CDF note 8803

Top Quark Mass Measurement in Dileptonic Channel using Template Method in 1.2 fb^{-1}

The CDF Collaboration
URL http://www-cdf.fnal.gov

(Dated: July 15, 2007)

We present a top quark mass measurement in the dilepton channel using a template method. The unconstrained system of dilepton events is solved using the $t\bar{t}$ longitudinal momentum. In this measurement, the sample is separated into b-tagged and non-tagged samples. In 1.2 fb^{-1} of data, we measure the top mass $M_{top} = 169.7^{+5.2}_{-4.9}(stat.) \pm 3.1(syst.)$ GeV/c^2 from the 70 events passing the event selection and the mass reconstruction.

Preliminary Results

I. INTRODUCTION

The top quark mass measurement is one of the most important measurements at Tevatron. Its mass is close to the scale of electroweak symmetry breaking which raises a possibility that perhaps the top mass is not generated by the SM Higgs mechanism in the same way as postulated with other fermions. The top quark mass has a significant contribution to radiative terms in theoretical calculations of many observables. This is especially important on the predictions of the Higgs boson mass. The top quark mass also enters into various alternative models of particle physics. A more precise measurement can be used to further refine or reject many of those theories.

At the Tevatron collider, the main production mechanism for the top quark is the pair production via quark annihilation or gluon fusion. The top quark decays into a b-quark and a W-boson with a branching ration nearly 100%. In dileptonic channel, the two resulting W-bosons decay leptonically. Such events have little background contamination and only two possible parton-jet assignments, but they also have low statistics and two neutrinos escaping detector.

This note describes a measurement of the top quark mass in dileptonic channel using data collected by the the CDF detector [1] at the Fermilab Tevatron. The data sample corresponds to a total integrated luminosity of 1.2 fb^{-1}.

II. DATA SAMPLE & EVENT SELECTION

The signiture of dileptonic $t\bar{t}$ events consist of two high-p_T leptons , missing transverse energy ($\not\!\!E_T$) from the undetected neutrinos, and two jets coming from the hadronization of the b-quarks. Additional jets are often produced via initial state (ISR) and final state radiation (FSR).

The data are collected with an inclusive lepton trigger that requires an electron with E_T >18 GeV/c^2 or a muon with P_T >18 GeV/c. After full event reconstruction, the electron is required to have E_T >20 GeV/c^2 and the muon is required to have P_T >18 GeV/c.

One lepton must pass strict lepton identification requirements and be isolated. A lepton is isolated if the total transverse energy within cone $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.4$, minus the candidate lepton E_T, is less than 10% of the candidate lepton E_T. Tight electron candidates have a well-measured track pointing at an energy deposition in the calorimeter. In addition, the candidate's electromagnetic shower profile must be consistent with that expected for electrons. Similarly, tight muon candidates must have a well-measured track linked to hits in the muon chambers and energy deposition in the calorimeters consistent with that expected for muons.

The second lepton candidate is selected the same way as the first lepton, with the exceptions that it does not need to be isolated and muon identification requirements are relaxed. The tight and loose lepton have to be oppositely charged.

We require candidate events to have $\not\!\!E_T > 25$ GeV after correcting for all escaping muon momenta. If $\not\!\!E_T \leq 50$ GeV), we further require that the $\not\!\!E_T$ vector is at least 20° apart from the closest lepton or jet. The event is required to have at least two jets with $E_T > 15$ GeV and which are detected in $|\eta| < 2.5$, where the jet is defined as a fixed-cone cluster with a cone size of R = 0.4.

Moreover, we require the scalar transverse energy sum $H_T > 200$ GeV, and the events with cosmic rays, conversions or Z events are rejected. The expected number of events passing the event selection is shown in Table I.

	signal	background
Total	55.95±4.26	25.56±5.54
b-tagged	30.3±2.5	2.8±1.1
non-tagged	25.7±2.1	22.8±5.0

TABLE I: Expected sample composition in 1.2 fb^{-1}.

III. FULL KINEMATIC METHOD

In dilepton channel, the system is underconstrained for top mass fitting.

The momentum components of the charged leptons and jets can be measured with the detector, as well as the two components of $\not\!\!E_T$. The masses of the final state b-quarks and leptons are known, and the neutrinos are assumed to be massless. By further assuming $W^{\pm}$ boson masses, as well as that t and $\bar{t}$ masses are equal, we can write 23 equations for 24 quantities.

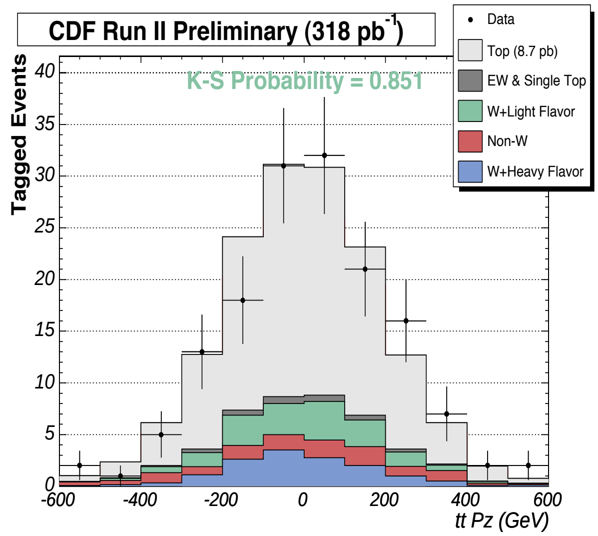

FIG. 1: Longitudinal momentum of data $t\bar{t}$ events in lepton+jets channel.

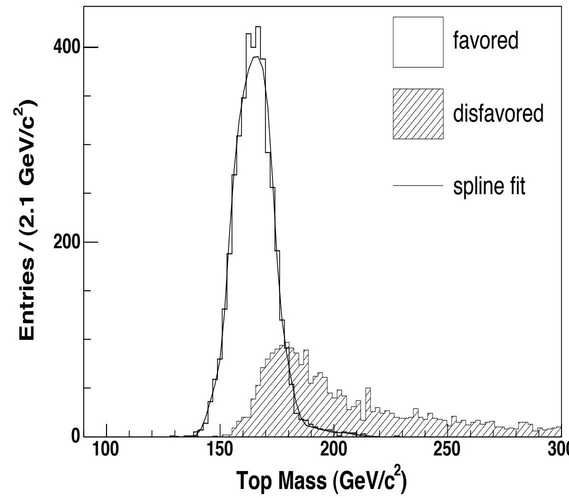

FIG. 2: Reconstructed top mass distributions for one event. Two histograms correspond to two possibilities to pair a lepton and a jet.

The system can be solved by using a distribution which is independent of the top mass. In this analysis, we take the momentum z-component of the $t\bar{t}$ system, $P_z^{t\bar{t}}$. This distribution is almost independent of the top mass, zero-centered Gaussian with $\sigma \approx$ 195 GeV/c.

The measured quantities have experimental uncertainties. We smear jet energies, $\not{E}_T$ and $P_z^{t\bar{t}}$ by randomly drawing these quantities 10000 times from a distribution which is centered at the measured value and has the width of the expected error. The kinematic equations are solved for each set of smeared quantities, yielding to a distribution of possible reconstructed top masses.

There can be up to 8 solutions to the kinematic equations. Four possible solutions correspond to different neutrino solutions. We select the neutrino solution which has the smallest effective mass of the $t\bar{t}$ system. The other two possible solutions correspond to two possible ways to associate a b-jet to a b-quark. We select the combination which has higher mass reconstruction probability. The reconstructed top mass distributions for the favoured and disfavoured b-jet − b-quark combinations are shown in Figure 2. The most probable value from the favoured distribution is considered as the reconstructed top mass for the given event. The normalized distribution of reconstructed top masses is called a template.

The final step in top mass reconstruction chain is the likelihood fit. We fit the reconstructed top mass values of the data candidates to the Monte Carlo templates.

A. Top Mass Templates for Signal and Background

We create signal templates using Pythia Monte Carlo samples with the input top masses from 150 GeV to 200 GeV with 2 GeV steps. The fitting function $f(m_{Top}^{rec}, m_{Top}^{orig})$ is a combination of Gaussian and Landau functions and depends on reconstructed mass, m_{Top}^{rec}, and the given true top mass, m_{Top}^{orig}. The parameters of this function are assumed to have linear dependence on the true top quark mass. The signal template for b-tagged events is shown in Figure 3, and the signal template for non-tagged events is shown in Figure 4.

The dilepton channel has the advantage of a good signal-to-background fraction compared to the other decay channels. The dominant backgrounds in the dilepton channel are diboson (WW,WZ,ZZ) production, Drell-Yan ($q\bar{q} \to Z/\gamma^* \to \ell^+\ell^-$) production and fake leptons coming from $W \to \ell\nu + jets$ events where a jet is falsely reconstructed as a lepton candidate.

To take into account the contribution from background events to the top mass reconstruction, we create background templates from each background process. These background templates are combined according to their expected yield. The combined background template is shown in Figure 5. We use the same background template for b-tagged and non-tagged events.

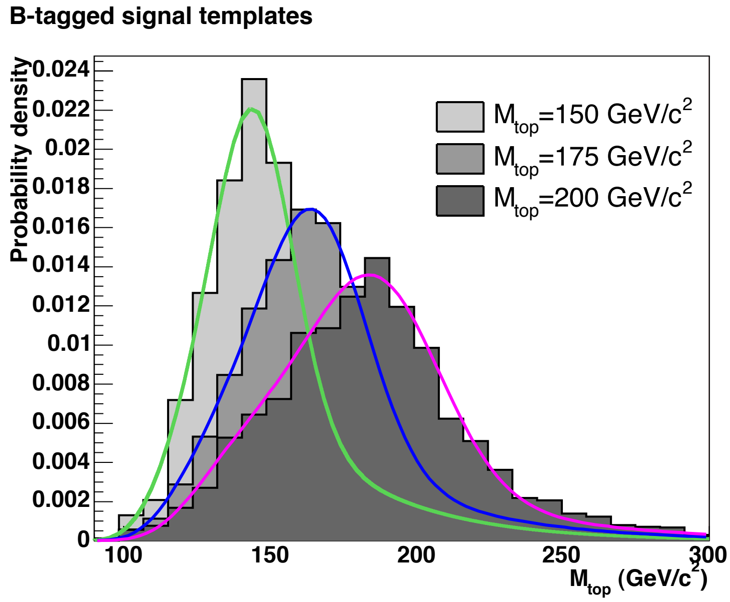

FIG. 3: B-tagged signal templates and parametrisations

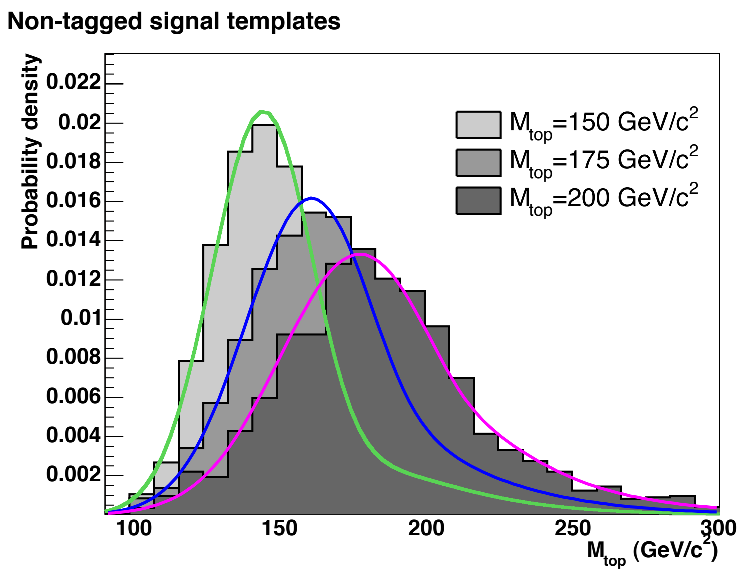

FIG. 4: Non-tagged signal templates and parametrisations

B. Likelihood fit

We use a maximum likelihood method to get the final top quark mass estimate. The reconstructed top mass distribution from data is compared to the parametrisations of the signal and background templates using the formula:

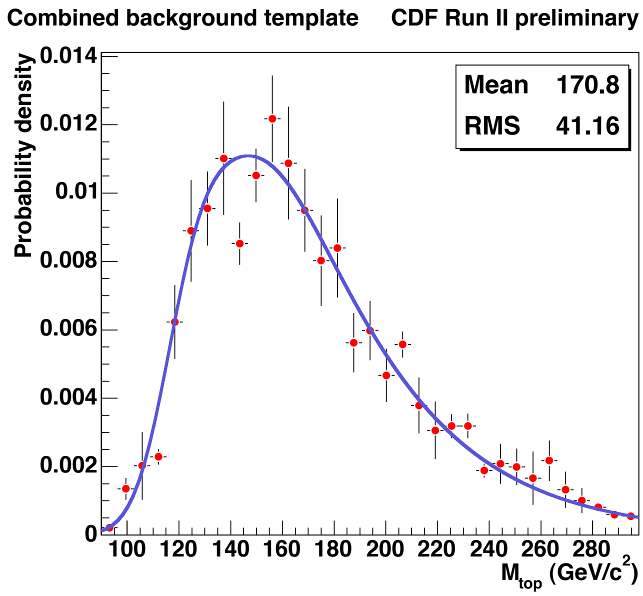

FIG. 5: Combined background template.

$$
\begin{aligned}
\mathcal{L} &\equiv \mathcal{L}_{b-tagged} \times \mathcal{L}_{non-tagged} \\
\mathcal{L}_{sub-sample} &\equiv \mathcal{L}_{shape} \times \mathcal{L}_{nev} \times \mathcal{L}_{bg} \\
\mathcal{L}_{shape} &\equiv \prod_{i=1}^{n} \frac{n_s \times f_s({m_{t_i}}^{rec}, m_t^{orig}) + n_b \times f_b({m_{t_i}}^{rec})}{n_s + n_b} \\
\mathcal{L}_{nev} &\equiv \frac{e^{-(n_s+n_b)}(n_s+n_b)^N}{N!} \\
-\ln \mathcal{L}_{bg} &\equiv \frac{(n_b - n_b^{exp})^2}{2\sigma_{n_b}^2}
\end{aligned}
$$

where the top mass m_t^{orig}, number of background events n_b, and number of signal events n_s are fit parameters. The number of background events is Gaussian constrained by $\mathcal{L}_{bg}$ and the Poisson term $\mathcal{L}_{nev}$ makes sure the total number of events $n_s + n_b$ is in agreement with the number of data events.

IV. TESTING THE METHOD WITH THE PSEUDOEXPERIMENTS

We perform pseudoexperiments to see if our method gives the correct estimate of the top mass and its error. We randomly choose expected number of signal and background events from signal and background templates, and perform likelihood fit on the selected set of events. We repeat this procedure 10000 times. The number of events in each pseudo-experiment is Poisson fluctuated around the expected value.

To see if our estimate of the error on top mass is correct we looked at the pull variable. The pull is defined as the ratio of the difference between original and reconstructed mass and the average of the estimated positive and negative error on the top mass. The plots for mass dependence of pull mean and pull width are shown in Figures 8 and 9.

V. SYSTEMATIC UNCERTAINTIES

Top mass determination using this method is sensitive to the Monte Carlo templates as well as to the systematics with respect to the jet reconstructions algorithms and different corrections applied to them. We have performed systematics studies with respect to each known source based on the studies of the pseudo-experiments.

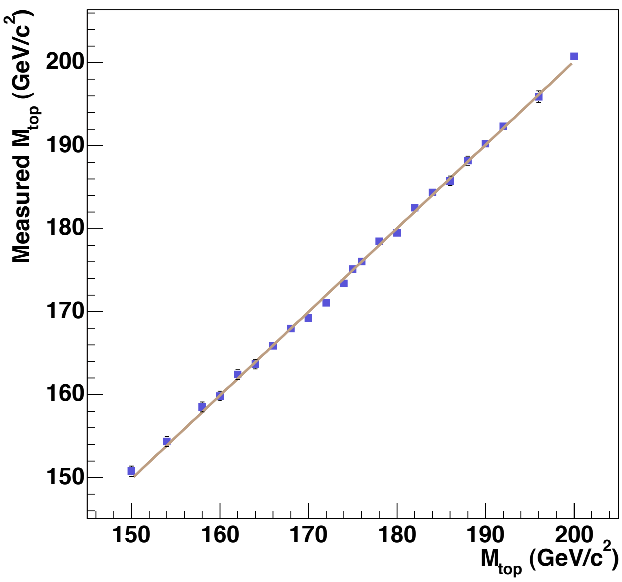

FIG. 6: The dependence of reconstructed mass on original mass.

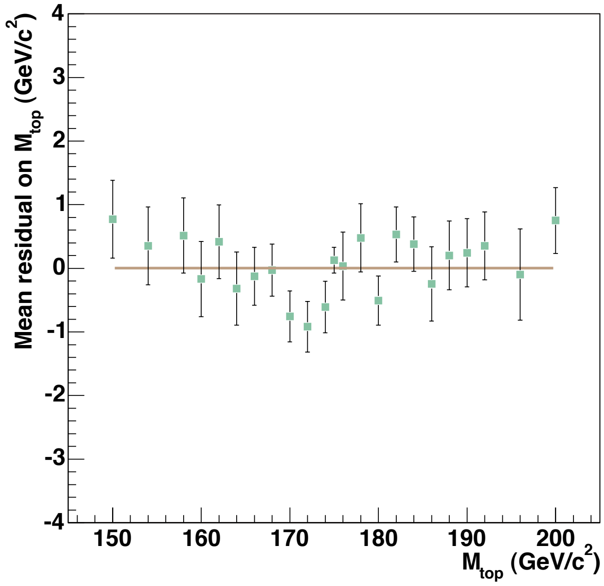

FIG. 7: The mass dependence of difference between original and reconstructed mass.

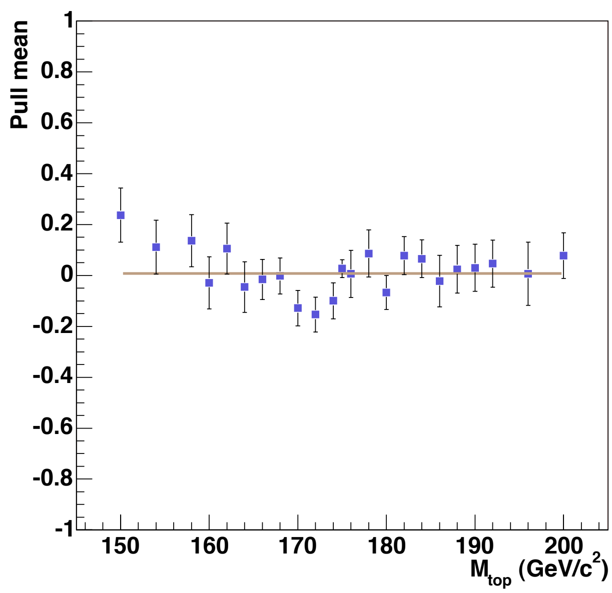

FIG. 8: The mass dependence of pull mean.

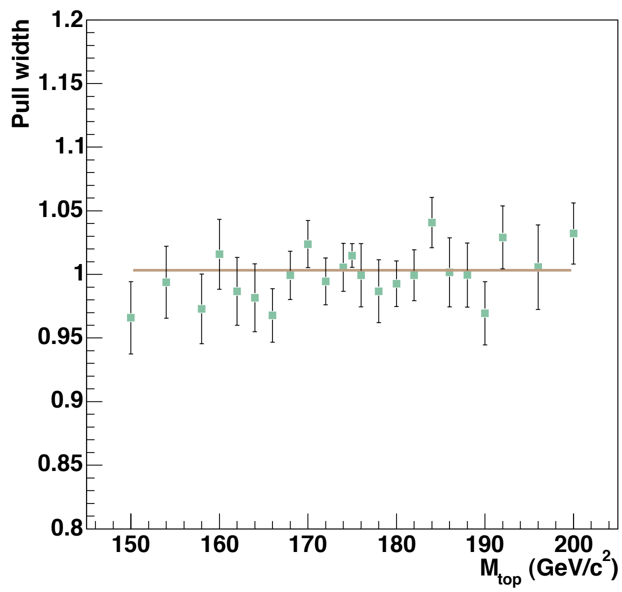

FIG. 9: The mass dependence of pull width.

The largest source of systematic error comes from the uncertainties on the jet energy corrections. We have studied it by changing the corrections by $\pm 1\sigma$. The total systematics coming from this source is 2.9 GeV/c^2. This error is related to light quark jets, so we evaluated an additional systematic error due to b-jet energy scale and found it to be 0.5 GeV/c^2. Lepton p_T was varied by $\pm 1\%$, which leads to the systematic uncertainty of 0.2 GeV/c^2.

We compared the HERWIG [2] and the PYTHIA [3] and assigned the systematic error due to different types of generators to be 0.3 GeV/c^2. The initial and final state radiations (ISR and FSR) uncertainties are estimated using PYTHIA Monte Carlo samples in which QCD parameters for parton shower evolution are varied based on the studies of the CDF Drell-Yan data. The systematic uncertainty due to the ISR and FSR is 0.2 GeV/c^2 and 0.4 GeV/c^2 respectively. Another systematic error is coming from using different parton distribution functions (PDF). We have considered 20 pairs of CTEQ6 $\pm 1\sigma$ uncertainty sets, two sets of MRSTs for different Λ_{OCD} values, and the difference in CTEQ and MRST PDF groups. The total systematic from the PDFs is 0.6 GeV/c^2. Kinematics of the background processes was compared between data and Monte-Carlo. The difference was propagated to systematic error on top mass, 0.3 GeV/c^2. We also evaluated the systematic error related to the template statistics. We Poisson fluctuated number of events in each bin, and performed pseudo-experiments with the changed template. The procedure was repeated 100 times. The systematic uncertainty due to the signal template statistics is 0.2 GeV/c^2 and due to the background template statistics is 0.5 GeV/c^2. We also estimate an uncertainty coming from possible imperfections in modelling the Drell-Yan and fake background, as well as an uncertainty coming from using the general background template for the b-tagged events.

The systematic uncertainties are summarized in Table II.

CDF Run II preliminary

Source	ΔM_{top} (GeV/c^2)
Jet Energy Scale	2.9
B-jet energy scale	0.5
Lepton energy scale	0.4
Generators	0.3
ISR	0.2
FSR	0.4
PDFs	0.6
B-tagging E_T dependence	0.1
Sample composition	0.3
Signal template statistics	0.2
Background template shape	0.3
Background template statistics	0.5
Total	3.1

TABLE II: Summary of systematic uncertainties.

VI. RESULTS

We apply the loglikelihood fit to the 70 data events which pass the event selection and the mass reconstruction requirements. 31 of the events are b-tagged and 39 are non-tagged. Setting the number of b-tagged background events to 2.2±1.1 and the number of non-tagged background events to 17.8±4.7, the result of the likelihood fit is

$$M_{top} = 169.7^{+5.2}_{-4.9}(stat.) \quad \text{GeV}/c^2$$

where the amount of background as a result of the fit is 2.2±1.2 b-tagged background events and 15.5±4.6 non-tagged background events.

Figure 10 shows the reconstructed top mass distribution from the 70 data events together with the signal and background parametrisation functions. The expected statistical error distribution is shown in Figure 11.

VII. CONCLUSION

We have measured the top quark mass in dileptonic channel using 1.2 fb^{-1} of data collected by the CDF experiment. The measured top mass is $M_{top} = 169.7^{+5.2}_{-4.9}(stat.) \pm 3.1(syst.)$ GeV/c^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

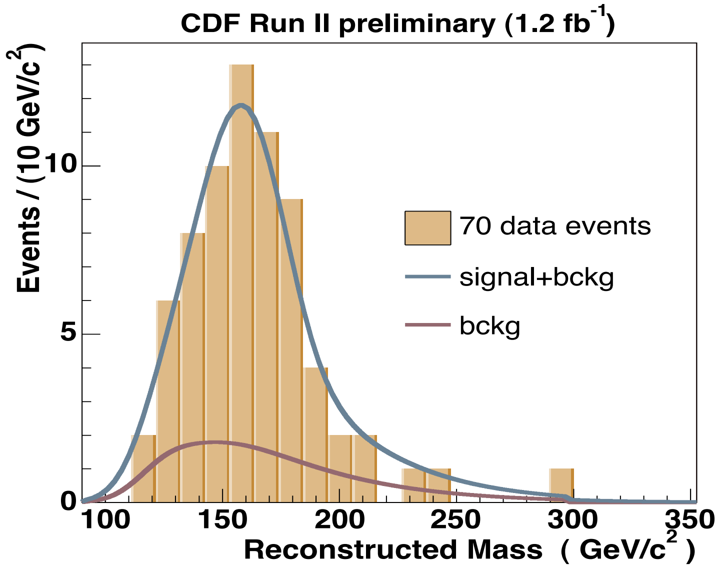

FIG. 10: Reconstructed top mass distribution from data together with signal and background template parametrisations.

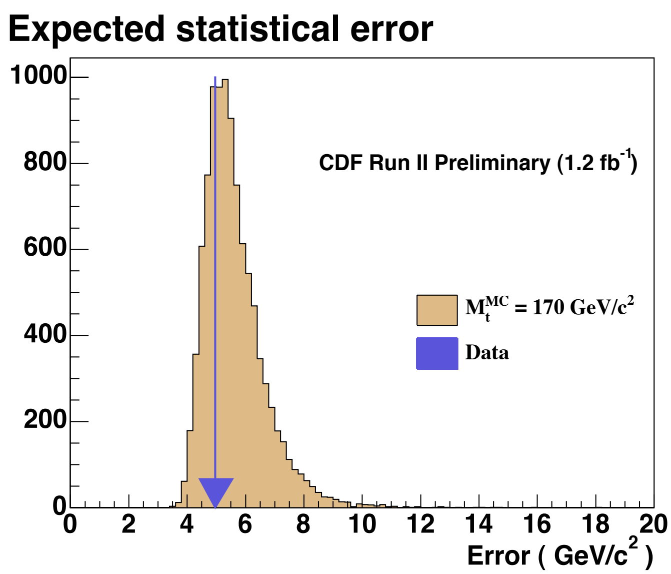

FIG. 11: Expected statistical error and error obtained from the data.

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 Particle Physics and Astronomy Research Council and the Royal Society, UK; the Russian Foundation for Basic Research; the Comisión Interministerial de Ciencia y Tecnología, Spain; in part by the European Community's Human Potential Programme under contract HPRN-CT-2002-00292; and the Academy of Finland.

---

[1] The CDFII Detector Technical Design Report, Fermilab-Pub-96/390-E (1996)

[2] G. Corcella et al., HERWIG 6: An Event Generator for Hadron Emission Reactions with Interfering Gluons (including supersymmetric processes), JHEP **01**, 10 (2001).
[3] T. Sjostrand et al., High-Energy-Physics Event Generation with PYTHIA 6.1, Comput. Phys. Commun. **135**, 238 (2001).