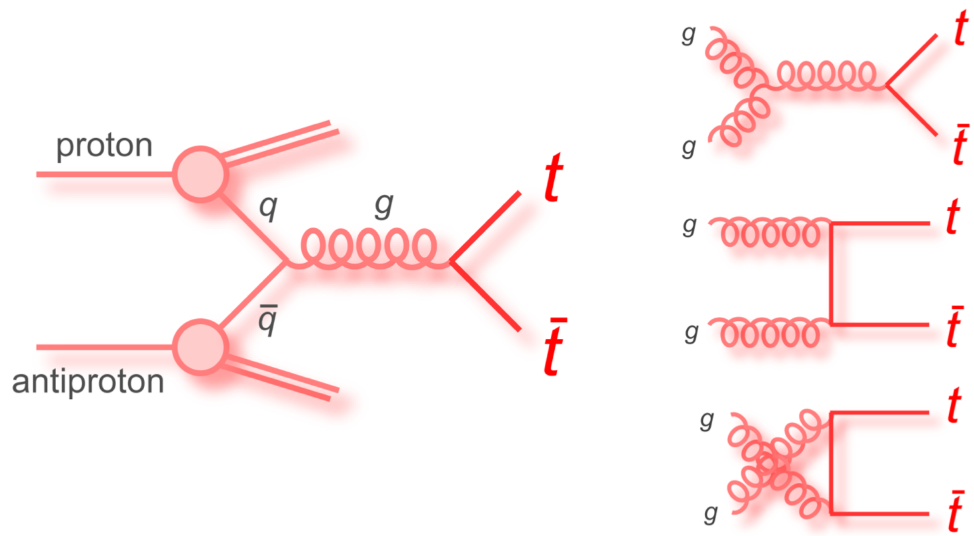
Evidence for
Single Top Quark Production at DØ

Ann Heinson
University of California, Riverside
for the DØ Collaboration

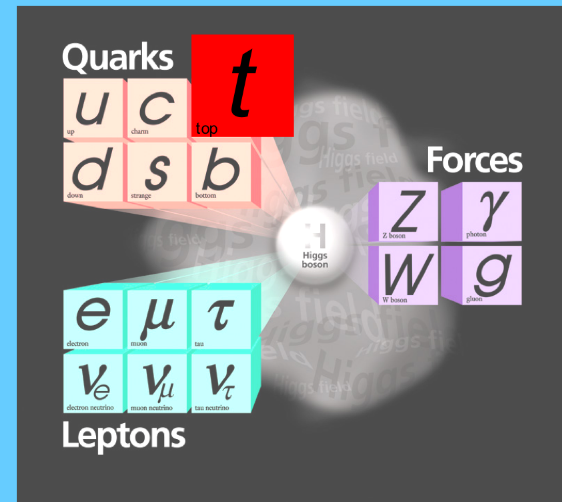
CERN Particle Physics Seminar
Tuesday January 30, 2007

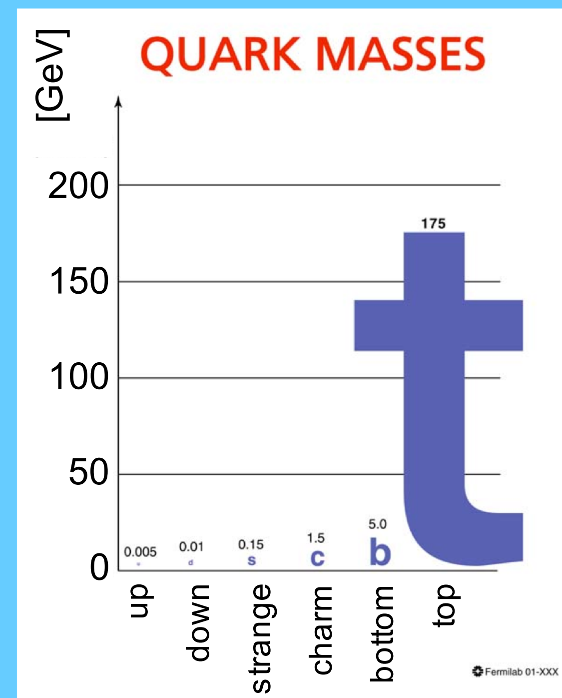
Top Quarks

- Spin 1/2 fermion,   charge +2/3
- Weak-isospin partner of the bottom quark
- ~40x heavier than its partner
- **M_{top} = 171.4 ± 2.1 GeV**
- Heaviest known fundamental particle



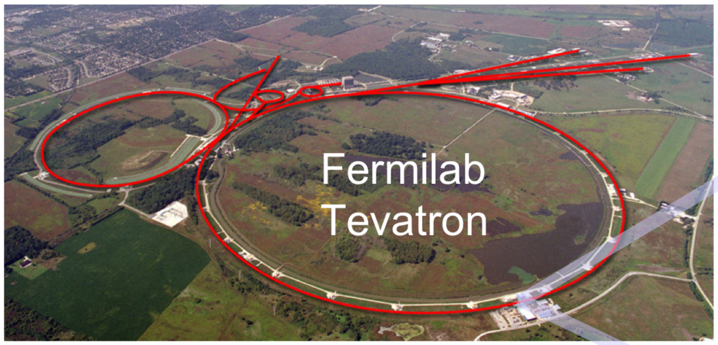
- Produced mostly in $t\bar{t}$ pairs at the Tevatron
- 85% $q\bar{q}$,   15% gg
- Cross section = 6.8 ± 0.6 pb at NNLO
- Measurements consistent with this value

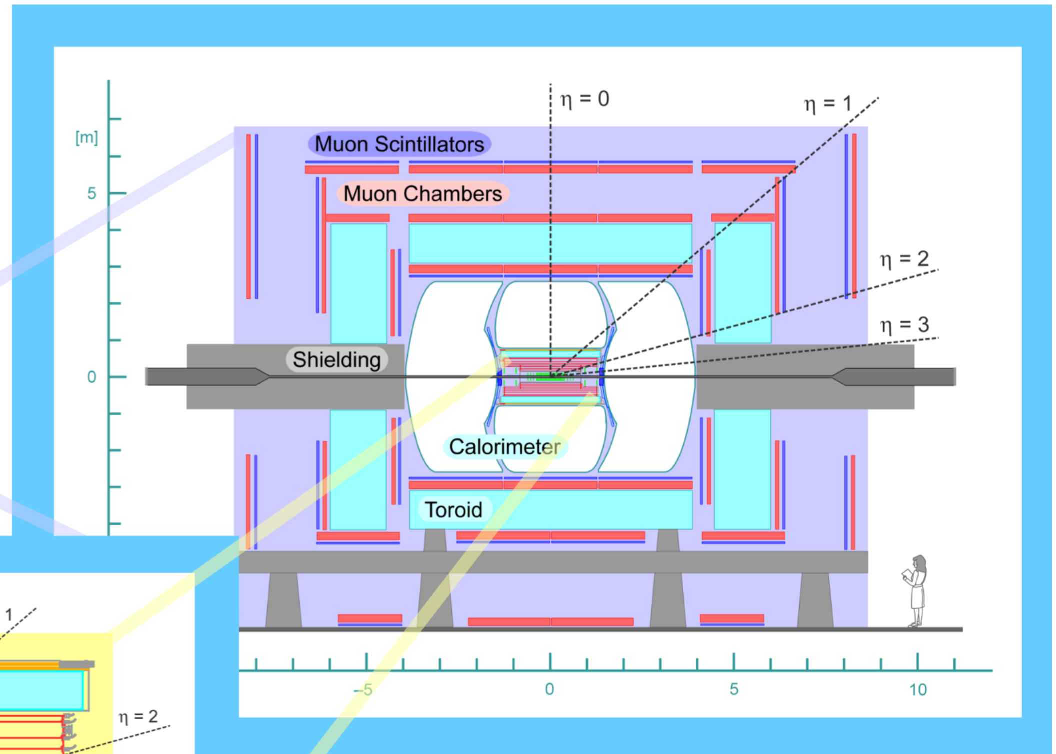


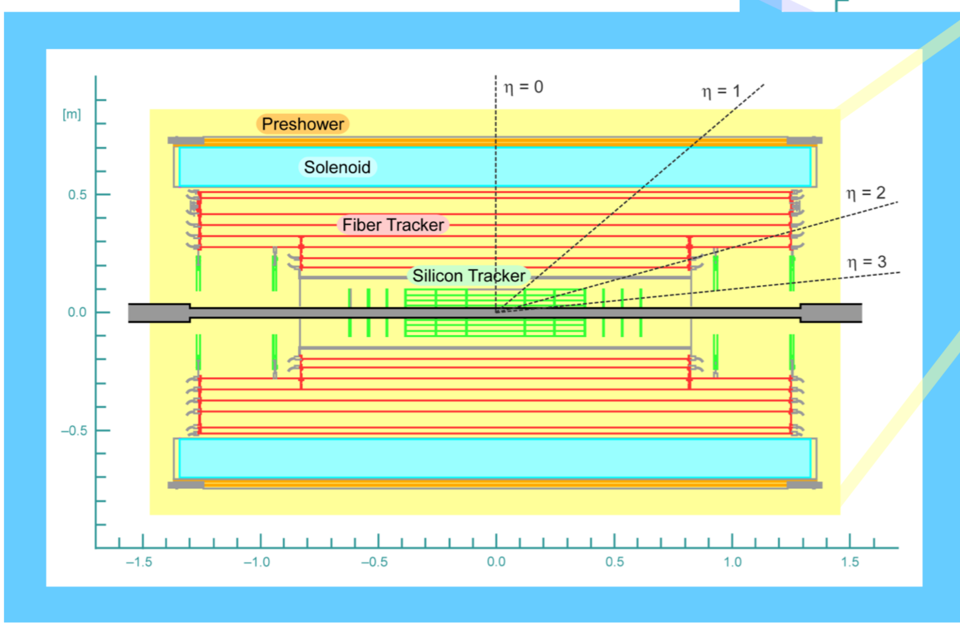


The DØ Experiment

- Top quarks observed by DØ and CDF in 1995 with ~50 pb^{-1} of data



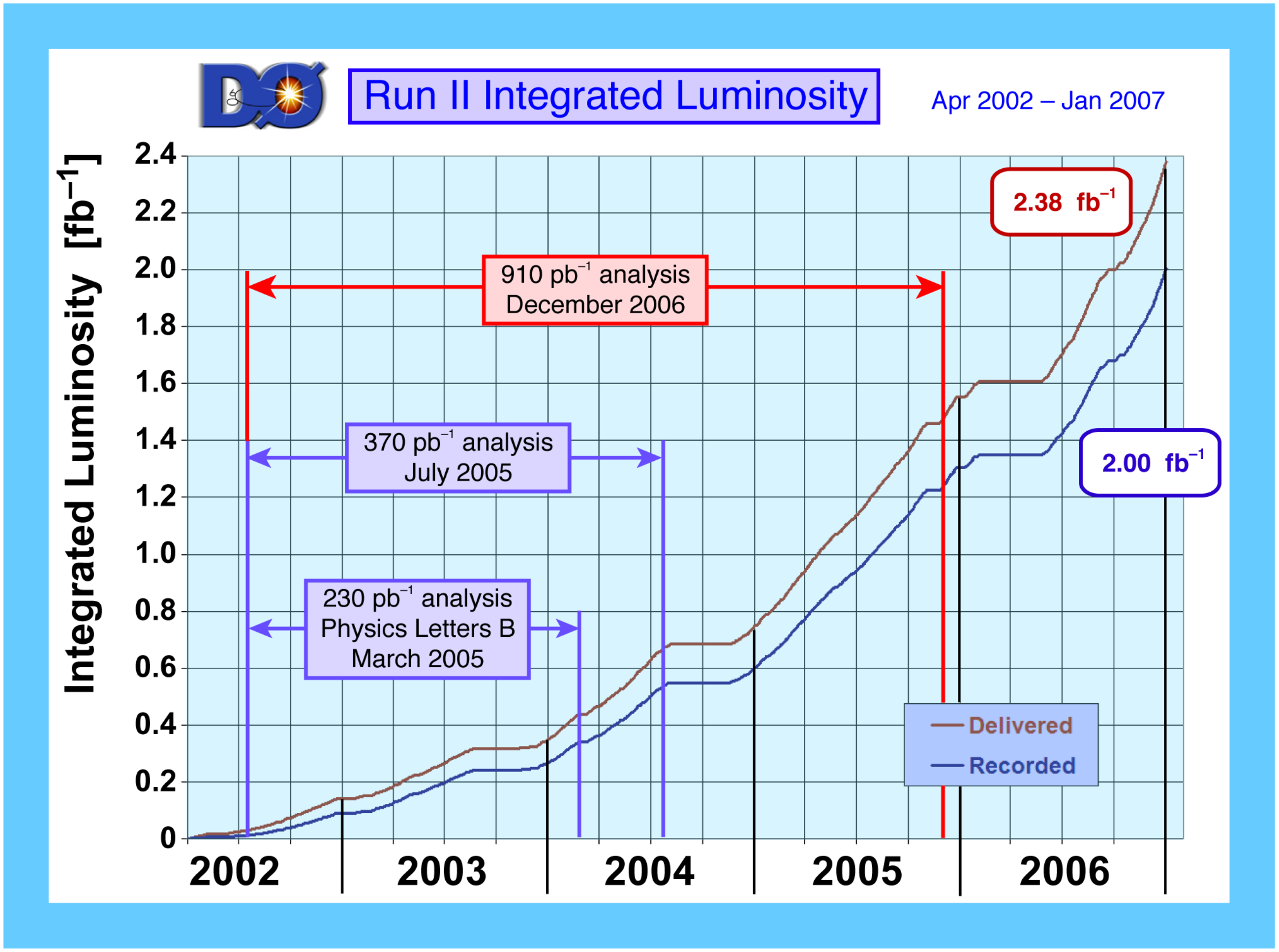




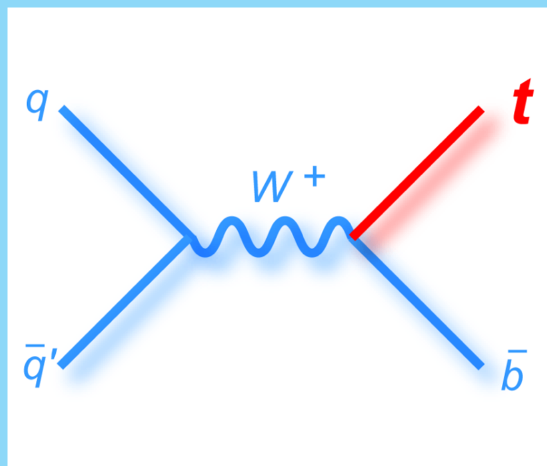
- Still the only place to see top
- Now have 40x more data → precision measurements

Dataset

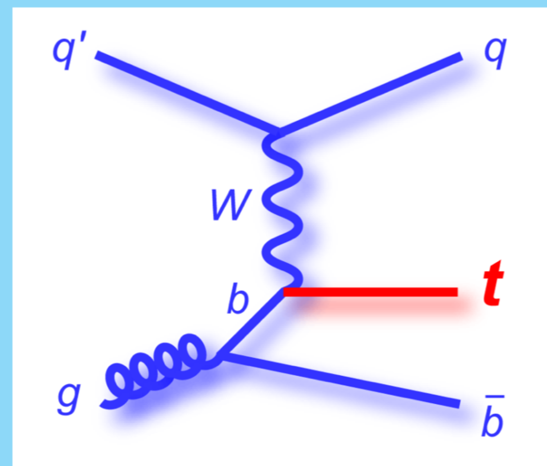
- DØ has 2 fb^{-1} on tape
- Many thanks to the Fermilab accelerator division!
- This analysis uses 0.9 fb^{-1} of data collected from 2002 to 2005



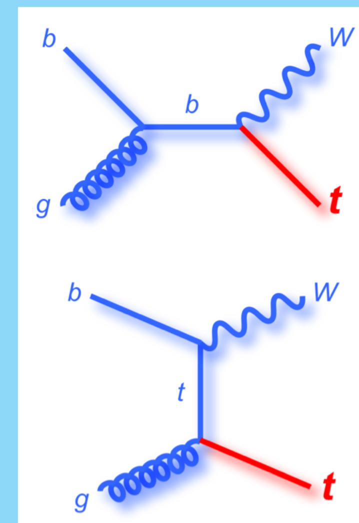
Single Top Overview



s-channel: "*tb*"
$\sigma_{NLO} = 0.88 \pm 0.11$ pb



t-channel: "*tqb*"
$\sigma_{NLO} = 1.98 \pm 0.25$ pb



"*tW* production"
$\sigma_{NLO} = 0.21$ pb
(Too small to see at the Tevatron)

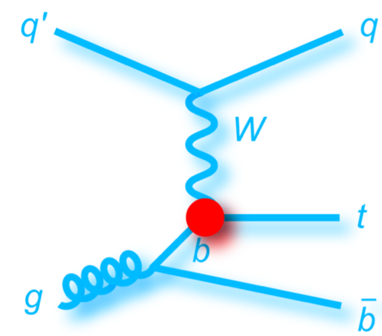
Experimental results (95% C.L.)

- **DØ** ***tb* < 5.0 pb** (370 pb^{-1})
- **CDF** ***tb* < 3.2 pb** (700 pb^{-1})

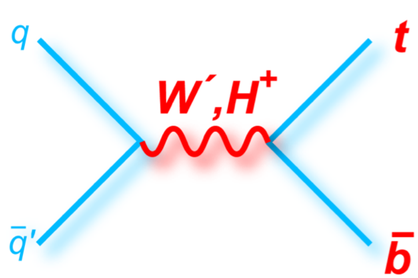
- **DØ** ***tqb* < 4.4 pb** (370 pb^{-1})
- **CDF** ***tqb* < 3.1 pb** (700 pb^{-1})

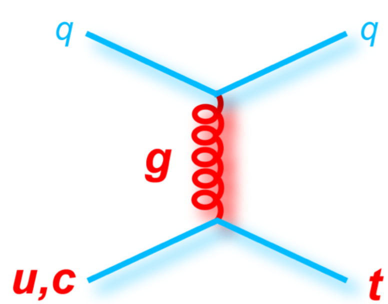
- **CDF** ***tb*+*tqb* < 2.7 pb** Likelihoods (960 pb^{-1})
  ***tb*+*tqb* < 2.6 pb** Neural networks
  ***tb*+*tqb* = 2.7 $^{+1.5}_{-1.3}$ pb** Matrix elements (significance of 2.3 σ)

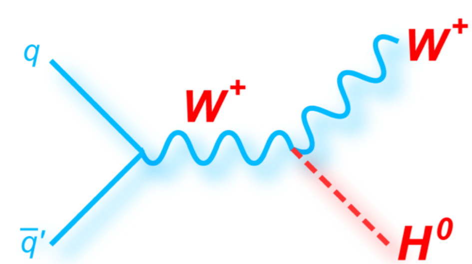
Motivation



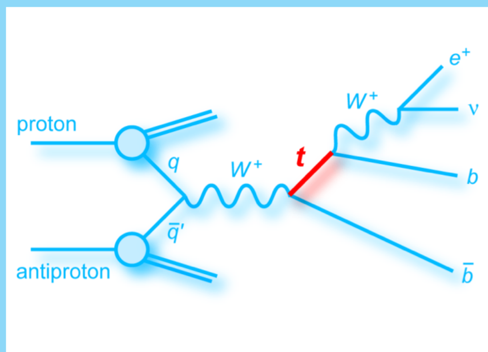
- **Study *Wtb* coupling in top production**
  - Measure $|V_{tb}|$ directly (more later)
  - Test unitarity of CKM matrix
  - Anomalous *Wtb* couplings
- **Cross sections sensitive to new physics**
  - s-channel: resonances (heavy W' boson, charged Higgs boson, Kaluza-Klein excited W_{KK}, technipion, etc.)
  - t-channel: flavor-changing neutral currents ($t - Z / \gamma / g - c / u$ couplings)
  - Fourth generation of quarks
- **Polarized top quarks – spin correlations measurable in decay products**
- **Measure top quark partial decay width and lifetime**
- **CP violation (same rate for top and antitop?)**
- **Similar (but easier) search than for *WH* associated Higgs production**
  - Backgrounds the same – must be able to model them successfully
  - Test of techniques to extract a small signal from a large background

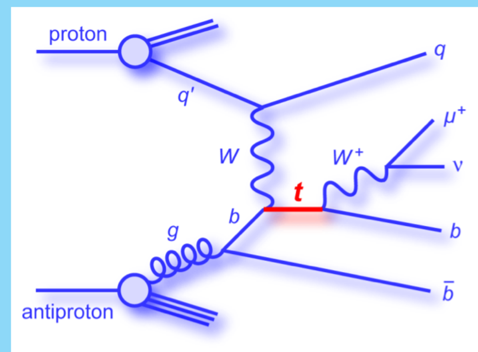




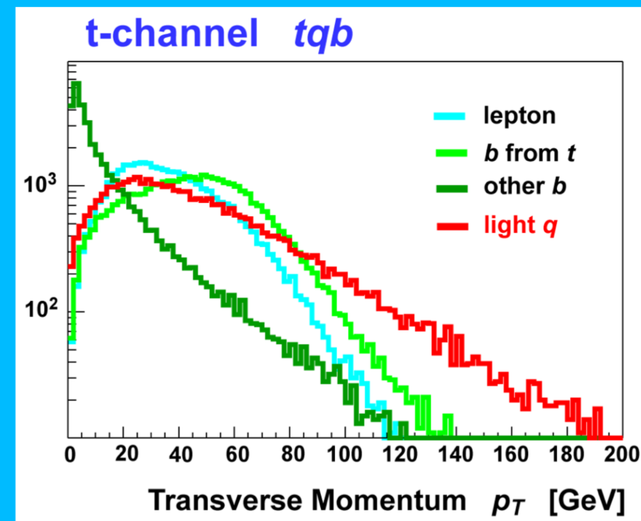


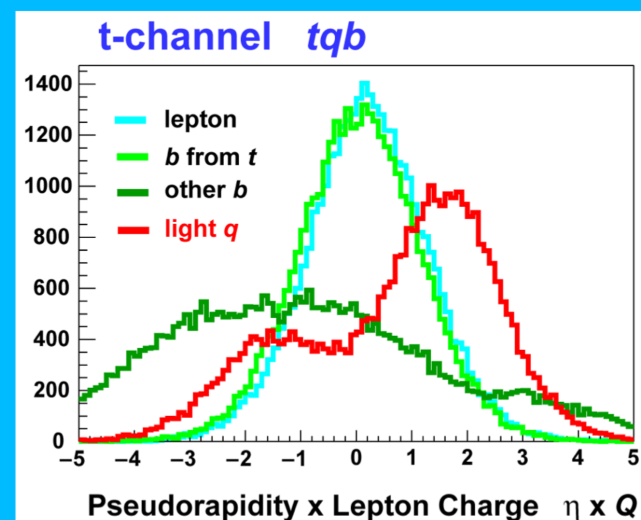
Event Selection





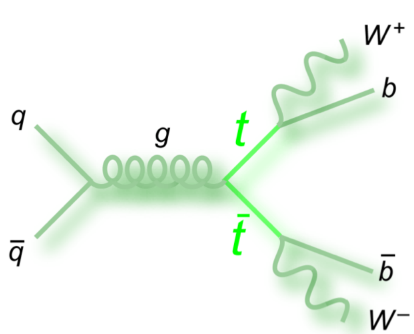
- **One isolated electron or muon**
  - Electron p_T > 15 GeV, $|\eta|$ < 1.1
  - Muon p_T > 18 GeV, $|\eta|$ < 2.0
- **Missing transverse energy**
  - $\not{E}_T$ > 15 GeV
- **One b-tagged jet and at least one more jet**
  - 2–4 jets with p_T > 15 GeV, $|\eta|$ < 3.4
  - Leading jet p_T > 25 GeV, $|\eta|$ < 2.5
  - Second leading jet p_T > 20 GeV

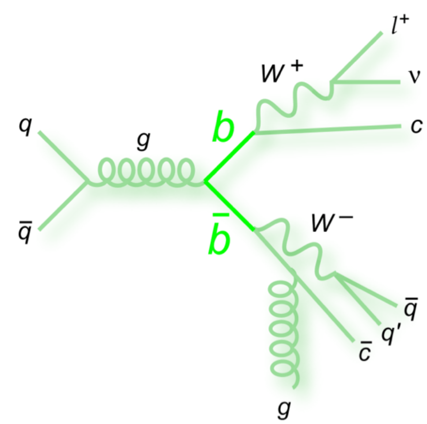


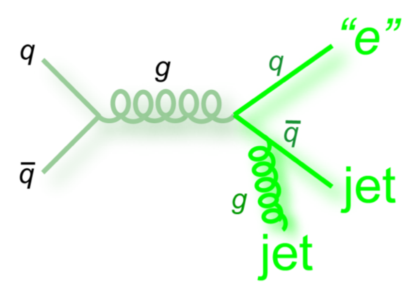


Signal and Background Models

- **Single top quark signals modeled using SINGLETOP**
  - By Moscow State University theorists, based on COMPHEP
  - Reproduces NLO kinematic distributions
  - PYTHIA for parton hadronization
- **$t\bar{t}$ pair backgrounds modeled using ALPGEN**
  - PYTHIA for parton hadronization
  - Parton-jet matching algorithm used to avoid double-counting final states
  - Normalized to NNLO cross section
  - 18% uncertainty includes component for top mass
- **Multijet background modeled using data with a non-isolated lepton and jets**
  - Normalized to data before *b*-tagging (together with *W*+jets background)

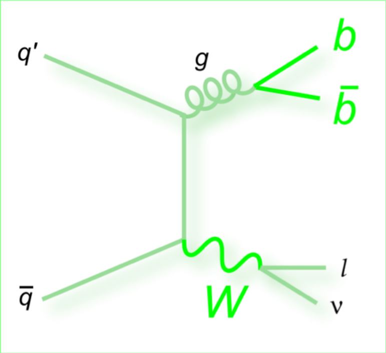


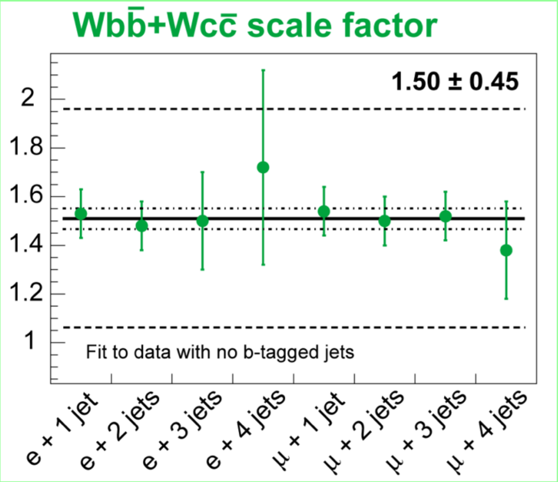


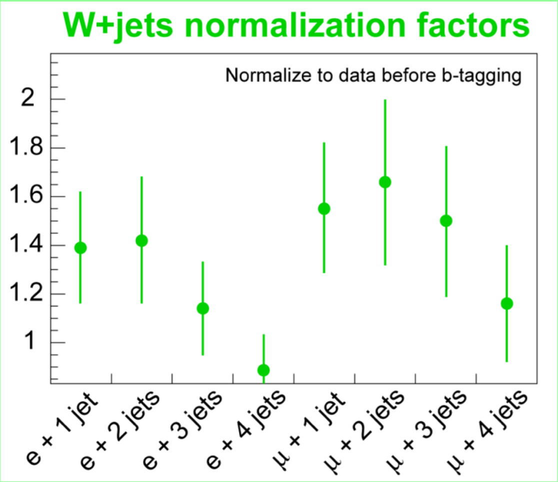


W+jets Background

- *W*+jets background modeled using ALPGEN
- PYTHIA for parton hadronization
- Parton-jet matching algorithm used to avoid double-counting final states
- $Wb\bar{b}$ and $Wc\bar{c}$ fractions from data to better represent higher-order effects
- 30% uncertainty for differences in event kinematics and assuming equal for $Wb\bar{b}$ and $Wc\bar{c}$
- *W*+jets normalized to data before *b*-tagging (with multijet background)
- *Z*+jets, diboson backgrounds very small, included in *W*+jets via normalization



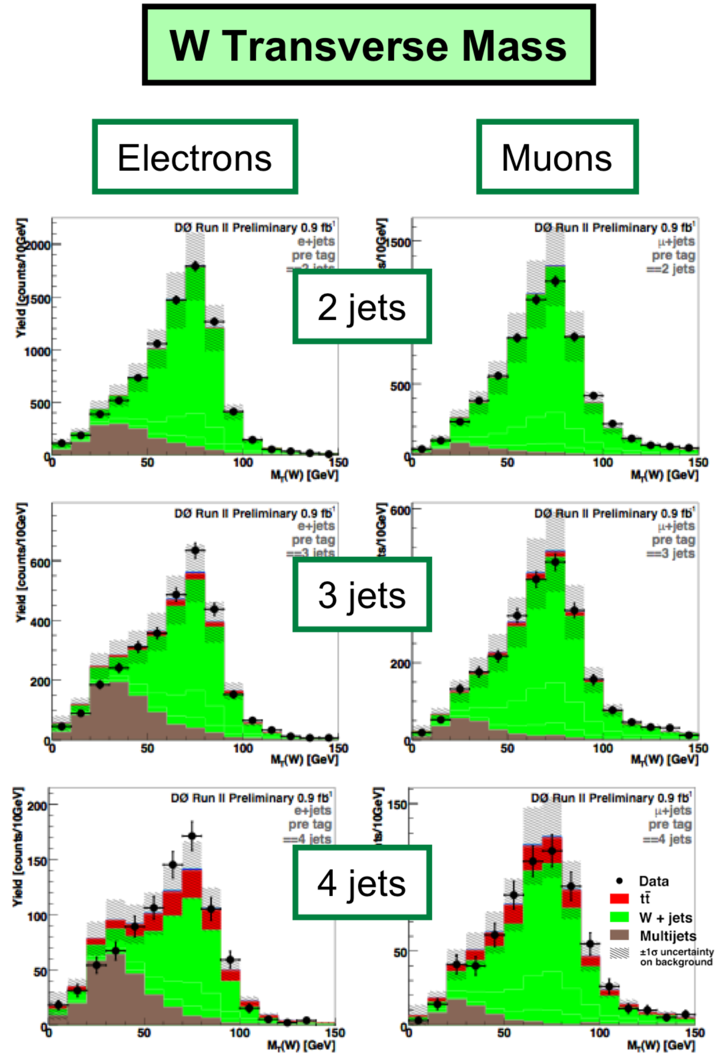




Event Yields Before *b*-Tagging

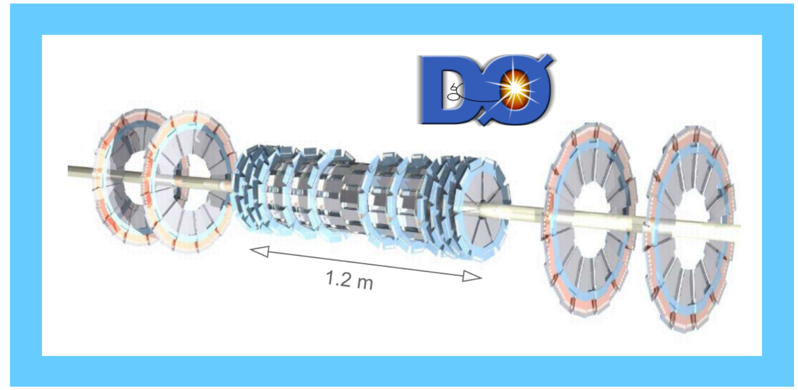
- Signal acceptances: *tb* = 5.1%, *tqb* = 4.5%
- **S:B ratio for *tb*+*tqb* = 1:180**
- Need to improve S:B to have a hope of seeing a signal → select only events with *b*-jets in them

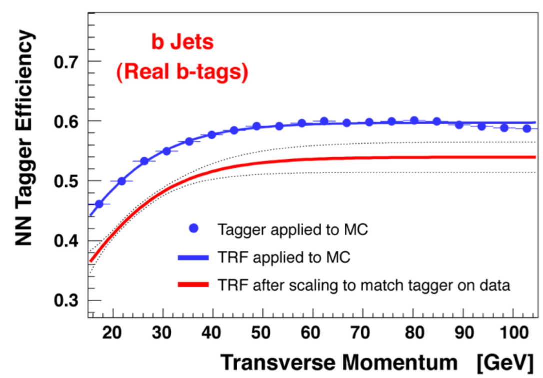
Source	Event Yields in 0.9 fb^{-1} Data Electron+muon combined, before *b*-tagging		
	2 jets	3 jets	4 jets
tb	25	12	3
tqb	47	25	8
$t\bar{t} \rightarrow ll$	62	50	18
$t\bar{t} \rightarrow l$+jets	40	175	227
$W+b\bar{b}$	670	310	89
$W+c\bar{c}$	1,959	912	224
$W+jj$	10,160	3,138	728
Multijets	1,762	1,083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

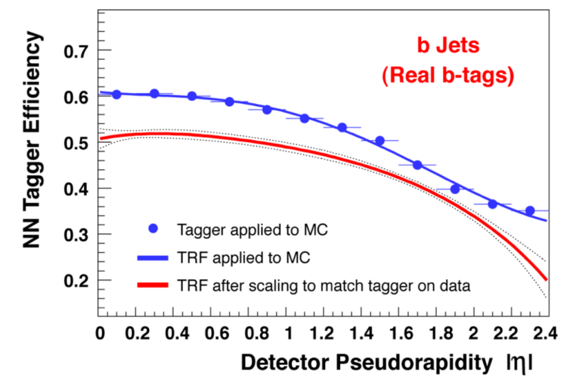


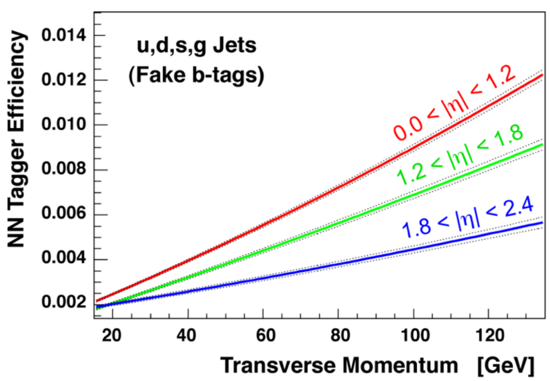
b-Jet Identification

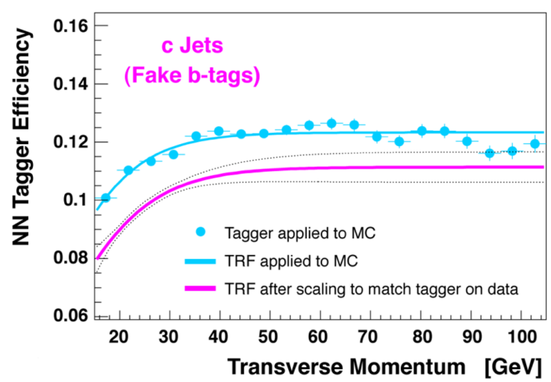
- Separate *b*-jets from light-quark and gluon jets to reject most *W*+jets background
- **DØ uses a neural network algorithm**
  - 7 input variables based on impact parameter and reconstructed vertex
- **Operating point:**
  - *b*-jet efficiency ≈ 50%
  - *c*-jet efficiency ≈ 10%
  - light-jet effic. ≈ 0.5%

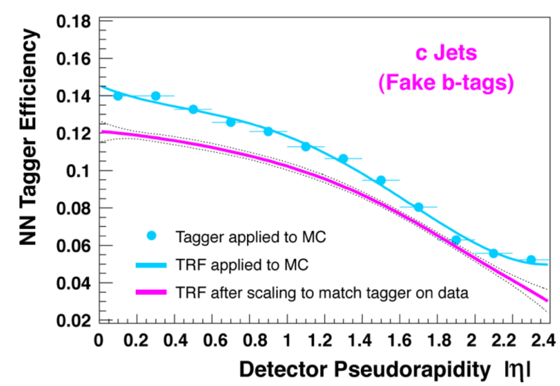












Event Yields after *b*-Tagging

Source	Event Yields in 0.9 fb^{-1} Data Electron+muon, 1tag+2tags combined 2 jets	3 jets	4 jets
tb	16 ± 3	8 ± 2	2 ± 1
tqb	20 ± 4	12 ± 3	4 ± 1
$t\bar{t} \rightarrow ll$	39 ± 9	32 ± 7	11 ± 3
$t\bar{t} \rightarrow l$+jets	20 ± 5	103 ± 25	143 ± 33
$W+b\bar{b}$	261 ± 55	120 ± 24	35 ± 7
$W+c\bar{c}$	151 ± 31	85 ± 17	23 ± 5
$W+jj$	119 ± 25	43 ± 9	12 ± 2
Multijets	95 ± 19	77 ± 15	29 ± 6
Total background	686 ± 41	460 ± 39	253 ± 38
Data	697	455	246

- Signal acceptances: *tb* = (3.2 ± 0.4)%, *tqb* = (2.1 ± 0.3)%
- **Signal:background ratios for *tb*+*tqb* are 1:10 to 1:50**
  - Most sensitive channels have 2jets/1tag, S:B = 1:20
- **Single top signal is smaller than total background uncertainty**
  - counting events is not a sensitive enough method
  - use a multivariate discriminant to separate signal from background

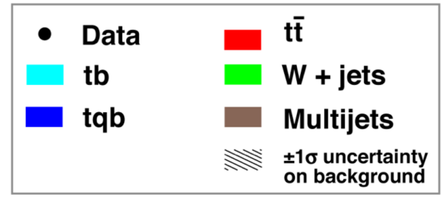
Search Strategy Summary

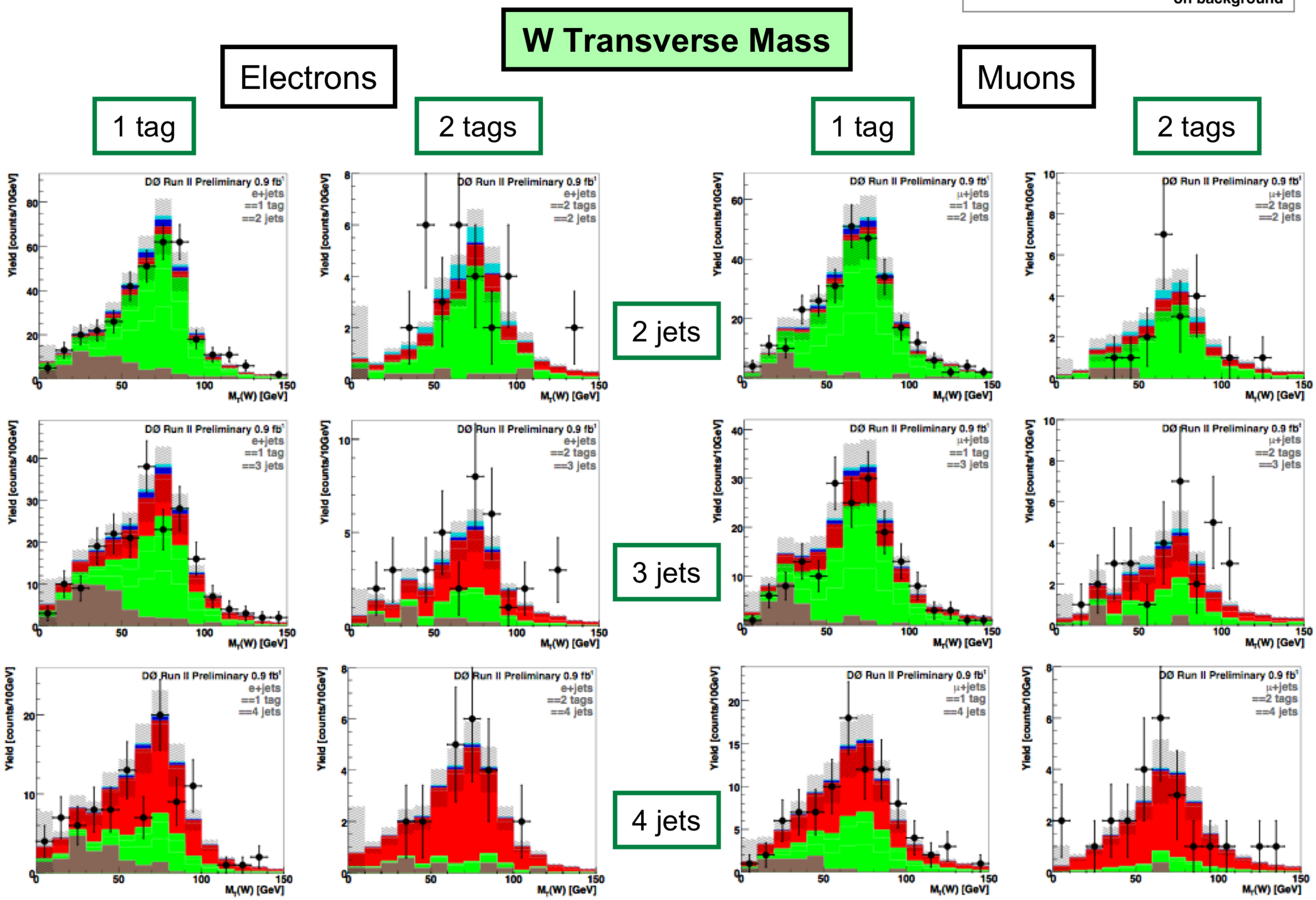
- **Maximize the signal acceptance**
  - Particle ID definitions set as loose as possible (i.e., highest efficiency, separate signal from backgrounds with fake leptons later)
  - Transverse momentum thresholds set low, pseudorapidities wide
  - As many decay channels used as possible – this analysis shown in red box
  - Channels analyzed separately since S:B and background compositions differ
- **Separate signal from background using multivariate techniques**

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

12 Analysis Channels





Systematic Uncertainties

- Uncertainties are assigned for each signal and background component in all analysis channels
- Most systematic uncertainties apply only to normalization
- **Two sources of uncertainty also affect the shapes of distributions**
  - jet energy scale
  - tag-rate functions for *b*-tagging MC events
- Correlations between channels and sources are taken into account
- **Cross section uncertainties are dominated by the statistical uncertainty,** the systematic contributions are all small

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%

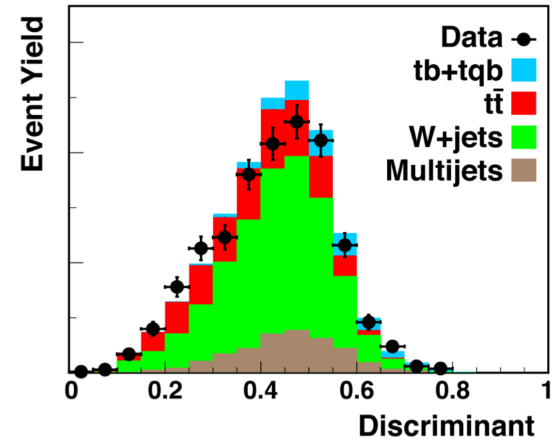
Final Analysis Steps

- We have selected 12 independent sets of data for final analysis
- Background model gives good representation of data in ~90 variables in every channel
- **Calculate discriminants that separate signal from background**
  - Boosted decision trees
  - Matrix elements
  - Bayesian neural networks
- Check discriminant performance using data control samples
- Use ensembles of pseudo-data to test validity of methods
- **Calculate cross sections using binned likelihood fits of (floating) signal + (fixed) background to data**

Measuring a Cross Section

$$d = S + B = \sigma \mathcal{A} \mathcal{L} + B = \sigma a + \sum_{i=1}^{\mathrm{Nbkgds}} b_i$$

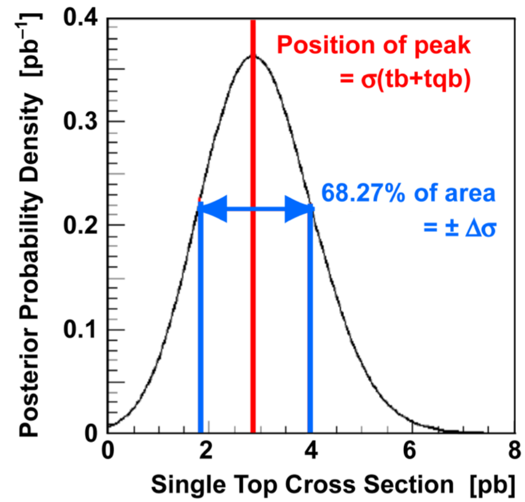
d = Predicted number of data events
S = Predicted number of signal events
B = Predicted number of background events
σ = Cross section
$\mathcal{A}$ = Signal acceptance
$\mathcal{L}$ = Integrated luminosity
a = Effective luminosity
b_i = No. of events in each background component



$$\mathrm{Prob}(D|d) \equiv \mathrm{Prob}(D|\sigma, a, \mathbf{b}) = \prod_{i=1}^{\mathrm{Nbins}} \mathrm{Prob}(D_i|d_i)$$

D = Observed number of data events
$\mathbf{b}$ = Vector of background components

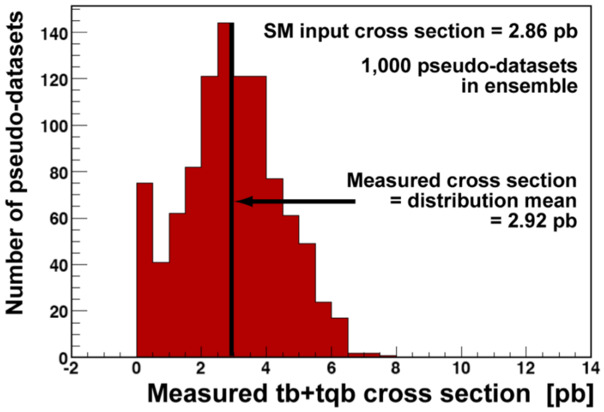
$$\text{Posterior Probability Density}(\sigma|D) \propto \int_a \int_{\mathbf{b}} \mathrm{Prob}(D|\sigma, a, \mathbf{b})\ \mathrm{Prior}(a, \mathbf{b})\ \mathrm{Prior}(\sigma)\ da\ d\mathbf{b}$$



- Nbkgds = 6 (*ttll, ttlj, Wbb, Wcc, Wjj,* multijets), Nbins = 12 chans x 100 bins = 1,200
- Cross section obtained from peak position of Bayesian posterior probability density
- Shape and normalization systematic uncertainties treated as nuisance parameters
- Correlations between uncertainties are properly accounted for
- Signal cross section prior is non-negative and flat

Testing with Pseudo-Data

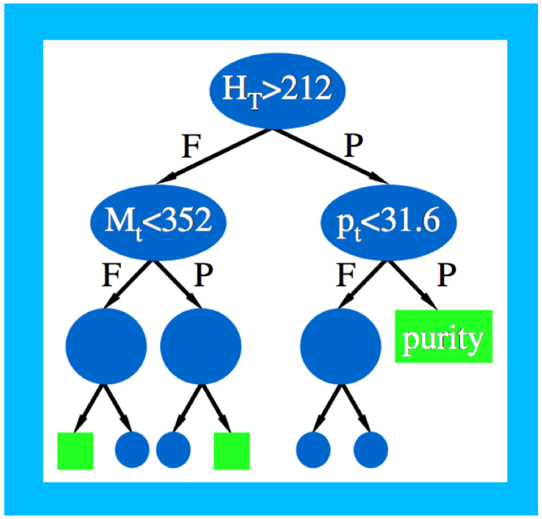
- To verify that the calculation methods work as expected, we test them using several sets ("ensembles") of pseudo-data
- Wonderful tool to test the analyses! Like running DØ many 1,000's of times
- **Select subsets of events from total pool of MC events**
  - Randomly sample a Poisson distribution to simulate statistical fluctuations
  - Background yields fluctuated according to uncertainties to reproduce correlations between components from normalization
- **Ensembles we used:**
  - Zero-signal ensemble, $\sigma(tb+tqb)$ = 0 pb
  - SM ensemble, $\sigma(tb+tqb)$ = 2.9 pb
  - "Mystery" ensembles, $\sigma(tb+tqb)$ = ? pb
  - Measured Xsec ensemble, $\sigma(tb+tqb) = \sigma_{meas}$



- Each pseudo-dataset is like one DØ experiment with 0.9 fb^{-1} of "data", up to 68,000 pseudo-datasets per ensemble

Signal-Background Separation using Decision Trees

- Machine-learning technique, widely used in social sciences, some use in HEP
- Idea: recover events that fail criteria in cut-based analyses



- Start at first "node " with "training sample" of 1/3 of all signal and background events
  - 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 to produce two "branches ⟶ " with corresponding events, (F)ailed and (P)assed cut
- Repeat recursively on each node
- Stop when improvement stops or when too few events are left (100)
- Terminal node is called a "leaf " with purity = $N_{signal}/(N_{signal}+N_{background})$
- Run remaining 2/3 events and data through tree to derive results
- Decision tree output for each event = leaf purity (closer to 0 for background, nearer 1 for signal)

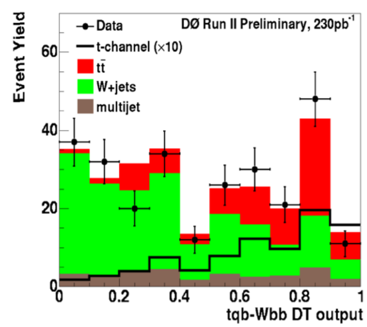
Boosting the Decision Trees

- Boosting is a recently developed technique that improves any weak classifier (decision tree, neural network, etc)
- Recently used with decision trees by GLAST and MiniBooNE
- Boosting averages the results of many trees, dilutes the discrete nature of the output, improves the performance

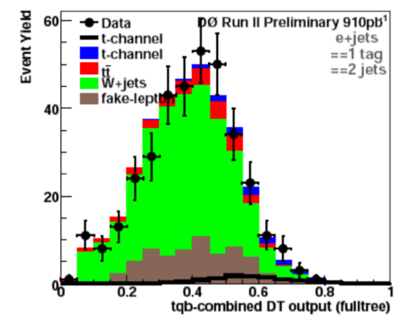
This analysis:

- Uses the "adaptive boosting algorithm":
  - Train a tree T_k
  - Check which events are misclassified by T_k
  - Derive tree weight w_k
  - Increase weight of misclassified events
  - Train again to build T_{k+1}
  - Boosted result of event i : $T(i) = \sum_{n=1}^{N_{tree}} w_k T_k(i)$
- 20 boosting cycles

Before boosting



After boosting



- Trained 36 sets of trees: (*tb+tqb, tb, tqb*) x (e,μ) x (2,3,4 jets) x (1,2 *b*-tags)
  - Separate analyses for *tb* and *tqb* allow access to different types of new physics
  - Search for *tb+tqb* has best sensitivity to see a signal – results presented here

Decision Tree 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)
- $\cos(\text{best1,lepton})_{\text{besttop}}$
- $\cos(\text{best1,notbest1})_{\text{besttop}}$
- $\cos(\text{tag1,alljets})_{\text{alljets}}$
- $\cos(\text{tag1,lepton})_{\text{btaggedtop}}$
- $\cos(\text{jet1,alljets})_{\text{alljets}}$
- $\cos(\text{jet1,lepton})_{\text{btaggedtop}}$
- $\cos(\text{jet2,alljets})_{\text{alljets}}$
- $\cos(\text{jet2,lepton})_{\text{btaggedtop}}$
- $\cos(\text{lepton}, Q(\text{lepton}) \times z)_{\text{besttop}}$
- $\cos(\text{lepton}_{\text{besttop}}, \text{besttop}_{\text{CofM}})$
- $\cos(\text{lepton}_{\text{btaggedtop}}, \text{btaggedtop}_{\text{CofM}})$
- $\cos(\text{notbest,alljets})_{\text{alljets}}$
- $\cos(\text{notbest,lepton})_{\text{besttop}}$
- $\cos(\text{untag1,alljets})_{\text{alljets}}$
- $\cos(\text{untag1,lepton})_{\text{btaggedtop}}$

Event Kinematics

- 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)
- $Q(\text{lepton}) \times \eta(\text{untag1})$
- $\sqrt{\hat{s}}$
- Sphericity(alljets, W)

Most discrimination power:

M(alljets)

M(W,tag1)

$\cos(\text{tag1,lepton})_{\text{btaggedtop}}$

Q(lepton) x η(untag1)

- 49 input variables
- Adding more variables does not degrade the performance
- Reducing the number of variables always reduces sensitivity of the analysis
- Same list of variables used for all analysis channels

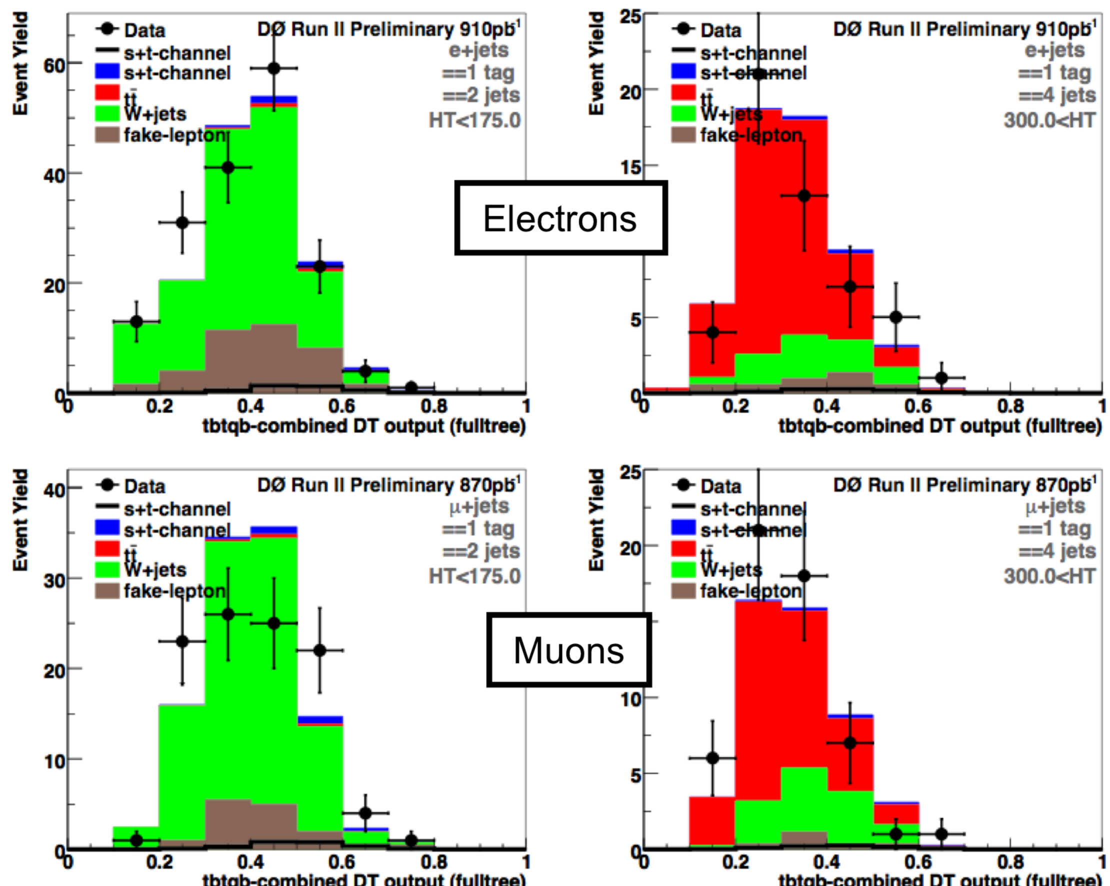
Decision Tree Cross Checks

- Select two background-dominated samples:
  - "*W*+jets": = 2 jets, H_T(lepton, $\not{E}_T$, alljets) < 175 GeV, =1 tag
  - "$t\bar{t}$": = 4 jets, H_T (lepton, $\not{E}_T$, alljets) > 300 GeV , =1 tag
- Observe good data-background agreement

Decision Tree Outputs

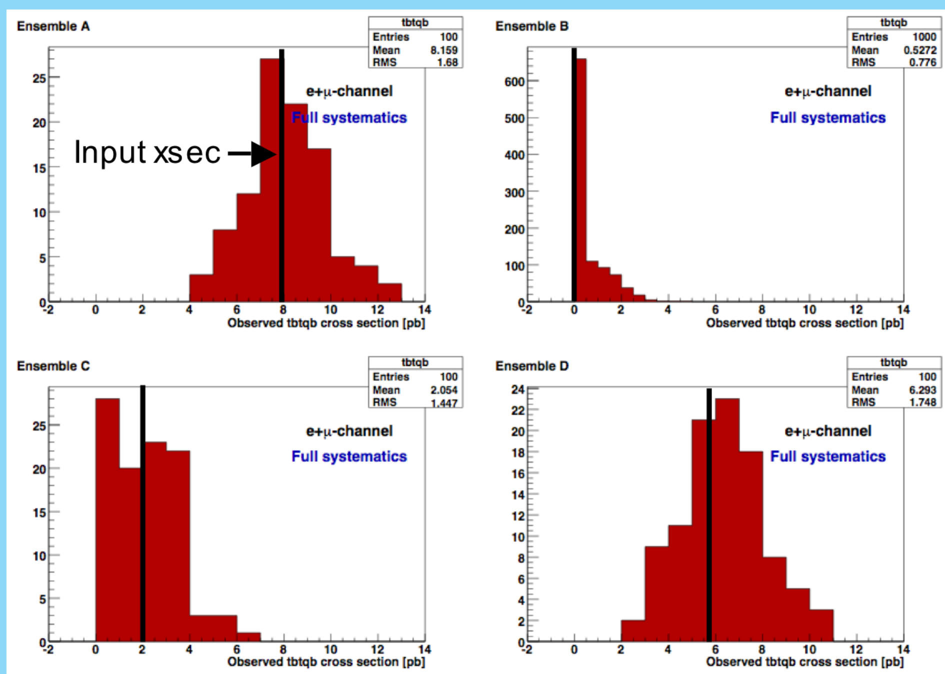
"*W*+jets"

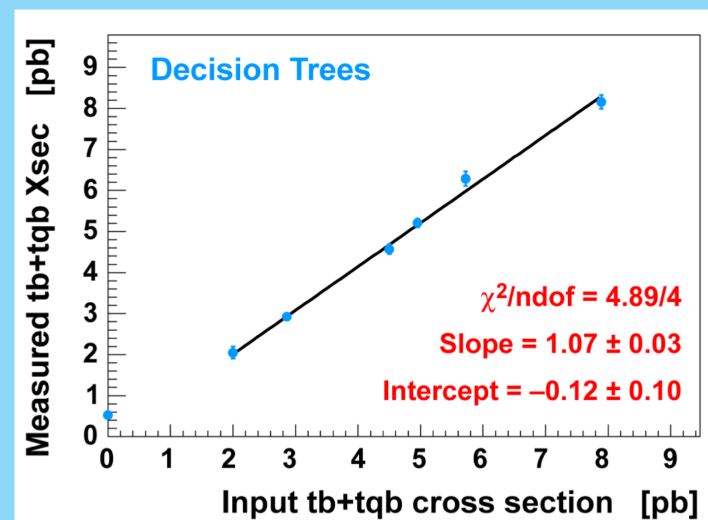
"$t\bar{t}$"



Decision Tree Verification

- Use “mystery” ensembles with many different signal assumptions
- Measure signal cross section using decision tree outputs
- Compare measured cross sections to input ones
- **Observe linear relation close to unit slope**





ME ME ME ME ME ME ME ME ME ME ME

Signal-Background Separation using Matrix Elements

- Method pioneered by DØ for top quark mass measurement
- Use the 4-vectors of all reconstructed leptons and jets
- Use matrix elements of main signal and background Feynman diagrams to compute an event probability density for signal and background hypotheses
- Goal: calculate a discriminant:

$$D_s(\vec{x}) = P(S|\vec{x}) = \frac{P_{Signal}(\vec{x})}{P_{Signal}(\vec{x}) + P_{Background}(\vec{x})}$$

- Define P_{Signal} as a normalized differential cross section:

$$P_{Signal}(\vec{x}) = \frac{1}{\sigma_S} d\sigma_S(\vec{x}) \quad \sigma_S = \int d\sigma_S(\vec{x})$$

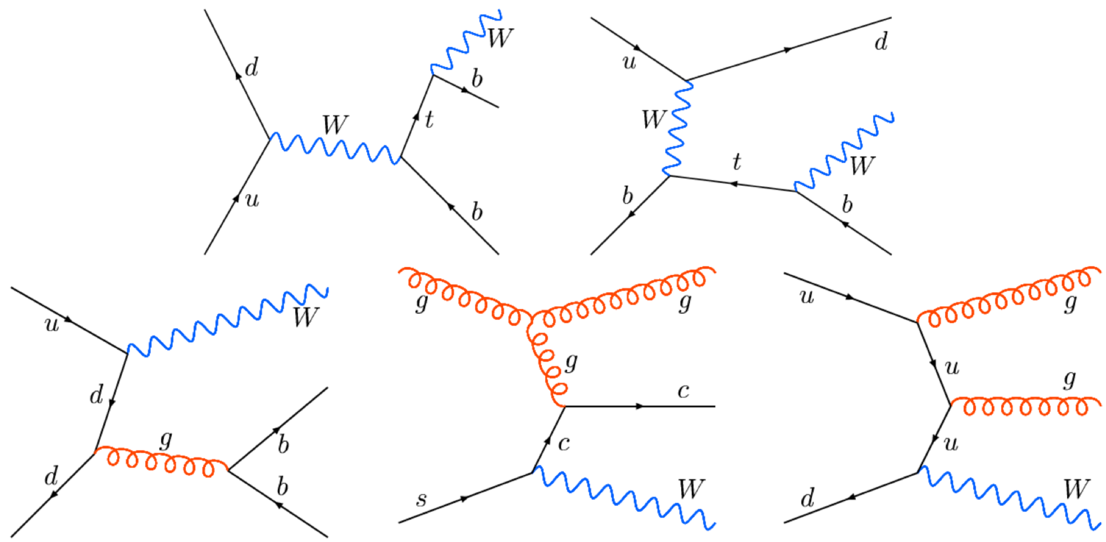
- Performed in 2-jets and 3-jets channels only
- No matrix element for $t\bar{t}$ so no discrimination between signal and top pairs yet
- Matrix element verification with ensembles shows good linearity, unit slope, near-zero intercept

Matrix Element Method
Feynman Diagrams

2-jet channels

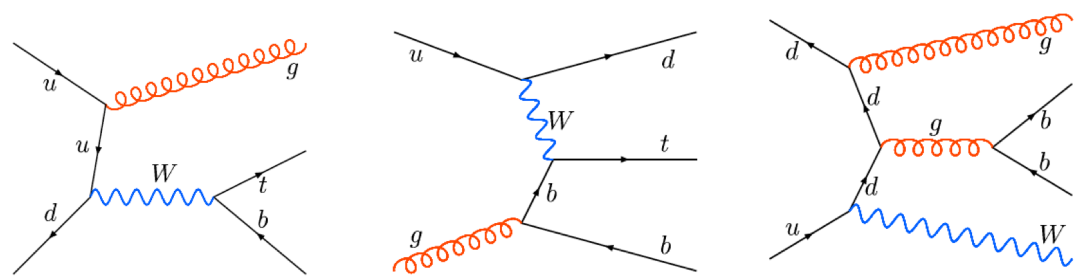
tb *tq*

Wbb *Wcg* *Wgg*



3-jet channels

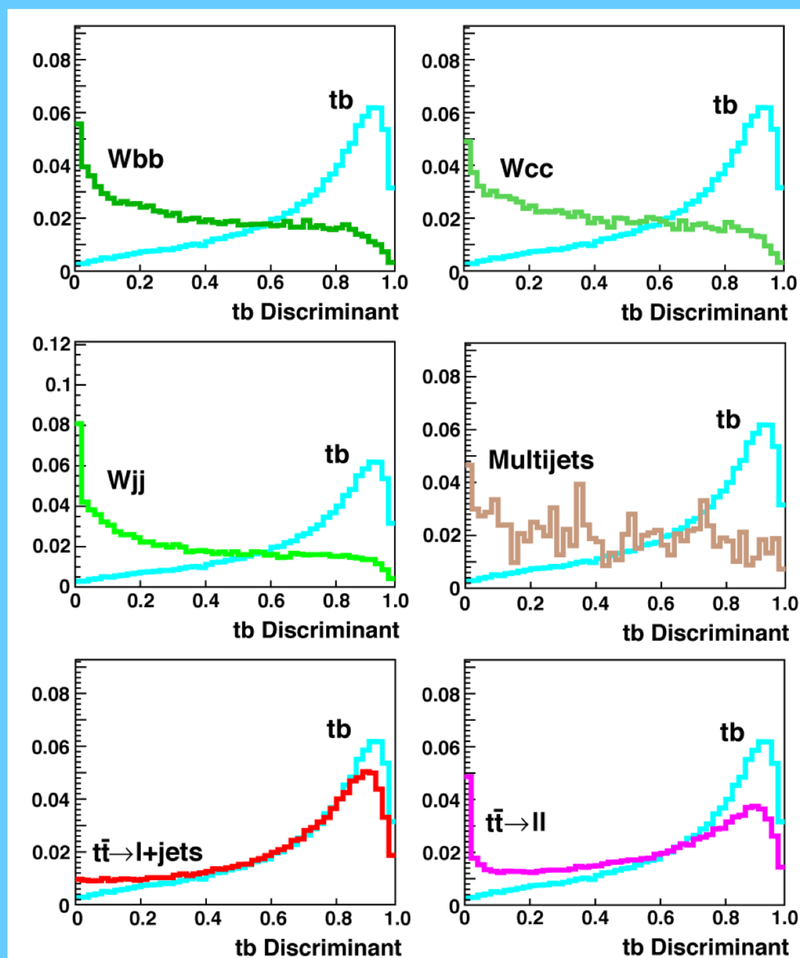
tbg *tdb* *Wbbg*



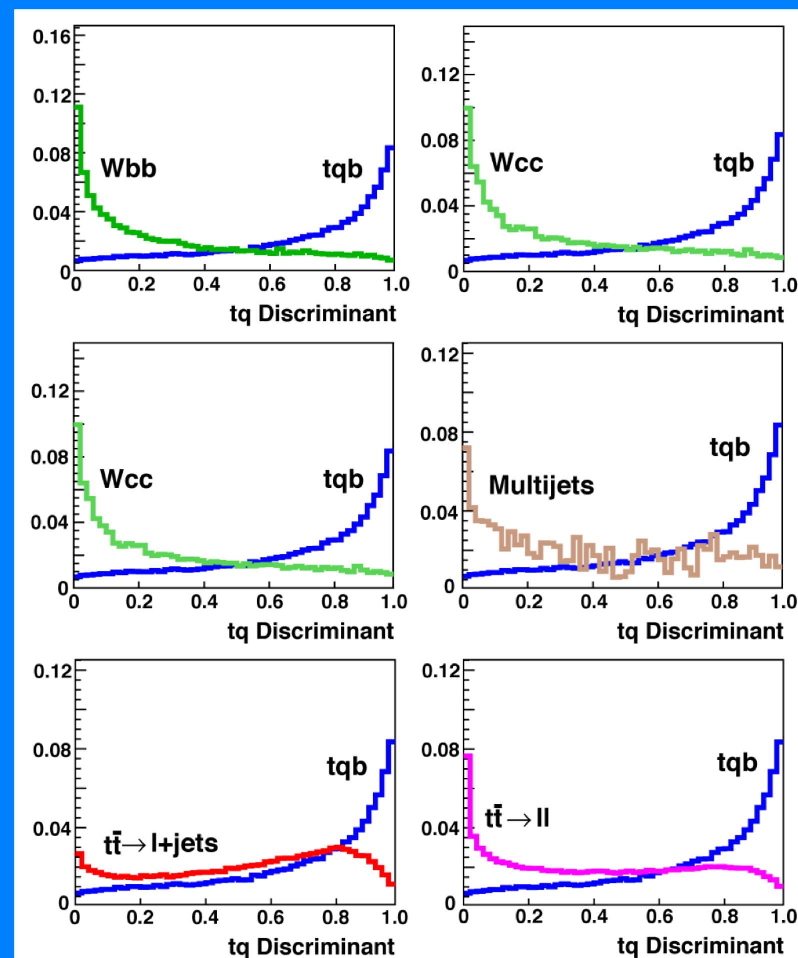
Matrix Element S:B Separation

2-jet channels

tb discriminant

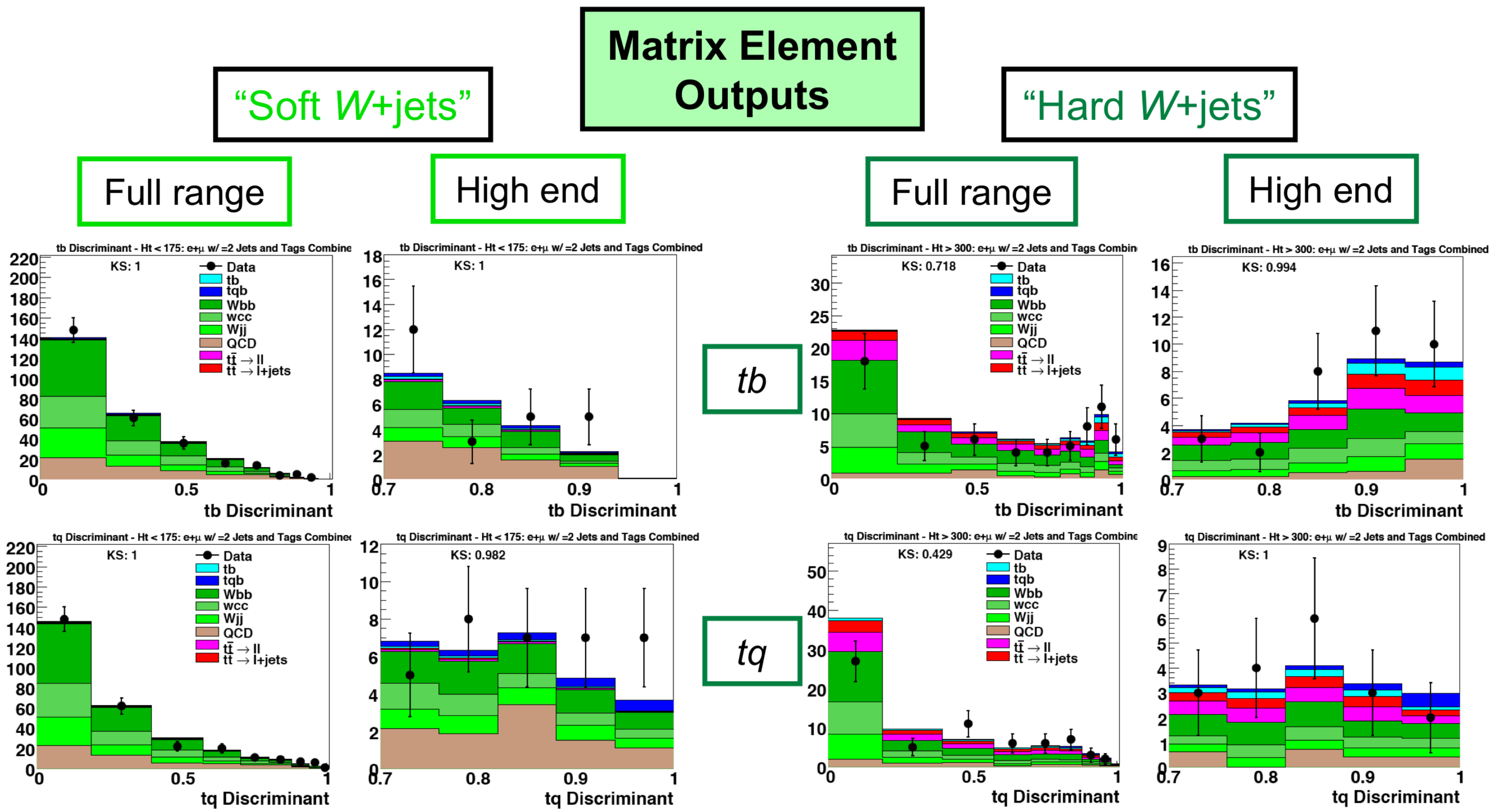


tq discriminant



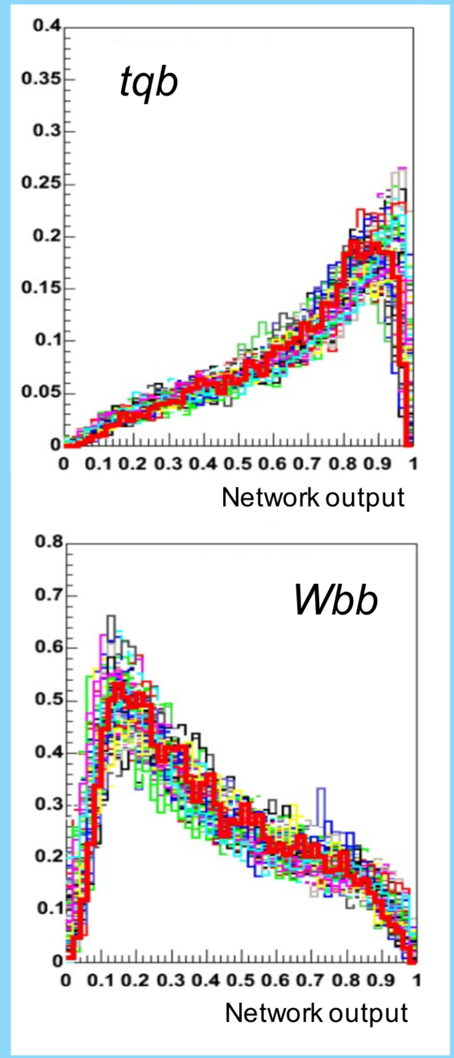
Matrix Element Cross Checks

- Select two background-dominated samples:
  - "Soft *W*+jets": = 2 jets, H_T(lepton, $\not{E}_T$, alljets) < 175 GeV, =1 tag
  - "Hard *W*+jets": = 2 jets, H_T(lepton, $\not{E}_T$, alljets) > 300 GeV, =1 tag
- Observe good data-background agreement



Signal-Background Separation using Bayesian Neural Networks

- Neural networks use many input variables, train on signal and background samples, produce one output discriminant
- **Bayesian neural networks improve on this technique:**
  - Average over many networks weighted by the probability of each network given the training samples
  - Less prone to over-training
  - Network structure is less important – can use larger numbers of variables and hidden nodes
- **For this analysis:**
  - 24 input variables (subset of 49 used by decision trees)
  - 40 hidden nodes, 800 training iterations
  - Each iteration is the average of 20 training cycles
  - One network for each signal (*tb+tqb, tb, tqb*) in each of the 12 analysis channels
- Bayesian neural network verification with ensembles shows good linearity, unit slope, near-zero intercept



Statistical Analysis

Before looking at the data, we want to know two things:

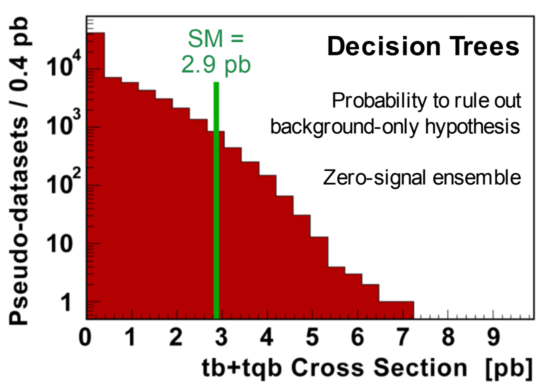
- **By how much can we expect to rule out a background-only hypothesis?**
  - Find what fraction of the ensemble of zero-signal pseudo-datasets give a cross section at least as large as the SM value, the "expected p-value"
  - For a Gaussian distribution, convert p-value to give "expected signficance"
- **What precision should we expect for a measurement?**
  - Set value for "data" = SM signal + background in each discriminant bin (non-integer) and measure central value and uncertainty on the "expected cross section"

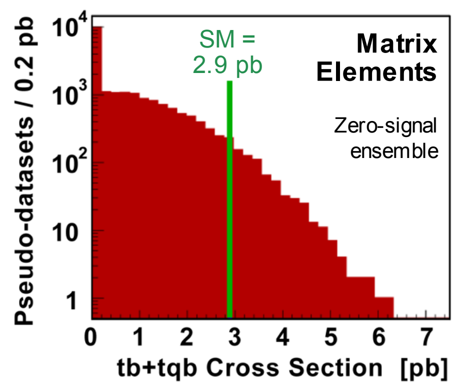
With the data, we want to know:

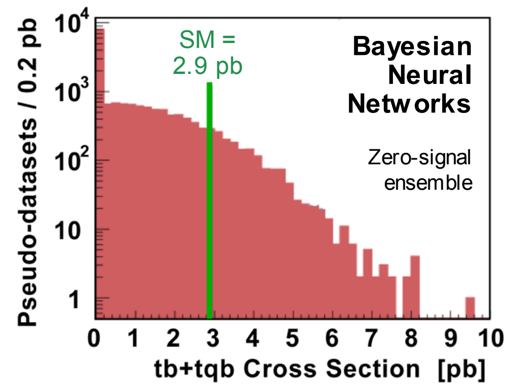
- **How well do we rule out the background-only hypothesis?**
  - Use the ensemble of zero-signal pseudo-datasets and find what fraction give a cross section at least as large as the measured value, the "measured p-value"
  - Convert p-value to give "measured signficance"
- **What cross section do we measure?**
  - Use (integer) number of data events in each bin to obtain "measured cross section"
- **How consistent is the measured cross section with the SM value?**
  - Find what fraction of the ensemble of SM-signal pseudo-datasets give a cross section at least as large as the measured value to get "consistency with SM"

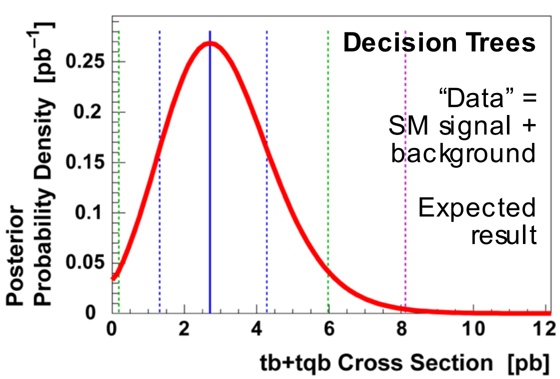
Expected Results

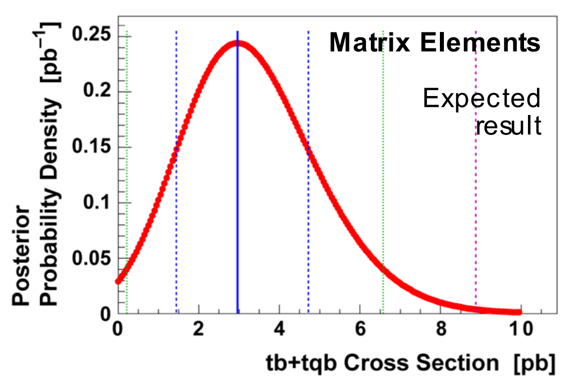
	Decision Trees	Matrix Elements	Bayesian NNs
Expected p-value	1.9 %	3.7 %	9.7 %
Expected significance	2.1 σ	1.8 σ	1.3 σ
Expected cross section	$2.7^{+1.6}_{-1.4}$ pb	$3.0^{+1.8}_{-1.5}$ pb	$3.2^{+2.0}_{-1.8}$ pb

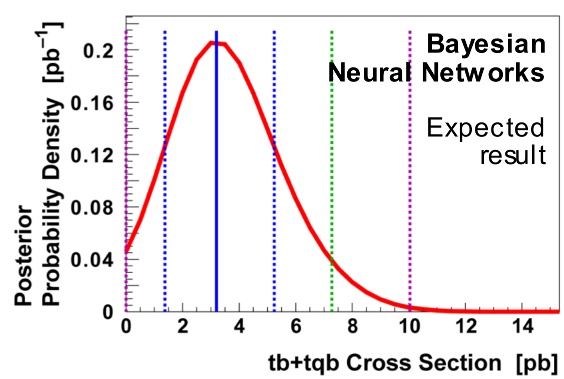




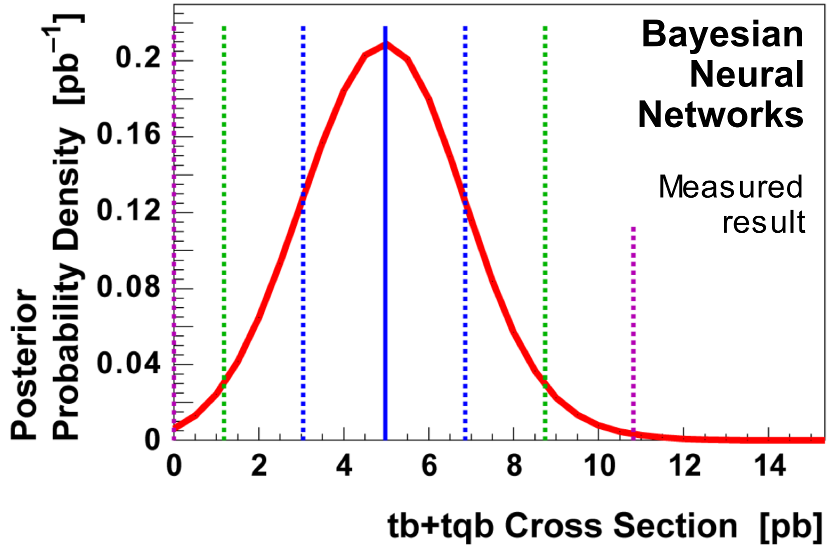








Bayesian NN Results

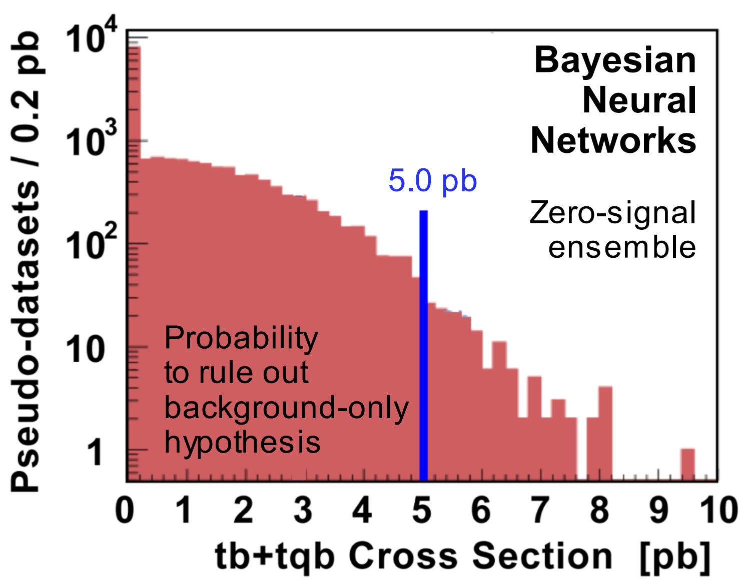


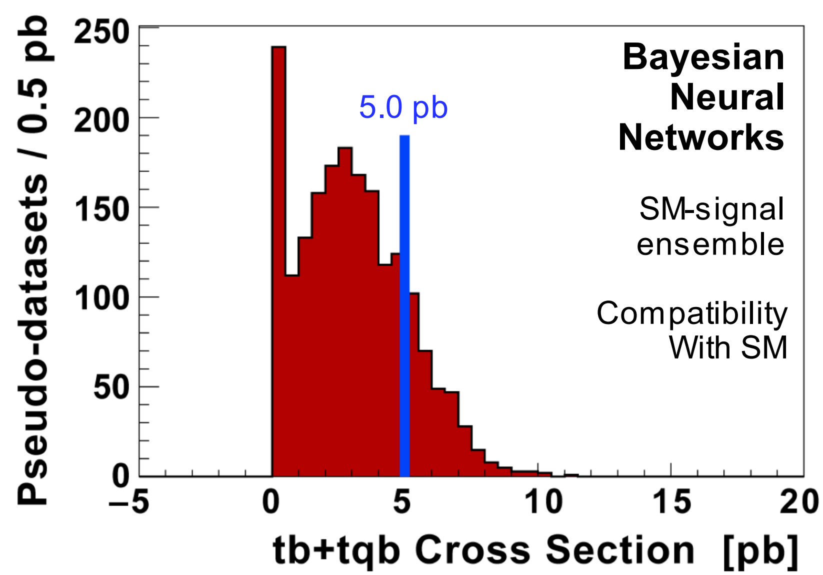
$\sigma(tb+tqb)$ = 5.0 ± 1.9 pb

Measured p-value = 0.89 %

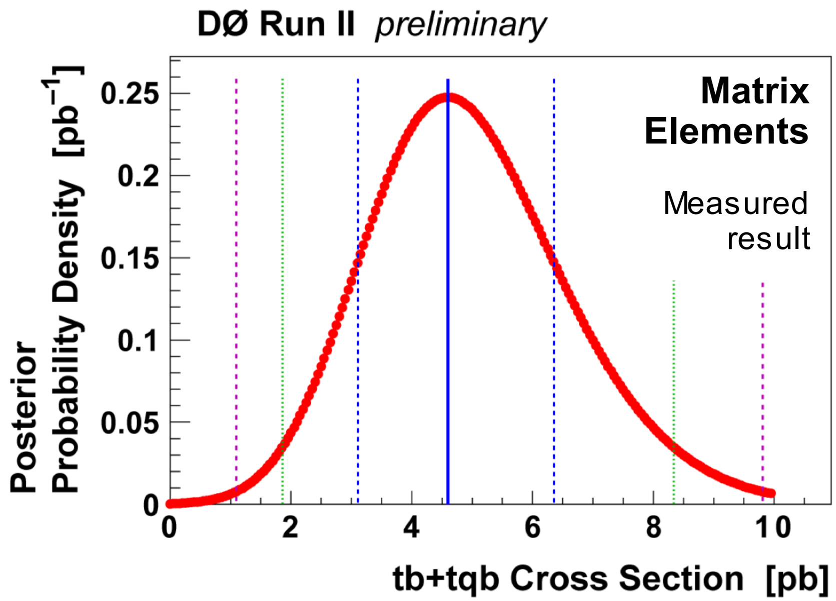
Measured significance = 2.4 σ

Compatibility with SM = 18%





Matrix Element Results

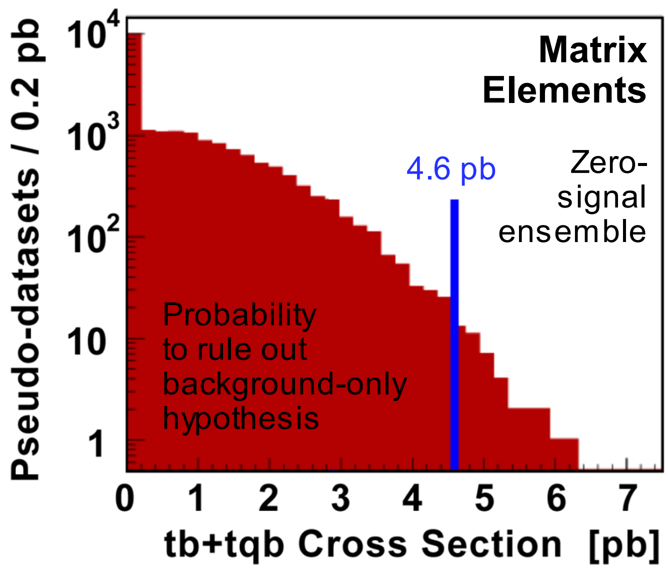


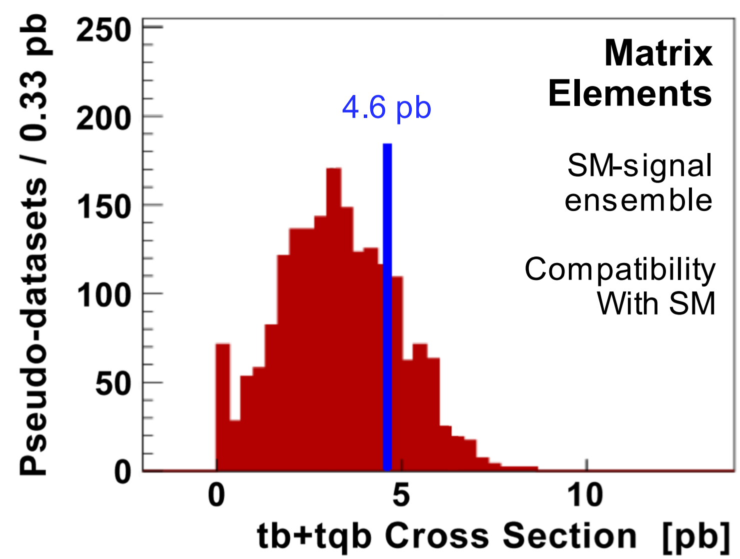
$\sigma(tb+tqb) = 4.6\ ^{+1.8}_{-1.5}$ pb

Measured p-value = 0.21 %

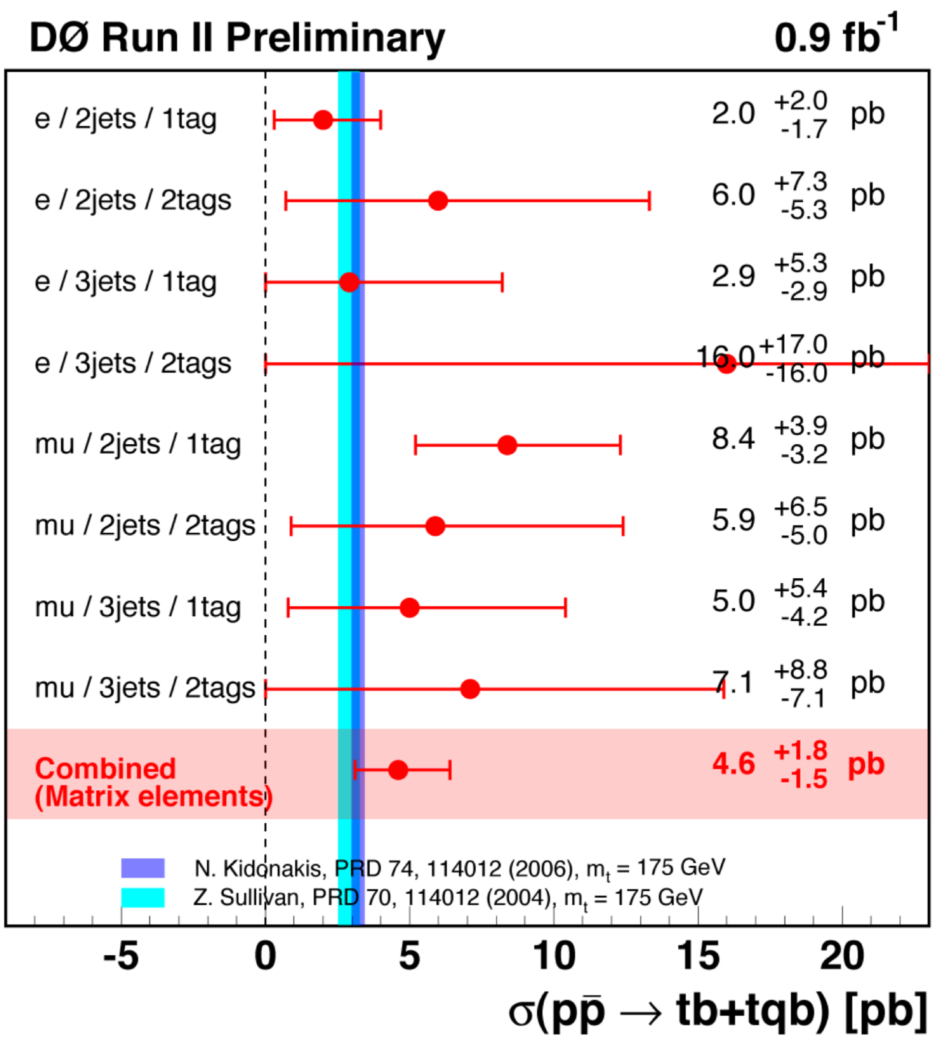
Measured significance = 2.9 σ

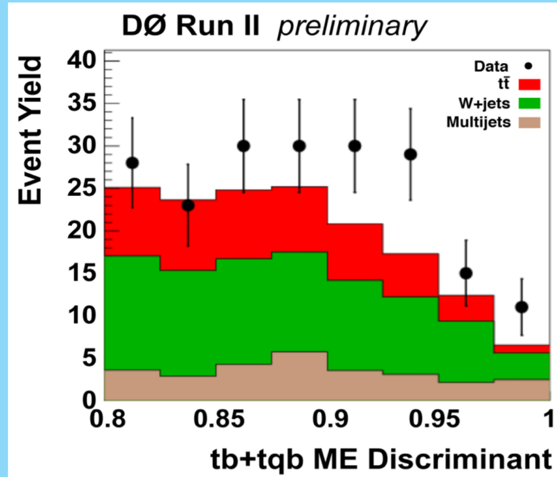
Compatibility with SM = 21%



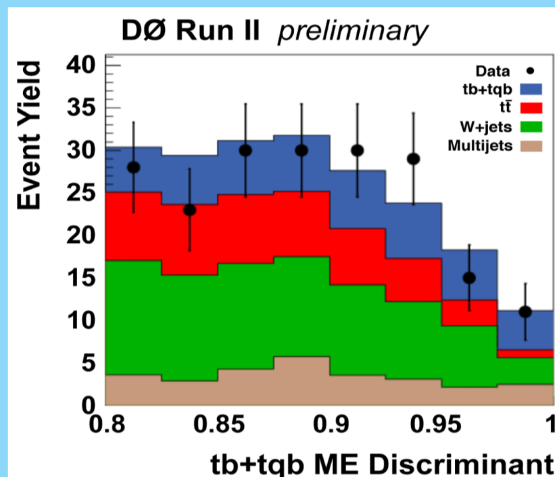


Matrix Element Results

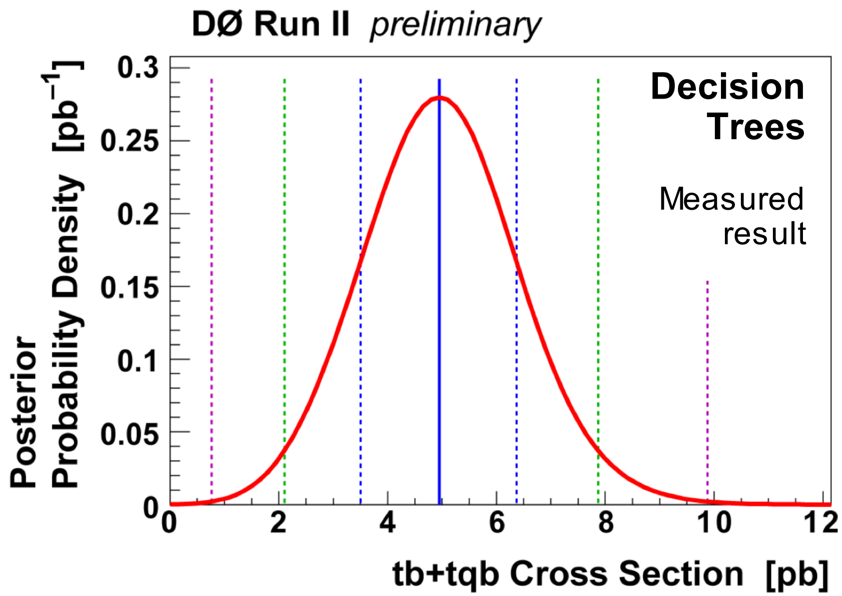




Discriminant output without and with signal component (all channels combined to "visualize" excess)



Decision Tree Results

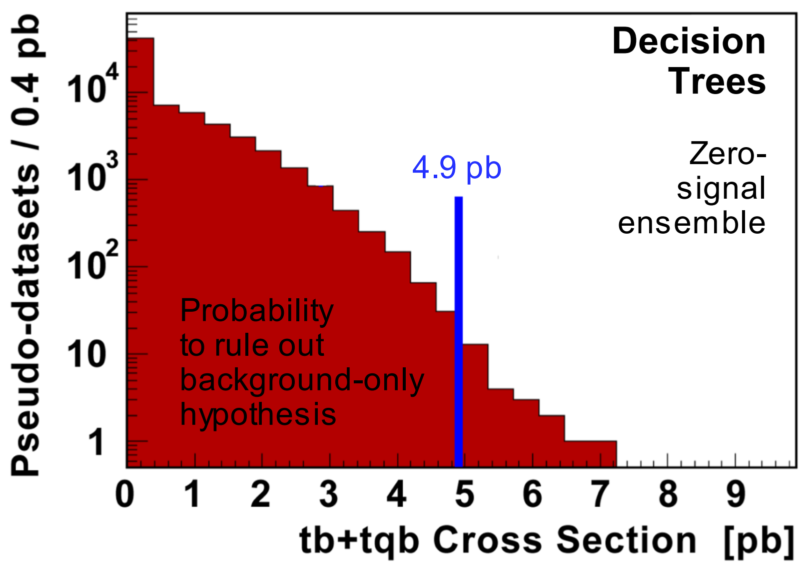


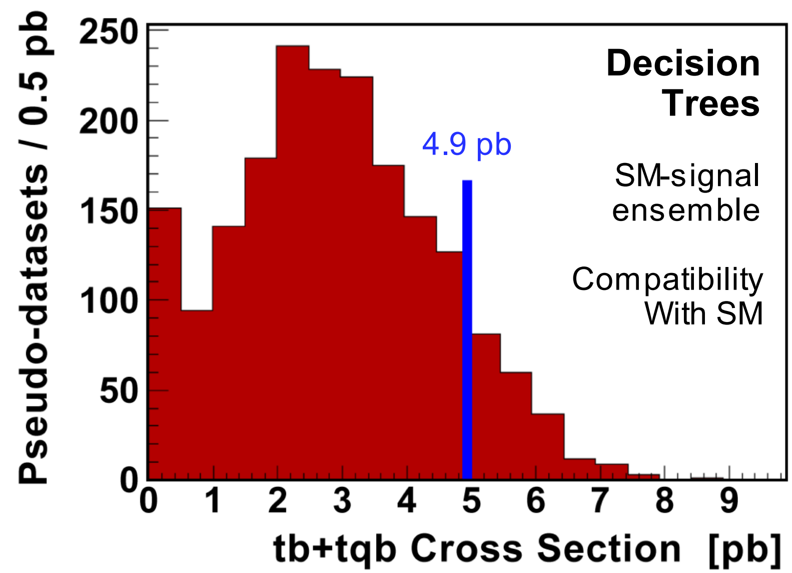
$\sigma(tb+tqb)$ = 4.9 ± 1.4 pb

Measured p-value = 0.035 %

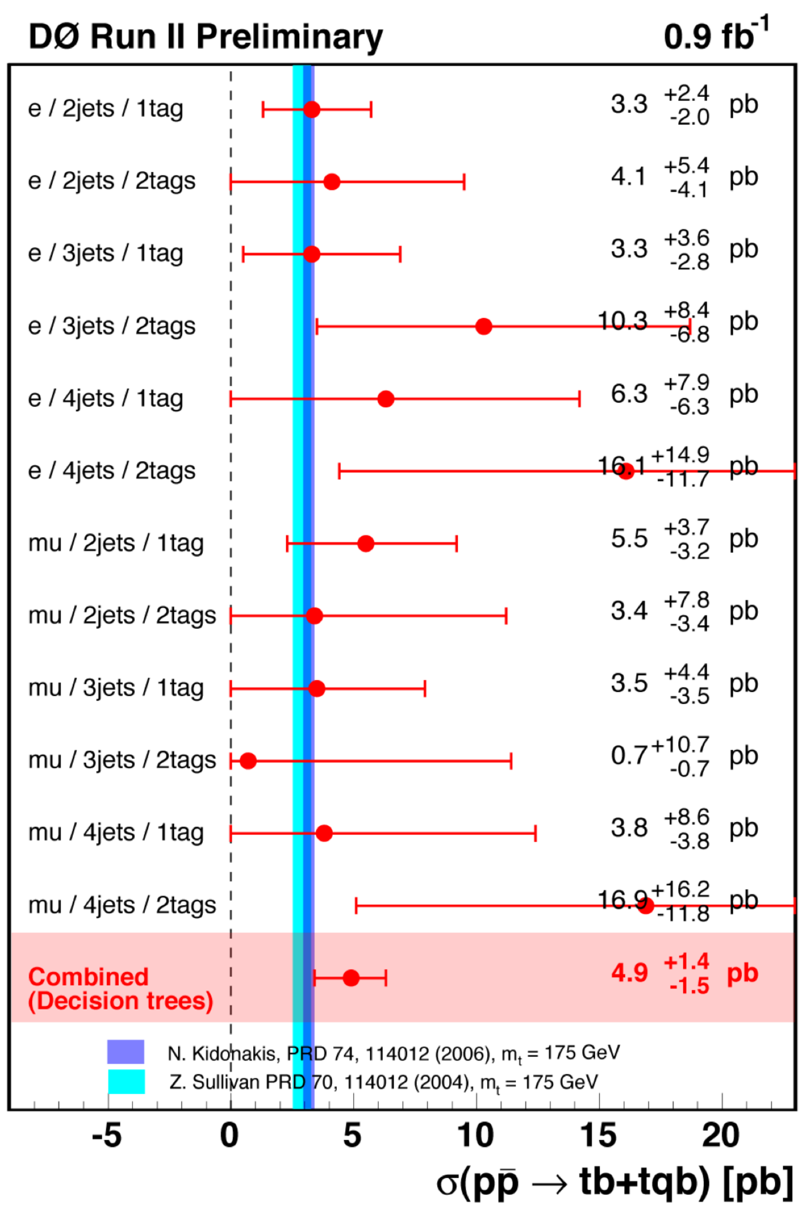
Measured significance = 3.4 σ

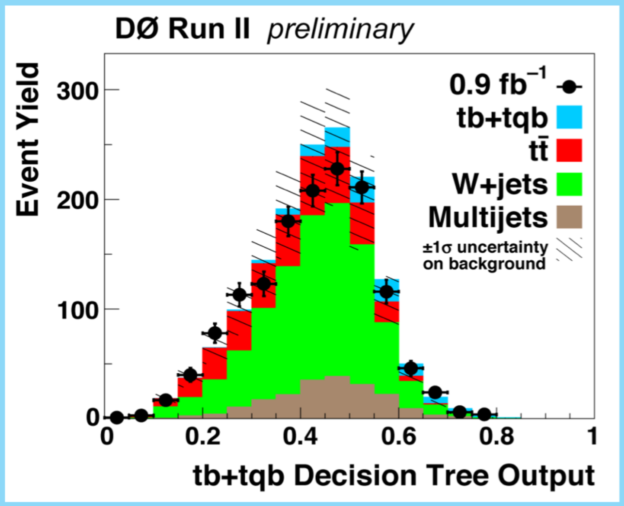
Compatibility with SM = 11%



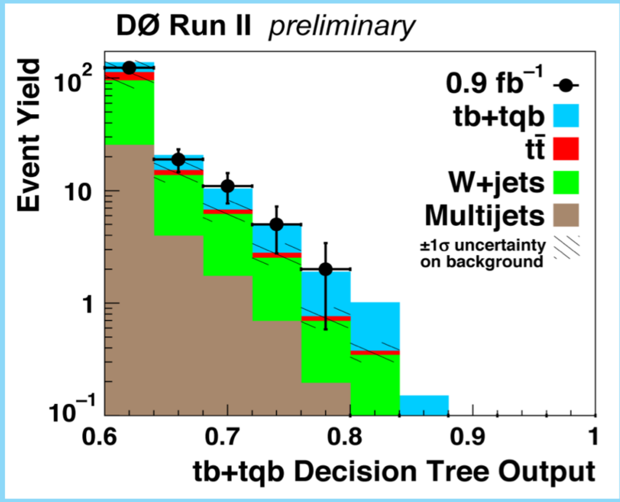


Decision Tree Results





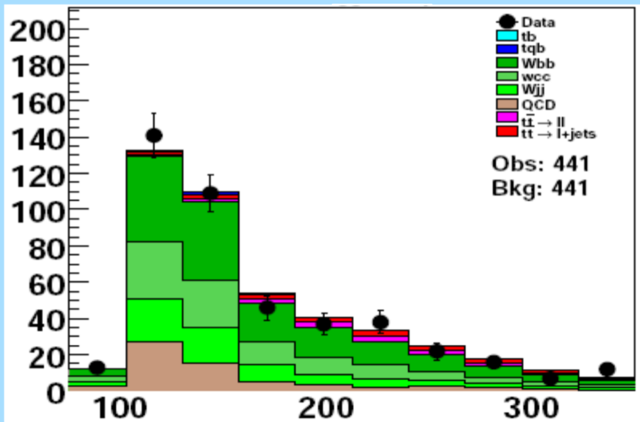
Discriminant output (all channels combined) over the full range and a close-up on the high end



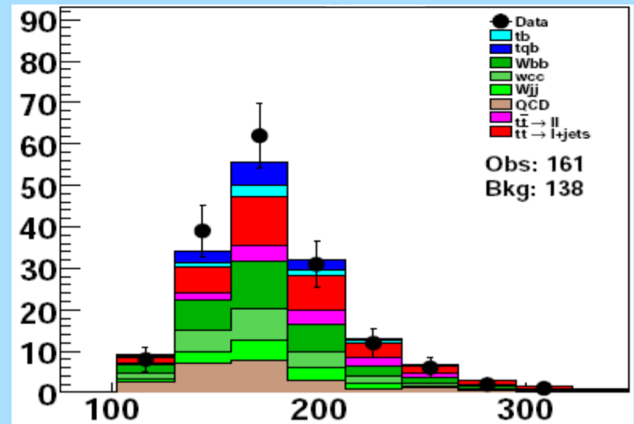
ME Event Characteristics

ME Discriminant < 0.4

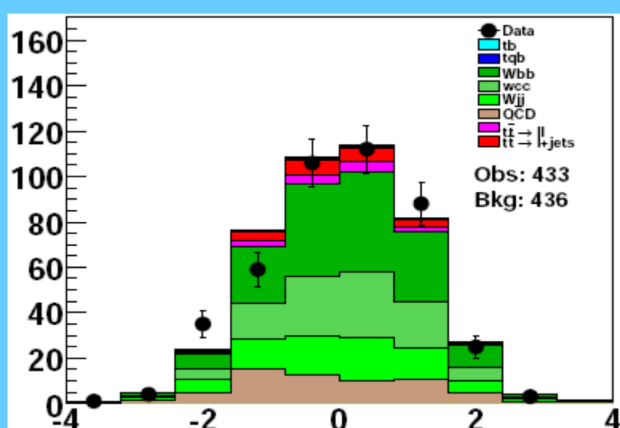
ME Discriminant > 0.7



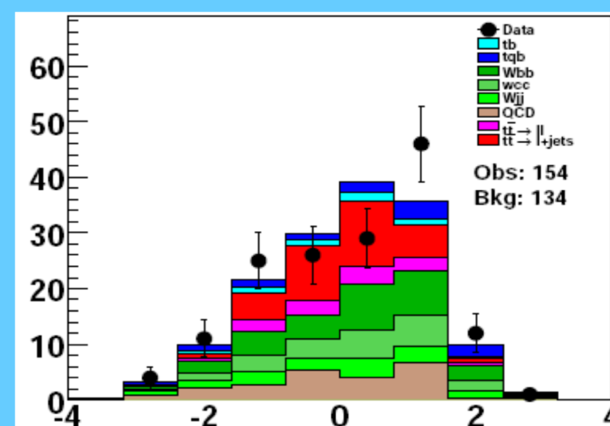
Mass (lepton,$\not{E}_T$,btagged-jet) [GeV]



Mass (lepton,$\not{E}_T$,btagged-jet) [GeV]



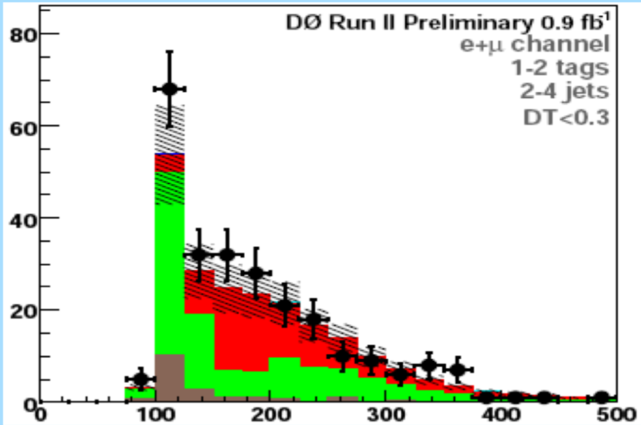
Q(lepton) x η(untagged-jet)



Q(lepton) x η(untagged-jet)

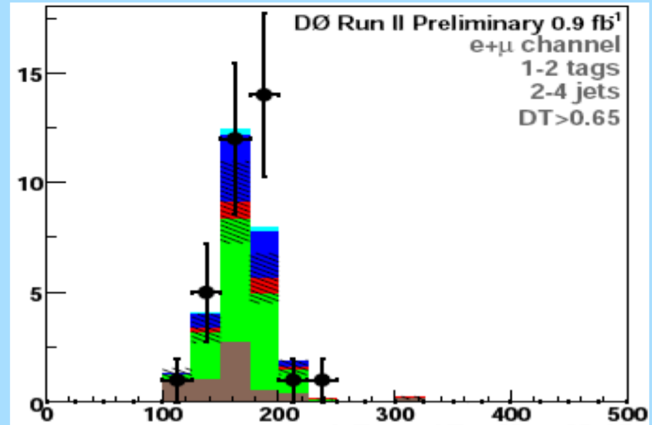
DT Event Characteristics

DT Discriminant < 0.3

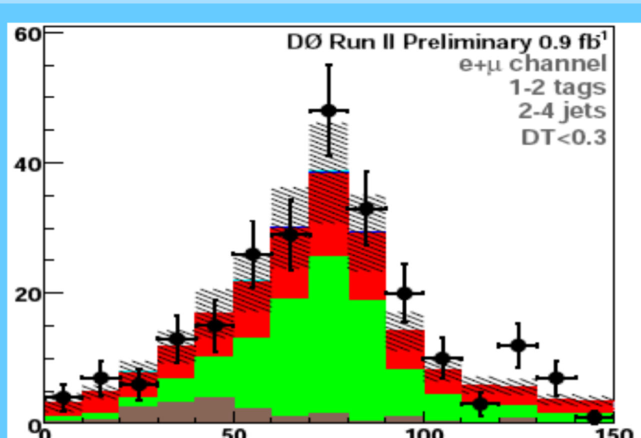


Mass (lepton,$\not{E}_T$,btagged-jet) [GeV]

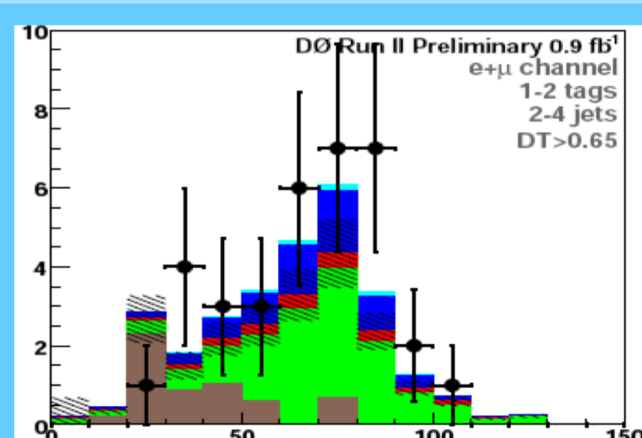
DT Discriminant > 0.65



Mass (lepton,$\not{E}_T$,btagged-jet) [GeV]

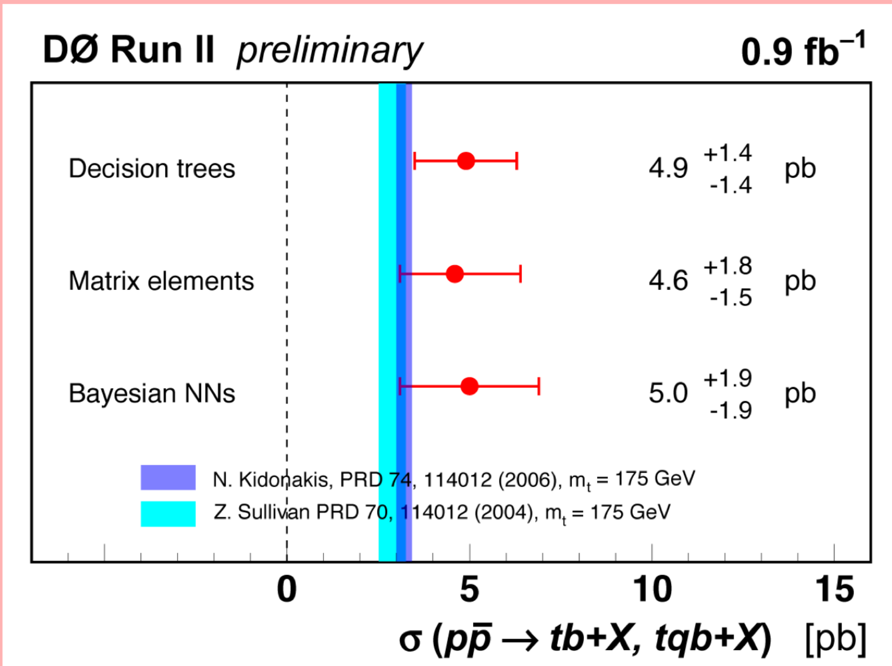


W Transverse Mass [GeV]



W Transverse Mass [GeV]

Correlation Between Methods



Results from the three methods are consistent with each other

Choose the 50 highest events in each discriminant and count overlapping events

Overlap of signal-like events

		DT	ME	BNN
Electrons	DT	100 %	52 %	56 %
	ME		100 %	46 %
	BNN			100 %
Muons	DT	100 %	58 %	48 %
	ME		100 %	52 %
	BNN			100 %

Measure cross section in 400 pseudo-datasets of SM-signal ensemble and calculate linear correlation between each pair of results

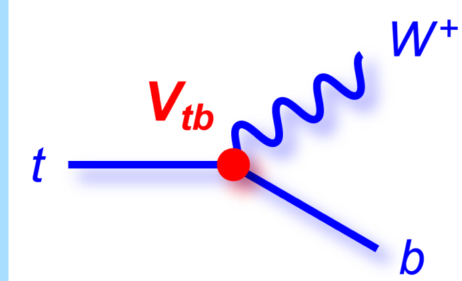
Correlation between measured cross sections

	DT	ME	BNN
DT	100 %	39 %	57 %
ME		100 %	29 %
BNN			100 %

CKM Matrix Element V_{tb}

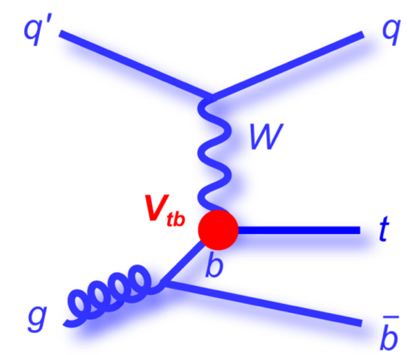
$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{\mathrm{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix}$$

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



- Weak interaction eigenstates and mass eigenstates are not the same: there is mixing between quarks, described by CKM matrix
- **In the SM, top must decay to *W* and *d, s,* or *b* quark**
  - $V_{td}^2 + V_{ts}^2 + V_{tb}^2 = 1$
  - Constraints on V_{td} and V_{ts} give $V_{tb} = 0.999100^{+0.000034}_{-0.000004}$
- **If there is new physics, then**
  - $V_{td}^2 + V_{ts}^2 + V_{tb}^2 < 1$
  - No constraint on V_{tb}
  - Interactions between top quark and gauge bosons are very interesting

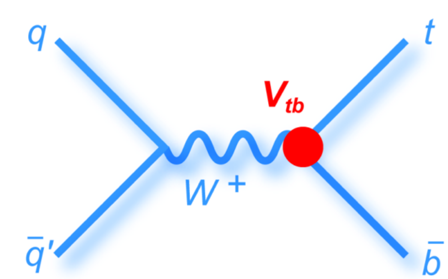
Measuring $|V_{tb}|$



- Use the measurement of the single top cross section to make the first direct measurement of $|V_{tb}|$
- Calculate a posterior in $|V_{tb}|^2$ $(\sigma(tb, tqb) \propto |V_{tb}|^2)$
- **General form of *Wtb* vertex:**

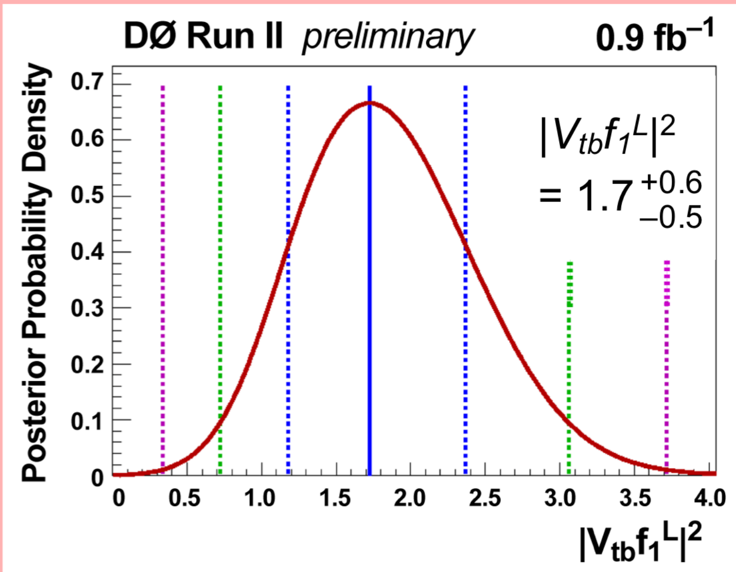
$$\Gamma^{\mu}_{Wtb} = -\frac{g}{\sqrt{2}} V_{tb} \left\{ \gamma^{\mu} \left[f_1^L P_L + f_1^R P_R \right] - \frac{i\sigma^{\mu\nu}}{M_W} (p_t - p_b)_{\nu} \left[f_2^L P_L + f_2^R P_R \right] \right\}$$

- **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

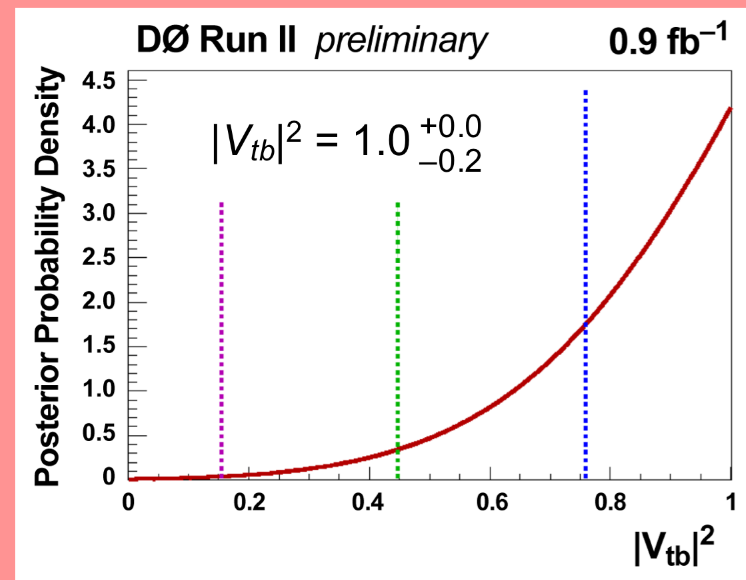


Additional theoretical uncertainties		
	tb	*tqb*
Top mass	13 %	8.5 %
Scale	5.4 %	4.0 %
PDF	4.3 %	10 %
α_s	1.4 %	0.01 %

First Direct Measurement of $|V_{tb}|$



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



$0.68 < |V_{tb}| \leq 1$ at 95% C.L.
(assuming $f_1^L = 1$)

Summary: Evidence for Single Top Quark Production at DØ

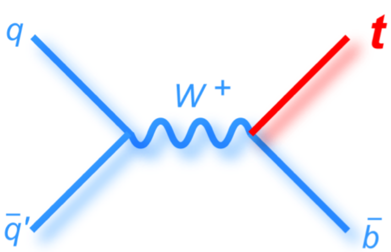
- Challenging measurement – small signal hidden in huge complex background
  Much time spent on tool development (*b*-tagging) and background modeling
- Three multivariate techniques applied to separate signal from background
- Boosted decision trees give result with 3.4 σ significance

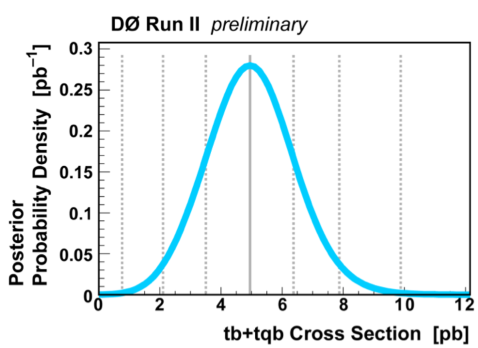
$$\sigma\left(p\bar{p} \rightarrow tb + X, tqb + X\right) \ = \ 4.9 \pm 1.4 \text{ pb}$$

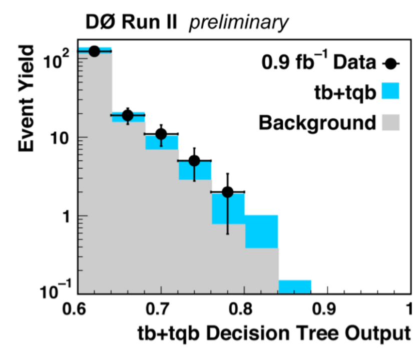
- First direct measurement of $|V_{tb}|$

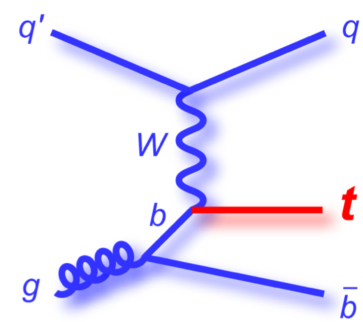
$$0.68 < |V_{tb}| \leq 1 \quad \text{at 95\% C.L.}$$

- Result submitted to *Physical Review Letters*
- Door is now open for studies of *Wtb* coupling and searches for new physics



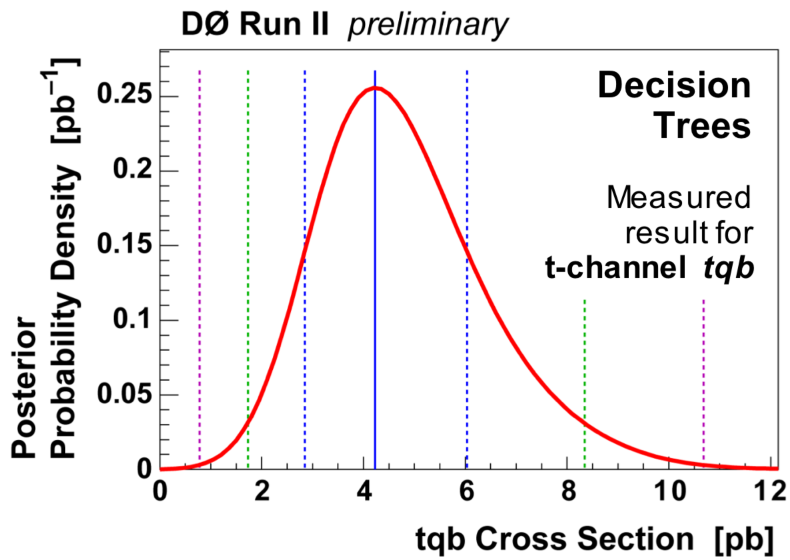


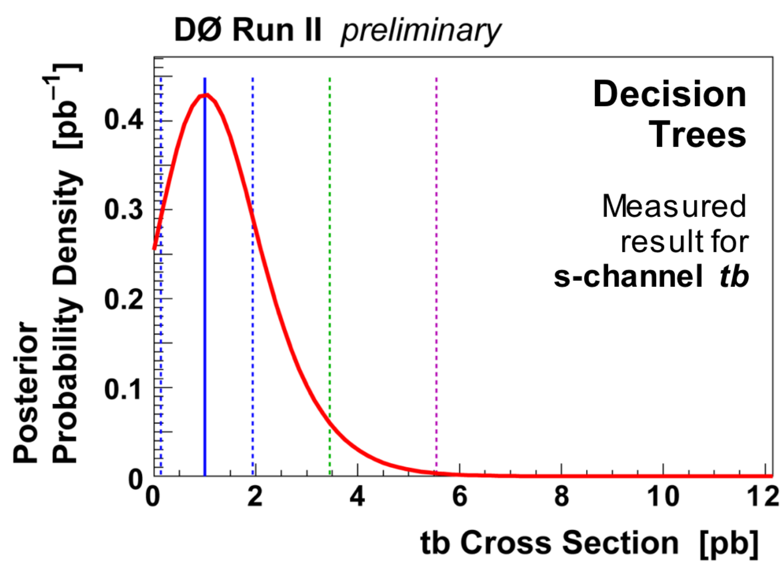




Additional Material

Results for *tb* and *tqb* Separately





$\sigma(tqb) = 4.2\ ^{+1.8}_{-1.4}$ pb

$\sigma(tb) = 1.0 \pm 0.9$ pb