CDF note 10033

Combined template-based Top Quark Mass Measurement in the Lepton+Jets and Dileptons Channels Using 4.8 fb^{-1} of data

The CDF Collaboration
URL http://www-cdf.fnal.gov
(Dated: February 18, 2010)

We report on a simultaneous measurement of the top quark mass ($\mathrm{M_{top}}$) in the Lepton+Jets and Dilepton channels using $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV from 4.8 fb $^{-1}$ of data collected with the CDF detector at the Fermilab Tevatron. In the Lepton+Jets channel, a top quark mass ($\mathrm{m}_t^{\mathrm{reco}}$) 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 ($\mathrm{m_{jj}}$) 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 (Δ_{JES}) in the detector. To bring up more information of top quark mass, we include a reconstructed top quark mass from 2nd best χ^2 fit ($m_t^{\mathrm{reco}(2)}$). In the underconstrained dilepton channel, the Neutrino Weighting Algorithm (NWA) is used to integrate over the pseudorapdities of the neutrinos and give a single reconstrcted mass per event ($\mathrm{m}_t^{\mathrm{NWA}}$). In addition, m_{T2} variable, which is a quantity of transverse mass in two missing particles system is used to improve mass resolution. The values of $\mathrm{m}_t^{\mathrm{reco}}$, $\mathrm{m_{jj}}$, and $m_t^{\mathrm{reco}(2)}$ for 970 Lepton+Jets candidate events with at least 1 b-tag and the values of $\mathrm{m}_t^{\mathrm{NWA}}$ and m_{T2} for 344 dilepton candidate events are compared to three-dimensional and two-dimensional, respectively, probability density function derived by applying kernel density estimation to fully simulated MC events with different values of the top quark mass and Δ_{JES} in the detector. We measure $\mathrm{M_{top}} = 171.9 \pm 1.1$ (stat.) $\pm\, 0.9$ (syst.) GeV/c^2.

Preliminary Results of TMT using 4.8 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 (lν) or to a pair of quarks (qq'). We look for events consistent with $t\bar{t}$ production and decay involving at least one lν pair (we do not consider events with taus). Events are split into Lepton+Jets $t\bar{t}$ candidates, in which one W boson decays hadronically and the other W boson decays leptonically, and Dilepton events, in which both W bosons decay leptonically. Lepton+Jets and Dilepton events are combined into a single likelihood, allowing for a more robust combination of results from the two channels into one measurement of $\mathrm{M_{top}}$. By using a single likelihood for both measurements, all correlations in systematics between the two channels are fully taken into account. 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. In the Lepton+Jets channel, we use two variables (m_t^{reco}, $m_t^{\mathrm{reco(2)}}$) that is strongly correlated with the true top quark mass ($\mathrm{M_{top}}$), and the other variable (m_{jj}) that is sensitive to shifts in the jet energy scale (Δ_{JES}) in the detector. The value of m_t^{reco} in each event is derived from a χ^2 minimization that uses knowledge of the overconstrained kinematics of the $t\bar{t}$ system [2–4]. Because m_t^{reco} do not bring 100 % of the $\mathrm{M_{top}}$ information, we include another reconstructed top quark mass from 2nd best combinatoric from kinematic fit by choosing the 2nd smallest χ^2 combination. The dijet mass (m_{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. In the Dilepton channel, we use the Neutrino Weighting Algorithm [2, 5, 6] to form at estimator for $\mathrm{M_{top}}$ by integrating over the unknown neutrino pseudoraptidies. In addition, m_{T2} variable [7–9], which is a quantity of transverse mass in two missing particle system, is used as an additional estimator for $\mathrm{M_{top}}$. In the dilepton channel, we do not have an observable constraint jet energy scale, however, in the combined measurement the *in situ* calibration from the Lepton+Jets channel keeps the JES systematic under control.

Monte Carlo samples generated with 76 different $\mathrm{M_{top}}$ are run through a full CDF detector simulation assuming 29 possible shifts in Δ_{JES}. The values of observables in data are compared to each point in the MC grid using a non-parametric approach based on Kernel Density Estimation (KDE) [10]. Local Polynomial Smoothing [11] is used to smooth out these points and calculate the probability densities at any arbitrary value of $\mathrm{M_{top}}$ and Δ_{JES}. An unbinned likelihood fit is used to measure $\mathrm{M_{top}}$ and profile out Δ_{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). In addition, large $\not\!\!E_T$ + two jets requirement is used to increase more muon acceptance. Offline, the events are required to have a single energetic lepton, large $\not\!\!E_T$ due to the escaping neutrino from the leptonic 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. Dilepton events used a high-E_T (or high-pt) lepton trigger only, but offline require two oppositely charged lepton candidates, only two energetic jets, and $\not\!\!E_T$. The value of H_T is required to be greater than 200 GeV. Events consistent with cosmic ray muons, photon conversions or Z boson decays are rejected.

Jets are reconstructed with the JETCLU [12] cone algorithm using a cone radius of $R \equiv \sqrt{\eta^2 + \phi^2} = 0.4$. To improve the statistical power of the method, both the Lepton+Jets and Dileptons samples are each divided into two subsamples, depending on the number of jets identified as arising from the hadronization and decay of b quarks. The SECVTX [13] 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 for Lepton+Jets events. In Lepton+Jets events with exactly one tag, we require exactly four jets with $E_T > 20$ GeV/c^2. For Lepton+Jets 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. Dilepton events are divided into untagged and tagged samples, which have very different signal-to-background ratios.

For Lepton+Jets events, we make a cut on the χ^2 out of the kinematic fitter described in Section IV, requiring it to

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

	1-tag	2-tag
b-tags	==1	> 1
Leading 3 jets E_T(GeV/c^2)	>20	>20
MET (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)	$100 < m_t^{\rm reco} < 350$	$100 < m_t^{\rm reco} < 350$
$m_{\rm jj}$ boundary cut (GeV/c^2)	$50 < m_{\rm jj} < 120$	$50 < m_{\rm jj} < 125$
$m_t^{\rm reco(2)}$ boundary cut (GeV/c^2)	$100 < m_t^{\rm reco(2)} < 350$	$100 < m_t^{\rm reco(2)} < 350$
Observed # events	699	271
Expected background	146.8±56.2	19.4±7.9

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

	0-tag	Tagged
b-tags	==0	> 0
Leading 2 jets E_T(GeV/c^2)	>15	>15
MET (GeV/c^2)	>25	>25
$m_t^{\rm NWA}$ boundary cut (GeV/c^2)	$100 < m_t^{\rm NWA} < 350$	$100 < m_t^{\rm NWA} < 350$
$m_{\rm T2}$ boundary cut (GeV)	$30 < m_{\rm T2} < 200$	$30 < m_{\rm T2} < 200$
Observed # events	215	129
Expected background	105.6±15.5	9.5±2.1

be less than 9.0. For Dilepton events, we require successful reconstruction by the NWA algorithm. Finally, in order to properly normalize our probability density functions, we define hard boundaries on the values of the observables. Events with values of an observable falling outside the boundary are rejected. Event selection is summarized in Tables I and II.

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 [14]. 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.
- 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.

IV. LEPTON+JETS TOP MASS RECONSTRUCTION

The value of the reconstructed mass in each Lepton+Jets event ($m_t^{\rm reco}$) is determined by minimizing a χ^2 describing the overconstrained kinematics of the $t\bar{t}$ system. The reconstructed mass 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

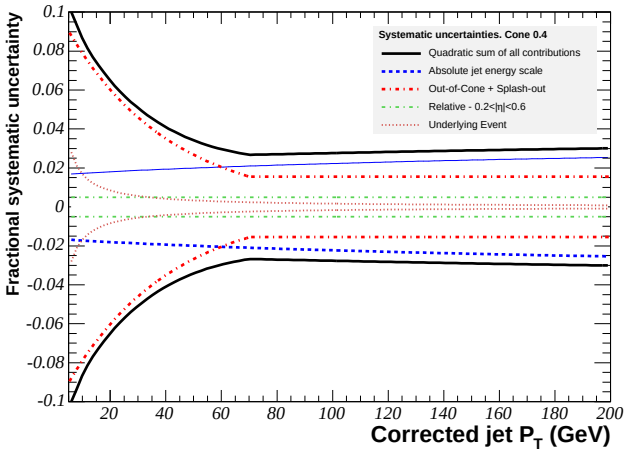

FIG. 1: Jet energy scale uncertainty as a function of the corrected jet p_T for the underlying event (dotted red), relative response (dashed green), out-of-cone energy (dashed red) and absolute response (dashed blue). The contribution of all sources are added in quadrature (full black) to form the total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

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,

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

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 remain near 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.

Although $\rm m_t^{reco}$ carry a lot of information about $\rm M_{top}$, it is still possible to be carried $\rm M_{top}$ information by another combination of jet-parton assignment. Therefore, we add another $\rm M_{top}$ from the 2nd smallest χ^2 combination which is $m_t^{\rm reco(2)}$. We use this value as 3rd observable in the 3d KDE machinery.

V. DIJET MASS

The value of $m_{\rm jj}$ in each Lepton+Jets 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.

VI. NEUTRINO WEIGHTING ALGORITHM

In the dilepton decay channel of $t\bar{t}$ events, there is not enough information to reconstruct the masses of the top quarks. The 4-momenta of jets and leptons and an overall imbalance in transverse energy are measured, but there is no way to disentangle the 4-vectors of the two escaping neutrinos. To form an estimator for $\rm M_{top}$, we need to integrate

over unknown quantities, taking the probability density functions from MC. In the NWA algorithm, we integrate over the pseudorapidities η_1 and η_2 of the two neutrinos. As inputs, we use jets corrected to reflect partons energies, the charged lepton momenta amd the $\not{E}_T$. The approach is as follows:

- Assume a value of the top quark mass.
- Choose a particular jet-to-b quark assignment (there are two possibilities).
- Assume values for η_1 and η_2.
- Using the world average masses of the W boson, b quark and charged leptons, solve for the p_x and p_y of each of the neutrinos. Solutions may not exist, but when they do exist, this gives two solutions for each neutrino.
- Compare each combination of neutrino solutions to the meausred $\not{E}_T$, assigning Gaussian weights. Since the correct combination is not known, sum the four weights.
- Integrate over η_1 and η_2, obtaining a total weight for the assumed top mass. The integration distribution for the neutrino pseudorapidities is taken from $t\bar{t}$ MC, and is a Gaussian with width approximately 1.0. The integration is performed by summing a grid of η values with 0.2 spacing.
- Obtain the weights corresponding to the other jet-to-b quark assignment.
- Sum the weights from the two jet-to-quark assignments, giving a handle on the probability that the true top quark mass is the assumed mass.
- Scan the top mass in units of 3 GeV.
- The point of maximum weight is found, and the scan is repeated with decreasing step size until it converages.
- The assumed top quark mass that yields the highest total weight is taken as the value of $\mathrm{m}_t^{\mathrm{NWA}}$.

VII. $m_{\mathbf{T2}}$

We can calculate m_{T2} variables in the $t\bar{t}$ dilepton events followed by Ref. [8, 9]. In the dilepton channel, we can consider the transverse mass of top decay leptonically $t \to bl\nu$

$$m_T^2 = m_{bl}^2 + m_\nu^2 + 2(E_T^{bl} E_T^\nu - \mathbf{p}_T^{bl} \cdot \mathbf{p}_T^\nu) \tag{VII.1}$$

where m_{bl} and $\mathbf{p}_T^{bl}$ denote the invariant mass and transverse momentum of the bl system, respectively. m_ν and $\mathbf{p}_T^\nu$ are the mass and transverse momentume of neutrino. The transerve energies of the bl system and neutrino are defined as

$$E_T^{bl} = \sqrt{|\mathbf{p}_T^{bl}|^2 + m_{bl}^2} \;\; and \;\; E_T^\nu = \sqrt{|\mathbf{p}_T^\nu|^2 + m_\nu^2} \tag{VII.2}$$

In the dilepton channel each top quark can define the m_T, which can be denoted by $m_T^{(1)}$ and $m_T^{(2)}$. The m_{T2} variable of each event is defined as

$$m_{T2} = min\left[max\{m_T^{(1)}, m_T^{(2)}\}\right] \tag{VII.3}$$

where $\mathrm{m}_T^{(i)}$ (i = 1,2) is the transverse mass of $t^i \to b^i l^i \nu^i$. And the minimization is performed over the trial neutrino momenta $\mathbf{p}_T^{\nu(i)}$ constrainted as

$$\mathbf{p}_T^{\nu(1)} + \mathbf{p}_T^{\nu(2)} = \mathbf{p}_T^{miss}. \tag{VII.4}$$

TABLE III: Expected number of Lepton+Jets background events after event selection, χ^2 and boundary cuts.

	CDF II Preliminary 4.8 fb^{-1}	
	1-tag	2-tag
W$b\bar{b}$	36.1 ± 11.6	8.8 ± 2.9
W$c\bar{c}$	22.8 ± 7.6	1.4 ± 0.5
Wc	11.4 ± 3.6	0.7 ± 0.2
W+light mistag	26.8 ± 5.9	0.7 ± 0.2
Z+jets	5.7 ± 1.3	0.6 ± 0.2
single top	5.7 ± 0.5	2.2 ± 0.3
Diboson	7.7 ± 1.0	1.0 ± 0.2
QCD	30.7 ± 24.9	4.0 ± 3.5
Total	146.8 ± 56.2	19.4 ± 7.9
$t\bar{t}$ (172.5 GeV/c^2, 7.4 pb)	557.0 ± 69.5	272.0 ± 42.0

VIII. BACKGROUNDS

An *a priori* estimate for the Lepton+Jets background composition is used to derive background shapes for $m_t^{\rm reco}$, $m_t^{\rm reco(2)}$, and $m_{\rm jj}$. ALPGEN [15] combined with PYTHIA [16] is used to model W+jets. Contributions include W$b\bar{b}$, W$c\bar{c}$, Wc and W+light flavor (LF) jets. Non-isolated leptons are used to model the QCD background. The relative fractions of the different W+jets samples are determined in MC, but the absolute normalization is derived from the data. The MC are combined using their relative cross sections and acceptances, and we remove events overlapping in phase space and flavor across different samples. MC and theoretical cross-sections are used to model the single-top and diboson backgrounds. The expected number of background from different sources is shown in Table III. The backgrounds are assumed to have no $\rm M_{top}$ dependence, but all MC-based backgrounds are allowed to have $\Delta_{\rm JES}$ dependence.

The main sources of background in the dilepton channel are fake events where a jet is misidentified as a lepton, Drell-Yan events, and diboson production. To model the fakes background, we select events in the data with one charged lepton and an object likely to be a jet faking a lepton. All other selection requirements are kept the same as for Dilepton events. Events are weighted by the probability of such an events being a fake events. The probability is calculated using QCD-enriched samples collected using a jet trigger. The Drell-Yan model comes from a set of matched Alpgen+Pythia samples. Included are contributions from Z→ $ee, \mu\mu, \tau\tau$+0,1,2,3, and ≥4 partons, as well as Z→ $ee, \mu\mu, \tau\tau, +b\bar{b}, c\bar{c}$+0,1, and ≥ 2 partons. Both off Z-peak and on Z-peak samples are used. We remove b and c quarks appearing in Pythia showering from light flavor and Z→ $ee, \mu\mu, \tau\tau + c\bar{c}$ samples. Dibosons are modeled using Pythia. The expected number of background from different sources is shown in Table IV.

	CDF II Preliminary 4.8 fb^{-1}	
	0-tag	tagged
WW	14.6 ± 2.7	0.5 ± 0.1
WZ	3.5 ± 0.1	0.1 ± 0.0
ZZ	2.2 ± 2.0	0.2 ± 0.1
Wγ	0.4 ± 0.4	0.0 ± 0.0
Drell Yan ($\tau\tau$)	11.0 ± 2.3	0.6 ± 0.1
Drell Yan (ee or $\mu\mu$)	28.7 ± 4.9	1.6 ±0.3
Fakes	45.3 ± 14.2	6.6 ±2.1
Total	105.6 ± 15.5	9.5 ± 2.1
$t\bar{t}$ (172.5 GeV/c^2, 7.4 pb)	108.4 ± 13.9	134.1 ± 17.4

TABLE IV: Backgrounds for the DIL tagged and non-tagged subsamples.

IX. KERNEL DENSITY ESTIMATES

Probability density functions for $m_t^{\rm reco}$-$m_{\rm jj}$-$m_t^{\rm reco(2)}$ and $\rm m_t^{NWA}$-$m_{\rm T2}$ at every point in the $\rm M_{top} - \Delta_{JES}$ grid and for backgrounds are derived using a Kernel Density Estimate (KDE) approach [10]. 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 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:

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

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:

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

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 three (two) dimensions by multiplying the three (two) kernels together for lepton+jets (dilepton) channel:

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

Because of a difficulty to draw 3d density estimates as plot, we make 2d projection of three combination of observables in Fig.s 2 for lepton+jets channel. Figure 3 show the 2d density estimates for dilepton signal events. Figures 4 and 5 show the 2d projection (estimates) of background events for lepton+jets (dilepton) channel. To estimate the background density, we combined each background sample with the appropriate weights taking into account the sample size and cross section, and then had KDE estimation.

X. LIKELIHOOD FIT

The values of m_t^{reco}-m_{jj}-$m_t^{\text{reco}(2)}$ and m_t^{NWA}-m_{T2} observed in data are compared to points in $\text{M}_{\text{top}} - \Delta_{\text{JES}}$ space. An extended maximum likelihood fit is performed to maximize the likelihood with respect to the expected number of signal (n_s) and background events (n_b) in each of the four subsamples. A Gaussian constraint on the expected number of background events is applied to each of the subsamples. The likelihood for subsample k with N events is given by:

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

where m_i is the value of m_t^{reco} (m_t^{NWA}) for the ith Lepton+Jets (Dilepton) event, and y_i is the value of m_{jj} (m_{T2}) for the ith Lepton+Jets (Dilepton) event. (z_i) is only defined for lepton+jets channel with the value of $m_t^{\text{reco}(2)}$. The overall likelihood is a product over the four individual subsample likelihoods, with a Gaussian constraint on Δ_{JES}, constraining it to the nominal $0 \pm 1\ \sigma_c$:

$$\mathcal{L} = \exp\left(-\frac{\Delta_{\text{JES}}^2}{2{\sigma_c}^2}\right) \times \mathcal{L}_{\text{1-tag,LJ}} \times \mathcal{L}_{\text{2-tag,LJ}} \times \mathcal{L}_{\text{0-tag,DIL}} \times \mathcal{L}_{\text{tagged,DIL}}. \tag{X.2}$$

The above gives values of $-\ln\mathcal{L}$ only for points in the $\mathrm{M_{top}} - \Delta_{\mathrm{JES}}$ grid, and not as a continuous function. To obtain density estimates for an arbitrary point in the $\mathrm{M_{top}} - \Delta_{\mathrm{JES}}$ grid, we use local polynomial smoothing [11] 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 $\mathrm{M_{top}} - \Delta_{\mathrm{JES}}$ space, where the values of $\mathrm{M_{top}}$ and Δ_{JES} far away from the point being estimated are deweighted. This allows for a smooth likelihood that can be minimized. The measured uncertainty on $\mathrm{M_{top}}$ comes from the largest possible shift in $\mathrm{M_{top}}$ on the $\Delta\ln\mathcal{L} = 0.5$ contour.

Our primary measurement is a combined fit using both Lepton+Jets and Dilepton data, but we also run fits and evaluate systematics for a Lepton+Jets-only measurement and a Dilepton-only measurement. The Dilepton-only measurement does not include a dijet resonance that can measure Δ_{JES}, so its likelihood is evaluated in 1d as a function of $\mathrm{M_{top}}$, with the unmeasured Δ_{JES} taken as a full systematic.

XI. METHOD CHECK

We test our machinery by running pseudoexperiments with varying values of $\mathrm{M_{top}}$ between 160 and 182.5 GeV/c^2 and varying values of Δ_{JES} between -1.0 and 1.0 σ_c. Figure 6 shows the $\mathrm{M_{top}}$ residuals as a function of true top quark mass. The dilepton channel is unbiased while the combined and lepton+jets measurement are biased. However, it can be well fitted by linear function we then correct bias to be matched with zero. This correction affect pull widths to be off from unity, so we need to inflate our measured statistical errors by 4.1% and 4.8% for lepton+jets and combined fit respectively. Figure 7 shows $\mathrm{M_{top}}$ residuals after correction for lepton+jets and combined measurement. Pull width with errrors inflation are shown in Figure 8. The dilepton channel is unbiased in its prediction for both residuals and pull widths so, these figures are not corrected for dilepton channel. For the all of measurement including systematic study, we apply these corrections.

XII. RESULTS

The likelihood procedure when applied to the data yields $\mathrm{M_{top}} = 171.9 \pm 1.1$ GeV/c^2. The Lepton+Jets-only fit yields $\mathrm{M_{top}} = 172.0 \pm 1.2$ GeV/c^2. The Dilepton-only fit yields $\mathrm{M_{top}} = 170.6 \pm 2.2$ GeV/c^2. The log-likelihood contours for our combined measurement and Lepton+Jets-only measurements are shown in Figures 9 left and middle. The 1d Δlog-likelihood for the Dilepton-only measurement is shown in Figure 9 right. As shown in Figure 10, 39 %

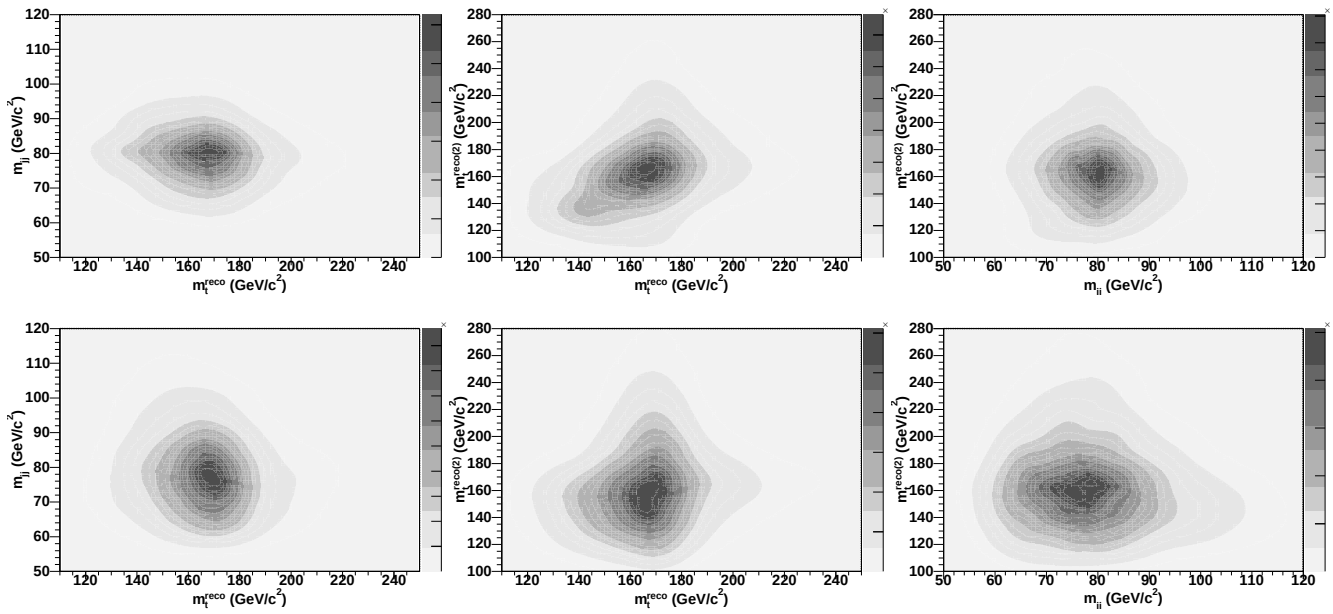

FIG. 2: 2d projection of 3d density estimates for input mass of 172 GeV/c^2 and $\Delta_{\mathrm{JES}} = 0.0$ for Lepton+Jets 1-tag events (top) and 2-tag events (bottom).

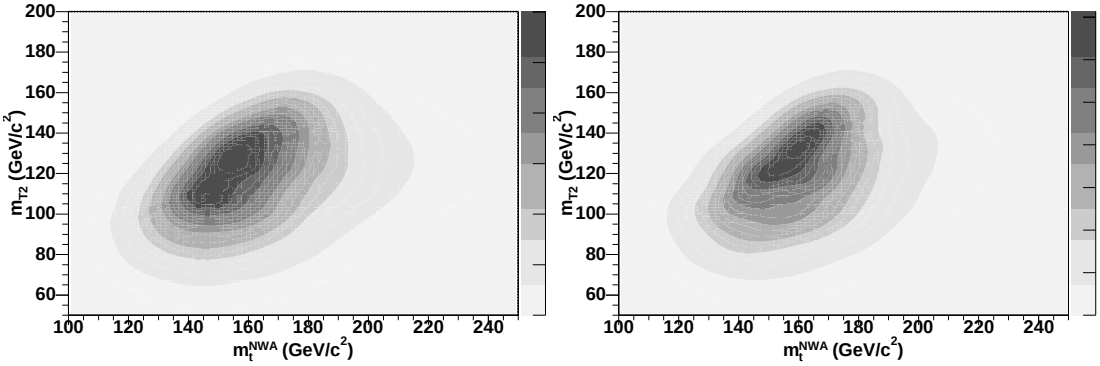

FIG. 3: 2d density estimates for input mass of 172 GeV/c^2 and $\Delta_{\rm JES} = 0.0$ for dilepton untagged events (left) and tagged events (right).

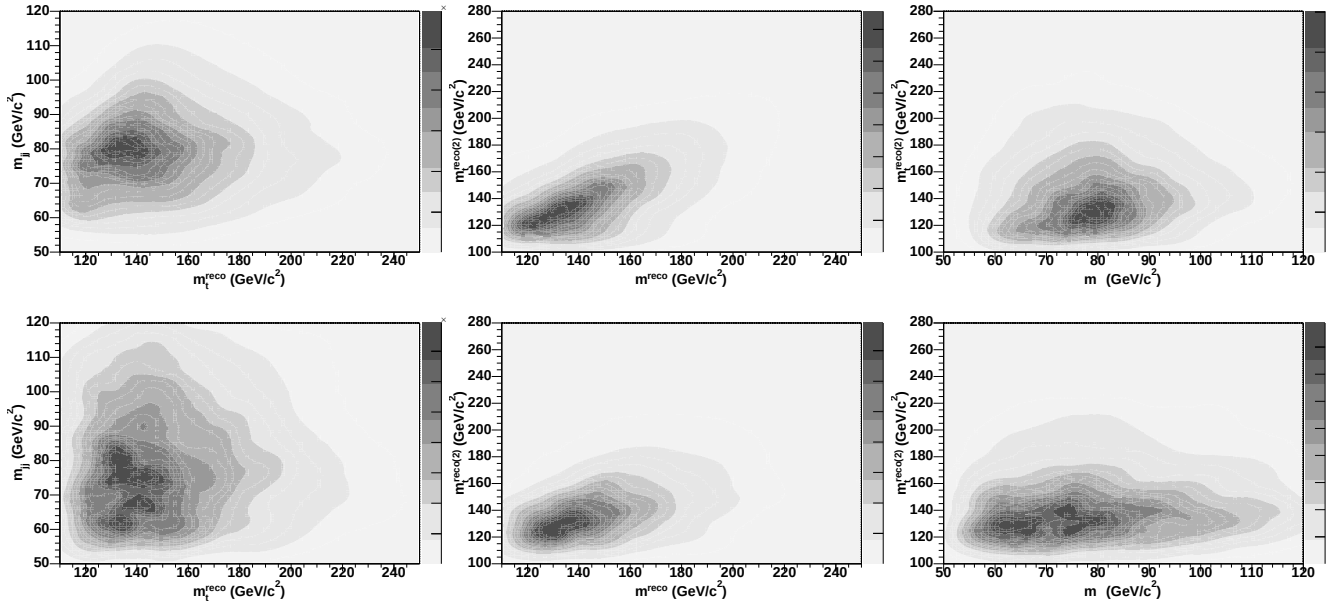

FIG. 4: 2d projection of 3d density estimates for the combined background for Lepton+Jets 1-tag events (top) and 2-tag events (bottom).

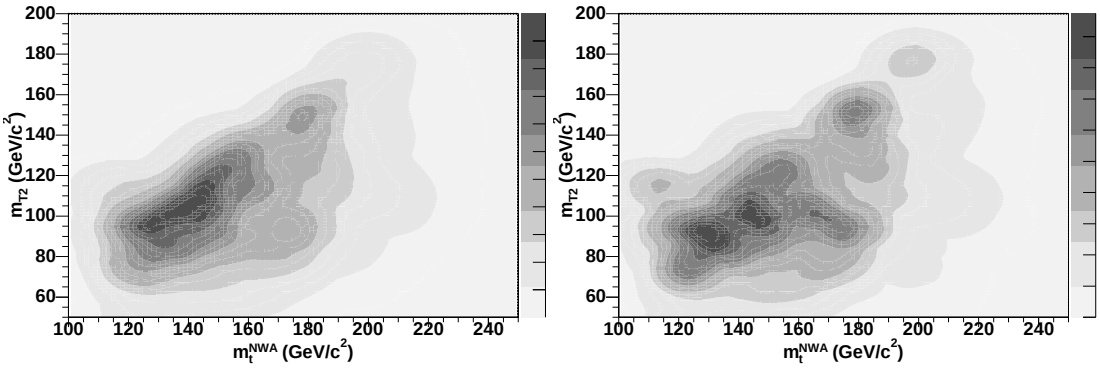

FIG. 5: 2d density estimates for the combined background for Dilepton untagged events (left) and tagged events (right).

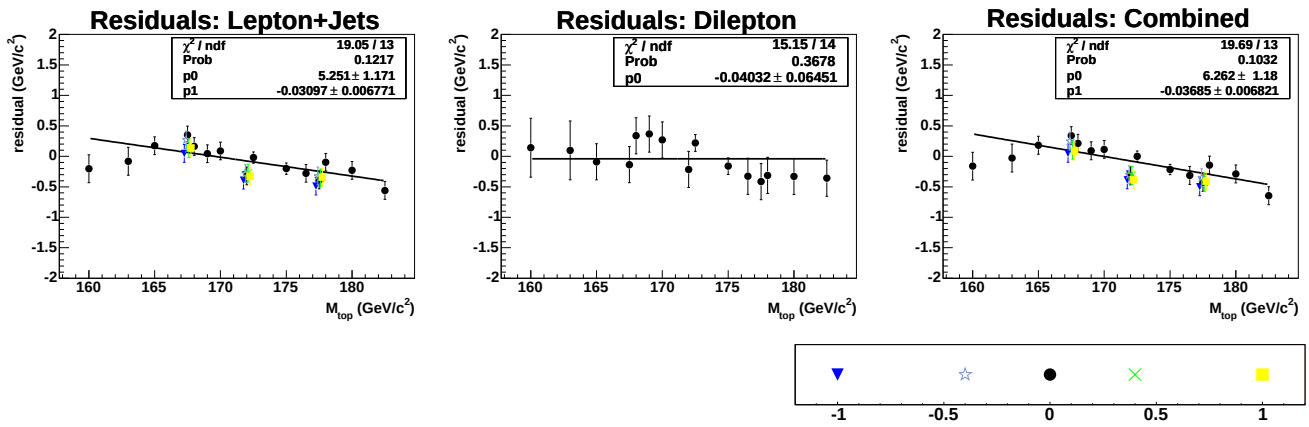

FIG. 6: Residual mass shift as a function of input mass from pseudoexperiments for Lepton+Jets-only pseudoexperiments (left), dilepton-only pseudoexperiments (middle) and fits using combined machinery (right) before corrections.

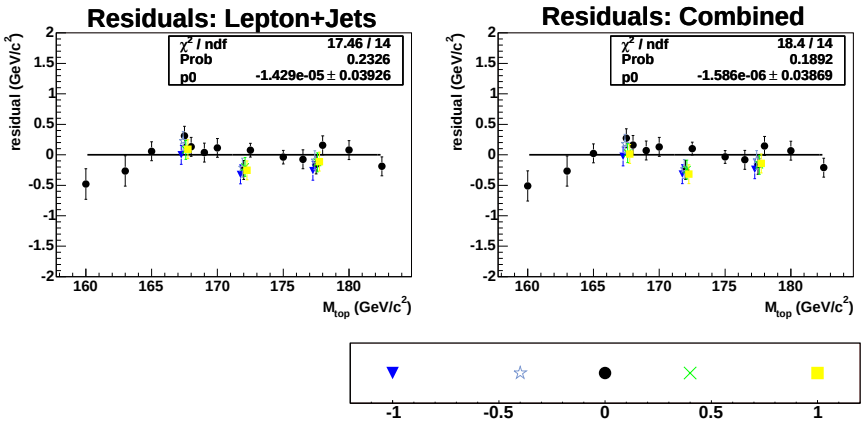

FIG. 7: Residual mass shift as a function of input mass from pseudoexperiments for Lepton+Jets-only pseudoexperiments (left) and fits using combined machinery (right) after corrections.

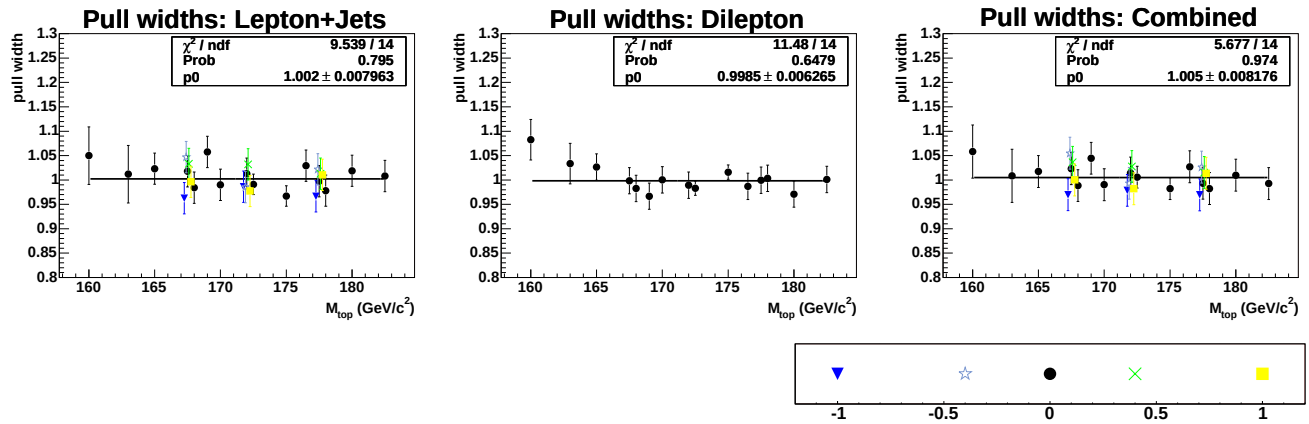

FIG. 8: Pull widths as a function of input mass from pseudoexperiments for Lepton+Jets-only pseudoexperiments (top), dilepton-only pseudoexperiments (middle) and fits using combined machinery (bottom). Lepton+jets and combined fit have inflation of errrors by 4.1% and 4.8% respectively.

of pseudoexperiments have a smaller error than the value measured in the combined fit in data. The p-values for the Lepton+Jets-only fit is 52%; the value for the Dilepton-only fits 15%.

The combined fit returns $\Delta_{\rm JES} = 0.05 \pm 0.26\ \sigma_{\rm c}$, and the Lepton+Jets fit returns $\Delta_{\rm JES} = 0.09 \pm 0.27\ \sigma_{\rm c}$.

We run a combined fit separately without the JES and background constraints and measure $\rm M_{top} = 171.9 \pm 1.1$ GeV/c^2 and $\rm M_{top} = 171.9 \pm 1.2$ GeV/c^2, showing that these priors do not significantly affect our result.

XIII. 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 $\rm M_{top}$ from pseudoexperiments, giving a residual JES uncertainty. For the dilepton-only measurement, which has no *in situ* calibration, these systematics dominate. 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 the p_T of Drell-Yan events to the $t\bar{t}$ mass region, resulting in a radiation systematics. Comparing pseudoexperiments generated with HERWIG [17] 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%. Systematics due to lepton energy scales are estimated by propagating 1% shifts on electron and muon energies scales. Background composition systematics are obtained by varying the fraction of the different types of backgrounds in pseudoexperiments. For Lepton+Jets backgrounds, varying the uncertain Q^2 of background events results in a background shape systematic, and using a different model for QCD events gives an additional QCD modeling systematic. For Dilepton backgrounds, varying the composition of the Drell-Yan sample between low and high jet multiplicities gives one systematic effect. We also shift the fake model in ways expected to maximally correlate with the reconstructed mass. It has been suggested that color reconnection effects could cause a bias in the top quark mass measurement [18]. We test this effect by generating MCs with and without CR and take the difference as systematics.

The total systematic error is 0.9 GeV/c^2 for both the combined and the Lepton+Jets measurement, and 3.1 GeV/c^2 for the Dilepton-only measurement. The systematics are summarized in Table V.

XIV. CONCLUSIONS

We present a simultaneous measurement of the mass of the top quark in the Lepton+Jets and Dileptons channels using a template-based technique with an *in situ* JES calibration. Using 3d and 2d templates derived from Kernel Density Estimation and 4.8 fb^{-1} of data collected by the Tevatron, we measure

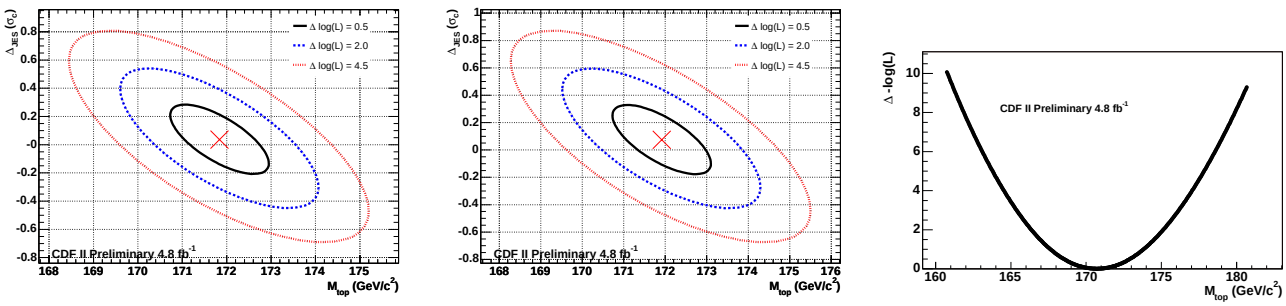

FIG. 9: Negative log-likelihood contours for the combined fit (left) and lepton+jets only fit (middle) are shown. Also negative log-likelihood for the Dilepton only fit (right) is shown.

TABLE V: Summary of systematics. All numbers have units of GeV/c^2.

	CDF II Preliminary 4.8 fb^{-1}		
Systematic	LJ	DIL	Combination
Residual JES	0.6	2.9	0.6
Generator:	0.6	0.6	0.6
PDFs	0.1	0.3	0.1
b jet energy	0.3	0.3	0.3
Background shape	0.1	0.3	0.1
gg fraction	0.1	0.3	<0.1
Radiation	0.1	0.3	0.1
MC statistics	0.1	0.3	0.1
Lepton energy	<0.1	0.3	<0.1
MHI	0.1	0.2	0.1
Color Reconnection	0.2	0.6	0.2
Total systematic	0.9	3.1	0.9

$$\mathrm{M_{top}} = 171.9 \pm 1.1 \text{ (stat.)} \pm 0.9 \text{ (syst.) GeV}/c^2 = 171.9 \pm 1.5 \text{ GeV}/c^2 \text{ (combined)}$$

$$\mathrm{M_{top}} = 172.0 \pm 1.2 \text{ (stat.)} \pm 0.9 \text{ (syst.) GeV}/c^2 = 172.0 \pm 1.5 \text{ GeV}/c^2 \text{ (Lepton+Jets-only)}$$

$$\mathrm{M_{top}} = 170.6 \pm 2.2 \text{ (stat.)} \pm 3.1 \text{ (syst.) GeV}/c^2 = 170.6 \pm 3.8 \text{ GeV}/c^2 \text{ (Dilepton-only)}$$

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 für Bildung und Forschung, Germany; the Korean Science and Engineering Foundation and the Korean Research Foundation; the Science and Technology Facilities Council and the Royal Society, UK; the Institut National de Physique Nucleaire et Physique des Particules/CNRS; the Russian Foundation for Basic Research; the Ministerio de Ciencia e Innovación,

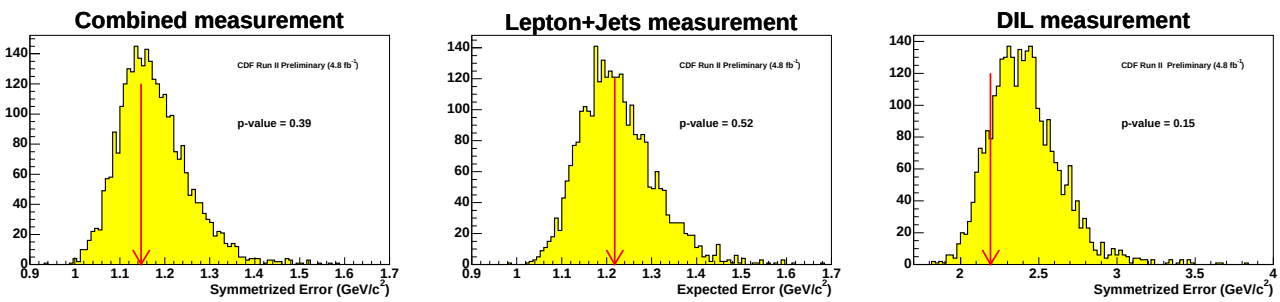

FIG. 10: Reported error from pseudoexperiments with the observed number of events for combined fits (top), Lepton+Jets fit (middle) and Dilepton fit (bottom).

and Programa Consolider-Ingenio 2010, Spain; the Slovak R&D Agency; and the Academy of Finland.

---

[1] F. Abe, et al., Nucl. Instrum. Methods Phys. Res. A **271**, 387 (1988);
[2] T. Aaltonen, el al., Phys. Rev. D **79**, 092005 (2009);
[3] A. Abulenci et al., Phys. Rev. D 73, 032003 (2006);
[4] A. Abulenci et al., Phys. Rev. Lett. 96, 022004 (2006);
[5] B. Abbott, et al., Phys. Rev. D 60, 052001 (1999);
[6] A. Abulenci, et al., Phys. Rev. D 73, 112006 (2006);
[7] C. Lester and D. Summers, Phys. Lett. B **463**, 99 (1999); A. Barr, C. Lester, and P. Stephens, J. Phys. G **29**, 2343 (2003);
[8] W. S. Cho, et al., Phys. Rev. D **78**, 034019 (2008);
[9] T. Aaltonen, et al., arXiv:0911.2956 (submitted to Phys. Rev. D);
[10] K. S. Cranmer, Comput. Phys. Commun. **136**, 198 (2001);
[11] C. Loader, *Local Regression and Likelihood*, New York, USA, (Springer, 1999);
[12] F. Abe, et al., Phys. Rev. D 45, 1448 (1992);
[13] T. Affolder, et al., Phys. Rev. D 64, 032002 (2001);
[14] A. Bhatti, et al., Nucl. Instrum. Methods Phys. Res. A **566**, 375 (2006);
[15] M. L. Mangano, et al., J. High Energy Phys. **0307**, 001 (2003);
[16] T. Sjostrand, S. Mrenna, and P. Skands, J. High Energy Phys. **0605**, 026 (2006);
[17] G. Corcella, et al., J. High Energy Phys. **0101**, 010 (2001);
[18] P. Z. Skands, D. Wicke, Eur. Phys. J. C **52**, 133 (2007);