Single Top Quark Production at the Tevatron

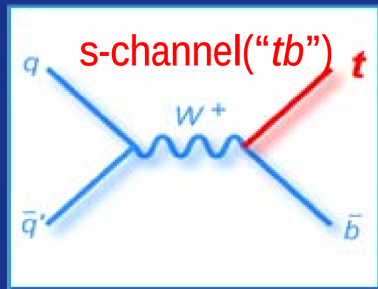
Shabnam Jabeen
Boston University

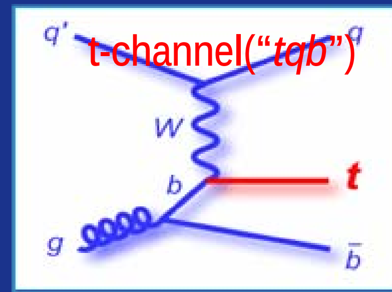
DIS 2007,April 16 – 21, Munich, Germany

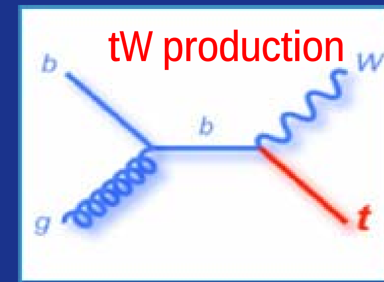
Outline

- Single Top
  - Introduction and Motivation
  - Event Signatures and Selection
- Analysis Steps
  - Search Strategy
  - Multivariate Techniques
  - Statistical Analysis
- Results & Conclusions

Single Top Production



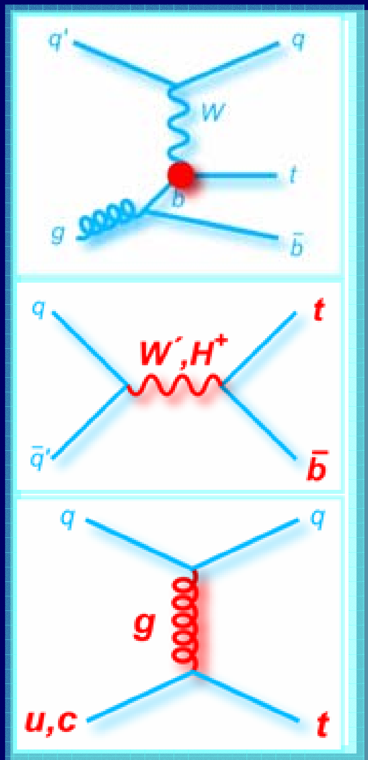




	s-channel	t-channel	tW production
Tev:	σ ~1 pb	σ ~ 2pb	σ ~0.25 pb
LHC:	σ ~11 pb	σ ~ 235pb	σ ~95 pb

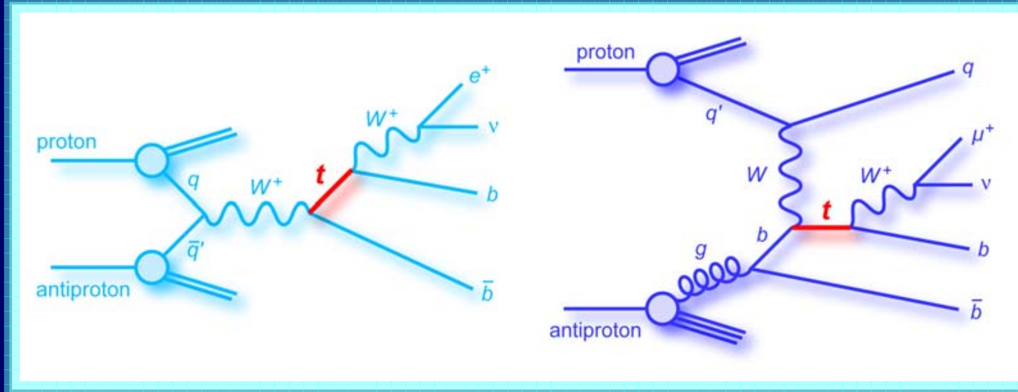
Motivation:

- Prediction of SM not observed so far
- Study *Wtb* coupling in top production
  Measure $|V_{tb}|$ directly: $\sigma \propto |V_{tb}|^2$
- Cross sections sensitive to new physics
  s-channel: resonances (heavy *W′* boson, charged Higgs boson, Kaluza-Klein excited W_{KK}, technipion, etc.), t-channel: flavor-changing neutral currents ($t - Z / \gamma / g - c / u$ couplings), Fourth generation of quarks
- Top properties
  Polarized top quarks – spin correlations measurable in decay products, Measure top quark partial decay width and lifetime, CP violation (same rate for top and antitop?)
- Similar search for *WH* associated Higgs production



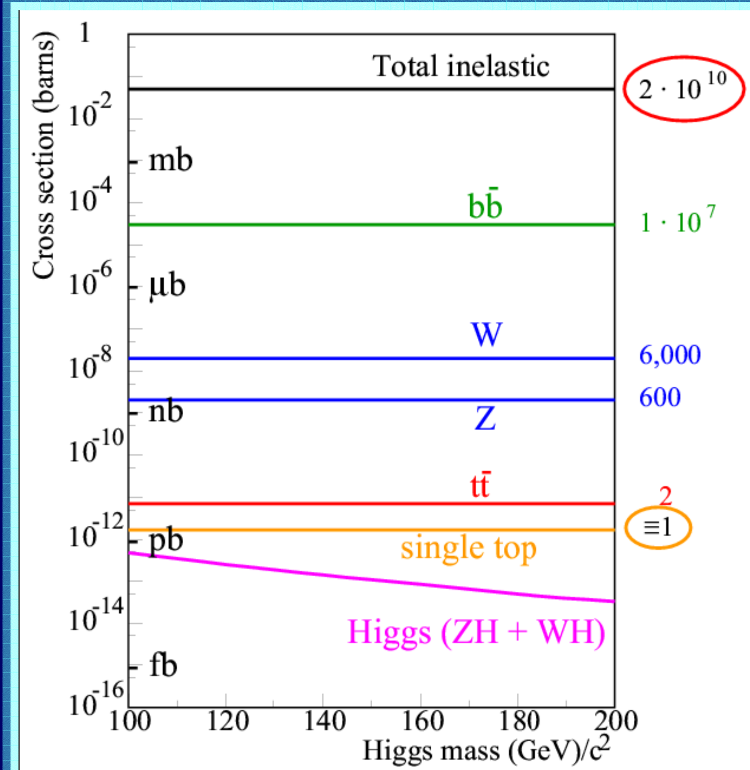
Event Signatures and Event Selection

- One isolated electron or muon
- Missing Transverse Energy
- At least two or more jets with at least one b-tagged jet



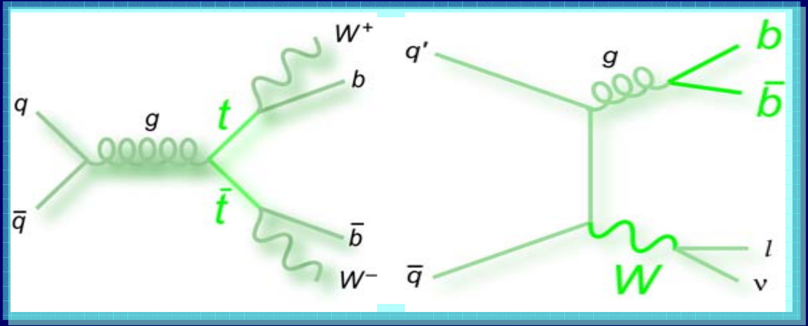
	CDF	DØ
e:	p_T>20 GeV, |η|<2.0	p_T>15 GeV, |η|<1.1
μ:	p_T>20 GeV, |η|<1.1	p_T>18 GeV, |η|<2.0
Missing E_T	MET>25 GeV	15<MET<200 GeV
Jets	=2, p_T^{uncorr}>15 GeV, |η|<2.8	2-4, p_T>15 GeV, |η|<3.4 $p_{T,1}$>25 GeV, |η_1|<2.5 $p_{T,2}$>20 GeV
B-jet	1 or 2	

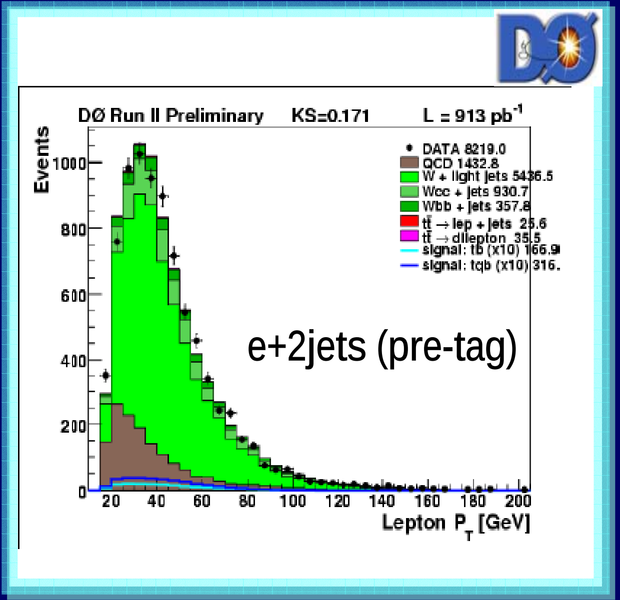
- Top pairs, W+jets, and Multijets are the main processes that can mimic these signatures



Background Modeling

- *W+jets* background:
  - Event kinematics and flavor composition modeled using Alpgen generator
  - Normalized to data before b tagging and after subtracting other backgrounds
  - Additional scale factor α~1.5 for Wbb and Wcc
- Multijet background:
  - Modeled using data with a non-isolated lepton and jets
- *Top pair backgrounds* modeled using ALPGEN
  - Normalized to NNLO theoretical prediction
- *Diboson (WW,WZ,ZZ)*
  - Very small contribution
  - CDF: normalized to theoretical prediction
  - D0: included in W+jets via normalization to data





Event Yields After B-Tagging

Sensitivity significantly increased after requiring b-tagging, e.g. in 2 jets channel:

	S/B	S/√B
Before b-tagging	~1/200	~ 0.6
After b-tagging (1+2 b-tags)	~1/18	~ 1.5

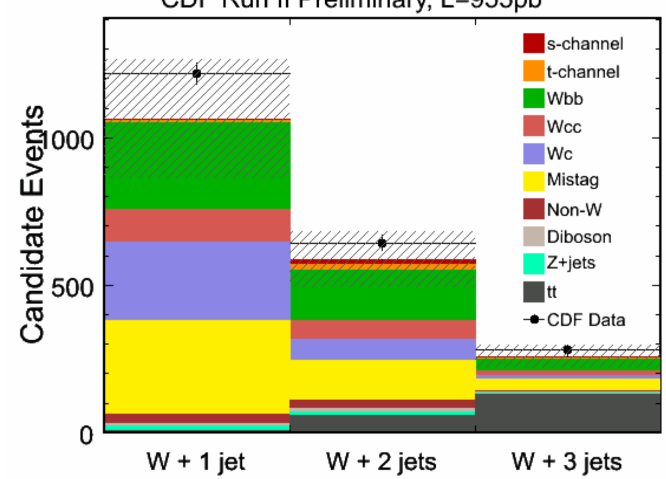
	Event Yields in 0.9 fb^{-1} Data Electron+muon, 1tag+2tags combined		
Source	2 jets	3 jets	4 jets
tb	16 ± 3	8 ± 2	2 ± 1
tqb	20 ± 4	12 ± 3	4 ± 1
$t\bar{t} \rightarrow ll$	39 ± 9	32 ± 7	11 ± 3
$t\bar{t} \rightarrow l$+jets	20 ± 5	103 ± 25	143 ± 33
$W+b\bar{b}$	261 ± 55	120 ± 24	35 ± 7
$W+c\bar{c}$	151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

- Cross section uncertainties are dominated by the statistical uncertainty
- Single top signal is smaller than total background uncertainty
- Counting events is not a sensitive enough method

Event Yields in 955 pb-1 Data
e+mu, 1tag+2tags combined,2 jets

s-channel	15.4 ± 2.2
t-channel	22.4 ± 3.6
tt	58.4 ± 13.5
Diboson	13.7 ± 1.9
Z + jets	11.9 ± 4.4
Wbb	170.9 ± 50.7
Wcc	63.5 ± 19.9
Wc	68.6 ± 19.0
Non-*W*	26.2 ± 15.9
Mistags	136.1 ± 19.7
Single top	37.8 ± 5.9
Total background	549.3 ± 95.2
Total prediction	587.1 ± 96.6
Observed	644

CDF Run II Preliminary, L=955pb^{-1}



Search Strategy and Analysis Steps

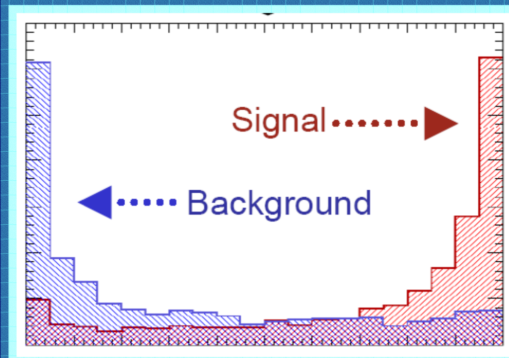
Signal acceptances (including BR)

	tb	tqb
CDF (2 jets)	~1.9%	~1.3%
DØ (2-4 jets)	~3.2%	~2.1%

Percentage of single top *tb+tqb* selected events and S:B ratio (white squares = no plans to analyze)

Electron + Muon	1 jet	2 jets	3 jets	4 jets	≥ 5 jets
0 tags	10% 1 : 3,200	25% 1 : 390	12% 1 : 300	3% 1 : 270	1% 1 : 230
1 tag	6% 1 : 100	21% 1 : 20	11% 1 : 25	3% 1 : 40	1% 1 : 53
2 tags		3% 1 : 11	2% 1 : 15	1% 1 : 38	0% 1 : 43

- Calculate discriminants that separate signal from background:
  - Boosted Decision Trees — *D0*
  - Matrix Elements — *D0, CDF*
  - Likelihood Discriminants — *CDF*
  - Bayesian Neural Networks — *D0, CDF*



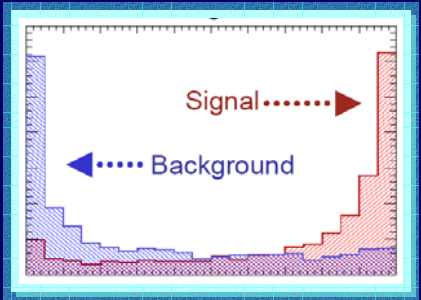
Event Discriminant

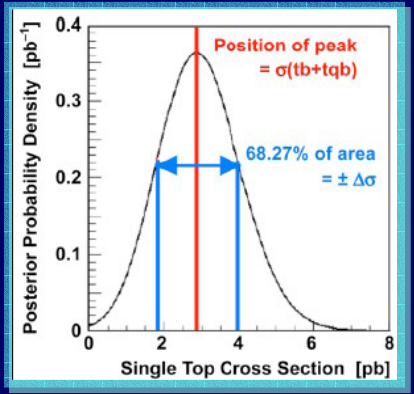
- Check discriminant performance using data control samples
- Use discriminant output to measure cross section and significance

Statistical Analysis

Cross Section Measurement

- Binned likelihood from discriminant distribution
- Compute posterior probability density of tb+tqb using Bayes' theorem:
  - Flat positive-defined prior for the cross section
  - Systematic uncertainties are treated as Gaussian nuisance parameters
- For example: for D0 there are 12 distributions with 100 bins each that go into this calculation



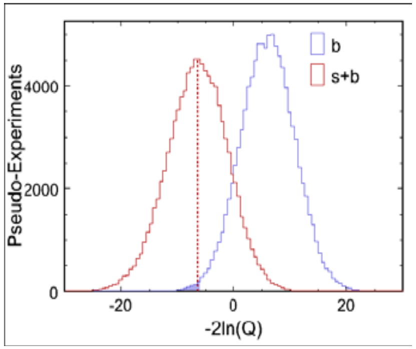


Significance

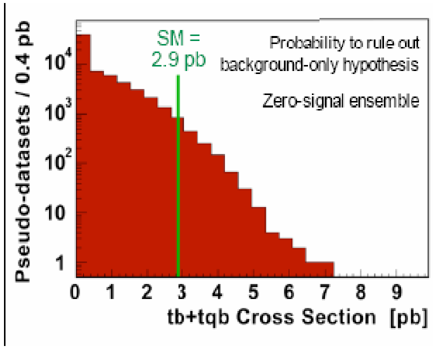
CDF and D0 are using different measures of significance which can not be compared directly

$$Q = \frac{L(data \mid s+b)}{L(data \mid b)}$$

$$p-value = 1 - CL_b = \int_{-\infty}^{obs} dQ P_b(Q)$$



$$p-value = \int_{\sigma_{tb+tqb}^{obs}}^{\infty} d\sigma_{tb+tqb} P_b(\sigma_{tb+tqb})$$



Signal-Background Separation Using Matrix Elements

- Use the 4-vectors of all reconstructed leptons and jets
- Use matrix elements of main signal and background Feynman diagrams to compute an event probability density for signal and background hypotheses

differential cross section (LO matrix element)

parton distribution functions

$$P_i(\vec{x}) = \frac{1}{\sigma} \int \ldots \int \sum_{comb} d^n\sigma_i(\vec{y})\, dq_1 dq_2\, f(q_1)\, f(q_2)\, W(\vec{x}\,|\,\vec{y})$$

transfer function: mapping from parton-level variables (y) to reconstruction-level variables (x)

- Calculate a discriminant using above probability:

CDF

$$EPD = \frac{b \cdot P_{tb+tqb}(\vec{x})}{b \cdot P_{tb+tqb}(\vec{x}) + b \cdot P_{Wbb}(\vec{x}) + (1-b)\cdot(P_{Wcc}(\vec{x}) + P_{Wcj}(\vec{x}))}$$

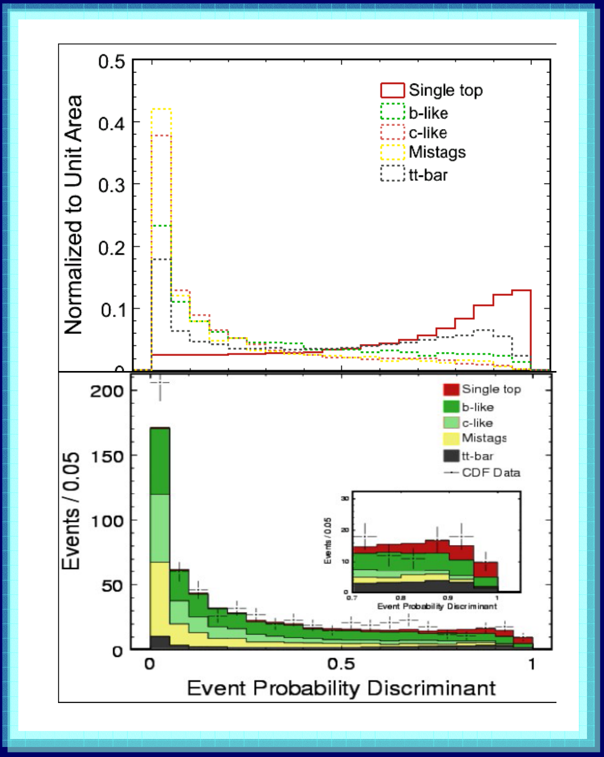
DØ

$$D_S(\vec{x}) = \frac{P_S(\vec{x})}{P_S(\vec{x}) + P_{bckg}(\vec{x})}; \qquad S = tb \text{ or } tqb$$

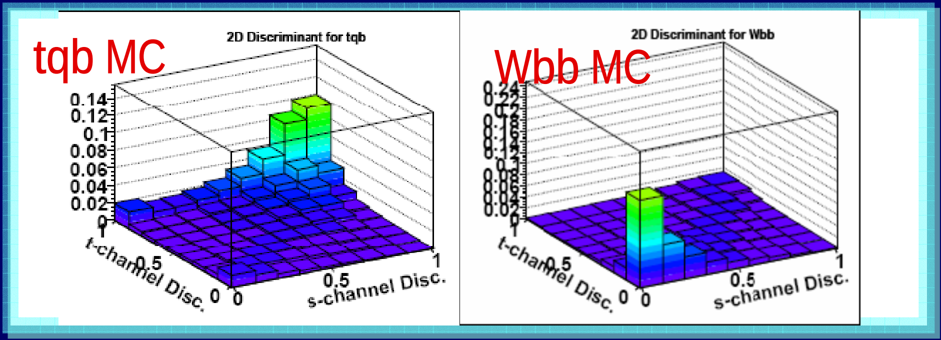
b = b-tagging NN probability (event-by-event)

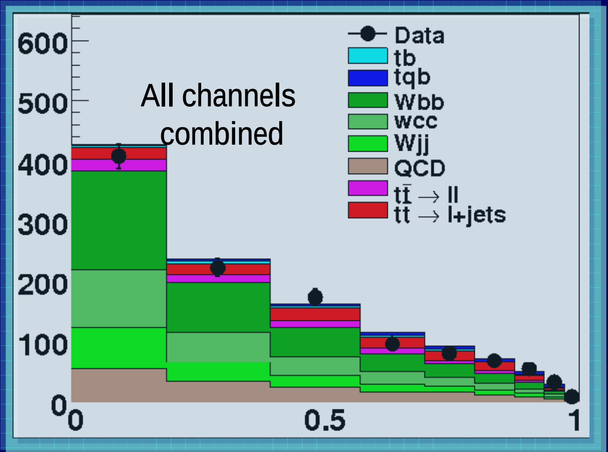
Matrix Elements Analysis

- Consider only 2-jet events.
- Single channel search
- (e+mu ,1tag and 2tags combined)
- Combined search based on 1D disc:



- Consider 2-jet and 3-jet events
- Six separate search channels
- (e, mu) x (2,3 jets) x (1,2 tags)
- Combined search based on 2D disc:

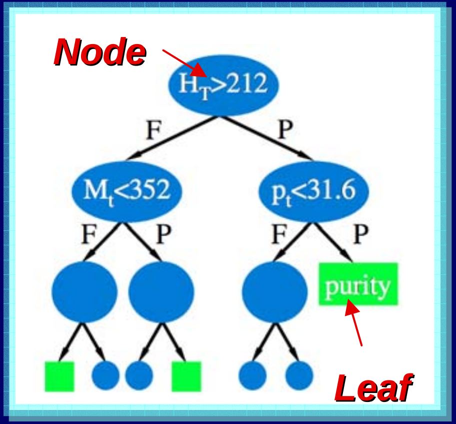




Expected Performance	median p-value (CDF)	p-value (DØ)
CDF ME analysis	0.5% (2.6σ)	
DØ ME analysis		3.6% (1.8σ)

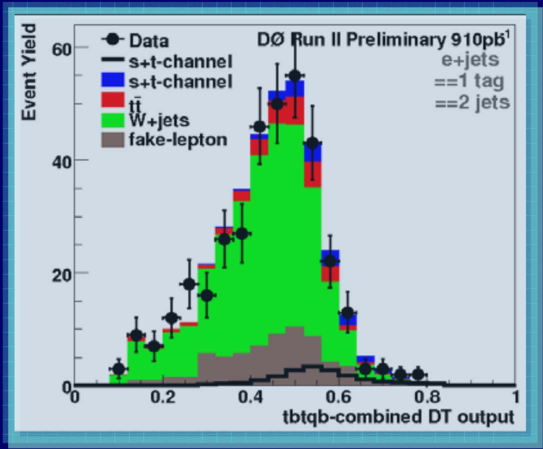
Boosted Decision Trees Analysis

- Machine-learning technique, widely used in social sciences, some use in HEP
  - Start at first "node " : For each variable, find splitting value with best separation between two children (mostly signal in one, mostly background in the other)
  - Select variable and splitting value with best separation to produce two "branches". Repeat recursively on each node
  - Stop when improvement stops or when too few events are left Decision tree output for each event = leaf purity closer to 1(0) for signal (background)



$$Purity = \frac{N_{Signal}}{N_{Signal} + N_{Background}}$$

- Improve performance of DT by using adaptive boosting , which averages over many trees, diluting the piecewise nature of the DT output



Expected Performance	p-value (DØ)
Combined search	1.8% (2.1σ)

Cross Checks

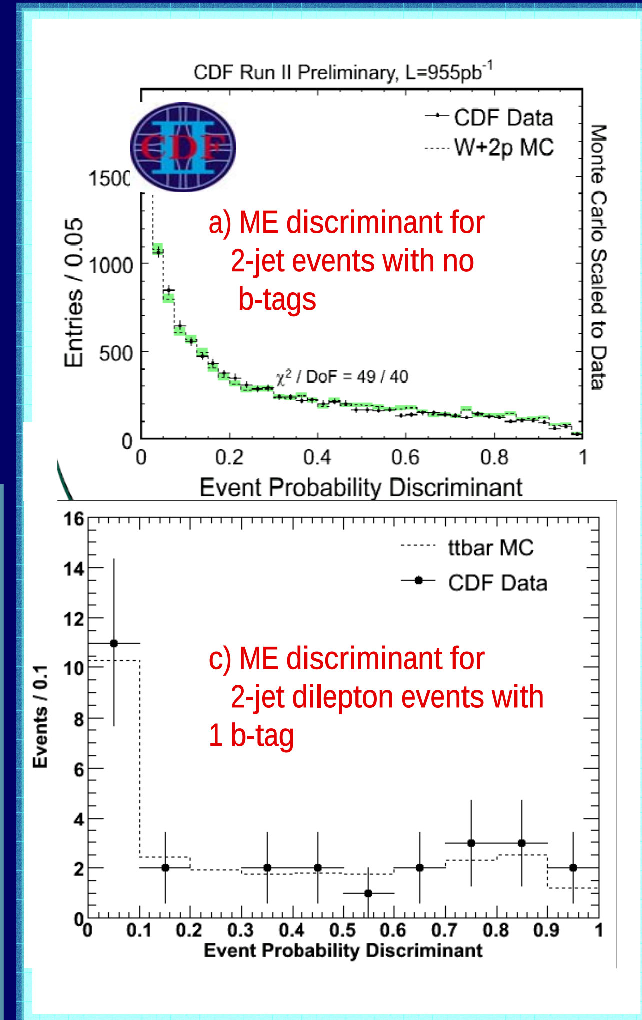
- validate a background model in side-band regions without looking at single top candidates.
- Examples:

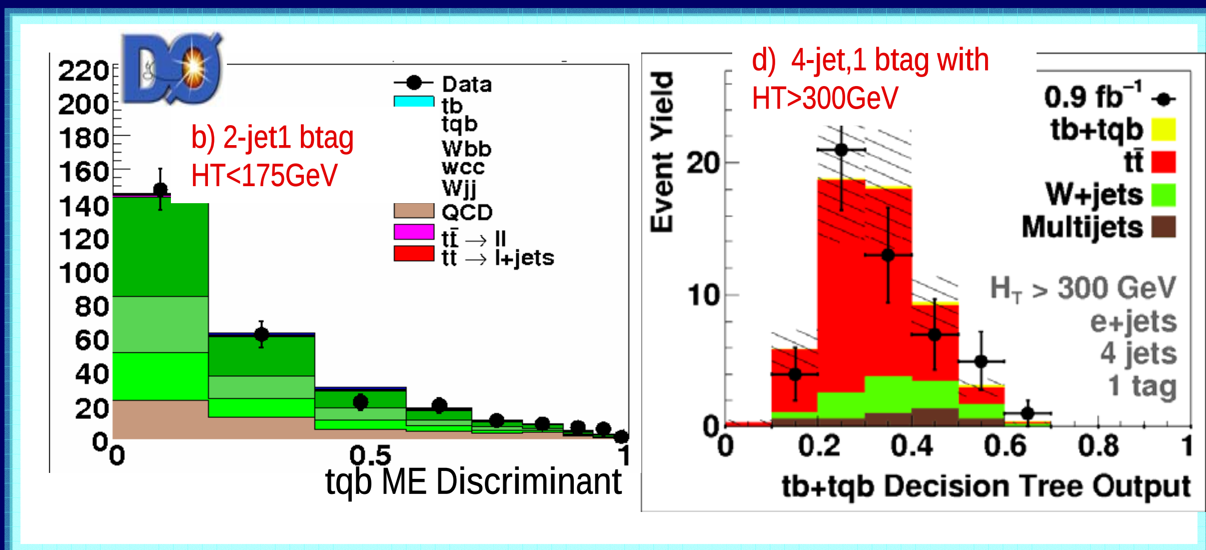
a) An enriched W+light jets background

b) An enriched W+heavy jets background

c) A top pair (dilepton) background

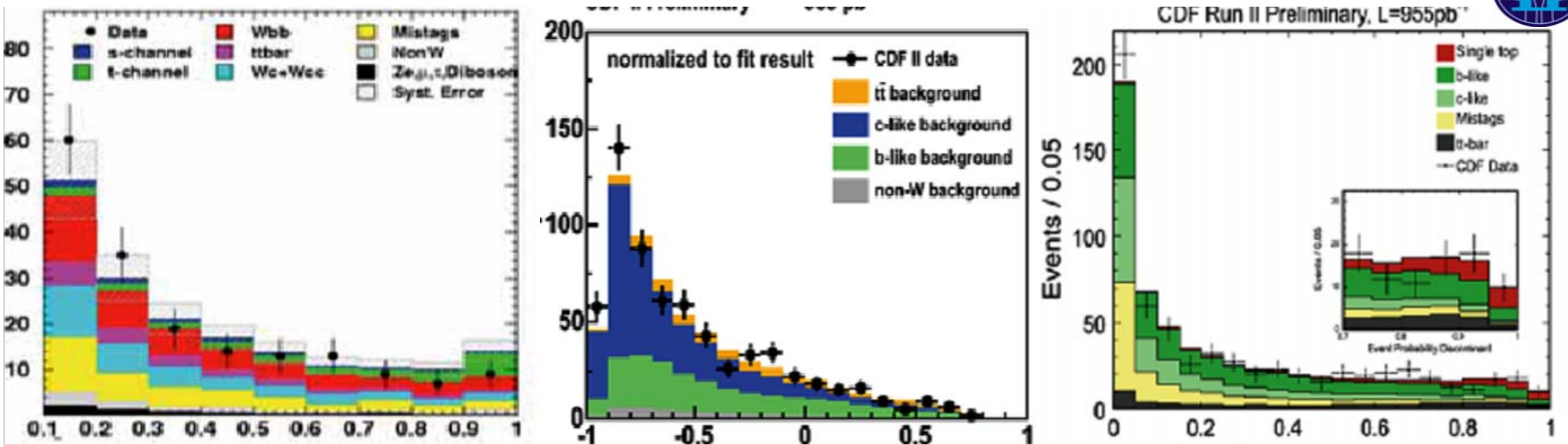
d) An enriched top pair background





CDF Results

Likelihood | Neural networks | Matrix Element



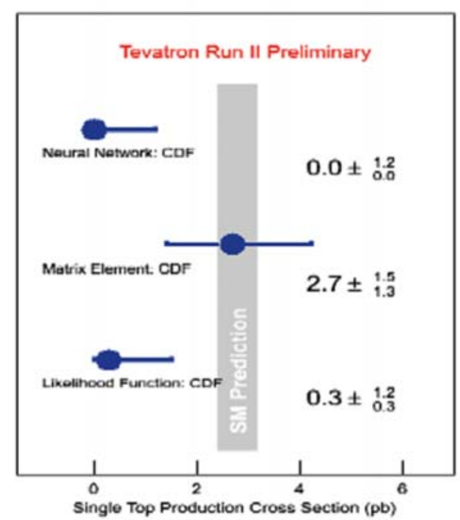
No evidence of signal
σ_{s+t}<2.7pb at 95% C.L.

No evidence of signal
σ_{s+t}<2.6pb at 95% C.L.

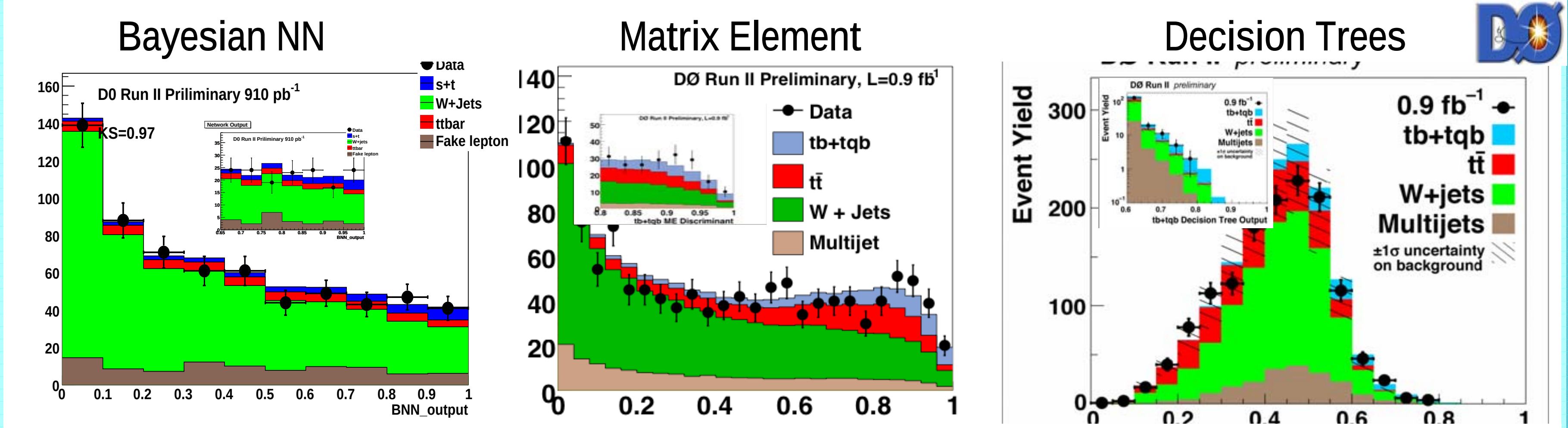
p-value = 1.0% (2.3σ)
σ_{s+t}=2.7(+1.5/-1.3)pb

Compatibility of CDF Results

- Performed common pseudo-experiments
- Correlation among analyses: ~60-70%
- 1.2% of the pseudo-experiments had an outcome as different as the one observed in data



D0 Results



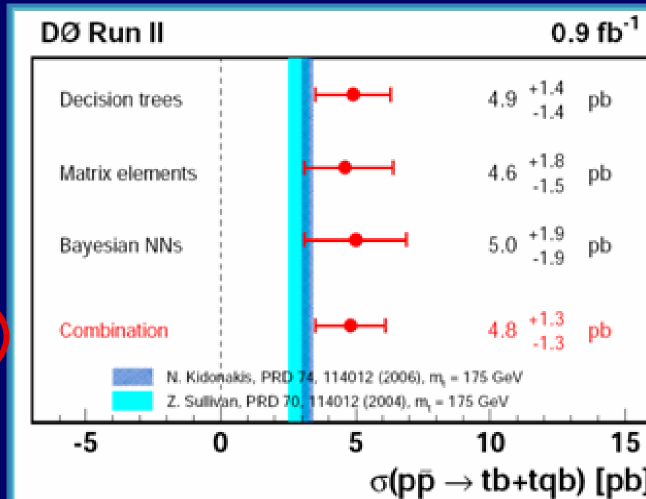
Observed p-value: 1.15% (2.3σ) *Observed p-value:* 0.20% **(2.9σ)** *Observed p-value:* 0.034% (3.4σ)

D0 Combination

Three measurements are not 100% correlated
A combination of these measurements gives:

$$\sigma_{tb+tqb} = 4.8 \pm 1.3 \Rightarrow$$ Significance of 3.5σ

Evidence for single top production!!

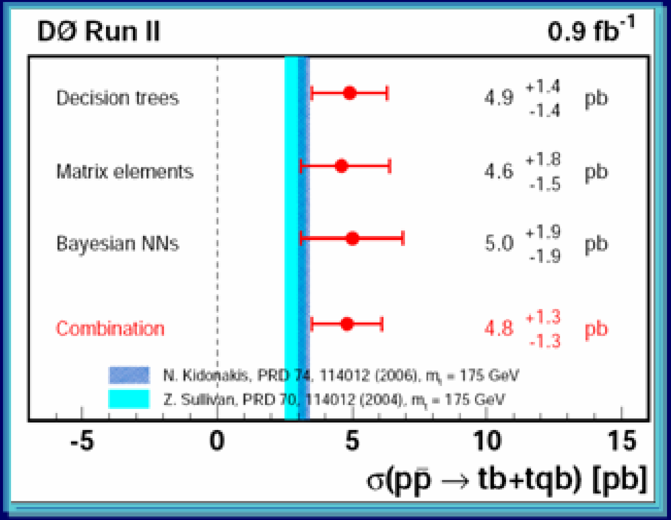


Summary:

First Evidence for Single Top Production

Three multivariate techniques applied
to separate signal from background
Boosted decision trees gives result
with 3.4 σ significance

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



First direct measurement of |*Vtb*|

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

Outlook:

- We have already collected more than twice the data used for this analysis
Hopefully, evidence will turn into observation soon!

- With Tevatron breaking its own Luminosity records every week, and with all the experience gained in search for single top, we are already closing in on many things, including Higgs – so,

stay tuned for more exciting news from the Tevatron!!!

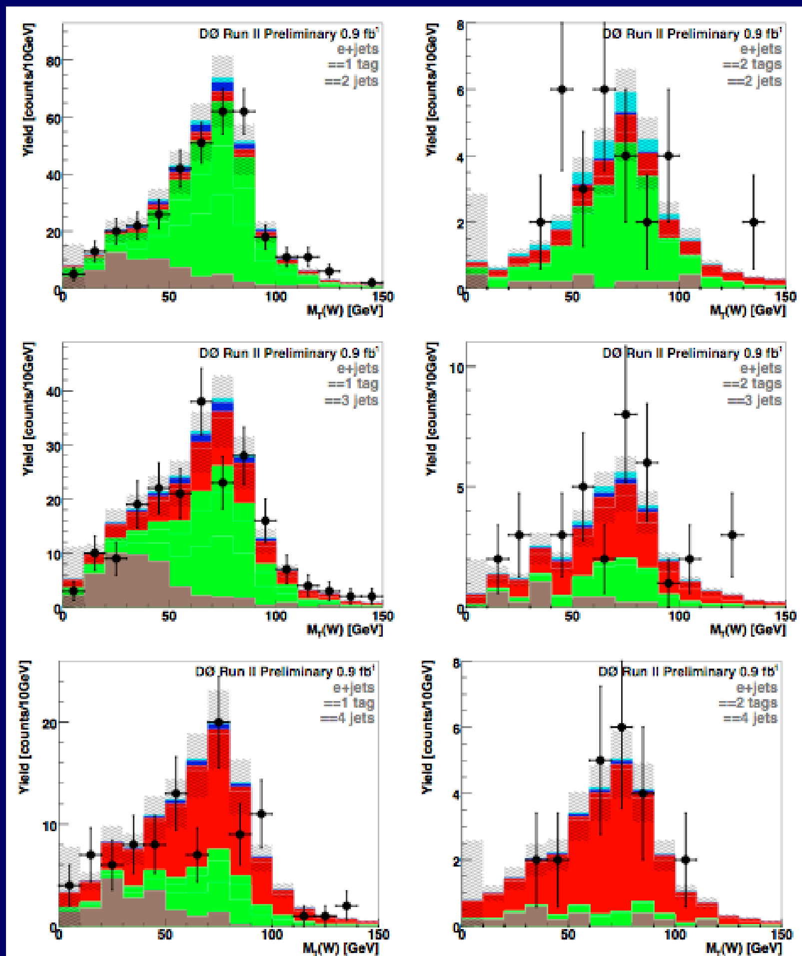
12 Analysis Channels

- ● Data
- tb
- tqb
- $t\bar{t}$
- W + jets
- Multijets
- ±1σ uncertainty on background

W Transverse Mass

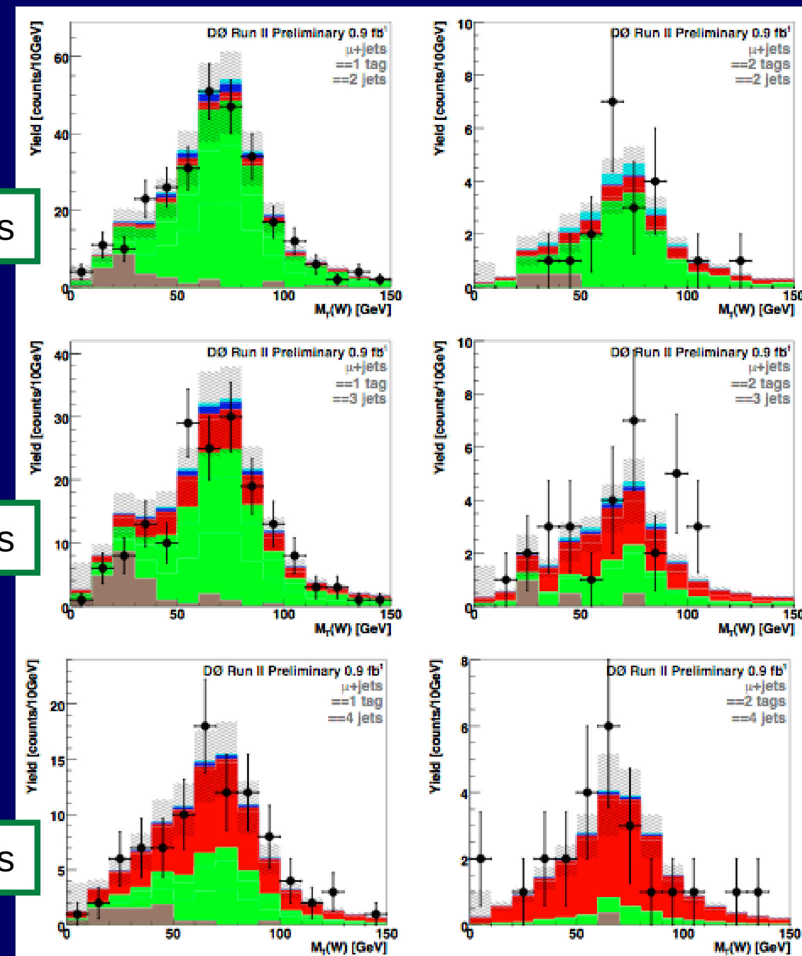
Electrons

1 tag | 2 tags



Muons

1 tag | 2 tags



2 jets

3 jets

4 jets

Systematic Uncertainties

- Uncertainties are assigned for each signal and background component in all analysis channels
- Correlations between channels and sources are taken into account

CDF

Uncertainties affecting single top

Systematic uncertainty	Rate	Shape
Jet energy scale	+1.6% / -2.0%	X
Initial state radiation	+2.0% / + 0.3%	X
Final state radiation	+2.6% / +1.9%	X
Parton distribution functions	+1.4% / -0.4%	X
Monte Carlo generator	±1.6%	
Event detection efficiency	±7.4%	
Luminosity	±6%	
Neural-net b tagger	N/A	X
Mistag model	N/A	X
Non-W model	N/A	X
Q^2 scale in Alpgen MC	N/A	X
Total rate uncertainty	±10.5%	

Source of Uncertainty	Size
Top pairs normalization	18%
W+jets & multijets normalization	18–28%
Integrated luminosity	6%
Trigger modeling	3–6%
Lepton ID corrections	2–7%
Jet modeling	2–7%
Other small components	Few %
Jet energy scale	1–20% ★
Tag rate functions	2–16% ★

★ Uncertainties that also affect the shapes of the distributions

DØ

- Cross section uncertainties are dominated by the statistical uncertainty, the systematic contributions are all small

Decision Tree Variables

Object Kinematics

p_T(jet1)
p_T(jet2)
p_T(jet3)
p_T(jet4)
p_T(best1)
p_T(notbest1)
p_T(notbest2)
p_T(tag1)
p_T(untag1)
p_T(untag2)

Angular Correlations

ΔR(jet1,jet2)
$\cos(\text{best1,lepton})_{\text{besttop}}$
$\cos(\text{best1,notbest1})_{\text{besttop}}$
$\cos(\text{tag1,alljets})_{\text{alljets}}$
$\cos(\text{tag1,lepton})_{\text{btaggedtop}}$
$\cos(\text{jet1,alljets})_{\text{alljets}}$
$\cos(\text{jet1,lepton})_{\text{btaggedtop}}$
$\cos(\text{jet2,alljets})_{\text{alljets}}$
$\cos(\text{jet2,lepton})_{\text{btaggedtop}}$
$\cos(\text{lepton}, Q(\text{lepton}) \times z)_{\text{besttop}}$
$\cos(\text{lepton,besttopframe})_{\text{besttopC}}$
$\cos(\text{lepton,btaggedtopframe})_{\text{btagg}}$
$\cos(\text{notbest,alljets})_{\text{alljets}}$
$\cos(\text{notbest,lepton})_{\text{besttop}}$
$\cos(\text{untag1,alljets})_{\text{alljets}}$
$\cos(\text{untag1,lepton})_{\text{btaggedtop}}$

Event Kinematics

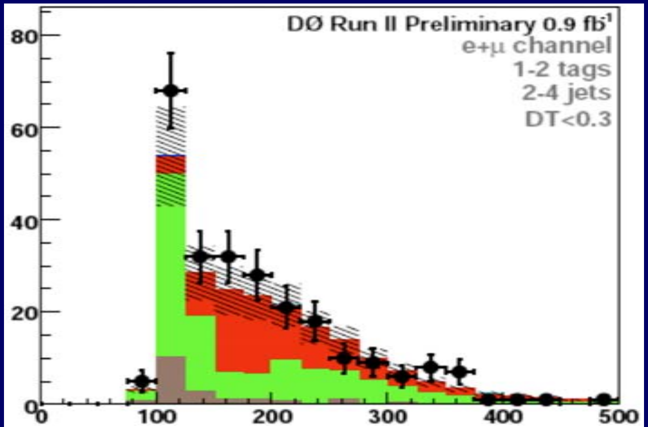
Aplanarity(alljets, W)
$M(W,\text{best1})$ ("best" top mass)
$M(W,\text{tag1})$ ("b-tagged" top mass)
H_T(alljets)
H_T(alljets−best1)
H_T(alljets−tag1)
H_T(alljets, W)
H_T(jet1,jet2)
H_T(jet1,jet2, W)
M(alljets)
M(alljets−best1)
M(alljets−tag1)
M(jet1,jet2)
M(jet1,jet2, W)
M_T(jet1,jet2)
$M_T(W)$
Missing E_T
p_T(alljets−best1)
p_T(alljets−tag1)
p_T(jet1,jet2)
$Q(\text{lepton}) \times \eta(\text{untag1})$
$\sqrt{\hat{s}}$
Sphericity(alljets, W)

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

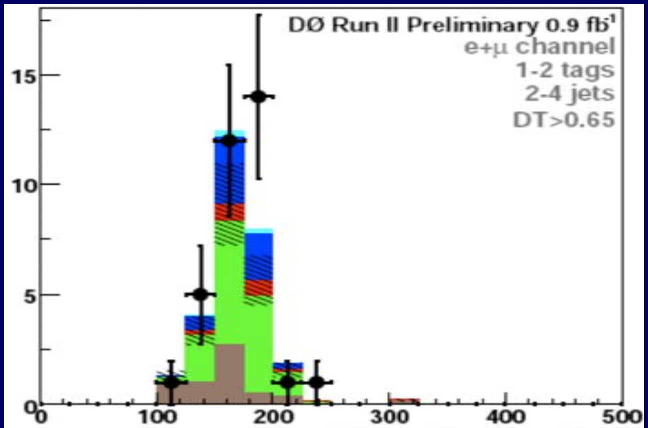
DT Event Characteristics

DT Discriminant < 0.3

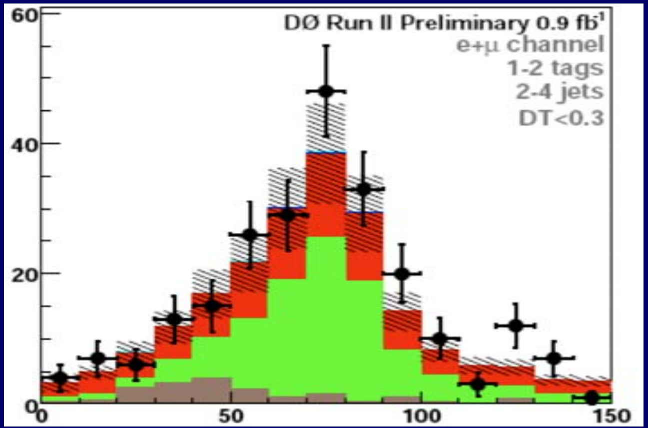
DT Discriminant > 0.65



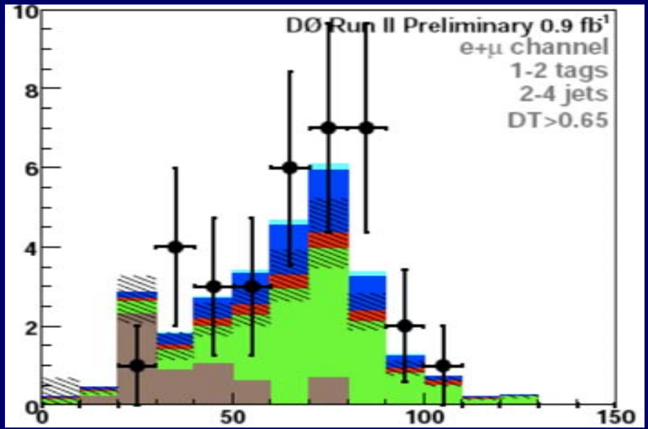
Mass (lepton,$\not\!\!E_T$,btagged-jet) [GeV]



Mass (lepton,$\not\!\!E_T$,btagged-jet) [GeV]

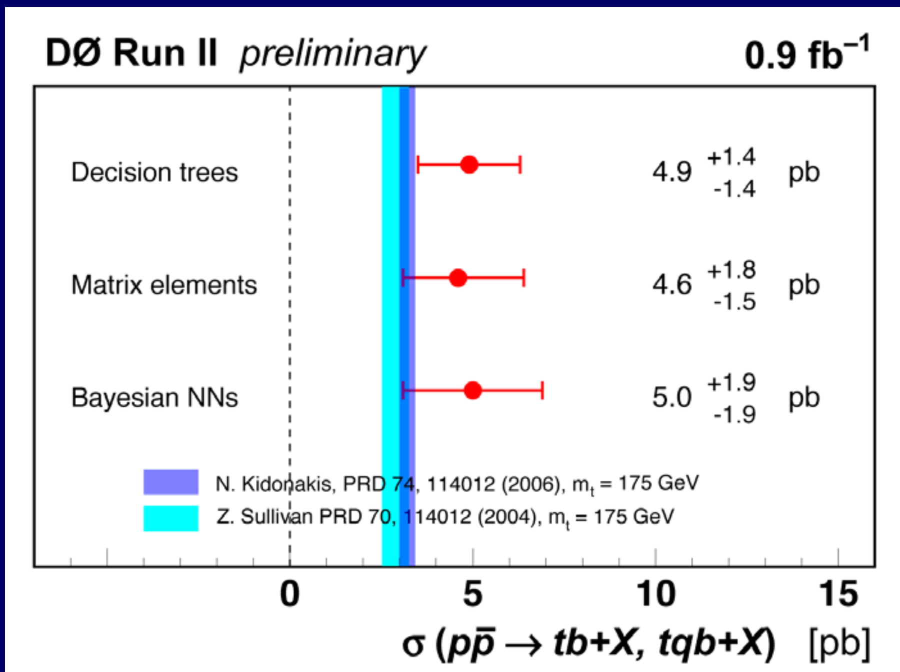


W Transverse Mass [GeV]



W Transverse Mass [GeV]

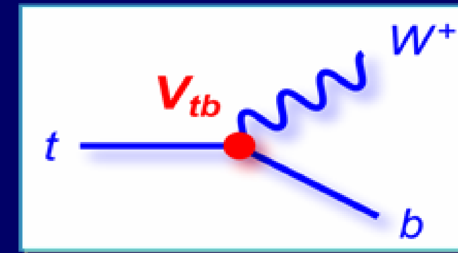
Correlation Between 3 Methods



$$\rho = \begin{pmatrix} DT & ME & BNN & \\ 1 & 0.57 & 0.51 & DT \\ 0.57 & 1 & 0.45 & ME \\ 0.51 & 0.45 & 1 & BNN \end{pmatrix}$$

Results from the three methods are consistent with each other

Measuring $|V_{tb}|$



Most general *tbW* vertex:

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

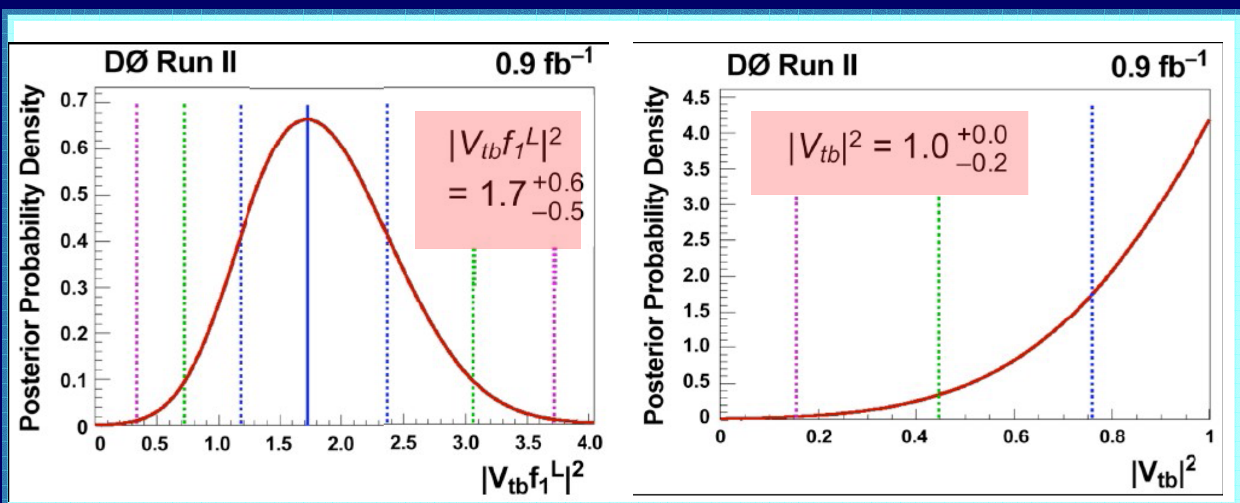
$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}$$

Within the SM:

- 3-generation and unitary CKM matrix $V_{tb} \sim 1$
- CP conserving pure V–A interaction: $f_1^L = 1,\ f_1^R = f_2^L = f_2^R = 0$

Measure V_{tb} assuming:

- No constraint on number of generation and unitarity of CKM matrix
- CP conserving pure V-A interaction, but not necessarily of SM strength



Strength of $V-A$ interaction
$|V_{tb} f_1^L| = 1.3 \pm 0.2$

assuming $f_1^L = 1$:

$0.68 \le |V_{tb}| \le 1$ @ 95% CL