Measurement of the $t\bar{t}$ Production Cross Section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV using Lepton + Jets Events with Soft Electron b-Tagging

The CDF Collaboration
URL http://www-cdf.fnal.gov
(Dated: May 23, 2008)

We present a measurement of the $t\bar{t}$ production cross section using soft electron tags in the lepton+jets decay channel using 1.7 fb^{-1} of CDF data. A soft electron tagger ($\texttt{SLT}_\texttt{e}$) is used to identify electrons resulting from the semileptonic decay of heavy flavor. Requiring at least one soft electron tag, we measure a cross section of 7.8 $\pm$ 2.4 (stat) $\pm$ 1.5 (syst) $\pm$ 0.5 (lumi) pb. This is the first measurement of the $t\bar{t}$ cross section in Run II with a soft electron tagger.

I. INTRODUCTION

This note presents the first measurement of the $t\bar{t}$ production cross section with soft electron tags at CDF in Run II. Approximately 1.7 fb^{-1} of $p\overline{p}$ collisions at $\sqrt{s} = 1.96$ TeV has been collected with the CDF detector at the Fermilab Tevatron for this result (see Reference [2] for a detailed explanation of the CDF detector).

The electron tagger is designed to identify central electrons with P_T as low as 2 GeV/c embedded in bottom and charm jets. This method for tagging heavy flavor (HF) is possible because of the high semileptonic branching fraction – approximately 10% per electron and per muon – for bottom and charm. Additionally, soft leptons from b-quarks can arise from either direct ($B \to \ell\nu X$) or cascade ($B \to D \to \ell\nu X$) decays, effectively doubling the tagger's acceptance for bottom.

Assuming $BF(t \to Wb) \sim .999$, we measure $\sigma_{t\bar{t}}$ in the lepton+jets channel, where one W decays hadronically and the other decays leptonically. We measure a cross section of 7.8 ± 2.4 (stat) ± 1.5 (syst) ± 0.5 (lumi) pb. This measurement assumes that the top mass is 175 GeV/c^2.

II. DATA SAMPLE & EVENT SELECTION

This analysis is based on an integrated luminosity of 1.7 fb^{-1} collected with the CDFII detector between March 2002 and May 2007. The data is 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 $E_T > 30$ GeV and at least 3 jets with $E_T > 20$ GeV. Events with only 1 or 2 jets with $E_T > 20$ GeV are considered as a control sample.

The data selected above is dominated by real W bosons (which subsequently decay to a lepton and neutrino) with associated light flavor production. To enhance the $t\bar{t}$ content of the sample, we require that at least one jet in the event contain an $\texttt{SLT}_\texttt{e}$ tag. We also suppress the backgrounds kinematically by requiring that the scalar sum of all transverse energy in the event, called H_T, exceeds 250 GeV (this requirement is not enforced in the control sample). For convenience, we define 'pretag' to refer to the sample after the lepton+jet selection but before $\texttt{SLT}_\texttt{e}$ tagging. We find 2196 pretag events in the signal region and 121 $\texttt{SLT}_\texttt{e}$ tagged events. 5 events have two $\texttt{SLT}_\texttt{e}$ tags.

A. Soft Electron Tags

1. *Tagger Algorithm*

We tag electrons embedded in jets by extrapolating tracks with $P_T > 2$ GeV/c and a beamline corrected impact parameter $|d_0'| < 0.3$ cm to the central wire and strip chambers located approximately six radiation lengths within the electromagnetic calorimeter (see [3] for details on this sub-detector). Candidate tracks must be fiducial to the wire and strip chambers and be away from the edges of chambers. Tracks must also be within a cone of $\Delta R \equiv \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.4$ of a jet axis.

A calorimeter cluster is then formed with the closest two calorimeter towers, adjacent in η-space, to the track extrapolation. We require that the calorimeter cluster has an energy deposition consistent with the presence of an electron: $0.6 < E_{EM}/P < 2.5$ and $E_{Had}/E_{EM} < 0.2$ where E_{EM} and E_{Had} are the total electromagnetic and hadronic energy in the cluster, respectively, and P is the momentum of the track. Because the calorimeter is coarsely segmented, these variables have a strong environmental dependence. It is also for this reason that we typically refer to the P_T rather than the E_T of the $\texttt{SLT}_\texttt{e}$ tag, since the latter is obfuscated by the environment.

Another cluster is formed about the extrapolated track in the wire and strip chambers, consisting of 2-7 wires and 2-7 strips. Clusters with fewer wires or strips are rejected to suppress minimum ionizing particles. The wire and strip chambers measure the transverse profile of the electromagnetic shower in orthogonal directions. The fine segmentation of the chambers mitigates the effects of the local jet environment. Two types of variables are formed from the clusters. One is a χ^2 comparing the shape of the transverse profile to that of a test beam of 10 GeV electrons. The other is the relative distance in cm between the extrapolated track and the center of energy of the cluster. These four variables (χ^2_{wire}, χ^2_{strip}, Δ_{wire}, Δ_{strip}) are combined into a likelihood. The likelihood is selected at three different operating points (called 'tag levels') to allow for future optimization. For this analysis, we find that the loosest operating point, that is tag level 1, results in the smallest expected uncertainty.

2. Performance

To determine the performance of the SLT$_e$, an independent sample of electrons from photon conversions is constructed. Figure 1 shows the per track tagging efficiency of the SLT$_e$ for conversion electrons as a function of track P_T. This figure overstates the average tagging efficiency for HF electrons, however, because conversion electrons are typically more isolated. The environmental dependence of the tagging efficiency is primarily the result of the coarse calorimeter segmentation. The average tagging efficiency for HF electrons from b-quark decay is approximately 40% per fiducial track using the loosest SLT$_e$ operating point.

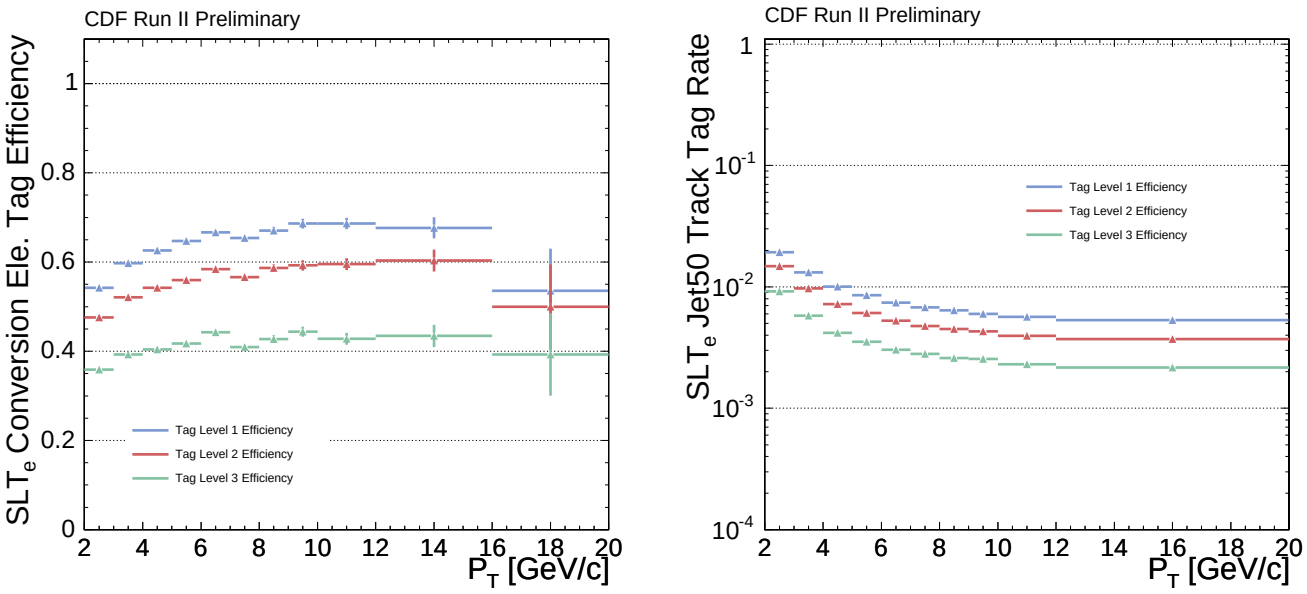

FIG. 1: Left: Per track SLT$_e$ tagging efficiency for electrons from conversions as a function of the track P_T. Right: Per track SLT$_e$ tag rate for tracks from the jet50 sample as a function of the track P_T. The tag rates for the three operating points are shown.

Similarly, an admixture of pions, kaons, and protons is constructed out of tracks from generic jets. Generic jet events are triggered on a jet with thresholds of 20, 50, 70, and 100 GeV. Figure 1 also shows the per track tagging rate for tracks from the jet50 sample. This figure also overstates the average tagging efficiency for non-electrons because of the presence of real electrons from photon conversions, heavy flavor, Dalitz decay of π^0, and other sources. We estimate that real electrons enhance the SLT$_e$ tag rate roughly 40% in this sample, so that the average tag rate for fake electrons in $t\bar{t}$ events is approximately 0.5% per fiducial track using the loosest SLT$_e$ operating point.

3. SLT$_e$ in Monte Carlo Simulation

When tagging soft electrons in simulation, we do not preserve the full SLT$_e$ algorithm. The track extrapolation and calorimeter clustering components are maintained to properly model the local environmental dependence. A 2.5% relative systematic uncertainty is applied to address the observed differences between comparable samples in simulation and data.

The wire and strip chamber variables, however, only have a loose dependence on the track's environment. Therefore, we parameterize the dependence of the tagging efficiency – after an 'electron-like' calorimeter cluster has been selected – as a function of the track's P_T, η, and isolation. This parametrization, called the '*tag matrix*,' is built from conversion electron data and is applied to electrons embedded in jets in simulation. A P_T dependent systematic is applied to the *tag matrix* to account for differences between the predicted tagging efficiency and the measured tagging efficiency in a sample of high P_T electrons from Z boson decay. A small correction is also applied to the *tag matrix* to account for the fact that the electrons from the conversion sample are rarely embedded in jets.

A similar parametrization, called the '*fake matrix*,' is made as a function of P_T, η, and isolation for non-electron tracks. The *fake matrix* is built from the generic jet samples jet20, jet50, jet70, and jet100, and it is corrected for the

real electron content. To verify that the *fake matrix* is well-behaved and that the electron contamination estimation is correct, the matrix is applied to a sample of charged pions from K_S decay, and the predicted and measured tag rates are shown to agree. Two sources of systematics enter into the *fake matrix* prediction: a 4% relative uncertainty from the agreement between the four generic jet samples, and a 6% relatively uncertainty from the estimate of the sample's electron contamination.

4. Photon Conversion Rejection

A heretofore neglected background of this analysis is electrons from photon conversions. The dominant source of conversion electrons embedded in jets is from neutral pion decay. Without removal, this background overwhelms the low P_T electron spectrum: in $t\bar{t}$ events prior to tagging, approximately three times as many candidate electron tracks are due to photon conversions than from semileptonic decay of heavy flavor. To suppress this background, we attempt to locate the partner electron track from the photon conversion using geometric selection techniques. Because of the low P_T threshold for candidate $\texttt{SLT}_\texttt{e}$ tracks and the asymmetric energy sharing between electrons from conversions, the partner track often has too little P_T to be reconstructed.

To further reduce this background, we rely on the material interaction behavior of conversions. Tracks are extrapolated through the silicon detector, and we count layers where silicon hits are expected within a band of the extrapolated path but not found. If more than three of the double-sided layers of silicon expect hits on each side, but none are found, then we reject that candidate track as a conversion even without a reconstructed partner track. The combined effort of this conversion filter algorithm results in an 70% rejection efficiency while suffering only a 7% over-efficiency for non-conversion tracks in the jet environment.

Due to differences in tracking and the material description, the conversion filter efficiency is different in data than in Monte Carlo simulation. Therefore, we measure a data-MC scale factor in the four generic jet datasets and the corresponding PYTHIA [4] di-jet MC filtered on matching jet thresholds. We find a combined efficiency SF of 0.93 ± 0.03. We also measure a combined over-efficiency SF of 1.0 ± 0.3.

B. Acceptance and Efficiency

The total acceptance for $t\bar{t}$ is measured in a combination of data and Monte Carlo. The geometric times kinematic acceptance of the basic lepton+jets event selection is measured using a PYTHIA $t\bar{t}$ sample with a $M_{top} = 175$ GeV/c^2. The efficiency for identifying the isolated, high P_T lepton is scaled to the value measured in the data using the unbiased leg in Z boson decays.

We measure the $\texttt{SLT}_\texttt{e}$ tagging efficiency for top events by applying the *tag matrix* and *fake matrix* to all the candidate tracks with an electron-like calorimeter cluster described in Section II A 1. We interpret the result of the matrices as the probability to tag that track. Tracks that are rejected by the conversion filter have their probabilities adjusted by the efficiency and over-efficiency scale factors. Table I shows the expected number of events before and after $\texttt{SLT}_\texttt{e}$ tagging. Approximately 25% of the tags result from non-electron tracks and 10% result from conversion electron tracks.

$t\bar{t}$ Expectation ($\sigma_{t\bar{t}} = 6.7\ pb$, $\int \mathcal{L} = 1.7\ fb^{-1}$)					
	1 Jet	2 Jets	3 Jets	4 Jets	$\geq$ 5 Jets
Pretag Expectation	34.0 ± 2.1	180.5 ± 11.2	295.1 ± 18.2	313.3 ± 19.4	108.3 ± 6.7
Tag Expectation	1.2 ± 0.1	11.4 ± 1.0	22.5 ± 2.0	26.2 ± 2.1	10.6 ± 0.9

TABLE I: Pretag and tag $t\bar{t}$ expectation assuming a top cross section of 6.7 *pb*. 'Pretag' refers to events after the lepton+jets event selection but before $\texttt{SLT}_\texttt{e}$ tagging. Uncertainties shown include the uncertainties associated with $\texttt{SLT}_\texttt{e}$ tagging, the acceptance calculation, and the luminosity.

III. BACKGROUNDS

We consider three different classes of backgrounds to top production. The first class includes those backgrounds which have a small enough contribution to the measurement that even large uncertainties on their production cross section can be tolerated. These backgrounds are di-boson (WW, WZ, ZZ), single top production, Z boson production

Systematic	Relative Uncertainty (%)
Jet Energy Scale	8.4
QCD Fit	5.1
K-Factor	3.0
Herwig	2.2
Lepton ID	1.7
Background Cross Section	0.9
PDF	0.9
FSR	0.6
ISR	0.5
Conversion (over-)efficiency	10.7
Calorimeter Modeling	7.4
Tag Matrix	6.6
Fake Matrix	6.0
Jet Environment correction	5.4
Total tagger syst.	16.5
Total	19.9

TABLE II: Summary of systematic uncertainties.

associated with jets, and Drell-Yan to di-lepton. Their estimation with Monte Carlo follows in the same manner from the $t\bar{t}$ signal estimation.

The second class of backgrounds is due to QCD events where a W signature is faked. The lepton can be produced through mis-reconstruction or via semileptonic decay of $b\bar{b}$ events. This background is estimated using a data-driven method: we use missing E_T templates from $b\bar{b}$ Monte Carlo and electron-like objects which fail the final electron selection to model the sideband, and extrapolate them into the signal region by fitting them with missing E_T templates from the other backgrounds and $t\bar{t}$.

The third class of backgrounds is the production of a W boson associated with jets. This is the dominant background to top production. The theoretical prediction of such a process is difficult because the next-to-leading-order (NLO) corrections are large compared to the leading-order (LO) calculation. We use a combined data and ALPGEN[6] plus PYTHIA Monte Carlo estimate for this background.

Our W+jet measurement technique assumes that the number of observed events in the pretag sample not already accounted for by the first two classes of backgrounds and by $t\bar{t}$ must be due to W+jet production. This sets the pretag estimation of the W+jet background. To estimate the number of $\texttt{SLT}_\texttt{e}$ tags due to W+jet production, we measure the tagging efficiency from the corresponding MC. To address the possibility that the heavy flavor composition in MC might be mis-estimated, we calibrate the fraction of events due to heavy flavor in generic jet data/MC. The correction ('K') factor is 1.0 ± 0.3.

IV. SYSTEMATIC UNCERTAINTIES

The dominant systematics for this analysis are those associated with the $\texttt{SLT}_\texttt{e}$, followed by the jet energy scale corrections, the QCD estimation, and the heavy flavor K-factor in W+jet MC (see table II for a summary). Each of the uncertainties from the $\texttt{SLT}_\texttt{e}$ algorithm have been discussed in Section II A. Other uncertainties include variations in the acceptance modeling by replacing the PYTHIA generator with HERWIG [5], identification of the primary lepton, theoretical/measured uncertainties on the various MC driven backgrounds, different PDFs, and more or less final state and initial state radiation. A luminosity systematic includes the theoretical uncertainty on the $p\bar{p}$ inelastic cross section and the calibration of the CDF luminosity counter.

V. RESULTS

We measure the top pair production cross section from the formula:

$$\sigma_{t\bar{t}} = \frac{N_{obs} - N_{bkg}}{\mathcal{A}\epsilon_{tag} \int \mathcal{L}dt} \tag{1}$$

where $\mathcal{A}$ is the geometric and kinematic acceptance of $t\bar{t}$Ẇe measure 7.8 ± 2.4 (stat) ± 1.5 (syst) ± 0.5 (lumi) pb. This measurement is consistent with the theoretical expectation and previous measurements of the cross section at

CDF Run II Preliminary (1.7 fb^{-1})

Process	1 jets	2 jets	3 jets	4 jets	$\geq$ 5 jets
Pretag	120599	19695	1358	645	193
Pretag $t\bar{t}$ (7.8 pb)	39.73 ± 2.11	210.70 ± 11.18	344.53 ± 18.29	365.76 ± 19.42	126.37 ± 6.71
WW	12.79 ± 1.26	12.22 ± 1.10	1.52 ± 0.14	0.63 ± 0.05	0.24 ± 0.02
WZ	1.37 ± 0.12	3.02 ± 0.26	0.42 ± 0.04	0.21 ± 0.02	0.06 ± 0.01
ZZ	0.16 ± 0.02	0.17 ± 0.02	0.05 ± 0.01	0.02 ± 0.00	0.01 ± 0.00
Single Top (s)	1.66 ± 0.16	6.96 ± 0.69	1.38 ± 0.14	0.52 ± 0.05	0.16 ± 0.02
Single Top (t)	5.65 ± 0.51	8.01 ± 0.73	1.08 ± 0.10	0.26 ± 0.02	0.04 ± 0.01
Z+Jets	51.26 ± 10.65	27.37 ± 4.87	2.49 ± 0.44	1.11 ± 0.20	0.28 ± 0.05
Drell-Yan	9.95 ± 2.26	6.45 ± 1.45	1.16 ± 0.26	0.36 ± 0.08	0.10 ± 0.02
QCD	39.37 ± 10.77	27.93 ± 9.42	3.89 ± 1.92	1.85 ± 0.91	0.55 ± 0.27
$W + b\bar{b}$	28.45 ± 10.95	22.14 ± 8.47	2.34 ± 0.90	1.00 ± 0.42	0.22 ± 0.10
$W + c\bar{c}, W + c$	103.64 ± 29.55	45.37 ± 14.06	3.64 ± 1.25	1.59 ± 0.60	0.35 ± 0.15
W+Light Flavor	952.40 ± 72.75	270.10 ± 20.93	17.76 ± 2.00	5.35 ± 1.12	1.17 ± 0.32
Total W+Jets	1084.48 ± 80.93	337.61 ± 21.89	23.74 ± 1.95	7.94 ± 1.37	1.74 ± 0.43
Background	1206.68 ± 78.86	429.73 ± 24.10	35.74 ± 2.27	12.89 ± 1.29	3.19 ± 0.40
$t\bar{t}$ (7.8 pb)	1.42 ± 0.10	13.26 ± 0.95	26.21 ± 1.91	30.59 ± 2.13	12.39 ± 0.85
Tags	1314	432	56	45	20

TABLE III: Sample Composition with $\texttt{SLT}_\texttt{e}$ tag level 1. Uncertainties include luminosity, lepton SFs, cross section uncertainties, tagger systematics, K-factor, and the QCD fit.

CDF. Our results are tabulated in Table III. Plots of event expectations versus jet multiplicity are shown in Figure 2. The P_T distribution for selected events is shown in Figure 3.

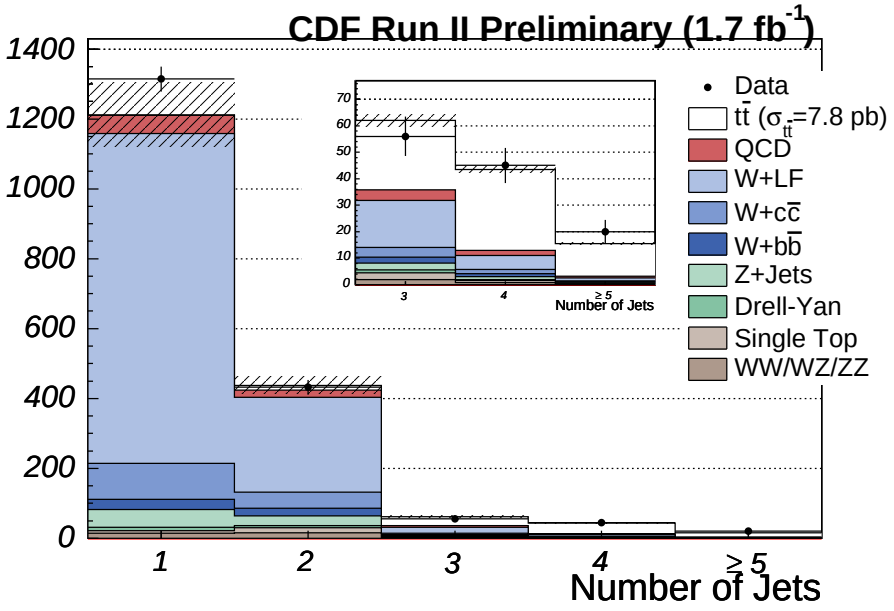

FIG. 2: Sample Composition of $\texttt{SLT}_\texttt{e}$ tags. The embedded plot is the $\geq$ 3 jet tag composition.

VI. CONCLUSIONS

We have measured the top pair production cross section at the Fermilab Tevatron at $\sqrt{s}$ = 1.96 TeV with 1.7 fb^{-1} of W+jets data. We have found a cross section consistent with previous measurements at CDF Run II in the

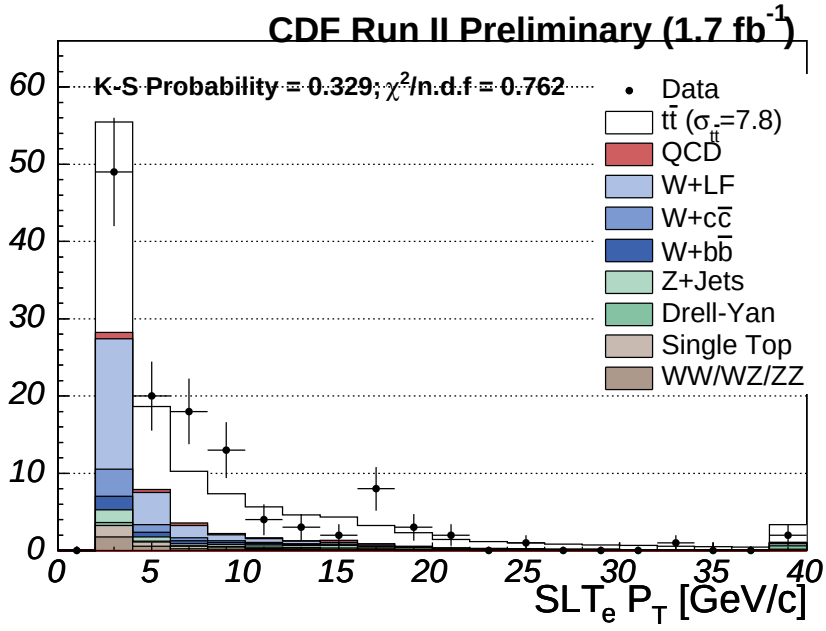

FIG. 3: Tag distribution by the $\texttt{SLT}_\texttt{e}$ track P_T (GeV/c).

lepton+jets channel. This is the first measurement of the cross section using a soft electron tagging technique at CDF.

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

[1] M. Cacciari, et al., The $t\bar{t}$ Cross-section at 1.8 TeV and 1.96 TeV: A Study of the Systematics due to Parton Densities and Scale Dependence, JHEP **404**, 68 (2004).
[2] 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
[3] L. Balka, et al., Nucl. Instrum. Methods Phys. Res. A **267**, 272 (1988).
[4] T. Sjostrand et al., High-Energy-Physics Event Generation with PYTHIA 6.1, Comput. Phys. Commun. **135**, 238 (2001).
[5] G. Corcella et al., HERWIG 6: An Event Generator for Hadron Emission Reactions with Interfering Gluons (including supersymmetric processes), JHEP **01**, 10 (2001).
[6] M.L. Mangano et al, JHEP **07**, 1 (2003).