Combination of CDF top quark pair production cross section measurements with 2.8 fb^{-1}.

Preliminary results for summer conferences 2008.

CDF Collaboration.

We present a combination of measurements of the top quark pair production cross section using a data sample with an integrated luminosity of up to 2.8 fb^{-1} collected by the CDF II detector at the Fermilab Tevatron. Three of five measurements carry most of the weight in the combination: the lepton+jets channel secondary vertex b-tagging with a weight of 44%, lepton+jets channel artificial neural network with a weight of 28%, followed closely by the dilepton channel with 25%. We also include the two soft lepton tagging results in the lepton+jets channel, which carry weights of 4% for muon and -1% for electron.

We obtain an improved estimate of:

$$\sigma_{t\bar{t}} = 7.02 \pm 0.63 \text{ pb for } m_{top} = 175 \text{ GeV}/c^2.$$

The statistical uncertainty is 0.30 pb, the systematic uncertainty is 0.38 pb, and the luminosity uncertainty is 0.41 pb. The result is in good agreement with the theoretical prediction. We find good agreement among five measurements in the dilepton and lepton+jets channels, with a probability of about 69% to find a less consistent set of measurements.

I. INTRODUCTION

In $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV, the strong interaction allows pair production of top quarks, at first order via the processes $q\bar{q} \rightarrow t\bar{t}$ and $gg \rightarrow t\bar{t}$. The subsequent decay of the massive top quark via the weak interaction process, $t \rightarrow W^+b$, with a predicted branching fraction of approximately 0.99 leads to three detector signatures: dilepton, where both W bosons decay leptonically, $W^+ \rightarrow \ell^+\nu$ and $W^- \rightarrow \ell^-\bar{\nu}$; lepton+jets, where one W boson decays leptonically and one W boson decays to $q\bar{q}'$; and all-hadronic, where both W bosons decay to $q\bar{q}'$. New processes beyond the standard model, for example top quark pair production via a new massive resonance [4] or top quark decay via a charged Higgs boson [5], could alter the experimentally observed rate in these different signatures. Each of these signatures is dominated by a standard model background, respectively from $Z \rightarrow \ell^+\ell^-$ with associated jets, $W \rightarrow \ell\nu$ with associated jets, and QCD multi-jet production.

CDF has performed many measurements of the top quark pair production cross-section. In this paper, we combine five experimental measurements with up to 2.8 fb^{-1}in order to check the consistency of the different measurements and to reduce the experimental uncertainty for an improved test of the theoretical prediction. We combine one measurement from the dilepton channel (DIL) with four measurements from the lepton+jets channel that take advantage of different characteristics to discriminate against background, such as the higher transverse energy and angular separation of the decay products of the massive top quark and the presence of two highly boosted B hadrons. The four measurements in the lepton+jets channel use an artificial neural network based on seven kinematic and topological observables (ANN), displaced secondary vertex b-tagging (SVX), and soft lepton b-tagging where the lepton is a muon (SLT-μ) or an electron (SLT-e). We summarize the five measurements in Table I.

We form a best linear unbiased estimate (BLUE) [6–8] for the top quark pair production cross section. We do this by constructing a covariance matrix from the statistical and systematic uncertainties for each result, taking into account the statistical and systematic correlations. We invert this covariance matrix to obtain a weight for each result and combine the results with these weights to obtain our best estimate. We have checked that our technique with pseudo-experiments. The consistency of the results can also be evaluated, with the statistical and systematic correlations taken into account. Although it is not done here, the BLUE technique also allows the evaluation of a cross section for each decay channel.

II. ERROR MATRIX AND TREATMENT OF UNCERTAINTIES

Having introduced the combination technique and the measurements, we now discuss the error matrix in more detail. The diagonal elements of the covariance matrix are the variances of the individual measurements. The correlations leading to off-diagonal elements will be discussed in the next section III.

For each measurement, the variance is computed from the sum in quadrature of the statistical and systematic uncertainties. For systematic uncertainties, we classify each uncertainty by its physical source and group these various sources of systematic uncertainty into categories that are uncorrelated for an individual measurement, though they may be correlated between different measurements. We consider the following sources of systematic uncertainty to be uncorrelated with each other for an individual measurement: trigger efficiency and lepton identification efficiency, jet energy calibration, SVX b-tag efficiency, SLT muon b-tag efficiency, SLT electron b-tag efficiency, $t\bar{t}$ model for the detector acceptance (from QCD ISR/FSR, fragmentation and hadronization, and the parton distribution function), integrated luminosity, and background estimate split into three categories: fake, mis-tag, W+jets Q^2 and EWK; W+heavy flavor; and QCD background. We summarize the systematic uncertainties for each result in Table II.

The categories of acceptance-like systematic uncertainty are:

- **Lepton Trigger and Identification efficiency:** The lepton identification efficiency is calibrated by comparison of a PYTHIA sample of Z boson decays with a data sample acquired with an single high p_T lepton trigger.

- **$t\bar{t}$ model:** This systematic includes the effects from QCD ISR/FSR, PDFs, and PYTHIA and HERWIG generator.

- **Jet energy calibration:** This is derived from the standard uncertainty on the jet energy scale [9]. The acceptance for the PYTHIA MC sample is recalculated after adjusting the jet energy by the estimated systematic uncertainty.

- **SVX b-tag efficiency:**The uncertainty on the efficiency between data and Monte Carlo is measured using a di-jet sample with a low p_T electron or muon. The dominant uncertainties are from the extrapolation of the scale-factor to the higher jet E_T typical of b-jets from top quark decay. There is also a component from the extrapolation to high instantaneous luminosity, as the control trigger is prescaled.

- **SLT muon b-tag efficiency:** The efficiency to identify a displaced muon with p_T above 3 GeV/c in the dense environment of a jet has systematics associated to the difference between tracking efficiency in Monte Carlo and data and the isolation dependence of the muon identification. Therefore, this systematic is not expected to be correlated with the SVX efficiency determination.

- **SLT electron b-tag efficiency:** The efficiency to identify a displaced electron with p_T above 2 GeV/c in the dense environment of a jet has systematics associated to the rejection of conversions and calorimeter modelling of electron identification. Therefore, this systematic is not expected to be correlated with the SVX or SLT muon efficiency determinations.

- **Luminosity:** The standard uncertainty on the luminosity of 5.8%, with 4% from the total inelastic $p\bar{p}$ cross section and 4.2% from detector effects [10]. The dominant contributions to the detector uncertainty are 3% from the simulation of material in CDF and 2% from choice of generator. As these effects are not expected to be time-dependent, it is reasonable to take the luminosity uncertainty across different periods of data-taking as 100% correlated.

The size in picobarns of the acceptance-like uncertainties depends on the measured cross-section. In order to avoid measurements that fluctuate low from carrying undeservedly large weights and this producing a bias in the combination towards a lower value than the true value, we evaluate acceptance-like uncertainties at the value of the combined cross-section for the current iteration (out of three total iterations). In contrast, the uncertainty on the cross section due to the background is kept fixed in picobarns; it does not depend on the value of the cross section.

III. THE CORRELATIONS

A. Systematic correlations

We treat the trigger and lepton identification uncertainties as correlated between the techniques using a lepton trigger and lepton identification, that is DIL, ANN, SVX, and SLTs. We treat the uncertainties from jet energy calibration, $t\bar{t}$ model and luminosity, which are estimated in a common way for all the techniques, as correlated across all measurements.

For the SLT and SVX measurements, the component that depends on W+heavy flavor and QCD is considered to be correlated between the measurements. In contrast, the soft lepton tagger fake matrix component is not considered to be correlated with the SVX mistag matrix. Note for the ANN, the QCD background is determined by the variation of the fit results from choosing different source for the QCD ANN distribution, after imposing $\not\!\!E_T > 35$ GeV. This is considered to be uncorrelated with the uncertainty on the absolute normalization that is used in the SLT and SVX results.

B. Statistical correlations

The four results in the lepton+jets channel each use a subset of events from those that pass the common lepton+jets selection. We assume the events that pass the common lepton+jets selection come from one of three processes: $t\bar{t}$, W+jets, and multi-jet QCD. We construct 1,000 pseudo-experiments that will track the primary source of correlation between the samples of events, which is due to statistical fluctuations in the base sample for the numbers of $t\bar{t}$ or background events having correlated consequences for each of the subsamples. For each pseudo-experiment, we first generate the number of events for each process that pass the event selection requirements common to all three samples from a Poisson distribution with mean equal to the expected number of events for an integrated luminosity of $\mathcal{L}$=2800 pb^{-1}. We then determine the number of events that survive the cuts for each subsample by applying process-dependent efficiencies for the additional requirements beyond the common selection, such as b-tagging and H_T cuts. An event in the base sample is defined to have passed a cut if a random number generated between 0 and 1 is less than that cut's efficiency.

In addition, we also model for the $t\bar{t}$ process the correlation between the value of the ANN output and the H_T, $\not\!\!E_T$, leading jet E_T, and SVX b-tag requirements. We do this using a PYTHIA Monte Carlo sample for the $t\bar{t}$ process. From previous studies with 300 pb^{-1}, we found the statistical correlations between ANN and SLT, and the SVX and SLT to be insensitive to whether we tracked the correlation between the SLT tag and the ANN output or SVX tag in $t\bar{t}$ events. For the background processes, we neglect the correlation between the the value of the ANN output and the H_T, $\not\!\!E_T$, leading jet E_T, and SVX b-tag requirements The ANN background rate is about twenty (forty) times larger than the rate of the SVX (SLT) background rates, so the impact of this effect would be small. While the SVX background is dominated by processes with heavy flavor, the SLT background is dominated by processes without heavy flavor. Thus the SVX and SLT backgrounds are not expected to be very correlated with each other either.

For each pseudoexperiment, we determine the $t\bar{t}$ cross section for each analysis. For ANN, we perform a maximum likelihood fit to the ANN distribution for each pseudoexperiment. For SVX and SLT, we use the observed number of tags and notags in each pseudoexperiment. We estimate a statistical correlation between ANN and SVX of 32%, between ANN and SLT of 8%, and between SVX and SLT of 12%. The statistical accuracy is about 2%. When we do not account for the correlation between the ANN output, H_T, $\not\!\!E_T$, leading jet E_T, and SVX b-tag information for $t\bar{t}$, we find a smaller statistical correlation between ANN and SVX of about 25%. We do take into account the smaller luminosity for SLT results here.

Therefore, for the combination, we use a statistical correlation between ANN and SVX as $\rho = 32\%$, between ANN and SLT as $\rho = 8\%$, and between SVX and SLT as $\rho = 12\%$. For studies of the stability of the cross section combination to systematic shifts in the statistical correlations, we will use a conservative uncertainty of $\pm 10\%$. For instance, we will decrease the statistical correlation between ANN and SVX to 42% and increase it to 22%.

IV. THE DATA RESULT

We have performed many checks of the combination technique and found that it is unbiased to within 0.1 pb. We now combine five CDF measurements of the top quark pair production cross section, assuming a top quark mass of 175 GeV/c^2. The results and their systematic uncertainties were summarized in Tables I and II. In the following, we keep two decimal places for the sake of comparison of different categories. We find:

$$\sigma(p\bar{p} \rightarrow t\bar{t}) = 7.02 \pm 0.30_{(stat)} \pm 0.38_{(syst)} \pm 0.41_{(lumi)} \text{ pb.}$$

The SVX result carries the largest weight of 44%, followed by ANN with 28%, then DIL with 25%, then SLT-μ with 4%, and SLT-electron with -1%

The total uncertainty is 0.63 pb. The systematic uncertainty of 0.38 pb contains the following contributions: trigger and lepton identification 0.06 pb, jet energy calibration 0.19 pb, $t\bar{t}$ modeling 0.20 pb, SVX b-tag efficiency 0.16 pb, SLT-μ b-tag efficiency 0.015 pb, SLT-electron b-tag efficiency 0.007 pb, and background estimation 0.10 pb from mistag and W+jets Q^2, 0.11 pb correlated between SVX and SLT W+HF, and 0.12 from QCD background. The contributions to the total uncertainty from acceptance-like uncertainties have been evaluated with respect to the combined value of 7.02 pb.

The total correlation matrix, including statistical and systematic effects, is given in Table III. The largest correlation is 47% between the ANN and SVX measurements, due to the statistical correlation and the correlation between the acceptance and luminosity systematics.

The χ^2 of these five measurements is 2.27 for 4 degrees of freedom, corresponding to a probability of 69% to have a less consistent set of measurements. The weight and pull of each measurement in the combination are given in Table IV. The pre-BLUE total uncertainty is calculated with respect to the value of the individual measurement. The post-BLUE total uncertainty is calculated with respect to the value of the combination.

The agreement between the results is shown graphically in Figure 1. For each measurement, the statistical uncertainty is shown by the magenta line, which is superimposed on the total uncertainty including systematic and luminosity uncertainties as shown by the blue line. As also shown in Figure 1, we find excellent agreement with the theoretical prediction for a top quark mass of 175 GeV/c^2 and centre-of-mass energy of $\sqrt{s}$ = 1.96 TeV from Cacciari *et al.* of $6.73\pm^{0.25}_{0.47}\pm^{0.46}_{0.32} = 6.73\pm^{0.71}_{0.79}$ pb on adding scale (first) and CTEQ6.5 PDF (second) uncertainties ($6.73\pm^{0.52}_{0.57}$ pb in quadrature) [1], from Kidonakis and Vogt of $6.73\pm0.27\pm^{0.02}_{0.03}\pm^{0.43}_{0.33} = 6.73\pm^{0.72}_{0.63}$ pb on adding kinematics choice, scale and CTEQ6.6 PDF uncertainties ($6.73^{+0.53}_{-0.46}$ pb in quadrature) [2], and finally a NNLO approximate calculation from Moch and Uwer of $6.90\pm^{0.07}_{0.25}\pm^{0.39}_{0.39} = 6.90\pm^{0.46}_{0.64}$ pb on adding scale and CTEQ6.6 PDF uncertainties ($6.90^{+0.40}_{-0.46}$ pb in quadrature) [3]. Note in particular the reduced scale dependence of the Moch and Uwer calculation compared to the previous calculations. It is worth noting that all theoretical groups see that the central value shifts up by about 0.3 pb on using the MRST2006nnlo set. As the probable source of this shift, the theoretical references mention that CTEQ6 has a smaller gluon pdf at large x than MRST.

A comparison of the combined cross-section result against theoretical predictions is shown in Figure 2. At $\sqrt{s}$ = 1960 GeV, the central values of the predictions from [1] are shown at a top quark mass of 170 and 175 GeV/c^2 for CTEQ6.5 and MRST2006nnlo. The uncertainty (due to scale and PDF variations) of the 175 CTEQ6.5 prediction is shown by the filled blue region. At $\sqrt{s}$ = 1800 GeV, the predictions use CTEQ6M and MRST 2002.

The combined value of the cross-section at 175 GeV/c^2 is shown in Figure 3 with the theoretical predictions as a function of top quark mass.

TABLE I: Summary of results, for m_{top}=175 GeV/c^2.

Quantity	DIL	ANN	SVX	SLT-μ	SLT-e
N_{tag} (events)	–	–	952	248	121
N_{notag} (events)	162	5267	3336	3459	2065
Background (events)	52	–	225	87	52
$\mathcal{A}$ (%)	0.59	6.2	6.6	6.7	6.2
$\epsilon_{t\bar{t}}$	–	–	0.59	0.135	0.08
$\mathcal{L}$ (fb^{-1})	2.8	2.8	2.7	2.0	1.7
$\sigma_{t\bar{t}}$ (pb)	6.67	6.79	7.21	8.7	7.8
$\delta_{t\bar{t}}^{stat}$ (pb)	0.77	0.38	0.37	1.1	2.4
$\delta_{t\bar{t}}^{syst}$ (pb)	0.53	0.74	0.65	0.8	1.5

TABLE II: Systematic uncertainties on the **measured cross section in %**. Sources correlated between results are indicated by †.

Quantity	DIL	ANN	SVX	SLT-μ	SLT-e
Lepton id/trigger/zvtx	$1.5^{\dagger}$	$0.67^{\dagger}$	$0.8^{\dagger}$	$0.8^{\dagger}$	$0.8^{\dagger}$
Jet energy	$3.2^{\dagger}$	$3.1^{\dagger}$	$2.2^{\dagger}$	$4.1^{\dagger}$	$8.4^{\dagger}$
$t\bar{t}$ model	$1.9^{\dagger}$	$4.7^{\dagger}$	$2.1^{\dagger}$	$2.8^{\dagger}$	$2.5^{\dagger}$
SVX b-tag	-	-	5.3	-	-
SLT-μ b-tag	-	-	-	5.1	-
SLT-e b-tag	-	-	-	-	15.5
Luminosity	$5.8^{\dagger}$	$5.8^{\dagger}$	$5.8^{\dagger}$	$5.8^{\dagger}$	$5.8^{\dagger}$
Background (W+jets, mis-tag)	2.8	2.8	2.1	3.8	6.0
Background (W+HF)	-	-	$3.2^{\dagger}$	$3.9^{\dagger}$	$6.0^{\dagger}$
Background (QCD)	-	6.4	$0.2^{\dagger}$	$1.2^{\dagger}$	-

TABLE III: Total correlation matrix, including statistical and systematic correlations. Diagonal symmetry is implied.

Correlation	ANN	SVX	DIL	SLT-μ	SLT-e
ANN	1	0.47	0.34	0.30	0.20
SVX		1	0.32	0.36	0.24
DIL			1	0.20	0.12
SLT-μ				1	0.14

TABLE IV: For top quark mass value of 175 GeV/c^2, the cross section, uncertainty, weight, and pull of each result in the combination.

Quantity	DIL	ANN	SVX	SLT-μ	SLT-e
$\sigma_{t\bar{t}}$ (pb)	6.67	6.79	7.21	8.7	7.8
pre-BLUE total uncertainty (pb)	0.92	0.80	0.76	1.44	2.89
post-BLUE total uncertainty (pb)	0.93	0.83	0.75	1.36	2.82
Weight (%)	25	28	44	4	-1
Pull	-0.37	-0.27	0.26	1.23	0.28

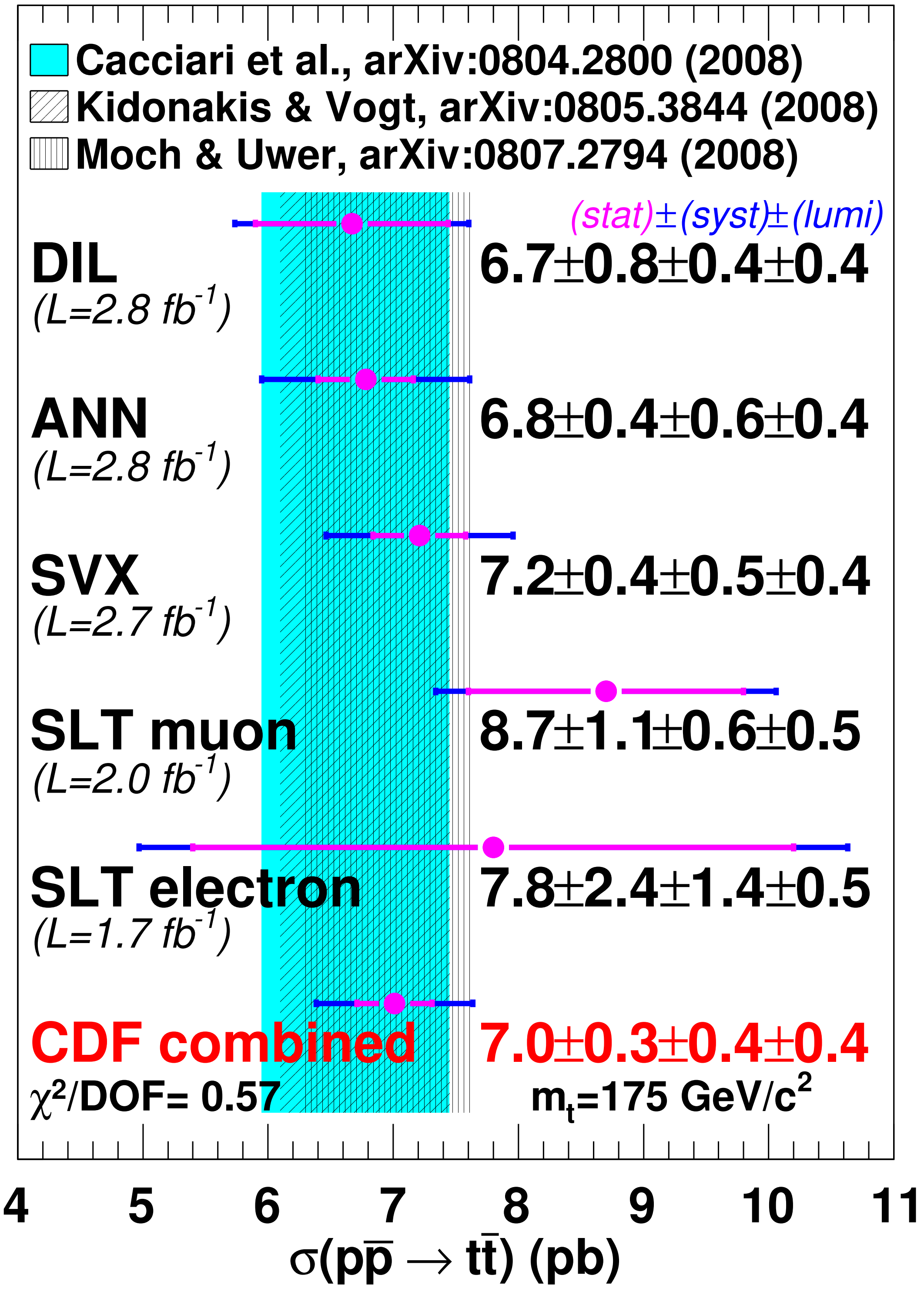

FIG. 1: Combination of CDF measurements of the top quark pair production cross section. The theoretical predictions use the CTEQ6.5 set for Cacciari et al., CTEQ6.6M for Kidonakis and Vogt, and CTEQ6.6M for Moch and Uwer.

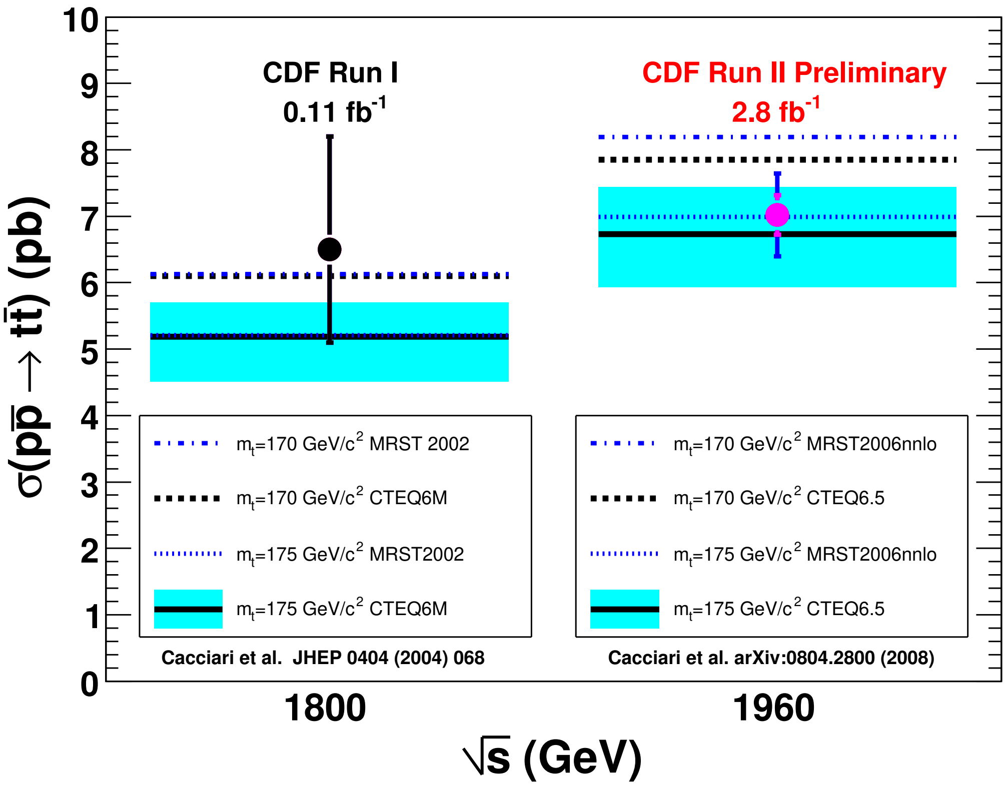

FIG. 2: Combination of CDF measurements of the top quark pair production cross section versus the center-of-mass energy of the Tevatron. The predictions from [1] are shown to compare the predictions at different energies, top quark masses and PDF sets. The other predictions are quite similar, so they are not shown here as otherwise this plot would become extremely busy.

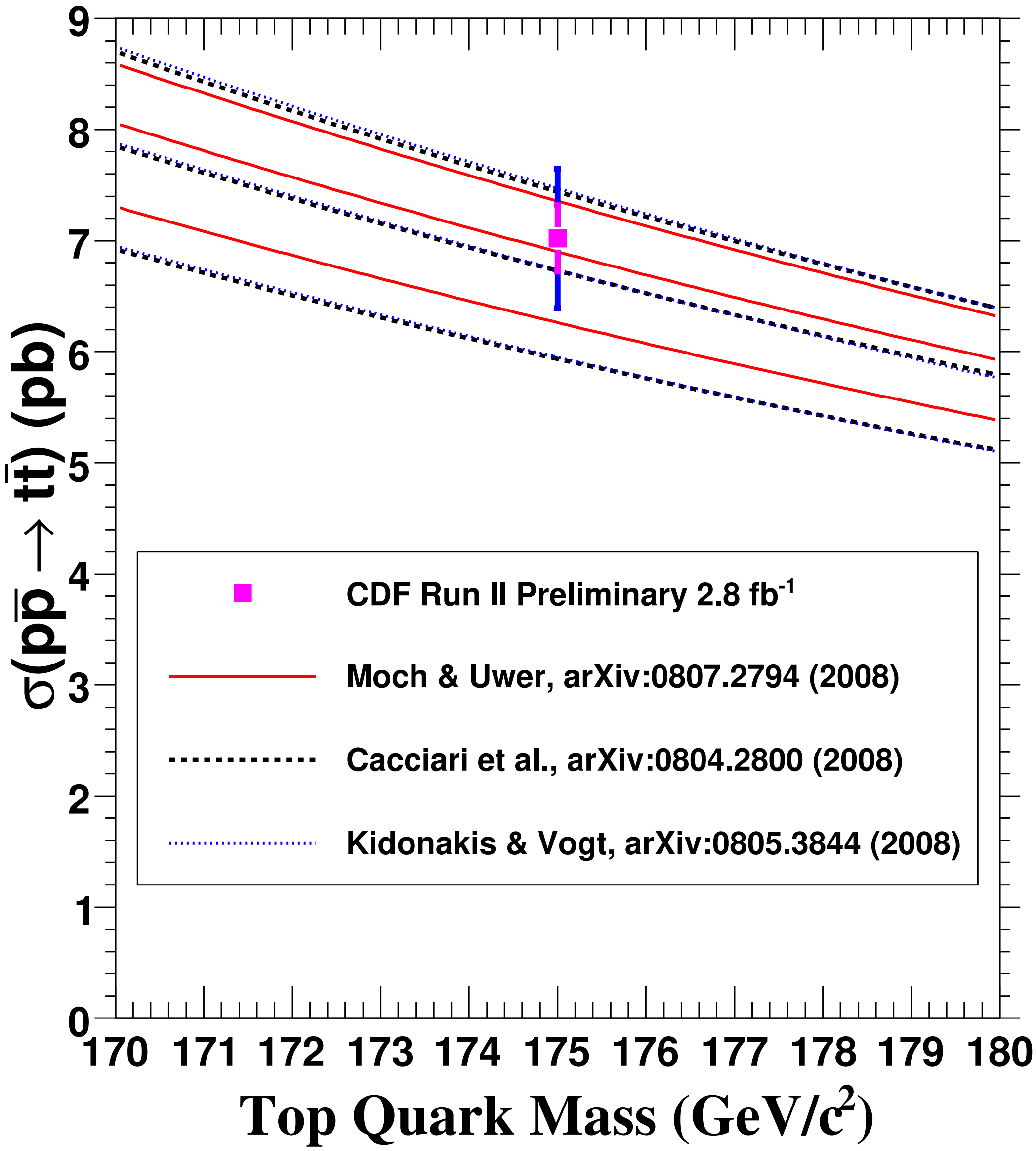

FIG. 3: Combination of CDF measurements of the top quark pair production cross section at 175 GeV/c^2 versus the theoretical predictions as a function of top quark mass.

A. Stability checks of data result

We checked the stability of the result for an increase or a decrease of 10% in the assumed ANN-SVX statistical correlation from 32% to 42% or 22%. The result and the uncertainty change by less than 0.01 pb.

We checked the stability of the result if the QCD background uncertainty is also assumed to be correlated between ANN and SVX and SLT. The result changes by -0.01 pb and the uncertainty changes by less than 0.01 pb.

V. CONCLUSIONS

We present a combination of five measurements of the top quark pair production cross section using a data sample with an integrated luminosity of up to 2800 pb^{-1} collected by the CDF II detector at the Fermilab Tevatron.

The data result is:

$$\sigma_{t\bar{t}} = 7.02 \pm 0.63 \text{ pb for } m_{top} = 175 \text{ GeV}/c^2.$$

The statistical uncertainty is 0.30 pb, the systematic uncertainty is 0.38 pb, and the luminosity uncertainty is 0.41 pb. The relative uncertainty on this determination is 9%. We find good agreement among five measurements in the dilepton and lepton+jets channels, with a probability of about 69% to find a less consistent set of measurements. In comparison to the single measurement with the smallest total uncertainty, that is SVX with 7.2±0.75 pb, we also find an improvement in precision of 16%.

The precision of this combined measurement of the top quark pair production measurement is close to that of the theoretical prediction. We find good agreement with the theoretical prediction for a top quark mass of 175 GeV/$c^2$and centre-of-mass energy of $\sqrt{s} = 1.96$ TeV of $6.73^{+0.71}_{-0.79}$ pb from [1], $6.73^{+0.72}_{-0.63}$ pb from [2], and $6.90^{+0.46}_{-0.64}$ pb from [3]. Note that the scale and PDF uncertainties have been added linearly here.

[1] M. Cacciari, S. Frixione, G. Ridolfi, M. Mangano and P. Nason, arXiv:0804.2800; M. Cacciari, S. Frixione, G. Ridolfi, M. Mangano and P. Nason, JHEP **404**, 68 (2004).

[2] N. Kidonakis and R. Vogt, arXiv:0805:3844 (2008);N. Kidonakis and R. Vogt, Phys. Rev. D **68**, 114014 (2003).

[3] S. Moch and P. Uwer, arXiv:0807.2794 (2008) and arXiv:0804.1476 (2008).

[4] R. Rossin *et al.*, CDF note 8044.

[5] A. Abulencia *et al.* (CDF Collaboration), Phys. Rev. Lett. 96, 042003 (2006).

[6] L. Lyons, D. Gibaut and P. Clifford, NIM A**270**, 110-117 (1988).

[7] L. Lyons, A. Martin and D. Saxon, Phys. Rev. D **41**, 3 (1990).

[8] A. Valassi, NIM A**500**, 391 (2003).

[9] A. Bhatti *et al.*, arXiv:hep-ex/0510047.

[10] S.Jindariani *et al.*, CDF note 7446.