Single top quark physics

Peter Dong, UCLA
on behalf of the CDF and D0 collaborations

Les Rencontres de Physique de la Vallee d’Aoste

Wednesday, February 27, 2008
The Tevatron

- World’s highest-energy operational particle accelerator
- Two multi-purpose particle detectors, CDF and DØ
- The only place in the world to study top quarks (for now)
Electroweak single top production

- Top quarks are predominantly produced in pairs by the strong force.
  - Can study properties like mass, charge, $W$ helicity.

Electroweak single top production allows in
- Direct measurement of CKM matrix element $|V_{tb}|$.
- Chance to measure polarized top quarks.

Quoted cross-sections at $M_{top} = 175$ GeV/c$^2$ — Cacciari et al. JHEP 0404:068; Z. Sullivan, hep-ph/0408049
Sensitivity to new physics

- Single top has similar experimental challenges to WH
  - WH is preferred channel for low-mass Higgs production at the Tevatron
  - Have to find single top to find Higgs in this channel!

Also sensitive to heavy $W'$, flavor-changing neutral currents, Kaluza-Klein heavy $W^{kk}$, ...
Experimental challenge

- Top quark pairs produce a final state with a lepton, missing transverse energy, and four jets — a distinct signature
  - Signal : background of about 3.4 : 1

- Single top production, with a smaller cross section and two fewer jets, is much harder to find
  - Signal : background of about 1 : 15
  - A simple counting experiment is impossible!
Although searching for something so tiny may seem hopeless...

...We are still sensitive to the effects if we know what to look for
General analysis method

- Data
- Analysis Event Selection
- Apply MC Corrections
- Signal/Background Monte Carlo
- Any technique to separate signal from background:
  - Decision tree (D0)
  - Likelihood function (CDF)
  - Matrix element (CDF, D0)
  - Neural network (CDF)
  - Bayesian NN (D0)

- Result
- Cross Section
- Statistical analysis
- Discriminant
  - Signal
  - Background

Cross Section: 68%
D0 single top search in 910 pb$^{-1}$
Event selection and background estimate

<table>
<thead>
<tr>
<th>Electron + Muon</th>
<th>1 jet</th>
<th>2 jets</th>
<th>3 jets</th>
<th>4 jets</th>
<th>( \geq 5 ) jets</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 tags</td>
<td>10%</td>
<td>25%</td>
<td>12%</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1 : 3,200</td>
<td>1 : 390</td>
<td>1 : 300</td>
<td>1 : 270</td>
<td></td>
</tr>
<tr>
<td>1 tag</td>
<td>6%</td>
<td>21%</td>
<td>11%</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1 : 100</td>
<td>1 : 20</td>
<td>1 : 25</td>
<td>1 : 27</td>
<td></td>
</tr>
<tr>
<td>2 tags</td>
<td>3%</td>
<td>2%</td>
<td>0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1 : 11</td>
<td>1 : 15</td>
<td>1 : 38</td>
<td>1 : 43</td>
<td></td>
</tr>
</tbody>
</table>

Percentage of single top \( tb+tqb \) selected events and S:B ratio (white squares = no plans to analyze)

- Exactly one isolated lepton
- Large missing transverse energy
- Two to four jets
- At least one \( b \)-tagged jet
  - Neural net tagger increases acceptance
  - Apply tag rate functions, derived in data, to Monte Carlo

Estimate large \( W + \text{jets} \) background by applying heavy flavor fraction to pretagged data, then applying tagging rate functions.

Use data events with non-isolated leptons as multijet model.
Matrix element method

- Calculate probability density of an event resulting from a given process

\[
P(p_{i1}, p_{j1}, p_{j2}) = \frac{1}{\sigma} \int dp_{j1} dp_{j2} dp_{v} \sum_{\text{comb}} \phi_4 |M(p_{i1})|^2 \frac{f(q_1)}{|q_1|} \frac{f(q_2)}{|q_2|} W_{\text{jet}}(E_{\text{jet}}, E_{\text{part}})
\]

- 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)

Inputs:
- lepton and jet 4-vectors - no other information needed!

Parton distribution functions
- Matrix element: Different for each process. Leading order, obtained from MadGraph
- Transfer functions: Account for detector effects in measurement of jet energy

Integrate over parton-level quantities
Bayesian neural network

- 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

Input Nodes → Hidden Nodes → Output Node

Protected against overtraining

Wednesday, February 27, 2008

Peter Dong, UCLA
Boosted decision tree

- 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
**D0 results**

Matrix element method

Bayesian neural network

Boosted decision tree

**Measured cross section:**

- **Matrix element method:** $4.8^{+1.6}_{-1.4} \text{ pb}$
- **Bayesian neural network:** $4.4^{+1.6}_{-1.4} \text{ pb}$
- **Boosted decision tree:** $4.9^{+1.4}_{-1.4} \text{ pb}$
**D0 combination**

- Results are not fully correlated — can gain sensitivity by combining
- Combine results with best linear unbiased estimator (BLUE) method

Measured cross section:

\[ 4.7 \pm 1.3 \text{ pb} \]

Expected sensitivity:

\[ 0.011 \ (2.3\sigma) \]

Observed p-value:

\[ 0.00014 \ (3.6\sigma) \]
Derive a limit on $|V_{tb}|$ based on boosted decision tree result

Assume that $|V_{td}|^2 + |V_{ts}|^2 << |V_{tb}|^2$

$|V_{tb}| > 0.68$ at 95% C.L.
CDF single top search
CDF search in 1.5 fb$^{-1}$

- In summer 2007, CDF also reported evidence for single top production with a two analysis methods: a matrix element method and a multivariate likelihood function
- D0 and CDF use similar approaches for the matrix element
- Expected sensitivity of 3.0$\sigma$; observed significance of 3.1$\sigma$
Jet flavor separator

- CDF analyses use a neural net to separate jets from $b$-quarks from light-quark jets.
- Produces a continuous output – not a cut.
- Improves sensitivity by 10–20%!
Multivariate likelihood function

- Multivariate binned likelihood combines several sensitive variables into a single discriminant
- Pioneered at LEP
- Seven variables in this analysis

$$N_{\text{sig}}$$

$$N_{\text{bkg}}$$

$$P_i^{\text{sig}} = \frac{N_i^{\text{sig}}}{N_i^{\text{sig}} + N_i^{\text{bkg}}}$$

$$L(\vec{x}) = \frac{\prod_{i=1}^{n_{\text{var}}} p_i^{\text{sig}}(x_i)}{\prod_{i=1}^{n_{\text{var}}} p_i^{\text{sig}}(x_i) + \prod_{i=1}^{n_{\text{var}}} p_i^{\text{bkg}}(x_i)}$$
CDF results with 1.5 fb$^{-1}$

Likelihood function

- Measured cross section: $2.7^{+1.3}_{-1.1}$ pb

Matrix element

- Measured cross section: $3.0^{+1.2}_{-1.1}$ pb
More improvements

- Want to improve precision of measurement as much as possible
  - Add more data (another 700 pb$^{-1}$)
  - Update neural network analysis
  - Increase acceptance
  - Improve separation of signal and background

- These will reduce expected uncertainty on the single top cross section, and thus $|V_{tb}|$
Neural network

- Update neural network measurement last used for 955 pb\(^{-1}\)
- Train separate networks for different tag multiplicities, based on the \(b\)-quark composition of each signal sample:
  - Train for t-channel for single-tagged events
  - Train for s-channel for double-tagged events
Extended muon coverage

- Increase acceptance by taking muons from complementary trigger (missing $E_T$ plus two jets)
- New muons “fill in the cracks” left by triggered muons
- Improve overall sensitivity by 7%!
Three-jet events

- Can recover another ~20% of signal from three-jet events
- Three-jet events have large tt-bar background — much more difficult to separate signal from background
- Sensitivity improvement of ~3%
Acceptance improvement

- Acceptance increases gives us more signal
- Luminosity scaling would give 88 expected events
- Now 125 single top events expected – 40% gain in signal acceptance!

CDF Run II Preliminary
Predicted event yield with 2.2 fb⁻¹

<table>
<thead>
<tr>
<th>Process</th>
<th>2 jets</th>
<th>3 jets</th>
</tr>
</thead>
<tbody>
<tr>
<td>s-channel</td>
<td>41.2 ± 5.9</td>
<td>13.5 ± 1.9</td>
</tr>
<tr>
<td>t-channel</td>
<td>62.1 ± 9.1</td>
<td>18.3 ± 2.7</td>
</tr>
<tr>
<td>Single top</td>
<td>103.3 ± 15.0</td>
<td>21.8 ± 4.6</td>
</tr>
<tr>
<td>tt</td>
<td>146.0 ± 20.9</td>
<td>338.7 ± 48.2</td>
</tr>
<tr>
<td>Diboson</td>
<td>63.2 ± 6.3</td>
<td>21.5 ± 2.2</td>
</tr>
<tr>
<td>Z + jets</td>
<td>26.7 ± 3.9</td>
<td>11.0 ± 1.6</td>
</tr>
<tr>
<td>W + bottom</td>
<td>461.6 ± 139.1</td>
<td>141.1 ± 42.6</td>
</tr>
<tr>
<td>W + charm</td>
<td>395.0 ± 121.8</td>
<td>108.8 ± 33.5</td>
</tr>
<tr>
<td>W + light</td>
<td>339.8 ± 56.1</td>
<td>101.8 ± 16.9</td>
</tr>
<tr>
<td>Multijet</td>
<td>59.5 ± 23.8</td>
<td>21.3 ± 8.5</td>
</tr>
<tr>
<td>Total background</td>
<td>1491.8 ± 268.6</td>
<td>754.8 ± 91.3</td>
</tr>
<tr>
<td>Total prediction</td>
<td>1595.1 ± 269.0</td>
<td>776.6 ± 91.4</td>
</tr>
<tr>
<td>Observed</td>
<td>1535</td>
<td>712</td>
</tr>
</tbody>
</table>
tt-bar matrix element

- Added $tt$-bar matrix element to matrix element method
- Assume a final-state $W$ that has been missed (whether it decays leptonically or hadronically)
- Integrate over the unmeasured $W$ particle to extract $tt$-bar probability
- Improves sensitivity by about 10%!
Sensitivity improvement

- Several small improvements add up
- Expected uncertainty on $|V_{tb}|$ for matrix element analysis increases by ~12% compared to simple luminosity scaling
- Equivalent to a 56% increase in data!

Matrix Element Analysis, CDF Run II Preliminary

![Graph showing expected uncertainties for matrix elements](image)
CDF results

CDF Run II Preliminary, L=2.2 fb⁻¹

- Single top
- b-like
- c-like
- Mistags
- tt-bar

Measured cross section:
- 1.8 ± 0.9 pb
- 2.0 ± 0.9 pb
- 2.2 ± 0.8 pb

Candidate Events

Event Probability Discriminant

Normalized to Prediction
CDF measurement of $|V_{tb}|$

- Assume:
  - Flat prior in $|V_{tb}|^2$
  - $0 < |V_{tb}|^2 \leq 1$
  - $|V_{td}|^2 + |V_{ts}|^2 << |V_{tb}|^2$

- Combined measurement and limit in preparation

$|V_{tb}| = 0.88^{+0.14}_{-0.12}$ (experiment) $\pm 0.07$ (theory)
# Tevatron summary

- **CDF combination in preparation**

## Tevatron Single Top Summary

<table>
<thead>
<tr>
<th>Method</th>
<th>Cross Section (pb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likelihood Function: CDF</td>
<td>1.8 ± 0.9</td>
</tr>
<tr>
<td>(2200 pb⁻¹)</td>
<td>0.8</td>
</tr>
<tr>
<td>Matrix Element: CDF</td>
<td>2.2 ± 0.8</td>
</tr>
<tr>
<td>(2200 pb⁻¹)</td>
<td>0.7</td>
</tr>
<tr>
<td>Neural Network: CDF</td>
<td>2.0 ± 0.9</td>
</tr>
<tr>
<td>(2200 pb⁻¹)</td>
<td>0.8</td>
</tr>
<tr>
<td>Decision Tree: D0</td>
<td>4.9 ± 1.4</td>
</tr>
<tr>
<td>(900 pb⁻¹)</td>
<td>1.4</td>
</tr>
<tr>
<td>Matrix Element: D0</td>
<td>4.8 ± 1.6</td>
</tr>
<tr>
<td>(900 pb⁻¹)</td>
<td>1.4</td>
</tr>
<tr>
<td>Neural Network: D0</td>
<td>4.4 ± 1.6</td>
</tr>
<tr>
<td>(900 pb⁻¹)</td>
<td>1.4</td>
</tr>
<tr>
<td>BLUE Combination: D0</td>
<td>4.7 ± 1.3</td>
</tr>
<tr>
<td>(900 pb⁻¹)</td>
<td>1.3</td>
</tr>
</tbody>
</table>

*Z. Sullivan, PRD 70, 114012 (2004)*

![Single Top Production Cross Section (pb)](image-url)
Conclusion

- Measuring the single top cross section is very challenging!
- Multivariate analysis techniques are essential to increase sensitivity
- Both CDF and D0 have seen evidence of single top quark production
- We are able to directly measure $|V_{tb}|$
- Now we enter the next stage:
  - Evidence for s- and t-channel production separately
  - Top quark polarization measurement
  - And, of course, onward to the Higgs!
Backup slides
Heavy flavor calibration

- Calibrate $W +$ heavy flavor in $W + 1$ jet data
- Three-parameter fit to bottom/charm/light templates of jet-flavor-separating distributions
- Correct for EWK/Top contributions
- Cross-check light flavor yield with prediction from mistag matrix

\[ K_{HF} = 1.4 \pm 0.4 \]
Non-W background

- Build non-W model from data
- Invert non-kinematic lepton identification cuts
- Data is superposition of non-W and W+jets contribution
- Do likelihood fit to data

**Electrons + 2 Jets**

**Muons + 2 Jets**

Residual non-W:

- Electrons + 2 Jets: 5.1% ± 2.0%
- Muons + 2 Jets: 1.4% ± 0.6%

Data is superposition of non-W and W+jets contribution.
Matrix element cross-check

- Check the discriminant in the orthogonal untagged region to show that we model the data well
- Region is dominated by light jets — validates our modeling of $W + \text{jet}$ kinematics
Matrix element method

- Similar approach for both experiments
- CDF also includes jet flavor separator in discriminant to reduce contamination from light flavor jets:

\[
EPD = \frac{b \cdot P_{\text{Singletop}}}{b \cdot P_{\text{Singletop}} + b \cdot P_{Wbb} + (1-b) \cdot P_{W+\text{charm}}}
\]
# Systematic errors

## 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 functions</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 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>