A Precision Top Quark Mass Measurement In the Lepton + Jets Channel at CDF II

The CDF Collaboration
URL http://www-cdf.fnal.gov
(Dated: July 20, 2005)

We report measurements of the top quark mass with the CDF II detector at Tevatron using pair-produced top quark events in a 318 pb^{-1} data sample observed in the lepton+jets final state. One technique uses an event-based likelihood technique that incorporates the full matrix element for the production and decay process, resulting in a top quark mass measurement of 173.8 $^{+2.7}_{-2.5}$ (stat.) $\pm$ 3.3 (syst.) GeV/c^2. The second technique determines the reconstructed top quark mass from each event, and then extracts the top quark mass from a fit to the distribution of these masses. The measured invariant mass of the hadronically decaying W boson is used to constrain the jet energy scale, JES. The latter technique results in the most precise measurement of the top quark mass of 173.5 $^{+3.7}_{-3.6}$ (stat. + JES) $\pm$ 1.5 (other syst.) GeV/c^2.

PACS numbers: Valid PACS appear here

The top quark is the heaviest known elementary particle, with a mass that is 35 times that of the next-heaviest fermion. Because of this anomalously large mass, top quark studies may provide insight into our understanding of mass in general, and test theories that explain the large mass heirarchy among quarks and leptons and the interactions that give rise to such disparate properties. Within the context of the Standard Model of particle physics, as the largest fermion mass in the theory, the top quark mass is correlated with the masses of the W intermediate vector boson and the Higgs boson, the latter object being the key to our understanding of the origin of mass [1]. Precision measurements of the top quark and W boson masses test the consistency of the Standard Model, and in particular the Higgs mechanism for spontaneous symmetry breaking in the theory. Improved measurement of the top quark mass is therefore one of the key goals of the experiments being performed at the Fermilab Tevatron Collider using 1.96 TeV proton-antiproton collisions.

This Letter reports two measurements of the top quark mass in the lepton + jets decay channel using the Collider Detector at Fermilab (CDF) with 318 pb^{-1} of collision data collected during run II between February 2002 and August 2004. We focus on the lepton + jets decay channel, which results from $t\bar{t}$ pair production and the subsequent fully hadronic decay of one heavy quark and the semi-leptonic decay of the other. This is the largest decay channel with good signal-to-background, and previous work has shown that the most accurate mass measurements are possible in this decay mode. In general, $t\bar{t}$ events observed in the lepton + jets channel contain an electron or muon, and a neutrino, both from the leptonic W decay, two light-quark jets from the hadronic W decay, and two additional jets arising from b quarks. We select events consistent with this decay topology, and then analyze them using two complementary techniques.

The first technique uses an event-by-event likelihood analysis employing the full matrix element for the production and decay to extract a likelihood distribution as a function of the true top quark mass, $\mathrm{M_{top}}$, for each event. This technique, known as the "dynamical likelihood method" or DLM, was developed by the CDF collaboration [2] and is similar to the method used by the DØ collaboration to make the previous most precise measurement of the top quark mass [3]. A second technique, developed by the CDF collaboration in run I [4], reconstructs a top quark mass, m_t^{reco}, in each event and uses the distribution of m_t^{reco} compared with template distributions derived from model calculations with differing top quark masses to estimate $\mathrm{M_{top}}$. We have improved this technique, known as the "template method," by using the fact that the W boson daughters should form a dijet final state that is consistent with the known W boson mass. This gives an independent constraint on the dominant systematic uncertainty in this measurement, the jet energy scale, that allows us to improve the overall accuracy of the measurement.

The CDF detector is a general-purpose charged and neutral particle detector located at one of the two interaction points along the Tevatron Collider, and is described in detail elsewhere [5]. The detector comprises a solenoidal charged particle spectrometer, consisting of a seven-layer silicon microstrip detector array and a cylindrical drift chamber immersed in a 1.4 T magnetic field, and a segmented sampling calorimeter with scintillator tile readout measuring energy flow up to $|\eta| = 3.6$ and providing electron and photon identification. A set of charged particle detectors outside the calorimeter are used to identify muon candidates. Due to the high intrinsic data rate of collisions, the experiment employs an on-line, three-level trigger system that selects approximately one collision in 20,000 to record for subsequent analysis. Events for this analysis were selected by the

TABLE I: The expected background composition and number of identified events for the $t\bar{t}$-enriched sample used by the template analysis, the subset of those events with ≥ 1 b tag, and the subset of the tagged events used in the DLM analysis.

	≥ 1 b tag	DLM sample
Source	Expected Background	
W + jets	19.6 ± 2.4	5.3 ± 1.1
QCD	4.7 ± 0.7	3.1 ± 1.0
Other	2.3 ± 0.2	0.8 ± 0.1
Total	26.6 ± 3.0	9.2 ± 1.8
	Identified $t\bar{t}$ Candidates	
Data	121	63

presence of a charged electron or muon candidate with $p_T > 20$ GeV/c and $|\eta| < 1$.

The resulting event sample was subsequently reduced by requiring in each event the presence of four or more jets, and by requiring that the missing transverse energy in the event exceed 20 GeV, corresponding to a high-energy neutrino candidate. To reduce backgrounds further, we required either that one of the jets be identified as a bottom quark through the presence of a displaced vertex in the middle of the jet that arises from the decay of the long-lived bottom quark (b tag), or that four jets with $E_T > 21$ GeV be present. In the former case, we also required that there be at least a fourth jet in the event with $E_T > 8$ GeV. In all cases, the jets were required to have pseudorapidity $|\eta| < 2.0$. This selection resulted in 165 events that, based on our background estimates, are primarily $t\bar{t}$ events.

The two analysis methods employ subsets of these data. The DLM technique uses the subset consisting of four and only four jets with $E_T > 15$ GeV with at least one of the four jets having a b tag. This results in 63 events. We have estimated the various sources of background contamination in this sample, summarized in Table I, and find that they total 9.2 ± 1.8 events. The template technique subdivides the data into four subsamples that have different top mass distributions and background levels. Ordered by the resulting statistical power, the four subsamples are 1) events with at least four jets with $E_T > 15$ GeV and one b-tagged jet (1-tag Tight sample with 57 events), 2) events with two b-tagged jets (2-tag sample with 16 events), 3) events with a fourth jet with 8 GeV $< E_T <$ 15 GeV and one b-tagged jet (1-tag Loose sample with 25 events), and 4) events with four jets with $E_T > 21$ GeV and no b-tagged jets (0-tag sample with 40 events). The estimated background rates in the samples with a b tag are summarized in Table I. We used the HERWIG event generator and a detailed detector simulation to model the signature of signal $t\bar{t}$ events [6].

The DLM method, described in detail in [7], defines a likelihood for each event based on the differential cross section for the event as a function of $\mathrm{M_{top}}$, taking into account detector resolution and combinatorial factors by introducing "transfer functions" that reflect these experimental effects. The actual parton kinematics of the $t\bar{t}$ final state are statistically reconstructed by (a) generating a random value for the virtual mass squared of the W boson in the leptonic channel, s_W, according to the Breit-Wigner form; (b) identifying the momentum of the e or μ with the measured value, and the neutrino transverse momentum with the measured transverse energy flow; and (c) generating random values for the momenta of final state quarks according to the transfer function probabilities. The transfer function $w(\boldsymbol{x}, \boldsymbol{y})$ correlates the quark and observed jet transverse energies, denoted by $\boldsymbol{x}$ and $\boldsymbol{y}$ respectively, which we obtain using simulated $t\bar{t}$ events. We determine the z component of the neutrino momentum, with a two-fold ambiguity, by using energy conservation. Thus, for a given set of $\boldsymbol{x}$ and s_W, we unambiguously determine the parton kinematics, and the event likelihood as a function of $\mathrm{M_{top}}$ is given by

$$\boldsymbol{L}(\mathrm{M_{top}}) = N \sum_{I_j} \sum_{I_\nu} \frac{d\sigma_{t\bar{t}}}{d\Phi}(\mathrm{M_{top}}; \boldsymbol{x}, s_W), \qquad (.1)$$

where the normalization factor, N, is constant for a given event, independent of $\mathrm{M_{top}}$, Φ are phase space variables and the indices, I_j and I_ν, run over the parton-jet assignments and the two neutrino solutions, respectively. A numerical integration is performed by repeating this process for each event by generating a large set of random values for s_W and $\boldsymbol{x}$.

A joint likelihood is formed by multiplying the event likelihoods together. We take into account the presence of background events by modeling their effect as a shift on the top quark mass. Figure 1 shows the distribution of the top quark mass value at the point of maximum likelihood in each event, comparing the 63 data events to the expectation from simulated events. An inset shows the jointlikelihood as a function of true top quark mass for the 63 events, from which we infer the maximum likelihood estimate $\mathrm{M_{top}} = 173.8\,^{+2.7}_{-2.5}$ (stat.) GeV/c^2, where the uncertainties are only statistical.

The template method is described in detail in [8]. We perform a χ^2 minimization to fit the parton momentum from the $t\bar{t}$ daughters and determine m_t^{reco} for each event, assuming that the final state arises from the decay of a particle-antiparticle pair into a pair of W bosons and b quarks. We use only the four leading jets in the mass reconstruction. In the χ^2 fit, both sets of W decay daughters are constrained to have the invariant mass of the W boson, and both Wb states are constrained to have the same mass. The ambiguity arising from the different ways of assigning the four jets to the four quarks, taking into account the b-tagging information, is resolved by selecting the assignment with the lowest χ^2. We construct a histogram of m_t^{reco} for each subsample. At the same time,

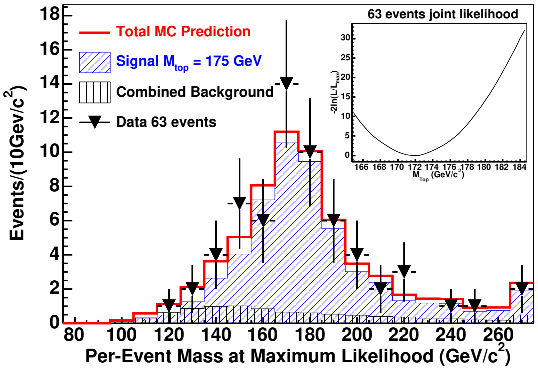

FIG. 1: For each event, the value of the top quark mass at the maximum of the DLM likelihood curve is plotted. Data events are compared to an expected distribution comprising simulated $t\bar{t}$ ($\mathrm{M_{top}}$ =175 GeV/c^2) and background events. The inset shows the joint likelihood for the 63 events.

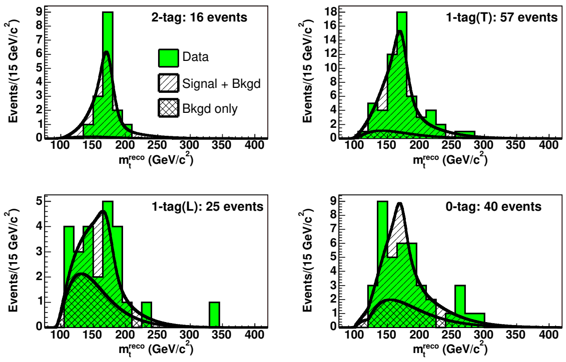

FIG. 2: The reconstructed top quark mass distribution for each subsample is shown overlaid with the expected distribution using the top quark mass, jet energy scale, signal normalization, and background normalization from the combined fit.

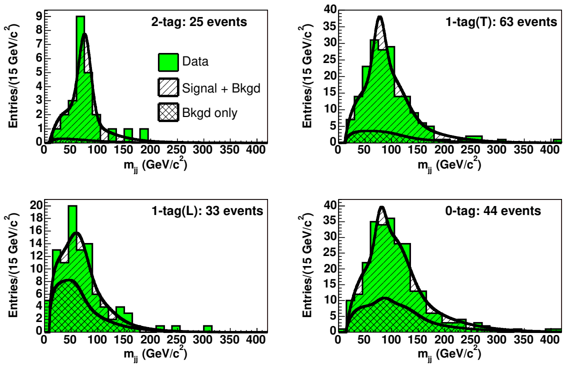

FIG. 3: The reconstructed dijet mass distribution for each subsample is shown overlaid with the expected distribution using the top quark mass, jet energy scale, signal normalization, and background normalization from the combined fit.

we analyze the events in a slightly different way, removing the W boson mass constraints, identifying for each event all combinations of parton-jet assignments that are consistent with the b-tagging information and determining all possible dijet invariant masses arising from the W boson decay. We use these masses to create a second histogram for each subsample, which, given the precisely known W boson mass [9], is sensitive to a jet energy scale correction factor, JES.

We use these eight histograms to measure simultaneously the true top quark mass and JES. An unbinned likelihood fit is performed to parameterized signal templates taken from simulated $t\bar{t}$ events generated using different values of $\mathrm{M_{top}}$ and JES, and background templates derived from simulations of the relevant background processes. We include in the fit a constraint on JES from the jet energy calibrations done using *in situ* data and instrument calibration [10], and we constrain the background rates in the 2-tag, 1-tag Tight, and 1-tag Loose samples to the estimated background rates. The background rate in the 0-tag sample is determined in the fit using the difference in predicted signal and background mass distributions.

The four reconstructed top quark mass distributions, and the results of the fit are shown in Fig. 2, where we also note the contributions from background sources. We show in Fig. 3 the distributions of the dijet invariant mass for the four subsamples, as well as the result of the fits. In all cases, we see agreement between the observed data distributions and the predictions.

We obtain $\mathrm{M_{top}} = 173.5\,^{+3.7}_{-3.6}$ (stat.) GeV/c^2, where the uncertainty is only statistical but incorporates the uncertainties on the JES measurement. Figure 4 shows the likelihood in the $\mathrm{M_{top}}$-JES plane for the combined measurement. The jet energy scale factor is defined as the difference between the observed and nominal jet energy scale, normalized by the JES calibration uncertainty. If we do not constrain JES to the nominal value of zero, we obtain JES $= -0.25 \pm 1.22\ \sigma$, which indicates our nominal jet energy calibrations are in good agreement with the energy scale information provided by the W boson mass peak in the $t\bar{t}$ decay.

The results from the two methods have various sources of systematic uncertainty. The dominant systematic uncertainty for both methods arises from our understanding of JES. In the DLM method, the jet energy scale is estimated as the shift in $\mathrm{M_{top}}$ arising from a 1 σ change in JES. In the template measurement, σ is used as a constraint on the jet energy scale, JES, in the simultaneous measurement of $\mathrm{M_{top}}$ and JES from the m_t^{reco} and dijet invariant mass histograms, so that the uncertainty due to the jet energy scale is improved with respect to DLM, and is ultimately reported as part of the uncertainty from the likelihood fit. We estimate the contribution of the jet

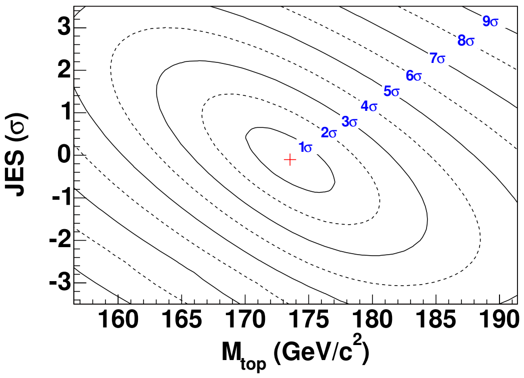

FIG. 4: The contours of the likelihood in the M_{top}-JES plane for the combined fit to all four subsamples. At each point in the plane, the likelihood is maximized with respect to the other free parameters. The crosshair shows the best fit point, and contours are given at intervals of 1 σ, where the $k\ \sigma$ curve is defined by $\Delta \ln L = 0.5k^2$.

TABLE II: The systematic uncertainties for the two analyses are summarized.

Systematic	DLM ΔM_{top} (GeV/c^2)	Template ΔM_{top} (GeV/c^2)
Jet Energy Scale	3.0	$[\sim 2.5]^a$
ISR/FSR	0.6	0.7
PDFs	0.5	0.3
Modeling	0.9	1.0
Method	0.6	0.6
Total	3.3	1.5^a

[a]The JES systematic is included in the uncertainty reported by the likelihood fit.

energy scale uncertainty to the uncertainty from the likelihood fit as 2.5 GeV/c^2.

There are a number of sources of additional systematic uncertainty that affect both analyses: initial state and final state radiation (ISR/FSR) uncertainties affect the extra jet activity in $t\bar{t}$ events and modify the jet kinematics; uncertainties arising from the parton distribution functions (PDFs); the modeling of b-jet fragmentation, decays, and color connections; modeling of the background processes; and uncertainties arising from modeling of the entire event, estimated by comparing the results when we replace the HERWIG Monte Carlo calculations with the PYTHIA generator [11]. Table II summarizes these uncertainties.

The DLM method has additional uncertainties that arise from the use of transfer functions and from the procedure that corrects the measured mass for the presence of background. Together with the common sources noted above, the systematic uncertainty on the DLM mass measurement is 3.3 GeV/c^2.

The template method has additional uncertainties arising from the statistical precision of the templates themselves and approximations made in treating the jet energy scale as a single parameter affecting all jets coherently. The systematic uncertainties for the template method, combined in quadrature, are 1.5 GeV/c^2.

In summary, we have made two new measurements of the top quark mass. An analysis using the DLM technique results in $M_{top} = 173.8\ ^{+2.7}_{-2.5}$ (stat.) $\pm$ 3.3 (syst.) GeV/c^2; the analysis using the template technique results in $M_{top} = 173.5\ ^{+3.7}_{-3.6}$ (stat. + JES) $\pm$ 1.5 (other syst.) GeV/c^2. There is a large statistical correlation between these measurements, so that we choose to quote as a result only the single most accurate measurement, the mass obtained using the template method. This provides the most precise single constraint on this important physical parameter, exceeding the accuracy of the previous world average for the top quark mass [9].

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] J. Erler and P. Langacker, Phys. Lett. **B592**, 1 (2004), hep-ph/0407097.

[2] K. Kondo, J. Phys. Soc. Jap. **57**, 4126 (1988).

[3] V. M. Abazov et al. (D0), Nature **429**, 638 (2004), hep-ex/0406031.

[4] F. Abe et al. (CDF), Phys. Rev. **D50**, 2966 (1994).

[5] D. Acosta et al. (CDF), Phys. Rev. **D71**, 032001 (2005), hep-ex/0412071. We employ a cylindrical coordinate system where θ is the polar angle with respect to the proton beam, ϕ is the azimuthal angle, and pseudorapidity is $\eta = -\ln\tan(\theta/2)$. Transverse energy and momentum are $E_T = E\sin\theta$ and $p_T = p\sin\theta$, respectively, where E and p are energy and momentum.

[6] G. Corcella, I. G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M. H. Seymour, and B. R. Webber, JHEP **01**, 010 (2001), [hep-ph/0011363]; hep-ph/0210213.

[7] ?? (CDF), Phys. Rev. **D??**, ?? (200?), placeholder for run II DLM top mass measurement.

[8] ?? (CDF), Phys. Rev. **D??**, ?? (200?), placeholder for run II TMT+JES top mass measurement.
[9] S. Eidelman et al. (Particle Data Group), Phys. Lett. **B592**, 1 (2004).
[10] ?? (CDF), NIM **??**, ?? (200?), placeholder for run II JES NIM.
[11] T. Sjostrand, L. Lonnblad, and S. Mrenna, TP 01-21, LU (2001), hep-ph/0108264.