Template-based Top Quark Mass Measurement in the Lepton+Jets Channel Using 1.7 $\mathbf{fb}^{-1}$ of data

The CDF Collaboration
URL http://www-cdf.fnal.gov
(Dated: August 9, 2007)

We report on a measurement of the top quark mass in the lepton plus jets channel of $t\bar{t}$ events from $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV using 1.7fb $^{-1}$ of data collected with the CDF detector at the Fermilab Tevatron. A top quark mass is reconstructed for every event by minimizing a χ^2-like function to the overconstrained kinematics of the $t\bar{t}$ system. The dijet mass of the hadronically decaying W boson is used to constrain *in situ* the largest systematic on top quark mass measurements, the uncertain jet energy scale (JES) in the detector. The reconstructed top quark masses and dijet masses for 307 events with at least one b-tag are compared to two-dimensional probability density function derived by applying kernel density estimation to fully simulated MC events with different values of the top quark mass and JES in the detector. We measure $\mathrm{M_{top}} = 171.6 \pm 2.1$ (stat.) ± 1.1 (syst.) GeV/c^2.

Preliminary Results for Summer 2007 Conferences

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 consistent with $t\bar{t}$ production and decay via the lepton+jets channel, in which one W boson decays hadronically, and the other W boson decays leptonically. The event signature is then 4 jets (one from each of the two b quarks and two from the hadronically decaying W), a high-p_T electron or muon, and missing energy in the transverse plane consistent with production of a neutrino. The CDF detector is described in detail 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 parent top quark mass distribution. We use one variable ($m_t^{\rm reco}$) that is strongly correlated with the true top quark mass ($\rm M_{top}$), and one variable ($m_{\rm jj}$) that is strongly correlated with shifts in the jet energy scale ($\Delta_{\rm JES}$) in the detector. We take 69 sets of MC with generated with different top quark masses, and run through full CDF detector simulation and the reconstructed described below assuming 29 possible shifts in $\Delta_{\rm JES}$. The value of $m_t^{\rm reco}$ in each event is derived from a χ^2 minimization that uses knowledge of the overconstrained kinematics of the $t\bar{t}$ system. The dijet mass ($m_{\rm jj}$) that we use in each event is chosen such that it often comes from the decay of the W resonance, and is sensitive to possible miscalibration of JES in the CDF detector. The values of $m_t^{\rm reco}$ and $m_{\rm jj}$ in data are compare to each point in the MC grid using a non-parametric approach based on kernel density estimation, and a two-dimensional parabola is fit to the log-likelihood values in the grid to both measure $\rm M_{top}$ and profile out the JES.

II. EVENT SELECTION

At the trigger level, lepton+jets candidate events are selected by requiring a high-E_T electron (or high-p_T muon). Offline, the events are required to have a single energetic lepton, large $\not\!\!E_T$ due to the escaping neutrino from the W decay, and at least four jets in the final state. Electron candidates are identified as a high-momentum track in the tracking system matched to an electromagnetic cluster reconstructed in the calorimeters with $E_T > 20$ GeV. We also require that energy shared by the towers surrounding the cluster is low. Muon candidates are reconstructed as high-momentum tracks with $p_T > 20$ GeV/c matching hits in the muon chambers. Energy deposited in the calorimeter is required to be consistent with a minimum ionizing particle. The $\not\!\!E_T$ is required to be greater tham 20 GeV.

Jets are reconstructed with the JETCLU cone algorithm using a cone radius of $R \equiv \sqrt{\eta^2 + \phi^2} = 0.4$. To improve the statistical power of the method, the sample is divided into two subsamples, depending on the number of jets identified as arising from the hadronization and decay of b quarks. The SECVTX algorithm uses the transverse decay length of tracks inside jets to tag jets as coming from b quarks. We require at least one tagged jet per event. In events with exactly one tag, we require exactly four jets with $E_T > 20$ GeV/c^2. For events with more than one tag, which have more statistical power and less background, we loosen these cuts, and allow events with more than four jets. We also loosen the cut on the 4th jet to $E_T > 12$ GeV/c^2 to increase the number of such events.

Two additional requirements are made on events. We make a cut on the χ^2 out of the kinematic fitter described in Section IV, requiring it to be less than 9.0. Finally, in order to properly normalize our probability density functions, we define hard boundaries on the values of the two observables, $m_t^{\rm reco}$ and $m_{\rm jj}$. Events with values of $m_t^{\rm reco}$ or $m_{\rm jj}$ falling outside the boundary are rejected. The combined χ^2 and boundary cut efficiency for the background is roughly 60% for 1-tag background events and 25-30% for 2-tag background events. These numbers are roughly 85% and 65% for signal events. The event selection and the observed number of events for both categories of events are summarized in Table I.

III. 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 [2]. There are many sources of uncertainties related to jet energy scale at CDF:

TABLE I: Event selection and observed numbers of events for the two event categories

	1-tag	2-tag
b-tags	==1	>1
Leading 3 jets E_T(GeV/c^2)	>20	>20
4th jet E_T(GeV/c^2)	>20	>12
Extra jets E_T(GeV/c^2)	<20	Any
χ^2	<9	<9
$m_t^{\rm reco}$ boundary cut (GeV/c^2)	$m_t^{\rm reco} > 110$ && $m_t^{\rm reco} < 350$	$m_t^{\rm reco} > 110$ && $m_t^{\rm reco} < 350$
$m_{\rm jj}$ boundary cut (GeV/c^2)	$m_{\rm jj} > 50$ && $m_{\rm jj} < 115$	$m_{\rm jj} > 30$ && $m_{\rm jj} < 120$
Observed # events	218	89
Expected background	36.6±7.1	6.4±1.7

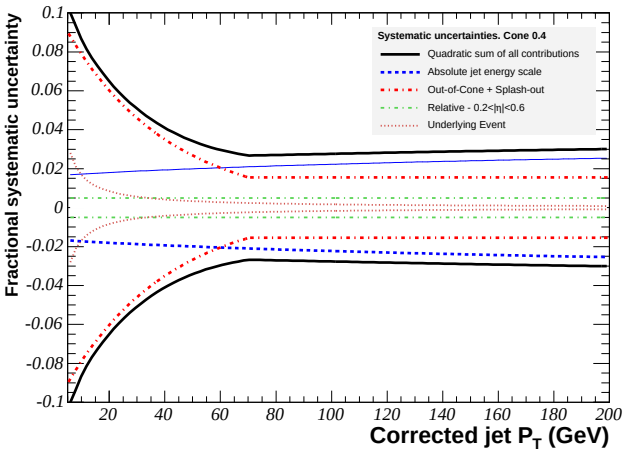

FIG. 1: Jet energy scale uncertainty as a function of the corrected jet p_T for the underlying event (dotted red), relative response (dashed green), out-of-cone energy (dashed red) and absolute response (dashed blue). The contribution of all sources are added in quadrature (full black) to form the total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

- Relative response of the calorimeters as a function of pseudorapidity.
- Single particle response linearity in the calorimeters.
- Fragmentation of jets.
- 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 of jet energy scale miscalibration (Δ_{JES}) in this analysis.

IV. TOP MASS RECONSTRUCTION

The value of the reconstructed mass in each event ($m_t^{\rm reco}$) is determined by minimizing a χ^2 describing the overconstrained kinematics of the $t\bar{t}$ system. $m_t^{\rm reco}$ is a number that distills all the kinematic information in each event into one variable that is a good estimator for the true top quark mass. The kinematic fitter uses knowledge of the lepton and jet four-vectors, b-tagging information and the measured $\not\!\!E_T$. The invariant masses of the lepton-neutrino pair and the dijet mass from the hadronic W decay are constrained to be near the well-known W mass, and the two top quark masses per event are constrained to be equal within the narrow top width. The χ^2,

$$\chi^2 = \Sigma_{i=\ell,4jets}\frac{(p_T^{i,fit}-p_T^{i,meas})^2}{\sigma_i^2} + \Sigma_{j=x,y}\frac{(U_j^{fit}-U_j^{meas})^2}{\sigma_j^2} + \frac{(M_{jj}-M_W)^2}{\Gamma_W^2} + \frac{(M_{\ell\nu}-M_W)^2}{\Gamma_W^2} + \frac{(M_{bjj}-m_t^{\rm reco})^2}{\Gamma_t^2} + \frac{(M_{b\ell\nu}-m_t^{\rm reco})^2}{\Gamma_t^2} \quad \text{(IV.1)}$$

is minimized for every jet-parton assignment consistent with b-tagging. The first sum constrains the p_T of the jets and lepton, within their uncertainties, to remain close to their measured values. The second term constrains the unclustered energy in the event to its measured value, providing a handle on the neutrino 4-vector. The W boson has a small width, and the two W mass terms provide the most powerful constraints in the fit. The last two terms in the χ^2 constrain the three-body invariant masses of each top decay chain to remain close to a single top quark mass, $m_t^{\rm reco}$. The single jet-parton assignment with the lowest χ^2 that is consistent with b-tagging gives the value of $m_t^{\rm reco}$ for the event. Events where the lowest $\chi^2 > 9.0$ are rejected. Distributions for $m_t^{\rm reco}$ in MC with three different input top quark masses are shown in in Figure 2.

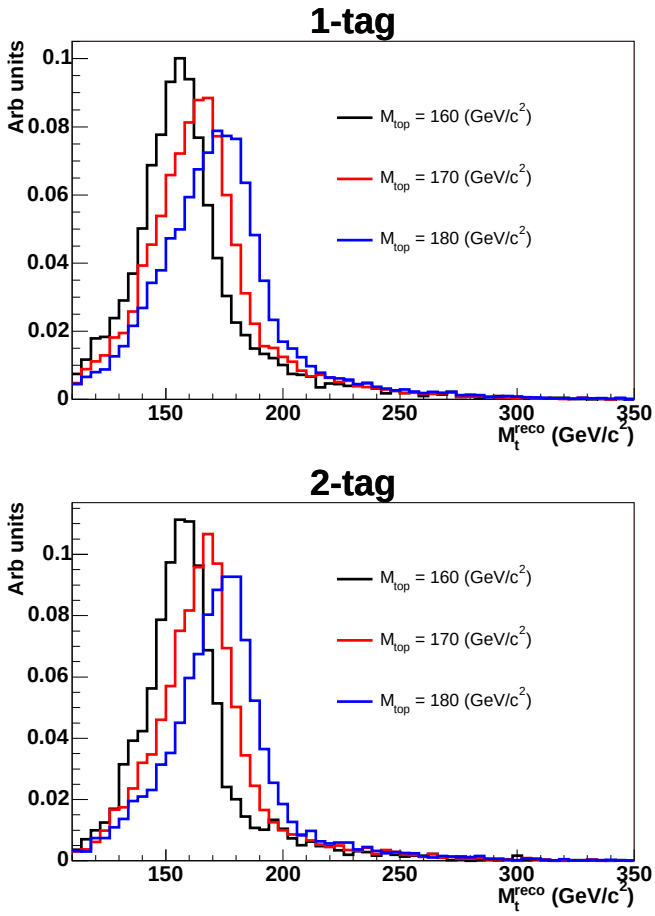

FIG. 2: Reconstructed top quark mass distributions for 1-tag(top) and 2-tag(bottom) events

V. DIJET MASS

The value of m_{jj} in each event can have an ambiguity due to not knowing which two jets came from a hadronic W decay. In 2-tag events, the value is chosen as the invariant mass of the two non-tagged jets in the leading 4 jets. In single-tag events, there are 3 dijet masses that can be formed from the 3 non-tagged jets among the 4 jets in the

event. We chose the single dijet mass that is closest to the well known W mass. Distributions for m_{jj} in MC with three different Δ_{JES} values are shown in Figure 3.

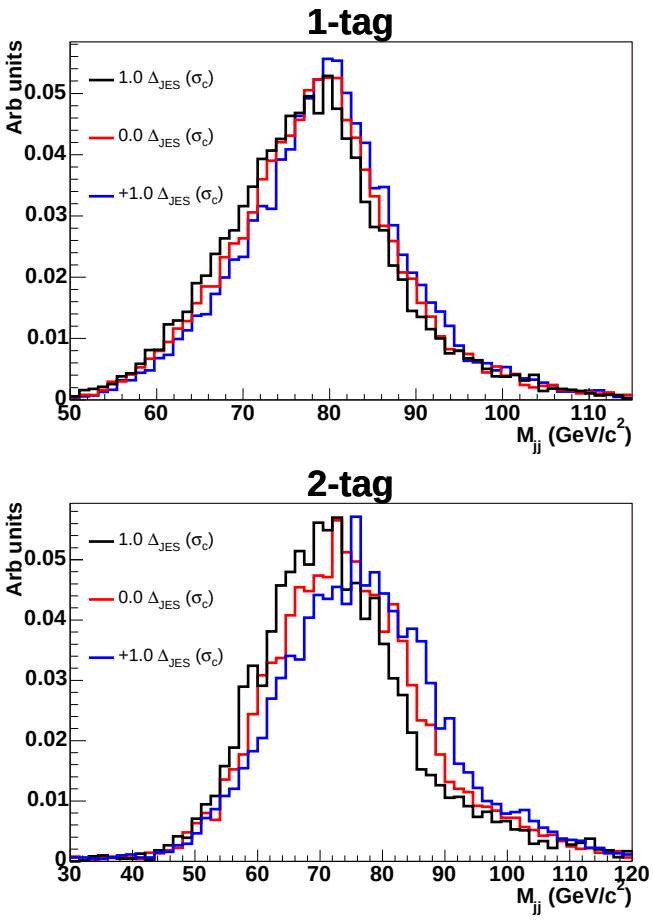

FIG. 3: Dijet mass distributions for 1-tag(top) and 2-tag(bottom) events for MC mass of 170 GeV/c^2.

VI. BACKGROUNDS

An *a priori* estimate for the background composition is used to derive background shapes for $m_t^{\rm reco}$ and m_{jj}. ALPGEN is used to model W+jets. Contributions include Wbb, Wcc, Wc and W+light flavor (LF) jets. The W+LF sample is used to model the QCD background, though a second, independent model using a non-isolated lepton sample is used to evaluate a systematic due to the uncertain QCD shape. The relative fractions of the different W+jets samples are determined in MC, but the absolute normalization is derived from the data. MC and theoretical cross-sections are used to model the single-top and diboson backgrounds. The fractions of the total estimated background from each individual background are shown in Table II. The background is assumed to have no M_{top}dependence. The small Δ_{JES} dependence is taken as a systematic uncertainty.

VII. KERNEL DENSITY ESTIMATES

Probability density functions for $m_t^{\rm reco}$ and m_{jj} at every point in the $M_{top} - \Delta_{JES}$ grid and for backgrounds are derived using a Kernel Density Estimate (KDE) approach. KDE is a non-parametric approach to forming density estimates that can easily be generalized to more than one dimension, making it useful for this analysis, which has two 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:

TABLE II: Background fractions from different sources after event selection and boundary cuts

	1-tag	2-tag
W+LF(mistags)	0.49	0.28
W+bb+jets	0.21	0.44
W+cc+jets	0.12	0.06
W+c+jets	0.08	0.02
s-channel single-top	0.02	0.09
t-channel single-top	0.03	0.06
WW	0.04	0.02
WZ	0.01	0.03
ZZ	0.001	0.002

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} K(\frac{x - x_i}{h}) \tag{VII.1}$$

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 varying probability to a measurement at x depending on its distance from x_i. 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 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:

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

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). Figures 4 shows the effect of using adaptive KDE – the tails are smoothed out more, and the core of the distribution is better described.

We extend KDE to two dimensions by multiplying the two kernels together:

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

Figures 5 and 6 show the 2d density estimates for both signal and background.

VIII. LIKELIHOOD FIT

The values of m_t^{reco} and m_{jj} observed in data are compared to every point in the $\text{M}_{\text{top}} - \Delta_{\text{JES}}$ grid. An extended maximum likelihood fit is performed at each point to minimize the likelihood with respect to the expected number of signal (n_s) and background events (n_b) in each of the two subsamples. A Gaussian constraint in the expected number of background events is applied to each subsamples. The likelihood for subsample k with N events is given by:

$$\mathcal{L}_k = \exp\left(-\frac{(n_b - n_b^0)^2}{2\sigma_{n_b}^2}\right) \times \prod_{i=1}^{N} \frac{n_s P_{sig}(m_{t,i}^{reco}, m_{jj,i}; M_{top}, \Delta_{JES}) + n_b P_{bg}(m_{t,i}^{reco}, m_{jj,i})}{n_s + n_b} \tag{VIII.1}$$

The overall likelihood is a product over the two individual subsample likelihoods, with a Gaussian constraint on Δ_{JES}, constraining it to the nominal $0 \pm 1\ \sigma_c$:

$$\mathcal{L} = \exp\left(-\frac{\Delta^2_{\rm JES}}{2{\sigma_c}^2}\right) \times \mathcal{L}_{\text{1-tag}} \times \mathcal{L}_{\text{2-tag}} \tag{VIII.2}$$

Given the $-\ln\mathcal{L}$ in the $M_{top} - \Delta_{JES}$ grid, we fit a 2d-parabola (with a cross-term to account for the correlation between the two variables) to measure M_{top} and profile out the JES. The uncertainty on M_{top} comes from the largest possible shift in M_{top} on the $\Delta\ln\mathcal{L} = 0.5$ contour.

IX. METHOD CHECK

We test our machinery by running pseudoexperiments with varying values of M_{top} between 165 and 180 GeV/c^2 and varying values of Δ_{JES} between -1.0 and 1.0 σ_c. Figure 7 shows the M_{top} and Δ_{JES} residuals as a function of true top quark mass. We conclude that the method is unbiased in its prediction of M_{top}. A small bias in Δ_{JES} may by present, but the reported value is not used in measuring M_{top}.

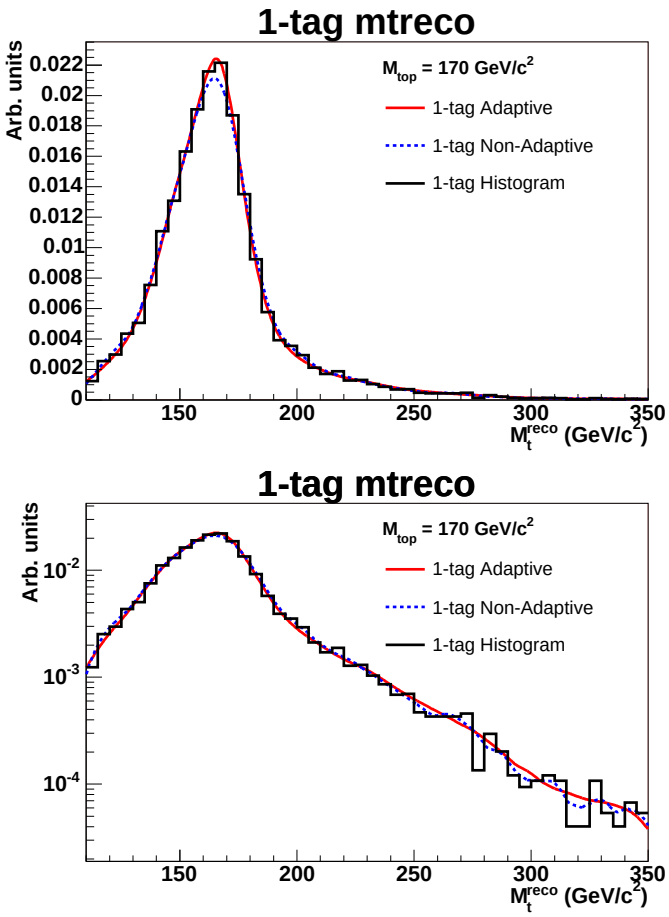

FIG. 4: 1-tag $m_t^{\rm reco}$ distributions using histograms for density estimation (black), nominal KDE (dotted blue lines) and adaptive KDE (red line). The two plots are the same, but the one on the bottom is on a log scale. The input mass used is 170 GeV/c^2.

X. RESULTS

The likelihood procedure when applied to the data yields $M_{top} = 171.6 \pm 2.0$ GeV/c^2. We examine whether any scaling of our error is needed by examining the width of differential pull distributions, which show the pull width as a function of reported error. Using 3 input masses close to the measured mass to increase statistics, we find the distribution shown in Figure 8. The value of the pull width on the best-fit line at our measured error is 1.047, which we use to correct our statistical uncertainty. The log-likelihood contours for our measurement are shown in Figure 9. As shown in Figure 10, 23.4% of pseudoexperiments have a smaller error than the value measured in the data.

We run a variety of cross-checks on the data. We fit without the JES constraint, and also without the background constraint, showing that these priors do not significantly change our result. A 1d fit for only the top quark mass gives an uncorrected result of $M_{top} = 171.7 \pm 1.5$ GeV/c^2 (stat only), allowing us to separate the statistical part of the 2.0 GeV/c^2 stat+JES error from the part due to allowed variations of Δ_{JES} (also 1.5 GeV/c^2). We also fit the 1-tag and 2-tag samples separately, and perform individual measuresments dividing by lepton type. The results of our cross-checks are summarized in Table III.

XI. 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

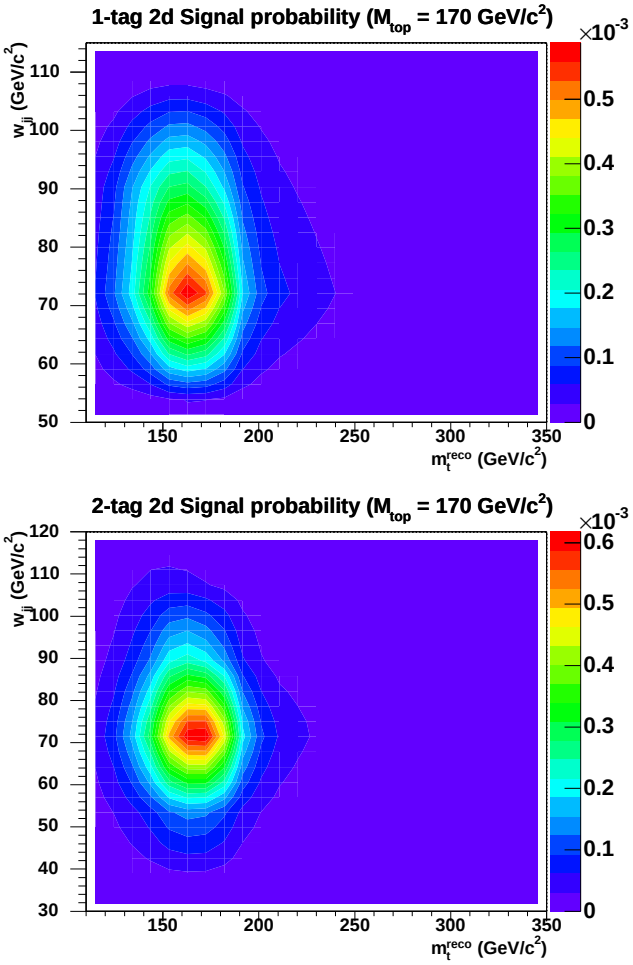

FIG. 5: Full 2d density estimates for input mass of 170 GeV/c^2 and $\Delta_{JES} = 0.0$ for 1-tag events (top) and 2-tag events (bottom).

TABLE III: Measurement cross-checks. All errors are uncorrected.

	Mass (GeV/c^2)
Nominal	171.6 ± 2.0
No background prior constraint	171.7 ± 2.0
No JES prior constraint	171.6 ± 2.1
1-tag only	167.5 ± 3.1
2-tag only	175.0 ± 2.6
Electron-only (no bkgd prior)	170.6 ± 2.9
Muon-only (no bkgd prior)	172.3 ± 3.0
1d-only fit (no JES)	171.7 ± 1.5

those with different η and p_T dependence. Fitting for the 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 and measure resulting shifts in M_{top} from pseudoexperiments. The shifts added in quadrature result in a residual JES systematic of 0.55 GeV/c^2. Varying the energy of b jets, which have different fragmentation than light quarks jets, as well as semi-leptonic decays and different color flow, results in an additional b-JES systematic of 0.6 GeV/c^2. An additional JES uncertainty due to not shifting JES in our background PDFs is estimated to be 0.38 GeV/c^2. Effects due to uncertain modeling of initial-state radiation (ISR) and final-state radiation (FSR) are studied by extrapolating uncertainties in the p_T of Drell-Yan events to the $t\bar{t}$ mass region, resulting in an ISR systematic of 0.37 GeV/c^2

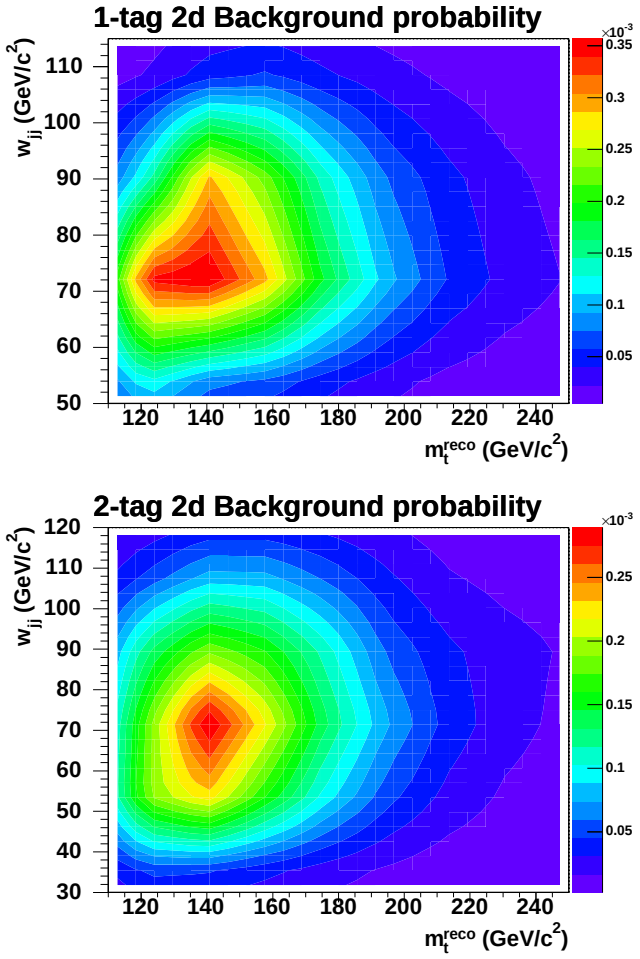

FIG. 6: Full 2d density estimates for the combined background for 1-tag events (top) and 2-tag events (bottom).

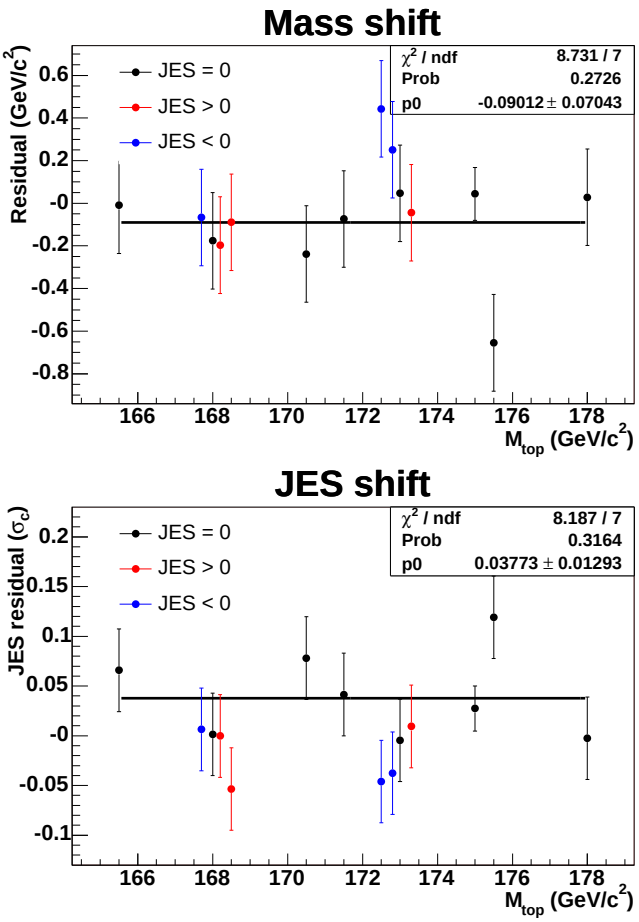

FIG. 7: Residual mass shift as a function of input mass from pseudoexperiments. For input mass = 168 GeV/c^2 we use $\Delta_{\rm JES}$ = -0.6, +0.4 and +1.0 σ_c and for input mass = 173 GeV/c^2 we use $\Delta_{\rm JES}$ = -1.0, -0.4 and +0.6 σ_c. Straight-line fits are to JES = 0.0 points only.

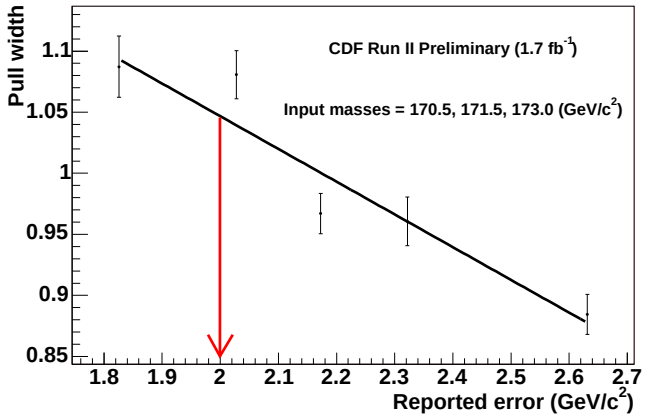

FIG. 8: Differential pull width as a function of reported error.

TABLE IV: Summary of systematics.

Effect	Systematic(GeV/c^2)
ISR	0.37
FSR	0.23
Generator	0.25
Residual JES	0.55
Background JES	0.38
PDFs	0.17
Background composition	0.24
Background shape	0.2
QCD modeling	0.11
B-JES	0.6
MC statistics	0.1
Total	1.1

and an FSR systematic of 0.25 GeV/c^2. Varying the uncertain Q^2 of background events results in a background shape systematic of 0.2 GeV/c^2. An additional background composition systematic of 0.24 GeV/c^2 comes from drawing from single types of backgrounds in pseudoexperiments instead of from the estimated background fractions. A systematic of 0.1 GeV/c^2 on QCD modeling is derived by examining an independent model of QCD events. Comparing pseudoexperiments generated with HERWIG and PYTHIA gives a generator systematic of 0.25 GeV/c^2. Finally, a systematic of 0.17 GeV/c^2 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. The total systematic error is 1.1 GeV/c^2.

XII. CONCLUSIONS

We presented a measurement of the mass of the top quark in the Lepton+Jets channel using a template-based measurement with an *in situ* JES calibration. Using 2d templates derived from Kernel Density Estimation and 1.7 fb^{-1} of data collected by the Tevatron, we measure

$$\mathrm{M_{top}} = 171.6 \pm 2.1\ (\mathrm{stat.}) \pm 1.1\ (\mathrm{syst.})\ \mathrm{GeV}/c^2 \tag{XII.1}$$

Adding the statistical and systematic errors in quadrature by assuming Gaussian errors, we measure

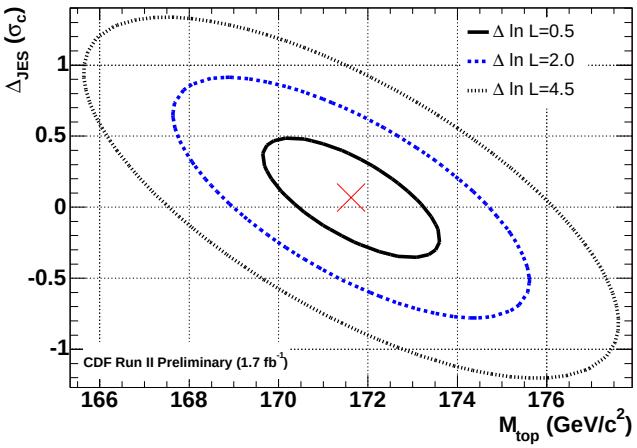

FIG. 9: Negative log-likelihood contours for the data.

$$M_{top} = 171.6 \pm 2.4 \text{ GeV}/c^2 \quad \text{(XII.2)}$$

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. Bhatti, et al., Nucl. Instrum. Methods Phys. Res. A **566**, 375 (2006);

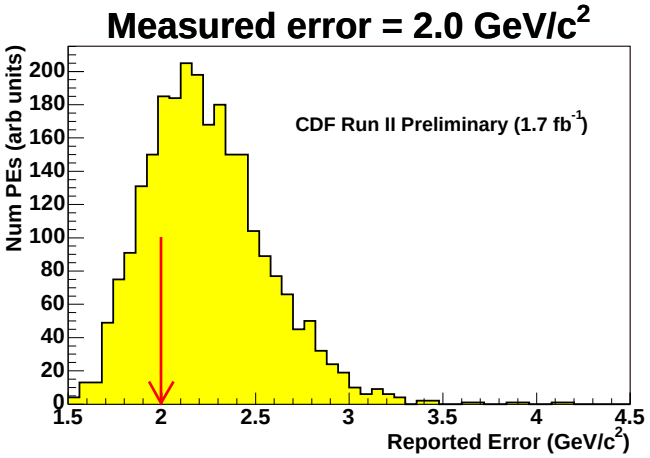

FIG. 10: Reported error from pseudoexperiments with the observed number of events.