Observation of Single Top Quark Production at the Tevatron

Ann Heinson
University of California, Riverside
for the CDF and DØ Collaborations

XXIth Rencontres de Blois: Windows on the Universe
Tuesday June 23, 2009
Single Top Quark Production

Electroweak production of single top quarks is predicted to occur at about half the rate of strong production of top-antitop pairs.

- **s-channel:** \( tb \)
  - \( m_{\text{top}} = 175 \text{ GeV} \)
    - \( \sigma_{\text{NLO}} = 0.88 \pm 0.11 \text{ pb} \)
  - \( m_{\text{top}} = 170 \text{ GeV} \)
    - \( \sigma_{(N)\text{NLO}} = 1.12 \pm 0.04 \text{ pb} \)

- **t-channel:** \( tqb \)
  - \( \sigma_{\text{NLO}} = 1.98 \pm 0.25 \text{ pb} \)
  - \( \sigma_{(N)\text{NLO}} = 2.34 \pm 0.12 \text{ pb} \)

Low rate at the Tevatron
(Important at the LHC)

- Observation of this process allows us to measure the CKM matrix element \( |V_{tb}| \)
- Can also study the \( Wtb \) coupling and search for heavy resonance production
- Same final state particles as for Higgs boson production \( (WH \rightarrow l\nu b\bar{b}) \), development of
  - accurate background modeling
  - powerful signal-background separation
  essential for Higgs boson discovery

- Observation analyses performed at \( m_{\text{top}} = 175 \text{ GeV} \) (CDF), 170 GeV (DØ), and assume SM \( tb:tqb \) ratio
- \( \sigma_{\text{NLO}} = \text{Sullivan, PRD 70, 114012 (2004)} \), \( \sigma_{(N)\text{NLO}} = \text{Kidonakis, PRD 74, 114012 (2006)} \)
Why Didn’t We See It Till Now?

- Predicted ~10 years before the discovery of the top quark in pair production
  - t-channel: Willenbrock and Dicus, PRD 34, 155 (1986)
  - s-channel: Cortese and Petronzio, PLB 253, 494 (1991)

- Observed 14 years after the top quark discovery (CDF and DØ, 1995)

- Single top ($tb+tqb$) has nearly half the $t\bar{t}$ cross section but S:B is 1:20 after selection compared with 5:1 for $t\bar{t}$ – backgrounds to single top are very difficult to deal with

- Signal is one isolated high-$p_T$ electron or muon and/or missing transverse energy from the $W$ decay, 2, 3, or 4 jets, and 1 or 2 $b$-tags
Backgrounds and Event Yields

Singlet Top Observation – Event Yields

<table>
<thead>
<tr>
<th></th>
<th>DØ 2.3 fb⁻¹</th>
<th>CDF 3.2 fb⁻¹</th>
<th>CDF 2.1 fb⁻¹</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lepton+#γ+jets / b-tagged</td>
<td>#γ+jets / b-tagged</td>
<td></td>
</tr>
<tr>
<td>(tb+\overline{t}qb) signal (*1, *2)</td>
<td>223 ± 30</td>
<td>191 ± 28</td>
<td>64 ± 10</td>
</tr>
<tr>
<td>(W+\text{jets})</td>
<td>2,647 ± 241</td>
<td>2,204 ± 542</td>
<td>304 ± 116 (*4)</td>
</tr>
<tr>
<td>(Z+\text{jets}, \text{dibosons})</td>
<td>340 ± 61</td>
<td>171 ± 15</td>
<td>171 ± 54</td>
</tr>
<tr>
<td>(t\overline{t} \text{ pairs } *1, *2, *3)</td>
<td>1,142 ± 168</td>
<td>686 ± 99</td>
<td>185 ± 30</td>
</tr>
<tr>
<td>(\text{Multijets})</td>
<td>300 ± 52</td>
<td>125 ± 50</td>
<td>679 ± 28 (*5)</td>
</tr>
<tr>
<td>Total prediction</td>
<td>4,652 ± 352</td>
<td>3,377 ± 505</td>
<td>1,403</td>
</tr>
<tr>
<td>Data</td>
<td>4,519</td>
<td>3,315</td>
<td>1,411</td>
</tr>
</tbody>
</table>

*1 DØ’s \(tb+\overline{t}qb\) signal and \(t\overline{t}\) background use \(m_{\text{top}} = 170\) GeV (and signal \(\sigma_{\text{NLO}}\))

*2 CDF’s \(tb+\overline{t}qb\) signal and \(t\overline{t}\) background use \(m_{\text{top}} = 175\) GeV (and signal \(\sigma_{\text{NLO}}\))

*3 DØ’s analysis includes 4-jet events, so the \(t\overline{t}\) yield is higher

*4 CDF’s \#γ+jets channel \(W+\text{jets}\) yield does not include \(Wj\) where \(j\) = light jet

*5 CDF’s \#γ+jets channel Multijets yield includes \(Wj\) events
# Development of Analysis Strategy

## Searches, upper limits

<table>
<thead>
<tr>
<th>Collaboration</th>
<th>Journal</th>
<th>Volume</th>
<th>Year</th>
<th>Luminosity</th>
<th>Methodology</th>
</tr>
</thead>
<tbody>
<tr>
<td>DØ</td>
<td>PRD 63</td>
<td>031101</td>
<td>2000</td>
<td>0.09 fb⁻¹</td>
<td>Cuts search</td>
</tr>
<tr>
<td>DØ</td>
<td>PLB 517</td>
<td>282</td>
<td>2001</td>
<td>0.09 fb⁻¹</td>
<td>Neural networks (28 variables)</td>
</tr>
<tr>
<td>CDF</td>
<td>PRD 65</td>
<td>091102</td>
<td>2002</td>
<td>0.11 fb⁻¹</td>
<td>Cuts + likelihood fit to $H_T$</td>
</tr>
<tr>
<td>CDF</td>
<td>PRD 69</td>
<td>052003</td>
<td>2004</td>
<td>0.11 fb⁻¹</td>
<td>Neural networks (18 variables) + likelihood fit</td>
</tr>
<tr>
<td>CDF</td>
<td>PRD 71</td>
<td>012005</td>
<td>2005</td>
<td>0.16 fb⁻¹</td>
<td>Cuts + likelihood fit to $Q\times\eta$</td>
</tr>
<tr>
<td>DØ</td>
<td>PLB 622</td>
<td>265</td>
<td>2005</td>
<td>0.23 fb⁻¹</td>
<td>Neural networks (25 variables) + Bayesian likelihoods</td>
</tr>
<tr>
<td>DØ</td>
<td>PRD 75</td>
<td>092007</td>
<td>2007</td>
<td>0.23 fb⁻¹</td>
<td>Cuts, Neural networks (25 variables) + Bayes. l’hoods</td>
</tr>
</tbody>
</table>

## >3σ Evidence

<table>
<thead>
<tr>
<th>Collaboration</th>
<th>Journal</th>
<th>Volume</th>
<th>Year</th>
<th>Luminosity</th>
<th>Methodology</th>
</tr>
</thead>
<tbody>
<tr>
<td>DØ</td>
<td>PRL 98</td>
<td>181802</td>
<td>2007</td>
<td>0.9 fb⁻¹</td>
<td>Boosted decision trees (49 variables), Bayesian NNs, Matrix elements, + Bayesian likelihoods</td>
</tr>
<tr>
<td>DØ</td>
<td>PRD 78</td>
<td>012005</td>
<td>2008</td>
<td>0.9 fb⁻¹</td>
<td>Boosted decision trees (49 variables), Bayesian NNs, Matrix elements, + BLUE combination + Bayesian likelihoods</td>
</tr>
<tr>
<td>CDF</td>
<td>PRL 101</td>
<td>252001</td>
<td>2008</td>
<td>2.2 fb⁻¹</td>
<td>Neural networks (18 variables), Likelihoods, Matrix elements, + NN-combination + Bayesian likelihoods</td>
</tr>
</tbody>
</table>

## 5σ Observation

<table>
<thead>
<tr>
<th>Collaboration</th>
<th>ArXiv</th>
<th>Year</th>
<th>Luminosity</th>
<th>Methodology</th>
</tr>
</thead>
<tbody>
<tr>
<td>DØ</td>
<td>arXiv:0903.0850</td>
<td>2009</td>
<td>2.3 fb⁻¹</td>
<td>Boosted decision trees (64 variables), Bayesian NNs, Matrix elements, + Bayesian-NN combination + Bayesian likelihoods</td>
</tr>
<tr>
<td>CDF</td>
<td>arXiv:0903.0885</td>
<td>2009</td>
<td>3.2 fb⁻¹</td>
<td>Boosted decision trees (22 variables), Neural networks (18 variables), Likelihoods, Matrix elements, + NN-combination + Bayesian likelihoods</td>
</tr>
</tbody>
</table>
Analysis Strategy Visualized

Data from all possible triggers

Select signal-like events

Check data is reproduced in all variables

Separate signal from background

~100 million pseudo-datasets

Monte Carlo for signal and backgrounds

Measure significance

Cross section from Bayesian likelihood

Combine results

Boosted Decision Trees

Matrix Elements

Neural Networks
Checking the Background Model

Cross-check samples with mostly $W^+$ light-jets, mostly $W^+$ jets, and mostly top pairs.

Background model must match data for hundreds of variables in 24 independent analysis channels.
Checking the Background Model

Background model must match data for scores of variables in all 9 independent analysis channels.

Also check all variables in cross-check samples with mostly $W+\text{jets}$ and mostly top pairs.

CDF Run II 2.2 fb$^{-1}$

CDF Run II 3.2 fb$^{-1}$

2 jets 0 tags

4 jets 1,2 tags
Systematic Uncertainties

Total error on $tb+tqb$ cross section is ±22%
Statistics-only error is ±18%
So systematics contribute ±13%
Separate Signal from Background

**Boosted Decision Trees**

DØ and CDF

Apply sequential cuts, keep events that fail

Boosting: averages over many trees, improves performance by ~20%

**Neural Networks**

CDF: one network for each analysis channel

DØ: average over 100 networks for each analysis channel (“Bayesian NNs”)

**Matrix Elements**

DØ and CDF

One likelihood for each signal and background

Measure \( tb + tqb \) and s-channel \( tb \) separately

**Likelihoods**

CDF only

One likelihood for each signal and background

Measure \( tb + tqb \) and s-channel \( tb \) separately

\[
p_{ik} = \frac{f_{ij,k}}{\sum_{m=1}^{5} f_{ij,m}},
\]

\[
L_k(\{x_i\}) = \frac{\prod_{i=1}^{5} \prod_{l=1}^{n_{var}} p_{ik}}{\sum_{m=1}^{5} \prod_{l=1}^{n_{var}} p_{ik}}.
\]
Combine Analyses

CDF have 5 analyses, DØ have 3, for l+jets

Not 100% correlated, so combining the results can improve the expected significance

NEAT Neural Networks

“NEAT = Neuro-evolution of augmenting topologies”

Eight networks, one per analysis channel

Training optimized on expected significance

Improves expected significance by 13% over best individual analyses (BDT and NN)

Bayesian Neural Networks

24 sets of networks, 1 per analysis channel

Markov Chain MC technique used with 500 networks, average over the last 100

Improves expected significance by 4% over best individual analysis (BDT)
Single Top Observation – Results

<table>
<thead>
<tr>
<th>Single Top Cross Section</th>
<th>Signal Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DØ</strong> 2.3 fb⁻¹</td>
<td></td>
</tr>
<tr>
<td>arXiv:0903.0850</td>
<td></td>
</tr>
<tr>
<td>$m_{\text{top}} = 170$ GeV</td>
<td></td>
</tr>
<tr>
<td>3.94 ± 0.88 pb</td>
<td>4.5 σ</td>
</tr>
<tr>
<td><strong>CDF</strong> 3.2 fb⁻¹</td>
<td></td>
</tr>
<tr>
<td>arXiv:0903.0885</td>
<td></td>
</tr>
<tr>
<td>$m_{\text{top}} = 175$ GeV</td>
<td></td>
</tr>
<tr>
<td>2.3 $^{+0.6}_{-0.5}$ pb</td>
<td>&gt;5.9 σ</td>
</tr>
</tbody>
</table>

CDF Run II 3.2 fb⁻¹

Events

Normalized to Prediction

Neural Network Output

S/B Ratio

DØ 2.3 fb⁻¹
Single Top Signal Plots

CDF’s signal is red

DØ’s signal is blue

Signal normalized to measured cross section

2 Jets 1 Tag 

MC normalized to SM prediction

M (tvb) [GeV/c^2]

Q (lep) • η (l-jet)

Yield [Events]

Top Quark Mass [GeV]

Yield [Events/0.8]

Q(lep) • η(light-quark jet)
CKM Matrix Element $V_{tb}$

- $\sigma(tb, tqb) \propto |V_{tb}|^2$ \therefore calculate a posterior in $|V_{tb}|^2$
- Assume 1. SM top quark decay: $V_{td}^2 + V_{ts}^2 \ll V_{tb}^2$
  2. Pure $V$–$A$ coupling with $CP$ conservation
- No need to assume only three quark families or CKM matrix unitarity
- $|V_{tb}| > 0.71$ (CDF), $|V_{tb}| > 0.78$ (DØ), at 95% CL

**Additional Systematic Uncertainties for the $|V_{tb}|$ Measurement**

- DØ 2.3 fb$^{-1}$
- For the $tb+tqb$ theory cross section
  - Top quark mass: 4.2%
  - Parton distribution functions: 3.0%
  - Factorization scale: 2.4%
  - Strong coupling $\alpha_s$: 0.5%

NB CDF and DØ use different theory cross sections and top quark mass points. Combination of $|V_{tb}|$ results is in progress.
Challenging measurements
- small signal hidden in large complex background

It took 14 years since top quark pair observation to collect enough data and develop the background models and analysis techniques to be able to achieve this

Single Top Observation Summary

- 5σ observation of single top quark production!