Observation of Single Top Quark Production at DØ

Liang Li
University of California at Riverside

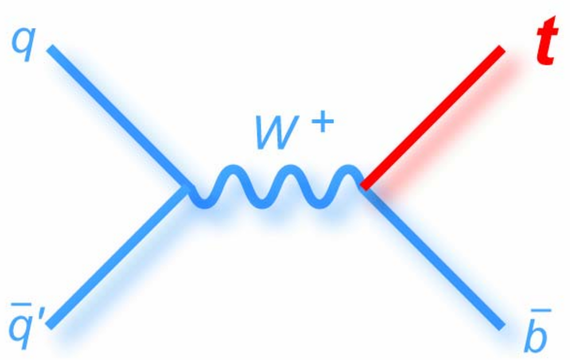
On Behalf of the D0 collaboration

Single Top Quark Production

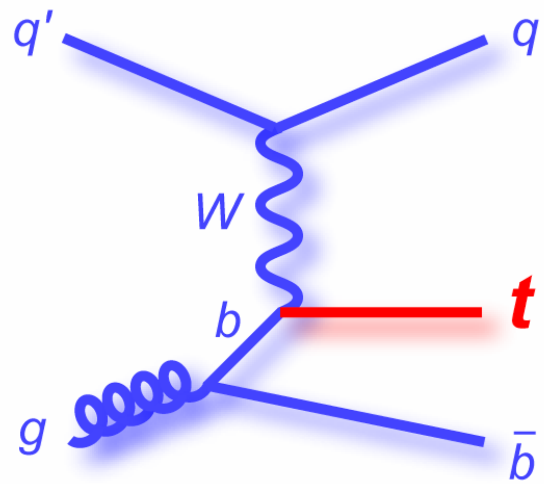
Single top electroweak interaction

- **Study Wtb coupling**
  - Direct measurement of the $|V_{tb}|$ CKM matrix element
  - Test of CKM unitarity
  - Anomalous *Wtb* couplings
- **New physics, example:**
  - s-channel is sensitive to W', H^+
  - t-channel is sensitive to FCNC
  - 4^{th} quark generation?
- **Study top properties:**
  - Polarization, decay width, lifetime, CP violation …
- **Experimental:**
  - Test of advanced searching techniques
  - Background study helps new physics searches, e.g. SM Higgs
  - Same final state as WH

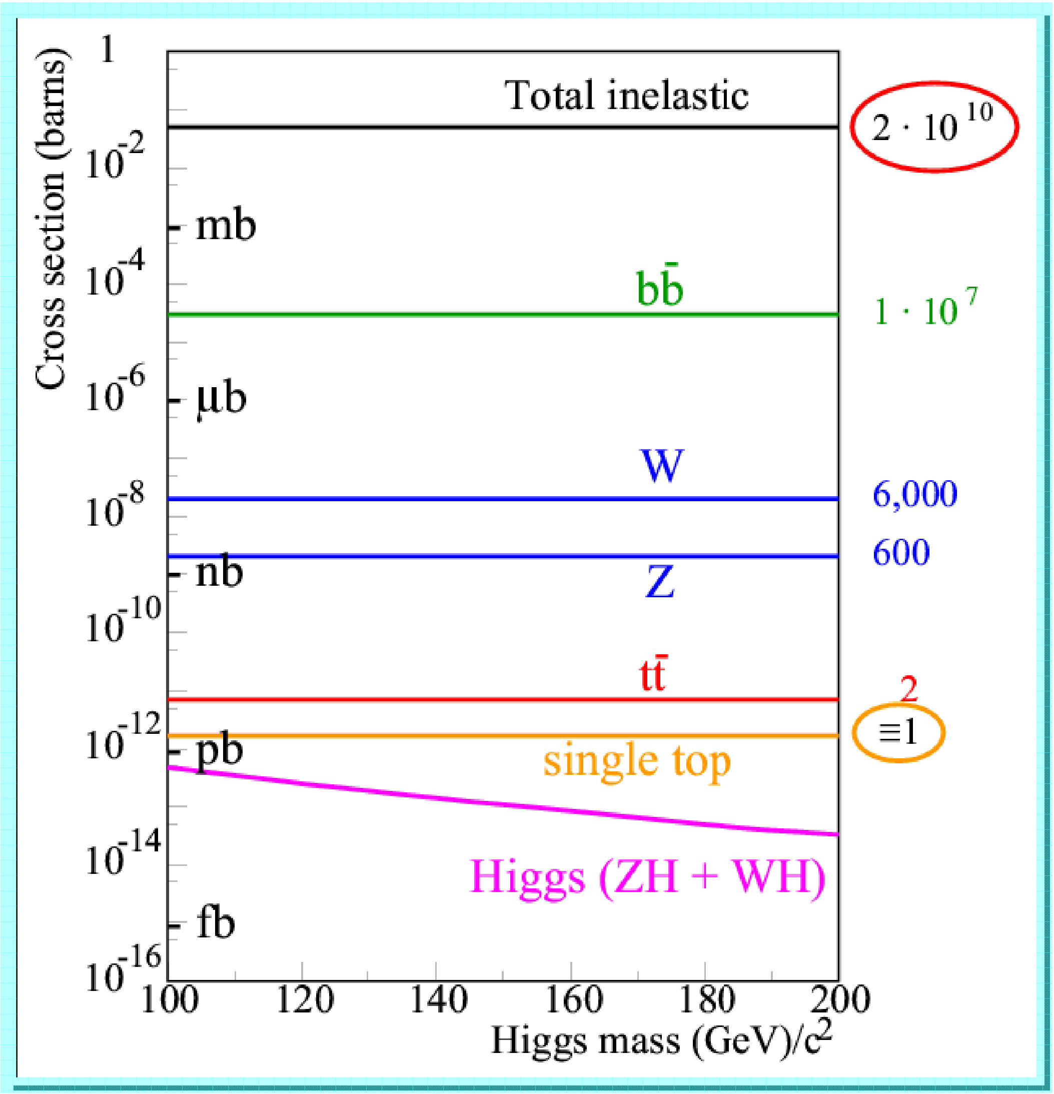
s-channel σ_{SM}=1.12±0.05 pb

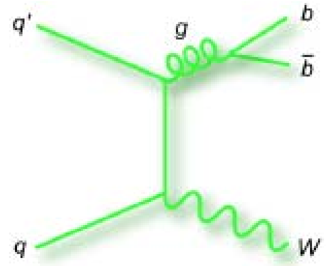


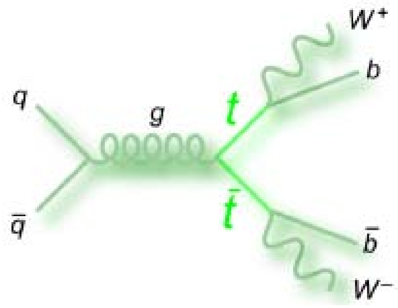
t-channel σ_{SM}=2.34±0.13 pb

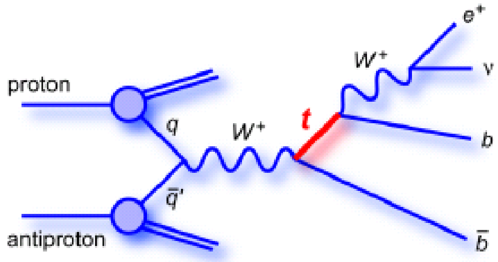


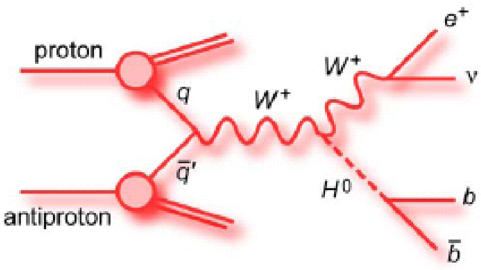
Experimental Challenge



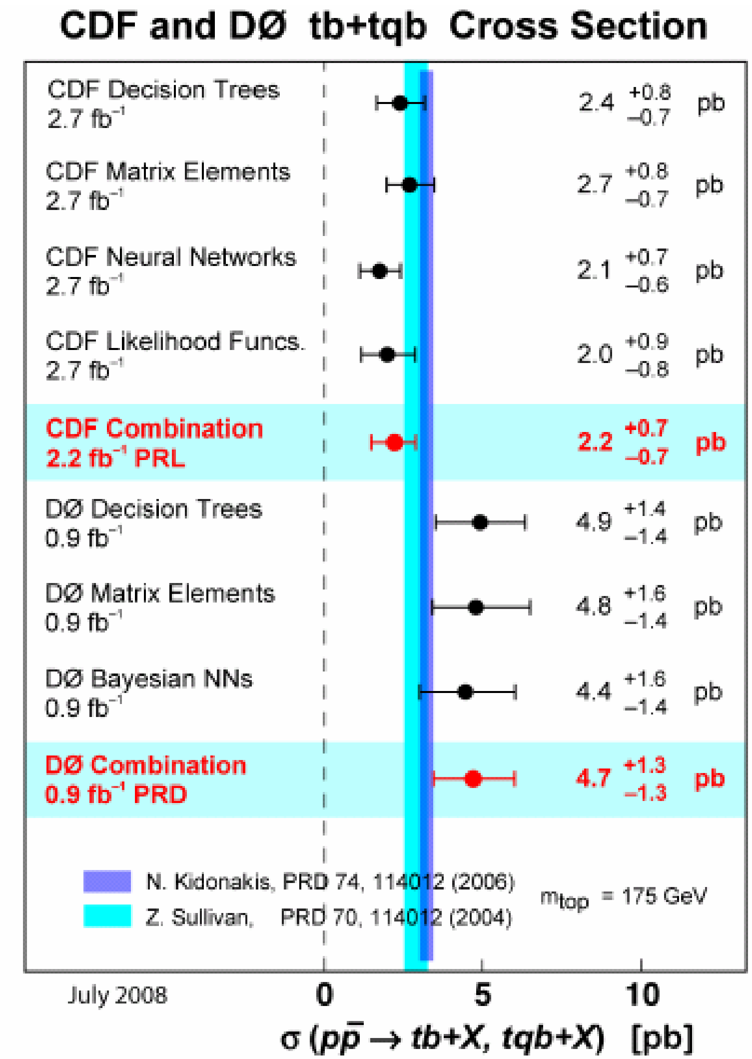








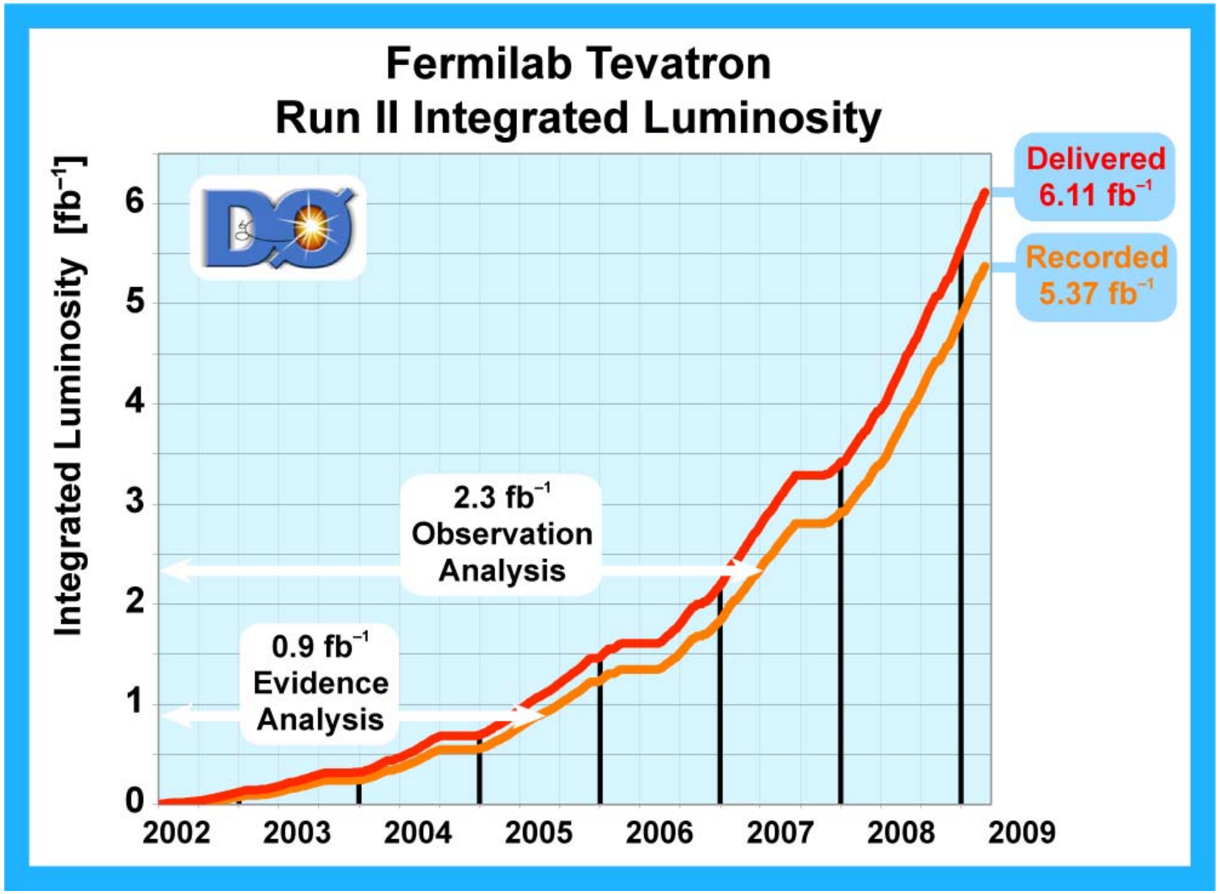
Previous Results: Evidence for Single Top Production



Single Top Cross Section	Signal Significance Expected	Signal Significance Observed	CKM Matrix Element V_{tb}
December 2006 DØ (0.9 fb^{-1})			PRL 98, 181802 (2007)
4.7 ± 1.3 pb	2.3 σ	3.6 σ	$\|V_{tb}f_1^L\| = 1.31^{+0.25}_{-0.21}$ $\|V_{tb}\| > 0.68$ at 95% CL
September 2008 CDF (2.2 fb^{-1})			PRL 101, 252001 (2008)
2.2 ± 0.7 pb	4.9 σ	3.7 σ	$\|V_{tb}f_1^L\| = 0.88^{+0.13}_{-0.12}$ $\|V_{tb}\| > 0.66$ at 95% CL

New chapter: march 4, 2009 both D0 and CDF independently present first observation of single top, 14 years after discovery of top pair production

Dataset



Tevatron performs very well, many thanks!

- **2.3 fb^{-1} for the observation analysis**
  - **1.1 fb^{-1} Run IIa dataset, 1.2 fb^{-1} Run IIb dataset**

Analysis Flow

Event Selection

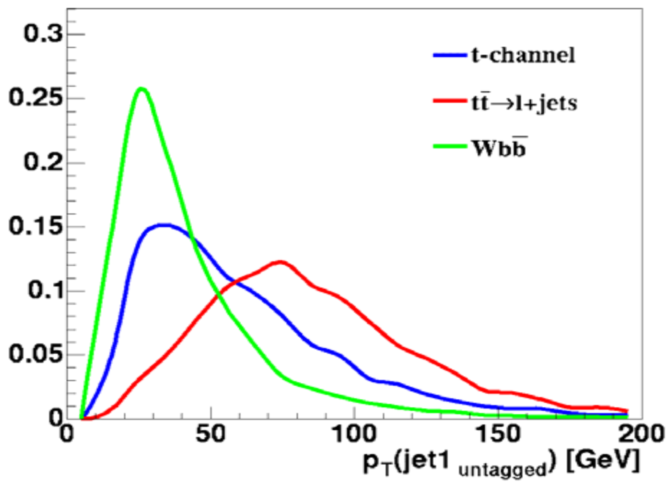
- Select W-like events
- Remove background-like events
- Apply b-tagging and maximize signal acceptance

Separate Signal from Background

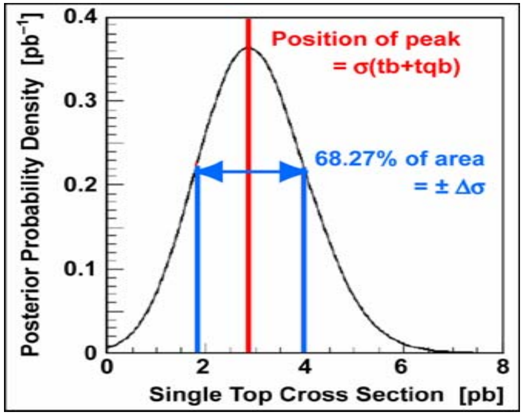
- Compare data with Monte Carlo
- Find discriminating variables
- Multivariate analyses

Determine Cross Section

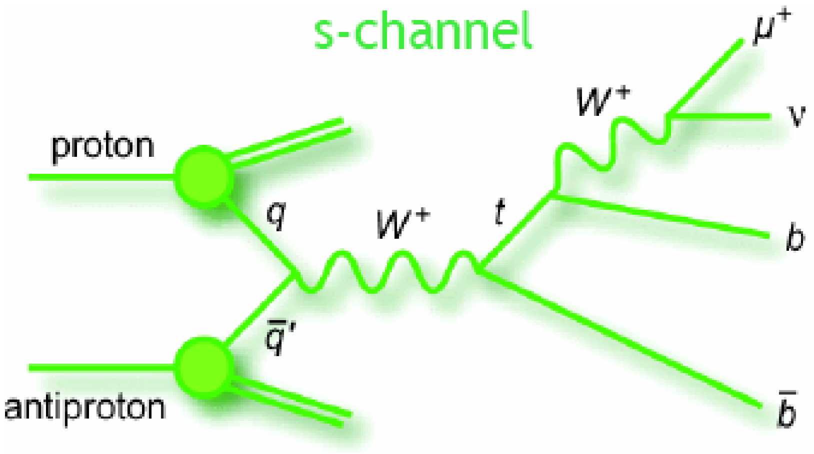
- Build binned likelihood
- Use shape information
- Bayesian approach

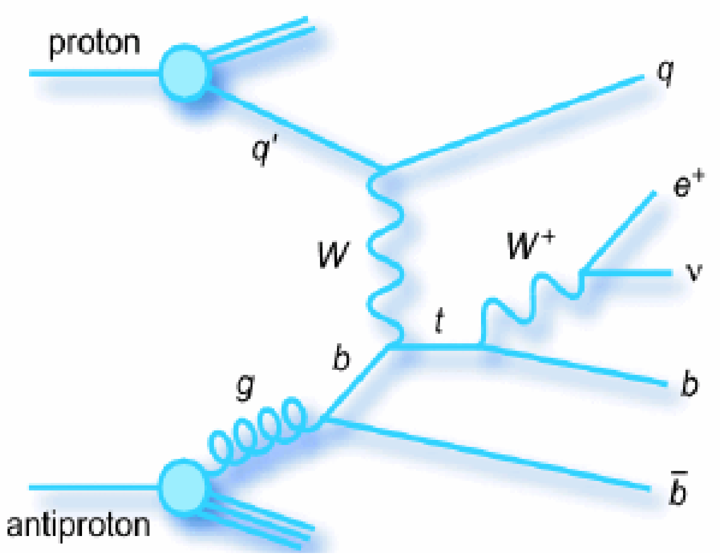


$$L = \frac{P(S)}{P(S) + P(B)}$$



Event Selection





Event Topology:

- **High energy isolated lepton (*e* or *mu* from *W*)**
- **Missing E_T (ν from *W*)**
- **One b-quark jet (from *t*)**
- **A light flavor jet and/or another b-jet (One or two b-tagged jets)**

Analysis done in 24 separate channels

- **Run IIa, Run IIb**
- **Two Lepton flavors: electron, muon**
- **Three jet bins: 2 jet, 3 jet, 4 jet**
- **Two tag bins: 1 b-tag, 2 b-tag**

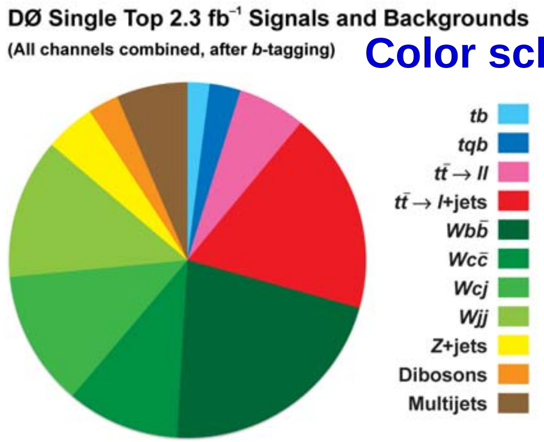
Signal and Background Modeling

Signal:

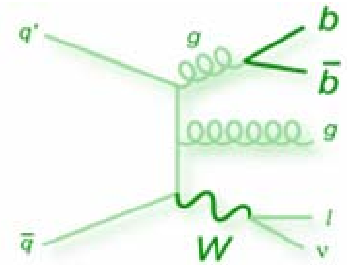
- CompHEP-SINGLETOP
- Distributions agree well with ZTOP & MCFM (NLO)

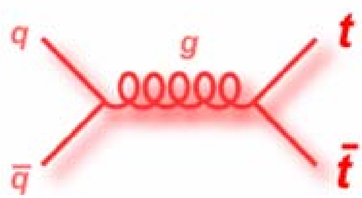
Background:

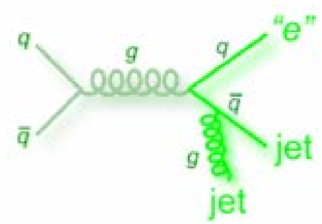
- W+jets production
  - Estimated from data & MC
  - Distribution shapes from ALPGEN
  - Normalization, Wcc and Wbb factor from data
- Top pair production
  - ALPGEN+PYTHIA
  - Normalized to NNLO cross section
- Multijet events
  - Misidentified lepton
  - Estimated from data



Color scheme





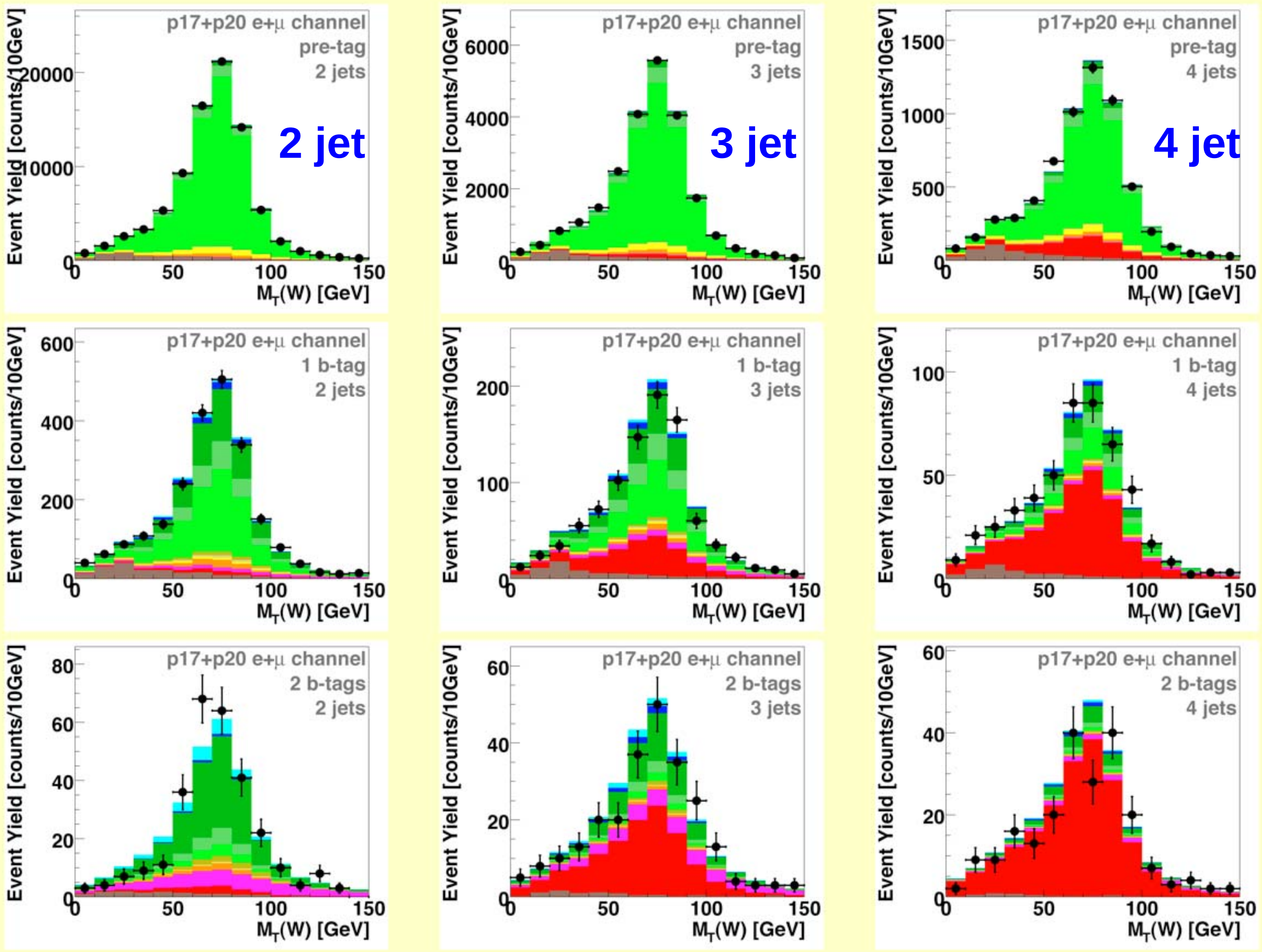


Data/MC Comparison

before b-tagging

1 b-tag

2 b-tags



Event Yields and Systematics

Event Yields in 2.3 fb^{-1} of DØ Data

Electron + muon, 1 tag + 2 tags combined

Source	2 jets	3 jets	4 jets
s-channel *tb*	62 ± 9	24 ± 4	7 ± 2
t-channel *tqb*	77 ± 10	39 ± 6	14 ± 3
$W+b\bar{b}$	678 ± 104	254 ± 39	73 ± 11
$W+c\bar{c}$	303 ± 48	130 ± 21	42 ± 7
$W+cj$	435 ± 27	113 ± 7	24 ± 2
$W+jj$	413 ± 26	140 ± 9	41 ± 3
$Z+$jets	141 ± 33	54 ± 14	17 ± 5
Dibosons	89 ± 11	32 ± 5	9 ± 2
$t\bar{t} \to \ell\ell$	149 ± 23	105 ± 16	32 ± 6
$t\bar{t} \to \ell+$jets	72 ± 13	331 ± 51	452 ± 66
Multijets	196 ± 50	73 ± 17	30 ± 6
Total prediction	2,615 ± 192	1,294 ± 107	742 ± 80
Data	2,579	1,216	724

Expected number of signal events less than background uncertainties

Must use multivariate discriminant to separate signal from background

Signal acceptance
s-channel, tb: 3.7 ± 0.5 %
t-channel, tqb: 2.5 ± 0.3 %

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%

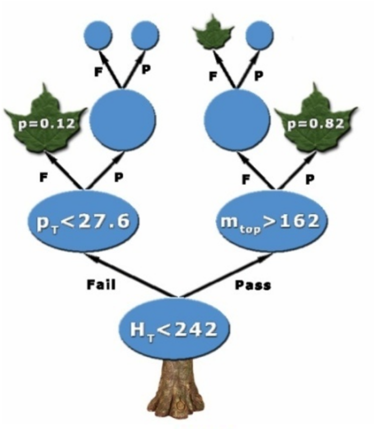
Most systematic uncertainties apply only to normalization, except jet energy scale and b-tagging which affect shapes

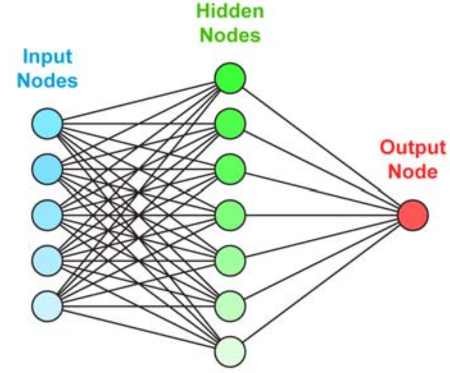
Cross section uncertainties are dominated by the statistical uncertainty

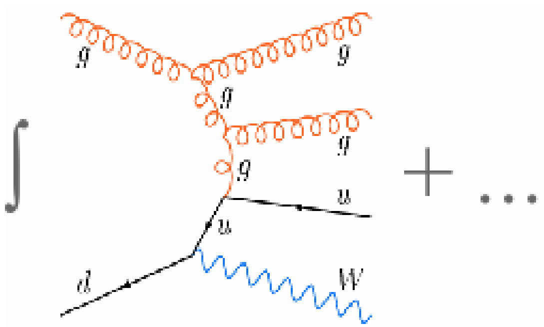
Multivariate Analyses

Three methods – "Blind Analyses": optimizing on expected sensitivity

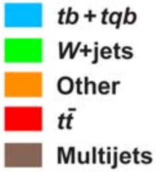
- **Boosted Decision Trees (BDT)**
  - Recover events that fail in cut-based analysis
  - Common object and event kinematics, angular correlations, jet reconstruction and top quark reconstruction variables
  - Use highest ranked common 64 variables
- **Bayesian Neural Network (BNN)**
  - Average over many neural networks, improving performance
  - Uses best 18-28 variables
- **Matrix Element (ME)**
  - Use Feynman diagrams to compute event probability density for signal and background (2jet: tb, tq, tt, WW, WZ, ggg, wbb, wcg, wgg; 3jet: tbg, tqb, tqg, wbbg, Wugg)
  - Split sample H_T > 175 GeV improves performance
- **Serve also as cross check of each other**

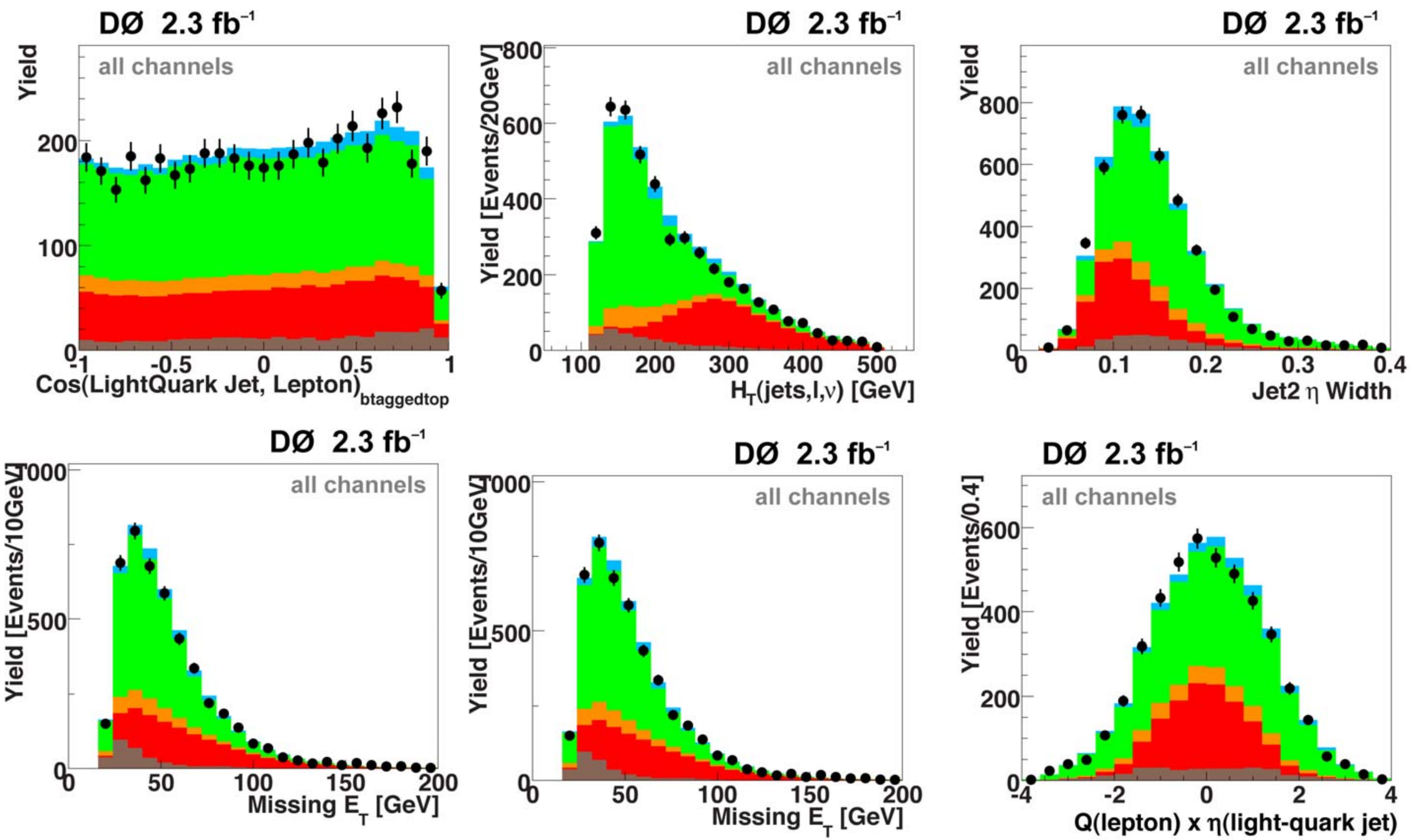




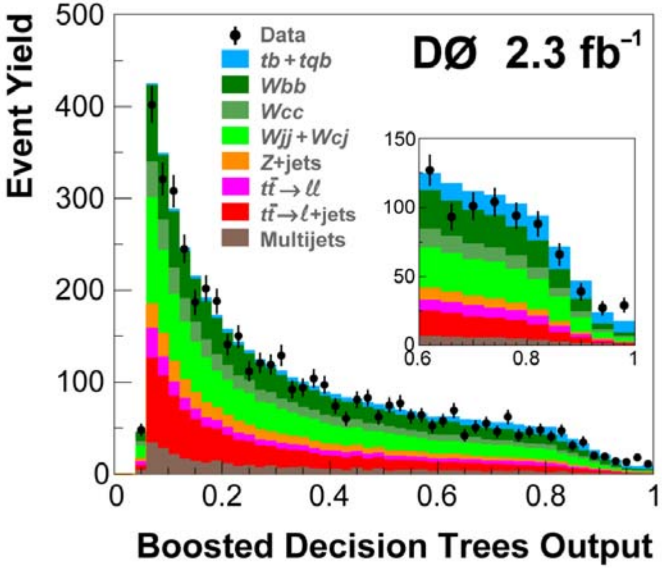


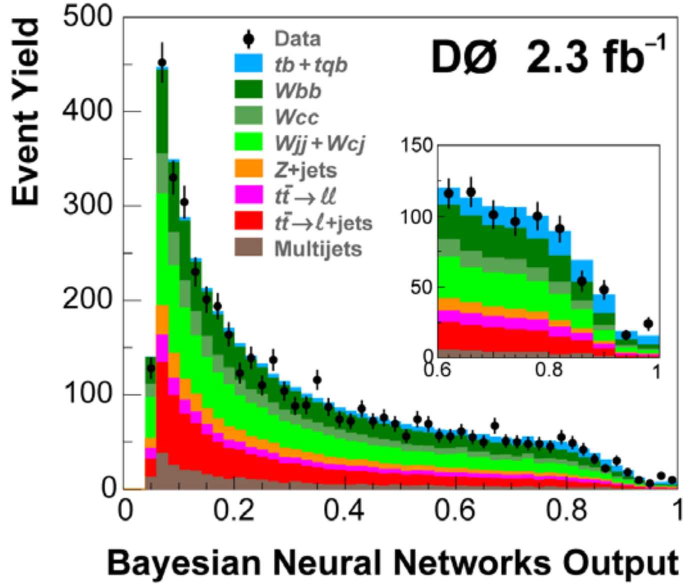
Discriminating Variables



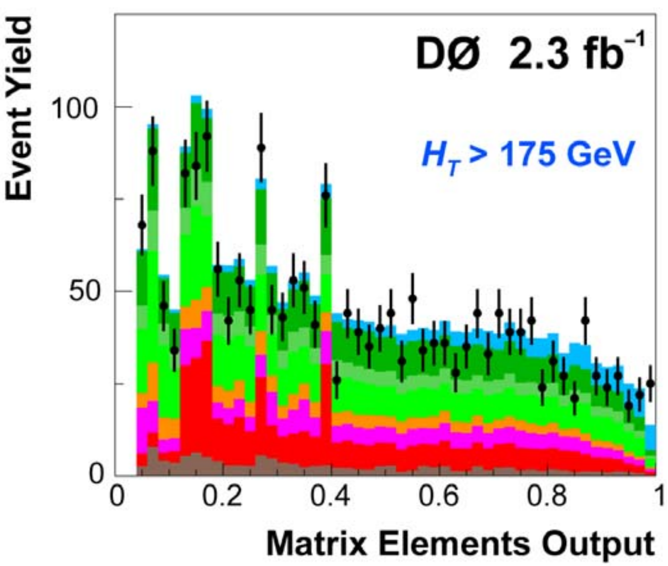


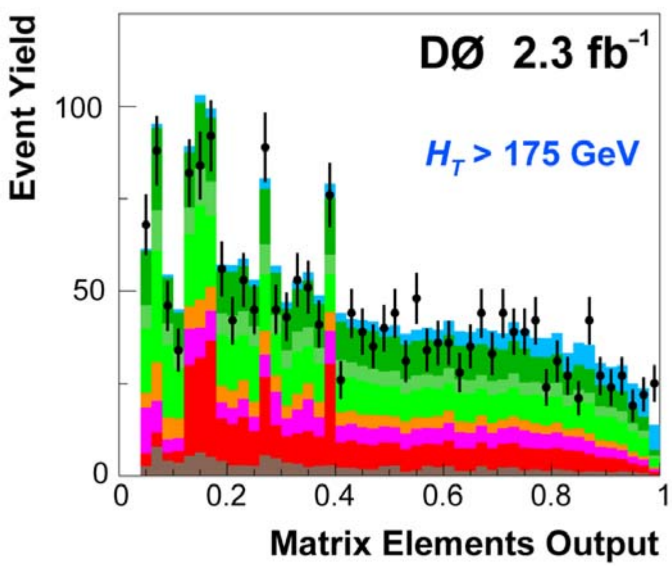
Multivariate Discriminant Outputs





Sum of 24 channels (illustration purpose, not used to measure cross section)

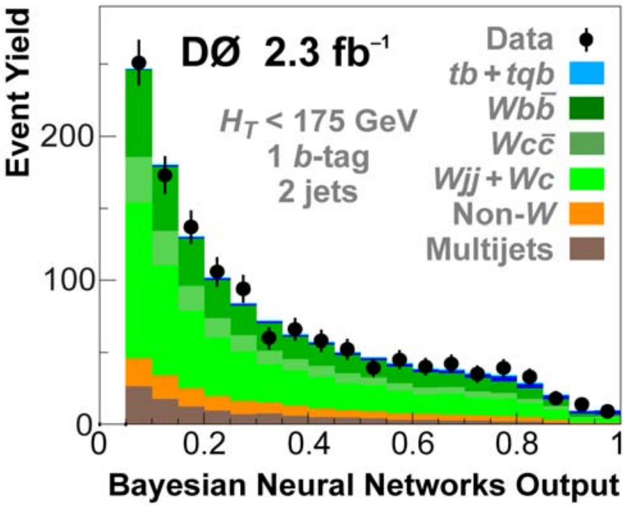




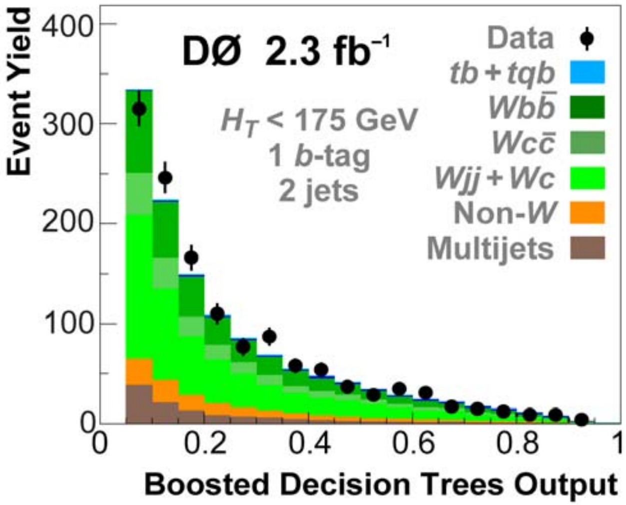
Cross Check Sample

Many cross checks, a few examples shown

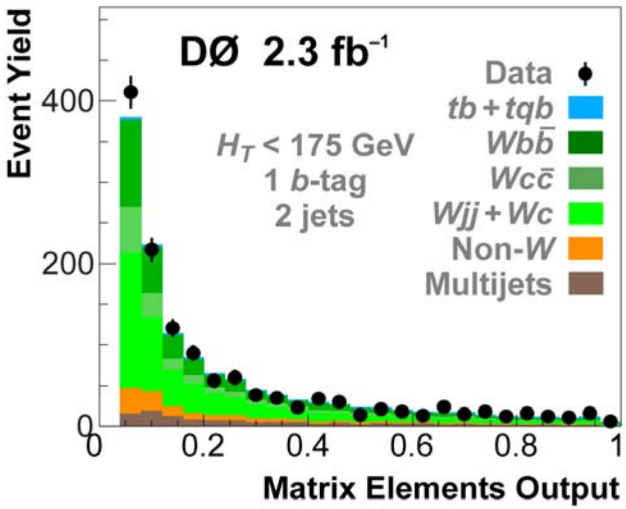
W+Jets Cross-Check Sample



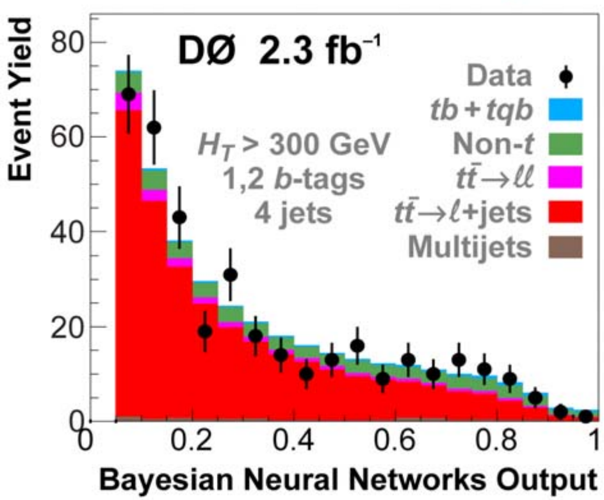
W+Jets Cross-Check Sample



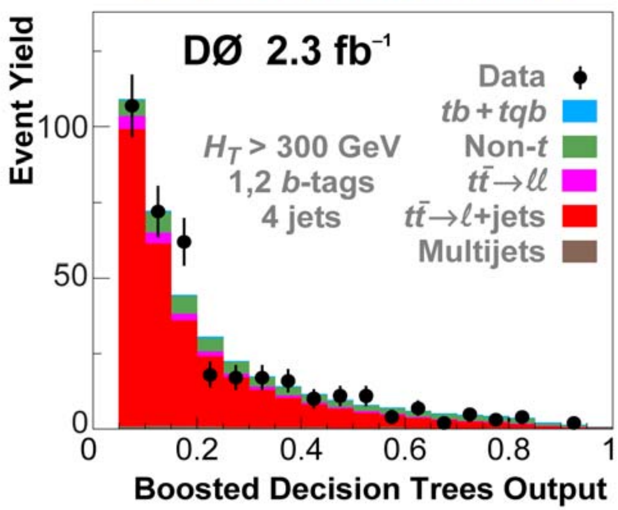
W+Jets Cross-Check Sample



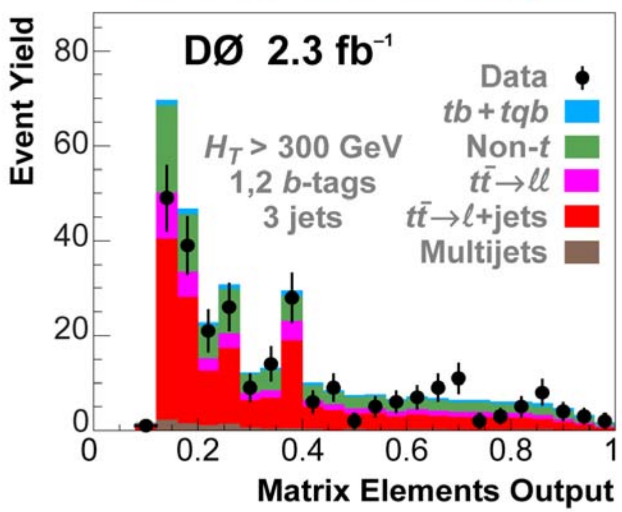
$t\bar{t}$-Pairs Cross-Check Sample



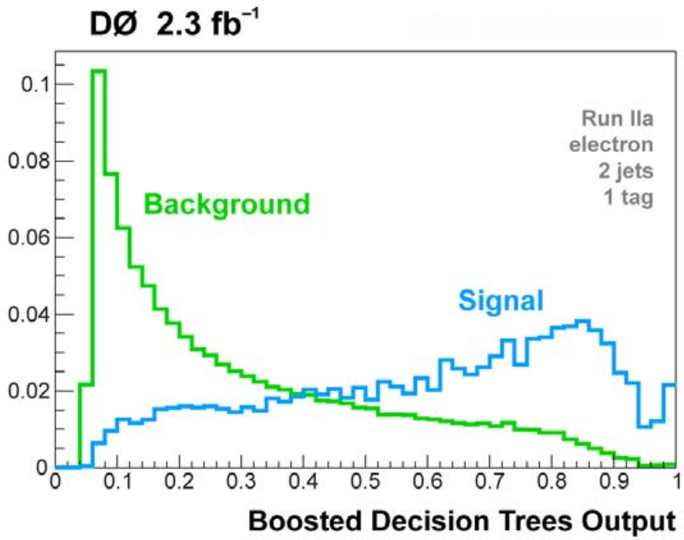
$t\bar{t}$-Pairs Cross-Check Sample

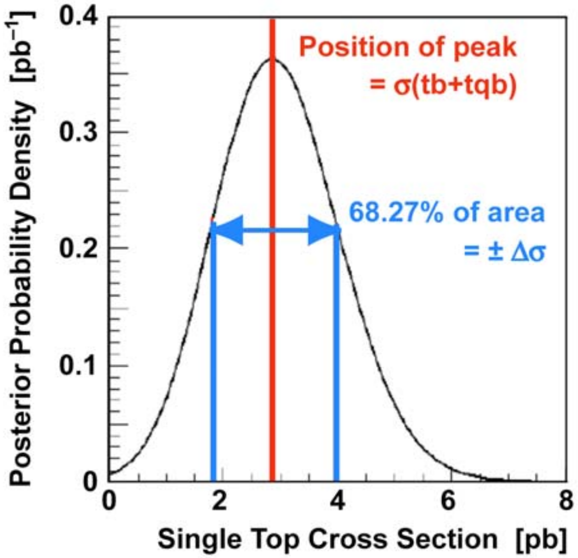


$t\bar{t}$-Pairs Cross-Check Sample



Cross Section Measurement



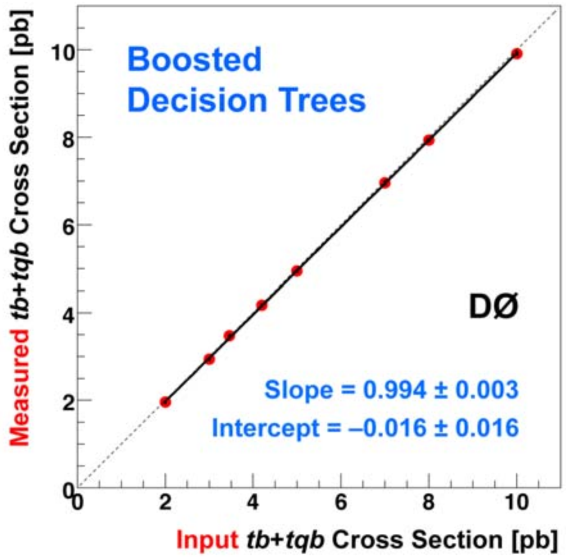


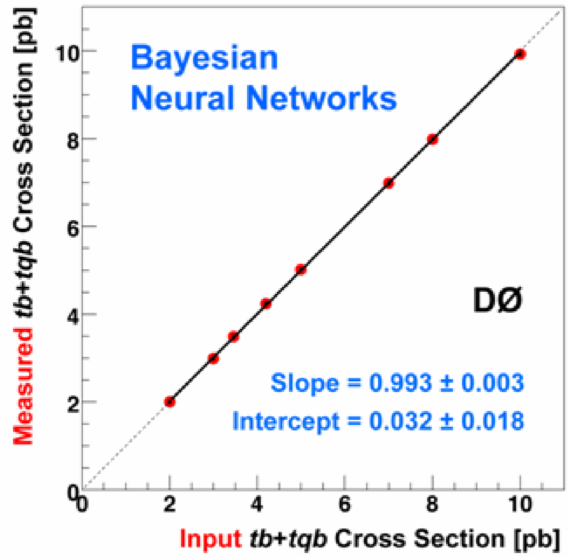
Compute Bayesian posterior probability density as a function of σ(tb+tqb)

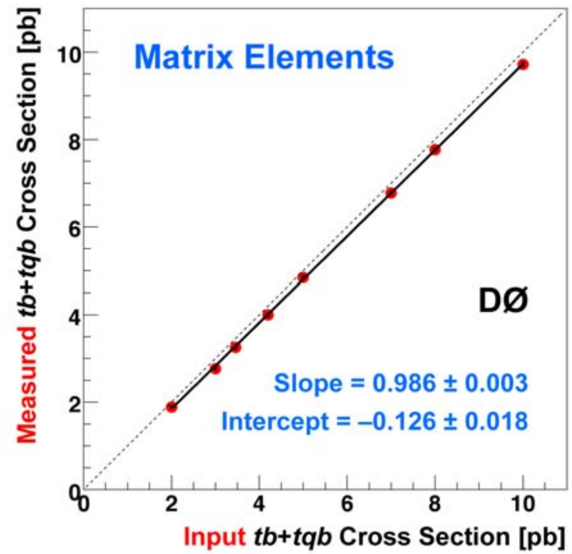
- **Using binned likelihood from discriminant distribution**
  - **Product of 24 channels**
- **Flat prior for the cross section**
- **Integrate over all systematic uncertainties**
- **Single top cross section is given by the position of the posterior density peak, with 68% asymmetric interval as uncertainty**

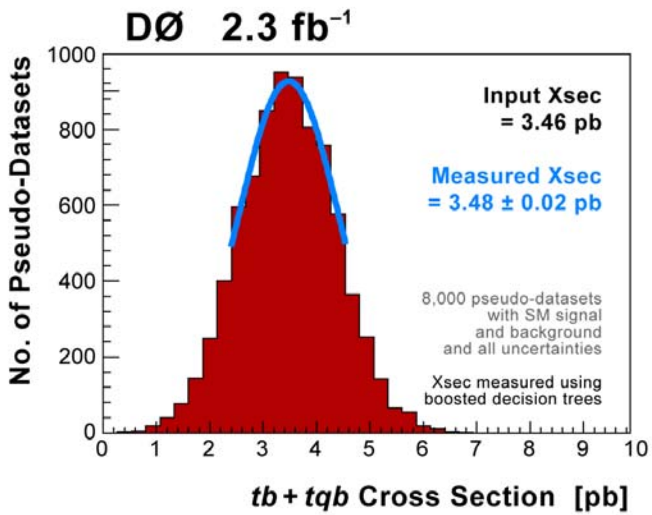
Linearity Check

Yet another cross check









Check discriminant performance

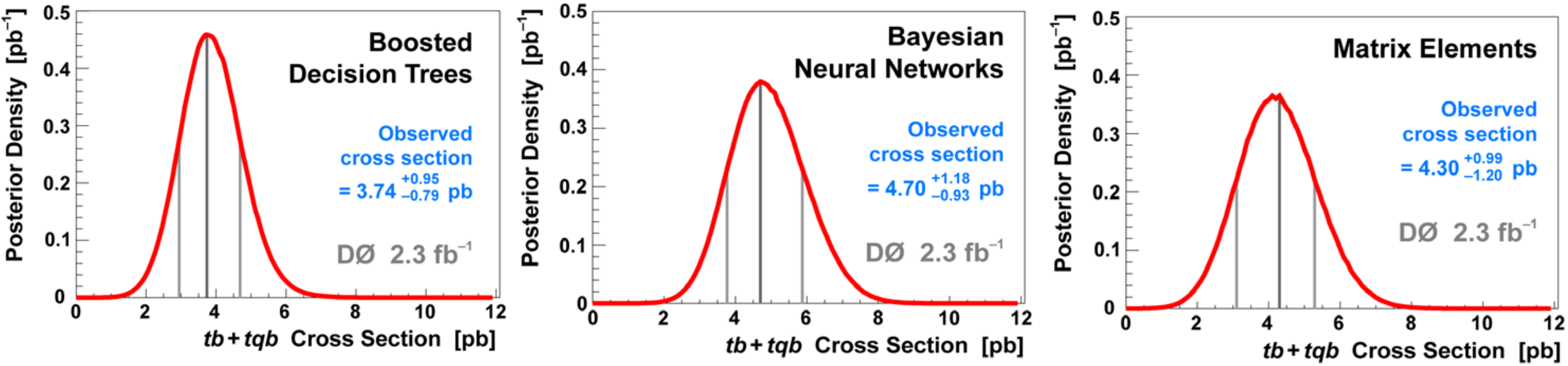
- **Measure cross section from ensemble tests**
  - **Many pseudo-datasets each contains signal and background events and uncertainties as in real data**
- **Input cross section reproduced with linear response**

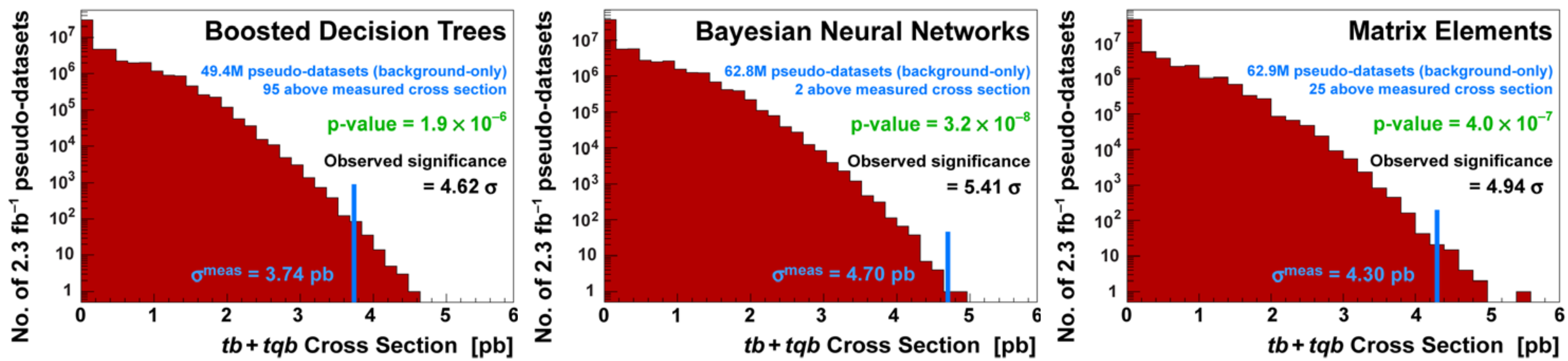
Significance

- P-value: assuming a null hypothesis, what's the probability to get a value equal to or greater than the value observed (cross section)
- Use a large ensemble of background-only pseudo-datasets. Each such dataset corresponds to 2.3 fb^{-1} data without any single top including all systematic uncertainties and all correlations
- The single top cross section was measured in each such pseudo-dataset in exactly the same way as we measure in our real dataset.
- Measure the fraction of background-only datasets in which we derive at least the SM cross section (expected significance), or at least the observed cross section (observed significance).

Expected and Observed Results

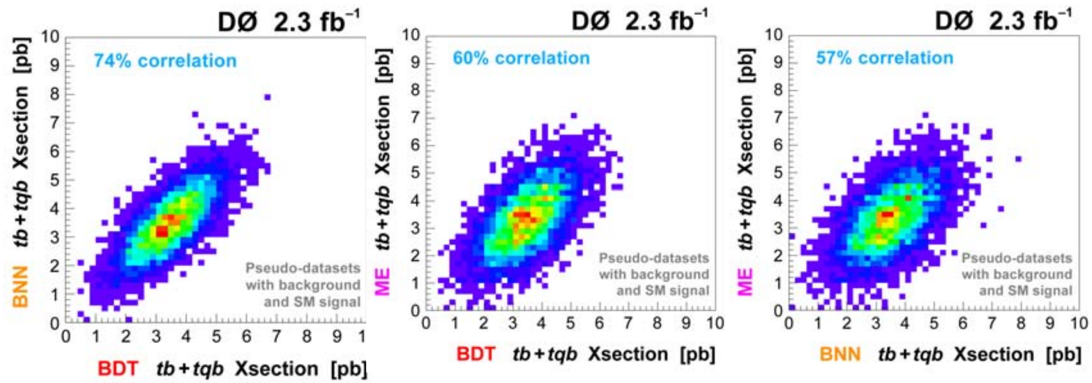
	Boosted Decision Trees		Bayesian NN		Matrix Element	
	Exp.	Obs.	Exp.	Obs.	Exp.	Obs.
σ(s+t)[pb]	$3.61^{+0.95}_{-0.89}$	$3.74^{+0.95}_{-0.79}$	$3.60^{+1.02}_{-0.90}$	$4.70^{+1.18}_{-0.93}$	$3.60^{+1.10}_{-0.96}$	$4.30^{+0.99}_{-1.20}$
significance	4.33	4.62	4.08	5.41	4.11	4.94

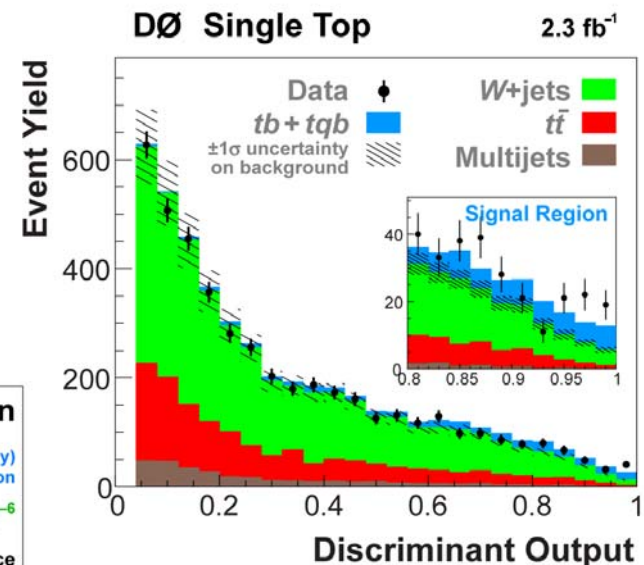


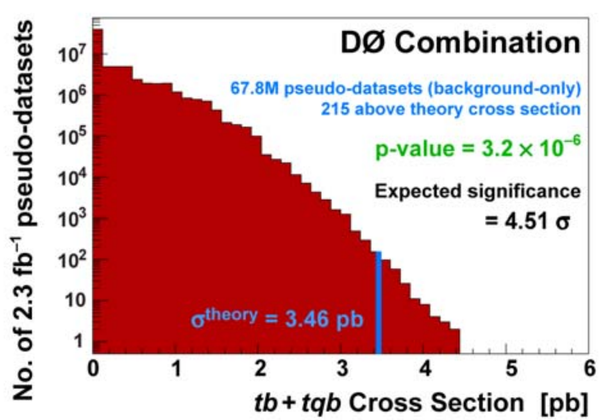


Combination of Three Analyses

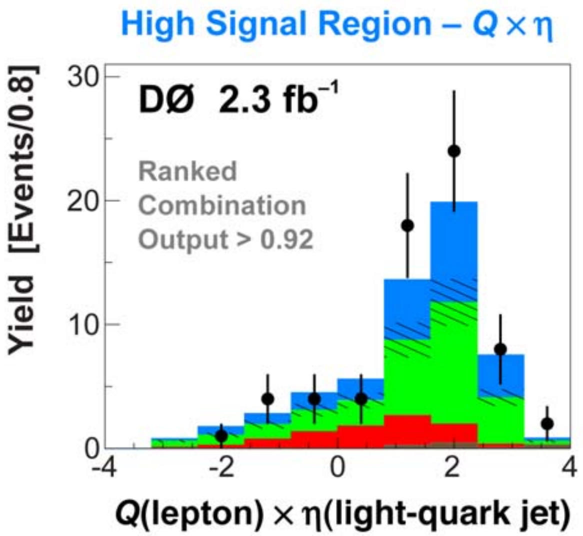
- Three analyses give consistent results
- Taking advantage of all information: a combination should lead to a more precise measurement and increased signal sensitivity
- Use a BNN to perform the combination. It takes BDT, BNN and ME discriminants as input and builds a "super discriminant"
- Best sensitivity: 4.5 σ expected significance!

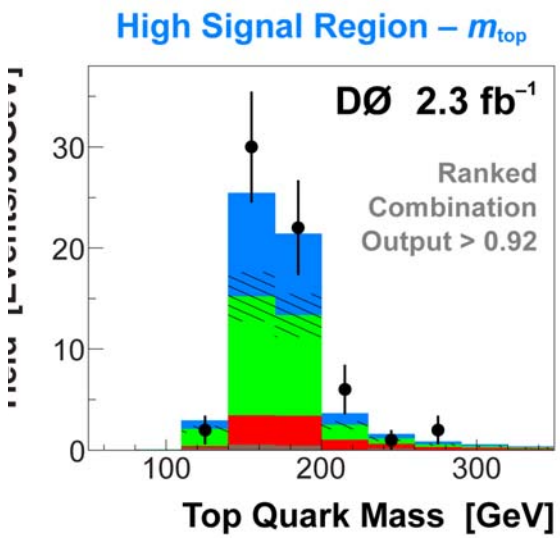


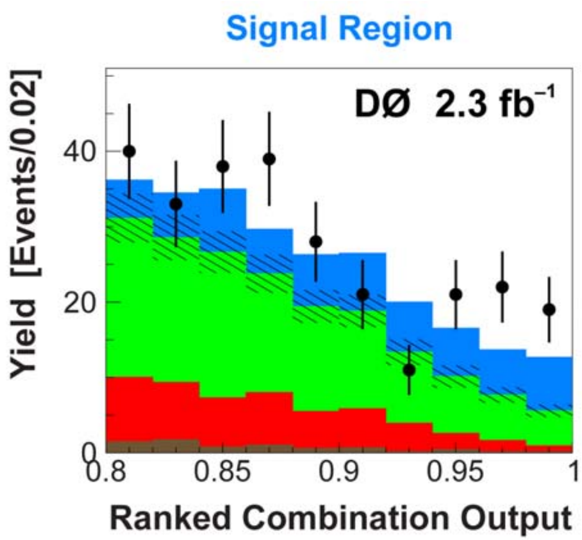


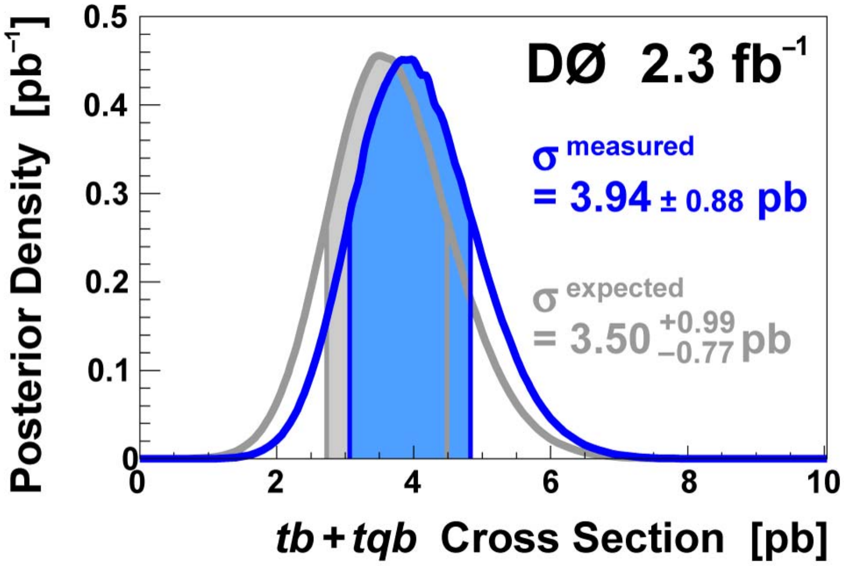


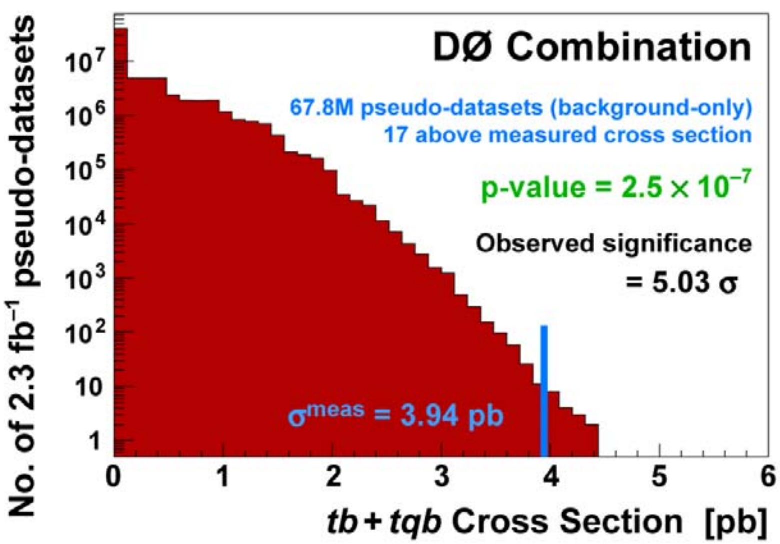
Combination Result



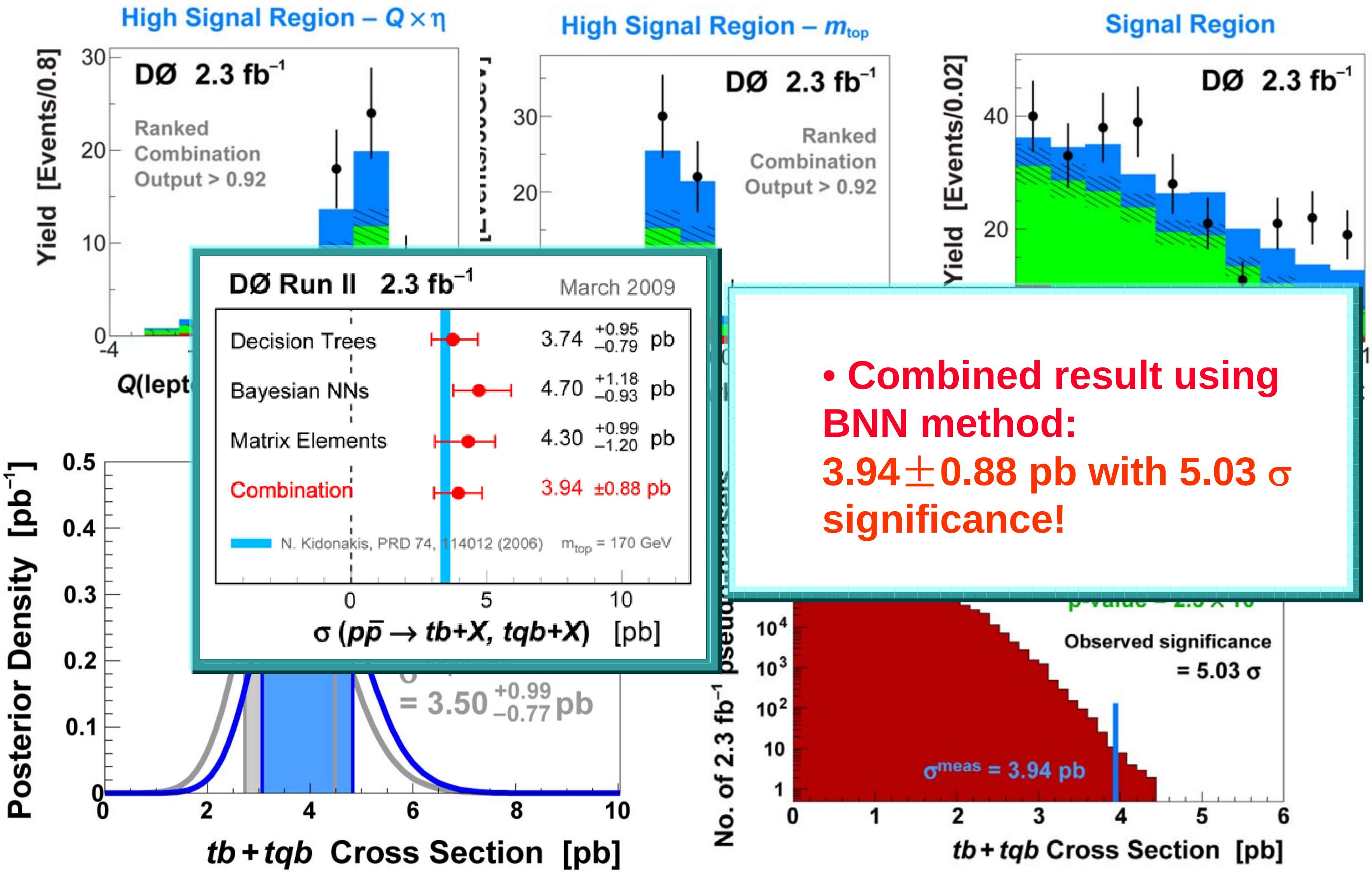








Combination Result



- **Combined result using BNN method: 3.94 ± 0.88 pb with 5.03 σ significance!**

Direct measurement of $|V_{tb}|$

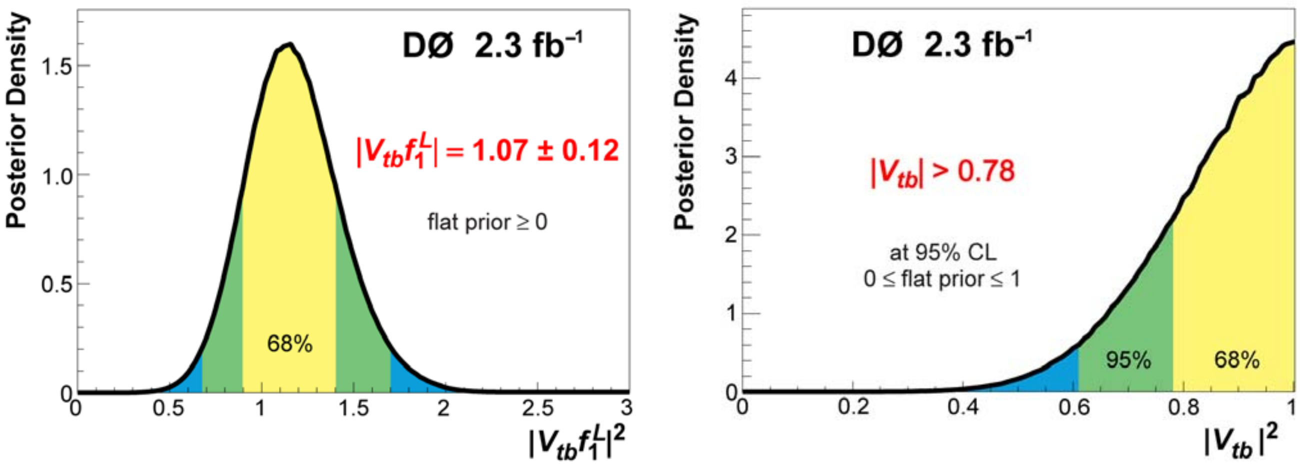
Once we have a cross section measurement, we can make direct measurement of $|V_{tb}|$

- **Calculate posterior in $|V_{tb}|^2$: $\sigma \propto |V_{tb}|^2$**

Assume standard model production:

- **Pure V-A and CP conserving interaction: $f_1^R = f_2^L = f_2^R = 0$**
- **$|V_{td}|^2 + |V_{ts}|^2 << |V_{tb}|^2$**
- **Additional theoretical errors included (top mass, scale, PDF etc…)**

Measurement does not assume 3 generations or unitarity



$0.78 < |V_{tb}| <= 1$ at 95% CL (assuming $f_1^L = 1$)

Summary

Observation for single top production at DØ !

- **Submitted to PRL (arXiv:0903.0850)**

Summary of results:

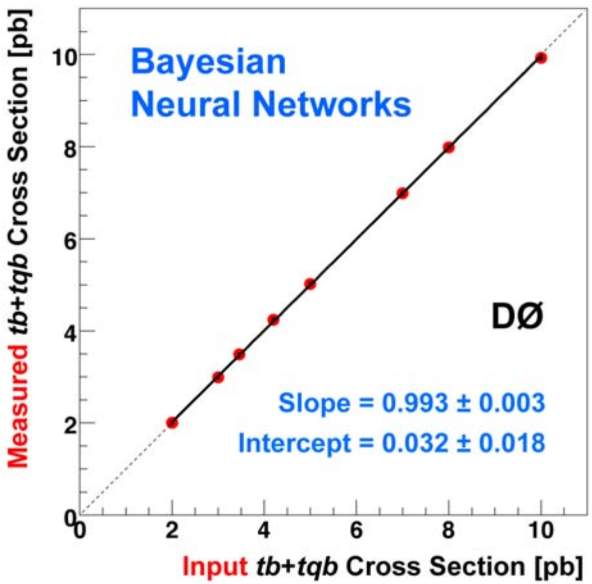
DØ 2.3 fb^{-1} Single Top Results			
Analysis Method	Single Top Cross Section	Significance	
		Expected	Measured
Boosted Decision Trees	$3.74^{+0.95}_{-0.79}$ pb	4.3 σ	4.6 σ
Bayesian Neural Networks	$4.70^{+1.18}_{-0.93}$ pb	4.1 σ	5.4 σ
Matrix Elements	$4.30^{+0.99}_{-1.20}$ pb	4.1 σ	4.9 σ
Combination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

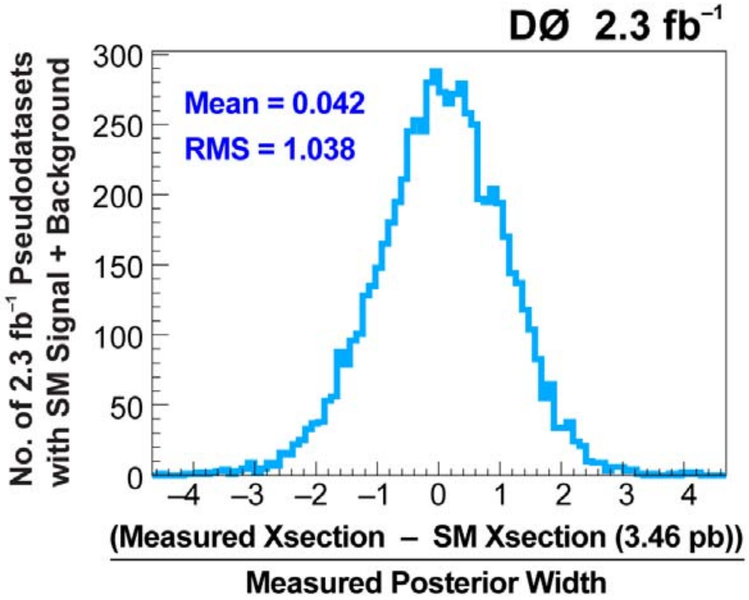
$|V_{tb} f_1^L| = 1.07 \pm 0.12$

$|V_{tb}| > 0.78$ at 95% CL

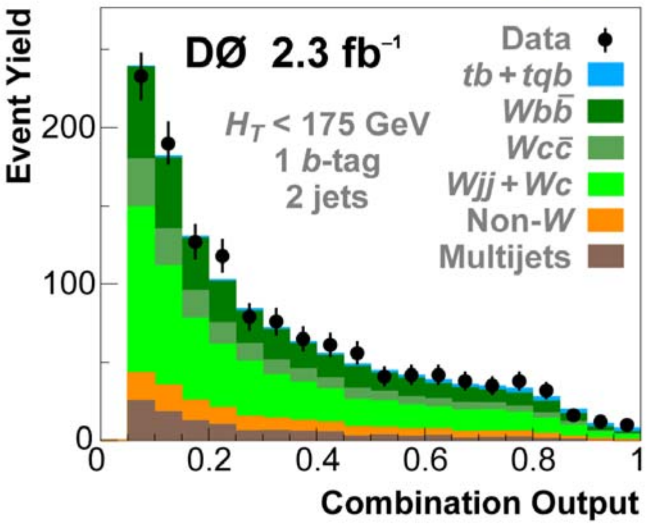
Backup

Combination Cross Check

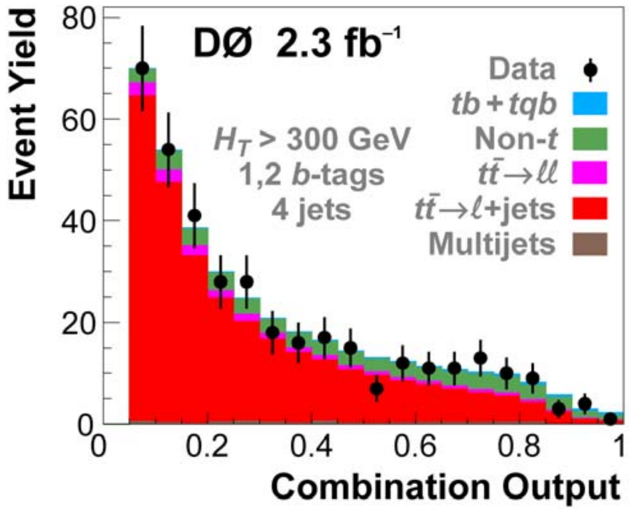




W+Jets Cross-Check Sample

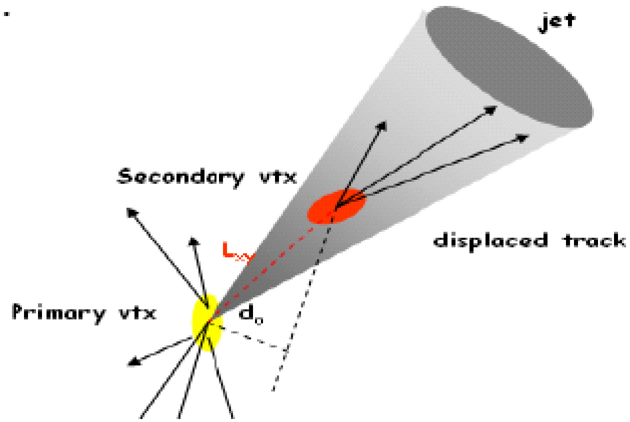


tt-Pairs Cross-Check Sample



b-jet identification (b-tagging)

- Separate *b*-jets from light-quark and gluon jets to reject most W+jets background
- DØ uses a neural network algorithm with seven input variables based on impact parameter and reconstructed vertex
- Two operating points used in analysis:
  - TIGHT ($\varepsilon_b = 40\%$, $\varepsilon_c = 9\%$, $\varepsilon_\ell = 0.4\%$)
  - LOOSE ($\varepsilon_b = 50\%$, $\varepsilon_c = 14\%$, $\varepsilon_\ell = 1.5\%$)
- Leading *b*-jet $p_T > 20$ GeV
- Define two exclusive samples
  - EqOneTag: 1T, no L
  - EqTwoTag: 2L (was 2T; ≈ 50% gain)



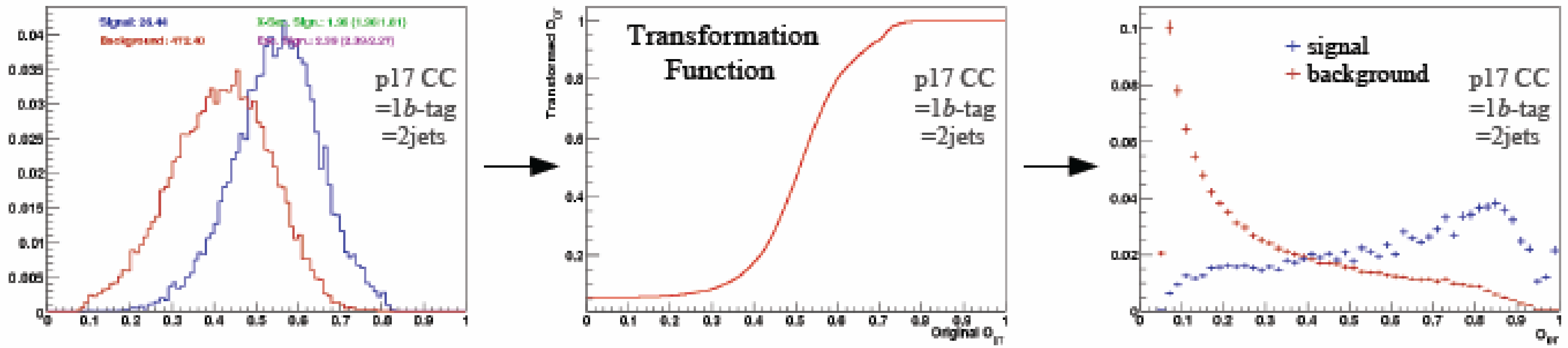
- Uncertainties dominated by variation in data samples used to measure the efficiencies.
- Smaller contribution from MC sample dependence

MC splitting and Binning Transformation

MC samples were split into three equal-sized independent subsets

- training sample : used to train the BDT, BNN and ME
- testing sample : used for testing purposes (and also to train the combination BNN)
- yield sample : used to measure the cross section and make all plots

BINNING TRANSFORTATION: Used for BDT, BNN & Combination



Same-sized bins (left) suffer from low entries, a bin can have signal but no background. Transformation (center) ensures a minimum number of background events in each bin. Transformed quantity (right) stays proportional to the probability of being signal.

ME: bins are ordered in S/B, and sample is split in two H_T regions