Top-quark mass measurement in the dilepton channel using *in situ* jet energy scale calibration

Hyun Su Lee

Department of Physics, Ewha Womans University, Seoul 120-750, Korea
(Received 6 June 2012; published 11 September 2012)

We employ a top-quark mass measurement technique in the dilepton channel with *in situ* jet energy scale calibration. Three variables having different jet energy scale dependences are used simultaneously to extract not only the top-quark mass but also the energy scale of the jet from a single likelihood fit. Monte Carlo studies with events corresponding to an integrated luminosity of 5 fb^{-1} proton-proton collisions at the Large Hadron Collider $\sqrt{s} = 7$ TeV are performed. Our analysis suggests that the overall jet energy scale uncertainty can be significantly reduced and the top-quark mass can be determined with a precision of less than 1 GeV/c^2, including jet energy scale uncertainty, at the Large Hadron Collider.

In the standard model, the top quark (t) is the heaviest known elementary particle [1]. The heavy mass significantly affects the electroweak radiative correction that relates the top-quark mass and the W boson mass to the Higgs boson mass [2]. In addition, the heavy mass may have implications for new physics theories, including the minimal supersymmetric standard model and technicolor-like models. Because of its importance, top-quark mass (M_{top}) measurements have been performed using various methods in different decay channels. The precision of the M_{top} measurement already surpasses 0.5% at the Tevatron [3].

It is important to measure M_{top} using different techniques and independent data samples in different decay channels. Significant differences in the measurements of M_{top} using different decay channels may indicate contributions from new physics beyond the standard model [4]. The dilepton decay channel is particularly interesting because the signature of this channel can be mimicked with supersymmetric partner stop pairs [5] as well as charged Higgs boson signals [6]. However, the M_{top} precisions of the dilepton channel measurements [7–9] have been limited because the branching fraction is much smaller and systematic uncertainty from the jet energy scale (JES) is much larger than those of the lepton + jets channel measurements [10–13]. At the Tevatron, a single measurement already achieved 1.2 GeV/c^2 precision in the lepton + jets channel [10]. However, the total uncertainty in the dilepton channel in a standalone measurement was 3.0 GeV/c^2 [8], with the uncertainty from JES being dominant. In the Large Hadron Collider (LHC), the cross section of $t\bar{t}$ pair production is approximately 20 times larger than that in the Tevatron. Therefore, the small branching fraction corresponding to data obtained in LHC experiments up to 2011 (more than 5 fb^{-1} integrated luminosity) is not an issue. As an example, the recent M_{top} measurement in the dilepton channel performed by the CMS Collaboration had 1.2 GeV/c^2 statistical uncertainty using 2.2 fb^{-1} data [9]. However, this measurement returned a larger systematic uncertainty, ($^{+2.5}_{-2.6}$ GeV/c^2), which is dominated by the

overall JES uncertainty (2.0 GeV/c^2). The JES systematic uncertainty cannot generally be reduced by using a larger data sample.

A similar issue was encountered in the lepton + jets channel. However, *in situ* JES calibration using hadronic decaying W bosons [14] resolved this issue. In this method, the overall JES uncertainty was absorbed in the statistical uncertainty; therefore, a larger data sample could reduce not only the pure statistical uncertainty but also the overall JES systematic uncertainty. However, in the dilepton channel, no single variable can be used to calibrate the JES uncertainty *in situ* for the M_{top} measurement. A recent measurement performed by the D0 Collaboration employed a new technique for JES calibration in this channel [15]. They used the result of the JES measurement in the lepton + jets channel. In this way, they significantly reduced overall JES systematic uncertainty in the dilepton channel. However, it is not a standalone measurement and the use of the lepton + jets channel result introduced an additional systematic uncertainty.

In this paper, we propose a novel technique for the M_{top} measurement in the dilepton channel, in which we perform the *in situ* JES calibration using three variables. The selected variables, which have already been used for the M_{top} measurements, have different dependences on M_{top} and JES, so we can perform simultaneous measurement of both M_{top} and JES. This is the first time the M_{top} measurement has been performed using *in situ* JES calibration in the dilepton channel even though it is a Monte Carlo (MC) simulated experiment. For a realistic estimation of the precision, we consider the environment of the LHC experiment with 5 fb^{-1} pp collisions.

There are different definitions of quark masses in the theoretical framework [16]. M_{top} of the $\overline{\text{MS}}$ renormalization scheme, $M_{\overline{\text{MS}}}$, differs from the pole mass, M_{ploe}, by about 10 GeV/c^2 [16,17]. The direct M_{top} measurements are all calibrated with MC simulations; the measured quantities therefore correspond to the M_{top} scheme used in the MC simulations, M_{MC}. Even though it is usually

assumed that $M_{\rm MC}$ is the same as $M_{\rm pole}$ [3], there are a number of theoretical questions in relating $M_{\rm pole}$ to $M_{\rm MC}$ [16,18]. An accurate relation is still under theoretical investigation [19]. As in all other direct $M_{\rm top}$ measurements, we measure $M_{\rm MC}$ in this paper.

We generate simulated $t\bar{t}$ samples using the leading order (LO) MC generator MADGRAPH/MADEVENT package [20]. In the MADGRAPH/MADEVENT generation, we vary the parameter $M_{\rm top}$ between 160 and 190 GeV/c^2 in steps of 2 GeV/c^2. A total of 500 000 $t\bar{t}$ pair events in the dilepton final state are produced for each sample. Showering and hadronization are performed by PYTHIA [21]. To take into account the detector effect and perform event reconstruction under realistic conditions, we use the fast detector simulation package DELPHES [22]. The conditions used for the simulated detector are the ones generally associated with the LHC detector. We assume the coverage of the tracker to be within $|\eta| = 2.5$ with 100% efficiency and that of the calorimeter to be within $|\eta| = 3$ with tower segment $\Delta\eta \sim 0.1$ and $\Delta\phi \sim 0.1$. The resolution of the electromagnetic calorimeter and hadronic calorimeter are parameterized by

$$\frac{\sigma_{EM}}{E} = 0.005 + \frac{0.25}{E} + \frac{0.05}{\sqrt{E}}, \qquad \frac{\sigma_{HA}}{E} = 0.05 + \frac{1.5}{\sqrt{E}}.$$

All physics objects such as leptons, jets, and missing transverse energy are reconstructed in the fast simulation. Jets originating from b quarks are tagged using a reliable b-tagging algorithm with an efficiency of approximately 40%.

In the dilepton decay channel, the production of a $t\bar{t}$ pair is followed by the decay of each top quark to a W boson and a b quark, where both W bosons decay to charged leptons (electron or muon) and neutrinos ($t\bar{t} \rightarrow ll'\nu\nu'b\bar{b}$). Events in this channel thus contain two leptons, two b quark jets, and two undetected neutrinos. To select the candidate events of $t\bar{t}$ dilepton topology, we require two oppositely charged lepton candidates with $p_T > 20$ GeV/c. We also require a missing transverse energy exceeding 25 GeV and at least two tagged jets with $E_T > 30$ GeV. The expected signal and background events are 5961 ± 545 and 320 ± 97, respectively, taken from Ref. [9] with scaling according to the respective integrated luminosity for the 5 fb^{-1} data. The expected background contribution is approximately 5%. Because of its small contribution, we do not include background events in our further analysis, for simplicity.

We use three reconstructed variables as observables of $M_{\rm top}$ as well as the nuisance parameter $\Delta_{\rm JES}$. $\Delta_{\rm JES}$ is the relative energy scale of the jet with respect to the nominal JES as used in the lepton + jets channel measurements [10–13]. All three variables have already been used in the $M_{\rm top}$ measurements.

(1) $m_t^{\rm NW}$—The reconstructed top-quark mass obtained using the neutrino weighting algorithm [23]: $m_t^{\rm NW}$ is still widely used in the dilepton channel [7–9] as a

template method. A study by the CDF Collaboration indicates that $m_t^{\rm NW}$ has the most-precise statistical uncertainty among the variables used in its study [24].

(2) $m_{\rm T2}$—The collider variable related to the missing transverse mass of the system of two missing particles: $m_{\rm T2}$ was initially developed for new physics particles as a variable sensitive to the mass of new particles in pair production [25]. Because of the two missing particles (neutrinos) in the dilepton channel, this variable was suggested for the $M_{\rm top}$ measurement [26]. The CDF Collaboration performed the $M_{\rm top}$ measurement using $m_{\rm T2}$ and improved the precision with the simultaneous usage of $m_t^{\rm NW}$ and $m_{\rm T2}$ [24]. They also extensively studied the statistical as well as systematical uncertainties in the different variables. The expected statistical uncertainty obtained using $m_t^{\rm NW}$ had the smallest value; however, the systematic uncertainty was larger than that in the measurement performed using $m_{\rm T2}$. This difference was caused by different JES systematics due to the different dependence of each variable on $M_{\rm top}$ and JES.

(3) $p_T^{\rm lepton}$—The average $p_T^{\rm lepton}$ of two leptons: $p_T^{\rm lepton}$ has been suggested as a good variable for the $M_{\rm top}$ measurement with large statistics in the LHC [27]. $p_T^{\rm lepton}$, in general, does not depend on JES, and it negligibly contributed to the systematic uncertainty from JES. The CDF Collaboration has performed measurements using this variable in the lepton + jets [28] and the dilepton [29] channels, obtaining negligible JES systematic uncertainty but relatively large statistical uncertainty. This is basically caused by the low sensitivity of $M_{\rm top}$ and the insensitivity of JES to the lepton p_T variable.

The three variables discussed above have very different $M_{\rm top}$ and JES dependences. Therefore, it is possible to extract both $M_{\rm top}$ and JES information together if we use the three variables simultaneously. To simulate and extract JES information from the MC pseudoexperiments, we vary the scale of the jet energy relative to nominal JES ($\Delta_{\rm JES}$) in the MC simulations from -15% to 15% in 1.5% steps. Even though the overall JES uncertainties depend on the jet η and p_T, we use the overall variation for simplicity in this paper. The overall jet energy uncertainty in the CMS experiment [30] is expected to be from 2 to 3%. The plots of each variable are shown in Fig. 1 with different $M_{\rm top}$ as well as $\Delta_{\rm JES}$. $m_t^{\rm NW}$ and $m_{\rm T2}$ depend on both $M_{\rm top}$ and $\Delta_{\rm JES}$. However, it is clear that $p_T^{\rm lepton}$ does not depend on $\Delta_{\rm JES}$. Therefore, using three variables simultaneously in a single likelihood function, we can extract $\Delta_{\rm JES}$ information *in situ* as a nuisance parameter of the $M_{\rm top}$ measurement.

The technique of using three variables in a single likelihood function was developed by the CDF Collaboration and has been used in lepton + jets channel measurements. We follow the procedure discussed in Ref. [7]. We estimate
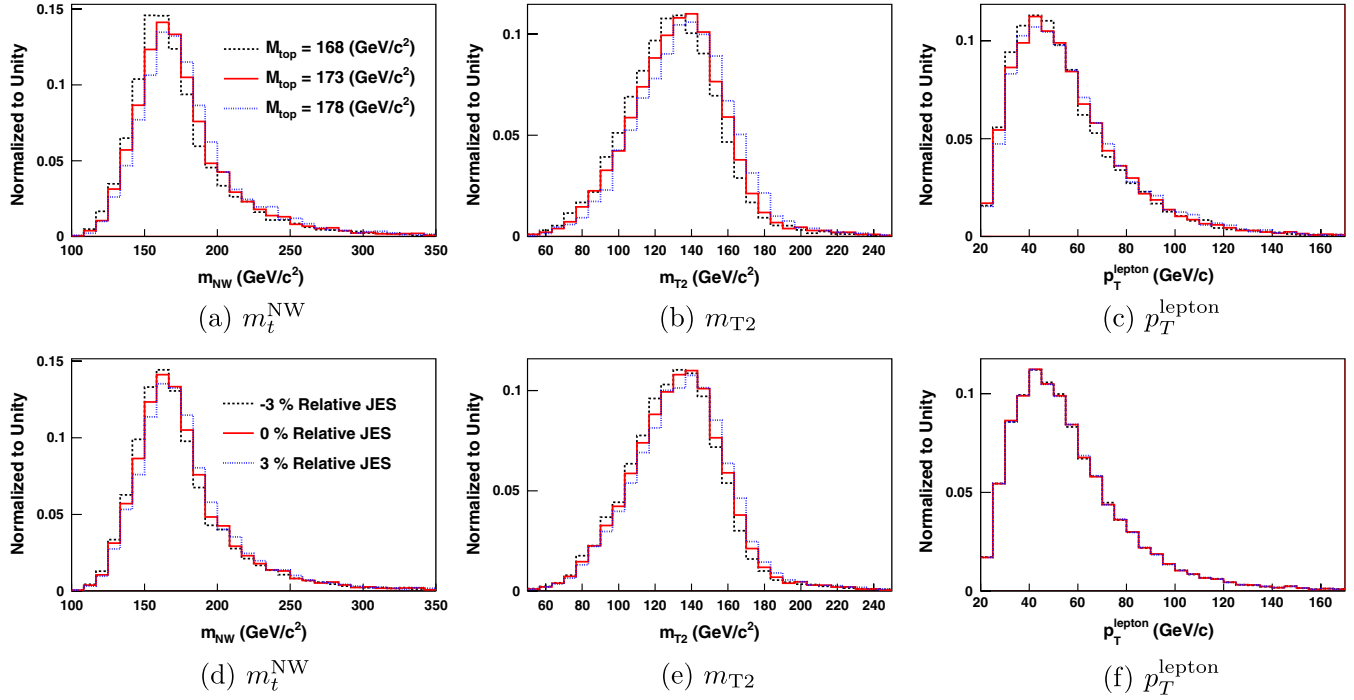
FIG. 1 (color online).   Plots of the three variables for different M_{top} (a, b, c) and Δ_{JES} (d, e, f).

the probability density functions (PDFs) of signals, $P_{\text{sig}}(m_t^{\text{NW}}, m_{\text{T2}}, p_T^{\text{lepton}}; M_{\text{top}}, \Delta_{\text{JES}})$, using kernel density estimation [31]. For the discrete values of M_{top} and Δ_{JES}, we estimate the PDFs for the observables. We smooth and interpolate the MC distributions to find PDFs for arbitrary values of M_{top} and Δ_{JES} using the local polynomial smoothing method [32]. We then build the unbinned maximum likelihood [33] for N events (and n_s expected signal events) with the Poisson fluctuation:

$$\mathcal{L} = \frac{e^{-n_s} n_s^N}{N!} \Sigma_{i=1}^N P_{\text{sig}}(m_t^{\text{NW}}, m_{\text{T2}}, p_T^{\text{lepton}}; M_{\text{top}}, \Delta_{\text{JES}}).$$

The quantity $P_{\text{sig}}(m_t^{\text{NW}}, m_{\text{T2}}, p_T^{\text{lepton}}; M_{\text{top}}, \Delta_{\text{JES}})$ denotes the signal PDFs as determined by kernel density estimation and local polynomial smoothing as a function of M_{top} and Δ_{JES}. We minimize the negative logarithm of the likelihood using MINUIT [34] with respect to all three parameters (n_s, M_{top}, and Δ_{JES}). The uncertainty on M_{top} or Δ_{JES} is found by searching for the points where the negative logarithm of the likelihood minimized with respect to all other parameters deviates by $1/2$ from the minimum. The uncertainty on the M_{top} measurement obtained in this way includes the statistical uncertainty as well as the systematic uncertainty of the overall JES owing to the allowed variation of Δ_{JES}. We also perform the likelihood fit without varying Δ_{JES} (M_{top}-only fit) and set it to zero. The M_{top}-only fit measurement allows us to compare the performance of the new technique with the *in situ* JES calibration (2D fit).

We test the likelihood procedure using MC pseudoexperiments. We construct pseudodata from a certain value of

M_{top} and Δ_{JES}. We select the number of signal events from a Poisson distribution with a mean equal to the expected number of signal events, 5961 ± 545. We perform the maximum likelihood fit described in the previous section. In principle, the likelihood fit, on average, returns the value of the top quark mass used to generate the pseudoexperiments. Figure 2 shows an example of a likelihood fit contour from a single pseudoexperiment of the 2D fit. We use a $M_{\text{top}} = 173$ GeV/c^2 and $\Delta_{\text{JES}} = 0\%$ sample
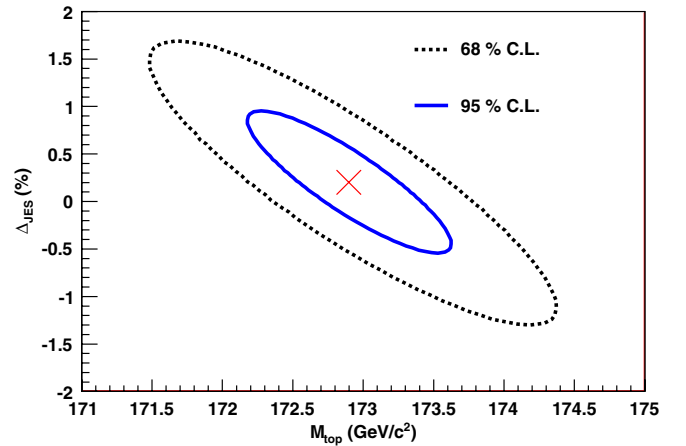

FIG. 2 (color online).   Negative log-likelihood contours of a single pseudoexperiment using a $M_{\text{top}} = 173$ GeV/c^2 and $\Delta_{\text{JES}} = 0.0\%$ sample. The minimum is indicated by the "x" and corresponds to the most probable top-quark mass and Δ_{JES}. The contours are drawn at 68% confidence level and 95% confidence level, corresponding to 1σ and 2σ uncertainty of the measurement, respectively.
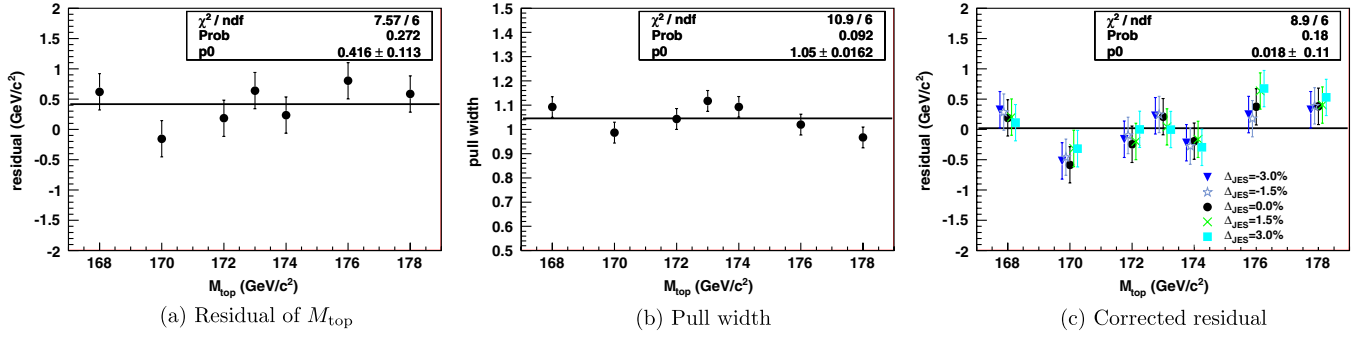
FIG. 3 (color online).   Plots for checking the bias in the fitted top-quark mass (a) and width of the pull distribution (b) for the 2D fit before corrections. With the corrections of the residual and pull width described in the text, we obtain the mass bias for various $\Delta_{\rm JES}$ points at (c) for the 2D fit.

for this experiment and obtain $M_{\rm top} = 172.89 \pm 0.57$ GeV/c^2 and $\Delta_{\rm JES} = 0.22 \pm 0.64\%$. We verify the $M_{\rm top}$-only fit using the same pseudodata and obtain $M_{\rm top} = 173.12 \pm 0.37$ GeV/c^2. We perform this experiment 3000 times for seven different $M_{\rm top}$ values ranging from 168 to 178 GeV/c^2. Figures 3(a) and 3(b) show the average residual (deviation from input $M_{\rm top}$) and the width of the pull (the ratio of the residual to the uncertainty reported by MINUIT), respectively, in the 2D fit using samples of $\Delta_{\rm JES} = 0.0\%$ without corrections. The small positive bias,

0.43 GeV/c^2, is corrected and the uncertainty is correspondingly increased by 5% to correct for the width of the pull distribution. The residual and pull width for the $\Delta_{\rm JES}$ parameter are also investigated using the same pseudoexperiments. We obtain a small negative bias of 0.33% and a pull width of 1.08. We apply suitable corrections for the nuisance parameter $\Delta_{\rm JES}$. These corrections enable us to test the $M_{\rm top}$ measurement with different $\Delta_{\rm JES}$ parameters. We vary the input $\Delta_{\rm JES}$ from -3.0% to 3.0%, corresponding to approximately 1–1.5 times the JES uncertainty

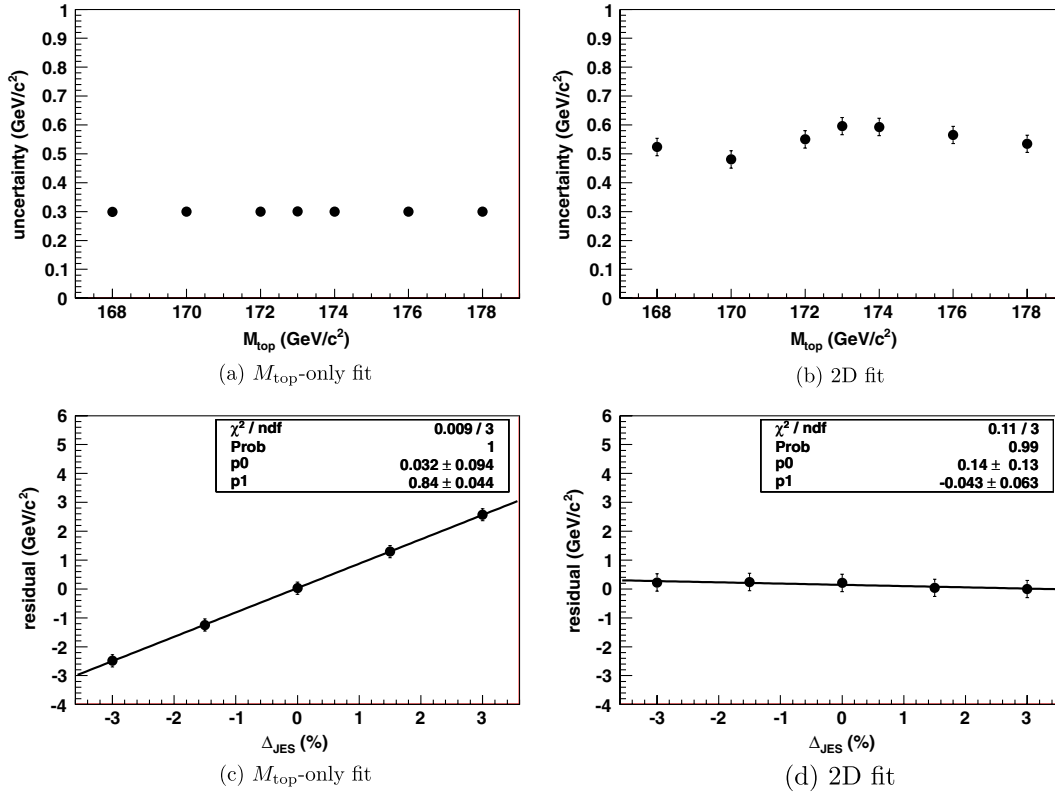FIG. 4 (color online).   The expected statistical uncertainty for the $M_{\rm top}$-only fit (a) and the $M_{\rm top} - \Delta_{\rm JES}$ 2D fit (b). To evaluate the systematic uncertainty from the overall JES uncertainty, the mass residuals as a function of $\Delta_{\rm JES}$ are shown for the $M_{\rm top} = 173$ GeV/c^2 sample in (c) ($M_{\rm top}$-only fit) and (d) (2D fit).

in the LHC experiment [30]. Figure 3(c) shows the residual distribution for various Δ_{JES} samples. There is no significant effect of Δ_{JES} on the mass residual. Therefore, we do not apply a correction of M_{top} for the Δ_{JES} parameter. The same procedures for the M_{top}-only fit are performed, and we find no bias. The width of the pull distribution is also consistent with unity.

With the corrections of the residual and the width of the pull, we can obtain the expected uncertainty. In Fig. 4, the expected statistical uncertainty for the M_{top}-only fit (a) and the 2D fit (b) have been plotted. Because the 2D fit absorbs the overall uncertainty from JES, it has a larger statistical uncertainty than the M_{top}-only fit. If we choose $M_{top} = 173$ GeV/c^2, close to the world average of M_{top} [3], the 2D fit has an expected statistical uncertainty of 0.60 ± 0.03 GeV/c^2, whereas that of the M_{top}-only fit is 0.36 ± 0.02 GeV/c^2. However, the M_{top}-only fit may have a much larger JES systematic uncertainty. To obtain the systematic uncertainty of JES, we check the mass residual as a function of Δ_{JES}, as shown in Fig. 4(c) (M_{top}-only fit) and (d) (2D fit). If we consider an optimistic 2% overall JES uncertainty [30], the M_{top}-only fit gives an expected JES systematic uncertainty of 1.68 GeV/c^2, whereas that in the case of the 2D fit is 0.08 GeV/c^2. If we consider a realistic JES systematic uncertainty, which will be 2%–3% depending on jet p_T and η, the M_{top}-only fit may have a JES systematic uncertainty slightly larger than 1.68 GeV/c^2, which is consistent with the recent CMS measurement of 2.0 GeV/c^2 without *in situ* JES calibration. However, the JES systematic uncertainty for the 2D fit will still be very small. We calculate the expected uncertainty from statistics and overall JES systematics together with a quadrature sum. The M_{top}-only fit gives an expected uncertainty of 1.72 GeV/c^2, whereas that in the case of the 2D fit is much smaller, 0.61 GeV/c^2. The expected statistical uncertainty of Δ_{JES} in the 2D fit is $0.61 \pm 0.05\%$ if we use $M_{top} = 173$ GeV/c^2 and a $\Delta_{JES} = 0.0\%$ sample.

In conclusion, we have presented a novel technique for top-quark mass measurement in the dilepton channel by performing *in situ* JES calibration using three different variables. Our study shows an approximately 0.61 GeV/c^2 statistical uncertainty including overall JES systematic uncertainty, with 5 fb^{-1} LHC data. This technique significantly improves the precision of the M_{top} measurement as compared to the M_{top}-only fit. In particular, the overall JES uncertainty, which would be approximately 1.7 GeV/c^2 without additional calibration of JES with a 2% overall uncertainty assumed, is significantly reduced with the *in situ* JES calibration method. To obtain precision below 1 GeV/c^2, one still needs to improve other important systematic uncertainties in the LHC experiments. However, the experience at the Tevatron presented a well-controlled systematic uncertainty of around 0.86 GeV/c^2 from the other systematic sources [10]. If the other systematic uncertainties are controlled to the level of the Tevatron, we eventually can reach a precision of 1 GeV/c^2 in the dilepton channel.

[1] K. Nakamura *et al.* (Particle Data Group), J. Phys. G **37**, 075021 (2010).

[2] ALEPH, CDF, D0, DELPHI, L3, OPAL, SLD Collaborations, LEP Electroweak Working Group, Tevatron Electroweak Working Group, SLD Electroweak, and Heavy Flavor Groups, arXiv:1012.2367.

[3] Tevatron Electroweak Working Group (CDF and D0 Collaborations).

[4] G. L. Kane and S. Mrenna, Phys. Rev. Lett. **77**, 3502 (1996).

[5] C. Balazs, M. Carena, and C. E. M. Wagner, Phys. Rev. D **70**, 015007 (2004); T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **104**, 251801 (2010).

[6] J. F. Gunion, H. E. Haber, F. E. Paige, W.-K. Tung, and S. S. D. Willenbrock, Nucl. Phys. **B294**, 621 (1987).

[7] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. D **83**, 111101 (2011).

[8] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. Lett. **107**, 082004 (2011).

[9] CMS Collaboration, CMS Report No. PAS TOP-11-016.

[10] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **105**, 252001 (2010).

[11] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. D **84**, 032004 (2011).

[12] CMS Collaboration, CMS Report No. PAS TOP-11-015.

[13] G. Aad *et al.* (ATLAS Collaboration), Eur. Phys. J. C **72**, 2046 (2012).

[14] A. Abulencia *et al.* (CDF Collaboration), Phys. Rev. D **73**, 032003 (2006).

[15] V. M. Abazov *et al.* (D0 Collaboration), arXiv:1201.5172.

[16] A. H. Hoang and J. W. Stewart, Nucl. Phys. B, Proc. Suppl. **185**, 220 (2008).

[17] U. Langenfeld, S. Moch, and P. Uwer, Phys. Rev. D **80**, 054009 (2009).

[18] A. H. Hoang, A. Jain, I. Scimemi, and I. W. Stewart, Phys. Rev. Lett. **101**, 151602 (2008).

[19] S. Fleming, A. H. Hoang, S. Mantry, and I. W. Stewart, Phys. Rev. D **77**, 074010 (2008).

[20] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, J. High Energy Phys. 06 (2011) 128.

[21] T. Sjöstrand, S. Mrenna, and P. Skands, J. High Energy Phys. 05 (2006) 026.

[22] S. Ovyn, X. Rouby, and V. Lemaitre, arXiv:0903.2225.

[23] B. Abbott *et al.* (D0 Collaboration), Phys. Rev. D **60,** 052001 (1999); A. Abulencia *et al.* (CDF Collaboration), Phys. Rev. D **73,** 112006 (2006).

[24] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. D **81,** 031102 (2010).

[25] C. Lester and D. Summers, Phys. Lett. B **463,** 99 (1999); A. Barr, C. Lester, and P. Stephens, J. Phys. G **29,** 2343 (2003).

[26] W. S. Cho, K. Choi, Y. G. Kim, and C. B. Park, Phys. Rev. D **78,** 034019 (2008).

[27] N. Giokaris *et al.*, Report No. JINR-E1-2005-104, 2005 (to be published).

[28] T. Aaltonen *et al.* (CDF Collaboration), Phys. Lett. B **698,** 371 (2011).

[29] CDF Collaboration, CDF Report No. Public Note 9831.

[30] S. Chatrchyan *et al.* (CMS Collaboration), JINST **6,** P11002 (2011).

[31] K. Cranmer, Comput. Phys. Commun. **136,** 198 (2001).

[32] C. Loader, *Local Regression and Likelihood* (Springer, New York, 1999).

[33] R. Barlow, Nucl. Instrum. Methods Phys. Res., Sect. A **297,** 496 (1990).

[34] F. James and M. Roos, Comput. Phys. Commun. **10,** 343 (1975).