Searches for Flavor-Changing Neutral Currents and Single Top Quarks at D0

Supriya Jain

Homer L. Dodge Dept. of Physics and Astronomy, University of Oklahoma, Norman, OK 73019

(Dated: April 18, 2008)

I present brief descriptions of two top quark analyses performed using D0 data from $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV: the search for flavor-changing neutral-current couplings (FCNC) of the top quark with a charm or an up quark with the exchange of a gluon, and the search for the standard model production of single top quarks. The search for FCNC couplings of a gluon to the top quark is the first such analysis at hadron colliders. I describe the application of the Bayesian approach for comparing our physics models to the observed data in both analyses and, in one of them, to separate small signals from large backgrounds.

INTRODUCTION

It has been over a decade since the top quark was discovered at the Fermilab Tevatron by the CDF and D0 collaborations [1]. Since then, several properties of the top quark have been examined and measured. These include studies and measurements of the kinematical properties of top quark production and measurements of the top quark mass and the production cross section. The most striking feature of the top quark is its large mass, which is intriguingly close to the scale of electroweak (EW) symmetry breaking. This fact raises a number of interesting questions that remain unanswered. Is the top quark mass generated by the Higgs mechanism as postulated by the standard model (SM) and is its mass related to the top to Higgs Yukawa coupling? Or, does the top quark play a fundamental role in the breaking of EW symmetry? If there exist new particles lighter than the top quark, does the top quark decay into them? Could non-SM physics manifest itself in non-standard couplings of the top quark and show up as anomalies in top quark production and decays? My goal in studying top quark physics is to help answer these questions in the most effective manner possible by developing and applying state of the art methods of analysis.

I have worked on many different aspects of top quark physics, but in this paper I describe the application of the Bayesian approach to two different analyses: the search for non-SM flavor-changing-neutral-current (FCNC) couplings of the top quark with a charm or an up quark, and the search for the SM production of single top quarks. These searches were performed using data collected by the D0 detector [2] in Run II.

SEARCH FOR NON-SM SINGLE TOP QUARKS

The top quark was discovered in the top-antitop ($t\bar{t}$) pair production mode via the strong interactions. But the SM also predicts the production of single top quarks via the electroweak interactions. There are two dominant modes of production at the Tevatron, the s-channel in which a top quark is created in association with a bottom quark (tb) and the t-channel in which a light quark accompanies the t and b quarks (tqb). The cross sections for the s-channel and t-channel are predicted to be 0.88 ± 0.14 pb and 1.98 ± 0.30 pb [3], respectively. The main tree-level Feynman diagrams for these processes are shown in Fig. 1.
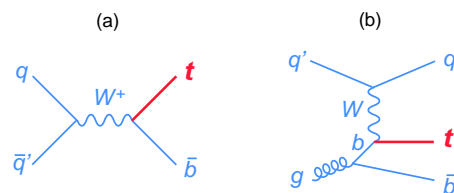

FIG. 1: Main tree-level Feynman diagrams for (a) s-channel single top quark production, and (b) t-channel production.

Besides the SM, several other models make testable predictions about the production of single top quarks, typically, via new interactions [4, 5]. In the SM, FCNC interactions of the top (t) quark with a charm (c) or an up (u) quark can occur via the exchange of a photon, Z boson or gluon through higher-order radiative corrections, but the effect is too small to be observed. Consequently, any observable signal arising from FCNC couplings would be unequivocal evidence of physics beyond the SM and would shed additional light on flavor physics in the top

quark sector. At present, strong constraints exist for FCNC processes via photon or Z boson exchange [6–8], while the search for FCNC couplings of a gluon (g) to the top quark is the first such analysis at hadron colliders [9]. To date, the best constraints on these processes are from the DESY ep Collider (HERA): $\kappa_g/\Lambda < 0.4$ TeV^{-1}, at 95% C.L. [10], where κ_g are the anomalous couplings and Λ is the new physics cut-off scale.

I developed an analysis to search for FCNC [9] that was based upon a previous search for SM single top quarks [11] using 230 pb^{-1} of D0 data from $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV. The FCNC couplings involving the gluon are predicted to be dominant in top quark production rather than in its decay for values of $\kappa_g/\Lambda \lesssim 0.2$ TeV^{-1} [12]. Therefore, for this search, we assume that the top quark decays into a b quark and a W boson as in SM, and the W boson subsequently decays leptonically ($W \to \ell\nu$, where $\ell = e$, μ or τ, with the τ decaying either to an electron or a muon, and two neutrinos). This gives rise to an event with a charged lepton of high transverse momentum (p_T), significant missing transverse energy ($\not{E}_T$) from the neutrinos, and at least two jets, one of which is a b jet (from the top quark decay), and the other a jet from a c quark, u quark, or a gluon. Displaced secondary vertices are used to identify b jets. The largest physics backgrounds to these events are from SM production of W+jets and $t\bar{t}$, along with smaller contributions from SM production of single top quarks (tb and tqb) and dibosons (WW and WZ). An additional source of background is from multijet events in which a jet is incorrectly identified as an electron or in which a muon from a heavy flavor decay appears isolated.

The FCNC signal kinematics for the top-charm-gluon (tcg) and top-up-gluon (tug) processes are modeled using COMPHEP [13], a leading order (LO) parton-level event generator. The LO cross sections are scaled to next-to-leading (NLO) order by a K-factor (the NLO/LO cross section ratio evaluated at some renormalization and factorization scale) of 1.6 [14]. Representative Feynman diagrams for the FCNC production of single top quarks are shown in Fig. 2. The production cross section depends quadratically on the coupling κ_g/Λ and can be significantly larger than that for the SM process as seen in Table I. Since there is no interference between the tcg
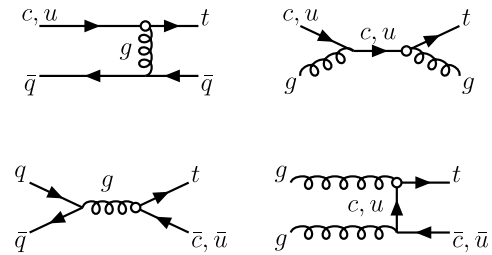

FIG. 2: Representative 2→2 Feynman diagrams for single top quark production through flavor-changing neutral currents involving the gluon.

and tug processes two sets of signal events are generated: one for the tcg process only, in which κ_g^u is set to zero, and the other for the tug process only, in which κ_g^c is set to zero. A value of 0.03 TeV^{-1} is used for κ_g/Λ in each case. Since the effect of these couplings is negligible in top quark decay, the kinematics of signal processes at any other value of κ_g/Λ are obtained by appropriate scaling of the generated events. The parton-level samples from COMPHEP are processed with PYTHIA [15] to model fragmentation, hadronization, and the underlying event, and use the CTEQ5L [16] parton distribution functions. The generated events are processed through a GEANT-based [17] simulation of the D0 detector. The modeling of non-single top backgrounds as well as SM single top quarks is the same as in Ref. [11].

The event selections applied to the simulated signals and backgrounds and to the D0 data are summarized in Table II. The resulting numbers of events in all samples, along with their systematic uncertainties [9], are shown in Table III. The observed numbers of events agree with the predicted numbers for the SM backgrounds within uncertainties in both the electron and muon channels. Hence, we set limits on the anomalous coupling parameters, κ_g/Λ. But note that the FCNC signals are predicted to be a tiny fraction of the SM backgrounds, therefore, optimal signal to background separation is needed to maximize the sensitivity to a putative signal. This separation is done with a neural network based multivariate discriminant.

TABLE I: The production cross sections of single top quarks through a gluon exchange in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV for different values of κ_g/Λ, as obtained from COMPHEP and scaled to NLO by a K-factor of 1.6.

κ_g/Λ [TeV^{-1}]	$\sigma(t)$ [pb]	
	tcg ($\kappa_g^u = 0$)	tug ($\kappa_g^c = 0$)
0.01	0.05	0.88
0.03	0.45	7.92
0.07	2.40	42.61
0.11	5.86	104.78

TABLE II: Summary of event selections.

	Electron channel	Muon channel
Lepton	$E_T > 15$ GeV	$p_T > 15$ GeV
	$|\eta| < 1.1$	$|\eta| < 2.0$
$\not{E}_T$	$15 < \not{E}_T < 200$ GeV	
Jets	2, 3 or 4 jets, $E_T > 15$ GeV, $|\eta| < 3.4$	
	$E_T(\text{jet1}) > 25$ GeV, $|\eta(\text{jet1})| < 2.5$	
	exactly one b-tagged jet	

The neural network is constructed using the ten input variables listed in Table IV. The distribution of one representative variable is plotted in Fig. 3a. The com-

TABLE III: Event yields after all selections for the electron and muon channels. The signal yields are evaluated at $\kappa_g/\Lambda = 0.03$ TeV^{-1}. The yields for $t\bar{t}$ include both lepton+jets and dilepton final states, and those from W+jets also include the diboson backgrounds.

Source	Electron channel	Muon channel
tcg	0.6 ± 0.2	0.6 ± 0.2
tug	8.4 ± 2.1	9.8 ± 2.7
SM single top ($tb+tqb$)	6.4 ± 1.4	6.1 ± 1.4
$t\bar{t}$	31.8 ± 6.9	31.4 ± 7.0
W+jets	84.6 ± 10.2	76.8 ± 8.5
Multijets	13.7 ± 4.3	17.2 ± 1.5
Total SM background	136.5 ± 13.4	131.5 ± 12.7
Observed no. of events	134	118

bination of several variables in this manner and training the FCNC signals (the sum of the tcg and tug processes)

against all backgrounds permits separation of the FCNC signals not only from the dominant backgrounds (W+jets and $t\bar{t}$) but also from the SM single top quark processes, as can be seen in Fig. 3b. Here the neural network outputs for the combined electron and muon channels are shown from different sources normalized to unity. The FCNC signals are evaluated at $\kappa_g/\Lambda = 0.03$ TeV^{-1}. Figure 3c shows the output distributions normalized to the D0 data with backgrounds summed. There is general agreement between the observed spectrum and the predicted SM background.

We use a Bayesian approach to set upper limits [20] on the FCNC coupling parameters simultaneously in a two-dimensional plane of $(\kappa_g^c/\Lambda)^2$ and $(\kappa_g^u/\Lambda)^2$. Given N observed events, we compute the posterior probability density as follows

$$p([\kappa_g^c/\Lambda]^2, [\kappa_g^u/\Lambda]^2 \mid N) \; \propto \; \int \int \int \; L(N \mid n) \; p_1(f_c, f_u, b) \; p_2([\kappa_g^c/\Lambda]^2) \; p_3([\kappa_g^u/\Lambda]^2) \; \mathrm{d}f_c \mathrm{d}f_u \mathrm{d}b, \qquad (1)$$

where L is a Poisson likelihood with mean n, and p_i ($i = 1, 2, 3$) are prior probability densities of the respective parameters. The likelihood L is a product of the likelihoods over all bins of the neural network output distributions, and n is the predicted number of events, equal to the sum of signal (s) and background (b) yields:

$$\begin{aligned} n \; &= \; s \; + \; b \qquad\qquad\qquad\qquad (2)\\ &= \; f_c \; \times \; (\kappa_g^c/\Lambda)^2 \; + \; f_u \; \times \; (\kappa_g^u/\Lambda)^2 \; + \; b, \end{aligned}$$

where the constant factors f_c and f_u are obtained from the simulated signal samples at $\kappa_g/\Lambda = 0.03$ TeV^{-1}. The prior probability density p_1 is a multivariate Gaussian with the mean and standard deviation defined by the estimated yields and their uncertainties, to take into account correlations among the different samples and bins. Since the signal cross sections depend quadratically on κ_g/Λ, for p_2 and p_3 we choose priors flat in $(\kappa_g^c/\Lambda)^2$ and $(\kappa_g^u/\Lambda)^2$ respectively, in order to be consistent with the conventional choice of a prior flat in cross section.

From the two-dimensional posterior probability density, allowed regions at different levels of confidence (k) are defined as contours of equal probability that enclose a volume k around the peak of the posterior density. These contours are shown in Fig. 4, using data from both electron and muon channels. The one-dimensional posterior probability density over any dimension is obtained by integrating the two-dimensional posterior over the other dimension. The resulting limits, translated to κ_g/Λ, using data (observed limits) as well as the expected limits for which the "observed" count is set to the predicted

background yield in any bin, are summarized in Table V.

In summary, this search for the production of single top quarks via FCNC interactions in 230 pb^{-1} of lepton+jets data collected at D0 yielded no deviation from SM predictions. Hence limits are set on anomalous coupling parameters, κ_g^c/Λ and κ_g^u/Λ, based on distributions of multivariate neural network discriminants. The observed (expected) limits are 0.15 (0.16) TeV^{-1} on κ_g^c/Λ, and 0.037 (0.041) TeV^{-1} on κ_g^u/Λ at 95% C.L. These limits on FCNC couplings of a gluon to the top quark and a charm or up quark, the first from a hadron collider, are a factor 3–11 times better than those from HERA.

SEARCH FOR SM SINGLE TOP QUARKS

Using a much larger dataset than was available for the search described above (0.9 fb^{-1} of lepton+jets data), D0 reported evidence for the production of SM single top quarks at a significance of 3.4 standard deviations [21]. Since the single top quark signals have a cross section roughly half that of $t\bar{t}$ production and larger backgrounds, due to the presence of only one massive object in the final state, they are extremely challenging to identify. At D0, three different multivariate techniques were used to create a discriminant

$$D(\mathbf{x}) = \frac{p(\mathbf{x}|S)}{p(\mathbf{x}|S) + p(\mathbf{x}|B)}, \qquad (3)$$

TABLE IV: Input variables used in the neural network analysis.

$p_T(\text{jet1})$	Transverse momentum of the leading jet
$p_T(\text{jet1}_{\text{tagged}})$	Transverse momentum of the b-tagged jet
$\eta(\text{lepton})$	Pseudorapidity [18] of the lepton
$\not{E}_T$	Missing transverse energy
$p_T(\text{jet1}, \text{jet2})$	Transverse momentum of the two leading jets
$H_T(\text{jet1}, \text{jet2})$	Scalar sum of the transverse momenta of the two leading jets
$p_T(W)$	Transverse momentum of the reconstructed W boson
$M(W, \text{jet1}_{\text{tagged}})$	Invariant mass of the reconstructed top quark using the W boson [19] and the b-tagged jet
$M(\text{alljets})$	Invariant mass of all jets
$\cos(\text{jet1}, \text{lepton})_{\text{lab}}$	Cosine of the angle between the leading jet and lepton in the laboratory frame of reference
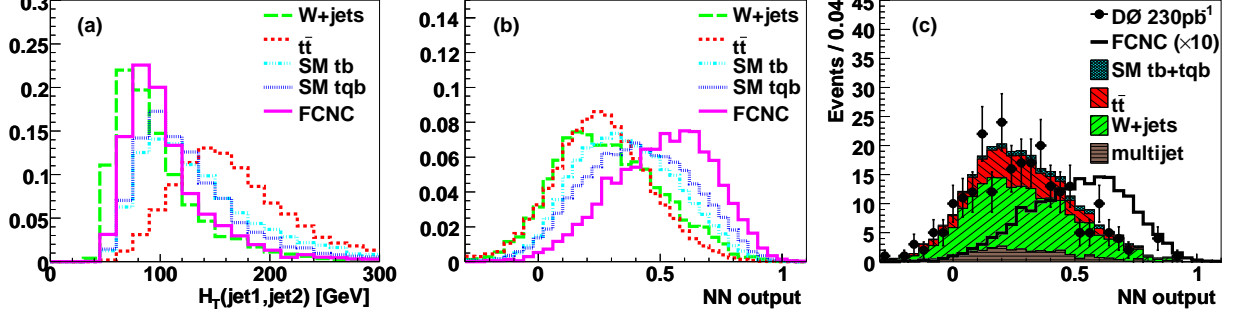
FIG. 3: Distributions of (a) an input variable to neural networks, and outputs normalized to (b) unity, and (c) 230 pb^{-1} of data with backgrounds summed. The FCNC signal is for the summed tcg and tug processes, each evaluated at $\kappa_g/\Lambda = 0.03$ TeV^{-1}.
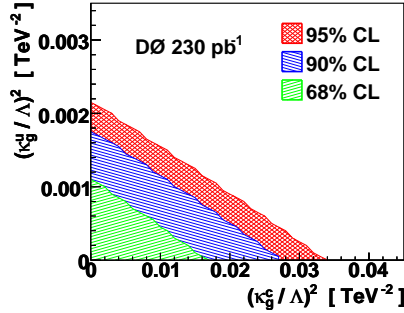
FIG. 4: Allowed regions at various levels of confidence using 230 pb^{-1} of D0 data in both the electron and muon channels.

TABLE V: Upper limits on κ_g^c/Λ and κ_g^u/Λ, at 95% C.L.

	Observed (expected) limits [TeV^{-1}]	
	κ_g^c/Λ	κ_g^u/Λ
Electron channel	0.16 (0.19)	0.046 (0.052)
Muon channel	0.21 (0.21)	0.049 (0.050)
Combined	**0.15 (0.16)**	**0.037 (0.041)**

to separate the small signals ($tb+tqb$) from the large backgrounds. Here $p(\mathbf{x}|S)$ and $p(\mathbf{x}|B)$ are the probability density functions for the signal and background, respec-

tively, with each event described by the variables $\mathbf{x}$. The multivariate techniques used were: boosted decision trees (DT) [22, 23], Bayesian neural networks (BNN) [24], and matrix elements (ME) [25, 26]. The DT analysis approximates the discriminant $D(\mathbf{x})$ using an average of many piece-wise approximations to $D(\mathbf{x})$. The BNN analysis uses nonlinear functions that approximate $D(\mathbf{x})$ directly, that is, without first approximating the densities $p(\mathbf{x}|S)$ and $p(\mathbf{x}|B)$. Both methods used several input variables based on object kinematics, global event variables and angular variables. The ME method approximates the densities $p(\mathbf{x}|S)$ and $p(\mathbf{x}|B)$ semi-analytically, starting with leading-order matrix elements, and computes $D(\mathbf{x})$ from them using the basic observables:

1. missing transverse energy 2-vector ($\not{E}_T, \phi$),

2. lepton 4-vector (E_T, η, ϕ), assuming massless leptons,

3. jet 4-vector (E_T, η, ϕ), assuming massless jets, and jet-type, that is, whether it is a b jet or not, for each jet.

Since the first evidence reported in 2007 [21], several improvements were made to the BNN and ME analyses, and also the results from the three techniques (DT, BNN and ME) were combined, resulting in an enhanced significance of the single top cross section measurement (3.6 standard deviations) [27]. My principal contribution was

in improving the BNN analysis and developing, testing, and performing the combined analysis of the results of the DT, BNN, and ME analyes as well as performing exhaustive ensemble studies of the entire Bayesian machinery, to which I also made significant contributions.

Bayesian Neural Network Analysis

The current D0 BNN analysis is a considerable improvement over the previous one and is as sensitive as the DT and ME analyses, as can be seen from the power curves in Fig. 5. The principal reasons for the improve-
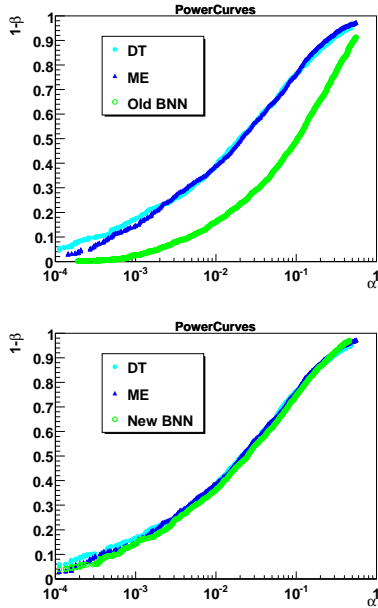

FIG. 5: The power computed from the SM signal+background ensemble (power=1-β on the y-axis) versus the significance from the background-only ensemble (α on the x-axis) for reference cross sections varying monotonically from $0 - 10$pb. Power is the probability to accept the signal+background hypothesis, if it is true, while significance is the probability to reject the background only hypothesis, if instead it is true. For a given significance, one wants the power to be as large as possible.

ment are the following:

1. systematic selection of input variables for each analysis channel (based on lepton flavor: electron or muon, jet multiplicity: 2-jet, 3-jet or 4-jet, and number of b-tags: 1-tag or 2-tags)

2. reduced sensitivity of the Bayesian posterior density of the network parameters to noise in the training sample

The variables for each channel are selected using an algorithm called RuleFit [28] that orders them according to

their discrimination importance (on a scale of 1 to 100). Variables with discrimination importance greater than 10 are used, which results in the selection of between 18 and 25 variables in the different channels. For example, the variables for the electron+2jets/1tag channel are shown in Fig. 6.

For the network parameters, a single hidden layer with 20 nodes is used, with 10,000 events in the training sample. The background MC events used turned out to be especially noisy because of the large dynamic range of their event weights. An obvious way to reduce the effect of events with abnormally large weights is simply to discard them from the training sample. However, we chose to keep all events, but to reduce the influence of those with very large weights by limiting the growth of the network parameter values. This was done by fixing the width of the (Gaussian) prior for each network parameter to one of a small set of values determined from a study of the distribution of parameter values. The use of this non-adaptive, and therefore noise-insensitive, prior led to a dramatic improvement with respect to the previous analysis.
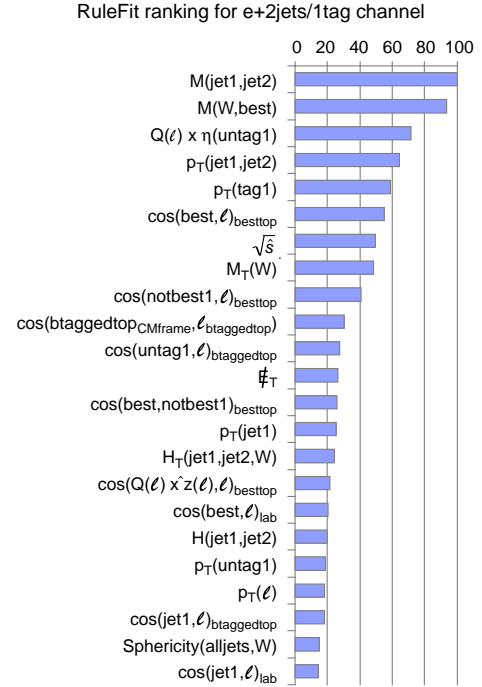

FIG. 6: BNN input variables according to their RuleFit ranking for the electron+2jets/1tag channel.

Figure 7 shows the observed BNN distribution summed over all channels, superimposed on the summed signal + background model. There is excellent agreement between observed data and model, with an excess (although within uncertainties) over SM backgrounds in consecutive bins at large BNN output values.
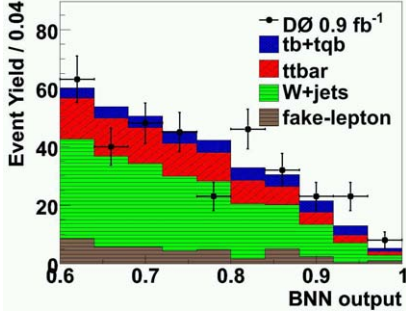
FIG. 7: The observed BNN output distribution summed over all 12 channels superimposed on the summed signal + background model, in the high BNN output region.

After performing a one-dimensional Bayesian statistical analysis [20] using the BNN output distribution, and with the $tb+tqb$ cross section as the parameter of interest, we obtain the following result:

$$\sigma\,(p\bar{p} \rightarrow tb + tqb + X) = 4.4^{+1.6}_{-1.4} \text{ pb}.$$

The Bayesian posterior densities using data (observed) and the signal+background hypothesis (expected) are shown in Fig. 8. From an ensemble study, this results in a p-value of 0.08%, corresponding to a 3.2 standard deviation significance. The p-value from the SM value of 2.86 pb for the signal cross section is 1.6%, which corresponds to a 2.2 standard deviation expected significance. It is interesting to note that with fewer input variables, 18-25 across the different channels, the BNN has an expected significance as good as the DT analysis (2.1 standard deviation) which uses 49 input variables.
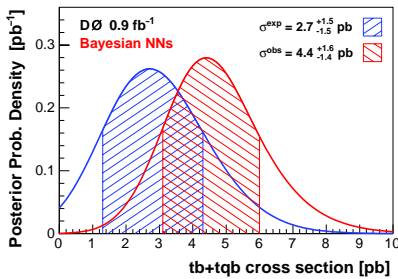

FIG. 8: Expected (SM) and observed Bayesian posterior density distributions for the BNN analysis. The shaded regions indicate one standard deviation above and below the peak positions.

Combination of Results

Since each multivariate analysis uses the same dataset to measure the single top quark cross section, their results

are highly correlated. However, because the correlation is rather less than 100%, one can still gain some additional sensitivity by combining the results. The three cross section measurements, σ_i (i = DT, BNN, ME) are combined using the best linear unbiased estimate (BLUE) method [29–31]; that is, a new estimate of the cross section is defined by the weighted sum

$$\sigma = \sum_i w_i\,\sigma_i, \qquad (4)$$

with $\sum_i w_i = 1$, and with the weights chosen so as to minimize the variance

$$\text{Var}(y) = \sum_i \sum_j w_i\,w_j\,\text{Cov}(\sigma_i, \sigma_j), \qquad (5)$$

where $\text{Cov}\,(\sigma_i, \sigma_j) \equiv \langle \sigma_i \sigma_j \rangle - \langle \sigma_i \rangle \langle \sigma_j \rangle$ are the matrix elements of the covariance matrix of the measurements. The variance is minimized when

$$w_i = \frac{\sum_j \text{Cov}^{-1}(\sigma_i, \sigma_j)}{\sum_i \sum_j \text{Cov}^{-1}(\sigma_i, \sigma_j)}, \qquad (6)$$

where $\text{Cov}^{-1}(\sigma_i, \sigma_j)$ denotes the matrix elements of the inverse of the covariance matrix.

Weights, Correlations, and Combined Measurements

We use an ensemble of simulated cross sections results, based on the SM signal+background hypothesis, to determine the weights w_i and the correlation matrix. The cross section measurements from this ensemble are shown in Fig. 9 for the individual and combined analyses. The mean and square root of the variance obtained from these distributions give the following expected measurements:

$$\sigma^{\text{SM-ens}}(p\bar{p} \rightarrow tb + tqb + X)$$
$$= 2.9 \pm 1.6 \text{ pb}\ \ (\text{DT})$$
$$= 2.7 \pm 1.5 \text{ pb}\ \ (\text{BNN})$$
$$= 3.2 \pm 1.4 \text{ pb}\ \ (\text{ME})$$
$$= 3.0 \pm 1.3 \text{ pb}\ \ (\text{Combined}).$$

The weights w_i for the three analyses are found to be

- $w_{\text{DT}} = 0.127,$

- $w_{\text{BNN}} = 0.386,$

- $w_{\text{ME}} = 0.488.$

with correlation matrix

$$\text{Correlation matrix} = \begin{pmatrix} & \text{DT} & \text{BNN} & \text{ME} \\ 1 & 0.66 & 0.64 \\ 0.66 & 1 & 0.59 \\ 0.64 & 0.59 & 1 \end{pmatrix} \begin{matrix} \text{DT} \\ \text{BNN} \\ \text{ME} \end{matrix}. \qquad (7)$$
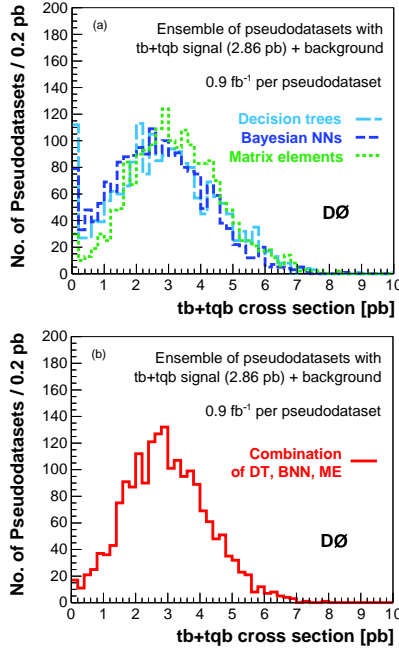
FIG. 9: Distributions of the measured cross sections from (a) the individual analyses, and (b) the combined analysis, using the SM ensemble.

The one-standard-deviation coverage probability of the (Bayesian) confidence interval is 0.67.

The observed measurements for the individual as well as combined analyses are summarized in Fig. 10.
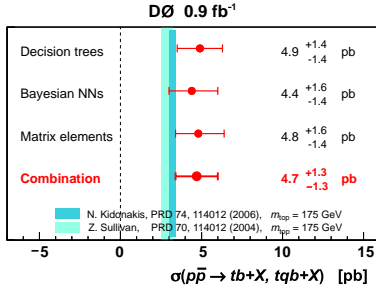

FIG. 10: The measured single top quark cross sections from the individual analyses and their combination.

Measurement Significance

The background-only ensemble is used to determine the signal significance corresponding to the observed measurement and the SM value of 2.86 pb for the single top quark production cross section. Distributions of the ensemble-results from all the analyses are shown in Fig. 11. The observed (expected) p-value, and the associated significance in Gaussian-like standard deviations, is
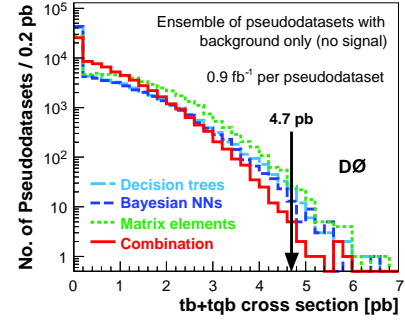

FIG. 11: Distributions of the cross sections measured from background-only ensemble by the three analyses and their combination. The arrow shows the combined cross section measurement, 4.7 pb.

obtained by counting how many background-only pseudodatasets yield a measured cross section greater than the observed (2.86 pb) cross section measurement. The results are tabulated in Table VI for the individual as well as combined analyses.

TABLE VI: The observed (expected) p-values, and significances for the individual and combined analyses, using the observed measurement (SM value of 2.86 pb) for the single top quark production cross section as the reference point in Fig. 11.

	p-value	significance (std. dev.)
Analysis	Observed (Expected)	Observed (Expected)
DT	0.00037 (0.018)	3.4 (2.1)
BNN	0.00083 (0.016)	3.1 (2.2)
ME	0.00082 (0.031)	3.2 (1.9)
Combined	0.00014 (0.011)	3.6 (2.3)

Using the SM (signal+background) ensemble, we also quantify the compatibility of our result with the SM expectation by counting how many pseudodatasets result in a cross section with the observed value or higher for each of the analyses. The probabilities for the different analyses are 10% for the DT analysis, 13% for the ME analysis, 13% for the BNN analysis, and 10% for the combined analysis.

In the method described above, we have used the SM ensemble to determine the weights for the individual analyses in the combined result. But one could have used a different ensemble, for example, one generated with a signal value different from 2.86 pb. This may yield a different set of results using the BLUE method. A more ideal approach would be to perform a three-dimensional probability density estimate of the results of the three analyses (DT, BNN and ME) for sets of ensembles, each generated with a different value of the $tb+tqb$ cross section between zero and a sufficiently large upper bound. This density estimate, $p(\sigma_{\rm DT}, \sigma_{\rm BNN}, \sigma_{\rm ME}|\sigma)$, evaluated

at the triplet of measured cross sections would serve as a likelihood function in a Bayesian analysis [20] to measure the cross section using the combined results. The merit of this approach is that all possible hypotheses for the signal cross section are taken into account in order to derive the combined result. I am currently exploring this technique to compare its results with those from the BLUE method.

Single Top Projections

For the measurement of the single top production cross section discussed so far, a one-dimensional Bayesian posterior density was constructed as a function of the $tb+tqb$ cross section, assuming the SM ratio between the s and t channel cross sections. I also extended the BNN analysis to build a two-dimensional posterior density in the plane of tb versus tqb cross sections where the s and t channel cross sections are allowed to vary independently of each other. A prior density flat in both variables is chosen, again by convention. Fig. 12 shows the expected allowed regions at 68% and 95% confidence levels for different values of integrated luminosity. No systematic effects are included in these projections but their effect is expected to be compensated by improvements to the analysis that have not yet been incorporated. These studies indicate that at about 7 fb^{-1} it is possible to rule out several non-SM processes.
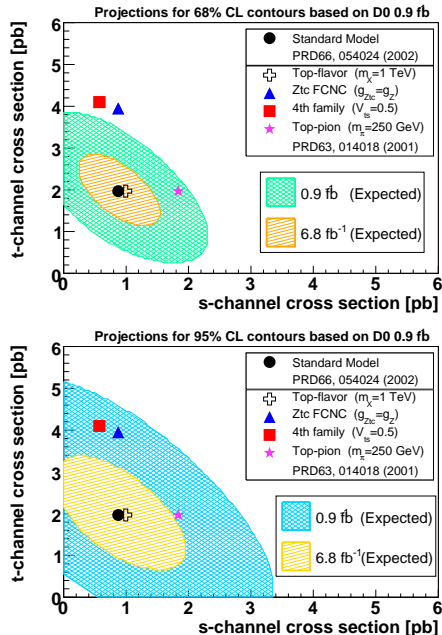

FIG. 12: Projections at different values of integrated luminosity the expected allowed regions at 68% and 95% confidence levels based on 0.9 fb^{-1} of D0 data in both the electron and muon channels.

Finally, as discussed in Ref. [27], the one-dimensional Bayesian posterior density estimation can be extended to measure the CKM matrix element $|V_{tb}|$ also since the single top cross section depends quadratically on this parameter. Fig. 13 shows projections made for the expected resolution of the $|V_{tb}|$ measurement using the BNN outputs. All systematic effects are included. We see that at D0 it is possible to reach a resolution of 15% at about 7 fb^{-1}, which, interestingly, is comparable to the reach at the LHC.
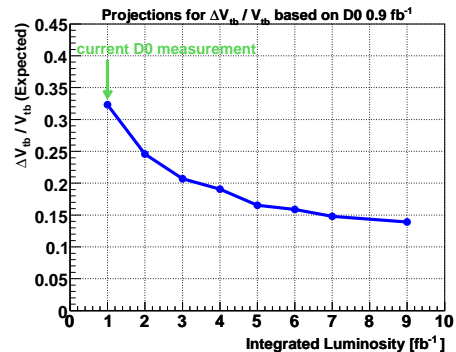

FIG. 13: Projections at different values of integrated luminosity for the expected resolution of the $|V_{tb}|$ measurement based on 0.9 fb^{-1} of D0 data in both the electron and muon channels.

FUTURE PLANS

I plan to explore further the Bayesian approach, both as a multivariate technique as well as a statistical inference tool. In the former, I would like to test the performance using the the four-vector variables directly (taking care to break the azimuthal symmetry of events) as inputs to construct the discriminant, $D(\mathbf{x})$. We note that the derived variables, multi-object variables or angular correlations, contain no more information than is contained in the original degrees of freedom. The reason we use them is that for some numerical approximation methods, it may prove easier to construct an accurate approximation to $D(\mathbf{x})$ if it is built using carefully chosen derived variables than one constructed directly in terms of the underlying degrees of freedom. However, the greater the number of variables used, the larger is the effect on the outputs from certain types of systematics, for example the jet energy scale. This is one of the dominant sources of uncertainty in our analysis ($6 - 10\%$ across the different channels), and a large number of jet-related variables could result in large fluctuations in the discriminant outputs. Secondly, the networks are expected to converge faster with fewer input variables.

In the Bayesian method used for comparing our model

(background-only hypothesis, or signal+background hypothesis) to the observed numbers of events, a necessary input is a prior probability density, both for the nuisance parameters such as backgrounds and acceptances, as well as for the parameters of interest, namely the signal cross sections. So far, we have assumed Gaussian densities for the systematic effects on the nuisance parameters. But small values of expected yields can result in truncation of the Gaussian density, since the numbers of events are constrained to be positive. I would like to explore different possibilities for modeling the systematics, for example, using the Gamma distribution, which goes to zero for zero expected counts. As for the prior for the signal cross section, we have used a density flat in cross section. This has been the convention so far in our field, but I would like to explore building a prior, in a more formal way, in order to maximize the expected sensitivity of our measurement.

To conclude, I have used the Bayesian approach extensively to separate small signals from large backgrounds and I have also extended it from the usual one-dimensional case to two-dimensions, and from cross sections to parameters of physical interest (eg. FCNC couplings κ_g/Λ or $|V_{tb}|$). I am currently exploring new ideas and techniques in this approach in order to maximize the sensitivity of the single top analysis at D0 . Although this work is being done in the single top context, it is an extremely useful test-bed for several other analyses in the field of high energy physics, specially those in which the signal processes are dominated by huge backgrounds.

This paper has been originally prepared as part of Fermilab's Tollestrup award competition for outstanding post-doctoral research work, and I would like to thank all my collaborators at D0, specially those within the single top group, without whom the work described here would not have been possible.

[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], Nucl. Instrum. Methods A **565**, 463 (2006).

[3] Z. Sullivan, Phys. Rev. D **70**, 114012 (2004).

[4] T. Han *et al.*, Phys. Rev. D **58**, 073008 (1998).

[5] T. Tait and C.P. Yuan, Phys. Rev. D **63**, 014018 (2001).

[6] F. Abe *et al.* [CDF Collaboration], Phys. Rev. Lett. **80**, 2525 (1998).

[7] P. Achard *et al.* [L3 Collaboration], Phys. Lett. B **549**, 290 (2002).

[8] S. Chekanov *et al.* [ZEUS Collaboration], Phys. Lett. B **559**, 153 (2004).

[9] V.M. Abazov *et al.* [D0 Collaboration], Phys. Rev. Lett. **99**, 191802 (2007).

[10] A.A. Ashimova and S.R. Slabospitsky, "The constraint on FCNC coupling of the top quark with a gluon from *ep* collisions," arXiv:hep-ph/0604119 (2006).

[11] V.M. Abazov *et al.* [D0 Collaboration], Phys. Lett. B **622**, 265 (2005); V.M. Abazov *et al.* [D0 Collaboration], Phys. Rev. D **75**, 092007 (2007).

[12] M. Hosch *et al.*, Phys. Rev. D **56**, 5725 (1997).

[13] E. Boos *et al.* [CompHEP Collaboration], Nucl. Instrum. Methods A **534**, 250 (2004).

[14] J.J. Liu *et al.*, Phys. Rev. D **72**, 074018 (2005).

[15] T. Sjöstrand *et al.*, Comput. Phys. Commun. **135**, 238 (2001).

[16] H.L. Lai *et al.* [CTEQ Collaboration], Eur. Phys. J. C **12**, 375 (2002).

[17] R. Brun *et al.*, CERN Program Library Long Writeup W5013 (1994).

[18] Pseudorapidity is defined as $\eta = -\ln(\tan\frac{\theta}{2})$, where θ is the polar angle with respect to the beam axis, with the origin at the primary vertex.

[19] The longitudinal component of the momentum of the neutrino is obtained using a SM W boson mass constraint, choosing the smaller of the two possible solutions.

[20] I. Bertram *et al.*, FERMILAB-TM-2104 (2000).

[21] V.M. Abazov *et al.* [D0 Collaboration], Phys. Rev. Lett. **98**, 181902 (2007).

[22] L. Breiman, J. Friedman, C.J. Stone, and R.A. Olshen, *Classification and Regression Trees* (Wadsworth, Stamford, 1984).

[23] D. Bowser-Chao and D.L. Dzialo, Phys. Rev. D **47**, 1900 (1993).

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

[25] V.M. Abazov *et al.* (D0 Collaboration), Nature **429**, 638 (2004).

[26] A. Abulencia *et al.* (CDF Collaboration), Phys. Rev. Lett. **96**, 152002 (2006).

[27] V.M. Abazov *et al.* (D0 Collaboration), *Evidence for Production of Single Top Quarks*, submitted to Phys. Rev. D, arXiv:0803.0739

[28] J.H. Friedman, in *Statistical Problems in Particle Physics, Astrophysics and Cosmology*, edited by L. Lyons and M. K. Ünel (Imperial College Press, London, England, 2006), p. 127, http://www-stat.stanford.edu/~jhf/R-RuleFit.html.

[29] L. Lyons, D. Gibaut, and P. Clifford, Nucl. Instrum. Methods Phys. Res., Sect. A **270**, 110 (1988).

[30] R.J. Barlow, *Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences* (The Manchester Physics Series, John Wiley and Sons, New York, 1989).

[31] G. Cowan, *Statistical Data Analysis* (Oxford University Press, Oxford, 1998).