Evidence for single top quark production at D0 using boosted decision trees

John BackusMayes
University of Washington

APS April Meeting
Jacksonville, FL

17 April, 2007

On behalf of the D0 collaboration

Single Top Production

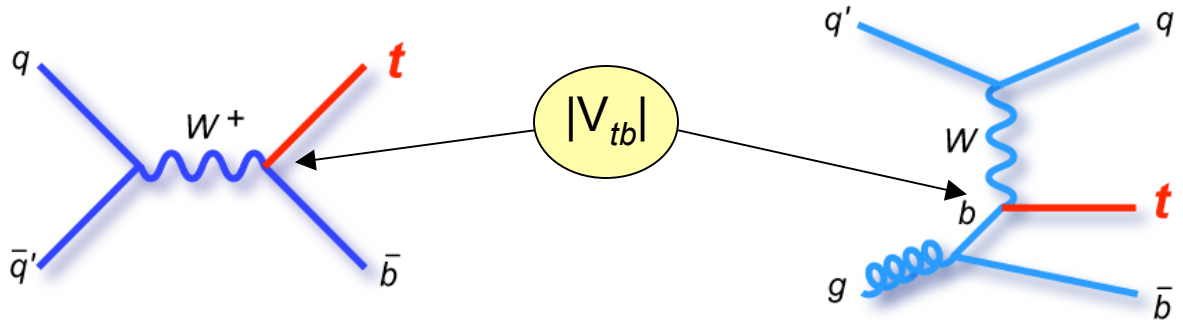
Weak interaction vertex
produces *one* top quark

→ Direct measurement of $|V_{tb}|$

→ Sensitive to a fourth quark family

Also important:

Single top is background
to WH search!



s-channel (*tb*):

- SM cross-section 0.88 ± 0.11 pb
- Sensitive to new resonances

t-channel (*tqb*):

- SM cross section 1.98 ± 0.25 pb
- Sensitive to FCNC

What the Detector Sees

Experimental Signature:

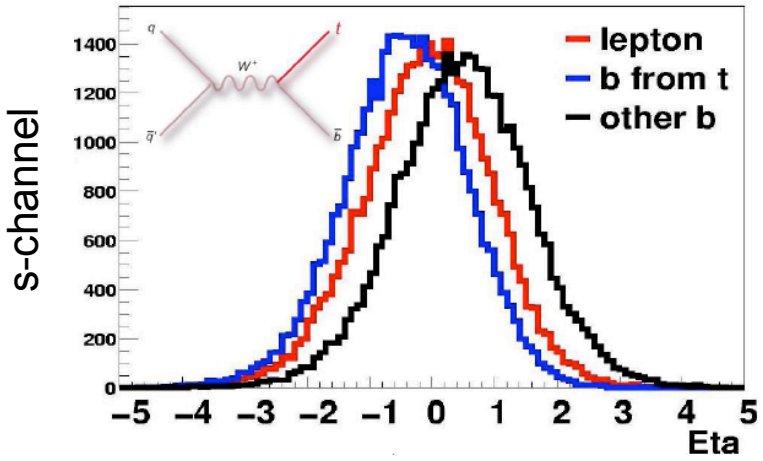
- One high-p_T central lepton
- Significant missing E_T
- 2-4 jets, with at least one *b* jet
- For specifics on event selection, see talk by Monica Pangilinan

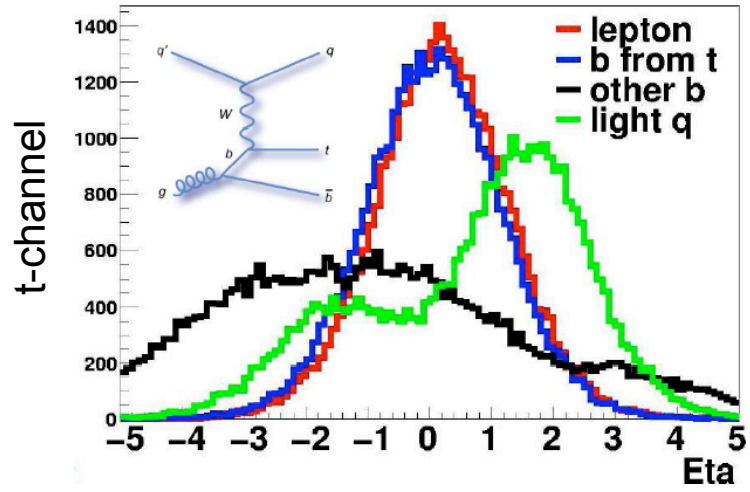
Backgrounds:

- *W* + jets
- top quark pair production
- fake-lepton multijet events

Multivariate Analyses:

- Bayesian NN (Monica Pangilinan)
- Matrix Element (Chris Potter)
- **Boosted Decision Trees**

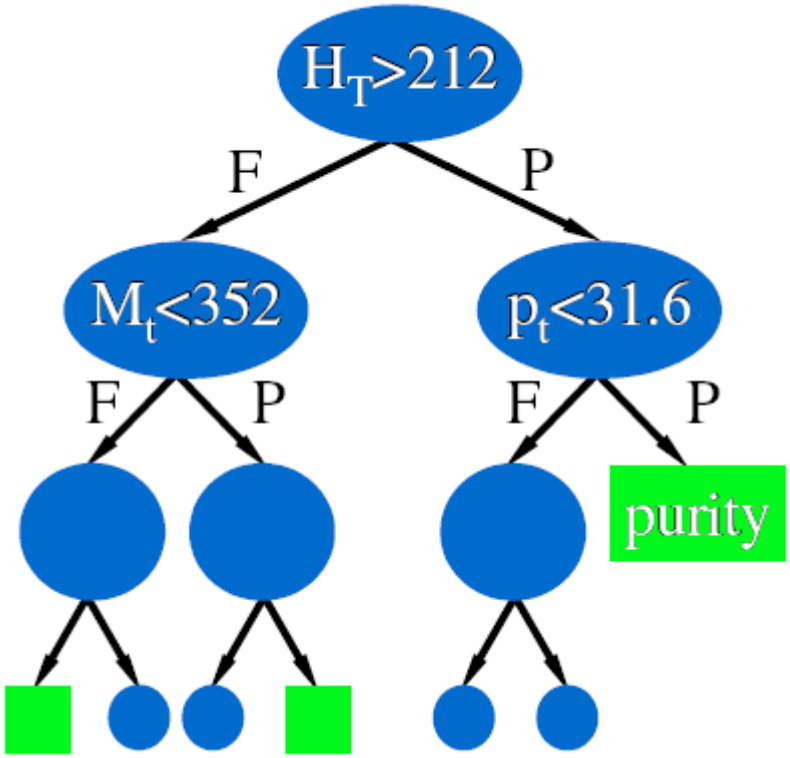




Decision Trees in Particle Experiment

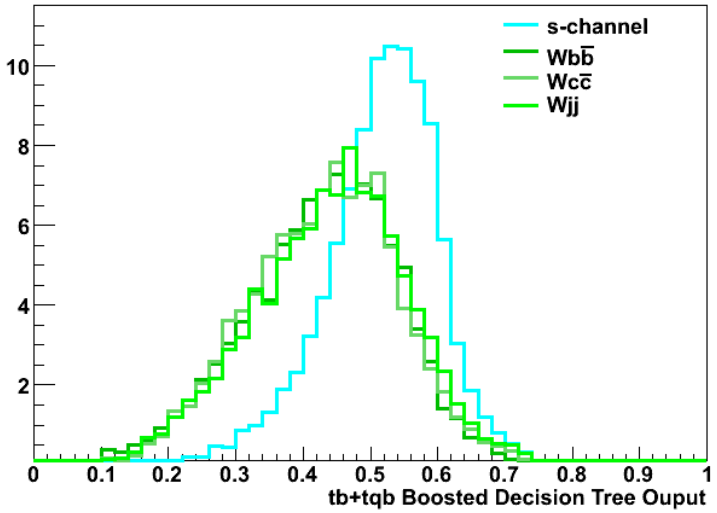
mathematically, they're *rooted binary trees*…

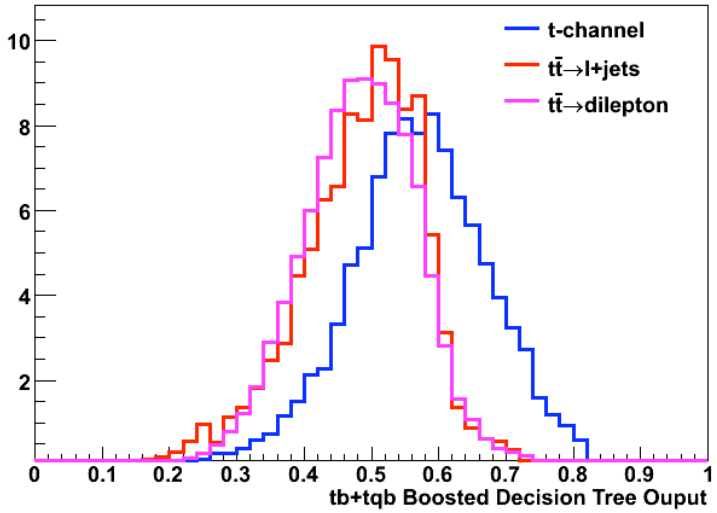
- Each *node* (blue) simultaneously represents a set of events *and* a cut
- Two child nodes split each node into events that pass cut and events that fail
- Cuts chosen to achieve maximal separation between signal and background (see slide 6)
- Splitting terminates eventually, forming *leaves* (green)
- Associated with each leaf is its *purity*, the weighted fraction of events in that leaf that are signal



Analysis via Decision Trees

- The decision tree sorts data into many disjoint sets, corresponding to its leaves
- Output for each event is the purity of the corresponding leaf
- Decision tree output serves as *discriminant*—by design, output distributions vary greatly between signal and background
- We use discriminant distributions to perform a *shape-based* analysis on the data





Decision Tree Training

Need a measure of *impurity* for nodes—such a function must be:

- Maximal for equal mix of signal and background
- Minimal for nodes with only signal or background
- Symmetric about purity $p = ½$

→ Use *Gini* measure:

$$Gini \propto p(1-p)$$

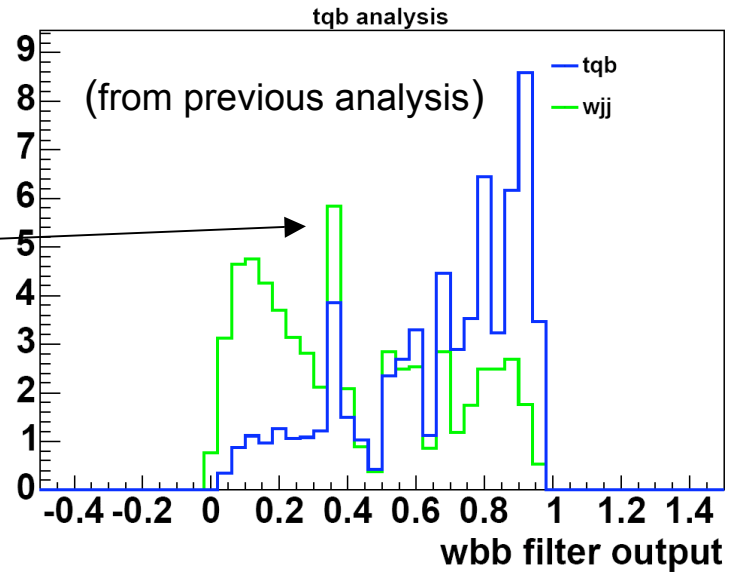
Training Procedure:

- Training sample: 1/3 of signal and background models
- List of 49 discriminating variables
- For each variable, choose cut to **maximize decrease in impurity**
- Choose best cut from list
- Split sample based on cut, then repeat process on resulting subsets
- Stop when split would produce subset with <100 events
- Test performance with other 2/3 of signal and background model

Boosting

Limitations of *one* tree:

- Small changes in training sample → big changes in decision tree
- Decision tree output often has weird spikes
- Many events *misclassified*—e.g. signal event in leaf that's primarily background



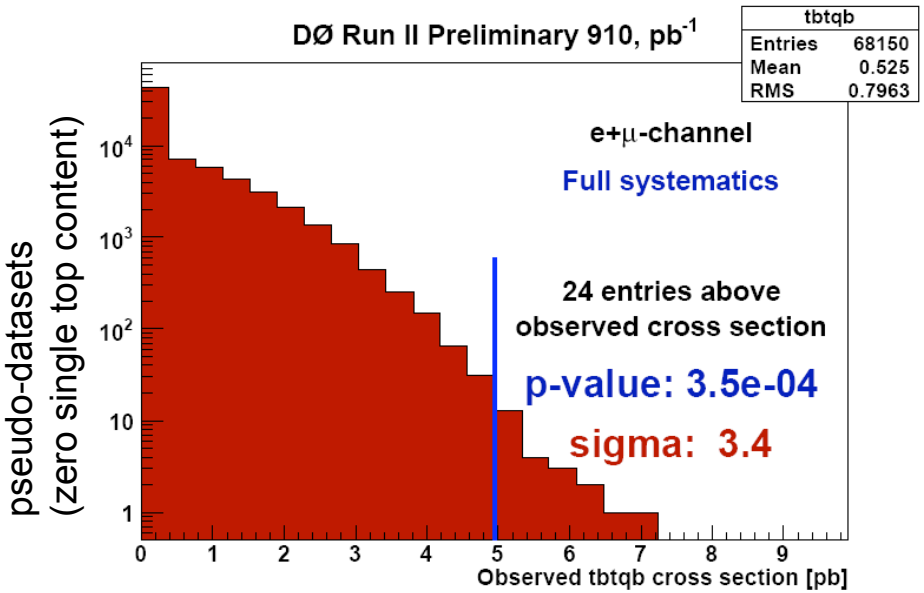
Improvement through *Boosting*:

- Train sequence of decision trees—we use 20
- Reweight training sample for each tree to focus on misclassified events
- For discriminant, use weighted average of decision tree outputs
- Weight trees according to fraction of events that were misclassified

Results

Expected: $\sigma_{tb+tqb} = 2.7^{+1.6}_{-1.4}$ pb

Observed: $\sigma_{tb+tqb} = 4.9^{+1.4}_{-1.4}$ pb



- p-value significance: 24/68150 = 0.00035
- Gaussian equivalent: **3.4 sigma evidence!**
- SM Compatibility: 11%

(for more details on determining significance, see talk by Monica Pangilinan)

$|V_{tb}|$ Measurement

$\sigma_{tb+tqb} \propto |V_{tb}|^2 \rightarrow$ Can determine $|V_{tb}|$ with same procedure as used for cross-section

We assume:

- SM top quark decay ($V_{td}^2 + V_{ts}^2 << V_{tb}^2$)
- CP-conserving V-A coupling (all form factors zero except f_1^L)

Results:

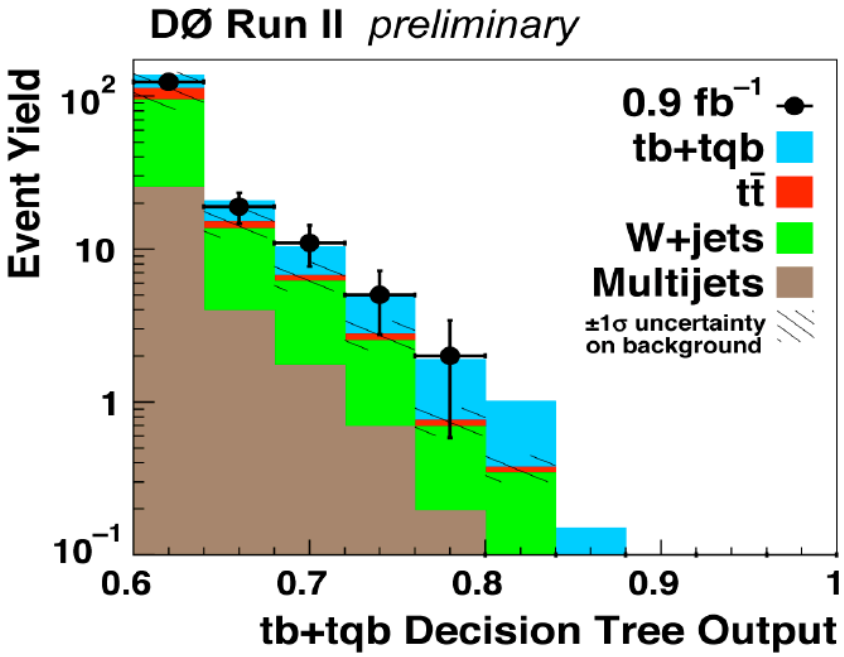
- $|V_{tb}f_1^L| = 1.3 \pm 0.2$
- **$|V_{tb}| > 0.68$ at 95% C.L.** assuming $f_1^L = 1$ (SM)

Not assumed:

- 3 quark families
- CKM unitarity

Summary

Using boosted decision trees, we see **3.4 sigma evidence** for single top!



$\sigma_{tb+tqb} = 4.9 \pm 1.4$ pb

First direct measurement of $|V_{tb}|$:

$0.68 < |V_{tb}| \leq 1$ at 95% C.L.

(assuming SM form factors)