Measurement of Single Top Production at the Tevatron

Dominic Hirschbühl
University of Karlsruhe
On behalf of the and Collaborations

DIS08  09.04.2008  London
Proton-Antiproton-Collisions with a rate of 1.7 MHz
Center-of-mass energy: $\sqrt{s} = 1.96 \text{ TeV}$

Record luminosity: $\mathcal{L} = 3.2 \times 10^{32} \text{ cm}^{-2}\text{s}^{-1}$

World Record for hadron colliders!
Top quark production – Standard Model

At the Tevatron, top quarks are primarily produced in pairs via the strong interaction.

\[ \sigma_{tt} = 6.7 \pm 0.8 \text{ pb} \]

Top quark production via the weak interaction:

- **t-channel**
- **s-channel**

\[ \sigma_t = 2.0 \pm 0.3 \text{ pb} \]
\[ \sigma_s = 0.9 \pm 0.1 \text{ pb} \]


assumed top quark mass

\[ m_{\text{top}} = 175 \text{ GeV} \]
Why measure single top?

- Source of single ~100% polarized top quarks:
  - Test V-A structure of Wtb vertex
  - Access to the top quark spin

- Test of the SM prediction. Does it exist?
  - Cross section $\propto |V_{tb}|^2$
    - allows direct measurement of $|V_{tb}|$
  - Test unitarity of the CKM matrix
  - Hints for existence of a 4th generation?
  - Vector t' is also a possibility

- Test of $b$ quark structure function: DGLAP evolution
- Can test CP-violation – single t vs. single tbar
Sensitivity to new physics

- Same final state signature as Higgs: WH, H \rightarrow bb\bar{b}. Understanding single-top backgrounds is a prerequisite for Higgs searches at the Tevatron. Same tools can be applied for Higgs searches.

Test non-SM phenomena
- Search W' or H^+ (s-channel signature)
- Search for FCNC, e.g. ug \rightarrow t
- ...

$$\sigma_{WH} \sim \frac{1}{10} \sigma_{\text{single top}}$$
Event signature and selection

Event Selection:

- 1 lepton (electron / muon)
  - $E_T > 20$ GeV, $|\eta| < 2.0$
  - electron: $E_T > 15$ GeV, $|\eta_e| < 1.1$
  - muon: $p_T > 18$ GeV, $|\eta_\mu| < 2.0$

- Missing $E_T$
  - MET > 25 GeV
  - MET > 15 GeV

- 2 - 4 jets
  - $E_T > 20$ GeV, $|\eta| < 2.8$
  - $p_T > 15$ GeV, $|\eta| < 3.4$
  - 1. jet: $p_T > 25$ GeV, $|\eta| < 2.5$
  - 2. jet: $p_T > 20$ GeV

Main Backgrounds:

W/Z+heavy flavor jets, Mistags, ttbar, Multijet events
Why it is challenging?

Single top hidden behind background uncertainty! \((s \ll \sigma_{bkg,syst})\)

⇒ Makes counting experiment impossible!

⇒ \(S/\sqrt{B}\) is an inadequate estimate of sensitivity
Improved b jet identification

About 50% of the background in the W + 2 jets sample do NOT contain b quarks even though a secondary vertex was required!

Jet and track variables, e.g. vertex mass, decay length, track multiplicity, ...

⊕

neural network

↓

powerful discriminant

Fit to NN output for W + 2 jets events with one secondary vertex

~20% gain in sensitivity for all three analyses
Search strategy

Combined Search

- t-channel and s-channel regarded as one single top signal.
- Cross section ratio is fixed to SM value.

Signal Model
- MadEvent + Pythia
- CompHep + Pythia

Background Model
- MC & Data

Event Selection

Matrix elements
- Neural Networks

Multivariate Analysis
- Likelihood discriminants
  \[ L_k(x_i) = \frac{\prod_{i=1}^{n_{var}} p_{ik}}{\sum_{m=1}^{5} \prod_{i=1}^{n_{var}} p_{im}} \]

Boosted Decision Trees

Data

Extraction of cross section

Ensemble tests
Matrix Element Analysis

Idea: Compute an event probability $P$ for signal and background hypotheses:

\[
P(p_1^\mu, p_{j1}^\mu, p_{j2}^\mu) = \frac{1}{\sigma} \int d\rho_{j1} d\rho_{j2} dp_v \sum_{\text{comb}} \phi_4 |M(p_i^\mu)|^2 \frac{f(q_1) f(q_2)}{|q_1||q_2|} W_{\text{jet}}(E_{\text{jet}}, E_{\text{part}})
\]

- Use full kinematic information of an event
- Calculate probability densities for
  - $s$- and $t$-channel single top production
  - Main background contributions
- Create Discriminant
  \[
  \text{EPD} = \frac{b \cdot P_{\text{signal}}(\hat{x})}{b \cdot P_{\text{signal}}(\hat{x}) + b \cdot P_{b\text{-bkg}}(\hat{x}) + (1-b) P_{\text{nonb-bkg}}(\hat{x})}
  \]
  \[
  D(\hat{x}) = \frac{P_{\text{signal}}(\hat{x})}{P_{\text{signal}}(\hat{x}) + P_{\text{background}}(\hat{x})}
  \]
Use large number of input variables (49)
- Non-discriminating variables are automatically ignored, but don't degrade the performance
- Optimize series of binary cuts with training sample
- Calculate for each leaf purity \( p = \frac{s}{s+b} \)
- Sort events by output purity
- Create series of “boosted“ trees by reweighting based on value of misclassification

Idea: Effective extension of a cut-based analysis
Neural Network Analysis

- uses 10-15 variables
- Network: NeuroBayes

- uses 18-25 variables (subset of those in the Decision Tree analysis)
- Bayesian Neural Network
  \( \rightarrow \) weighted average over many networks
Multivariate Likelihood Analysis

- Multivariate binned likelihood combines several sensitive variables into a single variable
- Pioneered at LEP
- Seven variables used in the 2 jet bin
- Ten variables used in the 3 jet bin

\[ L(x) = \frac{\prod_{i=1}^{n_{\text{var}}} p_{i}^{\text{sig}}(x_i)}{\prod_{i=1}^{n_{\text{var}}} p_{i}^{\text{bkg}}(x_i) + \prod_{i=1}^{n_{\text{var}}} p_{i}^{\text{bkg}}(x_i)} \]
Measured cross sections

Matrix element method  Bayesian neural network  Boosted decision tree

Measured using a Bayesian binned likelihood calculation

\[ \sigma_{t+s} = 4.8^{+1.6}_{-1.4} \text{ pb} \]

\[ \sigma_{t+s} = 4.4^{+1.6}_{-1.4} \text{ pb} \]

\[ \sigma_{t+s} = 4.9^{+1.4}_{-1.4} \text{ pb} \]

SM prediction : \[ \sigma_{t+s} = 2.9^{+0.4}_{-0.4} \text{ pb} \]
Sensitivity

Matrix element method
Bayesian neural networks
Boosted decision trees

Determined using ensemble tests without signal contribution

Expected p-value: 3% (1.9 \( \sigma \))
Observed p-value: 0.081% (3.2 \( \sigma \))

Expected p-value: 1.6% (2.2 \( \sigma \))
Observed p-value: 0.08% (3.1 \( \sigma \))

Expected p-value: 1.9% (2.1 \( \sigma \))
Observed p-value: 0.035% (3.4 \( \sigma \))
Combination & Limit on $|V_{tb}|$

Combine results with best linear unbiased estimator (BLUE) method
Correlations ~60%

- Derive a limit on $|V_{tb}|$ based on boosted decision tree result
- Assume $|V_{td}|^2 + |V_{ts}|^2 \ll |V_{tb}|^2$

$t$-channel cross section [pb]

$D\bar{O}$ 0.9 fb$^{-1}$
- 1 std. dev.
- 2 std. dev.
- 3 std. dev.

SM

$s$-channel cross section [pb]

Measured cross section:

$\sigma_{t+s} = 4.7 \pm 1.3$ pb

Expected p-value: 1.1% (2.3 $\sigma$)

Observed p-value: 0.014% (3.6 $\sigma$)

$|V_{tb}| > 0.68 @ 95\% C.L.$

Measured cross sections

**Measured using a binned likelihood fit**

\[
\sigma_{t+s} = 1.8^{+0.9}_{-0.8} \text{ pb}
\]

\[
\sigma_{t+s} = 2.2^{+0.8}_{-0.7} \text{ pb}
\]

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

**SM prediction:** \[
\sigma_{t+s} = 2.9 \pm 0.4 \text{ pb}
\]
Sensitivity

Determined using a hypothesis test

\[ Q = -2 \left( \ln L_{\text{red.}}(\beta_1 = 1) - \ln L_{\text{red.}}(\beta_1 = 0) \right) \]

- **Likelihood Discriminant**
  - Matrix element method
  - Neural networks

**Expected p-value:**
- 0.03\% (3.4 \( \sigma \))
- 0.0003\% (4.5 \( \sigma \))
- 0.0005\% (4.4 \( \sigma \))

**Observed p-value:**
- 2.5\% (2.0 \( \sigma \))
- 0.03\% (3.4 \( \sigma \))
- 0.06\% (3.2 \( \sigma \))
Two approaches to combine results:

- **Extended BLUE method**
  - Uses results and correlations as input
  - Correlations ~60-70%

- **Superdiscriminant (NEAT neural network)**
  - "Neuro-Evolution of Augmenting Topologies"
  - Uses output of individual analysis as input
  - Candidate networks compete against each other.
  - Automatically optimizes
    - Network topology, weights, output histogram binning, includes systematic errors in optimization procedure

**NEAT**

- Measured cross section
  \[ \sigma_{t+s} = 2.1^{+0.7}_{-0.6} \text{ pb} \]
- Expected p-value: 5.1 \( \sigma \)
- Observed p-value: 3.7 \( \sigma \)
- \(|V_{tb}| > 0.66@95 \text{ C.L.}\)

**BLUE**

- Measured cross section
  \[ \sigma_{t+s} = 2.2 \pm 0.7 \text{ pb} \]
- Expected p-value: 4.7 \( \sigma \)
- Observed p-value: 3.7 \( \sigma \)

\(\sigma\) 

\(|V_{tb}|\) 

CDF Run II Preliminary, \( L = 2.2 \text{ fb}^{-1} \)

\(|V_{tb}| > 0.66 \text{ (95\% C.L.)}\)
### Tevatron Single Top Summary

<table>
<thead>
<tr>
<th>Method</th>
<th>CDF 2200 pb⁻¹</th>
<th>DØ 500 pb⁻¹</th>
<th>Combination: CDF 2200 pb⁻¹</th>
<th>Combination: DØ 500 pb⁻¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likelihood Function</td>
<td>1.8 ± 0.9</td>
<td>4.9 ± 1.4</td>
<td>2.2 ± 0.7</td>
<td>4.4 ± 1.6</td>
</tr>
<tr>
<td>Matrix Element</td>
<td>2.2 ± 0.8</td>
<td>4.8 ± 1.6</td>
<td>2.2 ± 0.7</td>
<td>4.4 ± 1.6</td>
</tr>
<tr>
<td>Neural Network</td>
<td>2.0 ± 0.9</td>
<td></td>
<td>2.0 ± 0.9</td>
<td>4.4 ± 1.6</td>
</tr>
<tr>
<td>Combination: CDF</td>
<td></td>
<td></td>
<td>2.2 ± 0.7</td>
<td>4.4 ± 1.6</td>
</tr>
</tbody>
</table>

**Note:**
- CDF: Collider Detector at Fermilab (2200 pb⁻¹)
- DØ: DØ Collaboration (500 pb⁻¹)
- SM?: Standard Model

*Z. Sulikin, PRD 76, 112012 (2007)*
Background composition

**W+HF jets (Wbb/Wcc/Wc)**

W+jets normalization from data and heavy flavor (HF) fractions from ALPGEN Monte Carlo, calibrated in generic multijet data.

**Mistags (W+2jets)**

- Falsely tagged light quark or gluon jets
- Mistag probability parameterization obtained from inclusive jet data

**Non-W (QCD)**

- Multijet events with semileptonic $b$-decays or mismeasured jets
- Fit low missing $E_T$ data and extrapolate into signal region

**Top/EWK (WW/WZ/Z→bb, ttbar)**

- MC normalized to theoretical cross-section
# Event Yield

<table>
<thead>
<tr>
<th></th>
<th>CDF Run II Preliminary Predicted 2-jet event yield with 2.2 fb⁻¹</th>
<th>CDF Run II Preliminary Predicted 3-jet event yield with 2.2 fb⁻¹</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>s-channel</td>
<td>41.2 ± 5.9</td>
</tr>
<tr>
<td></td>
<td>t-channel</td>
<td>62.1 ± 9.1</td>
</tr>
<tr>
<td></td>
<td>Single top</td>
<td>103.3 ± 15.0</td>
</tr>
<tr>
<td></td>
<td>(tt)</td>
<td>146.0 ± 20.9</td>
</tr>
<tr>
<td></td>
<td>Diboson</td>
<td>63.2 ± 6.3</td>
</tr>
<tr>
<td></td>
<td>(Z + \text{jets})</td>
<td>26.7 ± 3.9</td>
</tr>
<tr>
<td></td>
<td>(W + \text{bottom})</td>
<td>461.6 ± 139.1</td>
</tr>
<tr>
<td></td>
<td>(W + \text{charm})</td>
<td>395.0 ± 121.8</td>
</tr>
<tr>
<td></td>
<td>(W + \text{light})</td>
<td>339.8 ± 56.1</td>
</tr>
<tr>
<td></td>
<td>Non-W</td>
<td>59.5 ± 23.8</td>
</tr>
<tr>
<td></td>
<td>Total background</td>
<td>1491.8 ± 268.6</td>
</tr>
<tr>
<td></td>
<td>Total prediction</td>
<td>1595.1 ± 269.0</td>
</tr>
<tr>
<td></td>
<td>Observed</td>
<td>1535</td>
</tr>
</tbody>
</table>
# Event Yield

**Event Yields in 0.9 fb\(^{-1}\) Data**

Electron+muon, 1tag+2tags combined

<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\bar{t} \to ll)</td>
<td>39 ± 9</td>
<td>32 ± 7</td>
<td>11 ± 3</td>
</tr>
<tr>
<td>(t\bar{t} \to 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>
<tr>
<td>Total background</td>
<td>686 ± 41</td>
<td>460 ± 39</td>
<td>253 ± 38</td>
</tr>
<tr>
<td>Data</td>
<td>697</td>
<td>455</td>
<td>246</td>
</tr>
</tbody>
</table>
Search for $W' \rightarrow tb$ events

- $W'$ occurs in some extensions of the SM with higher symmetry.
- Complementary to searches in $W' \rightarrow ev/\mu\nu$ (e.g. $W'$ of leptophobic nature).
- Same event selection and background estimate as single top analysis.
- Use $M(lvjj)$ as discriminant
- Neglect interference with SM $W$ boson.
Observe no evidence for resonant $W'$ production.


- $M(W') > 800$ GeV if $M(W'_R) > M(\nu_R)$
- $M(W') > 825$ GeV if $M(W'_R) < M(\nu_R)$
- $M(W') > 731$ GeV if $M(W'_R) > M(\nu_R)$
- $M(W') > 739$ GeV if $M(W'_R) < M(\nu_R)$