Top-Quark Mass Measurement Using Events with Missing Transverse Energy and Jets at CDF

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>As Published</td>
<td><a href="http://dx.doi.org/10.1103/PhysRevLett.107.232002">http://dx.doi.org/10.1103/PhysRevLett.107.232002</a></td>
</tr>
<tr>
<td>Publisher</td>
<td>American Physical Society (APS)</td>
</tr>
<tr>
<td>Version</td>
<td>Final published version</td>
</tr>
<tr>
<td>Accessed</td>
<td>Sun Feb 03 19:18:26 EST 2019</td>
</tr>
<tr>
<td>Citable Link</td>
<td><a href="http://hdl.handle.net/1721.1/69208">http://hdl.handle.net/1721.1/69208</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>Article is made available in accordance with the publisher’s policy and may be subject to US copyright law. Please refer to the publisher’s site for terms of use.</td>
</tr>
<tr>
<td>Detailed Terms</td>
<td></td>
</tr>
</tbody>
</table>
Top-Quark Mass Measurement Using Events with Missing Transverse Energy and Jets at CDF


1Institute of Physics, Academia Sinica, Taipei, Taiwan 11529, Republic of China
2Argonne National Laboratory, Argonne, Illinois 60439, USA
3University of Athens, 157 71 Athens, Greece
4Institut de Fisica d’Altes Energies, ICREA, Universitat Autonoma de Barcelona, E-08193, Bellaterra, Barcelona, Spain
5Baylor University, Waco, Texas 76798, USA
6aIstituto Nazionale di Fisica Nucleare Bologna, I-40127 Bologna, Italy
6bIstituto Nazionale di Fisica Nucleare Bologna, I-40127 Bologna, Italy
7University of California, Davis, Davis, California 95616, USA
8Instituto de Fisica de Cantabria, CSIC-University of Cantabria, 39005 Santander, Spain
9Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA
10Enrico Fermi Institute, University of Chicago, Chicago, Illinois 60637, USA
11Comenius University, 842 48 Bratislava, Slovakia; Institute of Experimental Physics, 040 01 Kosice, Slovakia
12Joint Institute for Nuclear Research, RU-141980 Dubna, Russia
13Duke University, Durham, North Carolina 27708, USA
14Fermi National Accelerator Laboratory, Batavia, Illinois 60510, USA
15University of Florida, Gainesville, Florida 32611, USA
16Laboratori Nazionali di Frascati, Istituto Nazionale di Fisica Nucleare, I-00044 Frascati, Italy
17University of Geneva, CH-1211 Geneva 4, Switzerland
18Glasgow University, Glasgow G12 8QQ, United Kingdom
19Harvard University, Cambridge, Massachusetts 02138, USA
20Division of High Energy Physics, Department of Physics, University of Helsinki and Helsinki Institute of Physics, FIN-00014, Helsinki, Finland
21University of Illinois, Urbana, Illinois 61801, USA
22The Johns Hopkins University, Baltimore, Maryland 21218, USA
23Institut für Experimentelle Kernphysik, Karlsruhe Institute of Technology, D-76131 Karlsruhe, Germany
24Center for High Energy Physics: Kyungpook National University, Daegu 702-701, Korea;
Seoul National University, Seoul 151-742, Korea;
Seoul National University, Seoul 151-742, Korea;
Sungkyunkwan University, Suwon 440-746, Korea;
Korea Institute of Science and Technology Information, Daejeon 305-806, Korea;
Chonnam National University, Gwangju 500-757, Korea;
Chonbuk National University, Jeonju 561-756, Korea
25Ernest Orlando Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA
26University of Liverpool, Liverpool L69 7ZE, United Kingdom
27University College London, London WC1E 6BT, United Kingdom
28Centro de Investigaciones Energeticas Medioambientales y Tecnologicas, E-28040 Madrid, Spain
29Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA
We present a measurement of the top-quark mass using a sample of $t\bar{t}$ events in 5.7 $fb^{-1}$ of integrated luminosity from $p\bar{p}$ collisions at the Fermilab Tevatron with $\sqrt{s} = 1.96$ TeV and collected by the CDF II Detector. We select events having large missing transverse energy, and four, five, or six jets with at least one jet tagged as coming from a $b$ quark, and reject events with identified charged leptons. This analysis considers events from the semileptonic $t\bar{t}$ decay channel, including events that contain tau leptons. The measurement is based on a multidimensional template method. We fit the data to signal templates of varying top-quark masses and background templates, and measure a top-quark mass of $M_{top} = 172.32 \pm 2.4$(stat) $\pm 1.0$(syst) GeV/$c^2$.

DOI: 10.1103/PhysRevLett.107.232002 PACS numbers: 14.65.Ha, 12.15.Ff, 13.85.Qk

The top quark ($t$) is the heaviest known elementary particle, with a mass approximately 40 times heavier than the mass of the bottom quark ($b$), its partner in the weak isospin doublet. The top-quark mass can be used as a consistency check for standard model (SM) parameters. For example, the top quark contributes significantly to electroweak radiative corrections relating the mass of the top quark, $M_{top}$, and that of the $W$ boson to the mass of the predicted Higgs boson within the SM and in new physics models [1,2]. Therefore, combined precision measurements of the $W$ mass and $M_{top}$ provide an important constraint on the Higgs boson mass. At the Tevatron, the top quark is predominantly produced in $t\bar{t}$ pairs. As in the SM the top quark decays almost exclusively to a $W$ boson and a $b$ quark, the expected signature of a $t\bar{t}$ production event is $t\bar{t} \rightarrow W^+bW^-\bar{b}$, assuming unitarity of the three-generation
Cabibbo-Kobayashi-Maskawa matrix [3]. Since the W boson subsequently decays either to a quark-antiquark pair or to a lepton-neutrino pair, the final state of \( t\bar{t} \) production can be classified by the number of charged leptons produced. In this Letter, we focus on events with large missing transverse energy (\( \sum \vec{E}_T \)) as expected for undetected energetic neutrinos, accompanied by jets. We explicitly veto events with identified high \( p_T \) electrons or muons (lepton + jets events) as well as multijet events where both \( W \) bosons decay hadronically (all-hadronic events). This ensures that our result is statistically independent from other CDF top-quark mass measurements [5–8] and allows for a future combination with them. A previous measurement of \( M_{t}\bar{t} \) in this final state used an integrated luminosity of 311 pb \(^{-1} \) [9] and yielded \( M_{t}\bar{t} = 172.3 \pm 15.3\text{(stat)} \pm 14.4\text{(syst)} \text{GeV/c}^2 \). Although no identified leptons are explicitly required, our measurement is sensitive to all \( W \) leptonic decays. This includes decays to \( \tau \) leptons, which constitute approximately 40% of the signal sample in our final selection. The frequency of observed \( \tau \) lepton final states is predicted to be enhanced by new physics models such as a charged Higgs decay [10], therefore, a significantly different measurement of \( M_{t}\bar{t} \) in this decay channel could indicate contributions from non-SM [11] physics processes.

We use data corresponding to an integrated luminosity of 5.7 fb \(^{-1} \) of \( pp \) collisions at the Fermilab Tevatron, and collected by the CDF II Detector [12]. The sample of events used in this measurement is a subset of events that initially passed a trigger requirement, which accepted events with at least four calorimetric clusters [13] of \( E_{T} > 15 \text{ GeV} \) and a scalar sum \( \sum E_{T} \) of these clusters greater than \( 175 \text{ GeV} \) [14]. After the trigger selection, event observables of physical interest are computed. Jets are reconstructed with the JETCLU [15] algorithm using a cone radius of \( \Delta R = \sqrt{\Delta \eta^2 + \Delta \phi^2} = 0.4 \) [16]. Jets are corrected [17] for nonuniformities of the calorimeter response as a function of \( \eta \), energy contributed by multiple \( p\bar{p} \) interactions in the event, and calorimeter nonlinear response. To determine if a jet originated from a \( b \) quark, a secondary vertex algorithm [18] is applied. This algorithm identifies jets that are likely to come from \( b \) quark hadronization through the presence of a displaced vertex within the jets (\( b \) tag). We require at least one jet to be identified as a \( b \) jet (\( b \) tagged). We divide the sample of candidate events into two, separating events with one \( b \)-tagged jet (1 tag) from events with two or more \( b \)-tagged jets (2 tag). Events are required to have at least four and at most six jets with transverse energy \( E_{T} > 15 \text{ GeV} \) and \( |\eta| < 2.0 \). To avoid overlap with other CDF top-quark mass measurements, we reject events with reconstructed electrons or muons with \( p_T > 20 \text{ GeV/c} \) and \( |\eta| < 1.0 \) (lepton + jets final state), and events with \( \sum E_{T} \) significance below 3 \text{ GeV/}^{1/2} \ (all-hadronic final state), where the \( \sum E_{T} \) significance is defined as \( \sum E_{T}/\sqrt{\sum \text{jet} E_{T}} \). For further rejection of multijet backgrounds from QCD processes, we require \( \Delta \phi_{\text{min}}(\sum E_{T}, \text{jet}) > 0.4 \), where \( \Delta \phi_{\text{min}}(\sum E_{T}, \text{jet}) \) is the smallest separation in the angle \( \phi \) between jets and \( \sum E_{T} \) [14].

Background events with \( b \) tags arise from QCD multijet and electroweak production of \( W \) bosons associated with heavy flavor jets. In order to improve the ratio of \( t\bar{t} \) signal in the semileptonic channel to the background of this analysis, an artificial neural network is trained to identify the kinematic and topological characteristics of SM \( t\bar{t} \) events using eight input variables [14]. We apply the neural network to all events passing the above selections, and make a cut on neural network score that retains 81% of the \( t\bar{t} \) signal events while rejecting 91% of background events. The selection criteria listed above define the “signal region.” We follow Ref. [14] and estimate the background rate using a data-driven method. The method uses events with exactly three jets and employs a per-jet parametrization of the \( b \)-tagging probability. Because of the presence of \( t\bar{t} \) events in samples with higher jet multiplicity, we extrapolate the \( b \)-tagging probability of the three-jet event sample to higher jet multiplicity events by iteratively removing the \( t\bar{t} \) content from the sample [14]. We estimate the background for the 1-tag and 2-tag samples separately. A \( b \)-tagging correction factor [19] is applied to take into account the fact that most of the heavy flavor jets are produced in pairs. With this procedure we obtain the estimated number of background events in the signal region shown in Table I. We also show the estimated number of \( t\bar{t} \) signal events, assuming a \( t\bar{t} \) production cross section of 7.5 pb at \( M_{t}\bar{t} = 172.5 \text{ GeV/c}^2 \) [20], together with the number of observed events in the data.

### Table I. Number of expected signal and background events

<table>
<thead>
<tr>
<th>Events</th>
<th>1-tag</th>
<th>2-tag</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected ( t\bar{t} ) signal</td>
<td>644.3 ± 118.7</td>
<td>262.9 ± 50.3</td>
</tr>
<tr>
<td>Expected background</td>
<td>410.6 ± 31.7</td>
<td>43.8 ± 11.0</td>
</tr>
<tr>
<td>Total expectation</td>
<td>1054.9 ± 122.9</td>
<td>306.7 ± 51.5</td>
</tr>
<tr>
<td>Observed data</td>
<td>1147</td>
<td>285</td>
</tr>
</tbody>
</table>

Monte Carlo (MC)-simulated \( t\bar{t} \) samples are generated by PYTHIA [21] of 76 different \( M_{t}\bar{t} \) values ranging from \( 150 \text{ GeV/c}^2 \) to \( 240 \text{ GeV/c}^2 \), with increments from \( 0.5 \text{ GeV/c}^2 \) in the region immediately above and below \( 172.5 \text{ GeV/c}^2 \) near the extreme mass regions. For each of them we reconstruct the events with different values of \( \Delta_{\text{RES}} \), the difference between the jet energy scale (JES) in the MC simulation and the data.
total of 27 different $\Delta_{\text{JES}}$ values, ranging from $-3.0\sigma_c$ to $+3.0\sigma_c$, where $\sigma_c$ is the uncertainty on the JES [22], are used to reduce the systematic effects due to the jet energy uncertainty, as described later.

For each MC-simulated sample in this analysis, we reconstruct three variables using the leading four or five jets to form the templates. The first variable, $m_{jj}$, defined as the invariant mass of the two jets from the hadronically decaying W boson, serves as an in situ constraint of the JES through the likelihood fit described later. We calculate $m_{jj}$ from the two non-$b$-tagged jets whose invariant mass produces the closest value to the world average W boson mass of 80.40 GeV/c$^2$ [23]. We also reconstruct the top-quark mass ($m_t^{\text{had,rec}}$) from the invariant mass of the three jets whose momentum sum yields the largest $p_T$ since the invariant mass constructed this way has a large correlation with the hadronically decaying top-quark mass. To enhance the probability that the jets used to compute $m_t^{\text{had,rec}}$ come from the hadronically decaying top quark, we add two constraints to the calculation of $m_t^{\text{had,rec}}$: first, $m_t^{\text{had,rec}_1}$ must contain the two jets that form $m_{jj}$, and second, in 2-tag events the third jet of $m_t^{\text{had,rec}_1}$ must be $b$ tagged. A third variable, $m_t^{\text{had,rec}_2}$, is defined as the invariant mass of three jets, two of which are the pair which defines $m_{jj}$. The third jet of $m_t^{\text{had,rec}_2}$ is required to be the most energetic jet of those not forming $m_t^{\text{had,rec}_1}$, and is also required to be $b$ tagged in a 2-tag event. The variable $m_t^{\text{had,rec}_2}$ plays a complementary role to $m_t^{\text{had,rec}_1}$ in extracting information on the top-quark mass, and is particularly important in events where the three jets used to compute $m_t^{\text{had,rec}_1}$ were not the actual decay products of the hadronically decaying top quark.

The template method used in the extraction of $M_{\text{top}}$ requires that a probability density function (pdf) be built for each template. For each MC signal and background sample, we estimate the pdfs using the kernel density estimation (KDE) [24,25] that employs a nonparametric method to construct pdfs. For each sample, we build a three dimensional pdf from the reconstructed observables $(m_{jj}, m_t^{\text{had,rec}_1},$ and $m_t^{\text{had,rec}_2}$), taking their correlations into account. To measure $M_{\text{top}}$, we fit the signal and background pdfs to the distributions of the observables in the data using an unbinned maximum likelihood fit [26] where we minimize the negative logarithm of the likelihood using MINUIT [27]. The likelihood fits $M_{\text{top}}$ and $\Delta_{\text{JES}}$ simultaneously, and is built separately for each subsample, 1-tag and 2-tag events, in order to improve the usage of statistical information. References [25,28] provide detailed information about this technique.

The mass fitting procedure is tested with pseudoexperiments for a set of MC $t\bar{t}$ samples with 14 different $M_{\text{top}}$ values ranging from 159 GeV/c$^2$ to 185 GeV/c$^2$. For each pseudoexperiment we select the number of background events from a Poisson distribution with a mean equal to the expected total number of background events in the sample and the number of signal events from a Poisson distribution with a mean equal to the expected number of signal events normalized to a $t\bar{t}$ production cross section of 7.5 pb [20]. The distributions of the average mass residual (the deviation from the input top-quark mass) and the width of the pull (the ratio of the residual to the uncertainty) for simulated experiments are corrected to be zero and unity, respectively. The correction is $M_t^{\text{corr}} = 1.24 \times M_t^{\text{meas}} - 40.6$ GeV/c$^2$, where $M_t^{\text{meas}}$ is the raw value from the likelihood fit and $M_t^{\text{corr}}$ is the corrected value of the measurement. The measured uncertainty is correspondingly increased by 20% to correct the width of the pull distribution.

We examine various sources of systematic uncertainties that could affect the measurement by comparing the results of pseudoexperiments in which we vary relevant parameters within their uncertainties. One of the dominant sources of systematic uncertainty is the residual JES [6,22]. We vary the JES components within their uncertainties in the generated signal MC events and interpret the shifts in the returned top-quark mass as uncertainties. The $b$ jet energy scale systematic uncertainty that arises from the modeling of $b$ fragmentation, $b$ hadron branching fractions, and calorimeter response captures the additional uncertainty not taken into account in the light quark jet energy scale [6]. The uncertainty arising from the choice of MC generator is estimated by comparing results from MC-simulated samples generated with PYTHIA and HERWIG [29]. We estimate the systematic uncertainty due to imperfect modeling of initial-state gluon radiation (ISR) and final-state gluon radiation (FSR) by varying the amounts of ISR and FSR in simulated events [30]. We estimate the systematic uncertainty due to parton distribution functions (PDF’s) of the proton by varying the independent eigenvectors of the CTEQ6M [31] PDF’s, varying $\Lambda_{\text{QCD}}$ (228 MeV vs 300 MeV), and comparing CTEQ5M [32] with MRST72 [33] PDF’s. To estimate the systematic uncertainty associated with uncertainties in the top-quark production mechanism, we vary the fraction of the top quarks produced by gluon-gluon annihilation from the default 6% to 20%, corresponding to a 1 standard deviation upper bound on the gluon fusion fraction [34]. We also evaluate the uncertainty due to background modeling effects by reweighting the background shape up and down and comparing the resulting measurements. We apply an additional uncertainty to account for the effect of the trigger simulation in the signal MC samples, in a similar way to the background shape systematic uncertainty estimation. We also estimate an uncertainty due to the effect of multiple hadron interactions, which takes into account the increasing instantaneous luminosity in this data set. The color reconnection (CR) systematic uncertainty [35] is evaluated using MC samples generated with
and without CR effects adopting different tunes [36] of PYTHIA. Table II summarizes the individual systematic uncertainties considered, giving a total systematic uncertainty of 1.0 GeV/c² for the measurement of M_{top}.

By applying a likelihood fit to the data using the three observables described above and the corrections obtained from the simulated experiments, the top-quark mass is measured to be

\[
M_{\text{top}} = 172.3 \pm 2.4(\text{stat}) \pm 1.0(\text{syst}) \text{ GeV}/c^2 = 172.3 \pm 2.6 \text{ GeV}/c^2.
\]  

(1)

Figure 1 shows the distribution of the observables used for the M_{top} measurement overlaid with their probability density functions from \( \bar{t}t \) signal events with \( M_{\text{top}} = 172.5 \text{ GeV}/c^2 \) and the estimated background.

In conclusion, we have performed a measurement of the top-quark mass in events with large \( E_T \) and jets, corresponding to an integrated luminosity of 5.7 fb⁻¹. The data sample has been chosen in such a way as to exclude events used in other CDF top-quark mass measurements. The result, \( M_{\text{top}} = 172.3 \pm 2.6 \text{ GeV}/c^2 \), is approximately a factor of 6 improvement from the previous measurement in this channel [9], and is in agreement with other measurements which contribute to the world average of \( M_{\text{top}} = 173.2 \pm 0.9 \text{ GeV}/c^2 \) [37].

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 für Bildung und Forschung, Germany; the Korean World Class University Program, the National Research Foundation of Korea; the Science and Technology Facilities Council and the Royal Society, U.K.; the Institut National de Physique Nucléaire et Physique des Particules/CNRS; the Russian Foundation for Basic Research; the Ministerio de Ciencia e Innovación, and Programa Consolider-Ingenio 2010, Spain; the Slovak R&D Agency; the Academy of Finland; and the Australian Research Council (ARC).

![Figure 1](image-url)
[4] The transverse momentum $p_T$ and transverse energy $E_T$ of a particle are defined as $|\vec{p} \sin \theta|$ and $E \sin \theta$, respectively, where $\theta$ is the polar angle of the particle momentum with respect to the proton beam direction. The missing transverse energy, an imbalance of energy in the plane transverse to the beam direction, is defined as $E_T = \sum_{tower} E_T \hat{n}_T$, where $\hat{n}_T$ is the unit vector normal to the beam and pointing to a given calorimeter tower and $E_T$ is the transverse energy measured in that tower.


[16] CDF uses a cylindrical coordinate system with the $z$ axis along the proton beam axis. Pseudorapidity is defined as $\eta = -\ln[\tan(\theta/2)]$, where $\theta$ is the polar angle relative to the proton beam direction, and $\phi$ is the azimuthal angle while $p_T = |\vec{p} \sin \theta|$, $E_T = E \sin \theta$.


