CDF note 9460

# Search For Single Top Quark Production Optimized for s-channel Production with 2.7 fb$^{-1}$ Data


The CDF Collaboration

*URL http://www-cdf.fnal.gov*



(Dated: July 29, 2008)



This note describe s-channel-optimized search for single top quark production using 2.7 fb$^{-1}$ of data accumulated with the CDF detector. This note reports an update result of the s-channel optimized search for the singletop combination. We are using events with one high-$P_T$ lepton, large missing $E_T$ and two b-quark-identified jets using secondary vertex tagger(SecVtx). In this analysis we have developed a kinematics fitter and a likelihood-based separator between signal and background. We are still working to improve sensitivity using the same luminosity data. So we didn't calculate s-channel sensitivity and didn't set a limit.


*Preliminary Results for Winter 2008 Conferences*

## I. INTRODUCTION

Top quarks are predominantly produced in pairs via the strong interaction, in $\bar{p}p$ collisions at Tevatron with $\sqrt{s}$ = 1.96 TeV. Top quark has a weak isospin 1/2 composing a weak isospin doublet with a bottom quark. It predict not only top quark pair production via strong interaction but also singly produced together with a bottom quark via weak interaction. However, finding single top quark production is challenging since it is rarely produced($\sigma_{\mathrm{singletop}}$=2.9 pb [2]) against background processes with the same final state like $W$+jets and $t\bar{t}$. Since the signal to background ratio is small, The mesurement of electroweak single top production probes the $W-t-b$ vertex, provides a direct determination of Cabbibo-Kobayashi-Maskawa (CKM) matrix element $|V_{tb}|$ and offers a source of almost 100% polarized top quarks [3] The CDF detector is described in detail in [4].

## II. DATA SAMPLE & EVENT SELECTION

This analysis is based on an integrated luminosity of 2.7 fb$^{-1}$ collected with the CDFII detector between March 2002 and April 2008. Candidate events for this analysis are selected by requiring a $W$ + 2jets event topology where $W$ decays leptonically, $W \to e\nu_e$ or $W \to \mu\nu_\mu$. The data are collected with an inclusive lepton trigger that requires an electron or muon with $E_T$ >18 GeV ($P_T$ >18 GeV/c for the muon). From this inclusive lepton dataset we select events offline with a reconstructed isolated electron $E_T$ (muon $P_T$) greater than 20 GeV, missing $E_T$ >25 GeV and 2 jets with $E_T$ >20 GeV and $|\eta|$ <2.8. Both of the two jets should be identified as a b-jet using secondary-vertex information(SecVtx) [5] The background surviving these selection are $t\bar{t}$, $W$ + heavy-flavorjets, 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 nonW + jets events (multi-jet events where one jet is erroneously identified as a lepton).

## III. 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 sample [7]. The heavy-flavor fraction were calibrated in the $b$-tagged $W$+1jet sample using data distributions which are sensitive to distinguish light-flavor from heavy-flavor jets, e.g. the mass of the secondary-vartex 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 $K_{HF}$=1.44±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 are 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 $W$+2jet and $W$+3jet sample is shown in Table I.

## IV. ANALYSIS METHOD

### A. Kinematics Fitter

There are two signigicant ambiguity in calculating the reconstructed $M_{lnb}$. One comes from the uncertainty on assigning a b-jet correctly to the $b$-jet from top quark decay. The other is the neutrino energy. Transverse energy of the neutorino can be measured indirectly as the missing transverse energy, but nutrino momentum $z$-component $P_z(\nu)$ can not be measured. The purpose of the kinematics fitter is to evaluate the neutrino momentum and find the right b-jet assignment. Neutrino momentum $z$-component can be caliculated as the solution in a quadratic equation using a $W$ mass constraint and the missing transverse energy. Since we have two solutions at most, we constructed two types of likelihood-based kinematics fitter for 2 solutions case and the other case. We minimized the negative log likelihood value by varying three parameters, $P_Z(\nu Z)$, $\not{E}_x$ and $\not{E}_y$. MINUIT was used this minimization. We construct a likelihood for b-jet from top decay as follows. $P_T(b)$, $M_{lb}$, $\cos\theta_b \times Q_{lep}$, where $P_T(b)$ is b-jet $P_T$, $M_{lb}$ is an invariant mass of lepton + b-jet and $\cos\theta_b$, $Q_{lep}$ is a scattering angle of b-jet in initial quarks ($q\bar{q}$) rest frame and $Q_{lep}$ is a lepton charge. We pick up a b-jet which have larger likelihood as the b-jet from top decay.

| Process | 2jets | 3jets |
|---|---|---|
| s-channel | 8.9 ± 1.5 | 3.1 ± 0.5 |
| t-channel | 1.7 ± 0.3 | 2.4 ± 0.4 |
| $WW$ | 0.3 ± 0.1 | 0.3 ± 0.0 |
| $WZ$ | 3.4 ± 0.4 | 1.0 ± 0.1 |
| $ZZ$ | 0.1 ± 0.0 | 0.0 ± 0.0 |
| $t\bar{t}$(lep+jets) | 14.5 ± 2.4 | 58.3 ± 9.5 |
| $t\bar{t}$(di-lepton) | 9.9 ± 1.6 | 9.5 ± 1.6 |
| $Z$+jets | 1.0 ± 0.2 | 0.6 ± 0.1 |
| $Wbb$ | 57.7 ± 18.0 | 19.8 ± 6.2 |
| $Wcc$ | 2.8 ± 0.9 | 1.7 ± 0.6 |
| $Wcj$ | 2.4 ± 0.8 | 0.9 ± 0.3 |
| Mistags | 1.6 ± 0.4 | 1.2 ± 0.3 |
| Non-$W$ | 1.7 ± 0.7 | 0.2 ± 0.1 |
| Total Prediction | 105.9 ± 20.3 | 99.0 ± 13.0 |
| Observed | 102.0 ± 0.0 | 117.0 ± 0.0 |

TABLE I: The number of expected single top and background events in 2.7fb$^{-1}$ of CDF data passing all event selection for $W$+2b-taggaed jets.

### B. Likelihood Function Technique

In order to discover single top production, and to measure its rate with a highest precision, we must take advantage of as many differences between the signal and the background as possible. To this end, a variety of quantities which can be computed from the reconstructed event variables have been investigated for their ability to separate the signal from the background. No single variable encodes all conceivable separation, and so a likelihood is proposed to combine several variables together into one discriminant to compute the cross section limits or to discover the signal.

The Likelihood $\mathcal{L}$ is constructed by first forming template histograms of each variable, separately for the signal and for the several background, denoted $P_i^m$ for variable $i$ for the background event class $m$ and signal. Since about 75 % of the background are $W$+HF and $t\bar{t}$, we used only $W$+HF and $t\bar{t}$ for the background template as the first try. These two background classes are separately used to construct the likelihood. These histograms are normalized so that area is 1. The likelihood is defined as :

$$\mathcal{L}^{signal} = \frac{\prod_{i=1}^{n_{val}} P_i^{\text{signal}}}{\prod_{i=1}^{n_{val}} P_i^{\text{signal}} + \sum_{m=1}^{n_{bkg}} \prod_{i=1}^{n_{val}} P_i^m \times \alpha^m} \tag{1}$$

where,

$$P_i^{\text{process}} = \frac{f_i^{\text{process}}}{f_i^{\text{signal}} + \sum_{m=0}^{n_{bkg}} f_i^{\text{m}}}, \quad \alpha^m = \frac{N_{\text{eve}}^m}{\sum_{m=1}^{n_{bkg}} N_{\text{eve}}^m} \tag{2}$$

In order to obtain the likelihood we used the following six valiables:

- $M_{bb}$ : di-jet invariant mass
- $Pt_{bb}$ : transverse momentum for bb system
- $M_{l\nu b}$ : invariant mass of lepton, neutorino and b-jet from top quark
- $Pt(j1)$ : leading jet transverse momentum
- $Ht$ : transverse momentum scalar sum of lepton, neutrino and two jets
- $\not\!\!E_T$ : Missing transverse energy
- $\mathcal{L}_{\nu\text{sol}} \times \mathcal{L}_{bb\text{sol}}$: kinematics fitter likelihood output times the likelihood for b-jet from top decay.

## V. 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 function, (5) the event generator, (6) the uncertainty in the event detection efficiency, (7)the uncertainty on the integrated luminosity, (8) ALPGEN Monte Carlo Factorization/Renormalization scale uncertainty, (9) uncertainty on the Monte Carlo Modeling. (10) uncertainty on the mistag model, Systematic uncertainties can influence both, the expected event yield(normalization) and the shape of the discriminant distribution. Mormalization uncertainties are estimated by calculating the variation in expected event yield due to a systematic effect. The range of systematic rate and shape variation across signal and background processes are shown in Table II. Shape uncertainties are estimated by producing shifted template histogram for each process du to the systematic effect. The bin-by-bin relative variations are used as shape systematics in the likelihood function. The letter "$\surd$" in Table II indicates that a shape systematic uncertainty 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 contributes $G(\beta|1,\Delta_j)$ in the likelihood function :

$$\mathcal{L}(\beta_1,...,\beta_5;\delta_1,...,\delta_{10}) = \prod_{k=1}^{B} \frac{e^{-\mu_k}\cdot\mu_k^{n_k}}{n_k!}\cdot\prod_{j=2}^{5} G(\beta_j|1,\Delta_j)\cdot\prod_{j=1}^{10} G(\delta_j,0,1) \tag{3}$$

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

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

The systematic normalization and shape uncertainties are incorporated into the likelihood as nuisance parameters, conforming with a fully Bayesian treatment [8]. We take the correlation between normalization and shape uncertainties for a given source into account [9]. The relative strength of a systematic effect due to the source $i$ is parameterized by the nuisance parameter $\delta_i$ in the likelihood function, constrained to a unit-width Gausian(last term in Equation 3). The $\pm1\sigma$ 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 $\pm1\sigma$ changes in bin $\kappa$ of discriminant templates for process $j$ due to the $i^{th}$ source of systematic uncertainty are quantified by $\kappa_{jik+}$ and $\kappa_{jik-}$ (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 $\delta_i$ appears in both the term for the normalization (Equation 4) and the shape uncertinaty (Equation 5), which is how correlations between both effects are taken into account. We reduce the likelihood function to the parameter of interest (single top cross-section) by the standard Bayesian marginalizing procedure [11]

## VI. RESULTS

We apply the analysis to 2.7 fb$^{-1}$ of CDF Run II Data. The distribution of likelihood discriminant is shown in Figure 1. To evaluate singletop cross section, we combined our analysis and another Likelihood analysis which is optimized to t-channel [10]. We fit single top s-channel cross section and t-channel cross section simultaneously. For the category which have $W$+2jet and both of 2 jets identified as b-jet, our $L_s$ discriminant are used, and for the others category, $L_t$ discriminant which is discribed in [10] are used. Figure 2 shows best fit cross section and it's $1\sigma$ and $2\sigma$ uncertainty as a two dimensional distribution. Best fit cross sections for the s-channel and t-channel are 0.9 pb and 1.2 pb respectively.

## VII. CONCLUSIONS

We have searced s-channel electroweak single top quark production using double tag method with likelihood based signal background separator tequnique. We have analyzed 2.7 fb$^{-1}$ of CDF Run II data. Best fit cross sections for the s-channel and t-channel are 0.9 pb and 1.2 pb respectively.

CDF RUN II Preliminary, L=2.7 fb$^{-1}$

| Source | Rate Uncertainties for expected events | shape variation |
|---|---|---|
| Jet Energy Scale | 4.7 % | $\surd$ |
| Initial State Radiation | 1.7 % | $\surd$ |
| Final State Radiation | 1.3 % | $\surd$ |
| Parton Distribution Fanction | 0.5 % | $\surd$ |
| MC generator | 0.5 % | |
| Event Selection Acceptance | 1.5 % | |
| Integrated Luminosity | 6.0 % | |
| Theoritical Cross Section | 3.1 % | |
| Top Mass | 2.3 % | $\surd$ |
| Data Based Background Estimation | 23 % | |
| Mistag Modeling | | $\surd$ |
| QCD Modeling | | $\surd$ |
| $Q^2$ scale in Alpgen MC | | $\surd$ |
| MC Mismodeling | | $\surd$ |

TABLE II: Systematic uncertainties for the total prediction normalization estimated across all different processes and analysis input channels. the $\surd$ indicates that a template shape uncertainty has been evaluated for that particular nuisance parameter and has been included in the likelihood function.

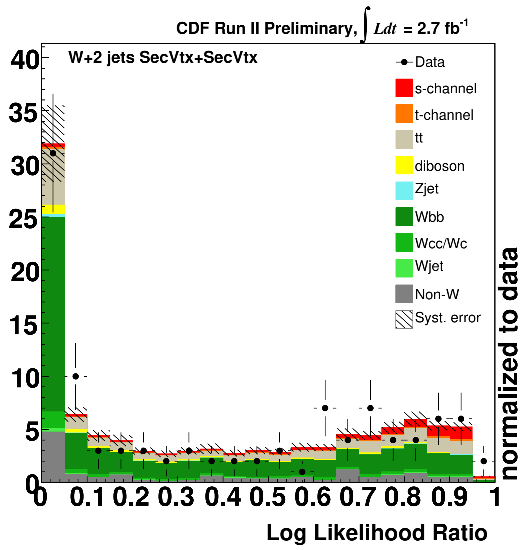


FIG. 1: Distribution of the discriminant for data and Monte Carlo.

## Acknowledgments

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; the Italian Istituto Nazionale di Fisica Nucleare; the Ministry of Education, Culture, Sports, Science and Technology of Japan; the Natural Sciences and Engineering Research Council of Canada; the National Science Council of the Republic of China; the Swiss National Science Foundation; the A.P. Sloan Foundation; the Bundesministerium fuer Bildung und Forschung, Germany; the Korean Science and Engineering Foundation and the Korean Research Foundation; the Particle Physics and Astronomy Research Council and the Royal Society, UK; the Russian Foundation for Basic Research; the Comision Interministerial de Ciencia y Tecnologia, Spain; and in part by the European Community's

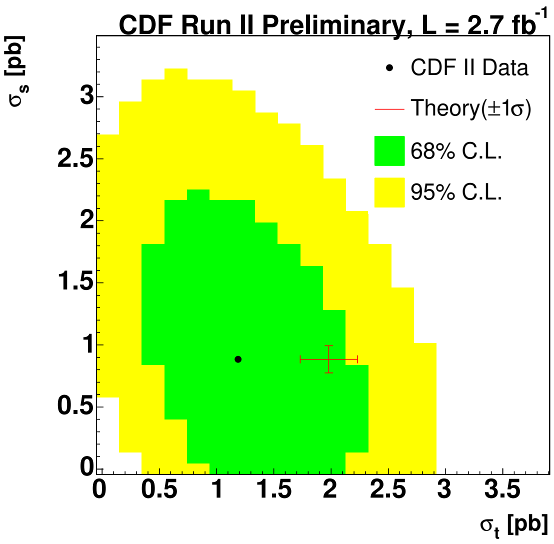


FIG. 2: best fit cross section(black dot) and it's $1\sigma$(green) and $2\sigma$(yellow) uncertainty. Red mark shows theoretical cross section with errors.

Human Potential Programme under contract HPRN-CT-20002, Probe for New Physics.


[1] K.Nakamura *et al.*,*cdf public note* 9341(2008)
[2] B.W.Harris *et al.*,*Phys. Rev.* D **66**, 054024 (2002).
Z.Sullivan, *Phys. Rev.* D **70**, 114012 (2004).
[3] G.Mahlon,*Phys. Rev.* D **55**, 7249 (1997), hep-ph/0011349.
[4] F. Abe, et al., Nucl. Instrum. Methods Phys. Res. A **271**, 387 (1988); D. Amidei, et al., Nucl. Instum. Methods Phys. Res. A **350**, 73 (1994); F. Abe, et al., Phys. Rev. D **52**, 4784 (1995); P. Azzi, et al., Nucl. Instrum. Methods Phys. Res. A **360**, 137 (1995); The CDFII Detector Technical Design Report, Fermilab-Pub-96/390-E
[5] D.Acosta *et al.*(CDF Collaboration),*Phys. Rev.* D **71**, 052003 (2005).
[6] D.Buskulic *et al.*(ALEPH Collaboration),*Phys. Lett.* B **313**, 535 (1993);
D.Abe *et al.*(CDF Collaboration),*Phys. Rev.* D **53**, 1051 (1996);
A.Affolder *et al.*(CDF Collaboration),*Phys. Rev.* D **64**, 032002 (2001), *Phys. Rev.* D **67**, 119901 (2003)
[7] F.Caravaglios *et al.*, M. L. Mangano *et al.*, JHEP 0307:001 (2003)
[8] L. Demortier, *Bayesian treatments of Systematic Uncertainties*, Proceeding of Advanced Statistical Techniques in Particle Physics, Grey College, Durham, 18-22 March 2002
[9] 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)
[10] The CDF Collaboration, *Multivariate Likelihood Function Measurement of Single-Top-Quark Production with 2.7 fb*$^{-1}$, CDF Note 9451 (2004)
[11] Particle Data Group, *The Review of Particle Physics*, http://pdg.lbl.gov (2004)
[12] 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).