CDF Note 9223

Measurement of Single Top Quark Production in 2.2 fb^{-1} of CDF II Data using the Matrix Element Technique

The CDF Collaboration
URL http://www-cdf.fnal.gov
(Dated: February 29, 2008)

We present an improved search for electroweak single top quark production using 2.2 fb^{-1} of CDF II data collected between February 2002 and August 2007 at the Tevatron in proton-antiproton collisions at a center-of-mass energy of 1.96 TeV. The analysis employs a matrix element technique which is used to calculate event probability densities for the signal and background hypothesis. The ratio of signal and background event probabilities is used as a discriminant variable which we fit to the data. We search for a combined single top s- and t-channel signal and measure a cross section of $2.2^{+0.8}_{-0.7}$ pb assuming a top quark mass of 175 GeV/c^2. The probability that the observed excess originated from a background fluctuation (p-value) is 0.0003 and the expected (median) p-value in pseudo-experiments is 0.000003 corresponding to a 3.4 σ signal significance observed in data and 4.5 σ expected.

Preliminary Results for the Winter Conferences 2008

INTRODUCTION

In proton anti-proton collisions at the Tevatron with a center-of-mass energy of 1.96 TeV, top quarks are predominantly produced in pairs via the strong force. In addition, the Standard Model predicts single top quarks to be produced through an electroweak t- and s-channel exchange of a virtual W boson as shown in Figure 1. The production cross sections have been calculated at Next-to-Leading-Order (NLO). For a top quark mass of 175 GeV/c^2 the results are 1.98±0.25 pb and 0.884±0.11 pb for the t-channel and s-channel process respectively [1]. The combined cross section is about 40% of the top anti-top pair production cross section ($\sigma_{singletop} \sim 2.9$ pb). The measurement of electroweak single top production probes the $W-t-b$ vertex, provides a direct determination of the Cabbibo-Kobayashi-Maskawa (CKM) matrix element $|V_{tb}|$ and offers a source of almost 100% polarized top quarks [2]. Moreover, the search for single top also probes exotic models beyond the Standard Model. New physics, like flavor-changing neutral currents or heavy W' bosons, could alter the observed production rate [3]. Finally, single top processes result in the same final state as the Standard Model Higgs boson process $WH \to Wb\bar{b}$, which is one of the most promising low mass Higgs search channels at the Tevatron [4]. Essentially, all analysis tools developed for the single top search can be used for this Higgs search.

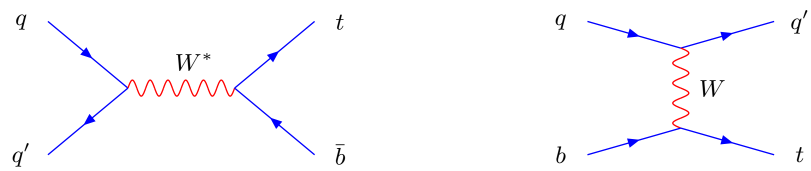

FIG. 1: Leading order Feynman diagrams for s-channel (left) and t-channel (right) single top quark production.

Finding single top quark production is challenging since it is rarely produced in comparison with other processes with the same final state like W+jets and $t\bar{t}$. The signal to background ratio of the analysis is small, typically on the order of less than S/B∼1/15. This calls for a better discrimination of signal and background events which can be achieved by using more information to characterize each event.

In this analysis, we have employed an analysis technique that attempts to make optimal use of information in the data, the Matrix Element technique. In this method, improved sensitivity is achieved by exploiting matrix element calculations for the signal and background hypothesis. Although the strategy of this method is derived from a precision measurement of the top quark mass in $t\bar{t}$ lepton + jets events [5], a novel feature of this analysis is the application of this technique to a search [6].

DATA SAMPLE & EVENT SELECTION

Our single top event selection exploits the kinematic features of the signal final state, which contains a real W boson, one or two bottom quarks, and possibly additional jets. To reduce multi-jet backgrounds, the W originating from the top quark decay is required to have decayed leptonically. We demand therefore a high-energy electron or muon ($E_T(e) > 20$ GeV, or $P_T(\mu) > 20$ GeV/c) and large missing transverse energy (MET) from the undetected neutrino MET>25 GeV. Electrons are measured in the central and in the forward calorimeter, $|\eta| < 1.6$. Exactly two or three jets with $E_T > 20$ GeV and $|\eta| < 2.8$ are required to be present in the event. A large fraction of the backgrounds is removed by demanding at least one of these two jets to be tagged as a b-quark jet by using displaced secondary vertex information from the silicon vertex detector. The secondary vertex tagging algorithm identifies tracks associated with the jet originating from a vertex displaced from the primary vertex indicative of decay particles from relatively long lived B mesons. The backgrounds surviving these selections are $t\bar{t}$, W + heavy-flavor jets, i.e. $W + b\bar{b}$, $W + c\bar{c}$, $W + c$ and diboson events WW, WZ, and ZZ. Instrumental backgrounds originate from mis-tagged W + jets events (W events with light-flavor jets, i.e. with u, d, s-quark and gluon content, misidentified as heavy-flavor jets) and from non-W + jets events (multi-jet events where one jet is erroneously identified as a lepton).

BACKGROUND ESTIMATE

Estimating the background contribution after applying the event selection to the single top candidate sample is an elaborate process. NLO cross section calculations exist for diboson and $t\bar{t}$ production, thereby making the estimation of their contribution a relatively straightforward process. The main background contributions are from $W+b\bar{b}$, $W+c\bar{c}$ and $W+c+$ jets, as well as mis-tagged W + light quark jets. We determine the W + jets normalization from the data and estimate the fraction of the candidate events with heavy-flavor jets using ALPGEN Monte Carlo samples [7]. The heavy-flavor fractions were calibrated in the b-tagged W + 1 jet sample using data distributions which are sensitive to distinguish light-flavor from heavy-flavor jets, e.g. the mass of the secondary-vertex and, more sophisticated, the output of the Neural Network jet-flavor separator. Based on these studies, the heavy flavor content was corrected by a factor $K_{HF} = 1.4 \pm 0.4$. The probability that a W + light-flavor jet is mis-tagged is parameterized using large statistics generic multi-jet data. The instrumental background contribution from non-W events is estimated using side-band data with low missing transverse energy, devoid of any signal, and we subsequently extrapolate the contribution into the signal region with large missing transverse energy, MET>25 GeV. The expected signal and background yield in the W + 2 jet and W + 3 jet sample is shown in Table I and graphically as a function of W + jet multiplicity next to the table.

Process	Number of Events in 2.2 fb^{-1}	
	W + 2 jets	W + 3 jets
s-channel	41.2 ± 5.9	13.5 ± 1.9
t-channel	62.1 ± 9.1	18.3 ± 2.7
$Wb\bar{b}$	461.6 ± 139.7	141.1 ± 42.6
$Wc\bar{c}, Wcj$	395.0 ± 121.8	108.8 ± 33.5
$Mistags$	339.8 ± 56.1	101.8 ± 16.9
$non-W$	59.5 ± 23.8	21.3 ± 8.5
Diboson	63.2 ± 6.3	21.5 ± 2.2
$Z+jets$	26.7 ± 3.9	11.0 ± 1.6
$t\bar{t}$	146.0 ± 20.9	338.7 ± 48.2
Total signal	103.3 ± 15.0	31.8 ± 4.6
Total background	1491.8 ±268.6	754.8 ± 91.3
Total prediction	1595.1 ± 269.0	776.6 ± 91.4
Observed in data	1535	712

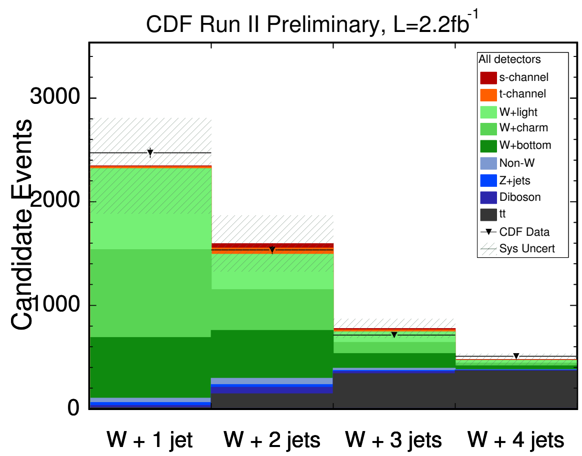

TABLE I: Number of expected single top and background events in 2.2 fb^{-1} of CDF II data passing all event selection cuts (left). Graphical representation of the predicted and observed W+jets yield (right).

ANALYSIS METHOD

This analysis is based on a Matrix Element method in order to maximize the use of information in each event. We calculate event probability densities under the signal and background hypotheses as follows. Given a set of measured variables of each event (the 4-vectors of the lepton and the two jets), we calculate the probability densities that these variables could result from a given underlying interaction (signal or background). The probability density is constructed by integrating over the parton-level differential cross-section $d\sigma$, which includes the matrix element for the process (we use MadEvent for this calculation [8]), the parton distribution functions $f(x_i)$, and the detector resolutions parameterized by transfer functions $W(y,x)$:

$$P(x) = \frac{1}{\sigma} \int d\sigma(y) dq_1 dq_2 f(x_1) f(x_2) W(y,x) \quad (1)$$

This analysis calculates probability densities for seven different underlying processes: s-channel, t-channel, $Wb\bar{b}$, $Wc\bar{c}$, Wc + jet, W+gg and $t\bar{t}$. The transfer functions $W(x,y)$ are used to include detector resolution effects. Lepton quantities and jet angles are considered to be well measured. However, jet energies are not, and their resolution is parameterized from Monte Carlo simulation to create a jet resolution transfer function. We integrate over the quark energies and over the unobserved z-component of the neutrino four momentum to create a final probability density.

We use these probability densities to construct a discriminant variable for each event. The two single-top channels are combined to form a single signal probability density. We also introduce extra non-kinematic information by using the output (b) of a neural network b-tagger which assigns a probability ($0 < b < 1$) for each b-tagged jet of originating from a b quark. The discriminant variable is then constructed as:

$$EPD = \frac{b \cdot P_{singletop}}{b \cdot P_{singletop} + b \cdot (P_{Wbb} + P_{t\bar{t}}) + (1-b) \cdot (P_{Wcc} + P_{Wcj} + P_{Wgg})} \tag{2}$$

We construct template histograms for signal and background. The combined s- and t-channel discriminants for all signal and background processes are shown in Figure 2.

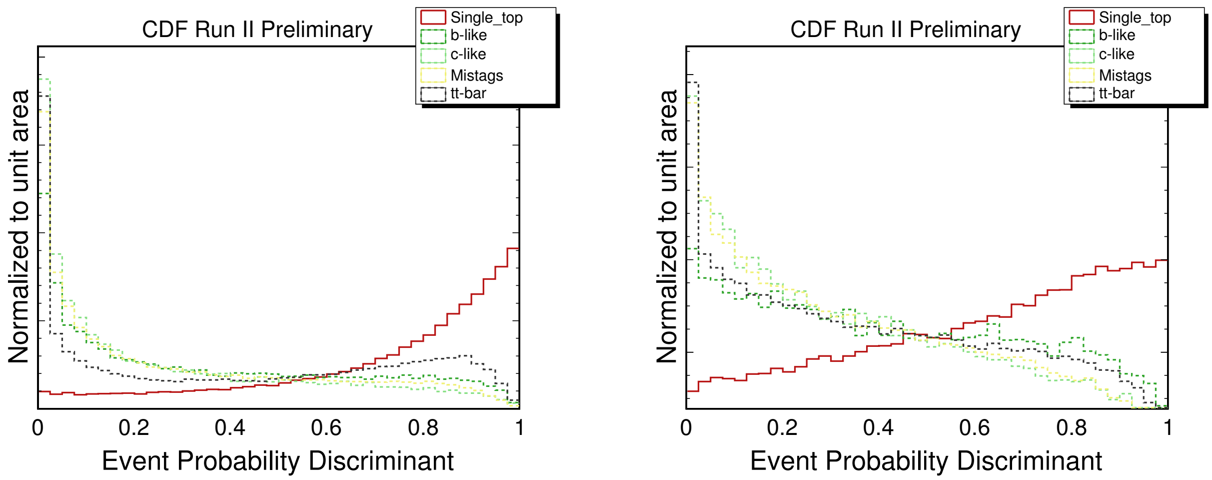

FIG. 2: Event probability discriminant distribution for single top and background processes in W + 2 jets events (left) and W + 3 jets events (right). All template histograms are normalized to unit area.

We perform a binned maximum likelihood fit to the data, in which the background templates are Gaussian constrained (within their respective uncertainties) to the predicted background yield while the signal template is free floating in the fit. The likelihood fit result determines the most probable value of the single-top cross section. Sources of systematic uncertainty are accounted for in the definition of the likelihood function shown in Equation 3.

ANALYSIS IMPROVEMENTS

Since our last update in Summer 2007 for the Lepton Photon conference, we have improved the analysis in several ways. First, we have added additional acceptance by including muon + jets events triggered via a MET + 2 jets trigger complementary to the default inclusive high p_T-muon trigger used in the analysis as shown in Figure 3. The gain in sensitivity due to the additional acceptance is roughly 9%. Moreover, we added W + 3 jets candidate events to the analysis as a separate analysis channel which resulted in a $\sim$ 1-3% gain in sensitivity. To increase the separation of signal and background in the discriminant distribution, we added Mistag (W+gg) and $t\bar{t}$ Matrix Element probability densities which resulted in an additional $\sim$ 10% gain in sensitivity.

SYSTEMATIC UNCERTAINTIES

We address systematic uncertainty from several different sources: (1) jet energy scale (2), initial state radiation (ISR) (3), final state radiation (FSR), (4) parton distribution functions, (5) the event generator, (6) the uncertainty in the event detection efficiency, (7) the uncertainty on the integrated luminosity, (8) Neural Network b-tagger uncertainty, (9) ALPGEN Monte Carlo Factorization/Renormalization scale uncertainty, (10) uncertainty on the mistag model, (11) uncertainty on the non-W model, and (12) uncertainty on the Monte Carlo Modeling. Systematic uncertainties can influence both, the expected event yield (normalization) and the shape of the discriminant distribution.

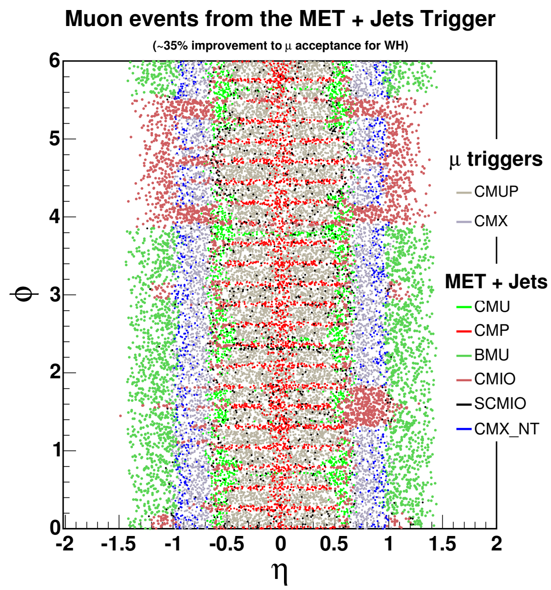

FIG. 3: This plot shows the increased acceptance of muon + jets events triggered through MET + 2jets trigger in addition to muon + jets events triggered through the default inclusive high p_T-muon trigger.

Systematic Uncertainty	Rate Uncertainty	Shape Uncertainty
Jet energy scale	0...16%	X
Initial state radiation	0...11%	X
Final state radiation	0...15%	X
Parton distribution functions	2...3%	X
Monte Carlo generator	1...5%	
Event detection efficiency	0...9%	
Luminosity	6%	
Neural Network Jet-Flavor Separator	N/A	X
Fact. Ren. Scale in Alpgen MC	N/A	X
Mistag model	N/A	X
non-W	N/A	X
MC mis-modeling	N/A	X
W+bottom normalization	30%	
W+charm normalization	30%	
Mistag normalization	17...29%	
$t\bar{t}$ normalization	23%	

TABLE II: Minimum to maximum range of observed systematic normalization variations estiamted across all different processes and analysis input channels. The X indicates that a template shape uncertainty has been evaluated for that particular nuisance parameter and has been included in the likelihood function.

Normalization uncertainties are estimated by calculating the variation in the expected event yield due to a systematic effect. The range of systematic rate and shape variations across signal and background processes are shown in Table II. Shape uncertainties are estimated by producing shifted template histograms for each process due to the systematic effect. The bin-by-bin relative variations are used as shape systematics in the likelihood function. The letter 'X' in Table II indicates that a shape systematic has been evaluated for the particular nuisance parameter and included in the likelihood function.

For all backgrounds the normalization uncertainties are represented by the uncertainty on the predicted number of background events and are incorporated in the analysis as Gaussian constraints $G(\beta_j|1,\Delta_j)$ in the likelihood function:

$$\mathcal{L}(\beta_1,\ldots,\beta_5;\delta_1,\ldots,\delta_{10}) = \underbrace{\prod_{k=1}^{B} \frac{e^{-\mu_k}\cdot\mu_k^{n_k}}{n_k!}}_{Poisson\ term} \cdot \underbrace{\prod_{j=2}^{5} G(\beta_j|1,\Delta_j)}_{Gauss\ constraints} \cdot \underbrace{\prod_{i=1}^{12} G(\delta_i,0,1)}_{Systematics} \quad (3)$$

$$where,\ \mu_k = \sum_{j=1}^{5} \beta_j \ \cdot \ \underbrace{\left\{ \prod_{i=1}^{12} \left[1 + |\delta_i| \cdot (\epsilon_{ji+} H(\delta_i) + \epsilon_{ji-} H(-\delta_i))\right] \right\}}_{Normalization\ Uncertainty} \tag{4}$$

$$\underbrace{\cdot\ \alpha_{jk}}_{Shape\ P.} \ \cdot \ \underbrace{\left\{ \prod_{i=1}^{12} \left(1 + |\delta_i| \cdot (\kappa_{jik+} H(\delta_i) + \kappa_{jik-} H(-\delta_i))\right) \right\}}_{Shape\ Uncertainty} \tag{5}$$

The systematic normalization and shape uncertainties are incorporated into the likelihood as nuisance parameters, conforming with a fully Bayesian treatment [9]. We take the correlation between normalization and shape uncertainties for a given source into account [10]. The relative strength of a systematic effect due to the source i is parameterized by the nuisance parameter δ_i in the likelihood function, constrained to a unit-width Gaussian (last term in Equation 3). The $\pm 1\sigma$ changes in the normalization of process j due to the i^{th} source of systematic uncertainty are denoted by $\epsilon_{ji}+$ and $\epsilon_{ji}-$ (see Equation part 4). The $\pm 1\sigma$ changes in bin k of the EPD templates for process j due to the i^{th} source of systematic uncertainty are quantified by κ_{jik+} and κ_{jik-} (see Equation part 5). $H(\delta_i)$ represents the Heaviside function, defined as $H(\delta_i) = 1$ for $\delta_i > 0$ and $H(\delta_i) = 0$ for $\delta_i < 0$. The Heaviside function is used to separate positive and negative systematic shifts (for which we have different normalization and shape uncertainties). The variable δ_i appears in both the term for the normalization (Equation 4) and the shape uncertainty (Equation 5), which is how correlations between both effects are taken into account. We reduce the likelihood function to the parameter of interest (the single top cross-section) by the standard Bayesian marginalizing procedure [12].

VALIDATION OF THE METHOD

In order to validate the analysis method we have performed checks in several data control samples. For example, we can evaluate the event probability discriminant in the taggable, but untagged W + 2 jets control sample, i.e. we select W + 2 jets events according to our nominal event selection and require that both jets are *not* tagged. This event selection is orthogonal to the single top candidate selection but represents a very high statistics sample of events with very similar kinematic event topology. The contribution from top is very small, <0.5%. Figure 4 shows good agreement of the data to the Monte Carlo prediction of the event probability discriminant.

We also evaluate the EPD distribution in the b-tagged W + 4 jets sample (using only the two most energetic jets), which agrees well with $t\bar{t}$ Monte Carlo as shown on the right of Figure 4.

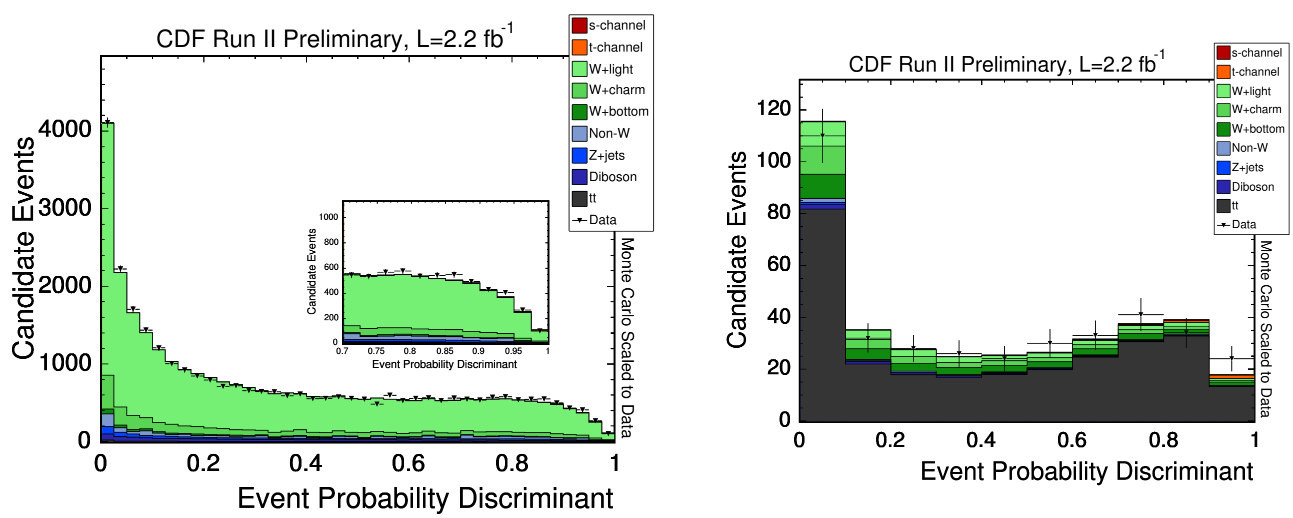

FIG. 4: Left: Evaluation of the event probability discriminant in the high statistics taggable but untagged W + 2 jets control sample. The error bars on the data points are Gaussian errors. Right: Evaluation of the event probability discriminant in the tagged W + 4 jets sample using only the two jets with the highest transverse momentum as input to the discriminant calculation. This control sample is enriched in $t\bar{t}$ events ($\sim$ 85%).

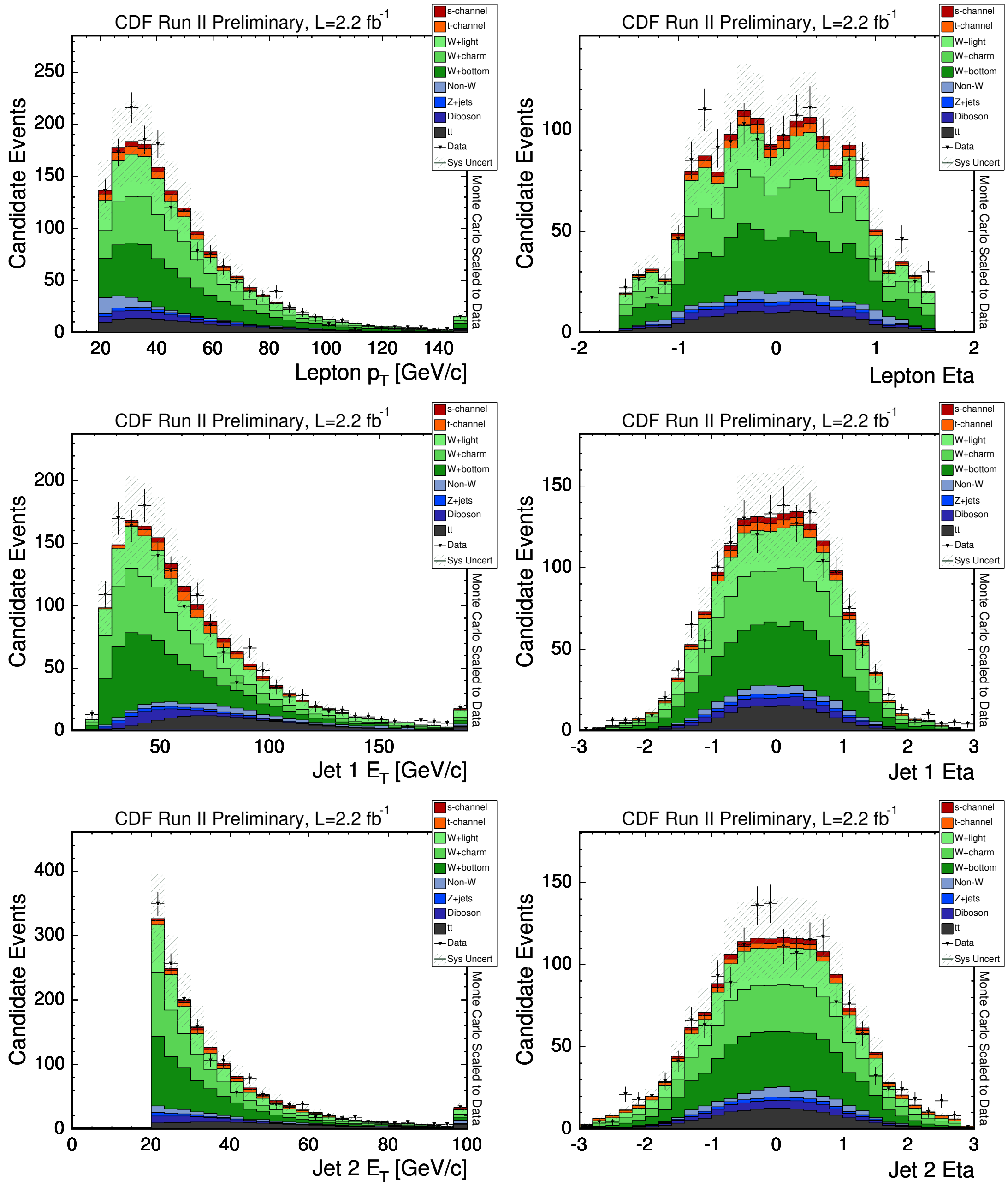

FIG. 5: Shape comparison of the event probability input variables for Monte Carlo prediction and data.

In Figure 5, we compare the distributions of the lepton and jets kinematic distributions, which serve as input to the analysis. All distributions show good agreement between data and the Monte Carlo prediction.

RESULTS

We apply the analysis to 2.2 $fb^{-1}$of CDF Run II Data. In order to extract the most probable single top content in the data we perform a maximum likelihood fit of the event probability discriminant distributions. The posterior p.d.f is obtained by using Bayes' theorem:

$$p(\beta_1|data) = \frac{\mathcal{L}^*(data|\beta_1)\pi(\beta_1)}{\int \mathcal{L}^*(data|\beta_1')\pi(\beta_1')d\beta_1'}$$

where $\mathcal{L}^*(data|\beta_1)$ is the marginalized likelihood and $\pi(\beta_1)$ is the prior p.d.f. for β_1. We adopt a flat prior, $\pi(\beta_1) = H(\beta_1)$, in this analysis, with H being the Heaviside step function.

The most probable value (MPV) corresponds to the most likely single top production cross section given the data. The uncertainty corresponds to the range of highest posterior probability density which covers 68.27%. Performing the likelihood fit with all systematic rate and shape uncertainties included in the likelihood function, we measure a single top cross section of $2.2^{+0.8}_{-0.7}$ pb. The posterior probability density is shown on the left of Figure 6. On the right of Figure 6 the EPD distribution for signal and background, normalized to the Standard Model prediction is shown and compared to the distribution in data.

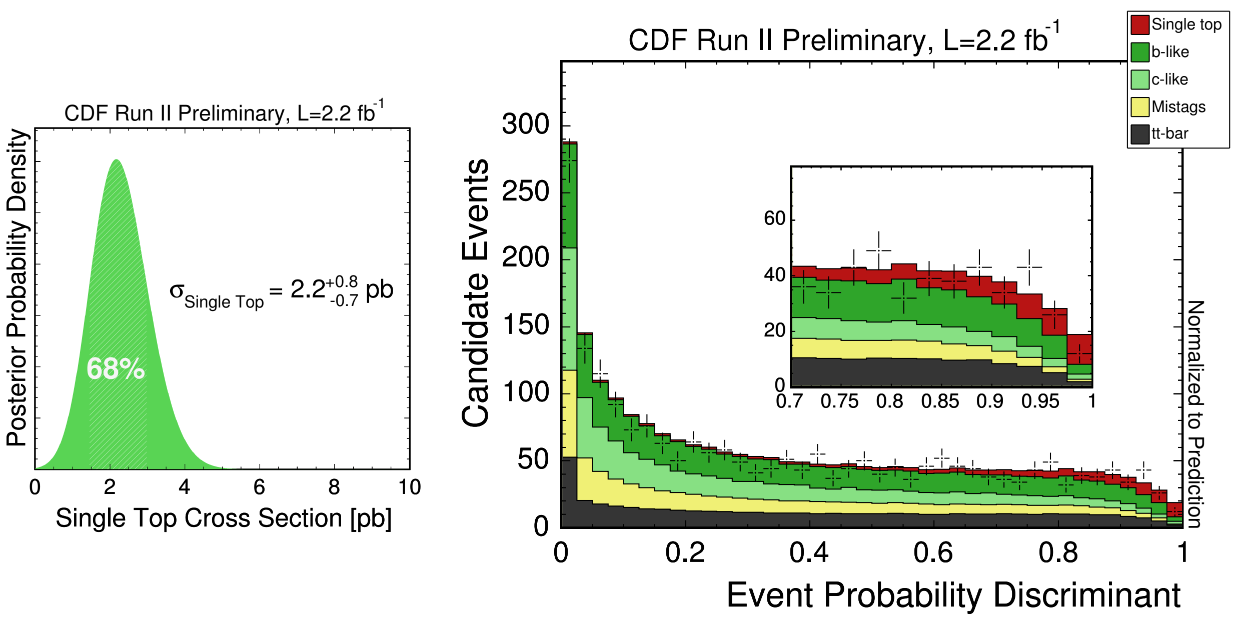

FIG. 6: Left: Cross section result using 2.2 $fb^{-1}$of CDF II data. The error band shows the 68% uncertainty (all systematics included) on the measurement. Right: Distribution of the event probability discriminant (EPD) for data and Monte Carlo, normalized to the Standard Model predicted yield. The insert shows a zoom in the signal region, EPD>0.7.

We have calculated the signal significance of this result using a standard likelihood ratio technique [11]. In this approach, pseudo-experiments are generated from background only events. We define the likelihood ratio test statistic $Q = -2\ln\frac{P(data)|(s+b)}{P(data)|(b)}$ and calculate the p-value, i.e. the probability of the background only hypothesis (b) to fluctuate to the observed result in data or higher. We estimate the expected p-value, by taking the median of the test hypothesis (s+b) distribution as the 'observed' value (dashed red line in Figure 7). We expect a p-value of 0.000003 (4.5 σ) and observe a p-value of 0.0003 (3.4 σ) in the data. All sources of systematic uncertainty are included in our statistical treatment and we consider correlation between normalization and discriminant shape changes due to sources of systematic uncertainty (e.g. the jet-energy-scale uncertainty) as described in the previous section.

SINGLE TOP SIGNAL FEATURES

We enrich our candidate sample with single top events by making increasing cuts on our event probability discriminant and look for single signal features for a few sensitive variables like $Q_{lepton} \cdot \eta_{untagged\ jet}$ which shows

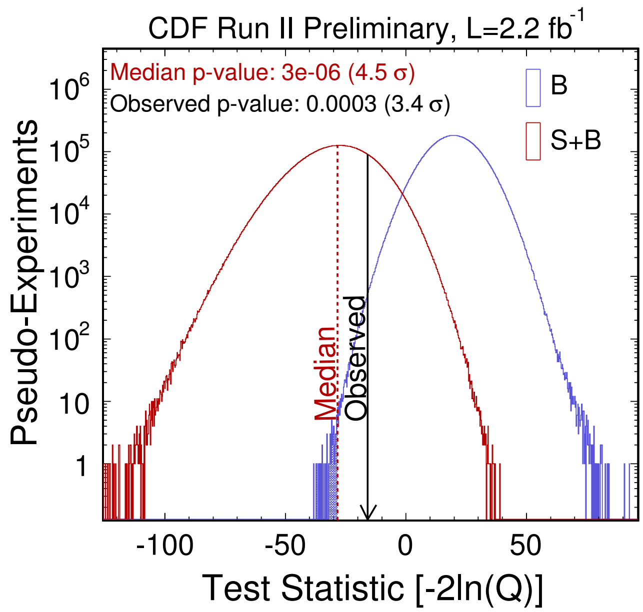

FIG. 7: Distribution of the likelihood ratio test statistic for the signal + background (s+b) and background only hypothesis. The arrow indicates the result observed in data and the red dashed line indicates the expected median result.

a distinct asymmetry for t-channel single top events and the invariant mass of the $W-b$ system, a quantity which should be close to the top quark mass. Although the uncertainties are large, there is a good shape agreement between data and the Monte Carlo prediction including single top (all plots normalized to the observed data).

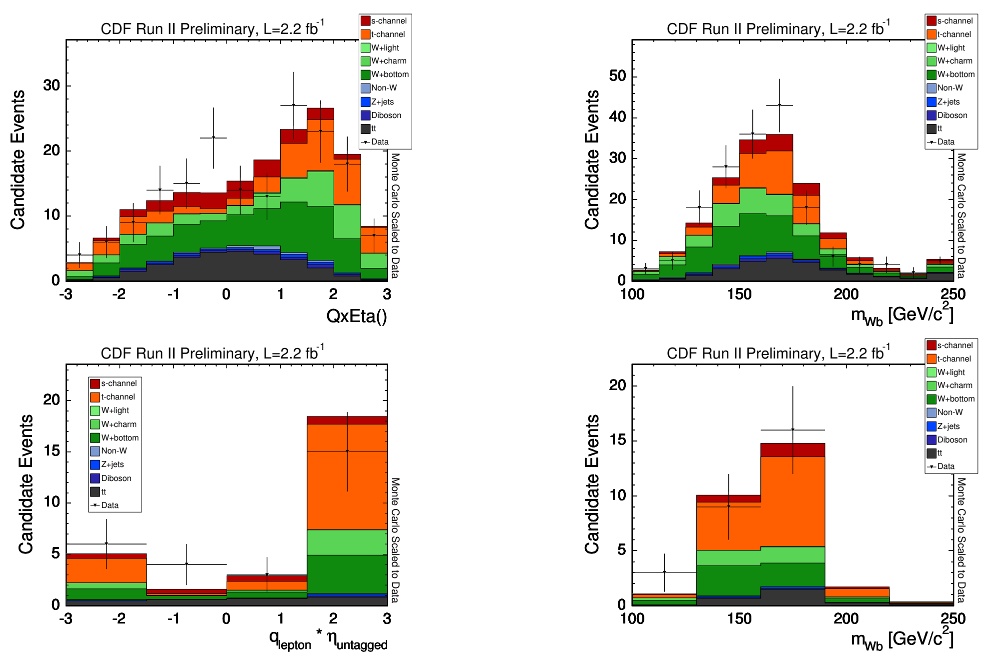

FIG. 8: Data and Monte Carlo comparison of the $Q_{lepton} \cdot \eta_{untagged\ jet}$ and m_{Wb} distributions for increasing cuts on the EPD discriminant. The top row includes events with EPD discriminant values (EPD>0.8) and the bottom row includes events with EPD discriminant values (EPD>0.95).

CONCLUSIONS

We report a measurement of electroweak single top quark production at CDF II using 2.2 $fb^{-1}$of proton-antiproton collisions recorded at the Tevatron. We employ a Matrix Element Analysis technique for this search and measure a combined s-channel and t-channel single top cross-section of $\sigma_{singletop} = 2.2^{+0.8}_{-0.7}$ pb assuming a top quark mass of 175 GeV/c^2. We use a standard likelihood ratio technique to calculate the signal significance. The observed p-value is 0.0003 (3.4 σ) and the expected (median) p-value in pseudo-experiments is 0.000003 (4.5 σ). Since the Summer 2007 analysis [13], we have improved our analysis in serveral ways which resulted in a 15-20% improved expected sensitivity.

We thank the Fermilab staff and the technical staffs of the participating institutions for their vital contributions. This work was supported by the U.S. Department of Energy and National Science Foundation and the A. v. Humboldt Foundation, Germany.

[1] B.W. Harris *et. al.*, *Phys. Rev.* D **66**, 054024 (2002)
Z. Sullivan *Phys. Rev.* D **70**, 114012 (2004).
[2] G. Mahlon, *Phys. Rev.* D **55**, 7249 (1997), hep-ph/0011349.
[3] T. M. P. Tait and C.-P. Yuan, *Phys. Rev.* D **63**, 014018 (2002).
[4] CDF and DØ Collaborations, FERMILAB-PUB-03/320-E (2003).
[5] D0 Collaboration, V.M. Abazov, et, al., Nature 429 (2004); D0 Collaboration, V.M. Abazov, et. al., Phys. Lett. **B** 617 (2005); F. Canelli, PhD thesis, University of Rochester (2003);
[6] B. Stelzer, PhD Thesis, University of Toronto, FERMILAB-THESIS-2005-79.
`http://www.slac.stanford.edu/spires/find/hep/www?r=umi-nr07812`
[7] F. Caravaglios *et. al.*, M. L. Mangano *et. al.*, JHEP 0307:001 (2003).
[8] F. Maltoni, T. Stelzer, JHEP 0302:027, (2003).
[9] L. Demortier, *Bayesian treatments of Systematic Uncertainties*, Proceedings of Advanced Statistical Techniques in Particle Physics, Grey College, Durham, 18 - 22 March 2002
[10] C. Ciobanu, T. Junk, T. Müller, P. Savard, B. Stelzer, W. Wagner, T. Walter, *Likelihood Function for Single Top Search with 162* pb^{-1}, CDF Note 7106 (2004)
[11] L. Read, J.Phys G **28**, 2693 (2002) and T. Junk, Nucl. Instum. Meth. **434**, 435 (1999). See also P. Bock *et al.* (The LEP Collaborations), CERN-EP-98-046 (1998) and CERN-EP-2000-055 (2000).
[12] Particle Data Group, *The Review of Particle Physics*, http://pdg.lbl.gov (2004)
[13] F. Canelli, P. Dong, B. Stelzer, R. Wallny *Evidence for Single Top Quark Production in L=1.51* fb^{-1} *of CDF Run II Data using the Matrix Element Technique*
`http://www-cdf.fnal.gov/physics/new/top/2007/singletop/ME1_5fb/`