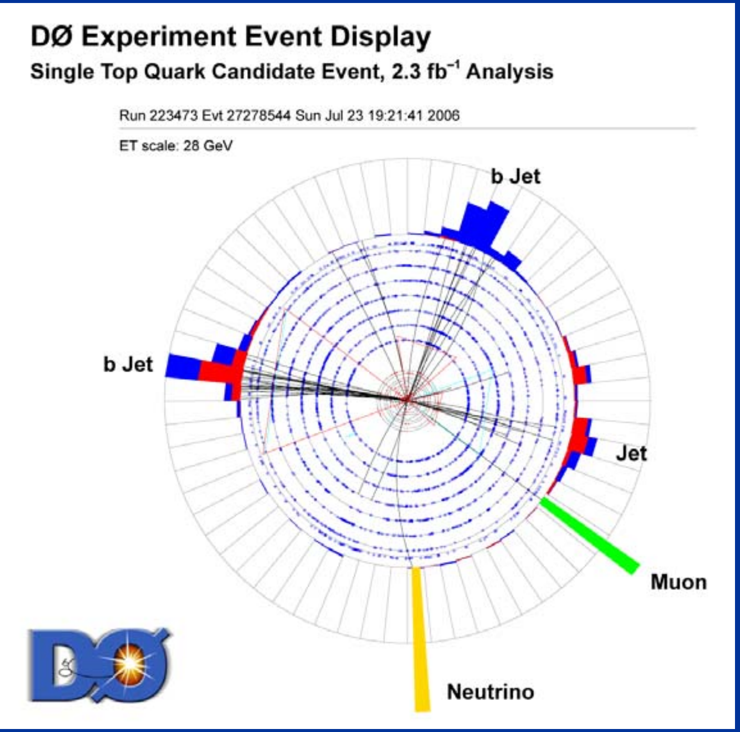
Single Top Production using Bayesian Neural Networks at DØ

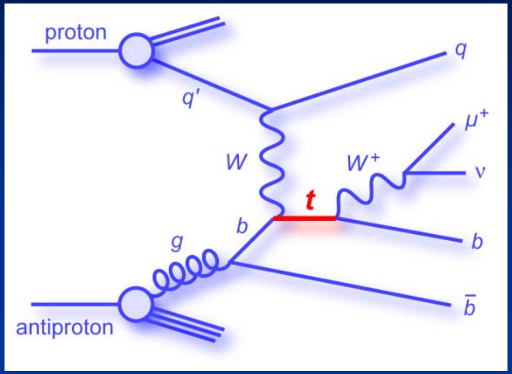


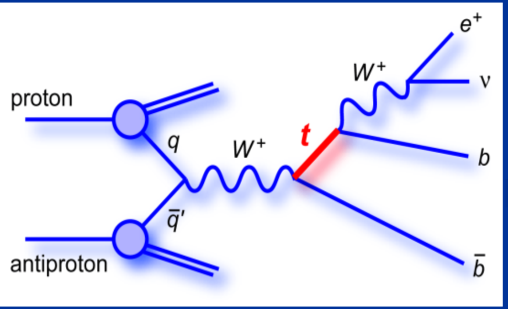
Cecilia E. Gerber
University of Illinois at Chicago

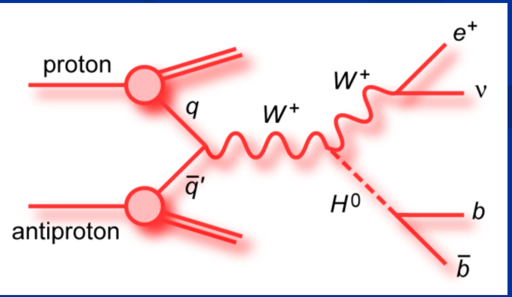
APS April meeting, May 2-5 2009

Single Top Production

- Predicted by the Standard Model, and observed for the first time in May 2009, 14 years after the observation of the top quark pair production
- Probe of the Wtb interaction with no assumption on the number of quark families or unitarity of the CKM matrix
- Cross sections sensitive to beyond-the-SM processes
  - s-channel:
    - Resonances: heavy W' boson, charged Higgs boson, Kaluza-Klein excited W_{KK}, technipion, etc.
  - t-channel
    - flavor-changing neutral currents
  - Fourth generation of quarks
- Same final state as WH
  - Same backgrounds
  - Test techniques to extract small signal

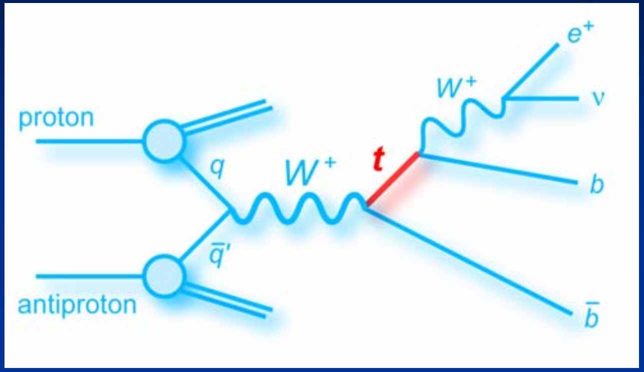






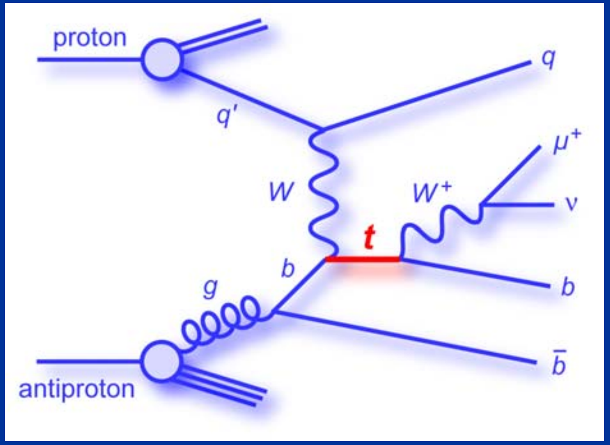
Experimentally Very Challenging

s-channel ("tb")



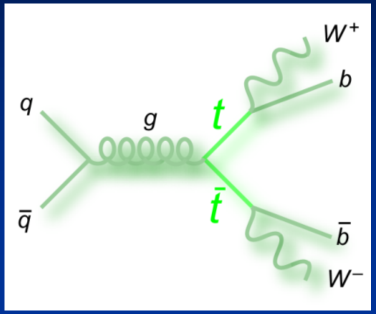
$\sigma_{SM} = 1.12 \pm 0.05$ pb

t-channel ("tqb")

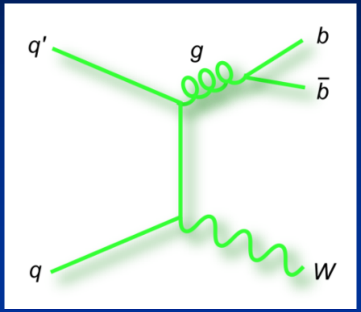


$\sigma_{SM} = 2.34 \pm 0.13$ pb

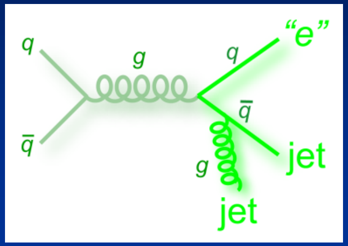
Top pairs



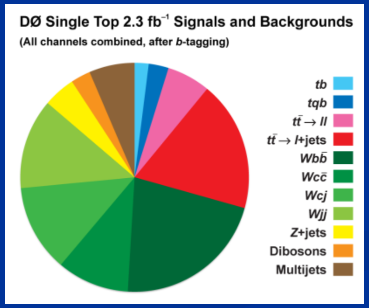
W+jets



Multijets



- Event Selection (24 channels)
  - One high-p_T isolated electron or muon
  - Large missing transverse energy
  - A b-jet from the top quark decay
  - A second b-jet or a light jet

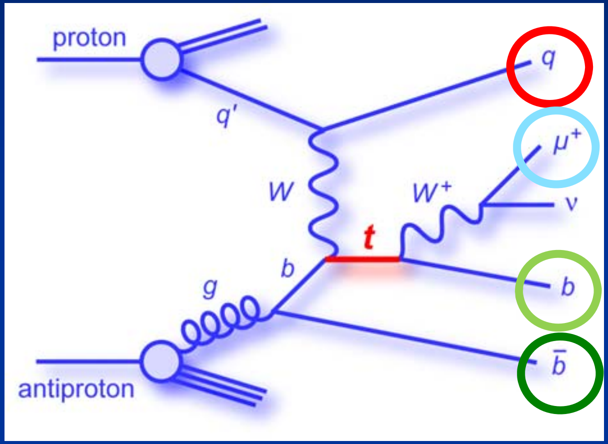


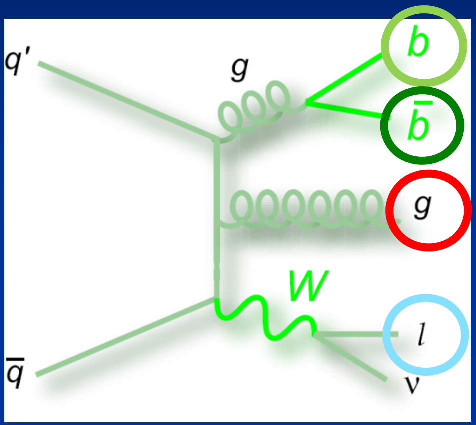
S:B = 1:21 in 1Tag
S:B = 1:15 in 2Tag

Single top cross sections: Kidonakis and Vogt, PRD 68, 114014 (2003) for m_t =170 GeV

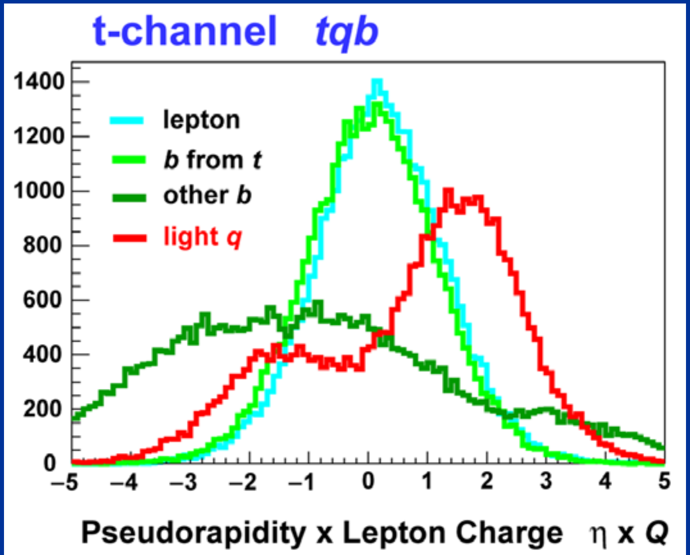
Multivariate Analysis

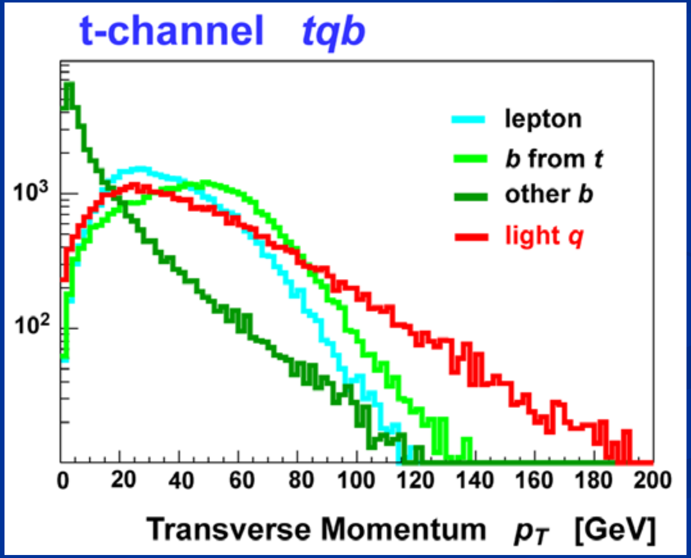
- Exploit kinematic differences between signal and background





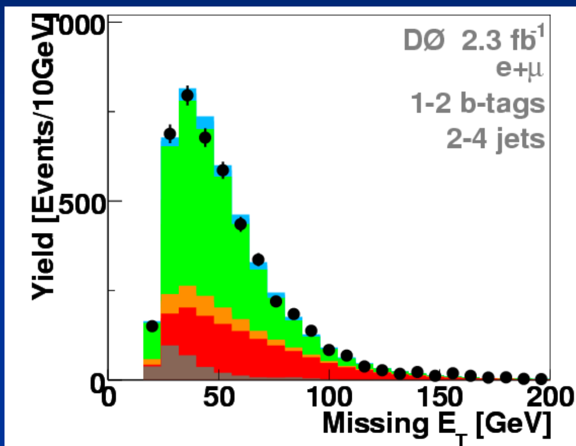
Even though final state is identical, MVA can extract the signal due to characteristic shape of variables with high discriminating power



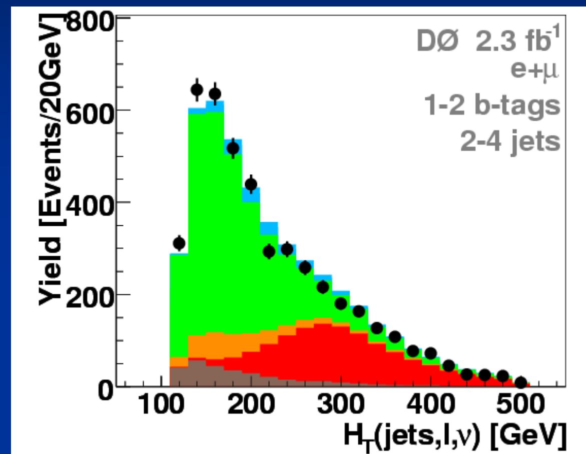


Discriminating Variables

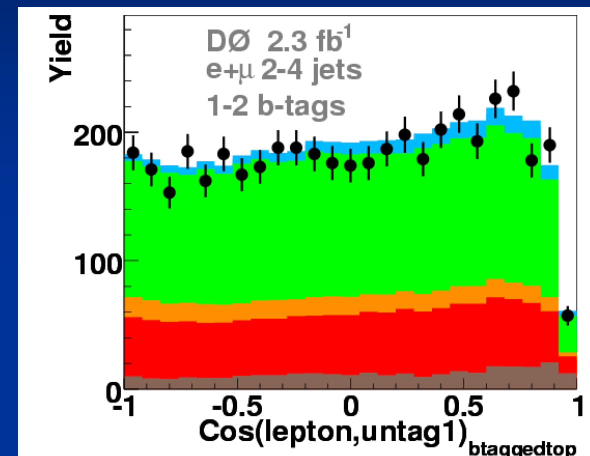
OBJECT KINEMATICS



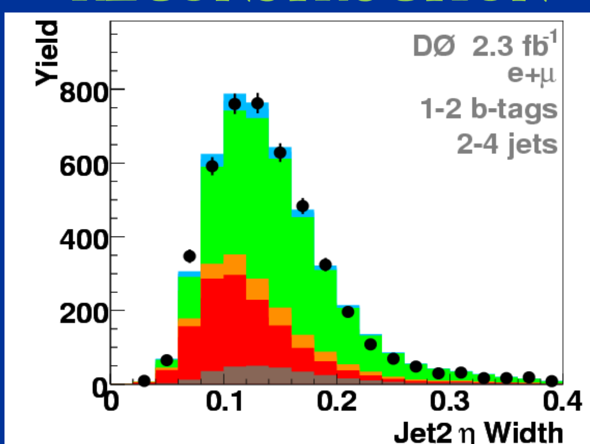
EVENT KINEMATICS



ANGULAR CORRELATIONS

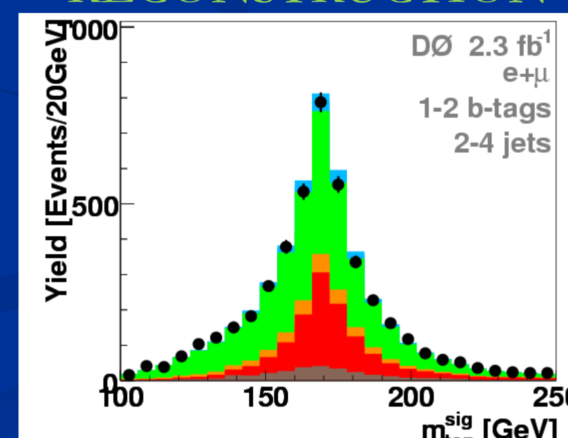


JET RECONSTRUCTION



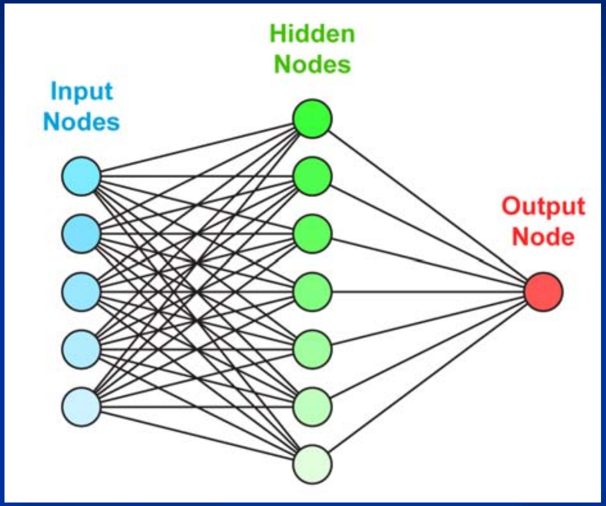
New categories of variables added since 2006 improve performance

TOP QUARK RECONSTRUCTION

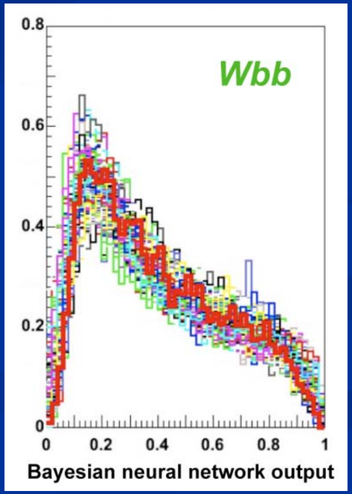


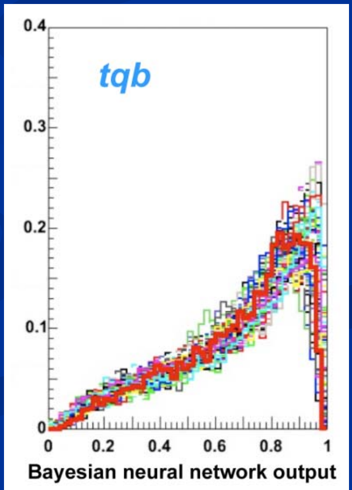
Bayesian Neural Networks

- 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
  - 20 hidden nodes
  - Output: probability for the event to be signal



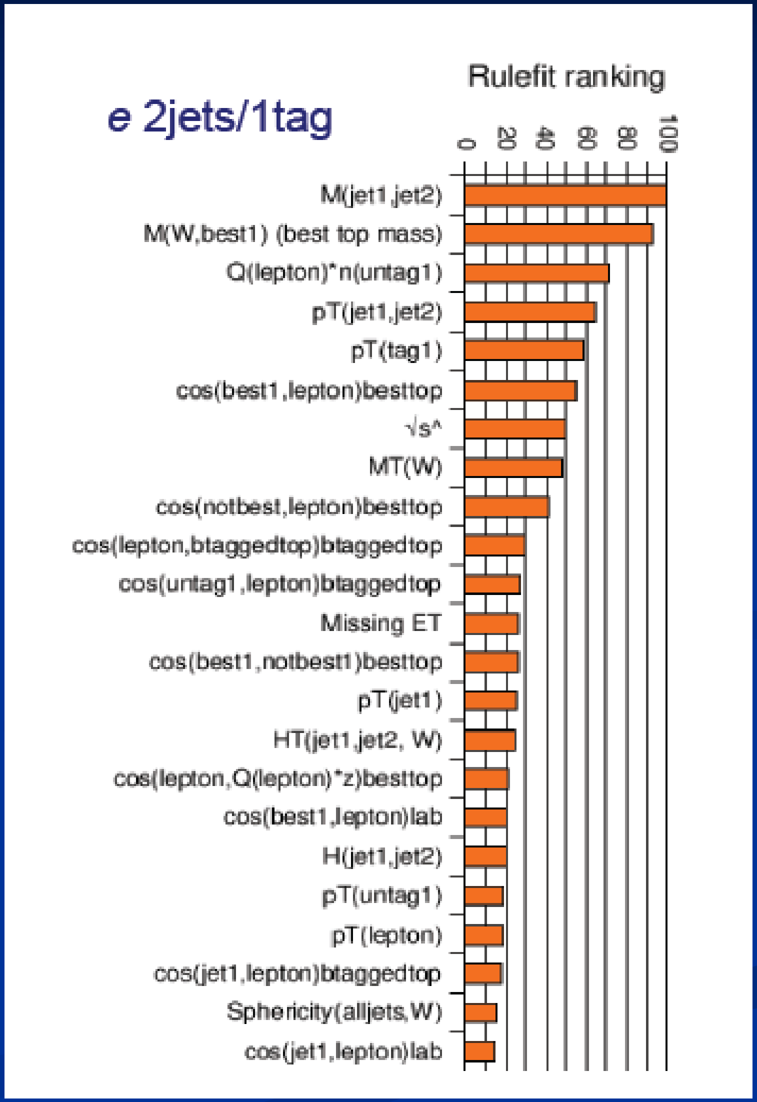
- A Bayesian Neural Network (BNN) is an average over the output of many NN trained iteratively
- It "averages out" statistical fluctuations and avoids over-training.





Selection of Variables

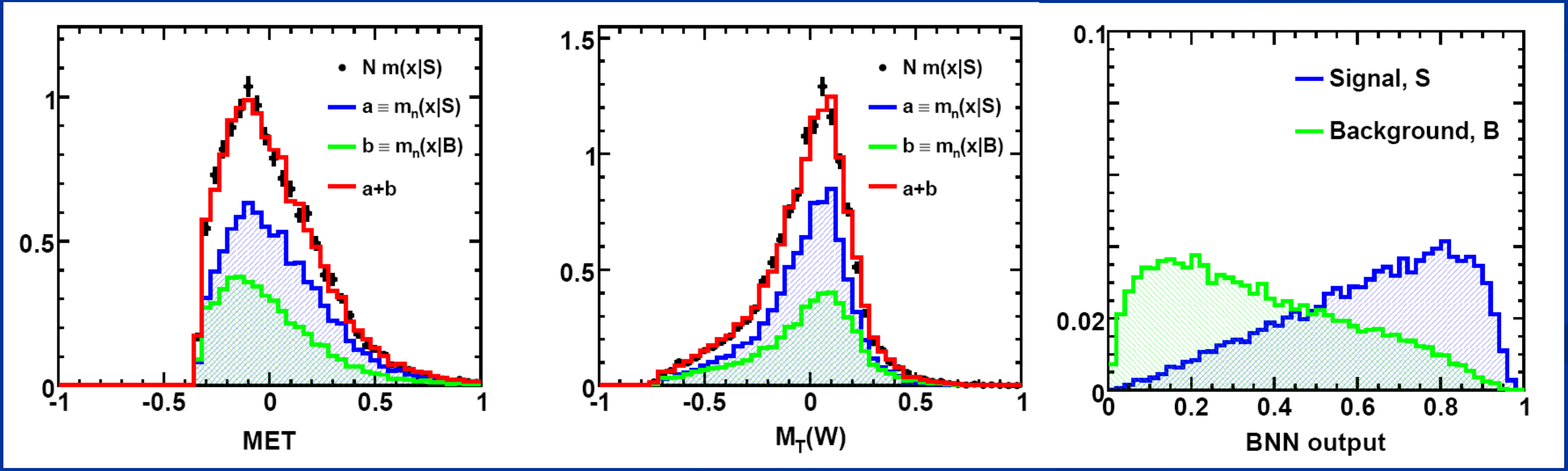
- Start from a set of ~150 well modeled variables
- Use the highest ranked variables for each channel
  - Ranking determined by Rulefit* - a MVA based on Decision Trees (DT)
  - Uses 1/3 of the available MC samples. These samples are later not used for the measurement
  - Importance of each variable 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



* http://www-stat.stanford.edu/~jhf/ftp/RuleFit.pdf

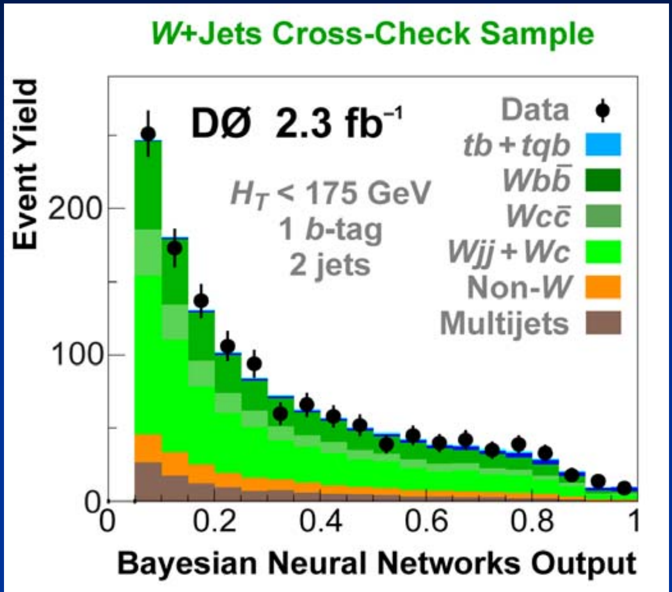
BNN Training & Verification

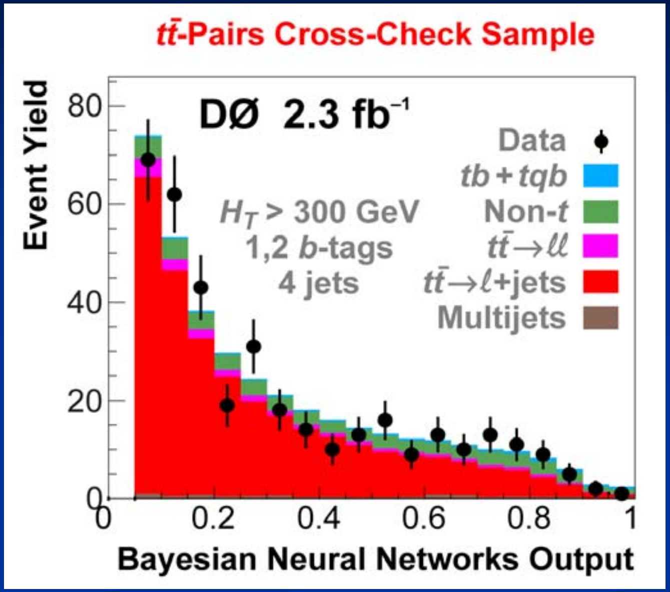
- MC samples were divided into 3 independent subsets: first used for training, second for verification, third to measure cross section
- BNN was trained on a sample consisting of an admixture of signal (tb+tqb) and background (W+jets, $t\bar{t}$, multijets, Z+jets & dibosons)
- Verify that the BNN has converged



Cross Checks

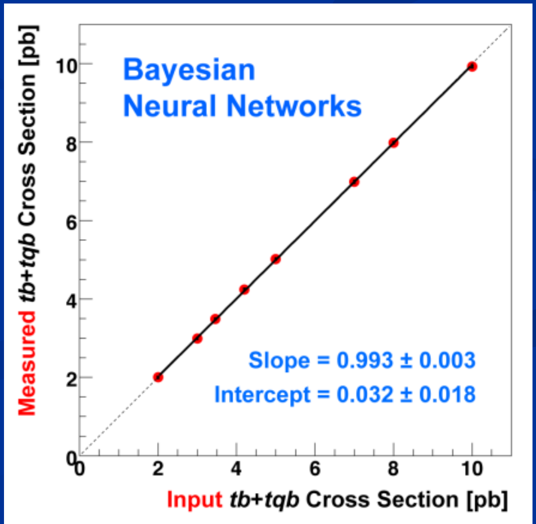
- Check BNN output discriminant in regions dominated by one type of background: tt or W+jets.





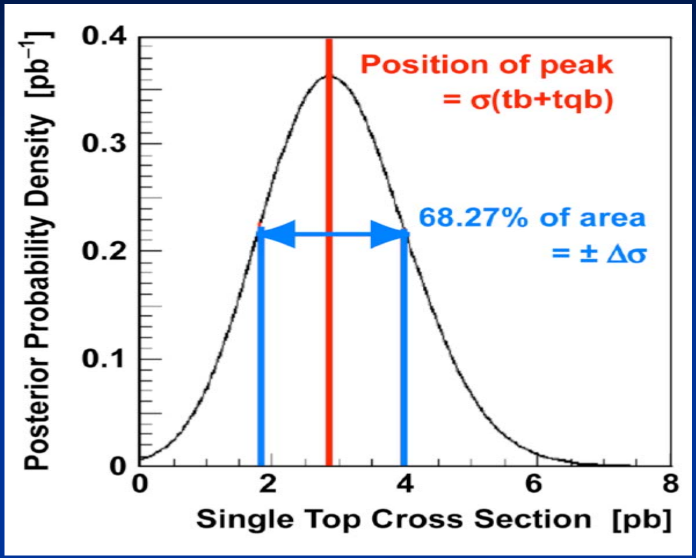
- Check linearity of single top cross section extraction procedure

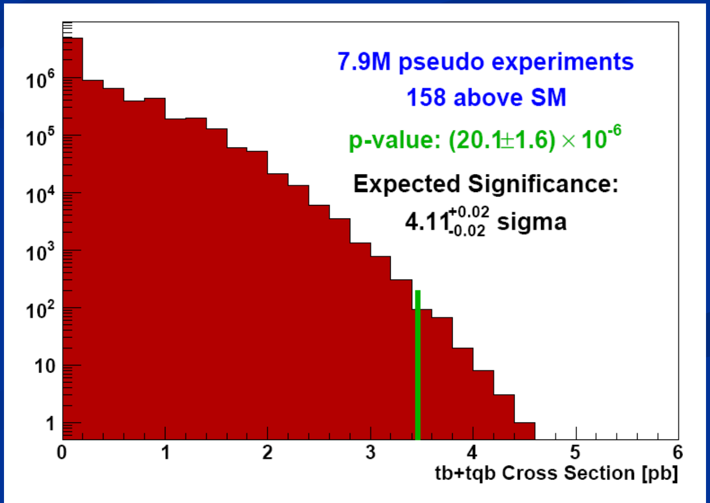
BNN behaves as expected



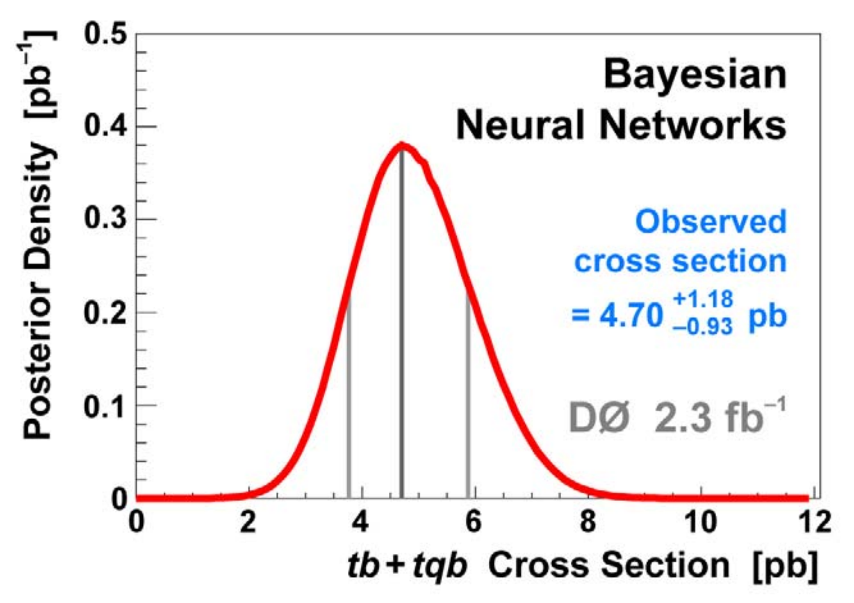
Cross Section & Significance

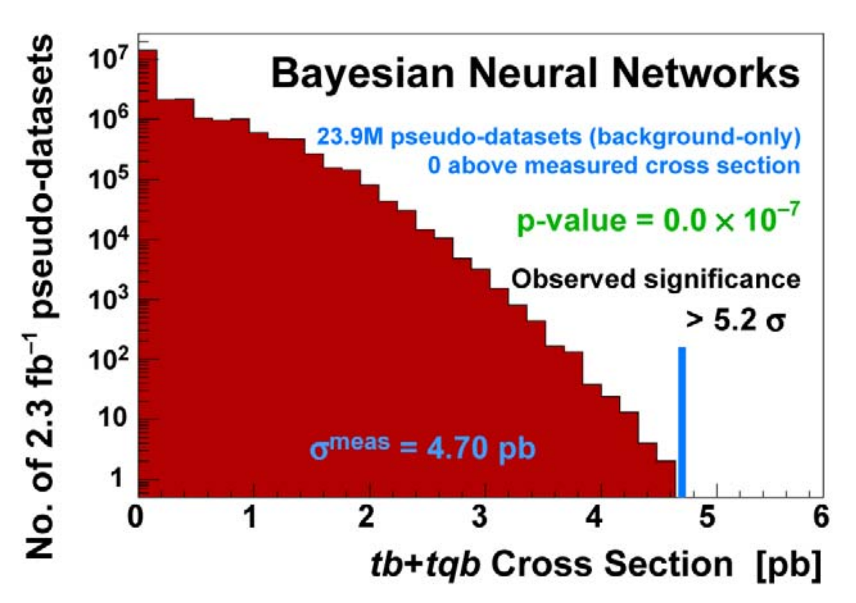
- Cross sections are measured by building a Bayesian posterior probability density
- For each analysis, the single top cross section is given by the position of the posterior density peak, with 68% asymmetric interval as uncertainty
- Gaussian prior for systematic uncertainties
  - Correlations of uncertainties properly taken into account
  - Flat prior in signal cross sections
- Significance derived from background-only pseudo-datasets
  - Expected/Observed: SM/Measured x-sec





BNN Results



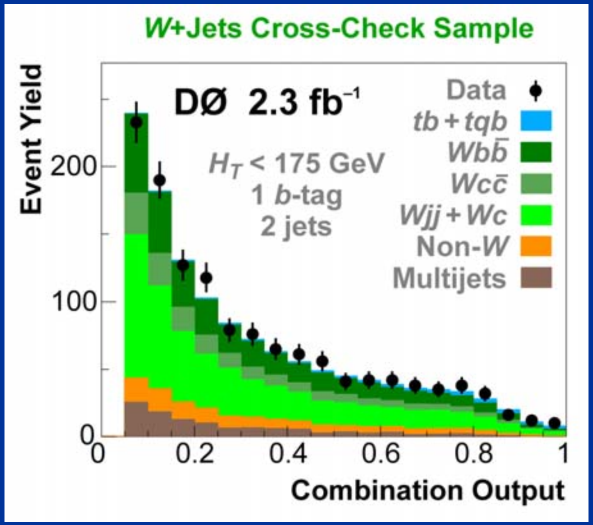


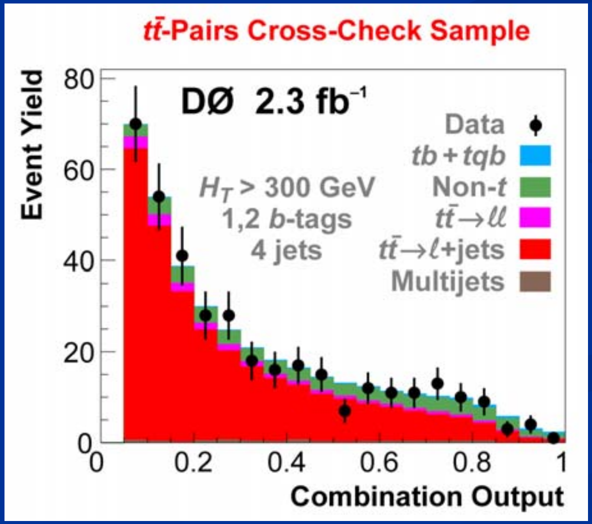
$\sigma \pm \Delta\sigma$(pb)	Expected Sensitivity	Observed Significance
$4.70 \pm^{1.18}_{0.93}$	4.1σ	5.2σ

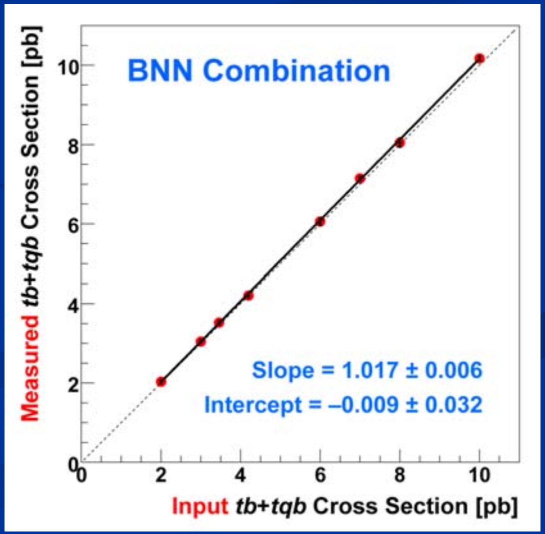
Combination of Results

- Even though all MVA analyses use the same data, they are not 100% correlated
  - BNN&BDT are 75% correlated with each other, 60% with ME
- We use a BNN to combine the three methods. The BNN takes as input variables the output discriminants of the individual methods
- Expected sensitivity for the BNN Combination: 4.5 σ

CROSS CHECK SAMPLES AND LINEARITY



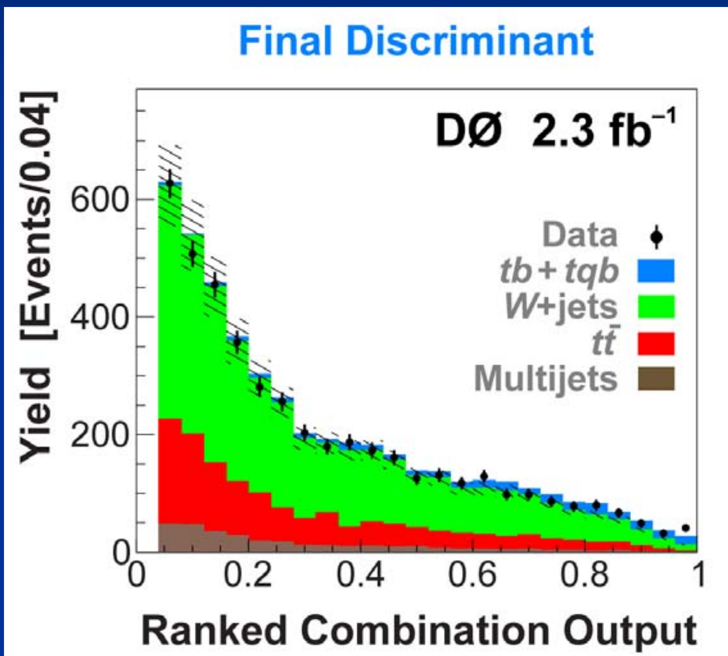


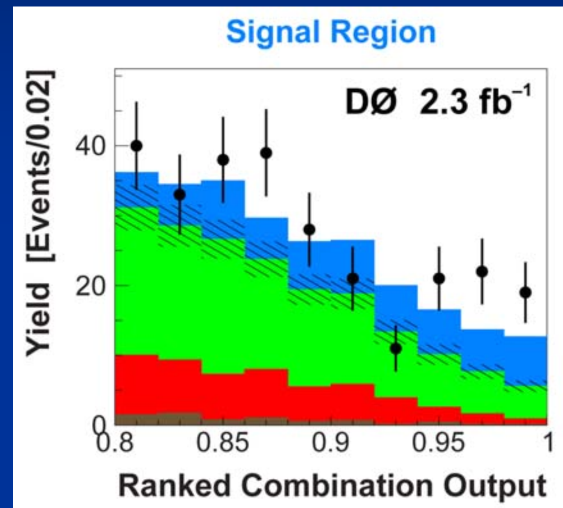


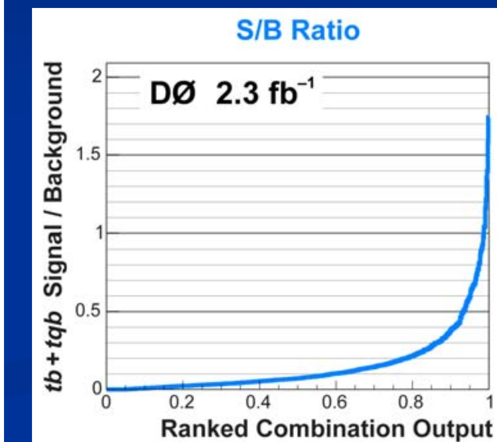
Combined Results

$$\sigma(p\bar{p} \rightarrow tb + X, tqb + X) = 3.94 \pm 0.88 \text{ pb}$$

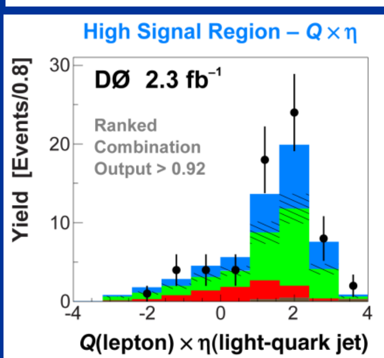
(m_t=170GeV)

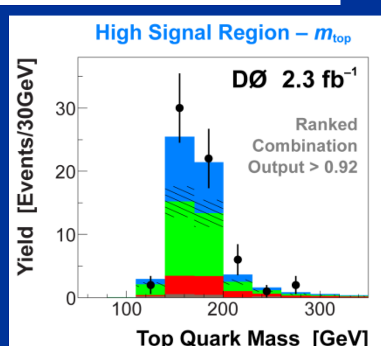




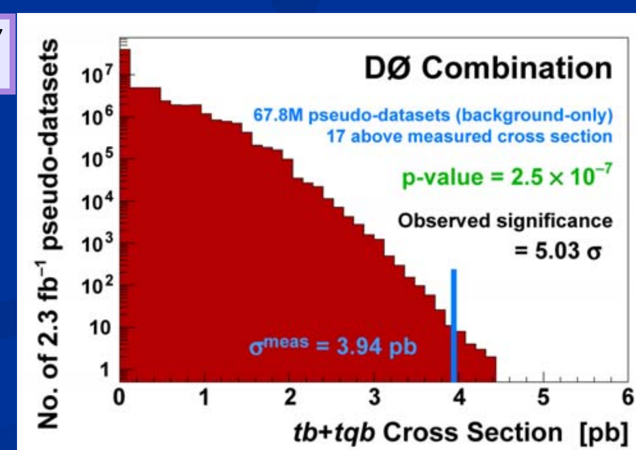


$$p-value = 2.5\times10^{-7}$$





Measured Significance = 5.03σ



Conclusions

- The DØ collaboration observes single top quark production in 2.3 fb^{-1} of Run II data

$$\sigma(p\bar{p} \rightarrow tb + X, tqb + X) = 3.94 \pm 0.88 \text{ pb}$$

Measured Significance 5.03σ

- Bayesian Neural Network (BNN) used as one of the three MVA techniques and for the combination.

http://arxiv.org/abs/0903.0850 submitted to PRL