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

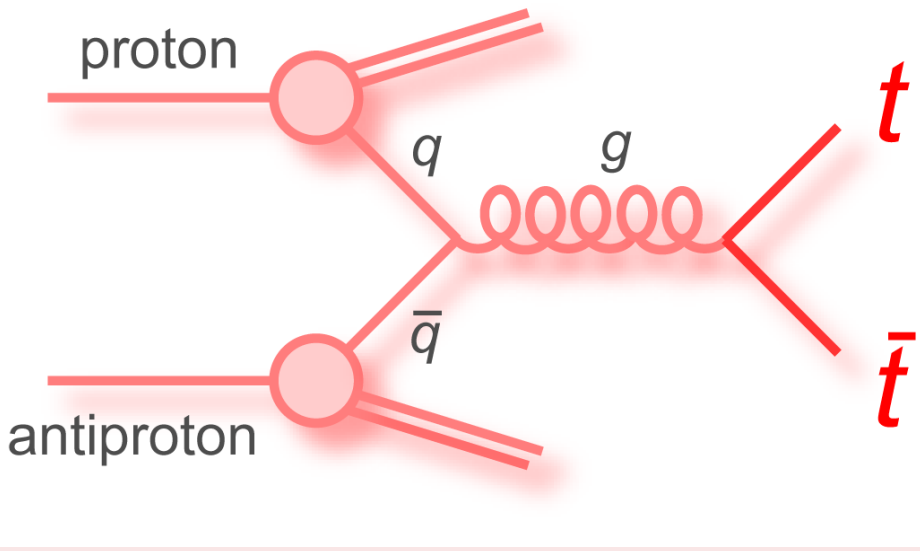
Dag Gillberg
For the DØ Collaboration

SIMON FRASER UNIVERSITY

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 ± 0.6 pb (incl. uncert. on the mass $\longrightarrow \pm 1.2$ pb)

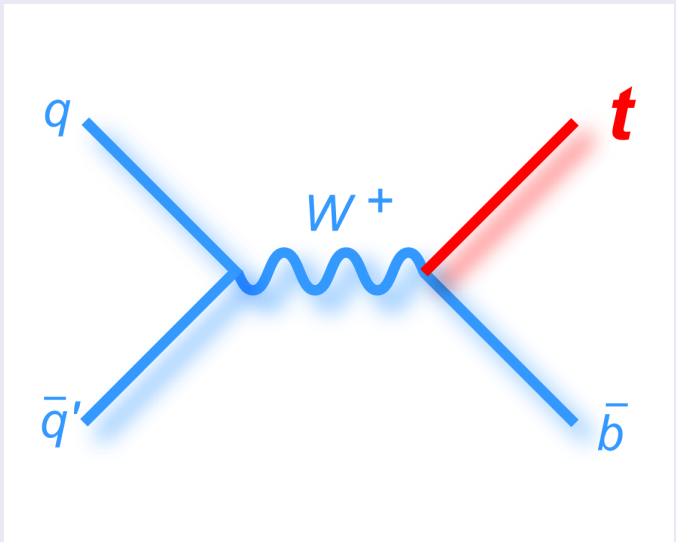
$t\bar{t}$ production channels



Electroweak single top production

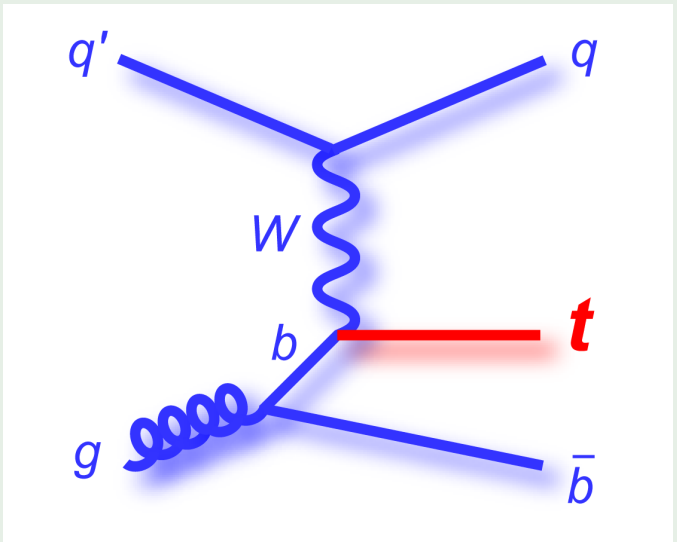
- Never observed before
- Dominant production channels at the Tevatron:

s-channel (tb)



$\sigma_{\mathbf{NLO}} = \mathbf{0.88} \pm 0.11$ pb (*)

t-channel (tqb)



$\sigma_{\mathbf{NLO}} = \mathbf{1.98} \pm 0.25$ pb(*)

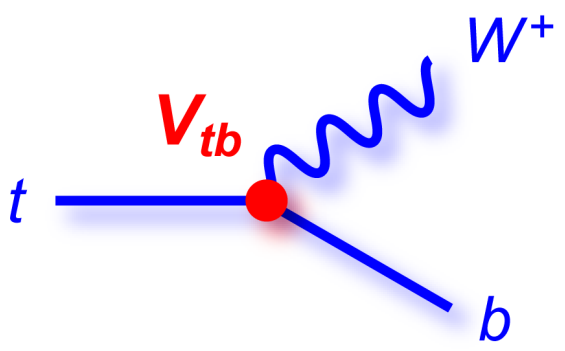
(*) $m_t = 175$ GeV, Phys.Rev. D70 (2004) 114012

- Total SM cross section: $\sigma_{\mathbf{s+t}} = \mathbf{2.9\ pb}$

Motivation

1. Predicted but not observed before
2. **Study the Wtb coupling**

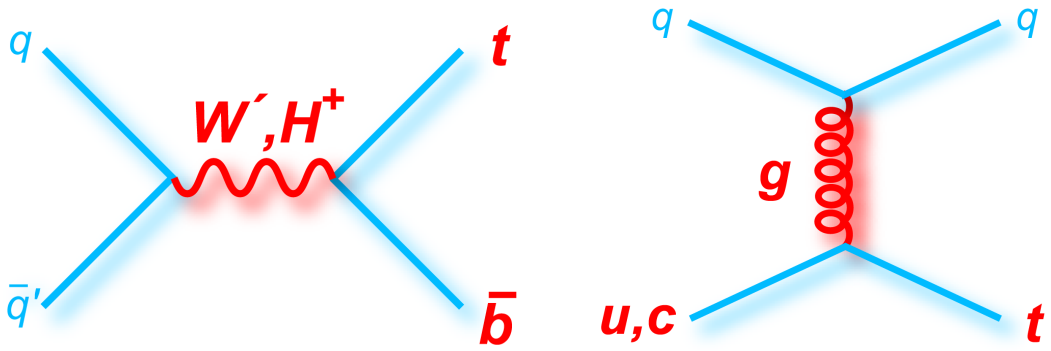
$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}$$



   - Direct measurement of $|V_{tb}|$ (more later)
   - Unitarity test of CKM matrix
   - Anomalous Wtb couplings
3. Measurement of top quark properties:
   polarization, decay width, lifetime, mass...

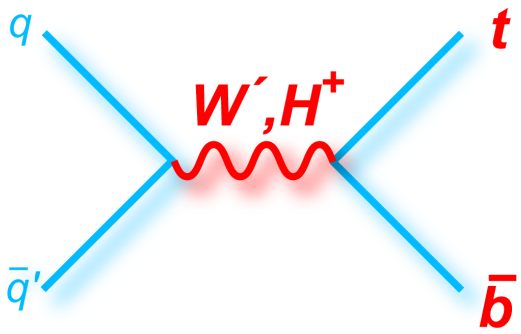
Motivation

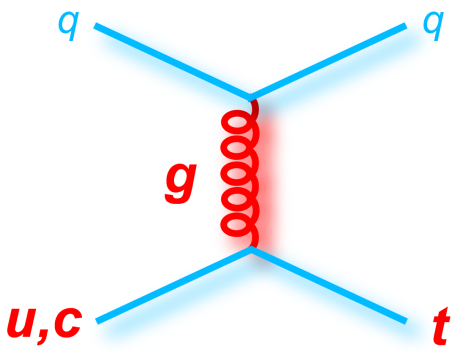
4. **New physics**, example:
   - s-channel: resonances like heavy W', charged Higgs
   - t-channel: Flavour-changing neutral currents
   - 4-th quark generation?



Motivation

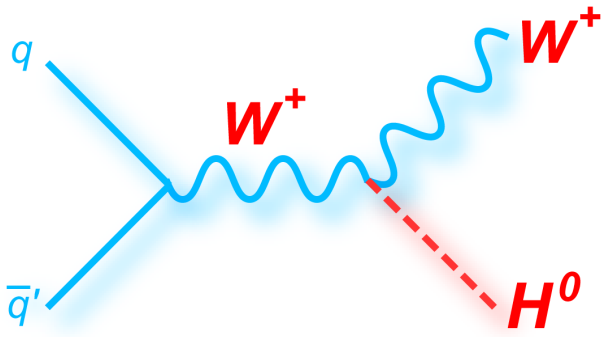
4. **New physics**, example:
   - s-channel: resonances like heavy W', charged Higgs
   - t-channel: Flavour-changing neutral currents
   - 4-th quark generation?





5. Very similar to SM Higgs production
   - Same final state and bkg as

$$WH \rightarrow l\nu b\bar{b}$$

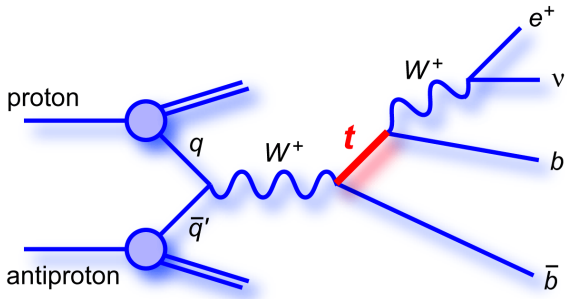


6. Test of advanced separation techniques

Signal and backgrounds samples

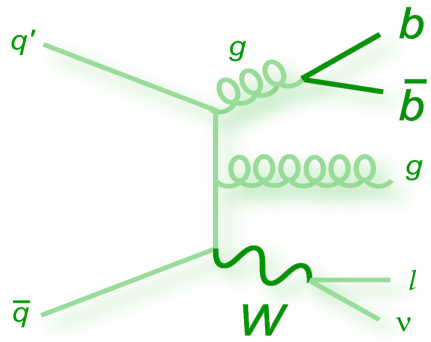
Single top signal, MC

- COMPHEP-SINGLETOP



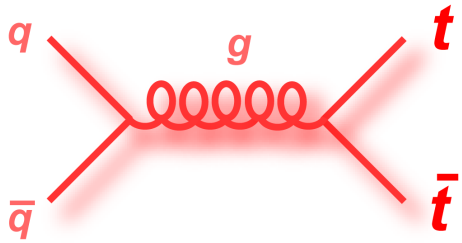
W+jets, MC

- Most difficult background
- Shapes from ALPGEN
- Normalization and heavy flavour fraction from data



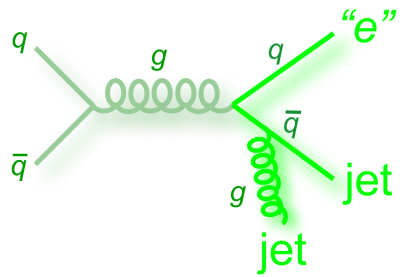
$t\bar{t}$, MC

- ALPGEN
- Normalized to $\sigma_{NNLO} = 6.8$ pb



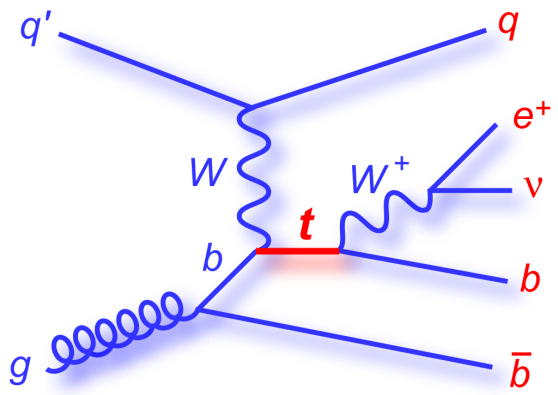
Multijet events, from data

- misidentified lepton



Analysis Strategy

Strategy: Maximize signal acceptance – extract signal using multivariate techniques

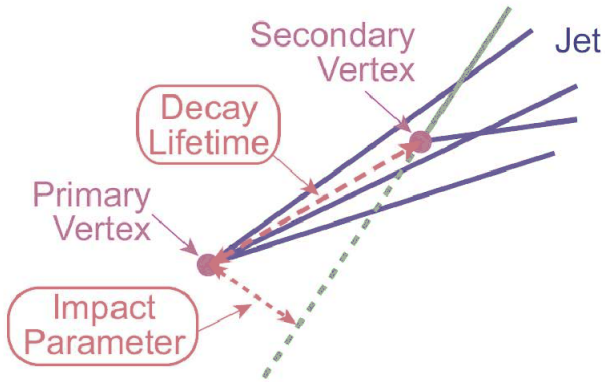


Signature

- isolated lepton
- $\not\!\!E_T$ (from ν)
- jets
- at least 1 b-jet

Event selection

- One lepton:
  - electron: $p_T > 15$ GeV, $|\eta_{det}| < 1.1$
  - muon: $p_T > 18$ GeV, $|\eta_{det}| < 2$
- $\not\!\!E_T > 15$ 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
  - **1-2 jets b-tagged**



b-tagging

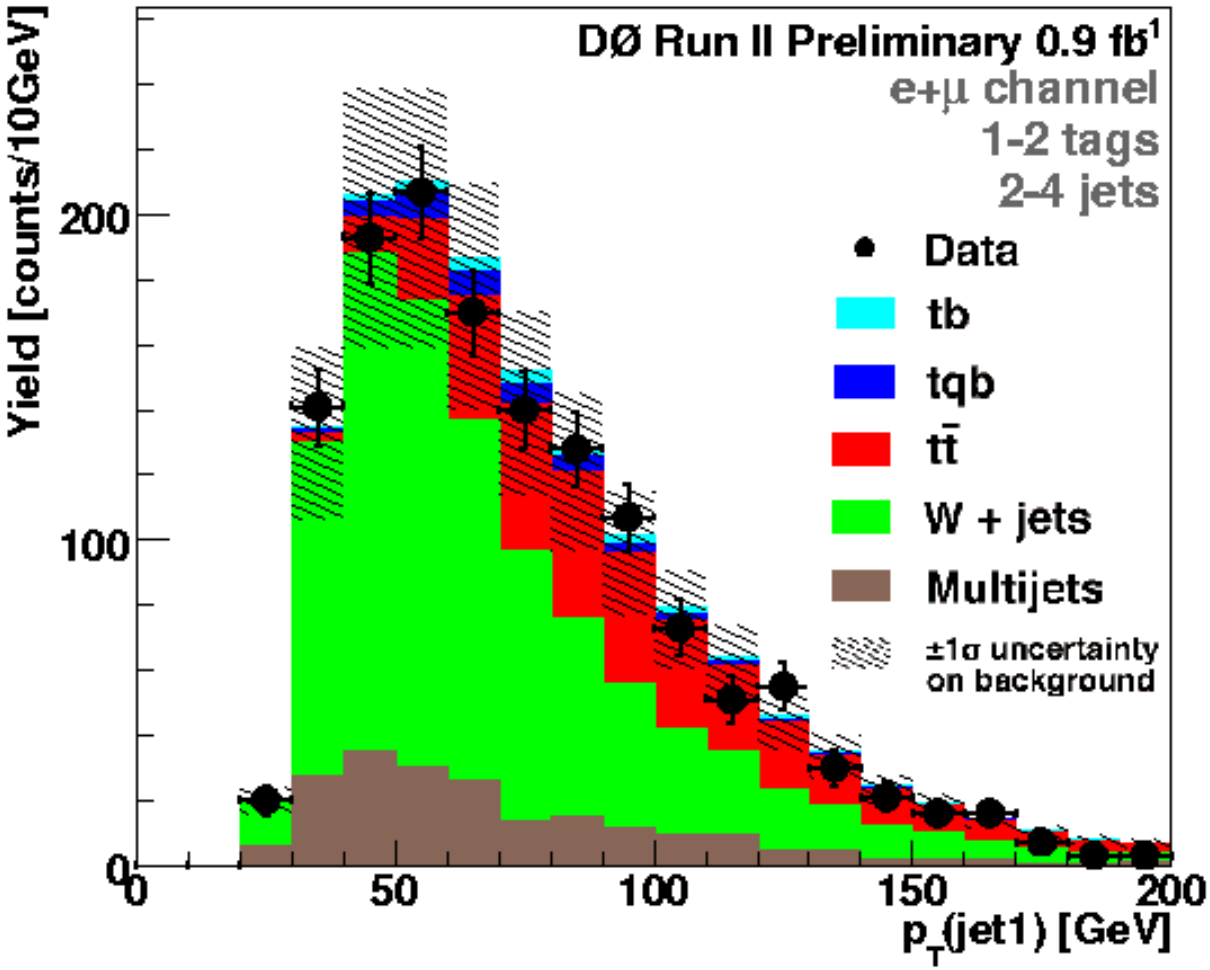
Data-background model agreement

Expected N_{events}

$s + t$ **signal**	**62**
$t\bar{t}$	348
Wjj	174
Wbb & Wcc	675
multijet	201

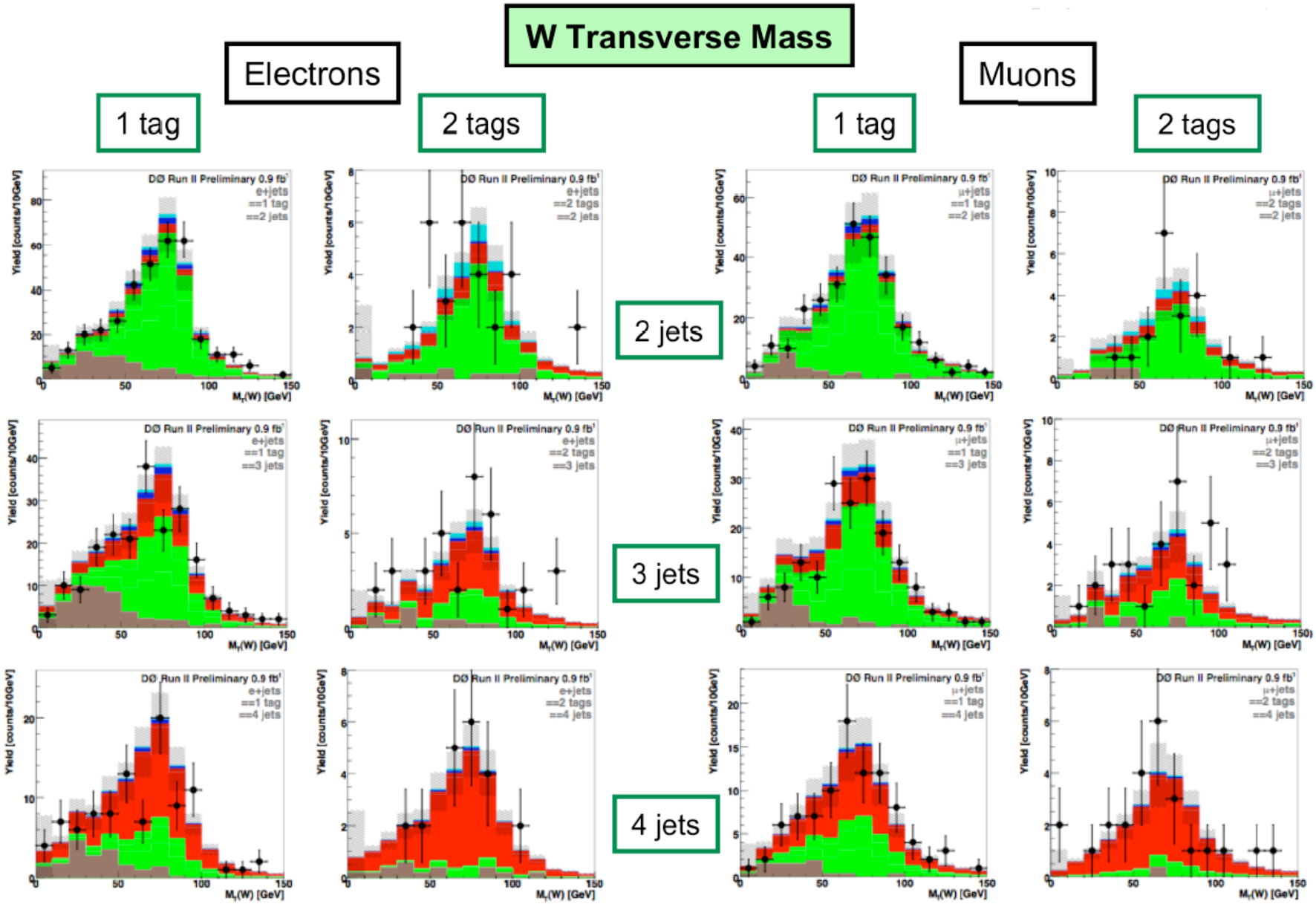
data/bkg N_{events}

signal	**62**
background	**1399**
data	**1398**



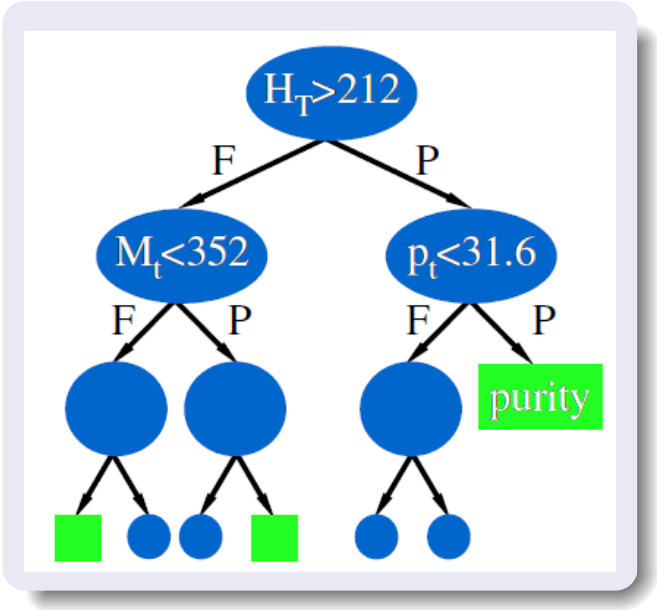
We have hundreds of these kind of plots for other kinematic variables

The 12 (lep, N_{tag}, N_{jet})-bins



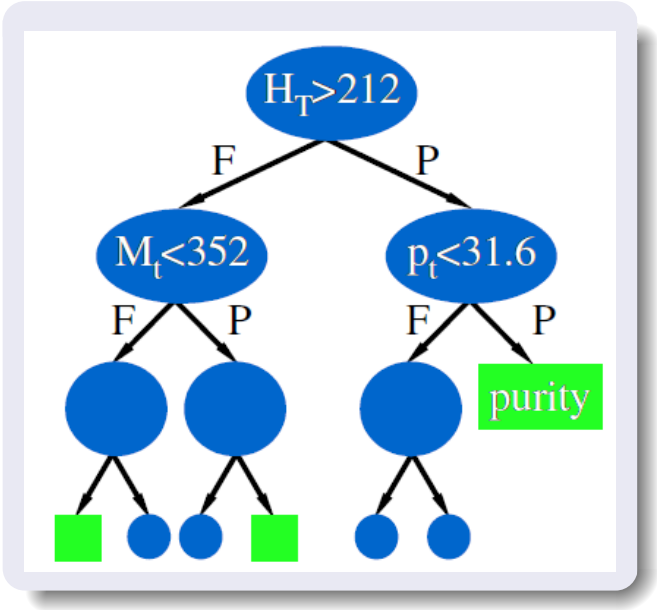
The Decision Tree Analysis

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



The Decision Tree Analysis

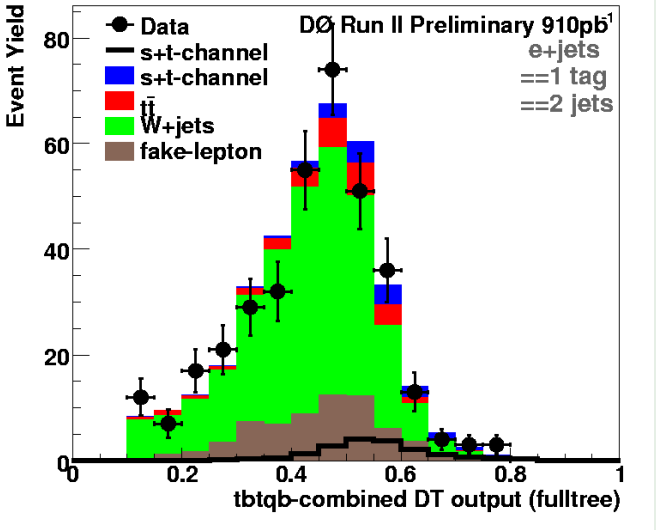
- Single top signal sample trained against our background samples
- Idea: recover events that fail criteria in cut-based analysis
- Separate training per (lep, N_{tag}, N_{jet})-bin
- **Boosting** – give mis-classified events higher weight and re-train! (20 times)



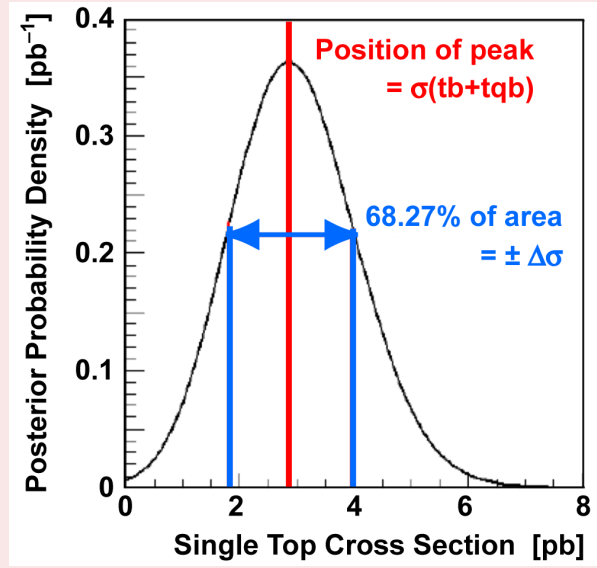
Discriminating variables

- 49 variables used, categories:
  object kinematics, global event kinematics, angular correlations
- Most discriminating power:
  m(alljets), $m(W, b\text{-jet1})$, cos(b-jet1,lep), $q_{lep} \times \eta_{\text{un-tagged jet1}}$

The Decision Tree Analysis



Decision Tree Output



Bayesian posterior

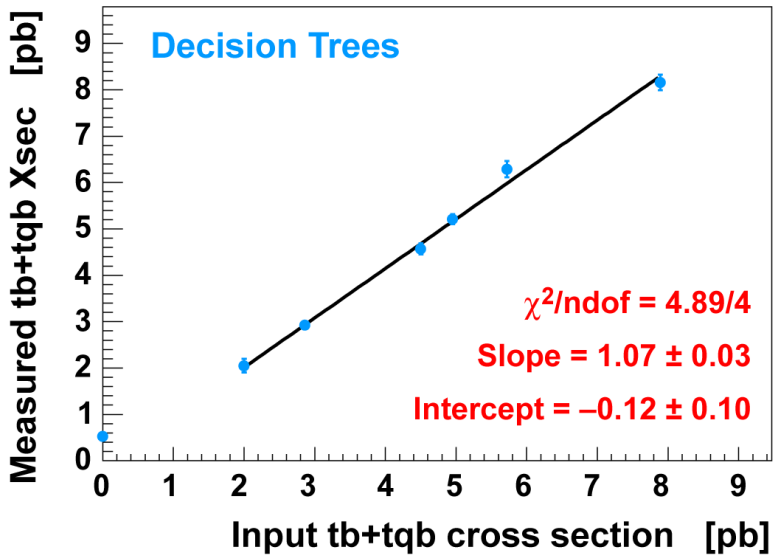
- Bayesian posterior probability density:

  $P(\sigma|D) \int_{a,b} P(D|\sigma, a, b) Prior(\sigma) Prior(a, b)$
- Shape and normalization systematics treated as nuisance parameters
- Correlations between uncertainties properly accounted for

Ensemble Testing

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

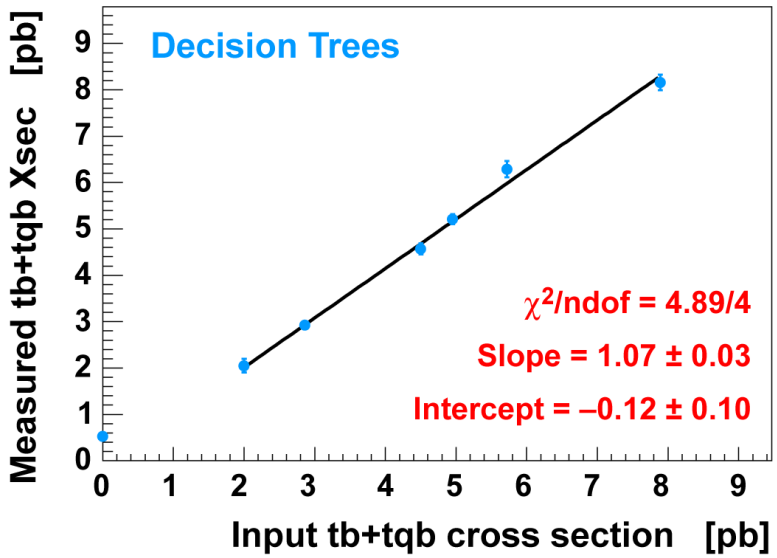
- Subset of our total pool of background events
- Individual statistical and systematical fluctuations
- Wonderful tool – like running DØ 1000s of times!
- Generated several ensembles with different single top content $\longrightarrow$



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 $\longrightarrow$



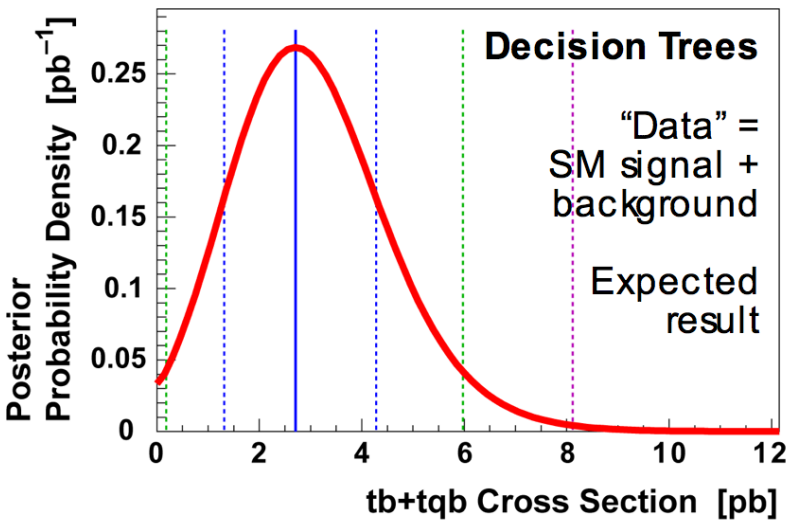
Significance

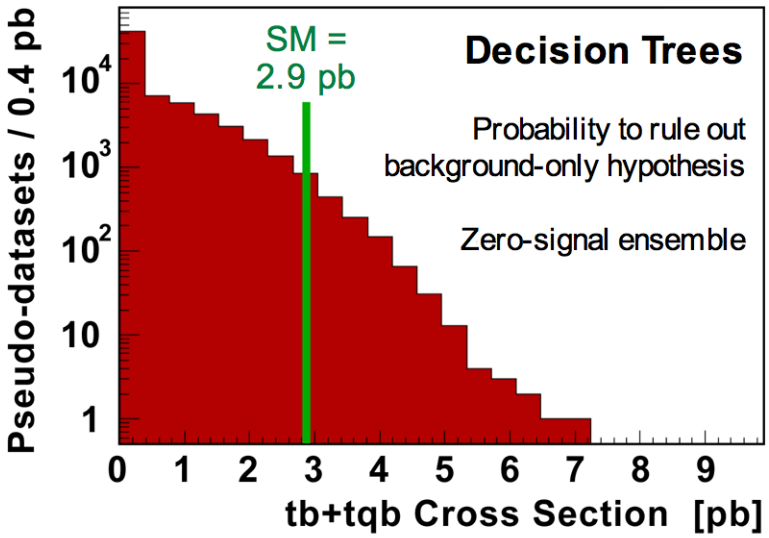
- Used 70,000 pseudo-datasets with **no single top content**
- We measure the fraction of zero-signal datasets in which we derive at least the SM cross section (**expected significance**), or at least the observed cross section (**observed significance**)

Expected results

Expected Decision Tree Results

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





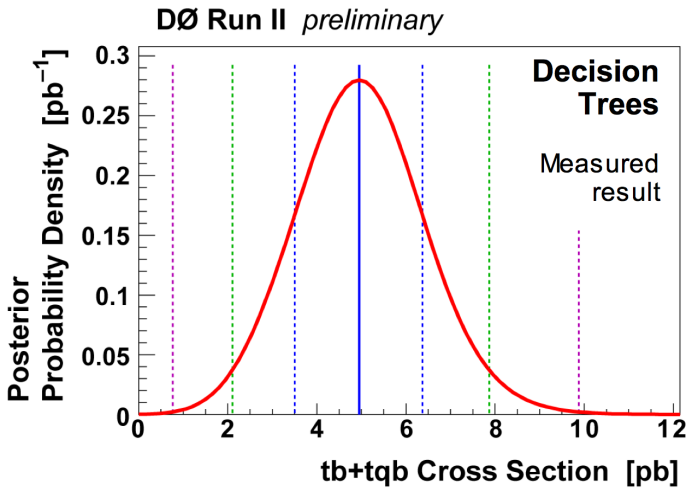
Boosted decision tree observed results

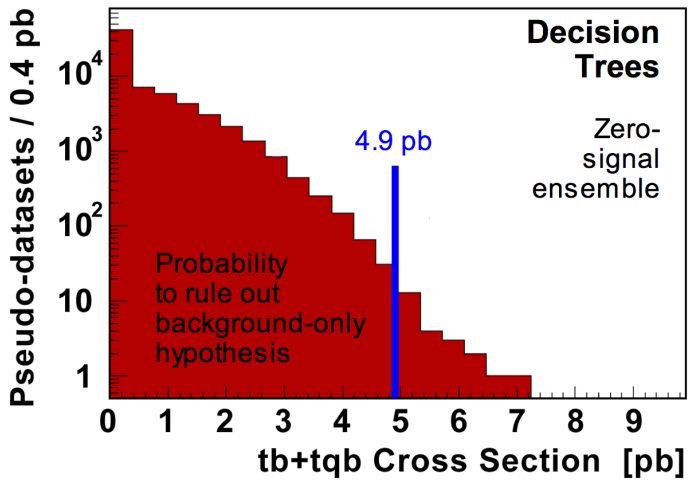
$\sigma_{s+t} = 4.9 \pm 1.4$ pb

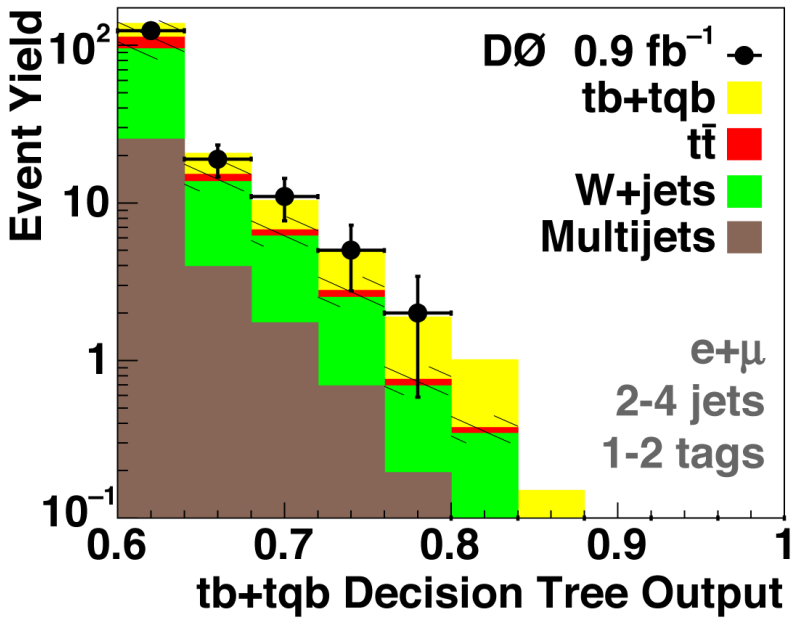
p-value $= 0.035\%$ (3.4σ)

SM compatibility: 11% (1.3σ)

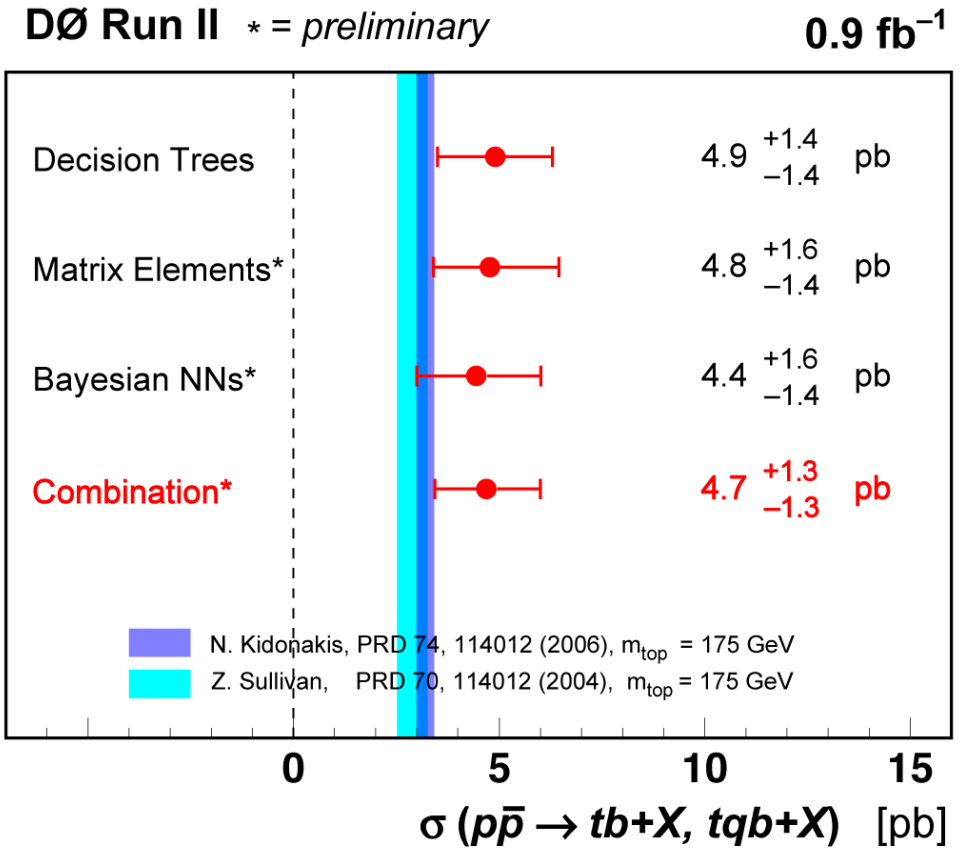
Evidence!







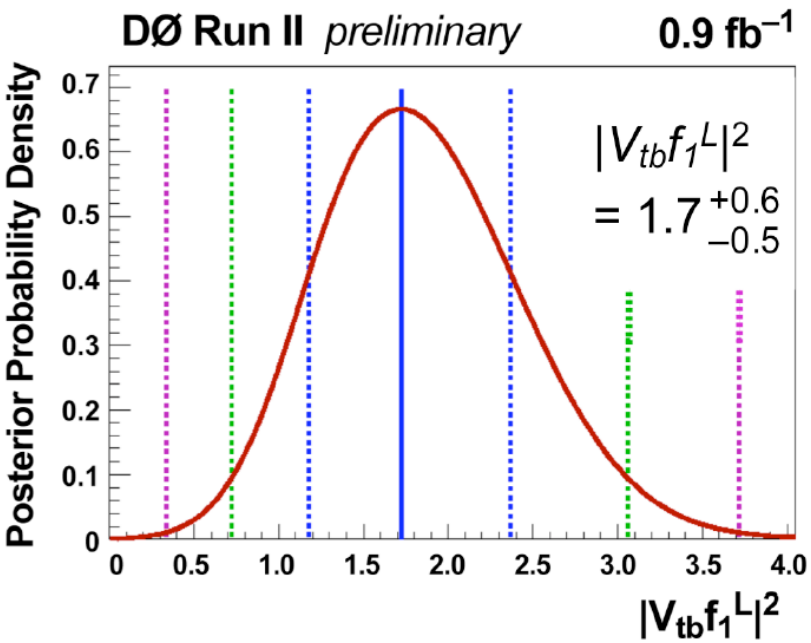
Observed σ_{s+t} for all three analyses



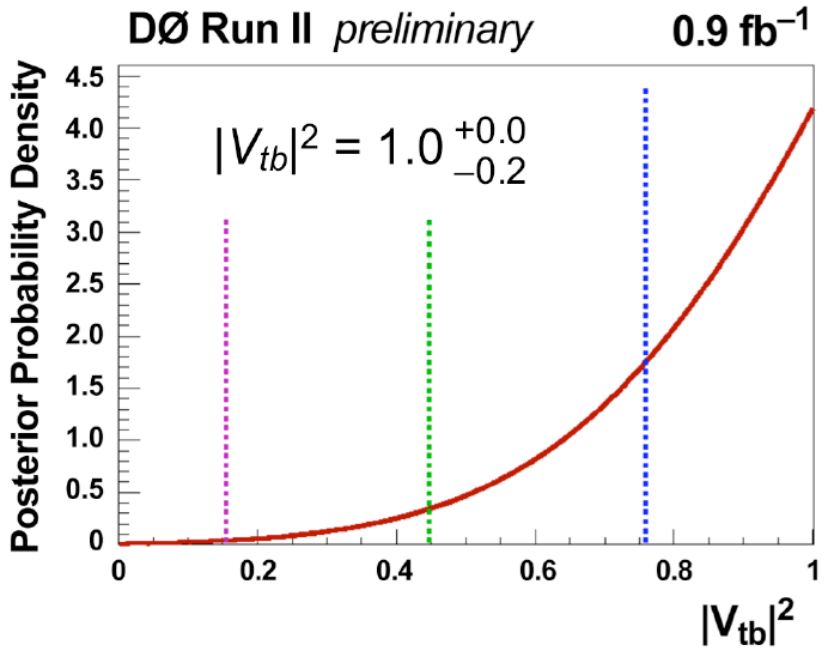
- The three analyses give consistent results!
- **Combined result using the BLUE method:**
  4.7 ± 1.3 pb with **3.6 σ** significance!
- Same data used (0.9 fb^{-1}) – therefore correlated

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

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



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



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

- No assumption about number of quark families or CKM unitarity

Conclusion

Boosted Decision Trees result published: PRL 98, 181802

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 σ significance

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

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

assuming $f_1^L = 1$ (SM): $0.68 < |V_{tb}| \leq 1$ @ 95% CL

Latest result:

Combined result, DT + New ME + New BNN

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

3.6 σ significance

BACKUP SLIDES

Systematics Uncertainties

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%

Expected signal yields and S:B ratios

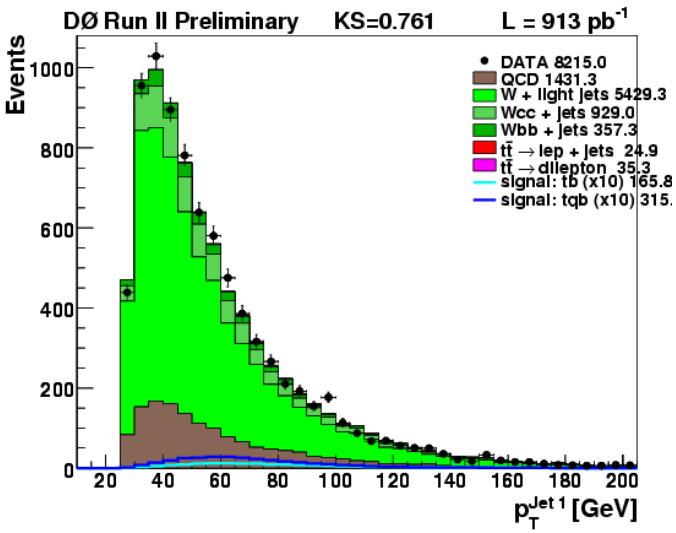
Percentage of single top *tb+tqb* selected events and S:B ratio (white squares = no plans to analyze)

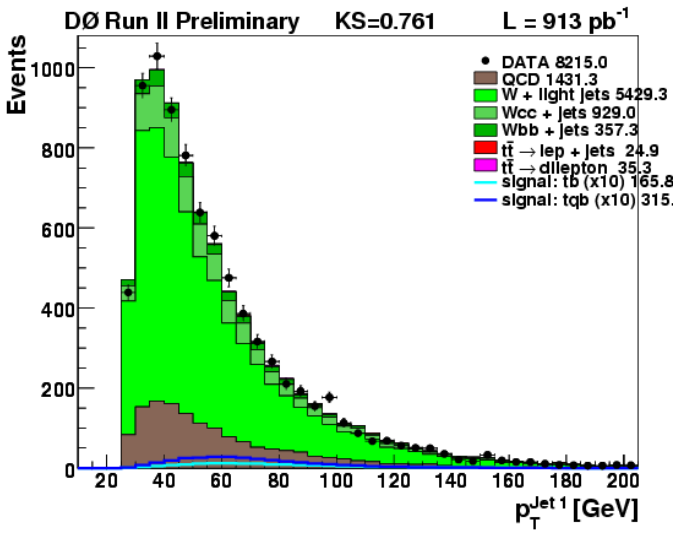
Electron + Muon	1 jet	2 jets	3 jets	4 jets	≥ 5 jets
0 tags	10% 1 : 3,200	25% 1 : 390	12% 1 : 300	3% 1 : 270	1% 1 : 230
1 tag	6% 1 : 100	21% 1 : 20	11% 1 : 25	3% 1 : 40	1% 1 : 53
2 tags		3% 1 : 11	2% 1 : 15	1% 1 : 38	0% 1 : 43

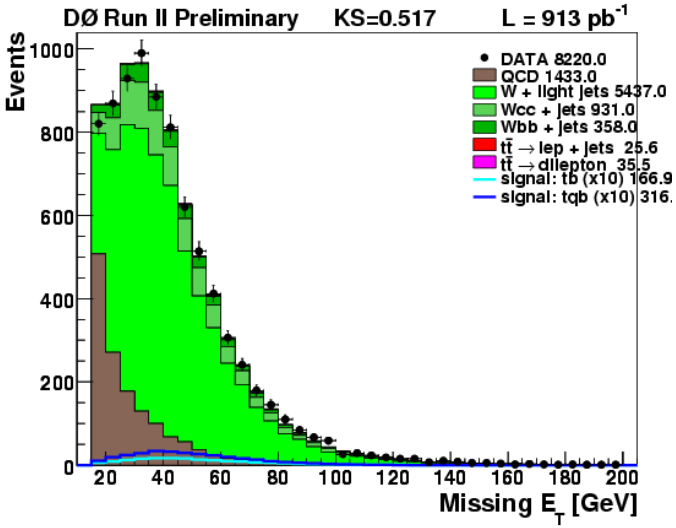
Yields before b-tagging

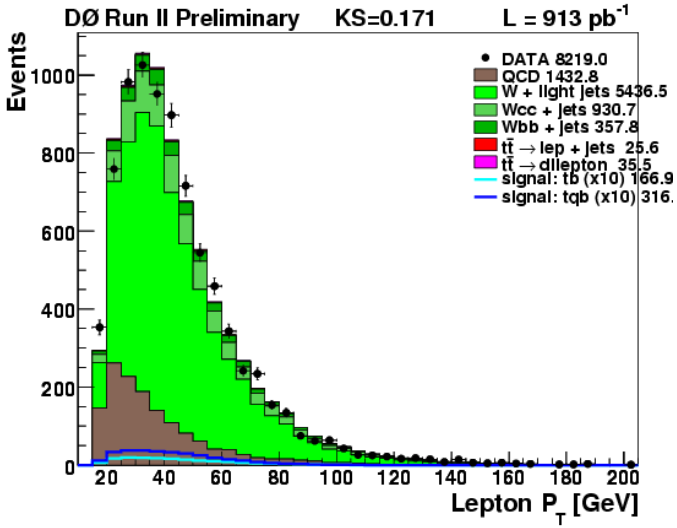
Source	Event Yields in 0.9 fb^{-1} Data Electron+muon combined, before b-tagging 2 jets	3 jets	4 jets
tb	25	12	3
tqb	47	25	8
$t\bar{t} \rightarrow ll$	62	50	18
$t\bar{t} \rightarrow l$+jets	40	175	227
$W+b\bar{b}$	670	310	89
$W+c\bar{c}$	1,959	912	224
$W+jj$	10,160	3,138	728
Multijets	1,762	1,083	314
Total background	14,654	5,667	1,601
Data	14,652	5,665	1,601

Data-background agreement before b-tagging







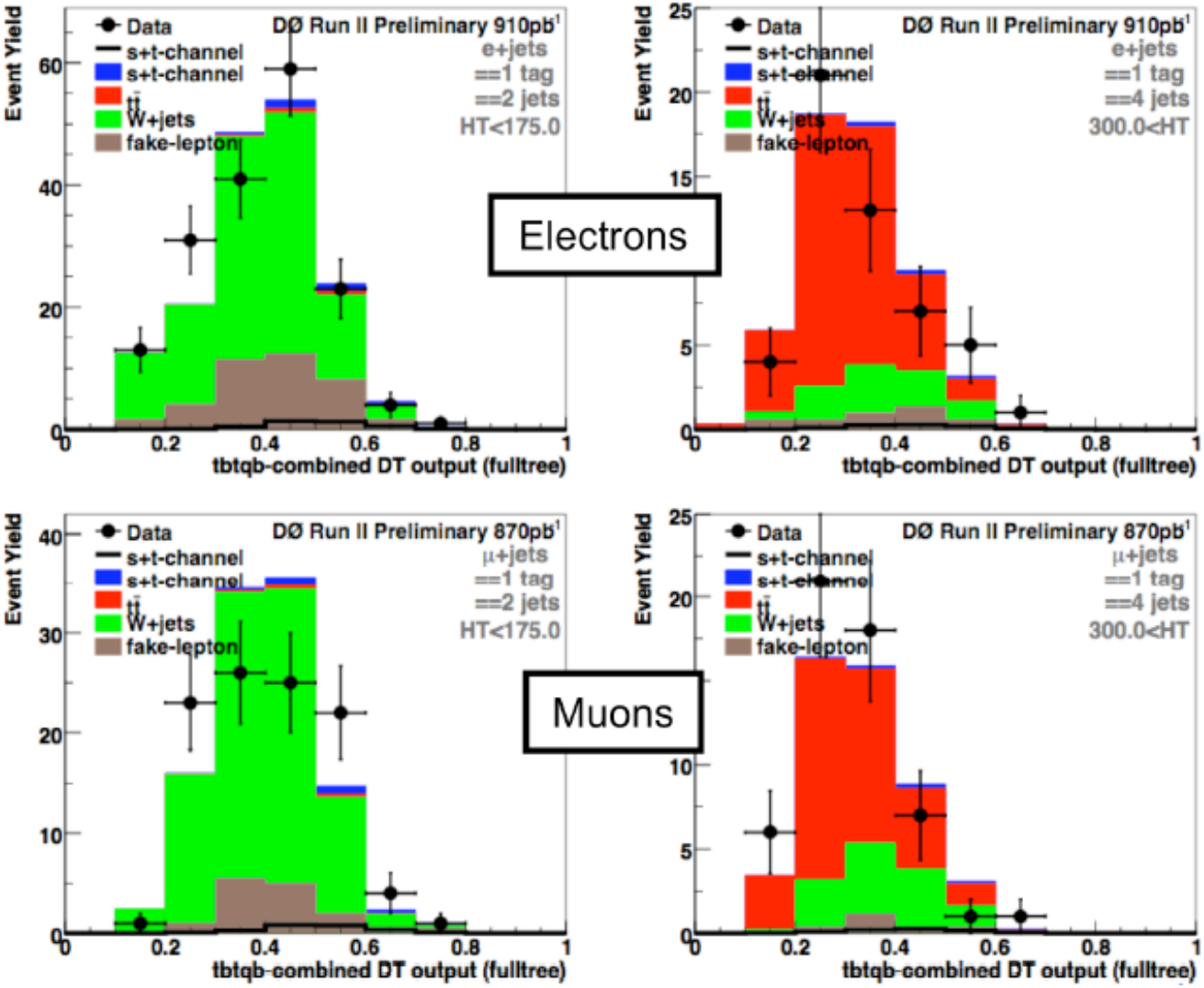


- **Background model show good agreement with data!**

Cross-check samples

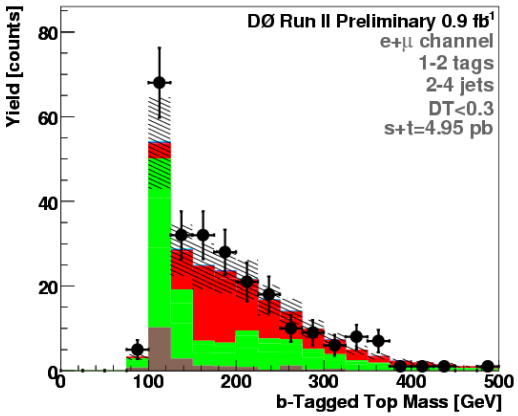
"W+jets"

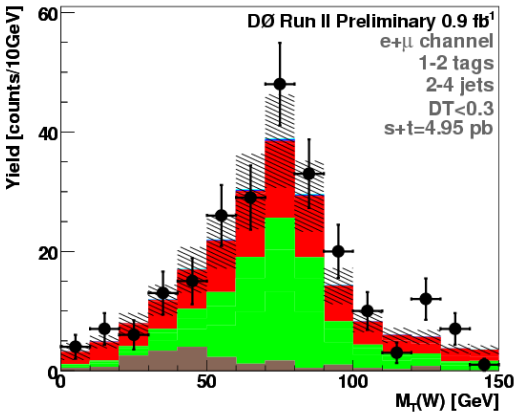
"$t\bar{t}$"



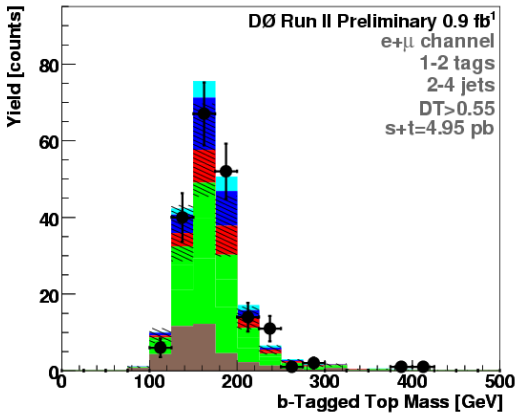
Boosted decision tree event characteristics

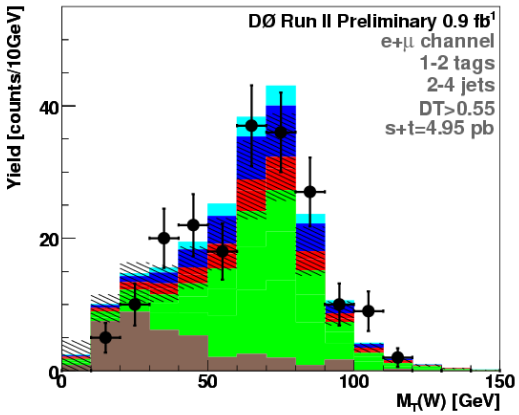
$DT < 0.3$



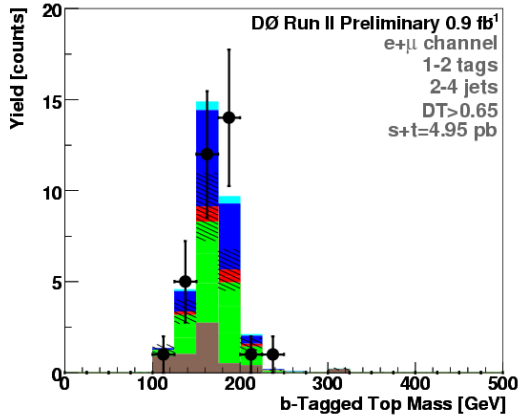


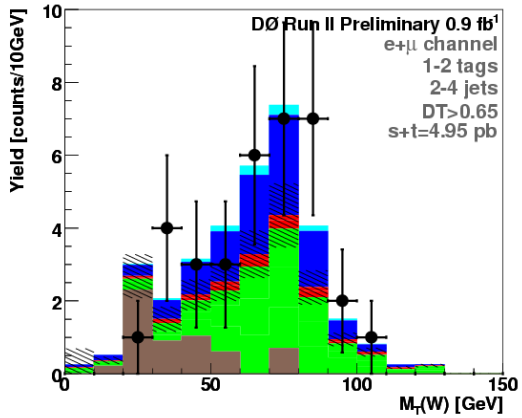
$DT > 0.55$





$DT > 0.65$





Decision Trees - 49 input 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$_{\mathrm{besttop}}$,besttop$_{\mathrm{CMframe}}$)
- cos(lepton$_{\mathrm{btaggedtop}}$,btaggedtop$_{\mathrm{CMframe}}$)
- 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, lost some sensitivity
- Same list used for all channels

Sensitivity determination

Using 70,000 pseudo-datasets without any single top content:

Expected p-value

Fraction of 0-signal pseudo-datasets in which we measure at least 2.9 pb (SM σ)

Observed p-value

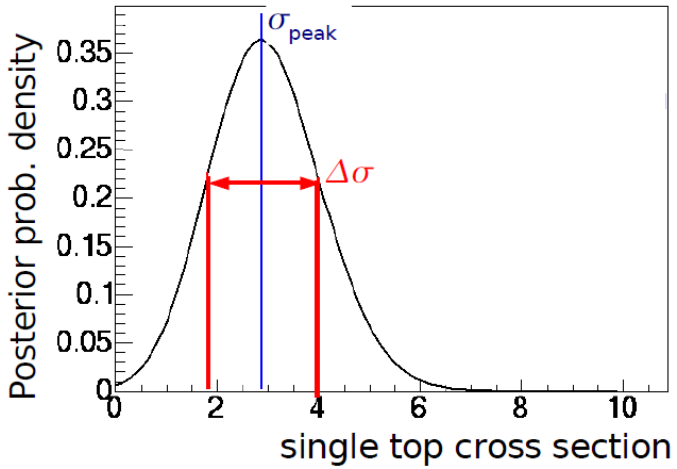
Fraction of 0-signal pseudo-datasets in which we measure at least the observed cross section

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

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

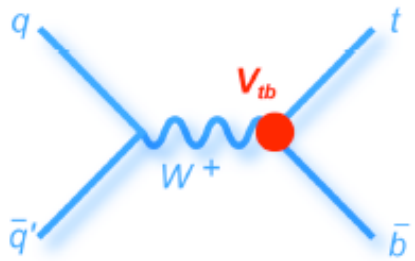
- Bayesian posterior probability density
- Shape and normalization systematics treated as nuisance parameters
- Correlations between uncertainties properly accounted for
- Flat prior in signal cross section

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

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

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

- **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		
	tb	*tqb*
Top mass	13 %	8.5 %
Scale	5.4 %	4.0 %
PDF	4.3 %	10 %
α_s	1.4 %	0.01 %