Les Rencontres de Physique de la Vallee d'Aoste,
March 4-10, 2007, La Thuile, Aosta Valley, Italy

Single Top Quark Production at the Tevatron

Aurelio Juste
Fermi National Accelerator Laboratory

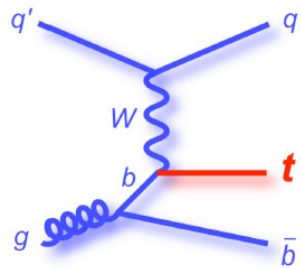
OUTLINE

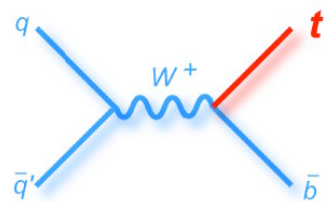
- Motivation
- Experimental Challenge
- Search Strategy
- Results:
  - Evidence for Single Top Quark Production
  - First Direct Measurement of $|V_{tb}|$
- Summary

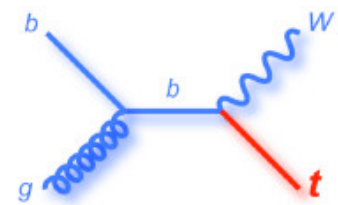
Single Top Quark Production

- Main SM production mechanisms at a hadron collider:

t-channel ("tqb") | *s-channel ("tb")* | *associated production ("tW")*







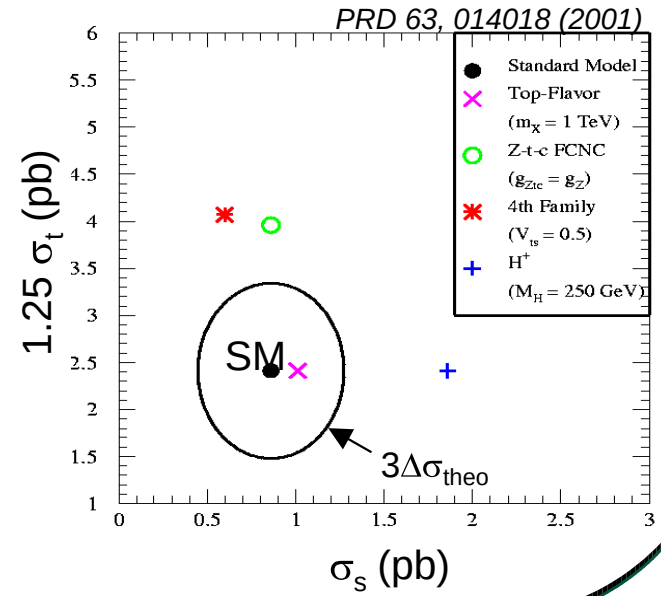
	t-channel ("tqb")	s-channel ("tb")	associated production ("tW")
Tevatron ($\sqrt{s}$=1.96 TeV):	~2.2 pb	~1.0 pb	~0.25 pb
LHC ($\sqrt{s}$=14 TeV):	~235 pb	~11 pb	~95 pb

(for m_t=175 GeV)
hep-ph/0609287
hep-ph/0701080

- Not discovered yet. It has been the subject of intense search at the Tevatron since Run I. Here will discuss the experimental status based on ~1 fb^{-1} of Run II data.

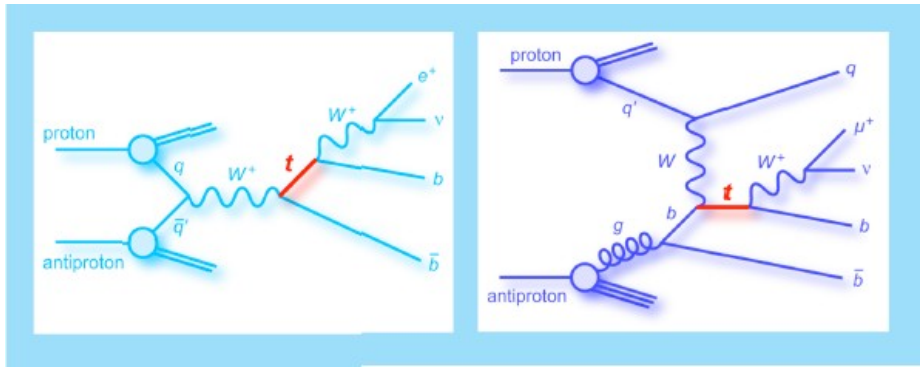
- Motivation:
  - Study of the *tbW* interaction:
    - Direct measurement of $|V_{tb}|$: $\sigma \propto |V_{tb}|^2$
    - Anomalous couplings
  - Sensitivity to different New Physics:
    - s-channel: W', $H^{\pm}$, W_{KK},…
    - t-channel: FCNC interactions, 4th family,..
  - Top spin physics (~100% polarized top quark)
  - Develop/exercise techniques to extract small signal in a large background (e.g. Higgs search).

PRD 63, 014018 (2001)



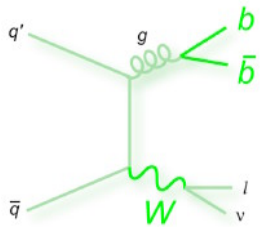
Experimental Signature and Event Selection

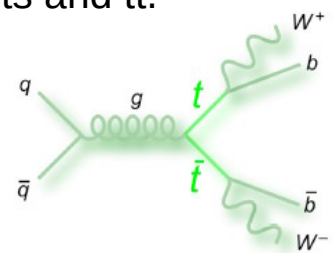
- Experimental signature:
  - One high p_T isolated lepton (e or μ)
  - High missing transverse energy
  - $\geq$2 jets ($\geq$1 b-tag)

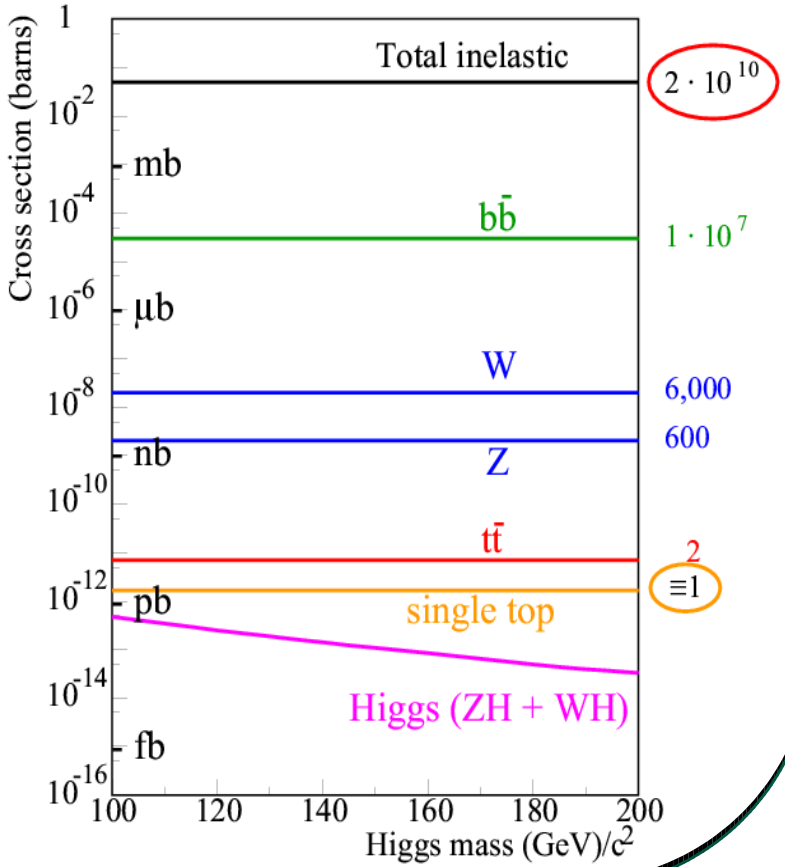


e:	p_T>20 GeV, $\lvert\eta\rvert$<2.0	p_T>15 GeV, $\lvert\eta\rvert$<1.1
μ:	p_T>20 GeV, $\lvert\eta\rvert$<1.1	p_T>18 GeV, $\lvert\eta\rvert$<2.0
Missing E_T	MET>25 GeV	15<MET<200 GeV
Jets	=2, p_T^{uncorr}>15 GeV, $\lvert\eta\rvert$<2.8	2-4, p_T>15 GeV, $\lvert\eta\rvert$<3.4 $p_{T,1}$>25 GeV, $\lvert\eta_1\rvert$<2.5 $p_{T,2}$>20 GeV
B-jet	1 or 2	

- Experimental signature similar to tt→lepton+jets but lower jet multiplicity.
- Main backgrounds: W+jets and tt.







Backgrounds

Diboson (WW/WZ/ZZ), Z->ττ

- Rather small contribution
- CDF: normalized to theoretical prediction
- DØ: included in W+jets estimate via normalization

Top pair production

- Sizable especially for events with ≥3 jets.
- Normalized to NNLO theoretical prediction

Multijets (or "non-W"): jets faking electrons or muons

- Estimated using data with leptons failing one or more isolation criteria.

W+jets: *dominant background*

- Event kinematics and flavor composition modeled using ALPGEN ME generator with PS-ME matching to avoid double/under-counting of radiation.
- Normalized to data before tagging and after subtracting other backgrounds:

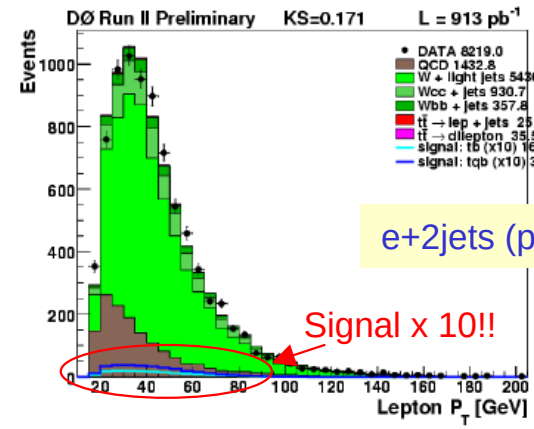
$$N^{data}_{W+jets} = N^{data}_{Candidates} - N_{non-W} - N_{t\bar{t}/EWK}$$

- Wbb/Wcc require additional scaling factor α~1.5 to match data (30% systematic uncertainty assigned)

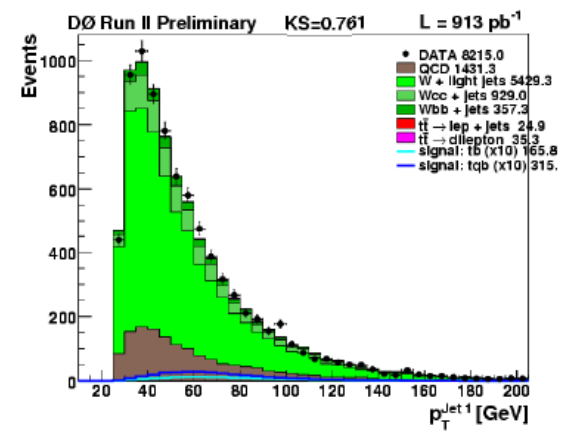
$$N^{data}_{Wbb} = \left(N^{MC}_{Wbb} / N^{MC}_{W+jets}\right) \times \alpha \times N^{data}_{W+jets}$$

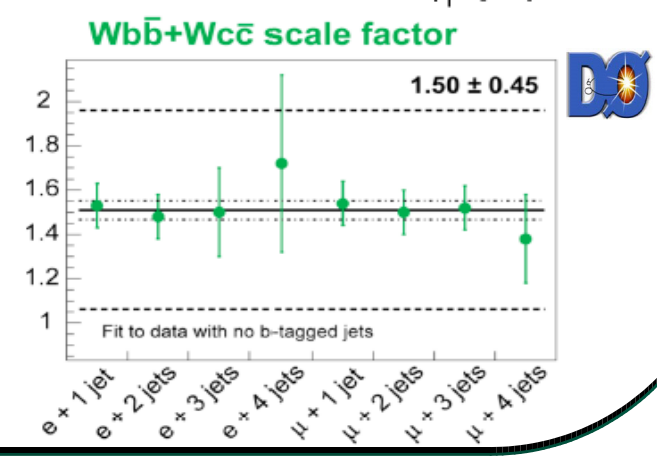
CDF: estimated in inclusive jet data

DØ: estimated in W+jets data with 0-tags



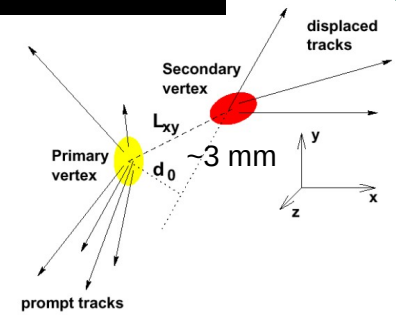
e+2jets (pre-tag)



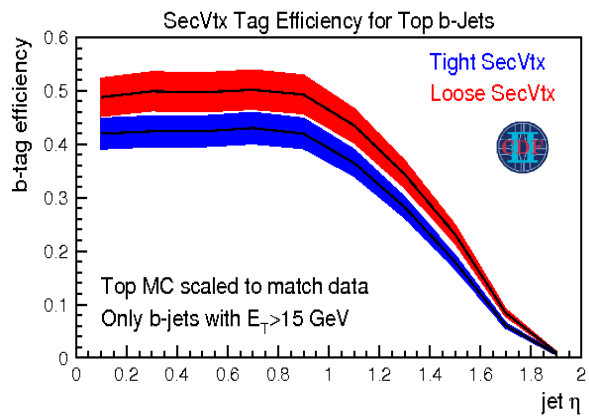


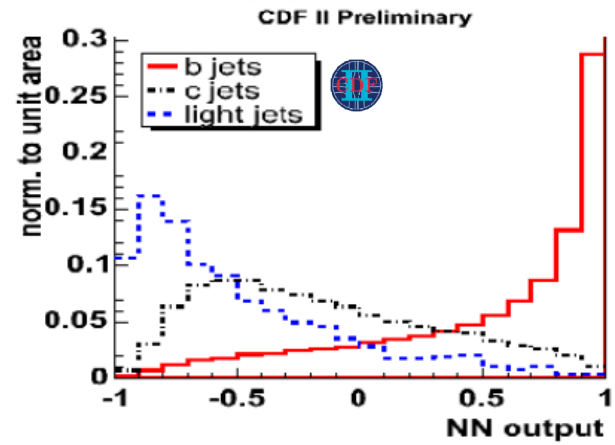
B-Jet Identification

- Exploiting b-quark content critical to improve S:B.
- CDF and DØ have developed high performance lifetime-based taggers.

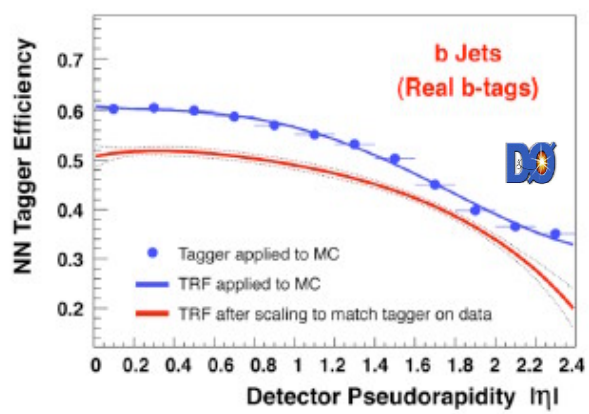


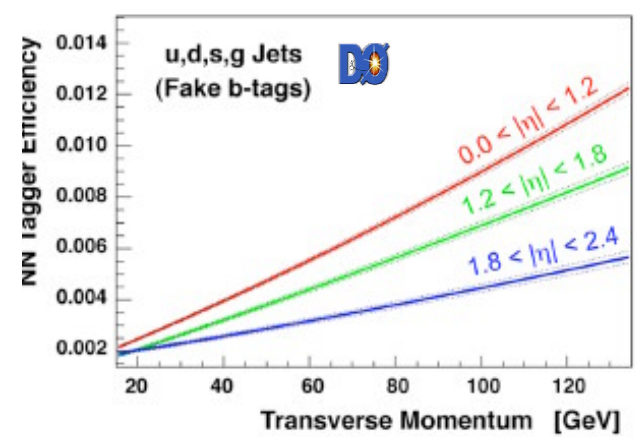
- Jets tagged using secondary vertex (SVX) :
  ε_b~40%, ε_{light}~0.8%
- NN trained using SVX information (25 variables) to further discriminate between b/c/light jets.
  Independent of b-source!





- Jets tagged using NN trained on 7 variables from existing SVX and impact parameter-based taggers.
- Operating point:
  ε_b~50%, ε_c~10%, ε_{light}~0.5%





- Tagged background predictions based on MC (tt, diboson, W+jets) use tagging efficiencies either calibrated to data (b/c) or directly from data (mistags).

Event Yields – After B-tagging

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

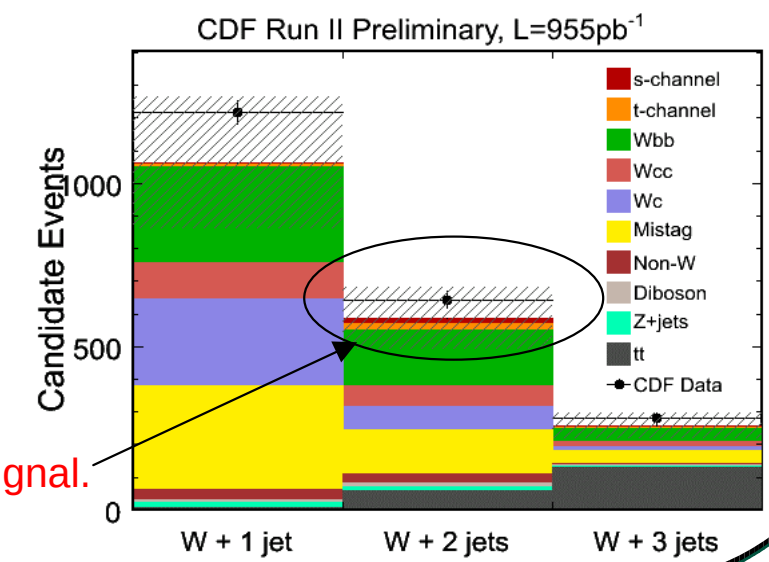
	S/B	S/√B
Before b-tagging	~1/200	~ 0.6
After b-tagging (1+2 b-tags)	~1/18	~ 1.5

Event Yields in 955 pb^{-1} Data
Electron+muon, 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

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
$t\bar{t} \to ll$	39 ± 9	32 ± 7	11 ± 3
$t\bar{t} \to l$+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



- Good agreement between expectation and observation after b-tagging.
- Total background uncertainty larger than expected signal. **Counting experiment hopeless!!**

Systematic Uncertainties

- Systematic uncertainties are assigned for each signal and background component in each of the analysis channels.
- Correlations between channels and sources are taken into account.
- Most uncertainties affect only the background normalization. Some uncertainties can in addition affect the shape of distributions.

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%	

Also affect (some) bkgs

Source of Uncertainty	Size	Shape
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%	X
Tag rate functions	2–16%	X

Background	Rate Uncertainty
W+bottom	25%
W+charm	28%
Mistag	15%
tt	23%

- Uncertainties included as nuisance parameters during the statistical phase of the analysis.

Search Strategy Overview

- In order to achieve the highest possible sensitivity, analyses underwent careful optimization:
  - Maximize acceptance (loose lepton identification, low p_T thresholds, wide η range, improved b-tagging performance,…)
  - Include as many channels as possible:
    DØ: 2-4 jets
  - Perform analysis in separate channels since S/B different and combine at the end:
    DØ: (e,μ) x (2,3,4 jets) x (1,2 tags) = 12
  - Develop sophisticated multivariate analysis techniques for best possible signal-to-background discrimination.
  - Optimize analysis for combined (tb+tqb) search (also perform separate searches).

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

Signal acceptances (including BR)

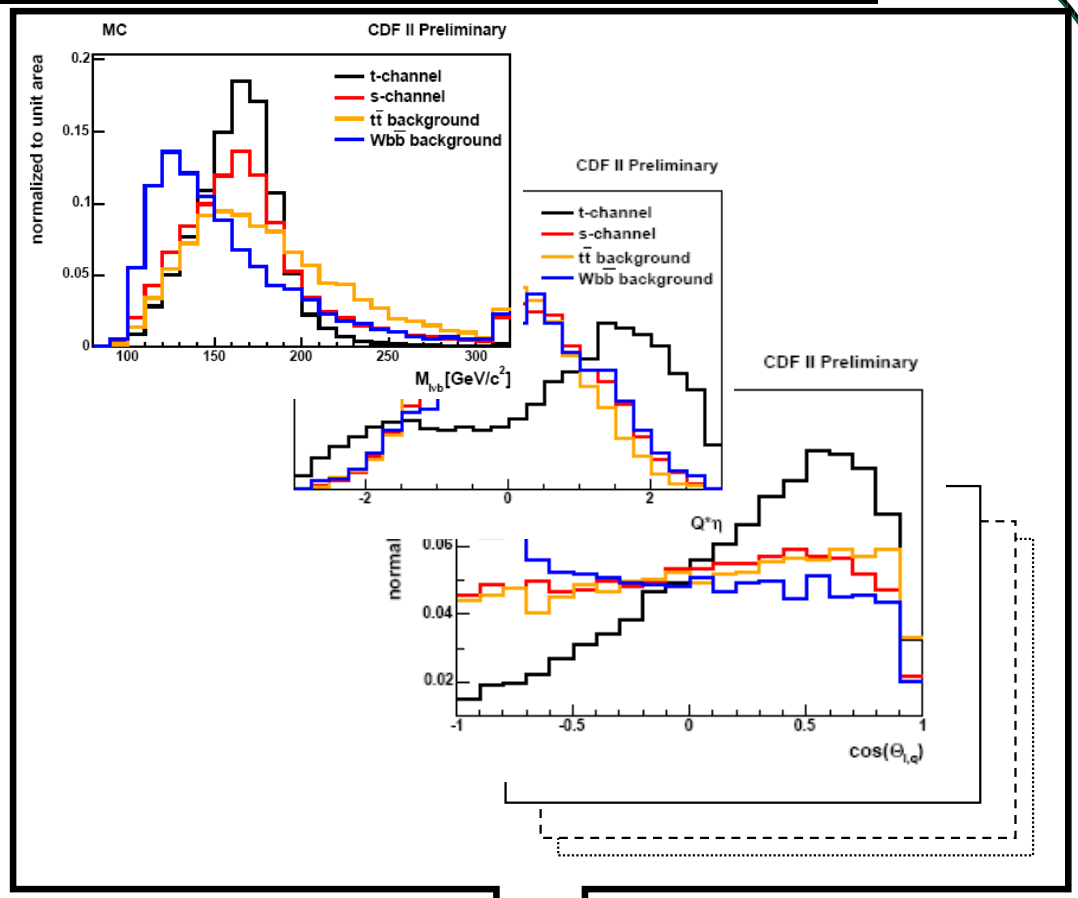
	tb	tqb
CDF (2 jets)	~1.9%	~1.3%
DØ (2-4 jets)	~3.2%	~2.1%

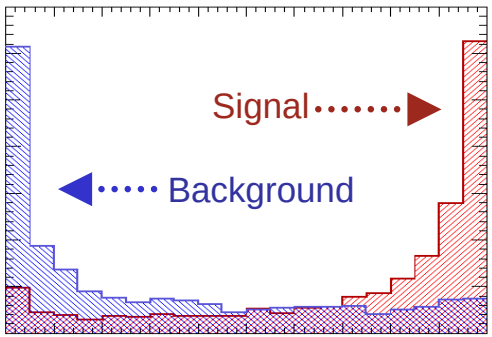
Multivariate Analysis Techniques

- A number of discriminant variables between signal and background can be identified:
  - B-tagging NN
  - Reconstructed top mass
  - Q(lepton)•η(untagged jet)
  - Top spin-related angular variables
  - ….

  but no single variable is powerful enough to cut on it.

  ⇒ Combine a number of variables into a single more powerful discriminant variable by using a multivariate analysis technique.





Event Discriminant

- Standard Neural Network
- Bayesian Neural Network (*)
- Likelihood Discriminants (*)
- Matrix Element Discriminants
- Boosted Decision Trees

(*) Not discussed in detail

Statistical Analysis

Cross section measurement:

- Binned likelihood from discriminant distribution using Poisson statistics.
- Compute posterior probability density of σ_{tb+tqb} using Bayes' theorem.
  - Flat positive-defined prior for σ_{tb+tqb}.: $\Theta(\sigma_{tb+tqb})$
  - Systematic uncertainties (α) treated as Gaussian nuisance parameters.

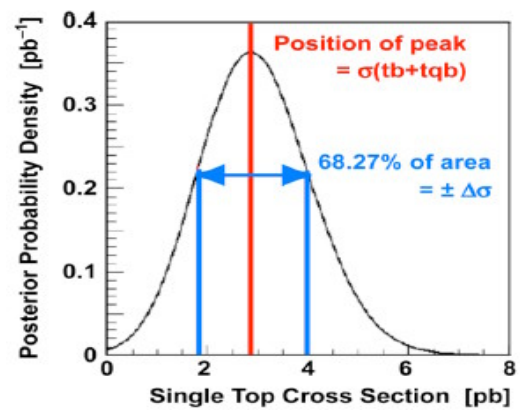
$$P(\sigma_{tb+tqb} \mid data) \propto L^*(data \mid \sigma_{tb+tqb})\Theta(\sigma_{tb+tqb})$$

"Profile" likelihood (maximized over systematics)

$$L^*(data \mid \sigma_{tb+tqb}) = L(data \mid \sigma_{tb+tqb}, \hat{\vec{\alpha}})$$

$$P(\sigma_{tb+tqb} \mid data) \propto \int \cdots \int d\vec{\alpha} L(data \mid \sigma_{tb+tqb}, \vec{\alpha}) \text{Prior}(\sigma_{tb+tqb}, \vec{\alpha})$$

$$\text{Prior}(\sigma_{tb+tqb}, \vec{\alpha}) = \left[\prod_{Systematic\ s} G(0, \sigma_{\alpha_i}) \right] \Theta(\sigma_{tb+tqb})$$



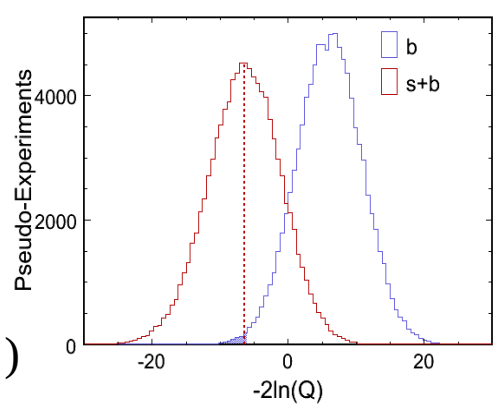
- Extract most probable value and CL intervals from $P(\sigma_{tb+tqb}|data)$.

Significance:

- CDF: fraction of background-only pseudo-experiments with $Q \leq Q^{obs}$

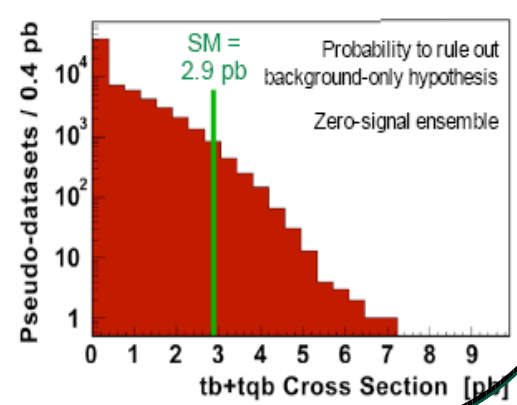
$$Q = \frac{L(data \mid s+b)}{L(data \mid b)}$$

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



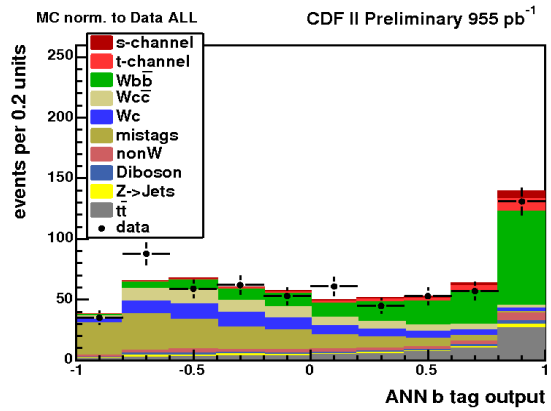
- DØ: fraction of background-only pseudo-experiments with $\sigma \geq \sigma^{obs}$

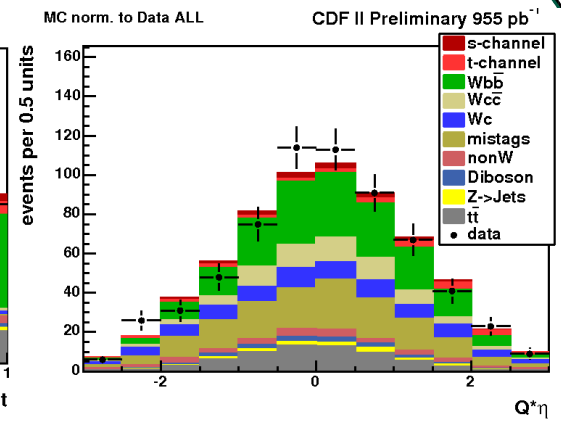
$$p-value = \int_{\sigma^{obs}_{tb+tqb}}^{\infty} d\sigma_{tb+tqb} P_b(\sigma_{tb+tqb})$$

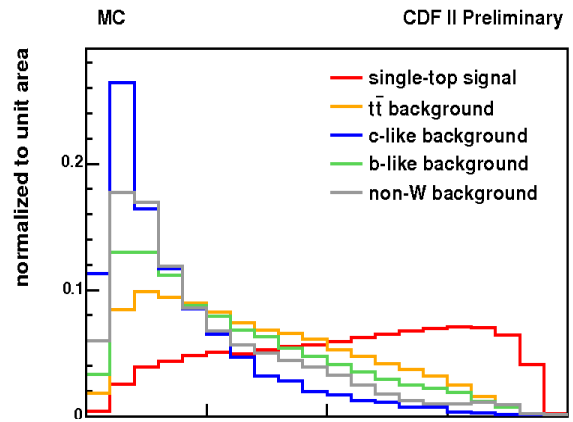


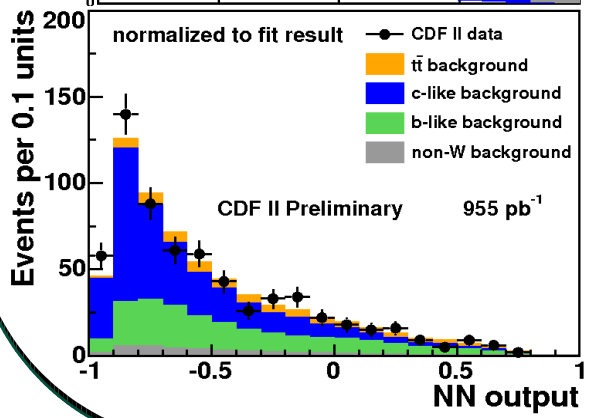
Neural Network Analysis

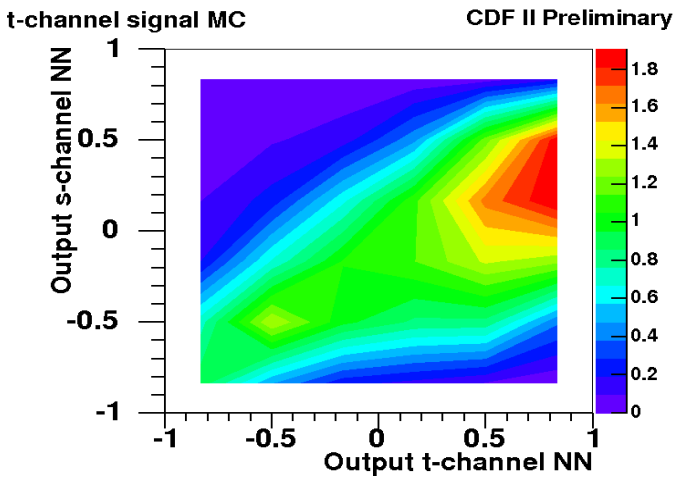
- Consider 26 kinematic or event-shape variables: b-tagging NN, reconstructed top mass, Qxη, etc.
- Three NNs trained for combined (tb+tqb) and separate (tb and tqb) searches.
- Build templates for five categories: signal, tt, c-like, b-like and non-W.

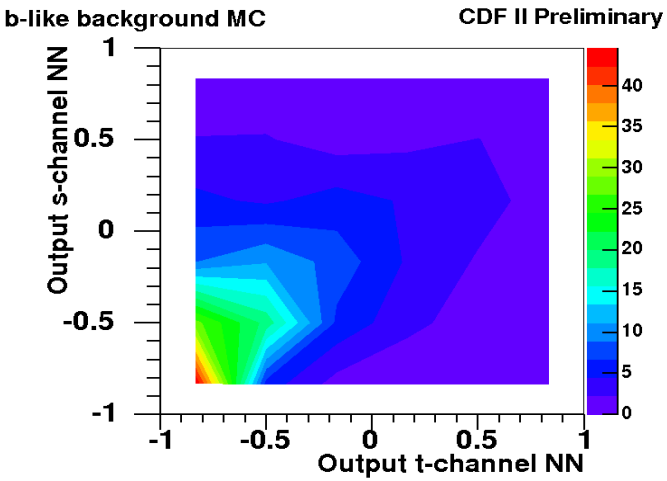












Expected Performance	Median p-value (CDF)	σ_{95} (*)
Combined search	0.5% (2.6σ)	2.6 pb
Separate search	0.4% (2.7σ)	1.3 pb (tqb) 1.5 pb (tb)

(*) Assuming no single top

Matrix Element Discriminants

- Pioneered by DØ in Run I top mass measurement. Now being used in a search.
- Attempt to make an optimal use of the kinematic information in the event.
- Principle: compute event probability density for a given hypothesis (e.g. single top) making use of all reconstructed objects in the event (integrate over unknowns).

E.g. 2 jet events: $\vec{x} = \{p_\ell^\mu, p_{j1}^\mu, p_{j2}^\mu\}$ — Only 6 "discriminant variables" used!!

differential cross section (LO matrix element)

parton distribution functions

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

$$d^n\sigma_i(\vec{y}) = \frac{(2\pi)^4}{2s}\,|M_i(\vec{y})|^2\, d\Phi^n(\vec{y})$$

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

Maximize sensitivity by:

- summing over all permutations of jets and neutrino solutions
- allowing better measured events to contribute more (via the transfer function)
- implementing b-tagging information

- Matrix element discriminant defined as:

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

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

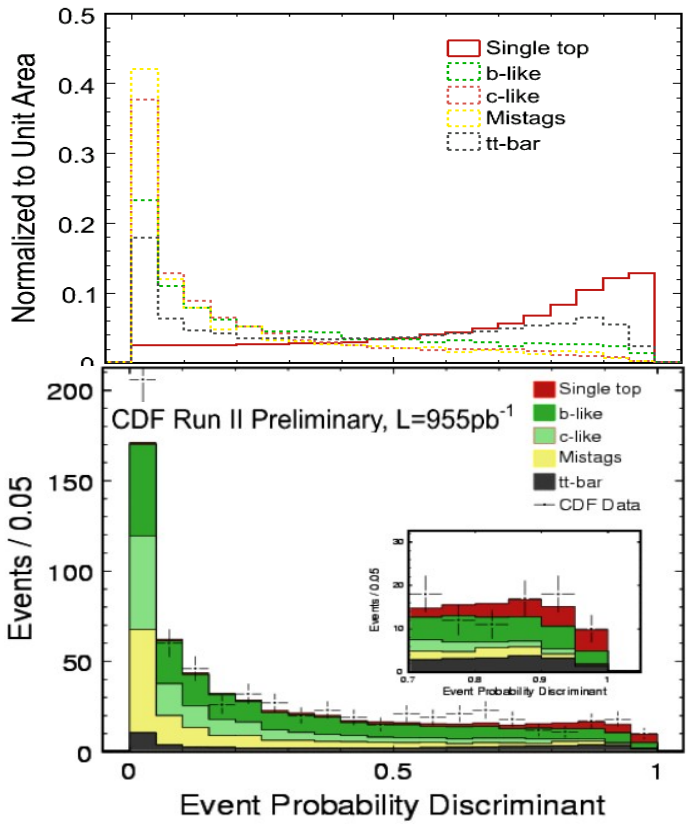
$$D_S(\vec{x}) = \frac{P_S(\vec{x})}{P_S(\vec{x}) + P_{bckg}(\vec{x})}; \quad S = tb \text{ or } tqb$$

$$P_{bckg}^{2j}(\vec{x}) = c_{Wbb}P_{Wbb}(\vec{x}) + c_{Wcg}P_{Wcg}(\vec{x}) + c_{Wgg}P_{Wgg}(\vec{x})$$

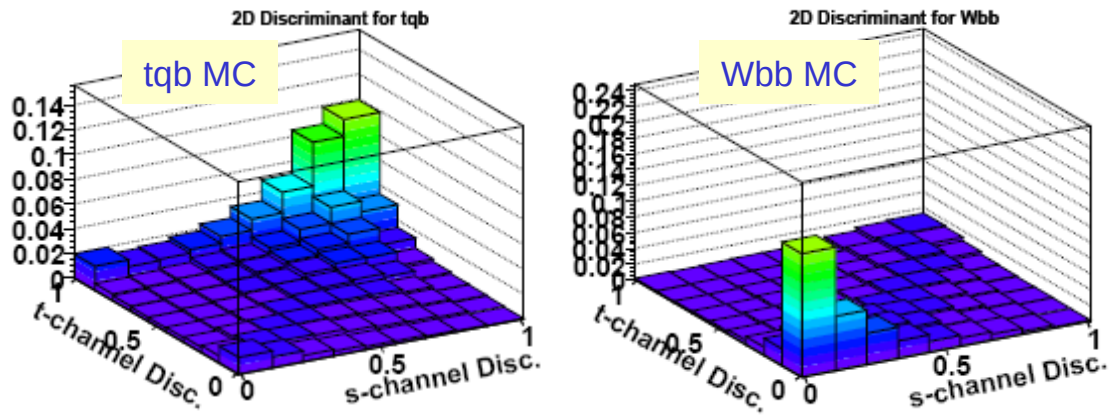
$$P_{bckg}^{3j}(\vec{x}) = P_{Wbbg}(\vec{x})$$

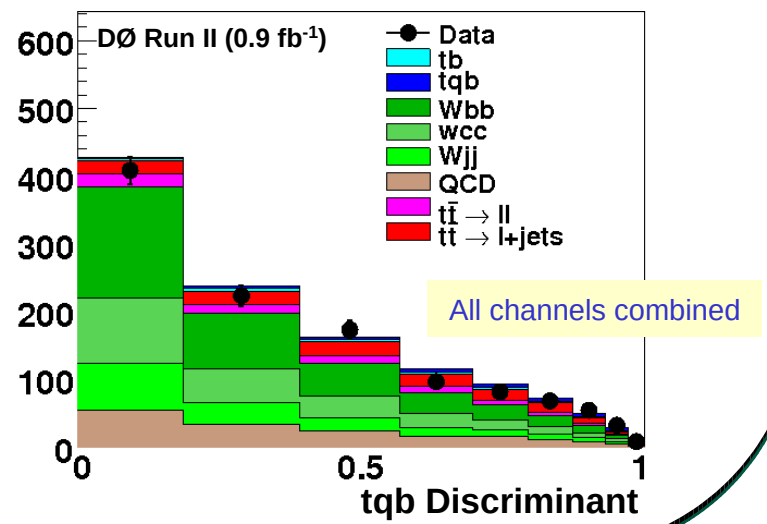
Matrix Element Analyses

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



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

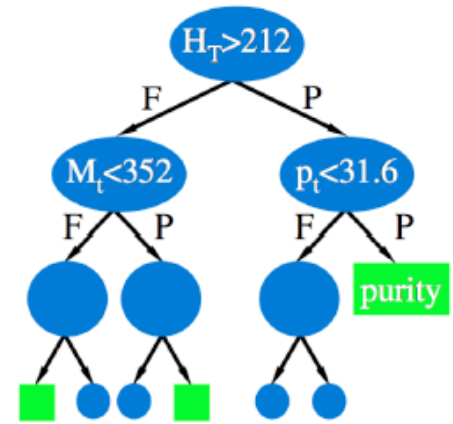




Not comparable!

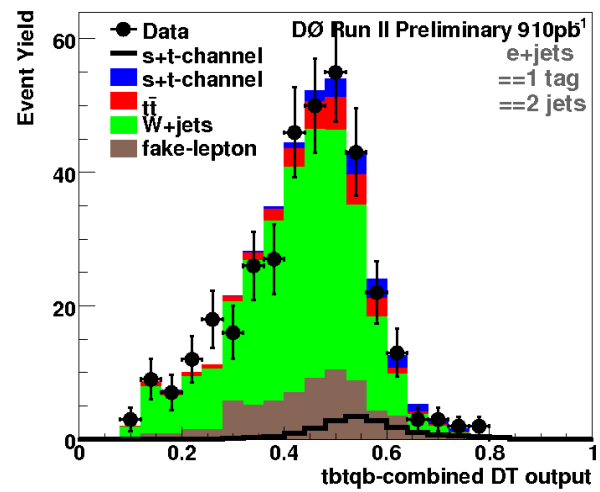
Expected Performance	median p-value (CDF)	p-value (DØ)
CDF ME analysis	0.5% (2.6σ)	
DØ ME analysis		3.6% (1.8σ)

Boosted Decision Trees Analysis



- Machine learning technique, widely used in social sciences, some use in HEP (e.g. MiniBooNe).
- Idea: recover events that fail criteria in cut-based analyses.

- Start at first "node " with "training sample" of all signal and background events.
- Select variable and splitting value with best separation to produce two "branches".
- Repeat recursively at each node. Stop when there is no improvement or too few events are left.
- DT output = "leaf " purity, close to 1(0) for signal(background)
- Improve performance of DT by using "adaptive boosting", which averages over many trees, diluting the piecewise nature of the DT output.



- Consider a total of 49 variables to discriminate between signal and backgrounds.
  - Most sensitive variables: M(all jets), M(W,b_1) ("top mass"), Qxη, cosθ(lepton,b_1)$|_{\text{top rest-frame}}$
  - Adding more variables does not reduce discrimination.
  - Reducing number of variables always reduces sensitivity.
  - Same list of variables used for all analysis channels.
- Trained 36 sets of trees:

  (tb+tqb, tb, tqb) x (e,μ) x (1,2,3,4 jets) x (1tag, 2tags)
  - Signal trained against sum of all backgrounds.
  - Combined search (tb+tqb) has best sensitivity.

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

Cross-Check Samples

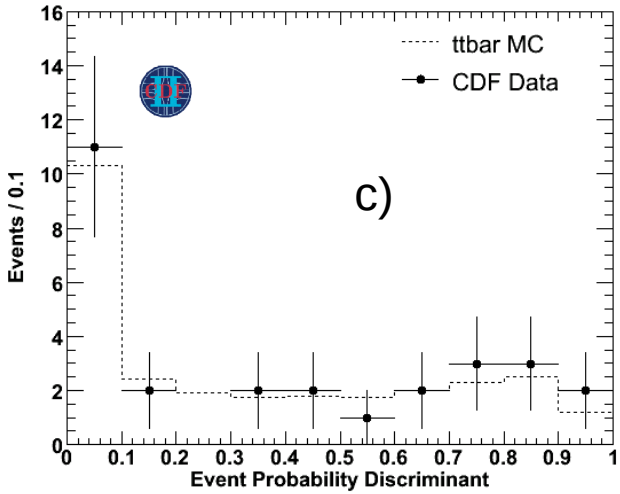
- Crucial to validate background model in “side-band” regions without looking at single top candidates.
- Examples:

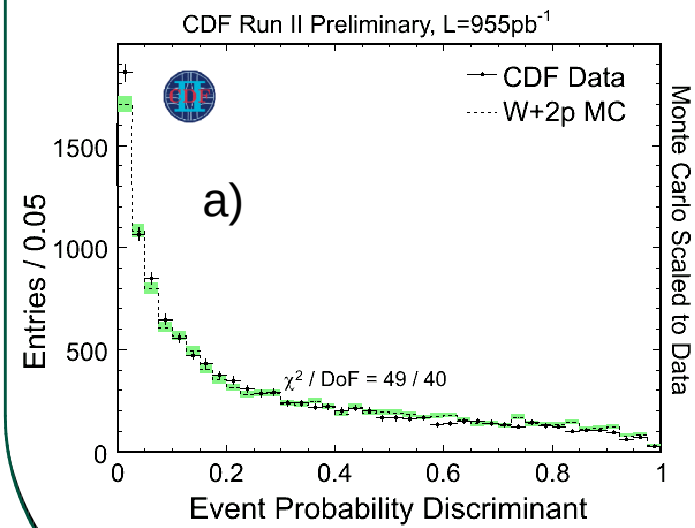
a) ME discriminant for 2-jet events with no b-tags (enriched W+light jets bckg)

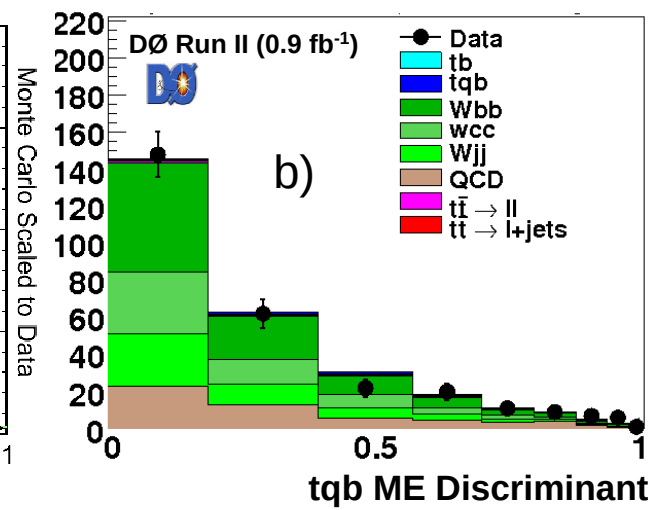
b) ME discriminant for 2-jet events with ≥1 b-tag and H_T<175 GeV (enriched W+heavy jets bckg)

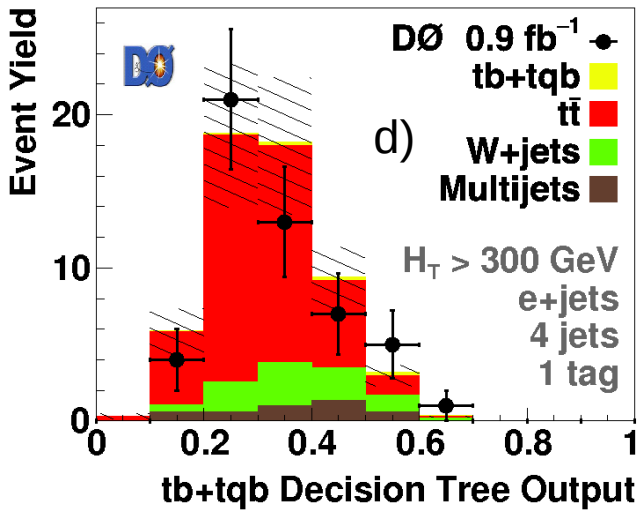
c) ME discriminant for b-tagged dilepton+2-jet events (~pure tt→dilepton bckg)

d) DT discriminant for 4-jet events with ≥1 b-tag and H_T>300 GeV (enriched tt→lepton+jets bckg)





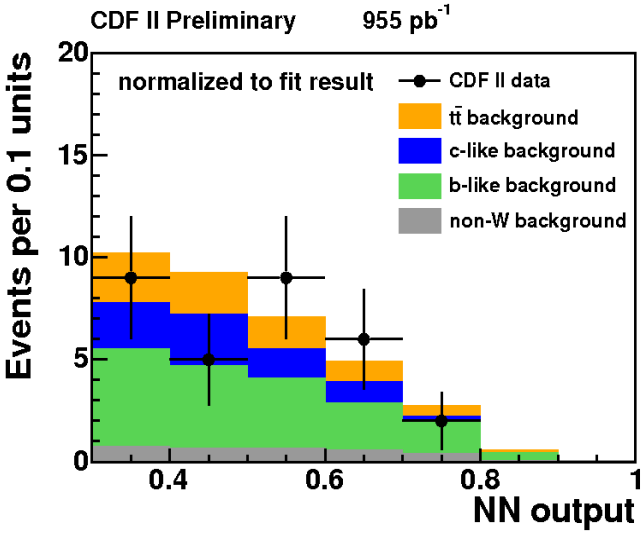




⇒ Good agreement observed.

CDF Results

NN analysis

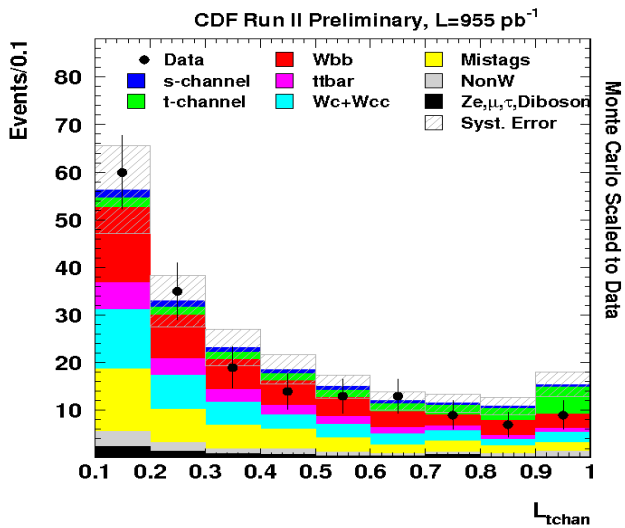


$\sigma_{tb} = 0.7^{+1.5}_{-0.7}\ pb;\ \sigma_{tqb} = 0.2^{+1.1}_{-0.2}\ pb$

$\sigma_{tb+tqb} < 2.6\ pb\ @\ 95\%\ CL$

No evidence of signal

Likelihood Function analysis

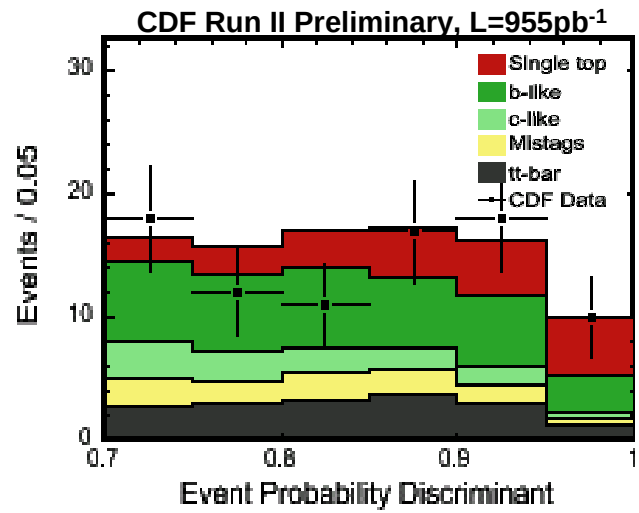


$\sigma_{tb} = 0.1^{+0.7}_{-0.1}\ pb;\ \sigma_{tqb} = 0.2^{+0.9}_{-0.2}\ pb$

$\sigma_{tb+tqb} < 2.7\ pb\ @\ 95\%\ CL$

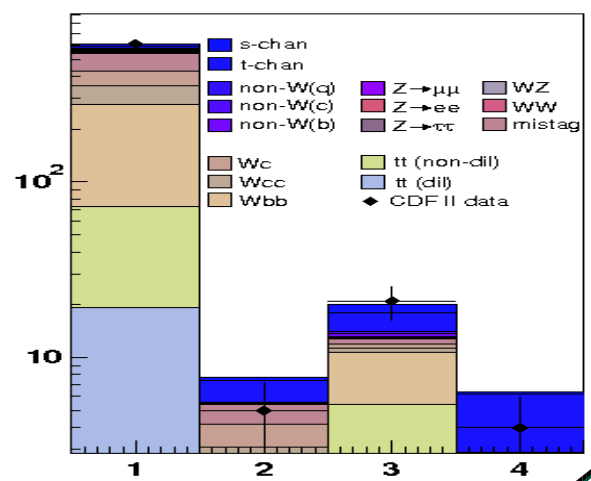
No evidence of signal

Matrix Element analysis

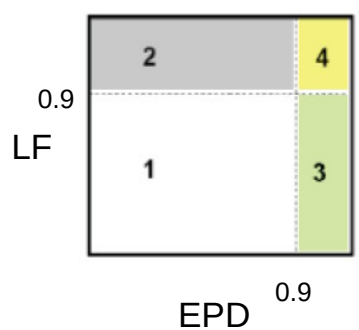


$\sigma_{tb+tqb} = 2.7^{+1.5}_{-1.3}\ pb$

Observed p-value: 1.0% *(2.3σ)*



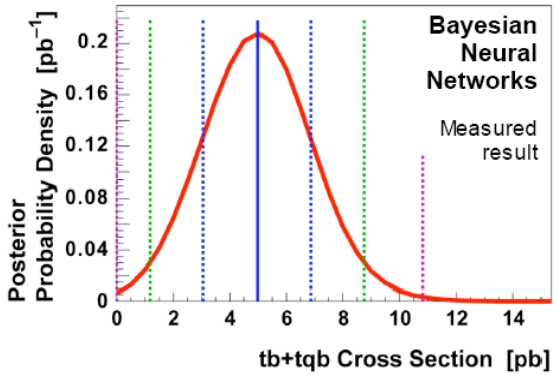
- Studies ongoing to quantify compatibility of results. ME-LF:
  - correlation of fit results: ~53%
  - compatibility of measurement in data: ~4-6%
- Analyzing more data should shed some light.



Prob(null hyp) = 33.7%
Prob(signal hyp) = 49.8%

DØ Results

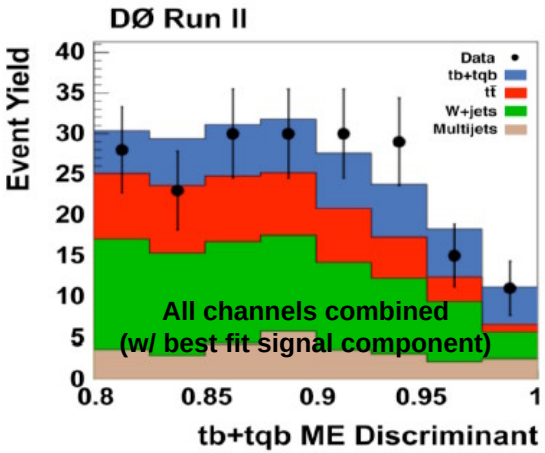
Bayesian NN analysis



$\sigma_{tb+tqb} = 5.0 \pm 1.9\ pb$

Observed p-value: 1.15% *(2.3σ)*

Matrix Element analysis

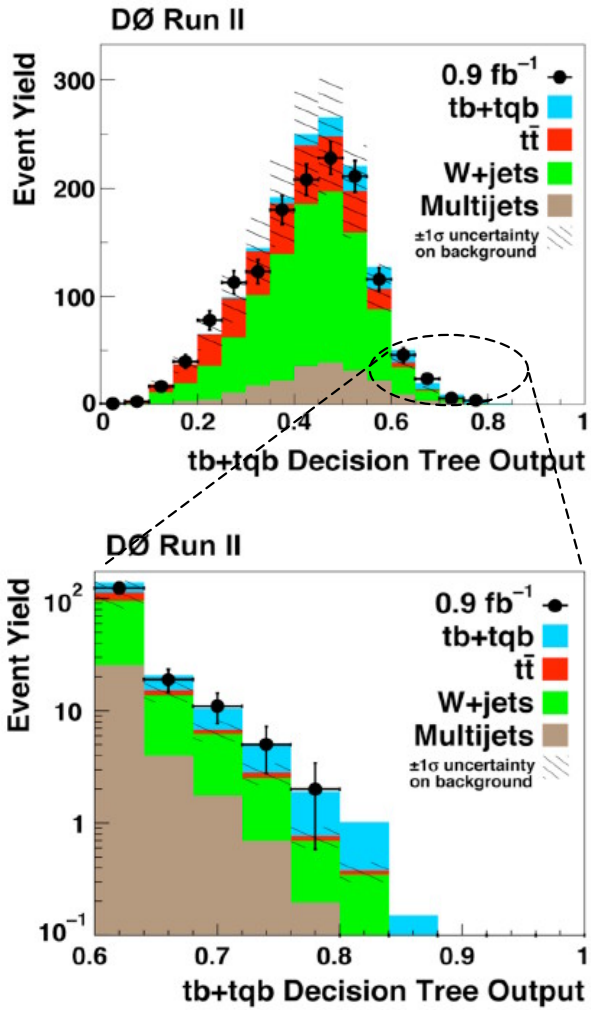


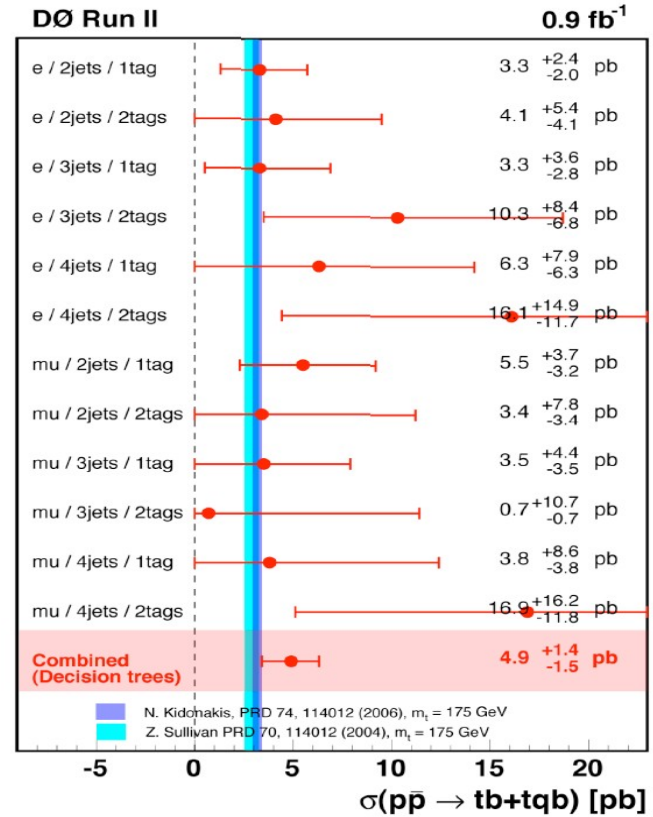
$\sigma_{tb+tqb} = 4.8^{+1.8}_{-1.5}\ pb$

Observed p-value: 0.20% ***(2.9σ)!***

Boosted Decision Trees analysis

All channels combined
(w/ best fit signal component)





$\sigma_{tb+tqb} = 4.9 \pm 1.4\ pb$

Observed p-value: 0.034% ***(3.4σ) !!***

Compatibility with SM: 11%

hep-ex/0612052
Submitted to Phys. Rev. Lett.

Evidence for single top production!

First Direct Measurement of $|V_{tb}|$

- Large $m_t \Leftrightarrow$ New Physics (EWSB-related)??
  $\Rightarrow$ interactions between the top quark and weak gauge bosons extremely interesting!!
- Most general *tbW* vertex:

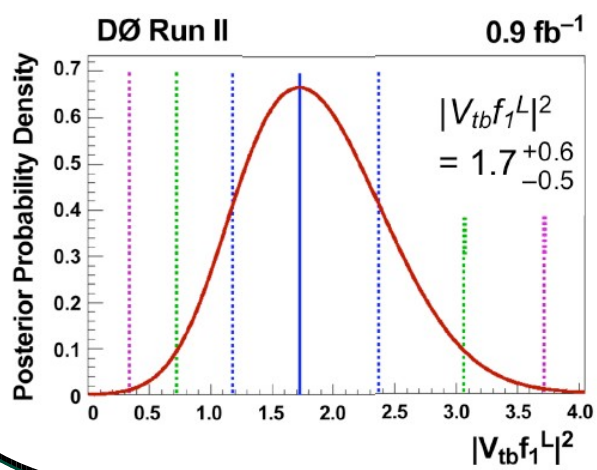
$$\Gamma^{\mu}_{tbW} = -\frac{g}{\sqrt{2}} V_{tb} \left\{ \gamma^{\mu} \left[f_1^L P_L + f_1^R P_R \right] - \frac{i\,\sigma^{\mu\nu}}{M_W} (p_t - p_b)_{\nu} \left[f_2^L P_L + f_2^R P_R \right] \right\}$$

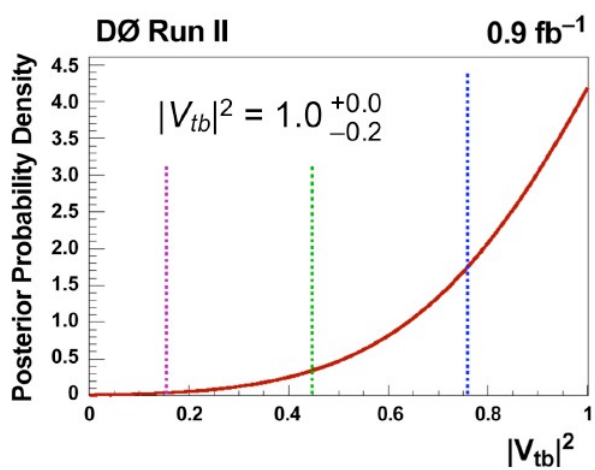
$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}$$

Within the SM:

- 3-generation and unitary CKM matrix: $V_{tb} \sim 1$ (if assumption released, V_{tb} essentially unconstrained)
- CP conserving pure V-A interaction: $f_1^L = 1,\ f_1^R = f_2^L = f_2^R = 0$

- Use (Decision Trees) single top cross section to obtain <u>first direct measurement</u> of $|V_{tb}|$.
- <u>Assumptions</u>:
  - No other source of single top except via the *tqW* interaction
  - $|V_{td}|^2 + |V_{ts}|^2 << |V_{tb}|^2$ (i.e. *tbW* production mechanism dominates and SM top decay B(t→Wb)~1)
  - CP conserving pure V-A interaction (but not necessarily of SM strength!)
  - No assumption about number of generations or unitarity of CKM matrix.





Strength of the V-A interaction:

$$|V_{tb} f_1^L| = 1.3 \pm 0.2$$

Assuming $f_1^L = 1$:

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

hep-ex/0612052
Submitted to Phys. Rev. Lett.

Summary and Conclusions

- Single top search very challenging: small signal hidden in large complex background.

- CDF and DØ have invested much time and effort in developing optimized selections, detailed background understanding and sophisticated analysis techniques.
  Such developments benefit the whole Tevatron physics program and are of great importance for the Higgs search.

- DØ's Boosted Decision Trees analysis (0.9 fb^{-1}) finds 3.4σ evidence for single top production:

$$\sigma_{tb+tqb} = 4.9 \pm 1.4\ pb$$

First direct measurement of the strength of the *tbW* interaction (*):

$$|V_{tb} f_1^L| = 1.3 \pm 0.2$$

$$|V_{tb}| \geq 0.68\ @95\%CL\ (\text{assuming } f_1^L = 1)$$

(*) See previous slide for details on assumptions

⇒ Major milestone achieved that opens the door to detailed studies of the *tbW* interaction and searches for new physics (W', FCNC,…).

- Both experiments have already a factor of two more data available and many ideas on how to improve further the analyses!

Stay Tuned!!

Backup

For More Information…

- DØ's "Single top evidence" public webpage (contains links to individual single top analyses):

  http://www-d0.fnal.gov/Run2Physics/top/public/fall06/singletop/

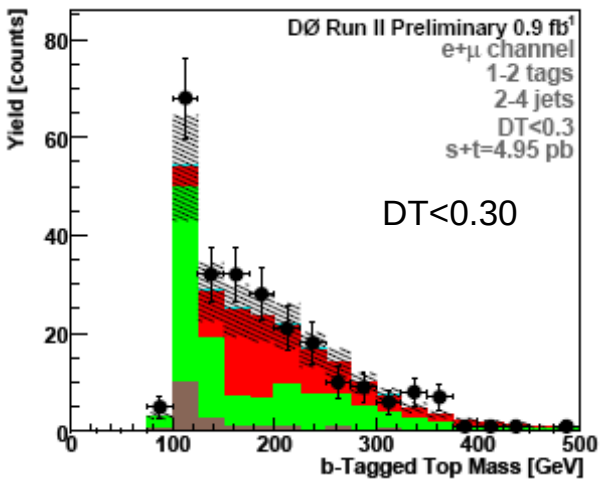
- CDF Top Group's public webpage (contains links to individual single top analyses):

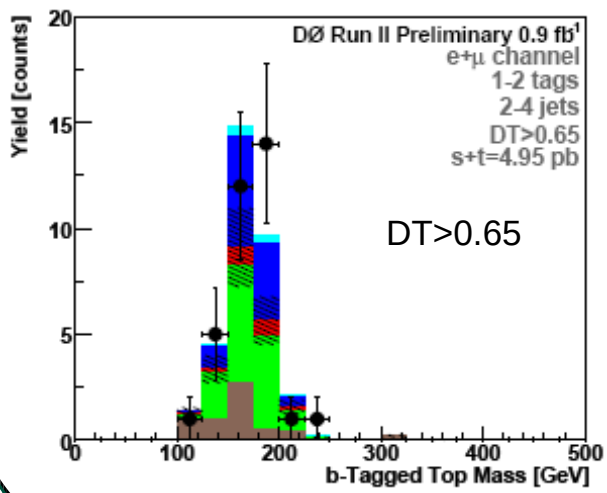
  http://www-cdf.fnal.gov/physics/new/top/top.html

DØ Results

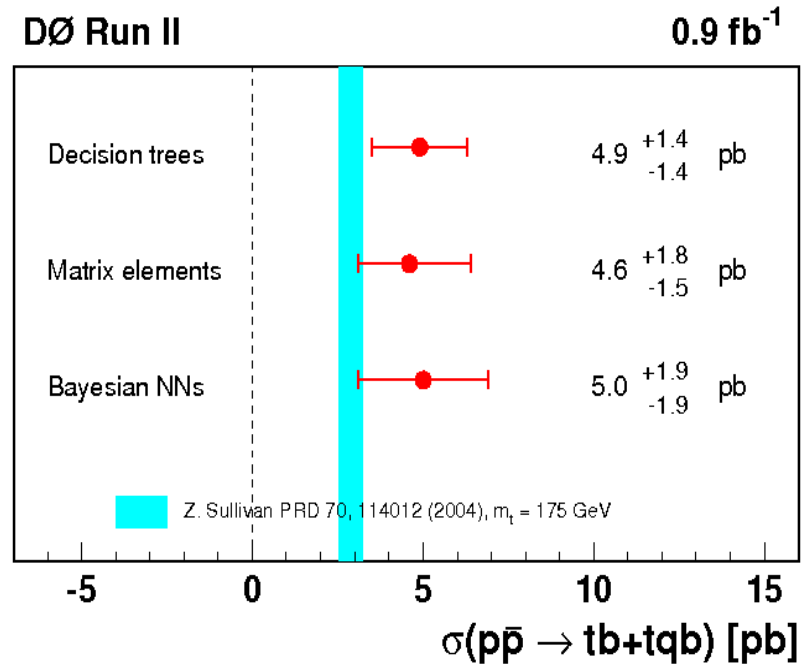
Event Characteristics

All channels combined
(w/ best fit signal component)





Summary of Measurements



- Results from the three analyses are consistent with each other.
- Overlap within the subset of 50 highest-discriminant events per analysis: ~50%.
- Preliminary estimate of correlation in measured cross section between analyses using pseudo-experiments (incl. syst. uncertainties): ~50%.

Combined cross section and significance will soon become available.

This document was created with Win2PDF available at http://www.daneprairie.com.
The unregistered version of Win2PDF is for evaluation or non-commercial use only.