Measurement of Single Top-Quark Production at CDF

Jan Lueck
Karlsruhe Institute of Technology
On behalf of the CDF Collaboration

top2008 - Isola d’Elba, Italy - May 18-24, 2008
Single Top-Quark Production

- test of Standard Model (V-A, …)
- direct measurement of $|V_{tb}|^2$
- test of b-quark structure function
- same channel as WH @ Tevatron
- proving ground for WH

Standard Model prediction:

\[ \sigma_{t\text{-channel NLO}} = 2.0 \pm 0.3 \text{ pb} \]
\[ \sigma_{s\text{-channel NLO}} = 0.9 \pm 0.1 \text{ pb} \]
\[ \sigma_{\text{single top NLO}} = 2.9 \pm 0.4 \text{ pb} \]

Harris et al., Phys. Rev. D 66, 054024
Sullivan, Phys. Rev. D 70, 114012

Tait, Phys. Rev. D 61, 034001
Event Signature

top quark decay \( \sim 100\% \) in W-boson and b-quark

\[ W \rightarrow l + \nu \]

- exact 1 isolated charged lepton \((e/\mu)\) \(E_T > 20\) GeV
- missing transverse energy \(\text{MET} > 25\) GeV
- 2 or 3 jets \((E_T > 20\) GeV, \(|\eta| < 2,8\))
  with 1 or 2 jets with reconstructed secondary vertex (b-tag)
- complex QCD-veto, Z-veto, …
Single-Top and Bkg Estimate

Determined from MC
Single-Top, Diboson, Z+jets
and top pair production:

\[ N_{\text{pred}} = \sigma_{\text{theo}} \cdot \epsilon_{\text{evt}} \cdot \int L \, dt \]

Determined from MC and Data
W+HF jets: W+jets normalization from data,
HF-fraction from ALPGEN MC
(calibrated in W+1jet side-band)
tagging eff. from ALPGEN MC

Determined from Data
W+LF jets (negative tag eff rates)
and QCD events (MET fit)
Jet Flavor Separator

~50% of all events do NOT contain b quarks even though a secondary vertex was required!

Train neural network with input from jet and track variables, e.g. vertex mass, decay length, track multiplicity, …

up to 20% gain in sensitivity!
W + Heavy Flavor Calibration

Use ALPGEN predictions of kinematic distribution but mistrust ALPGEN rates

Use b-tagged W+1-jet data control sample to estimate the HF fractions

Three-parameter fit to **bottom/charm/light** templates of KIT flavor separator distribution

Apply correction factor for Wbb and Wcc:

\[ K_{HF} = 1.4 \pm 0.4 \]
QCD Rate Estimate

- Apply complex QCD-veto

- MC models of inclusive jet+MET production not precise enough

- Use data samples to model kinematic:
  - “anti-lepton”: lepton candidates which pass all but two of the non-kinematic lepton-ID requirements
  - “jet-leptons”: Multijet events where one jet fakes the lepton

- Perform fits to MET distributions and extract QCD content with MET > 25 GeV

- Uncertainty on QCD rate: 40%
Single-Top- and Bkg Estimate

CDF Run II Preliminary, L=2.2 fb⁻¹

- Single-top signal hidden behind large bkg uncertainty
- tt: counting experiment possible

Signal region divided into 1 or 2 b-tags, e.g. 4 independent channels
optimization of apriori separation power done individually remaining treatment identical, e.g.
• event selection (QCD-veto, divide into 4 independent channels, …)
• signal + bkg predictions (Wbb/cc rate calibration, QCD rate MET fits, …)
• signal + bkg models
• systematic rate and shape uncertainties
Neural Network Analysis

Training in 4 channels
- NeuroBayes® package
- Signal:Bkg ratio 50:50
- Bkg mixed as predicted by SM
- up to ~80K training events
- ~10 to 20 input variables
- 15 hidden nodes
- 1 output node: “NN Output”
Fit Templates in 4 Channels
## Systematic Uncertainties

<table>
<thead>
<tr>
<th>Syst. Uncertainty</th>
<th>Rate</th>
<th>Shape</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 Function</td>
<td>2...3%</td>
<td></td>
</tr>
<tr>
<td>MC 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>NN Flavor-Separator</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mistag model</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q2 scale in ALPGEN MC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Input variable mismodeling</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wbb+Wcc normalization</td>
<td>30%</td>
<td></td>
</tr>
<tr>
<td>Wc normalization</td>
<td>30%</td>
<td></td>
</tr>
<tr>
<td>Mistag normalization</td>
<td>17...29%</td>
<td></td>
</tr>
<tr>
<td>Top-pair normalization &amp; mtop</td>
<td>23%</td>
<td></td>
</tr>
</tbody>
</table>
Systematic Uncertainties (e.g. JES)

**Estimate of syst. uncertainties of the**

**a) rates:** difference in acceptance between syst. and default sample

<table>
<thead>
<tr>
<th>process</th>
<th>2jets 1tag</th>
<th>2jets 2tags</th>
<th>3jets 1tag</th>
<th>3jets 2tags</th>
</tr>
</thead>
<tbody>
<tr>
<td>t-ch</td>
<td>-1.1/-0.6 %</td>
<td>4.8/-3.5 %</td>
<td>-10.4/10.6 %</td>
<td>-5.7/4.3 %</td>
</tr>
<tr>
<td>s-ch</td>
<td>-0.1/-0.6 %</td>
<td>1.2/-1.9 %</td>
<td>-8.3/9.4 %</td>
<td>-7.2/7.4 %</td>
</tr>
<tr>
<td>single-top</td>
<td>-0.8/0.2 %</td>
<td>1.8/-2.2 %</td>
<td>-9.1/9.9 %</td>
<td>-6.6/6.1 %</td>
</tr>
<tr>
<td>t\bar{t}</td>
<td>9.8/-9.4 %</td>
<td>8.1/-7.5 %</td>
<td>4.6/-5.1 %</td>
<td>5.4/-5.2 %</td>
</tr>
<tr>
<td>W+c+c+Wb</td>
<td>7.0/-6.9 %</td>
<td>10.8/-10.6 %</td>
<td>8.4/-7.7 %</td>
<td>11.0/-12.1 %</td>
</tr>
<tr>
<td>Wc</td>
<td>7.0/-6.3 %</td>
<td>11.3/-10.3 %</td>
<td>8.2/-6.9 %</td>
<td>13.9/-15.8 %</td>
</tr>
<tr>
<td>Z+jets</td>
<td>-5.3/5.4 %</td>
<td>5.0/-5.0 %</td>
<td>-10.8/14.0 %</td>
<td>-5.0/5.0 %</td>
</tr>
<tr>
<td>Diboson</td>
<td>-2.7/1.7 %</td>
<td>-3.0/1.5 %</td>
<td>-12.4/11.9 %</td>
<td>-11.0/11.0 %</td>
</tr>
</tbody>
</table>

**b) shape of the “NN Output”:**

bin-wise difference of normalized syst. and default shape

rate and shape uncertainties of same syst. source are treated correlated in the fit.
Simultaneous Likelihood-Fit to Data

Candidate Events vs. NN Output for W+2Jets 1Tag and CDF II Preliminary 2.2 fb⁻¹

Candidate Events vs. NN Output for W+2Jets 2Tag and CDF II Preliminary 2.2 fb⁻¹

Candidate Events vs. NN Output for W+3Jets 1Tag and CDF II Preliminary 2.2 fb⁻¹

Candidate Events vs. NN Output for W+3Jets 2Tag and CDF II Preliminary 2.2 fb⁻¹
Result

\[ \sigma_{\text{single top}} = 2.0^{+0.9}_{-0.8} \, \text{(stat. + syst.) pb} \]
- Modified frequentist approach
- Perform pseudo experiments (PE) with and without SM single top
  - Fluctuate all syst. uncertainties in PE
  - Binned likelihood fit for each PE:
    reduced likelihood as a function of the single-top cross section
- \( Q = \frac{L_{\text{reduced}}(\sigma_{s+t} = \text{SM})}{L_{\text{reduced}}(\sigma_{s+t} = 0)} \)

exp. Significance: \( 4.4\sigma \)
obs. Significance: \( 3.2\sigma \)
Events with "NN Output" > 0.8

![Graph showing candidate events vs NN output with different categories such as single top, tt, Wb+Wc, Wc, Mistags, Diboson, Z+jets, and data.]
Golden Electron Event with „NN Output“ = 0.93

Event taken 2007/05/27

light jet $E_T = 52$ GeV

electron $P_T = 66$ GeV/c

b-tagged jet $E_T = 38$ GeV
Multivariate Likelihood Function Analysis

histogram based binned likelihood function (LEP style) with 7-10 variables

Optimized for t-channel

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

\[
L(x) = \frac{\prod p_i^{\text{sig}}(x_i)}{\prod p_i^{\text{sig}}(x_i) + \prod p_i^{\text{bkg}}(x_i)}
\]
Multivariate Likelihood Function Analysis

CDF Run II Preliminary, 2.2 fb⁻¹

Events

$\sigma_{s+}\approx 1.8^{+0.9}_{-0.8}$ pb

exp. Significance: 3.4σ
obs. Significance: 2.0σ
Matrix Element Analysis

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

$$P(s_{i}^{\mu}, p_{j1}^{\mu}, p_{j2}^{\mu}) = \frac{1}{\sigma} \int d\rho_{j1} d\rho_{j2} d\rho_{\nu} \sum_{\text{comb}} \phi_{4} |M(p_{i}^{\mu})|^{2} \frac{f(q_{1}) f(q_{2})}{|q_{1}| |q_{2}|} W_{\text{jet}}(E_{\text{jet}}, E_{\text{part}})$$

Probability of measuring a jet energy $E_{j}$ if $E_{p}$ was produced.

Combine the signal and background probability densities into a single Event Probability Discriminant:

$$EPD = \frac{b \cdot P_{\text{Single top}}}{b \cdot P_{\text{Single top}} + b \cdot P_{\text{Wb}} + b \cdot P_{\text{tt}} + (1-b) \cdot P_{\text{W+charm}} + (1-b) \cdot P_{\text{W+light}}}$$

$b$ is b-probability from KIT flavor separator
Matrix Element Analysis

**W + 2 Jets**: exp. Significance: $4.5\sigma$, obs. Significance: $3.4\sigma$

**W + 3 Jets**: exp. Significance: $4.5\sigma$, obs. Significance: $3.4\sigma$

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

### Posterior Probability Density

- **Single Top Cross Section [pb]**
  - $\sigma_{\text{Single Top}} = 2.2^{+0.8}_{-0.7}\text{ pb}$
  - $68\%$ confidence interval

CDF Run II Preliminary
Use many input variables (>20), 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 \( p = s/(s+b) \) and sort events by output purity

Create series of “boosted” trees by re-weighting based on value of misclassification
Boosted Decision Trees Analysis

CDF Run II Preliminary, L=2.2 fb$^{-1}$

Candidate Events

BDT$_{2j1t}$

Normalized to Prediction

exp. Significance: 4.6$\sigma$

obs. Significance: 2.8$\sigma$

CDF Run II Preliminary, L=2.2 fb$^{-1}$

Candidate Events

BDT$_{2j2t}$

Normalized to Prediction

exp. Significance: 4.6$\sigma$

obs. Significance: 2.8$\sigma$

CDF Run II Preliminary, L=2.2 fb$^{-1}$

Posterior Probability Density

$\sigma_{\text{Single Top}} = 1.9^{+0.8}_{-0.7}$ pb

68%
Combination of NN, LF, ME

Combine Neural Network, Likelihood Function, and Matrix Elements analyses discriminants in each channel as inputs to a neural network.

**Combination:**

\[ \sigma_{\text{single-top}} = 2.2 \pm 0.7 \text{ pb} \]

B.L.U.E. cross check

\[ \sigma_{\text{single-top}} = 2.1^{+0.7}_{-0.6} \text{ pb} \]

**exp. Significance:** 5.1\sigma

**obs. Significance:** 3.7\sigma

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

**|V_{tb}| = 0.89 \pm 0.14 \text{ (exp.)} \pm 0.07 \text{ (theory)}**

**|V_{tb}| > 0.66 \text{ (95\% C.L.)}**
LF Separate s-channel Search

Use double b-tagged events with
- 2 secondary vertices
- 1 sec. vertex and a “loose” b-tag (jet probability tagger)

No QCD veto applied since QCD bkg small

Use only W+2jet events, only 1.9 fb⁻¹ of data

Use Likelihood function as multivariate analysis

Assume SM t-channel (very low)
NN Separate s- and t-channel Search

- Build t- and s-channel templates from the networks of the combined search for a 2 parameter likelihood fit to data
- Add an additional s-channel network in 2 Jets 1 Tag channel to build 2D discriminant apriori s-channel sensitivity improvement of ~15%

unwind 2D discriminant bin by bin
NN Separate s- and t-channel Search

\[ \sigma_{\text{t-channel}} = 0.8^{+0.7}_{-0.6} \text{ (stat. + syst.) pb} \]
\[ \sigma_{\text{s-channel}} = 1.6^{+0.9}_{-0.8} \text{ (stat. + syst.) pb} \]
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 ± 0.8</td>
</tr>
<tr>
<td>(2200 pb⁻¹)</td>
<td>0.8</td>
</tr>
<tr>
<td>Decision Tree: CDF</td>
<td>1.9 ± 0.8</td>
</tr>
<tr>
<td>(2200 pb⁻¹, not in combination)</td>
<td>0.7</td>
</tr>
<tr>
<td>Combination: CDF</td>
<td>2.2 ± 0.7</td>
</tr>
<tr>
<td>(2200 pb⁻¹)</td>
<td>0.7</td>
</tr>
</tbody>
</table>