Evidence for Single Top Using Bayesian Neural Networks (BNN) at DØ

Monica Pangilinan
Brown University
on behalf of

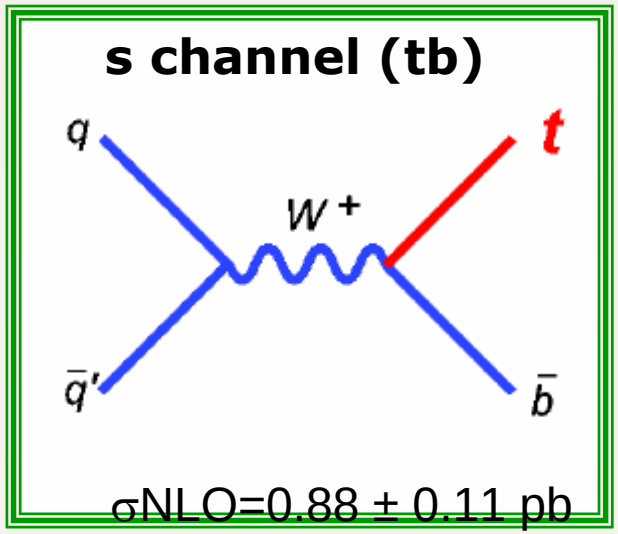
APS April Meeting 2007

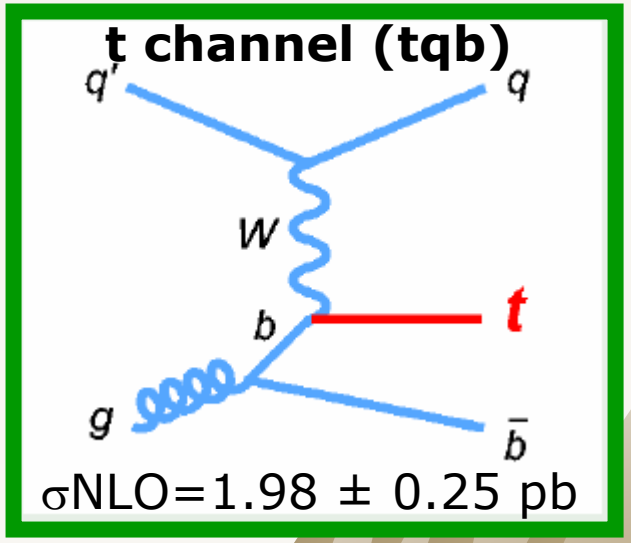
Outline

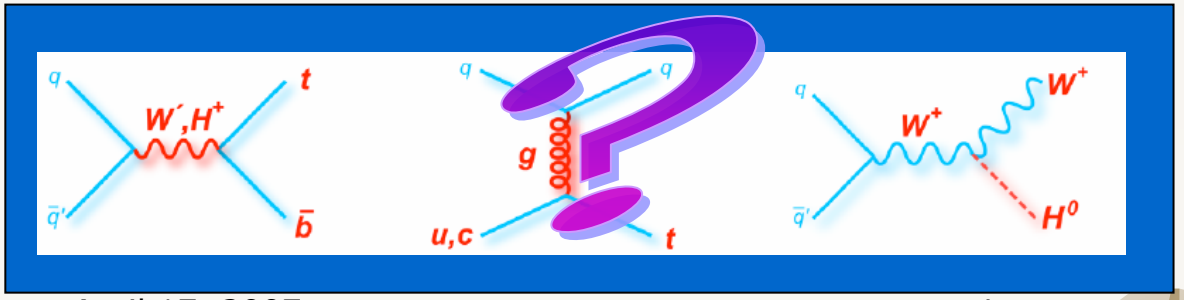
- Motivation
- Event Selection
- BNN
  - Description
  - Discriminant Output
  - Results
- Conclusion/Outlook

Motivation

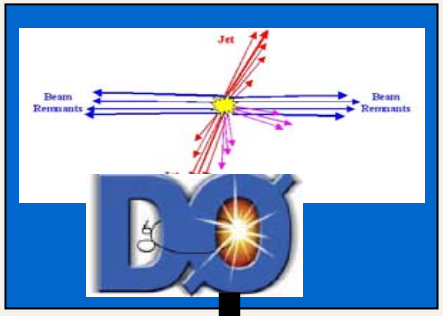
- investigate Wtb coupling
  - direct measurement of $|V_{tb}|$
  - checks unitarity of CKM matrix
  - anomalous Wtb couplings
- study top quark properties
  - polarization, lifetime, decay width, mass
- window to new physics
  - s channel: charged Higgs, heavy W'
  - t channel: flavor changing neutral currents
  - 4th quark generation?



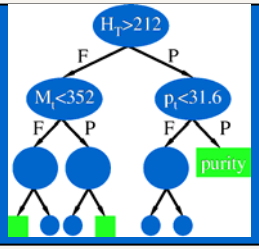




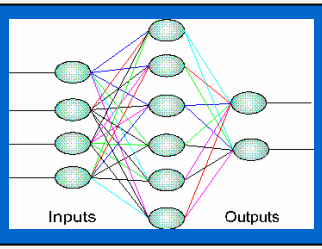
Analysis Chain



- Event Selection
- Sample divided based on lepton, b-tagged, number of jets
- Data/Monte Carlo (MC) Comparisons



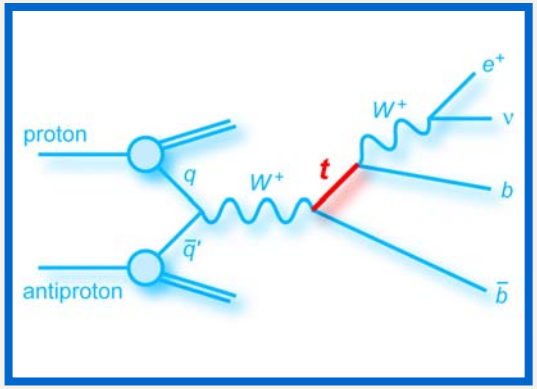
Multivariate Analysis



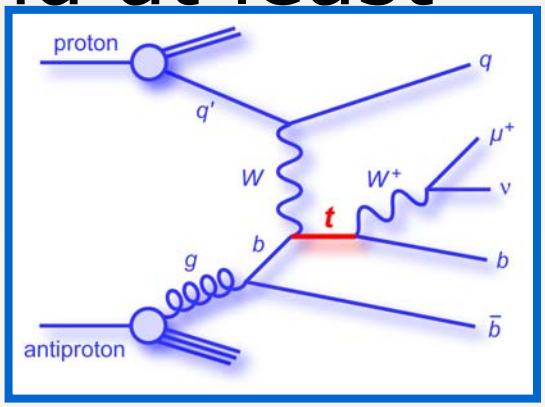
Cross Section by binned Likelihood

Event Selection

- 1 lepton
  - Electron: pT > 15GeV, |eta| < 1.1
  - Muon: pT > 18GeV, |eta| < 2.0



- Missing Transverse Energy
  - $\not{E}_T$ > 15 GeV
- At least one b-tagged jet and at least one more jet
  - 2–4 jets with p_T > 15 GeV, $|\eta|$ < 3.4
  - Leading jet p_T > 25 GeV, $|\eta|$ < 2.5
  - Second leading jet p_T > 20 GeV



- Efficiency of b-tagging greatly improved by using a neural network tagger

Why divide?

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

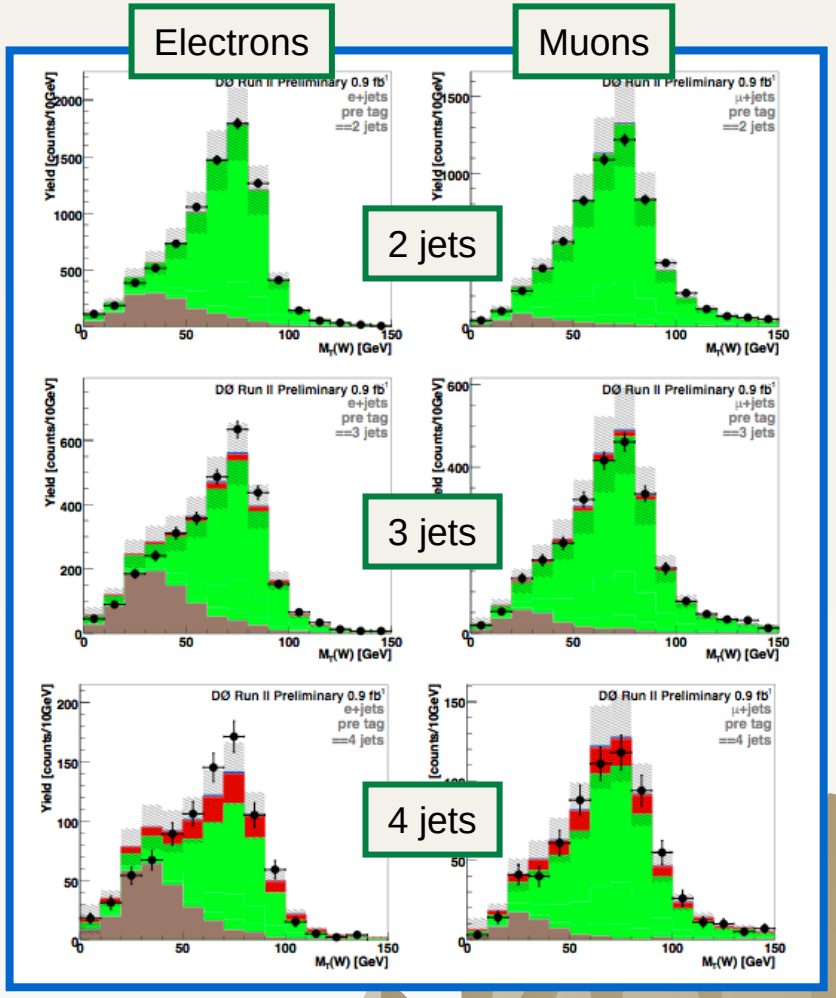
Electron + Muon	1 jet	2 jets	3 jets	4 jets	≥ 5 jets
0 tags	10% 1 : 3,200	25% 1 : 390	12% 1 : 300	3% 1 : 270	1% 1 : 230
1 tag	6% 1 : 100	21% 1 : 20	11% 1 : 25	3% 1 : 40	1% 1 : 53
2 tags		3% 1 : 11	2% 1 : 15	1% 1 : 38	0% 1 : 43

Data vs. Monte Carlo

Normalize MC samples to match data yields before b-tagging

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
$t\bar{t} \to ll$	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

W Transverse Mass



Make sure that variables used by our analysis methods are well modeled

BNN Overview

- BNN:

  $n(x) = \int n(x,w)p(w|T)\, dw$

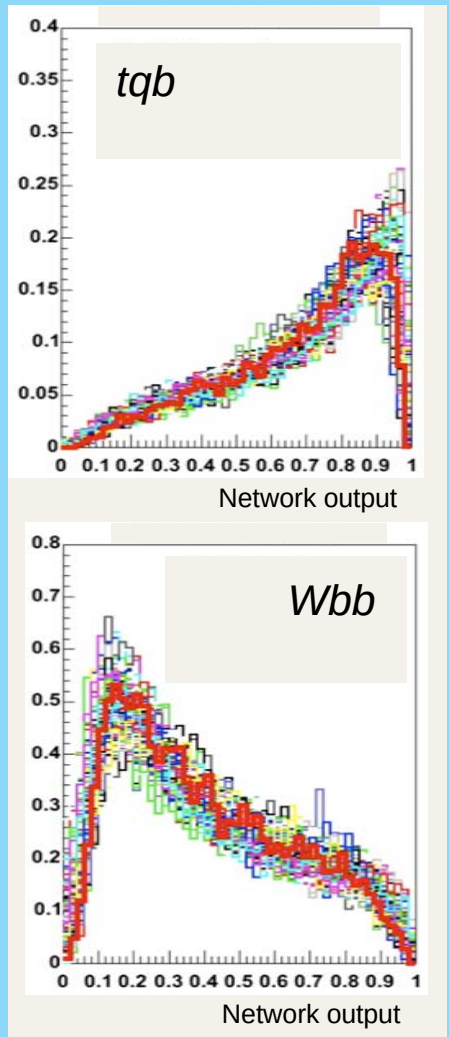
  where $n(x,w)$ is a specific neural network defined by parameters w

  $p(w|T)$ is a posterior probability density over $\{w\}$

- A NN is a single point w_0 in $\{w\}$
- Provides a more stable discriminant less prone to overtraining

BUT

how is one supposed to compute that integral?



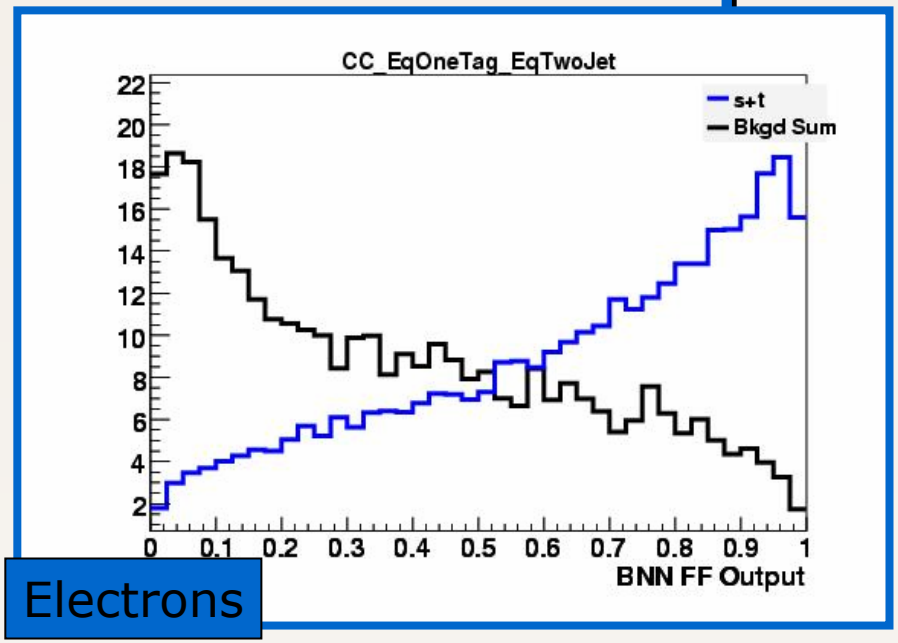
Statistical Physics?!

Markov Chain Monte Carlo

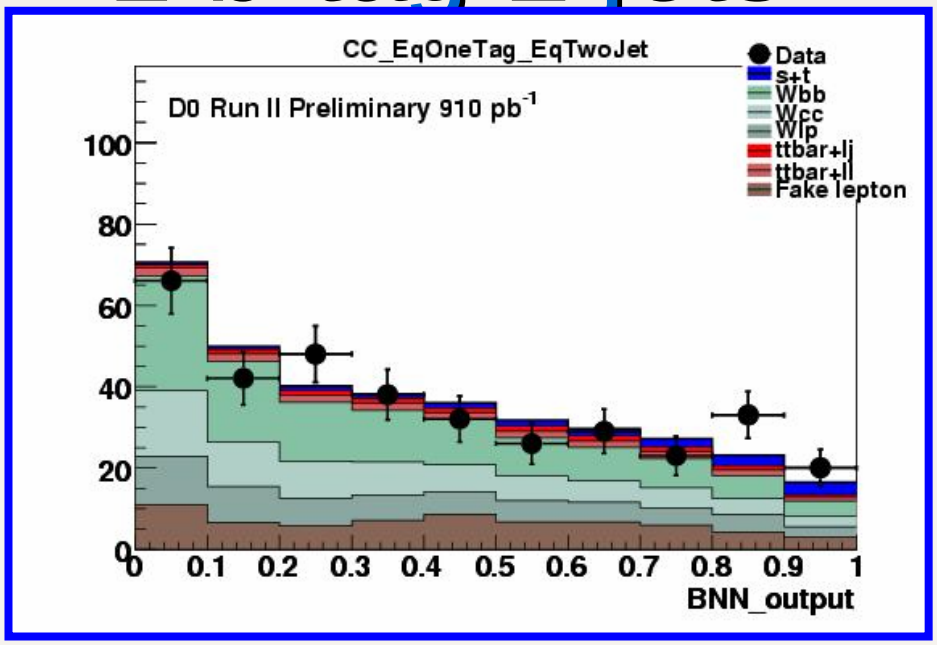
- Want to step through the {w} such that points are visited with a probability given by p(w|t,x)
- Recast the problem as a particle moving through a potential V(q) where p(w|T)=exp[-V(q)]
  - Add a kinetic energy component to V(q) so by Hamiltonian dynamics, will visit every place in w based on density exp[-(V(q)+K.E)]
  - Inject or remove energy with a certain probability -> move to different constant energy states
  - Keep only q values (which correspond to NN)

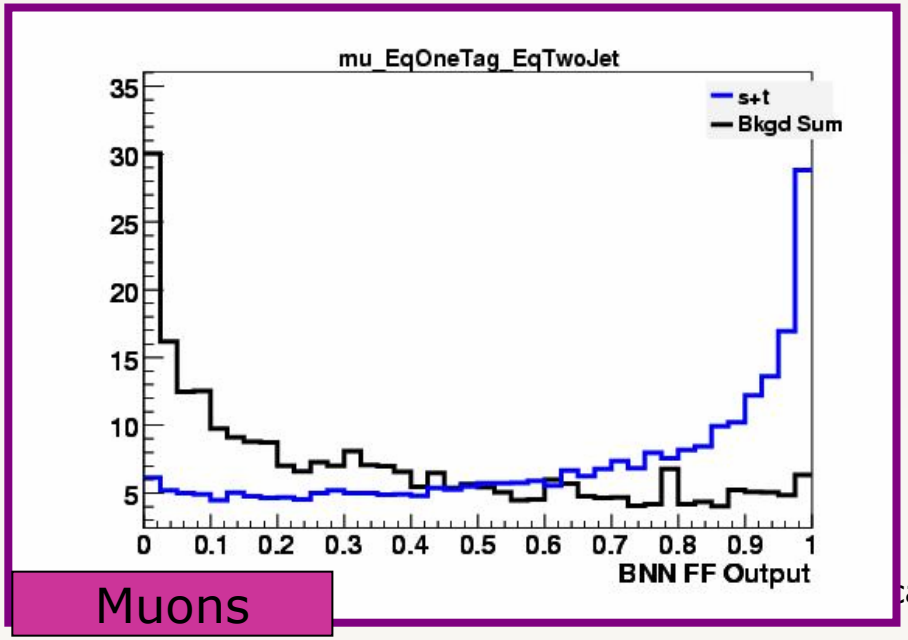
$$n(x) = \int n(x,w)p(w|T)\,dw \approx \frac{1}{N_{\text{net}}} \sum_{i=1}^{N_{\text{net}}} n(x,w_i)$$

BNN Output – 1 b-tag 2 jets

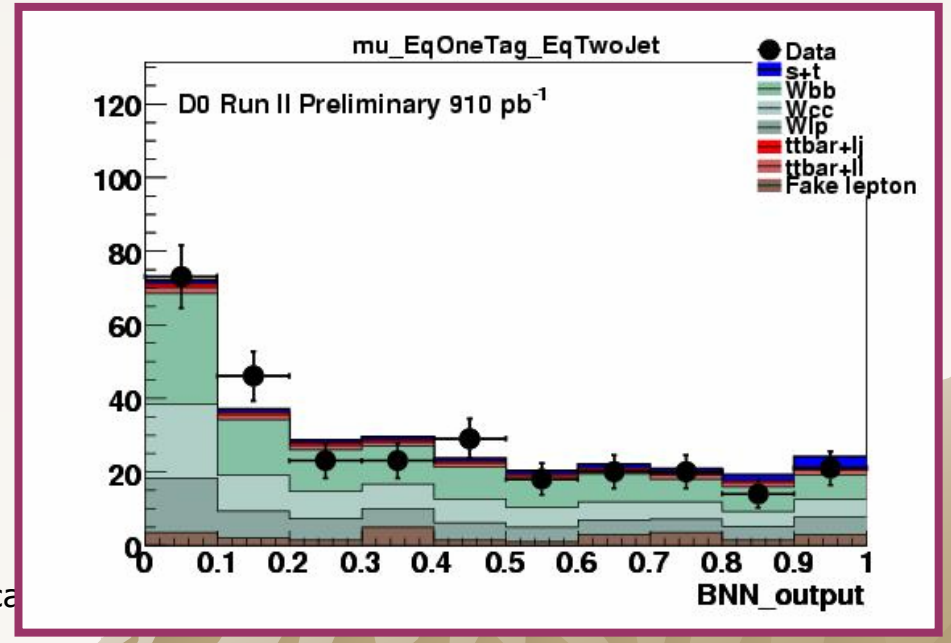


Electrons

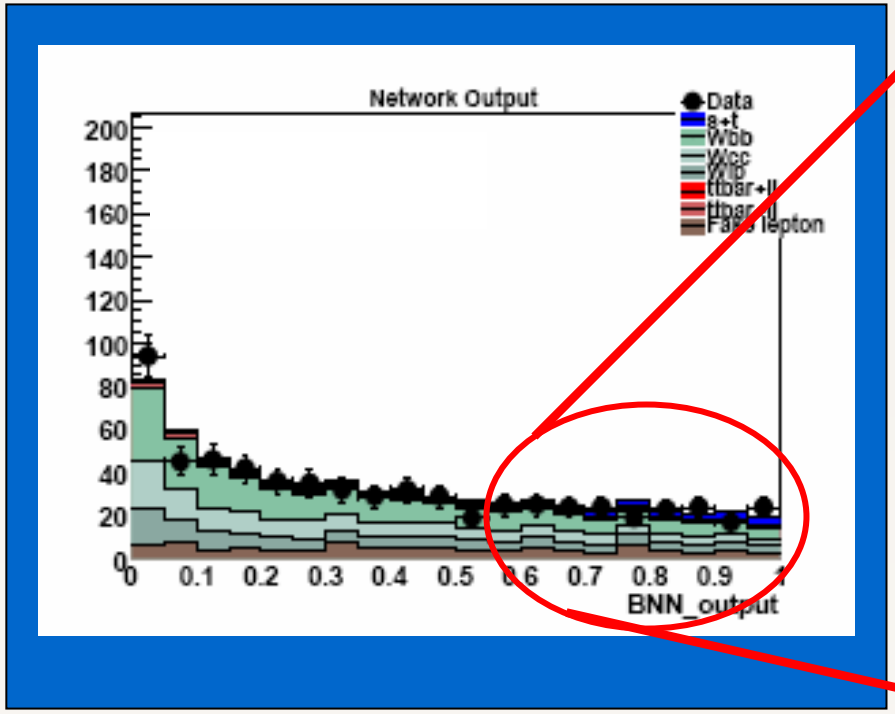


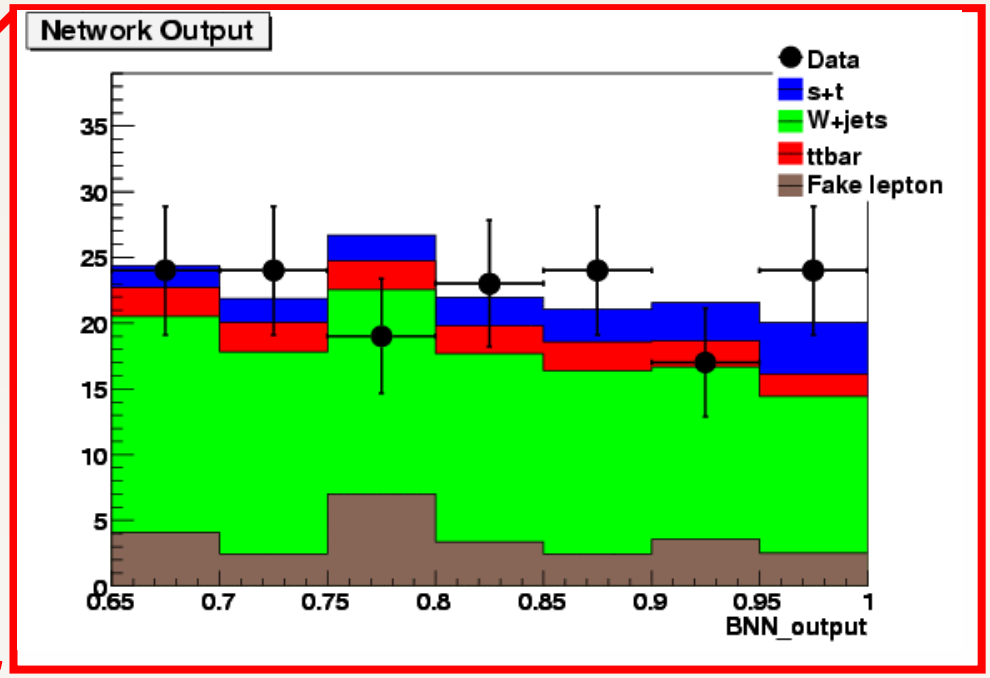


Muons



BNN Output – 1 b-tag 2 jets



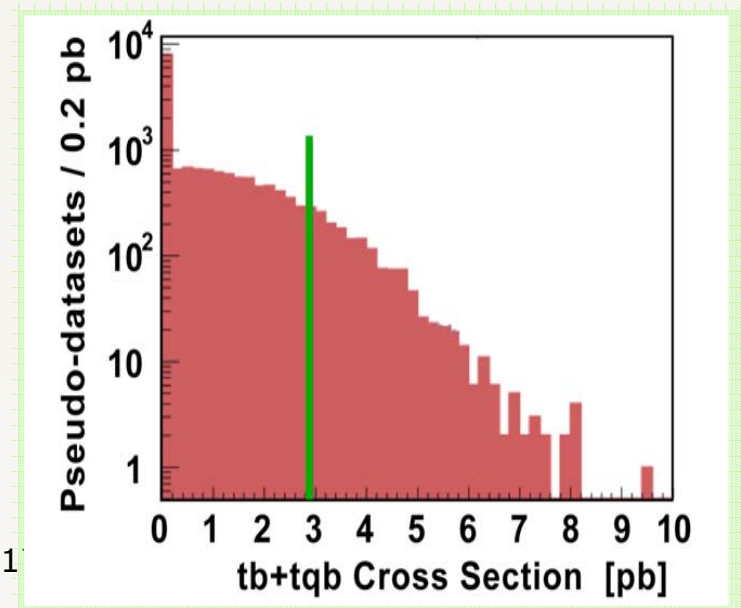


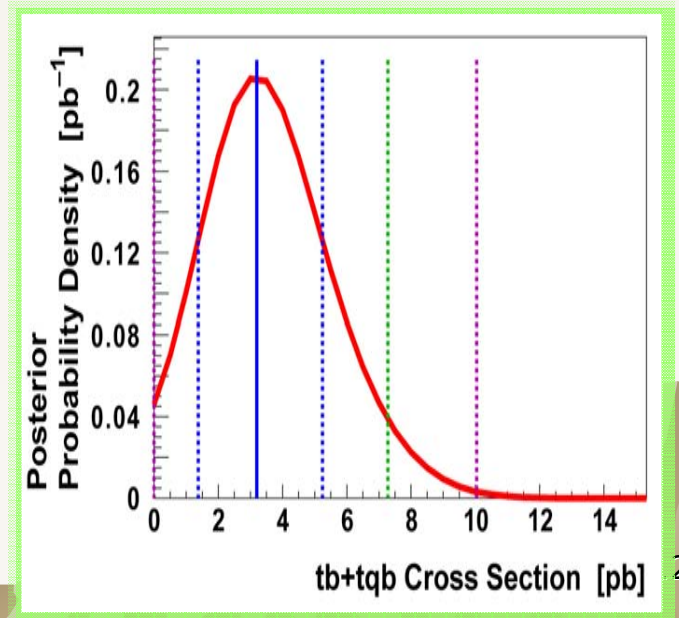
Electron and Muon channel

Expected Results

- Expected Cross Section:
- Expected p-value:
  - Based on generated DØ pseudo experiments, how many cross sections are above SM value
- Expected Significance:

  - Assuming a gaussian distribution, convert expected p-value to significance

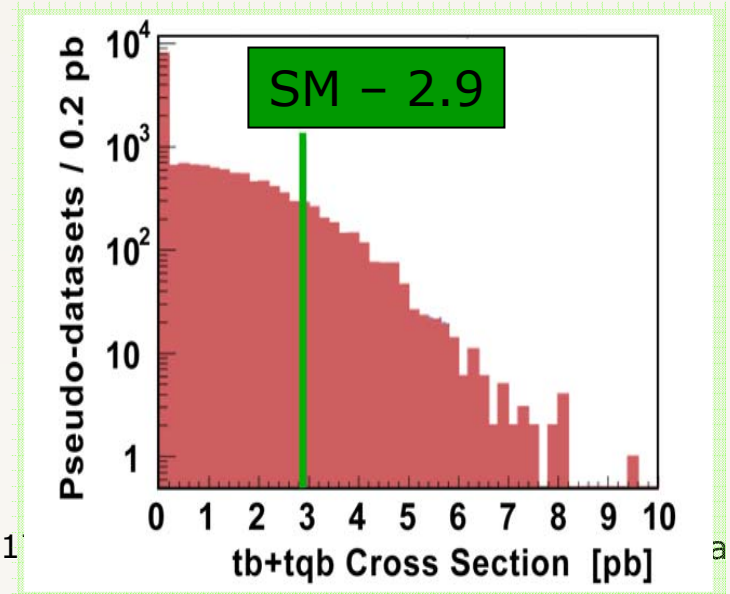


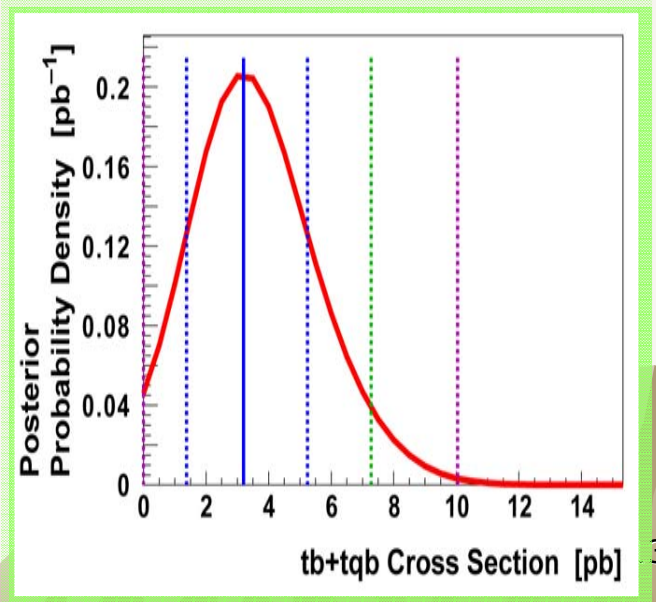


Expected Results

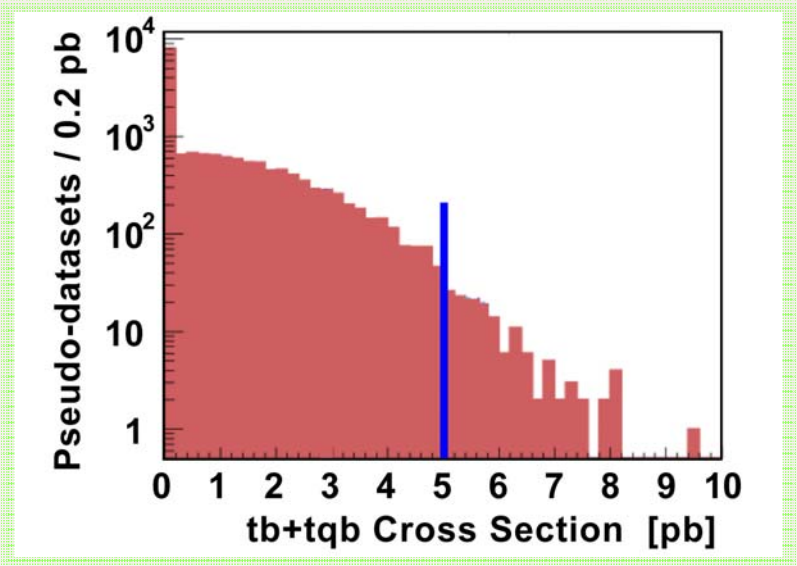
- Expected Cross Section: $3.2\ ^{+2.0}_{-1.9}$ pb
- Expected p-value: 9.7%
  - Based on generated DØ pseudo experiments, how many cross sections are above SM value
- Expected Significance: 1.3 σ

  - Assuming a gaussian distribution, convert expected p-value to significance

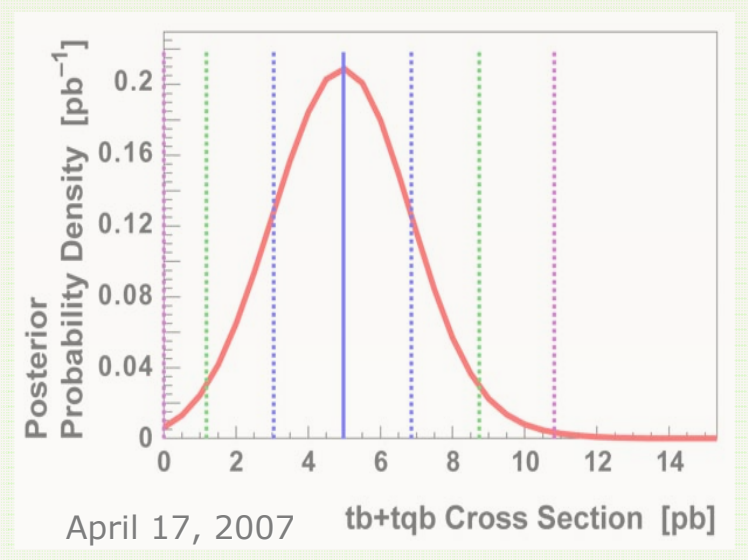


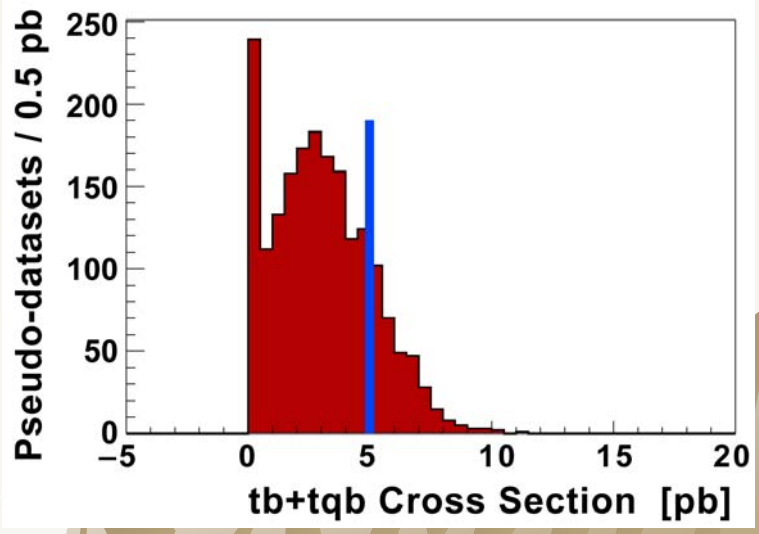


Measured Results

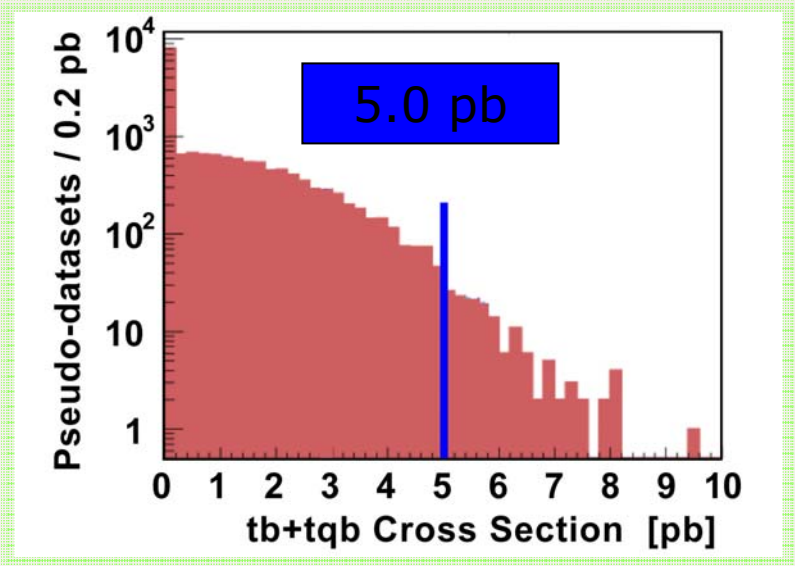


- Measured Cross Section:
- Measured p-value:
  - Based on generated DØ pseudo experiments, how many cross sections are above our measured value
- Measured Significance:
- Compatibility with SM:
  - Using SM ensembles, how many times do we measure a cross section above our measured cross section?

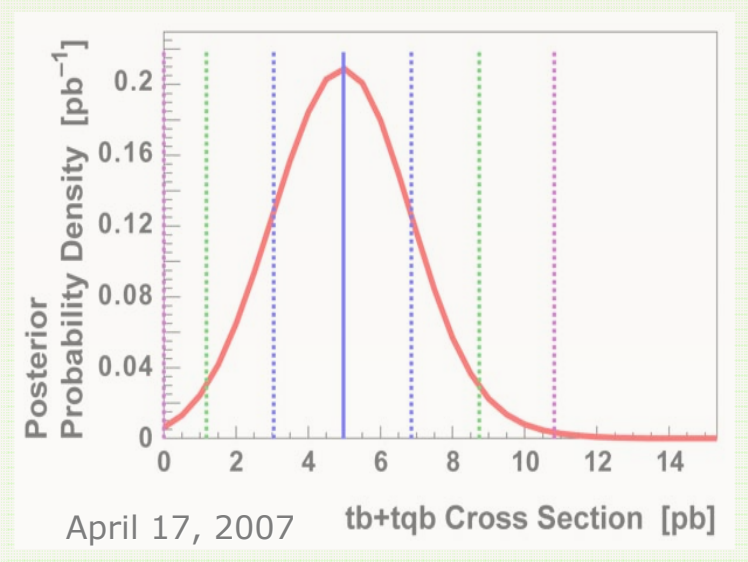


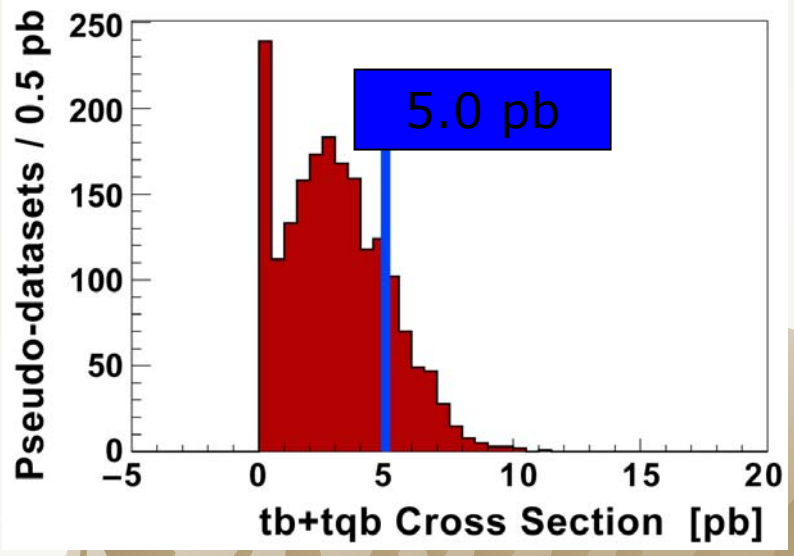


Measured Results



- Measured Cross Section: 5.0±1.9 pb
- Measured p-value: .89%
  - Based on generated D0 pseudo experiments, how many cross sections are above our measured value
- Measured Significance: 2.4 σ
- Compatibility with SM: 18%
  - Using SM ensembles, how many times do we measure a cross section above our measured cross section?





Conclusion/Outlook

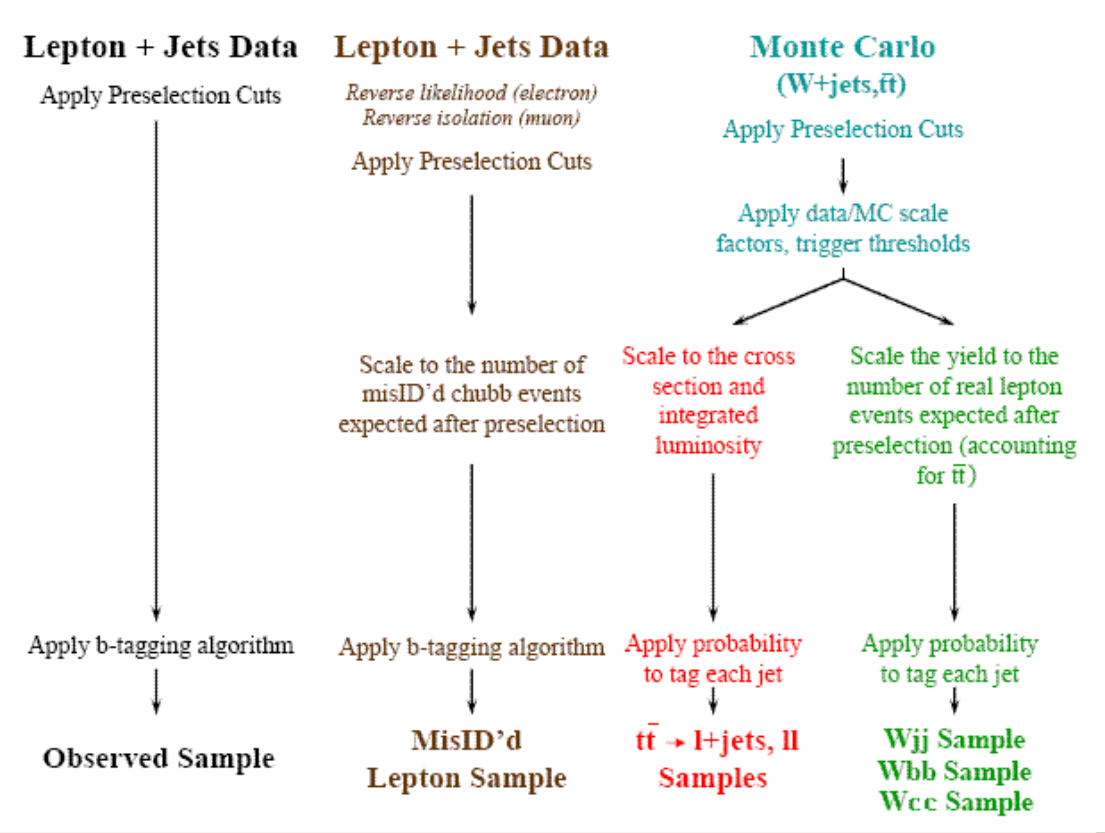
- BNN measures
  - Cross section: **5.0 ± 1.9pb**
  - a significance of **2.4σ**
  - Compatibility with SM: **18%**
- Planning to incorporate twice as much data
- Single top will be "observed" soon!

http://www.deviantart.com/deviation/31400428/

Backup Slides

Data vs Monte Carlo (II)

- Singletop:
  - Used CompHep
- W+jets
  - Algen generated
  - Wbb and Wcc scale factor obtained using data
- ttbar
  - Alpgen generated
  - Normalized to NNLO cross section of 6.8pb
- Fake jets
  - Generated from data



B-tagging

- uses vertexing information as variables in the NN
- can use different operating points
- For our operating point,
  - b jet efficiency ~50%
  - c jet efficiency ~10%
  - Light jet efficiency ~ .5%

Variables for Training

Bayesian NN Analysis Variables
Object Kinematics
p_T(jet1)
p_T(best1)
p_T(tag1)
Event Kinematics
Aplanarity(alljets,W)
M(W,best1) ("best" top mass)
M(W,tag1) ("b-tagged" top mass)
M_T(W)
H_T(alljets)
H_T(alljets,W)
p_T(alljets−best1)
H_T(jet1,jet2)
H_T(jet1,jet2,W)
M(alljets−best1)
M(alljets−tag1)
p_T(alljets−best1)
M_T(jet1,jet2)
M_T(W)
Angular Variables
cos(best1,notbest1)$_{\text{besttop}}$
cos(tag1,lepton)$_{\text{btaggedtop}}$
cos(lepton,besttopframe)$_{\text{besttopCMframe}}$
cos(lepton,btaggedtopframe)$_{\text{btaggedtopCMframe}}$
cos(lepton,Q(lepton)$\times z$)$_{\text{besttop}}$
cos(notbest,alljets)$_{\text{alljets}}$
Q(lepton) $\times\eta$(untagged jet)

Expected Cross Section

Expected Single Top Cross Sections [pb]								
	1,2tags + 2,3,4jets		e,μ + 2,3,4jets		e,μ + 1,2tags			All
	e-chan	μ-chan	1 tag	2 tags	2 jets	3 jets	4 jets	channels
Statistics only								
tb	$0.9^{+0.9}_{-0.9}$	$0.9^{+1.4}_{-0.9}$	$0.9^{+1.1}_{-0.9}$	$0.9^{+1.2}_{-0.9}$	$0.9^{+0.9}_{-0.8}$	$0.9^{+2.0}_{-0.9}$	$0.9^{+7.7}_{-0.9}$	$0.9^{+0.8}_{-0.8}$
tqb	$2.0^{+1.6}_{-1.5}$	$2.0^{+1.9}_{-1.8}$	$2.0^{+1.4}_{-1.3}$	$2.0^{+8.9}_{-2.0}$	$2.0^{+1.5}_{-1.4}$	$2.0^{+2.3}_{-2.0}$	$2.0^{+9.0}_{-2.0}$	$2.0^{+1.4}_{-1.3}$
$tbtqb$	$2.9^{+1.7}_{-1.6}$	$2.8^{+2.1}_{-2.0}$	$2.9^{+1.5}_{-1.4}$	$2.8^{+3.0}_{-2.8}$	$2.9^{+1.6}_{-1.5}$	$2.8^{+2.4}_{-2.3}$	$2.8^{+8.1}_{-2.8}$	$2.9^{+1.4}_{-1.4}$
With systematics								
tb	$0.6^{+1.9}_{-0.6}$	$0.0^{+3.4}_{-0.0}$	$0.9^{+2.2}_{-0.9}$	$0.6^{+2.4}_{-0.6}$	$0.9^{+1.4}_{-0.9}$	$0.0^{+6.0}_{-0.0}$	$0.0^{+18.5}_{-0.0}$	$\mathbf{1.0}^{+1.2}_{-1.0}$
tqb	$1.7^{+2.0}_{-1.7}$	$2.2^{+2.8}_{-2.2}$	$2.3^{+1.9}_{-1.6}$	$0.0^{+16.5}_{-0.0}$	$2.3^{+1.9}_{-1.7}$	$1.5^{+5.1}_{-1.5}$	$0.4^{+19.1}_{-0.4}$	$\mathbf{2.4}^{+1.8}_{-1.6}$
$tbtqb$	$2.5^{+2.4}_{-2.1}$	$2.1^{+3.4}_{-2.1}$	$3.1^{+2.2}_{-1.9}$	$2.0^{+6.7}_{-2.0}$	$3.1^{+2.2}_{-2.0}$	$2.1^{+5.4}_{-2.1}$	$1.0^{+19.3}_{-1.0}$	$\mathbf{3.2}^{+2.0}_{-1.8}$

Measured Cross Section

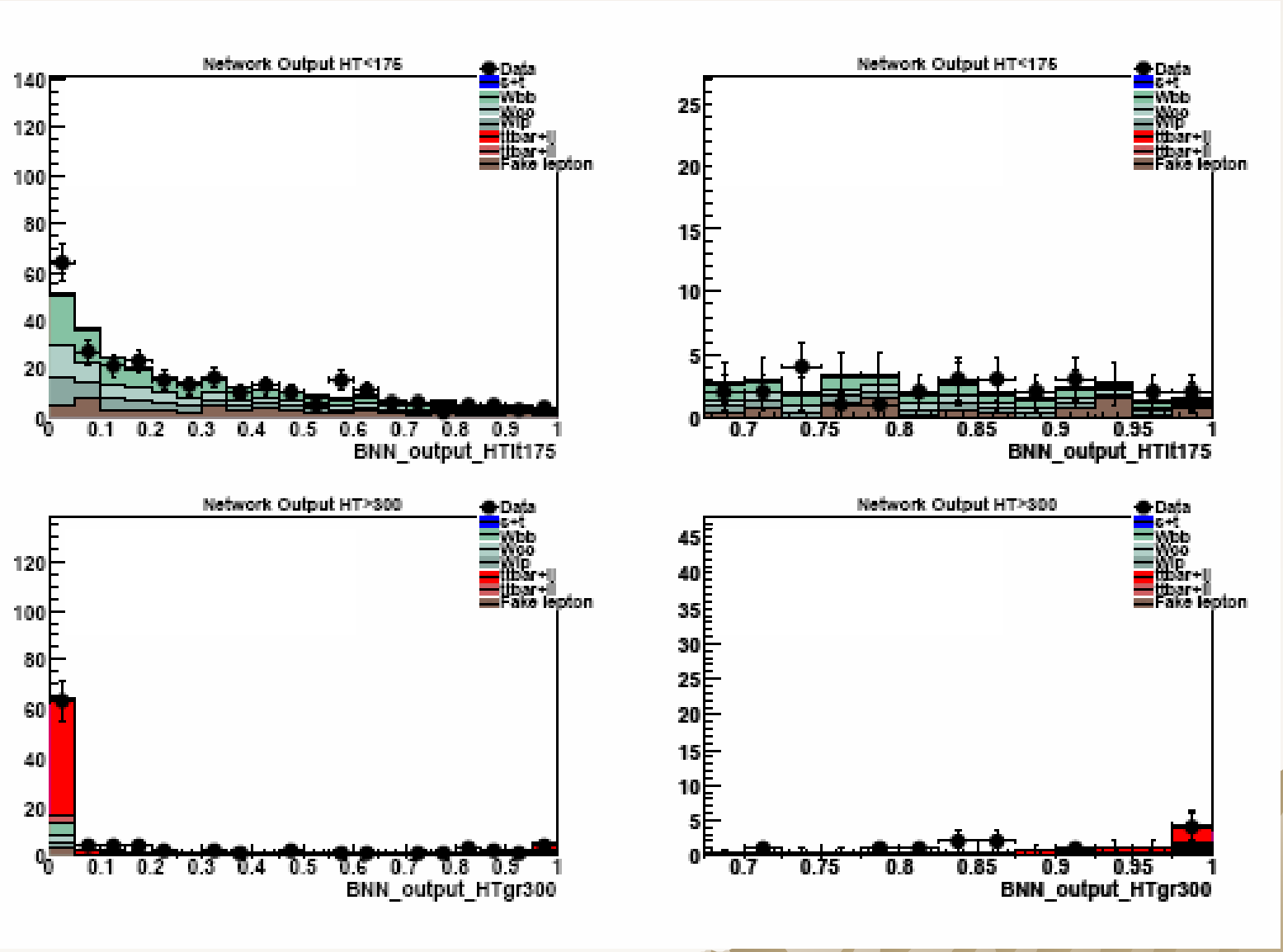
Measured Single Top Cross Sections [pb]								
	1,2tags + 2,3,4jets		e,μ + 2,3,4jets		e,μ + 1,2tags			All
	e-chan	μ-chan	1 tag	2 tags	2 jets	3 jets	4 jets	channels
Statistics only								
$t\bar{b}$	$2.1^{+1.3}_{-1.1}$	$1.4^{+1.3}_{-1.3}$	$1.0^{+1.0}_{-1.0}$	$2.8^{+1.7}_{-1.4}$	$2.2^{+1.2}_{-1.1}$	$0.7^{+2.0}_{-0.7}$	$0.0^{+7.4}_{-0.0}$	$1.9^{+1.0}_{-1.0}$
$tq\bar{b}$	$3.2^{+1.8}_{-1.6}$	$0.3^{+2.2}_{-0.3}$	$1.9^{+1.3}_{-1.2}$	$12.8^{+9.2}_{-8.6}$	$3.3^{+1.7}_{-1.6}$	$0.0^{+2.4}_{-0.0}$	$0.0^{+9.8}_{-0.0}$	$2.2^{+1.4}_{-1.3}$
$t\bar{b}tq\bar{b}$	$3.2^{+1.7}_{-1.6}$	$0.9^{+1.9}_{-0.9}$	$1.7^{+1.3}_{-1.2}$	$7.0^{+4.0}_{-3.6}$	$3.4^{+1.6}_{-1.5}$	$0.0^{+2.6}_{-0.0}$	$0.0^{+9.0}_{-0.0}$	$2.4^{+1.3}_{-1.3}$
With systematics								
$t\bar{b}$	$4.7^{+2.4}_{-2.1}$	$3.7^{+1.9}_{-2.8}$	$3.5^{+1.8}_{-2.2}$	$4.4^{+3.1}_{-2.5}$	$3.6^{+1.7}_{-1.7}$	$2.3^{+4.5}_{-2.3}$	$0.0^{+19.4}_{-0.0}$	$3.8^{+1.5}_{-1.5}$
$tq\bar{b}$	$5.3^{+2.7}_{-2.5}$	$0.6^{+3.8}_{-0.6}$	$3.8^{+1.9}_{-1.8}$	$15.9^{+17.8}_{-13.2}$	$4.4^{+2.3}_{-2.0}$	$0.0^{+4.6}_{-0.0}$	$2.6^{+18.8}_{-2.6}$	$3.7^{+2.0}_{-1.8}$
$t\bar{b}tq\bar{b}$	$6.7^{+2.9}_{-2.6}$	$2.3^{+2.7}_{-2.3}$	$4.0^{+1.9}_{-2.0}$	$13.0^{+8.9}_{-7.2}$	$5.5^{+2.2}_{-2.2}$	$0.4^{+5.2}_{-0.4}$	$1.7^{+19.7}_{-1.7}$	$5.0^{+1.9}_{-1.9}$

BNN Training

- 24 variables used
- trained with s+t channel as our signal and W+jets, ttbar combined as our background
- 40 hidden nodes
- 800 iterations of training (100 steps in equipotential x 20 reject/accept energy change = 1 iteration)
- last 100 NN were averaged

Relative Uncertainties

Contributions to the Cross Section Uncertainty	
Systematics component	
Luminosity	0.71 pb
$t\bar{t}$ cross section	0.84 pb
Matrix method	1.45 pb
Trigger	0.49 pb
Primary vertex	0.31 pb
Lepton ID	0.51 pb
Jet ID	0.17 pb
Jet fragmentation	0.65 pb
Jet energy scale	0.29 pb
Tag-rate functions	0.44 pb
Combined systematics	1.65 pb
Statistics	1.47 pb
Measured total uncertainty	1.9 pb

TABLE 7: Contribution of each systematic uncertainty to the total systematic uncertainty on the $tb+tqb$ cross section.

Cross Checks