GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

GEFÖRDERT VOM

Bundesministerium für Bildung und Forschung

Observation of the single top production at the Tevatron

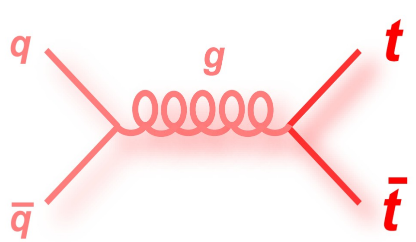
Elizaveta Shabalina
University of Goettingen, Germany

Outline

- Motivation
- Signal and background
- Event selection
  - Systematic uncertainties
- Statistical analysis
  - Cross section extraction
  - Significance calculation
- Multivariate Methods
- Combination
- $|V_{tb}|$ measurement
- Summary and outlook

The Top quark

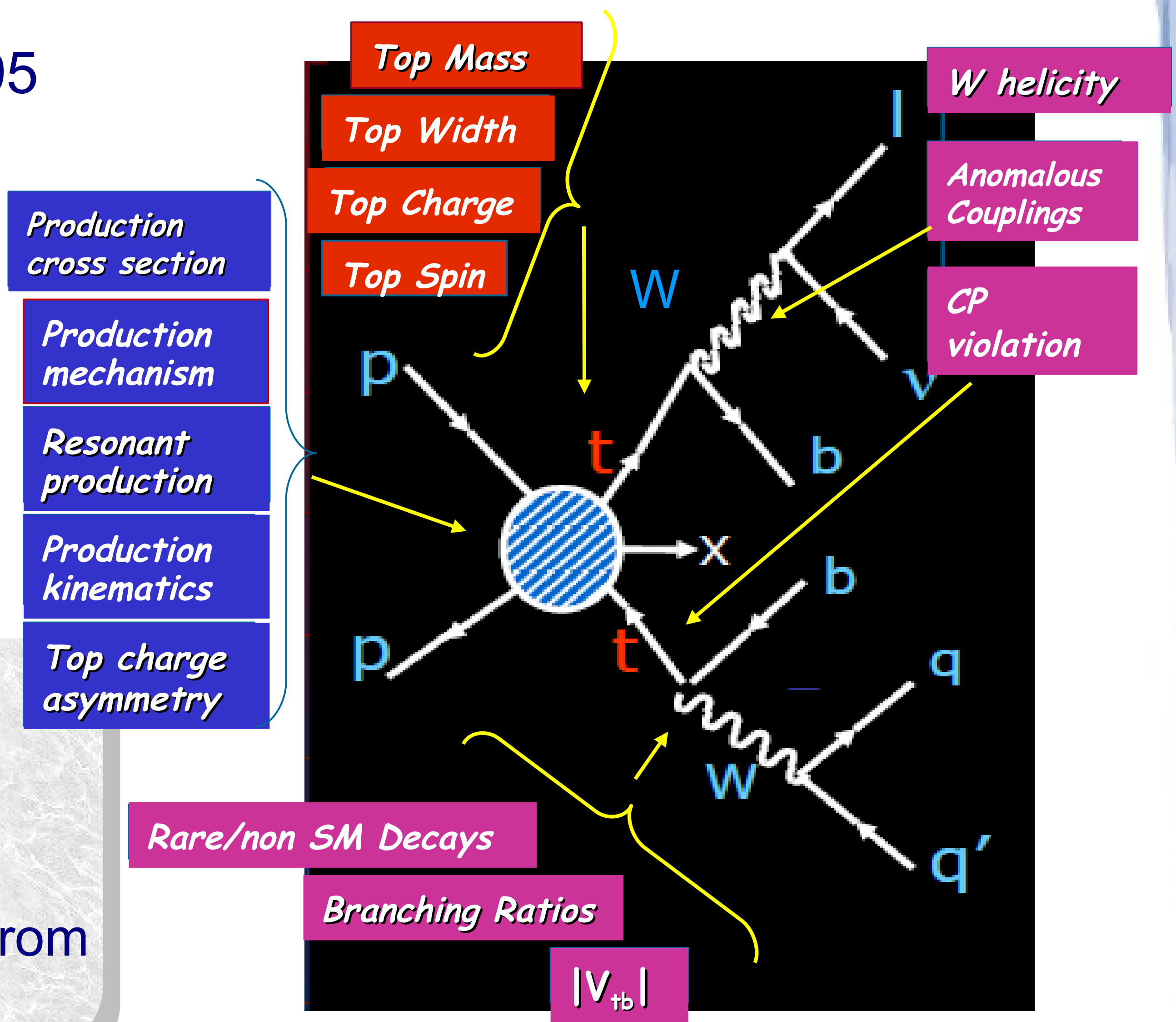
Discovered in March 1995



σ_{NLO}= 6.8±0.6 pb @175 GeV

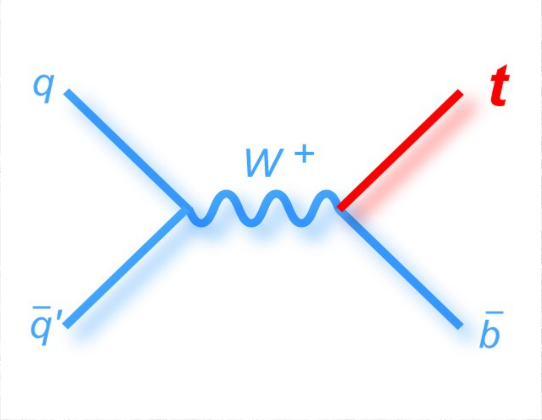
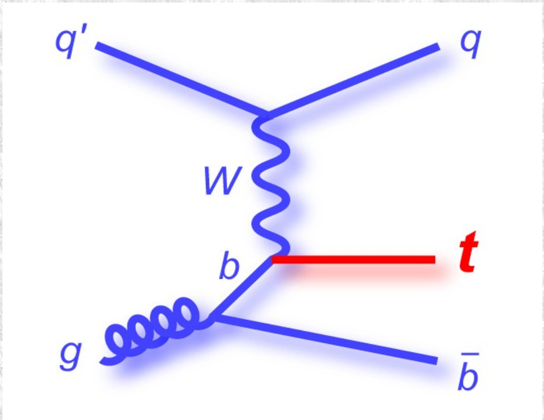
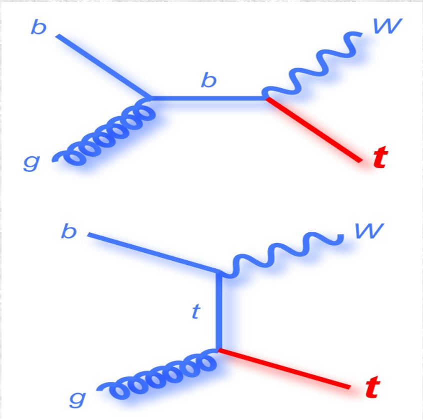
N.Kidonakis et. al., M.Cacciari et. al., S.Moch, P.Uwer et.al.

- Main mechanism
- Distinct signature
- Thoroughly studied
- All knowledge comes from strong production



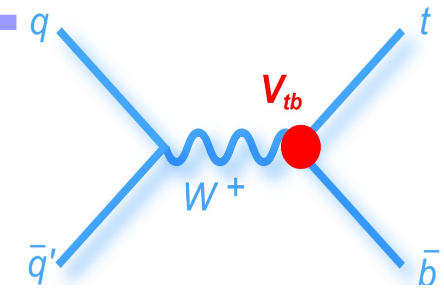
Electroweak production

- Predicted 10 years before top quark discovery
  S.Willenbrock, D. Dicus, Phys. Rev. D34, 155 (1986); S Cortese and R Petronzio, PLB 253, 494 (1991)
- Observed 14 years after top quark discovery...

m_t (GeV)		s-channel (tb)	t-channel (tqb)		Wt channel
					
175	σ_{NLO}	0.88±0.11 pb	1.98±0.25 pb	Z. Sullivan, Phys. Rev. D 70, 114012 (2004)	**Small at Tevatron**
170	$\sigma_{(N)NLO}$	1.12±0.05 pb	2.34±0.13 pb	N. Kidonakis, Phys. Rev. D 74, 114012 (2006)	**Important for LHC**

In observation analysis CDF (D0) assumes m_t=175 (170) GeV

What will we learn?



$\sigma \propto |V_{tb}|^2$

Access to W-t-b vertex

- Probe V-A structure
- Top quark spin

Direct measurement of $|V_{tb}|^2$

- Test unitarity of CKM matrix
- Is it 3×3 matrix?
- Is 4[th] generation possible?

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} & V_{uX}? \\ V_{cd} & V_{cs} & V_{cb} & V_{cX}? \\ V_{td} & V_{ts} & V_{tb} & V_{tX}? \\ V_{Yd}? & V_{Ys}? & V_{Yb}? & V_{YX}? \end{pmatrix}$$

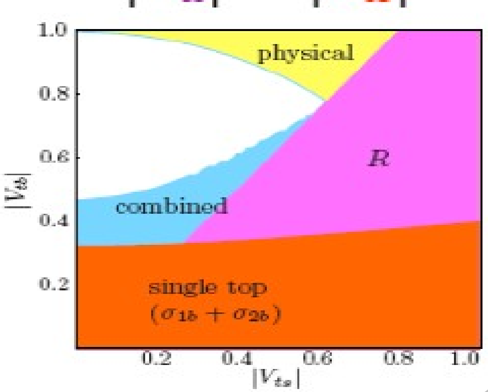
Small mixing with 4[th] family is favored
Quite large mixing is still not excluded

Constraints:
tree-level 3×3 CKM elements
FCNC processes (K-, D-, B_d-, B_s-mixing, b → s)
Assumption: unitary 4×4 CKM matrix
A. Lenz et al. in arXiv 0902.4883 [hep-ph]

Combination of R_b and s- & t-channel cross sections

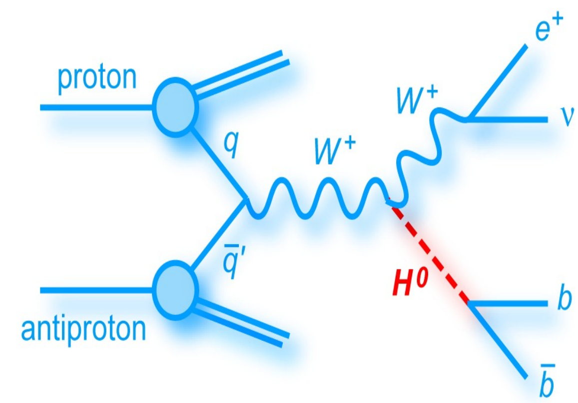
J. Alwall et. al., Eur. Phys. J. C49 791 (2007).:

$|V_{ts}|$ vs $|V_{tb}|$



SM and beyond

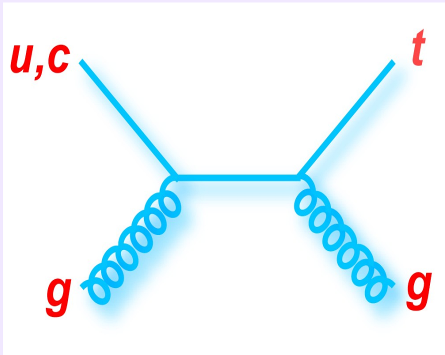
- Milestone for Higgs search in WH channel
  - Same signature, 10 times smaller σ
  - Background to Higgs search



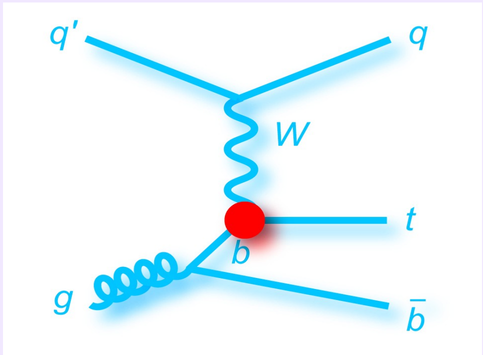
$p\bar{p} \to WH$

- s- and t-channels are sensitive to different processes beyond the standard model

t-channel

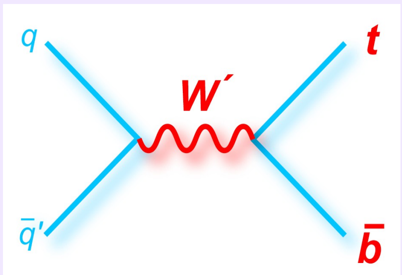


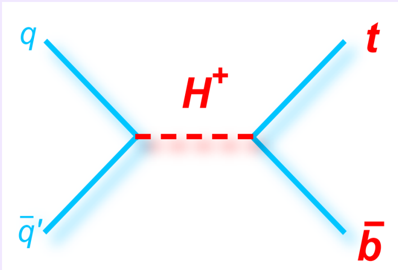
FCNC ($ug \to t$)



Anomalous couplings

s-channel





Resonances

Long way to discovery

DØ

· Search:	PRD 63, 031101 **(2000)**
· Search:	PLB 517, 282 **(2001)**
· Search:	PLB 622, 265 **(2005)**
· W':	PLB 641, 423 **(2006)**
· Search:	PRD 75, 092007 **(2007)**
· **Evidence:**	PRL 98, 181802 **(2007)**
· FCNC:	PRL 99, 191802 **(2007)**
· W':	PRL 100, 211802 **(2007)**
· Evidence:	PRD 78, 012005 **(2008)**
· Wtb:	PRL 101, 221801 **(2008)**
· Wtb:	PRL 102, 092002 **(2009)**
· H^+:	(PRL) arXiv:0807.0859
· **Observation:**	**(PRL) arXiv:0903.0850**

Run I

Run II

CDF

· Search:	PRD 65, 091102 **(2002)**
· W'	PRL 90, 081802 **(2003)**
· Search:	PRD 69, 052003 **(2004)**
· Search:	PRD 71, 012005 **(2005)**
· **Evidence:**	**PRL 101, 252001 (2008)**
· FCNC:	(PRL) arXiv:0812.3400
· W':	(PRL) arXiv:0902.3276
· **Observation:**	**(PRL) arXiv:0903.0885**

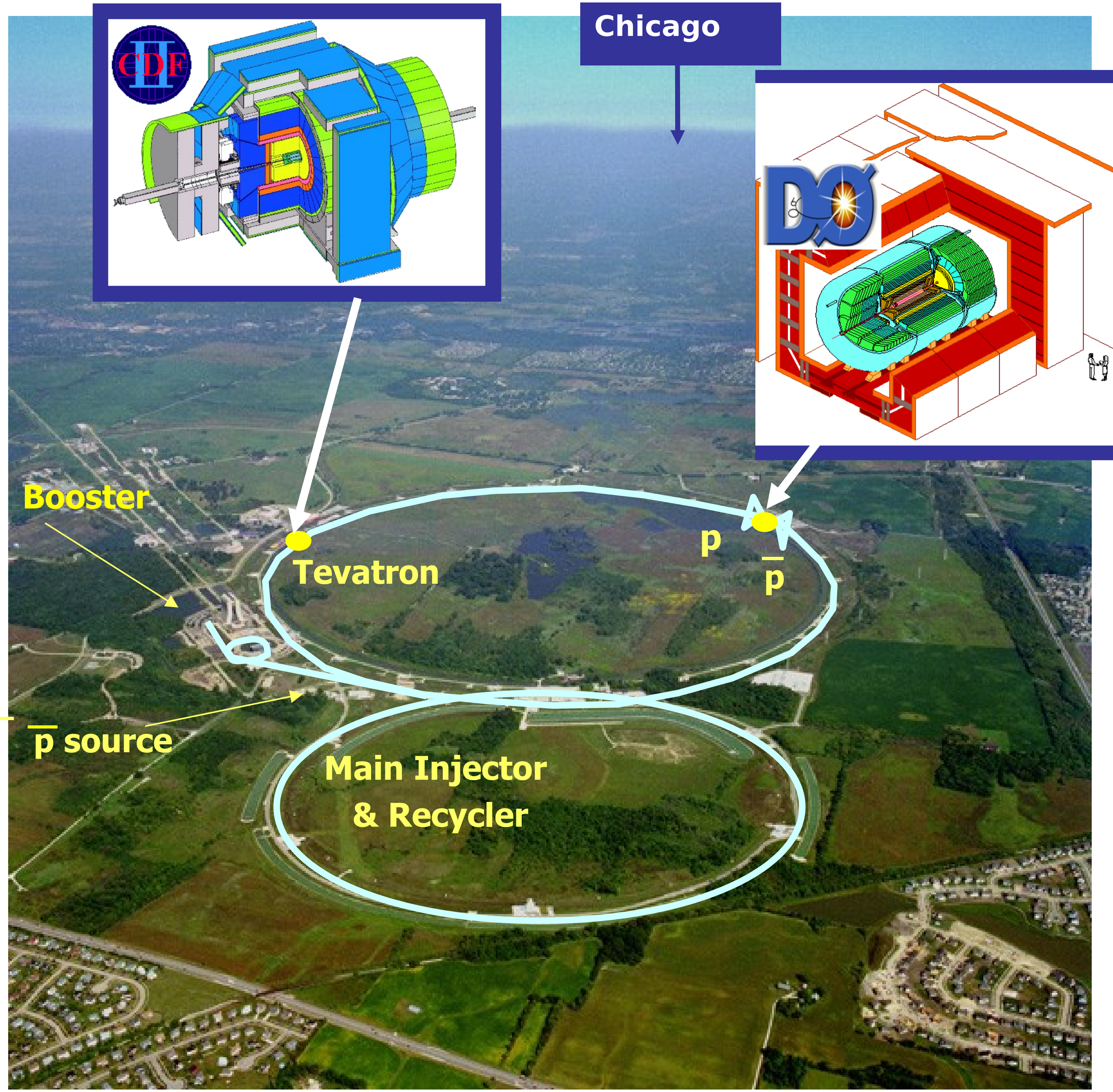
Single Top Cross Section	Signal Significance Expected	Signal Significance Observed	CKM Matrix Element V_{tb}
December 2006 DØ (0.9 fb^{-1})			PRL 98, 181802 (2007)
4.7 ± 1.3 pb	2.3 σ	3.6 σ	$\lvert V_{tb} f_1^L \rvert = 1.31^{+0.25}_{-0.21}$ $\lvert V_{tb} \rvert > 0.68$ at 95% CL
September 2008 CDF (2.2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

The Tevatron

- The highest energy particle accelerator in the world
- Proton-antiproton collider with $\sqrt{s}$ = 1.96 TeV

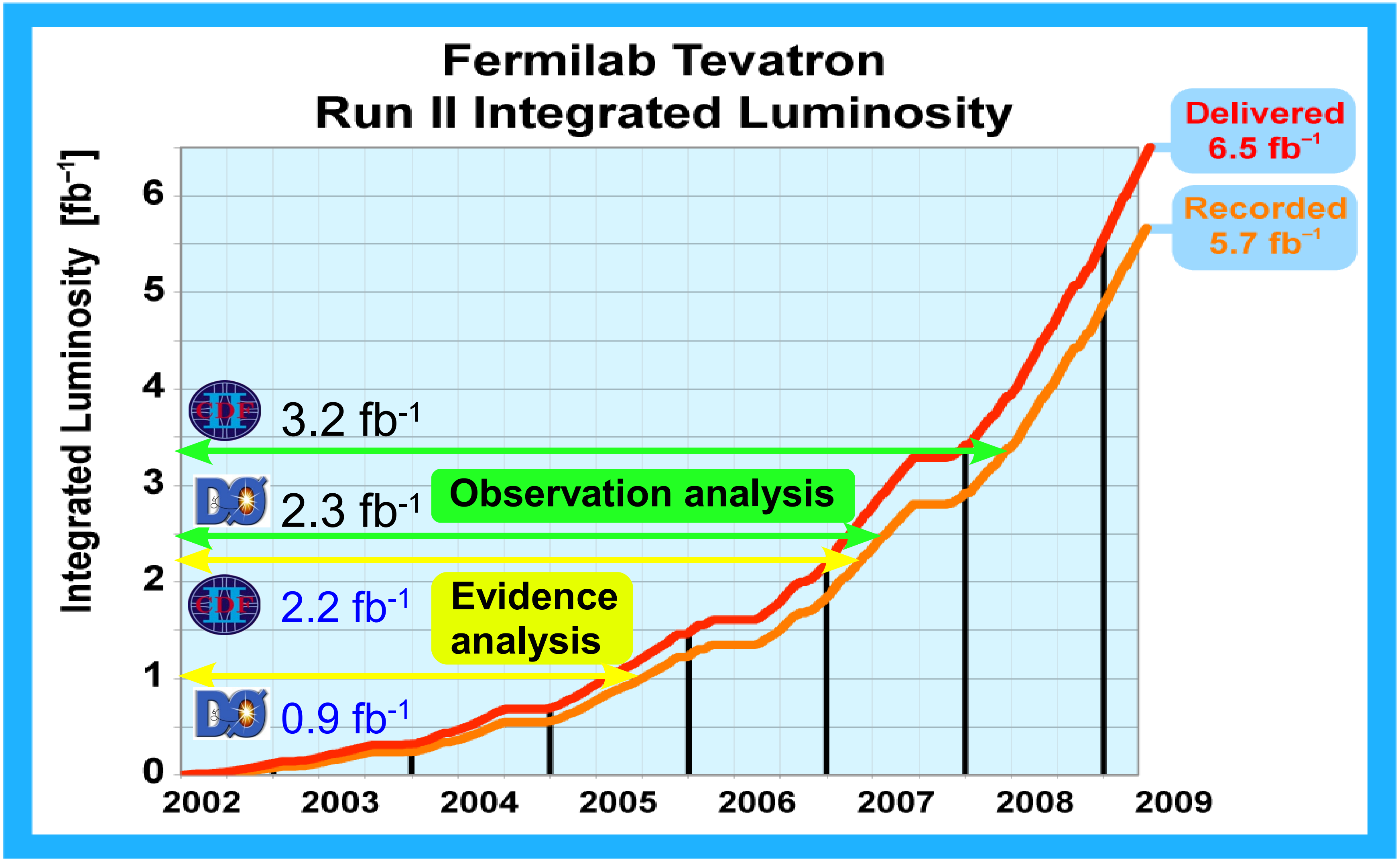
Run I 1992-1995
Top quark discovered!

Run II 2001-11(?)
Single top quark discovered!



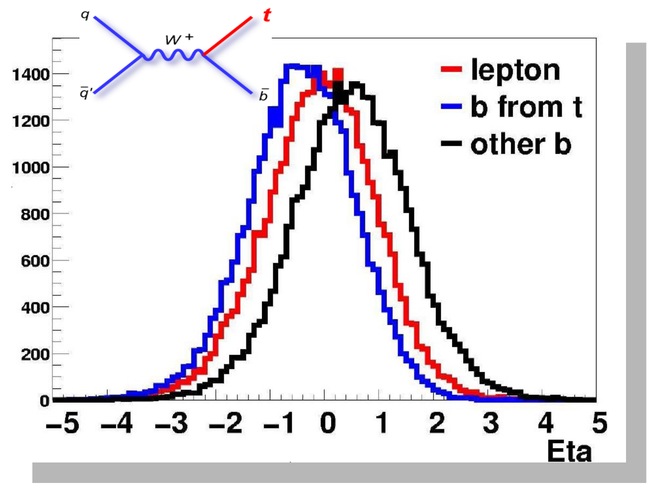
Climbing to the top...

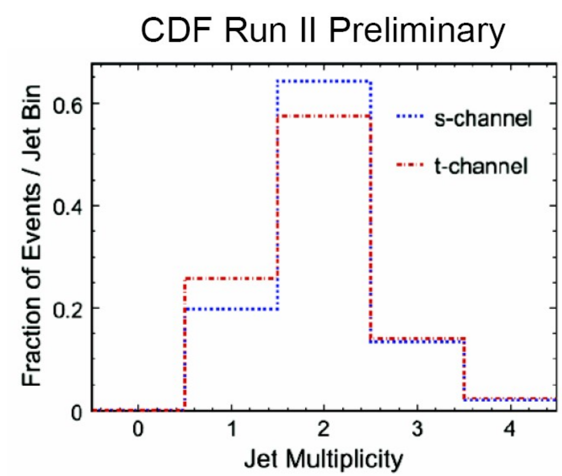
Outstanding performance of the Tevatron!



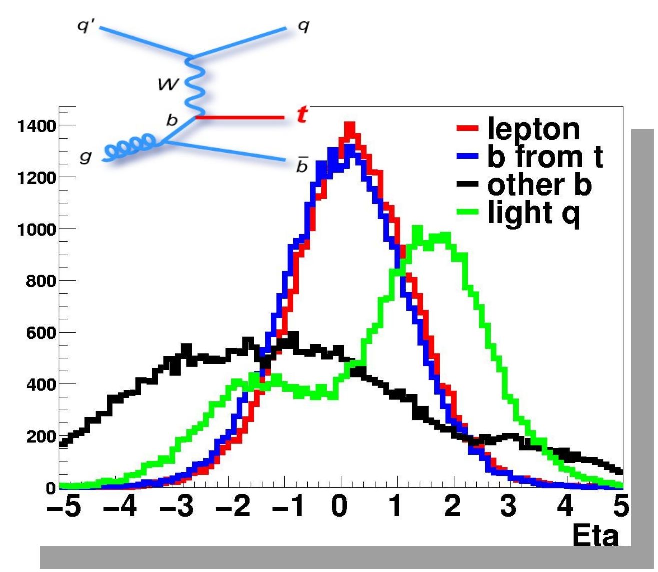
Signal

- s-channel
  2 *b*-jets
  Top quark decay products and the *b* tend to be all central



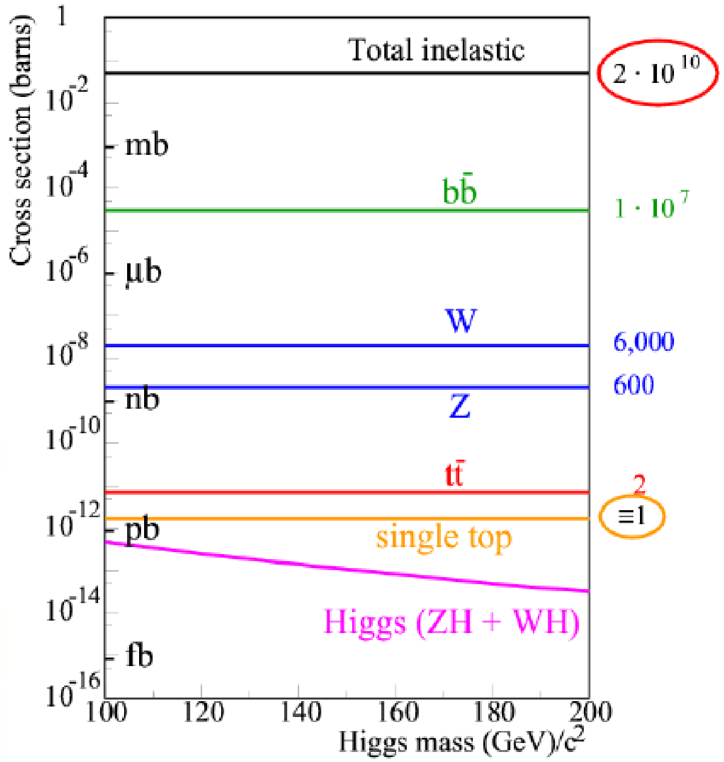


- t-channel
  2 *b*-jets and one light
  One of *b*'s tends to be very close to the beam pipe
- No striking signatures as for $t\bar{t}$
- Signal and background distributions look similar



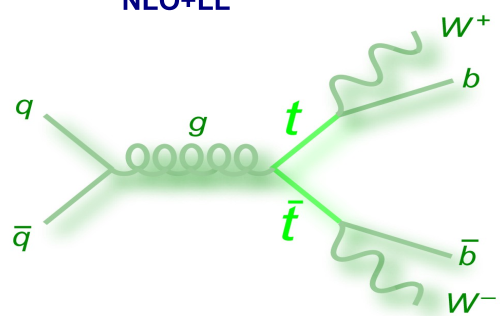
Simulated with
- CompHEP–SingleTop (D0)
- MadEvent (CDF)
- Matching of LO and NLO calculations

Backgrounds



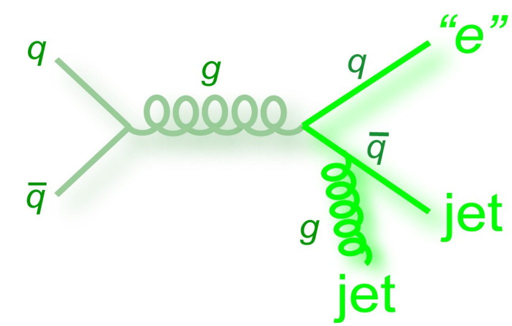
Top pairs

- **Alpgen or Pythia**
- **σ_{NLO+LL}**



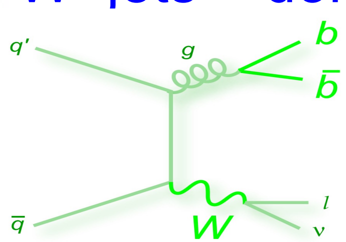
Multijet (instrumental)

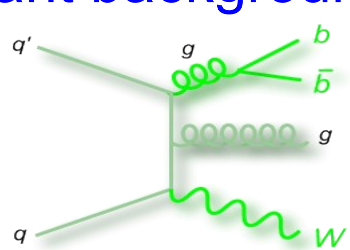
- **From data**



Other small backgrounds: Z+jets, diboson – MC

W+jets – dominant background



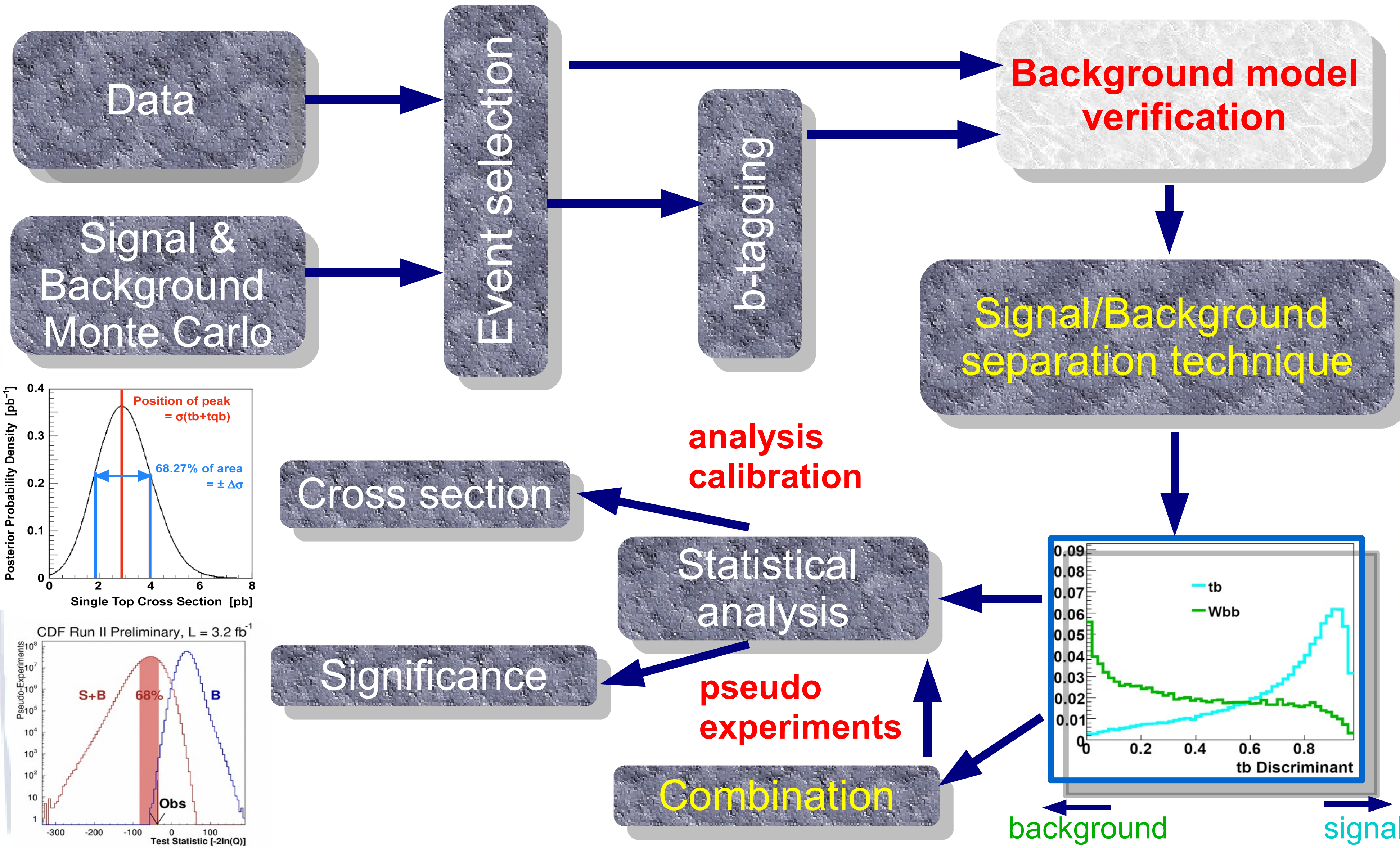


- **Estimated from data and MC**

Challenge:

- Backgrounds are large and similar in shape to signal
- Multivariate techniques are necessary to exploit all available information

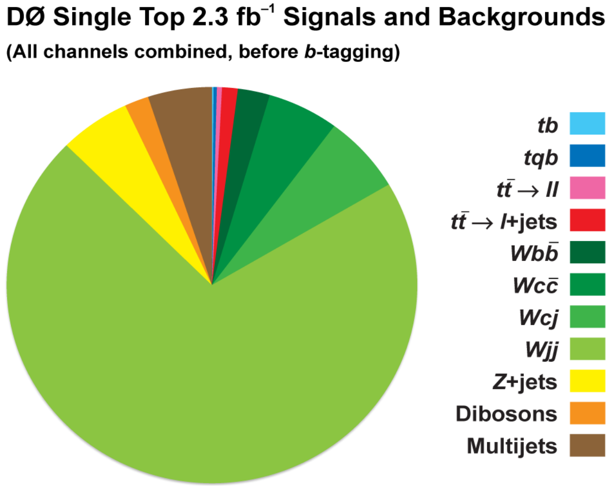
Building blocks

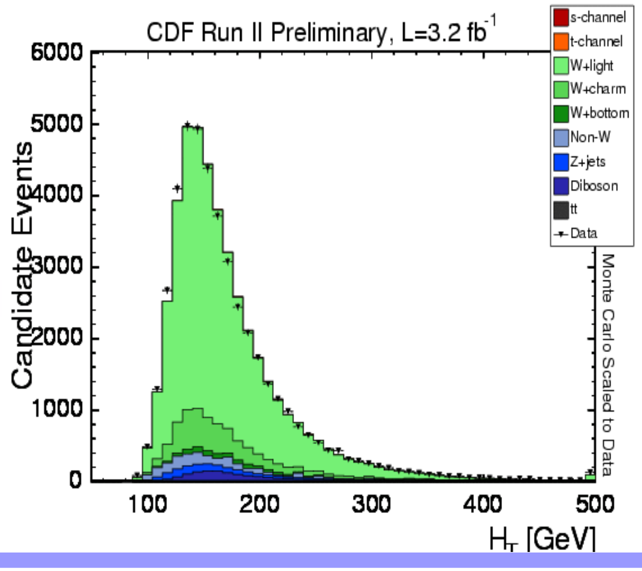


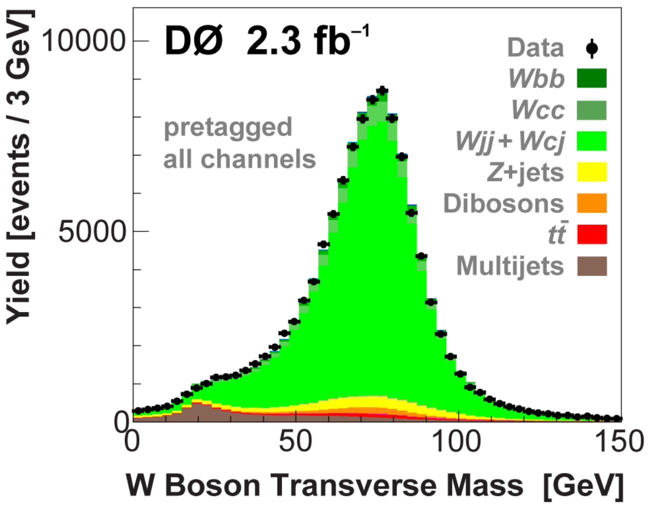
Selection I (l+jets)

Starting S:B = 1:10^9

- Single lepton (e, μ) & MET+ jets triggers
- One high p_T lepton
- MET and 2-4 (D0), 2-3 (CDF) high p_T jets
- Cuts to suppress multijet background
- Veto to suppress Z+jets and $t\bar{t}$ dilepton
- S:B ~ 1: 260







- Verify background model before *b*-jet tagging
- Dominated by W+ light jets

Selection II: *b*-tagging (l+jets)

DØ NN tagger

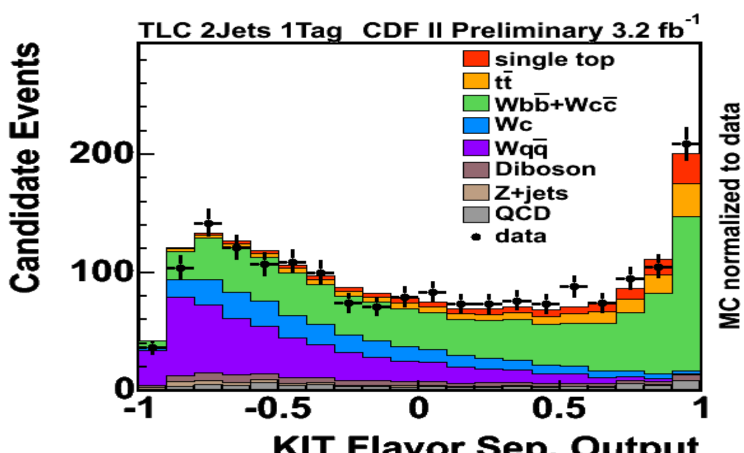
- 1 tight tag (40/9/0.4% *b*/*c*/light) or
- 2 loose (50/14/1.5% *b*/*c*/light)

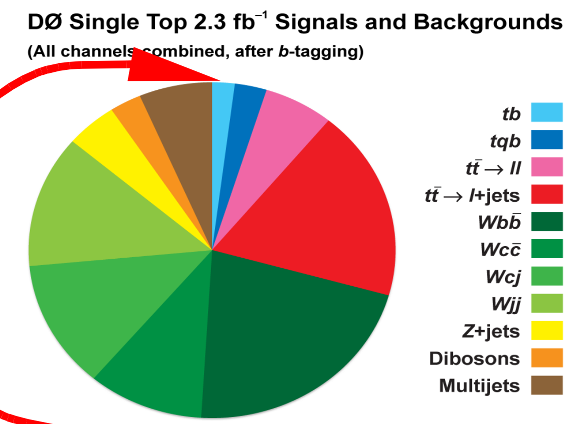
CDF 1 SVX tag

- 50/9/0.5÷1.0% b/c/light

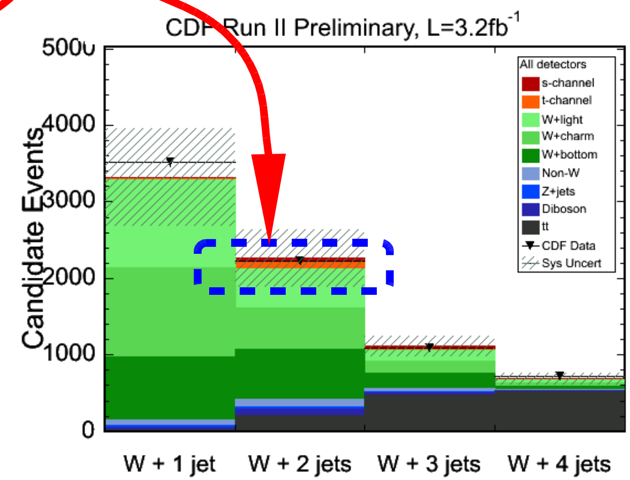
Additional flavor separation

- NN trained with 25 input variables
- Continuous variable
- Improves sensitivity by 10-15%





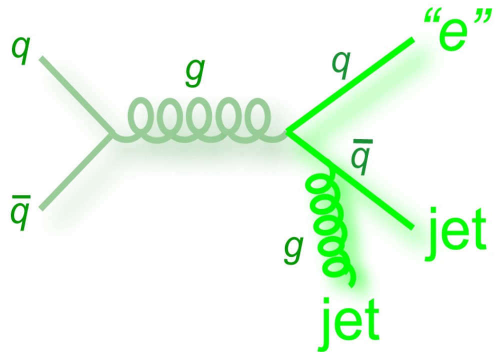
Single top signal



- **Signal smaller than background error**
- **Cut-based analysis will not work**
- **Exploit kinematic differences between signal and background**

Multijet background

- Jet misidentified as electron
- Muon in jet appeared isolated



Strategy:

- Reduce as much as possible using topological cuts (~5%)
- Determine before *b*-tagging

Modeling:

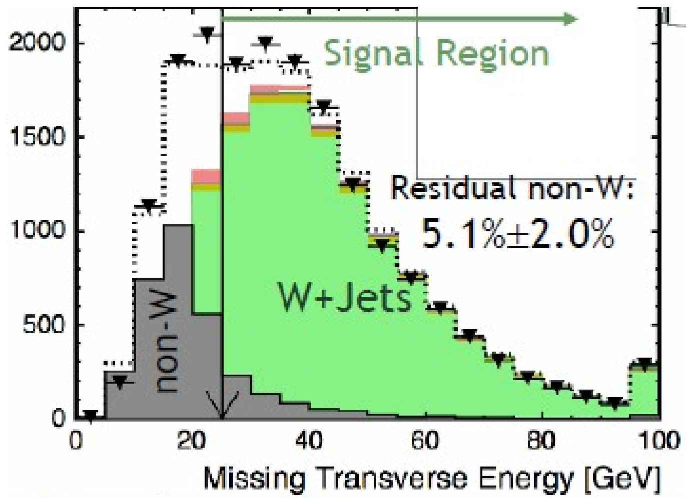
- Anti-Electrons (cuts reverted)
- Jet-Electrons (jets with high EM fraction)
- Anti-Muons (cuts reverted)

Method:

- Iterative template fits to data on three sensitive variables: $p_T(l)$, MET, $M_T(W)$

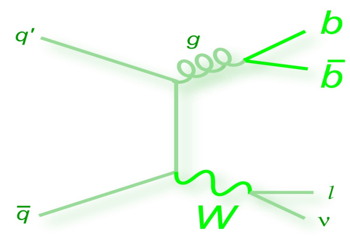
$$N^{data}_{pretag} - N^{MC}_{bkgd} = S_{W+\ jets} N^{MC}_{W+\ jets} + S_{multijet} N^{data}_{multijet}$$

- Fit to MET distribution in data

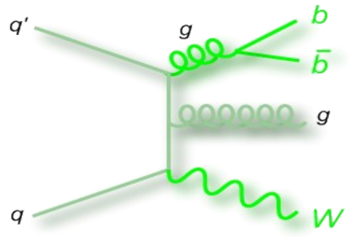


Large systematic uncertainties assigned (30÷50%)

W+jets background

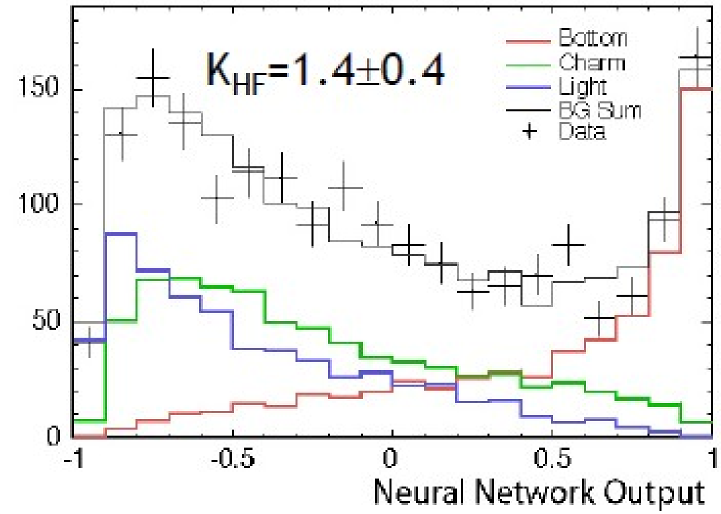


- Shapes simulated by Alpgen+Pythia MC
- Normalized to data before tagging



- W+HF normalized to theory (MCFM)
  - **1.47 (Wbb, Wcc), 1.38 (Wcj)**
- Additional empirical correction
  - **derived from two-jet data and simulation: includes 0-tag events**
  - **0.95 ± 0.13 (Wbb, Wcc)**
- Uncertainties considered
  - **Data statistics ± 9%**
  - **± 40% single top cross section → ±7% in SF**
  - **± 10% on Wcj theory SF → ± 8% in SF**
  - **± 10% Wbb/Wcc ratio → ± 5% in SF**

- HF fractions (Wbb:Wcc:Wc) from Alpgen+Pythia MC
- HF contribution is boosted by factor from fit to flavor separator output in tagged W+1 jets events



Selection II – MET+ jets

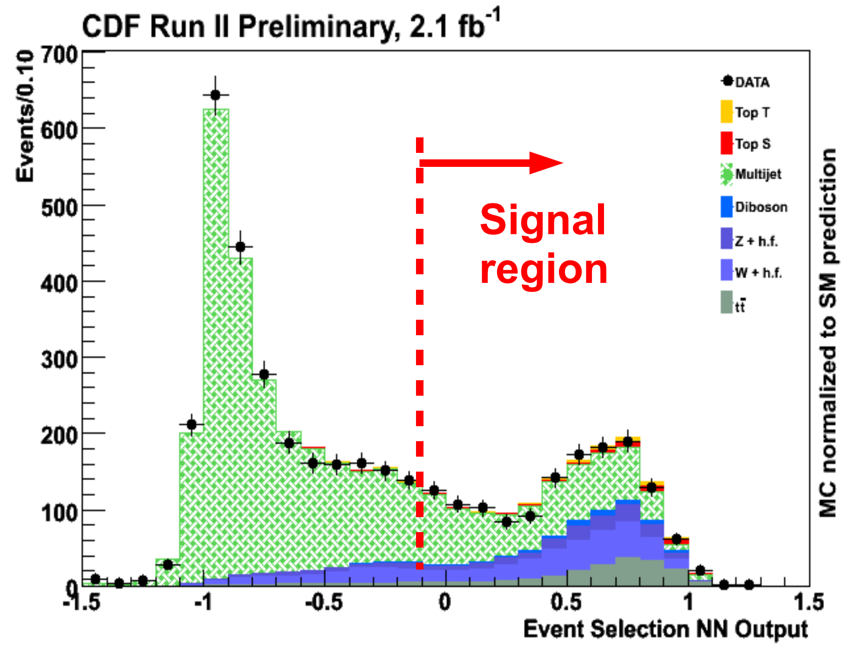
New channel

- Recover non-fiducial leptons and hadronic τ decay
  - Orthogonal to lepton+jets
- MET+ jets trigger

  Huge instrumental background from QCD multijets
- MET>50 GeV and veto leptons
- E_T>35 (25) GeV 1st (2nd) jet
- At least 1 *b*-tag
- NN to suppress multijet bckg

  Signal region: ANN>–0.1

  Control region: ANN<–0.1

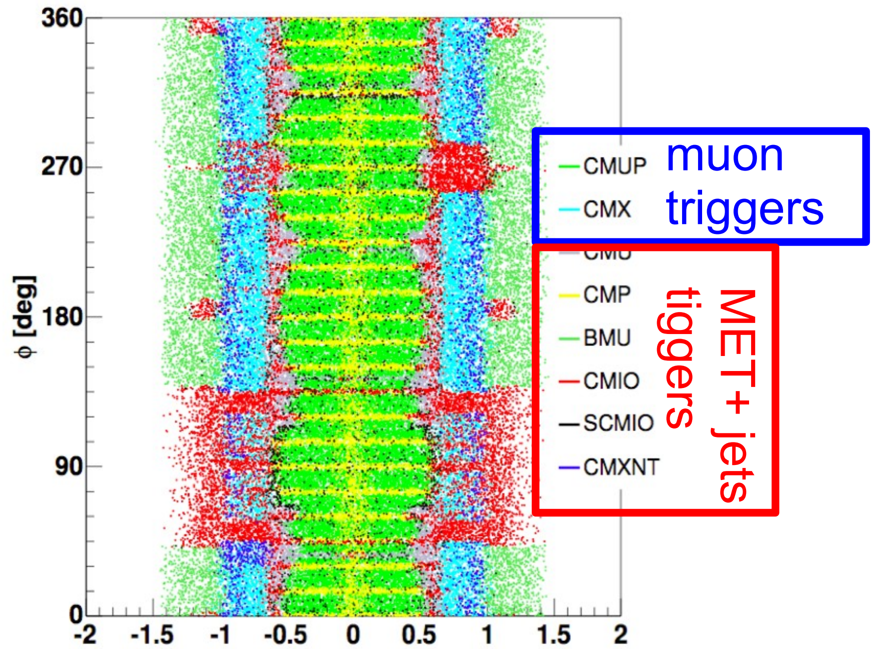


Quantity	Pre-selection	After QCDNN cut	Difference
Signal (S)	75	68	-9%
QCD Background	2960	675	-77%
Total Background (B)	3840	1350	-65%
S/√S+B	1.2	1.8	+50%
S/B	1/50	1/20	+150%

Improvements

- 3.2 fb^{-1} (2.2–2.7 fb^{-1} in summer)
- Extended muon coverage

30% gain in muon acceptance
10-14% gain in sensitivity



- Additional channel MET+ jets

33% increase of acceptance

- Separate s- and t-channel searches

- 2.6 times more data (2.3 fb^{-1})
- 18% larger acceptance
  - **Logical OR of many triggers**
  - **Looser cuts on 2nd jet and muon p_T**
  - **Increased $|\eta|$ for 1st jet (2.5→ 3.4)**
  - **Looser b-tagging requirements for 2 b-tag events**
- Additional cuts to reduce background
- Improved (more detailed) background modeling
  - **Data-based corrections to Alpgen model of W+jets**
- Improved treatment of multijet background

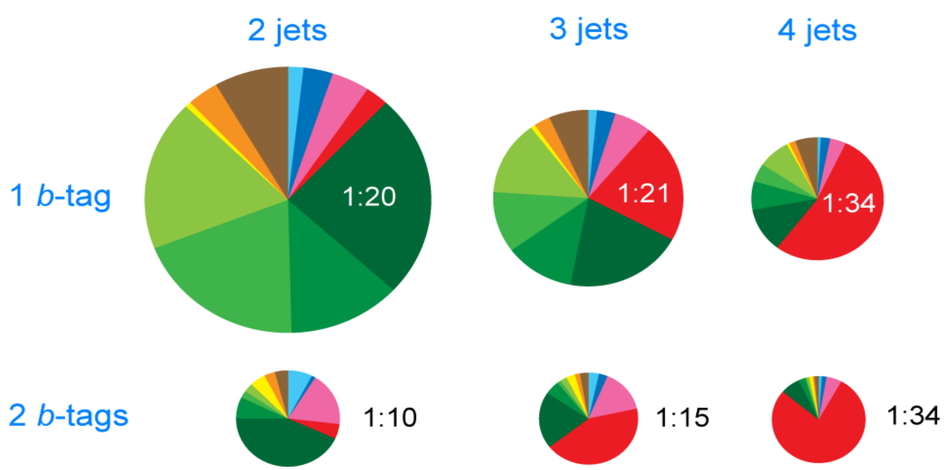
Yields

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

for m_t=170

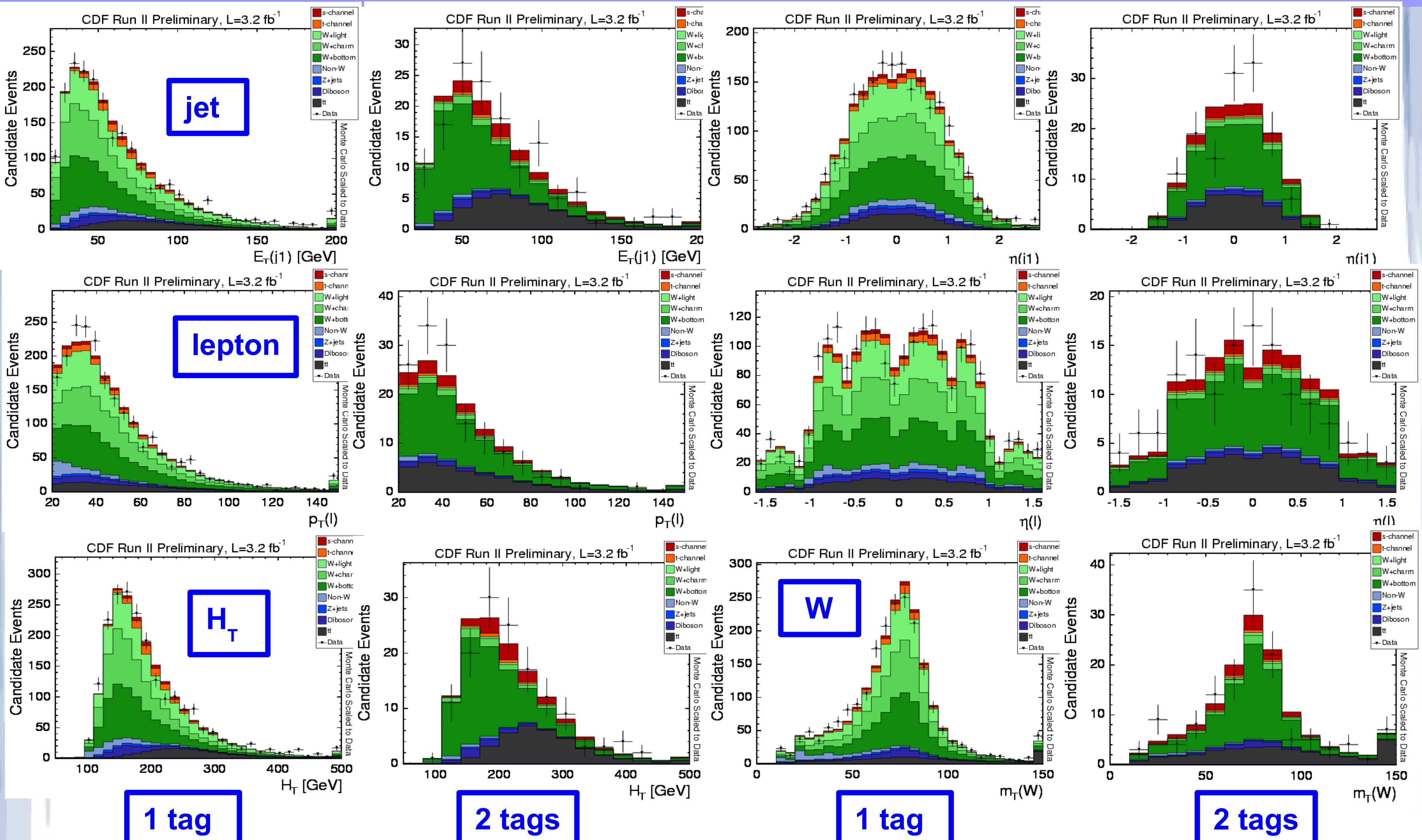
255 events for m_t=175

Process	$\ell + \not{E}_T$ + jets	$\not{E}_T$ + jets
s-channel signal	77.3 ± 11.2	29.6 ± 3.7
t-channel signal	113.8 ± 16.9	34.5 ± 6.1
$W + HF$	1551.0 ± 472.3	304.4 ± 115.5
$t\bar{t}$	686.1 ± 99.4	184.5 ± 30.2
Z+jets	52.1 ± 8.0	128.6 ± 53.7
Diboson	118.4 ± 12.2	42.1 ± 6.7
QCD+mistags	777.9 ± 103.7	679.4 ± 27.9
Total prediction	3376.5 ± 504.9	1404 ± 172
Observed	3315	1411



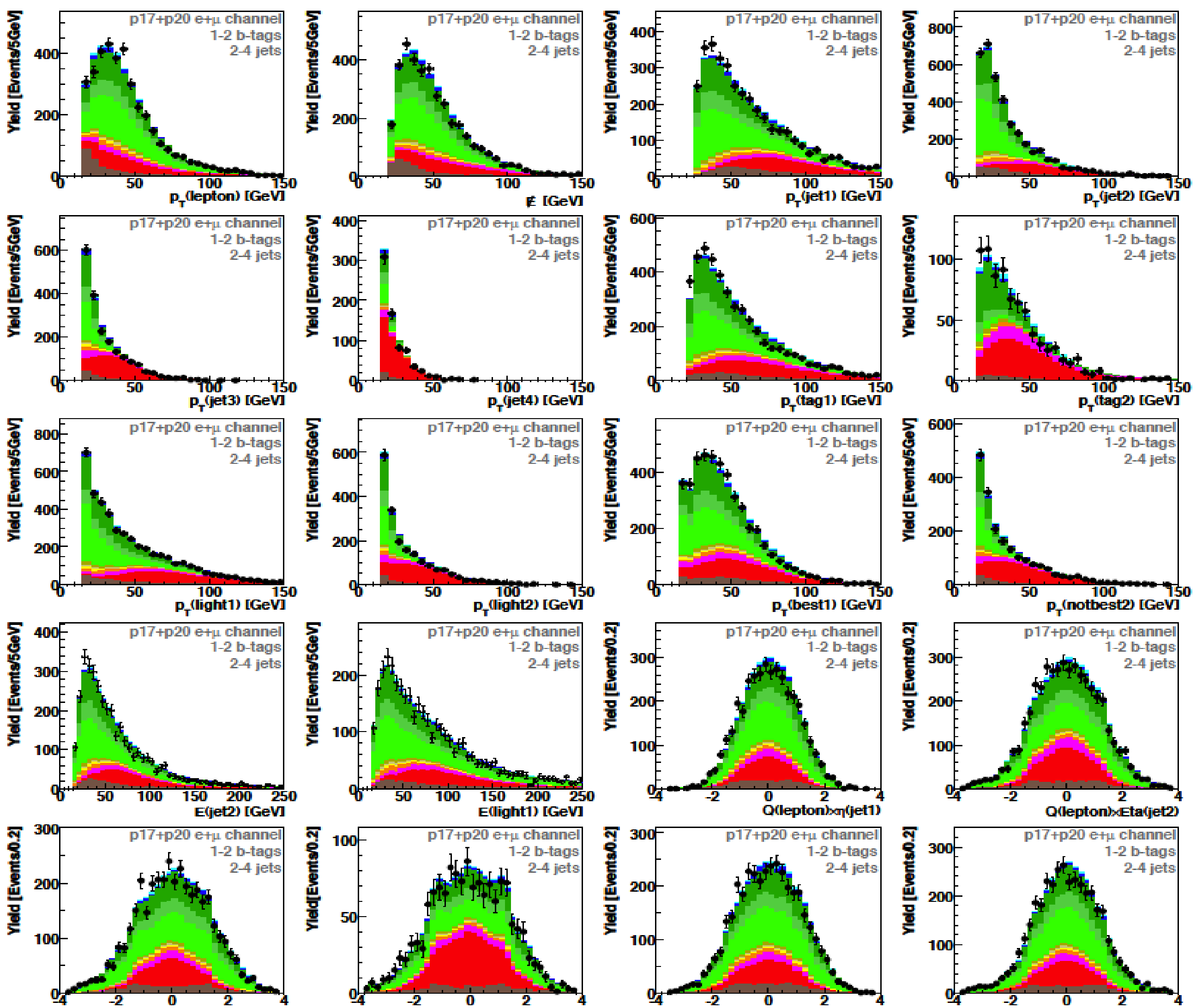
- S:B – 1:10 to 1:34 depending on number of jets and tags
- Most powerful - 2 jets/1 tag
- Keep channels separately in the analysis to take advantage of different S:B and different background composition

Background model validation



Background model validation

SINGLE OBJECT KINEMATICS



- Check thousands of distributions to verify background model before and after tagging
- Several classes of variables used in discriminants
  - Single object kinematics
  - Event kinematics
  - Jet reconstruction
  - Top quark reconstruction
  - Angular correlations

Systematics

- Statistically limited measurement
- But systematics is important
- Affects normalization and shapes

- Estimated for each background and signal source in each analysis channel
- Background uncertainty dominates

Systematic Uncertainties

Ranked from Largest to Smallest Effect on Single Top Cross Section

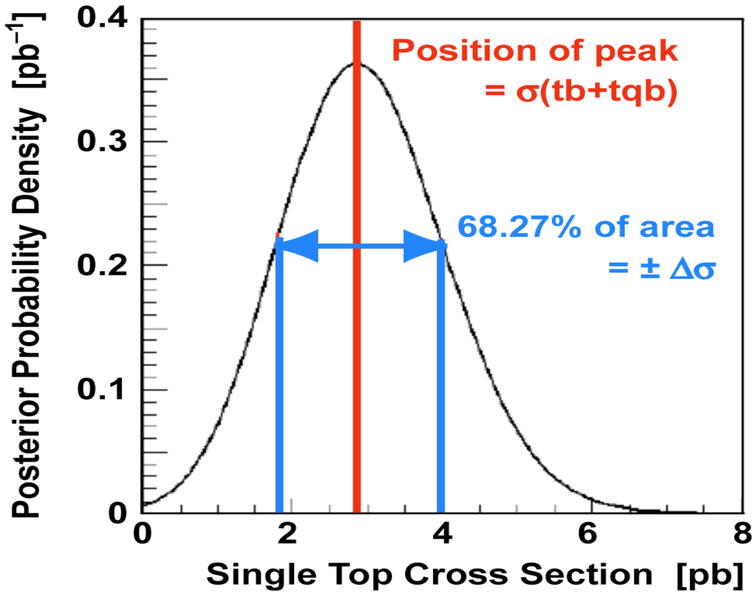
DØ 2.3 fb^{-1}

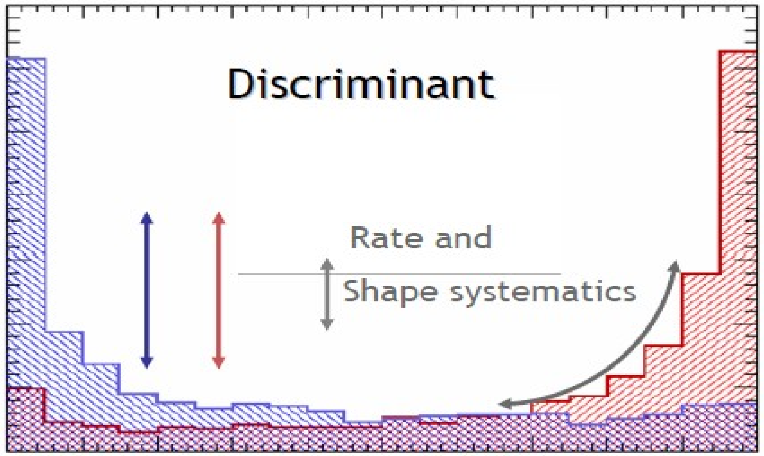
Larger terms	
b-ID tag-rate functions (includes shape variations)	(2.1–7.0)% (1-tag) (9.0–11.4)% (2-tags)
Jet energy scale (includes shape variations)	(1.1–13.1)% (signal) (0.1–2.1)% (bkgd)
W+jets heavy-flavor correction	13.7%
Integrated luminosity	6.1%
Jet energy resolution	4.0%
Initial- and final-state 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 Uncertainty	Rate	Shape
Jet Energy Scale	0...10%	✓
Initial + Final State Radiation	0...15%	✓
Parton Distribution Functions	2...3%	✓
Monte Carlo Generator	1...5%	
Event Detection Efficiency	0...9%	
Luminosity	6%	
Neural Net B-tagger		✓
Mistag Model		✓
Q^2 scale in ALPGEN MC		✓
Input variable mismodeling		✓
Wbb+Wcc normalization	30%	
Wc normalization	30%	
Mistag normalization	17...29%	
ttbar normalization & m_{top}	23%	✓

Cross section

- Discriminant outputs (from each analysis channel separately) are used to measure cross section
- Build Bayesian probability density with flat nonnegative prior for the cross section
- Peak of posterior distribution gives the cross section, 68% interval gives the uncertainty
- Shape and normalization systematic uncertainties are treated through nuisance parameters with Gaussian distribution
  - Correlations are properly taken into account

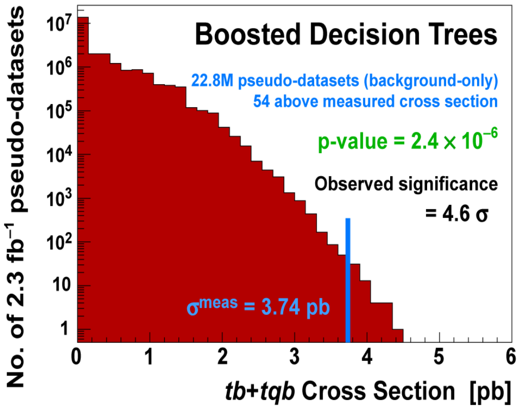




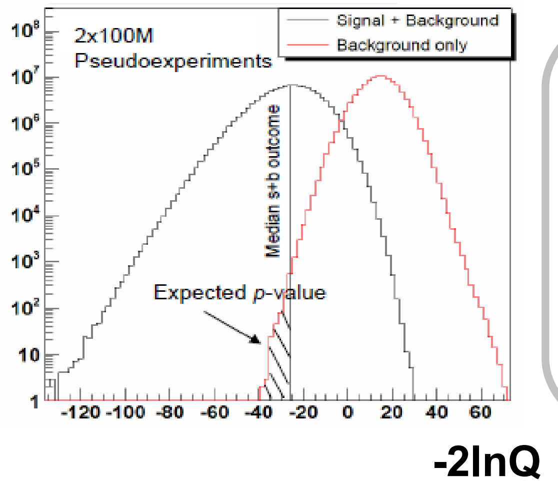
Statistical analysis

- Build ensembles of pseudo-data
  - Includes signal and background events or background only
  - Includes all systematic uncertainties
- Purpose before data
  - Test performance of different methods
  - Measure expected cross section uncertainty
  - Expected significance
- With data
  - Consistency of the measured cross section with the SM
  - Observed significance

Significance – probability of the upward background fluctuation that gives observed result in data



***p*-value:** **Fraction of zero-signal ensemble datasets that give $\sigma \geq \sigma_{meas}$**



Likelihood ratio technique

$$Q = \frac{P(\text{data} \mid s+b, \hat{\theta})}{P(\text{data} \mid b, \hat{\hat{\theta}})}$$

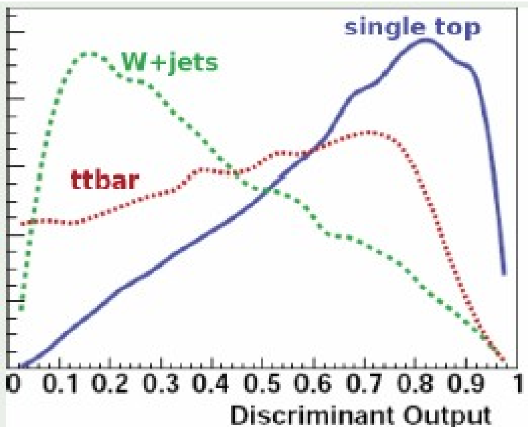
Signal from background separation

Signal model

Background model

Data

Multivariate separation technique

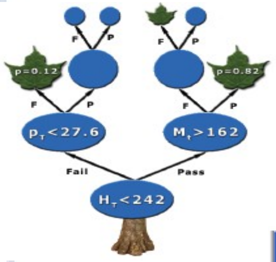


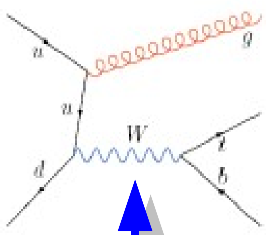
background signal

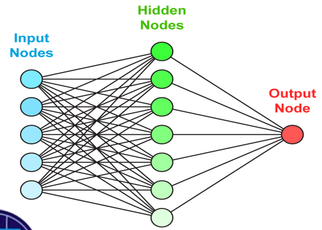
Approximations of discriminant:

$$D_S(\vec{x}) = P(S|\vec{x}) = \frac{P_S(\vec{x})}{P_S(\vec{x}) + P_B(\vec{x})}$$

Boosted decision trees	Matrix elements	Neural Networks	Likelihood function







$$\mathcal{L}^{signal} = \frac{\prod_{i=1}^{n_{val}} P_i^{signal}}{\prod_{i=1}^{n_{val}} P_i^{signal} + \sum_{m=1}^{n_{bkg}} \prod_{i=1}^{n_{val}} P_i^{m} \times a^{m}}$$

Combined analysis

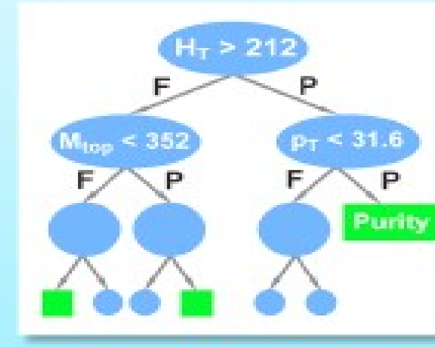
t-channel + s-channel = one single-top signal
cross section ratio is fixed to SM value.
important for „observation“ and $|V_{tb}|$ test

Separate search

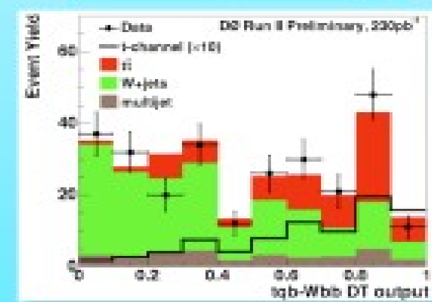
consider t-channel and s-channel as separate processes
important for new physics search

Boosted Decision Trees method

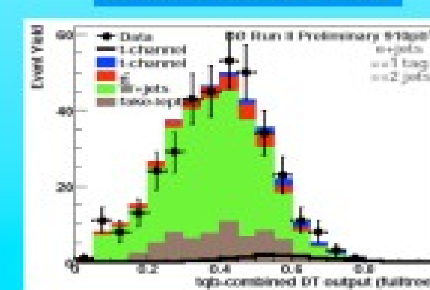
- Idea: recover events that fail criteria in cut-based analyses
- Start with all events (first node)
  - **For each variable, find splitting value with best separation**
  - **Select variable and cut: produce Pass and Failed branches**
- Repeat recursively on each node
- Stop when no improvement or too few (100) events left
- Terminal node: leaf with purity = $N_S/(N_S+N_B)$
- Decision tree output for each event = leaf purity value (closer to 0 for background, closer to 1 for signal)
- Boosting – averaging over many trees – improves stability and performance by ~20% by diluting discrete output
- Adaptive boosting algorithm with 50 cycles
- Run independent MC and data through the tree to obtain the result



Before boosting



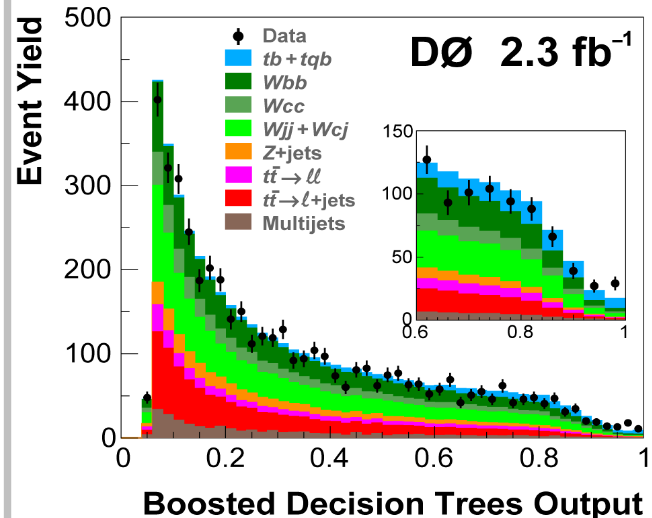
After boosting

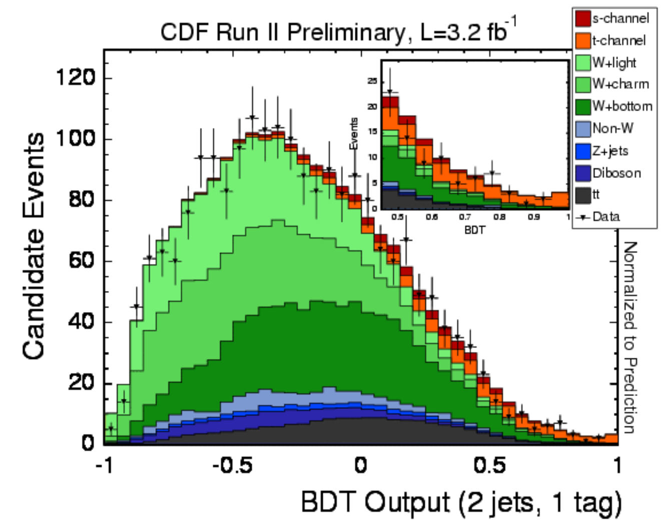


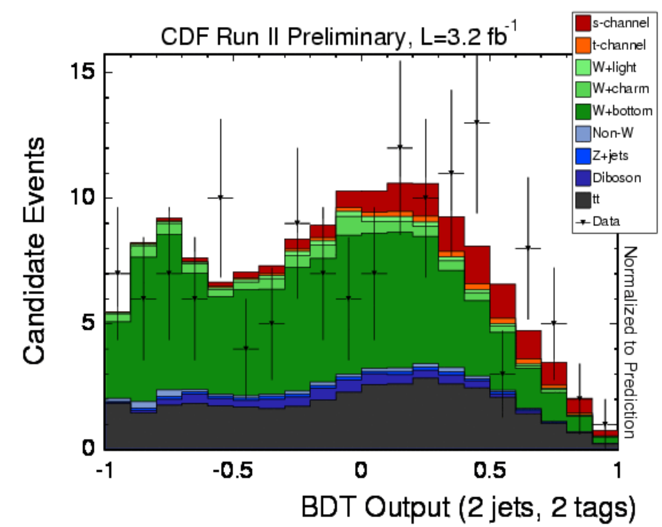
Boosted Decision Trees results

Advantages:

- Fast to train
- Not degraded by the addition of more input variables
- no need to optimize the choice
- use all sensitive variables with good agreement to data
- Input variables 64 (D0), 20 (CDF)



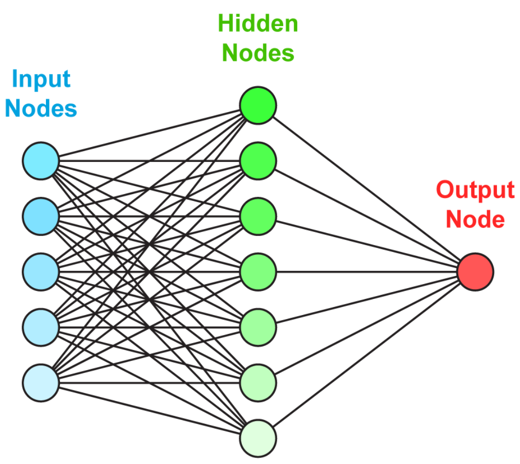




	$\mathcal{L}$ [fb^{-1}]	Significance Exp.	Significance Obs.	σ_{s+t} [pb]
DØ	2.3	4.3σ	4.6σ	$\mathbf{3.7^{+1.0}_{-0.8}}$
CDF	3.2	5.2σ	3.5σ	$\mathbf{2.1^{+0.7}_{-0.6}}$

Neural Networks method

- A Neural Network (NN) is an Interconnected group of nodes. It can be used to model complex relationships between inputs and outputs, or to find patterns in data
- For this analysis:
  - Inputs: variables with high discriminating power
  - Output: probability for the event to be signal

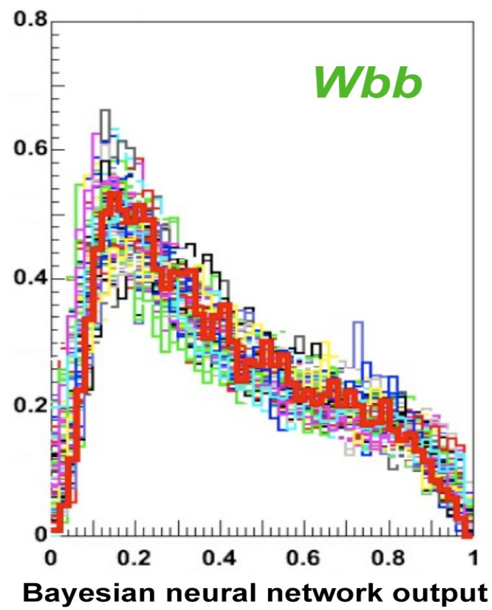


NEUROBAYES

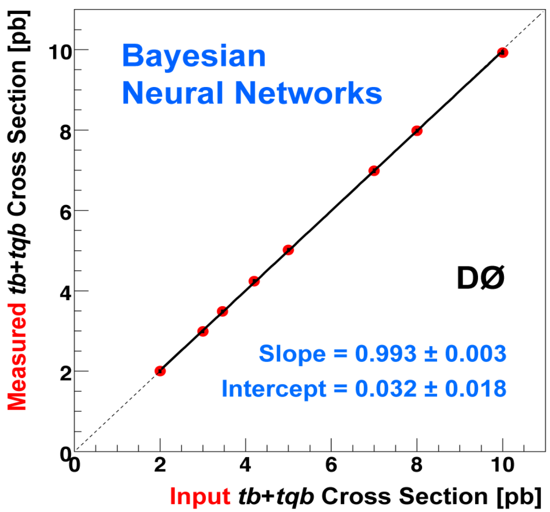
- 3 layer feed-forward network with complex and robust preprocessing of input variable
- Bayesian regularization to avoid over-training
- 4 networks, each divided into 2 channels based on trigger
- 14 variables

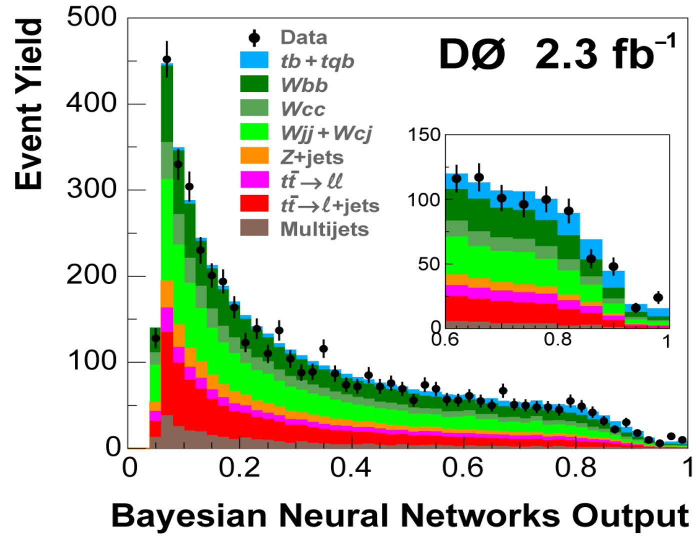
Bayesian NN

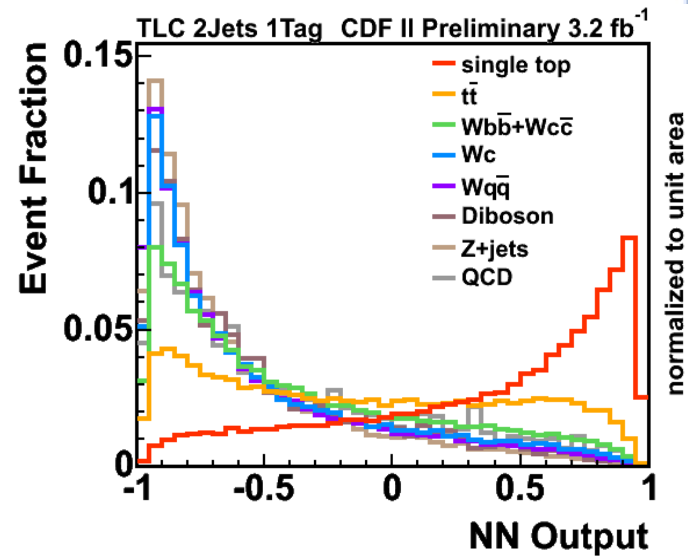
- Weighted average over hundreds of networks
- Better stability
- Immune to over-training
- 18-25 input variables



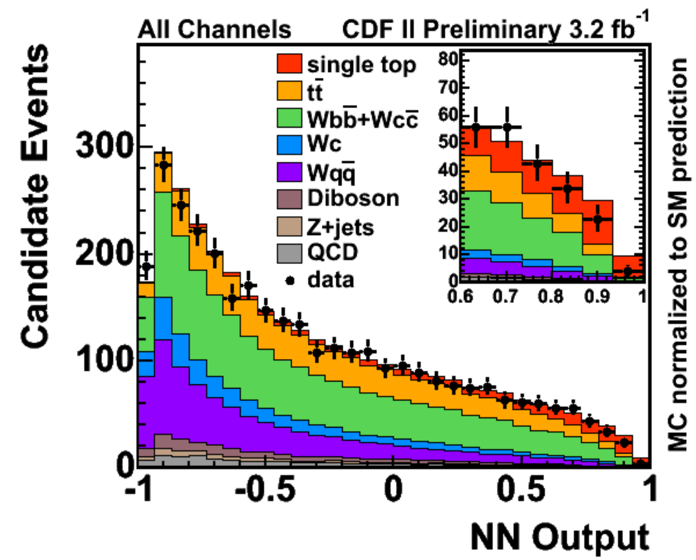
Neural Networks results







	$\mathcal{L}$ [fb^{-1}]	Significance Exp.	Obs.	σ_{s+t} [pb]
DØ	2.3	4.1σ	5.2σ	$4.7^{+1.2}_{-0.9}$
CDF	3.2	5.2σ	3.5σ	$1.8^{+0.6}_{-0.6}$



Matrix Elements method

- Given 4-vectors of reconstructed lepton and jets compute event probability density for signal and background hypothesis

$$P(\vec{x}) = \frac{1}{\sigma} \times \frac{\partial \sigma}{\partial \vec{x}}$$

$$d\sigma(\vec{x}) = \sum_{i,j} \int d\vec{y} \left[f_i(q_1, Q^2) dq_1 \times f_j(q_2, Q^2) dq_2 \times \frac{\partial \sigma_{hs,ij}(\vec{y})}{\partial \vec{y}} \times W(\vec{x}, \vec{y}) \times \Theta_{\mathrm{Parton}}(\vec{y}) \right]$$

Parton distribution functions for intial parton i,j carrying momentum q

Differential cross section for the hard scatter. Uses leading-order matrix element

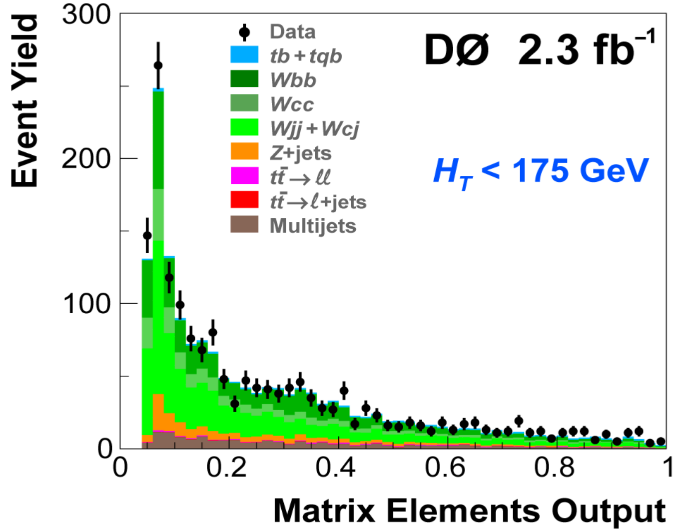
Transfer function which relates observed state in the detector (x) to original partons (y)

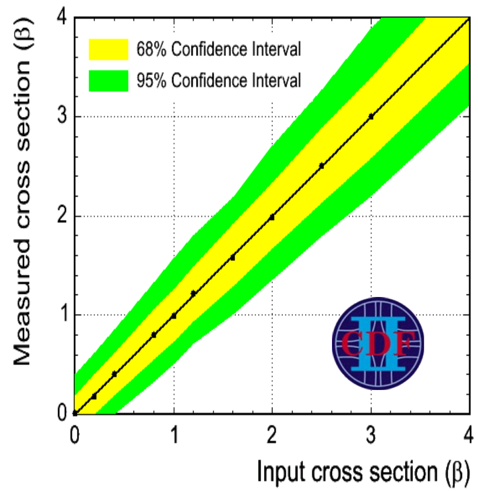
- Matrix elements:
  - D0: 2 jets: tb, tq, Wb$\bar{b}$, Wcg, Wgg, WW, WZ, ggg, t$\bar{t}$
  - D0: 3 jets: tbg, tqg, tqb, Wb$\bar{b}$g, t$\bar{t}$, Wugg
  - CDF: tb, tq, tbg, tqg, Wb$\bar{b}$, Wcg, Wgg, t$\bar{t}$, Wc$\bar{c}$

- Improve performance by
  - D0: splitting samples by H_T into t$\bar{t}$ and W+jets dominated regions
  - CDF&D0: weighting events by jet flavor probability

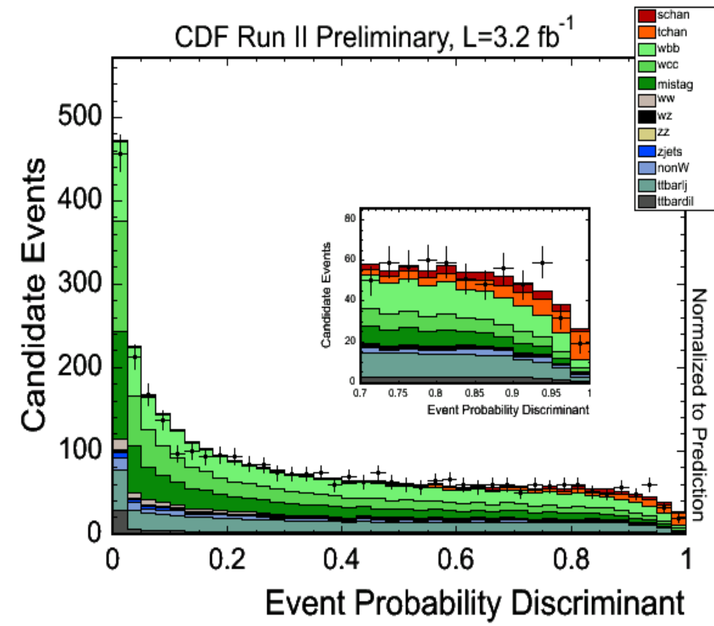
$$EPD = \frac{b \cdot P_{singletop}}{b \cdot P_{singletop} + b \cdot (P_{Wb\bar{b}} + P_{t\bar{t}}) + (1-b) \cdot (P_{Wc\bar{c}} + P_{Wcj} + P_{Wgg})}$$

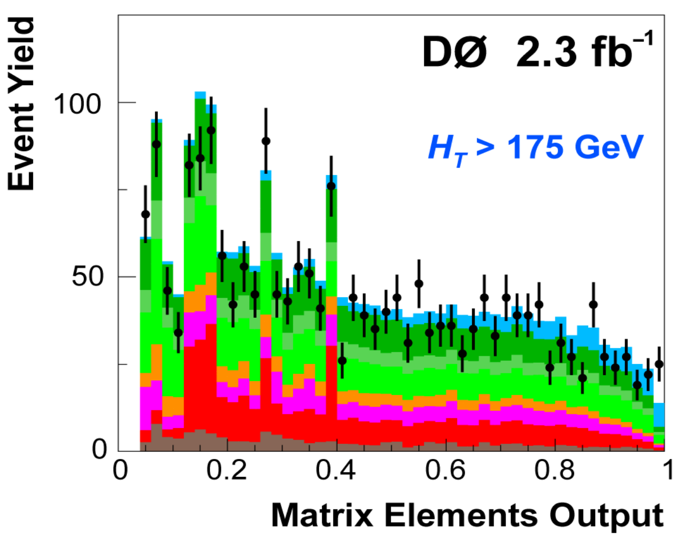
Matrix Elements results





Linearity test



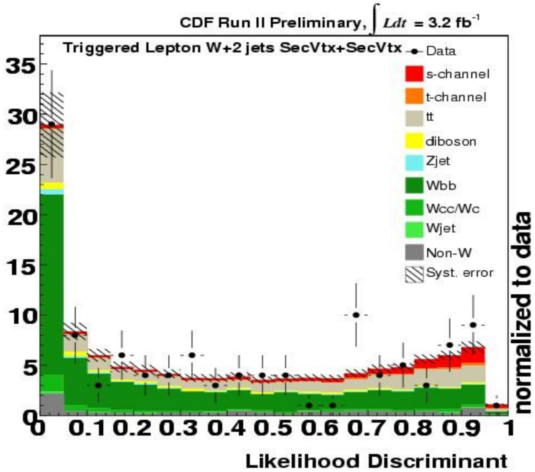


	$\mathcal{L}$ [fb^{-1}]	Significance Exp.	Significance Obs.	σ_{s+t} [pb]
DØ	2.3	4.1σ	5.0σ	**$4.3^{+1.0}_{-1.2}$**
CDF	3.2	4.9σ	4.3σ	**$2.5^{+0.7}_{-0.6}$**

Multivariate Likelihood Function

$$\mathcal{L}^{signal} = \frac{\prod_{i=1}^{n_{var}} P_i^{signal}}{\prod_{i=1}^{n_{var}} P_i^{signal} + \sum_{m=1}^{n_{bkg}} \prod_{i=1}^{n_{var}} P_i^{m} \times a^{m}}$$

s-channel

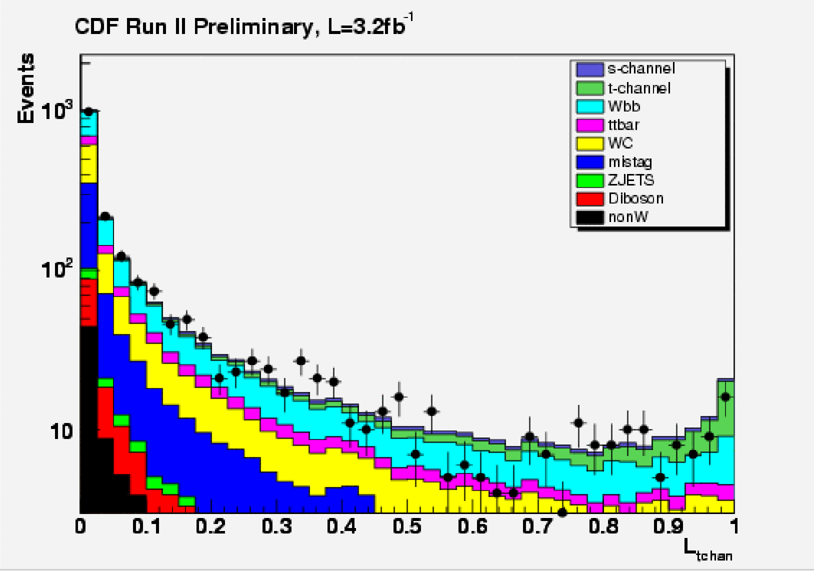


- **Events with 2 *b*-tags**
- **2 classes: $t\bar{t}$ and W+HF**
- **9-10 variables**

	$\mathcal{L}$ [fb^{-1}]	Significance Exp.	Significance Obs.	σ_{s+t} [pb]
LF	3.2	4.0σ	2.4σ	$\mathbf{1.6}^{+1.0}_{-0.8}$
s-channel	3.2	1.1σ	2.0σ	$\mathbf{1.5}^{+0.9}_{-0.8}$

Combine many variables into a likelihood function

t-channel likelihood

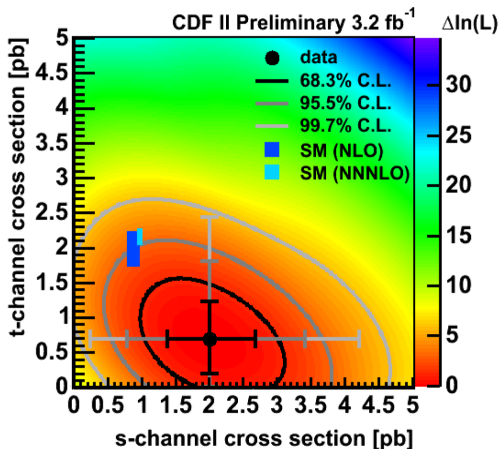


- Signal template built for t-channel
- 4 background classes: $Wb\bar{b}$, $Wc\bar{c}$/ Wc, $t\bar{t}$, mistags
- 7 (10) variables in 2 (3) jet bin to isolate t-channel contribution

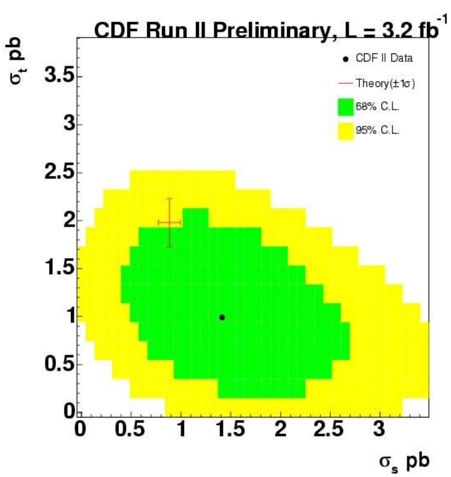
More results...

Separate search using NN

- 5 networks, 2 outputs combined into 2D discriminant in 2j1tag channel

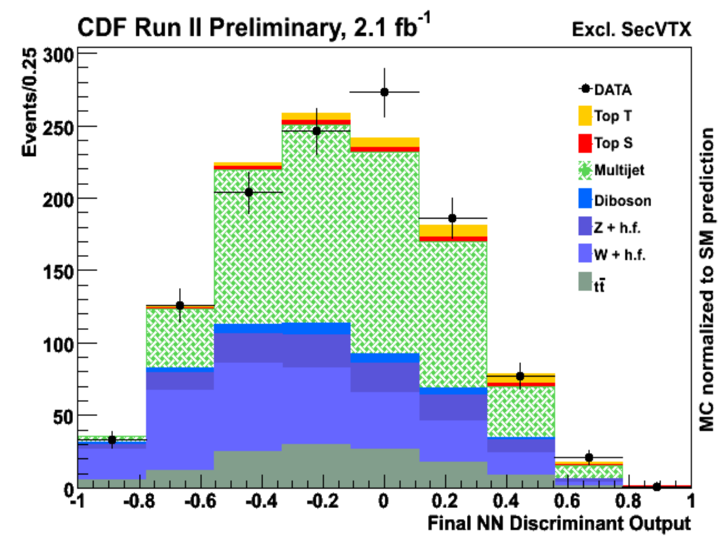


	σ (pb)	
s-ch	$2.0^{+0.7}_{-0.6}$	1.4
t-ch	$0.7^{+0.5}_{-0.5}$	1.0



Simultaneous fit of σ_s and σ_t (LF)

MET+ jets combined search



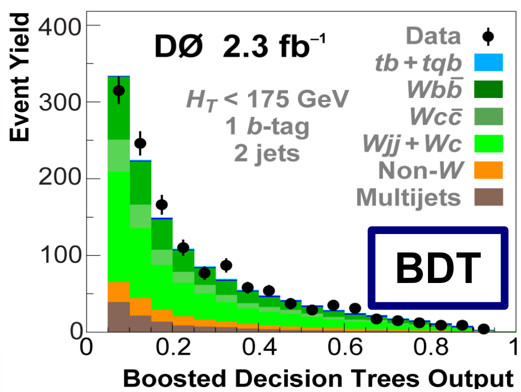
	$\mathcal{L}$ [fb^{-1}]	Significance Exp.	Significance Obs.	σ_{s+t} [pb]
$\not{E}_T$+jets	2.1	1.4σ	2.1σ	$4.9^{+2.5}_{-2.2}$

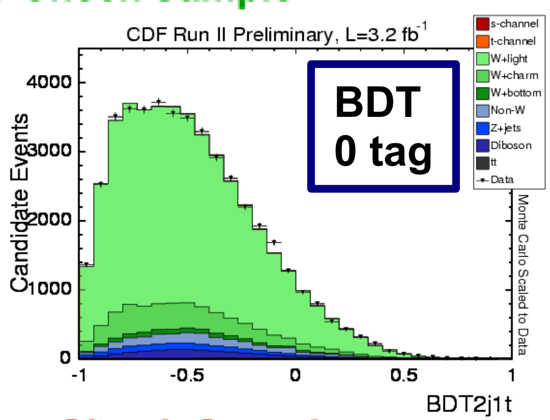
Cross check samples

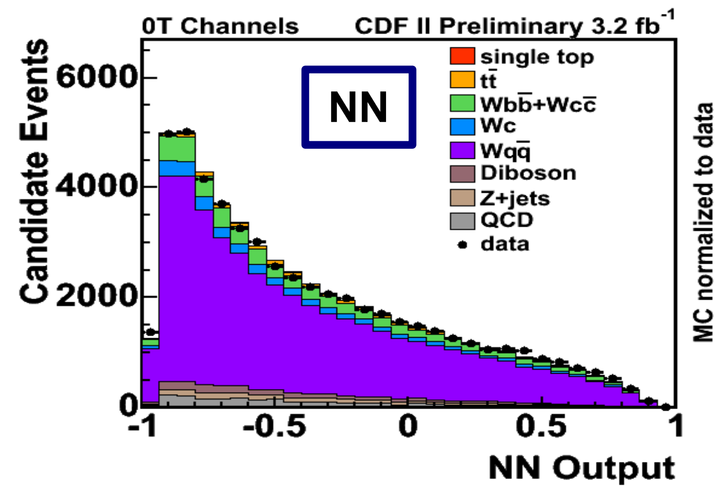
Cross checks of discriminant performance using samples depleted in signal

- Untagged (high statistics)
- W+jets (nj=2, 1 *b*-tag, H_T(l,ν,jets) < 175 GeV)
- $t\bar{t}$ dominated (nj=4, ≥1 *b*-tag, H_T > 300 GeV)

W+Jets Cross-Check Sample

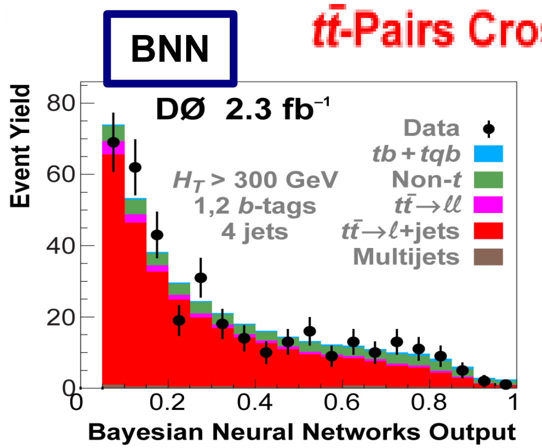


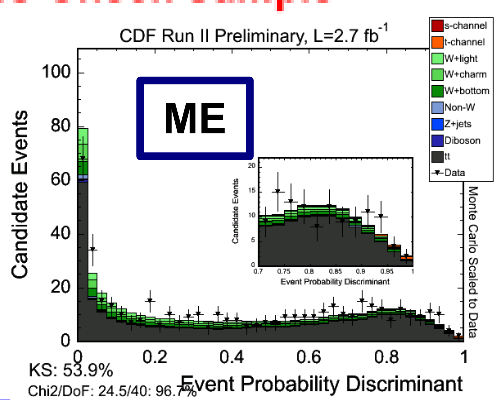


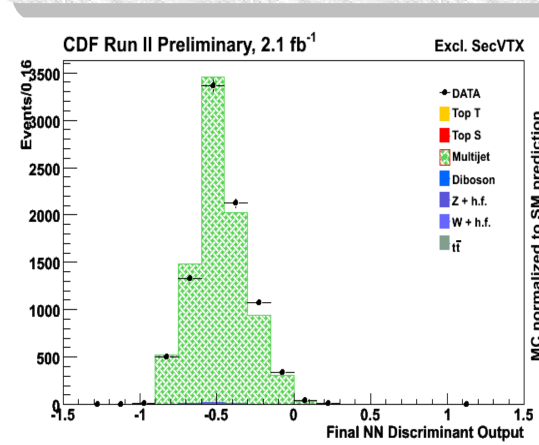


MET + jets
QCD enriched sample

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







Combination method

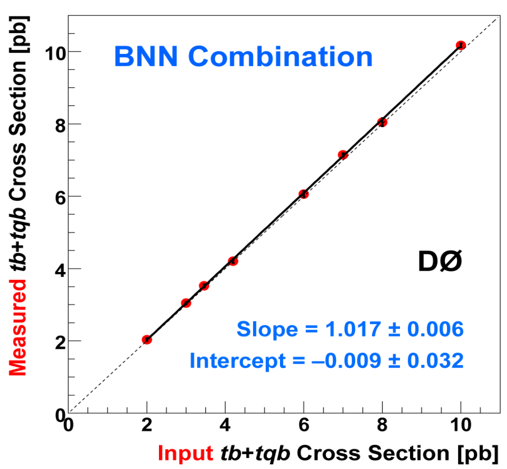
- Even though all (majority) D0 (CDF) MVA analyses use the same data they are not 100% correlated
- Choose a priori to quote combination result as main

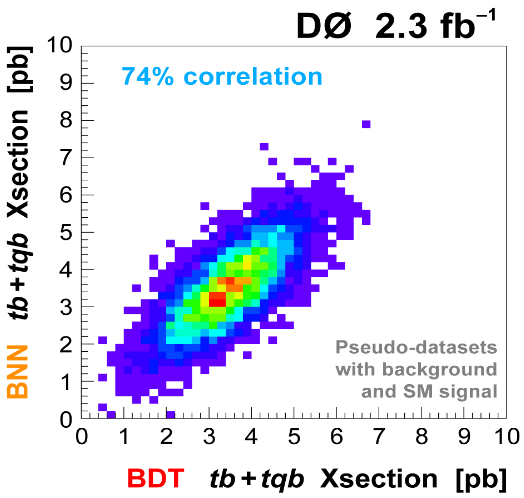
Technique: NeuroEvolution of Augmenting Topologies

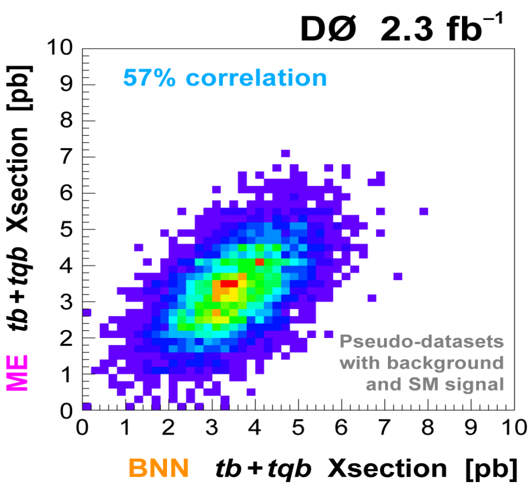
- Optimizes network topology, inter-node weights, output binning
- NN trained to give the best expected *p*-value
- Inputs:
  5 l+jets discriminants (BDT, ME, NN, LF, SLF), and MET+ jets discriminant (8+3=11 channels)
- Sensitivity: 5.2σ→>5.9σ

Technique: Bayesian Neural Network

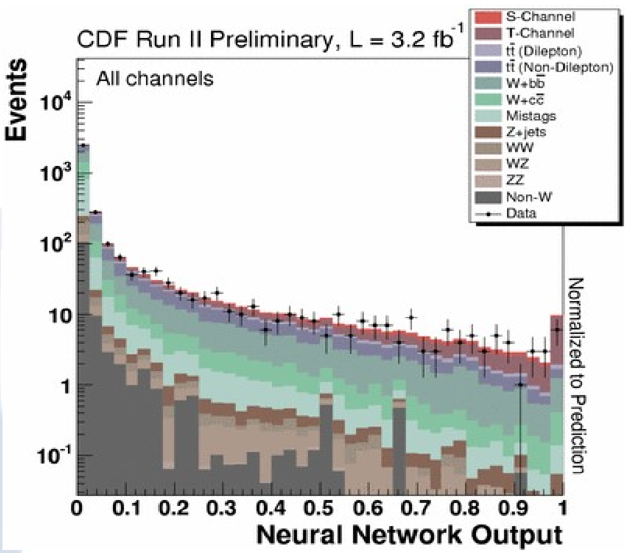
- Inputs: **3 discriminant output distributions (BDT, ME, NN)**
- Sensitivity: 4.3σ → 4.5σ

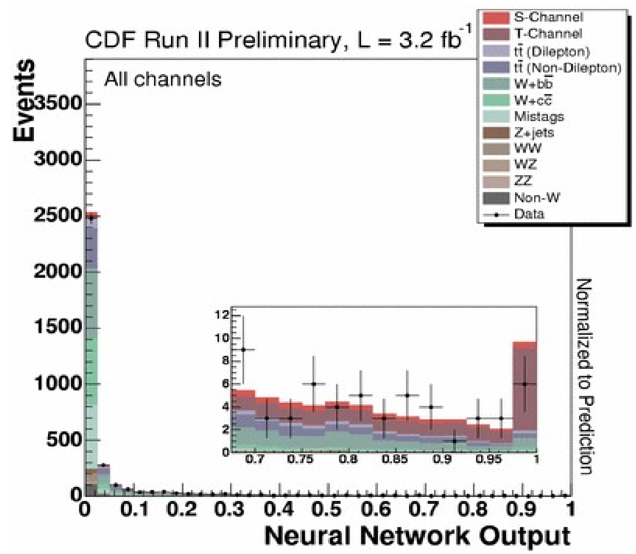




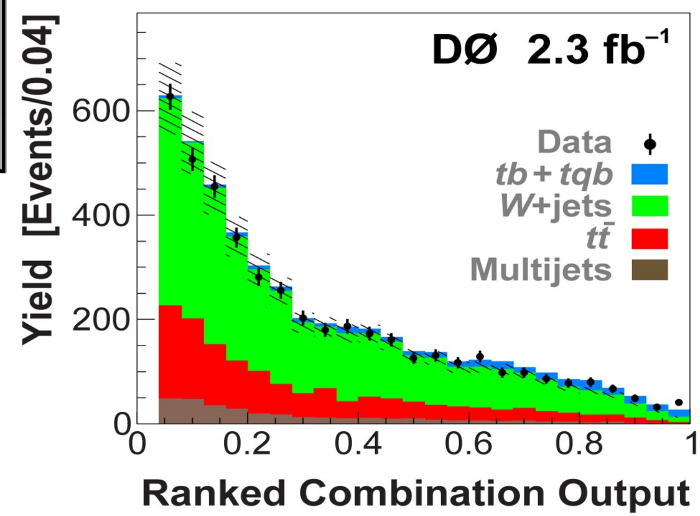


Combination results

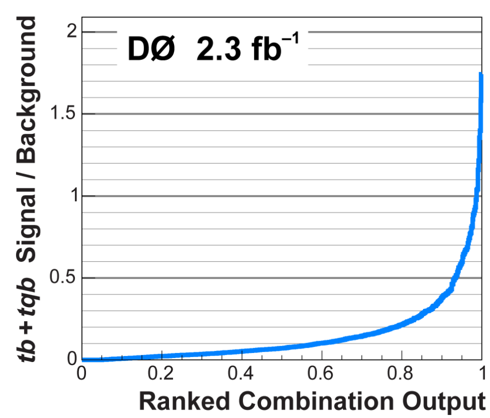




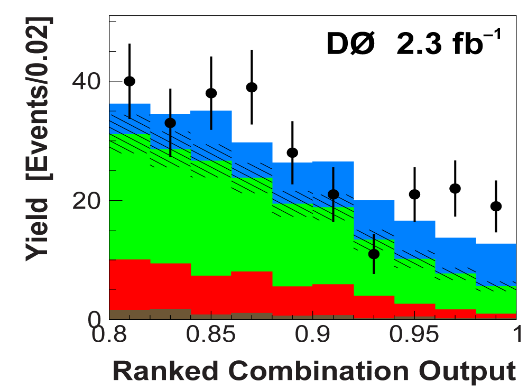
Final Discriminant



S/B Ratio



Signal Region



Observation !!

	$\mathcal{L}$ [fb^{-1}]	Significance Exp.	Significance Obs.	σ_{s+t} [pb]
DØ	2.3	4.5σ	5.0σ	$3.9^{+0.9}_{-0.9}$
CDF	3.2	5.9σ	5.0σ	$2.3^{+0.6}_{-0.5}$

Signal normalized to measured cross section

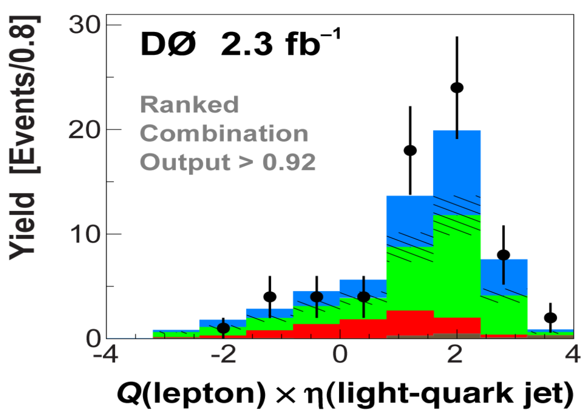
Can we see it?

- Look at high discriminant regions

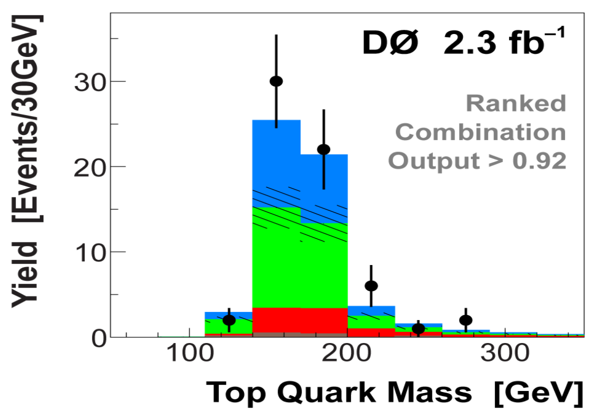
BNN combined D >0.92, all channels

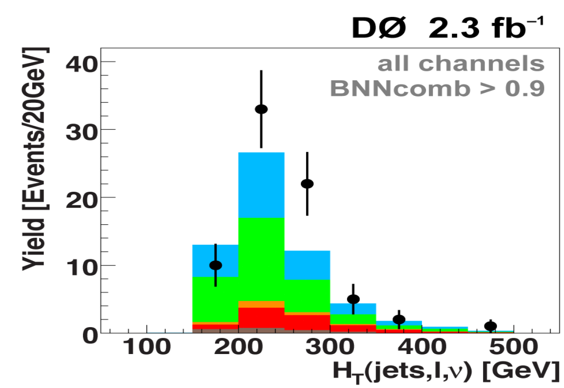
Signal normalized to measured σ

High Signal Region – $Q \times \eta$



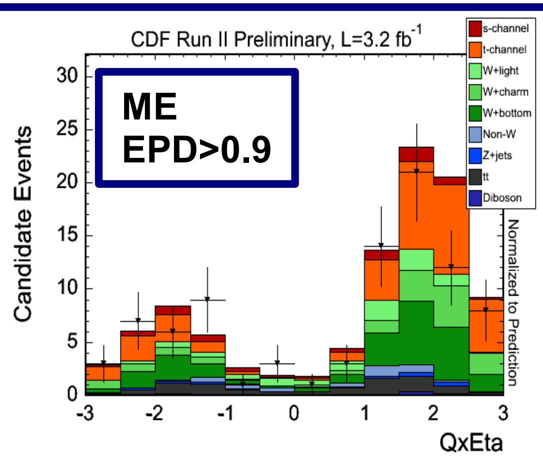
High Signal Region – m_{top}

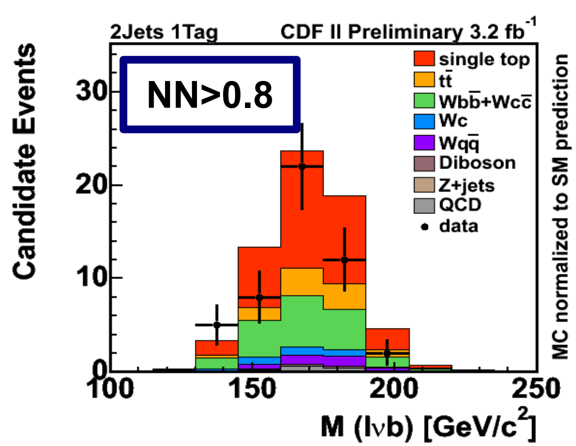


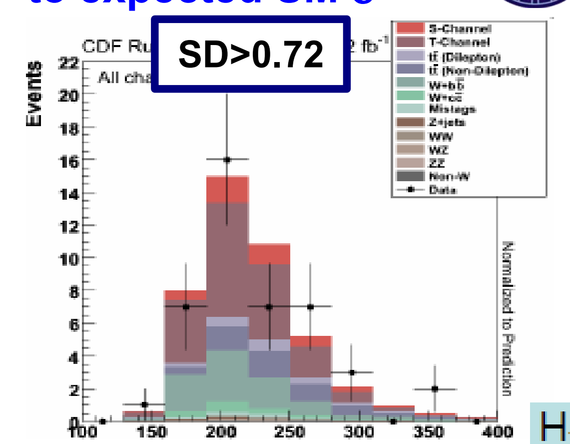


The most sensitive channel: 2 jets, 1 b-tag

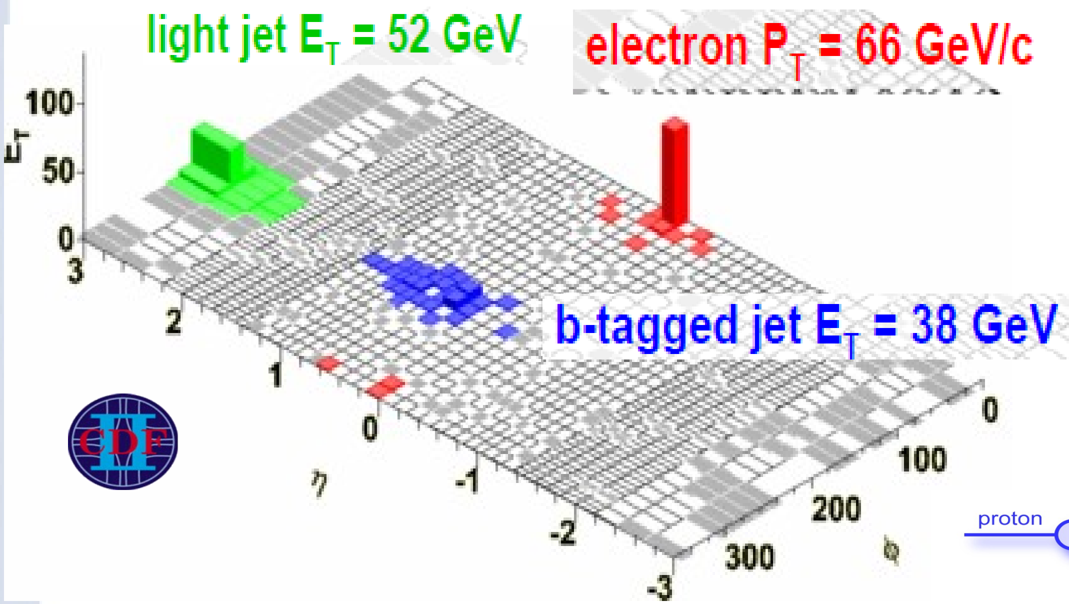
Signal normalized to expected SM σ

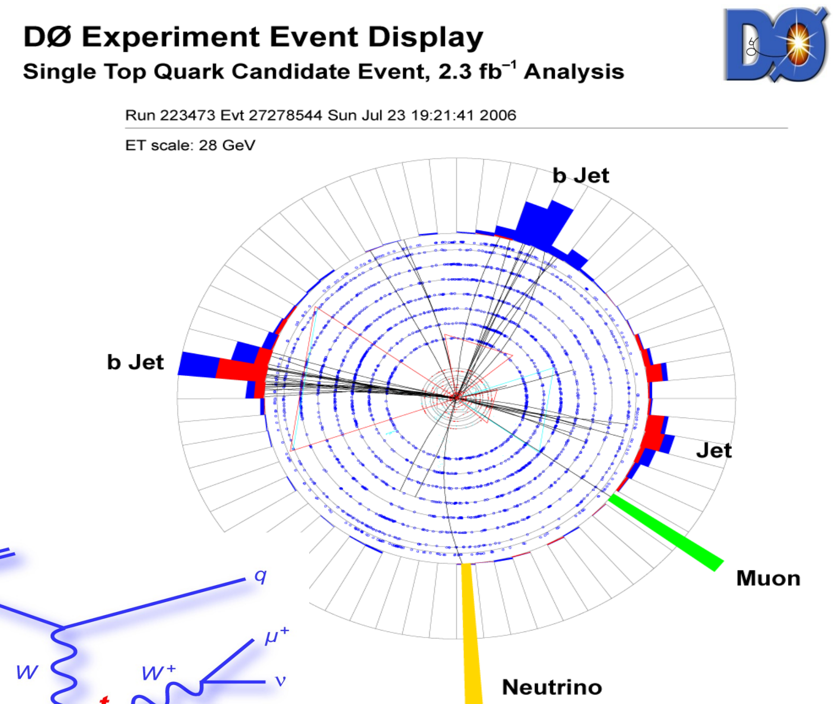


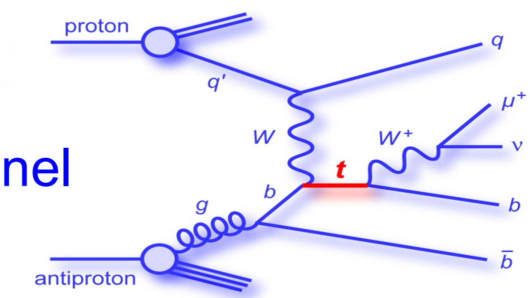




Events



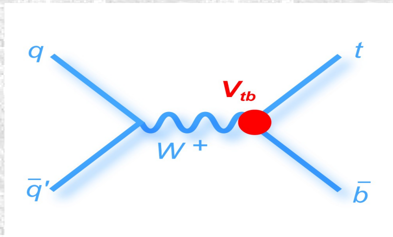




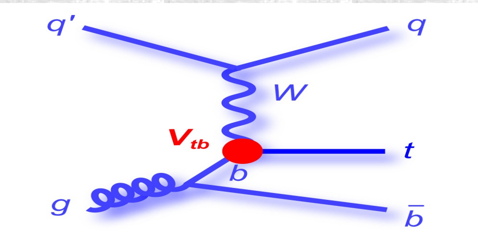
t-channel

Run	Event	Lepton	KIT NN	BDT	LF	ME	NN	SD	H_T	$M_{l\nu b}$	$Q \times \eta$
148916	792764	CEM	0.94	0.76	0.94	0.97	0.94	0.99	219.0	189.7	-2.15
206282	3294678	CMUP	0.99	0.76	0.80	0.98	0.94	1.0	307.3	178.0	1.31
229936	4412760	CMX	1.00	0.95	1.00	0.99	0.97	1.0	221.0	171.1	2.03
242557	1564229	CEM	0.54	0.85	1.00	0.99	0.93	1.0	189.8	164.5	2.84
262776	4920497	CEM	0.86	0.95	1.00	1.00	0.92	1.0	191.8	160.8	2.67

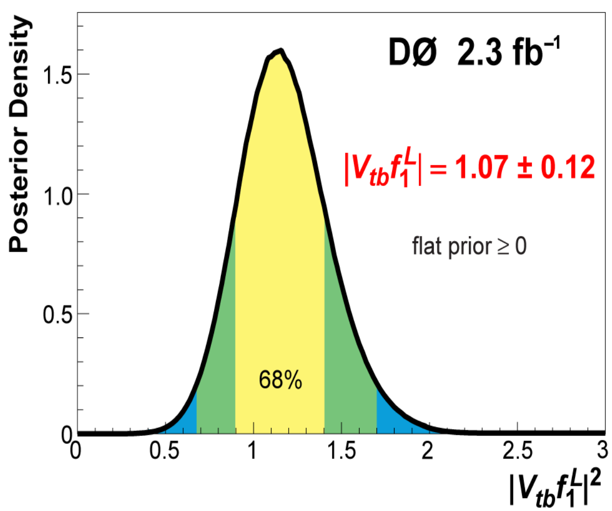
Measurement of $|V_{tb}|$



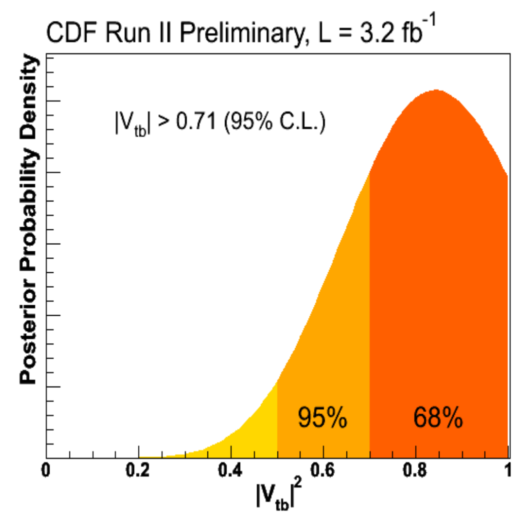
$$|V_{tb,meas}|^2 = \frac{\sigma_{meas}}{\sigma_{SM}} |V_{tb,SM}|^2$$



- Assume $|V_{td}|^2+|V_{ts}|^2<<|V_{tb}|^2$, SM (V–A) and CP conserving Wtb vertex
- No assumption on the number of quark families or CKM unitarity



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%



$|V_{tb}f_1^L|=1.07\pm0.12(\text{sys+th})$
$|V_{tb}|>0.78$ at 95% CL

$|V_{tb}|=0.91\pm0.11(\text{sys})\pm0.07(\text{th})$
$|V_{tb}|>0.71$ at 95% CL

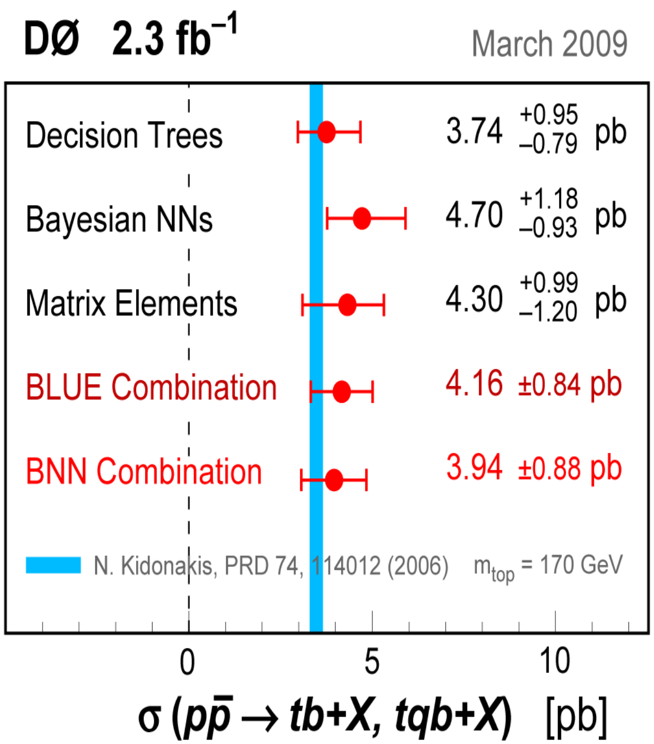
Summary

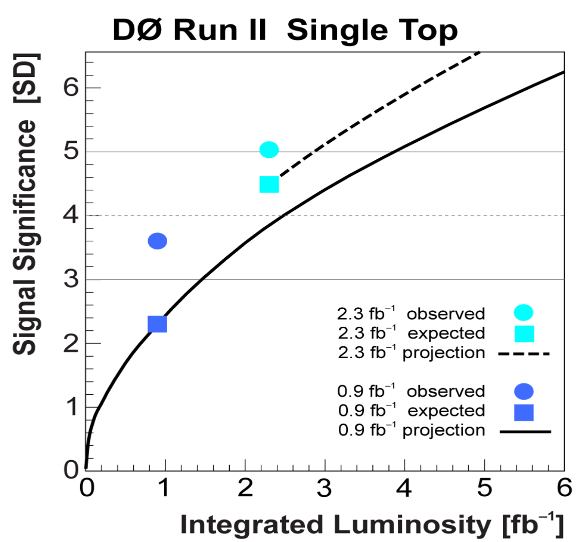
- Single top quark production has been observed at Tevatron by CDF and D0 with signal significance of 5σ
- Both cross section and $|V_{tb}|$ measurements agree with SM

Submitted to PRL

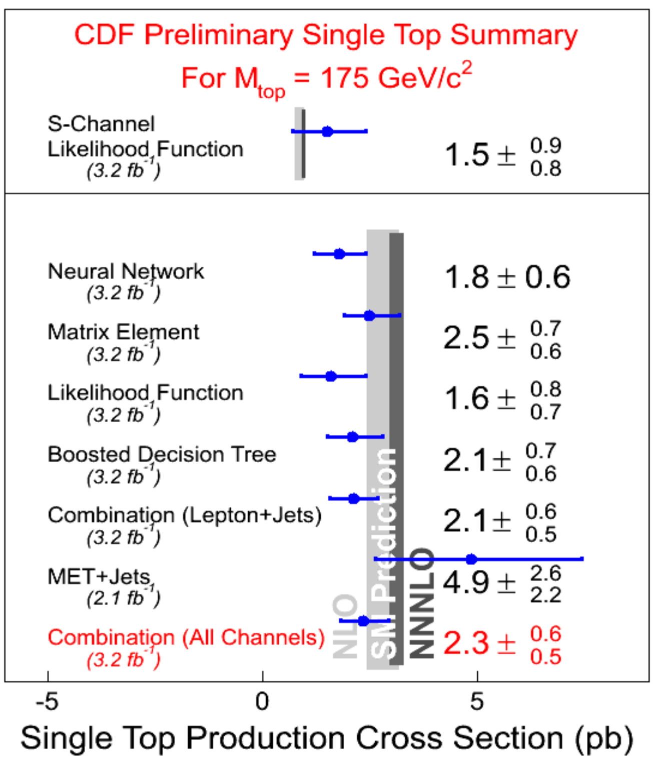
ArXiv:0903.0850

ArXiv:0903.0885





D0 projection

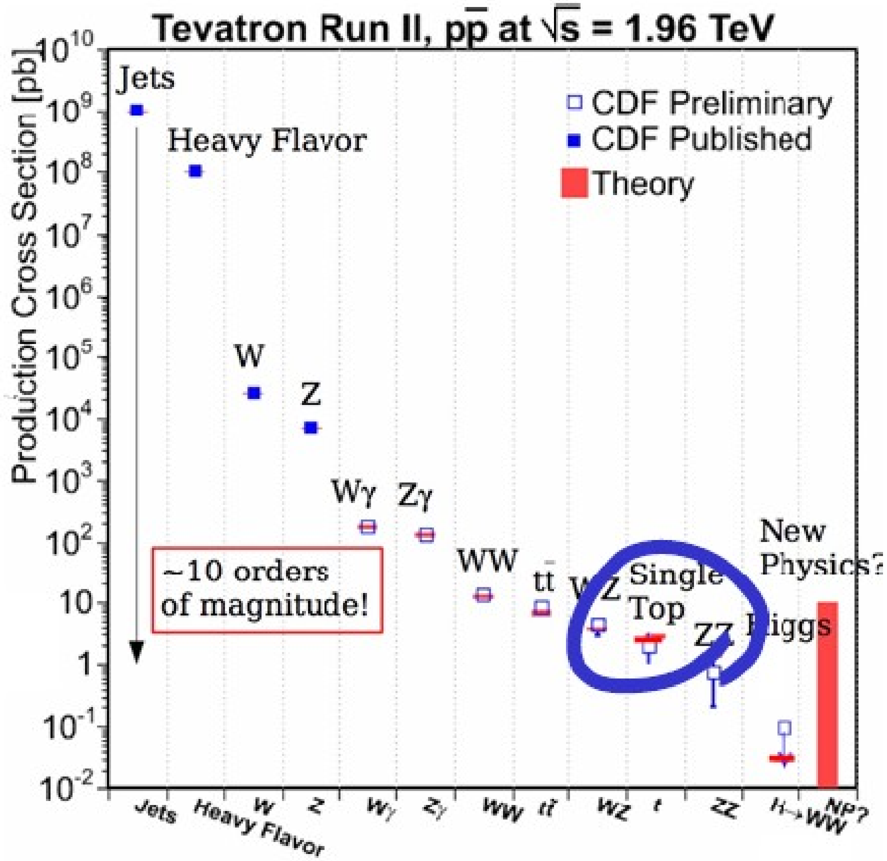


Outlook

This is just the beginning of the single top physics

- Precise measurements of σ_t and σ_s
- Top quark polarization
- Search for Anomalous Top quark couplings
- W ' and H^+ searches
- Top production through FCNC

From R.Wallny's Wine and Cheese talk, 03/10/2009



Milestone in the race for Higgs Boson !

Public web sites

More details can be found on the public pages of the experiments:

http://www-cdf.fnal.gov/physics/new/top/public_singletop.html

http://www-d0.fnal.gov/Run2Physics/top/singletop_observation

Backup

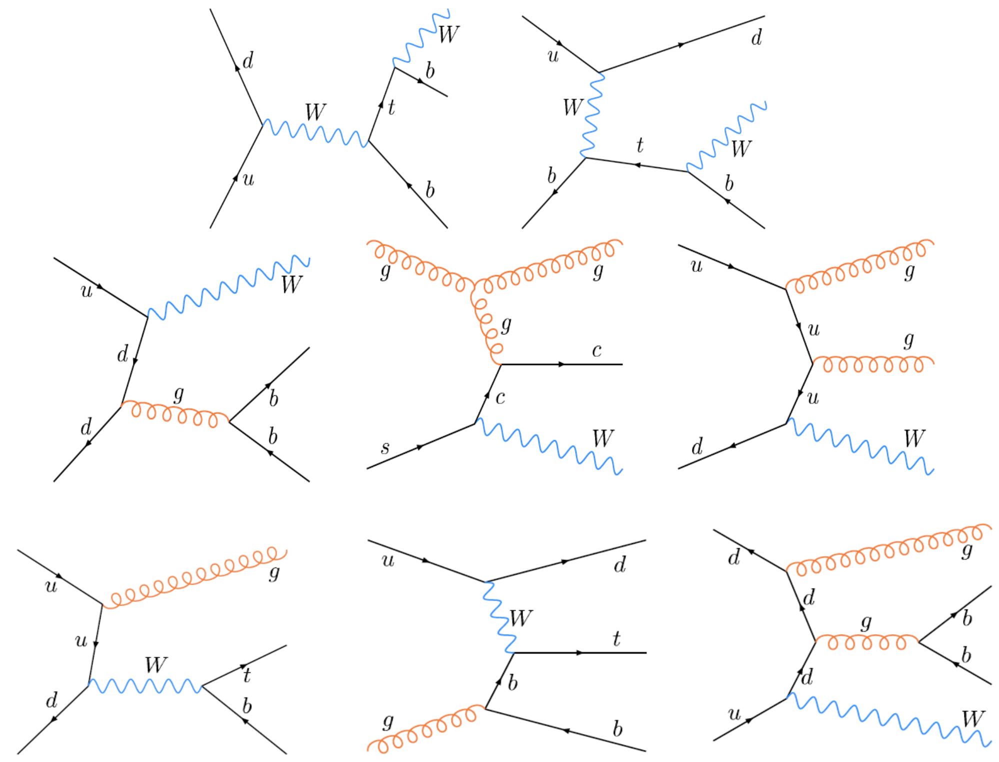
V_{tb} measurement

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

= 0

Summary of MEs

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

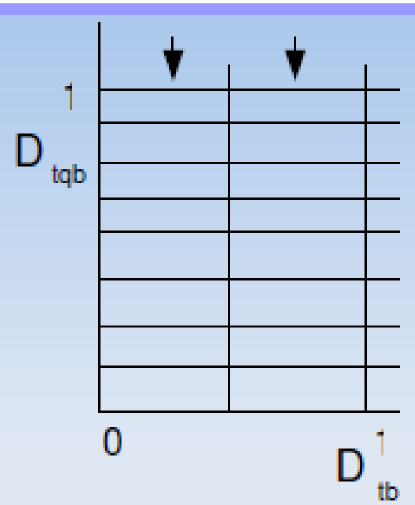


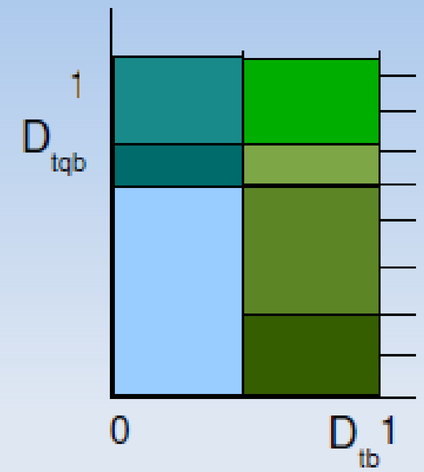
$$P_B(\vec{x}) = C_{Wb\bar{b}g} P_{Wb\bar{b}g}(\vec{x}) + C_{W ugg} P_{W ugg}(\vec{x}) + C_{t\bar{t}} P_{t\bar{t}}(\vec{x})$$

$$d\sigma_{Wcg}(\ell, j_1, j_2) = \varepsilon_c(j_1)(1-\varepsilon_l(j_2))d\sigma_{Wcg}(\ell, j_1\to c, j_2\to g) + \varepsilon_l(j_2))(1-\varepsilon_c(j_1))d\sigma_{Wcg}(\ell, j_2\to c, j_1\to g).$$

← Wcg example with one tagged and one not tagged jet

Presentation of discriminant



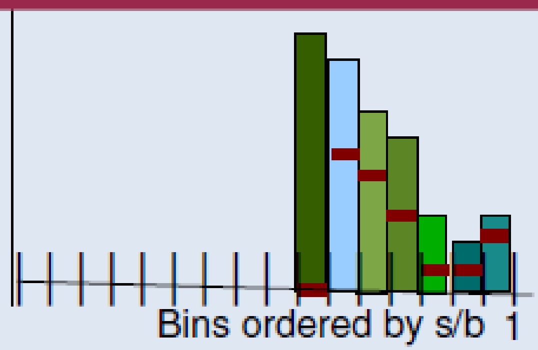


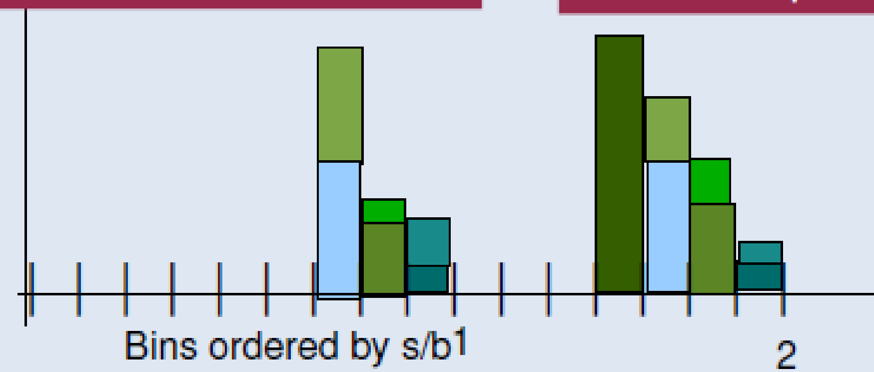
s/b = 0.90
s/b = 0.85
s/b = .80
s/b = .60
s/b = .34
s/b = .16
s/b = .15

Step 1: tb and tqb 2-D discriminant divided into 16 bins total

Step 2: Merge bins so that have sufficient background -> 7 bins total

Step 3: Merged bins s/b computed





Step 4: Merged bins ordered by s/b in 16 bins from 0 to 1. Red line signifies H_T cut where below the red line is yield < 175 GeV and above is yield > 175 GeV

Step 5: Histogram divided into low H_T region [0,1] and high H_T region [1,2]. Also these regions are rebinned to 16 bins from [0,2]

Variables in BDT and BNN

Object Kinematics

- p_T(lepton)
- $\not\!\!E_T$
- p_T(jet1)
- p_T(jet2)
- p_T(jet3)
- p_T(jet4)
- p_T(tag1)
- p_T(tag2)
- p_T(light1)
- p_T(light2)
- p_T(best1)
- p_T(notbest2)
- E(jet2)
- E(light1)
- Q(lepton)$\times\eta$(jet1)
- Q(lepton)$\times\eta$(jet2)
- Q(lepton)$\times\eta$(light1)
- Q(lepton)$\times\eta$(light2)
- Q(lepton)$\times\eta$(best1)
- Q(lepton)$\times\eta$(notbest1)

Event Kinematics

- Aplanarity(W,alljets)
- Centrality(alljets)
- Sphericity(W,alljets)
- H(alljets$-$tag1)
- H_T(alljets)
- H_T(alljets$-$tag1)
- H_T(alljets$-$best1)
- H_T(jet1,jet2)
- H_T(lepton,$\not\!\!E_T$)
- H_T(lepton,$\not\!\!E_T$,alljets)
- H_T(lepton,$\not\!\!E_T$,jet1,jet2)
- M(alljets)
- M(alljets$-$tag1)
- M(alljets$-$best1)
- M(jet1,jet2)
- M(jet3,jet4)
- M(light1,light2)
- M(W,jet1,jet2)
- M_T(jet1,jet2)
- p_T(jet1,jet2)
- $\sqrt{\hat{s}}$
- $M_T(W)$

Jet Reconstruction

- Width$_\eta$(jet1)
- Width$_\eta$(jet2)
- Width$_\eta$(jet4)
- Width$_\eta$(tag1)
- Width$_\eta$(tag2)
- Width$_\eta$(best1)
- Width$_\eta$(light2)
- Width$_\eta$(notbest2)
- Width$_\phi$(jet1)
- Width$_\phi$(jet2)
- Width$_\phi$(jet4)
- Width$_\phi$(tag2)
- Width$_\phi$(light1)
- Width$_\phi$(light2)
- Width$_\phi$(notbest1)
- p_T^{rel}(jet1, μ)
- M(jet1)
- M(tag1)
- M(best1)

Top Quark Reconstruction

- M(W,jet1) (leading jet top mass)
- M(W,jet1,$S2$) (with second neutrino solution)
- M(W,jet2)
- M(W,jet2,$S2$)
- M(W,jet3,$S2$)
- M(W,tag1) ("b-tagged" top mass)
- M(W,tag1,$S2$)
- M(W,tag2)
- M(W,light1,$S2$)
- M(W,best1) ("best" top mass)
- M(W,best1,$S2$)
- M(W,notbest1,$S2$)
- M(W,notbest2)
- M(W,notbest2,$S2$)
- $M_{\mathrm{top}}^{\Delta M^{\mathrm{min}}}$
- $M_{\mathrm{top}}^{\mathrm{sig}}$
- $\Delta M_{\mathrm{top}}^{\mathrm{min}}$
- Significance$_{\mathrm{min}}(M_{\mathrm{top}})$
- $p_z(\nu, S2)$

Angular Correlations

- ΔR(jet1,jet2)
- ΔR(lepton,jet1)
- ΔR(lepton,tag1)
- ΔR(lepton,best1)
- ΔR(lepton,light1)
- ΔR^{min}(alljets)
- ΔR^{min}(lepton,alljets)
- $\Delta\phi$(lepton, $\not\!\!E_T$)
- $\Delta\phi$(lepton,tag1)
- cos(lepton,jet1)$_{\mathrm{btaggedtop}}$
- cos(lepton,tag1)$_{\mathrm{btaggedtop}}$
- cos(lepton$_{\mathrm{btaggedtop}}$,btaggedtop$_{\mathrm{CMframe}}$)
- cos(lepton,light1)$_{\mathrm{btaggedtop}}$
- cos(lepton,best1)$_{\mathrm{besttop}}$
- cos(best1,notbest1)$_{\mathrm{besttop}}$
- cos(lepton$_{\mathrm{besttop}}$,besttop$_{\mathrm{CMframe}}$)
- cos(lepton,Q(lepton)$\times z$)$_{\mathrm{besttop}}$

Choice of variables – BDT

- Start with 600 variables expected to differ between signal and at least one of the background components
- Remove variables from list with low KS-test value between data and background model
- Remove variables with not much discrimination power to reduce computation time later (rank them after decision tree training)

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)$_{btaggedtop}$
	Δφ(lepton,$\not{E}_T$)
	Q(lepton) x η(light1)

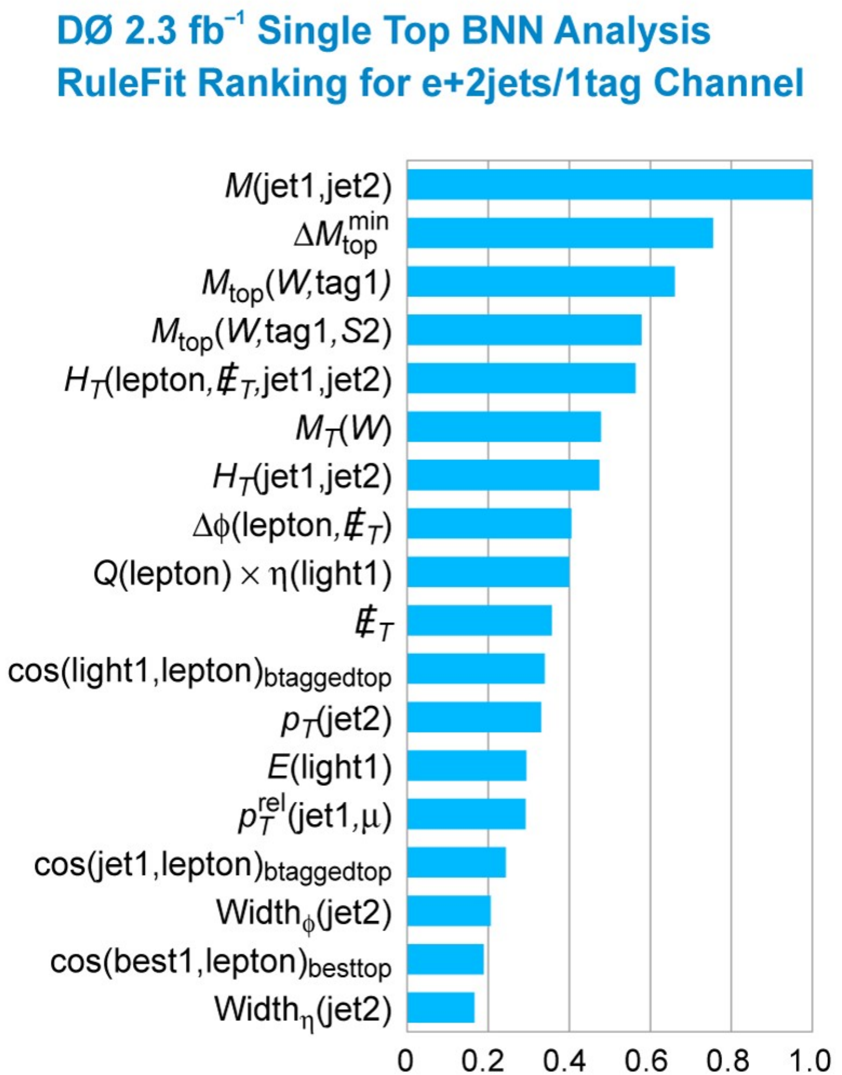
Best Variables to Separate Single Top from Top Pairs

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)

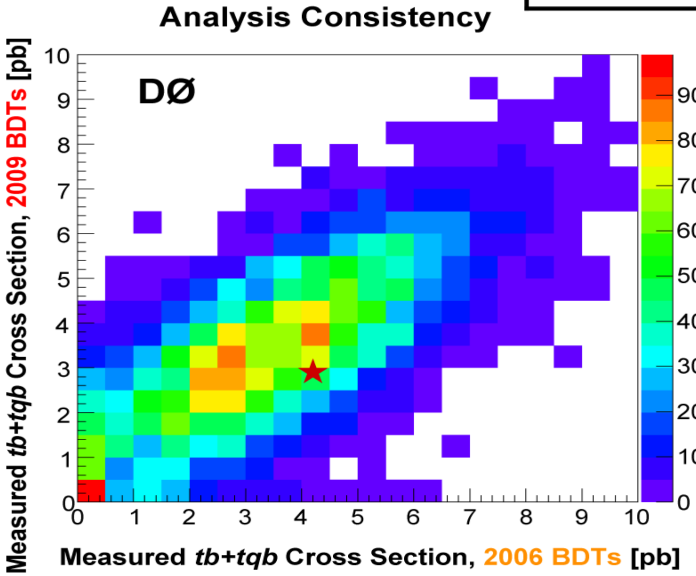
Choice of variables – BNN

- Start from a set of ~150 well modeled variables
- Use the highest ranked variables for each channel
- Ranking determined by Rulefit* -a MVA based on Decision Trees (DT)
- Uses 1/3 of the available MC samples. These samples are later not used for the measurement
- Importance of each variable given by how often it appears in the set of rules that define the DT
- Keep variables with Importance > 10
- Corresponds to 18-28 variables, depending on the channel



Consistency test

Measured *tb+tqb* Cross Section [pb] DØ 2.3 fb^{-1} Analysis					
Method	All Channels	Electron Only	Muon Only	Run IIa Only	Run IIb Only
BDT	3.74 +0.95 -0.79	4.4 +1.5 -1.2	3.3 +1.2 -1.0	2.5 +1.3 -1.2	4.9 +1.4 -1.2
BNN	4.70 +1.18 -0.93	5.3 +1.5 -1.4	4.0 +1.3 -1.2	3.4 +1.4 -1.3	5.7 +1.5 -1.3
ME	4.30 +0.98 -1.20	3.2 +1.9 -1.6	3.9 +1.2 -1.2	2.6 +1.3 -1.2	5.4 +1.7 -1.4
Combination	3.94 +0.88 -0.88	4.5 +1.3 -1.2	3.5 +1.1 -1.0	2.6 +1.2 -1.1	5.0 +1.2 -1.2



Measured single top cross sections using the 2009 and the 2006 decision trees on 5,000 pseudo-datasets generated from the 2009 Run IIa e+jets samples.

The red star shows the measurements in real data: 4.2 pb from the 2006 analysis and 2.9 pb from the 2009 analysis. This 1.3 pb shift is not uncommon, as seen from the width of the distribution from the pseudo-datasets.

Systematics on combination