Evidence for Single Top Quark Production at D0

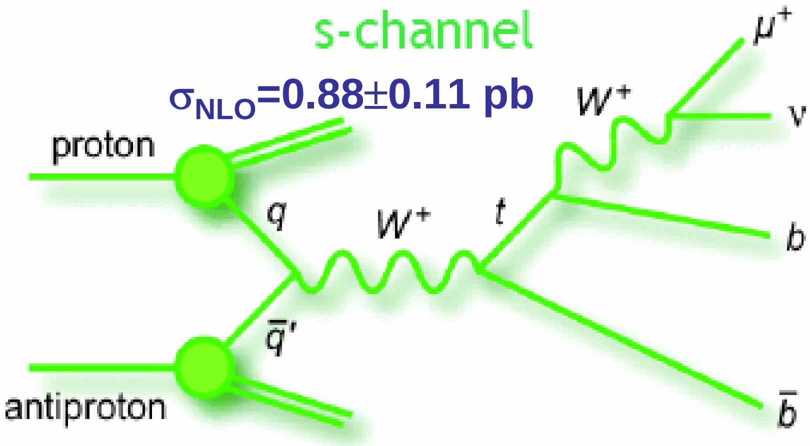
Liang Li
University of California, Riverside

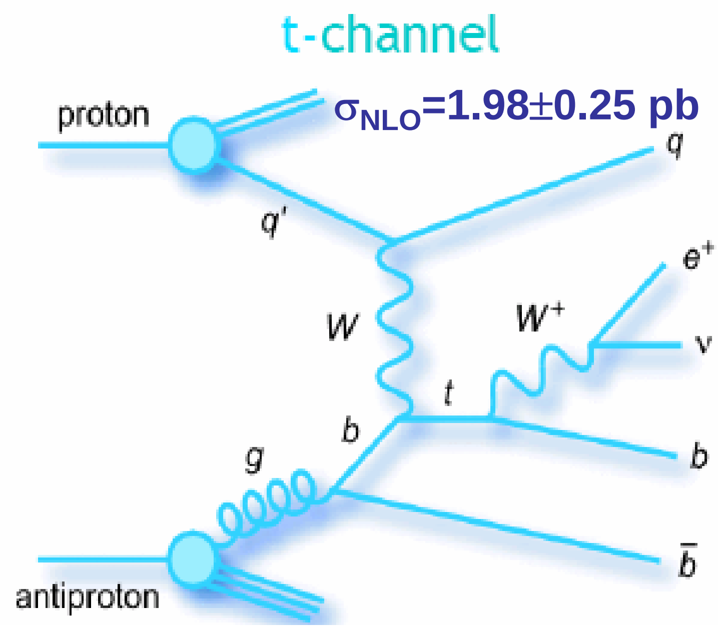
On Behalf of the DØ Collaboration

Single Top Quark Production

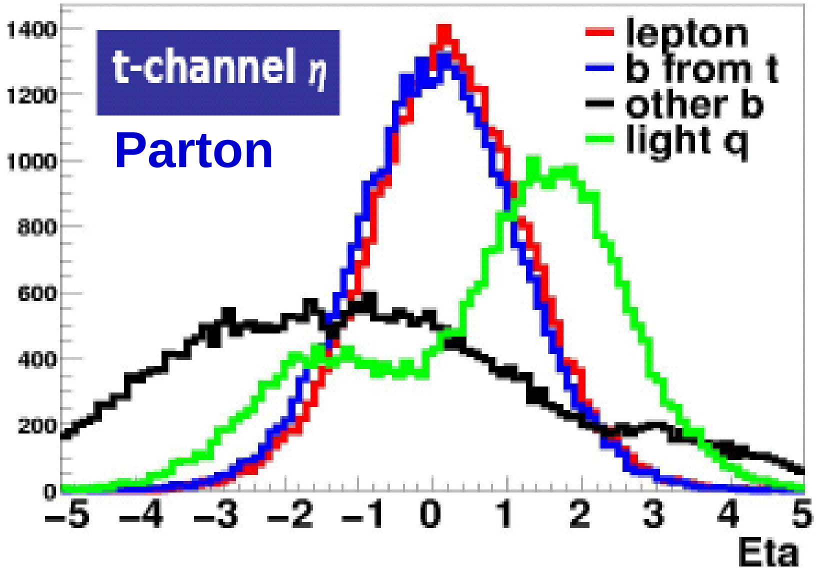
Single Top: Electroweak Interaction

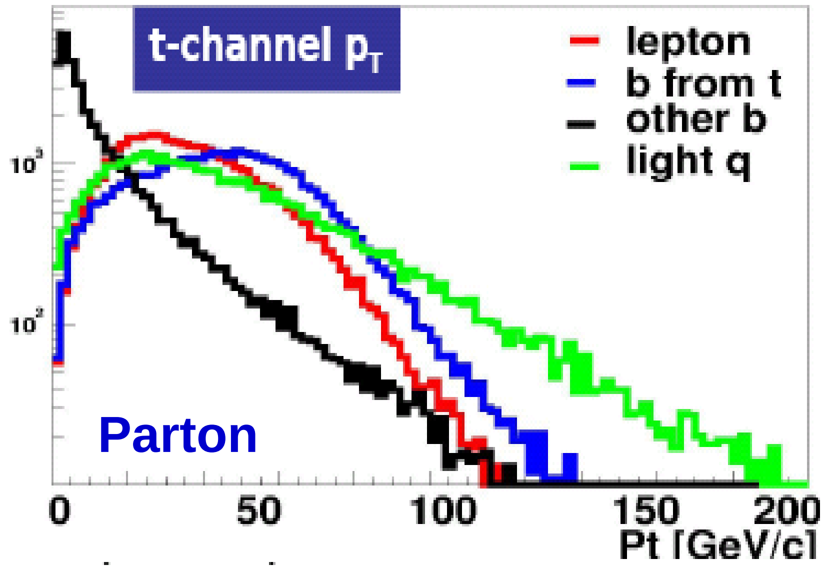
- **Study Wtb coupling**
  - Direct measurement of the |Vtb| CKM matrix element
  - Test of CKM unitarity
  - Anomalous *Wtb* couplings
- **New physics, example:**
  - s-channel is sensitive to W', H^+
  - t-channel is sensitive to FCNC
  - 4^{th} quark generation?
- **Study top properties:**
  - Polarization, decay width, lifetime…
- **Background study helps many physics searches, e.g. SM Higgs**





Event Selection





Event Topology:

- Energetic isolated lepton (from W)
- Missing E_T (ν from W)
- One b-quark jet (from top)
- A light flavor jet and/or another b-jet

One isolated lepton

- Electron p_T > 15 GeV, $|\eta|$ < 1.1
- Muon p_T > 18 GeV, $|\eta|$ < 2.0

Missing E_T > 15 GeV

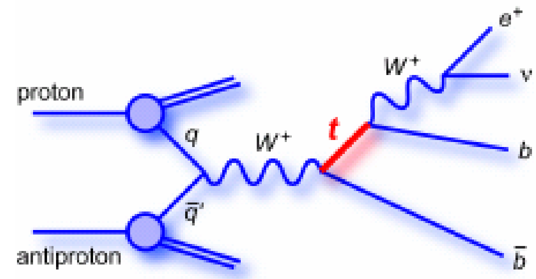
One or two b-tagged jet

- 2–4 jets: p_T > 15 GeV, $|\eta|$ < 3.4
- Leading jet: p_T > 25 GeV, $|\eta|$ < 2.5
- Second leading jet: p_T > 20 GeV

Signal and Background Modeling

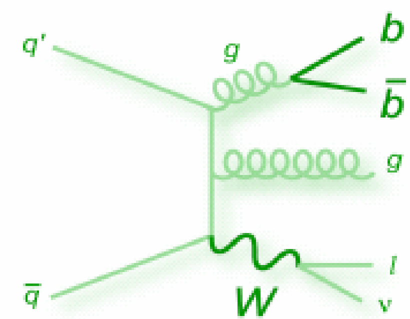
Signal:

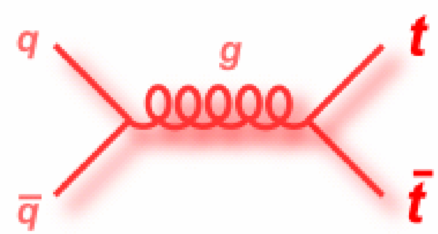
- CompHEP-SINGLETOP
- Distributions agree well with ZTOP & MCFM (NLO)

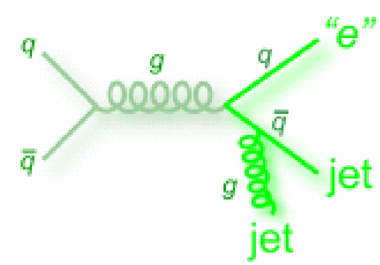


Background:

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







Event Yields and Systematics

Source	Event Yields in 0.9 fb^{-1} Data Electron+muon, 1tag+2tags combined		
	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

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

Component	Size
W+jets&QCD normalization	18 – 28%
top pair normalization	18%
Tag rate functions (shape)	2 – 16%
Jet energy scale (shape)	1 – 20%
Luminosity	6%
Trigger modeling	3 – 6%
Lepton ID	2 – 7%
Jet modeling	2 – 7%
Other small components	few%

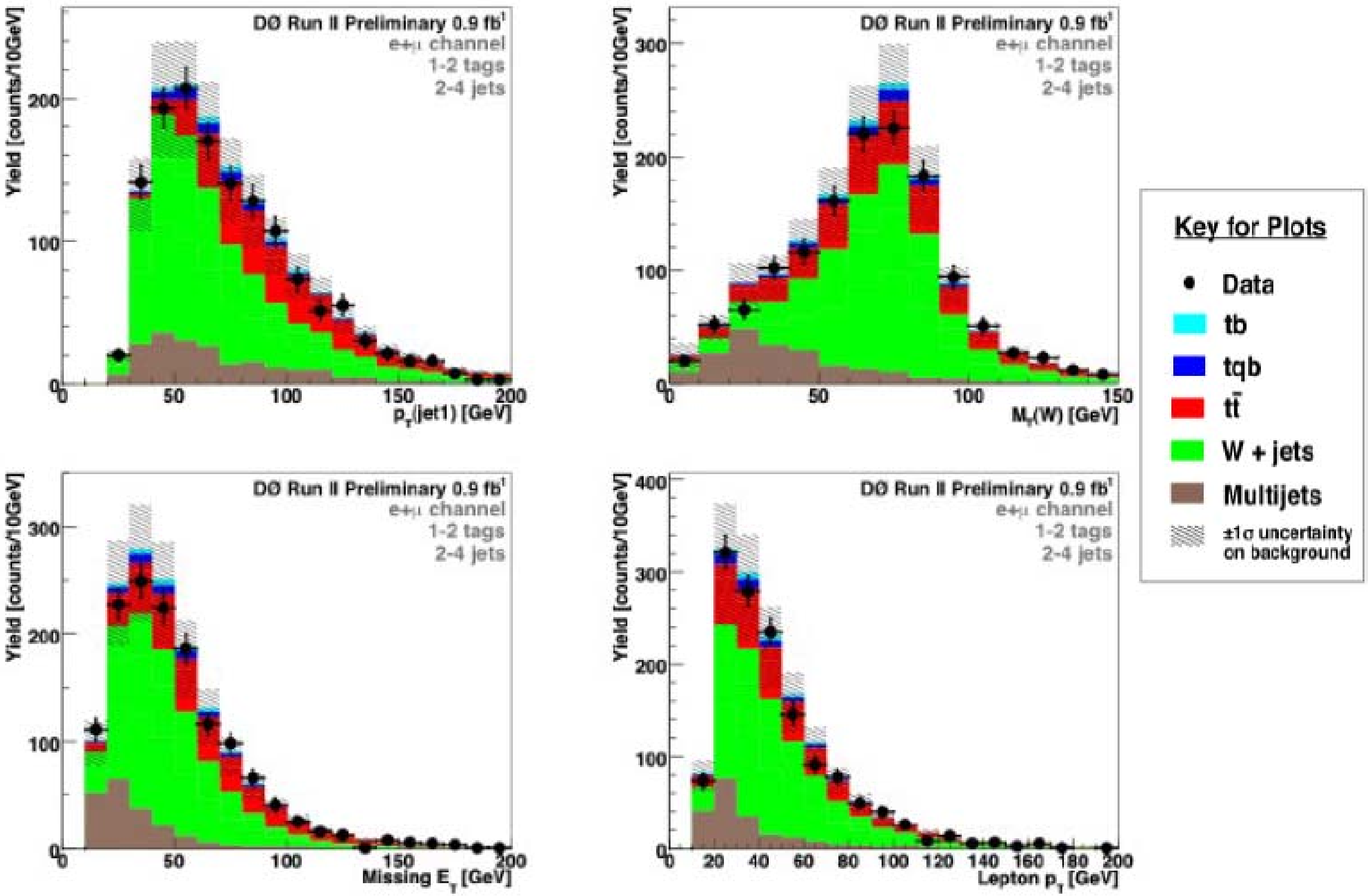
Expect 62 signal and 1398 bkgd events

Use multivariate discriminant to separate signal from background

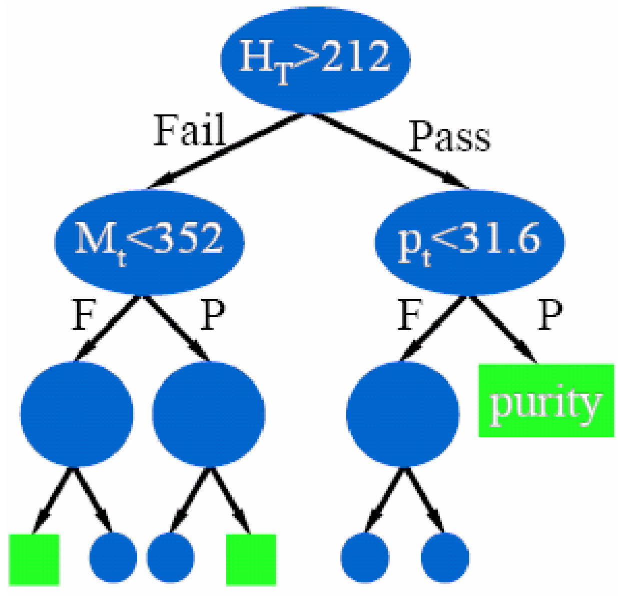
Most systematic uncertainties apply only to normalization, except jet energy scale and b-tagging which affect shapes

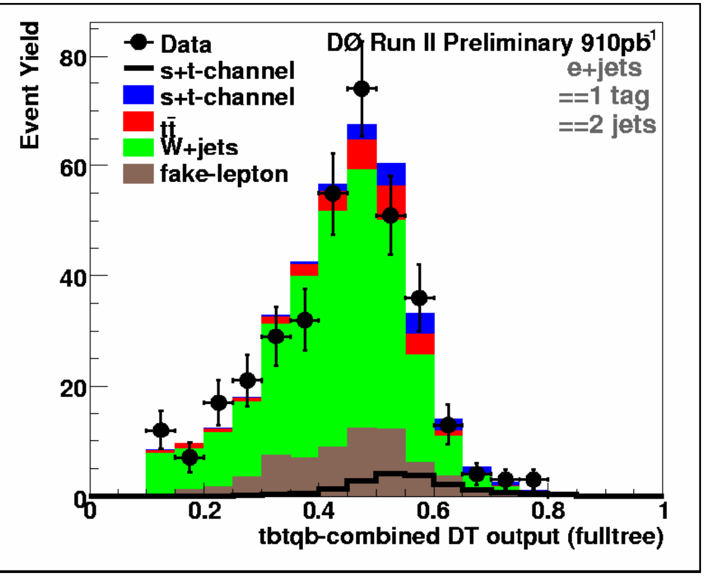
Cross section uncertainties are dominated by the statistical uncertainty

Background Modeling



Decision Tree Method





Train on 1/3 of bkgd and signal sample

- Creates a tree, with a simple straight cut at every branch point
- Select variable and splitting value with best separation, repeat recursively. Stop if improvement stops or too few events
- Each leaf classifies an event with a purity $N_S/(N_S+N_B)$
- Measurements done on the other 2/3's of signal and background sample

Boosting:

- Retrain 20 times to learn from misclassified events

Discriminating variables:

- 49 physics motivated variables
- $M_{alljets}$, $M_{W,b\text{-}jet1}$, $\cos_{b\text{-}jet1,lepton}$, $q_{lepton}*\eta_{light\text{-}jet}$

Matrix Element Method

Use full kinematical information contained in event: the four vectors from the reconstructed lepton and jets

Use matrix elements of main signal and background Feynman diagrams to compute an event probability density for signal and background hypotheses

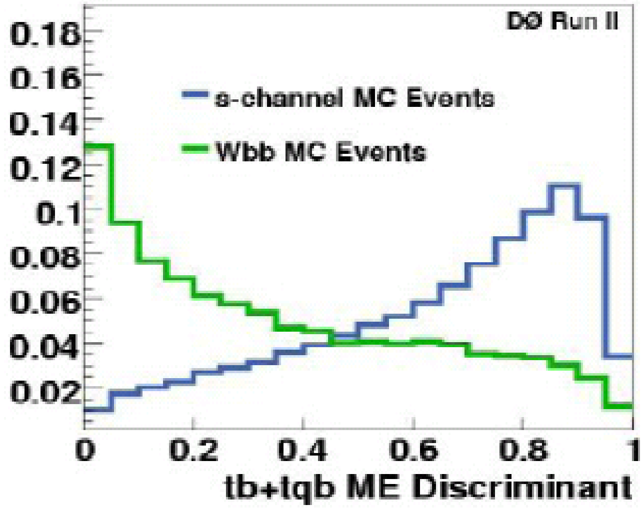
$$P(\vec{x}) = \frac{1}{\sigma} \int f(q_1;Q)\,dq_1\, f(q_2;Q)\,dq_2 \times |M(\vec{y})|^2 \phi(\vec{y})\,dy \times W(\vec{x},\vec{y})$$

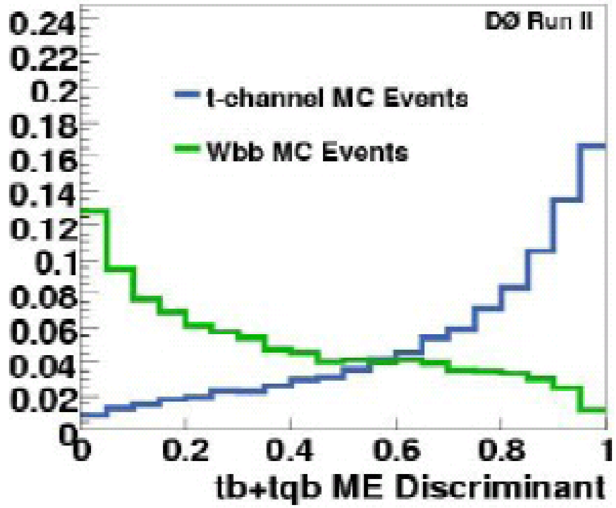
Parton distribution functions CTEQ6

Differential cross section (LO ME from Madgraph)

Transfer Function: maps parton level (y) to reconstructed variables (x)

$$D_s(\vec{x}) = P(S|\vec{x}) = \frac{P_{Signal}(\vec{x})}{P_{Signal}(\vec{x}) + P_{Background}(\vec{x})}$$



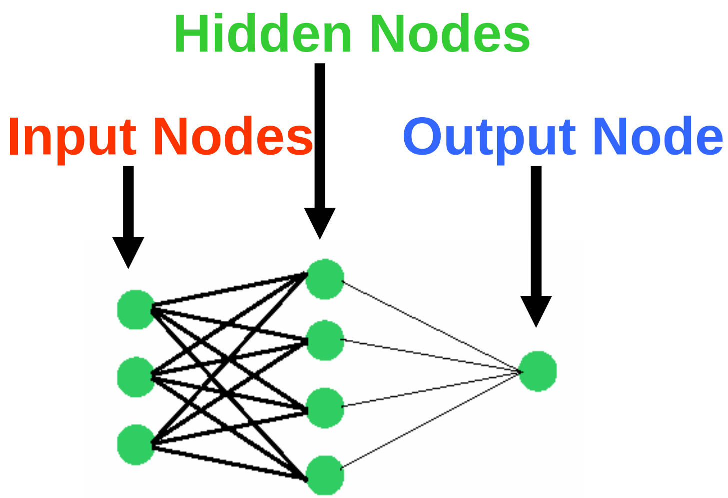


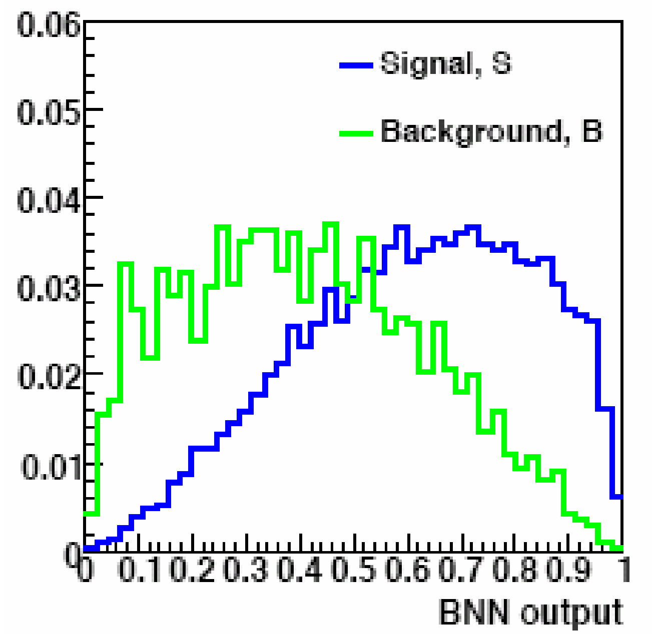
Bayesian Neural Networks Method

Neural networks use many input variables, train on signal and background samples, produce one output discriminant

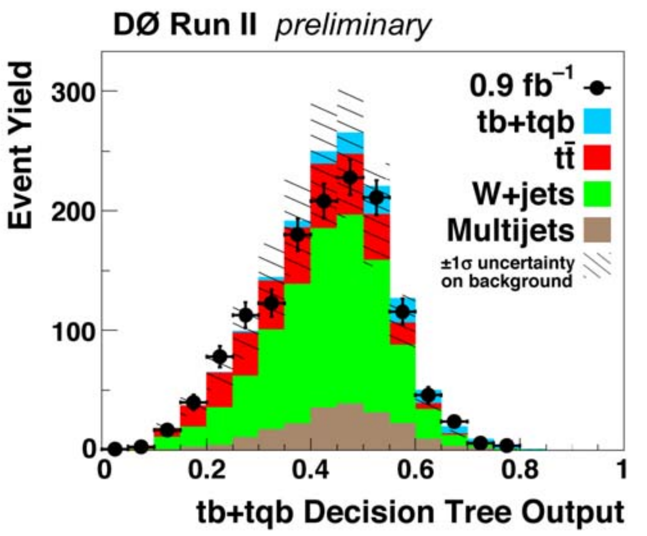
Bayesian neural networks improve on this technique:

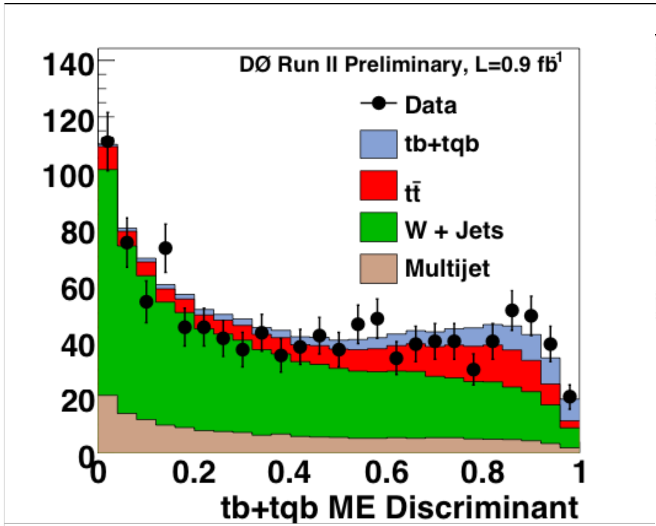
- Average over many networks weighted by the probability of each network given the training samples
- Less prone to over-training
- Network structure is less important
    - Can use larger numbers of variables and hidden nodes
    - 20 input variables (subset of DT)
    - 20 hidden nodes

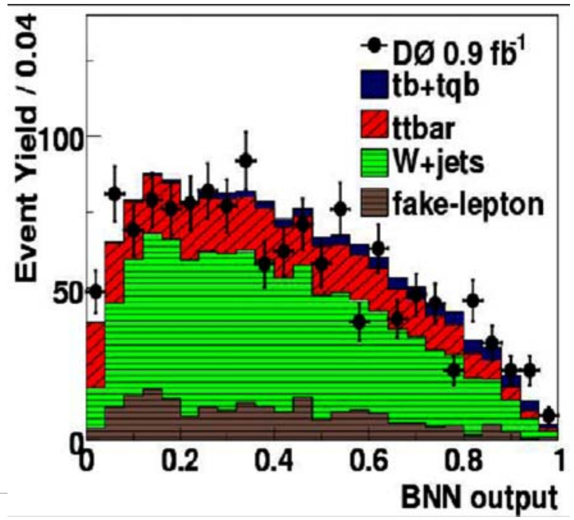


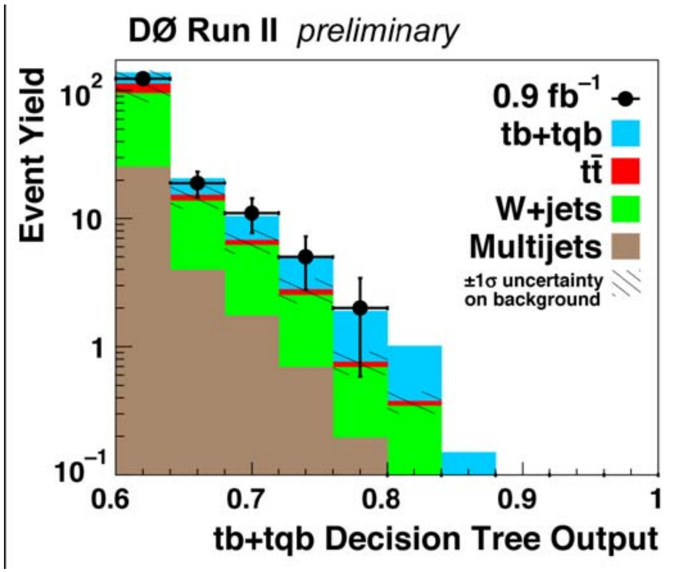


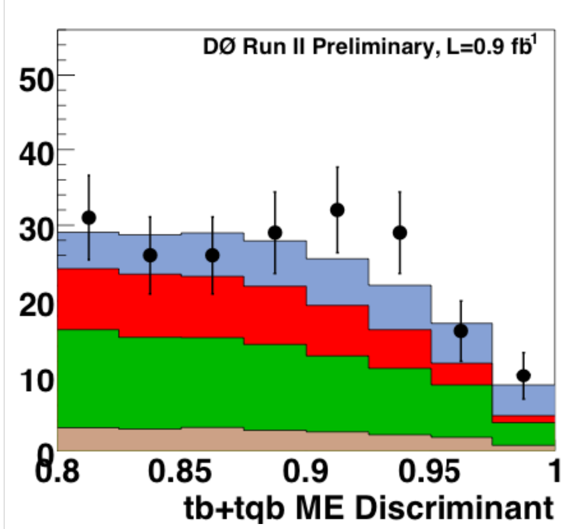
Discriminant Output

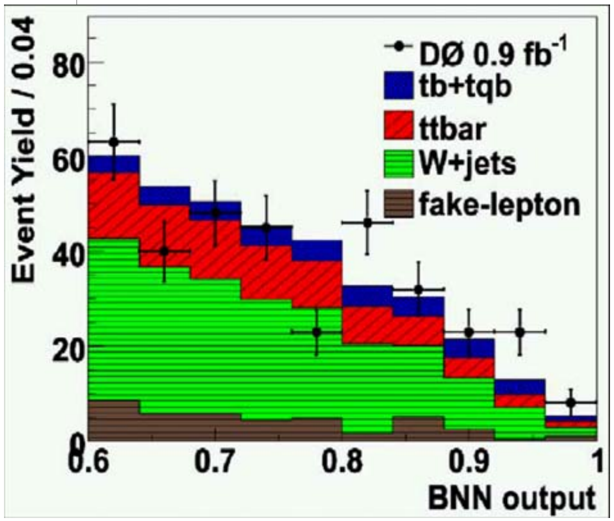




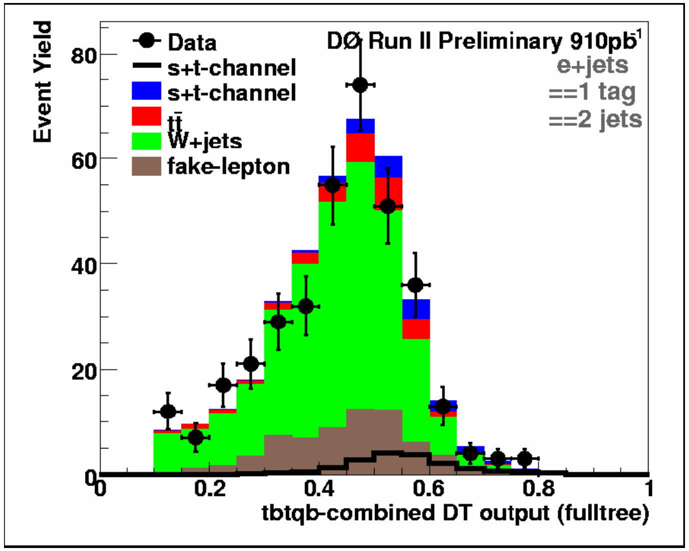


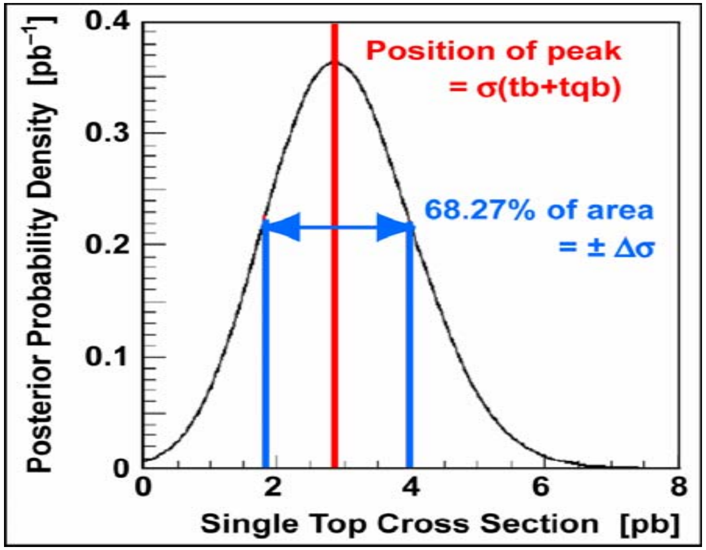






Cross Section Measurement





Binned likelihood from discriminant distribution

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

- Flat prior for the cross section
- Systematic uncertainties are treated as nuisance parameters

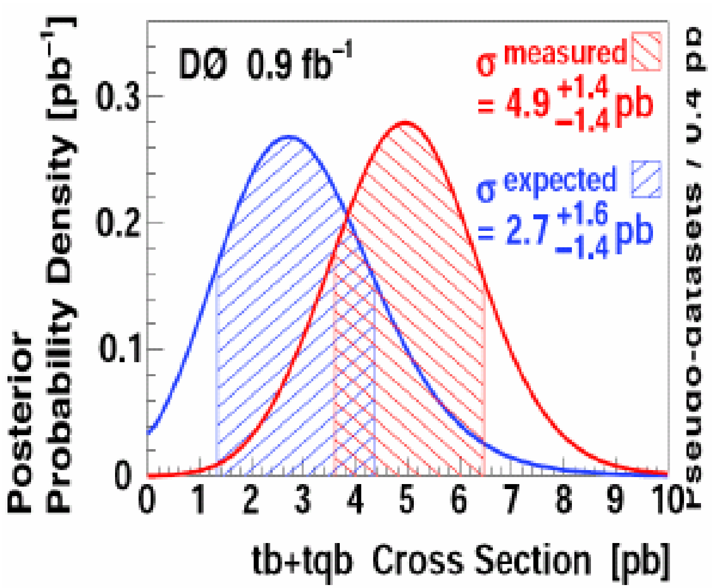
Significance as in "excess in data over background"

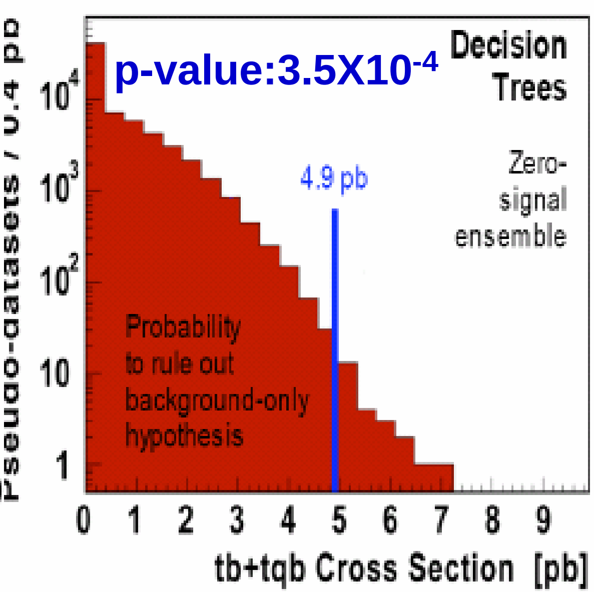
- P-value: assuming a null hypothesis, what's the probability to get a value equal to or greater than the value observed
- We measure the fraction of zero-signal datasets in which we derive at least the SM cross section (expected significance), or at least the observed cross section (observed significance)

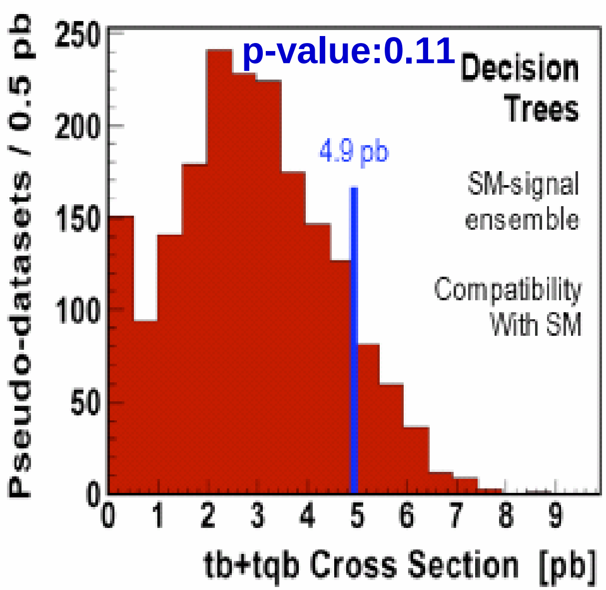
Expected and Observed Results

	Bayesian NN		Matrix Element		Decision Trees	
	Exp.	Obs.	Exp.	Obs.	Exp.	Obs.
σ(tb+tqb)[pb]	2.7 ± 1.4	$4.4^{+1.6}_{-1.4}$	$2.8^{+1.6}_{-1.4}$	$4.8^{+1.6}_{-1.4}$	$2.7^{+1.5}_{-1.4}$	4.9 ± 1.4
Significance	2.2σ	3.1σ	2.1σ	3.2σ	2.1σ	3.4σ

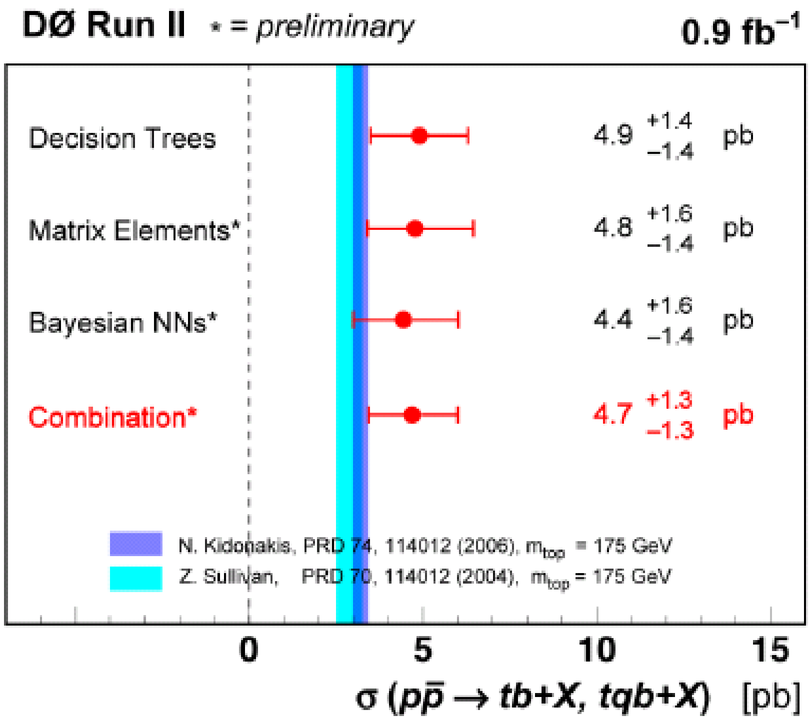
All three analyses show > 3.0 σ excess,
Evidence for single top quark production!
SM compatibility is 11%

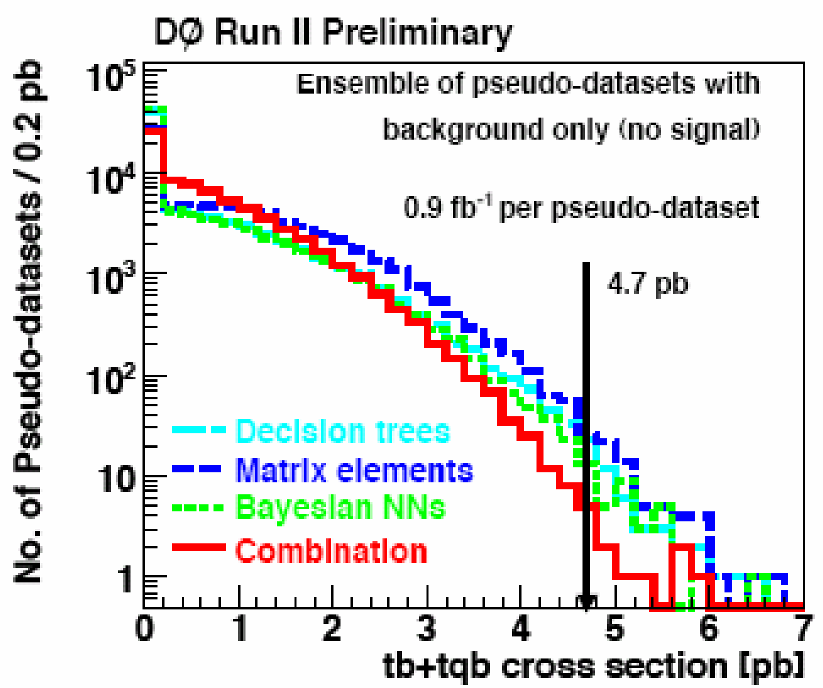






Combination for All Three Analyses





Three analyses give consistent results

- **Using same data set, thus highly correlated**

Combined result using BLUE method

- **4.7±1.3 pb with 3.6 σ significance!**

Direct measurement of $|V_{tb}|$

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

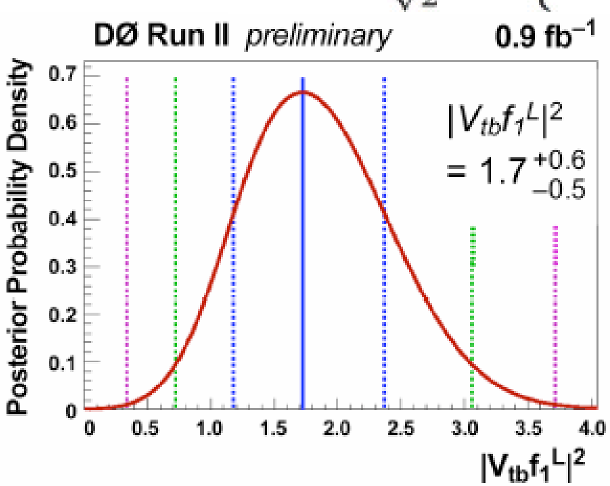
- **Calculate posterior in $|V_{tb}|^2$: $\sigma \propto |V_{tb}|^2$**

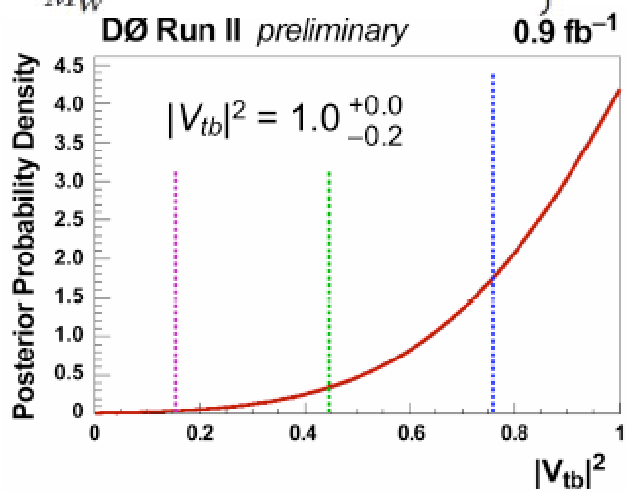
Assuming standard model production:

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

Measurement does not assume 3 generations or unitarity

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





$|V_{tb}\ f_1^L| = 1.3 \pm 0.2$ **$0.68 < |V_{tb}| < 1$ at 95%CL (assuming $f_1^L = 1$)**

Summary

Evidence for single top production found at D0!

σ (s+t) = 4.9 $\pm$ 1.4 pb with 3.4 σ significance!

Analysis published: PRL 98, 181802

Analysis webpage

First direct measurement of $|V_{tb}|$

$|V_{tb}|$>0.68 @ 95%C.L. (assuming $f_1^L = 1$)

Latest combined result: DT + ME + BNN

σ (s+t) = 4.7 $\pm$ 1.3 pb with 3.6 σ significance!

Outlook

Collected more than twice data used for this analysis

- Data delivered: >3fb^{-1} (goal of RunII is 4-9 fb^{-1})

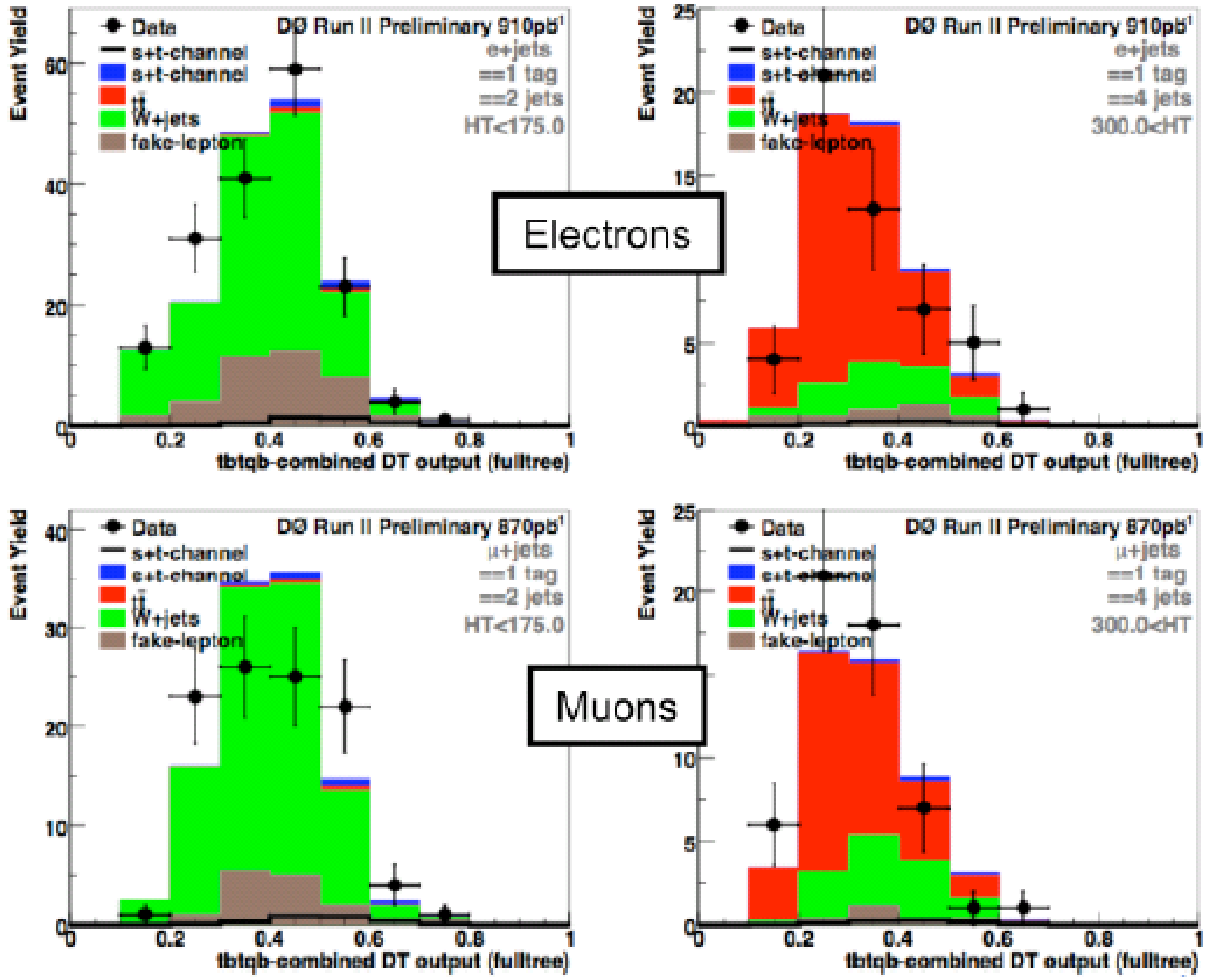
Expand to searches of new phenomena

- H^+ , anomalous Wtb coupling and more…

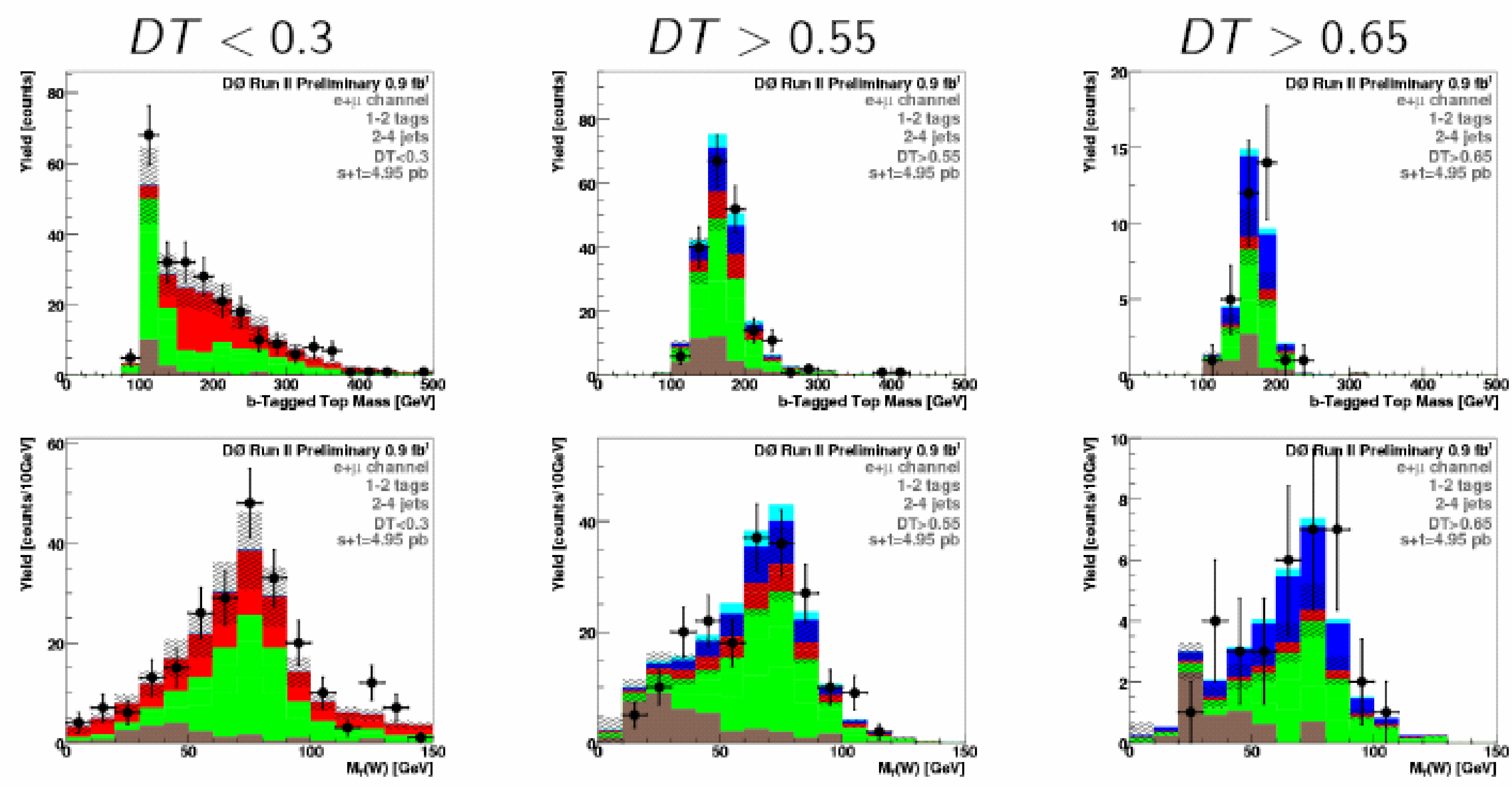
Backup Slides

Decision Tree: Cross-check Samples

"W+jets" | "$t\bar{t}$"



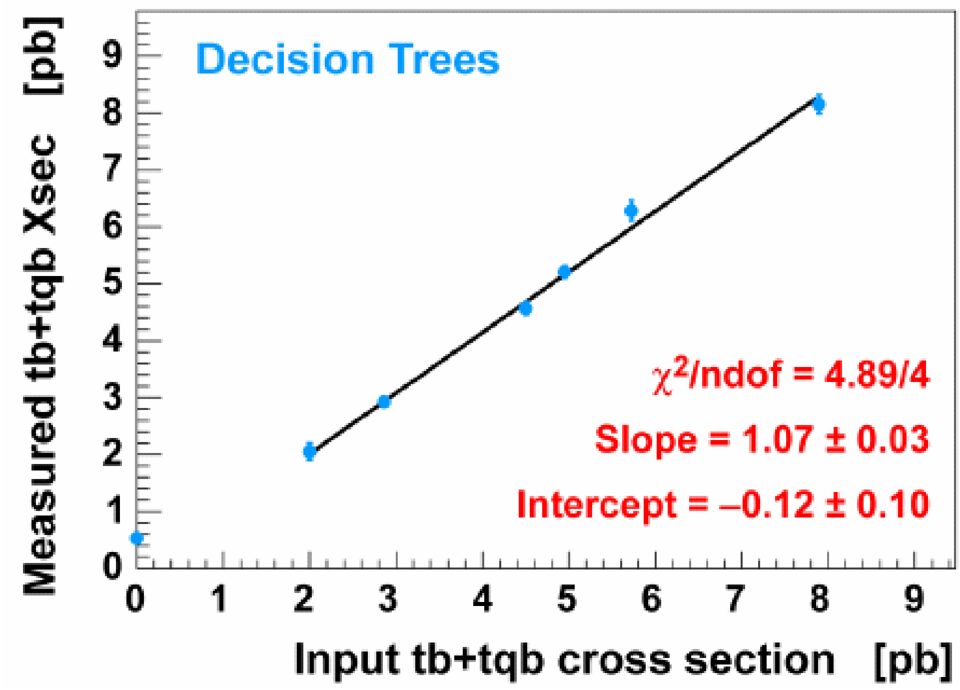
Decision Tree: Event Characteristics



Decision Tree: Ensemble Testing

Tested our machinery with many sets of pseudo-data

- Subset of our total pool of background events
- Individual statistical and systematical fluctuations
- Wonderful tool – like running D0 1000s of times!
- Generated several ensembles with different single top content
- Compare measured cross sections to input ones
- Linear response, negligible bias



NN b-jet Tagger

NN trained on 7 input variables from SVT, JLIP and CSIP taggers

Much improved performance!

- Fake rate reduced by 1/3 for same b-efficiency relative to previous tagger
- Smaller systematic uncertainty

Tag Rate Functions (TRFs) in η, P_T and z-PV derived in data are applied to MC

Our operating point:

- b-jet efficiency: ~50%
- c-jet efficiency: ~10%
- Light-jet efficiency: ~0.5%