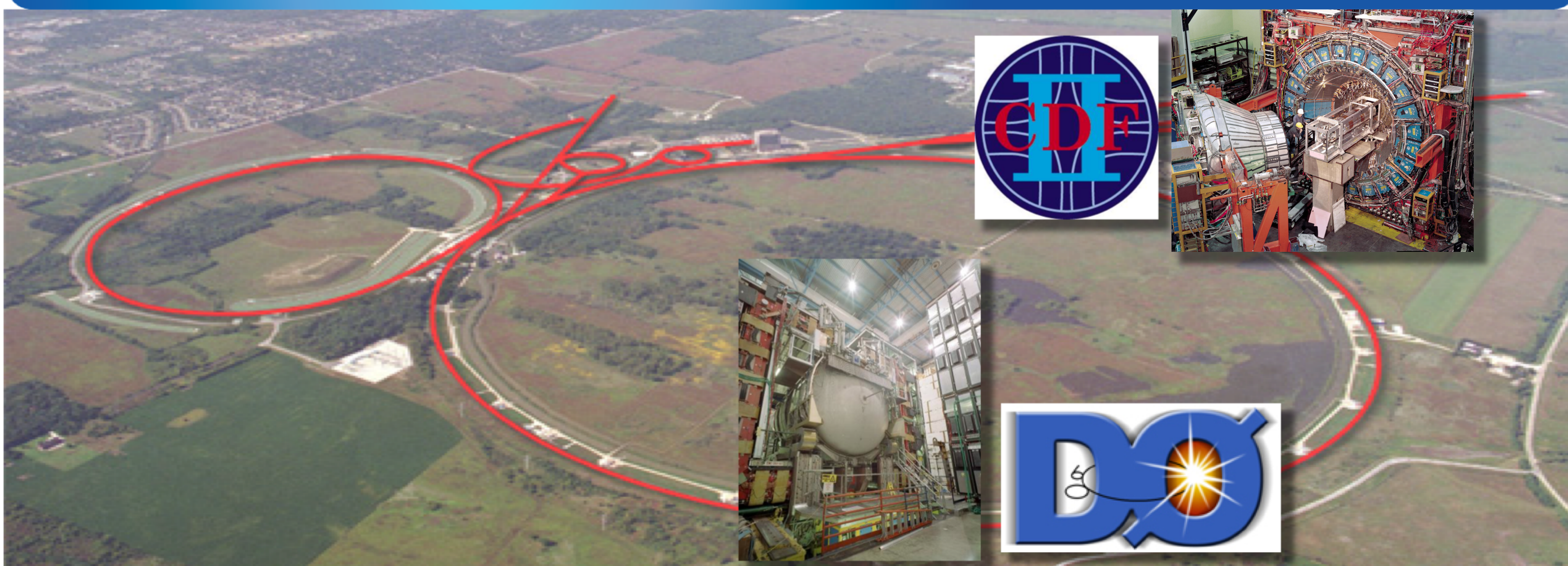
Observation of Single Top Quark Production at the Tevatron

Ann Heinson
University of California, Riverside
for the CDF and DØ Collaborations

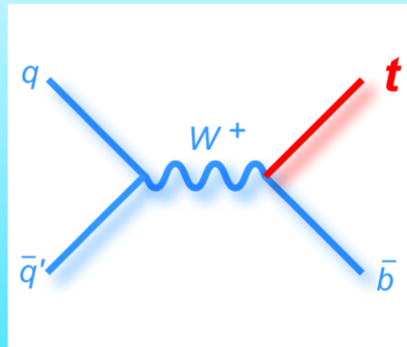
XXIth Rencontres de Blois: Windows on the Universe
Tuesday June 23, 2009



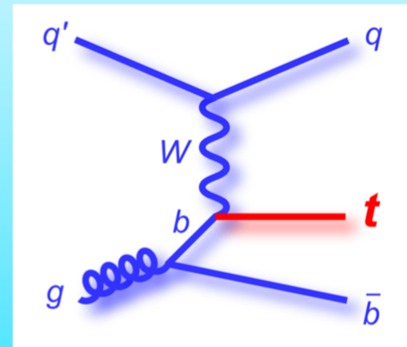
Single Top Quark Production

- Electroweak production of single top quarks is predicted to occur at about half the rate of strong production of top-antitop pairs

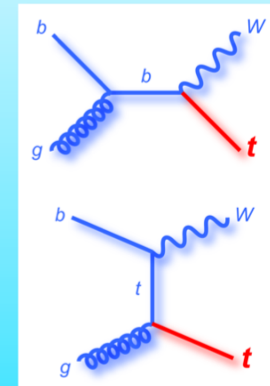
s-channel: "*tb*"



t-channel: "*tqb*"



"*tW*" production



m_{top}	s-channel	t-channel
175 GeV	σ_{NLO} = 0.88 ± 0.11 pb	σ_{NLO} = 1.98 ± 0.25 pb
170 GeV	$\sigma_{(N)NLO}$ = 1.12 ± 0.04 pb	$\sigma_{(N)NLO}$ = 2.34 ± 0.12 pb

Low rate at the Tevatron
(Important at the LHC)

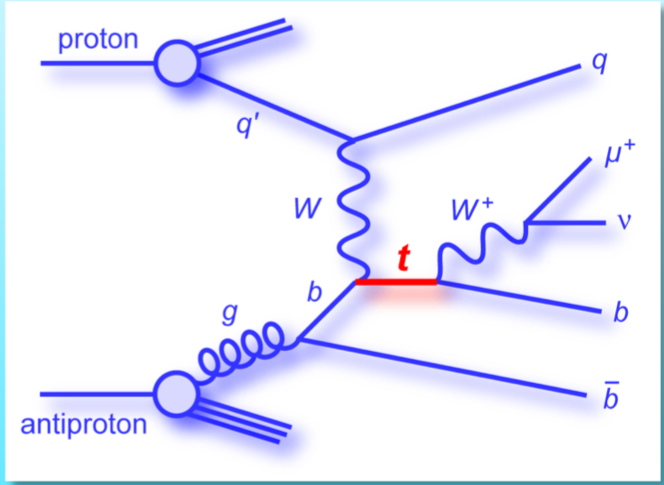
- Observation of this process allows us to measure the CKM matrix element $|V_{tb}|$
- Can also study the *Wtb* coupling and search for heavy resonance production
- Same final state particles as for Higgs boson production ($WH \rightarrow l\nu b\bar{b}$), development of
  - accurate background modeling
  - powerful signal-background separation

  essential for Higgs boson discovery

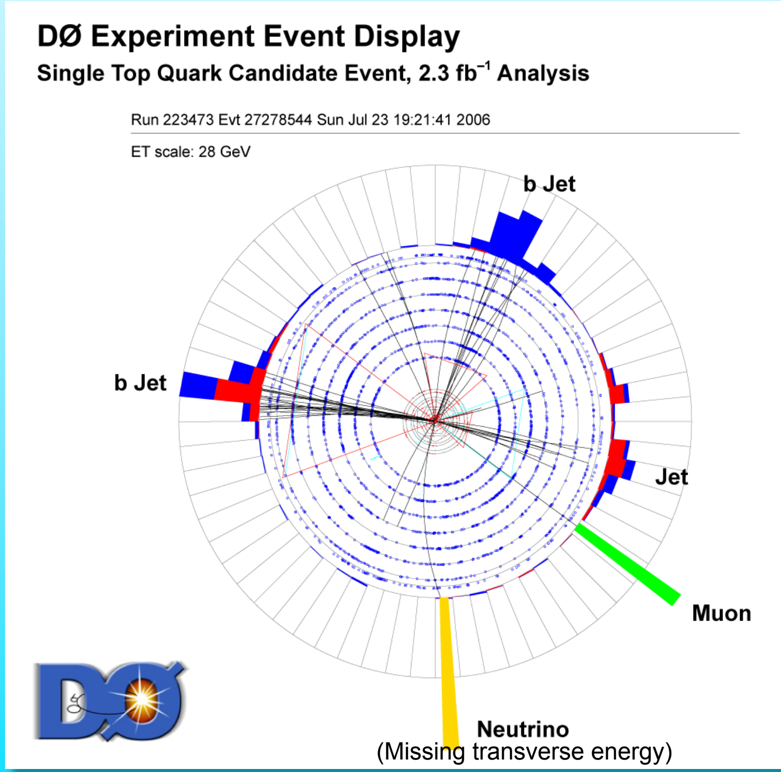
- Observation analyses performed at m_{top} = 175 GeV (CDF), 170 GeV (DØ), and assume SM *tb:tqb* ratio
- σ_{NLO} = Sullivan, PRD 70, 114012 (2004), $\sigma_{(N)NLO}$ = Kidonakis, PRD 74, 114012 (2006)

Why Didn't We See It Till Now?

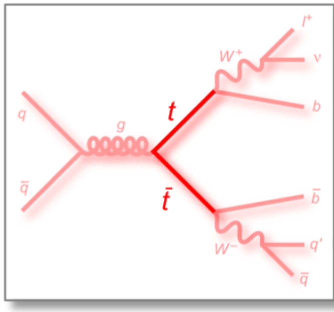
- Predicted ~10 years before the discovery of the top quark in pair production
  - t-channel: Willenbrock and Dicus, PRD 34, 155 (1986)
  - s-channel: Cortese and Petronzio, PLB 253, 494 (1991)
- Observed 14 years after the top quark discovery (CDF and DØ, 1995)
- Single top (*tb*+*tqb*) has nearly half 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

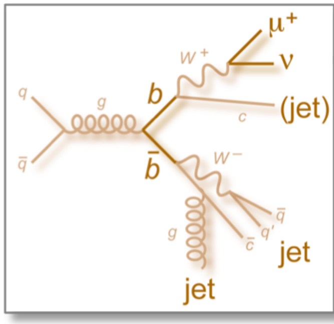


- Signal is one isolated high-p_T electron or muon and/or missing transverse energy from the *W* decay, 2, 3, or 4 jets, and 1 or 2 *b*-tags



Backgrounds and Event Yields

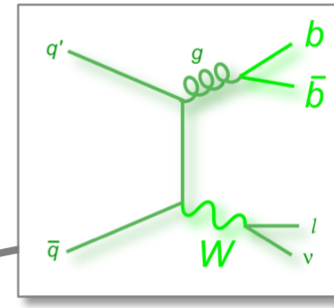


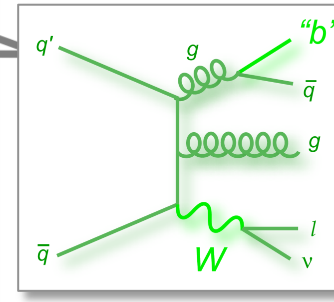


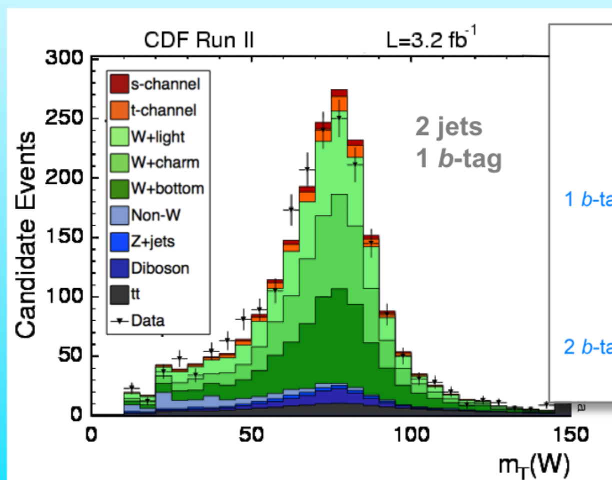
Single Top Observation – Event Yields

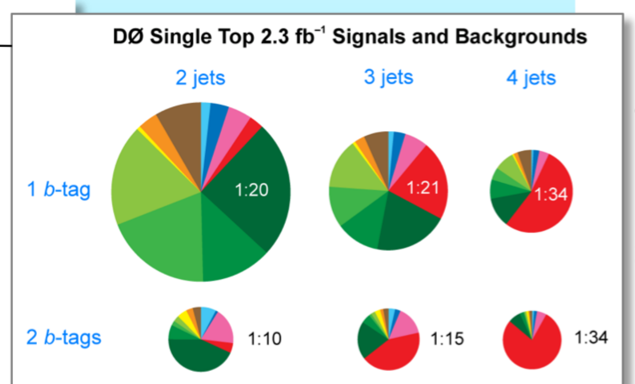
	DØ 2.3 fb^{-1}	CDF 3.2 fb^{-1}	CDF 2.1 fb^{-1}
	Lepton+$\not{E}_T$+jets / b-tagged		$\not{E}_T$+jets / b-tagged
tb + tqb signal *1,*2	223 ± 30	191 ± 28	64 ± 10
W+jets	2,647 ± 241	2,204 ± 542	304 ± 116 *4
Z+jets, dibosons	340 ± 61	171 ± 15	171 ± 54
$t\bar{t}$ pairs *1,*2, *3	1,142 ± 168	686 ± 99	185 ± 30
Multijets	300 ± 52	125 ± 50	679 ± 28 *5
Total prediction	4,652 ± 352	3,377 ± 505	1,403
Data	4,519	3,315	1,411

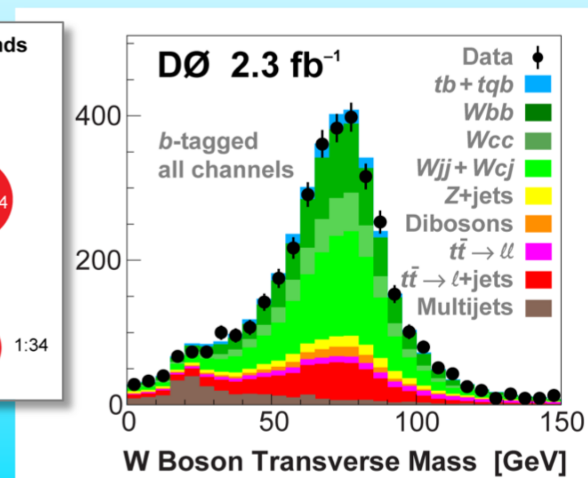
*1 DØ's tb+tqb signal and $t\bar{t}$ background use m_{top} = 170 GeV (and signal $\sigma_{(N)NLO}$)
*2 CDF's tb+tqb signal and $t\bar{t}$ background use m_{top} = 175 GeV (and signal σ_{NLO})
*3 DØ's analysis includes 4-jet events, so the $t\bar{t}$ yield is higher
*4 CDF's $\not{E}_T$+jets channel W+jets yield does not include Wjj where j = light jet
*5 CDF's $\not{E}_T$+jets channel Multijets yield includes Wjj events











Development of Analysis Strategy

Searches, upper limits

DØ:	PRD **63**, 031101	(2000)	0.09 fb^{-1}	Cuts search
DØ:	PLB **517**, 282	(2001)	0.09 fb^{-1}	Neural networks (28 variables)
CDF:	PRD **65**, 091102	(2002)	0.11 fb^{-1}	Cuts + likelihood fit to H_T
CDF:	PRD **69**, 052003	(2004)	0.11 fb^{-1}	Neural networks (18 variables) + likelihood fit
CDF:	PRD **71**, 012005	(2005)	0.16 fb^{-1}	Cuts + likelihood fit to $Q\times\eta$
DØ:	PLB **622**, 265	(2005)	0.23 fb^{-1}	Neural networks (25 variables) + Bayesian likelihoods
DØ:	PRD **75**, 092007	(2007)	0.23 fb^{-1}	Cuts, Neural networks (25 variables) + Bayes. l'hoods

>3σ Evidence

DØ:	PRL **98**, 181802	(2007)	0.9 fb^{-1}	Boosted decision trees (49 variables) , Bayesian NNs, Matrix elements, + Bayesian likelihoods
DØ:	PRD **78**, 012005	(2008)	0.9 fb^{-1}	Boosted decision trees (49 variables), Bayesian NNs, Matrix elements, + BLUE combination + Bayesian likelihoods
CDF:	PRL **101**, 252001	(2008)	2.2 fb^{-1}	Neural networks (18 variables), Likelihoods, Matrix elements, + NN-combination + Bayesian likelihoods

5σ Observation

DØ:	arXiv:0903.0850	(2009)	2.3 fb^{-1}	Boosted decision trees (64 variables), Bayesian NNs, Matrix elements, + Bayesian-NN combination + Bayesian likelihoods
CDF:	arXiv:0903.0885	(2009)	3.2 fb^{-1}	Boosted decision trees (22 variables), Neural networks (18 variables), Likelihoods, Matrix elements, + NN-combination + Bayesian likelihoods

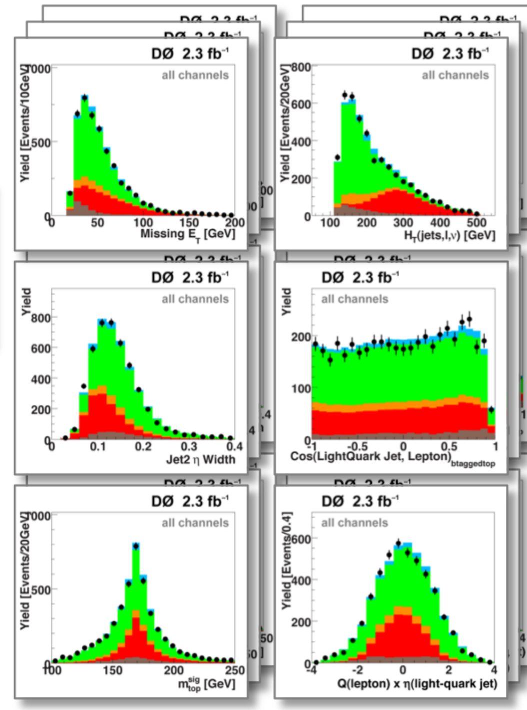
Analysis Strategy Visualized

Data from all possible triggers

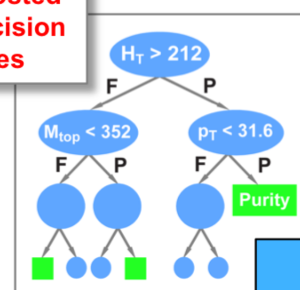
Monte Carlo for signal and backgrounds

Select signal-like events

Check data is reproduced in all variables



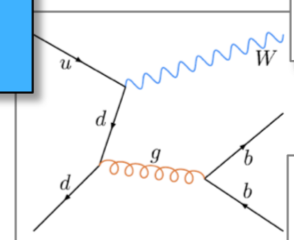
Boosted Decision Trees

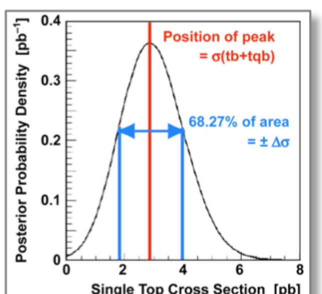


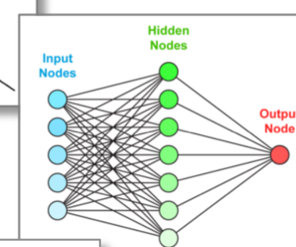
~100 million pseudo-datasets

Separate signal from background

Matrix Elements







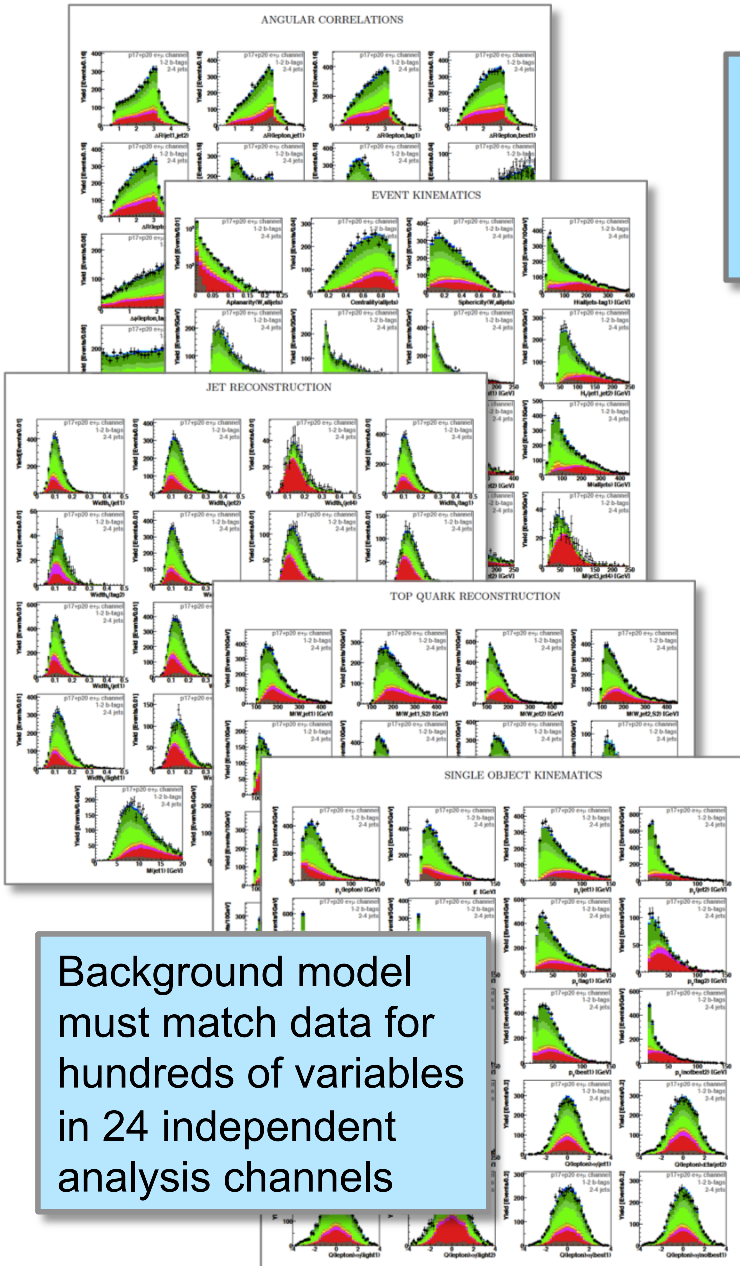
Neural Networks

Measure significance

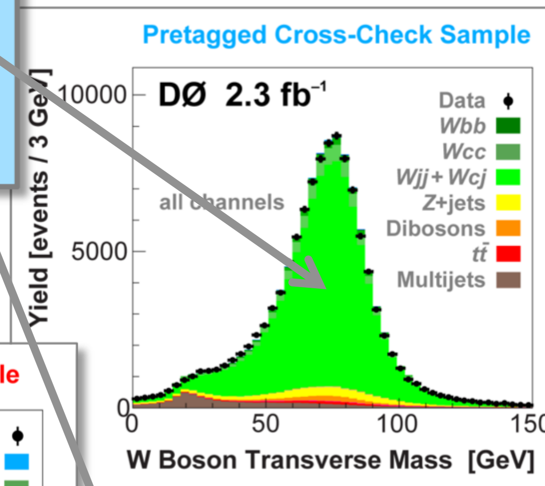
Cross section from Bayesian likelihood

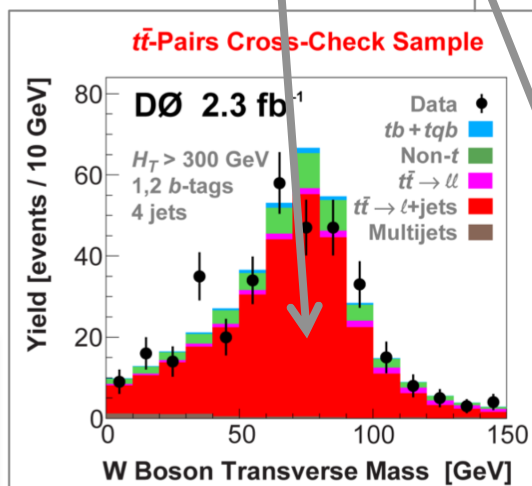
Combine results

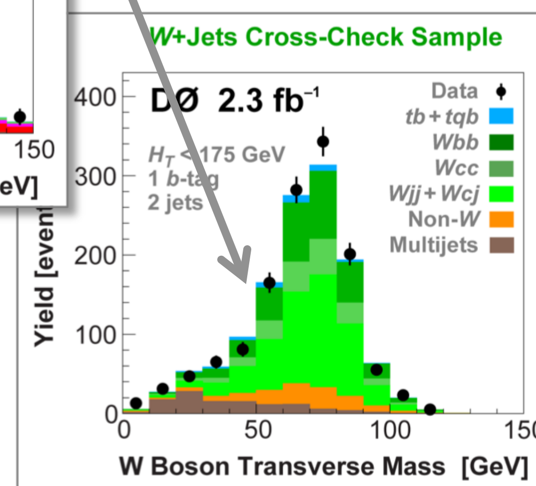
Checking the Background Model



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

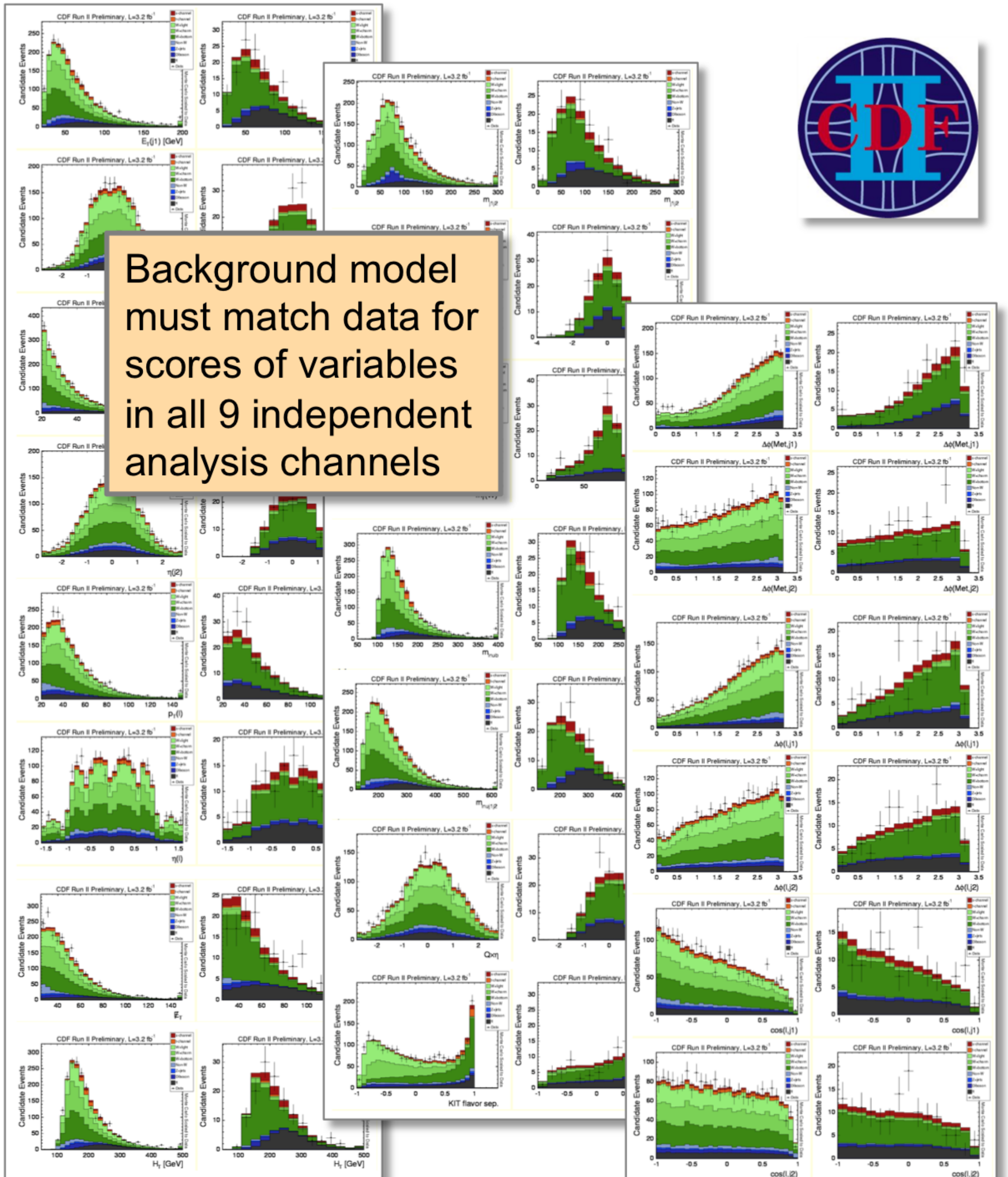




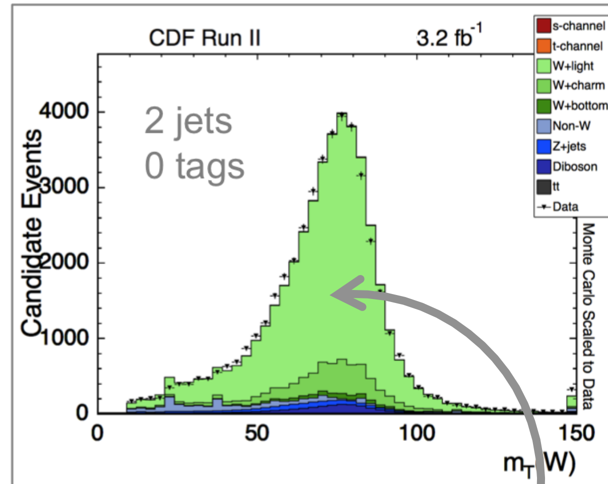


Background model must match data for hundreds of variables in 24 independent analysis channels

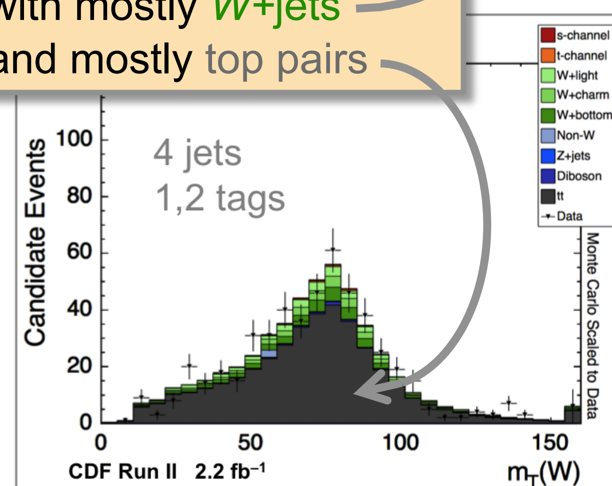
Checking the Background Model



Background model must match data for scores of variables in all 9 independent analysis channels



Also check all variables in cross-check samples with mostly *W*+jets and mostly top pairs

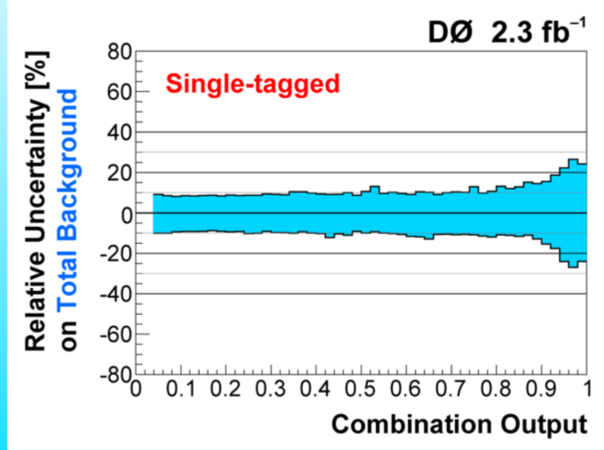


Systematic Uncertainties

Total error on *tb*+*tqb* cross section is ±22%

Statistics-only error is ±18%

So systematics contribute ±13%



Systematic Uncertaintes
CDF 3.2 fb–1

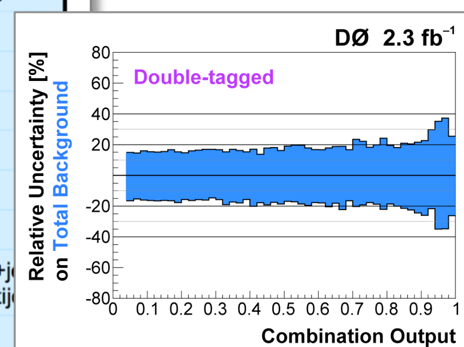
Systematic uncertainty	Rate	Shape
Jet energy scale	0...16%	X
Initial state radiation	0...11%	X
Final state radiation	0...15%	X
Parton distribution functions	2...3%	X
Monte Carlo generator	1...5%	
Event detection efficiency	0...9%	
Luminosity	6.0%	
Neural-net *b* tagger	N/A	X
Mistag model	N/A	X
Non-*W* model	N/A	X
Non-W normalization	40%	
Q^2 scale in Alpgen MC	N/A	X
Monte Carlo mismodeling	N/A	X
Wbb+Wcc normalization	30%	
Wcj normalization	30%	
Mistag normalization	17...29%	
tt-bar normalization	12%	
MC Statistics	bin-by-bin	

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 Uncerta
Ranked from Largest to Sm on Single Top Cross S
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*+j (30–54)% (multij
Diboson theory cross sections	5.8%
Alpgen *W*+jets shape corrections	shape only
Trigger	5%



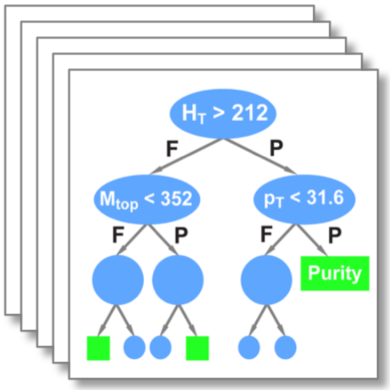
Separate Signal from Background

Boosted Decision Trees

DØ and CDF

Apply sequential cuts, keep events that fail

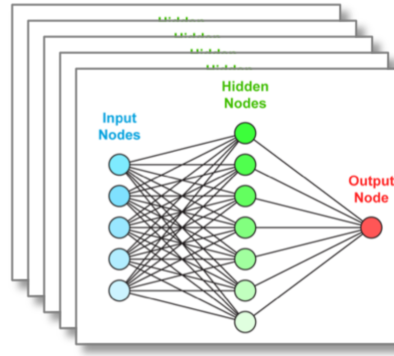
Boosting: averages over many trees, improves performance by ~20%



Neural Networks

CDF: one network for each analysis channel

DØ: average over 100 networks for each analysis channel ("Bayesian NNs")

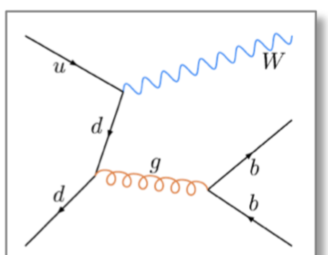


Matrix Elements

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

2 Jets		3 Jets	
$t\bar{b}$	$u\bar{d} \to t\bar{b}$	$t\bar{b}g$	$u\bar{d} \to t\bar{b}g$
tq	$ub \to td$ $\bar{d}\bar{b} \to t\bar{u}$	tqg	$ub \to tdg$ $\bar{d}\bar{b} \to t\bar{u}g$
		$tq\bar{b}$	$ug \to td\bar{b}$ $\bar{d}g \to t\bar{u}\bar{b}$
$Wb\bar{b}$	$u\bar{d} \to Wb\bar{b}$	$Wb\bar{b}g$	$u\bar{d} \to Wb\bar{b}g$
$Wc\bar{g}$	$\bar{s}g \to Wc\bar{g}$		
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	$gg \to ggg$		
$t\bar{t}$	$q\bar{q} \to t\bar{t} \to \ell^+ \nu b \ell^- \nu \bar{b}$		
$t\bar{t}$	$q\bar{q} \to t\bar{t} \to \ell^+ \nu b \bar{u} d \bar{b}$	$t\bar{t}$	$q\bar{q} \to t\bar{t} \to \ell^+ \nu b \bar{u} d \bar{b}$

DØ and CDF



Likelihoods

CDF only

One likelihood for each signal and background

Measure *tb+tqb* and s-channel *tb* separately

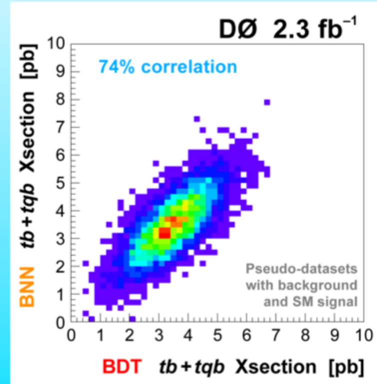
$$p_{ik} = \frac{f_{ij_ik}}{\sum_{m=1}^{5} f_{ij_im}},$$

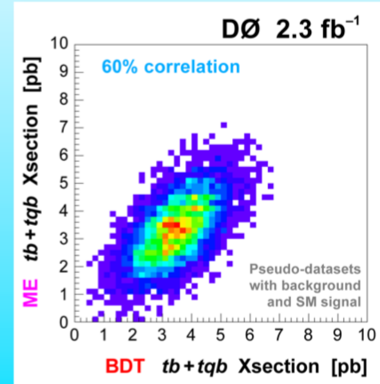
$$\mathcal{L}_k(\{x_i\}) = \frac{\prod_{i=1}^{n_{var}} p_{ik}}{\sum_{m=1}^{5} \prod_{i=1}^{n_{var}} p_{ik}}.$$

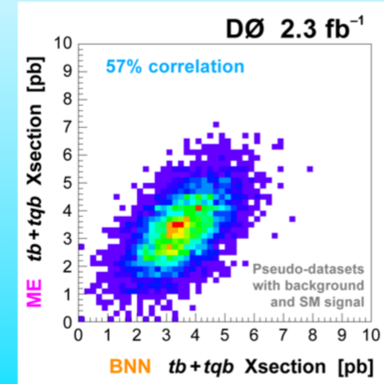
Combine Analyses

CDF have 5 analyses,
DØ have 3, for l+jets

Not 100% correlated,
so combining the
results can improve the
expected significance







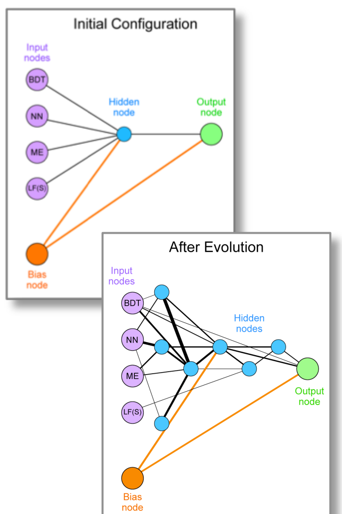
NEAT Neural Networks

"NEAT = Neuro-evolution of augmenting topologies"

Eight networks, one per analysis channel

Training optimized on expected significance

Improves expected significance by 13% over best individual analyses (BDT and NN)

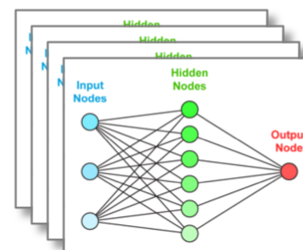


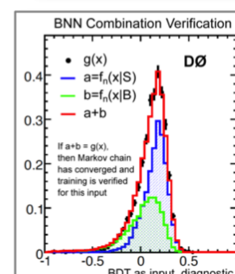
Bayesian Neural Networks

24 sets of networks, 1 per analysis channel

Markov Chain MC technique used with 500 networks, average over the last 100

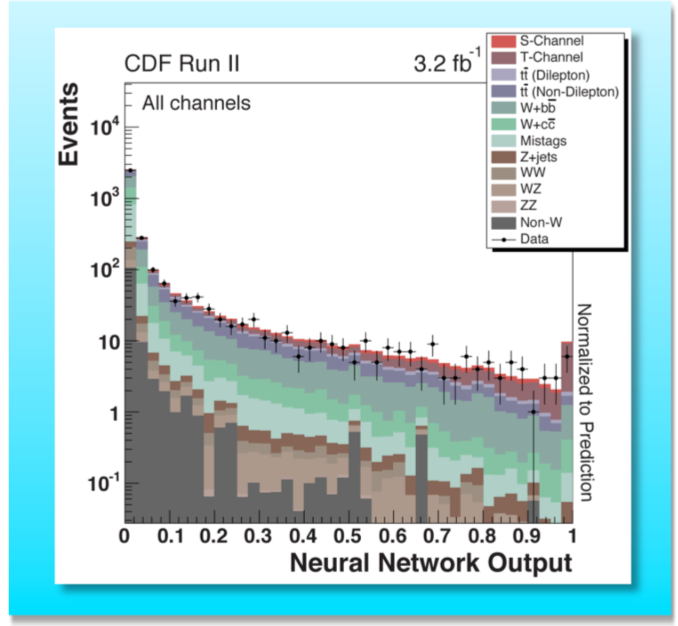
Improves expected significance by 4% over best individual analysis (BDT)

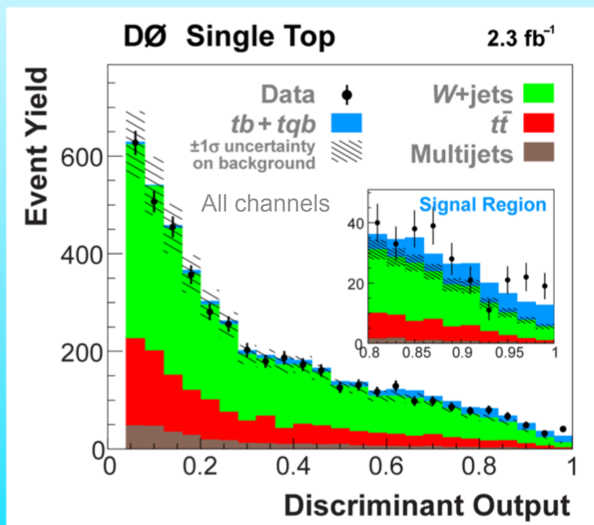


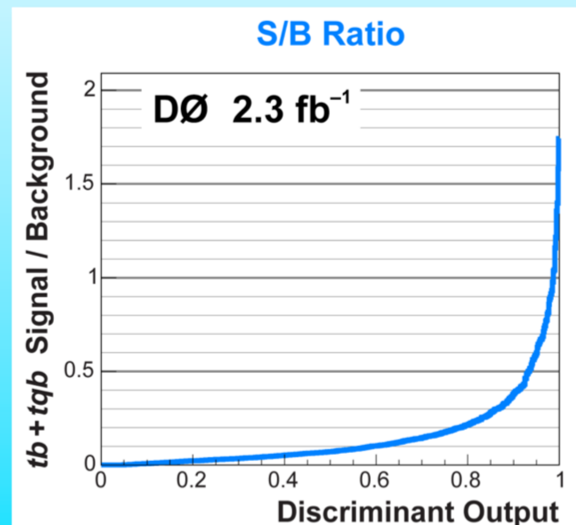


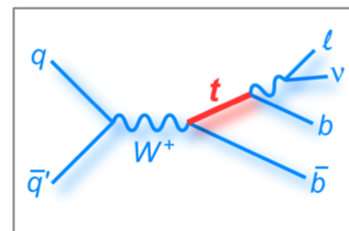
Single Top Observation – Results

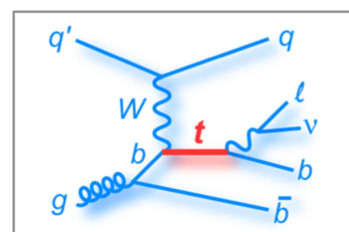
Single Top Cross Section	Signal Significance Expected	Signal Significance Observed
DØ 2.3 fb^{-1} arXiv:0903.0850 m_{top} = 170 GeV		
3.94 ± 0.88 pb	4.5 σ	5.0 σ
CDF 3.2 fb^{-1} arXiv:0903.0885 m_{top} = 175 GeV		
$2.3^{+0.6}_{-0.5}$ pb	>5.9 σ	5.0 σ





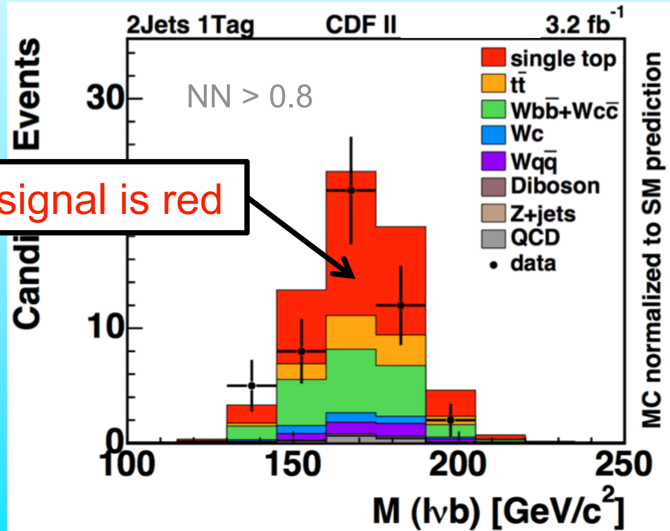


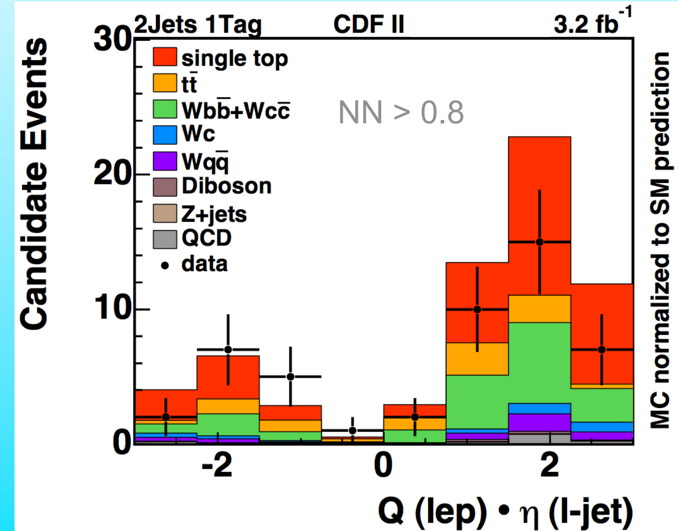




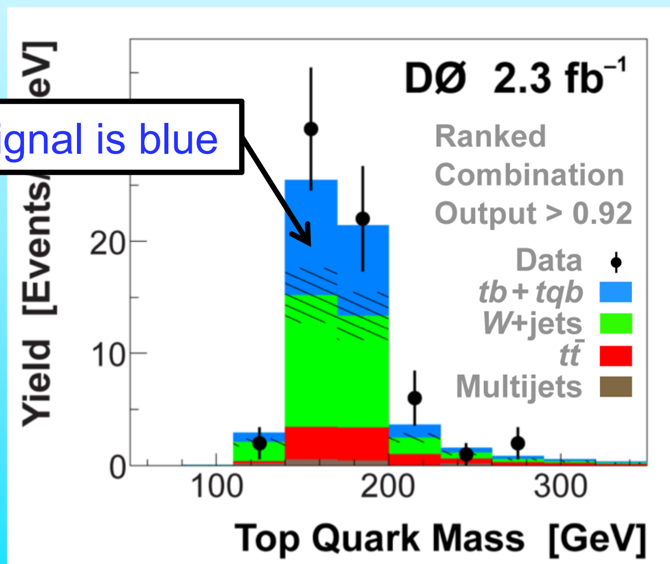
Single Top Signal Plots

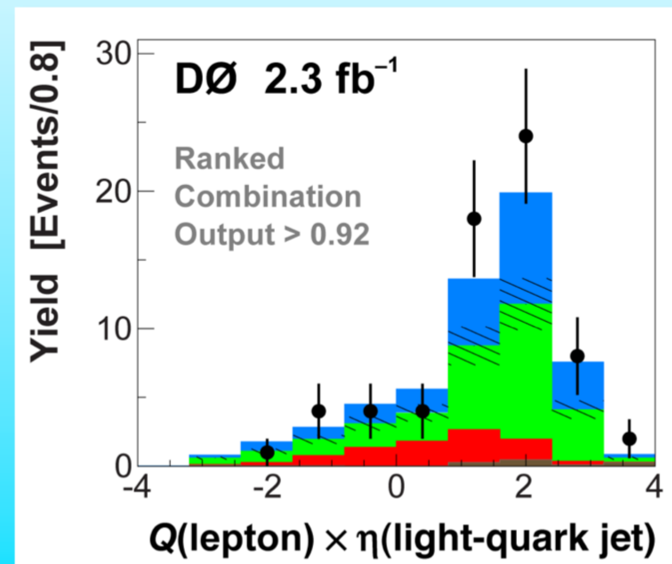
CDF's signal is red





DØ's signal is blue



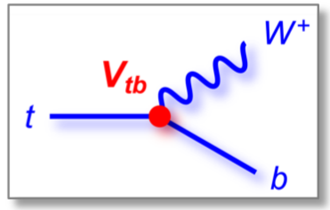


Signal normalized to measured cross section

CKM Matrix Element V_{tb}

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

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



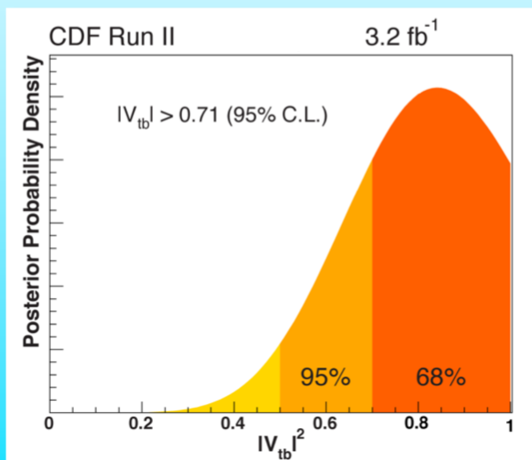
- $\sigma(tb, tqb) \propto |V_{tb}|^2$ ∴ calculate a posterior in $|V_{tb}|^2$
- Assume 1. SM top quark decay : $V_{td}^2 + V_{ts}^2 \ll V_{tb}^2$
  2. Pure *V–A* coupling with *CP* conservation
- No need to assume only three quark families or CKM matrix unitarity
- **$|V_{tb}| > 0.71$ (CDF), $|V_{tb}| > 0.78$ (DØ), at 95% CL**

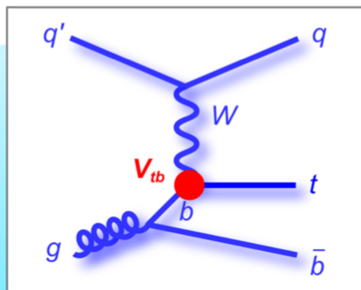
Additional Systematic Uncertainties for the $|V_{tb}|$ Measurement

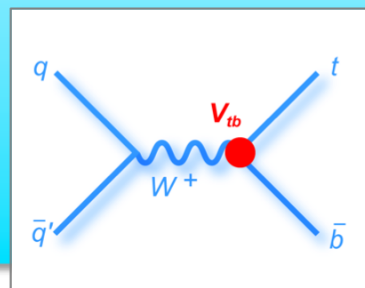
DØ 2.3 fb^{-1}

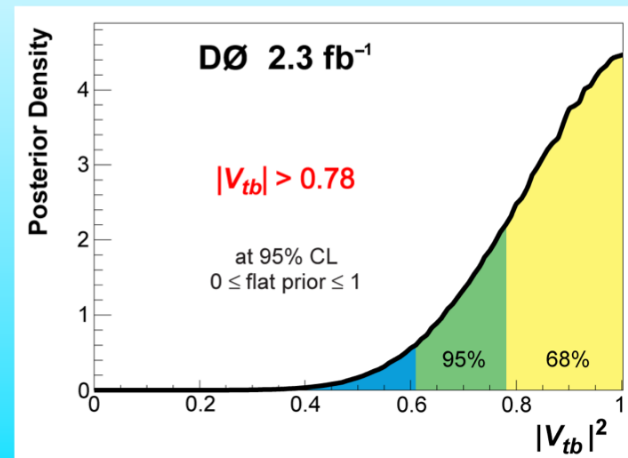
For the *tb+tqb* theory cross section	
Top quark mass	4.2%
Parton distribution functions	3.0%
Factorization scale	2.4%
Strong coupling α_s	0.5%

NB CDF and DØ use different theory cross sections and top quark mass points
Combination of $|V_{tb}|$ results is in progress





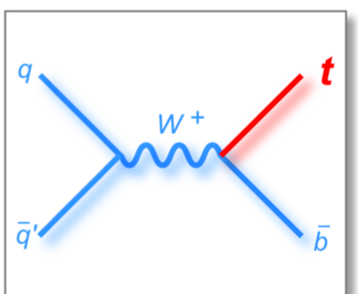


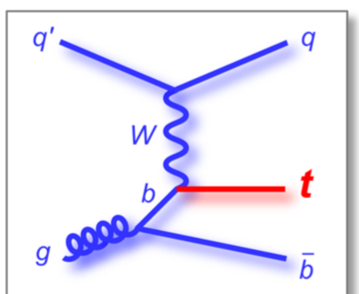


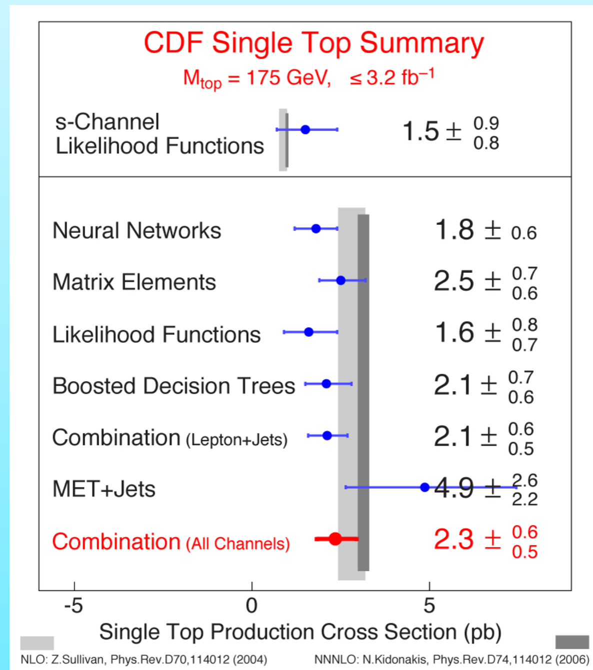
Single Top Observation Summary

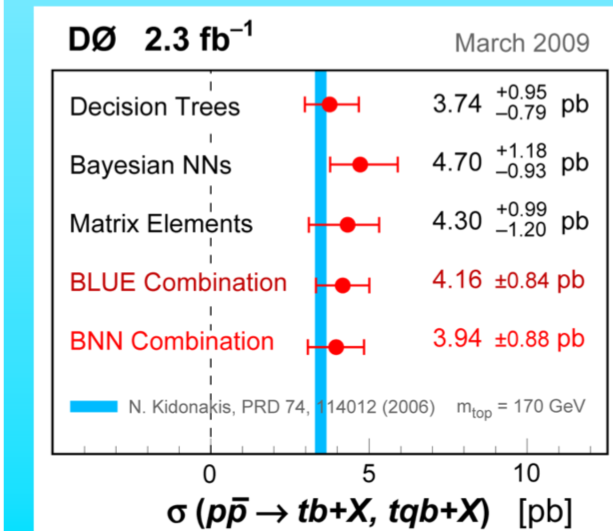
- Challenging measurements
  – small signal hidden in large complex background
- It took 14 years since top quark pair observation to collect enough data and develop the background models and analysis techniques to be able to achieve this

Single Top Cross Section	Signal Significance Expected	Signal Significance Observed	CKM Matrix Element V_{tb}
March 2009	**DØ (2.3 fb^{-1})**	arXiv:0903.0850	(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
March 2009	**CDF (3.2 fb^{-1})**	arXiv:0903.0885	(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









- 5σ observation of single top quark production !