



DØnote 5591-CONF

Measurement of the production cross section for $t\bar{t}$ pairs in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV

The DØ Collaboration
URL <http://www-d0.fnal.gov>
(Dated: February 24, 2008)

We measure the $t\bar{t}$ production cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV in the lepton+jets channel. Two complementary methods discriminate between signal and background, b -tagging and a kinematic likelihood discriminant. Based on about 0.9 fb^{-1} of data collected by the D0 detector, we measure $\sigma_{t\bar{t}} = 7.77 \pm 0.86 \text{ pb}$, assuming the current world average $m_{top} = 170.9 \text{ GeV}$.

The standard model fixes all properties of the top quark except its mass. The cross section with which top quarks are produced depends on the interaction of the top quark with the initial state particles and on its mass. By comparing the measured cross section to predictions we can test whether the top quark behaves as expected.

The Tevatron collides protons and antiprotons at $\sqrt{s} = 1.96$ TeV. Most top quarks at the Tevatron are created in $t\bar{t}$ pairs through the strong interaction, although evidence of single top quark production has been observed recently [1]. For a top quark mass of 175 GeV, the standard model prediction for the $t\bar{t}$ pair production cross section is about 6.7 pb [2, 3]. Past measurements [4, 5] are in agreement with this prediction within their precision of about 15%. Here we present the most precise measurement of the $t\bar{t}$ pair production cross section to date. It is based on data collected by the D0 detector [6] between August 2002 and December 2005 with an integrated luminosity of about 0.9 fb^{-1} .

The standard model predicts that the top quark decays always to a W boson and a b quark. The decay modes of the W boson define the possible final states. Here we focus on the lepton+jets decay channel in which one W boson decays to $e\nu$, $\mu\nu$, or $\tau\nu$ followed by $\tau \rightarrow e\nu\bar{\nu}$ or $\mu\nu\bar{\nu}$. We will refer to such leptons as prompt. The other W boson decays to jets or to $\tau\nu$ followed by a hadronic decay of the τ . The branching fraction for this channel is about 38%.

The event selection requires one isolated electron or muon with a momentum component transverse to the beam direction $p_T > 20$ GeV and $|\eta| < 1.1$ (for e) or $|\eta| < 2$ (for μ) [7], at least three jets with $p_T > 20$ GeV and $|\eta| < 2.5$ of which the leading jet has $p_T > 40$ GeV, and missing transverse momentum $\cancel{p}_T > 20$ GeV (for e +jets) or 25 GeV (for μ +jets). The lepton must point away from \cancel{p}_T in azimuth to reject background events with mismeasured particles. Jets are reconstructed using the Run II cone algorithm [8] with cone size $\sqrt{(\Delta\phi)^2 + (\Delta y)^2} = 0.5$. We call this the inclusive lepton+jets sample and Table I gives the number of events selected (N_{data}). The expected top signal accounts only for about 17% of this sample. Most events originate from other processes that produce prompt leptons and jets (mostly W +jets production) and from events with jets which mimic the signature of a lepton. We use two complementary techniques to distinguish the $t\bar{t}$ signal from these backgrounds, b -tagging and a kinematic likelihood discriminant.

The b -tagging analysis exploits the fact that every $t\bar{t}$ decay produces two b quarks. Therefore we enhance signal/background by requiring that at least one jet be tagged as a b -jet, i.e., identified to contain the decay of a long-lived particle such as a b -hadron [10].

Starting with the inclusive lepton+jets sample, we first determine the background from events without prompt leptons from the data. For this purpose we define loose data samples by relaxing the electron identification and the muon isolation requirements. We use simulated events to determine the probability ϵ_s for leptons from W boson decays that satisfy the loose selection to also pass the selection used for the measurement. We correct this efficiency for known differences between efficiencies observed in the Monte Carlo simulation and in data. We determine the corresponding efficiency ϵ_b for misidentified leptons using data selected with the criteria given above except for requiring $\cancel{p}_T < 10$ GeV to minimize contributions from leptons from W boson decays. We write the number of events in our selected sample as $N_{\text{data}} = N_{\ell+\text{jets}} + N_{\text{jj}}$, where $N_{\ell+\text{jets}}$ is the number of events with prompt leptons and N_{jj} the number of events without prompt leptons. The number of events in the corresponding loose sample is then $N_{\text{loose}} = N_{\ell+\text{jets}}/\epsilon_s + N_{\text{jj}}/\epsilon_b$. These two equations determine N_{jj} , given in Table I.

We model the $t\bar{t}$ signal and all backgrounds with prompt leptons using Monte Carlo simulations. W +jets and Z +jets production are generated using the ALPGEN [11] generator and PYTHIA [12] for showering. A matching algorithm [13] avoids double counting of final states. Single top production is generated using SINGLETOP [14] and COMHEP [15] as described in Ref. [1]. Diboson and $t\bar{t}$ production are generated by PYTHIA. All simulated events are processed by a detector simulation based on GEANT [16] and the standard D0 reconstruction programs.

We predict the number of events, N_{other} in Table I, from the smaller background processes (single top, Z +jets, and diboson production) using the Monte Carlo simulation and next-to-leading order (NLO) cross sections [17]. This leaves the number of $t\bar{t}$ events, $N_{t\bar{t}}$, and W +jets events, $N_{W+\text{jets}}$, to be determined. We start with the expected $t\bar{t}$ cross section to get a first estimate of the number of $t\bar{t}$ events in the sample. After we obtain a cross section from the

	$e+3$ jets	$e+\geq 4$ jets	$\mu+3$ jets	$\mu+\geq 4$ jets
N_{data}	1300	320	1120	306
N_{loose}	2592	618	1389	388
$\epsilon_s(\%)$	84.8 ± 0.3	84.0 ± 1.8	87.3 ± 0.5	84.5 ± 2.2
$\epsilon_b(\%)$	19.5 ± 1.7	19.5 ± 1.7	27.2 ± 5.4	27.2 ± 5.4
$N_{t\bar{t}}$	154 ± 17	132 ± 14	115 ± 13	109 ± 12
$N_{W+\text{jets}}$	746 ± 41	93 ± 18	824 ± 24	151 ± 14
N_{other}	132 ± 15	35 ± 4	139 ± 15	36 ± 4
N_{jj}	268 ± 34	60 ± 10	42 ± 14	10 ± 6

TABLE I: Event counts in the inclusive lepton+jets sample.

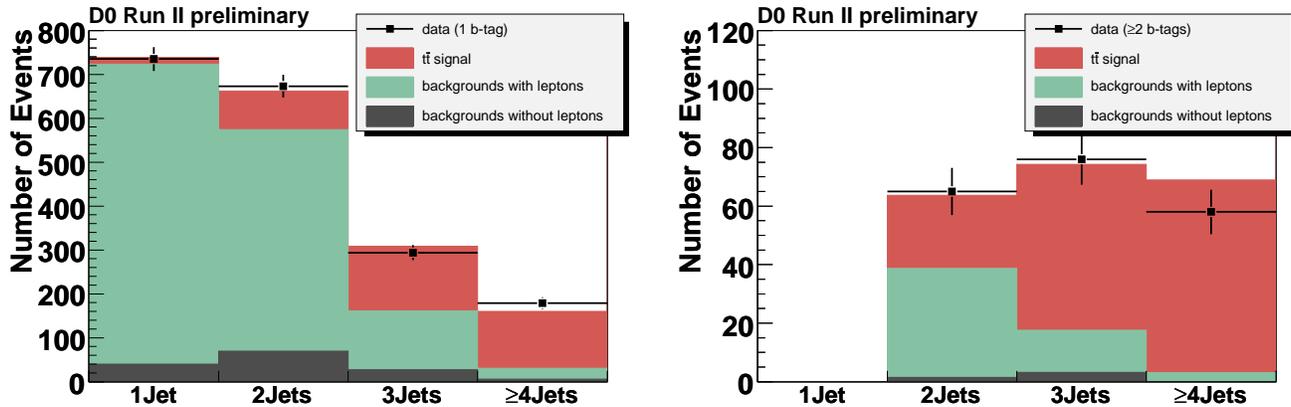


FIG. 1: Jet multiplicity spectra for e +jets and μ +jets events with one b -tagged jet (left) and with at least 2 b -tagged jets (right).

	$e+3$ jets	$e+\geq 4$ jets	$\mu+3$ jets	$\mu+\geq 4$ jets
$N_{t\bar{t}}$	71 ± 8	60 ± 7	53 ± 6	50 ± 5
$N_{W+\text{jets}}$	50 ± 4	8 ± 2	58 ± 3	12 ± 1
N_{other}	14 ± 2	4 ± 0	13 ± 1	4 ± 0
N_{jj}	23 ± 5	6 ± 3	4 ± 2	-1 ± 1
total	158 ± 10	78 ± 7	128 ± 7	65 ± 6
N_{data}	164	88	130	91

TABLE II: Numbers of events with one b -tagged jet.

	$e+3$ jets	$e+\geq 4$ jets	$\mu+3$ jets	$\mu+\geq 4$ jets
$N_{t\bar{t}}$	27 ± 4	30 ± 4	21 ± 3	26 ± 4
$N_{W+\text{jets}}$	5 ± 1	1 ± 0	5 ± 1	1 ± 0
N_{other}	2 ± 0	1 ± 0	2 ± 0	1 ± 0
N_{jj}	2 ± 2	1 ± 1	1 ± 1	-1 ± 1
total	36 ± 4	33 ± 4	29 ± 3	27 ± 4
N_{data}	41	26	35	32

TABLE III: Number of events with two or more b -tagged jets.

b -tagged samples as described below we change this to the measured cross section and iterate until the cross section result is stable. We choose the number of W +jets events in the inclusive sample so that the sum of all background and signal contributions equals the observed number of events.

Then we consider events that contain at least one b -tagged jet. We determine the number of background events without prompt leptons as above and the number of events from other background sources from the number of background events in the inclusive sample times their probability to be b -tagged. We obtain the b -tagging probability from the Monte Carlo simulation corrected for differences in the b -tagging efficiencies observed in the simulation and in data. In order for the Monte Carlo model to correctly predict the number of lepton+jets events with two jets in which at least one jet is b -tagged, we have to scale the number of W +jets events with heavy quarks (b , c) by a factor of 1.17 ± 0.18 relative to the rest of the W +jets events. We use the same scale factor for lepton+jets events with three or more jets. Figure 1 shows the jet multiplicity spectrum of events with b -tags, compared to the prediction from our model. The composition of the b -tagged samples is given in Tables II and III. The $t\bar{t}$ contribution in Fig. 1 and Tables I-III is based on the cross section measured in the b -tag analysis.

We calculate the cross section using a maximum likelihood fit to the number of events in eight different channels defined by lepton flavor (e , μ), jet multiplicity (3 , ≥ 4), and b -tag multiplicity (1 , ≥ 2). The likelihood is defined as $\mathcal{L} = \prod_i \mathcal{P}(N_i, \mu_i(\sigma_{t\bar{t}}))$, where i runs over the eight channels and $\mathcal{P}(N, \mu)$ is the Poisson probability to observe N events when μ are expected. The expected number of events is the sum of the number of events from all backgrounds plus the number of $t\bar{t}$ events as a function of $\sigma_{t\bar{t}}$. This procedure is fully described in Ref. [18]. We obtain $\sigma_{t\bar{t}} = 8.05 \pm 0.54(\text{stat}) \pm 0.70(\text{syst}) \pm 0.49(\text{lumi})$ pb, assuming $m_{\text{top}} = 175$ GeV. Table IV gives the breakdown of the systematic uncertainties. They are explained below.

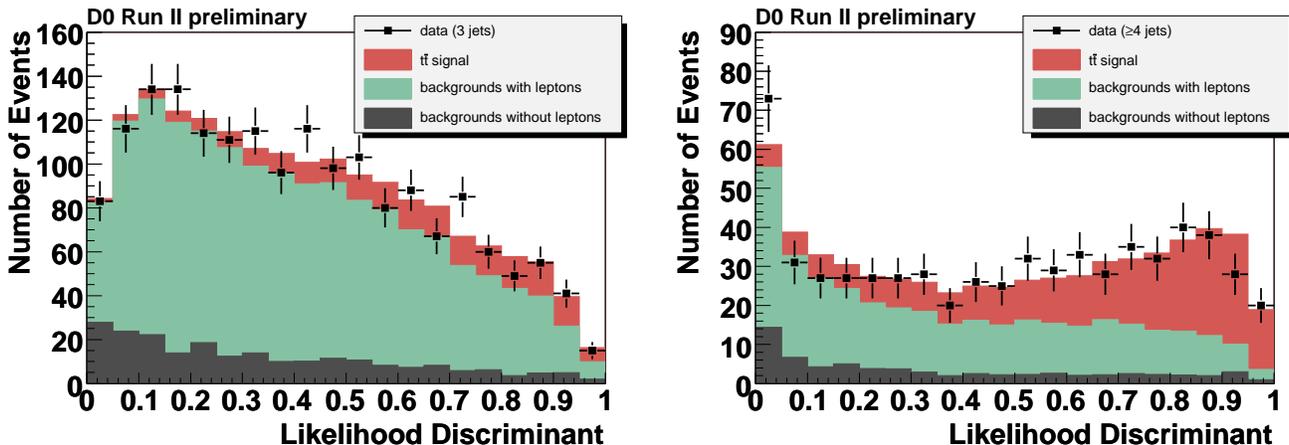
Events with $t\bar{t}$ decays also differ in their event kinematics from background events. No single kinematic quantity, however, can separate signal and background very well. We therefore build a likelihood discriminant from 5-6 such variables, listed in Table V, in each channel to discriminate the $t\bar{t}$ signal from the backgrounds. The variables were selected to be well modelled by the Monte Carlo simulation and to have good discrimination power. For this analysis, four channels are defined by lepton flavor and jet multiplicity (3 , ≥ 4). Events with three jets must satisfy $\sum_{i=1}^{N_j} p_T(i) > 120$ GeV in addition to the selection described above.

We determine the probability density functions of the likelihood discriminant for signal and prompt lepton back-

source	b -tag	likelihood	combined
selection efficiency	0.26 pb	0.25 pb	0.25 pb
jet energy calibration	0.30 pb	0.11 pb	0.20 pb
b -tagging	0.48 pb	—	0.24 pb
MC model	0.29 pb	0.11 pb	0.19 pb
N_{jj}	0.06 pb	0.10 pb	0.07 pb
likelihood fit	—	0.15 pb	0.08 pb

TABLE IV: Breakdown of systematic uncertainties.

variable	channel
$\sum_{i=3}^{N_j} p_T(i)$	all
$\sum_{i=1}^{N_j} p_T(i) / \sum_{i=1}^{N_j} p_z(i)$	$e+3$ jets, $e+\geq 4$ jets
$\sum_{i=1}^{N_j} p_T(i) + p_T(e/\mu) + \cancel{p}_T$	$e+3$ jets, $e+\geq 4$ jets
ΔR between lepton and jet 1	all
ΔR between jets 1 and 2	$e+\geq 4$ jets, $\mu+\geq 4$ jets
$\Delta\phi$ between lepton and \cancel{p}_T	$\mu+3$ jets, $\mu+\geq 4$ jets
$\Delta\phi$ between jet 1 and \cancel{p}_T	$e+3$ jets, $\mu+3$ jets
sphericity	all but $\mu+3$ jets
aplanarity	all but $\mu+3$ jets

TABLE V: Variables used for the likelihood discriminant, where i indexes the list of N_j jets with $p_T > 15$ GeV, ordered in decreasing p_T , and $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$.FIG. 2: Likelihood discriminant spectra for e +jets and μ +jets events with 3 jets (left) and events with at least 4 jets (right).

grounds from the simulation and for events without prompt leptons from a control data sample. We perform a maximum likelihood fit to the distribution of the likelihood discriminant from data in all four channels simultaneously with the $t\bar{t}$ pair production cross section as a free parameter and constraining the number of events without prompt leptons to the value obtained from the loose data sample as described above. Table VI gives the sample composition returned by the fit and Figure 2 shows the likelihood discriminant distributions for the best fit. We measure $\sigma_{t\bar{t}} = 6.62 \pm 0.78(\text{stat}) \pm 0.36(\text{syst}) \pm 0.40(\text{lumi})$ pb, assuming $m_{top} = 175$ GeV. The systematic uncertainties are listed in Table IV.

Systematic uncertainties arise from the following main categories. *Selection* covers the uncertainties in acceptance and efficiency for leptons and jets. *Jet energy calibration* describes the effect of uncertainties in the jet energy scale and resolution. The uncertainty in the b -tagging efficiencies for b , c , and light quark/gluon jets make up the *b -tagging uncertainty*. *Monte Carlo model* uncertainties originate from uncertainties in the cross sections used to normalize the simulated backgrounds, differences observed between $t\bar{t}$ samples generated with ALPGEN and PYTHIA, the factorization and renormalization scale in the W +jets simulation, and the parton distributions functions (pdf). N_{jj} covers the uncertainties in the determination of the number of events without prompt leptons. *Likelihood fit* gives

	$e+3$ jets	$e+\geq 4$ jets	$\mu+3$ jets	$\mu+\geq 4$ jets
N_{data}	948	320	812	306
$N_{t\bar{t}}$	141 ± 16	128 ± 15	104 ± 12	105 ± 12
$N_{W+\text{jets}} + N_{\text{other}}$	622 ± 36	131 ± 22	672 ± 32	190 ± 21
N_{jj}	193 ± 26	60 ± 10	34 ± 11	10 ± 6

TABLE VI: Sample composition from the likelihood fit.

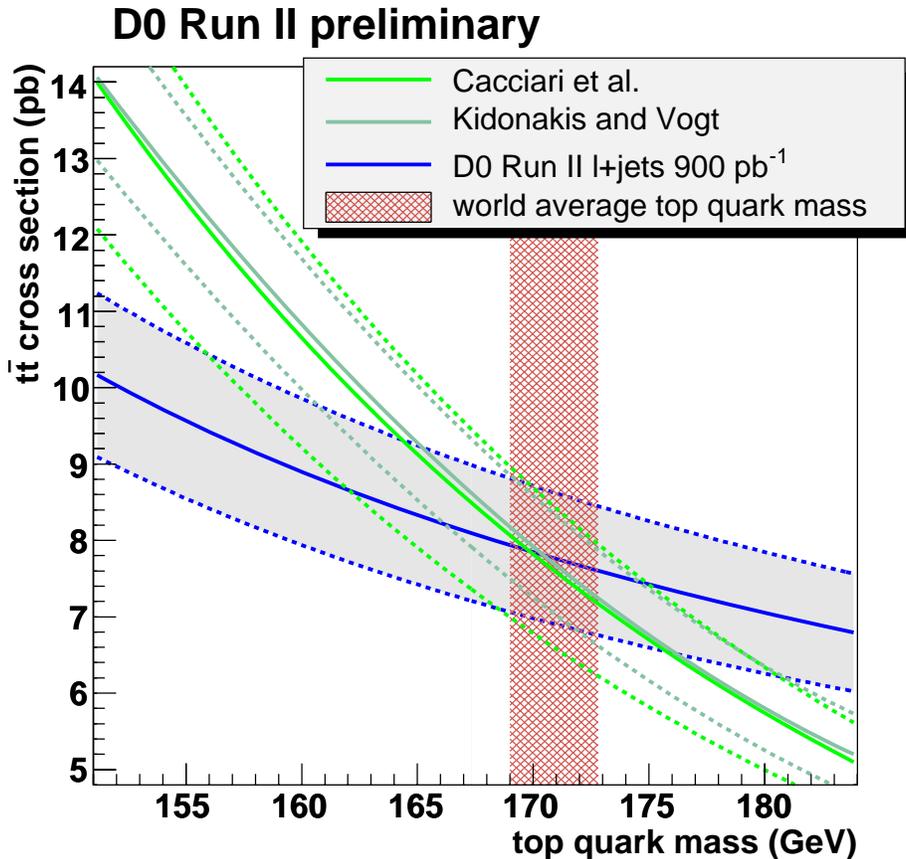


FIG. 3: Comparison of measured cross section and theory prediction versus top quark mass. The shaded region indicates the 68% C.L. band for our measurement.

the uncertainty in the result of the likelihood fit from statistical fluctuations in the likelihood discriminant shapes determined from the Monte Carlo simulation.

To obtain the best measurement of the $t\bar{t}$ pair production cross section, we combine the results from the two complementary analyses using the BLUE method [19]. The statistical correlation factor for the two analyses is 0.31, determined by Monte Carlo generated pseudodata sets that model the statistical correlation between the two analyses. The systematic uncertainties from each source are completely correlated between the two analyses. The combined result is $\sigma_{t\bar{t}} = 7.42 \pm 0.53(\text{stat}) \pm 0.46(\text{syst}) \pm 0.45(\text{lumi})$ pb for $m_{top} = 175$ GeV with $\chi^2 = 2$, corresponding to a p-value of 0.16.

We use simulated samples of $t\bar{t}$ events generated at different values of the top quark mass to determine the measured cross section as a function of top quark mass. A polynomial fit as a function of the top quark mass gives $\sigma_{t\bar{t}}/\text{pb} = 7.42 - 7.9 \times 10^{-2}\Delta m + 9.7 \times 10^{-4}(\Delta m)^2 - 1.7 \times 10^{-5}(\Delta m)^3$, where $\Delta m = m_{top}/\text{GeV} - 175$ GeV. Figure 3 shows this parametrization and the theoretical predictions versus top quark mass. Our measurement is in good agreement with the prediction at the current world average of direct measurements of the top quark mass of 170.9 ± 1.8 GeV [21]. We thus conclude that the strong interactions of the top quark are consistent with QCD predictions. If we assume $m_{top} = 170.9$ GeV we obtain the cross section $\sigma_{t\bar{t}} = 7.77 \pm 0.86$ pb. This is the most precise measurement of the $t\bar{t}$ pair production cross section to date.

-
- [1] D0 Collaboration, V.M. Abazov *et al.*, Phys. Rev. Lett. **98**, 181802 (2007).
 [2] M. Cacciari *et al.*, JHEP **404**, 068 (2004).
 [3] N. Kidonakis and R. Vogt, Phys. Rev. D **68**, 114014 (2003).

- [4] D0 Collaboration, V.M. Abazov *et al.*, Phys. Rev. D **74**, 112004 (2006) and Phys. Rev. D **76**, 052006 (2007).
- [5] CDF Collaboration, A. Abulencia *et al.*, Phys. Rev. Lett. **97**, 082004 (2006).
- [6] D0 Collaboration, V.M. Abazov *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **565**, 463 (2006).
- [7] Pseudorapidity $\eta = -\ln[\tan(\theta/2)]$, where θ is the polar angle with the proton beam direction.
- [8] G. Blazey *et al.*, in *Proceedings of the Workshop: "QCD and Weak Boson Physics in Run II,"* edited by U. Baur, R.K. Ellis and D. Zeppenfeld, FERMILAB-Pub-00/297 (2000).
- [9] D0 Collaboration, V.M. Abazov *et al.*, to be published in Phys. Rev. D, arXiv:0705.2788.
- [10] Tim Scanlon "*b*-tagging and the search for neutral supersymmetric Higgs bosons at DØ; Ph.D. thesis, FERMILAB-THESIS-2006-43.
- [11] M.L. Mangano *et al.*, JHEP **307**, 001 (2003).
- [12] T. Sjöstrand *et al.*, PYTHIA 6.3: PHYSICS AND MANUAL, hep-ph/0308153 (2003).
- [13] S. Höche *et al.*, arXiv:hep-ph/0602031 (2004).
- [14] E.E. Boos *et al.*, Phys. Atom. Nucl. **69**, 1317 (2006).
- [15] E.E. Boos *et al.* (CompHEPCollaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **534**, 250 (2004).
- [16] R. Brun and F. Carminati, CERN Program Library Long Writeup W5013 (1993).
- [17] E.E. Boos *et al.*, Phys. Atom. Nucl. **69**, 1317 (2006); Z. Sullivan, Phys. Rev. D **70**, 114012 (2004); J.M. Campbell and R.K. Ellis, Phys. Rev. D **60**, 113006 (1999).
- [18] D0 Collaboration, V.M. Abazov *et al.*, Phys. Rev. D **74**, 112004 (2006).
- [19] L. Lyons, D. Gibaut and P. Clifford, Nucl. Instrum. Meth. A **270**, 110 (1988); A. Valassi, Nucl. Instrum. Meth. A **500**, 391 (2003).
- [20] D. Stump *et al.*, JHEP **0310**, 046, (2003).
- [21] D0 and CDF Collaborations, "A Combination of CDF and D0 results on the mass of the top quark", arXiv:hep-ex/0703034.

Appendix: additional plots

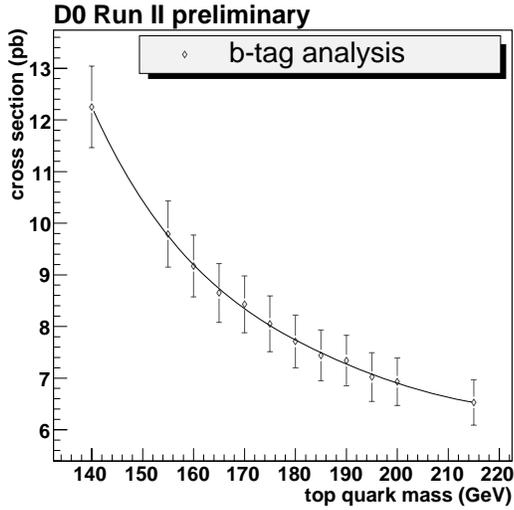


FIG. 4: Measured $t\bar{t}$ pair production cross section versus top quark mass from the b -tag analysis. Errors are statistical only.

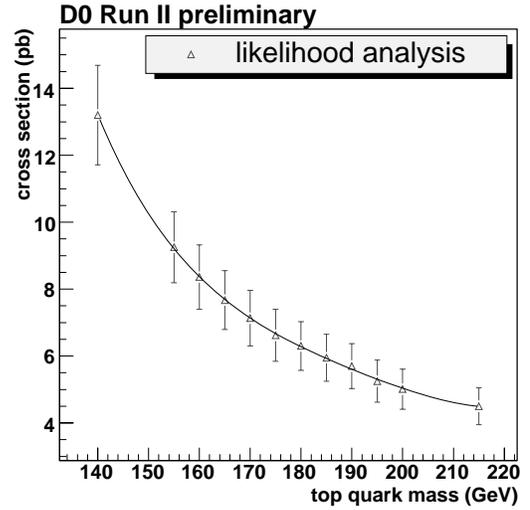


FIG. 5: Measured $t\bar{t}$ pair production cross section versus top quark mass from the likelihood analysis. Errors are statistical only.

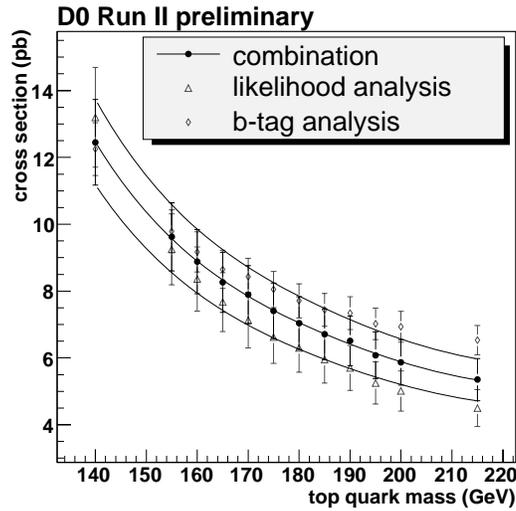


FIG. 6: Measured $t\bar{t}$ pair production cross section versus top quark mass. For the results of the individual analyses the statistical errors are plotted only. For the combined result the total error is plotted. The smooth lines are polynomial fits.