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

Dugan O’Neil
For the DØ Collaboration

Dec. 8, 2006
Outline

- Introduction and Motivation
- Preparing for the Measurement
- Multivariate Analysis Techniques
  - Decision Trees
  - Matrix Elements Method
  - Bayesian Neural Networks
- Expected Sensitivity
- Cross Sections and Significance
- First Direct Measurement of $|V_{tb}|$
- Conclusions
The Tevatron is still the only place to make top quarks.
Single top quark production

**s-channel (tb)**

- $\sigma_{NLO} = 0.88 \pm 0.11$ pb (*)
- current limits (95% C.L.):
  - Run II DØ: $< 5.0$ pb (370pb$^{-1}$)
  - Run II CDF: $< 3.1$ pb (700pb$^{-1}$)

**t-channel (tqb)**

- $\sigma_{NLO} = 1.98 \pm 0.25$ pb (*)
- current limits (95% C.L.):
  - Run II DØ: $< 4.4$ pb (370pb$^{-1}$)
  - Run II CDF: $< 3.2$ pb (700pb$^{-1}$)

And some very nice CDF results in W&C just last week!!

Motivation

- Directly measure $|V_{tb}|$ for the first time (more later)
- Cross section sensitivity to beyond the SM processes
- Source of polarized top quarks. Spin correlations measurable in decay products.
- Important background to Higgs search
- Test of techniques to extract a small signal out of a large background
It’s not like we haven’t been looking already...


plus 7 PhDs. (CDF has a similar list)
...but it is a challenge!

(Stolen from CDF W&C - THANKS!)
Improvements

Run II Integrated Luminosity

- 910 fb\(^{-1}\) analysis
  December 2006
- 370 fb\(^{-1}\) analysis
  July 2005
- 230 fb\(^{-1}\) analysis
  Physics Letters B
  March 2005
- 2.13 fb\(^{-1}\)
- 1.79 fb\(^{-1}\)

Many thanks to the Accelerator Division
More Improvements...

- Background model improvements
- Fully reprocessed dataset: new calibrations, jet threshold, etc.
- Neural network b-tagging
- Split analysis channels by numbers of jets (exclusive bins)
- Combined $s + t$ search added (SM $s:t$ ratio)
Neural Network b-jet Tagger

- NN trained on 7 input variables from SVT, JLIP and CSIP taggers.
- Much improved performance!
  - fake rate reduced by 1/3 for same b-efficiency relative to previous tagger
  - smaller systematic uncertainties
- Tag Rate Functions (TRFs) in $\eta$, $P_T$, z-PV applied to MC
- Our operating point:
  - b-jet efficiency $\sim 50\%$
  - c-jet efficiency $\sim 10\%$
  - Light jet efficiency $\sim 0.5\%$
Event Selection

**Signature**
- isolated lepton
- $E_T$
- 2-4 jets
- at least 1 b-jet
- Only one tight and no other loose lepton
  - electron: $p_T > 15$ GeV and $|\eta_{det}| < 1.1$
  - muon: $p_T > 18$ GeV and $|\eta_{det}| < 2$
- $15 < E_T < 200$ GeV
- 2-4 jets with $p_T > 15$ GeV and $|\eta_{det}| < 3.4$
  - Leading jet with $p_T > 25$ GeV and $|\eta_{det}| < 2.5$
  - Second leading jet $p_T > 20$ GeV
Event Selection - Agreement Before Tagging

- Normalize $W$+multijet to data before tagging
- Checked 90 variables, 3 jet multiplicities, 1-2 tags, electron + muon
- Shown: electron, 2 jets, before tagging
- Good description of data

Dugan O'Neil (SFU)
## Event Selection - Yields

<table>
<thead>
<tr>
<th>Source</th>
<th>Event Yields in 0.9 fb(^{-1}) Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2 jets</td>
</tr>
<tr>
<td>(tb)</td>
<td>16 ± 3</td>
</tr>
<tr>
<td>(tqb)</td>
<td>20 ± 4</td>
</tr>
<tr>
<td>(t\bar{t} → ll)</td>
<td>39 ± 9</td>
</tr>
<tr>
<td>(t\bar{t} → l+\text{jets})</td>
<td>20 ± 5</td>
</tr>
<tr>
<td>(W+b\bar{b})</td>
<td>261 ± 55</td>
</tr>
<tr>
<td>(W+c\bar{c})</td>
<td>151 ± 31</td>
</tr>
<tr>
<td>(W+jj)</td>
<td>119 ± 25</td>
</tr>
<tr>
<td>Multijets</td>
<td>95 ± 19</td>
</tr>
<tr>
<td>Total background</td>
<td>686 ± 131</td>
</tr>
<tr>
<td>Data</td>
<td>697</td>
</tr>
</tbody>
</table>
### 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>≥ 5 jets</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 tags</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10%</td>
<td>25%</td>
<td>12%</td>
<td>3%</td>
<td>1%</td>
<td></td>
</tr>
<tr>
<td>1 : 3,200</td>
<td>1 : 390</td>
<td>1 : 300</td>
<td>1 : 270</td>
<td>1 : 230</td>
<td></td>
</tr>
<tr>
<td>1 tag</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6%</td>
<td>21%</td>
<td>11%</td>
<td>3%</td>
<td>1%</td>
<td></td>
</tr>
<tr>
<td>1 : 100</td>
<td>1 : 20</td>
<td>1 : 25</td>
<td>1 : 40</td>
<td>1 : 53</td>
<td></td>
</tr>
<tr>
<td>2 tags</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3%</td>
<td>2%</td>
<td>1%</td>
<td>0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 : 11</td>
<td>1 : 15</td>
<td>1 : 38</td>
<td></td>
<td>1 : 43</td>
<td></td>
</tr>
</tbody>
</table>
Systematic Uncertainties

- Systematic uncertainties can be either “shaped” (jet energy scale, tag rate functions)
  - Shift inputs by $\pm 1\sigma$, redo analysis
- or “normalization”
  - Uncertainties assigned per background, jet multiplicity, lepton, number of tags

Examples of Relative Systematic Uncertainties

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$t\bar{t}$ cross section</td>
<td>18%</td>
</tr>
<tr>
<td>Luminosity</td>
<td>6%</td>
</tr>
<tr>
<td>Electron trigger</td>
<td>3%</td>
</tr>
<tr>
<td>Muon trigger</td>
<td>6%</td>
</tr>
<tr>
<td>Jet energy scale</td>
<td>wide range</td>
</tr>
<tr>
<td>Jet fragmentation</td>
<td>5–7%</td>
</tr>
<tr>
<td>Heavy flavor ratio</td>
<td>30%</td>
</tr>
<tr>
<td>Tag-rate functions</td>
<td>2–16%</td>
</tr>
</tbody>
</table>
Systematic Uncertainties

Key for Plots
- Data
- $tb$
- $tq_{b}$
- $tt$
- $W + jets$
- $\pm 1\sigma$ uncertainty on background
Measuring the Cross Section

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

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

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

The cross section is obtained

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

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

- To verify that all of this machinery is working properly we test with many sets of **pseudo-data**.
- Wonderful tool to test analysis methods! Run DØ experiment 1000s of times!
- Generated ensembles include:
  1. 0-signal ensemble \((s + t \sigma = 0pb)\)
  2. SM ensemble \((s + t \sigma = 2.9pb)\)
  3. “Mystery” ensembles to test analyzers \((s + t \sigma = ??pb)\)
  4. Ensembles at measured cross section \((s + t \sigma = \text{measured})\)
  5. A high luminosity ensemble
- Each analysis tests linearity of “response” to single top.
Train

- Start with all events (first node)
- For each variable, find the splitting value with best separation between children (best cut).
- Select best variable and cut and produce Failed and Passed branches
- Repeat recursively on each node
- Stop when improvement stops or when too few events left. Terminal node = leaf.
Measure and Apply

- Take trained tree and run on independent simulated sample, determine purities.
- Apply to Data
- Should see enhanced separation (signal right, background left)
- Could cut on output and measure, or use whole distribution to measure.
**Decision Trees - Boosting**

**Boosting**
- Recent technique to improve performance of a weak classifier
- Recently used on DTs by GLAST and MiniBooNE
- Basic principal on DT:
  - train a tree $T_k$
  - $T_{k+1} = \text{modify}(T_k)$

**AdaBoost algorithm**
- Adaptive boosting
- Check which events are misclassified by $T_k$
- Derive tree weight $\alpha_k$
- Increase weight of misclassified events
- Train again to build $T_{k+1}$
- Boosted result of event $i$:
  \[ T(i) = \sum_{n=1}^{N_{\text{tree}}} \alpha_k T_k(i) \]

- Averaging dilutes piecewise nature of DT
- Usually improves performance

Decision Trees - Application to this Analysis

**DT Choices**
- 1/3 of MC for training
- Adaboost $\beta = 0.2$
- Boosting cycles = 20
- Signal leaf if purity $> 0.5$
- Minimum leaf size = 100 events
- Same total weight to signal and background to start
- Goodness of split - Gini factor

**Analysis Strategy**
- Train 36 separate trees: $(s, t, s + t) \times (e, \mu) \times (2,3,4 \text{ jets}) \times (1,2 \text{ tags})$
- For each signal train against the sum of backgrounds
Decision Trees - 49 variables

Object Kinematics

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

Angular Correlations

\[ \Delta R(\text{jet1, jet2}) \]
\[ \cos(\text{best1, lepton})_{\text{besttop}} \]
\[ \cos(\text{best1, notbest1})_{\text{besttop}} \]
\[ \cos(\text{tag1, alljets})_{\text{alljets}} \]
\[ \cos(\text{tag1, lepton})_{\text{btaggedtop}} \]
\[ \cos(\text{jet1, alljets})_{\text{alljets}} \]
\[ \cos(\text{jet1, lepton})_{\text{btaggedtop}} \]
\[ \cos(\text{jet2, alljets})_{\text{alljets}} \]
\[ \cos(\text{jet2, lepton})_{\text{btaggedtop}} \]
\[ \cos(\text{lepton, Q(lepton) × z})_{\text{besttop}} \]
\[ \cos(\text{lepton, btaggedtopframe})_{\text{besttopCMframe}} \]
\[ \cos(\text{lepton, btaggedtopframe})_{\text{btaggedtopCMframe}} \]
\[ \cos(\text{notbest, alljets})_{\text{alljets}} \]
\[ \cos(\text{notbest, lepton})_{\text{besttop}} \]
\[ \cos(\text{untag1, alljets})_{\text{alljets}} \]
\[ \cos(\text{untag1, lepton})_{\text{btaggedtop}} \]

Event Kinematics

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

- Adding variables does not degrade performance
- Tested shorter lists, lose some sensitivity
- Same list used for all channels
Decision Trees - Ensembles

- SM input is returned by DTs
- “Mystery” ensembles are unraveled by the DTs
- Linear response is achieved
Matrix Elements Method - Introduction

A matrix elements analysis takes a very different approach:

- Use the 4-vectors of all reconstructed leptons and jets
- Use matrix elements of main signal and background diagrams to compute an event probability density for signal and background hypotheses.
- Goal: calculate a discriminant:

\[
D_{s}(\vec{x}) = P(S|\vec{x}) = \frac{P_{\text{Signal}}(\vec{x})}{P_{\text{Signal}}(\vec{x}) + P_{\text{Background}}(\vec{x})}
\]

- Define \( P_{\text{Signal}} \) as properly normalized differential cross section

\[
P_{\text{Signal}}(\vec{x}) = \frac{1}{\sigma_{S}} d\sigma_{S}(\vec{x}) \quad \sigma_{S} = \int d\sigma_{S}(\vec{x})
\]

- Shared technology with mass measurement in \( t\bar{t} \) (eg. transfer functions)
Matrix Elements Method - Introduction

2-jets:

3-jets:
Matrix Elements Method - Ensembles

**ME analysis**

- \( \chi^2/\text{ndof} = 10.15/4 \)
- Slope = 1.04 ± 0.02
- Intercept = 0.27 ± 0.10
A different sort of neural network:

- Instead of choosing one set of weights, find posterior probability density over all possible weights
- Averaging over many networks weighted by the probability of each network given the training data
- Less prone to overtraining
- For details see: http://www.cs.toronto.edu/radford/fbm.software.html

Use 24 variables (subset of DT variables)
Bayesian Neural Network - Ensembles

BNN analysis

$\chi^2/\text{ndof} = 3.60/2$

Slope = $0.99 \pm 0.05$

Intercept = $0.31 \pm 0.15$
EXPECTED SENSITIVITY
Significance/Sensitivity Determination

We use our 0-signal ensemble to determine a significance for each measurement.

**Expected p-value**

The fraction of 0-signal pseudo-datasets in which we measure at least 2.9 pb.
We use our 0-signal ensemble to determine a significance for each measurement.

**Expected p-value**

The fraction of 0-signal pseudo-datasets in which we measure at least 2.9pb.

**Observed p-value**

The fraction of 0-signal pseudo-datasets in which we measure at least the measured cross section.

We also can use the SM ensemble to see how compatible our measured value is with the SM.
Use a pool of weighted signal + background events (about 850k in each of electron and muon)

Fluctuate relative and total yields in proportion to systematic errors

Randomly sample from a Poisson distribution about the total yield

Generate a set of pseudo-data (a member of the ensemble)

Pass the pseudo-data through the full analysis chain (including systematic uncertainties)
Expected p-value $s + t$

### Decision Trees
- **p-value**: 1.9%

### Matrix Elements
- **p-value**: 3.7%

### Bayesian NN
- **p-value**: 9.7%

---

- **DØ Run II Preliminary 910 pb$^{-1}$**
  - **tbq$^b$**
  - **Entries**: 68126
  - **Mean RMS**: 0.525, 0.7062
  - **e+μ-channel**
  - **Full systematics**
  - 1300 entries above observed cross section
  - **p-value**: 1.9e-02
  - **sigma**: 2.1

- **DØ Run II Preliminary**
  - **p-Value = 0.037**
  - **Sig = 1.8σ**

- **Cross Section For Zero Signal Ensembles**
  - **Entries**: 1.66e+04
  - **p-Value**: 0.0965
  - **Sigma**: 1.31
  - **σ90% = 2.9 pb**

---

- **s+t-channels, tbq$^b$**
  - **DØ Run II Preliminary, 910 pb$^{-1}$**
  - **Measured Cross Section**: 2.7$^{+1.5}_{-1.4}$ pb
  - **Bayes Ratio**: 8.0

- **Posterior Density: e+μ, w/ 2+3 Jets and >1 Tag**
  - **σ$^{+1.8}_{-1.5}$**
  - **$\delta_{post} = 0.55$**

- **s+t-channels, tbq$^b$**
  - **DØ Run II Preliminary, 910 pb$^{-1}$**
  - **Measured Cross Section**: 3.2$^{+1.6}_{-1.0}$ pb
  - **Bayes Ratio**: 4.5
Cross-Checks on Data
Cross-check samples

- “$W$+jets”: $=2$ jets, $H_T(\text{lepton}, \not{E}_T, \text{all jets}) < 175$ GeV
- “ttbar”: $=4$ jets, $H_T(\text{lepton}, \not{E}_T, \text{all jets}) > 300$ GeV
- Shown: $tb + tqb$ DT output for $e+jets$

Good agreement of model with data
Matrix Elements Method - Cross-Checks

Look at $H_T$ "sidebands" in 2 and 3 jets
Cross Sections and Significance
Bayesian Neural Network - Observed

Least sensitive (a-priori) analysis sees $2.4\sigma$ effect!
Discriminant output with and without signal component (all channels combined in 1D to “visualize” excess)
Posterior Density: \( e^+\mu \) w/ 2+3 Jets and \( \geq 1 \) Tag

\[
\frac{\sigma_{s+t}}{\sigma_{s+t}} = 0.36
\]

\[
\sigma_{s+t} = 4.6^{+1.8}_{-1.5}
\]
Matrix Elements Method - Summary

DØ Run II Preliminary

<table>
<thead>
<tr>
<th>Process</th>
<th>Cross Section [pb]</th>
</tr>
</thead>
<tbody>
<tr>
<td>e / 2jets / 1tag</td>
<td>2.0^{+2.0}_{-1.7}</td>
</tr>
<tr>
<td>e / 2jets / 2tags</td>
<td>6.0^{+7.3}_{-5.3}</td>
</tr>
<tr>
<td>e / 3jets / 1tag</td>
<td>2.9^{+5.3}_{-2.9}</td>
</tr>
<tr>
<td>e / 3jets / 2tags</td>
<td>16.0^{+17.0}_{-16.0}</td>
</tr>
<tr>
<td>mu / 2jets / 1tag</td>
<td>8.4^{+3.9}_{-3.2}</td>
</tr>
<tr>
<td>mu / 2jets / 2tags</td>
<td>5.9^{+6.5}_{-5.0}</td>
</tr>
<tr>
<td>mu / 3jets / 1tag</td>
<td>5.0^{+5.4}_{-4.2}</td>
</tr>
<tr>
<td>mu / 3jets / 2tags</td>
<td>7.1^{+8.8}_{-7.1}</td>
</tr>
<tr>
<td>Combined (Matrix elements)</td>
<td>4.6^{+1.8}_{-1.5}</td>
</tr>
</tbody>
</table>

Z. Sullivan PRD 70, 114012 (2004), m_t = 175 GeV
p-value = 0.0021, 2.9σ !!
SM compatibility = 21%
Of course, we have 36 different Decision Trees, let’s look at electron, 2 jet, 1 tag:
Decision Trees - Event Characteristics $M(W, b)$

$DT < 0.3$

$DT > 0.55$

$DT > 0.65$

- Excess in high DT output region.
Decision Trees - Event Characteristics $M_{T_W}$

$DT < 0.3$

$DT > 0.55$

$DT > 0.65$

- Excess in high DT output region.
Decision Trees - Observed

s+t-channels, tbtqb  DØ Run II Preliminary, 910 pb⁻¹

- Measured Cross Section
  = $4.9^{+1.4}_{-1.4}$ pb

- Bayes Ratio > 10
A 3.4σ excess!!

DØ Run II Preliminary 910, pb⁻¹

<table>
<thead>
<tr>
<th>tbtqb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entries</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>RMS</td>
</tr>
</tbody>
</table>

Full systematics

24 entries above observed cross section

p-value: 3.5e-04

sigma: 3.4
Consistent with SM?

**SM Ensemble**

- **Entries**: 1910
- **Mean**: 2.87
- **RMS**: 1.604

**e+\(\mu\)-channel**

- **Full systematics**
- **201 entries above observed cross section**
- **p-value**: 1.1e-01
- **sigma**: 1.3
s + t Summary - All methods

DØ Run II

<table>
<thead>
<tr>
<th>Method</th>
<th>Sigma (pb)</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decision trees</td>
<td>4.9</td>
<td>+1.4, -1.4</td>
</tr>
<tr>
<td>Matrix elements</td>
<td>4.6</td>
<td>+1.8, -1.5</td>
</tr>
<tr>
<td>Bayesian NNs</td>
<td>5.0</td>
<td>+1.9, -1.9</td>
</tr>
</tbody>
</table>

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

sigma(p̅p → tb+tqb) [pb]
Correlations - All methods

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

<table>
<thead>
<tr>
<th>Technique</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>
Choose the 50 highest events in each discriminant and look for overlap

<table>
<thead>
<tr>
<th>Technique</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>

Also measured the cross section in 400 members of the 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>100%</td>
<td>100%</td>
<td>29%</td>
</tr>
<tr>
<td>BNN</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>
Measuring $V_{tb}$
Direct access to $V_{tb}$

$V_{CKM} = \begin{pmatrix}
V_{ud} & V_{us} & V_{ub} \\
V_{cd} & V_{cs} & V_{cb} \\
V_{td} & V_{ts} & V_{tb}
\end{pmatrix}$

- Weak interaction eigenstates are not mass eigenstates
- In SM: top must decay to a $W$ and $d$, $s$ or $b$ quark
  - $V_{td}^2 + V_{ts}^2 + V_{tb}^2 = 1$
  - Constraints on $V_{td}$ and $V_{ts}$: $V_{tb} > 0.998$
- New physics that couples to the top quark:
  - $V_{td}^2 + V_{ts}^2 + V_{tb}^2 < 1$
  - No constraint on $V_{tb}$
Given that we now have a measurement of the single top cross section, we can make the first direct measurement of $|V_{tb}|$.

Use the same infrastructure as cross section measurement but make a posterior in $|V_{tb}|^2$.

Caveat: assume SM top quark decays.

Additional theoretical errors are needed (see hep-ph/0408049)

<table>
<thead>
<tr>
<th>Source</th>
<th>$s$</th>
<th>$t$</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>10.0%</td>
</tr>
<tr>
<td>$\alpha_s$</td>
<td>1.4%</td>
<td>0.01%</td>
</tr>
</tbody>
</table>
Measuring $|V_{tb}|^2$

DØ Run II Preliminary, 910 pb$^{-1}$

Posterior Density

Measurement: $1.7^{+0.6}_{-0.5}$
Limiting $|V_{tb}|$

Constrain $|V_{tb}|$ to physical region and integrate:

$$|V_{tb}| = 1.00^{+0.12}_{-0.24}$$
Conclusions

Preliminary First Evidence for Single Top Quark Production!!

- $s + t$ cross section: $4.9 \pm 1.4 \text{pb}$
- $3.4\sigma$ significance!

- Three techniques in good agreement.
- First direct measurement of $|V_{tb}|$

$$|V_{tb}| = 1.00^{+0.12}_{-0.12}$$
Electron ID

- We require electrons to be within the central calorimeter:
  \[ |\eta^{\text{det}}| < 1.1. \]

- **Loose isolated electron**
  At least 90% of the energy of the cluster must be contained in the electromagnetic section of the calorimeter. The \( \chi^2 \) from the \( 7 \times 7 \) H-matrix must be less than 50. The energy deposition in the calorimeter must be matched with a charged particle track from the tracking detectors with \( p_t > 5 \) GeV. Isolation:
  \[ \frac{E_{\text{total}}(R < 0.4) - E_{\text{EM}}(R < 0.2)}{E_{\text{EM}}(R < 0.2)} < 0.15. \]

- **Tight isolated electron**
  A tight isolated electron must pass the loose isolation requirements above, and have a value of the seven-variable EM-likelihood \( \mathcal{L} > 0.85. \)
Muon ID

Loose muons must be of medium $|\text{nseg}| = 3$ quality and pass the loose cosmic ray rejection timing requirements: $|\Delta t(\text{A layer scint, } t_0)| < 10$ ns and $|\Delta t(\text{BC layer scints, } t_0)| < 10$ ns. The track reconstructed in the muon system must match a track reconstructed in the central tracker with $\chi^2/\text{ndof} < 4$. The central track is required to have distance of closest approach (dca) to the primary vertex of $|\text{dca}(x, y)| < 0.2$. Note that the previous analysis imposed a dca significance cut of 3 standard deviations that has been removed now.

Loose muons must be isolated from jets by $\Delta R > 0.5$.

**Tight isolated muon**

Tight isolated muons are loose muons with the additional isolation criteria: (a) the momenta of all tracks in a cone of radius $R < 0.5$ around the muon direction, except the track matched to the muon, add up to less than 20% of the muon $p_T$; and (b) the energy deposited in an annular cone of radius $0.1 < R < 0.4$ around the muon direction is less that 20% of the muon $p_T$. 
$\alpha(Wb\bar{b} + Wc\bar{c}) + Wjj + t\bar{t} + \text{QCD} = \text{Data}$

<table>
<thead>
<tr>
<th>Scale Factor $\alpha$ to Match Heavy Flavor Fraction to Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Electron Channel</strong></td>
</tr>
<tr>
<td>0 tags</td>
</tr>
<tr>
<td>1.53 ± 0.10</td>
</tr>
<tr>
<td>1.48 ± 0.10</td>
</tr>
<tr>
<td>1.50 ± 0.20</td>
</tr>
<tr>
<td>1.72 ± 0.40</td>
</tr>
<tr>
<td>1 tag</td>
</tr>
<tr>
<td>1.29 ± 0.10</td>
</tr>
<tr>
<td>1.58 ± 0.10</td>
</tr>
<tr>
<td>1.40 ± 0.20</td>
</tr>
<tr>
<td>0.69 ± 0.60</td>
</tr>
<tr>
<td>2 tags</td>
</tr>
<tr>
<td>–</td>
</tr>
<tr>
<td>1.71 ± 0.40</td>
</tr>
<tr>
<td>2.92 ± 1.20</td>
</tr>
<tr>
<td>-2.91 ± 3.50</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Muon Channel</strong></td>
</tr>
<tr>
<td>0 tags</td>
</tr>
<tr>
<td>1.54 ± 0.10</td>
</tr>
<tr>
<td>1.50 ± 0.10</td>
</tr>
<tr>
<td>1.52 ± 0.10</td>
</tr>
<tr>
<td>1.38 ± 0.20</td>
</tr>
<tr>
<td>1 tag</td>
</tr>
<tr>
<td>1.11 ± 0.10</td>
</tr>
<tr>
<td>1.52 ± 0.10</td>
</tr>
<tr>
<td>1.32 ± 0.20</td>
</tr>
<tr>
<td>1.86 ± 0.50</td>
</tr>
<tr>
<td>2 tags</td>
</tr>
<tr>
<td>–</td>
</tr>
<tr>
<td>1.40 ± 0.40</td>
</tr>
<tr>
<td>2.46 ± 0.90</td>
</tr>
<tr>
<td>3.78 ± 2.80</td>
</tr>
</tbody>
</table>
Heavy flavour scale factor $\alpha$ measured in the zero tag bins

$\alpha = 1.51 \pm 0.04$
1) Estimate generic jet heavy flavor fraction in ALPGEN Monte Carlo
2) Fit for bottom and charm fraction in generic jet data

Difference between the two outcomes suggests $K = 1.5 \pm 0.4$
Result supported by study using MCFM: J. M. Campbell, J. Houston,
Method 2 at NLO, hep-ph/0405276
Matrix Element Method

![Graphs showing s-channel and t-channel MC events compared to Wbb MC events for DØ Run II.](image-url)
Matrix Element Method
Motivation - New Physics

First Evidence for Single Top

Dec. 8, 2006

Standard Model

Top-flavor ($m_t=1$ TeV)

$Z$-c FCNC ($g_{Zc}=g_Z$

4th family ($V_{ts}=0.5$)

Top-pion ($m_{\pi}=250$ GeV)

PRD63, 014018 (2001)
## Uncertainties

### Relative Systematic Uncertainties

<table>
<thead>
<tr>
<th>Source</th>
<th>Uncertainty</th>
<th>Source</th>
<th>Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t\bar{t}$ cross section</td>
<td>18%</td>
<td>Primary vertex</td>
<td>3%</td>
</tr>
<tr>
<td>Luminosity</td>
<td>6%</td>
<td>Electron reco * ID</td>
<td>2%</td>
</tr>
<tr>
<td>Electron trigger</td>
<td>3%</td>
<td>Electron trackmatch &amp; likelihood</td>
<td>5%</td>
</tr>
<tr>
<td>Muon trigger</td>
<td>6%</td>
<td>Muon reco * ID</td>
<td>7%</td>
</tr>
<tr>
<td>Jet energy scale</td>
<td>wide range</td>
<td>Muon trackmatch &amp; isolation</td>
<td>2%</td>
</tr>
<tr>
<td>Jet efficiency</td>
<td>2%</td>
<td>$\varepsilon_{\text{real}-e}$</td>
<td>2%</td>
</tr>
<tr>
<td>Jet fragmentation</td>
<td>5–7%</td>
<td>$\varepsilon_{\text{real}-\mu}$</td>
<td>2%</td>
</tr>
<tr>
<td>Heavy flavor fraction</td>
<td>30%</td>
<td>$\varepsilon_{\text{fake}-e}$</td>
<td>3–40%</td>
</tr>
<tr>
<td>Tag-rate functions</td>
<td>2–16%</td>
<td>$\varepsilon_{\text{fake}-\mu}$</td>
<td>2–15%</td>
</tr>
</tbody>
</table>