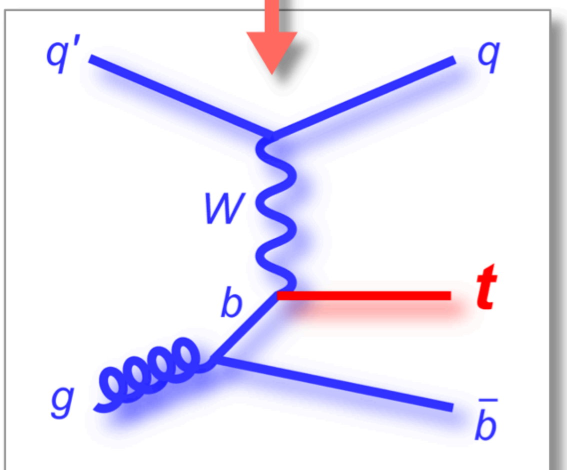
Observation of Single Top Quark Production at DØ, and Beyond

Ann Heinson
University of California, Riverside
for the DØ Collaboration

3rd International Workshop on Top Quark Physics
Bruges, Belgium
Tuesday, June 1, 2010

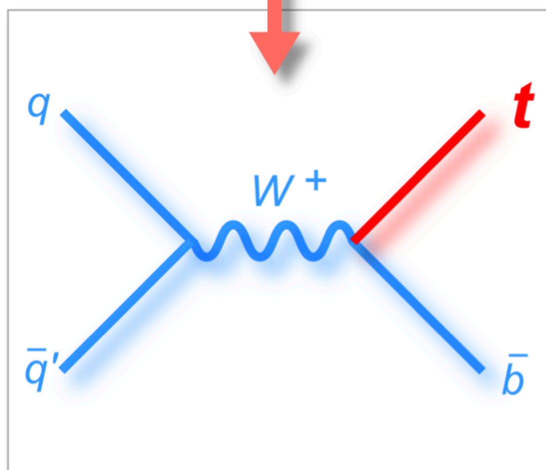
Single Top Quark Production

- **Three** electroweak modes for producing single top quarks
- **Two** have high enough rates to be observed



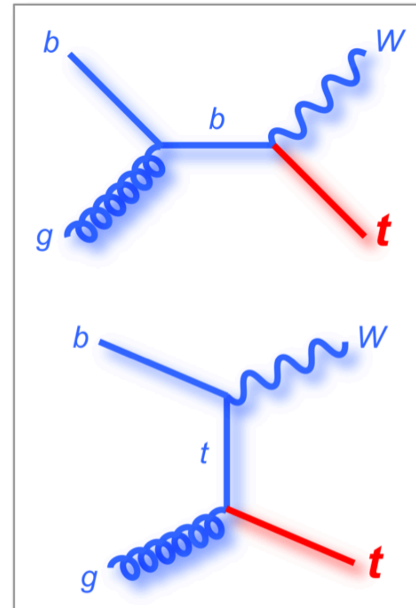
t-channel: "***tqb***"

σ = **2.34** ± 0.13 pb



s-channel: "***tb***"

σ = **1.12** ± 0.05 pb



"***tW*** production"

σ = **0.30** ± 0.06 pb

(Too small to see at the Tevatron)

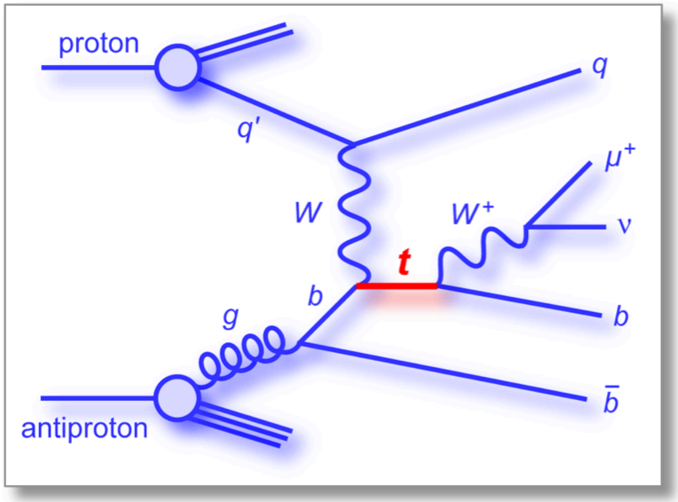
Top and antitop rates are the same as each other

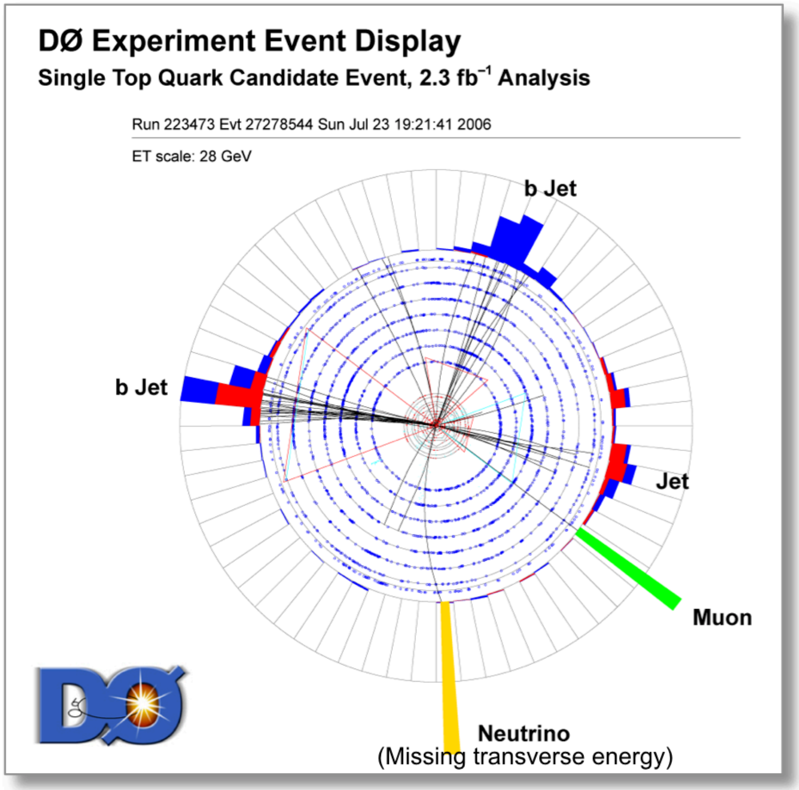
"tqb" = $tq\bar{b} + \bar{t}\bar{q}b$, *"tb"* = $t\bar{b} + \bar{t}b$

M_{top} = 170 GeV, σ = (N)NNLO
N. Kidonakis, PRD **74**, 114012 (2006)

Why Didn't We See It Till Now?

- Predicted ~10 years before the discovery of the top quark in pair production
  - t-channel: Willenbrock and Dicus, PRD 34, 155 (1986)
  - s-channel: Cortese and Petronzio, PLB 253, 494 (1991)
- Observed 14 years after the top quark discovery (CDF and DØ, 1995)
- Single top (*tb*+*tqb*) has nearly half the $t\bar{t}$ cross section but S:B is 1:20 after selection compared with 5:1 for $t\bar{t}$ – backgrounds to single top are very difficult to deal with

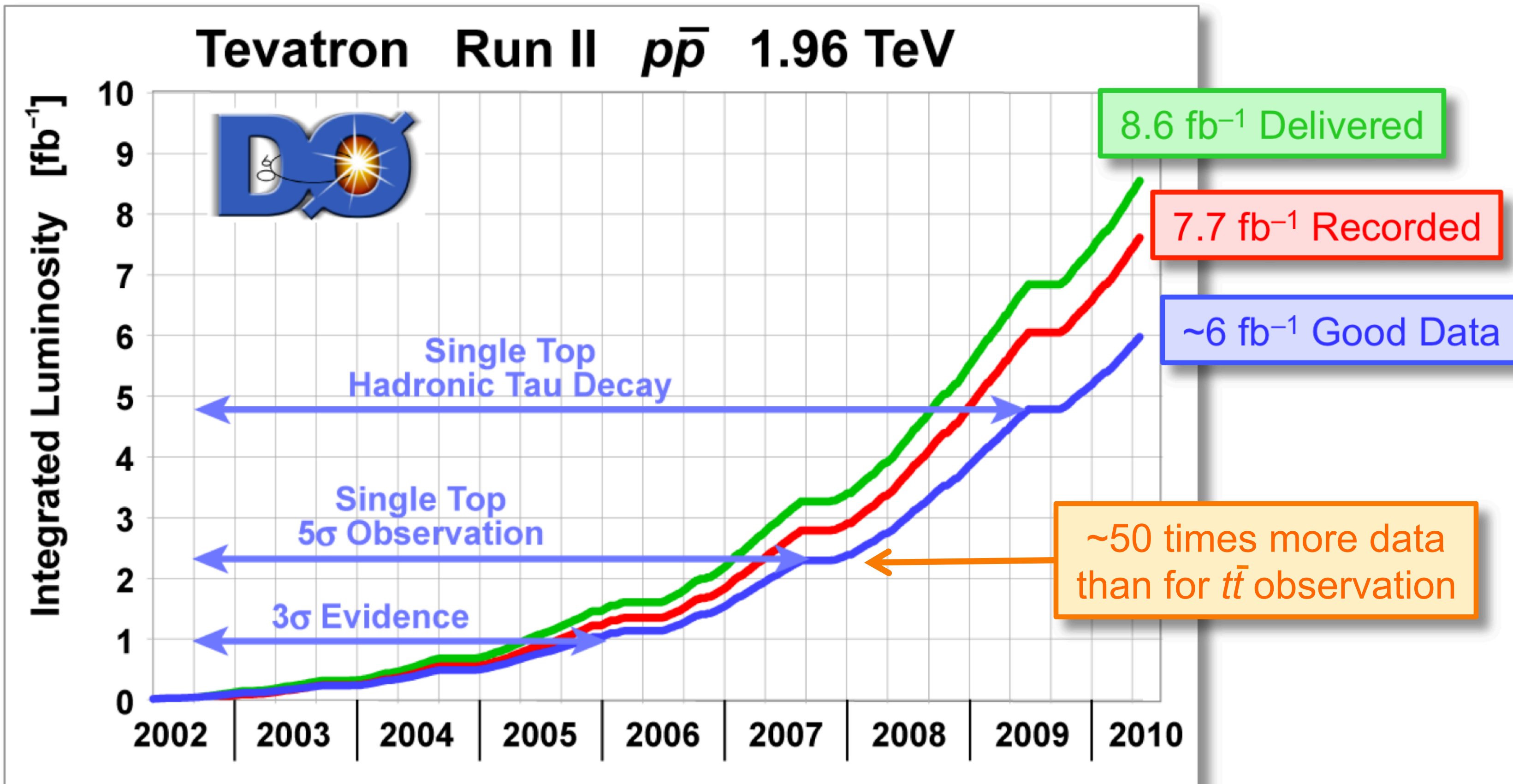




- Signal is one isolated high-p_T electron or muon and missing transverse energy from the *W* decay, 2, 3, or 4 jets, and 1 or 2 *b*-tags

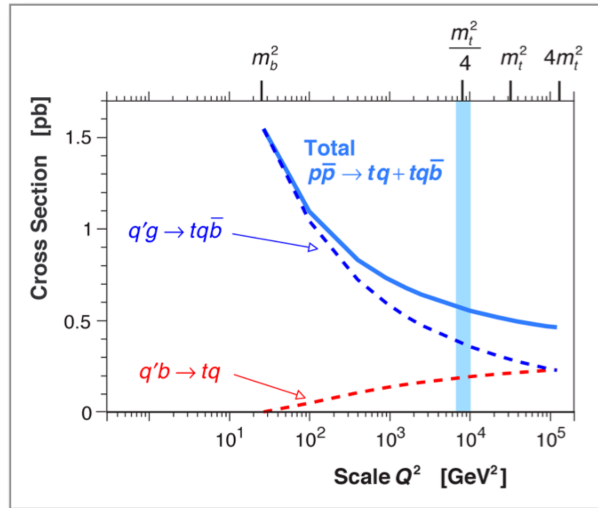
Dataset

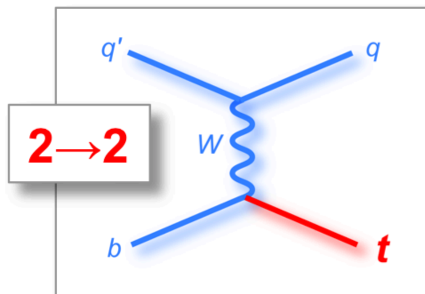
- DØ has 7.7 fb^{-1} on tape – Many thanks to Fermilab's Accelerator Division!
- The observation analysis uses 2.3 fb^{-1} of data collected from 2002 to 2007
- Select events passing any reasonable trigger
- Skim events containing an electron or muon ⇒ 1.2 billion events

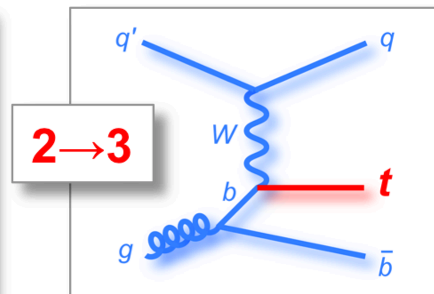


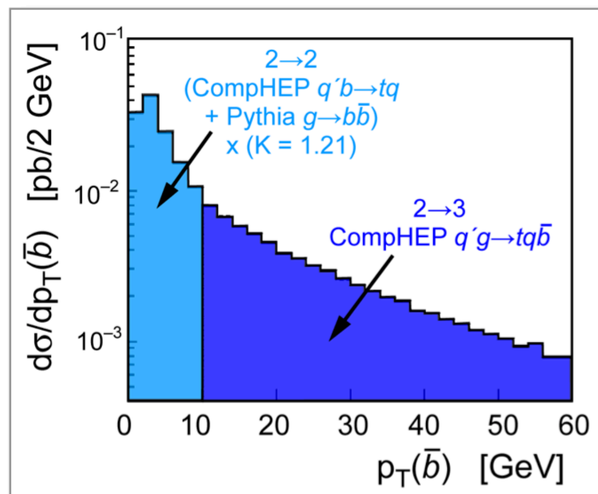
Signal Model

- **Single top quark signals modeled using SINGLETOP**
  - By Moscow State University theorists, based on COMPHEP
  - Reproduces NLO kinematic distributions (ZTOP by Sullivan)
  - M_{top} = 170 GeV
  - s-channel scale = M_{top}^2
  - t-channel scale = $\left(M_{top}/2\right)^2$
  - Decay top and W in SINGLETOP to preserve spin information
- CTEQ6M parton distributions
- PYTHIA for parton hadronization
- EVTGEN for b decays
- TAUOLA for τ decays
- Zero-bias data events overlaid to model multiple interactions (Poisson distribution, means = 2, 5)
- **2.7 million MC events**



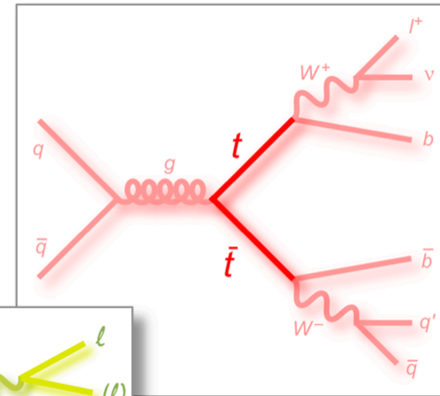


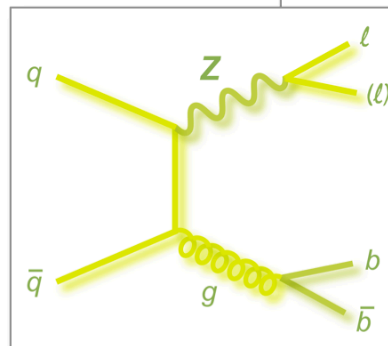


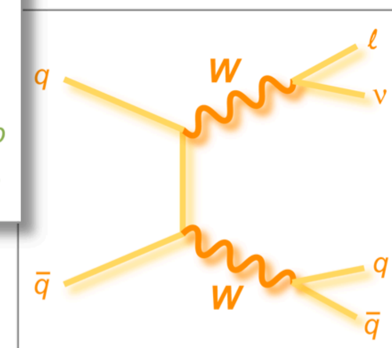


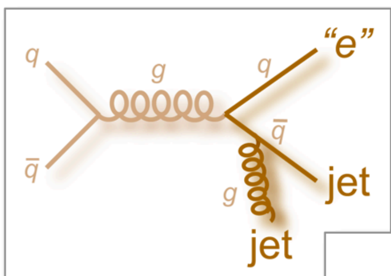
Background Models

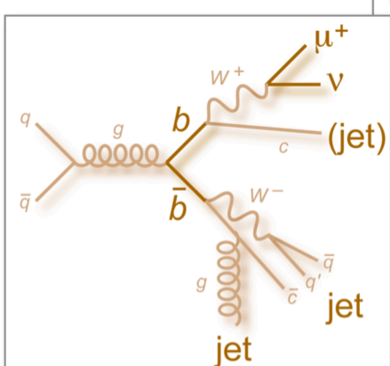
- ***tt*** **pairs, *Z*+jets, and diboson backgrounds modeled using ALPGEN**
  - PYTHIA for parton hadronization (m^2-ordered showering model)
  - Parton-jet matching algorithm used to avoid double-counting final states
  - **27 million MC events**

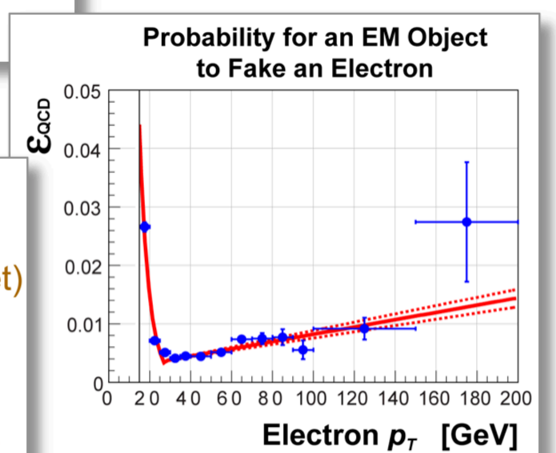








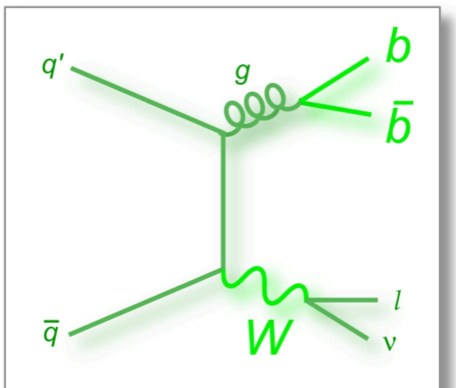


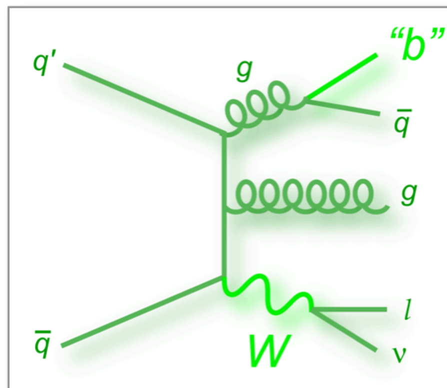


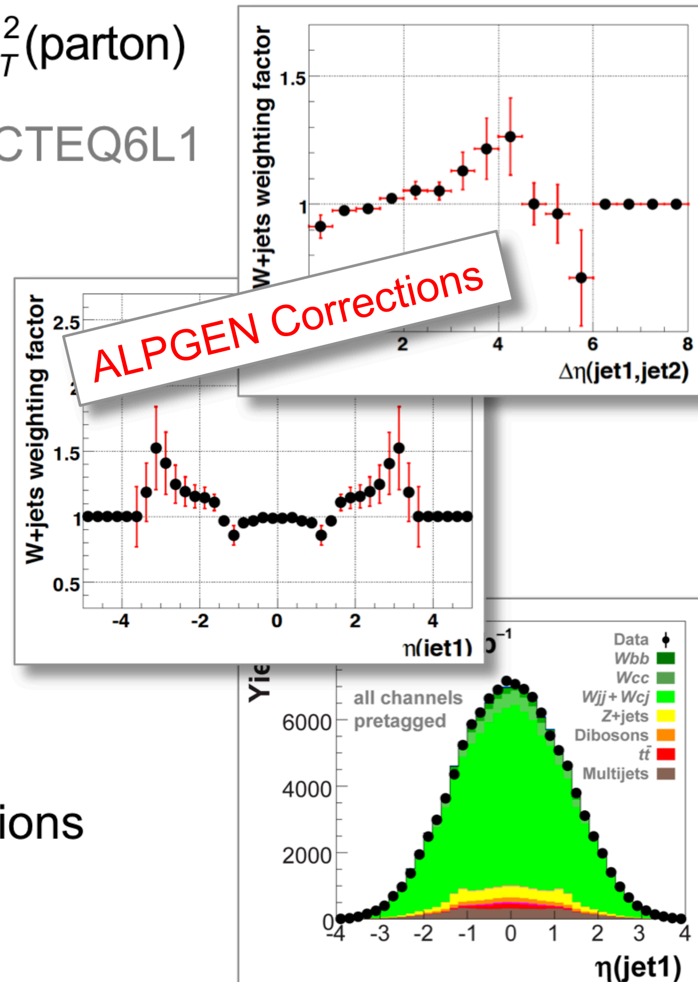
- **Multijet backgrounds modeled using data with a non-isolated lepton and jets**
  - Measure fake probabilities from low-$\not{E}_T$ data
  - Reshape low-p_T EM data to correct for e-γ mix
  - **800k events after selection**

More on the *W*+Jets Background

- **Dominant *W*+jets background modeled using ALPGEN**
  - PYTHIA for parton hadronization (m^2-ordered showering model)
  - Parton-jet matching algorithm used to avoid double-counting final states
  - Scale is $Q^2 = m^2(W) + \sum_{\text{partons}} m^2(\text{parton}) + p_T^2(\text{parton})$
  - Jets in *Wbb* and *Wcc* have mass, PDF is CTEQ6L1
  - **49 million MC events**







- The ALPGEN model of W+jets events have weight functions applied to make the distributions wider for: η(jet1), η(jet2), Δφ(jet1,jet2), Δη(jet1,jet2), η(jet3), η(jet4)

Background Normalization

- Small MC backgrounds ($t\bar{t}$, Z+jets, dibosons) use (N)NLO theory values
  - +8%–13% uncertainty for $t\bar{t}$ includes component for top mass
- Heavy flavor fraction of W+jets is leading-log in ALPGEN, needs adjusting to get NLO rates (similarly for Z+jets):

$$K'_{\mathrm{HF}} = \frac{\sigma^{\mathrm{HF}}_{\mathrm{NLO}}}{\sigma_{\mathrm{NLO}}} = 1.47 \text{ for } Wc\bar{c} \text{ and } Wb\bar{b} \text{ (from MCFM)}$$

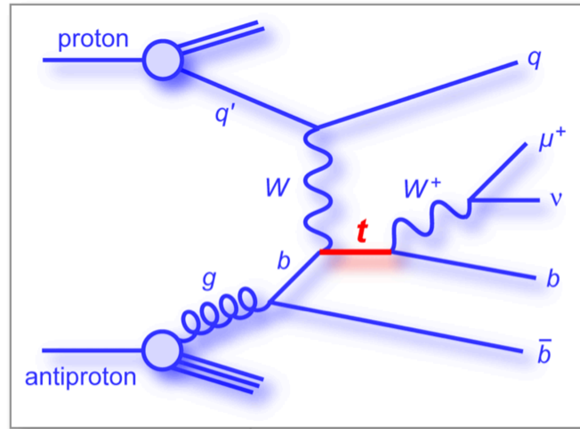
$$= 1.38 \text{ for } Wcj \text{ (from data)}$$

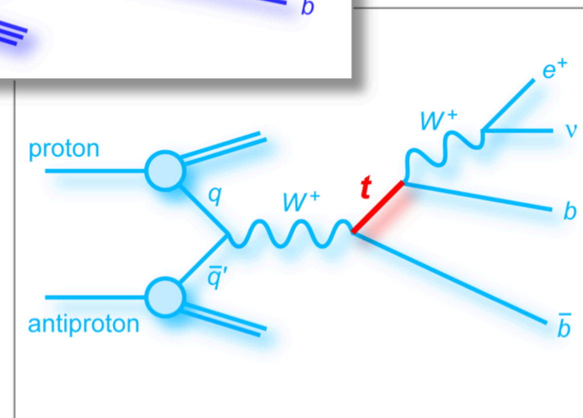
$$S_{\mathrm{HF}} = 0.95 \pm 0.13 \text{ for } Wc\bar{c} \text{ and } Wb\bar{b} \text{ (empirical correction)}$$

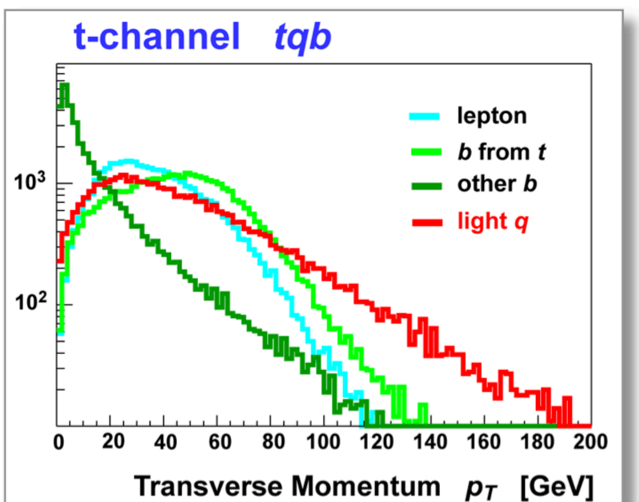
  - This is a 40% boost for Wbb and Wcc relative to Wjj, a large correction
  - The 14% uncertainty on S_{HF} is one of the largest on the final result
- W+jets MC and multijets data are normalized together to pretagged data
  - Subtract other backgrounds first
  - Iterative Kolmogorov-Smirnov technique
  - Separately in each of the 24 analysis channels
  - Normalized using three variables: p_T(lepton), $\not{E}_T$, and $M_T(W)$

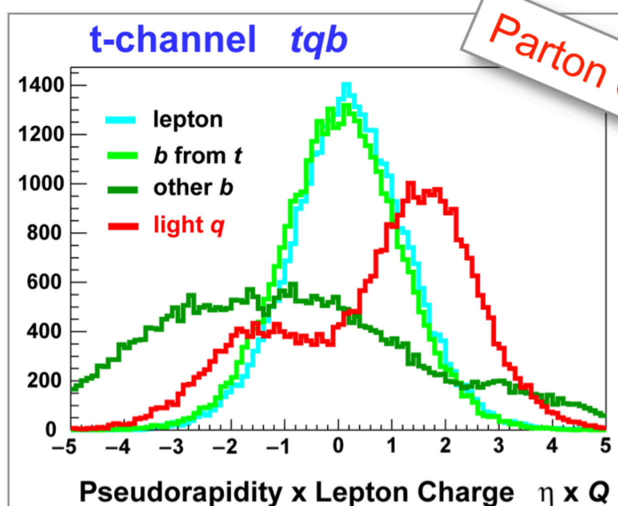
Event Selection

- **One isolated electron or muon**
  - Electron p_T > 15 GeV, |η| < 1.1
  - Muon p_T > 15 GeV, |η| < 2.0
- **Missing transverse energy**
  - $\not{E}_T$ > 20 GeV
- **One b-tagged jet and at least one more jet**
  - 2–4 jets with p_T > 15 GeV, |η| < 3.4
  - Leading jet p_T > 25 GeV
- **Split analysis into 24 independent channels:**
  - 2,3,4 jets; e,μ; 1,2 *b*-tags; Run IIa, Run IIb





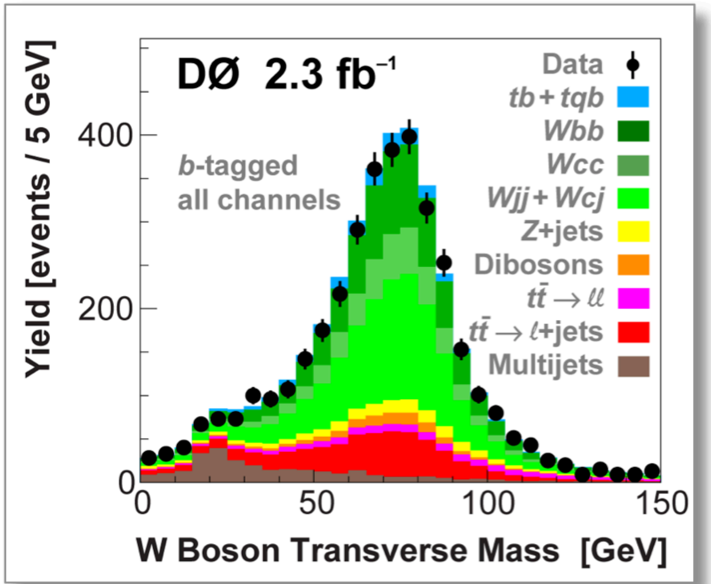




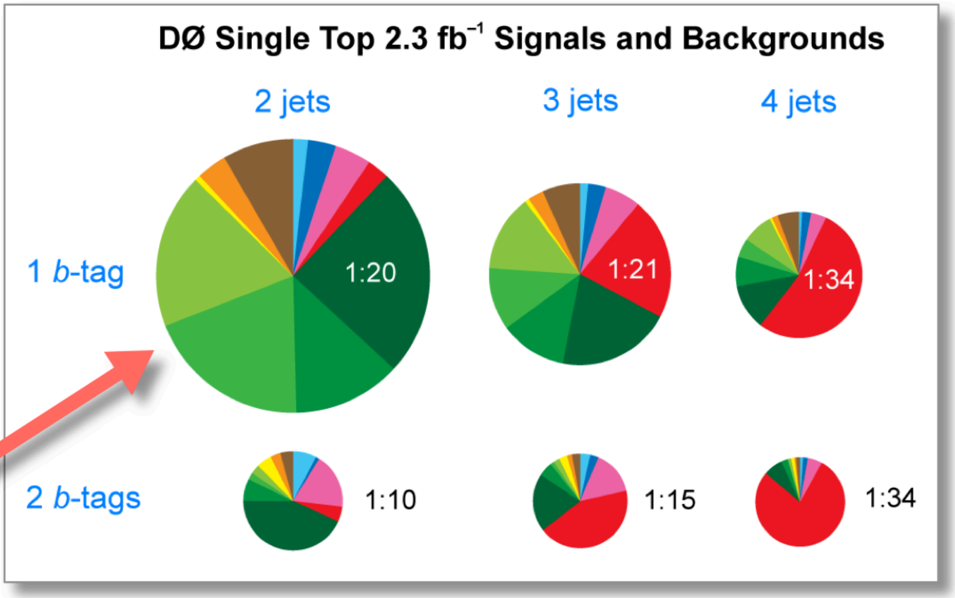
Parton distributions

Event Yields after *b*-Tagging

Event Yields in 2.3 fb^{-1} of DØ Data	
e,μ, 2,3,4-jets, 1,2-tags combined	
tb + *tqb*	223 ± 30
W+jets	2,647 ± 241
Z+jets, dibosons	340 ± 61
$t\bar{t}$ pairs	1,142 ± 168
Multijets	300 ± 52
Total prediction	4,652 ± 352
Data	4,519

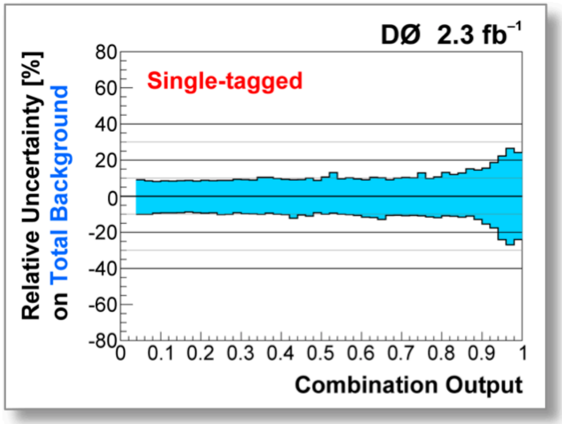


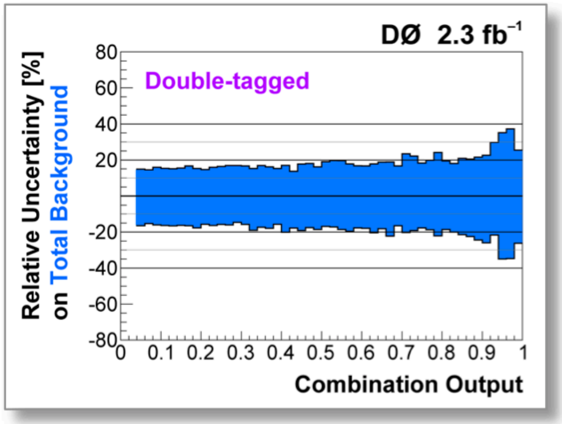
- **Signal acceptances:**
  - *tb* = (3.7 ± 0.5)%
  - *tqb* = (2.5 ± 0.3)%
- **Signal:Background**:
  - 1:10 to 1:34
- **Highest acceptance channel**:
  - 2-jets/1tag



Systematic Uncertainties

- Total error on *tb*+*tqb* cross section is ±22%
- Statistics-only error is ±18%
- So systematics contribute ±13%





Systematic Uncertainties
Ranked from Largest to Smallest Effect on Single Top Cross Section
DØ 2.3 fb^{-1}

Larger terms	
b-ID tag-rate functions (includes shape variations)	(2.1–7.0)% (1-tag) (9.0–11.4)% (2-tags)
Jet energy scale (includes shape variations)	(1.1–13.1)% (signal) (0.1–2.1)% (bkgd)
W+jets heavy-flavor correction	13.7%
Integrated luminosity	6.1%
Jet energy resolution	4.0%
Initial- and final-state radiation	(0.6–12.6)%
b-jet fragmentation	2.0%
$t\bar{t}$ pairs theory cross section	12.7%
Lepton identification	2.5%
Wbb/*Wcc* correction ratio	5%
Primary vertex selection	1.4%

Systematic Uncertainties
Ranked from Largest to Smallest Effect on Single Top Cross Section
DØ 2.3 fb^{-1}

Smaller terms	
Monte Carlo statistics	(0.5–16.0)%
Jet fragmentation	(0.7–4.0)%
Branching fractions	1.5%
Z+jets heavy-flavor correction	13.7%
Jet reconstruction and identification	1.0%
Instantaneous luminosity correction	1.0%
Parton distribution functions (signal)	3.0%
Z+jets theory cross sections	5.8%
W+jets and multijets normalization to data	(1.8–3.9)% (*W*+jets) (30–54)% (multijets)
Diboson theory cross sections	5.8%
Alpgen *W*+jets shape corrections	shape only
Trigger	5%

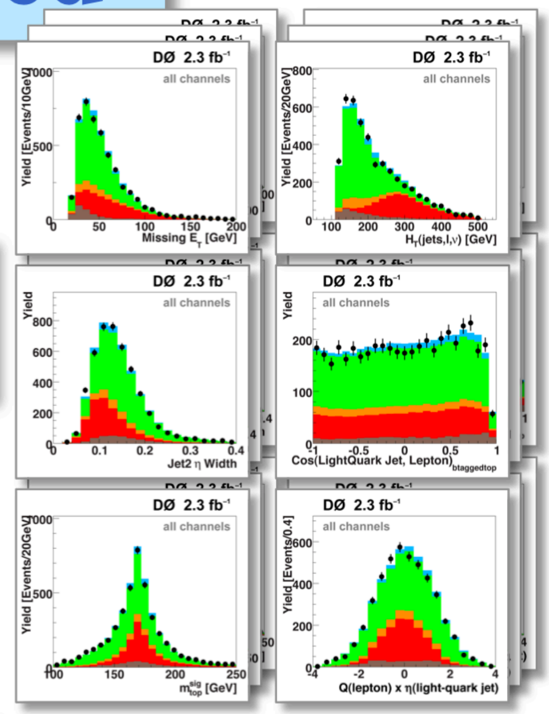
Analysis Strategy Visualized

Data from all reasonable triggers

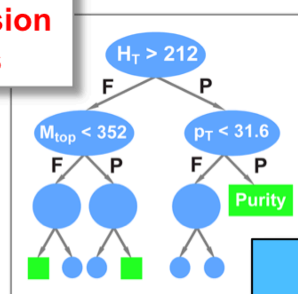
Monte Carlo for signal and backgrounds

Select signal-like events

Check data is reproduced in all variables



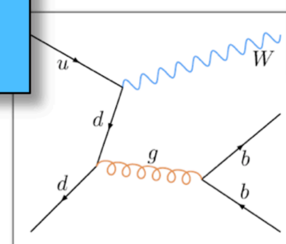
Boosted Decision Trees

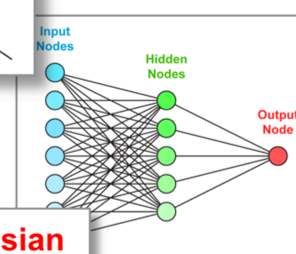


~100 million pseudo-datasets

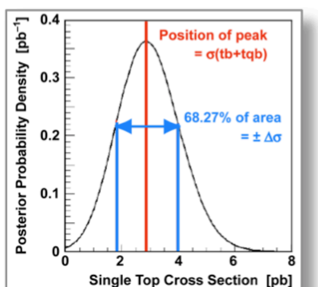
Separate signal from background

Matrix Elements





Bayesian Neural Networks



Measure significance

Cross section from Bayesian likelihood

Combine results

Sensitive Variables

- 97 variables used to separate signal from background
  - Each variable has a different distribution between *tb* or *tqb* and at least one background component and
  - Good agreement between background model and data

5 classes of variables

Best Variables to Separate Single Top from W+Jets

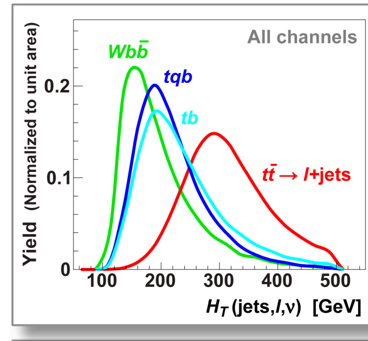
DØ 2.3 fb^{-1} Analysis

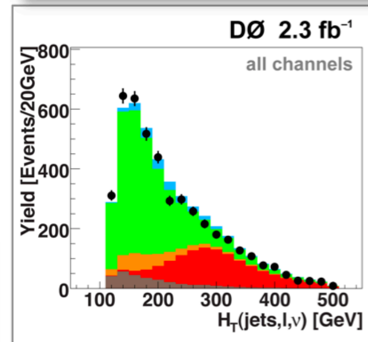
Object kinematics	$\not{E}_T$
	p_T(jet2)
	p_T^{rel}(jet1,tag-μ)
	E(light1)
Event kinematics	M(jet1,jet2)
	M_T(W)
	H_T(lepton,$\not{E}_T$,jet1,jet2)
	H_T(jet1,jet2)
	H_T(lepton,$\not{E}_T$)
Jet reconstruction	Width$_\phi$(jet2)
	Width$_\eta$(jet2)
Top quark reconstruction	M_{top}(W,tag1)
	ΔM_{top}^{min}
	M_{top}(W,tag1,S2)
Angular correlations	cos(light1,lepton)$_{btaggedtop}$
	$\Delta\phi$(lepton,$\not{E}_T$)
	Q(lepton) x η(light1)

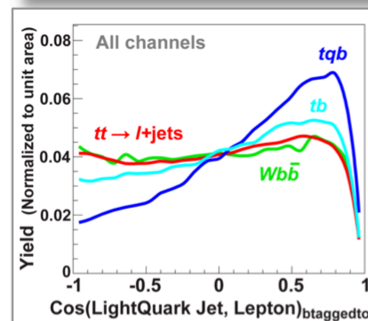
Best Variables to Separate Single Top from Top Pairs

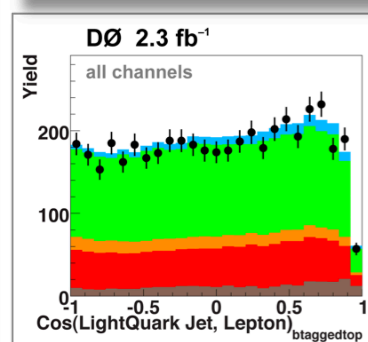
DØ 2.3 fb^{-1} Analysis

Object kinematics	pT(notbest2)
	pT(jet4)
	pT(light2)
Event kinematics	M(alljets–tag1)
	Centrality(alljets)
	M(alljets–best1)
	H_T(alljets–tag1)
	H_T(lepton,$\not{E}_T$,alljets)
	M(alljets)
Jet reconstruction	Width$_\eta$(jet4)
	Width$_\phi$(jet4)
	Width$_\phi$(jet2)
Angular correlations	cos(lepton$_{btaggedtop}$, btaggedtop$_{CMframe}$)
	Q(lepton) x η(light1)
	ΔR(jet1,jet2)





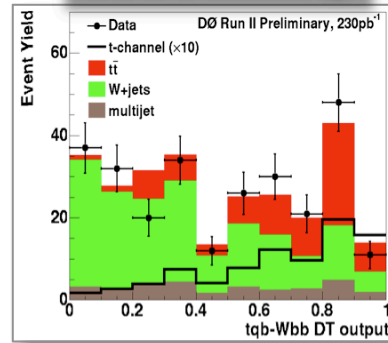




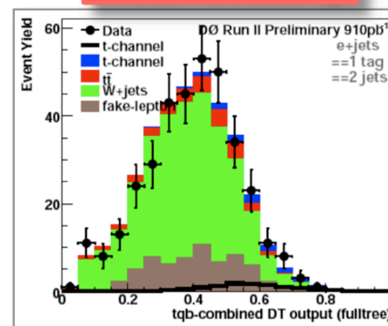
Boosted Decision Trees

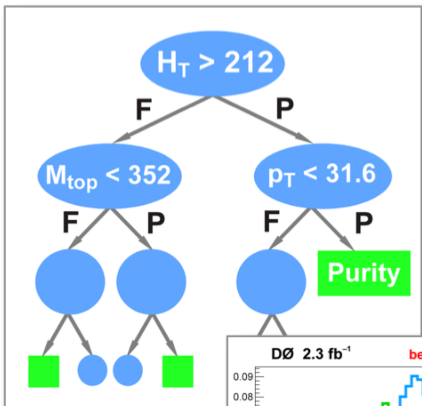
- A decision tree applies sequential cuts to events but does not reject ones that fail the cuts
- Trained on two large samples – signal, and all backgrounds combined
- Use the best 64 variables. Adding variables does not degrade performance – not useful ones are ignored
- Boosting averages the results over many trees, improves performance by 20%
- Monotonically transform the output so every bin has at least 40 background events
- One set of trees in each of the 24 analysis channels

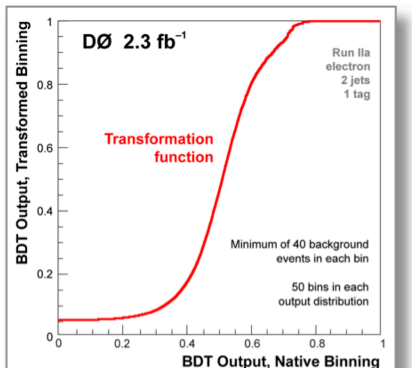
Before boosting

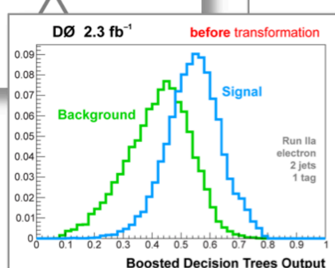


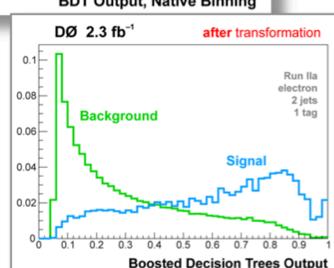
After boosting



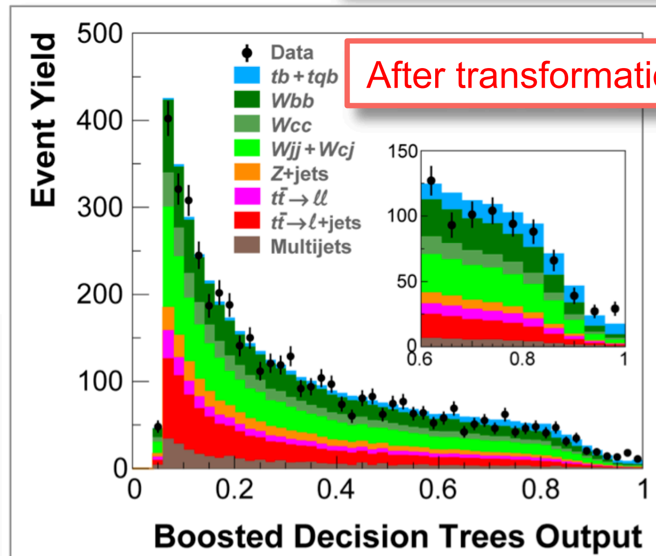






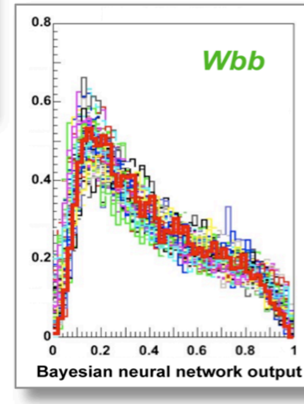


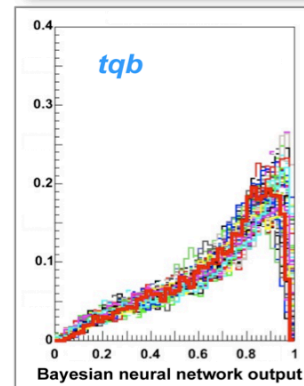
After transformation

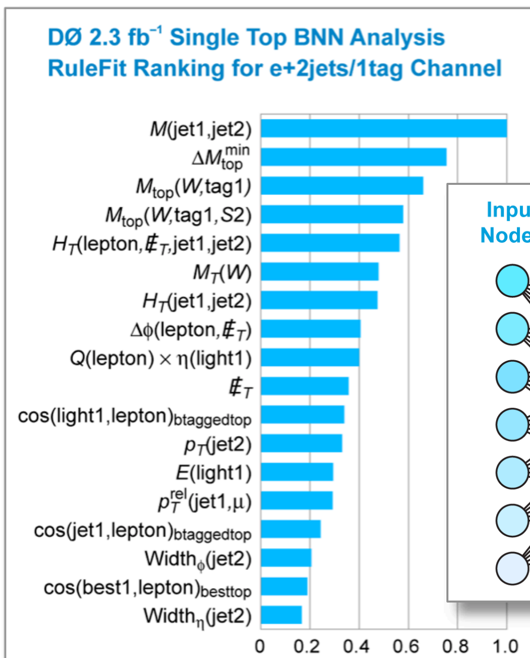


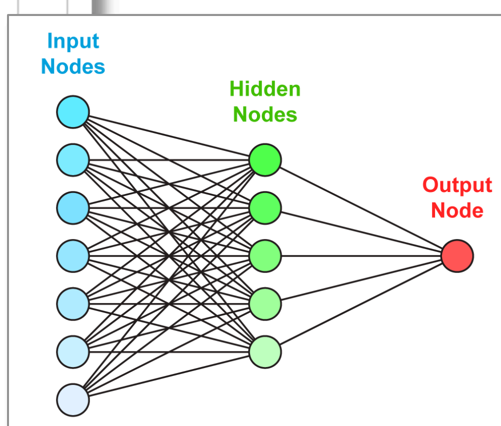
Bayesian Neural Networks

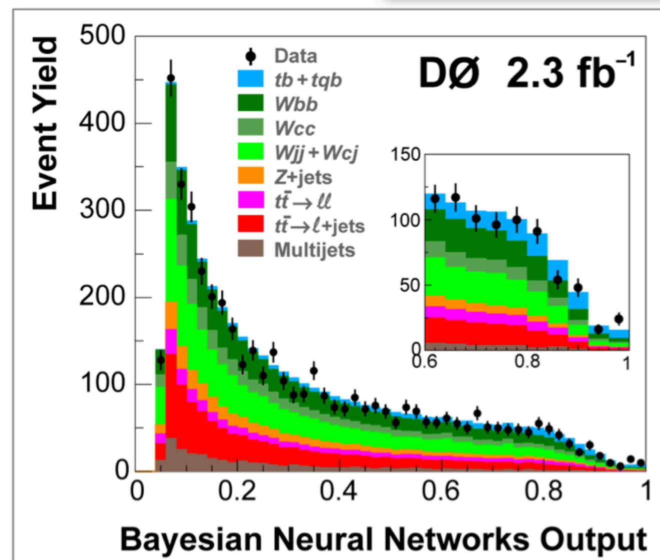
- A neural network performs a combination of input variables with weights between nodes and thresholds at nodes
- Trained on two large samples – signal, and all backgrounds combined
- Best 18-28 variables chosen using RuleFit. Using too many variables degrades performance – noise is added
- "Bayesian" means averaging over many networks using Markov Chain MC sampling technique – no overtraining
- One set of Bayesian NNs in each of the 24 analysis channels





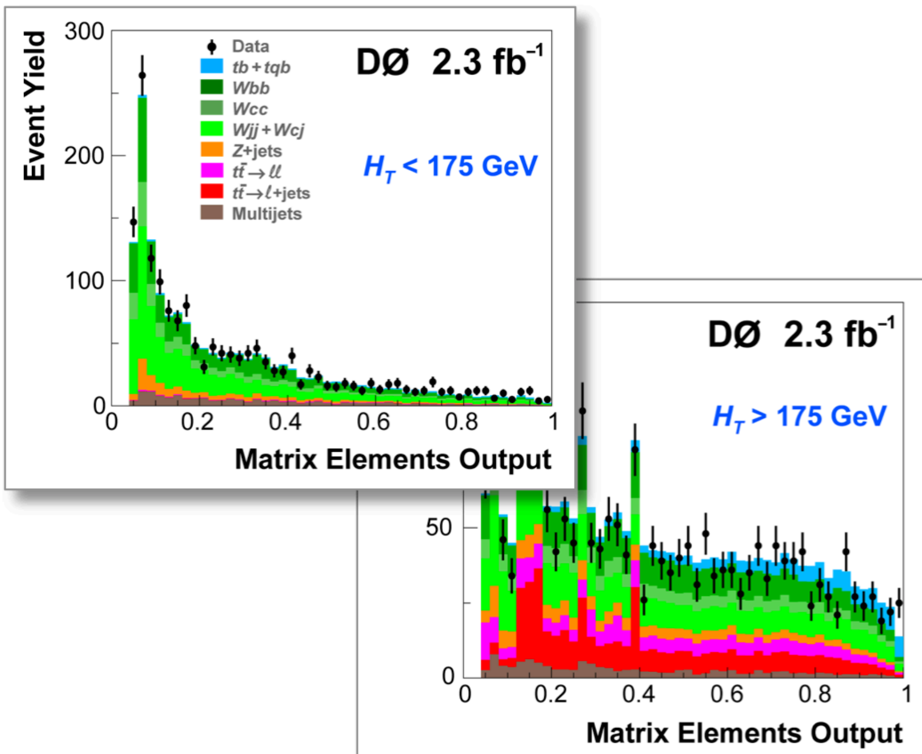






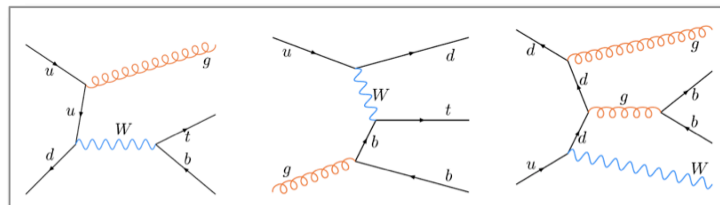
Matrix Elements

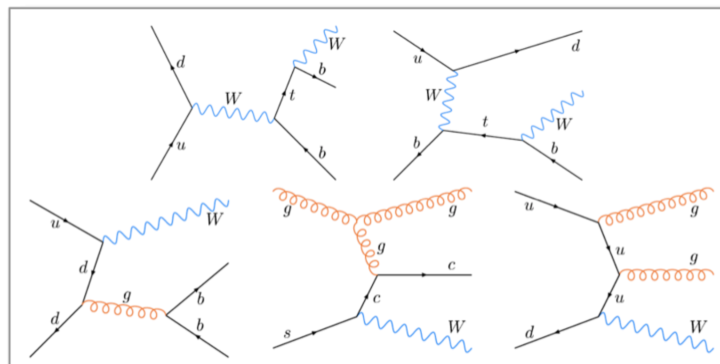
- Matrix elements corresponds to signal and background probabilities for each event
  - Every element calculated for every event (data, MC)
  - PDFs and jet resolution transfer functions used
  - Split samples with a cut on H_T to improve sensitivity



Matrix Elements used to Separate Single Top Signal from Background
DØ 2.3 fb^{-1}

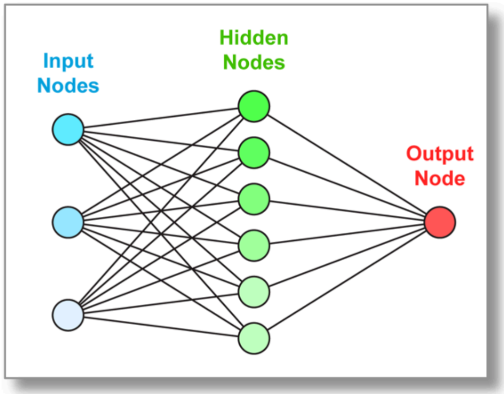
2 Jets		3 Jets	
tb	$u\bar{d}\to t\bar{b}$	tbg	$u\bar{d}\to t\bar{b}g$
tq	$ub\to td$ $\bar{d}b\to t\bar{u}$	tqg	$ub\to tdg$ $\bar{d}b\to t\bar{u}g$
		tqb	$ug\to td\bar{b}$ $\bar{d}g\to t\bar{u}\bar{b}$
$Wb\bar{b}$	$u\bar{d}\to Wb\bar{b}$	$Wb\bar{b}g$	$u\bar{d}\to Wb\bar{b}g$
Wcg	$\bar{s}g\to W\bar{c}g$		
Wgg	$u\bar{d}\to Wgg$	$W\bar{u}gg$	$\bar{u}g\to W\bar{u}gg$
WW	$q\bar{q}\to WW$		
WZ	$q\bar{q}\to WZ$		
ggg	$gg\to ggg$		
$t\bar{t}$	$q\bar{q}\to t\bar{t}\to \ell^+\nu b\ell^-\bar{\nu}\bar{b}$		
$t\bar{t}$	$q\bar{q}\to t\bar{t}\to \ell^+\nu b\bar{u}d\bar{b}$	$t\bar{t}$	$q\bar{q}\to t\bar{t}\to \ell^+\nu b\bar{u}d\bar{b}$

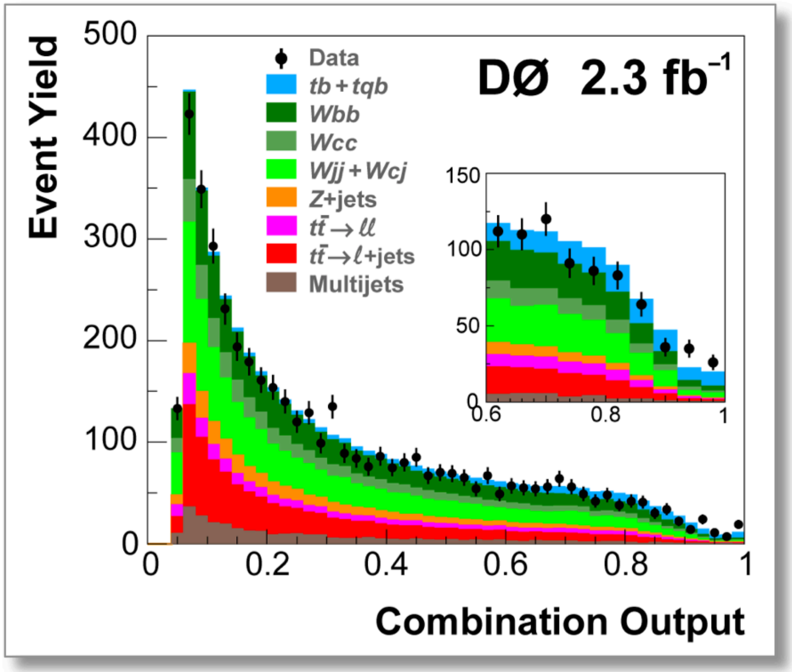




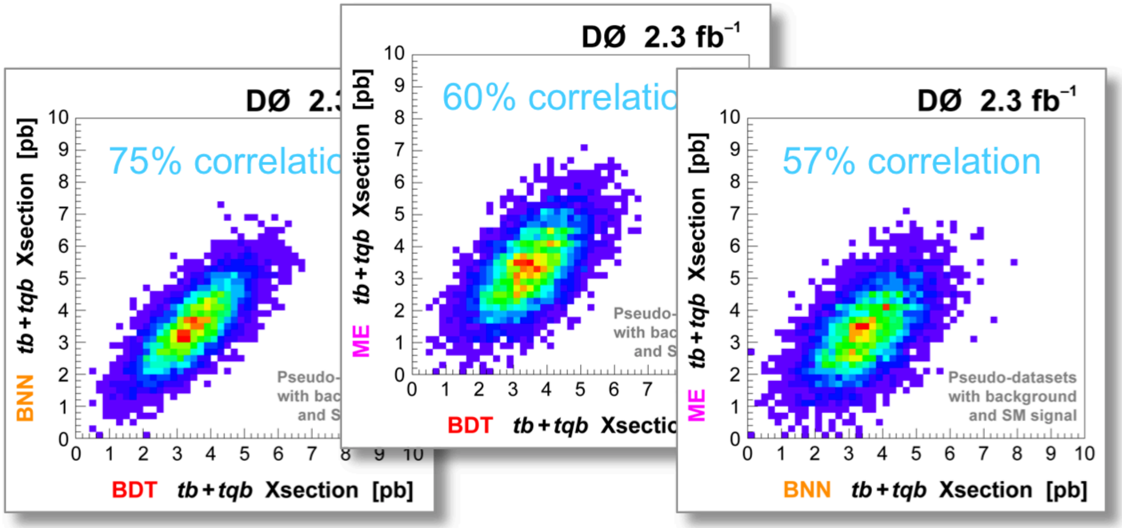
Combining the Results

- Combine the three analyses to improve the expected signal significance and cross section measurement precision
- This works because the three analyses are not 100% correlated
- Use a set of Bayesian Neural Networks
  - 3 inputs, 6 hidden nodes, 1 output
  - Markov Chain MC technique has 300 networks, average over the last 100



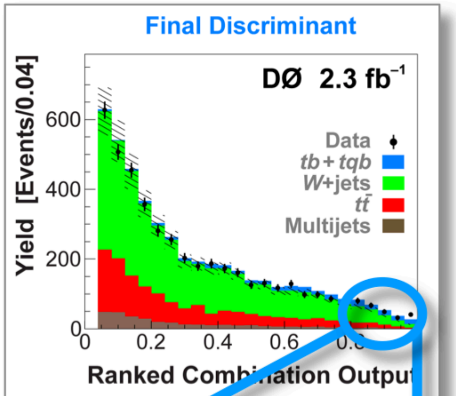


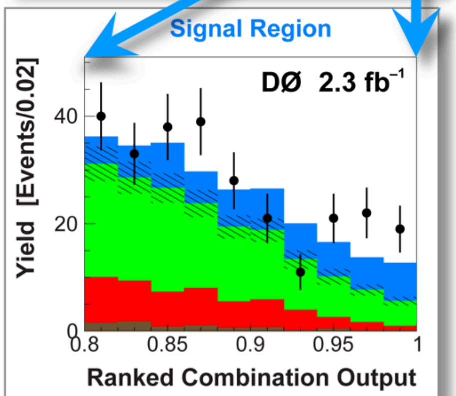
- Expected significance is improved by 4% over best individual analysis (BDT)

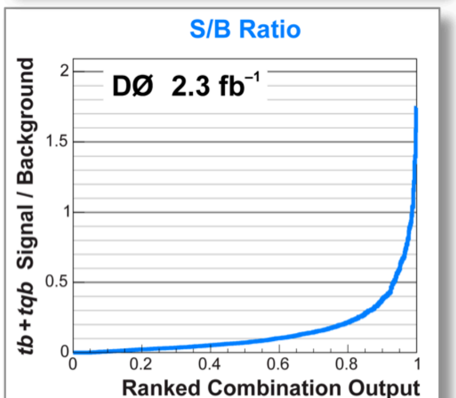


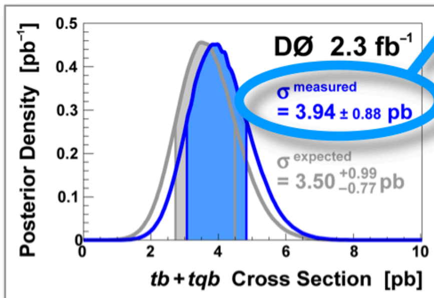
Single Top Observation – Results

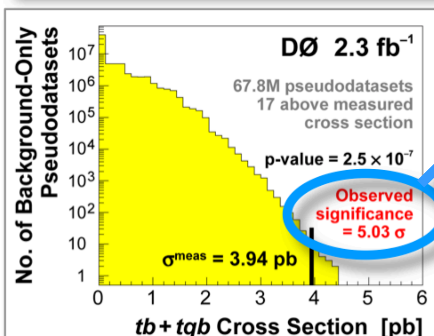
DØ 2.3 fb^{-1} Single Top Results			
Analysis Method	Single Top Cross Section	Significance: Expected	Significance: Measured
Boosted Decision Trees	$3.74^{+0.95}_{-0.79}$ pb	4.3 σ	4.6 σ
Bayesian Neural Networks	$4.70^{+1.18}_{-0.93}$ pb	4.1 σ	5.4 σ
Matrix Elements	$4.30^{+0.99}_{-1.20}$ pb	4.1 σ	4.9 σ
Combination	**3.94 ± 0.88 pb**	**4.5 σ**	**5.0 σ**

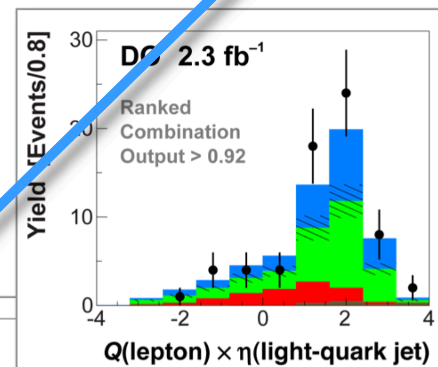


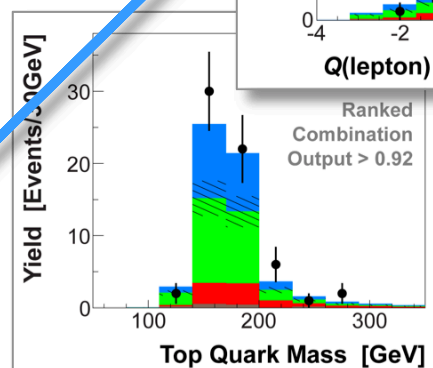










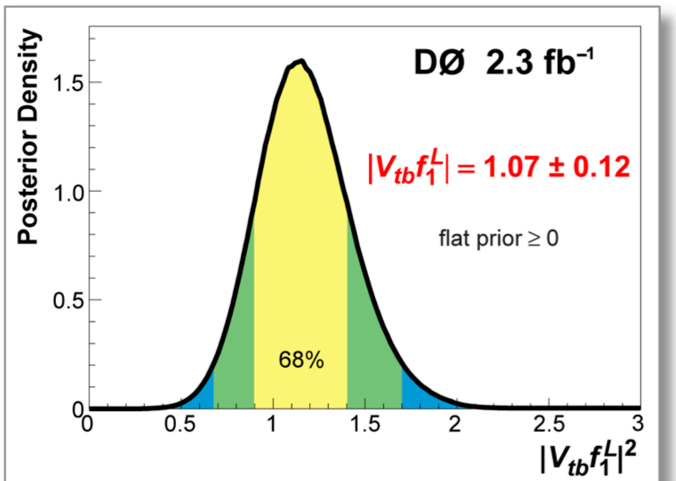


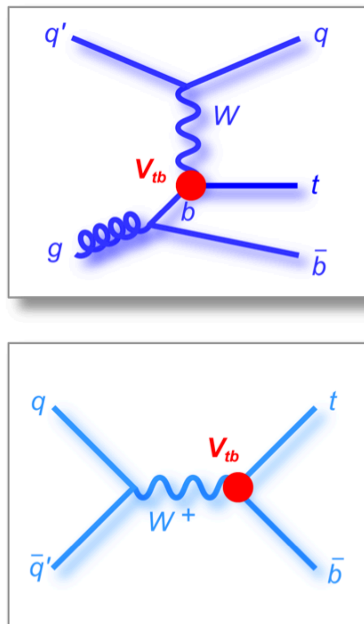
CKM Matrix Element V_{tb}

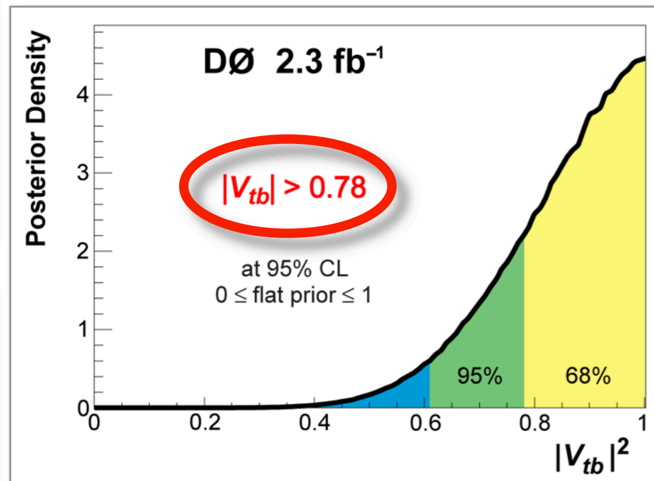
- **General form of *Wtb* vertex:**

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

- $\sigma(tb, tqb) \propto |V_{tb}|^2$ ∴ calculate a posterior in $|V_{tb}|^2$
- **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**

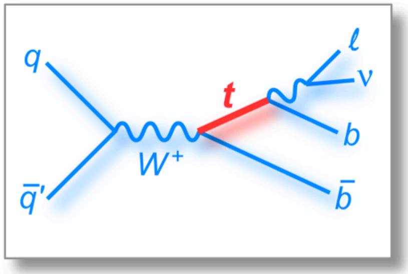


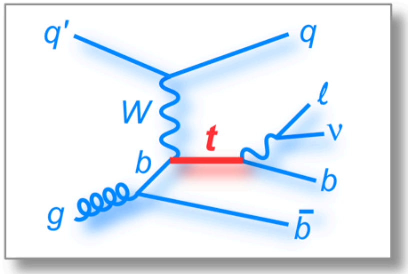




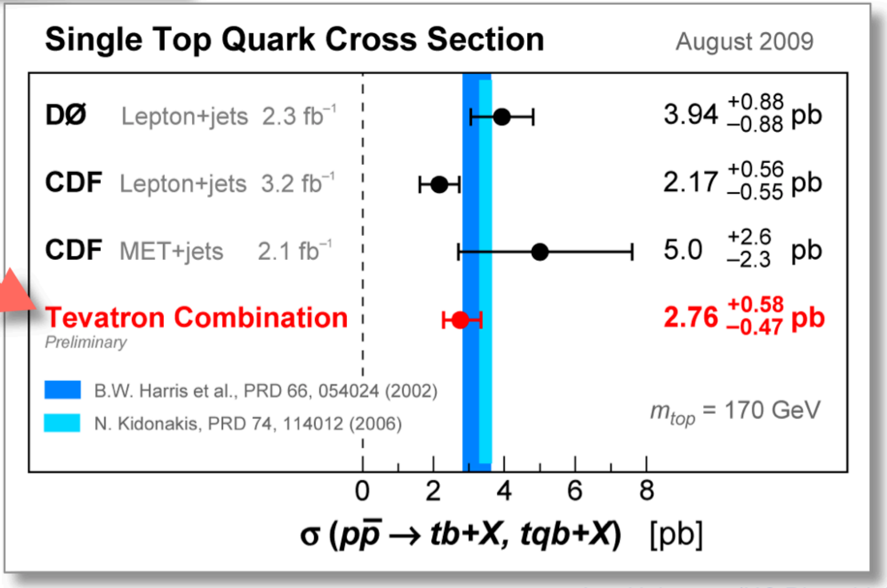
Combination with CDF's Result

Single Top Cross Section	Signal Significance Expected	Observed	CKM Matrix Element V_{tb}
DØ (2.3 fb^{-1}) March 2009 PRL **103**, 092001 (2009) (m_{top} = 170 GeV)			
3.94 ± 0.88 pb	4.5 σ	5.0 σ	$\lvert V_{tb} f_1^L \rvert = 1.07 \pm 0.12$ $\lvert V_{tb} \rvert > 0.78$ at 95% CL
CDF (3.2, 2.1 fb^{-1}) March 2009 PRL **103**, 092002 (2009) (m_{top} = 175 GeV)			
2.3 $^{+0.6}_{-0.5}$ pb	>5.9 σ	5.0 σ	$\lvert V_{tb} f_1^L \rvert = 0.91 \pm 0.13$ $\lvert V_{tb} \rvert > 0.71$ at 95% CL
DØ & CDF combined August 2009 FERMILAB-TM-2440-E (m_{top} = 170 GeV)			
2.76 $^{+0.58}_{-0.47}$ pb			$\lvert V_{tb} f_1^L \rvert = 0.88 \pm 0.07$ $\lvert V_{tb} \rvert > 0.77$ at 95% CL

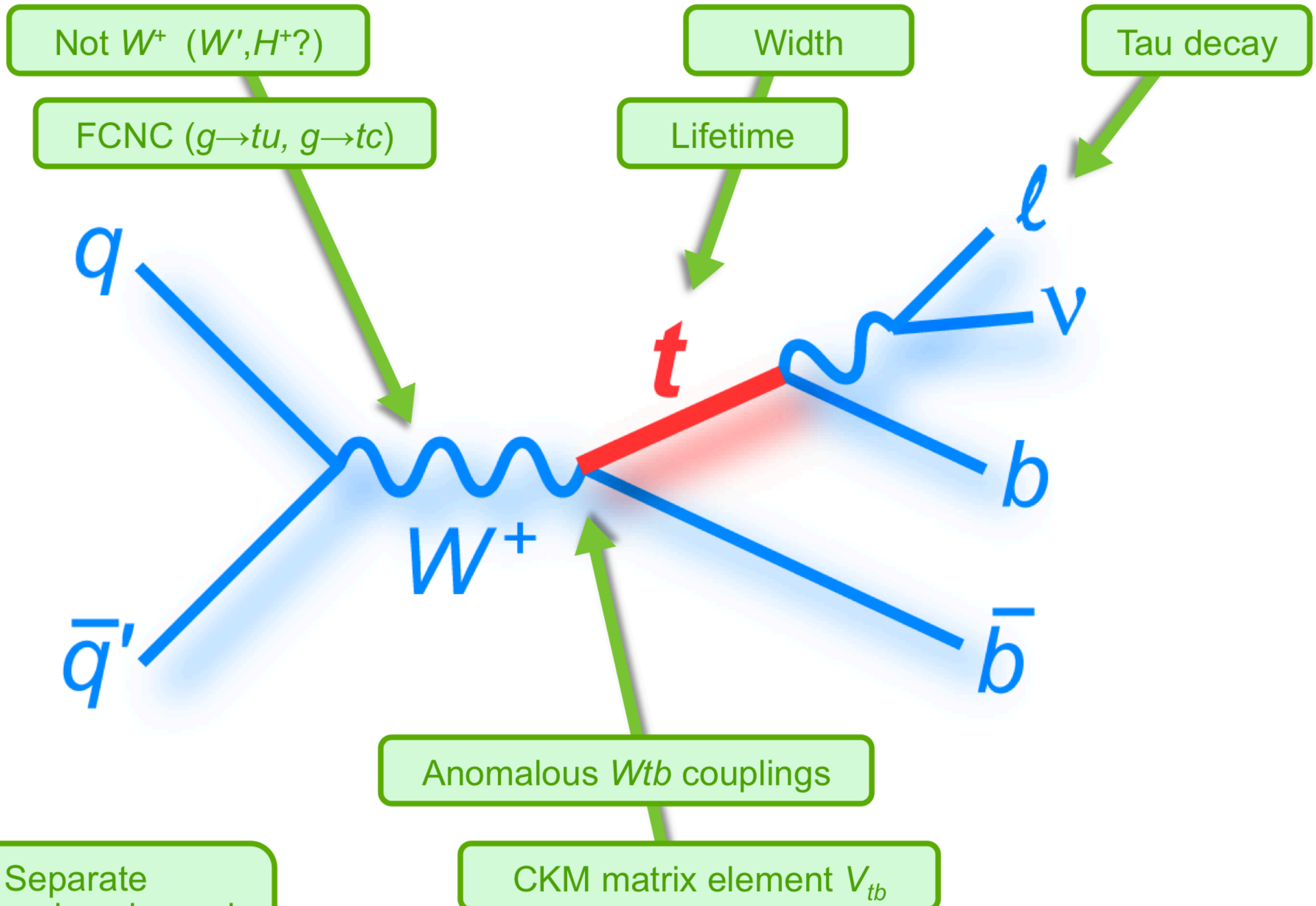




- A Bayesian statistical analysis is performed on all DØ and CDF analysis channels
- (Same method as used by DØ and CDF separately)
- All systematic uncertainties and their correlations are included



Top Quark Properties

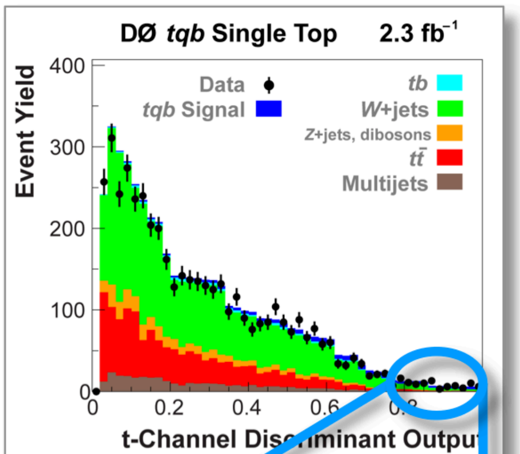


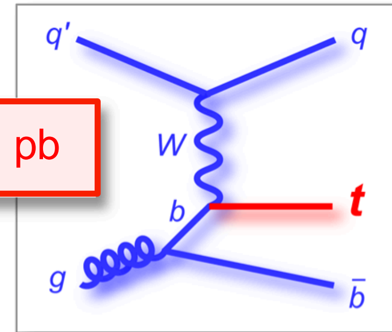
Separate t-channel, s-channel cross sections

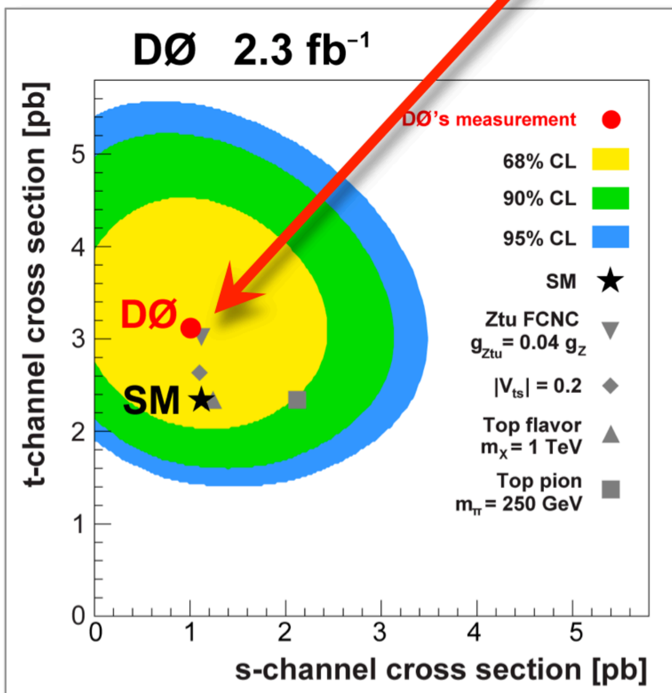
t-Channel *tqb* Measurement

- Use new BDTs, BNNs, and MEs on observation dataset, trained to find t-channel single top, then combine with BNNs
- First model-independent measurement of *tqb* production
- **Physics Letters B 682, 363 (2010)**

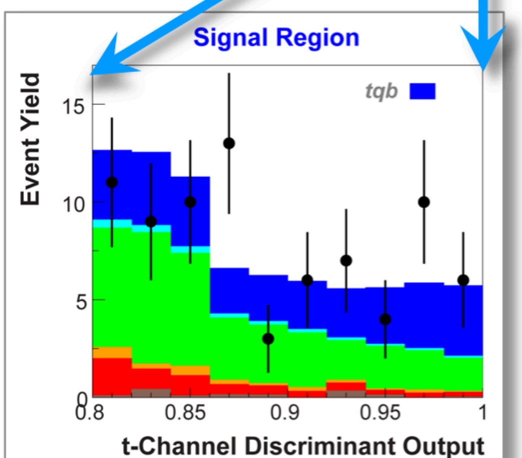
$$\sigma\left(p\bar{p} \rightarrow tqb + X\right) = 3.14^{+0.94}_{-0.80}\ \text{pb}$$

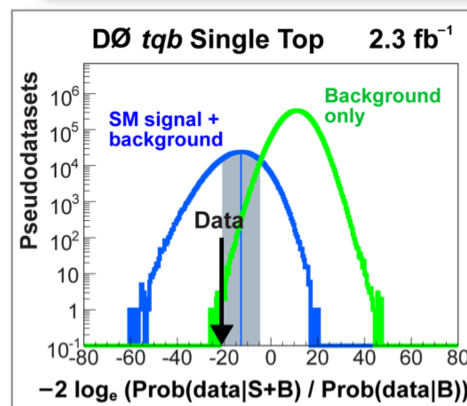






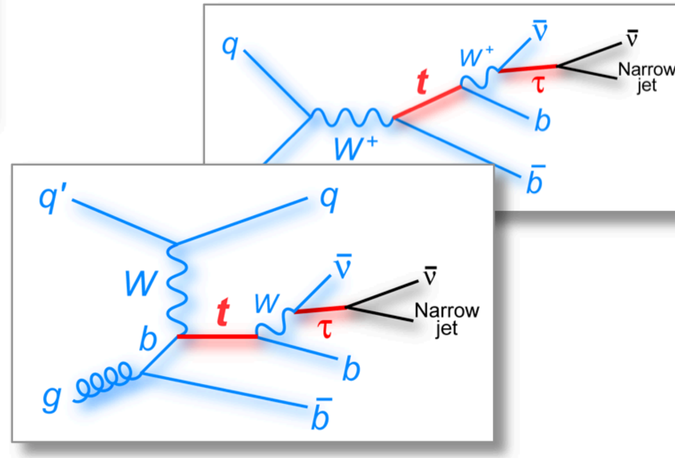
4.8σ significance

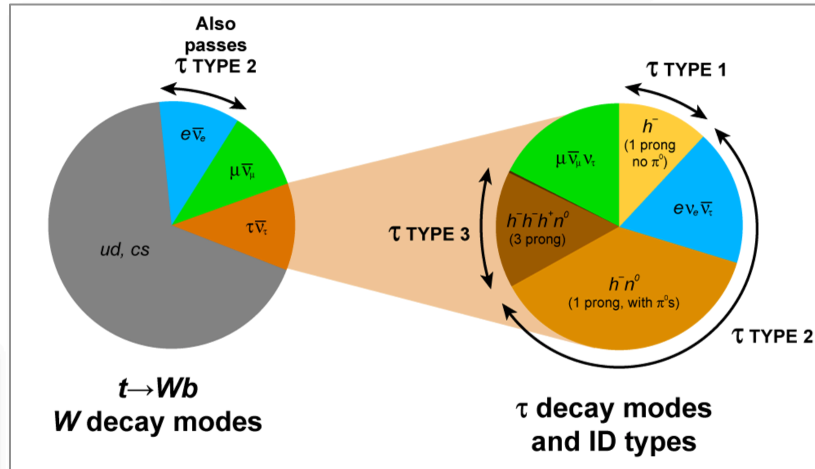


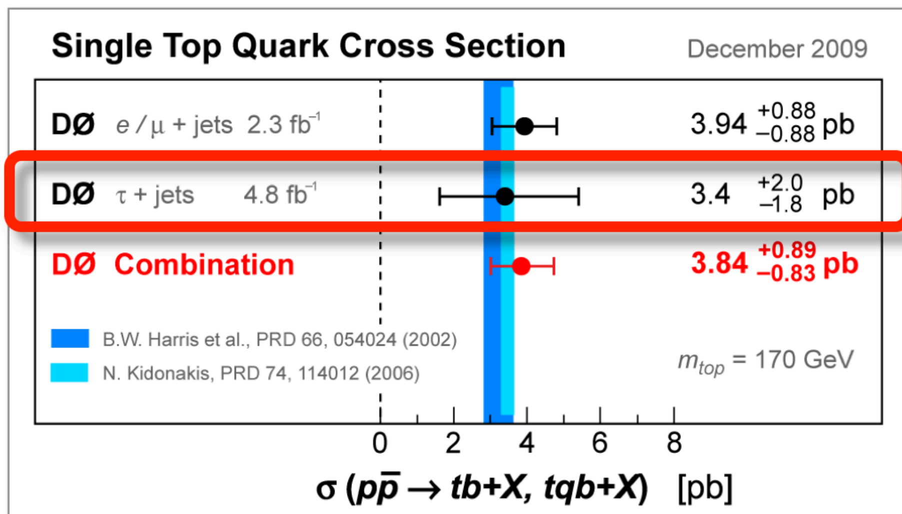


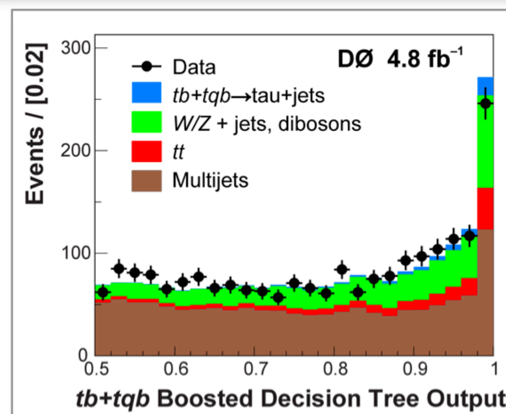
tqb+tb with τ Decay

- **Identify hadronically decaying tau leptons:**
  - TYPE 1 = calorimeter cluster + 1 track
  - TYPE 2 = cal cluster + 1 track + EM energy
  - TYPE 3 = cal cluster + 2 or 3 tracks
- **New tau ID for this analysis:**
  - boosted decision trees, 44–70 variables
- **Efficient for taus in events with jets:**
  - 76% (TYPE 1), 69% (2), 59% (3) for 98% background rejection
- **Phys. Lett. B 689, xxx (2010)**









Top Quark Width and Lifetime

- **Get the top quark partial width using:**
  - DØ's t-channel single top cross section measurement
  - DØ's branching ratio measurement from $t\bar{t}$ decay (i.e., don't assume $\mathcal{B}(t \to Wb) = 1$)
  - SM top quark partial width and t-channel cross section
- **Combine using Bayesian statistical analysis**

New Result !

$$\Gamma(t \to Wb)_{D\O} = \sigma(p\bar{p} \to tqb + X)_{D\O} \times \frac{\Gamma(t \to Wb)^{SM}_{NLO}}{\sigma(p\bar{p} \to tqb + X)^{SM}_{NLO}}$$

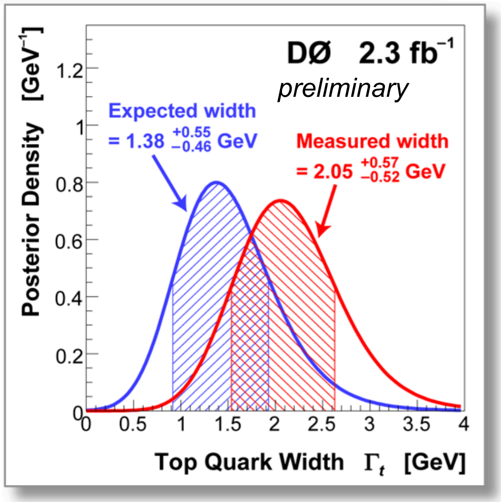
$$= \frac{3.14\ ^{+30\%}_{-25\%}\ \text{pb}}{0.962\ ^{+10\%}_{-9\%}} \times \frac{1.26\ \text{GeV}}{2.15 \pm 11\%\ \text{pb}} \Rightarrow 1.90\ ^{+0.58\ (+31\%)}_{-0.48\ (-25\%)}\ \text{GeV}$$

- **Convert partial width to full width**

$$\Gamma(t \to Wb)_{D\O} = \frac{\Gamma(t \to Wb)_{D\O}}{\mathcal{B}(t \to Wb)_{D\O}} \Rightarrow 2.05\ ^{+0.57\ (+28\%)}_{-0.52\ (-25\%)}\ \text{GeV}$$

- **Convert full width to lifetime**

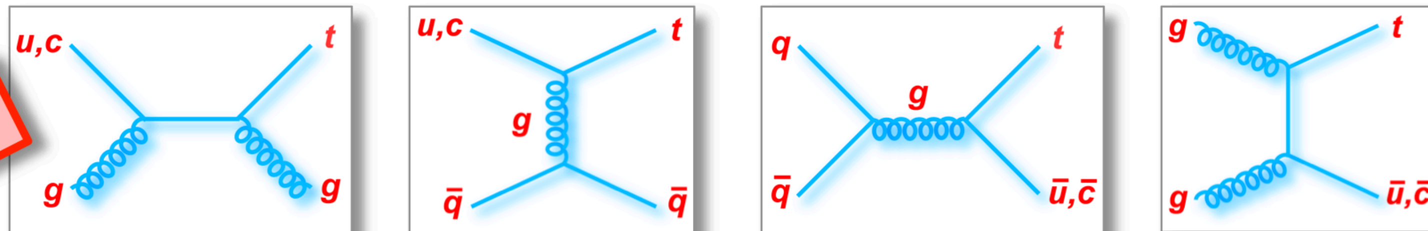
$$\tau(t) = \hbar / \Gamma(t) = \left(3.2\ ^{+1.1\ (+34\%)}_{-0.7\ (-22\%)}\right) \times 10^{-25}\ \text{s}$$

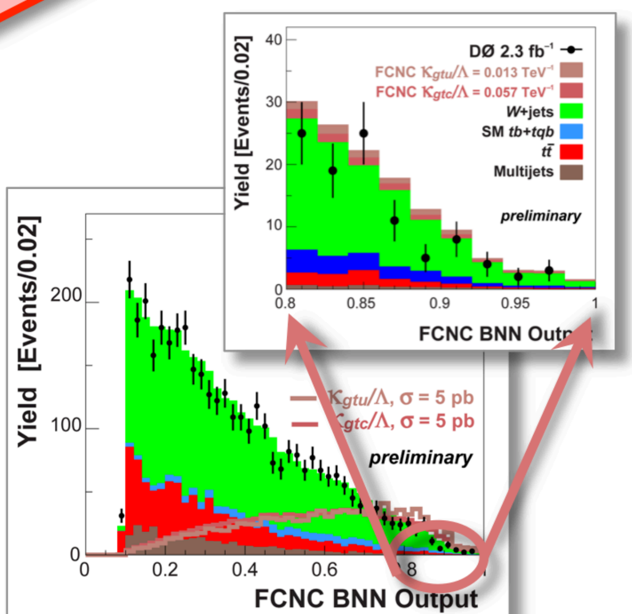


Flavor-Changing Neutral Currents

- Same selection as for cross section measurement, except:
  - require only one *b*-tagged jet (from decay of top) – no 2nd *b* present
- Bayesan neural networks used to combine ~24 kinematic variables per channel from a total of 54 to separate signal from background
- **Cross sections scale with** $\left(\kappa_{gtu}/\Lambda\right)^2$ **and** $\left(\kappa_{gtc}/\Lambda\right)^2$

New Result !



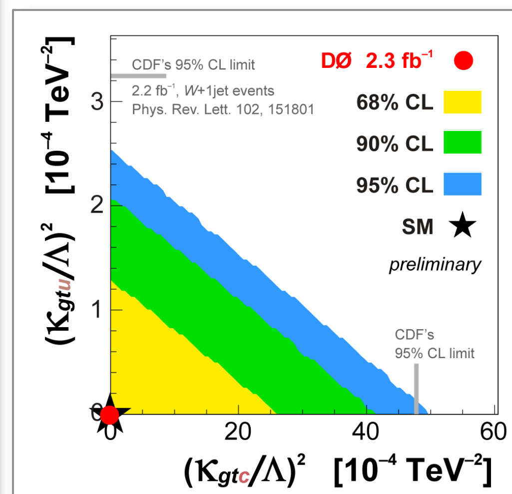


$$\kappa_{gtu}/\Lambda \;<\; 0.013\ \text{TeV}^{-1}$$
$$\kappa_{gtc}/\Lambda \;<\; 0.057\ \text{TeV}^{-1}$$

$$B\left(t \to gu\right) \;<\; 2.0\times10^{-4}$$
$$B\left(t \to gc\right) \;<\; 3.9\times10^{-3}$$

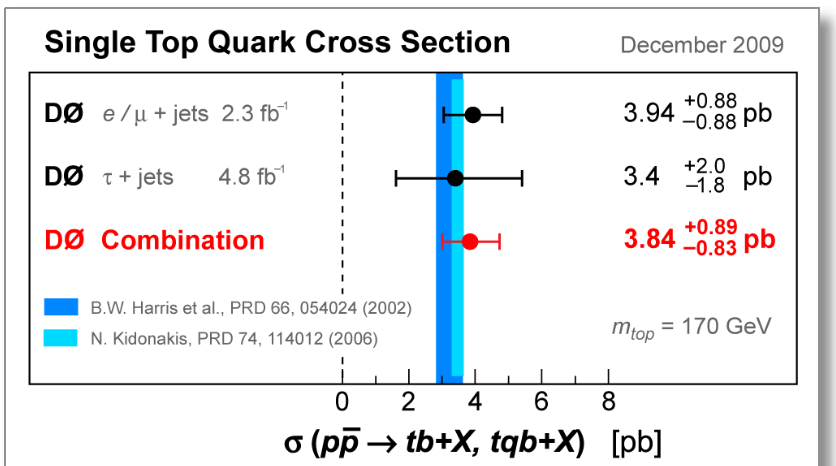
$$\sigma\left(gtu\right) \;<\; 0.20\ \text{pb}$$
$$\sigma\left(gtc\right) \;<\; 0.27\ \text{pb}$$

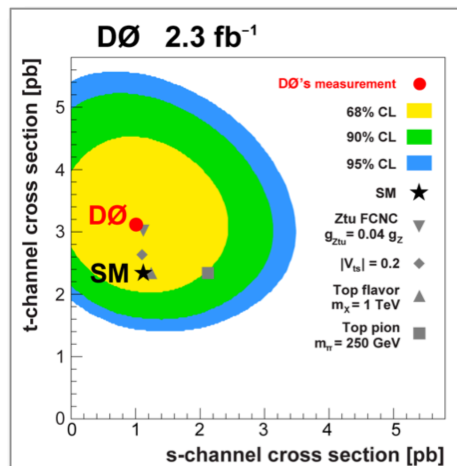
at 95% CL

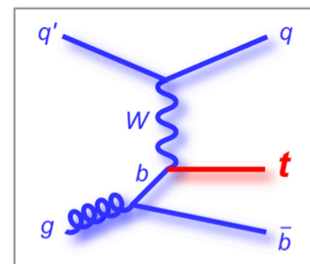


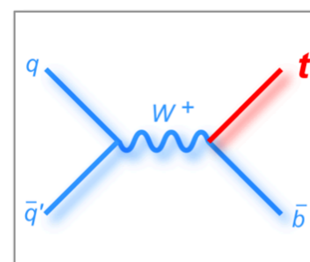
Single Top Quark Summary

- Challenging measurements
  – small signal hidden in large complex background (S:B = 1:20)
- It took 14 years since top quark pair observation to collect 50x more data and develop the background models and analysis techniques to be able to observe single top quark production with 5σ significance
- Tools and techniques are now being applied to the Tevatron Higgs boson searches and elsewhere
- The data are being used to make sensitive tests of the properties of the top quark and to search for physics beyond the SM
- 2.3 fb^{-1} of data needed for observation, 5.4 fb^{-1} of data will soon be analyzed, expect 10 fb^{-1} by October 2011, with discussion ongoing to run for 3 more years
- **A golden period for the physics of the top quark!**









Additional Material

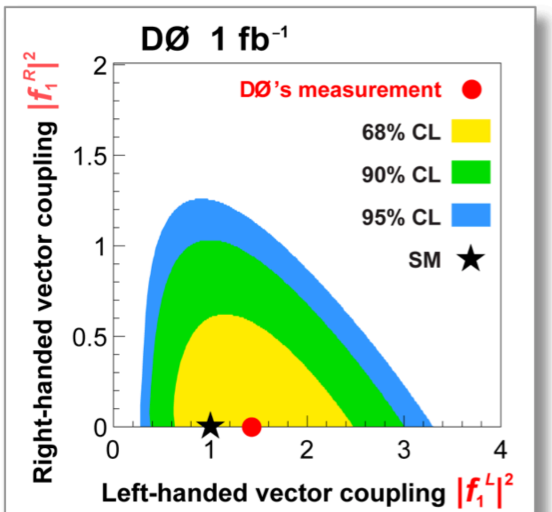
Wtb Anomalous Couplings

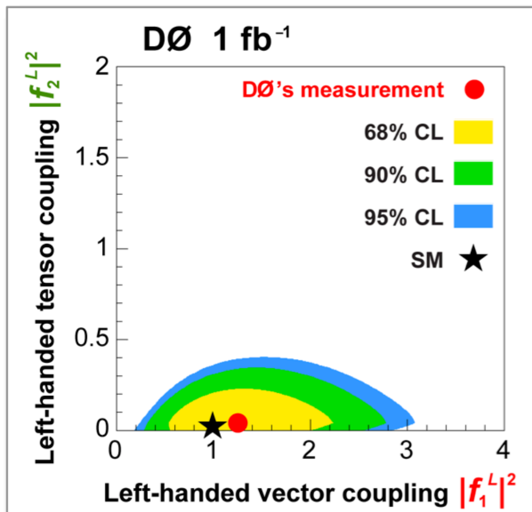
- Combined t-channel single top quark measurement with DØ's *W* helicity measurement in top pair decays

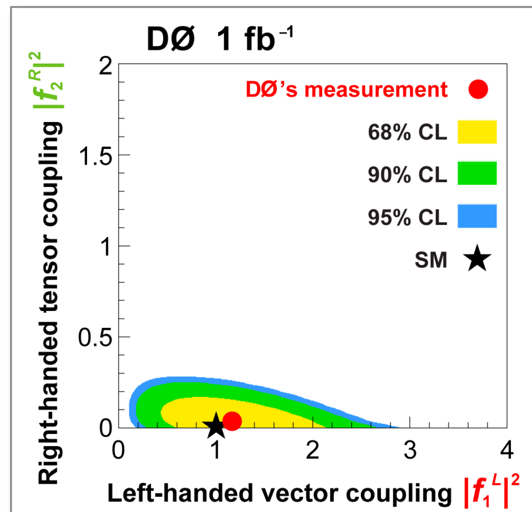
$$\Gamma^{\mu}_{Wtb} = -\frac{g}{\sqrt{2}} V_{tb} \left\{ \gamma^{\mu} \left[f_1^L P_L + f_1^R P_R \right] - \frac{i\sigma^{\mu\nu}}{M_W} (p_t - p_b)_{\nu} \left[f_2^L P_L + f_2^R P_R \right] \right\}$$

Wtb coupling — Vector form factors — Tensor form factors

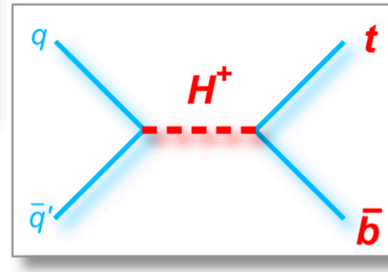
- If $f_1^L = 1$, then at 95% CL, $|f_1^R|^2 < 1.01$, $|f_2^L|^2 < 0.28$, and $|f_2^R|^2 < 0.23$
- **Phys. Rev. Lett. 102, 092002 (2009)**



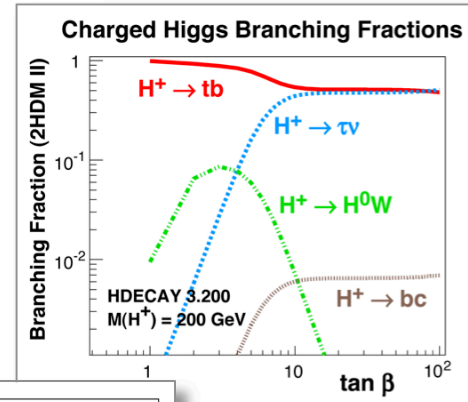


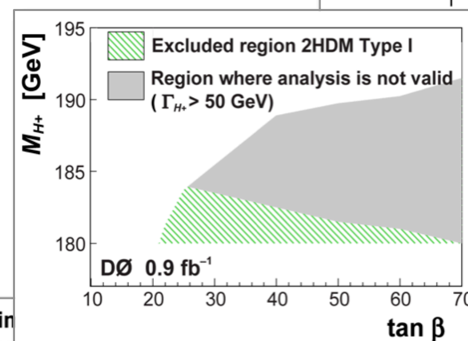


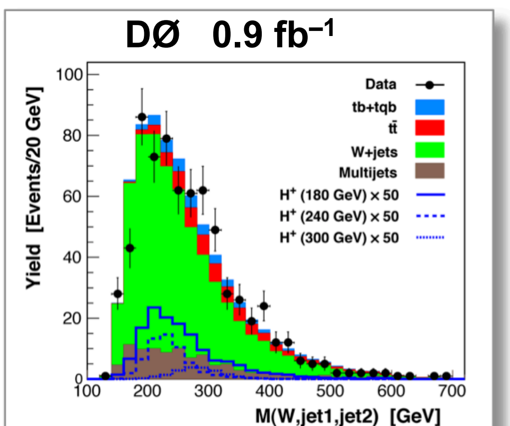
Charged Higgs Resonance

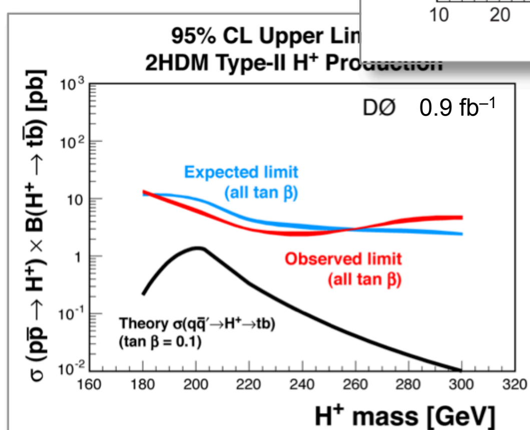


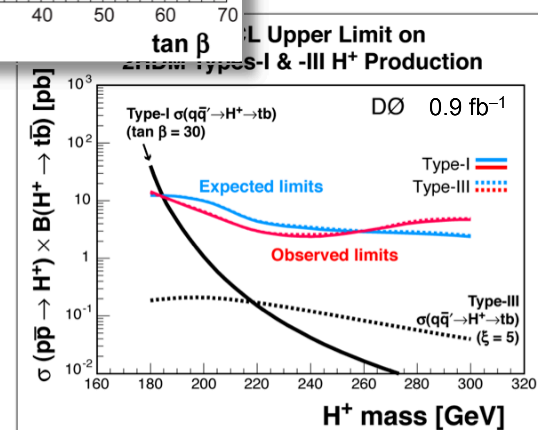
- **First search for heavy H^+ decaying to top**
  - s-channel *tb* resonance, use 2-jets, 1,2 *b*-tags
  - Binned likelihood calculation in *M*(*W*,jet1,jet2,)
- **Three two-Higgs-doublet models (2HDM) studied**
  - Type I – only one doublet couples to fermions
  - Type II – one doublet couples to up-type fermions, the other to down-type fermions
  - Type III – both doublets couple to fermions
- **Upper limits set close to predicted Xsecs**
  - Small region excluded for 2HDM Type-I H^+ between 180 and 184 GeV with 20 < tan β < 70
- **Phys. Rev. Lett. 102, 191802 (2009)**



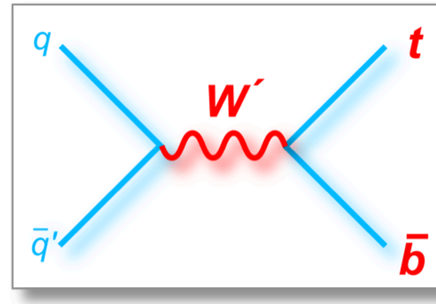




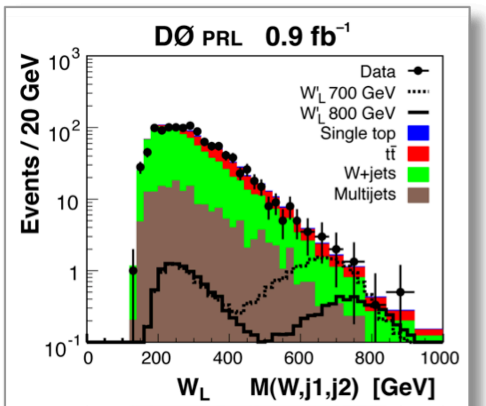


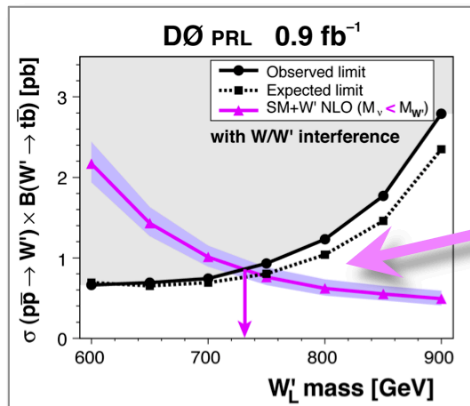


Heavy W' Resonance

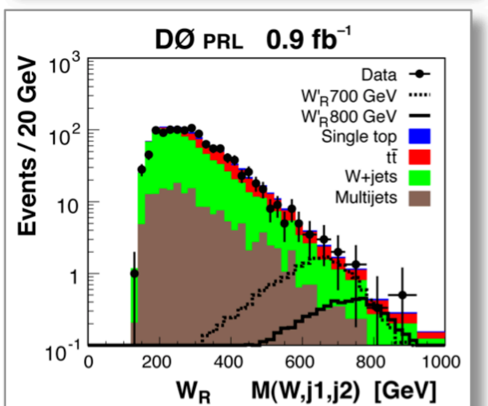


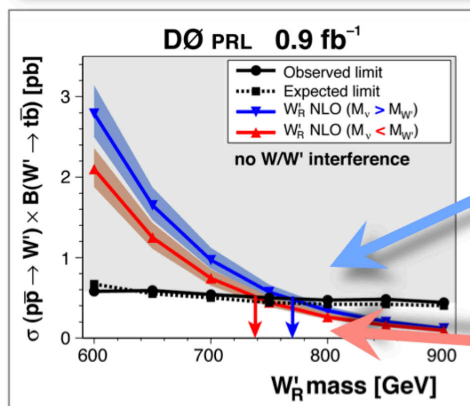
- **Search for s-channel *tb* resonance**
  - left-handed W'_L with SM couplings (including interference with SM W in signal model)
  - right-handed W'_R that decays to $\ell\nu$ and $q\bar{q}'$ $(M(\nu_R) < M(W'_R))$
  - right-handed W'_R that decays only to $q\bar{q}'$ $(M(\nu_R) > M(W'_R))$
- **Phys. Rev. Lett. 100, 211803 (2008)**





$M(W'_L) > 731$ GeV





$M(\nu_R) > M(W'_R)$

$M(W'_R) > 768$ GeV

$M(\nu_R) < M(W'_R)$

$M(W'_R) > 739$ GeV

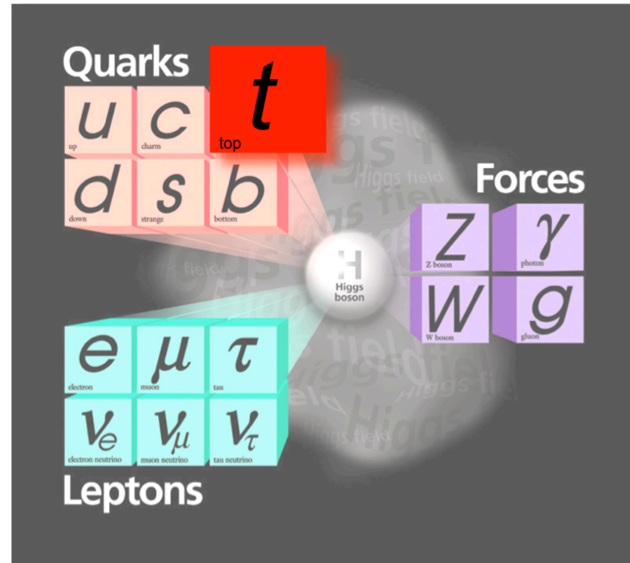
Top Quarks

- First observed in 1995 by DØ and CDF

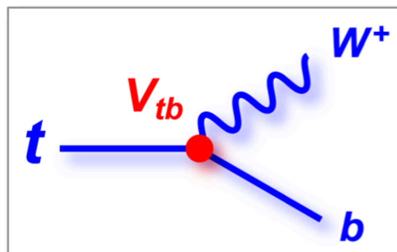
- Spin 1/2 fermion,   charge +2*e*/3
- Weak-isospin partner of the bottom quark
- ~40x heavier than its partner

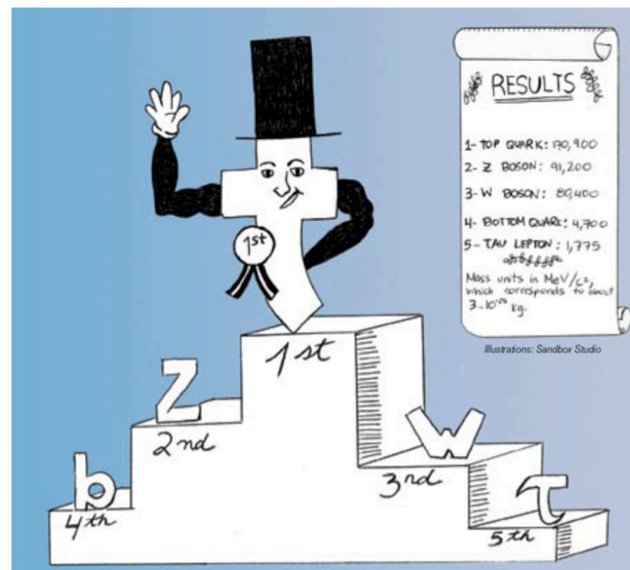
Mass = 173.1 ± 1.3 GeV

- Heaviest known fundamental particle



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





- Decays almost 100% of the time to *Wb*

Lifetime = 0.5 × 10^{-24} seconds

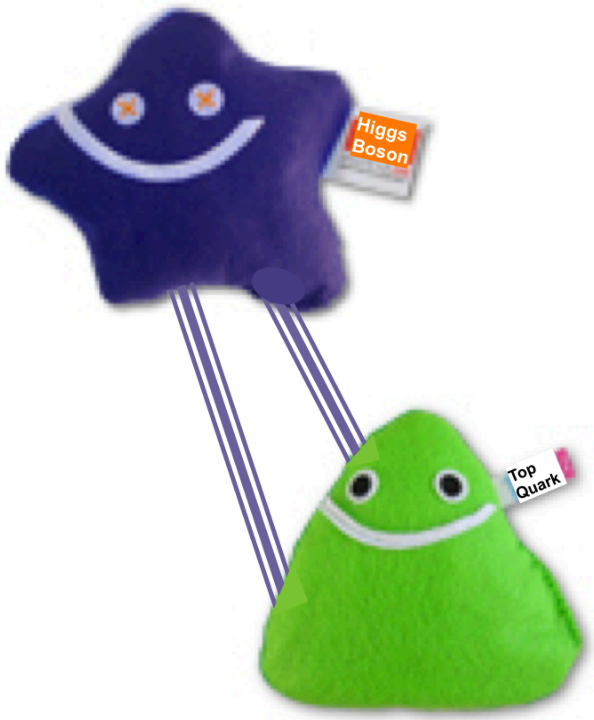
- Top quark decays before it hadronizes
- We are studying a naked quark

Electroweak Symmetry Breaking

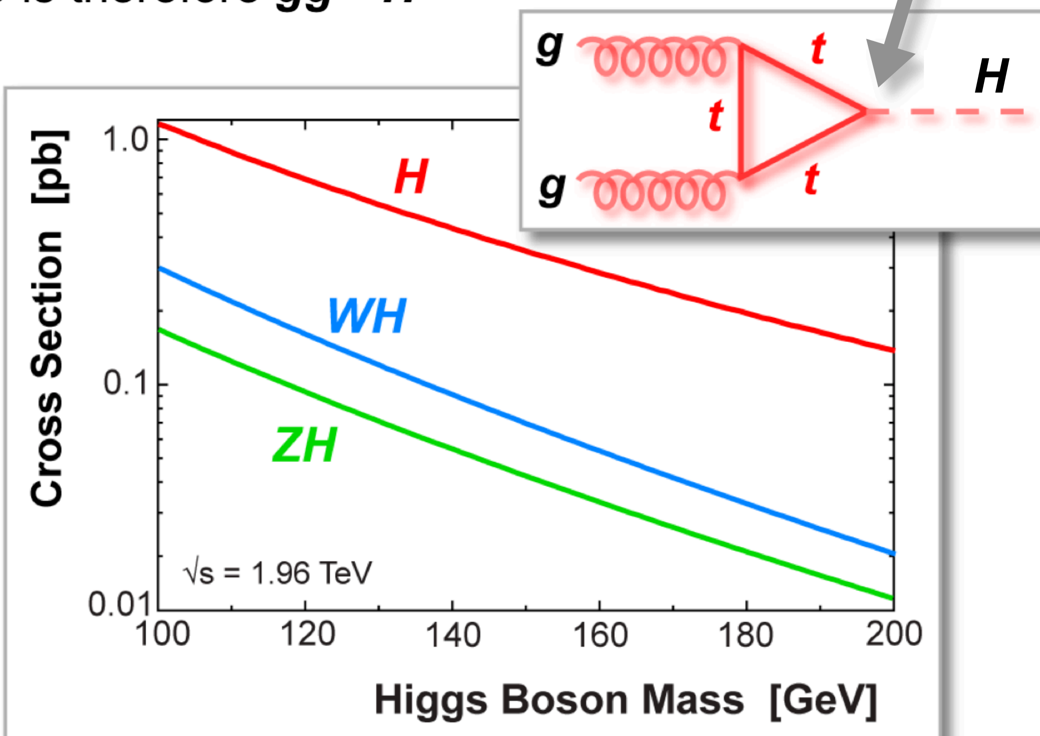
- The top quark is intimately linked with EWSB
- Higgs-top Yukawa coupling:

$$g_{Ht} = \sqrt{2}\ m_t / \mathrm{VEV} = \sqrt{2}\ 173.1\ \mathrm{GeV} / 246\ \mathrm{GeV} \approx 1$$

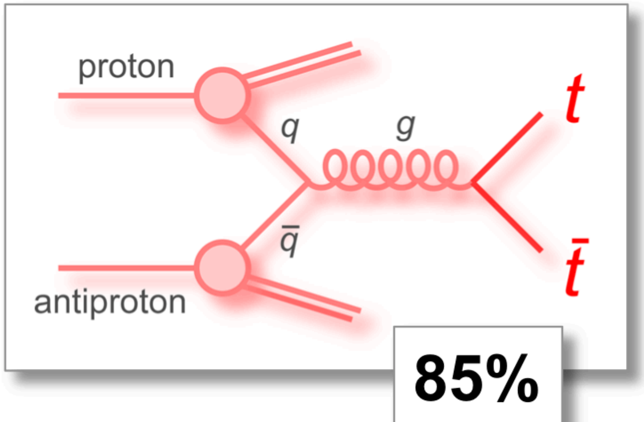
- … so Higgs will couple to the top quark rather than anything else
- Dominant production mode is therefore ***gg→H***



www.particlezoo.net



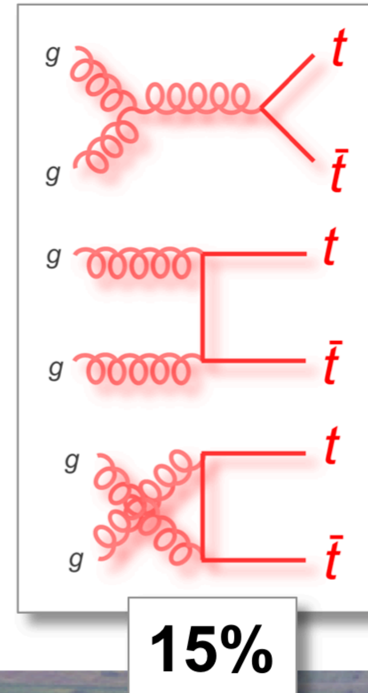
Top Pairs at the Tevatron

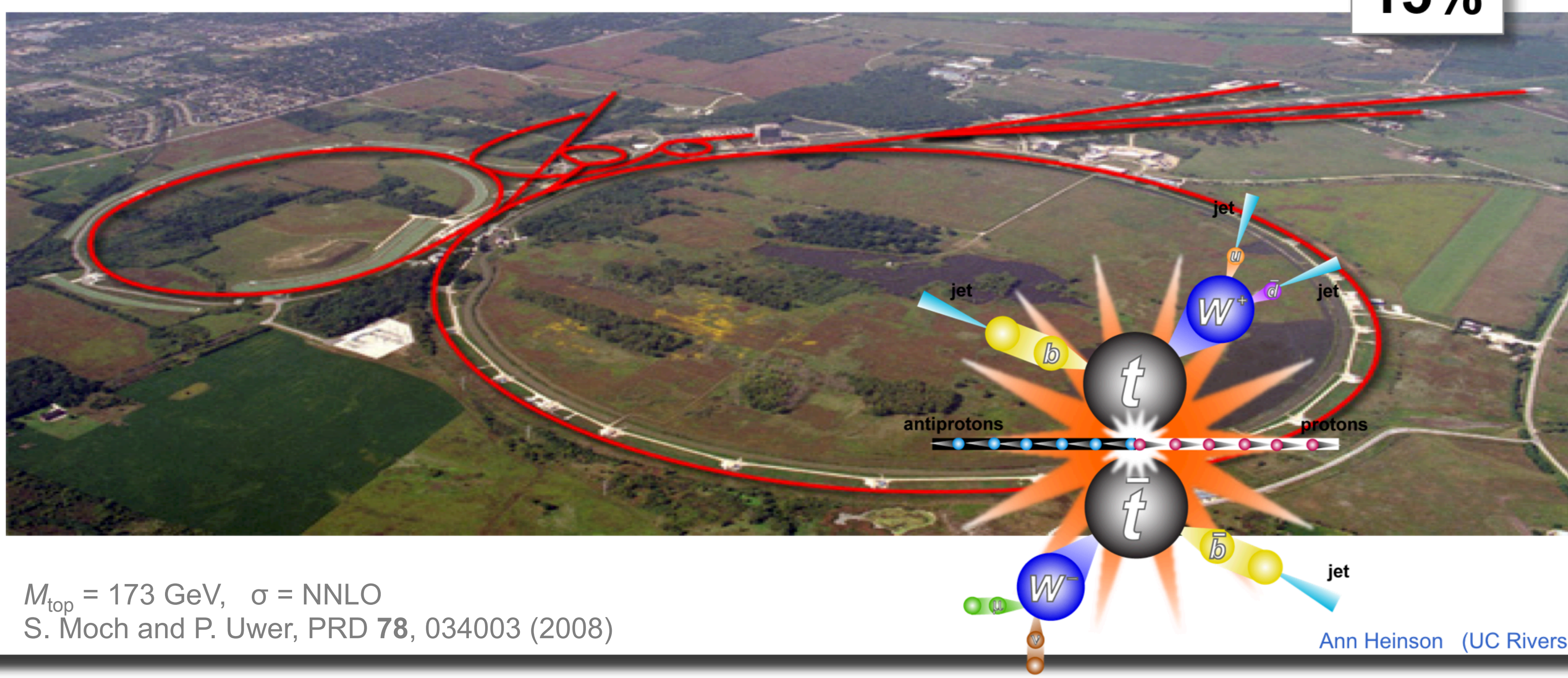


- Top produced mostly in $t\bar{t}$ pairs (*strong* coupling)

  σ_{NNLO} = 7.4 ± 0.6 pb

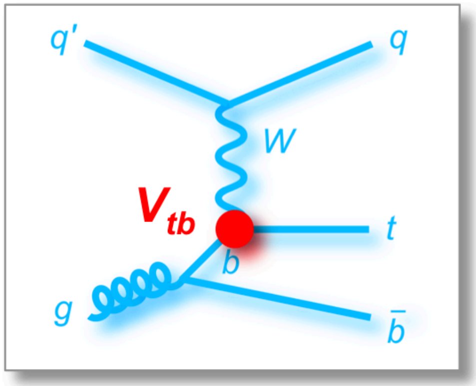
- Many top quark properties measured with these events



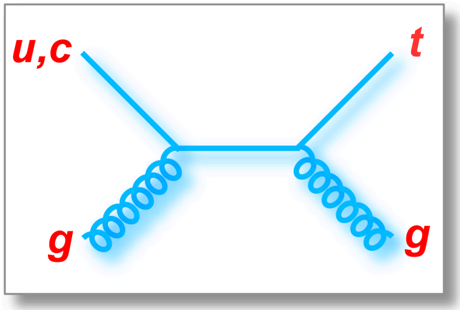


M_{top} = 173 GeV, σ = NNLO
S. Moch and P. Uwer, PRD **78**, 034003 (2008)

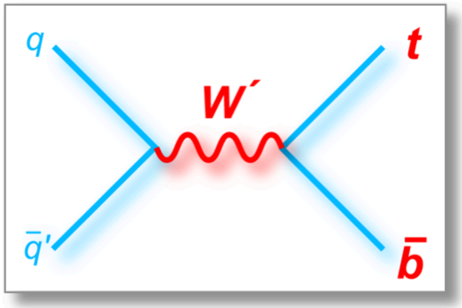
Motivation

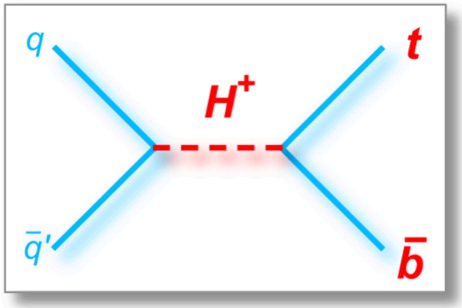


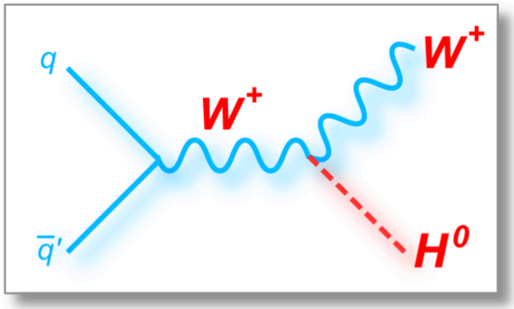
- **Study *Wtb* coupling in top production**
  - Measure $|V_{tb}|$ directly
  - Test unitarity of CKM matrix
  - Anomalous *Wtb* couplings
- **Cross sections sensitive to new physics**
  - **s-channel**: resonances (heavy *W'* boson, charged Higgs boson, Kaluza-Klein excited W_{KK}, technipion, top flavor *X*, etc.), flavor-changing neutral currents ($t - Z / \gamma / g - c / u$ couplings)
  - Fourth generation of quarks



- **Polarized top quarks – spin correlations 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



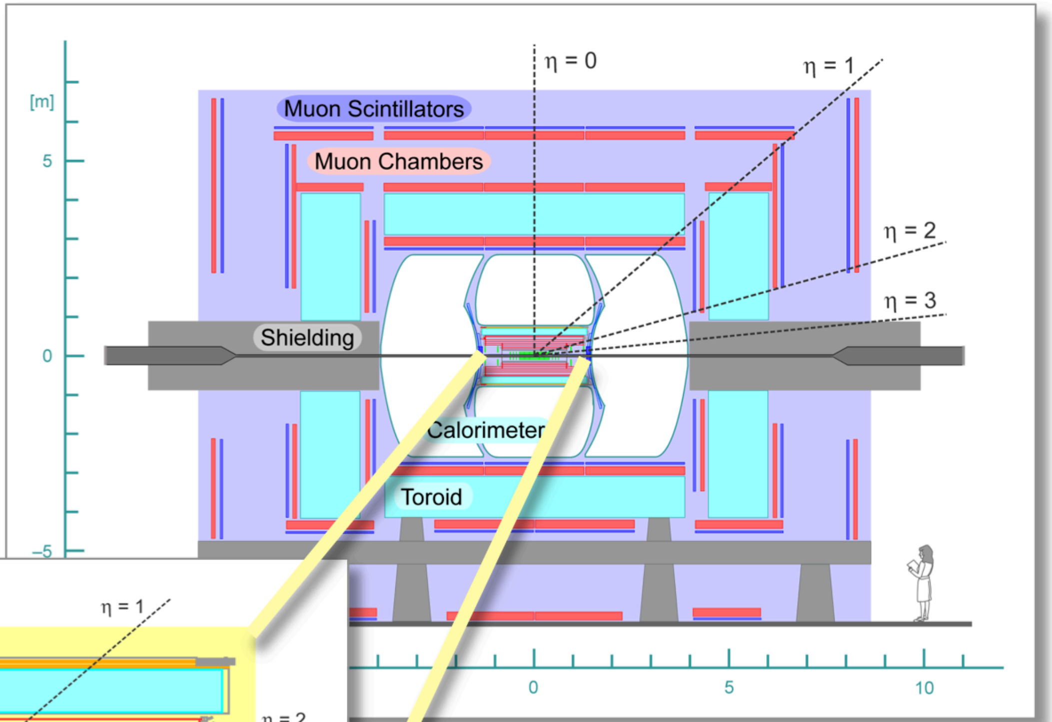


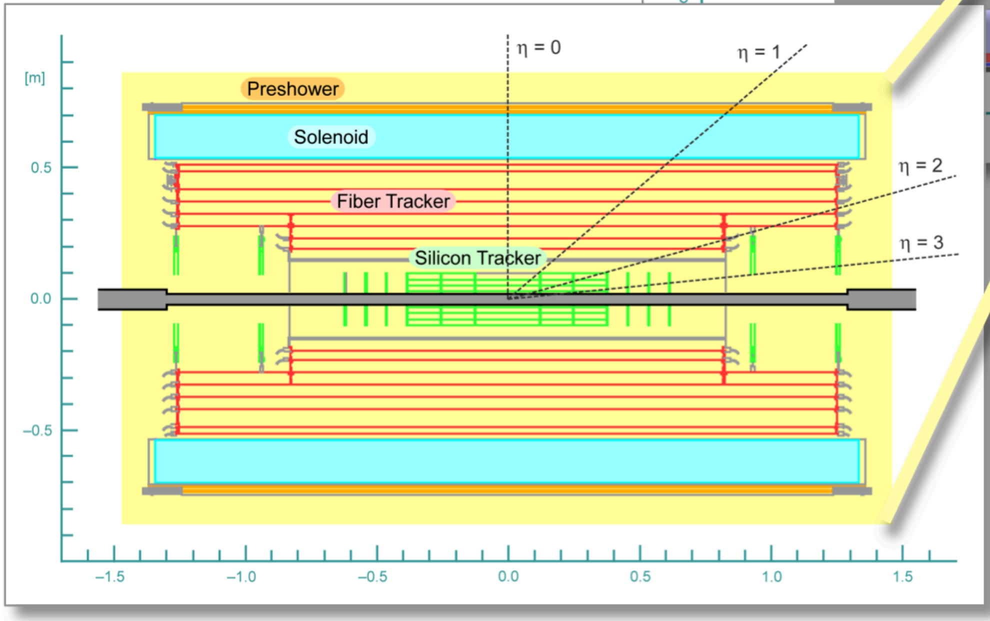


The DØ Experiment

Thick hermetic calorimeter:
- Excellent jet reconstruction
- Excellent $\not{E}_T$ resolution
- No punch-through

2 T central solenoid and 1.8 T outer toroids:
- Two muon p_T measurements





High resolution tracking
Low mass system
Close to beam-pipe
- Low photon emission
- Excellent vertex resolution

DØ's Search History

25x more data

Many improvements in analysis methods

Searches, upper limits

PRD **63**, 031101	(2000)	0.09 fb^{-1}	Cuts	TOPCITE = 50+
PLB **517**, 282	(2001)	0.09 fb^{-1}	Neural networks (28 variables)	TOPCITE = 50+
PLB **622**, 265	(2005)	0.23 fb^{-1}	NNs (25 variables) Bayesian likelihoods	TOPCITE = 50+
PRD **75**, 092007	(2007)	0.23 fb^{-1}	Long write-up	

>3σ Evidence

PRL **98**, 181802	(2007)	0.9 fb^{-1}	Boosted decision trees (49 variables) Bayesian neural networks (25 variables) Matrix elements Bayesian likelihoods	TOPCITE = 100+
PRD **78**, 012005	(2008)	0.9 fb^{-1}	Long write-up	TOPCITE = 50+

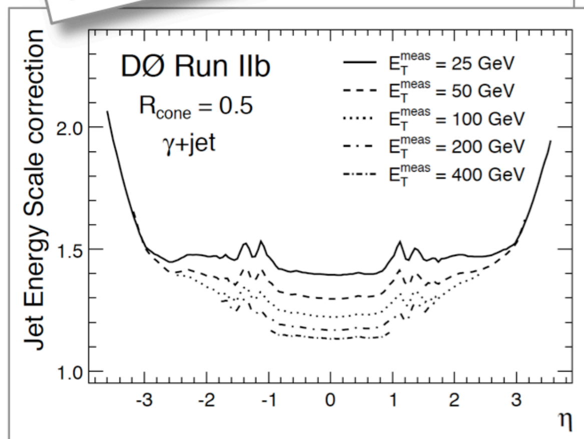
5σ Observation

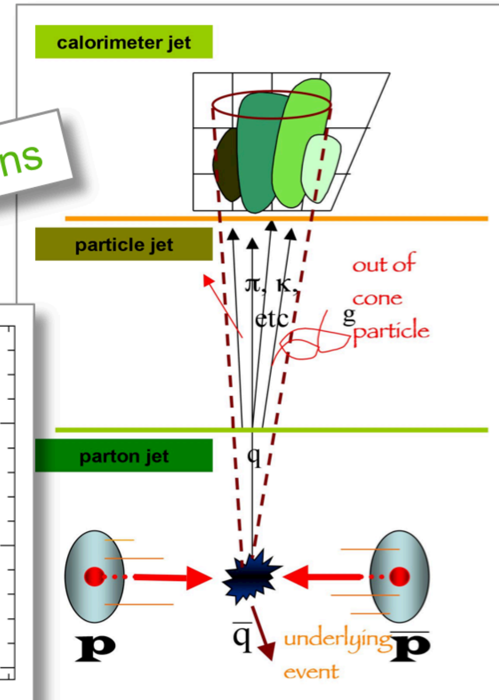
PRL **103**, 092001	(2009)	2.3 fb^{-1}	Boosted decision trees (64 variables) Bayesian NNs (18–28 variables) Matrix elements Bayesian-NN combination Bayesian likelihoods	TOPCITE = 50+

Event Reconstruction

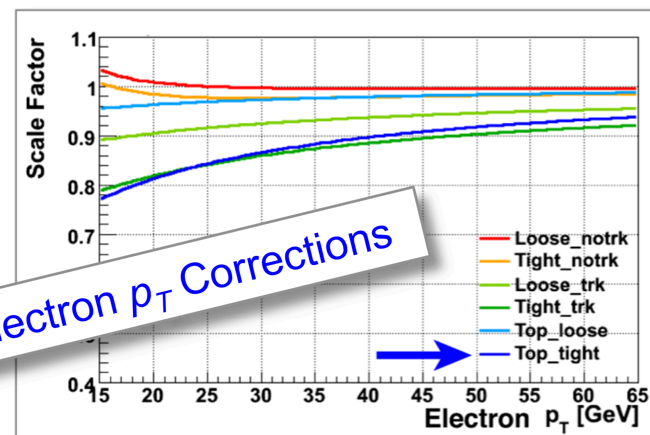
- MC events are processed through a GEANT model of the DØ detector
- Data and MC events are reconstructed to identify jets, EM objects, muons
- Jets' energies are corrected back to particle level, with ~1.3% uncertainty

Jet Energy Scale Corrections





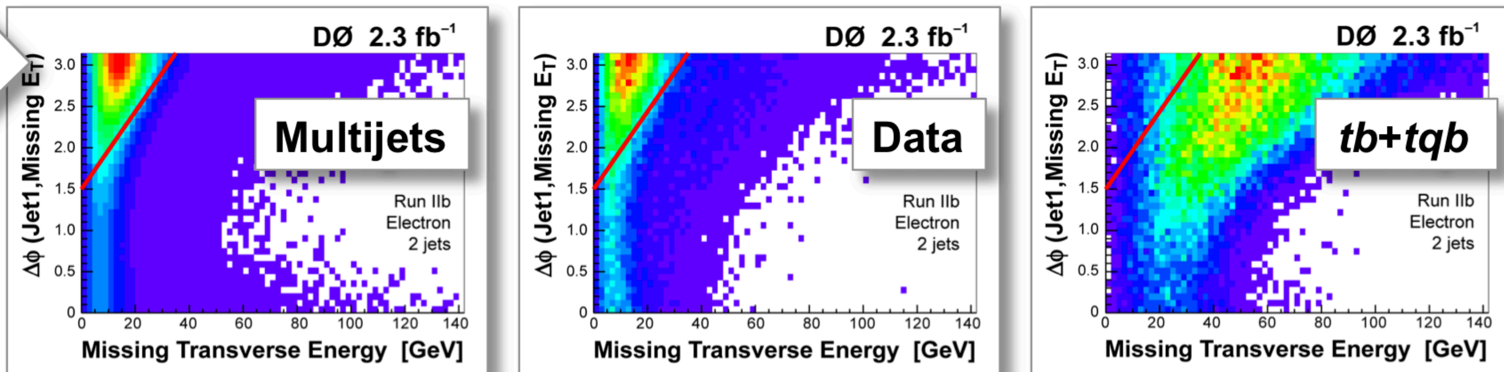
- Monte Carlo events need further reweighting to reproduce data
- Weights are applied to each MC event to correct the efficiencies and distributions:
  - Primary vertex, electron, and muon IDs
  - Jet energies are shifted, smeared, and some are removed
  - Missing transverse energy is adjusted
  - Taggability and *b*-tagging probabilities
  - Instantaneous luminosity distributions



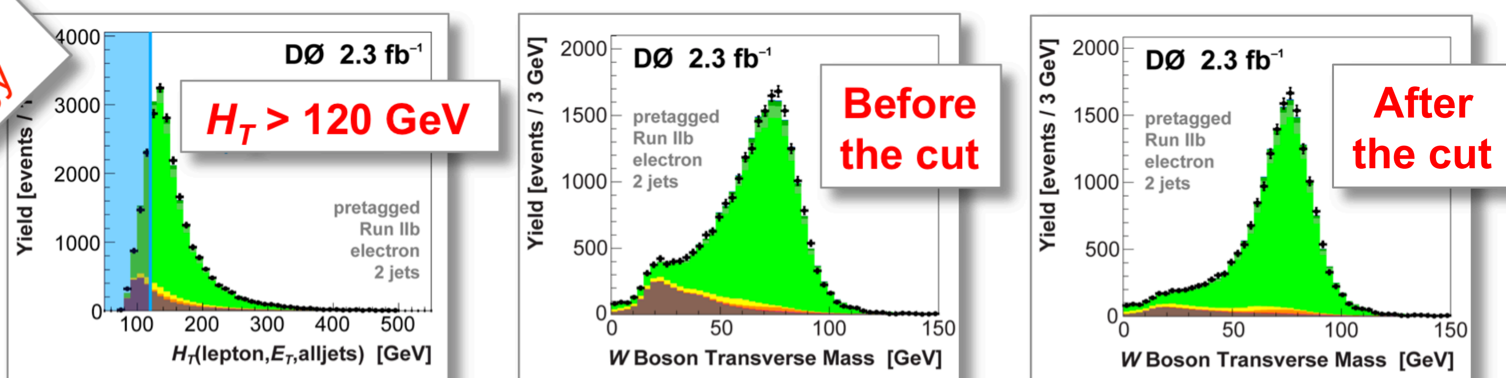
Event Selection – Fake Removal

- Multijets background is created when:
  - a jet is misidentified as an electron, or
  - a muon from a *b* decay travels wide of its jet or the jet is not found
- Difficult to model accurately and get the heavy-flavor jet fractions correct
- Strategy – remove as much multijet background as reasonably possible

Remove events with mismeasured $\not{E}_T$



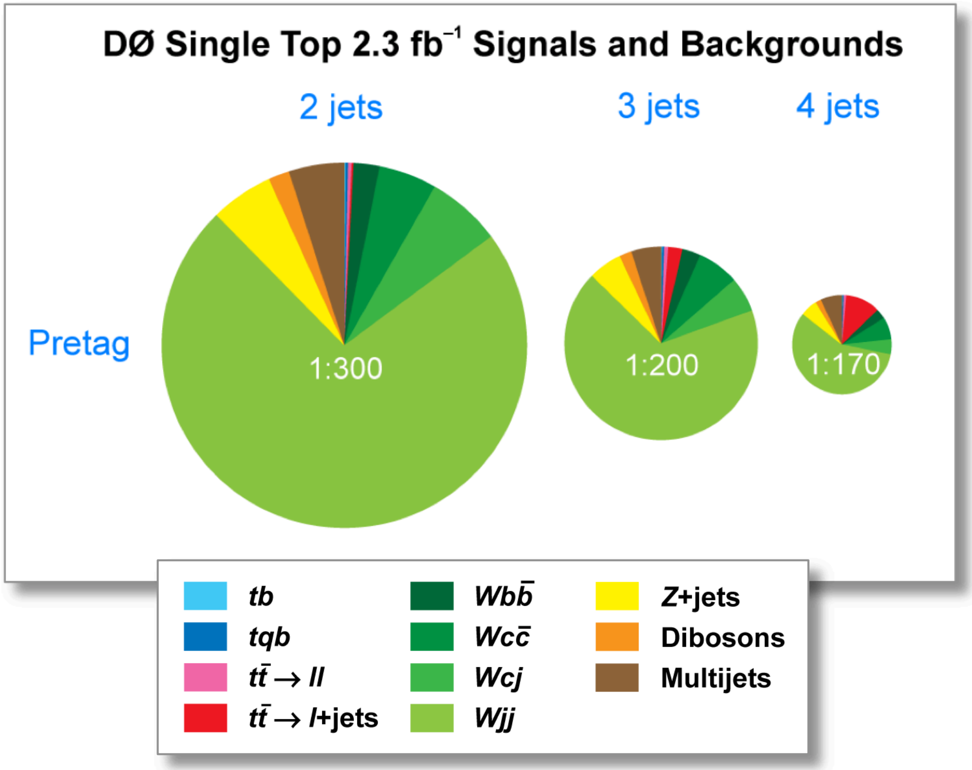
Remove events with low total transverse energy

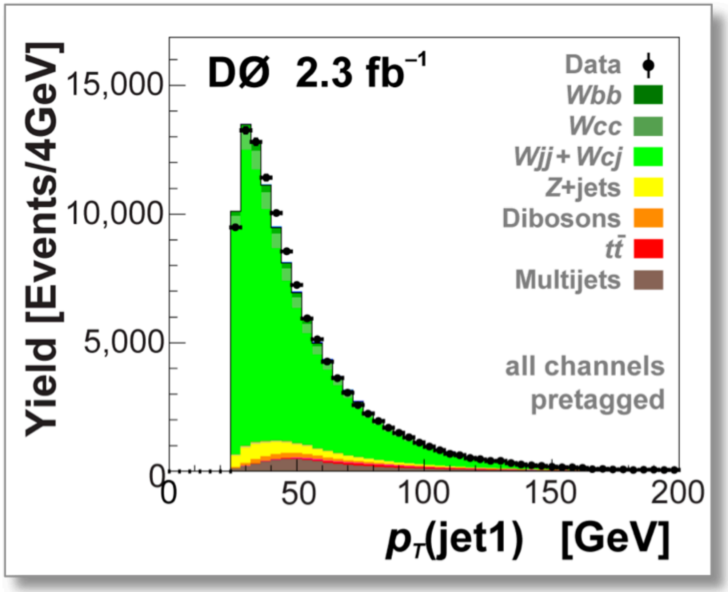


Event Yields Before *b*-Tagging

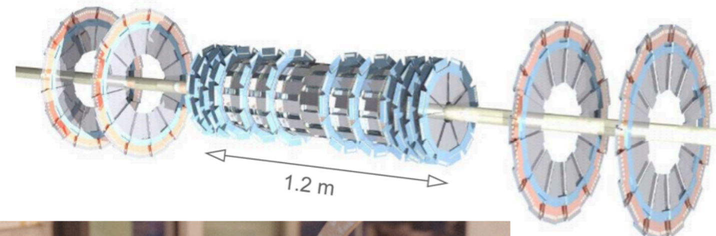
- Signal acceptances: *tb* = 5.9%, *tqb* = 5.6%
- **S:B ratio for *tb*+*tqb* = 1:260**
- Need to improve S:B to have a hope of seeing a signal → select only events with *b*-jets in them

Event Yields in 2.3 fb^{-1} of DØ Data	
e,μ, 2,3,4-jets, pretag	
tb + *tqb*	444
W+jets	98,444
Z+jets, dibosons	8,631
$t\bar{t}$ pairs	1,895
Multijets	5,798
Total background	114,777
Data	114,777

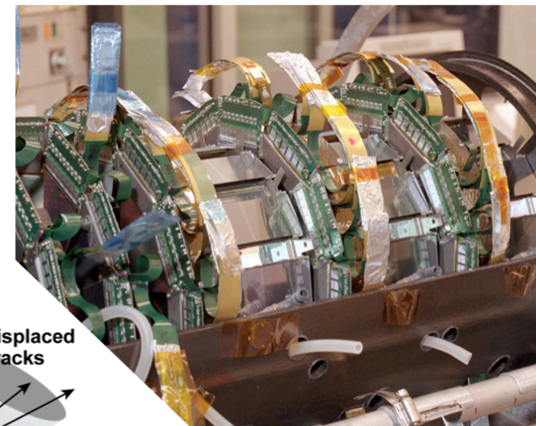


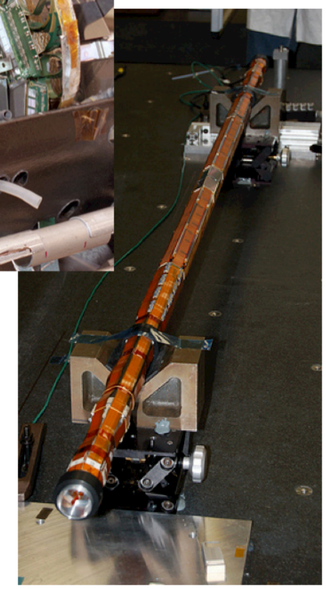


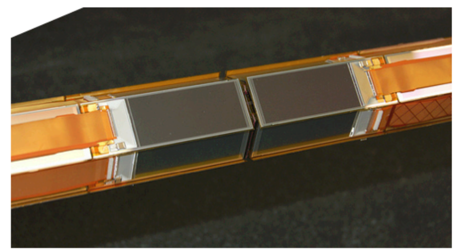
b-Jet Identification

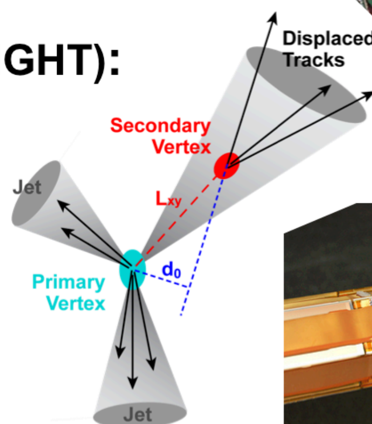


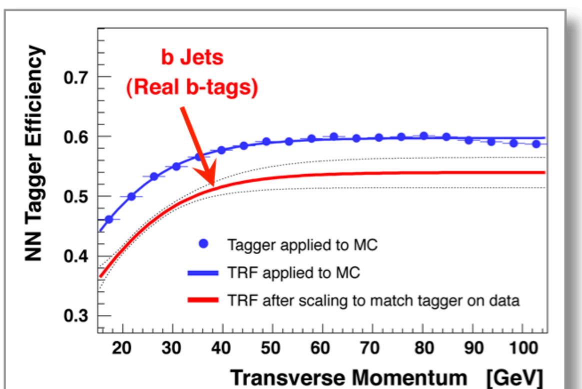
- Separate *b*-jets from light-quark and gluon jets to reject most *W*+jets background
- **DØ uses a neural network algorithm**
  - 9 input variables based on impact parameter and reconstructed vertex
- **Single-tag operating point (TIGHT):**
  - *b*-jet efficiency = 47%
  - *c*-jet efficiency = 10%
  - light-jet effic. = 0.5%
- **Double-tag uses LOOSE tags:**
  - 58% (*b*), 17% (*c*), 1.8% (*j*)

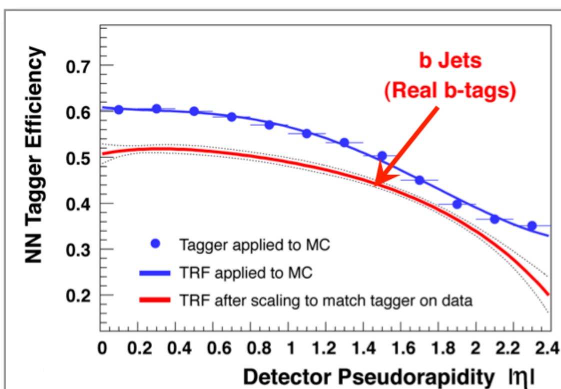


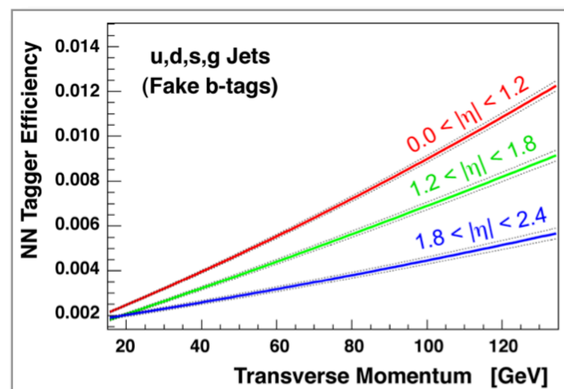




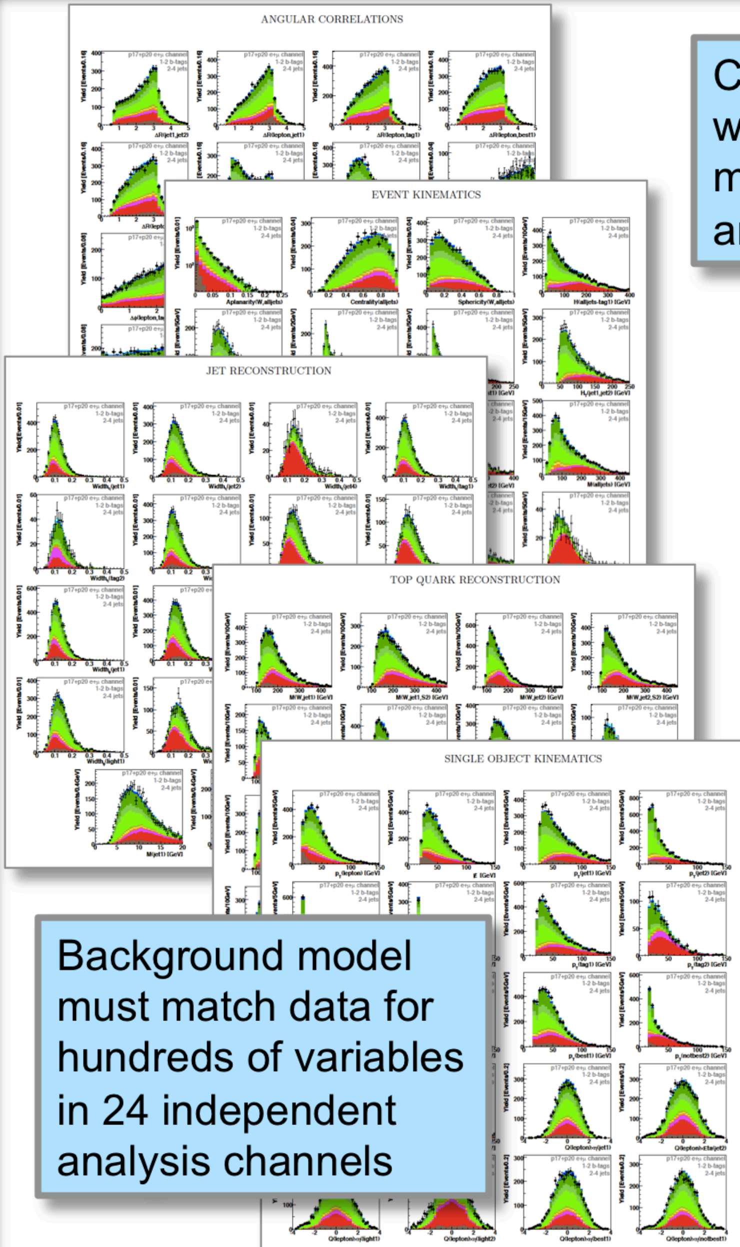




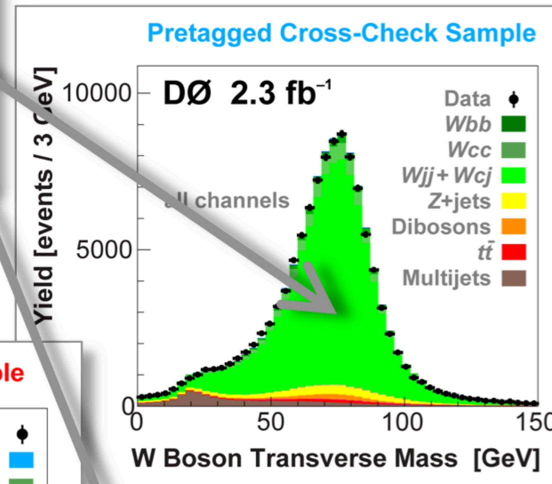


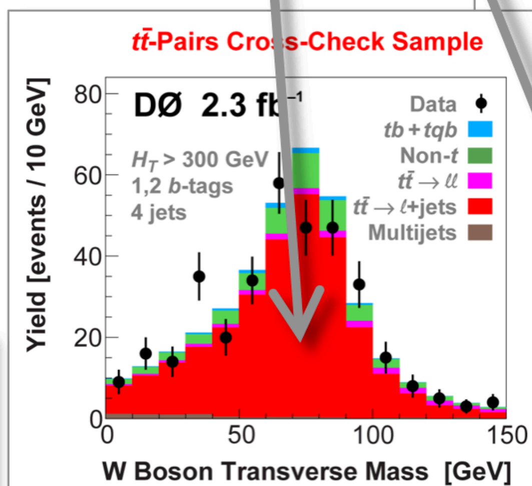


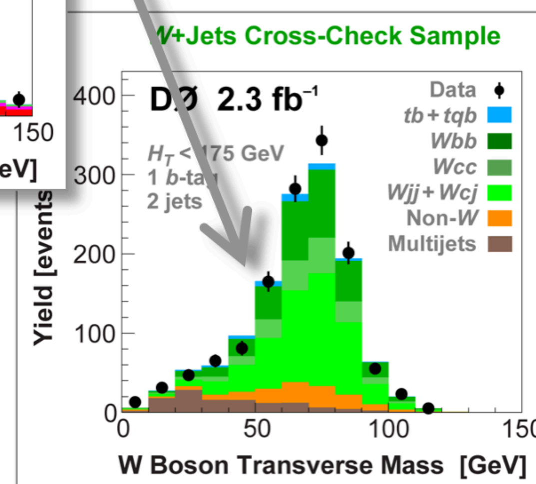
Checking the Background Model



Cross-check samples with mostly *W*+light-jets, mostly *W*+jets, and mostly top pairs





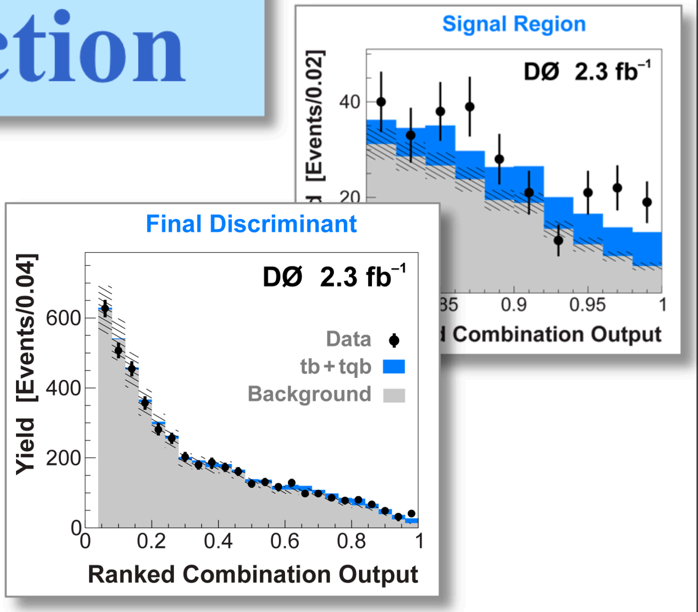


Background model must match data for hundreds of variables in 24 independent analysis channels

Measuring a Cross Section

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

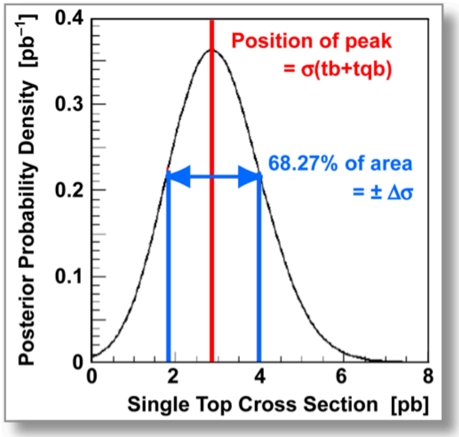
$$\begin{aligned} d &= \text{Predicted number of data events} \\ S &= \text{Predicted number of signal events} \\ B &= \text{Predicted number of background events} \\ \sigma &= \text{Cross section} \\ \mathcal{A} &= \text{Signal acceptance} \\ \mathcal{L} &= \text{Integrated luminosity} \\ a &= \text{Effective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



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

$$\begin{aligned} D &= \text{Observed number of data events} \\ \mathbf{b} &= \text{Vector of background components} \end{aligned}$$

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



- Nbkgds = 13 (*ttll, ttlj, Wbb, Wcc, Wcj, Wjj, Zbb, Zcc, Zjj, WW, WZ, ZZ,* multijets),
  Nbins = 24 analysis channels x 50 bins per channel = 1,200 bins
- Signal cross section prior is non-negative and flat
- 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

Pseudo-Data and Linearity Tests

- 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!
- Each pseudo-dataset is like one DØ experiment with 2.3 fb^{-1} of "data," up to 63 million pseudo-datasets per ensemble
- **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