Seeking Single Top at DØ

Dugan O'Neil

Simon Fraser University

Sambamurti Memorial Lecture (BNL) 17/07/07

Outline

- Introduction to Particle Physics
- Introduction to DØ
- Top Quark Physics
- Single Top - The Challenge
- Decision Trees
- Measuring the Cross Section
- First Direct Measurement of $|\mathrm{V_{tb}}|$
- Conclusions

Particle Physics

At the smallest distance scales, what is the world made of? How do those components interact?

By convention there is color,
By convention sweetness,
By convention bitterness,
But in reality there are atoms
and space.
-Democritus (c. 400 BCE)

An aside: Pie-ology

Pie-ology

At the smallest distance scales, what is a pie made of? How do those components interact to make a tasty dessert?

- Given no list of ingredients nor any description of its preparation, "understand the pie".

An aside: Pie-ology

Pie-ology

At the smallest distance scales, what is a pie made of? How do those components interact to make a tasty dessert?

- Given no list of ingredients nor any description of its preparation, "understand the pie".
- Use a photon-scattering experiment using light-sensitive detector. Process images through complex neural net. (ie. look)

Digitized image from visible-light scatter experiment.

An aside: Pie-ology

Pie-ology

At the smallest distance scales, what is a pie made of? How do those components interact to make a tasty dessert?

- Given no list of ingredients nor any description of its preparation, "understand the pie".
- Use a photon-scattering experiment using light-sensitive detector. Process images through complex neural net. (ie. look)
- Chemically analyze escaping gases. Apply complex neural net (ie. smell)

Digitized image from visible-light scatter experiment.

An aside: Pie-ology

Pie-ology

At the smallest distance scales, what is a pie made of? How do those components interact to make a tasty dessert?

- Given no list of ingredients nor any description of its preparation, "understand the pie".
- Use a photon-scattering experiment using light-sensitive detector. Process images through complex neural net. (ie. look)
- Chemically analyze escaping gases. Apply complex neural net (ie. smell)
- I have understood some features of the pie, but I could not make my own...

Digitized image from visible-light scatter experiment.

Elementary Pie-ology

- To understand it, we must eat it.

Pie-ology is a very competitive field.

Elementary Pie-ology

- To understand it, we must eat it.
- Break down the components and analyze each one (with ice cream if possible)

Pie-ology is a very competitive field.

Elementary Pie-ology

- To understand it, we must eat it.
- Break down the components and analyze each one (with ice cream if possible)
- Establish an ingredient list: apples, sugar, flour, water

Pie-ology is a very competitive field.

Elementary Pie-ology

- To understand it, we must eat it.
- Break down the components and analyze each one (with ice cream if possible)
- Establish an ingredient list: apples, sugar, flour, water
- Recreate the pie in the laboratory, understand relative abundances of ingredients, describe their interactions.

Pie-ology is a very competitive field.

Elementary Pie-ology

- To understand it, we must eat it.
- Break down the components and analyze each one (with ice cream if possible)
- Establish an ingredient list: apples, sugar, flour, water
- Recreate the pie in the laboratory, understand relative abundances of ingredients, describe their interactions.
- Write GEANT4-based pie-simulator

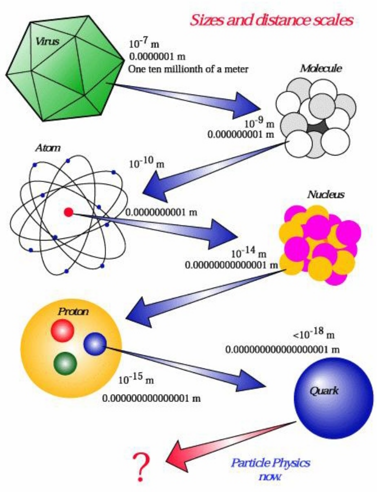
Pie-ology is a very competitive field.

Elementary Pie-ology

- To understand it, we must eat it.
- Break down the components and analyze each one (with ice cream if possible)
- Establish an ingredient list: apples, sugar, flour, water
- Recreate the pie in the laboratory, understand relative abundances of ingredients, describe their interactions.
- Write GEANT4-based pie-simulator
- Warning: closely-related types of apples are difficult to distinguish. Even the same apple types can taste different if prepared differently.

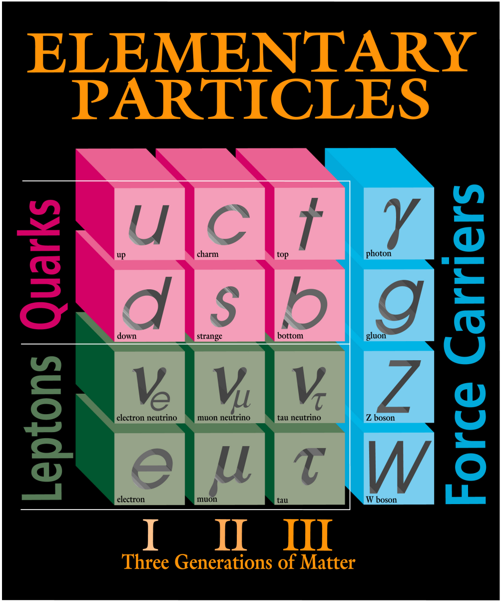
Pie-ology is a very competitive field.

Particle Physicists Don't Eat



- Of course, particle physicists don't probe by eating...we use accelerators and detectors to tell us about structure.
- Historically we go to ever smaller scales.
- On the way down in scale we have discovered hundreds of particles
- However, the fundamental ones are few...

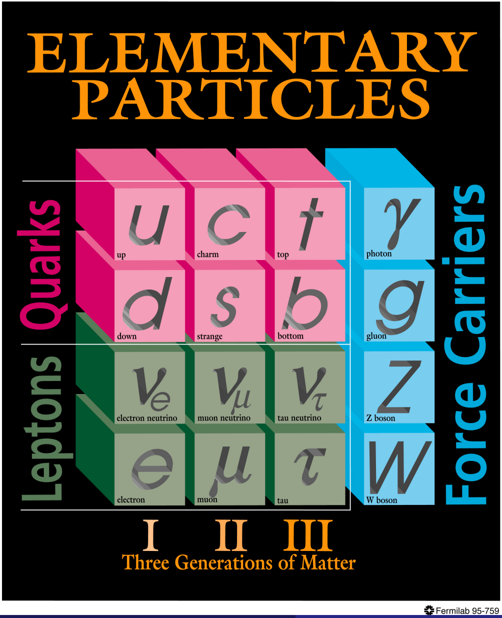
The Standard Model - List of Ingredients



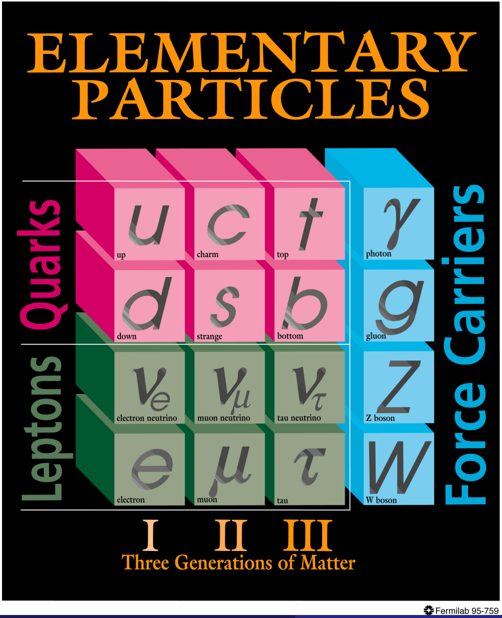
Fermilab 95-759

- Direct searches for what's not here (secret ingredients)

The Standard Model - List of Ingredients

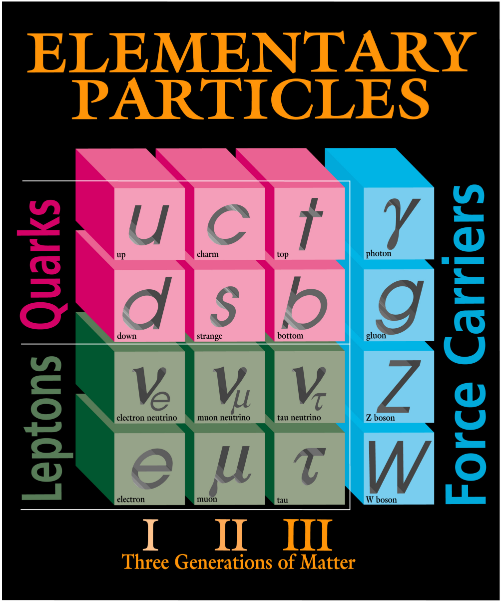


- Direct searches for what's not here (secret ingredients)
- Precision measurements of what's here



- Direct searches for what's not here (secret ingredients)
- Precision measurements of what's here
- Relationships between members of the chart

The Standard Model - List of Ingredients



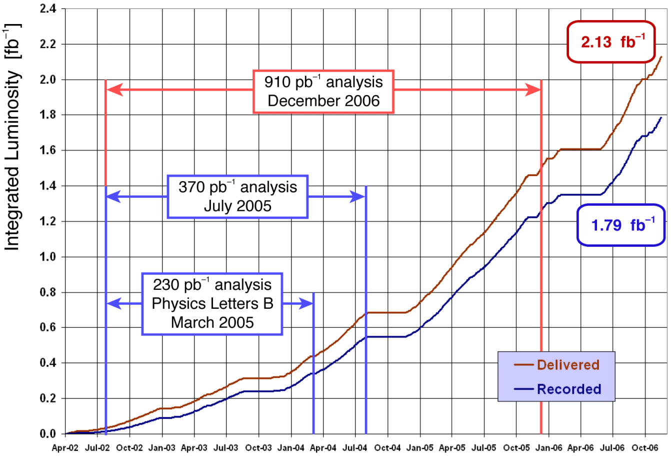
- Direct searches for what's not here (secret ingredients)
- Precision measurements of what's here
- Relationships between members of the chart
- New measurements of SM parameters

The Tools - Introduction to DØ

The Fermilab Tevatron

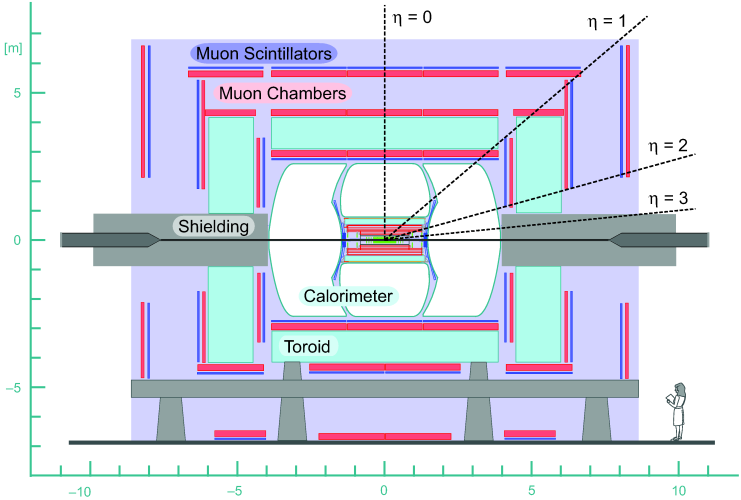
- Run II began in March 2001
- Proton-antiproton collisions at 1.96TeV
- Luminosity up to $2.5 \times 10^{32} cm^{-2} s^{-1}$ (so far)
- Integrated Luminosity (recorded) $>$2.7 fb^{-1} (billions of events recorded)

Tevatron Luminosity



Many thanks to the Accelerator Division

The DØ Cartoon

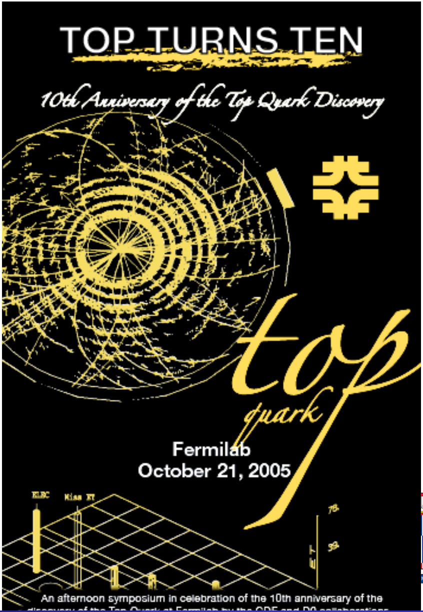


The Collaboration

The People

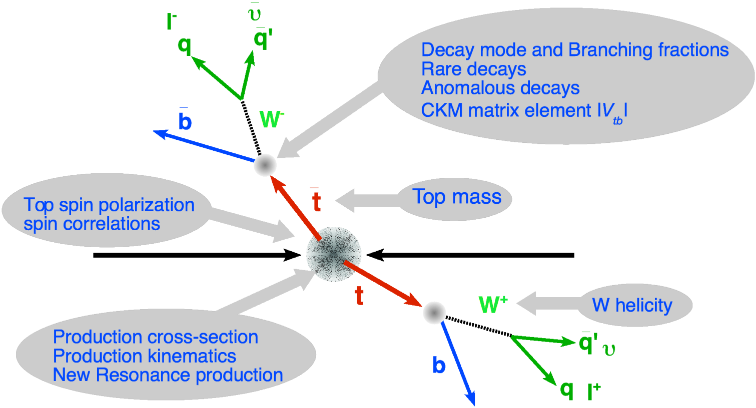
The Physics - Top Quarks

Top Quark Physics



Top Quark Physics

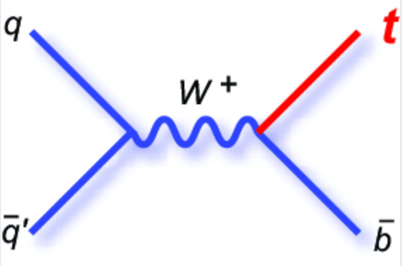
The Tevatron is still the only place to make top quarks. We learn a lot from pair production via the strong interaction.



Single top quark production

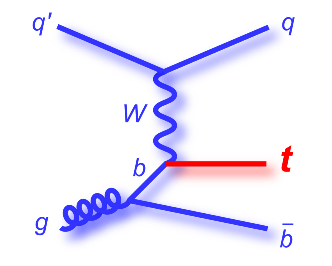
- But we have never observed electroweak production!!

s-channel



- $\sigma = 0.88 \pm 0.11$ pb
- published limits (95% C.L.):
  Run II DØ: < 5.0 pb
  Run II CDF: < 3.1 pb

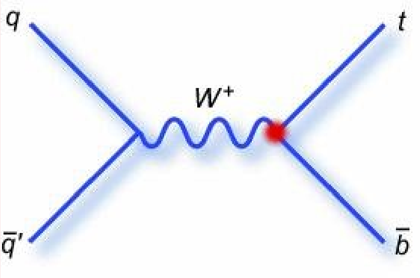
t-channel



- $\sigma = 1.98 \pm 0.25$ pb
- published limits (95% C.L.):
  Run II DØ: < 4.4 pb
  Run II CDF: < 3.2 pb

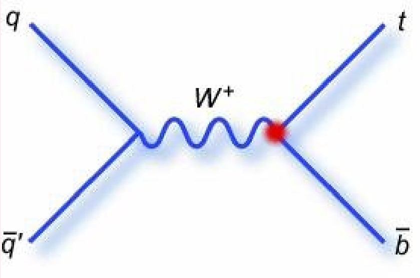
Why do we care? - $|V_{tb}|$

- Has never been observed before!



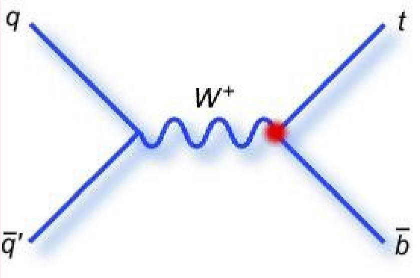
Why do we care? - $|V_{tb}|$

- Has never been observed before!
- It should happen (if SM is right)



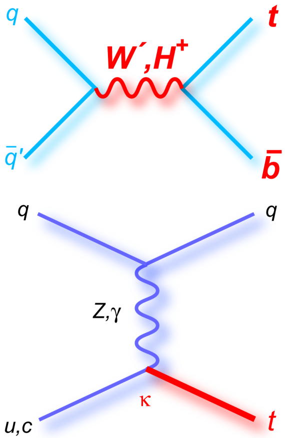
Why do we care? - $|V_{tb}|$

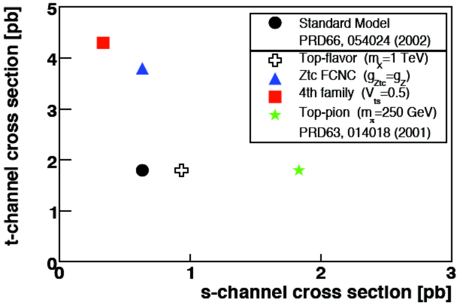
- Has never been observed before!
- It should happen (if SM is right)
- The value of the cross section is a SM test and the first measurement of $|V_{tb}|$ - more later



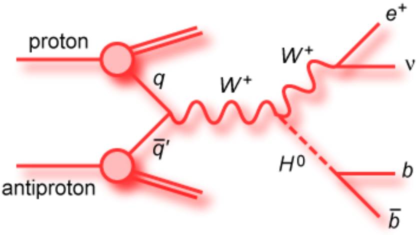
Why do we care? - New Physics

The s-channel and t-channel are sensitive to different new physics



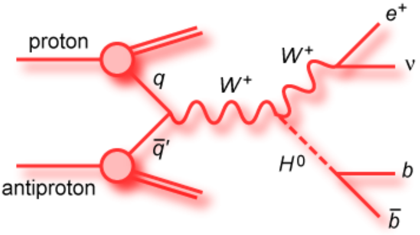


Why do we care? - Higgs Backgrounds, Top Spin

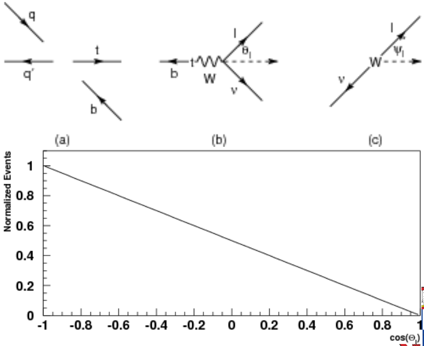


- This looks a lot like single top!
- As soon as we discover it, somebody will try to get rid of it....

Why do we care? - Higgs Backgrounds, Top Spin



- Top decays before it can hadronize (no top jets)
- First chance to measure the polarization of a bare quark!



- This looks a lot like single top!
- As soon as we discover it, somebody will try to get rid of it....

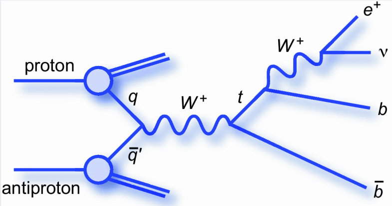
Single Top - The Challenge

Single Top - The Challenge

What precisely are we looking for??

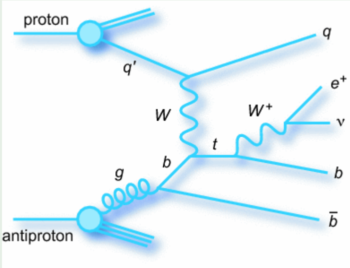
- Electroweak production in two main mechanisms at the Tevatron:

s-channel



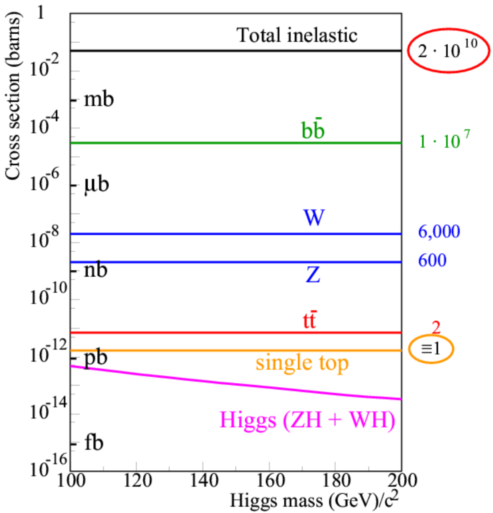
- What to look for (tb):
  - An isolated lepton
  - 2 b-jets
  - Missing transverse energy

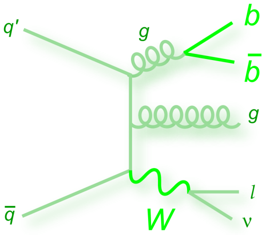
t-channel



- What to look for (tqb):
  - An isolated lepton
  - 1 or 2 b-jets
  - Missing transverse energy
  - A light-quark jet

So, just find it already!





We HAVE been looking!

1. V.M. Abazov *et al.*, "Search for Single Top Quark Production at DØ Using Neural Networks," Phys. Lett. **B 517**, 282 (2001).
2. The Single Top Working Group, "Search for Single Top Quark Production at DØ in Run II," DØ Note 4398 (2004).
3. The Single Top Working Group, "Improved Search for Single Top Quark Production," DØ Note 4670 (2005).
4. V.M. Abazov *et al.*, "Search for Single Top Quark Production in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV," Phys. Lett. **B 622**, 265 (2005).
5. V.M. Abazov *et al.*, "Multivariate Searches for Single Top Quark Production with the DØ Detector," submitted to Phys. Rev. D, hep-ex/0604020.

plus 7 PhDs.

Making the Background Model Agree with Data

- We have reconstructed data (electrons, muons and jets, etc).

Making the Background Model Agree with Data

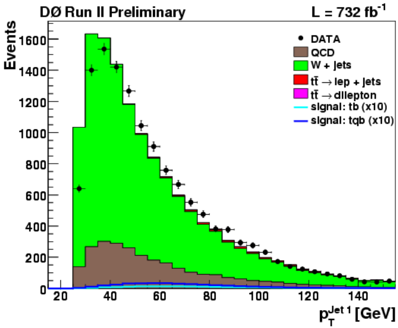
- We have reconstructed data (electrons, muons and jets, etc).
- We have produced a detailed background model.

Making the Background Model Agree with Data

- We have reconstructed data (electrons, muons and jets, etc).
- We have produced a detailed background model.
- We have built an analysis software suite to handle this new data.

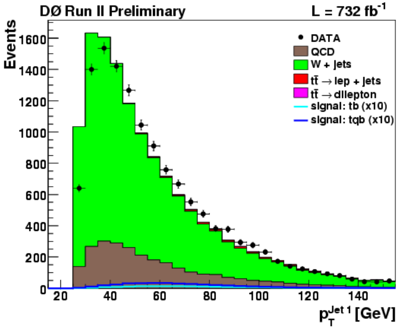
Making the Background Model Agree with Data

- We have reconstructed data (electrons, muons and jets, etc).
- We have produced a detailed background model.
- We have built an analysis software suite to handle this new data.
- We compare background model to data.



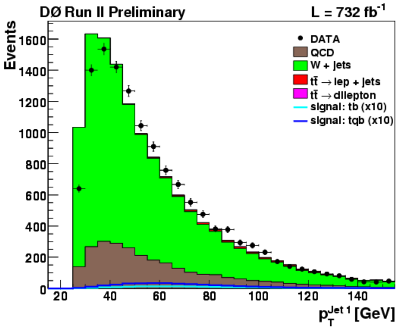
Making the Background Model Agree with Data

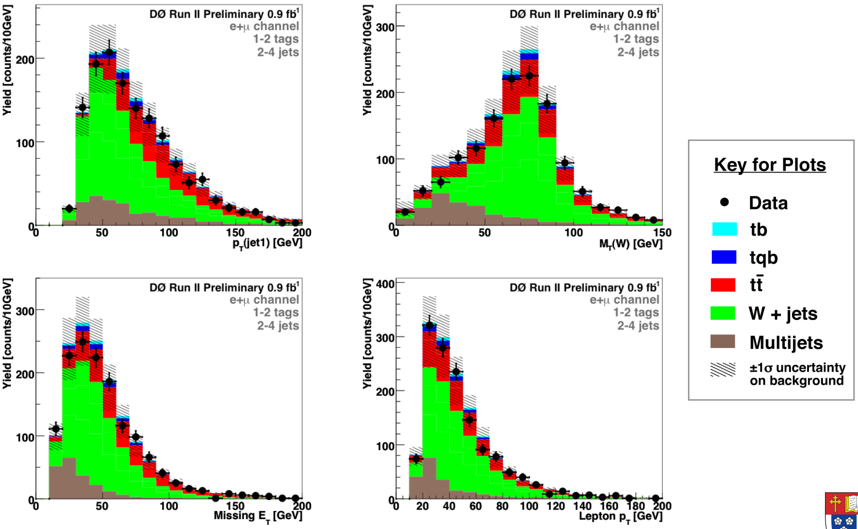
- We have reconstructed data (electrons, muons and jets, etc).
- We have produced a detailed background model.
- We have built an analysis software suite to handle this new data.
- We compare background model to data.
- We get upset. We go home and complain to spouse about the injustice of it all.



Making the Background Model Agree with Data

- We have reconstructed data (electrons, muons and jets, etc).
- We have produced a detailed background model.
- We have built an analysis software suite to handle this new data.
- We compare background model to data.
- We get upset. We go home and complain to spouse about the injustice of it all.
- Roughly 1 year later....





About 2000 of these to look at!

Modeling and Preselection Complete

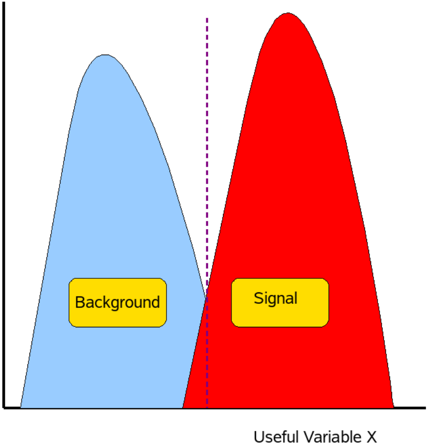
Percentage of single top *tb+tqb* selected events and S:B ratio (white squares = no plans to analyze)

Electron + Muon	1 jet	2 jets	3 jets	4 jets	≥ 5 jets
0 tags	10% 1 : 3,200	25% 1 : 390	12% 1 : 300	3% 1 : 270	1% 1 : 230
1 tag	6% 1 : 100	21% 1 : 20	11% 1 : 25	3% 1 : 40	1% 1 : 53
2 tags		3% 1 : 11	2% 1 : 15	1% 1 : 38	0% 1 : 43

Decision Trees

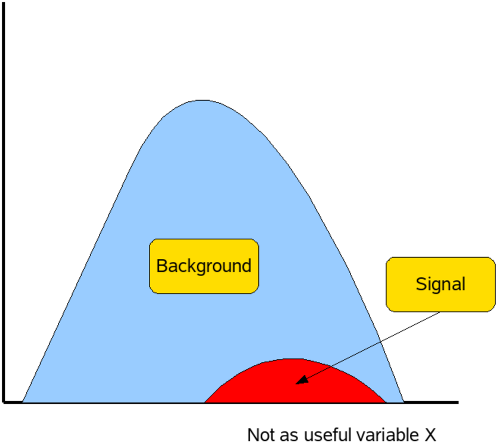
Improving S/B

- Normally, we look for variables to distinguish signal from background and make a “cut”:



Improving S/B

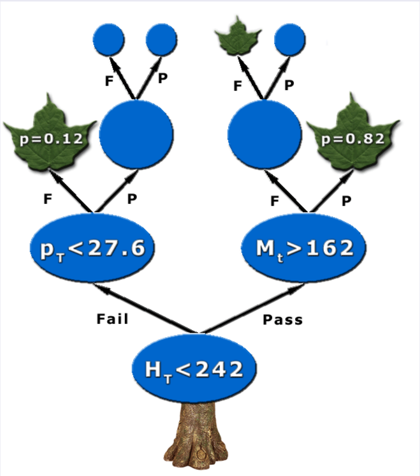
- Unfortunately, we have a whole lot of this:



Decision Trees

Train

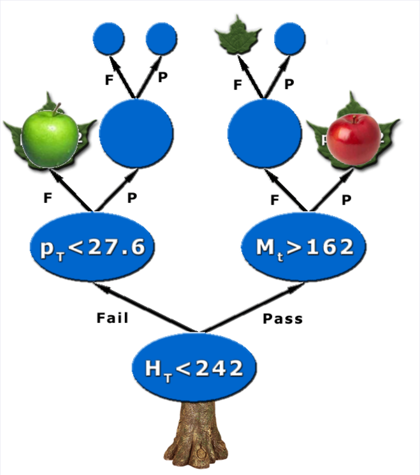
- Start with all events (first node)
- For each variable, find the splitting value with best separation between children (best cut).
- select best variable and cut and produce Failed and Passed branches
- Repeat recursively on each node
- Stop when improvement stops or when too few events left. Terminal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



Decision Trees

Train

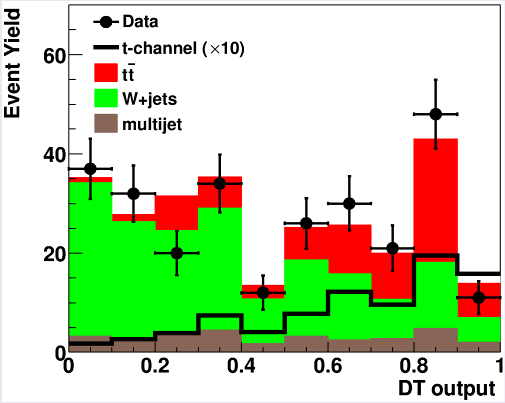
- Start with all events (first node)
- For each variable, find the splitting value with best separation between children (best cut).
- select best variable and cut and produce Failed and Passed branches
- Repeat recursively on each node
- Stop when improvement stops or when too few events left. Terminal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



Decision Trees

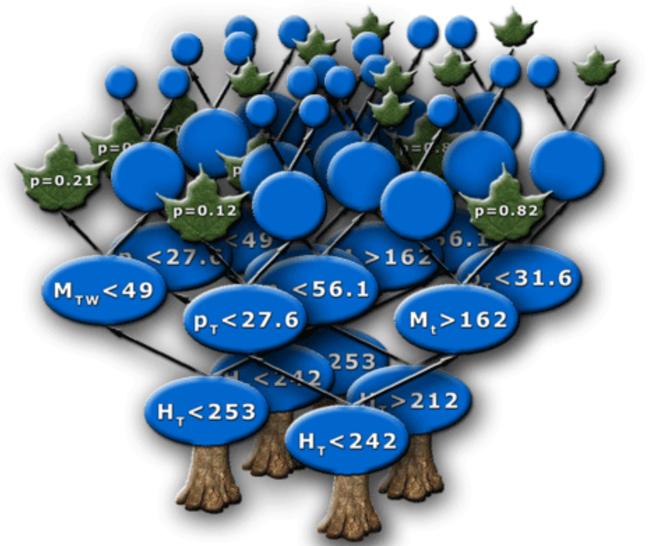
Measure and Apply

- Take trained tree and run on independent fake-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.

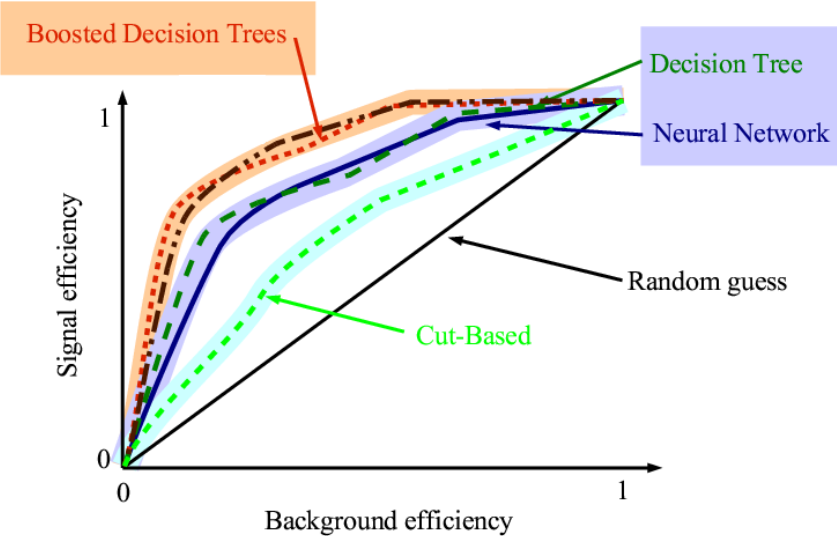


Boosting: Why Use a Tree When You Could Use a Forest?

Boosting: Why Use a Tree When You Could Use a Forest?



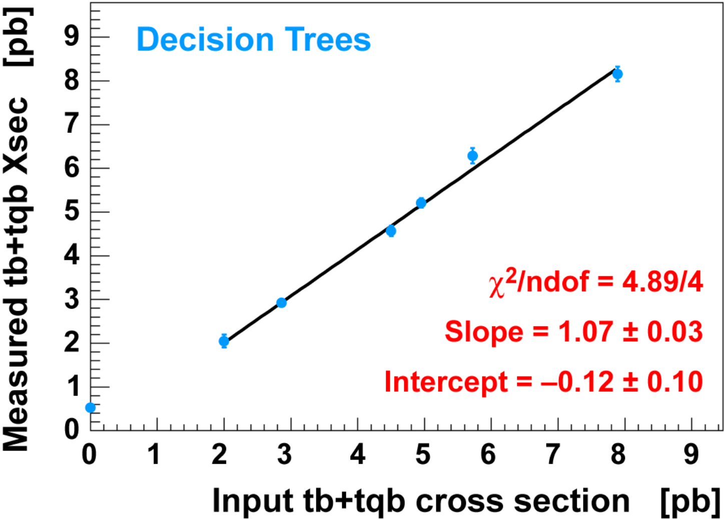
Decision Trees - Boosting



Decision Trees - Ensembles

- 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 ($s + t\ \sigma =$measured)
  5. A high luminosity ensemble
- Each analysis tests linearity of "response" to single top.

Decision Trees - Ensembles



Sensitivity Determination

Significance/Sensitivity Determination

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

Expected p-value

In a universe with no single top, how often do we measure at least the SM cross section? (what fraction of 0-signal pseudo-datasets do we measure at least 2.9pb)

Significance/Sensitivity Determination

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

Expected p-value

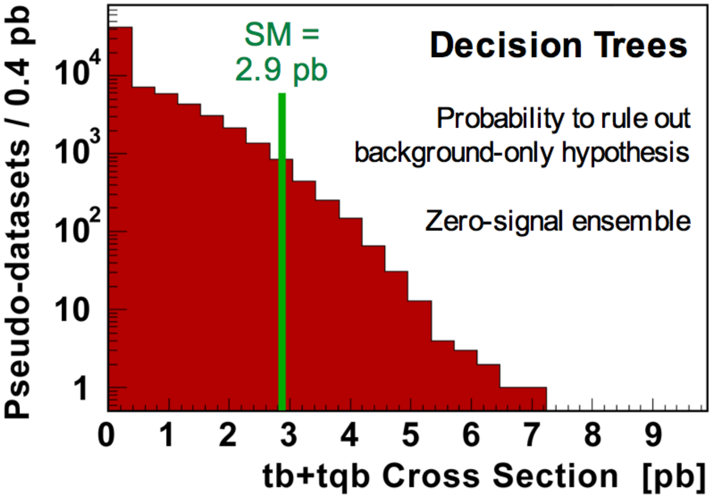
In a universe with no single top, how often do we measure at least the SM cross section? (what fraction of 0-signal pseudo-datasets do we measure at least 2.9pb)

Observed p-value

In a universe with no single top, how often do we measure at least the cross section we see in our data.

We also can use the SM ensemble to see how compatible our measured value is with the SM.

Significance/Sensitivity Determination

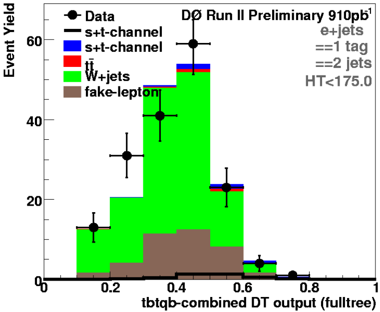


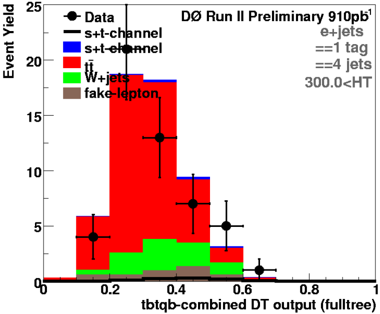
1.9%, 2.1 sigma

Looking at Data

Looking at Data: Cross-check samples

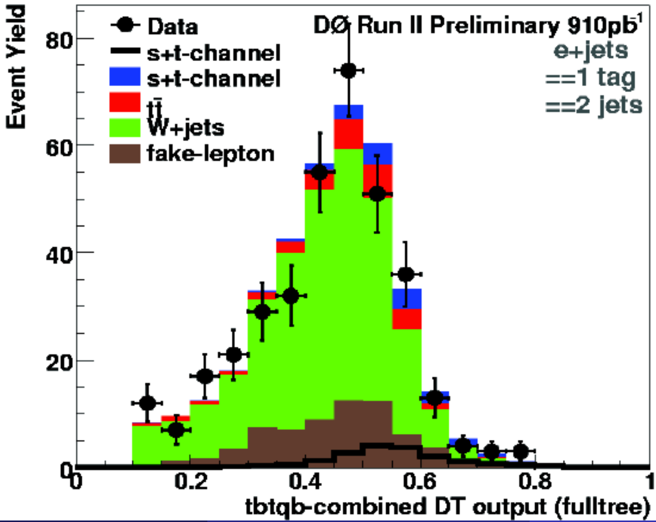
- "W+jets": =2jets, H_T(lepton,$\not\!\!E_T$,alljets) < 175 GeV
- "ttbar": =4jets, H_T(lepton,$\not\!\!E_T$,alljets) > 300 GeV
- Shown: $tb+tqb$ DT output for e+jets





- Good agreement of model with data

Of course, we have 36 different Decision Trees, let's look at electron, 2 jet, 1 tag:

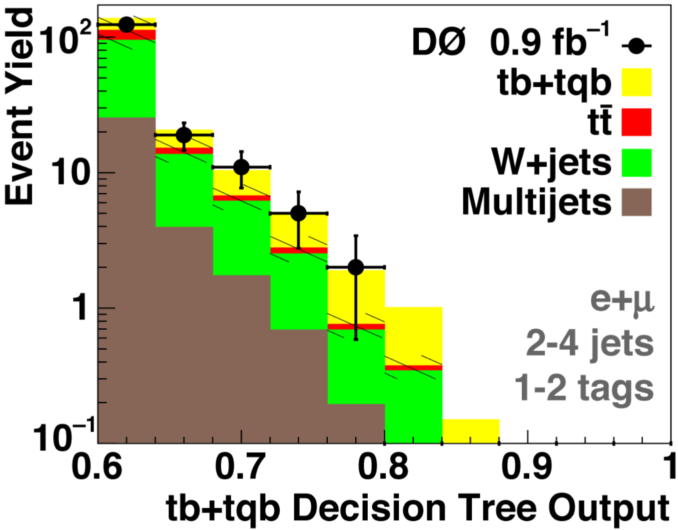


Looking at Data: All (zoom)

What if we just stack them up and zoom in?

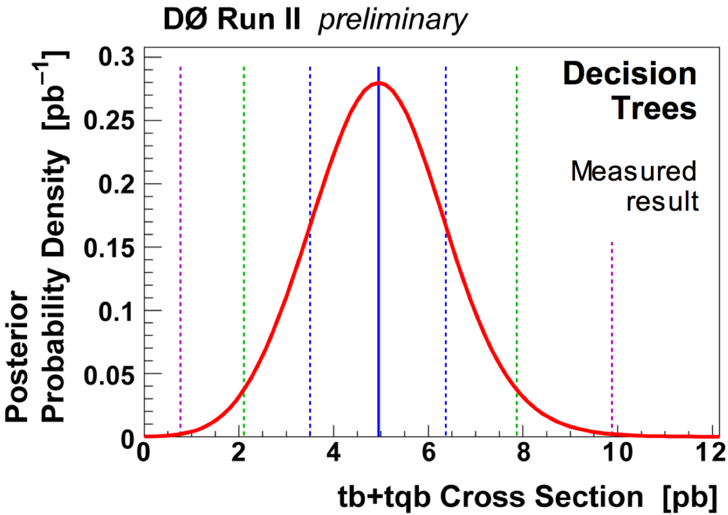
Looking at Data: All (zoom)

What if we just stack them up and zoom in?



Measuring the Cross-Section

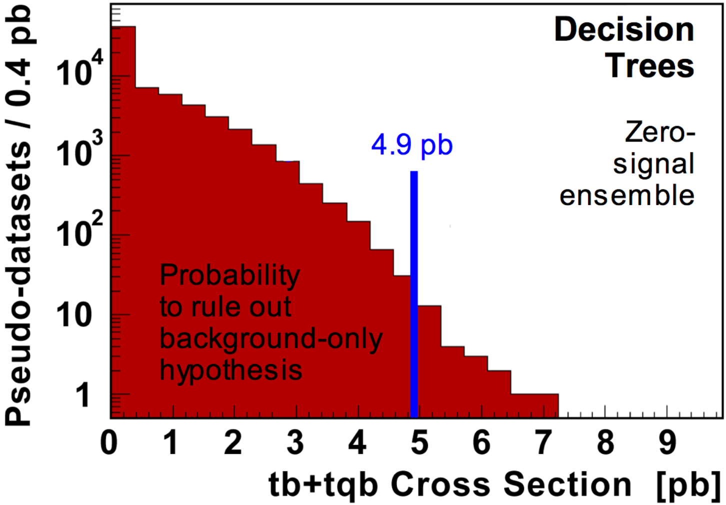
Measuring the Cross-Section



4.9 ± 1.4pb

Significance (p-value)

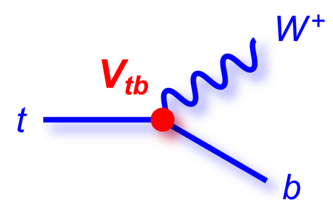
A 3.4σ excess!!



Measuring $|V_{tb}|$

Measuring $|V_{tb}|$

Direct access to V_{tb}



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

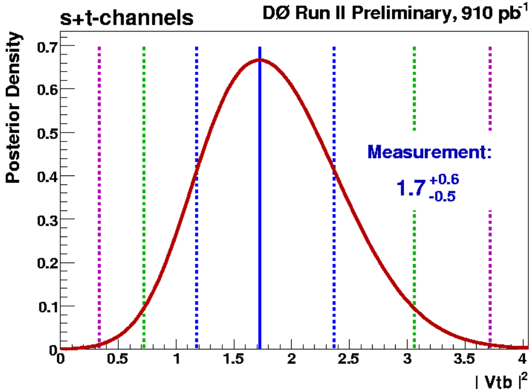
- Weak interaction eigenstates are not mass eigenstates
- In SM: top must decay to a W and d, s or b quark
  - $V_{td}^2 + V_{ts}^2 + V_{tb}^2 = 1$
  - constraints on V_{td} and V_{ts}: $V_{tb} = 0.9991$
- New physics that couples to the top quark:
  - $V_{td}^2 + V_{ts}^2 + V_{tb}^2 + V_{tx}^2 = 1$
  - no constraint on V_{tb}

Measuring $|V_{tb}|$

- Given that we now have a measurement of the single top cross section, we can make the first direct meassurement of $|V_{tb}|$.
- Use the same infrastructure as cross section measurement but make a measurement of $|V_{tb}|^2$ instead of cross section.
- Caveat: assume SM top quark decays.
- Additional theoretical errors are needed (see hep-ph/0408049)

	s	t
top mass	13%	8.5%
scale	5.4%	4.0%
PDF	4.3%	10.0%
α_s	1.4%	0.01%

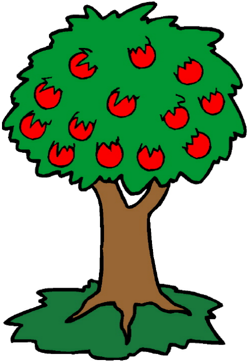
Limiting $|V_{tb}|$



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

$$0.68 < |V_{tb}| < 1.00$$

Conclusions



Brand new measurement!

- Top quark physics is a rich area to explore!
- We finally have evidence of EW production!
- Decision Trees helped us sort our apples and better understand the pie.
- $s + t$ cross section: 4.9 ± 1.4pb
- 3.4σ significance!

$|V_{tb}|$

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

$$0.68 < |V_{tb}| < 1$$

Phys.Rev.Lett.98:181802,2007

Shameless SFU Plug

BACKUP SLIDES

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

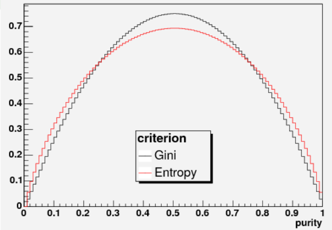
- Decrease of impurity for split s of node t into children t_L and t_R (goodness of split):
  $\Delta i(s,t) =$
  $i(t) - p_L \cdot i(t_L) - p_R \cdot i(t_R)$
- Aim: find split s^* such that:

$$\Delta i(s^*, t) = \max_{s \in \{\text{splits}\}} \Delta i(s,t)$$

Examples

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

$entropy = -\sum_{i=s,b} p_i \log p_i$



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 α_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

Ref: Freund and Schapire, "Experiments with a new boosting algorithm", in *Machine Learning: Proceedings of the Thirteenth International Conference*, pp 148-156 (1996)

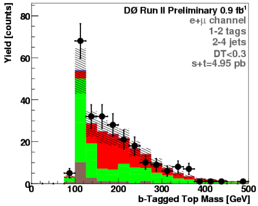
Decision Trees - In this analysis

Analysis strategy

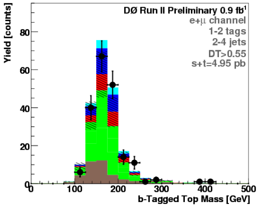
- Train 36 separate trees:
  - 3 signals (s,t,$s+t$)
  - 2 leptons (e,μ)
  - 3 jet multiplicities (2,3,4 jets)
  - 2 b-tag multiplicities (1,2 tags)
- For each signal train against the sum of backgrounds
- results shown are 12 $s+t$ trees

Looking at Data - Event Characteristics $M(W, b)$

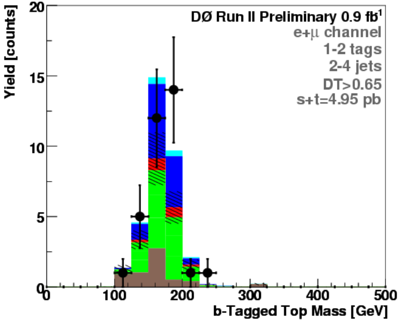
$DT < 0.3$



$DT > 0.55$

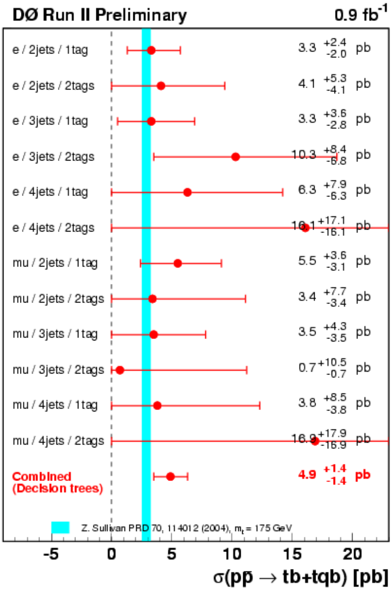


$DT > 0.65$

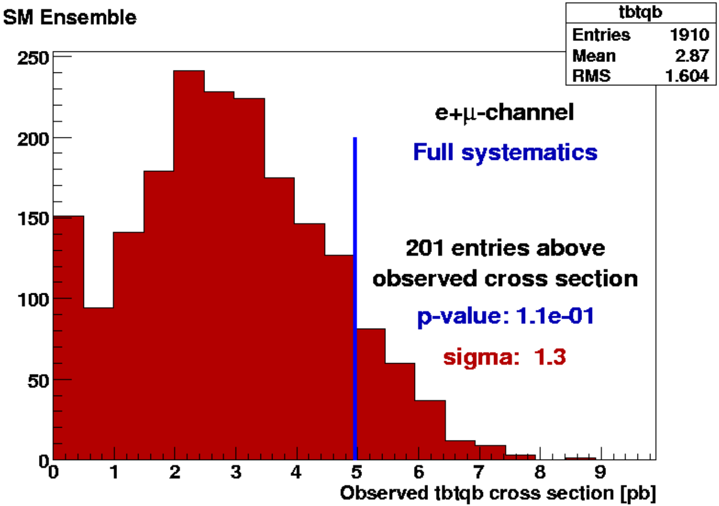


- Excess in high DT output region.

Measuring the Cross-Section - Summary



Consistent with SM?



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

Source of Uncertainty	Size
Top pairs normalization	18%
W+jets & multijets normalization	18–28%
Integrated luminosity	6%
Trigger modeling	3–6%
Lepton ID corrections	2–7%
Jet modeling	2–7%
Other small components	Few %
Jet energy scale	1–20%
Tag rate functions	2–16%

Measuring the Cross Section

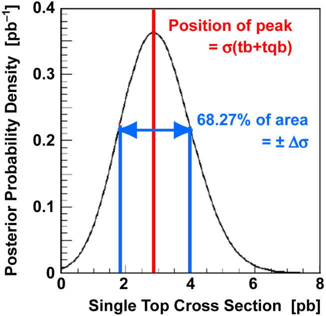
Probability to observe data distribution D, expecting y:

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

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

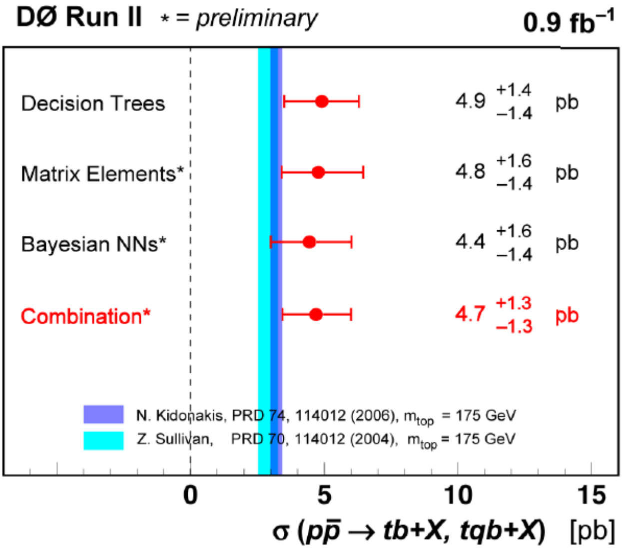
The cross section is obtained

$$Post(\sigma|D) \equiv P(\sigma|D) \propto \int_a \int_b P(D|\sigma, a, b) Prior(\sigma \ldots (a, b)$$



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

Other Results



Measuring $|\mathbf{V_{tb}}|$

W-t-b Vertex

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

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

- SM: $f_1^L = 1$, $f_1^R = f_2^L = f_2^R = 0$
- Effectively measuring strength of $V{-}A$ coupling $|V_{tb} f_1^L|$, can be > 1

Decision Trees - 49 variables

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

ΔR(jet1,jet2)
cos(best1,lepton)$_{\mathrm{besttop}}$
cos(best1,notbest1)$_{\mathrm{besttop}}$
cos(tag1,alljets)$_{\mathrm{alljets}}$
cos(tag1,lepton)$_{\mathrm{btaggedtop}}$
cos(jet1,alljets)$_{\mathrm{alljets}}$
cos(jet1,lepton)$_{\mathrm{btaggedtop}}$
cos(jet2,alljets)$_{\mathrm{alljets}}$
cos(jet2,lepton)$_{\mathrm{btaggedtop}}$
cos(lepton,Q(lepton)$\times z$)$_{\mathrm{besttop}}$
cos(lepton,besttopframe)$_{\mathrm{besttopCMframe}}$
cos(lepton,btaggedtopframe)$_{\mathrm{btaggedtopCMframe}}$
cos(notbest,alljets)$_{\mathrm{alljets}}$
cos(notbest,lepton)$_{\mathrm{besttop}}$
cos(untag1,alljets)$_{\mathrm{alljets}}$
cos(untag1,lepton)$_{\mathrm{btaggedtop}}$

Event Kinematics

Aplanarity(alljets,W)
M(W,best1) ("best" top mass)
M(W,tag1) ("b-tagged" top mass)
H_T(alljets)
H_T(alljets−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, lose some sensitivity
- Same list used for all channels