CDF note 7075

Measurement of the $t\bar{t}$ production cross section in the all-hadronic channel

The CDF Collaboration
URL http://www-cdf.fnal.gov

(Dated: August 17, 2004)

We present here the measurement of the $t\bar{t}$ production cross section in the all-hadronic channel, where both $W's$ decay hadronically. The analysis is performed using 165 pb^{-1} of $p\bar{p}$ collisions collected with a multijet trigger at $\sqrt{s} = 1.96$ GeV with the Collider Detector at Fermilab. After the application of an optimized kinematical selection we observe an excess of events with 6 or more jets, including one or more b jets, relative to background expectations. Based on this excess we measure the production cross section $\sigma_{t\bar{t}} = 7.8^{+5.3}_{-3.4}$ pb.

Preliminary Results for Summer 2004 Conferences

I. INTRODUCTION

This note describes a measurement of the $t\bar{t}$ production cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$TeV with the CDF[1] detector at Fermilab Tevatron. In this analysis we search for events in which both W bosons decay into a $q\bar{q}'$ pair, leading to an all-hadronic final state characterized by a nominal 6-jet topology. This final state has the advantage of a large branching ratio ($BR \approx 4/9$) and of being fully reconstructed. The major drawback is that it competes against a very large QCD multijet background which dominates by three order of magnitude even after the application of a specific multijet trigger.

To improve the signal-to-background ratio, S/B, we apply a set of kinematical requirements which favors $t\bar{t}$ candidates leading to $S/B \approx 1/24$ with a good efficiency ($\epsilon_k = 6.2\%$). This is clearly not enough and we need to require the presence of displaced secondary vertices (*tags*) in the event to increase the S/B and to extract a $t\bar{t}$ signal. The analysis presented here is one of the various strategies used by CDF in Run I for the observation of $t\bar{t}$ production and the measurement of the top mass in this final state [2],[3]. This measurement provides a test of the QCD calculations of the $t\bar{t}$ production cross section[4], and a significant deviation from the predicted cross section could signal beyond the standard model production mechanisms for $t\bar{t}$ pairs.

II. DATA SAMPLE & EVENT SELECTION

We summarize here the procedure used to isolate all-hadronic $t\bar{t}$ candidates in the $p\bar{p}$ collisions collected by CDF in the period 2002/2003 and corresponding to 165 pb^{-1}.

A special multijet trigger has been devised specifically for this analysis. This trigger requires at least four calorimeter clusters with $E_T \geq 15$ GeV (with respect to $z = 0$) and a total transverse energy exceeding 125 GeV. Such a trigger has a cross section of about 11 nb and an efficiency of about 88% on all-hadronic $t\bar{t}$ events. The signal-to-background ratio at this level amounts to about 1/2500 (assuming $\sigma_{t\bar{t}} = 6.7$ pb).

We define some pre-requisites in order to have a sample of events as well defined and as clean as possible:

1. at least one *good quality* vertex with $|z_{vert}| < 60$ cm;
2. $\not\!\!E_T/\sqrt{\sum E_T} < 6$ to reject badly reconstructed events;
3. at least 4 jets (fixed cone radius 0.4) with raw-$E_T > 15$ GeV (with respect to the real vertex) and $|\eta| < 3.0$;
4. total transverse energy of these jets, $\sum E_T \geq 125$ GeV;
5. removal of events with a *good, high-P_T central electron or muon* to avoid overlap with other leptonic analyses.

After such pre-requisites we remain with about 1113000 data events.

In this analysis, Monte Carlo simulations are used only for the generation of $t\bar{t}$ samples intended for the optimization of the kinematical selection and the evaluation of its efficiency and systematics. The main sample has been generated with HERWIG [5] 6.4(CTEQ5L as p.d.f.), while additional samples (PYTHIA [6] 6.203) generated under different conditions of ISR/FSR modeling and using different p.d.f. sets have been produced for the evaluation of the systematic uncertainty on the efficiency. All samples have been processed with a realistic simulation of the CDF detector response in particular in order to reproduce the performances of the Silicon detector.

A. Kinematical selection

After the initial preselection we seek an optimal kinematical selection based on the best $S/\sqrt{B+S}$ achievable. As signal we consider $t\bar{t}$ Monte Carlo events (all-hadronic events only), while as background we consider the multijet events themselves (this is possible because the S/B is very small). We consider jets with a transverse energy $E_T \geq 20$ GeV (corrected for detector and reconstruction effects), and pseudo-rapidity $|\eta| \leq 2.0$. The quantities used for the selection are, as suggested by [2], number of jets (N_{jet}), total transverse energy ($\sum E_T$), centrality (C), aplanarity (A) and the sub-leading energy ($\sum_3 E_T \equiv \sum E_T - E_T^1 - E_T^2$) obtained by removing the two jets with the highest E_T. The comparison of the data (background dominated) and signal MC for the chosen kinematical variables is shown in Figure 1.

The optimal selection is the following (in cascade):

- all the pre-requisites

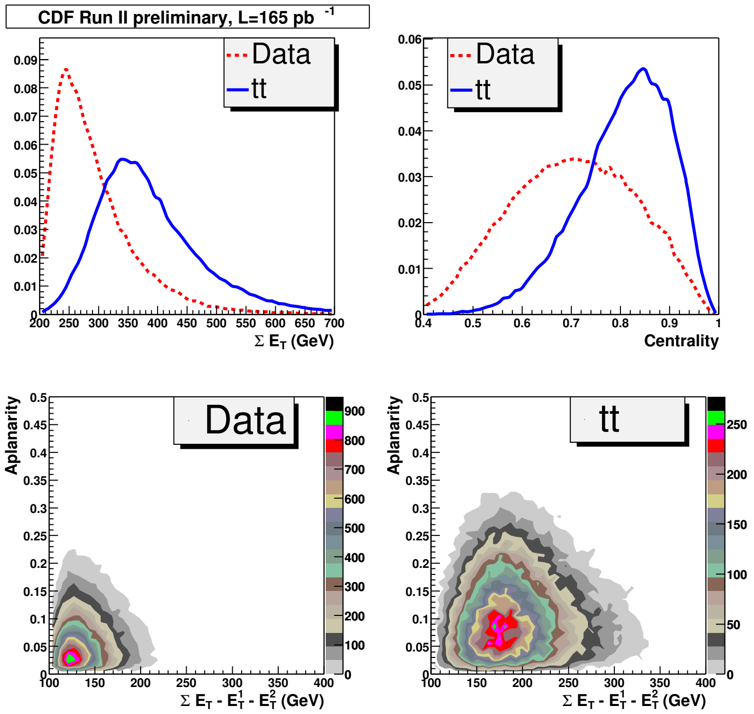

FIG. 1: Kinematical variables distributions for $t\bar{t}$ signal and multijet data for events $6 \leq N(jet) \leq 8$

- $6 \leq N_{jet} \leq 8$
- $A + 0.0037 \sum_3 E_T \geq 0.85$
- $C \geq 0.77$
- $\sum E_T \geq 320$ GeV

Such a selection isolates about 1700 candidates with an efficiency of $6.2 \pm 1.9\%$ (on all $t\bar{t}$ events) and $S/B \approx 1/24$. The relative systematic uncertainty on the efficiency is quite large (30%) and dominated by the uncertainties on the jet energy scale (see Table I): this will be reduced with the new versions of the calorimeter response simulation.

Source	$\delta\epsilon/\epsilon$ (%)
Generator	2.9
ISR/FSR modeling	4.0
p.d.f.	7.4
Jet energy scale	28.8
Total	30.1

TABLE I: Relative systematic uncertainties on the kinematical efficiency.

B. Displaced vertex tagging

In order to further improve the S/B we require the presence of secondary vertices reconstructed with at least two good quality tracks with hits in our Silicon vertex detector. To be called *tags* the jets with secondary vertices need to have a signed displacement from the primary vertex with a significance larger than 3.0. The cross section measurement is performed using the total number of tagged jets (not events) in order to avoid the explicit calculation of the background for double tagged events (this is the subject of another ongoing analysis). The effects due to the presence of double tags are included as a systematic contribution.

The average number of tags in a $t\bar{t}$ event passing the kinematical selection is $n_{tag}^{ave} = 0.763 \pm 0.005(\text{stat.}) \pm 0.065(\text{syst.})$, as derived from simulations and comparison to b-enriched events.

III. BACKGROUND ESTIMATE

The method we use to estimate the background is based on the verified assumption that even if the relative contribution from all QCD processes changes as a function of jet multiplicity, the probability for a *taggable* jet (that is a jet inside the vertex detector acceptance) to be tagged remains flat increasing the jet multiplicity. This allows to use the probability extracted from a control sample (events with $N_{jet} = 4$) on jets in events with higher jet multiplicity in order to estimate the number of tags expected in the signal sample. Such a probability (*tag rate*) is defined as:

$$\mathcal{P} = \frac{\#\ of\ tagged\ jets}{\#\ of\ taggable\ jets}$$

To account for the different conditions of the control and the signal samples, the tag rate is parametrized in terms of jet-E_T, jet-η, number of tracks reconstructed in the vertex detector (N_{trk}) and aplanarity (A).

Before the kinematical selection the multijet sample is composed, essentially, by background events only. The goodness of the parametrization and of the resulting estimates across different jet multiplicities is shown in Table II where we see that the expectations agree with the observed number of tags within a few %.

Jet multiplicity	4	5	6	7	8
Events	532634	249573	66807	13073	2134
Taggable jets	1274233	742349	238532	54664	10221
Observed Tags	38421	21313	6680	1415	293
Expected Tags	38421 ± 191	21150 ± 116	6524 ± 46	1464 ± 14	270 ± 3
$N_{obs} - N_{exp}$	0 ± 274	163 ± 186	156 ± 94	-49 ± 40	23 ± 17
$\frac{N_{obs}-N_{exp}}{N_{exp}}$ (%)	0 ± 0.7	0.8 ± 0.9	2.3 ± 1.4	-3.3 ± 2.7	8.5 ± 6.5

TABLE II: Observed and Expected Tags (statistical uncertainties only).

The agreement can be seen also in terms of average number of tags, observed and expected, per event, see Figure 2.

IV. SYSTEMATICS UNCERTAINTIES ON THE BACKGROUND ESTIMATE

The kinematical selection changes the event characteristics with respect to those found in the control sample (events with $N_{jet} = 4$) where the parametrization has been derived. Possible biases due to the selection are treated as systematic uncertainties on the background prediction and are evaluated with the help of suitable control samples which are essentially composed of multijet events as highly populated as possible but with the smallest possible presence of $t\bar{t}$ events. We define several samples where we exclude $t\bar{t}$ events by "negating,, the all-hadronic kinematical selection. With these samples we evaluate a relative systematic uncertainty of 5% on the background prediction after the kinematical selection, see Table III. We also checked for possible dependencies on instantaneous luminosity and detector effects and found those negligible.

Variable	Relative Uncertainty (%)
$6 \leq N_{jet} \leq 8$	0.2
$\sum E_T \geq 320$ GeV	1.4
Centrality ≥ 0.77	0.3
Aplanarity $+0.0037 \sum_3 E_T \geq 0.85$	4.7
Total	5

TABLE III: Systematic uncertainties on the kinematical requirements.

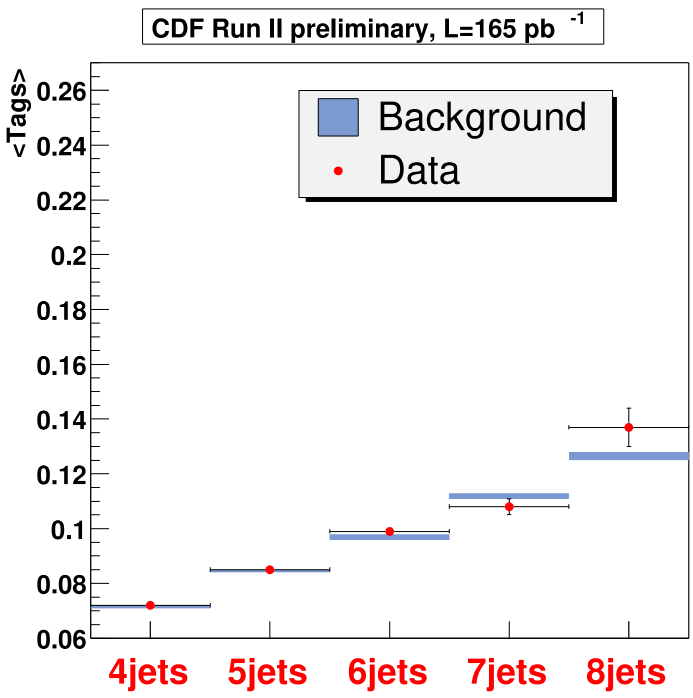

FIG. 2: Average number of tags (observed and expected) per event before kinematical selection as a function of the jet multiplicity.

V. BACKGROUND ESTIMATE AFTER THE KINEMATICAL SELECTION

The background estimate after the kinematical selection is summarized in Table IV. The presence of an excess in the signal region ($6 \leq N_{jet} \leq 8$) is visible in Figure 3, and also when we look at the average number of tags per event as a function of the jet multiplicity (see Figure 4).

N_{jet}	4	5	6	7	8
Events	60	420	773	630	253
Background	7.3 ± 0.7	61.3 ± 4.1	114 ± 7.0	103 ± 7.0	45.9 ± 2.4
$t\bar{t}$ ($\sigma = 6.7$ pb)	0.2 ± 0.1	7.1 ± 2.2	27.4 ± 8.5	19.3 ± 6.0	5.7 ± 1.8
All SM sources	7.5 ± 0.7	68.4 ± 4.7	141.4 ± 11.0	122.3 ± 9.2	51.6 ± 3.0
Observed Tags	11	70	170	116	40
$Data - Bkgd$	3.7 ± 3.4	8.7 ± 9.3	56 ± 15	13 ± 13	-5.9 ± 6.8

TABLE IV: Observed tags and background estimate after the kinematical selection

After the application of the kinematical selection to a multijet sample of 165 pb^{-1}, we are left with $N_{evt} = 1656$ events, out of which 284 events have at least one tag, for a total of $n_{obs} = 326$ candidate tags. The total background expectation in the signal region, provided by the tag rate parametrization, amounts to $n'_{exp} = 264.7 \pm 17.2$ tags after correcting the background for the presence of $t\bar{t}$ events in the pretag sample. The excess of tags $Data - Background = 61 \pm 25$ is ascribed to $t\bar{t}$ production and from this excess we measure the production cross section. In Figure 5 we show the number of tagged jets as a function of the jet multiplicity and compare it to background and signal (assuming $\sigma_{t\bar{t}} = 6.7$ pb) contributions.

The ingredients we use for the cross section measurement are shown in Table V and with these ingredients we build a likelihood function:

$$\mathcal{L} = e^{-(L-\overline{L})^2/2\sigma_L^2} \cdot e^{-(\epsilon_k-\overline{\epsilon_k})^2/2\sigma_{\epsilon_k}^2} \cdot e^{-(n_{tag}^{ave}-\overline{n_{tag}^{ave}})^2/2\sigma_{n_{tag}^{ave}}^2} \cdot e^{-(b-\overline{b})^2/2\sigma_b^2}.$$

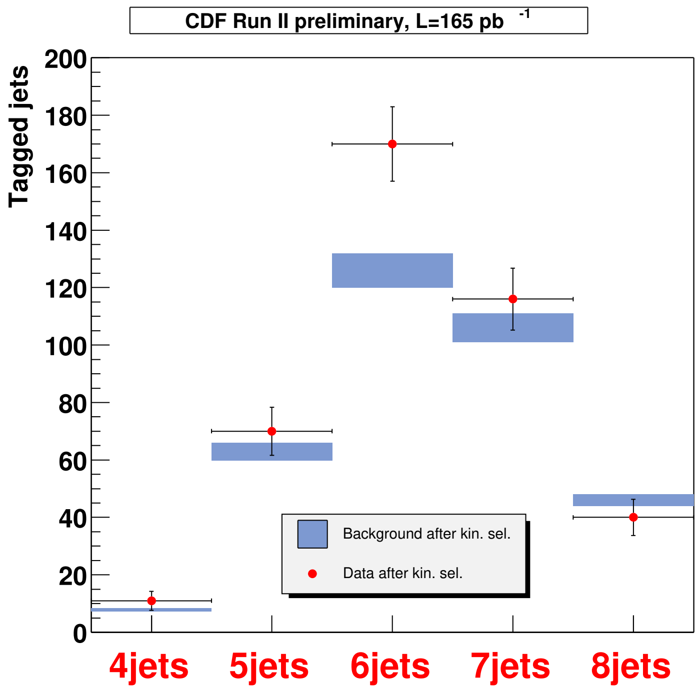

FIG. 3: Comparison between observed and expected tags in each multiplicity bin

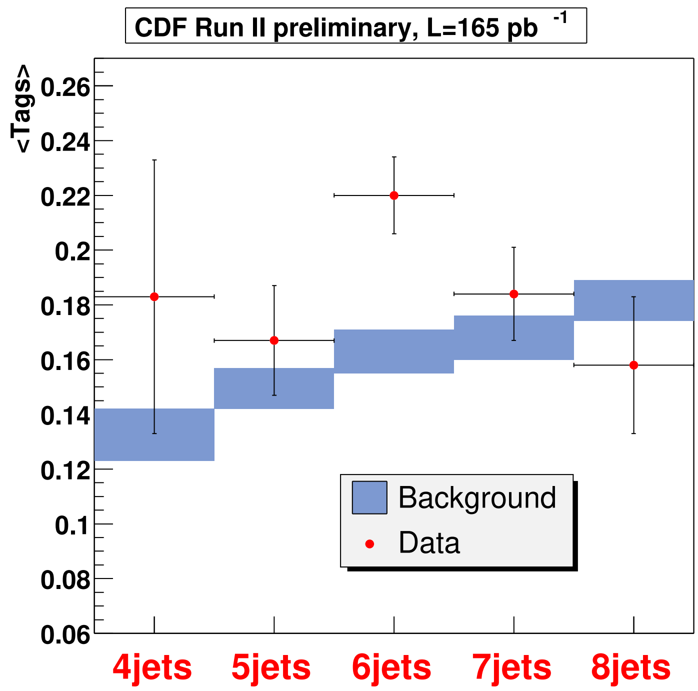

FIG. 4: Average number of tags per event as a function of the jet multiplicity

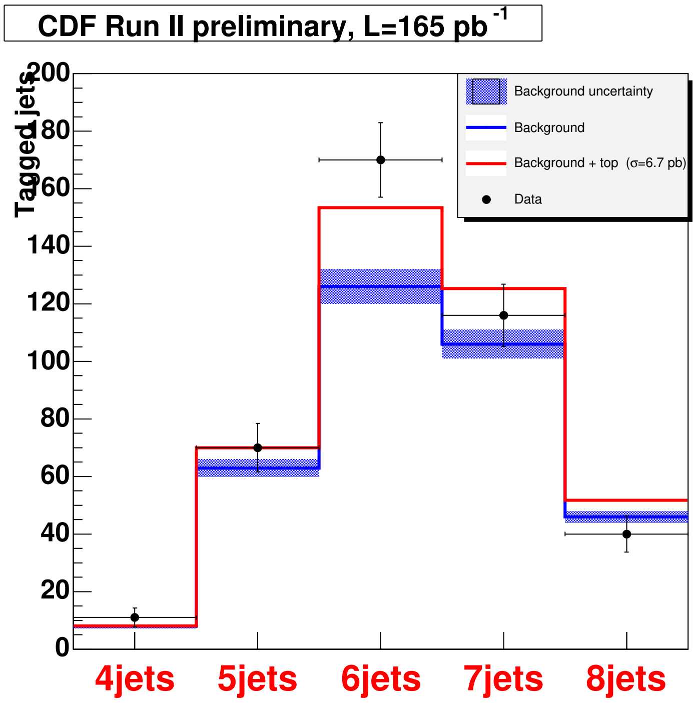

FIG. 5: Number of tagged jets *vs* jet multiplicity. Data (points), background with uncertainty band and $t\bar{t}$ expectation ($\sigma_{t\bar{t}} = 6.7$ pb)

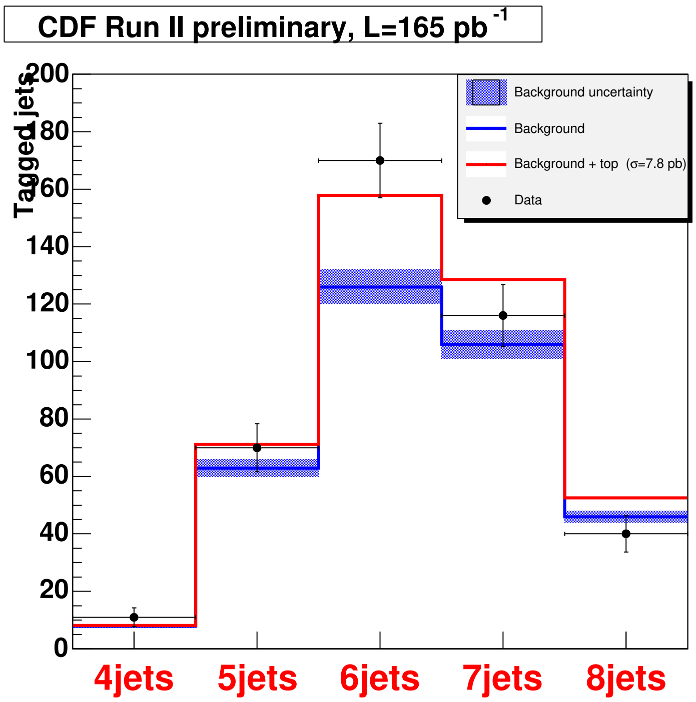

FIG. 6: Number of tagged jets *vs* jet multiplicity. Data (points), background with uncertainty band and $t\bar{t}$ expectation (assuming $\sigma_{t\bar{t}} = 7.8$ pb)

Variable	Symbol	Input value
Integrated luminosity (pb^{-1})	L	165 ± 10
Candidate tags	$n \equiv n_{obs}$	326
Expected background	$b \equiv n'_{exp}$	264.7 ± 17.2
Kinematical efficiency (%)	ϵ_k	6.2 ± 1.9
Average number of tags per $t\bar{t}$	n^{ave}_{tag}	0.763 ± 0.065

TABLE V: Input values for the Likelihood maximization

$$\cdot \frac{(\sigma_{t\bar{t}} \cdot \epsilon_k \cdot n^{ave}_{tag} \cdot L + b))^n}{n!} \cdot e^{-(\sigma_{t\bar{t}} \cdot \epsilon_k \cdot n^{ave}_{tag} \cdot L + b)}$$

The central value of the cross section is given by the maximum of $\log \mathcal{L}$, that is

$$\sigma_{t\bar{t}} = \frac{n-b}{\epsilon_k n^{ave}_{tag} L}$$

while the uncertainties are calculated as $\Delta \log \mathcal{L} = \frac{1}{2}$ points of the likelihood function. The value of the $t\bar{t}$ production cross section, measured in the all-hadronic channel is then:

$$\sigma_{t\bar{t}} = 7.8 \pm 2.5(stat)\ ^{+4.7}_{-2.3}(syst) = 7.8^{+5.3}_{-3.4}\ pb$$

Figure 6 shows the number of tagged jets as a function of the jet multiplicity and its comparison to the sum of background and signal (assuming our measured value $\sigma_{t\bar{t}} = 7.8$ pb) expectations.

VI. CONCLUSIONS

We have discussed here the measurement of the $t\bar{t}$ production cross section in the all-hadronic channel using a sample of 165 pb^{-1} events collected by CDF with a multijet trigger in the period 2002/2003. Starting from a very small signal-to-background ratio ($S/B \approx 1/2500$), we improve it to $S/B \approx 1/4$ by applying a specific kinematical selection ($\epsilon = 6.2 \pm 1.9\%$) and the request of presence of displaced secondary vertices ($n^{ave}_{tag} = 0.763 \pm 0.065$)). The background estimate is completely data driven based on a parametrization of the tag rate on a multijet sample that is signal depleted but kinematically similar to our signal region. The counting of tags shows an excess of about 61 tags with respect to a background of 265 ± 17 tags. Abscribing such an excess to $t\bar{t}$ production, we measure the cross section to be

$$\sigma_{t\bar{t}} = 7.8 \pm 2.5(stat)\ ^{+4.7}_{-2.3}(syst) = 7.8^{+5.3}_{-3.4}\ pb$$

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] 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

[2] F. Abe *et al.*, *First Observation of the All-Hadronic Decays of $t\bar{t}$ Pairs*, Phys. Rev. Lett. 79, 1992 (1997).
[3] A. Sidoti *et al*, *Investigation of $t\bar{t}$ in the full hadronic final state at CDF with a Neural Network approach*, Nucl.Instrum.Meth.A466:538-550 (2001)
[4] 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).
[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] T. Sjostrand et al., High-Energy-Physics Event Generation with PYTHIA 6.1, Comput. Phys. Commun. **135**, 238 (2001).