Evidence for Single Top Quark Production at DØ & A First Direct Measurement of $|V_{tb}|$

Dag Gillberg
For the DØ Collaboration
The Tevatron and DØ

The DØ detector

- Silicon and fiber tracker
- 2T Superconducting Solenoid
- Ur-LAr Calorimeter
- Muon system

Fermilab

[Diagram of Fermilab with Tevatron, CDF, DØ, 1.96 TeV, and other elements]

Dag Gillberg (SFU)
The Top Quark

- Top pair ($t\bar{t}$) production observed 1995 by the DØ and the CDF Collaborations
- By far the heaviest fundamental particle we know of
- Cross section: $6.8 \pm 1.2$ pb (incl. uncertainty on the mass)
Electroweak single top production

- Never observed before
- Dominant production channels at the Tevatron:

\[ \sigma_{NLO} = 0.88 \pm 0.11 \text{ pb} \] (*)

\[ \sigma_{NLO} = 1.98 \pm 0.25 \text{ pb}(*) \]

(*) \( m_t = 175 \text{ GeV} \), Phys.Rev. D70 (2004) 114012
Motivation

1. Predicted but not observed before

- Study the $Wtb$ coupling
- $Wtb$ measurement (more later)
- Unitarity test of CKM matrix
- Anomalous $Wtb$ couplings

3. Measurement of top quark properties: polarization, decay width, lifetime, mass...

4. New physics, example:
   - $s$-channel: heavy $W'$, charged Higgs
   - $t$-channel: Flavour-changing neutral currents
   - 4-th quark generation?

5. Also, very similar to $WH (\rightarrow l\nu b\bar{b})...$
Motivation

1. Predicted but not observed before
2. **Study the Wtb coupling**
   - Direct $|V_{tb}|$ measurement (more later)
   - Unitarity test of CKM matrix
   - Anomalous Wtb couplings
Motivation

1. Predicted but not observed before
2. **Study the Wtb coupling**
   - Direct $|V_{tb}|$ measurement (more later)
   - Unitarity test of CKM matrix
   - Anomalous Wtb couplings
3. Measurement of top quark properties:
   - **polarization**, decay width, lifetime, mass...

4. New physics, example:
   - $s$-channel: heavy $W'$, charged Higgs
   - $t$-channel: Flavour-changing neutral currents
   - 4-th quark generation?

5. Also, very similar to $WH (\rightarrow l\nu b\bar{b})...$
Motivation

1. Predicted but not observed before
2. **Study the Wtb coupling**
   - Direct $|V_{tb}|$ measurement (more later)
   - Unitarity test of CKM matrix
   - Anomalous Wtb couplings
3. Measurement of top quark properties: **polarization**, decay width, lifetime, mass...
4. **New physics**, example:
   - $s$-channel: heavy $W'$, charged Higgs
   - $t$-channel: Flavour-changing neutral currents
   - 4-th quark generation?
5. Also, very similar to $WH(\rightarrow l\nu b\bar{b})$...
Signal and backgrounds samples

**Single top signal**
- CompHEP-SingleTop

**W+jets**
- Most difficult background
- Shapes from Alpgen
- Normalization and heavy flavour fraction from data

**tt**
- Alpgen
- Normalized to $\sigma_{NNLO} = 6.8$ pb

**Multijet events**
- misidentified lepton, from data
**Analysis Strategy**

**Strategy:** Maximize signal acceptance – extract signal using multivariate techniques

**Signature**
- isolated lepton
- $\mathbb{E}_T$
- jets
- at least 1 b-jet

**Event selection**
- One lepton:
  - electron: $p_T > 15 \text{ GeV}, |\eta_{det}| < 1.1$
  - muon: $p_T > 18 \text{ GeV}, |\eta_{det}| < 2$
- $\mathbb{E}_T > 15 \text{ GeV}$
- 2-4 jets: $p_T > 15 \text{ GeV}, |\eta| < 3.4$
  - Leading jet: $p_T > 25 \text{ GeV}, |\eta_{det}| < 2.5$
  - Second leading jet: $p_T > 20 \text{ GeV}$
- 1-2 jets b-tagged
**Expected** $N_{\text{events}}$

<table>
<thead>
<tr>
<th>Component</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s + t$ signal</td>
<td>62</td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>348</td>
</tr>
<tr>
<td>$Wjj$</td>
<td>174</td>
</tr>
<tr>
<td>$Wbb$ &amp; $Wcc$</td>
<td>675</td>
</tr>
<tr>
<td>multijet</td>
<td>201</td>
</tr>
</tbody>
</table>

**data/bkg** $N_{\text{events}}$

<table>
<thead>
<tr>
<th>Component</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>signal</td>
<td>62</td>
</tr>
<tr>
<td>background</td>
<td>1399</td>
</tr>
<tr>
<td>data</td>
<td>1398</td>
</tr>
</tbody>
</table>
The Decision Tree Analysis

- Single top signal sample trained against our background samples
- Idea: recover events that fail criteria in cut-based analysis
- Separate training per \((\text{lep}, N_{\text{tag}}, N_{\text{jet}})\)-bin
- **Boosting** – give mis-classified events higher weight and re-train! (20 times)

Decision Tree Output

Bayesian posterior

\[
P(\sigma | D) \propto \int_{a,b} P(D | \sigma, a, b) \cdot \text{Prior}(\sigma) \cdot \text{Prior}(a, b)
\]
Ensemble Testing

Tested our machinery with many sets of pseudo-data

- Subset of our total pool of background events
- Individual statistical and systematical fluctuations
- Wonderful tool – like running DØ 1000s of times!
- Generated several ensembles with different single top content →

![Decision Trees graph](image)

Used 70,000 pseudo-datasets with no single top content

<table>
<thead>
<tr>
<th>Input (tb+tbq) cross section [pb]</th>
<th>Measured (tb+tbq) Xsec [pb]</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
</tr>
</tbody>
</table>

Dag Gillberg (SFU)
First Evidence for Single Top – Feb 19, 2007
Ensemble Testing

Tested our machinery with many sets of **pseudo-data**

- Subset of our total pool of background events
- Individual statistical and systematical fluctuations
- **Wonderful tool** – like running DØ 1000s of times!
- Generated several ensembles with different single top content →

**Significance**

- Used 70,000 pseudo-datasets with no single top content
- Defined by fraction of datasets in which we measure at least the SM cross section (**expected significance**) or at least the observed cross section (**observed significance**)

![Decision Trees Graph](image)
### Expected Decision Tree Results

- **Expected cross section**: $2.7^{+1.5}_{-1.4}$ pb
- **Expected p-value**: 1.9%
- **Expected significance**: $2.1\sigma$

---

#### Diagram

- **Posterior Probability Density [pb$^{-1}$]**
  - Decision Trees
  - "Data" = SM signal + background
  - Expected result

- **Pseudo-datasets / 0.4 pb**
  - SM = 2.9 pb

- **tb+tqb Cross Section [pb]**
  - Decision Trees
  - Probability to rule out background-only hypothesis
  - Zero-signal ensemble
Boosted decision tree observed results

\[ \sigma_{s+t} = 4.9 \pm 1.4 \text{ pb} \]

p-value = 0.035\% (3.4\(\sigma\))

SM compatibility: 11\% (1.3\(\sigma\))

Evidence!
The three analyses give consistent results!

Same data used (0.9 fb\(^{-1}\)) – therefore correlated
First direct measurement of $|V_{tb}|$

- Assuming $V_{td}^2 + V_{ts}^2 \ll V_{tb}^2$ and pure $V-A$ and CP-conserving $W_{tb}$ interaction

- $|V_{tb}^{f_1}| = 1.3 \pm 0.2$

No assumption about number of quark families or CKM unitarity

$0.68 < |V_{tb}| \leq 1 \text{ @ 95\% CL}$ (assuming $f_1^L = 1$, flat prior in [0,1])
Conclusion

Using Decision Trees:

First evidence for single top production found at DØ!

\[
\sigma(p\bar{p} \rightarrow tb + X, tqb + X) = 4.9 \pm 1.4 \text{ pb}
\]
3.4\(\sigma\) significance

First direct measurement of \(|V_{tb}|\)

\[
|V_{tb}f_1^L| = 1.3 \pm 0.2
\]
assuming \(f_1^L = 1\):

\[
0.68 < |V_{tb}| \leq 1 \text{ @ 95\% CL}
\]
(Assuming \(V_{td}^2 + V_{ts}^2 \ll V_{tb}^2\) and pure \(V-A\) and CP-conserving \(Wtb\) interaction)

hep-ex/0612052, submitted to PRL

- Working on combining our three multivariate analyses
- A lot more data (\(\times 2!\)) already at hand
## Percentage of single top $tb+tqb$ selected events and S:B ratio

(white squares = no plans to analyze)

<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></td>
<td></td>
<td></td>
<td></td>
<td>1%</td>
</tr>
<tr>
<td></td>
<td>10%</td>
<td>25%</td>
<td>12%</td>
<td>3%</td>
<td>1 : 230</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></td>
<td></td>
<td></td>
<td></td>
<td>1%</td>
</tr>
<tr>
<td></td>
<td>6%</td>
<td>21%</td>
<td>11%</td>
<td>3%</td>
<td>1 : 53</td>
</tr>
<tr>
<td></td>
<td>1 : 100</td>
<td>1 : 20</td>
<td>1 : 25</td>
<td>1 : 40</td>
<td></td>
</tr>
<tr>
<td>2 tags</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>3%</td>
<td>2%</td>
<td>1%</td>
<td>1%</td>
<td>1 : 43</td>
</tr>
<tr>
<td></td>
<td>1 : 11</td>
<td>1 : 15</td>
<td>1 : 38</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Yields before $b$-tagging

<table>
<thead>
<tr>
<th>Source</th>
<th>Event Yields in 0.9 fb$^{-1}$ Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Electron+muon combined, before $b$-tagging</td>
</tr>
<tr>
<td></td>
<td>2 jets</td>
</tr>
<tr>
<td>$tb$</td>
<td>25</td>
</tr>
<tr>
<td>$tqb$</td>
<td>47</td>
</tr>
<tr>
<td>$t\bar{t} \rightarrow ll$</td>
<td>62</td>
</tr>
<tr>
<td>$t\bar{t} \rightarrow l+\text{jets}$</td>
<td>40</td>
</tr>
<tr>
<td>$W+b\bar{b}$</td>
<td>670</td>
</tr>
<tr>
<td>$W+c\bar{c}$</td>
<td>1,959</td>
</tr>
<tr>
<td>$W+jj$</td>
<td>10,160</td>
</tr>
<tr>
<td>Multijets</td>
<td>1,762</td>
</tr>
<tr>
<td>Total background</td>
<td>14,654</td>
</tr>
<tr>
<td>Data</td>
<td>14,652</td>
</tr>
</tbody>
</table>
Data-background agreement before \( b \)-tagging

- Background model show good agreement with data!
Agreement after tagging

Sample | # of Events
--- | ---
s&t-channel Signal | 62
Wjj | 174
tt→l+jets | 266
Wbb & Wcc | 675
Mis-ID’s leptons | 201
Diboson, tt→ dileptons | 82

| Totals | 2 Jets | 3 Jets | 4 Jets |
--- | --- | --- | --- |
Data | 697 | 455 | 246 |
Total Background | 685 | 460 | 253 |
Signal | 36 | 20 | 6 |
<table>
<thead>
<tr>
<th>Source of Uncertainty</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top pairs normalization</td>
<td>18%</td>
</tr>
<tr>
<td>W+jets &amp; multijets normalization</td>
<td>18–28%</td>
</tr>
<tr>
<td>Integrated luminosity</td>
<td>6%</td>
</tr>
<tr>
<td>Trigger modeling</td>
<td>3–6%</td>
</tr>
<tr>
<td>Lepton ID corrections</td>
<td>2–7%</td>
</tr>
<tr>
<td>Jet modeling</td>
<td>2–7%</td>
</tr>
<tr>
<td>Other small components</td>
<td>Few %</td>
</tr>
<tr>
<td>Jet energy scale</td>
<td>1–20%</td>
</tr>
<tr>
<td>Tag rate functions</td>
<td>2–16%</td>
</tr>
</tbody>
</table>
Cross-check samples

“W+jets”

“tt̄”

Electrons

Muons
Measuring a cross section

Probability to observe data distribution $D$, expecting $y$:

$$y = \alpha l \sigma + \sum_{s=1}^{N} b_s \equiv a \sigma + \sum_{s=1}^{N} b_s$$

$$P(D|y) \equiv P(D|\sigma, a, b) = \prod_{i=1}^{nbins} P(D_i|y_i)$$

The cross section is obtained

$$Post(\sigma|D) \equiv P(\sigma|D) \propto \int_{a}^{b} \int_{\sigma}^{\sigma_{peak}} P(D|\sigma, a, b)Prior(\sigma)Prior(a, b)$$

- Bayesian posterior probability density
- Shape and normalization systematics treated as nuisance parameters
- Correlations between uncertainties properly accounted for
- Flat prior in signal cross section
Object Kinematics

\[ p_T(jet1) \]
\[ p_T(jet2) \]
\[ p_T(jet3) \]
\[ p_T(jet4) \]
\[ p_T(best1) \]
\[ p_T(notbest1) \]
\[ p_T(notbest2) \]
\[ p_T(tag1) \]
\[ p_T(untag1) \]
\[ p_T(untag2) \]

Angular Correlations

\[ \Delta R(jet1,jet2) \]
\[ \cos(best1,lepton)_{besttop} \]
\[ \cos(best1,notbest1)_{besttop} \]
\[ \cos(tag1,alljets)_{alljets} \]
\[ \cos(tag1,lepton)_{btaggedtop} \]
\[ \cos(jet1,alljets)_{alljets} \]
\[ \cos(jet1,lepton)_{btaggedtop} \]
\[ \cos(jet2,alljets)_{alljets} \]
\[ \cos(jet2,lepton)_{btaggedtop} \]
\[ \cos(lepton,Q(lepton) \times z)_{besttop} \]
\[ \cos(lepton_{besttop,besttop})_{CMframe} \]
\[ \cos(lepton_{btaggedtop,btaggedtop})_{CMframe} \]
\[ \cos(notbest,alljets)_{alljets} \]
\[ \cos(notbest,lepton)_{besttop} \]
\[ \cos(untag1,alljets)_{alljets} \]
\[ \cos(untag1,lepton)_{btaggedtop} \]

Event Kinematics

\[ Aplanarity(alljets,W) \]
\[ M(W,best1) \text{ ("best" top mass)} \]
\[ M(W,tag1) \text{ ("b-tagged" top mass)} \]
\[ H_T(alljets) \]
\[ H_T(alljets−best1) \]
\[ H_T(alljets−tag1) \]
\[ H_T(alljets,W) \]
\[ H_T(jet1,jet2) \]
\[ H_T(jet1,jet2,W) \]
\[ M(alljets) \]
\[ M(alljets−best1) \]
\[ M(alljets−tag1) \]
\[ M(jet1,jet2) \]
\[ M(jet1,jet2,W) \]
\[ M_T(jet1,jet2) \]
\[ M_T(W) \]
\[ Missing \ E_T \]
\[ p_T(alljets−best1) \]
\[ p_T(alljets−tag1) \]
\[ p_T(jet1,jet2) \]
\[ Q(lepton) \times \eta(untag1) \]
\[ \sqrt{\hat{s}} \]
\[ Sphericity(alljets,W) \]

- Adding variables does not degrade performance
- Tested shorter lists, lost some sensitivity
- Same list used for all channels
Boosted decision tree event characteristics

$DT < 0.3$

$DT > 0.55$

$DT > 0.65$
Matrix Elements Results

\[ \sigma(tb+tqb) = 4.6^{+1.8}_{-1.5} \text{ pb} \]

Measured p-value = 0.21 \%

Measured significance = 2.9 \sigma

Compatibility with SM = 21\%
Bayesian Neural Network Analysis Results

\[ \sigma(tb+tqb) = 5.0 \pm 1.9 \text{ pb} \]

Measured p-value = 0.89 %

Measured significance = 2.4 \( \sigma \)

Compatibility with SM = 18%
High discriminant correlation

Choose the 50 highest events in each discriminant and look for overlap

<table>
<thead>
<tr>
<th></th>
<th>Electron</th>
<th>Muon</th>
</tr>
</thead>
<tbody>
<tr>
<td>DT vs ME</td>
<td>52%</td>
<td>58%</td>
</tr>
<tr>
<td>DT vs BNN</td>
<td>56%</td>
<td>48%</td>
</tr>
<tr>
<td>ME vs BNN</td>
<td>46%</td>
<td>52%</td>
</tr>
</tbody>
</table>

Linear correlation

Measured cross section in 400 members of SM ensemble with all three techniques and calculated the linear correlation between each pair

<table>
<thead>
<tr>
<th></th>
<th>DT</th>
<th>ME</th>
<th>BNN</th>
</tr>
</thead>
<tbody>
<tr>
<td>DT</td>
<td>100%</td>
<td>39%</td>
<td>57%</td>
</tr>
<tr>
<td>ME</td>
<td>39%</td>
<td>100%</td>
<td>29%</td>
</tr>
<tr>
<td>BNN</td>
<td>57%</td>
<td>29%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Z. Sullivan PRD 70, 114012 (2004), m_t = 175 GeV
Ensemble Testing

To verify that the machinery is working properly we test with many sets of pseudo-data – Random subsets from our total pool of background events

- Each subset correspond to 0.9 fb$^{-1}$ data
- Individual statistical and systemtaical fluctuations for each pseudo-dataset according to our uncertainties
- Wonderful tool – like running DØ 1000s of times!
Ensemble Testing

To verify that the machinery is working properly we test with many sets of **pseudo-data** – Random subsets from our total pool of background events

- Each subset correspond to 0.9 fb$^{-1}$ data
- Individual statistical and systemtaical fluctuations for each pseudo-dataset according to our uncertainties
- **Wonderful tool** – like running DØ 1000s of times!

Generated ensembles include:

1. 0-signal ensemble ($s + t \sigma = 0 pb$)
2. SM ensemble ($s + t \sigma = 2.9 pb$)
3. “Mystery” ensembles to test analyzers ($s + t \sigma = ?? pb$)

Each analysis tests linearity of “response” to single top.
Sensitivity determination

- Use the 0-signal ensemble (68,000 pseudo-datasets)

<table>
<thead>
<tr>
<th>Expected p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fraction of 0-signal pseudo-datasets in which we measure at least 2.9 pb (SM $\sigma$)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Observed p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fraction of 0-signal pseudo-datasets in which we measure at least the observed cross section</td>
</tr>
</tbody>
</table>
First direct measurement of $|V_{tb}|$

- **General form of $Wtb$ vertex:**

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

- **Assume**
  - SM top quark decay: $V_{td}^2 + V_{ts}^2 \ll V_{tb}^2$
  - Pure $V-A$: $f_1^R = 0$
  - CP conservation: $f_2^L = f_2^R = 0$

- **No need to assume only three quark families or CKM matrix unitarity**
  (unlike for previous measurements using $t\bar{t}$ decays)

- **Measure the strength of the $V-A$ coupling**, $|V_{tb} f_1^L|$, which can be $> 1$

**Additional theoretical uncertainties**

<table>
<thead>
<tr>
<th></th>
<th>$tb$</th>
<th>$tqb$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top mass</td>
<td>13 %</td>
<td>8.5 %</td>
</tr>
<tr>
<td>Scale</td>
<td>5.4 %</td>
<td>4.0 %</td>
</tr>
<tr>
<td>PDF</td>
<td>4.3 %</td>
<td>10 %</td>
</tr>
<tr>
<td>$\alpha_s$</td>
<td>1.4 %</td>
<td>0.01 %</td>
</tr>
</tbody>
</table>