Evidence for single top quark production at DØ

Yann Coadou

CERN
(formerly at Simon Fraser University)

EP Seminar, CERN
30 July 2007
Located outside Chicago, Illinois

The world’s highest-energy accelerator

$\bar{p}p$ collider, centre-of-mass energy 1.96 TeV

Run I: 1992-1996 at 1.8 TeV

Started operating for Run II in March 2001

Upgraded for Run II
- 396 ns bunch spacing
- new Main Injector and Recycler
  ⇒ increased antiproton intensity

Peak luminosity
\[ > 2.5 \cdot 10^{32} \text{ cm}^{-2}\text{s}^{-1} \]
The DØ detector upgrade

- 2 T superconducting solenoid
- silicon detector
- fiber tracker
- preshower detector
- upgraded muon system
- new calorimeter electronics
- upgraded trigger and DAQ

Yann Coadou (CERN) — Evidence for single top quark production at DØ

EP Seminar, CERN 30 July 2007
The collaboration

- 600+ physicists, 89 institutes, 18 countries

The DØ Collaboration
Top quark physics

- Top quark discovered in 1995 by CDF and DØ at the Tevatron
- Heaviest of all fermions
- Couples strongly to Higgs boson
- So far only observed in pairs, only at the Tevatron

Decay mode and Branching fractions
- Rare decays
- Anomalous decays
- CKM matrix element $|V_{tb}|$

Top spin polarization
- Spin correlations

Production cross-section
- Production kinematics
- New Resonance production

Top mass

W helicity

©B.Vachon
Single top quark production

- Never observed before: electroweak production

$s$-channel (tb)

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

$t$-channel (tqb)

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

Why do we care? — $|V_{tb}|$

- Has never been observed before!
- Should happen in SM
- The value of the cross section is a SM test and the first measurement of $|V_{tb}|$

**Direct access to $|V_{tb}|$**

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

- 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 $V_{ts}$:
    - $|V_{tb}| = 0.9991^{+0.00034}_{-0.00004}$

- New physics:
  - $V_{td}^2 + V_{ts}^2 + V_{tb}^2 < 1$
  - no constraint on $V_{tb}$
  - e.g. 4th generation:
    - $0.07 < |V_{tb}| < 0.9993$
Why do we care? — New physics

- s and t cross sections differently sensitive to new physics

**s-channel: charged resonances**

- heavy $W'$ boson in topflavour model (separate interaction for 3rd family)
- charged Higgs boson $H^\pm$ in models with extra Higgs doublets (e.g. MSSM)
- charged top pion in topcolor-assisted technicolor
- 4th generation (reduced cross section from $|V_{tb}| < 1$)
- Kaluza-Klein excited $W_{KK}$, etc...

**t-channel: new interactions**

- flavour-changing neutral currents ($t-Z/\gamma/g-c$ and/or $t-Z/\gamma/g-u$ couplings)
- 4th generation (potentially strong enhancement from large $V_{ts}$)
**Why do we care? — Spin, Higgs, analysis techniques**

### Top quark spin
- Large mass $\Rightarrow$ top quark decays before it can hadronize (no top jets)
- First chance to study a bare quark!
- Top polarization reflected in angular distributions of decay products
- SM predicts high degree of left-handed tops $\Rightarrow$ possible sign of new physics, or help pin down what new physics

### Higgs searches
- Important background to $WH$ associated Higgs production
- As soon as we discover it, somebody will try to get rid of it....

### Advanced analysis techniques
- Test of techniques to extract small signal out of large background
- If tools don’t work for single top, forget about the Higgs and other small signals
- If tools don’t work at Tevatron, not much hope for LHC
It has been challenging for years...

- Several publications since Run I by DØ and CDF
- 7 DØ and 6 CDF PhDs (Dec '06)
- $\sigma_{t\bar{t}}$ only $\sim 2 \times \sigma_{\text{singletop}}$, but has striking signature

---

**Total inelastic**

<table>
<thead>
<tr>
<th>Cross section (barns)</th>
<th>mb</th>
<th>$b\bar{b}$</th>
<th>$\mu b$</th>
<th>W</th>
<th>nb</th>
<th>Z</th>
<th>$t\bar{t}$</th>
<th>pb</th>
<th>single top</th>
<th>Higgs (ZH + WH)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$10^{-2}$</td>
<td></td>
<td></td>
<td>$1 \cdot 10^7$</td>
<td></td>
<td></td>
<td>$6,000$</td>
<td></td>
<td>$600$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$10^{-4}$</td>
<td></td>
<td></td>
<td></td>
<td>$2 \cdot 10^{10}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$2 \equiv 1$</td>
<td></td>
</tr>
<tr>
<td>$10^{-6}$</td>
<td></td>
<td></td>
<td>$1 \cdot 10^7$</td>
<td></td>
<td></td>
<td>$6,000$</td>
<td></td>
<td>$600$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$10^{-8}$</td>
<td></td>
<td></td>
<td></td>
<td>$2 \cdot 10^{10}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$2 \equiv 1$</td>
<td></td>
</tr>
<tr>
<td>$10^{-10}$</td>
<td></td>
<td></td>
<td>$1 \cdot 10^7$</td>
<td></td>
<td></td>
<td>$6,000$</td>
<td></td>
<td>$600$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$10^{-12}$</td>
<td></td>
<td></td>
<td></td>
<td>$2 \cdot 10^{10}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$2 \equiv 1$</td>
<td></td>
</tr>
<tr>
<td>$10^{-14}$</td>
<td></td>
<td></td>
<td>$1 \cdot 10^7$</td>
<td></td>
<td></td>
<td>$6,000$</td>
<td></td>
<td>$600$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$10^{-16}$</td>
<td></td>
<td></td>
<td></td>
<td>$2 \cdot 10^{10}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$2 \equiv 1$</td>
<td></td>
</tr>
</tbody>
</table>
Event selection

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

**Event selection**

- Only one tight (no loose) lepton
  - electron: $p_T > 15$ GeV, $|\eta_{det}| < 1.1$
  - muon: $p_T > 18$ GeV, $|\eta_{det}| < 2$
- $15 < \mathbf{E}_T < 200$ GeV
- 2-4 jets: $p_T > 15$ GeV, $|\eta| < 3.4$
  - Leading jet: $p_T > 25$ GeV, $|\eta_{det}| < 2.5$
  - Second leading jet: $p_T > 20$ GeV
- Mis-reconstructed events: require $\mathbf{E}_T$ direction not aligned or anti-aligned in azimuth with lepton or jet
- One or two $b$-tagged jets
Signal and backgrounds

**Single top signal (m_t = 175 GeV)**
- CompHEP-SingleTop + Pythia

**W+jets**
- Most difficult background
- Alpgen+Pythia (MLM matching between matrix elements and parton shower)
- Heavy flavour fraction and normalization from data

**t\bar{t} (m_t = 175 GeV)**
- Alpgen+Pythia (MLM)
- Normalized to \(\sigma_{NNLO} = 6.8 \text{ pb}\)

**Multijet events**
- misidentified lepton, from data
Event selection — Agreement before $b$ tagging

- Normalize $W+\text{jets}$ and multijet to data before $b$ tagging
- Checked 90 variables, 4 jet multiplicities, electron + muon
- Good description of data

Yann Coadou (CERN) — Evidence for single top quark production at DØ
EP Seminar, CERN 30 July 2007
b-jet tagger

- NN trained on 7 input variables from existing taggers.
  - secondary vertices
  - impact parameter
- 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
- Operating point:
  - $b$-jet efficiency $\sim 50\%$
  - $c$-jet efficiency $\sim 10\%$
  - light jet efficiency $\sim 0.5\%$
### 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>10%</td>
<td>25%</td>
<td>12%</td>
<td>3%</td>
<td>1%</td>
</tr>
<tr>
<td></td>
<td>1:3200</td>
<td>1:390</td>
<td>1:300</td>
<td>1:270</td>
<td>1:230</td>
</tr>
<tr>
<td>1 tag</td>
<td>6%</td>
<td>21%</td>
<td>11%</td>
<td>3%</td>
<td>1%</td>
</tr>
<tr>
<td></td>
<td>1:100</td>
<td>1:20</td>
<td>1:25</td>
<td>1:40</td>
<td>1:53</td>
</tr>
<tr>
<td>2 tags</td>
<td>3%</td>
<td>2%</td>
<td>1%</td>
<td>0%</td>
<td>0%</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>
Systematic uncertainties

- Assigned per background, jet multiplicity, lepton flavour and number of tags
- Uncertainties that affect both normalisation and shapes: jet energy scale and tag rate functions ($b$-tagging parameterisation)
- All uncertainties sampled during limit-setting phase

Relative systematic uncertainties

<table>
<thead>
<tr>
<th>Component</th>
<th>Uncertainty</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 efficiency</td>
<td>2%</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>
<tr>
<td>Primary vertex</td>
<td>3%</td>
</tr>
<tr>
<td>$e$ reco * ID</td>
<td>2%</td>
</tr>
<tr>
<td>$e$ trackmatch &amp; likelihood</td>
<td>5%</td>
</tr>
<tr>
<td>$\mu$ reco * ID</td>
<td>7%</td>
</tr>
<tr>
<td>$\mu$ trackmatch &amp; isolation</td>
<td>2%</td>
</tr>
<tr>
<td>$\varepsilon_{\text{real} - e}$</td>
<td>2%</td>
</tr>
<tr>
<td>$\varepsilon_{\text{real} - \mu}$</td>
<td>2%</td>
</tr>
<tr>
<td>$\varepsilon_{\text{fake} - e}$</td>
<td>3–40%</td>
</tr>
<tr>
<td>$\varepsilon_{\text{fake} - \mu}$</td>
<td>2–15%</td>
</tr>
</tbody>
</table>
Agreement after tagging

![Graph showing yield versus missing transverse energy with data and signal predictions.]

<table>
<thead>
<tr>
<th>Sample</th>
<th># of Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>s&amp;t-channel Signal</td>
<td>62</td>
</tr>
<tr>
<td>Wjj</td>
<td>174</td>
</tr>
<tr>
<td>tt→l+jets</td>
<td>266</td>
</tr>
<tr>
<td>Wbb &amp; Wcc</td>
<td>675</td>
</tr>
<tr>
<td>Mis-ID’s leptons</td>
<td>201</td>
</tr>
<tr>
<td>Diboson,tt→ dileptons</td>
<td>82</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Totals</th>
<th>2 Jets</th>
<th>3 Jets</th>
<th>4 Jets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data</td>
<td>697</td>
<td>455</td>
<td>246</td>
</tr>
<tr>
<td>Total Background</td>
<td>685</td>
<td>460</td>
<td>253</td>
</tr>
<tr>
<td>Signal</td>
<td>36</td>
<td>20</td>
<td>6</td>
</tr>
</tbody>
</table>
Multivariate analysis techniques

- Boosted decision trees
- Matrix element
- Bayesian neural networks
Decision trees

- Machine-learning technique, widely used in social sciences
- Idea: recover events that fail criteria in cut-based analysis

- Start with all events = first node
  - sort all events by each variable
  - for each variable, find splitting value with best separation between two children (mostly signal in one, mostly background in the other)
  - select variable and splitting value with best separation, produce two branches with corresponding events ((F)ailed and (P)assed cut)
- Repeat recursively on each node
- Splitting stops: terminal node = leaf

- DT output = leaf purity, close to 1 (0) for signal (bkg)

**Splitting a node**

**Impurity $i(t)$**
- maximum for equal mix of signal and background
- symmetric in $p_{signal}$ and $p_{background}$
- minimal for node with either signal only or background only
- strictly concave $\Rightarrow$ reward purer nodes

- Decrease of impurity for split $s$ of node $t$ into children $t_L$ and $t_R$ (goodness of split):
  \[ \Delta i(s, t) = i(t) - p_L \cdot i(t_L) - p_R \cdot i(t_R) \]
- Aim: find split $s^*$ such that:
  \[ \Delta i(s^*, t) = \max_{s \in \{\text{splits}\}} \Delta i(s, t) \]

- Maximizing $\Delta i(s, t) \equiv$ minimizing overall tree impurity

**Examples**

- **Gini**
  \[ Gini = 1 - \sum_{i=s,b} p_i^2 = \frac{2sb}{(s+b)^2} \]

- **Entropy**
  \[ entropy = - \sum_{i=s,b} p_i \log p_i \]

\[ criterion \]
\[ \text{--- Gini} \]
\[ \text{--- Entropy} \]
Object Kinematics
\( p_T(jet1) \)
\( p_T(jet2) \)
\( p_T(jet3) \)
\( p_T(jet4) \)
\( p_T(best1) \)
\( p_T(notbest1) \)
\( p_T(false2) \)
\( p_T(tag1) \)
\( p_T(untag1) \)
\( p_T(untag2) \)

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

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

- Adding variables does not degrade performance
- Tested shorter lists, lost some sensitivity
- Same list used for all channels
Measure and apply

- Take trained tree and run on independent pseudo-data 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

Advantages

- DT has human readable structure (no black box)
- Training is fast
- Deals with discrete variables
- No need to transform inputs
- Resistant to irrelevant variables

Limitations

- Instability of tree structure
- Piecewise nature of output
Boosting a decision tree

**Boosting**
- Recent technique to improve performance of a weak classifier
- Recently used on decision trees by GLAST and MiniBooNE
- Basic principle 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 by $e^{\alpha_k}$
- Train again to build $T_{k+1}$
- Boosted result of event $i$: $T(i) = \sum_{k=1}^{N_{\text{tree}}} \alpha_k T_k(i)$

- Averaging $\Rightarrow$ dilutes piecewise nature of DT
- Usually improves performance

Decision tree parameters

<table>
<thead>
<tr>
<th>DT choices</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/3 of MC for training</td>
</tr>
<tr>
<td>AdaBoost parameter $\beta = 0.2$</td>
</tr>
<tr>
<td>20 boosting cycles</td>
</tr>
<tr>
<td>Signal leaf if purity $&gt; 0.5$</td>
</tr>
<tr>
<td>Minimum leaf size $= 100$ events</td>
</tr>
<tr>
<td>Same total weight to signal and background to start</td>
</tr>
<tr>
<td>Goodness of split - Gini factor</td>
</tr>
</tbody>
</table>

Analysis strategy

- Train 36 separate trees:
  - 3 signals ($s,t,s + t$)
  - 2 leptons ($e,\mu$)
  - 3 jet multiplicities (2,3,4 jets)
  - 2 $b$-tag multiplicities (1,2 tags)
- For each signal train against the sum of backgrounds
Matrix element method

- Pioneered by DØ top mass analysis. Now used in search
- Use the 4-vectors of all reconstructed leptons and jets
- Use matrix elements of main signal and bkgd diagrams to compute event probability density for signal and bkgd hypotheses
- Goal: calculate a discriminant:

\[ D_s(\vec{x}) = P(S|\vec{x}) = \frac{P_{\text{signal}}(\vec{x})}{P_{\text{signal}}(\vec{x}) + P_{\text{bkg}}(\vec{x})} \]

- Encoded in normalized differential cross section for process S:

\[ P_S(\vec{x}) = \frac{1}{\sigma_S} d\sigma_S(\vec{x}), \quad \sigma_S = \int d\sigma_S(\vec{x}) \]

Used only limited number of Feynman diagrams

- Sensitivity would increase (but so does computation time) if more diagrams were included. In particular, no \( t\bar{t} \) diagrams are computed (serious limitation for \( >2 \) jets)
Bayesian neural networks

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
- Used 25 variables (subset of DT variables)
- Same strategy as DT: 36 different BNN

Advantages
- Less prone to overtraining
- Details of each network not important

Limitation
- Darker black box
- Computationally demanding

Implementation: Flexible Bayesian Modeling (FBM) package
http://www.cs.toronto.edu/~radford/fbm.software.html
Analysis validation

Ensemble testing
- Test the whole machinery with many sets of pseudo-data
- Like running DØ experiment 1000s of times
- Generated ensembles with different signal contents (no signal, SM, other cross sections, higher luminosity)

Ensemble generation
- Pool of weighted signal + background events
- Fluctuate relative and total yields in proportion to systematic errors, reproducing correlations
- Randomly sample from a Poisson distribution about the total yield to simulate statistical fluctuations
- Generate pseudo-data set, pass through full analysis chain (including systematic uncertainties)

All analyses achieved linear response to varying input cross sections and negligible bias
Cross-check samples

- Validate methods on data in no-signal region
- \( W+\text{jets} \): 2 jets, \( H_T(\text{lepton}, \slash E_T, \text{all jets}) < 175 \text{ GeV} \)
- \( \text{ttbar} \): 4 jets, \( H_T(\text{lepton}, \slash E_T, \text{all jets}) > 300 \text{ GeV} \)
- Good agreement
Use the 0-signal ensemble

**Expected p-value**
Fraction of 0-signal pseudo-datasets in which we measure at least 2.9 pb (SM single top cross section)

**Observed p-value**
Fraction of 0-signal pseudo-datasets in which we measure at least the observed cross section.

Also use the SM ensemble to check compatibility of observed result with SM prediction
Expected sensitivity $s+t$

### Decision trees

**p-value 1.9% (2.1σ)**

### Matrix elements

**p-value 3.7% (1.8σ)**

### Bayesian NN

**p-value 9.7% (1.3σ)**
Matrix element

\[ \sigma = 4.6^{+1.8}_{-1.5} \text{ pb} \]
\[ \text{p-value} = 0.21\% \ (2.9 \sigma) \]
\[ \text{SM compatibility 21\%} \]

Bayesian NN

\[ \sigma = 5.0 \pm 1.9 \text{ pb} \]
\[ \text{p-value} = 0.89\% \ (2.4 \sigma) \]
\[ \text{SM compatibility 18\%} \]

New preliminary ME result

- Included \( t\bar{t} \rightarrow \ell + \text{jets} \) ME in 3-jet discriminant

\[ \sigma = 4.8^{+1.6}_{-1.4} \text{ pb} \]
\[ \text{exp. p-value} = 3.1\% \ (1.9 \sigma) \]
\[ \text{obs. p-value} = 0.082\% \ (3.2 \sigma) \]

New preliminary BNN result

- Better treatment of noisy training data

\[ \sigma = 4.4^{+1.6}_{-1.4} \text{ pb} \]
\[ \text{exp. p-value} = 1.6\% \ (2.2 \sigma) \]
\[ \text{obs. p-value} = 0.083\% \ (3.1 \sigma) \]

ME discriminant output, with and without signal content (all channels combined)
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 for single top production!

\[ \sigma_s = 1.0 \pm 0.9 \text{ pb} \]

\[ \sigma_t = 4.2^{+1.8}_{-1.4} \text{ pb} \]
Boosted decision tree event characteristics

$DT < 0.3$

$DT > 0.55$

$DT > 0.65$

Yann Coadou (CERN) — Evidence for single top quark production at DØ

EP Seminar, CERN 30 July 2007
Now that we have a cross section measurement, we can make the first direct measurement of $|V_{tb}|$

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

**Additional theoretical errors (hep-ph/0408049)**

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

Most general $Wtb$ coupling ($P_{L,R} = (1 \mp \gamma_5)/2$):

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

SM: $f_1^L = 1$, $f_1^R = 0$ (pure $V-A$), $f_2^L = f_2^R = 0$ (CP conservation)

Effectively measuring strength of $V-A$ coupling $|V_{tb}f_1^L|$, can be $> 1$
First direct measurement of $|V_{tb}|$

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

\[ |V_{tb}f_1^L| = 1.3 \pm 0.2 \]

\[ 0.68 < |V_{tb}| \leq 1 \text{ @ 95\% CL} \]

(assuming $f_1^L = 1$, flat prior in $[0,1]$)

- No assumption about number of quark families or CKM matrix unitarity
New: combination of s+t analyses

Correlations

- 3 analyses with similar performance on same dataset
- Combined using BLUE method

<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>64%</td>
<td>66%</td>
</tr>
<tr>
<td>ME</td>
<td>100%</td>
<td>59%</td>
<td></td>
</tr>
<tr>
<td>BNN</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

DØ Run II

* = preliminary

<table>
<thead>
<tr>
<th>Analysis</th>
<th>Cross Section [pb]</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decision Trees</td>
<td>4.9 ±1.4</td>
<td>3.4σ</td>
</tr>
<tr>
<td>Matrix Elements*</td>
<td>4.8 ±1.6</td>
<td>3.2σ</td>
</tr>
<tr>
<td>Bayesian NNs*</td>
<td>4.4 ±1.6</td>
<td>3.1σ</td>
</tr>
<tr>
<td>Combination*</td>
<td>4.7 ±1.3</td>
<td>3.6σ</td>
</tr>
</tbody>
</table>
Conclusion

First evidence for single top quark production (DØ decision trees)

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

3.4\( \sigma \) significance

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

\[ |V_{tb}f_1^L| = 1.3 \pm 0.2 \]

assuming \( f_1^L = 1 \): \[ 0.68 < |V_{tb}| \leq 1 \atop 95\% \text{ CL} \]

(Always assuming \( V_{td}^2 + V_{ts}^2 \ll V_{tb}^2 \) and pure \( V-A \) and CP-conserving \( Wtb \) interaction)


New preliminary combination of DT, ME and BNN

\[ \sigma(p\bar{p} \rightarrow tb + X, tqb + X) = 4.7 \pm 1.3 \text{ pb} \]

3.6\( \sigma \) significance

- A lot more data already at hand
Single top prospects — Tevatron and LHC

**Tevatron**
- By 2009 we should have observed single top production and measured its cross section to 15-20%.
- $|V_{tb}|$ is then known to $\sim 10\%$

**LHC**
- Much larger production rates:
  - $\sigma_{s}^{t/\bar{t}} = 6.6/4.1 \text{ pb (}\pm 10\%)$
  - $\sigma_{t}^{t/\bar{t}} = 156/91 \text{ pb (}\pm 5\%)$
  - $\sigma_{tW}^{t/\bar{t}} = 34/34 \text{ pb (}\pm 10\%)$
- Try to observe all three channels (s-channel challenging)
- $|V_{tb}|$ measured to percent level
- Large samples $\Rightarrow$ study properties
More information:
http://www-d0.fnal.gov/Run2Physics/top/public/fall06/singletop