

E1-2014-18

J. Budagov¹, V. Glagolev¹, I. Suslov¹, G. Velev²
(on behalf of the CDF Collaboration)

TOP-QUARK MASS MEASUREMENT
IN THE $t\bar{t}$ -DILEPTON CHANNEL USING
THE FULL CDF RUN II DATA SET

¹Joint Institute for Nuclear Research, Dubna

²Fermi National Accelerator Laboratory, Batavia, USA

Будагов Ю. А. и др. (от коллаборации CDF)

E1-2014-18

Измерение массы топ-кварка в эксперименте CDF
на дилептонной выборке $t\bar{t}$ -событий

На основе данных эксперимента CDF измерена масса топ-кварка на выборке дилептонных событий, $t\bar{t}$ -кандидатов. Полный набор данных, полученный при $p\bar{p}$ -взаимодействиях с энергией в системе центра масс 1,96 ТэВ и соответствующий интегральной светимости $9,1 \text{ фб}^{-1}$, содержит 520 событий, удовлетворяющих критериям отбора. Масса топ-кварка получена из фита распределения некоторой наблюдаемой переменной суммой сигнальной и фоновой функций. Данная переменная определена специальным образом с целью уменьшить систематическую ошибку измерения, связанную с погрешностью калибровки калориметра. Для сигнала и фона функции распределения плотности вероятностей определяются с помощью моделирования. Выполнив фит методом наибольшего правдоподобия, мы получили значение массы топ-кварка $(170,80 \pm 1,83 \text{ (стат.)} \pm 2,69 \text{ (сист.)}) \text{ ГэВ}/c^2$ (или $(170,80 \pm 3,25) \text{ ГэВ}/c^2$).

Работа выполнена в Лаборатории ядерных проблем им. В. П. Дзелепова ОИЯИ.

Сообщение Объединенного института ядерных исследований. Дубна, 2014

Budagov J. et al. (on behalf of the CDF Collaboration)

E1-2014-18

Top-Quark Mass Measurement in the $t\bar{t}$ -Dilepton Channel Using
the Full CDF Run II Data Set

We present a measurement of the top-quark mass with $t\bar{t}$ -dilepton events using the full CDF Run II data set, which corresponds to an integrated luminosity of 9.1 fb^{-1} collected from $\sqrt{s} = 1.96 \text{ TeV}$ $p\bar{p}$ collisions at the Fermilab Tevatron. A sample of 520 events is obtained after all selection requirements. The top-quark mass is estimated by a fit of the distribution of some variable to a sum of signal and background contributions. This variable is defined using special approach to reduce the systematic error due to the jet energy scale uncertainty. Templates are built from simulated $t\bar{t}$ and background events, and parameterized in order to provide probability distribution functions. A likelihood fit of the data returns the top-quark mass of $(170.80 \pm 1.83 \text{ (stat.)} \pm 2.69 \text{ (syst.)}) \text{ GeV}/c^2$ (or $(170.80 \pm 3.25) \text{ GeV}/c^2$).

The investigation has been performed at the Dzhelapov Laboratory of Nuclear Problems, JINR.

Communication of the Joint Institute for Nuclear Research. Dubna, 2014

In memory of Aldo Menzione

INTRODUCTION

The mass of the top quark (M_{top}) is a fundamental parameter of the Standard Model (SM), and its large value makes the top-quark contribution dominant in loop corrections to many observables, like the W -boson mass. The electroweak radiative corrections relate the top-quark mass and the W -boson mass to the mass of the Higgs boson [1]. Precision measurements of M_{top} are therefore important in assessing the internal consistency of the SM and of its extensions. In addition, it is important to measure M_{top} using independent data samples in different decay channels. Significant differences in the measurements of M_{top} in different decay channels could indicate contributions from new physics beyond the SM [2].

In the CDF Run II we study proton–antiproton collisions at a center-of-mass energy 1.96 TeV. The CDF detector [3] is a general-purpose apparatus designed to study $p\bar{p}$ collisions at the Tevatron. A charged-particle tracking system, consisting of a silicon microstrip tracker and a drift chamber, is immersed in a 1.4 T magnetic field. Electromagnetic and hadronic calorimeters surround the tracking system and measure particle energies. Drift chambers and scintillators, located outside the calorimeters, detect muon candidates. To describe the kinematics of reconstructed events, we use a cylindrical coordinate system, where θ is the polar angle with respect to the proton beam direction, ϕ is the azimuthal angle about the beam axis, and the pseudorapidity is defined as $\eta = -\ln \tan(\theta/2)$. We define transverse energy as $E_T = E \sin \theta$, where E is the energy measured in the calorimeter.

Top quarks are mostly produced in pairs ($t\bar{t}$) from quark–antiquark annihilations ($\sim 85\%$) or from gluon–gluon fusion. According to the SM, both top quarks decay almost exclusively as $t \rightarrow Wb$. The channels of $t(\bar{t})$ decay are classified according to the decay modes of the W boson. The *dilepton* channel, when both W decay to leptons (e, μ), gets only 5% of decays, but has the best signal-to-background ratio (S/B). Near 30% of decays go to the *lepton + jets* channel, with one W producing an electron or a muon, and the other decaying into a quark pair and producing jets. The *all-hadronic* decay channel collects 44% of events, but has a large QCD background.

This paper reports a measurement of the top-quark mass in the dilepton channel with full CDF Run II data set corresponding to 9.1 fb^{-1} of integrated

luminosity. This analysis updates the last CDF result [4] with additional data collected at the end of Run II and corresponding to integrated luminosity of about 3 fb^{-1} .

The top mass analysis in dilepton channel has new peculiarity with increased statistics: statistical error ceases to be the leading uncertainty of measurement. Main limitation in measurement arises from the systematic error, which is primarily due to the jet energy scale (JES) uncertainty. In contrast to lepton + jets and all-hadronic channels, in the dilepton events, there are no quarks originated from W -boson decay and as a consequence there is no a dijet mass constraint, which permits a precise calibration of the JES calorimeter. This fact requires directing the efforts towards searches of new possibilities for reduction of systematics uncertainty of the top mass due to the uncertainties in JES.

In our analysis we optimize the combined uncertainty ($stat \oplus JES$ error) that includes two main parts: statistical and jet energy scale uncertainties. Practically, we analyze the $t\bar{t}$ -dilepton events using two initial variables, which are sensitive to different kinematic properties of the $t\bar{t}$ system. The first variable is the reconstructed top mass (M_t^{reco}), which is calculated using all available experimental information and kinematic constraints and, as a result, is the most sensitive to the top-quark mass. In contrast to the reconstructed mass, our second variable is the most sensitive to the top-quark mass that we can build without using any jet energy information in the events. Therefore, this variable is insensitive to JES but still has some sensitivity to the top mass. The use of this variable allows reducing the systematic error due to the jet energy scale uncertainty. Finally, we define the «hybrid» variable for template analysis using the weighted sum of these two variables. The weight is chosen by requiring the smallest expected $stat \oplus JES$ error of measurement. This method allows us to reduce this error by about 12% with respect to when using only the reconstructed top mass as a mass-sensitive parameter.

1. DATA SAMPLE & EVENT SELECTION

The data were collected with an inclusive lepton trigger that require an electron with $E_T > 18 \text{ GeV}$ (or a muon with $P_T > 18 \text{ GeV}/c$) in the central region of the detector. The analyzed event sample was obtained with selection criteria developed for the $t\bar{t}$ -cross-section measurement in the dilepton channel [5]. In our analysis, we have introduced additional cuts to improve modeling and to reduce background.

Below we just list the basic selection requirements and refer for details to the above-cited note. We select events with two high- E_T leptons of opposite charge, one of which must be isolated. Here, we require $E_T > 20 \text{ GeV}$ for electrons or $P_T > 20 \text{ GeV}/c$ for muons. Missing transverse energy must be $\cancel{E}_T > 25 \text{ GeV}$ indicating the presence of neutrino. Z -veto cut eliminates ee and $\mu\mu$ events

with insufficient missing E_T significance and with dilepton invariant mass in the Z window. L cut designed to reject $Z \rightarrow \tau\tau$ and events with mismeasured \vec{E}_T requires that the angle between \vec{E}_T and the nearest jet is $\Delta\phi > 20^\circ$ if $E_T < 50$ GeV. Two (or more) jets with corrected $E_T > 15$ GeV and $|\eta| < 2.5$ are also required. The transverse energy sum, H_T , has to be more than 200 GeV, and the dilepton invariant mass has to be larger than 5 GeV.

Our additional cuts are presented below. We require the minimal distance in the η - ϕ space between any lepton and any jet in event, $\Delta R_{lj} = \sqrt{\Delta\eta_{lj}^2 + \Delta\phi_{lj}^2}$, to be more than 0.2. This cut reduces significantly the background of events with fake leptons. We reject events with M_t^{reco} more than 250 GeV. We introduce this requirement because simulation shows that the tail of the M_t^{reco} distribution contains mainly background events. The S/B ratio in the region $M_t^{\text{reco}} > 250$ GeV is expected to be about 1/3. We also tighten the cut on the dilepton invariant mass. We require the dilepton invariant mass to be larger than 10 GeV. This is done in order to reject events from processes which we do not model.

In total, we have 520 dilepton candidates after these selection requirements. The same cuts are applied to the Monte Carlo events generated for signal or background processes. The sensitivity of our measurement to the top-quark mass can be improved by analyzing separately events with beauty-flavored (b -tagged) jets. We divide the event sample into two mutually exclusive subsamples: b -tagged and nontagged ones. The first subsample contains events which have at least one tight SecVtx b -tagged jet [6]. The nontagged subsample contains events which have no tight SecVtx b -tagged jets and events for which the b -tagging algorithm cannot be applied. Table 1 gives the summary of expected contributions and observed events for the b -tagged and nontagged samples.

Table 1. Summary table of expected contributions and observed events in SecVtx b -tagged and nontagged dilepton data samples

CDF Run II Preliminary (9.1 fb^{-1})

$t\bar{t}$ -dilepton sample		
Source	Tagged events	0-tag events
WW	0.57 ± 0.15	16.4 ± 3.6
WZ	0.12 ± 0.03	5.2 ± 1.0
ZZ	0.20 ± 0.06	3.0 ± 0.5
DY/Z	4.4 ± 0.4	51.2 ± 8.0
Fakes	8.6 ± 2.7	21.4 ± 6.2
Total background	13.9 ± 2.8	97.2 ± 14.5
$t\bar{t}$ ($\sigma = 7.4 \text{ pb}$)	227.2 ± 16.2	173.2 ± 13.3
Total SM expectation	241.1 ± 16.4	270.3 ± 26.4
Observed	230	290

2. CALCULATING THE VARIABLE FOR TOP-QUARK MASS MEASUREMENT

2.1. «Hybrid» Variable’s Method. To measure the top-quark mass, typically, we can perform the template analysis using a variable sensitive to the top mass. The choice of this variable can be made by the requirement of minimal expected error of measurement. In our analysis, we optimize the combined error that includes two main terms: statistical and jet energy scale systematic errors. To achieve this optimization, we use two initial variables with different properties. The first variable is reconstructed mass (M_t^{reco}). We calculate it using a kinematic fit of dilepton events (see Subsec. 2.2), and it is the most sensitive variable to the true top mass. In contrast to M_t^{reco} , our second variable is the most sensitive to the top mass, which we can build without using information about jet energies. Therefore, this variable is insensitive to jet energy scale (JES) but it is less sensitive to the top mass if compared to M_t^{reco} . We denote it here as «alternative» mass, M_t^{alt} . Details of M_t^{alt} calculation can be found in Subsec. 2.3. As a next step, we define a «hybrid» variable using a weighted sum of these two variables. In the note we will denote this variable as an «effective» top mass (M_t^{eff}):

$$M_t^{\text{eff}} = w \cdot M_t^{\text{reco}} + (1 - w) \cdot M_t^{\text{alt}}, \quad (1)$$

where w is the weighting parameter. If we change w from 0 to 1, M_t^{eff} ’s properties are smoothly transforming from M_t^{alt} ’s to M_t^{reco} ’s properties. Therefore, we can choose w in our analysis by the requirement of the minimal expected $\text{stat} \oplus \text{JES}$ error of the measurement. The choice of the optimal value of $w = 0.7$ is discussed in Subsec. 2.4.

We choose the «hybrid» variable method as alternative to the best linear unbiased estimator (BLUE) method [7]. In contrast to BLUE, we do not need to combine correlated results because the template method framework automatically accounts for the usage of the right amount of information from both variables.

2.2. Calculating the Reconstructed Mass. The method implemented in this analysis for reconstructing the top-quark mass event by event is called the «Neutrino ϕ Weighting Method». This method was previously used for top-quark mass measurement on the lepton + track sample [8].

In contrast to the lepton + jets mode, for the dilepton case due to the existence of two neutrinos we have a nonconstrained kinematics. The number of independent variables is one more than the number of kinematic constraints: a total number of 24 unknown (b , \bar{b} , l^- , l^+ , ν and $\bar{\nu}$ 4-momenta) and only 23 equations to constrain the kinematics (measured 3-momenta for two b jets and two leptons, assumed known mass for 6 final particles, used two transverse components of calorimeter missing energy, constrained invariant mass for two W and assumed equal constrained mass of top and antitop quarks).

In order to constrain the kinematics, the scanning over the space of possibilities for the azimuthal angles of neutrinos ($\phi_{\nu_1}, \phi_{\nu_2}$) is used. A top-quark mass is reconstructed by minimizing a chi-squared function (χ^2) in the dilepton $t\bar{t}$ -event hypothesis. The χ^2 has two terms:

$$\chi^2 = \chi_{\text{reso}}^2 + \chi_{\text{constr}}^2. \quad (2)$$

The first term takes into account the detector uncertainties, whereas the second one constrains the parameters to the known physical quantities given their uncertainties. The first term is as follows:

$$\chi_{\text{reso}}^2 = \sum_{l=1}^2 \frac{(P_T^l - \tilde{P}_T^l)^2}{\sigma_{P_T}^l{}^2} - 2 \sum_{j=1}^2 \ln(\mathcal{P}_{\text{tf}}(\tilde{P}_T^j | P_T^j)) + \sum_{i=x,y} \frac{(UE^i - \tilde{UE}^i)^2}{\sigma_{UE}^2}. \quad (3)$$

With the use of the tilda ($\tilde{}$) we specify the parameters of the minimization procedure, whereas variables without tilda represent the measured values. \mathcal{P}_{tf} are the transfer functions between b quark and jets: they express the probability of measuring a jet transverse momentum P_T^j from a given b quark with transverse momentum \tilde{P}_T^j . The sum in the first term is over the two leptons in the event; the second sum loops over the two highest- E_T (leading) jets, which are assumed to originate from the b quarks. After candidate events are selected, leading jets momenta are further corrected for multiple interactions, underlying event, and out-of-cone energy loss.

The third sum runs over the transverse components of the unclustered energy (UE^x, UE^y), which is defined as the sum of the energy vectors from the towers not already associated with leptons or any leading jets.

The uncertainties (σ_{P_T}) on the lepton P_T used for electrons (e) and muons (μ) are calculated as [8]:

$$\frac{\sigma_{P_T}^e}{P_T^e} = \sqrt{\frac{0.135^2}{P_T^e[\text{GeV}/c]} + 0.02^2}, \quad (4)$$

$$\frac{\sigma_{P_T}^\mu}{P_T^\mu} = 0.0011 \cdot P_T^\mu[\text{GeV}/c]. \quad (5)$$

Uncertainty on the transverse components of the unclustered energy (σ_{UE}) is defined from phenomenological formula $\sigma_{UE} [\text{GeV}/c] = 0.4 \sqrt{\sum E_T^{\text{uncl}} [\text{GeV}/c]}$ [9], where $\sum E_T^{\text{uncl}}$ is the scalar sum of the transverse energy excluding two leptons and two leading jets.

The second term in Eq. (2), χ_{constr}^2 , constrains the parameters of the minimization procedure through the invariant masses of the lepton–neutrino and of

the lepton–neutrino-leading jet systems. This term is as follows:

$$\begin{aligned} \chi_{\text{constr}}^2 = & -2 \ln(\mathcal{P}_{\text{BW}}(m_{\text{inv}}^{l_1, \nu_1} | M_W, \Gamma_{M_W})) - 2 \ln(\mathcal{P}_{\text{BW}}(m_{\text{inv}}^{l_2, \nu_2} | M_W, \Gamma_{M_W})), \\ & - 2 \ln(\mathcal{P}_{\text{BW}}(m_{\text{inv}}^{l_1, \nu_1, j_1} | \tilde{M}_t, \Gamma_{\tilde{M}_t})) - 2 \ln(\mathcal{P}_{\text{BW}}(m_{\text{inv}}^{l_2, \nu_2, j_2} | \tilde{M}_t, \Gamma_{\tilde{M}_t})). \end{aligned} \quad (6)$$

\tilde{M}_t is the parameter giving the reconstructed top-quark mass. $\mathcal{P}_{\text{BW}}(m_{\text{inv}}; m, \Gamma) \equiv \frac{\Gamma^2 \cdot m^2}{(m_{\text{inv}}^2 - m^2)^2 + m^2 \Gamma^2}$ indicates the relativistic Breit–Wigner distribution function, which expresses the probability that an unstable particle of mass m and decay width Γ decays into a system of particles with invariant mass m_{inv} . We use the PDG values for M_W and Γ_{M_W} . For the top width we use the function

$$\Gamma_{M_t} = \frac{G_F}{8\sqrt{2}\pi} M_t^3 \left(1 - \frac{M_W^2}{M_t^2}\right)^2 \left(1 + 2\frac{M_W^2}{M_t^2}\right) \quad (7)$$

according to Ref. [10].

The longitudinal components of the neutrino momenta are free parameters of the minimization procedure, while the transverse components are related to \vec{E}_T and to the assumed $(\phi_{\nu_1}, \phi_{\nu_2})$ as follows:

$$\left\{ \begin{array}{l} P_x^{\nu_1} \equiv P_T^{\nu_1} \cdot \cos(\phi_{\nu_1}) = \frac{\vec{E}_{T_x} \cdot \sin(\phi_{\nu_2}) - \vec{E}_{T_y} \cdot \cos(\phi_{\nu_2})}{\sin(\phi_{\nu_2} - \phi_{\nu_1})} \cdot \cos(\phi_{\nu_1}), \\ P_y^{\nu_1} \equiv P_T^{\nu_1} \cdot \sin(\phi_{\nu_1}) = \frac{\vec{E}_{T_x} \cdot \sin(\phi_{\nu_2}) - \vec{E}_{T_y} \cdot \cos(\phi_{\nu_2})}{\sin(\phi_{\nu_2} - \phi_{\nu_1})} \cdot \sin(\phi_{\nu_1}), \\ P_x^{\nu_2} \equiv P_T^{\nu_2} \cdot \cos(\phi_{\nu_2}) = \frac{\vec{E}_{T_x} \cdot \sin(\phi_{\nu_1}) - \vec{E}_{T_y} \cdot \cos(\phi_{\nu_1})}{\sin(\phi_{\nu_1} - \phi_{\nu_2})} \cdot \cos(\phi_{\nu_2}), \\ P_y^{\nu_2} \equiv P_T^{\nu_2} \cdot \sin(\phi_{\nu_2}) = \frac{\vec{E}_{T_x} \cdot \sin(\phi_{\nu_1}) - \vec{E}_{T_y} \cdot \cos(\phi_{\nu_1})}{\sin(\phi_{\nu_1} - \phi_{\nu_2})} \cdot \sin(\phi_{\nu_2}). \end{array} \right. \quad (8)$$

The minimization procedure described above must be performed for all the allowed values of $\phi_{\nu_1}, \phi_{\nu_2}$ in the $(0, 2\pi) \times (0, 2\pi)$ region. Based on simulation, we choose a $\phi_{\nu_1}, \phi_{\nu_2}$ grid of 24×24 values as inputs for the minimization procedure. In building the grid we avoid the singular points at $\phi_{\nu_1} = \phi_{\nu_2} + k \cdot \pi$, where k is integer. Note from Eq. (8) that performing the transformation $\phi_\nu \rightarrow \phi_\nu + \pi$ leaves P_x^ν and P_y^ν unchanged, but reverses the sign of P_T^ν . We exclude unphysical solutions ($P_T^{\nu_1} < 0$ and/or $P_T^{\nu_2} < 0$) and choose the solution which leads to positive transverse momenta for both neutrinos. This decreases the number of grid points to 12×12 . At each point 8 solutions can exist, because of the two-fold ambiguity in the longitudinal momentum for each neutrino and of the ambiguity on the lepton–jet association. Therefore, for each event, we perform 1152 minimizations, each of which returns a value of M_{ijk}^{reco} and χ_{ijk}^2 ($i, j = 1, \dots, 12$; $k = 1, \dots, 8$). We define $\chi_{ij}^{\prime 2} = \chi_{ij}^2 + 4 \cdot \ln(\Gamma_{M_t})$,

which is obtained by using Eq. (6), where \mathcal{P}_{BW} is substituted with $\mathcal{P}'_{\text{BW}} \sim \frac{\Gamma \cdot m^2}{(m_{\text{inv}}^2 - m^2)^2 + m^2 \Gamma^2}$, and select the lowest χ'^2 solution for each point of the $(\phi_{\nu_1}, \phi_{\nu_2})$ grid, thereby reducing the number of the obtained masses to 144. Each mass is weighted according to the next formula:

$$w_{ij} = \frac{e^{-\chi'^2_{ij}/2}}{\sum_{i=1}^{12} \sum_{j=1}^{12} e^{-\chi'^2_{ij}/2}}. \quad (9)$$

In the next step, we build a mass distribution to define the most probable value (MPV). Masses below a threshold of 30% of the MPV bin content are discarded, and the remaining ones are averaged to compute the preferred top-quark mass for the event (M_t^{reco}).

2.3. Defining M_t^{alt} such as to Be Insensitive to the Jet Energy Scale. In order to define M_t^{alt} as a variable insensitive to JES, we should not use the values of the jet energies. In this case, we can exploit only the following experimental information: 4-momenta (l_1 and l_2) of the lepton and jet's directions. We can specify the track direction of particle p by the following vector c_p :

$$c_p \equiv (c_x^p, c_y^p, c_z^p), \quad (10)$$

where c_x^p , c_y^p , and c_z^p are the direction cosines of the particle momentum. Also, to operate within a 4-vector scheme, we define the following 4-vector c_p :

$$c_p \equiv (1, \mathbf{c}_p). \quad (11)$$

We can interpret c_p as 4-momenta of massless particle with energy 1 GeV that has the same flight direction as p in the laboratory coordinate system. And, using c_p , we can write our definition for M_t^{alt} by formula

$$M_t^{\text{alt}} \equiv \sqrt{\langle l_1, c_{b_1} \rangle \cdot \langle l_2, c_{b_2} \rangle} + 120 \text{ GeV}, \quad (12)$$

where c_{b_1} and c_{b_2} are 4-vectors giving the directions of the two leading jets. In formula (12), a notation like $\langle l, c \rangle$ means the scalar product of two 4-vectors l and c . From the two possible assignments of the leptons and jets in (12), we choose one with maximal value of $\langle c_{l_1}, c_{b_1} \rangle + \langle c_{l_2}, c_{b_2} \rangle$ (notation $\langle c_1, c_2 \rangle$ means the scalar product of two three-dimensional vectors c_1 and c_2). This criterion was chosen based on simulation and it gives correct assignment in about 60% of simulated $t\bar{t}$ events. For the variable M_t^{alt} in (12), we apply an additional shift of 120 GeV, which has no impact on our analysis. It is introduced for convenience to equalize the x -axes of the M_t^{reco} and M_t^{alt} variables in (1).

Equation (12) is equivalent to the next formula:

$$M_t^{\text{alt}} \equiv \sqrt{\frac{\langle l_1, b_1 \rangle \cdot \langle l_2, b_2 \rangle}{E_{b_1} E_{b_2}}} + 120 \text{ GeV}, \quad (13)$$

where b_1 and b_2 are 4-momenta of two leading jets, which are defined as massless particles with energies E_{b_1} and E_{b_2} .

2.4. Optimizing the Measurement. In order to find the optimal value of w (Eq. (1)), we scan the interval of $[0, 1]$ with a step of 0.05. In every point of the scan, we build the signal and background templates for M_t^{eff} , define the likelihood function and perform pseudo-experiments (PE's). The PE's procedure is described in the next sections of this note. We perform a check that the top mass and its error returned from PE's are correct by examining the PE's pulls. We define the expected statistical error as the mean of the error distribution obtained from PE's with $M_{\text{top}} = 172.5 \text{ GeV}$. We define the JES systematic error by applying the JES shifts according to the nominal procedure approved for the JES systematic uncertainty. The obtained results are presented in Fig. 1.

It can be seen that the JES systematic error does not equal to zero at $w = 0$ as we would expect. This is an effect from events entering or exiting the candidate sample after shift in JES due to the fact that JES is involved in the event selection. Instead of this, the JES systematics equals to zero with $w \cong 0.12$. At this point, we have a full compensation of two effects related to JES: the change in template due to the changing variable is compensated by the change in template due to events entering or exiting the candidate sample. These two mechanisms have different signs and impact differently on the top mass measurement. Typically,

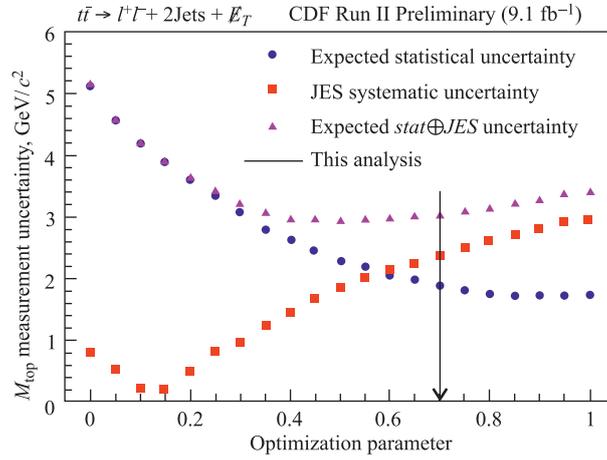


Fig. 1. The statistical, JES systematic and $\text{stat} \oplus \text{JES}$ uncertainties of top-quark mass measurement as a function of parameter w in Eq. (1)

the first effect is bigger and is compensated partially by the second one. In the region when we choose smaller values of the w parameter, the second effect can compensate or even overpower the first one.

From Fig. 1, we can conclude that the minimal value of the $stat \oplus JES$ uncertainty corresponds to a value of $w = 0.5$. In the analysis, we have decided to use a value of $w = 0.7$. At $w = 0.7$, the $stat \oplus JES$ uncertainty has unessential difference from minimum of about 2%, while the expected statistical uncertainty is notably better. As a general rule, we prefer to avoid increasing of the statistical uncertainty because it can increase also systematic uncertainties from the sources described in Sec. 4. By selecting $w = 0.7$, the expected $stat \oplus JES$ uncertainty is reduced by 12% in comparison with the case of choosing $w = 1$, i. e., exploiting only M_t^{reco} in the mass measurement.

3. TOP-QUARK MASS DETERMINATION

3.1. Templates. The selected data sample is a mixture of signal and background events. In order to extract the top-quark mass, the M_t^{eff} distribution (see Eq. (1)) in data is compared with probability density functions (p.d.f.'s) for signal and background by means of a likelihood fit. The p.d.f.'s are defined as the parametrizations of templates.

Signal templates are built from $t\bar{t}$ samples generated with Pythia for top-quark masses in the range from 160 to 185 GeV/ c^2 in 1 GeV/ c^2 steps. They are parameterized separately for b -tagged and nontagged events in a global fit by using a combination of two Landau and one Gaussian distribution functions, as

$$\begin{aligned}
P_s(M_t^{\text{eff}}|M_t) = & \frac{p_8 p_7}{\sqrt{2\pi p_1}} e^{-0.5(\frac{M_t^{\text{eff}} - p_2}{p_1} + \exp(-\frac{M_t^{\text{eff}} - p_2}{p_1}))} + \\
& + \frac{p_8(1 - p_7)}{\sqrt{2\pi p_3}} e^{-0.5(\frac{M_t^{\text{eff}} - p_4}{p_3})^2} + \\
& + \frac{(1 - p_8)}{\sqrt{2\pi p_5}} e^{-0.5(\frac{M_t^{\text{eff}} - p_6}{p_5} + \exp(-\frac{M_t^{\text{eff}} - p_6}{p_5}))}. \quad (14)
\end{aligned}$$

The p.d.f.'s for the signal, P_s^{tag} and P_s^{notag} , express the probability that M_t^{eff} is indicated by an event with true top-quark mass M_t . The parameters p_1, \dots, p_8 depend on the true top-quark mass M_t . These parameters are calculated as

$$p_k = \alpha_k + \alpha_{k+8} \cdot (M_t[\text{GeV}/c^2] - 175), \quad k = 1, \dots, 8. \quad (15)$$

The parameters α_k are defined by the signal templates fit.

Representative background templates are built separately for b -tagged and nontagged events by adding diboson, fakes, and Drell–Yan templates. These

templates have been normalized to the expected rates reported in Table 1. The fakes template is built from W +jets data events by weighting each event according to the probability for a jet to be misidentified as a lepton (fake rate). Drell–Yan and diboson templates are built from simulated samples. The combined background templates are fitted with a sum of two Landau and one Gaussian distribution functions, as

$$\begin{aligned}
P_b(M_t^{\text{eff}}) = & \frac{\beta_8\beta_7}{\sqrt{2\pi}\beta_1} e^{-0.5(\frac{M_t^{\text{eff}}-\beta_2}{\beta_1} + \exp(-\frac{M_t^{\text{eff}}-\beta_2}{\beta_1}))} + \\
& + \frac{\beta_8(1-\beta_7)}{\sqrt{2\pi}\beta_3} e^{-0.5(\frac{M_t^{\text{eff}}-\beta_4}{\beta_3})^2} + \\
& + \frac{(1-\beta_8)}{\sqrt{2\pi}\beta_5} e^{-0.5(\frac{M_t^{\text{eff}}-\beta_6}{\beta_5} + \exp(-\frac{M_t^{\text{eff}}-\beta_6}{\beta_5}))}, \quad (16)
\end{aligned}$$

where the fitted parameters $\beta_1 \cdots \beta_8$ are M_t -independent.

3.2. Likelihood Form. The top mass is extracted from the data sample by performing an unbinned likelihood fit. We define the likelihood function as a product of independent likelihood functions obtained for b -tagged and nontagged subsamples

$$\mathcal{L}^{\text{total}} = \mathcal{L}^{\text{tag}} \cdot \mathcal{L}^{\text{notag}}. \quad (17)$$

The likelihood functions, \mathcal{L}^{tag} and $\mathcal{L}^{\text{notag}}$, express the probability that a top-mass distribution from data is described by a mixture of background events and $t\bar{t}$ -dilepton events with an assumed top-quark mass. Inputs for the likelihood are the values of top mass from data events, the signal and background p.d.f.'s and the expected background. The background expectations and its errors are taken from Table 1. The likelihoods, \mathcal{L}^{tag} and $\mathcal{L}^{\text{notag}}$, have the same form

$$\mathcal{L} = \mathcal{L}_{\text{shape}} \cdot \mathcal{L}_{\text{backgr}}, \quad (18)$$

where

$$\mathcal{L}_{\text{shape}} = \frac{e^{-(n_s+n_b)} \cdot (n_s+n_b)^N}{N!} \cdot \prod_{n=1}^N \frac{n_s \cdot P_s(M_t^{\text{eff}}|M_{\text{top}}) + n_b \cdot P_b(M_t^{\text{eff}})}{n_s+n_b}, \quad (19)$$

$$\mathcal{L}_{\text{backgr}} = \exp\left(\frac{-(n_b - n_b^{\text{exp}})^2}{2\sigma_{n_b^{\text{exp}}}^2}\right). \quad (20)$$

The shape likelihood term, $\mathcal{L}_{\text{shape}}$ (Eq. (19)), expresses the probability of an event being signal with a top mass of M_{top} or background. The signal (P_s) and background (P_b) probabilities are weighted according to the number of signal (n_s)

and background (n_b) events, which are floated in the likelihood fit. In the fitting procedure, n_b is constrained to be Gaussian-distributed with mean value n_b^{exp} and standard deviation $\sigma_{n_b^{\text{exp}}}$, as shown by Eq. (20), while $(n_s + n_b)$ is the mean of a Poisson distribution of N selected events.

We perform the likelihood fit using the MINUIT [11] program. The fit returns an estimated top-quark mass (M_t^{fit}) and estimated numbers of signal ($n_s^{\text{tag fit}}$ and $n_s^{\text{notag fit}}$) and background events ($n_b^{\text{tag fit}}$ and $n_b^{\text{notag fit}}$). M_t^{fit} returned by the likelihood fit is the mass corresponding to the minimum of the $[-\ln \mathcal{L}^{\text{total}}]$ function. The positive and negative statistical uncertainties (σ^+ and σ^-) are the difference between M_t^{fit} and the mass values at $[-\ln(\mathcal{L})]_{\text{min}} + 0.5$. The positive and negative statistical uncertainties are returned by MINOS [11]. The final result is presented with symmetrized statistical error: $\sigma = (\sigma^+ + |\sigma^-|)/2$.

3.3. Bias Checks. We checked whether the fit with likelihood form (17) is able to return the correct mass. Checks are performed by running a large number of pseudo-experiments (PE's) on simulated background and signal events where the true top-quark mass is known. Each PE consists of determining the number of signal (N_s^{PE}) and background (N_b^{PE}) events in the sample, drawing N_s^{PE} masses from a signal template and N_b^{PE} from the background template, and likelihood fitting, as described in Subsec. 3.2. A top-quark mass (M_t^{fit}) and its symmetrized statistical error are returned by the fit. Numbers of signal and background events are generated according to Poisson distributions with means given in Table 1.

For each input top-quark mass the median of the M_t^{fit} distribution is chosen as the top-quark mass estimate (M_t^{out}). The bias, defined as $M_t^{\text{out}} - M_t$, is shown in Fig. 2. The error bars are determined by the limited statistics of the signal templates. The average bias (horizontal line in the plot) is consistent with zero. Therefore, the obtained top-quark mass M_t^{fit} can be considered as unbiased estimate of the true top mass and we do not apply any additional corrections for it.

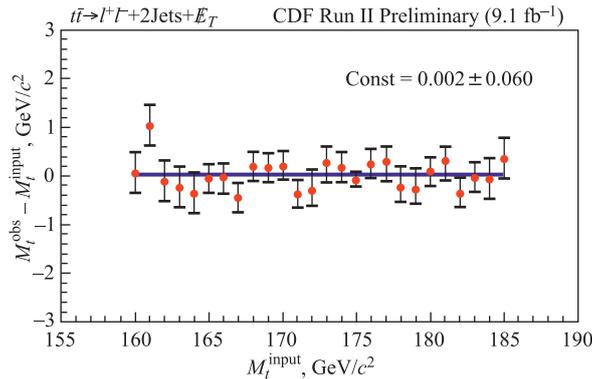


Fig. 2. Results from pseudo-experiments: the bias vs input masses

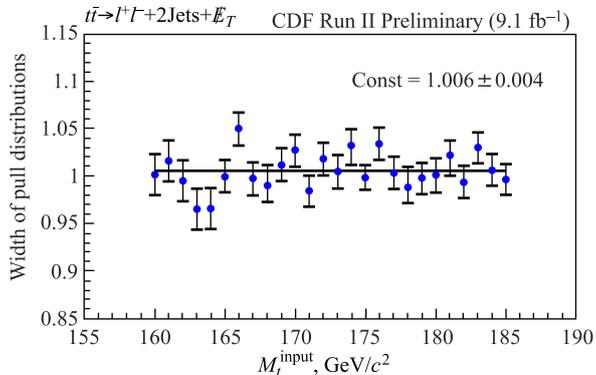


Fig. 3. Results from pseudo-experiments: width of pull distributions vs generated top-quark mass

In order to check the bias on the statistical error, we use pulls. Pulls are defined as follows:

$$\frac{M_t^{\text{fit}} - M_t}{\sigma}, \quad (21)$$

where $\sigma = (\sigma^+ + |\sigma^-|)/2$. For each generated top-quark mass, pull distributions are fitted by using Gaussian functions. The width of pull distributions versus generated top-quark mass is shown in Fig. 3. Error bars account for the limited statistics of signal templates. The average width of pull distributions can be considered compatible with one within uncertainties. Accordingly, there is no need to rescale the statistical uncertainty obtained from data.

4. SYSTEMATIC UNCERTAINTIES

Since our method compares findings to expectations estimated from Monte Carlo simulations, uncertainties in our models used to generate events cause systematic uncertainties. We calculate the contribution to the systematics from each source of uncertainty. The generic procedure for estimating a systematic uncertainty is as follows. The parameters used for the generation of events are modified by ± 1 standard deviation in their uncertainties and new templates are built. PE's from the modified templates are performed using the same p.d.f.'s as in the analysis. The difference between the median of the top-quark mass distribution from PE's and the nominal top-quark mass is used as the estimate of the systematic uncertainty.

The largest contribution comes from the uncertainty in the jet energy measurement, which includes uncertainties due to the following effects: non-uniformity in calorimeter response as a function of $|\eta|$, multiple $p\bar{p}$ collisions, hadronic jet

energy scale, underlying events, and out-of-cone energy lost in the clustering procedure.

Difference between data and MC luminosity profile is accounted by rescaling the top mass dependence on the number of interactions in the event by the difference in the number of interactions between data and MC.

The initial and final state radiation (IFSR) uncertainties are estimated using the Pythia Monte Carlo samples, in which the QCD parameters for parton shower evolution in the initial and final states are varied simultaneously. The amount of variation is based on the CDF studies of Drell–Yan data.

The uncertainty in reconstructing the top-quark mass due to the use of a particular parton distribution function (PDF) comes from three sources: PDF parametrization, PDF choice, and QCD scale (Λ_{QCD}).

The effect of the choice of a particular Monte Carlo generator is studied by comparing our default Pythia generator to Herwig. Also, we estimate the systematics due to the NLO effects by the comparison of the Pythia and Powheg generators.

In order to estimate the systematic uncertainty for the background composition, expected rates for fake, diboson, and Drell–Yan events are alternatively varied by ± 1 standard deviation without changing the total number of expected background events. We also studied the effect from changing the shape of the main background contributors: Drell–Yan and «fakes».

The uncertainty for the b -jet scale due to the heavy-quark fragmentation, semileptonic b -jet branching ratio, and b -jet calorimeter response is also taken into account. The effect of the fragmentation model on the top-quark mass is evaluated by reweighting events according to two different fragmentation models, while effects of the uncertainties on the semileptonic b -jet branching ratio (BR) and b -jet energy calorimeter response are estimated by shifting the BR and the b -jet energy scale.

The effect on the top mass from the uncertainty on lepton energy scale is studied by applying $\pm 1\%$ shifts to the P_T lepton.

The effect of color reconnection (CR) on our result is studied using the Pythia 6.4 MC generator, which includes CR effects.

Since Pythia is a leading-order MC generator, the number of $t\bar{t}$ events from gluon fusion in Pythia samples is approximately 6%. In case of the NLO framework, the gluon fusion fraction is expected to be $(15 \pm 5)\%$. We take into account the uncertainty in top mass due to this effect.

Also, the effect on the top mass from the uncertainty in b -tagging modeling is studied.

The source of each systematic uncertainty is assumed to be uncorrelated to the other ones, so that the overall systematic error is obtained by adding in quadrature the individual uncertainties. The systematic uncertainties along with

Table 2. Summary of systematic uncertainties on the top-quark mass measurement

CDF Run II Preliminary (9.1 fb ⁻¹)	
M_{top} Measurement in the $t\bar{t}$ -Dilepton Final State	
Source	Uncertainty, GeV/ c^2
Jet energy scale	2.42
NLO effects	0.64
Monte Carlo generators	0.49
Lepton energy scale	0.36
b -jet energy scale	0.34
Initial and final state radiation	0.33
Background modeling	0.33
Luminosity profile (pileup)	0.30
Color reconnection	0.24
gg fraction	0.24
Parton distribution functions	0.21
MC statistics	0.19
b -tagging	0.05
Total systematic	2.69
Statistical	1.83
Total	3.25

the total uncertainty are summarized in Table 2. The total systematic uncertainty is estimated as 2.69 GeV/ c^2 .

5. RESULTS

The background-constrained likelihood fit described in Sec. 3 is performed on the dilepton data sample and returns $M_{\text{top}} = (170.80 \pm 1.83)$ GeV/ c^2 . The

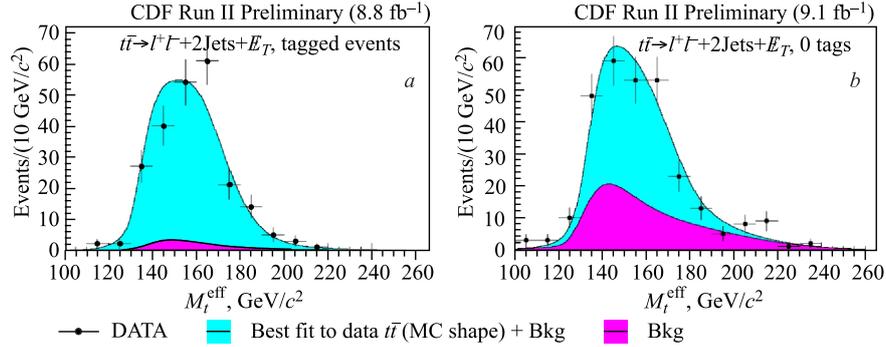


Fig. 4. Likelihood fit to the dilepton data sample. Background (purple solid) and signal + background (cyan solid) p.d.f.'s, normalized according to the numbers returned by the fitter, are superimposed on the top mass distribution from data (histogram). Plots a and b are for b -tagged and nontagged subsamples, respectively

experimental top mass distributions for b -tagged and nontagged subsamples are shown in Fig. 4. The fitted mass-dependent negative log-likelihood function from the likelihood fit to the dilepton data sample is presented in Fig. 5. Also, the post-fit plots for our initial variables (M_t^{reco} and M_t^{alt}) are presented in Figs. 6 and 7.

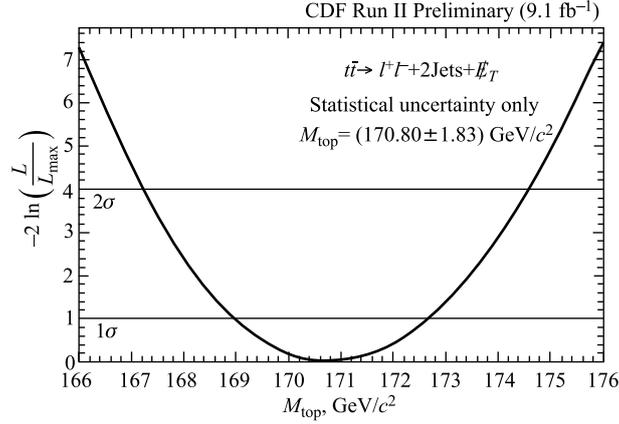


Fig. 5. The fitted mass-dependent negative log-likelihood function from the likelihood fit to the dilepton data sample

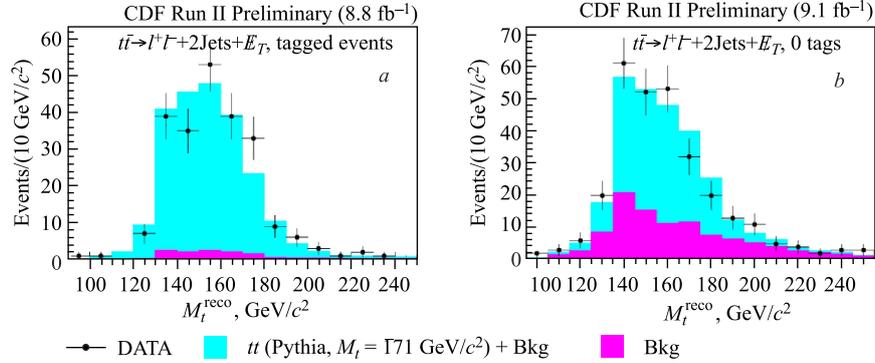


Fig. 6. Background (purple solid) and signal+background (cyan solid) templates for reconstructed top-quark mass, normalized according to the numbers returned by the fitter, are superimposed on the reconstructed mass distribution from data (histogram). Plots a and b are for b -tagged and nontagged subsamples, respectively. The value of $171 \text{ GeV}/c^2$ is assumed for the top-quark mass (close to the central value of the data fit)

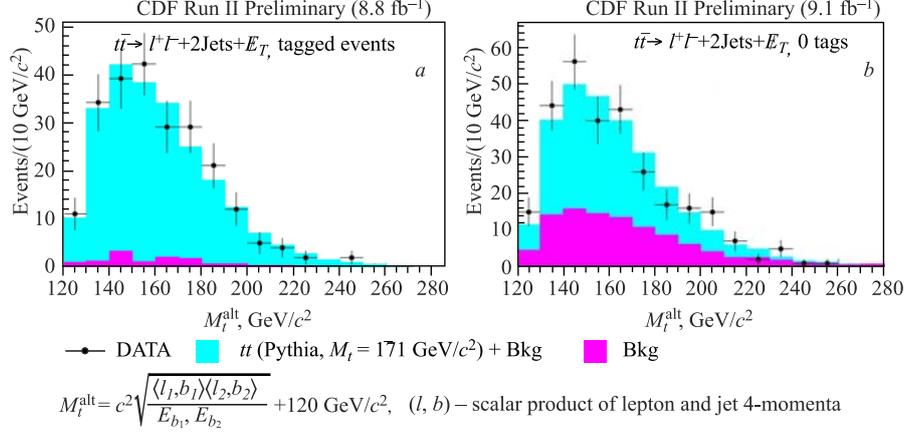


Fig. 7. Background (purple solid) and signal+background (cyan solid) templates for M_t^{alt} variable, normalized according to the numbers returned by the fitter, are superimposed on the M_t^{alt} variable distribution from data (histogram). Plots *a* and *b* are for *b*-tagged and nontagged subsamples, respectively. The value of 171 GeV/ c^2 is assumed for the top-quark mass (close to the central value of the data fit)

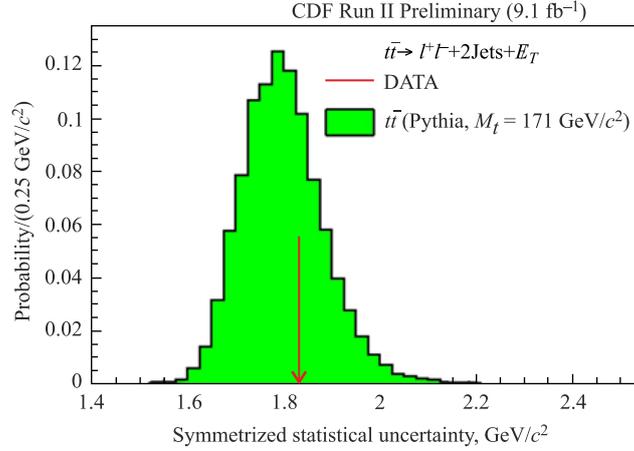


Fig. 8. Symmetrized statistical errors from pseudo-experiments generated with a top-quark mass of 171 GeV/ c^2 . The arrow indicates the error found in measurement

In order to check that the measured statistical error is reasonable, a set of PE's is performed on simulated background and signal events with $M_t = 171 \text{ GeV}/c^2$ (close to the central value of the constrained fit). The obtained symmetrized error distribution along with the observed value (represented by the arrow) are shown in Fig. 8. We estimate that the probability for obtaining a precision better

than that found in this experiment (p-value) is 68%. This value is obtained by comparing the measured statistical uncertainty with that expected from PE's.

CONCLUSION

Using the full CDF data set we measure on a dilepton sample a top-quark mass of

$$\begin{aligned} M_{\text{top}} &= (170.80 \pm 1.83 \text{ (stat.)} \pm 2.69 \text{ (syst.)}) \text{ GeV}/c^2 \\ &\text{or} \\ M_{\text{top}} &= (170.80 \pm 3.25) \text{ GeV}/c^2. \end{aligned} \tag{22}$$

This result is compatible with the Tevatron average top-quark mass ($M_{\text{top}} = (173.20 \pm 0.87) \text{ GeV}/c^2$ [12]), obtained by combining the main CDF and DØ Run I and Run II results. Compared with the last CDF result in this channel ($M_{\text{top}} = (170.3 \pm 3.7) \text{ GeV}/c^2$ [4]) an improvement of about 12% in the total error has been achieved.

Acknowledgements. We thank the Fermilab staff and the technical staff 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 Comisión Interministerial de Ciencia y Tecnología, Spain; the European Community's Human Potential Programme; the Slovak R&D Agency; and the Academy of Finland.

REFERENCES

1. *ALEPH, CDF, D0, DELPHI, L3, OPAL, SLD Collaborations, the LEP Electroweak Working Group, the Tevatron Electroweak Working Group, the SLD Electroweak and Heavy Flavour Groups.* Precision Electroweak Measurements and Constraints on the Standard Model. arXiv:1012.2367.
2. Kane G. L., Mrenna S. // *Phys. Rev. Lett.* 1996. V. 77. P. 3502.
3. Acosta D. et al. (*CDF Collab.*) // *Phys. Rev. D.* 2005. V. 71. P. 032001.
4. Aaltonen T. et al. (*CDF Collab.*) // *Phys. Rev. D.* 2011. V. 83. P. 111101.

5. Aaltonen T. *et al.* (CDF Collab.) // Phys. Rev. Lett. 2013. V. 88. P. 091103.
6. Acosta D. *et al.* (CDF Collab.) // Phys. Rev. D. 2005. V. 71. P. 052003.
7. Lyons L., Gibaut D., Clifford P. // Nucl. Instr. Meth. A. 1988. V. 270. P. 110;
Valassi A. // Nucl. Instr. Meth. A. 2003. V. 500. P. 391.
8. Aaltonen T. *et al.* (CDF Collab.) // Phys. Rev. D. 2009. V. 79. P. 072005.
9. Abulencia A. *et al.* (CDF Collab.) // Phys. Rev. D. 2006. V. 73. P. 032003.
10. Beneke M. *et al.* Top-Quark Physics. arXiv:hep-ph/0003033.
11. James F. MINUIT: Function Minimization and Error Analysis. CERN Program Library Long Writeup D506.
12. The Tevatron Electroweak Working Group (CDF and D0 Collab.). Combination of CDF and DØ Results on the Mass of the Top Quark Using up to 8.7 fb^{-1} at the Tevatron. arXiv:1305.3929.

Received on March 12, 2014.

Редактор *В. В. Булатова*

Подписано в печать 03.07.2014.

Формат 60 × 90/16. Бумага офсетная. Печать офсетная.

Усл. печ. л. 1,3. Уч.-изд. л. 1,8. Тираж 285 экз. Заказ № 58289.

Издательский отдел Объединенного института ядерных исследований
141980, г. Дубна, Московская обл., ул. Жолио-Кюри, 6.

E-mail: publish@jinr.ru

www.jinr.ru/publish/