Observation of Single Top Quark Production at DØ using Bayesian Neural Networks

Jyoti Joshi, Suman Beri (for the DØ Collaboration)
Panjab University, Chandigarh, India

We present observation of the electroweak production of single top quarks in 2.3 fb^{-1} of data recorded by the DØ experiment at the Fermilab Tevatron collider operating at 1.96 TeV center-of-mass energy. The cross section of single top quark production for the combined tb+tqb channels measured is $4.70\,^{+1.18}_{-0.93}$ pb using Bayesian neural networks (BNN). The probability to measure a cross section at this value or higher in the absence of signal is 3.2 x 10^{-8}, corresponding to a 5.4 standard deviation significance for the observation.

1 Introduction

The standard model (SM) predicts top quarks being produced in pairs via strong interaction or singly via the electroweak interaction. In 1995, discovery of the top quark via strong interaction was announced by DØ and CDF experiments[1]. The electroweak single top quark production has been observed recently[2]. The two main production modes of single top are illustrated in Figure 1: the s-channel (tb) process which proceeds via the decay of virtual W boson, and the t-channel (tqb) process which proceeds via the exchange of virtual W boson. The sum of their predicted cross sections for a top quark mass of 170 GeV is 3.46 ± 0.18 pb[3].

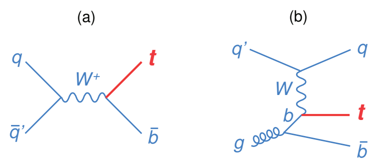

Figure 1: Feynman diagrams for s-channel (a) and t-channel (b) single top production.

2 Event Selection and Analysis Strategy

The 2.3 fb^{-1} dataset used in the analysis is collected from 2002 to 2007 using the DØ detector[4]. The events are selected with a final state consisting of one high transverse momentum (p_T) isolated electron or muon and missing transverse energy ($\not\!\!E_T$) together with a b-quark jet from the decay of the top quark ($t \rightarrow Wb \rightarrow \ell\nu b$), and an additional b antiquark in case of s-channel production, or an additional light-quark jet and a b-antiquark jet for t-channel production. Single top signal events are modeled using COMPHEP-based Monte Carlo (MC) event generator SINGLETOP[5]. $t\bar{t}$ and W+jets backgrounds are simulated using ALPGEN[6], and multijet background is modeled using data which contains non-isolated leptons. All MC events are passed through PYTHIA[7] and then through a GEANT-based full detector simulation. The selection criteria[2] gives 4,519 events, which are expected to contain 223 ± 30 single top quark events. The analysis is split into 24 separate analysis channels based on lepton flavor (e or mu), jet multiplicity (2,3, or 4) and number of identified b jets (1 or 2), to increase the search sensitivity. Systematic uncertainties are considered for all corrections applied to the background model. The total uncertainty on the background is (8–16)% depending on the analysis channel.

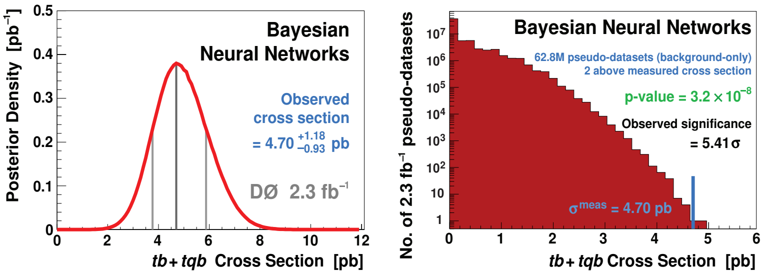

Figure 2: (Left)Posterior density distribution (Right)Cross section distributions on the pseudodatasets for BNN method.

3 Signal-background Separation using BNN

After event selection, we apply Bayesian neural networks (BNN)[8] to extract small single top signal from the large background. A BNN is an average over the output of many neural networks (NN) trained iteratively. Averaging makes the network training more efficient and less prone to overtraining.

We have improved and optimized BNN method from our previous analysis[9] in the choice of input variables and detailed tuning of some parameters. We started from 150 discriminating variables as input and then apply the RuleFitJF algorithm[10] to select the most sensitive kinematic variables, keeping between 18 and 28 of these as inputs, depending on the analysis channel.

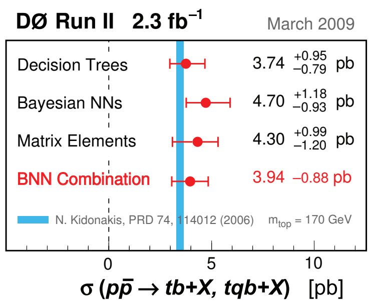

Figure 3: Summary of the measured cross sections compared to theoretical predictions.

4 Cross Section Measurement

We use the BNN output distributions of the 24 analysis channels to form a posterior probability density for the single top production cross section. The position of the peak of the resulting posterior density gives the cross section value and the 68% interval about the peak gives the ± 1 standard deviation (SD) uncertainty. We measure: $\sigma(p\bar{p} \to \text{tb+X,tqb+X}) = 4.70\ ^{+1.18}_{-0.93}$ pb. These measurements are consistent with the SM next-to-leading-order theory calculation.

The sensitivity of analysis to a contribution from single top quark production is estimated by generating an ensemble of pseudodatasets that samples the background model and its uncertainties in the absence of signal. We apply the BNN and measure the cross section for each pseudodataset in the same manner as for the real data which allows us to calculate the probability to measure the SM cross section ("expected significance") or the observed cross section ("observed significance"). The cross section measured by BNN has a $p-$value of 3.2 x 10^{-8} and a significance of 5.4 SD. Figure 2 shows (Left) the posterior density distribution and (Right) the cross section distributions on the pseudodatasets for the BNN method.

Along with BNN, two other MVA techniques are applied: boosted decision trees (BDT) and matrix elements (ME). As the methods are not 100% correlated, we combine them using additional BNN that take as input the output discriminants of the BNN, BDT and ME methods, and produces a single combination output discriminant. The combination leads to an increased expected sensitivity and a more precise measurement of the single top cross section. Figure 3 summarizes the cross sections measured by each of the analysis.

5 Summary

We have used Bayesian neural networks to separate single top quark signal from the background in a sample of lepton+jets events selected from 2.3 fb^{-1} of RunII data. The BNN output distributions across the 24 independent channels were combined using a Bayesian binned likelihood. The measured *tb*+*tqb* single top quark production cross section is:

$$\sigma\,(p\bar{p} \rightarrow tb + X, tqb + X) \quad = 4.70^{+1.18}_{-0.93}\text{ pb}$$

The observed p-value is 3.2×10^{-8}, which corresponds to a 5.4 standard deviation significance.

References

[1] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **74** 2626 (1995);
S. Abachi *et al.* (D0 Collaboration), Phys. Rev. Lett. **74** 2632 (1995).

[2] V.M. Abazov *et al.* (D0 Collaboration), Phys. Rev. Lett. **103** 092001 (2009).

[3] N. Kidonakis, Phys. Rev. D**74** 114012 (2006).

[4] V.M. Abazov *et al.* (D0 Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. **A565** 463 (2006).

[5] E.E. Boos *et al.*, Phys. Atom. Nucl. **69** 1317 (2006).

[6] M.L. Mangano *et al.*, J. High Energy Phys. 07 (2003) 001.

[7] T. Sjostrand, S. Mrenna and P. Skands, J. High Energy Phys. 05 (2006) 026.

[8] R.M. Neal, *Bayesian Learning for Neural Networks* **330** (Springer-Verlag, New York, 1996).

[9] V.M. Abazov *et al.* (D0 Collaboration), Phys. Rev. D**78** 012005 (2008).

[10] J.H. Friedman and B.E. Popescu, Ann. Stat. **2** 916 (2008).