Measurement of the Top Quark Mass in the All-Hadronic Mode at CDF

A measurement of the top quark mass ($M_{\text{top}}$) in the all-hadronic decay channel is presented. It uses 5.8 fb$^{-1}$ of pp data collected with the CDF II detector at the Fermilab Tevatron Collider. Events with six to eight jets are selected by a neural network algorithm and by the requirement that at least one of the jets is tagged as a b quark jet. The measurement is performed with a likelihood fit technique, which simultaneously determines $M_{\text{top}}$ and the jet energy scale (JES) calibration. The fit yields a value of $M_{\text{top}} = 172.5 \pm 1.4 \text{(stat)} \pm 1.0 \text{(JES)} \pm 1.1 \text{(syst)} \text{GeV}/c^2$.

PACS numbers: 14.65.Ha, 13.85.Ni,13.85.Qk

The mass of the top quark ($M_{\text{top}}$) is a fundamental parameter of the standard model (SM), and its large value makes the top quark contribution dominant in loop corrections to many observables, like the W boson mass $M_W$. Precise measurements of $M_W$ and $M_{\text{top}}$ allow one to set indirect constraints on the mass of the, as yet unobserved, Higgs boson [1].

In this Letter we present a measurement of $M_{\text{top}}$ using proton-antiproton collision events at a center-of-mass energy of 1.96 TeV. Top quarks are produced at the largest rate in pairs ($t\bar{t}$), with each top quark decaying immediately into a W boson and a b quark nearly 100% of the time [2]. In this analysis events...
where both the $W$’s decay to a quark-antiquark pair are considered. This all-hadronic final state has the largest branching ratio among the possible decay channels (46%), but it is overwhelmed by the QCD multijet background processes, which surpass $tt$ production by three orders of magnitude even after a dedicated trigger requirement. Nevertheless, it will be shown how this difficult background can be successfully controlled and significantly suppressed with a properly optimized event selection. Comparing to the previous result from CDF [3], that already represented the most precise measurement in this channel, the improvements in the analysis technique and a larger dataset allow us to decrease the total uncertainty on $M_{\text{top}}$ by 21%. The additional dataset has been acquired at higher instantaneous luminosity, which results in a higher number of background events in the data sample. Despite this fact, the introduction of significant improvements to the analysis resulted in the world best measurement of $M_{\text{top}}$ in the all-hadronic channel so far, also entering with the third largest weight in the $M_{\text{top}}$ world average calculation [4].

The data correspond to an integrated luminosity of 5.8 fb$^{-1}$. They have been collected between March 2002 and February 2010 by the CDF detector, a general-purpose apparatus designed to study $p\bar{p}$ collisions at the Tevatron and described in detail in [5]. Events used in this measurement are selected by a multijet trigger [3], and retained only if they are well contained in the detector acceptance, have no well identified energetic electron or muon, and have no significant ($< 3 \text{GeV/}{c}$) missing transverse energy $E_T$ [6]. Candidate events are also required to have from six to eight “tight” ($E_T \geq 15 \text{GeV}$ and $|\eta| \leq 2.0$) jets. After this preselection, a total of about 5.6M events is observed in the data, with less than 9 thousand expected from $tt$ events. To improve the signal-to-background ratio ($S/B$) a multivariate algorithm is implemented. An artificial neural network, based on a set of kinematic and jet shape variables [3], is used to take advantage of the distinctive features of signal and background events. The neural network was trained using simulated $tt$ events generated by PYTHIA [7] and propagated through the CDF detector simulation. At this level of the selection the fraction of signal events is still negligible and the data are used to represent the background. The value of the output node $N_{\text{out}}$ is used as a discriminant between signal and background. In order to further increase the signal purity, a $b$-tagging algorithm [8] is used to identify (“$b$-tag” or simply “tag”) jets that most likely resulted from the fragmentation of a $b$ quark. Only events with one to three tagged jets are then retained.

The background for the $tt$ multijet final state comes mainly from QCD production of heavy-quark pairs ($bb$ and $cc$) and events with false $b$-tags of light-quark and gluon jets. Given the large theoretical uncertainties on the QCD multijet production cross section, the background prediction is obtained from the data themselves. A tag rate per jet is used, defined as the probability of tagging a jet whose tracks are reconstructed in the tracking system (“fiducial” jet). This rate is evaluated from events passing the preselection with five tight jets ($S/B \approx 1/2000$), and it is parametrized in terms of the jet $E_T$, the number of tracks associated to the jet, and the number of reconstructed primary vertices. The rate of a fiducial jet in a candidate event selected before the $b$-tagging represents an estimate of the probability for that jet to come from background and to be tagged. This allows us to predict the number of background events with a given number of tagged jets as well as their distributions [3]. The background modeling is tested in background-dominated control regions with six to eight jets and small values of $N_{\text{out}}$. Small residual discrepancies are accounted for as systematic uncertainties.

This analysis employs the template method to measure $M_{\text{top}}$, with simultaneous calibration of the jet energy scale (JES). The latter is a multiplicative factor representing a correction applied to the raw energy of a reconstructed jet ($E_{T}^{\text{raw}}$), so that its corrected energy $E_T = \text{JES} \cdot E_{T}^{\text{raw}}$, is a better estimate of the energy of the underlying parton [9]. Discrepancies between data and simulation lead to an uncertainty on the JES used in Monte Carlo (MC) events, and, as a consequence, on the measurements of $M_{\text{top}}$. The MC distributions of the reconstructed top quark mass, $m_{t}^{\text{rec}}$, and $W$ boson mass, $m_{W}^{\text{rec}}$, are used as a reference (“template”) in the measurement, with the latter providing the information necessary to calibrate “in situ” the JES, by using the precisely measured value of the $W$ boson mass [2].

For each selected event, the six highest-$E_T$ jets are assumed to come from the quarks of a $tt$ all-hadronic final state. Each of the different combinations where the jets are arranged in two doublets (the $W$ bosons) and two triplets (the top quarks) is considered. To reduce the number of permutations, $b$-tagged jets are assumed to come from $b$ quarks only, resulting in 30, 6 or 18 permutations for events with one, two or three tagged jets, respectively [10]. For each permutation $m_{t}^{\text{rec}}$ is obtained through a constrained fit based on the minimization of the following $\chi^2$-like function:

$$
\chi^2_t = \frac{(m_{jj}^{(1)} - M_W)^2}{\Gamma_t^2} + \frac{(m_{jj}^{(2)} - M_W)^2}{\Gamma_t^2} + \frac{(m_{jjb}^{(1)} - m_{t}^{\text{rec}})^2}{\Gamma_t^2} + \frac{(m_{jjb}^{(2)} - m_{t}^{\text{rec}})^2}{\Gamma_t^2} + \frac{6}{\sigma_t^2} \left( \frac{p_{T,j}^{\text{fit}} - p_{T,j}^{\text{meas}}}{\sigma_{pT}} \right)^2
$$

where $m_{jj}^{(1,2)}$ are the invariant masses of the two pairs of jets assigned to light flavor quarks, $m_{jjb}^{(1,2)}$ are the invariant masses of the triplets including one pair and
a \(b\)-tagged jet, \(M_W = 80.4\ \text{GeV}/c^2\) and \(\Gamma_W = 2.1\ \text{GeV}/c^2\) are the measured mass and natural width of the \(W\) boson [2], and \(\Gamma_t = 1.5\ \text{GeV}/c^2\) is the assumed natural width of the top quark [11]. The jet transverse momenta are constrained in the fit to the measured values, \(p_T^j\), within their known resolutions, \(\sigma_T\). The fit is performed with respect to \(m_{t}^{\text{rec}}\) and the transverse momenta of the jets \(p_T^j\), and the permutation which gives the lowest value for the minimized \(\chi^2\) is selected. The variable \(m_{t}^{\text{rec}}\) is reconstructed by the same procedure considered for \(m_{t}^{\text{rec}}\), but with a \(\chi^2\) function, \(\chi_{W}^2\), where also the \(W\) mass is left free to vary in the fit. The selected values of \(m_{t}^{\text{rec}}\) and \(m_{W}^{\text{rec}}\) enter the respective distributions, built separately for events with exactly one or \(\geq 2\) tags.

Signal templates are built using MC events with \(M_{\text{top}}\) values from 160 to 185 GeV/c\(^2\), with steps of 2.5 GeV/c\(^2\). For each \(M_{\text{top}}\) the corrected jets’ \(E_T\) are changed to values \(E_T'\) given by \(E_T' = [1 + \Delta\text{JES} \cdot (\sigma_{\text{JES}}/\text{JES})] \cdot E_T\), where \(\sigma_{\text{JES}}\) is the absolute uncertainty on the JES and \(\Delta\text{JES}\) is a dimensionless number. This equivalently means that the applied JES differs by \(\Delta\text{JES} \times \sigma_{\text{JES}}\) from the default value. It should be noted that \(\sigma_{\text{JES}}/\text{JES}\) is a function of the jet \(E_T\) [9]. Values of \(\Delta\text{JES}\) between \(-2\) and \(+2\), in steps of 0.5, have been considered, and in the following we refer to this parameter to denote variations of the JES.

To construct the background templates we apply the fitting technique to the data events passing the neural network selection cut, omitting the \(b\)-tagging requirement ("pretag" sample). All possible combinations are considered where one to three fiducial jets are treated as tagged. The weight of each combination is given by the probability, evaluated by the tag rates, that those jets are tagged in the event by the \(b\)-tagging algorithm, and it is used for the corresponding values of \(m_{t}^{\text{rec}}\) and \(m_{W}^{\text{rec}}\) to build the templates. As the procedure is applied to data, signal contributions must be properly subtracted.

Sets of simulated experiments ("pseudoeperiments", PEs) have been performed to optimize the requirements on the values of \(N_{\text{out}}, \chi^2_{W}\) and \(\chi^2_{t}\) in order to minimize the statistical uncertainty on the \(M_{\text{top}}\) measurement. As an improvement with respect to [3], two different sets of events, denoted by \(S_{\text{JES}}\) and \(S_{M_{\text{top}}}\), are considered to build the \(m_{W}^{\text{rec}}\) and \(m_{t}^{\text{rec}}\) templates, respectively. This choice contributes in reducing the final total uncertainty on \(M_{\text{top}}\) with respect to [3] by about 12%. The set \(S_{\text{JES}}\) is selected by using cuts on \(N_{\text{out}}\) and \(\chi^2_{W}\), while \(S_{M_{\text{top}}}\) is selected by a further requirement on \(\chi^2_{t}\), so that \(S_{M_{\text{top}}} \subseteq S_{\text{JES}}\). The procedure gives \(\{N_{\text{out}} \geq 0.97, \chi^2_{W} \leq 2, \chi^2_{t} \leq 3\}\) and \(\{N_{\text{out}} \geq 0.94, \chi^2_{W} \leq 3, \chi^2_{t} \leq 4\}\) as the optimized selection requirements for 1-tag and \(\geq 2\)-tag events respectively. Correlations between \(m_{t}^{\text{rec}}\) and \(m_{W}^{\text{rec}}\) in events selected both in \(S_{M_{\text{top}}}\) and in \(S_{\text{JES}}\) are taken into account during the calibration procedure described below.

In order to measure \(M_{\text{top}}\) simultaneously with JES, a fit is performed in which an unbinned extended likelihood function is maximized to find the values of \(M_{\text{top}}, \Delta\text{JES}\), and the number of signal \((n_s)\) and background \((n_b)\) events for each tagging category which best reproduce the observed distributions of \(m_{t}^{\text{rec}}\) and \(m_{W}^{\text{rec}}\). The likelihood depends on the probability density functions (p.d.f.’s) expected for signal (s) and background (b): \(P_s(m_{t}^{\text{rec}} | M_{\text{top}}, \Delta\text{JES}),\ P_s(m_{W}^{\text{rec}} | M_{\text{top}}, \Delta\text{JES}),\ P_b(m_{t}^{\text{rec}})\), and \(P_b(m_{W}^{\text{rec}})\). The p.d.f.’s are obtained by fitting normalized functions to the templates, initially built as histograms. For the signal the continuous dependence of the p.d.f.’s on \(M_{\text{top}}\) and \(\Delta\text{JES}\) is obtained by fitting simultaneously the whole set of templates, corresponding to the large set of values simulated for those two variables. In the fit a linear dependence of the parameters of the p.d.f.’s on \(M_{\text{top}}\) and \(\Delta\text{JES}\) is assumed, so that the resulting fitted functions have continuous variable shapes [3]. Figure 1 shows examples of signal and background templates for the \(\geq 2\)-tag sample, with the corresponding p.d.f.’s superimposed.

![Templates of \(m_{t}^{\text{rec}}\) for events with \(\geq 2\) tags and corresponding probability density functions superimposed.](image)

**FIG. 1:** Templates of \(m_{t}^{\text{rec}}\) for events with \(\geq 2\) tags and corresponding probability density functions superimposed. Top plot: the signal p.d.f. \(P_s\) for various values of \(M_{\text{top}}\) and \(\Delta\text{JES} = 0\). Bottom plot: the background p.d.f., \(P_b\).

The likelihood function used for the measurement can be divided into three parts:

\[
\mathcal{L} = \mathcal{L}_{1\text{tag}} \times \mathcal{L}_{\geq 2\text{tags}} \times \mathcal{L}_{\Delta\text{JESconst}}
\]

where \(\mathcal{L}_{\Delta\text{JESconst}}\) is a gaussian term constraining the JES to the nominal value (i.e. \(\Delta\text{JES}\) to 0) within its
uncertainty. Terms $\mathcal{L}_{1\text{tag}}$ and $\mathcal{L}_{2\text{tags}}$ are defined as:

$$\mathcal{L}_{1,2\text{tags}} = \mathcal{L}_{\text{JES}} \times \mathcal{L}_{M_{\text{top}}} \times \mathcal{L}_{\text{evts}} \times \mathcal{L}_{N_{\text{bkg}}},$$

where, omitting the dependences on $M_{\text{top}}$ and $\Delta_{\text{JES}},$

$$\mathcal{L}_{\text{JES}} = \prod_{i=1}^{N_{\text{evts}}} \frac{n_s \mathcal{P}^{\text{rec}}_{s,i}(m_{W,i}) + n_b \mathcal{P}^{\text{rec}}_{b,i}(m_{W,i})}{n_s + n_b}$$

and

$$\mathcal{L}_{M_{\text{top}}} = \prod_{i=1}^{N_{\text{evts}}} \frac{A_s n_s \mathcal{P}^{\text{rec}}_{s,i}(m_{t,i}) + A_b n_b \mathcal{P}^{\text{rec}}_{b,i}(m_{t,i})}{A_s n_s + A_b n_b}.$$

Here the probability of observing the $N_{\text{evts}}$ values of $m_{t}^{\text{rec}}$ and the $N_{\text{evts}}$ values of $m_{W}^{\text{rec}}$ reconstructed in the data $S_{\text{JES}}$ and $S_{M_{\text{top}}}$ samples, respectively, is evaluated from the expected distributions as a function of the fit parameters $M_{\text{top}}, \Delta_{\text{JES}}, n_s,$ and $n_b.$ The factors $A_s$ and $A_b$ represent the acceptance of $S_{M_{\text{top}}}$ with respect to $S_{\text{JES}}$ for signal and background, respectively (i.e., the fraction of events selected by the requirements on $\chi^2$ only). For the signal this acceptance is parametrized as a function of the fit parameters $M_{\text{top}}$ and $\Delta_{\text{JES}}.$ The other two factors included in $\mathcal{L}_{1,2\text{tags}}$ are: $\mathcal{L}_{\text{evts}},$ which gives the probability of observing the number of events selected in the data, evaluated by Poisson and binomial distributions, and $\mathcal{L}_{N_{\text{bkg}}}$ which constrains the parameter $n_b$ to the a priori estimate of the expected background, obtained by the tag rate.

The possible presence of biases in the values returned by the likelihood fit has been investigated. Pseudo-experiments are performed assuming specific values for $M_{\text{top}}$ and $\Delta_{\text{JES}}$ and “pseudo-data” are therefore extracted from the corresponding signal and background templates. The results of these PEs have been compared to the input values, and calibration functions to be applied to the output from the fit have been defined in order to obtain, on average, a more reliable estimate of the true values and uncertainties.

Finally, the likelihood fit is applied to data. After the event selection described above, we are left with 4930 and 1196 events with one and $\geq 2$ tags (187 have 3 tags), respectively, in the $S_{\text{JES}}$ sample. The corresponding expected backgrounds amount to 3652 ± 181 and 718 ± 14 events, respectively. The tighter requirements used for the $S_{\text{M}_{\text{top}}}$ samples select 2256 with one tag and 600 with $\geq 2$ tags (76 have 3 tags), with average background estimates of 1712 ± 77 and 305 ± 22, respectively.

For these events the variables $m_{W}^{\text{rec}}$ and $m_{t}^{\text{rec}}$ have been reconstructed and used as the data inputs to the likelihood fit. Once the calibration procedure has been applied, the measurements of $M_{\text{top}}$ and $\Delta_{\text{JES}}$ are:

$$M_{\text{top}} = 172.5 \pm 1.4 \text{ (stat)} \pm 1.0 \text{ (JES)} \text{ GeV}/c^2$$

$$\Delta_{\text{JES}} = -0.1 \pm 0.3 \text{ (stat)} \pm 0.3 \text{ (M}_{\text{top}}).$$

FIG. 2: Negative log-likelihood contours for the likelihood fit performed for the $M_{\text{top}}$ and $\Delta_{\text{JES}}$ measurement. The minimum is shown along with the contours whose projections correspond to one, two, and three $\sigma$ uncertainties on the $M_{\text{top}}$ and $\Delta_{\text{JES}}$ measurements.

Various sources of systematic uncertainties affect the $M_{\text{top}}$ and JES measurements, as described in [3]. They are evaluated by performing PEs using templates built by signal samples where effects due to systematic uncertainties have been included. The differences in the average values of $M_{\text{top}}$ and JES with respect to the PEs performed with default templates are then considered. Possible residual biases existing after the calibration, and uncertainties on the parameters of the calibration functions are also taken into account. The largest contributions come from uncertainties on the modeling of the background, on the simulation of $tt$ events, and on the individual corrections which the JES depends on [9]. Table I shows a summary of all the systematic uncertainties.

In summary, we have presented a measurement of the top quark mass in the all-hadronic channel, using $pp$ collision data corresponding to an integrated luminosity of 5.8 fb$^{-1}$. The measured value is $M_{\text{top}} = 172.5 \pm 1.4 \text{ (stat)} \pm 1.0 \text{ (JES)} \pm 1.1 \text{ (syst)} \text{ GeV}/c^2,$ for a total uncertainty of 2.0 GeV/$c^2.$ This result complements and is consistent with the most recent measurements obtained in other channels by the CDF and D0 Collaborations, and also represents the only all-hadronic measurement for the Run II of the Tevatron [4].

We thank the Fermilab staff and the technical staffs of the participating institutions for their vital contri-
FIG. 3: Distributions of $m_{t}^{\text{rec}}$ (top plot) and $m_{W}^{\text{rec}}$ (bottom plot) as obtained in the selected data (black points) with $\geq 1$ tags, compared to the distributions from signal and background corresponding to the measured values of $M_{\text{top}}$ and $\Delta \text{JES}$. The expected distributions are normalized to the best fit yields.

TABLE I: Sources of systematic uncertainty affecting the $M_{\text{top}}$ and $\Delta \text{JES}$ measurements. The total uncertainty is obtained by the quadrature sum of each contribution.

<table>
<thead>
<tr>
<th>Source</th>
<th>$\delta M_{\text{top}}$ (GeV/$c^2$)</th>
<th>$\delta \Delta \text{JES}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residual bias</td>
<td>0.2</td>
<td>0.03</td>
</tr>
<tr>
<td>Calibration</td>
<td>0.1</td>
<td>0.01</td>
</tr>
<tr>
<td>Generator</td>
<td>0.5</td>
<td>0.21</td>
</tr>
<tr>
<td>Initial/final state radiation</td>
<td>0.1</td>
<td>0.04</td>
</tr>
<tr>
<td>$b$-jet energy scale</td>
<td>0.2</td>
<td>0.05</td>
</tr>
<tr>
<td>$b$-tag</td>
<td>0.1</td>
<td>0.01</td>
</tr>
<tr>
<td>Residual JES</td>
<td>0.4</td>
<td>--</td>
</tr>
<tr>
<td>Parton distribution functions</td>
<td>0.2</td>
<td>0.04</td>
</tr>
<tr>
<td>Multiple $p\bar{p}$ interactions</td>
<td>0.1</td>
<td>0.04</td>
</tr>
<tr>
<td>Color reconnection</td>
<td>0.3</td>
<td>0.12</td>
</tr>
<tr>
<td>Statistics of templates</td>
<td>0.3</td>
<td>0.05</td>
</tr>
<tr>
<td>Background</td>
<td>0.6</td>
<td>0.11</td>
</tr>
<tr>
<td>Trigger</td>
<td>0.2</td>
<td>0.04</td>
</tr>
<tr>
<td>Total</td>
<td>1.1</td>
<td>0.29</td>
</tr>
</tbody>
</table>


[6] We use a cylindrical coordinate system where $\theta$ is the polar angle with respect to the proton beam direction ($z$ axis), $\phi$ is the azimuthal angle about the beam axis, and the pseudorapidity is defined as $\eta = -\ln \tan(\theta/2)$. A particle’s transverse momentum $p_T$ and tranverse energy $E_T$ are given by $|p| \sin \theta$ and $E \sin \theta$ respectively. The missing $E_T$ vector, $\vec{E}_T$, is defined by $\vec{E}_T = -\sum \vec{E}_T,i \hat{n}_{T,i}$, where $\hat{n}_{T,i}$ is the unit vector in the $x-y$ plane pointing from the primary interaction vertex to a given calorimeter tower $i$, and $E_{T,i}$ is the $E_T$ measured in that tower. Finally, $\vec{E}_T = |\vec{E}_T|$, and its significance is defined as $\frac{|\vec{E}_T|}{\sqrt{\sum E_T}}$, where the sum runs over all the selected jets.


[10] If three $b$-tagged jets are present in the event, the three possible assignments of two out of three of them to $b$ quarks are also considered, while the remaining one is treated as a light flavor jet.