td>1 : 38</td>
<td>1 : 43</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Expected 62 single top events in 12 analysis channels
s+t cross section measurement (0.9 fb⁻¹)

- Measured $\sigma_{s+t}$
  - 4.9 $^{+1.4}_{-1.4}$ pb
  - 4.4 $^{+1.6}_{-1.4}$ pb
  - 4.8 $^{+1.6}_{-1.4}$ pb

- Expected $\sigma_{s+t}$
  - 2.7 $^{+1.6}_{-1.4}$ pb
  - 2.7 $^{+1.5}_{-1.5}$ pb
  - 2.8 $^{+1.6}_{-1.4}$ pb
D0 combination

- Combine results using BLUE (best linear unbiased estimator) method
- Determine correlations from pseudo-datasets

  - Measured $\sigma_{s+t}$
    (Expected $\sigma_{s+t}$)

    $\begin{bmatrix} 4.7 \pm 1.3 \text{ pb} \\ (3.0 \pm 1.3 \text{ pb}) \end{bmatrix}$

  - Measured significance
    [Expected significance]

    $0.00014 (3.6\sigma)$
    $[0.011 (2.3\sigma)]$
Other D0 measurements

- $V_{tb}$
  - no constraint on unitarity, # of generations
  - assuming $|V_{td}|^2 + |V_{ts}|^2 \ll |V_{tb}|^2$

- Allowed contours at different levels of confidence in $\sigma_s$ versus $\sigma_t$ plane
CDF Results

**CDF**

Expected uncertainty on $V_{ub}$ as a function of integrated luminosity.

The expected uncertainty is calculated assuming SM single top cross section. This plot is done with one of our analyses (matrix-element). The improvements in the new analysis translate in 15-20% in the expected sensitivity with respect to the old analysis.

---

**Color code:**

<table>
<thead>
<tr>
<th>Channel</th>
<th>Description (Link to web-page)</th>
<th>Measurement</th>
<th>Integrated Luminosity (pb$^{-1}$)</th>
<th>Publication</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lepton+jets <strong>New</strong></td>
<td>Neural Network</td>
<td>$s$+t channel = 2.0 +0.3-0.8 pb</td>
<td>2.2 fb$^{-1}$</td>
<td>02/27/2008 Conf. Note 9217</td>
</tr>
</tbody>
</table>
| Lepton+jets **New**      | Multivariate Likelihood Function| $s$+t channel = 1.8 +0.9 -0.8 pb

**: the expected uncertainty is calculated assuming SM single top cross section. This plot is done with one of our analyses (matrix-element). The improvements in the new analysis translate in 15-20% in the expected sensitivity with respect to the old analysis.

• CDF confirmed evidence using 1.5 fb\(^{-1}\)
  • Measured (expected) significance: 3.1 (3.0) \(\sigma\)
• New results using 2.2 fb\(^{-1}\)
  • Increased acceptance
    • extended muon coverage by adding new triggers
  • included 3-jet channel (besides 2-jet channel)
  • Improved performance of multivariate techniques
Likelihood method

- Likelihood functions built from several variables
- Kinematic variables, b-tag NN, t-channel ME, kinematic solver
Likelihood method

- Likelihood functions built from several variables
- Kinematic variables, b-tag NN, t-channel ME, kinematic solver

**Measured $\sigma_{s+t}$**

1.8 $^{+0.9}_{-0.8}$ pb
Neural networks

- 4 separate networks built in 2jet-1tag, 2jet-2tag, 3jet-1tag, and 3jet-2tag channels
- Train for t-channel in 1-tag events, and s-channel in 2-tag events
- Including b-tag NN, kinematic variables, angular correlations

\[ \sigma_{s+t} = 2.0^{+0.9}_{-0.8} \text{ pb} \]
Matrix element

- Include ttbar matrix element for both 2-jet and 3-jet events
- Include b-tag NN as weight in likelihood ratio

- Measured $\sigma_{s+t}$
Tevatron summary ($\sigma_{s+t}$)

Tevatron Single Top Summary

- Likelihood Function: CDF (2200 $pb^{-1}$) - $1.8 \pm 0.9$ ($0.8$)
- Matrix Element: CDF (2200 $pb^{-1}$) - $2.2 \pm 0.8$ ($0.7$)
- Neural Network: CDF (2200 $pb^{-1}$) - $2.0 \pm 0.9$ ($0.8$)
- Decision Tree: D0 (900 $pb^{-1}$) - $4.9 \pm 1.4$ ($1.4$)
- Matrix Element: D0 (900 $pb^{-1}$) - $4.8 \pm 1.6$ ($1.4$)
- Bayesian NN: D0 (900 $pb^{-1}$) - $4.4 \pm 1.6$ ($1.4$)
- BLUE Combination: D0 (900 $pb^{-1}$) - $4.7 \pm 1.3$ ($1.3$)

Tevatron projections

- Based on current measurement, CDF predicts 10% precision on $|V_{tb}|$ measurement at 3.5 fb$^{-1}$
Tevatron projections

- Based on current measurement, CDF predicts 10% precision on $|V_{tb}|$ measurement at 3.5 fb$^{-1}$

- Based on the Bayesian NN analysis, D0 predicts the following contours in $\sigma_s$ versus $\sigma_t$ plane at 95% CL and 68% CL with twice its data at 7 fb$^{-1}$
Conclusions

- The search for single top quark production is turning into measurements in the single top final state
  - Both experiments have seen 3 $\sigma$ evidence
  - First direct measurement of $|V_{tb}|$ performed
- Further improvements in progress
  - CDF combination
  - D0 update with larger dataset
- Several multivariate techniques, some new to our field, have been explored
  - Show similar performance in D0 analyses
Backup slides
### Systematics

**CDF Run II Preliminary**

<table>
<thead>
<tr>
<th>Systematic uncertainty</th>
<th>Range of Effect</th>
<th>Shape variations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jet energy scale</td>
<td>0...16%</td>
<td>✓</td>
</tr>
<tr>
<td>Initial state radiation</td>
<td>0...11%</td>
<td>✓</td>
</tr>
<tr>
<td>Final state radiation</td>
<td>0...15%</td>
<td>✓</td>
</tr>
<tr>
<td>Parton distribution</td>
<td>2...3%</td>
<td>✓</td>
</tr>
<tr>
<td>Monte Carlo generator</td>
<td>1...5%</td>
<td></td>
</tr>
<tr>
<td>Event detection efficiency</td>
<td>0...9%</td>
<td></td>
</tr>
<tr>
<td>Luminosity</td>
<td>6%</td>
<td></td>
</tr>
<tr>
<td>Neural net jet flavor</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>separator</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Mistag model</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Non-W model</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>$Q^2$ scale in Alpgen Monte Carlo</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Monte Carlo mismodeling</td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>

**D0**

**TABLE XI: Summary of the relative systematic uncertainties.** The ranges shown represent the different samples and channels.

<table>
<thead>
<tr>
<th>Relative Systematic Uncertainties</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Integrated luminosity</td>
<td>6%</td>
</tr>
<tr>
<td>$t\bar{t}$ cross section</td>
<td>18%</td>
</tr>
<tr>
<td>Electron trigger</td>
<td>3%</td>
</tr>
<tr>
<td>Muon trigger</td>
<td>6%</td>
</tr>
<tr>
<td>Primary vertex</td>
<td>3%</td>
</tr>
<tr>
<td>Electron reconstruction &amp; identification</td>
<td>2%</td>
</tr>
<tr>
<td>Electron track match &amp; likelihood</td>
<td>5%</td>
</tr>
<tr>
<td>Muon reconstruction &amp; identification</td>
<td>7%</td>
</tr>
<tr>
<td>Muon track match &amp; isolation</td>
<td>2%</td>
</tr>
<tr>
<td>Jet fragmentation</td>
<td>(5–7)%</td>
</tr>
<tr>
<td>Jet reconstruction and identification</td>
<td>2%</td>
</tr>
<tr>
<td>Jet energy scale</td>
<td>(1–20)%</td>
</tr>
<tr>
<td>Tag-rate functions</td>
<td>(2–16)%</td>
</tr>
<tr>
<td>Matrix-method normalization</td>
<td>(17–28)%</td>
</tr>
<tr>
<td>Heavy flavor ratio</td>
<td>30%</td>
</tr>
<tr>
<td>$\varepsilon_{\text{real}-e}$</td>
<td>2%</td>
</tr>
<tr>
<td>$\varepsilon_{\text{real}-\mu}$</td>
<td>2%</td>
</tr>
<tr>
<td>$\varepsilon_{\text{fake}-e}$</td>
<td>(3–40)%</td>
</tr>
<tr>
<td>$\varepsilon_{\text{fake}-\mu}$</td>
<td>(2–15)%</td>
</tr>
</tbody>
</table>
FIG. 3: The tag-rate functions (TRF’s) used to weight the MC events according to the probability that they should be $b$ tagged. In plots (a)–(d), the points show the neural network $b$ tagging algorithm (the “tagger”) applied directly to the MC events. The upper line that passes through the points is the result of the tag-rate functions, before scaling-to-data, being applied to the MC events to reproduce the result from the tagger. The lower line, with dotted error band, shows the tag-rate functions after they have been scaled to match the efficiency of the NN $b$ tagging algorithm applied to data. In plot (e), the lines show the (scaled) tag-rate functions that are applied to MC events.
D0 combination

TABLE I: Mean and square root of variance from the SM signal (2.9 pb) + background ensembles for the different analyses.

<table>
<thead>
<tr>
<th>Analysis</th>
<th>Mean $\sigma$ [pb]</th>
<th>$\sqrt{\text{Var}}$ $\Delta \sigma$ [pb]</th>
<th>$\sigma / \Delta \sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decision trees (DT)</td>
<td>2.9</td>
<td>1.61</td>
<td>1.8</td>
</tr>
<tr>
<td>Matrix elements (ME)</td>
<td>3.2</td>
<td>1.42</td>
<td>2.3</td>
</tr>
<tr>
<td>Bayesian neural networks (BNN)</td>
<td>2.7</td>
<td>1.48</td>
<td>1.8</td>
</tr>
<tr>
<td>Combined</td>
<td>3.0</td>
<td>1.28</td>
<td>2.3</td>
</tr>
</tbody>
</table>

TABLE II: The expected $p$-values and significances for the individual and the combined analyses, using the SM value of 2.9 pb for signal cross section as the reference point in Fig. 3.

<table>
<thead>
<tr>
<th>Analysis</th>
<th>Expected $p$-value</th>
<th>Expected significance [std. dev.]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decision trees (DT)</td>
<td>0.0177</td>
<td>2.1</td>
</tr>
<tr>
<td>Matrix elements (ME)</td>
<td>0.0307</td>
<td>1.9</td>
</tr>
<tr>
<td>Bayesian neural networks (BNN)</td>
<td>0.0155</td>
<td>2.2</td>
</tr>
<tr>
<td>Combined</td>
<td>0.0105</td>
<td>2.3</td>
</tr>
</tbody>
</table>
Matrix elements methodology

- Calculate probability density of an event resulting from a given process
  \[ P(p_1^L, p_1^R, p_2^L) = \frac{1}{\sigma} \int d\rho_{j1} d\rho_{j2} dp_1^L \sum_{comb} \phi_4 M(p_i^L)^2 \frac{f(q_1) f(q_2)}{|q_1| |q_2|} W_{jj}(E_{jet}, E_{part}) \]

  - Inputs: lepton and jet 4-vectors - no other information needed!
  - Parton distribution functions
  - Integrates over parton-level quantities
  - Transfer functions: Account for detector effects in measurement of jet energy
  - Matrix element: Different for each process. Leading order, obtained from MadGraph

- Uses full kinematic information of an event to discriminate signal events from background events
- Calculate probabilities for s- and t-channel, $Wbb$, $Wcj$, $Wgg$, and $tt$-bar (for three-jet events)
- Use matrix element probability densities to create a discriminant: signal / (signal + background)
- Neural networks are trained on Monte Carlo to discriminate signal from background
- A Bayesian neural network is a weighted average of many networks
- Protected against overtraining
Boosted decision trees methodology

- Start with large number of input variables (49)
- Optimize series of binary cuts in Monte Carlo
  - Automatically finds “interesting” variables
- Sort events by output purity
- Create series of “boosted” trees by reweighting based on misclassified events