Observation of Single Top Quark Production

Cecilia E. Gerber
University of Illinois at Chicago
for the DØ collaboration

Wine & Cheese Seminar
Fermilab
March 10, 2009
Outline

- Introduction
- Understanding the data
  - Event Selection
  - Background Modeling
- Multivariate Analysis Techniques
  - Boosted Decision Trees
  - Bayesian Neural Networks
  - Matrix Elements Method
  - Combination
- Expected Sensitivity
- Cross Sections and Significance
- Direct Measurement of $|V_{tb}|$
- Conclusions
Why Study the Top Quark?

- Predicted by the SM and discovered in 1995 by CDF and DØ
  - $m_t = 172.4 \pm 1.2$ GeV
- Top-Higgs Yukawa coupling $\approx 1$
  - may help identify the mechanism of EWSB and mass generation
  - may serve as a window to new physics that might couple preferentially to top

- Successful Tevatron top quark program
  - High precision measurements for the top quark mass, top pair production cross section and decay properties
- Some basic quantities still unmeasured: spin, width, lifetime
- Single top quark production predicted by the SM, had not been observed till now
Single Top Production

- Probe of the Wtb interaction with no assumption on the number of quark families or unitarity of the CKM matrix

- Cross sections sensitive to beyond-the-SM processes
  - s-channel:
    - Resonances: heavy W’ boson, charged Higgs boson, Kaluza-Klein excited W_{KK}, technipion, etc.
  - t-channel
    - flavor-changing neutral currents
    - Fourth generation of quarks

Single top cross sections from Kidonakis and Vogt, PRD 68, 114014 (2003) for m_t = 170 GeV
Experimentally very challenging

Simple counting experiment cannot extract the signal from the overwhelming background
Experimentally very challenging

- Same final state as WH
- Backgrounds are the same
- Test of techniques to extract small signal from a large background

Same final state as WH

- Backgrounds are the same
- Test of techniques to extract small signal from a large background
Evidence for Single Top Production

Status as of March 3, 2009

<table>
<thead>
<tr>
<th>Signal Significance</th>
<th>Cross Section</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Expected</strong></td>
<td><strong>Observed</strong></td>
</tr>
<tr>
<td>DØ (0.9 fb⁻¹) PRL 98, 181802 (2007)</td>
<td>2.3σ</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>CDF (2.2 fb⁻¹) PRL 101, 252001 (2008)</td>
<td>4.9σ</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Observed significance is a measure of how likely it is to measure the cross section in the absence of signal.
Culmination of a Long History

**DØ**
- **Search:** PRD 63, 031101 *(2000)*
- **Search:** PLB 517, 282 *(2001)*
- **Search:** PLB 622, 265 *(2005)*
- **W:** PLB 641, 423 *(2006)*
- **Search:** PRD 75, 092007 *(2007)*
- **Evidence:** PRL 98, 181802 *(2007)*
- **FCNC:** PRL 99, 191802 *(2007)*
- **W:** PRL 100, 211802 *(2007)*
- **Evidence:** PRD 78, 012005 *(2008)*
- **Wtb:** PRL 101, 221801 *(2008)*
- **Wtb:** PRL 102, 092002 *(2009)*
- **H:** (PRL) arXiv:0807.0859
- **Observation:** (PRL) arXiv:0903.0850

**CDF**
- **Search:** PRD 65, 091102 *(2002)*
- **W:** PRL 90, 081802 *(2003)*
- **Search:** PRD 69, 052003 *(2004)*
- **Search:** PRD 71, 012005 *(2005)*
- **Evidence:** PRL 101, 252001 *(2008)*
- **FCNC:** (PRL) arXiv:0812.3400
- **W:** (PRL) arXiv:0902.3276
- **Observation:** (PRL) arXiv:0903.0885
Dataset

- DØ has >5 fb⁻¹ on tape

Many thanks to the Fermilab Accelerator Division!

This analysis uses 2.3 fb⁻¹ of data collected from 2002 to 2007
- Full Run IIa dataset, 1.1 fb⁻¹ (20% increase w.r.t. 2006 analysis)
- Run IIb dataset, 1.2 fb⁻¹
Event Selection

**FINAL STATE CONTAINS:**
- One high-\(p_T\) isolated electron or muon
- Large missing transverse energy \(\not{E_T}\)
- A b-jet from the top quark decay \(t \to Wb \to l\nu b\)
- A second b-jet or a light jet

**Analysis is performed in 24 channels:**
- Run IIa or Run IIb
- e or mu
- 2, 3 or 4 jets
- 1 or 2 b-tags
Changes in Event Selection

- Logical OR of many trigger conditions \((\text{was lepton + jets})\)
- Leading Jet \(\eta\) acceptance increased to 3.4 (was 2.5)
- Non-leading Jet \(p_T\) cut reduced to 15 GeV (was 20 GeV)
- Muon \(p_T\) cut reduced to 15 GeV (was 18 GeV)
- Loosened the b-jet identification criteria for the 2 tag case
- \(H_T, E_T^\prime\) not aligned with lepton or leading jet

Signal acceptance
- s-channel \(tb = (3.7 \pm 0.5)\) %
- t-channel \(tqb = (2.5 \pm 0.3)\) %

Signal acceptance increased by 18% compared to 2006
Signal and Background Models

- Single top quark signals modeled using SINGLETOP
  - Based on COMPHEP
  - Reproduces NLO kinematic distributions
  - PYTHIA for parton hadronization

- Top pair backgrounds modeled using ALPGEN
  - PYTHIA for parton hadronization
  - Parton-jet matching algorithm used to avoid double-counting final states
  - Normalized $\sigma = 7.91\text{pb}$ from Kidonakis and Vogt, PRD 68, 114014 (2003) for $m_t = 170$ GeV
  - Uncertainties $+7.7\% -12.7\%$ includes theory, PDF and mass shift to $(172.4\pm1.2)$ GeV
Signal and Background Models (Cont.)

- **W+jets** modeled using ALPGEN
  - PYTHIA for parton hadronization
  - MLM parton-jet matching avoids double-counting final states
  - $\eta$(jets), $\Delta\phi$(jet1,jet2), $\Delta\eta$(jet1,jet2) corrected to match data

- **QCD Multijet**
  - Misidentified lepton, directly from data
  - Switched to 10x larger background samples compared to 2006
  - Kept small (~5%) with topological selection cuts

- **Z+jets** modeled using ALPGEN + PYTHIA
  - Z+ heavy flavor corrected to theory, with ±20% uncertainty

- **Dibosons** modeled using PYTHIA
Background Normalization

- $W$+jets and multijet normalized using iterative template fits to data BEFORE TAGGING on three sensitive variables: $p_T(\ell)$, $E_T$, $M_T(W)$

$$N_{\text{data}}^{\text{pretag}} - N_{\text{bkgd}}^{\text{MC}} = S_{W+jets} N_{W+jets}^{\text{MC}} + S_{\text{multijet}} N_{\text{multijet}}^{\text{MC}}$$

- Normalization given by KS-weighted average
- Uncertainty obtained from max. difference with 1-variable result
  - 30 to 54% for multijet, 1.8 to 3.9% for $W$’s

$S:B = 1:259$
**b-Jet Identification**

- Separate $b$-jets from light-quark and gluon jets to reject most $W+jets$ background
- DØ uses a neural network algorithm with seven input variables based on impact parameter and reconstructed vertex
- Two operating points used in analysis:
  - TIGHT ($\varepsilon_b = 40\%$, $\varepsilon_c = 9\%$, $\varepsilon_l = 0.4\%$)
  - LOOSE ($\varepsilon_b = 50\%$, $\varepsilon_c = 14\%$, $\varepsilon_l = 1.5\%$)
- Leading $b$-jet $p_T > 20$ GeV
- Define two exclusive samples
  - EqOneTag: 1T, no L
  - EqTwoTag: 2L (was 2T; $\approx 50\%$ gain)
- Uncertainties dominated by variation in data samples used to measure the efficiencies.
- Smaller contribution from MC sample dependence
W + Heavy—Flavor Scale Factor

- W + heavy flavor normalized to theory (MCFM-NLO)
  - 1.47 (Wbb, Wcc), 1.38 (Wcj)

- Additional empirical correction
  - derived from two-jet data and simulation: includes zero-tag events
  - 0.95 ± 0.13 (Wbb, Wcc)

- Uncertainties considered
  - Data statistics ± 9%
  - ±40% single top cross section → ±7% in SF
  - ±10% on the Wcj theory SF → ±8% in SF
  - Additional ±10% Wbb/Wcc → ±5% in SF

S:B = 1:21 in 1Tag
S:B = 1:15 in 2Tag

DØ Single Top 2.3 fb⁻¹ Signals and Backgrounds
(All channels combined, after b-tagging)
Data/MC agreement (for all channels combined)
# Expected and Observed Events

## Event Yields in 2.3 fb\(^{-1}\) of DØ Data

<table>
<thead>
<tr>
<th>Source</th>
<th>2 jets</th>
<th>3 jets</th>
<th>4 jets</th>
</tr>
</thead>
<tbody>
<tr>
<td>s-channel (tb)</td>
<td>62 ± 9</td>
<td>24 ± 4</td>
<td>7 ± 2</td>
</tr>
<tr>
<td>t-channel (tqb)</td>
<td>77 ± 10</td>
<td>39 ± 6</td>
<td>14 ± 3</td>
</tr>
<tr>
<td>(W+b\bar{b})</td>
<td>678 ± 104</td>
<td>254 ± 39</td>
<td>73 ± 11</td>
</tr>
<tr>
<td>(W+c\bar{c})</td>
<td>303 ± 48</td>
<td>130 ± 21</td>
<td>42 ± 7</td>
</tr>
<tr>
<td>(W+cj)</td>
<td>435 ± 27</td>
<td>113 ± 7</td>
<td>24 ± 2</td>
</tr>
<tr>
<td>(W+ij)</td>
<td>413 ± 26</td>
<td>140 ± 9</td>
<td>41 ± 3</td>
</tr>
<tr>
<td>(Z+\text{jets})</td>
<td>141 ± 33</td>
<td>54 ± 14</td>
<td>17 ± 5</td>
</tr>
<tr>
<td>Dibosons</td>
<td>89 ± 11</td>
<td>32 ± 5</td>
<td>9 ± 2</td>
</tr>
<tr>
<td>(tt → ℓℓ)</td>
<td>149 ± 23</td>
<td>105 ± 16</td>
<td>32 ± 6</td>
</tr>
<tr>
<td>(tt → ℓ+\text{jets})</td>
<td>72 ± 13</td>
<td>331 ± 51</td>
<td>452 ± 66</td>
</tr>
<tr>
<td>Multijets</td>
<td>196 ± 50</td>
<td>73 ± 17</td>
<td>30 ± 6</td>
</tr>
<tr>
<td><strong>Total prediction</strong></td>
<td><strong>2,615 ± 192</strong></td>
<td><strong>1,294 ± 107</strong></td>
<td><strong>742 ± 80</strong></td>
</tr>
<tr>
<td><strong>Data</strong></td>
<td>2,579</td>
<td>1,216</td>
<td>724</td>
</tr>
</tbody>
</table>
Cross Check Samples

- Selected to test background model in regions dominated by one type of background: W+jets or Top Pairs

**W + JETS**
- 2 jets, 1 b-tagged jet
- $H_T(\ell, E_T, \text{allJets}) < 175$ GeV

**Top Pairs**
- 4 jets, 1 or 2 b-tagged jets
- $H_T(\ell, E_T, \text{allJets}) > 300$ GeV
## Systematic Uncertainties

### Components for normalization

<table>
<thead>
<tr>
<th>Component</th>
<th>Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Integrated luminosity</td>
<td>6.1%</td>
</tr>
<tr>
<td>$t\bar{t}$ cross section</td>
<td>12.7%</td>
</tr>
<tr>
<td>$Z$+jets and dibosons cross section</td>
<td>5.8%</td>
</tr>
<tr>
<td>Branching fractions</td>
<td>1.5%</td>
</tr>
<tr>
<td>Parton distribution functions (signal only)</td>
<td>3.0%</td>
</tr>
<tr>
<td>Triggers</td>
<td>5.0%</td>
</tr>
<tr>
<td>Instantaneous luminosity reweighting</td>
<td>1.0%</td>
</tr>
<tr>
<td>Primary vertex selection</td>
<td>1.4%</td>
</tr>
<tr>
<td>Lepton identification</td>
<td>2.5%</td>
</tr>
<tr>
<td>Jet fragmentation</td>
<td>(0.7–4.0)%</td>
</tr>
<tr>
<td>Initial-state and final-state radiation</td>
<td>(0.6–12.6)%</td>
</tr>
<tr>
<td>$b$-jet fragmentation</td>
<td>2.0%</td>
</tr>
<tr>
<td>Jet reconstruction and identification</td>
<td>1.0%</td>
</tr>
<tr>
<td>Jet energy resolution</td>
<td>4.0%</td>
</tr>
<tr>
<td>$W$+jets and $Z$+jets heavy flavor correction</td>
<td>13.7%</td>
</tr>
<tr>
<td>Multijets normalization to data</td>
<td>(30–54)%</td>
</tr>
<tr>
<td>Monte Carlo and multijets statistics</td>
<td>(0.5–16)%</td>
</tr>
</tbody>
</table>

### Components for normalization and shape

<table>
<thead>
<tr>
<th>Component</th>
<th>Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jet energy scale for signal</td>
<td>(1.1–13.1)%</td>
</tr>
<tr>
<td>Jet energy scale for total background</td>
<td>(0.1–2.1)%</td>
</tr>
<tr>
<td>$b$ tagging for single-tagged</td>
<td>(2.1–7.0)%</td>
</tr>
<tr>
<td>$b$ tagging for double-tagged</td>
<td>(9.0–11.4)%</td>
</tr>
</tbody>
</table>

### Component for shape only

<table>
<thead>
<tr>
<th>Component</th>
<th>Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALPGEN reweighting</td>
<td>–</td>
</tr>
</tbody>
</table>

Components that most affect the cross section measurement are shown in *red*

Other important contributions are shown in *pink*
Analysis Strategy

- Maximize signal acceptance
  - Background model gives good representation of data in each of the 24 independent analysis channels

- Calculate discriminant functions that separate signal from background
  - Boosted Decision Trees (BDT)
  - Bayesian Neural Networks (BNN)
  - Matrix Elements (ME)

- Check discriminant performance using data control samples
- Use ensembles of pseudo-data to test validity of methods
- Cross sections measured using binned likelihood calculation of signal + background to data
Multivariate Analyses: BDT & BNN

Use common Object and Event Kinematics, Angular Correlations, Jet Reconstruction and Top Quark Reconstruction variables

Boosted Decision Trees (BDT)
- Recover events that fail criteria in cut-based analysis
- Boosting averages the results over many trees, improving the performance
- Uses highest ranked common 64 variables

Bayesian Neural Network (BNN)
- NN train on signal and background, producing one output discriminant
- Bayesian NN average over many networks, improving the performance
- Uses highest ranked 18-28 variables in each channel
Discriminating Variables – BDT/BNN

OBJECT KINEMATICS

EVENT KINEMATICS

ANGULAR CORRELATIONS

New categories of variables added since 2006 improve BDT & BNN performance
Multivariate Analyses: ME

**Matrix Element (ME)**

- Method pioneered by DØ for the top quark mass measurement in Run I
- Use the 4-vectors of all reconstructed leptons and jets
- Use Feynman diagrams to compute an event probability density for signal and background hypotheses
- Uses events with 2 and 3 jets only
- ME for signal (tb & tqb) and background
- Split the sample in high and low $H_T$ (W+jets and top quark pair dominated regions) improves the performance
Multivariate Analyses: ME

2-jet channels
- tb
- tq
- Wbb
- Wcg
- Wgg

3-jet channels
- tbg
- tqb
- tqg
- Wbbg

Added additional Matrix Elements since 2006
- 2jets: top pair, WW, WZ, ggg
- 3jets: top pair, Wugg
Cross Check Samples and Linearity

- **BDT**
- **BNN**
- **ME**
Multivariate Discriminant Outputs

(a) Data vs. Boosted Decision Trees Output
- $tb + tqb$
- $W_{bb}$
- $W_{cc}$
- $W_{jj} + W_{c}$

(b) Data vs. Bayesian Neural Networks Output
- Multi-jets
- $Z + jets$
- $t\bar{t}\rightarrow ll$
- $t\bar{t}\rightarrow l + jets$
- $H_{T} < 175$ GeV

$DØ 2.3$ fb$^{-1}$

Signal normalized to measured x-sec
Statistical Analysis

Before looking at the data, we can calculate:

- **“expected p-value”**: fraction of the ensemble of zero-signal pseudo-datasets that give a cross section at least as large as the SM value.
  - For a Gaussian distribution, convert p-value to give “expected significance”

With the data, we can calculate:

- **“measured cross section”**
- **“measured p-value”**: fraction of the ensemble of zero-signal pseudo-datasets that give a cross section at least as large as the measured value
  - For a Gaussian distribution, convert p-value to give “measured significance”
Cross Section Measurement

- Cross Sections are measured by building a Bayesian posterior probability density.
- For each analysis, the single top cross section is given by the position of the posterior density peak, with 68% asymmetric interval as uncertainty.
- Gaussian prior for systematic uncertainties
  - Correlations of uncertainties properly taken into account
- Flat prior in signal cross sections
- Significance derived from background-only pseudo-datasets
## Cross Section Results

<table>
<thead>
<tr>
<th>MVA</th>
<th>$\sigma \pm \Delta \sigma$ (pb)</th>
<th>Expected Sensitivity</th>
<th>Observed Sensitivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>BDT</td>
<td>$3.74 \pm 0.95$</td>
<td>4.3 $\sigma$</td>
<td>4.6 $\sigma$</td>
</tr>
<tr>
<td>BNN</td>
<td>$4.70 \pm 1.18$</td>
<td>4.1 $\sigma$</td>
<td>5.2 $\sigma$</td>
</tr>
<tr>
<td>ME</td>
<td>$4.30 \pm 0.99$</td>
<td>4.1 $\sigma$</td>
<td>4.9 $\sigma$</td>
</tr>
</tbody>
</table>

![Graphs showing cross section results for BDT, BNN, and ME](image)
Combination of Results

- Even though all MVA analyses use the same data, they are not 100% correlated
  - BNN&BDT are 75% correlated with each other, 60% with ME
- We use a BNN to combine the three methods. The BNN takes as input variables the output discriminants of the individual methods
- Expected sensitivity for the BNN Combination: 4.5 $\sigma$
- BLUE combination (used in 2006) now presented as a cross check

CROSS CHECK SAMPLES AND LINEARITY

![Event Yield vs Combination Output](image1)

![Event Yield vs Combination Output](image2)

![BNN Combination](image3)
Distributions for BNNComb > 0.9

**OBJECT KINEMATICS**

- **DØ**
  - 2.3 fb$^{-1}$
  - $e$+$μ$
  - 1-2 b-tags
  - 2-4 jets

**EVENT KINEMATICS**

- **DØ**
  - 2.3 fb$^{-1}$
  - $e$+$μ$
  - 1-2 b-tags
  - 2-4 jets

**ANGULAR CORRELATIONS**

- **DØ**
  - 2.3 fb$^{-1}$
  - $e$+$μ$
  - 1-2 b-tags
  - 2-4 jets

**JET RECONSTRUCTION**

- **DØ**
  - 2.3 fb$^{-1}$
  - $e$+$μ$
  - 1-2 b-tags
  - 2-4 jets

**TOP QUARK RECONSTRUCTION**

- **DØ**
  - 2.3 fb$^{-1}$
  - $e$+$μ$
  - 1-2 b-tags
  - 2-4 jets

**SINGLE TOP FINAL STATE**

- **DØ**
  - 2.3 fb$^{-1}$
  - $e$+$μ$
  - 1-2 b-tags
  - 2-4 jets
DØ Experiment Event Display
Single Top Quark Candidate Event, 2.3 fb⁻¹ Analysis


ET scale: 39 GeV
Combined Results

\[ \sigma(p\bar{p} \rightarrow tb + X, tqb + X) = 3.94 \pm 0.88 \text{ pb} \]

\((m_t=170\text{GeV})\)

\[\sigma^\text{meas} = 3.94 \text{ pb} \]

\(p\text{-value} = 2.5 \times 10^{-7}\)

Measured Significance = 5.03σ
Cross Section Summary

**DØ 2.3 fb⁻¹**

<table>
<thead>
<tr>
<th>Method</th>
<th>MVA</th>
<th>Expected Signif.</th>
<th>Observed Signif.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decision Trees</td>
<td>3.74</td>
<td>+0.95 -0.79 pb</td>
<td></td>
</tr>
<tr>
<td>Bayesian NNs</td>
<td>4.70</td>
<td>+1.18 -0.93 pb</td>
<td></td>
</tr>
<tr>
<td>Matrix Elements</td>
<td>4.30</td>
<td>+0.99 -1.20 pb</td>
<td></td>
</tr>
<tr>
<td>BLUE Combination</td>
<td>4.16</td>
<td>±0.84 pb</td>
<td></td>
</tr>
<tr>
<td>BNN Combination</td>
<td>3.94</td>
<td>±0.88 pb</td>
<td></td>
</tr>
<tr>
<td>BNNComb</td>
<td>4.5σ</td>
<td></td>
<td>5.0σ</td>
</tr>
</tbody>
</table>

N. Kidonakis, PRD 74, 114012 (2006)  
m_{top} = 170 GeV

σ (p̅p → tb+X, tqb+X) [pb]
Weak interaction eigenstates and mass eigenstates are not the same: there is mixing between quarks, described by CKM matrix

General form of the $Wtb$ vertex

$$\Gamma_{Wtb}^\mu = -\frac{g}{\sqrt{2}} V_{tb} \left\{ \gamma^\mu \left[ f_1^L P_L + f_1^R P_R \right] - \frac{i\sigma^{\mu\nu}}{M_W} (p_t - p_b)_\nu \left[ f_2^L P_L + f_2^R P_R \right] \right\}$$

Measurement assumes SM production mechanisms

- Pure V–A and CP-conserving interaction ($f_1^R = f_2^L = f_2^R = 0$)
  - $f_1^L$: strength of the left-handed $Wtb$ coupling, is allowed to be anomalous
- $|V_{td}|^2 + |V_{ts}|^2 << |V_{tb}|^2$ (supported by CDF & DØ “ratio” measurements)

Does not assume 3 generations or unitarity of the CKM matrix
Measurement of $|V_{tb}|$

- Use the measurement of the single top cross section to make a direct measurement of $|V_{tb}|$: $\sigma(tb, tqb) \propto |V_{tb}|^2$
  - Calculate a posterior in $|V_{tb}|^2$
  - Measure the strength of the V–A

<table>
<thead>
<tr>
<th>Additional theoretical uncertainties ( %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$tb+tqb$</td>
</tr>
<tr>
<td>Top quark mass</td>
</tr>
<tr>
<td>Factorization Scale</td>
</tr>
<tr>
<td>PDF</td>
</tr>
<tr>
<td>$\alpha_s$</td>
</tr>
</tbody>
</table>

**Graphs**

- $|V_{tb}f_L^1| = 1.07 \pm 0.12$
  - Flat prior $\geq 0$
  - 68% confidence

- $|V_{tb}| > 0.78$
  - At 95% CL
  - $0 \leq$ flat prior $\leq 1$
  - 95% confidence
  - 68% confidence

assuming $f_1^L = 1$
Conclusions

The DØ collaboration observes single top quark production in 2.3 fb\(^{-1}\) of Run II data

\[ \sigma(p\bar{p} \rightarrow tb + X, tqb + X) = 3.94 \pm 0.88 \text{ pb} \]

Measured Significance 5.03\(\sigma\)

Direct measurement of \(|V_{tb}|\)

\[ |V_{tb} f_1^{L}| = 1.07 \pm 0.12 \]

flat prior \(\geq 0\)

0.78 < \(|V_{tb}| < 1\) @ 95% CL

\(0 \leq \text{ flat prior } \leq 1\)