

Combination of CDF top quark pair production cross section measurements with up to 4.6 fb^{-1} .

Preliminary results for summer conferences 2009.

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 4.6 fb^{-1} collected by the CDF II detector at the Fermilab Tevatron. There are four measurements that carry the following weights in the combination: the lepton+jets channel artificial neural network with a weight of 70%, the lepton+jets channel secondary vertex b -tagging with 18%, the dilepton channel with 18%, and the all-hadronic channel with -6%. We obtain an improved estimate of:

$$\sigma_{t\bar{t}} = 7.50 \pm 0.48 \text{ pb for } m_{top} = 172.5 \text{ GeV}/c^2.$$

The statistical uncertainty is 0.31 pb, the experimental systematic uncertainty is 0.33 pb, Z boson theoretical cross section uncertainty is 0.13 pb, and the luminosity uncertainty is 0.06 pb. The result is in good agreement with the theoretical prediction. We find good agreement among four measurements in with a probability of about 90% 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 [5] or top quark decay via a charged Higgs boson [6], 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 four experimental measurements [7–10] with up to 4.6 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) [7] and one measurement from the all-hadronic channel (HAD) [10] with two 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 two measurements in the lepton+jets channel use an artificial neural network based on seven kinematic and topological observables (ANN) [8], and a displaced secondary vertex b -tagging (SVX) [9]. We summarize the four measurements in Table I.

We form a best linear unbiased estimate (BLUE) [11–13] 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.

TABLE I: Summary of results, for $m_{top}=172.5$ GeV/ c^2 . For all-hadronic the numbers are for 1-tag (2-tag) events.

Quantity	DIL	ANN	SVX	HAD
N_{tag} (events)	119	–	1390	3452(441)
N_{notag} (events)	–	7348	4239	–
Background (events)	7.3 ± 1.8	–	392	2785 (201)
\mathcal{A} (%)	0.52	5.8	7.4	3.3 (0.95)
$\epsilon_{t\bar{t}}$	0.68	–	0.58	above
\mathcal{L} (fb $^{-1}$)	4.4	4.6	4.3	2.9
$\sigma_{t\bar{t}}$ (pb)	7.27	7.63	7.14	7.21
$\delta_{t\bar{t}}^{stat}$ (pb)	0.71	0.37	0.35	0.50
$\delta_{t\bar{t}}^{syst}$ (pb)	0.46	0.35	0.58	1.10
$\delta_{t\bar{t}}^{lumi}$ (pb)	0.42	0.15	0.14	0.42

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 summarize the systematic uncertainties for each result in Table II.

The categories of acceptance-like systematic uncertainty are:

- **Primary vertex z -position < 60 cm:** uncertainty on fraction of luminous region inside the CDF detector acceptance
- **Trigger and lepton identification efficiency:** the electron (CEM) and muon (CMUP) estimates are added in quadrature where appropriate. 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.
- **Track identification:** only applicable for ANN and SVX measurements that make a ratio

TABLE II: Systematic uncertainties as a percentage of the **measured cross-section**. Sources correlated between results are indicated by †; anti-correlation is indicated by ‡. The abbreviation NA means not applicable to that measurement, while CA means it cancels in the ratio to Z production, and ? means the uncertainty has not been evaluated although it should be present.

Category	Quantity	DIL	ANN	SVX	HAD	COMBO
1	Primary vertex	0.21†	CA	CA	0.2†	0.026
2	Lepton id/trigger	1.69†	0.66†	0.73†	1.8	0.90
3	Track ID	NA	0.60†	0.60†	NA	0.53
4	Lepton energy	0.82†	0.08†	0.08†	NA	0.22
5	Jet energy	2.7†	2.91†	4.0†	9.4†	2.69
6	SVX b -tag	3.52†	NA	4.4†	5.4 †	1.12
7	b -jet energy scale	NA	NA	NA	2.8	0.17
8	$t\bar{t}$ PDF	0.8†	0.94†	1.4†	3.4†	0.85
9	$t\bar{t}$ ISR/FSR	1.58†	0.42†	0.7†	0.6†	0.67
10	$t\bar{t}$ generator	0.94†	2.68†	2.7†	4.2†	2.28
11	W decay BF	1.85†	0.58†	0.6†	0.8‡	0.89
12	Color reconnection	2.61†	0.16†	0.4†	0.8†	0.60
13	Pileup	?	?	?	2.5	0.15
14	Method	NA	0.14	0.14	1.1	0.12
15	Z σ theory	NA	2.0†	2.0†	NA	1.76
16	Luminosity $p\bar{p}$ σ	4.2†	NA	NA	4.2†	0.50
17	Luminosity: detector	4.0†	0.2†	0.3†	4.0†	0.67
18	Background (charm)	NA	NA	1.1	NA	0.19
19	Background (mis-tag)	NA	NA	1.1	NA	0.19
20	Background (W +HF)	NA	NA	4.0	NA	0.70
21	Background (QCD/Fake)	NA	0.94	0.2	NA	0.67
22	Background (EWK)	NA	1.00	NA	NA	0.71
23	Background (other)	1.7	1.33	NA	8.2	1.09
24	Background ($Z\sigma$)	NA	0.27 †	0.3†	NA	0.24

to Z boson production.

- **Lepton energy:** uncertainty on the lepton energy scale.
- **Jet energy:** derived from the standard uncertainty on the jet energy scale [14]. The acceptance for the PYTHIA MC sample is recalculated after adjusting the jet energy by the estimated systematic uncertainty.

- **SVX b -tag efficiency:** 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.
- **b -jet energy scale:** considers the difference in jet energy scale between light flavor/gluon jets and b -jets. Only applicable to all-hadronic channel where it alters the fit to the reconstructed mass distribution. Three different sources of systematic uncertainties have been considered, that is: the uncertainties on branching fractions of semi-leptonic decays of b and c -quarks; the uncertainties on b -quark fragmentation parameters; and the uncertainty on the calorimeter response to b and c hadrons.
- **$t\bar{t}$ PDF:** The uncertainties on the parton distribution functions inside the proton alter the probability to find a particular parton inside the proton carrying a particular fraction of the proton's momentum. This in turn alters the kinematics of final state particles from top quark pair production and decay, and thus changes the selection efficiency. The default acceptance is calculated with leading-order CTEQ5L, the systematic includes the uncertainties evaluated using next-to-leading-order error vectors from CTEQ6M and the difference between the central value for leading-order and next-to-leading order PDFs.
- **$t\bar{t}$ ISR/FSR:** The amount of QCD radiation from partons in the initial (ISR) or final (FSR) state is set by parameters of the PYTHIA generator used to simulate signal events. The variation in these parameters is taken from a study of ISR in Drell-Yan events, $q\bar{q} \rightarrow Z/\gamma \rightarrow \mu\mu$, that share the same initial $q\bar{q}$ state as the majority of the signal.
- **$t\bar{t}$ generator:** the difference between PYTHIA and HERWIG due to different models for hadronization and the underlying event.
- **W decay branching fraction:** uncertainty on the branching fractions from the PDG.
- **Color reconnection:** the difference in selection efficiency between simulation models, which vary parameters affecting the exchange of momentum and energy via gluons between the the top quark decay chain and the color-connected anti-top quark decay chain.
- **Pileup:** multiple $p\bar{p}$ interactions in a single bunch crossing add extra particles to the event, reducing the isolation of the lepton and increasing the amount of energy in the calorimeter.

The number of primary vertices with distinct z -positions is used as an estimator for the number of interactions. Only estimated by all-hadronic at this time.

- **Method:** uncertainty unique to a particular method that does not fit into any other category and that is expected to be uncorrelated with other measurements. For ANN and SVX, this uncertainty is due to MC statistics, while for HAD it is due to calibration and template statistics.
- **Z boson production theoretical uncertainty:** only applicable for ANN and SVX measurements that exploit the ratio of the observed $t\bar{t}$ production rate to the observed Z boson production rate, then normalize by the theoretical prediction for Z boson production. This reduces the overall uncertainty upon the integrated luminosity.
- **Luminosity $p\bar{p}$ total inelastic cross-section:** the standard uncertainty on the luminosity from the 4% contribution from the total inelastic $p\bar{p}$ cross-section. Only applicable for DIL and HAD measurements that do not use the ratio of $t\bar{t}$ to Z production.
- **Luminosity detector:** the standard uncertainty on the luminosity from the 4.2% contribution from detector effects [15]. 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. For measurements that do use the ratio of $t\bar{t}$ to Z production, this is the residual systematic.
- **Background SVX c -tag:** uncertainty on mis-tag charm background.
- **Background SVX mis-tag:** uncertainty on mis-tag light flavor background.
- **Background W+HF:** uncertainty on W+HF fraction.
- **Background QCD/Fake:** uncertainty on multi-jet QCD background.
- **Background EWK:** uncertainty on single top, diboson backgrounds.
- **Background other:** uncertainty on dilepton background (no decomposition available but largely uncorrelated with others); uncertainty on all-hadronic background; uncertainty for ANN on Q^2 scale for W +jets background.

- **Background Z:** only applicable for ANN and SVX measurements that make a ratio to Z boson production. This is the uncertainty on Q^2 scale for Z +jets and on the background in Z measurement.

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

Correlations between measurements for a particular category of systematic uncertainty are indicated by \dagger in Table II. We treat the lepton trigger and identification, and lepton energy scale uncertainties as correlated between the techniques using a lepton trigger and lepton identification, that is DIL, ANN, and SVX. We treat the uncertainty on the SVX b -tag efficiency as correlated between measurements that use SVX (DIL, SVX, and HAD). We treat the uncertainties on primary vertex, jet energy calibration, and $t\bar{t}$ model (categories 8-13), which are estimated in a common way for all the techniques, as correlated across all measurements. We treat the uncertainties on detector luminosity as correlated across all measurements, and on the $p\bar{p}$ total inelastic cross-section as correlated across measurements that do not use the normalization to the Z boson rate (DIL, HAD). The uncertainties related to the normalization to Z boson production are taken as correlated across the ANN and SVX measurements, these are from the Z boson theoretical prediction, track identification, and Z background.

The W branching fraction is taken as anti-correlated between all-hadronic and the other channels. The all-hadronic branching fraction is the largest, so if it is increased, then the branching fraction for the other two channels is decreased.

B. Statistical correlations

The two 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 $\mathcal{L}=4600 \text{ pb}^{-1}$. 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 , \cancel{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. For the background processes, we neglect the correlation between the the value of the ANN output and the H_T , \cancel{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.

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, we use the observed number of tags and notags in each pseudoexperiment. We estimate a statistical correlation between ANN and SVX of 32%. When we do not account for the correlation between the ANN output, H_T , \cancel{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%.

Therefore, for the combination, we use a statistical correlation between ANN and SVX as $\rho = 32\%$. 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 four CDF measurements of the top quark pair production cross section, assuming a top quark mass of $172.5 \text{ GeV}/c^2$. 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:

$$\sigma(p\bar{p} \rightarrow t\bar{t}) = 7.50 \pm 0.31_{(stat)} \pm 0.34_{(syst)} \pm 0.15_{(lumi)} \text{ pb.}$$

The total uncertainty is 0.48 pb. The ANN result carries the largest weight of 70%, followed by SVX with 18% and DIL with 18%, then HAD with -6%. The negative weight for HAD is due to the large correlation with other measurements and its larger uncertainty; thus the HAD value is likely to be further away from the true value and collects a negative weight. The complete breakdown of the uncertainty into the previously described categories is also shown in Table II. The contributions to the total uncertainty from acceptance-like uncertainties have been evaluated with respect to the combined value of 7.50 pb.

The total correlation matrix, including statistical and systematic effects, is given in Table III. The largest correlation is 50% 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 four measurements is 0.60 for 3 degrees of freedom, corresponding to a probability of 90% 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 black line. As also shown in Figure 1, we find excellent agreement with the theoretical prediction for a top quark mass of $172.5 \text{ GeV}/c^2$ and centre-of-mass energy of $\sqrt{s} = 1.96 \text{ TeV}$

Using the information given in the theoretical references about the dependence of the prediction on the value of the top quark mass, we re-evaluate here the predictions for a top quark mass

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

Correlation	ANN	SVX	DIL	HAD
ANN	1	0.50	0.17	0.34
SVX		1	0.30	0.48
DIL			1	0.38

TABLE IV: For top quark mass value of $172.5 \text{ GeV}/c^2$, the cross-section, uncertainty, weight, and pull of each result in the combination.

Quantity	DIL	ANN	SVX	HAD
$\sigma_{t\bar{t}}$ (pb)	7.27	7.63	7.14	7.21
pre-BLUE total uncertainty (pb)	0.95	0.53	0.69	1.28
post-BLUE total uncertainty (pb)	0.96	0.53	0.71	1.31
Weight (%)	17.9	69.7	18.4	-6.0
Pull	-0.24	+0.24	-0.51	-0.22

of value of $172.5 \text{ GeV}/c^2$. The NNLO approximate calculation from Moch and Uwer [1] yields $7.46^{+0.66}_{-0.80}$ pb, the prediction from Cacciari et al. [3] is $7.26^{+0.78}_{-0.86}$ pb, and from Kidonakis and Vogt [4] is $7.29^{+0.79}_{-0.85}$ pb, where uncertainties are added in all cases.

A comparison of the combined cross-section result against theoretical predictions as a function of the center-of-mass energy is shown in Figure 2. At $\sqrt{s} = 1960 \text{ GeV}$, the central values of the predictions from [3] are shown at a top quark mass of 170, 172.5, and 175 GeV/c^2 for CTEQ6.6. The uncertainty (due to scale and PDF variations) of the 172.5 CTEQ6.6 prediction is shown by the filled blue region. At $\sqrt{s} = 1800 \text{ GeV}$, the predictions are for 170 and 175 GeV/c^2 , using CTEQ6M.

The combined value of the cross-section at $172.5 \text{ GeV}/c^2$ is shown in Figure 3 with the theoretical predictions as a function of top quark mass.

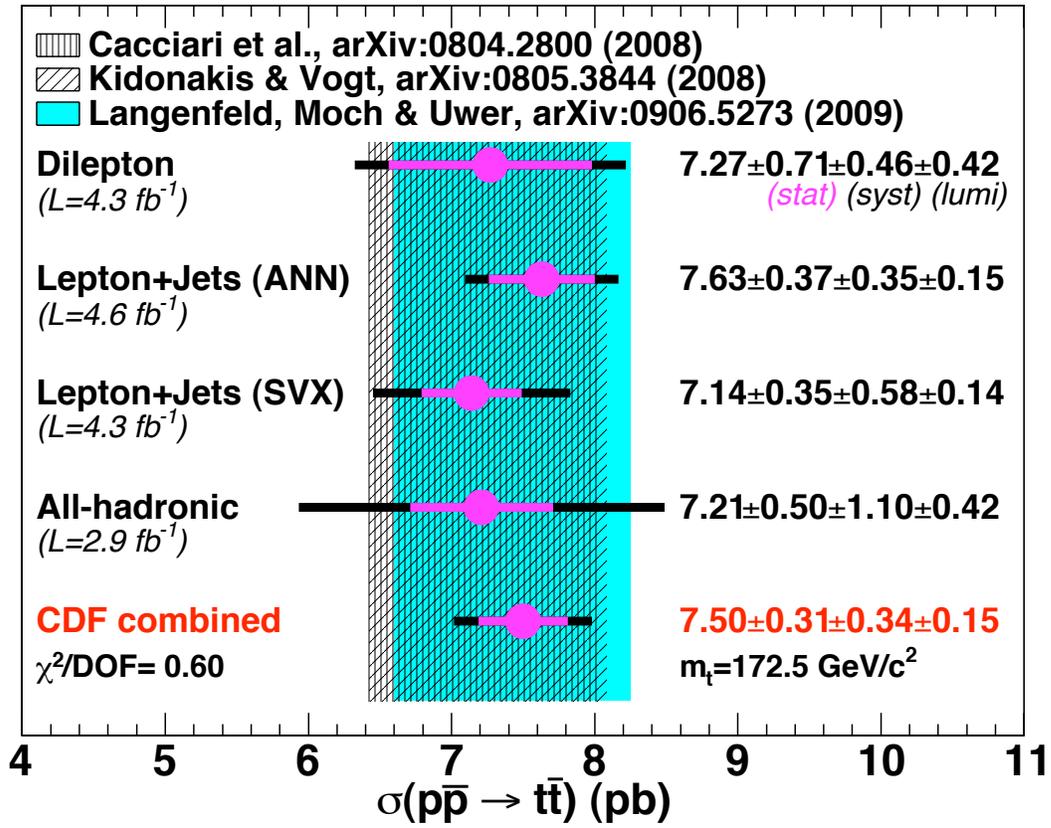


FIG. 1: Combination of CDF measurements of the top quark pair production cross section at a top quark mass of $172.5 \text{ GeV}/c^2$. The theoretical predictions use the CTEQ6.5 set for Cacciari et al., CTEQ6.6M for Kidonakis and Vogt, and CTEQ6.6M for Moch and Uwer.

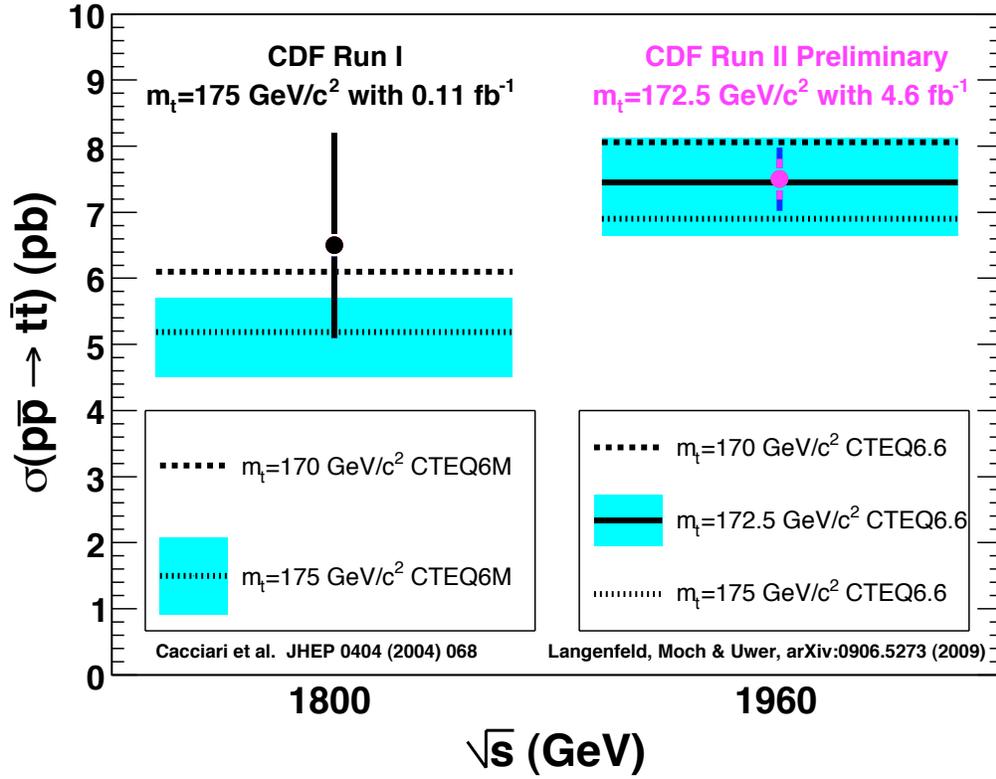


FIG. 2: Combination of CDF measurements of the top quark pair production cross section versus the center-of-mass energy of the Tevatron. The predictions from [3] are shown at $\sqrt{s}=1800$ GeV, and the most recent from Langerfeld, Moch & Uwe [1] at $\sqrt{s}=1960$ GeV.

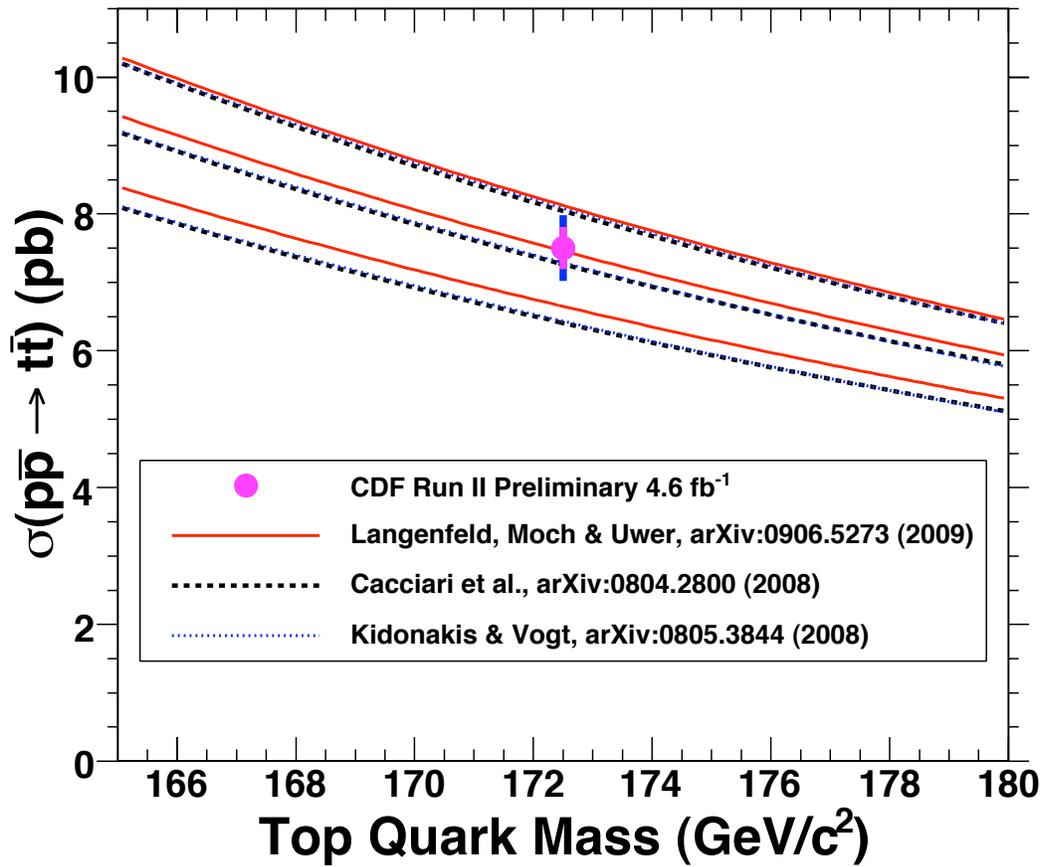


FIG. 3: Combination of CDF measurements of the top quark pair production cross section at 172.5 GeV/c² versus the theoretical predictions as a function of top quark mass.

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 4600 pb⁻¹ collected by the CDF II detector at the Fermilab Tevatron.

The data result is:

$$\sigma_{t\bar{t}} = 7.50 \pm 0.48 \text{ pb for } m_{top} = 175 \text{ GeV}/c^2.$$

The statistical uncertainty is 0.31 pb, the experimental systematic uncertainty is 0.34 pb, and the luminosity uncertainty is 0.15 pb. The relative uncertainty on this determination is 6.4%. We find good agreement among four measurements in the dilepton, lepton+jets, and all-hadronic channels, with a probability of about 90% to find a less consistent set of measurements. In comparison to the single measurement with the smallest total uncertainty, that is ANN with 7.63 ± 0.53 pb with a precision of 6.9%, we also find an improvement in precision of 9%.

The precision of this combined measurement of the top quark pair production measurement is now better than that of the theoretical prediction. We find good agreement with the theoretical prediction for a top quark mass of 172.5 GeV/ c^2 and centre-of-mass energy of $\sqrt{s} = 1.96$ TeV of $7.46_{-0.80}^{+0.66}$ pb from [1]. Note that the scale and PDF uncertainties have been added linearly here.

-
- [1] U. Langenfeld, S. Moch and P. Uwer, arXiv:0906.5273 (2009).
 - [2] S. Moch and P. Uwer, arXiv:0807.2794 (2008) and arXiv:0804.1476 (2008).
 - [3] 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).

- [4] N. Kidonakis and R. Vogt, arXiv:0805:3844 (2008);N. Kidonakis and R. Vogt, Phys. Rev. D **68**, 114014 (2003).
- [5] R. Rossin *et al.*, CDF note 8044.
- [6] A. Abulencia *et al.* (CDF Collaboration), Phys. Rev. Lett. **96**, 042003 (2006).
- [7] A. Varganov *et al.*, CDF conference note 9890.
- [8] A. Lister *et al.*, CDF conference note 9950.
- [9] T. Schwarz *et al.*, CDF conference note 9878.
- [10] A. Castro *et al.*, CDF conference note 9841.
- [11] L. Lyons, D. Gibaut and P. Clifford, NIM **A270**, 110-117 (1988).
- [12] L. Lyons, A. Martin and D. Saxon, Phys. Rev. D **41**, 3 (1990).
- [13] A. Valassi, NIM **A500**, 391 (2003).
- [14] A. Bhatti *et al.*, arXiv:hep-ex/0510047.
- [15] S.Jindariani *et al.*, CDF note 7446.