



A Measurement of Top Quark Mass Using MET + Jets Events With 5.7 fb^{-1}

The CDF Collaboration
URL <http://www-cdf.fnal.gov>
(Dated: February 28, 2011)

We present a measurement of top-quark mass using 5.7 fb^{-1} of Tevatron's $p\bar{p}$ collisions collected by the CDF detector at Fermilab. We use a neural network to select the events, which have a signature of $\cancel{E}_T + \text{jets}$ in the final state after all cuts. Data events are triggered by high jet-multiplicities, and the background is modeled from data set. We use a three-dimensional template method to build the probability density functions of both signal and background, where the three observables are the M_3 top mass of the underconstrained system, a second reconstructed M_3 top mass called M_3' , and the invariant mass m_{jj} of two jets from the hadronic W decays, which provides an *in situ* improvement in the determination of jet energy scale. We perform a minimum Log-likelihood fit to the data and measure the top-quark mass to be $172.3 \pm 2.6 \text{ GeV}/c^2$.

Preliminary Results of TMT using 5.7 fb^{-1}

I. INTRODUCTION

This note describes a measurement of the mass of the top quark using $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV with the CDF detector at the Tevatron. The mass of the top quark is of much interest to particle physicists, both because the top quark is the heaviest known fundamental particle, and also because a precise measurement of the top quark mass helps constrain the mass of the Higgs boson. Top quarks are produced predominantly in pairs at the Tevatron, and in the Standard Model decay nearly 100% of the time to a W boson and a b quark. The topology of a $t\bar{t}$ event is determined by the decay of the two W bosons, as each W boson can decay to a lepton-neutrino pair ($l\nu$) or to a pair of quarks (qq'). We look for events triggered by a large missing E_T and multiple jets, typically those with one W boson decaying hadronically and the other decaying leptonically. The CDF detector is described in [1].

Our measurement is a template-based measurement, meaning that we compare quantities in data with distributions from simulated MC events to find the most likely top quark mass. Due to the under-constrained feature of the event signature, fully reconstructing the events is impossible. However, we are able to reconstruct a variable called M3 which has high correlation with the top quark mass. Two such kinds of variables are created to increase the precision of the measurement. We also reconstruct a (m_{jj}) from the decay products of W resonance since it is sensitive to possible miscalibration of (jet energy scale) JES in the CDF detector.

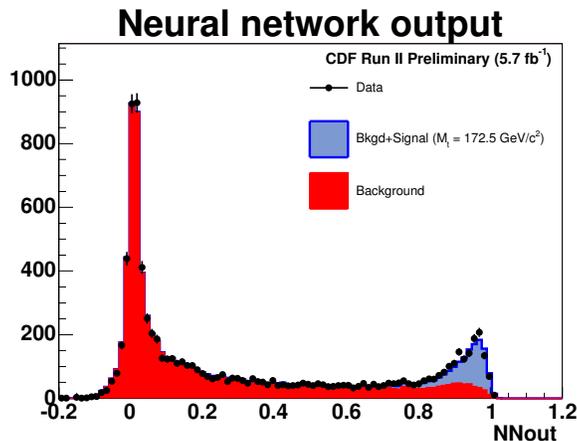
Monte Carlo samples generated with 76 different M_{top} are run through a full CDF detector simulation assuming 29 possible shifts in Δ_{JES} . The values of the three sets of observables in data are compared to each point in the MC grid using a non-parametric approach based on Kernel Density Estimation (KDE). Local Polynomial Smoothing is used to smooth out these points and calculate the probability densities at any arbitrary value of M_{top} and Δ_{JES} . An unbinned likelihood fit is used to measure M_{top} and profile out Δ_{JES} .

II. EVENT SELECTION

At the trigger level, the candidate events are selected by requiring high jet-multiplicities. Offline, a series of clean up cuts are applied first. The events are required to have a missing E_T significance $> 3 \text{ GeV}^{1/2}$ and events with tight or loose leptons are rejected. We also require at least four jets in the final state.

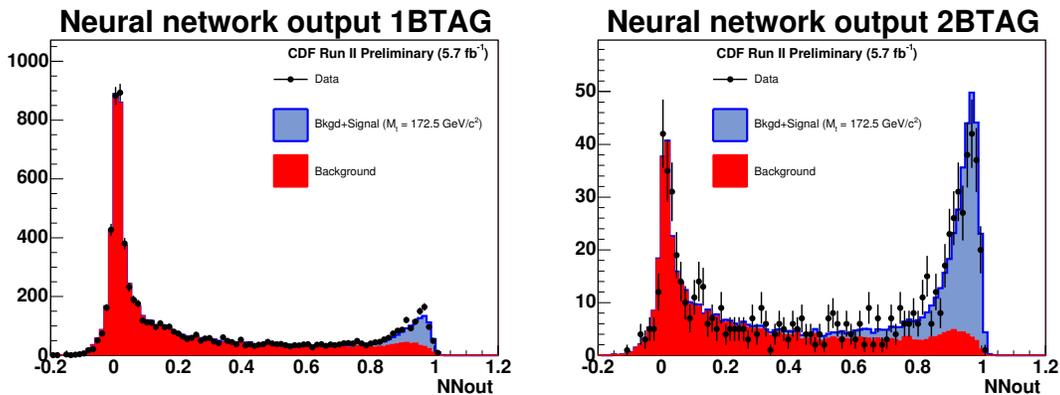
A neural network training with two hidden layers is performed after the clean up selection. Figure II shows the neural network output distribution, and a cut on the output will be applied to enhance the signal-to-background ratio.

To improve the statistical power of the method, we divide each event sample into two subsamples, depending on the number of jets identified as arising from the hadronization and decay of b quarks. The SECVTX [5] algorithm uses the transverse decay length of tracks inside jets to tag jets as coming from b quarks.



III. BACKGROUND ESTIMATION

We use a data driven background estimation in this analysis. Because the probability for a jet to be identified as a b-quark jet is different for the jets in the $t\bar{t}$ events and in the background processes, we can use this to distinguish



the two of them in the data set. We build a b-tag rate matrix from a $t\bar{t}$ -signal-negligible data sample, which consist of events with exactly three jets, and parameterize the b-tag probability as a function of jet characteristics such as jet transverse energy, jet number of tracks, and the MET projection along the jet direction.

At higher jet multiplicities, the signal contamination is not negligible, we use an iterative correction method to remove the sizable $t\bar{t}$ events in each jet bin in order to improve our background prediction. We validate our background estimation by comparing the expected and observed number of b-tagged events at different jet-multiplicities at the background dominant region $NN_{\text{output}} < 0.4$, as shown in Table I.

TABLE I: Expected and observed number of b-tagged events at different jet-multiplicities at the background dominant region $NN_{\text{output}} < 0.4$.

Jet Multiplicity	3 jets	4 jets	5 jets	6 jets	≥ 7 jets
Observed	7752	18998	11448	4498	2224
Expected	7705.6	19060.9	11332.1	4546.1	2263.3
Difference(%)	0.6	0.3	1.0	1.1	1.8

We calculate the number of background event of 1btag and 2btag respectively. The calculation involves a btagging correction factor that takes into account the fact that in QCD events heavy flavor quark tend to come in pairs, which enhances the 2tag probability of an event. After the correction factor we perform the neural network training for both 1btag and 2btag events, which are shown in Figure III, based on which we apply a cut on the neural network output $NN_{\text{out}} > 0.8$.

We calculate the estimated number of background events with the iterative correction and b-tagging correction. Table II shows the total number of both expected and observed events within $4 \leq n_{\text{jets}} \leq 6$, which are the candidate events for this top quark mass measurement.

IV. JET ENERGY SCALE

We describe in this section the *a priori* determination of the jet energy scale uncertainty by CDF that is used later in this analysis. More information on JES, calibration and uncertainty can be found in [6]. There are many sources of uncertainties related to jet energy scale at CDF:

- Relative response of the calorimeters as a function of pseudorapidity.
- Single particle response linearity in the calorimeters.
- Fragmentation of jets.

TABLE II: Expected and Observed number of events after all cuts within $4 \leq n_{\text{jets}} \leq 6$.

b-tagging	Signal	Background	Total Expected	Observed
1tag	644.29 ± 118.73	410.56 ± 31.65	1054.85 ± 122.88	1147
2tag	262.94 ± 50.32	43.77 ± 11.00	306.71 ± 51.51	285

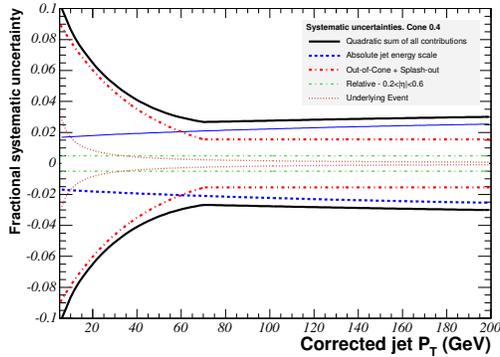


FIG. 1: Jet energy scale uncertainty as a function of the corrected jet p_T for the underlying event (dotted red), relative response (dashed green), out-of-cone energy (dashed red) and absolute response (dashed blue). The contribution of all sources are added in quadrature (full black) to form the total Δ_{JES} systematic σ_c .

- Modeling of the underlying event energy.
- Amount of energy deposited out of the jet cone.

The uncertainty on each source is evaluated separately as a function of the jet p_T (and η for the first uncertainty in the list above). Their contributions are shown in Fig. 1 for the region $0.2 < \eta < 0.6$. The black lines show the sum in quadrature (σ_c) of all contributions. This $\pm 1\sigma_c$ total uncertainty is taken as a unit of jet energy scale miscalibration (Δ_{JES}) in this analysis.

V. EVENT RECONSTRUCTION

The event signature after event selection has a large missing transverse energy and big jet-multiplicities, and based on studies from Monte-Carlo (MC) samples we found that most of the selected events come from the $t\bar{t}$ lepton + jets decay channel, in which one of the W bosons from the top quark decay decays into two jets and the other into a lepton and a neutrino. We reconstruct the W boson mass from its hadronically decay jets to constrain *in situ* the jet energy scale (JES). For each possible jet-to-quark assignment of an event we calculate the invariant mass of two non-btagged jets, then we choose the one that has the value closest to the world average W mass measurement, $80.40 \text{ GeV}/c^2$, to be the reconstructed W mass m_{jj} . The fact that this analysis does not identify leptons makes it hard to fully reconstruct the event, thus hard to reconstruct the top quark mass. However, we can reconstruct a variable called M3, which is defined as the invariant mass of three jets that give the largest summation of jet E_T . To enhance the chance that these three jets come from the same top quark decay, we take two of the three jets from the m_{jj} reconstruction.

We also reconstruct a third variable called $M3'$, similar to the M3 mentioned above, except that the third jet of $M3'$, also one of the leading four jets, is different from any of the three jets that construct M3. Figure 2 shows the distribution of m_{jj} from MC samples with different input JES values and the same input top mass ($M_{top} = 172.5 \text{ GeV}/c^2$). The distributions of reconstructed M3 and $M3'$ are shown in Figure 3 and Figure 4, from samples with different input top quark masses but the same input JES (JES=0).

VI. KERNEL DENSITY ESTIMATES

Probability density functions for $M3$ - $M3'$ - m_{jj} at every point in the signal $M_{top} - \Delta_{JES}$ grid and for backgrounds are derived using a Kernel Density Estimate (KDE) approach. KDE is a non-parametric method for forming density estimates that can easily be generalized to more than one dimension, making it useful for this analysis, which has three observables per event. The probability for an event with observable (x) is given by the linear sum of contributions from all entries in the MC:

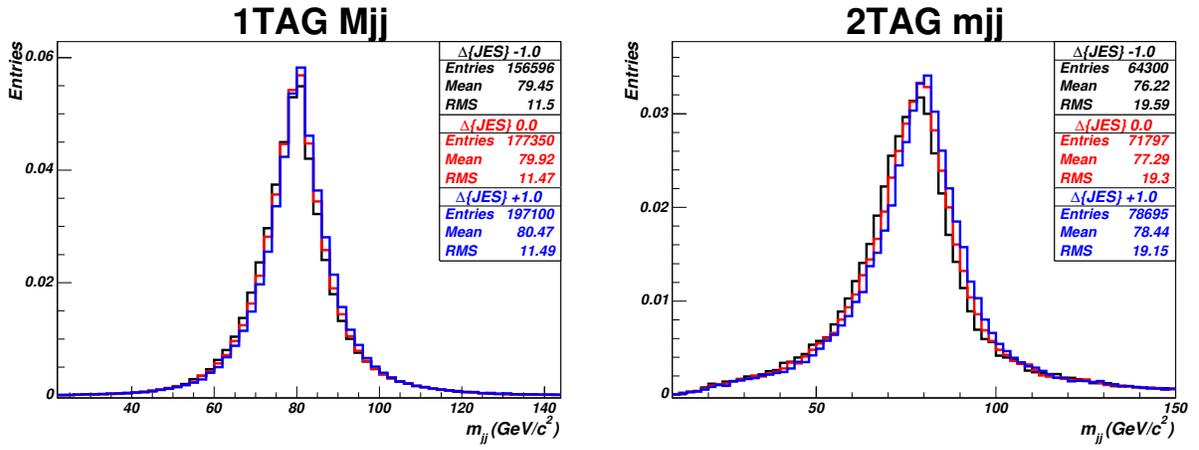


FIG. 2: Distribution of reconstructed m_{jj} with different input JES's, after all cuts are applied, both 1tag (left) and 2tag (right) are shown.

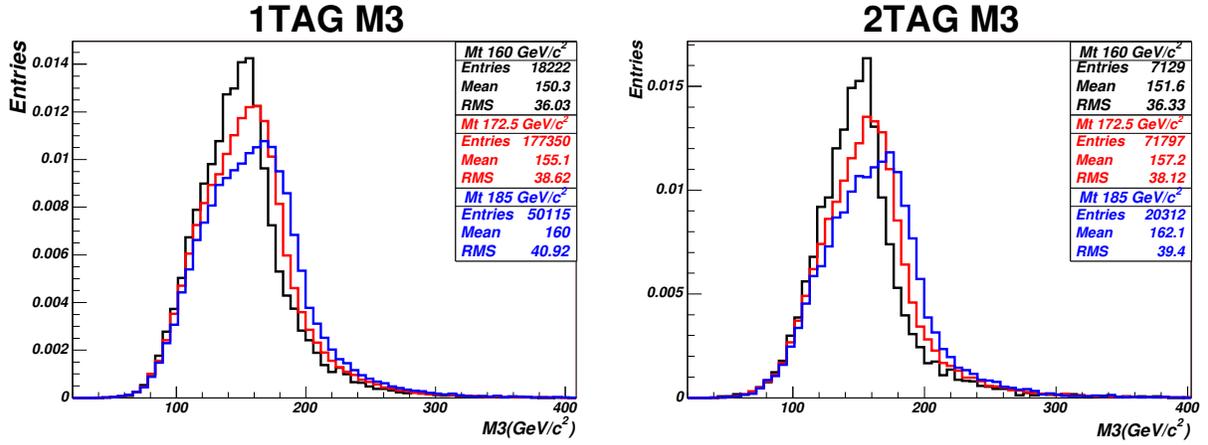


FIG. 3: Distribution of reconstructed M3 with different input top quark masses, after all cuts are applied, both 1tag (left) and 2tag (right) are shown.

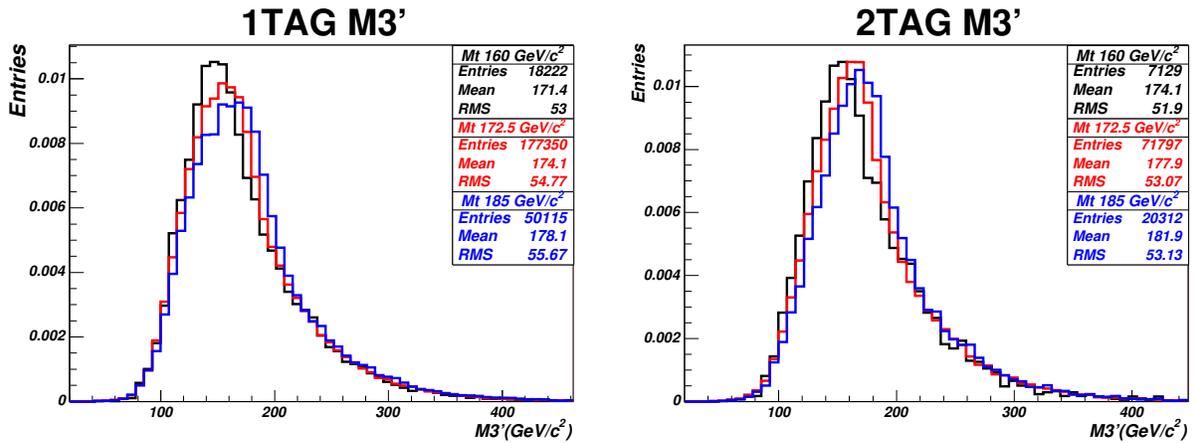


FIG. 4: Distribution of reconstructed M3' with different input top quark masses, after all cuts are applied, both 1tag (left) and 2tag (right) are shown.

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x-x_i}{h}\right). \quad (\text{VI.1})$$

In the above equation, $\hat{f}(x)$ is the probability to observe x given some MC sample with known mass and JES (or the background). The MC has n entries, with observables x_i . The kernel function K is a normalized function that adds less probability to a measurement at x as its distance from x_i increases. The smoothing parameter h (sometimes called the bandwidth) is a number that determines the width of the kernel. Larger values of h smooth out the contribution to the density estimate and give more weight at x farther from x_i . Smaller values of h provide less bias to the density estimate, but are more sensitive to statistical fluctuations. We use the Epanechnikov kernel, defined as:

$$K(t) = \frac{3}{4}(1-t^2) \text{ for } |t| < 1 \text{ and } K(t) = 0 \text{ otherwise,} \quad (\text{VI.2})$$

so that only events with $|x-x_i| < h$ contribute to $\hat{f}(x)$. We use an adaptive KDE method in which the value of h is replaced by h_i in that the amount of smoothing around x_i depends on the value of $\hat{f}(x_i)$. In the peak of the distributions, where statistics are high, we use small values of h_i to capture as much shape information as possible. In the tails of the distribution, where there are few events and the density estimates are sensitive to statistical fluctuations, a larger value of h_i is used. The overall scale of h is set by the number of entries in the MC sample (larger smoothing is used when fewer events are available), and by the RMS of the distribution (larger smoothing is used for wider distributions). We extend KDE to two dimensions by multiplying the two kernels together:

$$\hat{f}(x, y) = \frac{1}{n} \sum_{i=1}^n \frac{1}{h_{x,i} h_{y,i}} \left[K\left(\frac{x-x_i}{h_{x,i}}\right) \times K\left(\frac{y-y_i}{h_{y,i}}\right) \right]. \quad (\text{VI.3})$$

Figure 5 shows the 2d density estimates of signal events for each pair of the three observables.

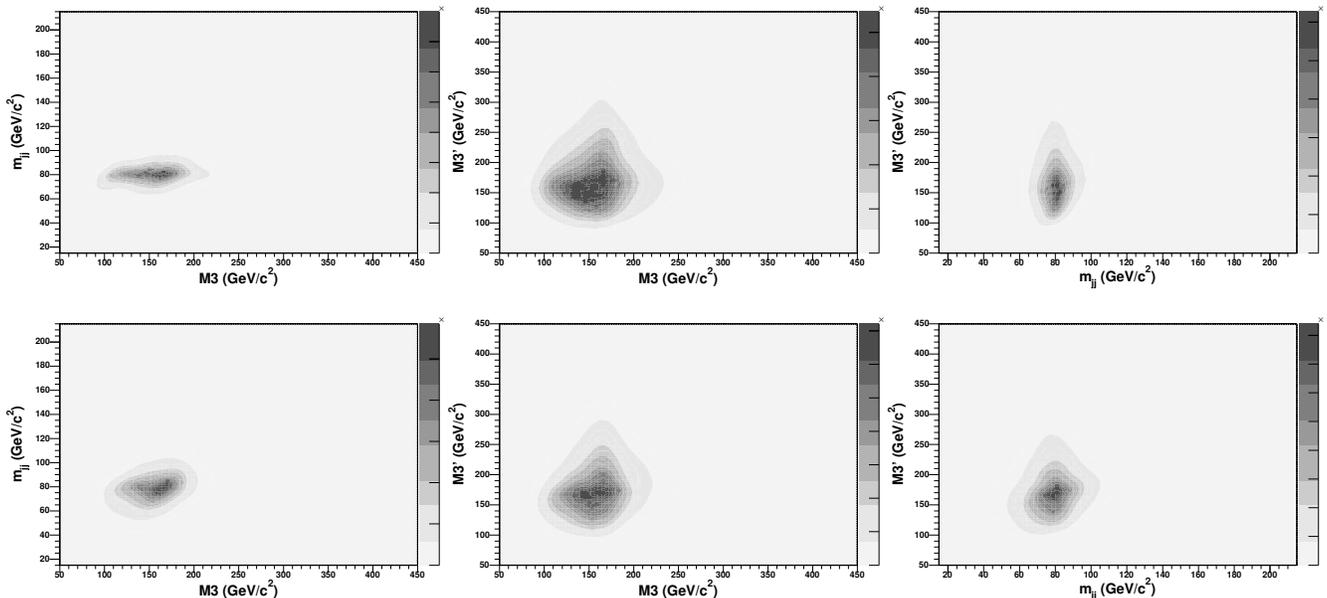


FIG. 5: Full 2d density estimates for input mass of $172.5 \text{ GeV}/c^2$ and $\Delta_{\text{JES}} = 0.0$ for 1-tag events (up) and 2-tag events (down).

VII. LIKELIHOOD FIT

The template method uses a multi-dimensional likelihood function fitting to find the most probable top mass given the data samples. We divide the original data samples into subsamples of 1btag and 2btag events separately, and for each subsample we define a likelihood function with the following format:

$$\mathcal{L}_{\text{shape}} = \frac{\exp(-(n_s + n_b))(n_s + n_b)^N}{N!} \times e^{-\frac{(n_{b0} - n_b)^2}{2\sigma_{n_{b0}}^2}}$$

$$\times \prod_{i=1}^N \frac{n_s P_s(M3, M3', m_{jj}; M_{\text{top}}, \Delta_{\text{JES}}) + n_b P_b(M3, M3', m_{jj})}{n_s + n_b}$$

where n_s and n_b are signal and background expectations and N is the number of events in the sample, P_s is the signal probability density function and P_b is the background probability density function. The first term in the likelihood is present because this is an extended maximum likelihood, in which the numbers of signal and background events obey Poisson statistics. The second term in the product expresses the Gaussian constraints on the background expectation. We use the *a-priori* estimate n_{b0} and its uncertainty $\sigma_{n_{b0}}$ to improve sensitivity. Shape information is used in the third term where probability density functions are used to discern between signal and background events and to extract mass information. A term like this exists for each of the two subsamples, and the final likelihood function is a product of them. We also impose a unit Gaussian constraint on Δ_{JES} .

The above gives the likelihood value only for points in the $M_{\text{top}} - \Delta_{\text{JES}}$ grid, and not as a continuous function. To obtain density estimates for an arbitrary point in the $M_{\text{top}} - \Delta_{\text{JES}}$ grid, we use local polynomial smoothing on a per-event basis. The value of the density estimate is obtained for an event at the available points, and a quadratic fit is performed in $M_{\text{top}} - \Delta_{\text{JES}}$ space, where the values of M_{top} and Δ_{JES} far away from the point being estimated are downweighted. This allows for a smooth likelihood that can be minimized. The measured uncertainty on M_{top} comes from the largest possible shift in M_{top} on the $\Delta \ln \mathcal{L} = 0.5$ contour.

VIII. METHOD CHECK

It is possible that our method has biases to the measured top quark mass. To estimate the bias, we run pseudo-experiments (PE) on signal samples with different input M_{top} and Δ_{JES} . Since the background sample is estimated from real data events, we use the same background events for different MC signal samples when running PE's. The number of background events used to run PE is the number we obtain from background estimation, while the number of signal events drawn to run PE is calculated from the theoretical calculation with cross section of $\sigma = 7.4$ pb, at $M_{\text{top}} = 172.5$ GeV/ c^2 .

Figure 6 shows that our method biases the top mass to a certain degree. We use a linear function to fit the measured top mass vs the input top mass, and apply this linear fit as a calibration to our measured top quark mass. The figure also shows the pullwidth plot of measured top mass, in which the pullwidth is defined, for each ensemble of PE's for a MC sample, as the sigma of the Gaussian fit of the mass pull distribution.

Note that the calibration of measured mass will also change the pullwidth of a sample, we therefore take similar steps to calibrate the pullwidth for each MC sample. After the pullwidth calibration, our expected statistical uncertainty on top mass increases about 20%. Figure 7 shows the biascheck after both calibrations, and the constant fitting lines on both plots tell us that after calibration the measured top mass well matches the input top mass and that we have a reasonable estimated statistical uncertainty on top mass measurement.

We then inspect the measured top mass dependencies on Δ_{JES} . Three top mass MC samples are chosen: $M_{\text{top}} = 167.5$ GeV/ c^2 , 172.0 GeV/ c^2 , and 177.5 GeV/ c^2 , each with five different Δ_{JES} : $0.0, \pm 0.4, \pm 1.0$. By applying the same calibration described above, we get the plots shown on Fig. 8, which tell us that our measured top mass after calibration has negligible dependency on Δ_{JES} . Figure 9, which presents the biascheck of measured Δ_{JES} vs input Δ_{JES} , after calibration, also shows a good agreement between the measured Δ_{JES} and input Δ_{JES} after calibration.

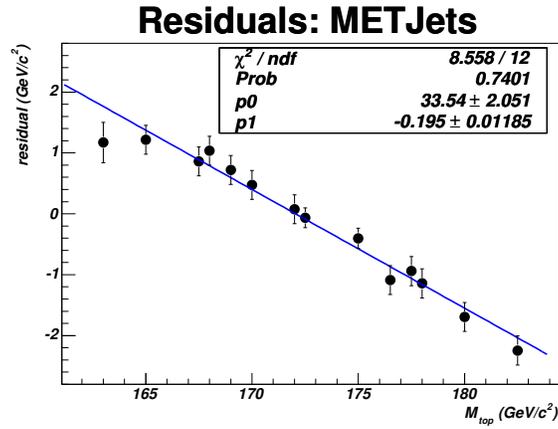


FIG. 6: Biaschecks of the measured top quark mass before any correction/calibration. All the MC samples have the same $\Delta_{\text{JES}} = 0.0$.

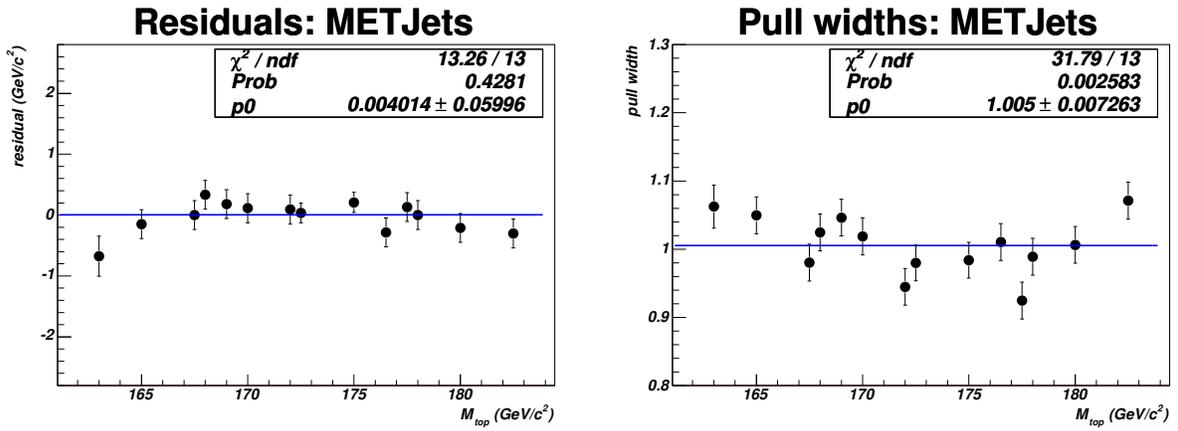


FIG. 7: Biaschecks of the measured top quark mass after calibrations. Residual plot is on the left, and the pullwidth plot on the right. All the samples have the same $\Delta_{\text{JES}} = 0.0$.

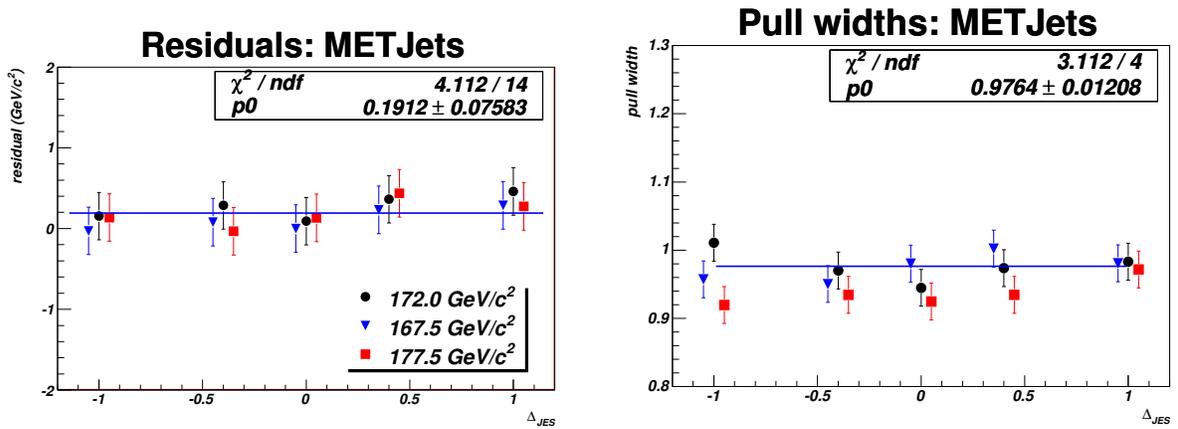


FIG. 8: Checks of the measured top mass dependencies on the Δ_{JES} . Plots shown here are after calibration.

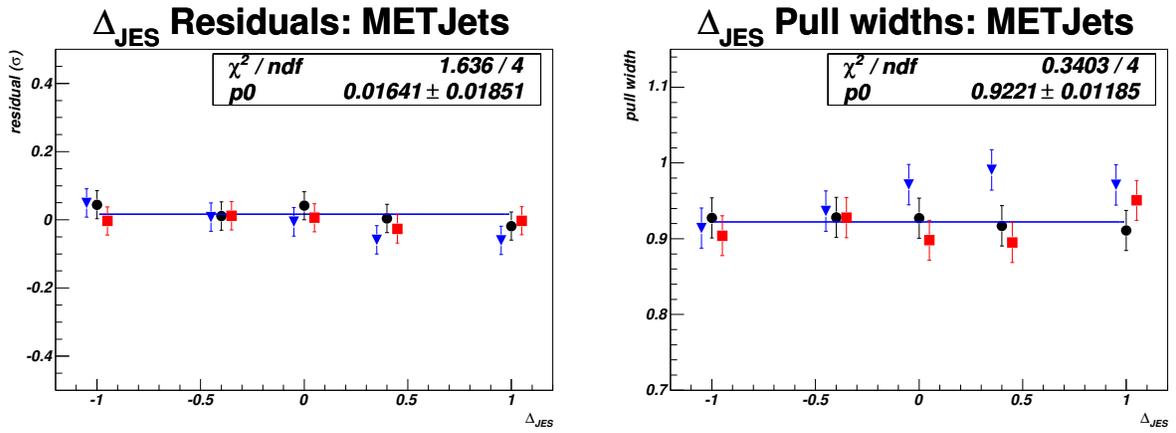


FIG. 9: Biaschecks of the measured Δ_{JES} vs the input Δ_{JES} . Different color represents different input top masses. Plots shown here are after calibration.

IX. FIT RESULTS

With the likelihood function we now can fit the data. Figure 10 is the two-dimensional likelihood fit for data, from which we measure (after the calibration)

$$M_{top} = 172.3 \pm 2.4 \text{ GeV}/c^2, \text{ calibrated}$$

$$\Delta_{JES} = 0.3 \pm 0.4 \text{ calibrated}$$

Figure 11 shows the three kinematic variables from data, overlaid with their corresponding one-dimensional p.d.f's from sample $t\text{top}25$ (with input $M_{top} = 172.5 \text{ GeV}/c^2$ and $\Delta_{JES} = 0.0$) plus the background fully modeled.

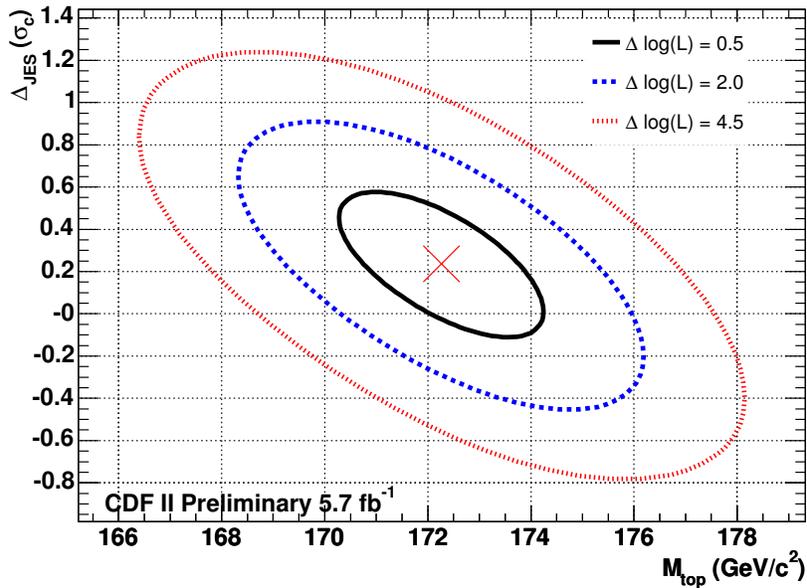


FIG. 10: 2D likelihood fit contours for the data fitting.

Figure 12 shows the p-value test of the measured statistical uncertainty.

X. SYSTEMATIC UNCERTAINTIES

We examine a variety of effects that could systematically shift our measurement. As a single nuisance parameter, the JES that we measure does not fully capture the complexities of possible jet energy scale uncertainties, particularly those with different η and p_T dependence. Fitting for the global JES removes most of these effects, but not all of them. We apply variations within uncertainties to different JES calibrations for the separate known effects in both signal and background pseudodata and measure resulting shifts in M_{top} from pseudoexperiments, giving a residual JES uncertainty. We also vary the energy of b jets, which have different fragmentation than light quarks jets, as well as semi-leptonic decays and different color flow, resulting in a b-JES systematic. Effects due to uncertain modeling of radiation including initial-state radiation (ISR) and final-state radiation (FSR) are studied by extrapolating uncertainties in

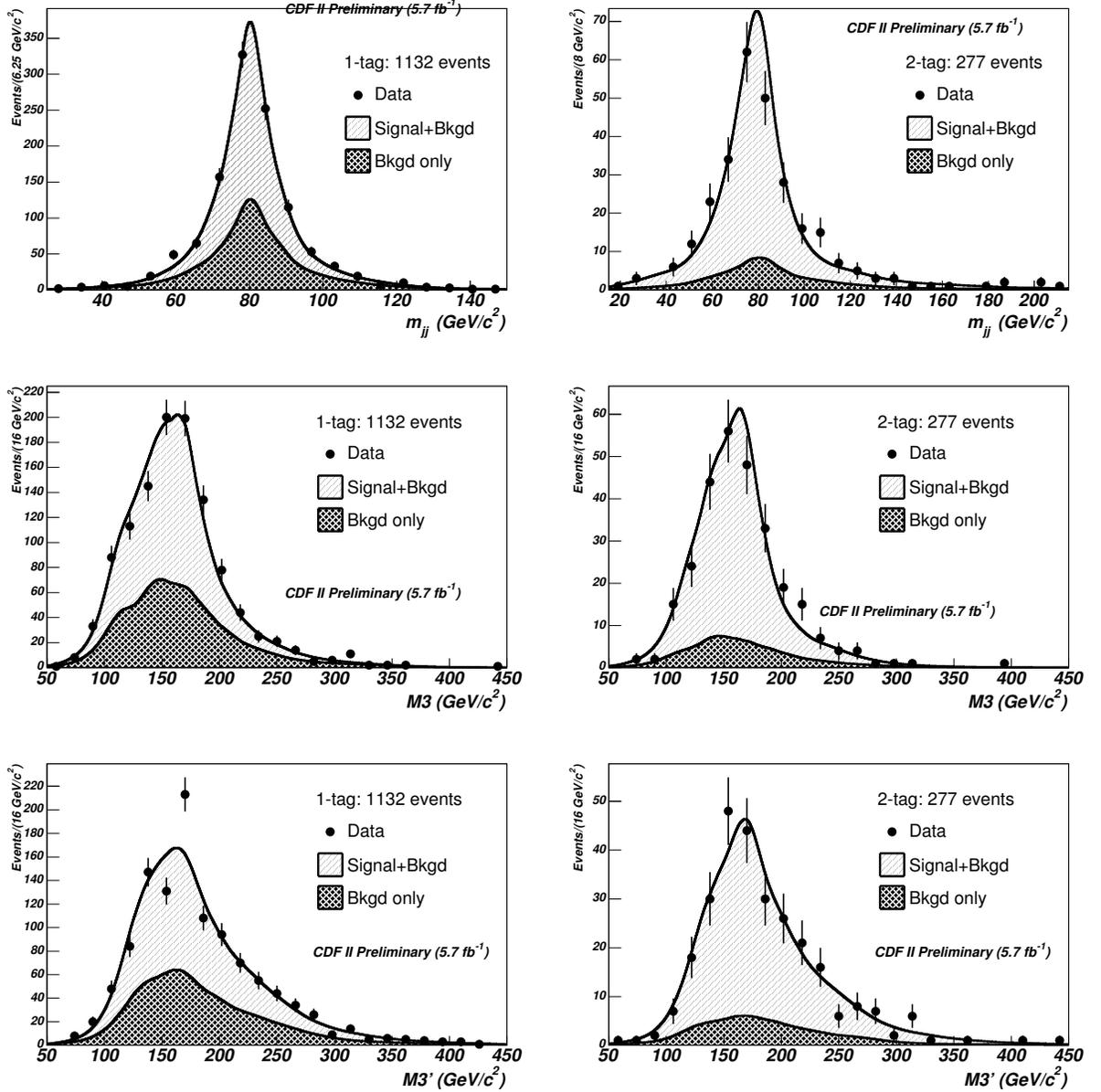


FIG. 11: Distribution of kinematic variables $M3$, m_{jj} , and $M3'$ from data, overlaid with their corresponding 1-d p.d.f's from MC sample $t\bar{t}p25$ (with input $M_{\text{top}} = 172.5 \text{ GeV}/c^2$ and $\Delta_{\text{JES}} = 0.0$) plus the background fully modeled. Both 1tag (left) and 2tag (right) events are displayed.

TABLE III: Summary of systematic effects. All numbers are after calibration.

Systematic (GeV/c^2)	ΔM_{top}
Residual JES	0.50
Generator	0.65
PDFs	0.20
b jet energy	0.29
Background shape	0.12
gg fraction	< 0.01
Radiation	0.21
Trigger simulation	0.14
Multiple Hadron Interaction	0.16
Color Reconnection	0.20
Total Effect	0.98

the p_T of Drell-Yan events to the $t\bar{t}$ mass region, resulting in a radiation systematics. Comparing pseudoexperiments generated with HERWIG and PYTHIA gives an estimate of the generator systematic. A systematic on different parton distribution functions is obtained by varying the independent eigenvector of the CTEQ6M set, comparing parton distribution functions with different values of Λ_{QCD} , and comparing CTEQ5L with MRST72. We also test the effect of reweighting MC to increase the fraction of $t\bar{t}$ events initiated by gg (vs qq) from the 6% in the leading order MC to 20%. Systematic effects due to the background modeling and trigger simulation are also taken into account. The color reconnection effects are accounted by generating new MC sample which have color effects.

The total systematic error is $1.0 \text{ GeV}/c^2$, summarized in Table III.

XI. CONCLUSIONS

We present a measurement of the top quark mass using events that have a signature of MET + high jet-multiplicities with a template-based technique. An *in situ* JES calibration is used to constrain the jet energy scale. Using three dimensional templates derived from Kernel Density Estimation and 5.7 fb^{-1} of data collected by the CDF II at Fermilab, we measure

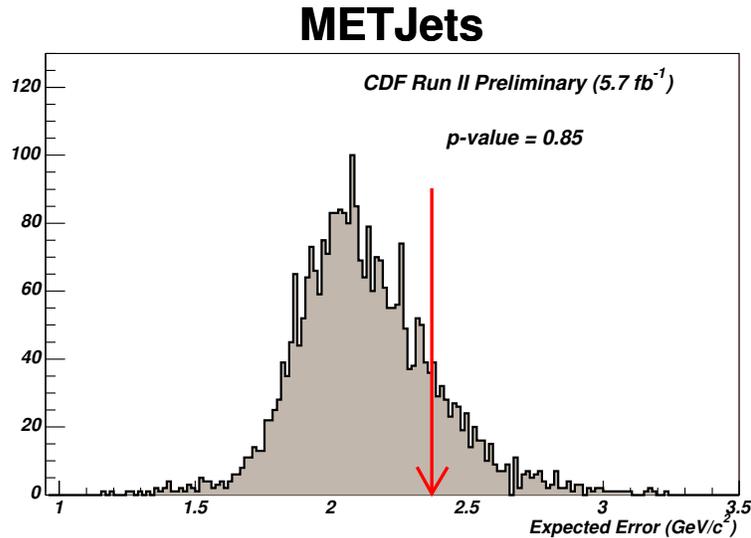


FIG. 12: P-value test of the measured statistical uncertainty (calibrated) of top mass.

$$\begin{aligned}
M_{top} &= 172.3 \pm 2.4 \text{ (stat. + JES)} \pm 1.0 \text{ (syst.) } GeV/c^2 \\
&= 172.3 \pm 2.6 GeV/c^2 \\
\Delta_{JES} &= 0.3 \pm 0.4 \text{ (stat. + } M_{top} \text{ only) } \sigma_c
\end{aligned}$$

Acknowledgments

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] F. Abe, et al., Nucl. Instrum. Methods Phys. Res. A **271**, 387 (1988);
 - [2] A. Abulenci et al., Phys. Rev. D **73**, 032003 (2006);
 - [3] A. Abulenci et al., Phys. Rev. Lett. **96**, 022004 (2006);
 - [4] F. Abe, et al., Phys. Rev. D **45**, 1448 (1992);
 - [5] T. Affolder, et al., Phys. Rev. D **64**, 032002 (2001);
 - [6] A. Bhatti, et al., Nucl. Instrum. Methods Phys. Res. A **566**, 375 (2006);