FERMILAB-Pub-98-261-E

Measurement of the Top Quark Mass in the Dilepton Channel

The DØ Collaboration*
Fermi National Accelerator Laboratory, Batavia, Illinois 60510
(August 26, 1998)

Abstract

We report a measurement of the top quark mass using six candidate events for the process $p\overline{p} \rightarrow t\overline{t} + X \rightarrow \ell^{+}\nu b\ell^{-}\overline{\nu}\overline{b} + X$, observed in the DØ experiment at the Fermilab $p\overline{p}$ collider. Using maximum likelihood fits to the dynamics of the decays, we measure a mass for the top quark of $m_t = 168.4 \pm 12.3$ (stat) $\pm$ 3.6 (syst) GeV. We combine this result with our previous measurement in the $t\overline{t} \rightarrow \ell +$ jets channel to obtain $m_t = 172.1 \pm 7.1$ GeV as the best value of the mass of the top quark measured by DØ.

Typeset using REVTEX

*Authors listed on the following page.

Submitted to Physical Review D.

B. Abbott,[40] M. Abolins,[37] V. Abramov,[15] B.S. Acharya,[8] I. Adam,[39] D.L. Adams,[48] M. Adams,[24] S. Ahn,[23] H. Aihara,[17] G.A. Alves,[2] N. Amos,[36] E.W. Anderson,[30] R. Astur,[42] M.M. Baarmand,[42] V.V. Babintsev,[15] L. Babukhadia,[16] A. Baden,[33] V. Balamurali,[28] B. Baldin,[23] S. Banerjee,[8] J. Bantly,[45] E. Barberis,[17] P. Baringer,[31] J.F. Bartlett,[23] A. Belyaev,[14] S.B. Beri,[6] I. Bertram,[26] V.A. Bezzubov,[15] P.C. Bhat,[23] V. Bhatnagar,[6] M. Bhattacharjee,[42] N. Biswas,[28] G. Blazey,[25] S. Blessing,[21] P. Bloom,[18] A. Boehnlein,[23] N.I. Bojko,[15] F. Borcherding,[23] C. Boswell,[20] A. Brandt,[23] R. Breedon,[18] R. Brock,[37] A. Bross,[23] D. Buchholz,[26] V.S. Burtovoi,[15] J.M. Butler,[34] W. Carvalho,[2] D. Casey,[37] Z. Casilum,[42] H. Castilla-Valdez,[11] D. Chakraborty,[42] S.-M. Chang,[35] S.V. Chekulaev,[15] L.-P. Chen,[17] W. Chen,[42] S. Choi,[10] S. Chopra,[36] B.C. Choudhary,[20] J.H. Christenson,[23] M. Chung,[24] D. Claes,[38] A.R. Clark,[17] W.G. Cobau,[33] J. Cochran,[20] L. Coney,[28] W.E. Cooper,[23] C. Cretsinger,[41] D. Cullen-Vidal,[45] M.A.C. Cummings,[25] D. Cutts,[45] O.I. Dahl,[17] K. Davis,[16] K. De,[46] K. Del Signore,[36] M. Demarteau,[23] D. Denisov,[23] S.P. Denisov,[15] H.T. Diehl,[23] M. Diesburg,[23] G. Di Loreto,[37] P. Draper,[46] Y. Ducros,[5] L.V. Dudko,[14] S.R. Dugad,[8] A. Dyshkant,[15] D. Edmunds,[37] J. Ellison,[20] V.D. Elvira,[42] R. Engelmann,[42] S. Eno,[33] G. Eppley,[48] P. Ermolov,[14] O.V. Eroshin,[15] V.N. Evdokimov,[15] T. Fahland,[19] M.K. Fatyga,[41] S. Feher,[23] D. Fein,[16] T. Ferbel,[41] G. Finocchiaro,[42] H.E. Fisk,[23] Y. Fisyak,[43] E. Flattum,[23] G.E. Forden,[16] M. Fortner,[25] K.C. Frame,[37] S. Fuess,[23] E. Gallas,[46] A.N. Galyaev,[15] P. Gartung,[20] V. Gavrilov,[13] T.L. Geld,[37] R.J. Genik II,[37] K. Genser,[23] C.E. Gerber,[23] Y. Gershtein,[13] B. Gibbard,[43] B. Gobbi,[26] B. Gómez,[4] G. Gómez,[33] P.I. Goncharov,[15] J.L. González Solís,[11] H. Gordon,[43] L.T. Goss,[47] K. Gounder,[20] A. Goussiou,[42] N. Graf,[43] P.D. Grannis,[42] D.R. Green,[23] H. Greenlee,[23] S. Grinstein,[1] P. Grudberg,[17] S. Grünendahl,[23] G. Guglielmo,[44] J.A. Guida,[16] J.M. Guida,[45] A. Gupta,[8] S.N. Gurzhiev,[15] G. Gutierrez,[23] P. Gutierrez,[44] N.J. Hadley,[33] H. Haggerty,[23] S. Hagopian,[21] V. Hagopian,[21] K.S. Hahn,[41] R.E. Hall,[19] P. Hanlet,[35] S. Hansen,[23] J.M. Hauptman,[30] D. Hedin,[25] A.P. Heinson,[20] U. Heintz,[23] R. Hernández-Montoya,[11] T. Heuring,[21] R. Hirosky,[24] J.D. Hobbs,[42] B. Hoeneisen,[4,*] J.S. Hoftun,[45] F. Hsieh,[36] Ting Hu,[42] Tong Hu,[27] T. Huehn,[20] A.S. Ito,[23] E. James,[16] J. Jaques,[28] S.A. Jerger,[37] R. Jesik,[27] T. Joffe-Minor,[26] K. Johns,[16] M. Johnson,[23] A. Jonckheere,[23] M. Jones,[22] H. Jöstlein,[23] S.Y. Jun,[26] C.K. Jung,[42] S. Kahn,[43] G. Kalbfleisch,[44] D. Karmanov,[14] D. Karmgard,[21] R. Kehoe,[28] M.L. Kelly,[28] S.K. Kim,[10] B. Klima,[23] C. Klopfenstein,[18] W. Ko,[18] J.M. Kohli,[6] D. Koltick,[29] A.V. Kostritskiy,[15] J. Kotcher,[43] A.V. Kotwal,[39] A.V. Kozelov,[15] E.A. Kozlovsky,[15] J. Krane,[38] M.R. Krishnaswamy,[8] S. Krzywdzinski,[23] S. Kuleshov,[13] S. Kunori,[33] F. Landry,[37] G. Landsberg,[45] B. Lauer,[30] A. Leflat,[14] J. Li,[46] Q.Z. Li-Demarteau,[23] J.G.R. Lima,[3] D. Lincoln,[23] S.L. Linn,[21] J. Linnemann,[37] R. Lipton,[23] F. Lobkowicz,[41] S.C. Loken,[17] A. Lucotte,[42] L. Lueking,[23] A.L. Lyon,[33] A.K.A. Maciel,[2] R.J. Madaras,[17] R. Madden,[21] L. Magaña-Mendoza,[11] V. Manankov,[14] S. Mani,[18] H.S. Mao,[23,†] R. Markeloff,[25] T. Marshall,[27] M.I. Martin,[23] K.M. Mauritz,[30] B. May,[26] A.A. Mayorov,[15] R. McCarthy,[42] J. McDonald,[21] T. McKibben,[24] J. McKinley,[37] T. McMahon,[44] H.L. Melanson,[23] M. Merkin,[14] K.W. Merritt,[23] C. Miao,[45] H. Miettinen,[48] A. Mincer,[40] C.S. Mishra,[23] N. Mokhov,[23] N.K. Mondal,[8] H.E. Montgomery,[23] P. Mooney,[4] M. Mostafa,[1] H. da Motta,[2] C. Murphy,[24] F. Nang,[16] M. Narain,[23] V.S. Narasimham,[8] A. Narayanan,[16] H.A. Neal,[36] J.P. Negret,[4] P. Nemethy,[40] D. Norman,[47] L. Oesch,[36] V. Oguri,[3] E. Oliveira,[2] E. Oltman,[17] N. Oshima,[23] D. Owen,[37] P. Padley,[48] A. Para,[23] Y.M. Park,[9] R. Partridge,[45] N. Parua,[8] M. Paterno,[41] B. Pawlik,[12] J. Perkins,[46]

M. Peters,[22] R. Piegaia,[1] H. Piekarz,[21] Y. Pischalnikov,[29] B.G. Pope,[37] H.B. Prosper,[21]
S. Protopopescu,[43] J. Qian,[36] P.Z. Quintas,[23] R. Raja,[23] S. Rajagopalan,[43] O. Ramirez,[24]
S. Reucroft,[35] M. Rijssenbeek,[42] T. Rockwell,[37] M. Roco,[23] P. Rubinov,[26] R. Ruchti,[28]
J. Rutherfoord,[16] A. Sánchez-Hernández,[11] A. Santoro,[2] L. Sawyer,[32] R.D. Schamberger,[42]
H. Schellman,[26] J. Sculli,[40] E. Shabalina,[14] C. Shaffer,[21] H.C. Shankar,[8] R.K. Shivpuri,[7]
M. Shupe,[16] H. Singh,[20] J.B. Singh,[6] V. Sirotenko,[25] E. Smith,[44] R.P. Smith,[23] R. Snihur,[26]
G.R. Snow,[38] J. Snow,[44] S. Snyder,[43] J. Solomon,[24] M. Sosebee,[46] N. Sotnikova,[14]
M. Souza,[2] A.L. Spadafora,[17] G. Steinbrück,[44] R.W. Stephens,[46] M.L. Stevenson,[17]
D. Stewart,[36] F. Stichelbaut,[42] D. Stoker,[19] V. Stolin,[13] D.A. Stoyanova,[15] M. Strauss,[44]
K. Streets,[40] M. Strovink,[17] A. Sznajder,[2] P. Tamburello,[33] J. Tarazi,[19] M. Tartaglia,[23]
T.L.T. Thomas,[26] J. Thompson,[33] T.G. Trippe,[17] P.M. Tuts,[39] V. Vaniev,[15] N. Varelas,[24]
E.W. Varnes,[17] D. Vititoe,[16] A.A. Volkov,[15] A.P. Vorobiev,[15] H.D. Wahl,[21] G. Wang,[21]
J. Warchol,[28] G. Watts,[45] M. Wayne,[28] H. Weerts,[37] A. White,[46] J.T. White,[47]
J.A. Wightman,[30] S. Willis,[25] S.J. Wimpenny,[20] J.V.D. Wirjawan,[47] J. Womersley,[23]
E. Won,[41] D.R. Wood,[35] Z. Wu,[23,†] H. Xu,[45] R. Yamada,[23] P. Yamin,[43] T. Yasuda,[35]
P. Yepes,[48] K. Yip,[23] C. Yoshikawa,[22] S. Youssef,[21] J. Yu,[23] Y. Yu,[10] B. Zhang,[23,†]
Y. Zhou,[23,†] Z. Zhou,[30] Z.H. Zhu,[41] M. Zielinski,[41] D. Zieminska,[27] A. Zieminski,[27]
E.G. Zverev,[14] and A. Zylberstejn[5]

(DØ Collaboration)

[1]*Universidad de Buenos Aires, Buenos Aires, Argentina*
[2]*LAFEX, Centro Brasileiro de Pesquisas Físicas, Rio de Janeiro, Brazil*
[3]*Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil*
[4]*Universidad de los Andes, Bogotá, Colombia*
[5]*DAPNIA/Service de Physique des Particules, CEA, Saclay, France*
[6]*Panjab University, Chandigarh, India*
[7]*Delhi University, Delhi, India*
[8]*Tata Institute of Fundamental Research, Mumbai, India*
[9]*Kyungsung University, Pusan, Korea*
[10]*Seoul National University, Seoul, Korea*
[11]*CINVESTAV, Mexico City, Mexico*
[12]*Institute of Nuclear Physics, Kraków, Poland*
[13]*Institute for Theoretical and Experimental Physics, Moscow, Russia*
[14]*Moscow State University, Moscow, Russia*
[15]*Institute for High Energy Physics, Protvino, Russia*
[16]*University of Arizona, Tucson, Arizona 85721*
[17]*Lawrence Berkeley National Laboratory and University of California, Berkeley, California 94720*
[18]*University of California, Davis, California 95616*
[19]*University of California, Irvine, California 92697*
[20]*University of California, Riverside, California 92521*
[21]*Florida State University, Tallahassee, Florida 32306*
[22]*University of Hawaii, Honolulu, Hawaii 96822*
[23]*Fermi National Accelerator Laboratory, Batavia, Illinois 60510*
[24]*University of Illinois at Chicago, Chicago, Illinois 60607*

[25]*Northern Illinois University, DeKalb, Illinois 60115*
[26]*Northwestern University, Evanston, Illinois 60208*
[27]*Indiana University, Bloomington, Indiana 47405*
[28]*University of Notre Dame, Notre Dame, Indiana 46556*
[29]*Purdue University, West Lafayette, Indiana 47907*
[30]*Iowa State University, Ames, Iowa 50011*
[31]*University of Kansas, Lawrence, Kansas 66045*
[32]*Louisiana Tech University, Ruston, Louisiana 71272*
[33]*University of Maryland, College Park, Maryland 20742*
[34]*Boston University, Boston, Massachusetts 02215*
[35]*Northeastern University, Boston, Massachusetts 02115*
[36]*University of Michigan, Ann Arbor, Michigan 48109*
[37]*Michigan State University, East Lansing, Michigan 48824*
[38]*University of Nebraska, Lincoln, Nebraska 68588*
[39]*Columbia University, New York, New York 10027*
[40]*New York University, New York, New York 10003*
[41]*University of Rochester, Rochester, New York 14627*
[42]*State University of New York, Stony Brook, New York 11794*
[43]*Brookhaven National Laboratory, Upton, New York 11973*
[44]*University of Oklahoma, Norman, Oklahoma 73019*
[45]*Brown University, Providence, Rhode Island 02912*
[46]*University of Texas, Arlington, Texas 76019*
[47]*Texas A&M University, College Station, Texas 77843*
[48]*Rice University, Houston, Texas 77005*

Contents

I. INTRODUCTION

The mass of the top quark is a free parameter in the standard model of the electroweak interactions [1]. It arises from the Yukawa coupling of the top quark to the Higgs field, which is not constrained by the model. Through radiative corrections, the value of the top quark mass affects predictions of the standard model for many processes. For example, the prediction for the mass of the W boson varies by approximately 7 MeV[1] for every 1 GeV change in the mass of the top quark [2]. Precise measurements of the masses of the top quark and the W boson constrain the mass of the Higgs boson. This dependence can be turned around and the top quark mass predicted from measurements of electroweak processes within the framework of the standard model. Such an analysis gives 158^{+14}_{-11} GeV for the top quark mass [3]. In this sense, a measurement of the top quark mass constitutes a consistency test of the standard model prediction.

The top quark is the only fermion with a mass close to the vacuum expectation value of the Higgs field, or equivalently, with a Yukawa coupling close to unity. It is therefore possible that by studying the properties of the top quark we can learn more about electroweak symmetry breaking.

The Fermilab Tevatron produces top quarks in collisions of protons and antiprotons at $\sqrt{s} = 1.8$ TeV. The Tevatron provided the first experimental confirmation of the existence of the top quark [4]. In $p\overline{p}$ collisions top quarks are produced predominantly in $t\overline{t}$ pairs. The standard model predicts the top quark primarily ($> 99\%$) to decay to Wb. The decay modes of the W boson then define the signatures of $t\overline{t}$ decays. If both W bosons decay leptonically the signature contains two charged leptons with high p_T. We call this the dilepton channel. Events in which one of the W bosons decays leptonically and the other into jets contain one high p_T charged lepton and high p_T hadron jets. We call this the lepton+jets channel. In the all-jets channel both W bosons decay into jets.

The DØ collaboration was first to measure the mass of the top quark in the dilepton channel [5,6]. In this article we present a more detailed account of this analysis. The most precise measurements of the top quark mass have been obtained using the lepton+jets channel [7,8]. Table I lists previously published measurements of the top quark mass.

The measurement described in this paper is based on an integrated luminosity of approximately 125 pb^{-1}, recorded by the DØ detector during the 1992–1996 collider runs. We first give a brief description of the experimental setup (Sect. II), data reconstruction (Sect. III) and calibration procedures (Sect. IV). We then describe the selection of the event sample (Sect. V), the mass analysis of the selected events (Sect. VI), the maximum likelihood fit to the data (Sect. VII), and the systematic uncertainties associated with the fit (Sect. VIII). Finally we summarize the results and combine them with the measurement in the lepton+jets channel (Sect. IX).

---

[1]We use natural units with $\hbar = c = 1$.

TABLE I. Published measurements of the top quark mass. The first uncertainty is statistical, the second systematic.

Experiment	Channel	Mass
DØ [7]	lepton+jets	173.3± 5.6± 5.5 GeV
DØ [5]	dilepton	168.4±12.3± 3.6 GeV
CDF [8]	lepton+jets	175.9± 4.8± 4.9 GeV
CDF [9]	dilepton	161 ±17 ±10 GeV
CDF [10]	all-jets	186 ±10 ±12 GeV

II. DETECTOR

DØ is a multipurpose detector designed to study $p\bar{p}$ collisions at high energies. The detector was commissioned at the Fermilab Tevatron during the summer of 1992. A full description of the detector can be found in Ref. [11]. Here, we describe only briefly the properties of the detector that are relevant for the mass measurement in the dilepton channel.

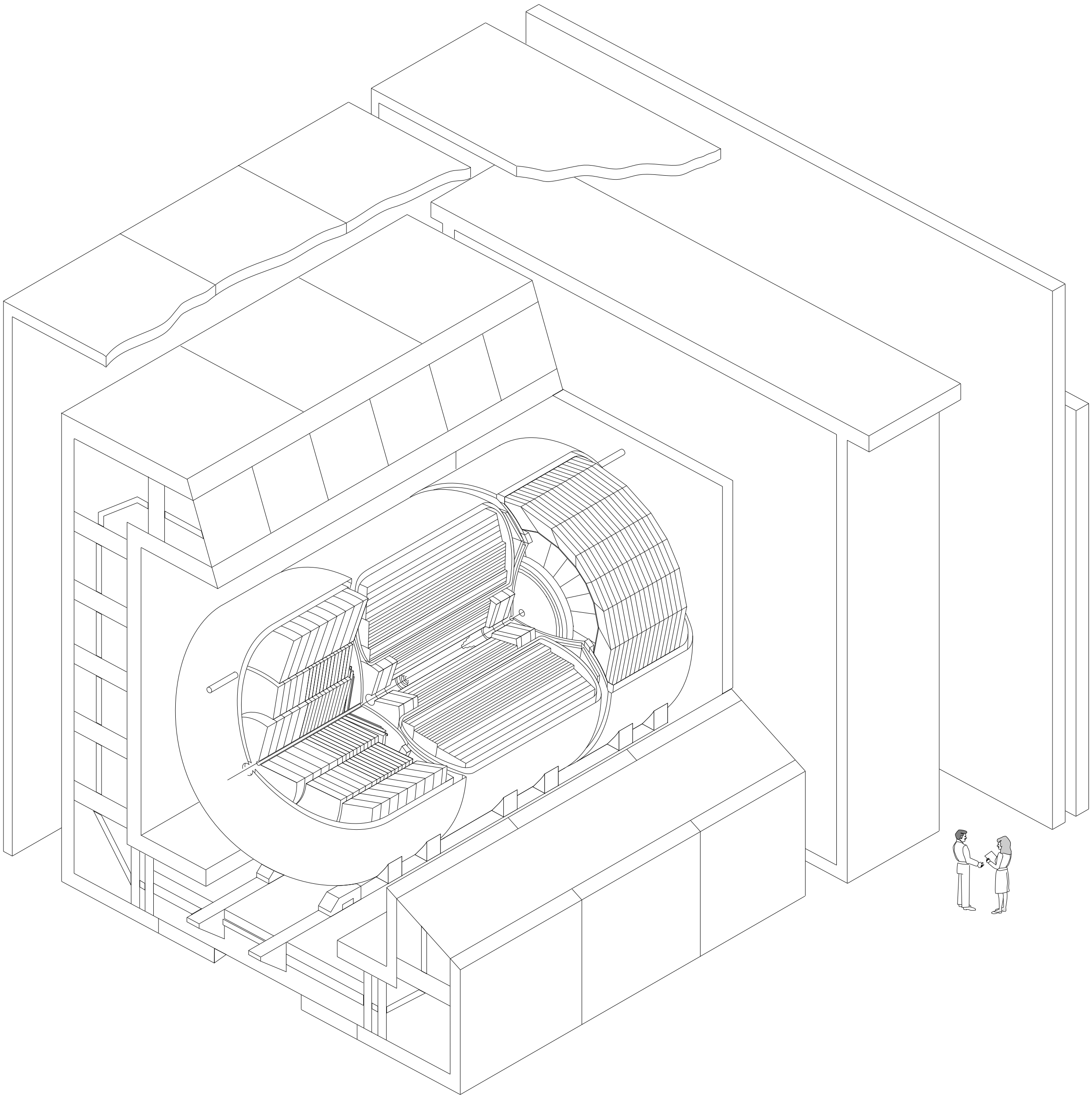
We specify detector coordinates in a system with its origin defined by the center of the detector and the z-axis defined by the proton beam. The x-axis points out of the Tevatron ring and the y-axis up. We use ϕ to denote the azimuthal coordinate and θ for the polar angle. Rather than θ, we often use the pseudorapidity $\eta = \tanh^{-1}(\cos\theta)$.

The detector consists of three primary systems: central tracking, calorimeter, and muon spectrometer. A cut away view of the detector is shown in Fig. 1.

The nonmagnetic central tracking system consists of four subdetectors that measure the trajectories of charged particles: a vertex drift chamber, a transition radiation detector, a central drift chamber, and two forward drift chambers. These chambers also measure ionization to identify tracks from single charged particles and e^+e^- pairs from photon conversions. The central tracking system covers the region $|\eta| < 3.2$.

The uranium-liquid argon calorimeter is divided into three parts, the central calorimeter and the two end calorimeters, and covers the pseudorapidity range $|\eta| < 4.2$. Longitudinally, the calorimeter is segmented into an electromagnetic (EM) section with fine sampling and a hadronic section with coarser sampling. The calorimeter is segmented transversely into quasiprojective towers with $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$. The third layer of the electromagnetic calorimeter, where EM showers are expected to peak, is segmented twice as finely in each direction. The hadronic calorimeter modules back up any cracks in the coverage of the EM calorimeter modules such that there are no projective cracks in the calorimeter, ensuring good resolution for the measurement of transverse momentum balance.

Since muons from top quark decays predominantly populate the central region, we use only the central portion of the muon system, which covers $|\eta| < 1.7$. This system consists of four planes of proportional drift tubes in front of magnetized iron toroids with a magnetic field of 1.9 T and two groups of three planes of proportional drift tubes behind the toroids. The magnetic field lines and the wires in the drift tubes are oriented transversely to the beam direction. The momentum is obtained from the deflection of the muon in the magnetic field of the toroid.



DØ Detector

FIG. 1. Cut away isometric view of the DØ detector.

III. PARTICLE IDENTIFICATION

The particle identification algorithms used for electrons, muons, and jets are the same as in previously published analyses [12]. We summarize them in the following sections.

A. Electrons

Electron candidates are first identified by finding isolated clusters of energy in the EM calorimeter along with a matching track in the central detector. We accept electron candidates with $|\eta| \leq 2.5$. Final identification is based on a likelihood test on the following five variables:

- The agreement of the shower shape with the expected shape of an electromagnetic shower, computed using the full covariance matrix of the energy depositions in the cells of the electromagnetic calorimeter.
- The electromagnetic energy fraction, defined as the ratio of the shower energy found in the electromagnetic calorimeter to the total shower energy.
- A measure of the distance between the track and the cluster centroid.
- The ionization dE/dx along the track.
- A variable characterizing the energy deposited in the transition radiation detector.

To a good approximation, these five variables are independent of each other for electron showers.

Electrons from W boson decay tend to be isolated. Thus, we make the additional cut

$$\frac{E_{\mathrm{tot}}(0.4) - E_{\mathrm{EM}}(0.2)}{E_{\mathrm{EM}}(0.2)} < 0.1, \tag{1}$$

where $E_{\mathrm{tot}}(0.4)$ is the energy within $\Delta R < 0.4$ of the cluster centroid and $E_{\mathrm{EM}}(0.2)$ is the energy in the EM calorimeter within $\Delta R < 0.2$. ΔR is defined as $\sqrt{\Delta\eta^2 + \Delta\phi^2}$.

B. Muons

Two types of muon selection are used in this analysis. The first is used to identify isolated muons from $W \to \mu\nu$ decay. The second type of muon selection is used to tag b-jets by identifying muons consistent with originating from $b \to \mu + X$ decay. We accept muons with $|\eta| < 1.7$. Besides cuts on the muon track quality, both selections require that the energy deposited in the calorimeter along a muon track be at least that expected from a minimum ionizing particle. For isolated muons, such as those from W boson decays, we require $\Delta R_{\mu,j} > 0.5$ for the distance $\Delta R_{\mu,j}$ in the $\eta - \phi$ plane between the muon and any jet. For soft muons in jets, such as those from $b \to \mu + X$ decay, we require $p_T \geq 4$ GeV and $\Delta R_{\mu,j} < 0.5$. The efficiency$\times$acceptance for either muon selection with these cuts is about 64%.

C. Jets

Jets are reconstructed in the calorimeter using a fixed-size cone algorithm. We use a cone size of $\Delta R = 0.5$. See Ref. [13] for a detailed description of the jet reconstruction algorithm.

D. Missing Transverse Momentum

The missing transverse momentum, $\vec{\not{p}}_T$, is the momentum required to balance the measured momenta in the event ($\sum \vec{p}_T + \vec{\not{p}}_T = 0$). In the calorimeter, we calculate $\vec{\not{p}}_T$ as

$$\vec{\not{p}}_T^{\,\mathrm{cal}} = -\sum_i E_i \sin\theta_i \begin{pmatrix} \cos\phi_i \\ \sin\phi_i \end{pmatrix}, \tag{2}$$

where i runs over all calorimeter cells, E_i is the energy deposited in the i^{th} cell, and ϕ_i is the azimuthal and θ_i the polar angle of the i^{th} cell. When there are muons present in the event we refine the calculation

$$\vec{\not{p}}_T = \vec{\not{p}}_T^{\,\mathrm{cal}} - \sum_k \vec{p}_T^{\,\mu_k}, \tag{3}$$

where $\vec{p}_T^{\,\mu}$ is the transverse momentum of the muon as measured by the muon system.

IV. ENERGY SCALE CALIBRATION

A. Electron Energy Scale

The measurement of the energy E of electromagnetic showers in the calorimeter is calibrated using $Z \to ee$, $J/\psi \to ee$, and $\pi^0 \to \gamma\gamma$ decays to a precision of 0.08% at $E = M_Z/2$ and to 0.6% at $E = 20$ GeV [14]. The electron energy scale calibration therefore does not give rise to any significant uncertainty in the top quark mass measurement.

B. Muon Momentum Scale

The muon momentum scale, calibrated using $J/\psi \to \mu\mu$ and $Z \to \mu\mu$ candidates, has an uncertainty of 2.5%. Its effect on our measurement of the top quark mass was determined by varying the muon momentum scale in Monte Carlo samples of $t\bar{t}$ events with $m_t = 170$ GeV. The tests indicate that the relation between muon scale and top quark mass error is given by

$$\delta m_t = 12 \text{ GeV} \frac{\delta p_T^{\mu}}{p_T^{\mu}}. \tag{4}$$

Hence, the 2.5% uncertainty in muon momentum scale leads to a systematic uncertainty of 0.3 GeV in our measurement of the top quark mass. This uncertainty is completely negligible compared to the effect of the jet energy scale.

C. Jet Energy Scale

The jet energy scale is calibrated relative to the electromagnetic energy scale by balancing the transverse momentum in events with jets and electromagnetic showers [15]. The exercise is carried out separately and symmetrically for both data and Monte Carlo.

In addition to the corrections in Ref. [15] we apply an η-dependent correction derived from a comparison between γ+jet events in data and Monte Carlo events created using the HERWIG [16] event generator and a GEANT [17] based detector simulation. We also correct jets that contain a muon, indicative of a semileptonic b quark decay, to compensate on average for the energy carried away by the undetected neutrino. These corrections are identical to those used and detailed in the mass analysis based on the lepton+jets final states [7] with the exception that no attempt is made to account for gluon radiation outside of the jet cone. Rather, the procedure in the dilepton analysis is to explicitly account for additional reconstructed jets, as described in Sect. VI C.

We estimate the degree of possible residual discrepancy between the jet energy response of the detector and the Monte Carlo simulation from the energy balance between electromagnetic energy clusters and jets from collider data, compared to photon+jets Monte Carlo, as a function of photon p_T. The data constrain the possible mismatch to less than $\pm(2.5\%+0.5$ GeV) in the jet energy [7]. This uncertainty gives rise to a significant systematic uncertainty in our top quark mass measurement (see Sect. VIII B).

V. EVENT SELECTION

A. Basic Event Selection Criteria

The event selection for the dilepton mass analysis is almost identical to that used for the measurement of the cross section [12]. We require two charged leptons (e, μ) and at least two jets in the events. In addition we cut on global event quantities like $\not{p}_T$ and H_T. The basic kinematic selection criteria are summarized in Table II. The variable H_T is defined as

$$H_T = \begin{cases} \sum p_T^j + p_T^{e_1} & \text{for the } ee \text{ and } e\mu \text{ channels;} \\ \sum p_T^j & \text{for the } \mu\mu \text{ channel,} \end{cases} \tag{5}$$

where e_1 is the leading electron in ee events. The sum is over all jets with $p_T>$ 15 GeV and $|\eta| < 2.5$. Muons are not included in the sum because their momenta are measured less precisely. H_T gives good rejection against background processes, which typically have less jet activity along with the dilepton signature.

The event selection criteria are designed to identify events with two charged leptons and additional jets in the final state as expected from $t\bar{t} \to \ell\ell + X$ decays. The background in the ee and $\mu\mu$ channels is dominated by $Z \to ee$ and $Z \to \mu\mu$ decays. We apply additional criteria, described in the following sections, that remove these particular backgrounds. Table III gives the number of background events expected in each dilepton channel after all selection criteria are applied. Instrumental backgrounds arise from particle misidentification, *e.g.* mistaking a jet for an electron.

TABLE II. Kinematic and fiducial cuts used in selecting dilepton events.

Objects		ee	$e\mu$	$\mu\mu$
2 Leptons	p_T^ℓ	> 20 GeV	> 15 GeV	> 15 GeV
	$\lvert\eta^\ell\rvert$	< 2.5	< 1.7	< 1.7
≥ 2 Jets	p_T^j	> 20 GeV	> 20 GeV	> 20 GeV
	$\lvert\eta^j\rvert$	< 2.5	< 2.5	< 2.5
Event	$\not\!p_T$	—	> 10 GeV	—
	$\not\!p_T{}^{\text{cal}}$	> 25 GeV	> 20 GeV	—
	H_T	> 120 GeV	> 120 GeV	> 100 GeV

TABLE III. Expected numbers of background events.

Background Source	ee	$e\mu$	$\mu\mu$
$Z \to \ell\ell$	0.058 ± 0.012	—	0.558 ± 0.21
$Z \to \tau\tau \to \ell\ell$	0.078 ± 0.022	0.099 ± 0.076	0.029 ± 0.017
WW	0.083 ± 0.023	0.074 ± 0.018	0.007 ± 0.004
Drell-Yan	0.054 ± 0.030	0.002 ± 0.003	0.066 ± 0.035
$t\bar{t} \to e + \text{jets}$	0.04	—	—
Instrumental	0.197 ± 0.046	0.035 ± 0.13	0.068 ± 0.010
Total Background	0.51 ± 0.09	0.21 ± 0.16	0.73 ± 0.25

B. $e\mu$ Channel

The $e\mu$ channel is the most powerful dilepton channel with twice the branching ratio of the ee and $\mu\mu$ channels and without the background from $Z \to ee$ or $Z \to \mu\mu$ decays. The largest background is $Z \to \tau\tau \to e\mu + X$, which is suppressed by both branching ratio and kinematics. Instrumental backgrounds arise from W bosons that decay to $\mu\nu$ which are produced in association with jets, one of which is mistaken for an electron.

We observe three events in this channel.

C. ee Channel

The primary source of physics background in the ee channel is Z boson production with associated jets. These events have no neutrinos and can be rejected effectively by cutting on $\not\!p_T$. We therefore require $\not\!p_T > 40$ GeV if the dielectron invariant mass is within 12 GeV of the Z boson mass peak. Instrumental backgrounds arise from W+jets production or multijet events in which jets fake the electron signature.

In this channel we extend our event selection criteria to include an additional event that was not part of the final sample for the measurement of the cross section. This event passes all selection criteria, except that one of the electron candidates has no matching track. This cluster is nevertheless consistent with originating from an electron because the trajectory connecting the vertex with the cluster passes only through the two inner layers of the CDC. The inner two layers do indeed have hits but to reconstruct a track, hits are required in at least three layers. The lack of a reconstructed track could indicate a higher probability for this electron to be misidentified. On the other hand one of the jets contains a muon, which passes all requirements for the muon-tag analyses reported in reference [12]. A muon tag indicates that the jet probably originates from the fragmentation of a b quark. The probability of tagging a jet from the fragmentation of a light quark or a gluon is quite small. The presence of a b jet reduces the likelihood that this event arises from instrumental background sources and we therefore include it in the event sample for the mass analysis.

We revise the background estimate for the ee channel to include an additional component due to the inclusion of this event. We compute the number of additional background events expected if events are admitted that are missing a matched track for one of the two electron candidates but have a muon tag. In our data we find 11 events with one electron candidate and three jets, one with muon tag. In these events, there are 22 jets that could fake a second electron. The probability for any one of these jets to mimic an electron signature without matched track requirement is 8×10^{-4} [18], so that we expect about 0.018 events due to the extension of the selection cuts. We also have to take into account that we specifically extended the selection criteria to add this event. The additional background only contributes to experiments in which at least one event satisfies the extended selection cuts. This is expected to happen only once every six experiments. The additional background component is therefore six times 0.018 or 0.11 events. The most significant source of these background events are $t\bar{t}$ decays to e+jets with a muon-tagged jet, in which one jet is misidentified as an electron.

In total, two ee events enter our final sample.

D. $\mu\mu$ Channel

The dimuon channel shares the $Z \to \ell\ell$ background with the dielectron channel. The less precise measurement of the muon momentum makes separation of the $t\bar{t}$ signal from this background more difficult. In order to reduce this background, a kinematic fit to the $Z \to \mu\mu$ hypothesis is applied, and the event is required to have χ^2 probability less than 1% for this fit. Even after this cut, Z boson production remains the dominant background source. Instrumental backgrounds arise from heavy quark jets with a high-p_T muon that is misidentified as an isolated muon.

One event survives all selection criteria.

E. Dilepton Events

Six events enter our dilepton event sample: three are $e\mu$ events, two are ee events, and one is a $\mu\mu$ event. Table IV lists the properties of these events.

TABLE IV. Kinematic properties of dilepton events (momenta in GeV) used in the reconstruction of the top quark mass. All corrections are included.

Event	Object	p_x	p_y	p_z	p_T	η	ϕ
$e\mu\#1$	e	12.3	−97.8	41.1	98.6	0.41	4.84
	μ	−68.3	272.5	95.1	280.0	0.33	1.82
	$\not{p}_T$	100.5	−152.7	—	182.9	—	5.29
	jet	−25.5	−9.9	−20.8	27.3	−0.70	3.51
	jet	−14.4	−20.5	32.3	25.1	1.07	4.10
$e\mu\#2$	e	−75.4	−1.1	−30.2	74.5	−0.39	3.16
	μ	−25.2	10.6	−12.8	27.4	−0.45	2.75
	$\not{p}_T$	62.0	5.2	—	62.3	—	0.08
	jet	38.9	−85.6	−16.0	94.0	−0.17	5.14
	jet	14.2	33.1	−11.4	36.0	−0.31	1.17
	jet	−1.6	29.3	11.9	29.4	0.39	1.63
$e\mu\#3$	e	−44.7	20.2	140.1	49.1	1.77	2.72
	μ	5.4	17.2	−3.3	18.1	−0.18	1.27
	$\not{p}_T$	−12.5	4.5	—	13.2	—	2.79
	jet	39.6	−29.9	11.3	49.7	0.22	5.64
	jet	19.8	−19.4	−31.0	27.7	−0.97	5.51
$ee\#1$	e	2.7	50.4	17.1	50.5	0.33	1.52
	e	−7.4	21.4	−47.6	22.6	−1.49	1.91
	$\not{p}_T$	41.3	−4.0	—	41.5	—	6.19
	jet	−29.2	−36.9	−37.0	47.1	−0.72	4.04
	jet	3.5	−27.1	−28.9	27.4	−0.92	4.84
$ee\#2$	e	52.3	−4.1	−34.4	52.5	−0.62	6.20
	e	−8.5	−26.6	27.0	27.9	0.86	4.40
	$\not{p}_T$	42.6	−11.3	—	44.1	—	6.02
	jet*	−92.4	−26.0	−61.6	96.0	−0.60	3.41
	jet	−23.5	25.3	−34.0	34.6	−0.87	2.32
	jet	0.0	27.7	18.3	27.7	0.62	1.57
$\mu\mu$	μ	−63.9	12.7	−21.4	65.1	−0.32	2.94
	μ	−16.0	31.0	1.9	34.9	0.05	2.05
	$\not{p}_T$	71.2	53.2	—	88.9	—	0.64
	jet	33.8	−103.1	−107.6	108.5	−0.88	5.03
	jet	−9.1	22.7	27.7	24.5	0.97	1.95
	jet	−8.4	−18.6	47.8	20.5	1.58	4.29

* tagged by a soft muon

VI. RECONSTRUCTION OF THE TOP QUARK MASS

A. Characteristics of Dilepton Events

The dilepton decay topology does not provide sufficient information to uniquely reconstruct the t and $\bar{t}$ quarks. In the simplest scenario, the decay $t \to W^+b$, $\bar{t} \to W^-\bar{b}$, followed by $W^+ \to \ell^+\nu$ and $W^- \to \ell^-\bar{\nu}$ produces six particles in the final state: two charged leptons, which we allow to be either electrons or muons (ee, $e\mu$, or $\mu\mu$); two neutrinos (ν,$\bar{\nu}$); and two b quarks (b,$\bar{b}$), as shown in Fig. 2. Given the identities of the particles, this final state is therefore completely specified by the momenta of these six particles, *i.e.* 18 numbers. We measure the momenta of the charged leptons and the jets from the hadronization of the b quarks directly. In addition, the observed $\not{p}_T$ provides the x and y components of the sum of the neutrino momenta for a total of 14 measurements. Assuming $m_t > M_W + m_b$ we can impose three constraints, two on the masses of the decaying W bosons, $m^{\ell^+\nu} = m^{\ell^-\bar{\nu}} = M_W$, and one on the masses of the top quarks, $m^{\ell^+\nu b} = m^{\ell^-\bar{\nu}\bar{b}}$. This leaves us with 17 equations and 18 unknowns so that a kinematic fit would be underconstrained. We have to develop a different procedure to obtain an estimate of the top quark mass from the available information. This is the fundamental difference between the mass determination in the dilepton channel and that in the lepton+jets channel, which allows a kinematic fit with two constraints.

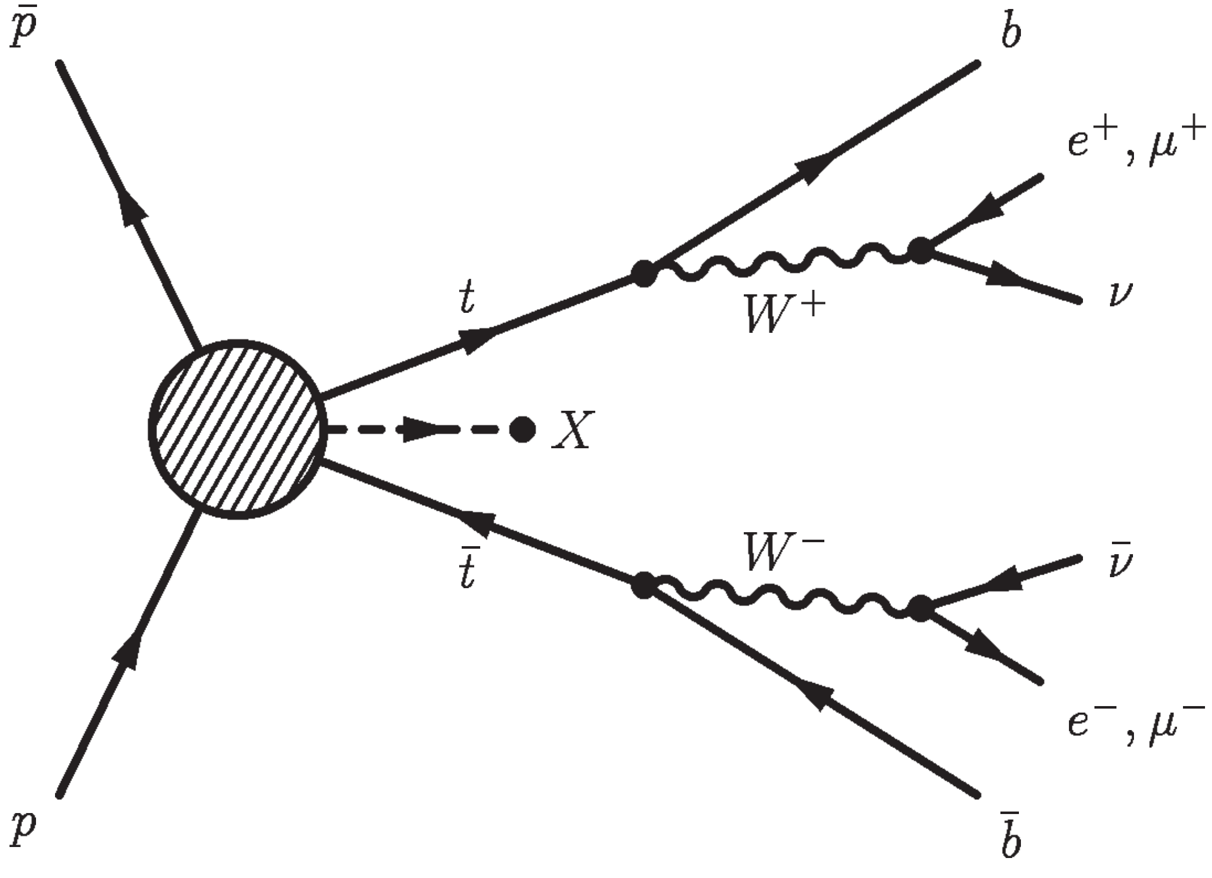


FIG. 2. Schematic representation of $t\bar{t}$ production and decay in the dilepton channels.

We solve this problem by fitting the dynamics of the decays [19]. For each event we derive a weight function, which is a measure of the probability density for a $t\overline{t}$ pair to decay to the observed final state, as a function of the top quark mass. We compare these weight functions to Monte Carlo simulations of $t\overline{t}$ decays for different values of the top quark mass and use a maximum likelihood fit to extract the mass value that yields the best agreement.

B. Computation of the Weight Function

Ideally we would like to compute analytically the probability density for a $t\overline{t}$ pair to decay to the observed final state for any given value of the top quark mass. This probability density is given by

$$\mathcal{P}(\{o\}|m_t) \propto \int f(x)f(\overline{x})|\mathcal{M}|^2 p(\{o\}|\{v\})\delta^4 \mathrm{d}^{18}\{v\}\,\mathrm{d}x\,\mathrm{d}\overline{x}, \tag{6}$$

where $\{o\}$ is the set of 14 measured quantities and $\{v\}$ is the set of 18 parameters that specify the final state. $\mathcal{M}$ is the matrix element for the process $q\overline{q}$ or $gg \to t\overline{t}+X \to \ell^+\nu b\ell^-\overline{\nu}\overline{b}+X$, $f(x)$ the parton density for quarks or gluons of momentum fraction x in the proton, and $f(\overline{x})$ that for antiquarks or gluons of momentum fraction $\overline{x}$ in the antiproton. The detector resolution function $p(\{o\}|\{v\})$ is the probability density to observe the values $\{o\}$ given the final state parameters $\{v\}$. The four-dimensional δ-function enforces the four mass constraints:

$$\delta^4 = \delta(m^{\ell^+\nu} - M_W) \times \delta(m^{\ell^-\overline{\nu}} - M_W) \times \delta(m^{\ell^+\nu b} - m_t) \times \delta(m^{\ell^-\overline{\nu}\overline{b}} - m_t). \tag{7}$$

Here we neglect the finite widths of the W boson and the top quark.

Unfortunately this expression involves a multidimensional integral that has to be evaluated numerically and is complicated by the need to include initial and final state gluon radiation. Such higher order effects complicate the reconstruction of the top quark mass substantially and cannot be neglected. We therefore do not attempt to compute the exact probability density given in Eq. 6. Rather, we construct simpler weights that retain sensitivity to the value of the top quark mass but can be evaluated with the available computing resources. We calibrate the effect of the simplifications by comparing the weight functions obtained from the collider data to Monte Carlo simulations (Sect. VII).

The calculation of the weight function proceeds in three steps. First we map the observed charged leptons and jets to the corresponding t and $\overline{t}$ decay products. There are ambiguities in this step because the fragmentation of the b quarks may result in more than one reconstructed jet or because a gluon radiated from the initial state may contribute a jet to the event. We cannot, in general, distinguish between jets originating from gluons and quarks. Furthermore, we do not measure the sign of the electron charge nor can we distinguish between jets originating from quarks and antiquarks. Therefore, there is an ambiguity in pairing the charged leptons and b jets that originate from the same top quark. We repeat the following two steps for each of the possible assignments and add the resulting weight functions.

Given the charged lepton and b quark momenta from the decay of the t and $\overline{t}$ quarks and the sum of the neutrino momentum components, $p_x^{\nu\overline{\nu}}$ and $p_y^{\nu\overline{\nu}}$, we compute a weight as

TABLE V. Possible assignments of three observed jets (j_1, j_2, and j_3) to the b quarks and initial state radiation (ISR).

Permutation	b-Jets		ISR
1	j_1	j_2	j_3
2	j_1	j_3	j_2
3	j_2	j_3	j_1
4	$j_1 + j_2$	j_3	—
5	$j_2 + j_3$	j_1	—
6	$j_1 + j_3$	j_2	—

a function of the top quark mass. We have developed two algorithms to compute the weight function which emphasize different aspects of top production dynamics. The first algorithm (matrix-element weighting) is an extension of the weight proposed in Ref. [20] and takes into account the parton distribution functions for the initial proton and antiproton and the decay distribution of the W bosons due to the V–A coupling of the charged current. The second (neutrino weighting) [6] is based on the available phase space for neutrinos from the decay of the $t\bar{t}$ pair.

Finally we average the weight function over the experimental resolution.

In the following, we first discuss the ambiguities in associating the observables with final state particles. Then we discuss the two algorithms that are used to compute the weight functions and finally the experimental resolutions.

C. Jet Combinatorics

In the calorimeter we detect the jets from the fragmentation of the two b quarks. The fragmentation of a b quark can produce more than one jet because of hard gluon radiation. This corresponds to final state radiation. Jets can also originate from gluons radiated by partons in the initial state. We refer to this as initial state radiation. It is not possible to tell whether a jet originates from the fragmentation of a quark or a gluon, unless a b quark decays semileptonically to a muon that we subsequently detect. Thus, reconstruction of the original partons from the observed jets presents some complication.

We consider jets with $p_T > 15$ GeV. If there are only two such jets we assign their measured momenta to the two b quarks. If there are more than two jets we have a range of possible assignments. To limit the possibilities, we restrict the procedure to the three leading jets in p_T. We assign two of them to the b quarks and the third jet either to initial state radiation, in which case we ignore it, or to final state radiation, in which case we add its momentum to that of one of the two b quarks. There are six possible permutations for three jets, as listed in Table V.

If there is a jet in the event that is tagged by a soft muon, we only allow permutations that assign this jet to a b quark. In the collider data sample this is the case for one ee event.

Not all permutations are equally likely to be correct. For each jet considered to be due to initial state radiation, we assign a weight factor

$$\mathcal{Q}_{\text{ISR}} = \exp\left(\frac{-p_T^j \sin\theta^j}{25 \text{ GeV}}\right). \tag{8}$$

Similarly, for every pair of jets that is assigned to a b quark, we define

$$\mathcal{Q}_{\text{FSR}} = \exp\left(\frac{-m^{jj}}{20 \text{ GeV}}\right), \tag{9}$$

where m^{jj} is the invariant mass of the two jets. These functional forms of the weights were derived empirically from a study of $t\overline{t}$ decays generated by ISAJET [21]. The factor $\mathcal{Q}_{\text{ISR}}$ favors assignments in which jets from initial state radiation are close to the beam direction, and $\mathcal{Q}_{\text{FSR}}$ favors the merging of jets which are soft or close together. The numerical coefficients of the exponents are chosen such that the mean reconstructed top quark masses for events with two-jet and multi-jet final states are the same.

After adding the four-momenta of the jets assigned to a b quark, we rescale the momentum components, keeping the energy fixed, so that the b quark four-momentum has an invariant mass of 5 GeV to put the outgoing quark momentum on the mass shell.

There are two ways to pair the momenta of the two charged leptons with the two b quark momenta. Since we cannot determine which b quark originated from the decay of the t quark and which from the decay of the $\overline{t}$ quark, we consider both pairings with equal probability.

D. Matrix-Element Weighting ($\mathcal{M}$WT) Algorithm

Assuming that we know the momenta of the charged leptons (p^{ℓ^+}, p^{ℓ^-}), the b quarks (p^b, $p^{\overline{b}}$), and the sum of the x and y components of the neutrino momenta ($p_x^{\nu\overline{\nu}}$, $p_y^{\nu\overline{\nu}}$) and that we impose the three constraints mentioned above, we are still one constraint short of being able to solve for the unknown components of the neutrino momenta. Assuming a fixed value for the top quark mass m_t supplies the required constraint to solve the problem, except for a fourfold ambiguity. Not all solutions are equally likely for any given value of m_t. We therefore assign a weight to the i^{th} solution [20]:

$$w_i^{\mathcal{M}}(m_t) = f(x)f(\overline{x})\, p(E_i^{\ell^-*}|m_t)\, p(E_i^{\ell^+*}|m_t), \tag{10}$$

where $f(x)$ and $f(\overline{x})$, the parton distribution functions, are evaluated at $Q^2 = m_t^2$, and $p(E^{\ell*}|m_t)$ is the probability density function for the energy of the charged lepton in the rest frame of the top quark ($E^{\ell*}$). This probability density is given by

$$p(E^{\ell*}|m_t) = \frac{4m_t E^{\ell*}(m_t^2 - m_b^2 - 2m_t E^{\ell*})}{(m_t^2 - m_b^2)^2 + M_W^2(m_t^2 + m_b^2) - 2M_W^4}. \tag{11}$$

We sum the weights for all solutions and normalize by a factor $\mathcal{A}(m_t)$ to obtain the weight for the event

$$w^{\mathcal{M}}(m_t) = \mathcal{A}(m_t) \sum_{i=1}^{4} w_i^{\mathcal{M}}(m_t). \tag{12}$$

The factor $\mathcal{A}(m_t)$ ensures that the average weight is independent of the top quark mass. We compute the weight function for $82 < m_t < 278$ GeV in steps of 4 GeV, where the lower limit is given by the requirement that the top quark decays into a real W boson and a b quark and the upper limit is placed well above the measurement of the top quark mass in the lepton+jets channel. The normalization factor is computed using a Monte Carlo simulation so that

$$\sum_1^N w^{\mathcal{M}}(m_t) = N, \tag{13}$$

where the sum is over the events that pass the selection cuts. We parametrize the factor $\mathcal{A}(m_t)$ at different values of m_t (in GeV) as

$$\mathcal{A}(m_t) = \left(5.86 - 0.044 m_t + 0.000084 m_t^2\right)^{-1}. \tag{14}$$

E. Neutrino Weighting (νWT) Algorithm

The neutrino weighting algorithm also computes a weight as a function of the top quark mass. In contrast to the $\mathcal{M}$WT algorithm it does not solve for the unknown neutrino momentum components, but rather samples the neutrino pseudorapidity space and computes a weight based on how much of the sampled space is consistent with the observed $\not{p}_T$.

For every value of the top quark mass, we sample the rapidities of neutrino (η^ν) and antineutrino ($\eta^{\overline{\nu}}$) from the $t\bar{t}$ decay. For each top decay we then know the momenta of the charged lepton and the b quark, the assumed neutrino pseudorapidity, and the top quark mass, which allows us to solve for the transverse momentum components of the neutrino (p_x^ν and p_y^ν) with a twofold ambiguity. The two solutions for each of the two top decays combine to give four solutions for the event. For the i^{th} solution we compute a weight based on the agreement between the observed $\not{p}_T$ and the sum of the calculated neutrino p_T values:

$$w_i^\nu(m_t) = \exp\left(\frac{-\left(\not{p}_x - p_x^\nu - p_x^{\overline{\nu}}\right)^2}{2\sigma^2}\right) \times \exp\left(\frac{-\left(\not{p}_y - p_y^\nu - p_y^{\overline{\nu}}\right)^2}{2\sigma^2}\right), \tag{15}$$

where $\sigma = 4$ GeV is the resolution for each component of $\vec{\not{p}}_T$ (Sect. VI F).

Not every value of the neutrino pseudorapidity is equally likely. Figure 3 shows the distribution of neutrino rapidities predicted by the HERWIG Monte Carlo program for several top quark masses. The distributions can be approximated by Gaussian curves. The width σ_η of the Gaussian varies as a function of the top quark mass. It can be parametrized by the second order polynomial

$$\sigma_\eta = 5.56 \times 10^{-6} m_t^2 - 2.16 \times 10^{-3} m_t + 1.314, \tag{16}$$

as shown in Fig. 4. We compute the weights w_i^ν for ten values of each of the neutrino rapidities, spaced such that they divide the Gaussian into slices of equal area.

To obtain the weight for the event we add the weights for all four solutions and all values of the neutrino rapidities,

$$w^\nu(m_t) = \sum_{\eta^\nu} \sum_{\eta^{\overline{\nu}}} \sum_{i=1}^{4} w_i^\nu(m_t). \tag{17}$$

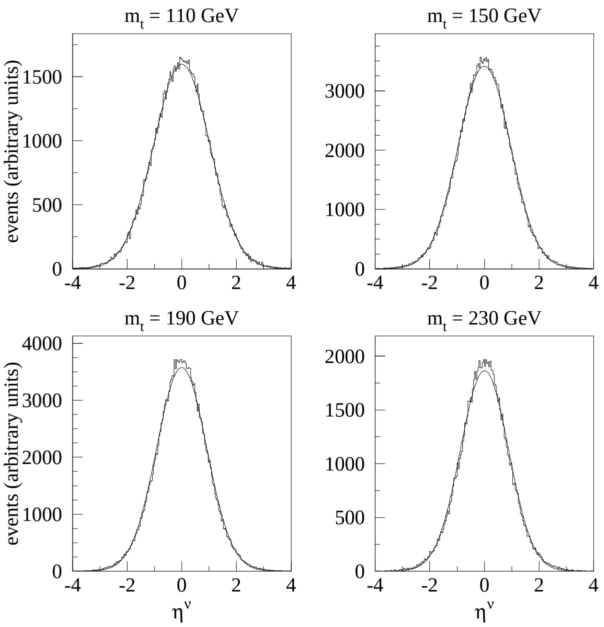


FIG. 3. Distributions of neutrino pseudorapidity from top quark decay, modeled by HERWIG, for several top quark masses. The smooth curves are fits to Gaussians.

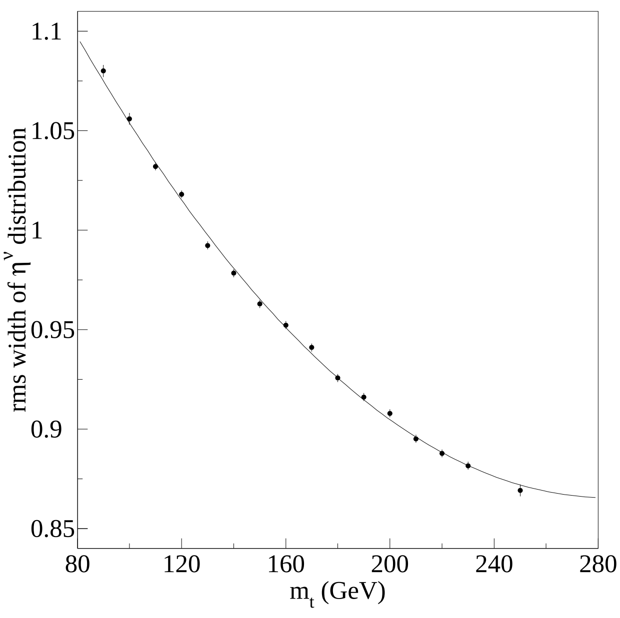


FIG. 4. Width of the Gaussian curves fit to the neutrino pseudorapidity distributions as a function of top quark mass. The smooth line is the polynomial parametrization used in the analysis.

F. Detector Resolution

The algorithms described in the two previous sections use as input the measured momenta of the charged leptons and b jets and the transverse components of the sum of the neutrino momenta. To account for finite resolution, we integrate the weights over the ranges of these quantities that are consistent with the measurements to smooth out the weight functions.

To evaluate this integral, we generate a large number of sets of event parameters over which we average the weights. These sets of event parameters derive from the observed events by adding normally distributed resolution terms to the observed values to populate the parameter space consistent with the measured values. The new values $\tilde{o}$ are given in terms of the observed value o, the resolution, σ, for the measurement of o, and a normally distributed random variable ξ:

$$\tilde{o} = o + \sigma\xi. \tag{18}$$

We apply such fluctuations to all momentum measurements. Directions are relatively precise and are therefore not fluctuated. This also reduces the number of numerical operations.

The energy resolution for electrons is

$$\sigma(E^e) = 0.15\ \mathrm{GeV}^{\frac{1}{2}}\sqrt{E^e}. \tag{19}$$

The resolution function for the inverse of the muon momentum is approximately Gaussian. We therefore fluctuate the inverse of the momentum with the resolution

$$\sigma\left(\frac{1}{p^\mu}\right) = \left\{\left(\frac{0.18(p^\mu - 2\ \mathrm{GeV})}{p^{\mu 2}}\right)^2 + \left(\frac{0.003}{\mathrm{GeV}}\right)^2\right\}^{\frac{1}{2}}. \tag{20}$$

The energy resolution for jets receives contributions from several effects. One is the intrinsic resolution of the calorimeter. The energy of the jet is measured as the energy in a cone of radius $\Delta R = 0.5$. This energy is not identical to that of the parton. Additional energy can be accrued from overlap with other jets and energy can be lost due to gluon radiation outside of the cone. These contributions to the resolution depend on the process and we therefore use Monte Carlo $t\bar{t}$ events to evaluate the jet energy resolution.

We compare the reconstructed jet p_T to that of the nearest cluster of hadrons generated by the Monte Carlo in a sample of $t\bar{t}$ events with top quark masses ranging from 110 to 190 GeV. Typically, the distribution in the fractional mismeasurement in p_T exhibits a narrow peak due to the intrinsic calorimeter resolution and broad tails due to ambiguity in the jet definition. We fit two Gaussian curves with equal means but different widths to the distribution, and parametrize the widths of the two Gaussians and their relative normalization as functions of p_T and η. Figure 5 shows a typical distribution along with the fit that we use as a resolution function. Figure 6 shows the rms resolution as a function of p_T.

The Monte Carlo simulation used to determine the jet energy resolution neither includes noise due to the intrinsic radioactivity of the uranium nor due to multiple interactions. We therefore add an additional uncorrelated constant noise term of 5–6 GeV, depending on η. These values were determined by balancing the p_T vectors in dijet events.

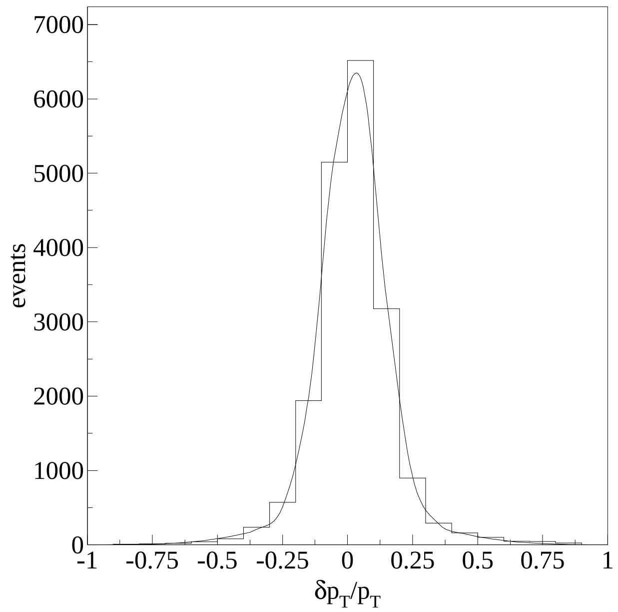


FIG. 5. Fractional p_T resolution for jets with $50 < p_T < 60$ GeV from $t\bar{t}$ decays generated with top quark masses between 110 and 190 GeV using the HERWIG program. The superimposed curve is the fit using two Gaussian curves.

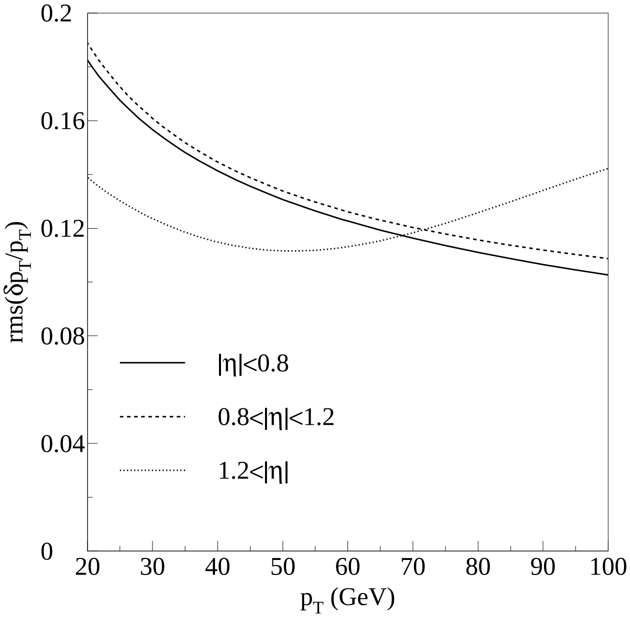


FIG. 6. Rms width of fractional jet p_T resolution functions versus jet p_T for three pseudorapidity regions.

Using a sample of random $p\overline{p}$ interactions, we measure the resolution for any component of $\vec{\not{p}}_T$ to be about 4 GeV. Both components of $\vec{\not{p}}_T$ are fluctuated by this resolution. The $\not{p}_T$ vector is also corrected for the fluctuations in the lepton and jet momenta.

The number of variations performed for each event is limited by the available computing power. We average over 100 variations per event for Monte Carlo samples and 5000 variations per event for the collider data.

The weight function for each event is then

$$W^x(m_t) = \frac{1}{N'N''} \sum_{j=1}^{N'} \sum_{k=1}^{2} \sum_{l=1}^{N''} \mathcal{Q}_{\rm ISR} \mathcal{Q}_{\rm FSR} w^x(m_t), \tag{21}$$

where $\mathcal{Q}_{\rm ISR}$ and $\mathcal{Q}_{\rm FSR}$ are the parametrized weights defined in Eqs. 8 and 9. The index j runs over the N' resolution fluctuations, k over the two lepton–b jet pairings, l over the N'' jet permutations, and x refers to the $\mathcal{M}$WT or νWT algorithms.

Figure 7 shows $W(m_t)$ for the dilepton events for the $\mathcal{M}$WT analysis and Fig. 8 shows the corresponding functions for the νWT analysis.

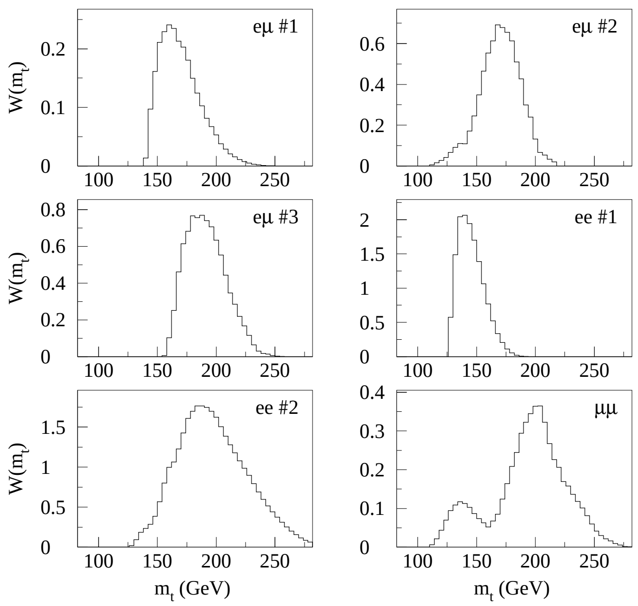


FIG. 7. $W(m_t)$ functions for the dilepton events from the $\mathcal{M}$WT analysis. The labels in the upper right hand corners identify the events (cf. Table IV).

G. Monte Carlo Tests

We now describe tests of the properties of the weight functions to demonstrate their sensitivity to the top quark mass and other parameters.

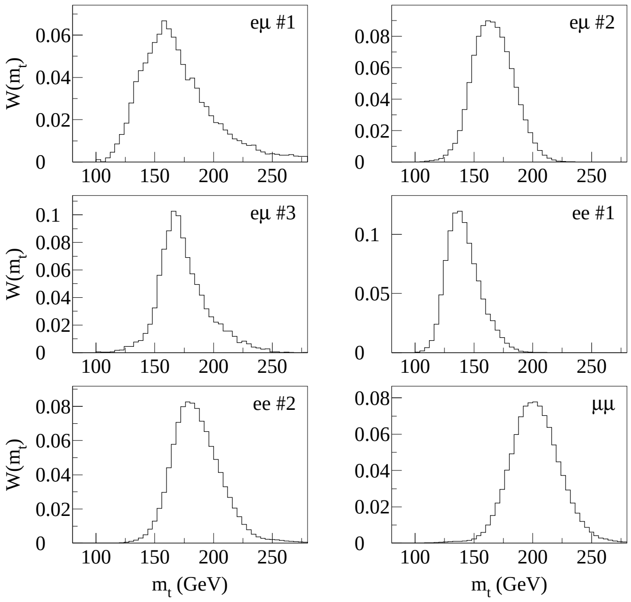


FIG. 8. $W(m_t)$ functions for the dilepton events from the νWT analysis. The labels in the upper right hand corners identify the events (cf. Table IV).

1. Parton-level Tests

Parton-level tests are based on the momenta of the partons generated by the Monte Carlo simulation. Tests at this level are neither subject to effects from detector resolution nor initial or final state radiation. To restrict the sample to events that are broadly similar to those which enter the collider data analysis, the event selection for these tests requires two b quarks and two leptons with $p_T > 20$ GeV and $|\eta| < 2.5$.

We examine the average weight function as a function of input top quark mass by normalizing the area of the weight function for each event to unity and then summing these normalized functions for a collection of Monte Carlo events. A sample of 10,000 events was used, about half of which passed the cuts. The results are shown in Fig. 9 for top quark masses of 130 and 190 GeV. On average, the weight function is sharply peaked within one GeV of the input mass. The tails of the function are asymmetric, with the high-end tail extending further than the low-end tail.

Figure 10 shows the impact of detector resolution, jet combinatorics, and radiation on the weight functions for 190 GeV Monte Carlo events. The distribution becomes significantly broader when resolution effects and both lepton-b jet pairings are considered, but the peak value remains unchanged. Initial state radiation increases the mean value and adds a high-mass tail, as expected. Final state radiation has the opposite effect. In total, the effect of resolution, combinatorics, and radiation is to broaden the distribution of the weight function and move the peak of the distribution away from the input mass.

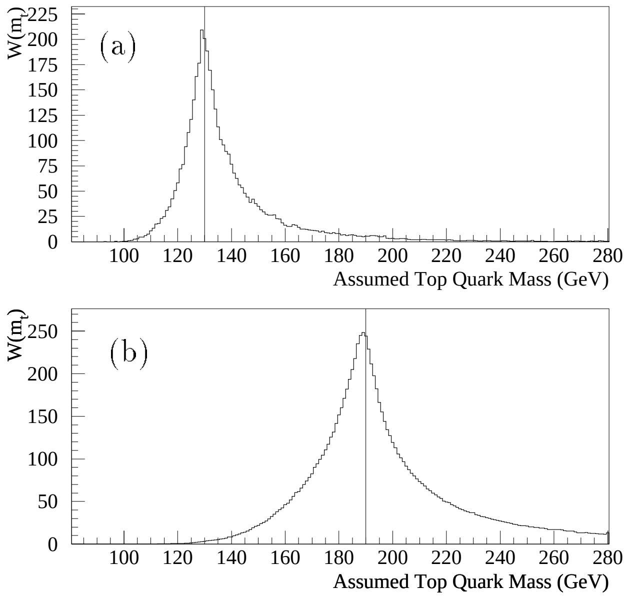


FIG. 9. Average parton-level weight $W(m_t)$ for $t\bar{t}$ decays with (a) $m_t = 130$ GeV and (b) $m_t = 190$ GeV for the νWT algorithm. The vertical lines indicate the input mass values.

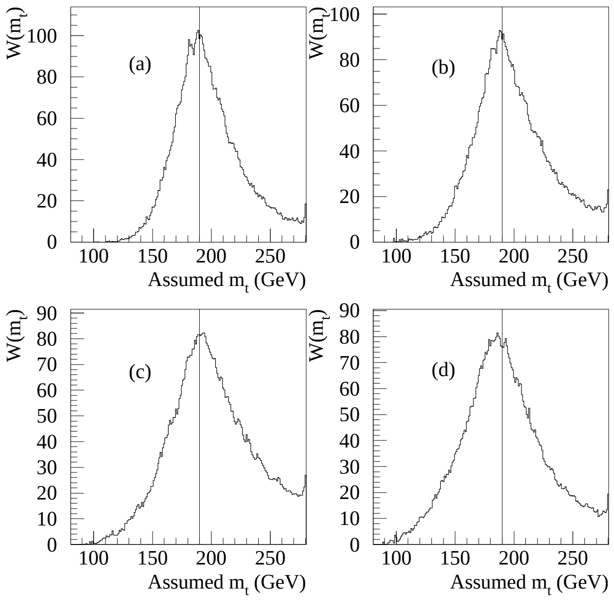


FIG. 10. Average parton-level weight functions for the νWT algorithm, obtained (a) with the parton momenta smeared by the detector resolutions, (b) with the two-fold ambiguity in lepton-jet pairings included, (c) with ISR but without FSR, and (d) without ISR but with FSR. The vertical lines indicate the input mass value of 190 GeV.

2. Tests using Full Simulation

To quantitatively assess the response of the fitting algorithm to events from the DØ data sample that pass the kinematic selection described in Sect. V, we use fully simulated samples of HERWIG $t\bar{t}$ decays. In contrast to the parametrized detector response used in the parton-level tests, these samples derive from a detailed detector model implemented using the GEANT program. The events are processed with the same reconstruction program and filtered using the same kinematic criteria as for the collider data.

Figures 11, 12, and 13 show the average weight functions for the full simulation of all three dilepton channels. Both the kinematic cuts and the additional complexity of the collider environment further degrade the resolution from that obtained in parton-level tests. In particular, for top quark masses less than 140 GeV, the distributions are distorted significantly by the H_T cut. This distortion reduces the precision with which a top mass value in this range can be measured. It does not, however, introduce any bias in our top mass determination since the effect of the H_T cut is modeled in the probability distribution functions used for the mass fits (Sect. VII).

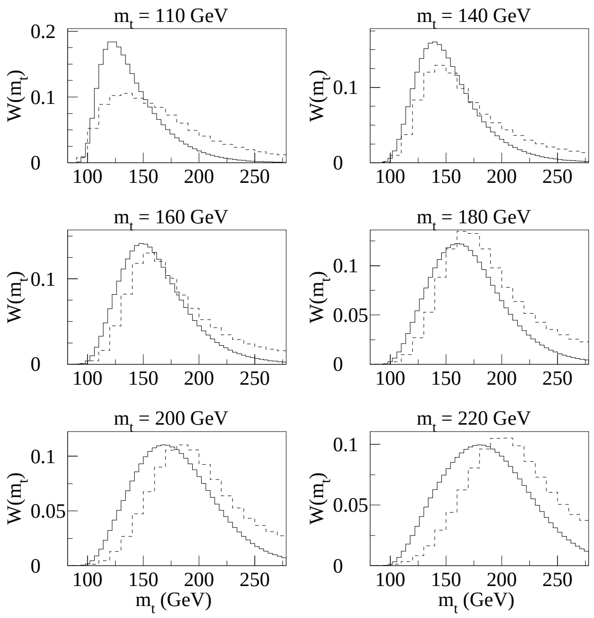


FIG. 11. Average weight functions for fully simulated $t\bar{t}$ decays events in the $e\mu$ channel from the $\mathcal{M}$WT analysis (solid line) and the νWT analysis (dashed line).

The weight distributions become less sharp as the number of muons in the final state increases, reflecting the relatively poor measurement of their momenta. This effect is more pronounced for the νWT analysis. For this reason, and also because the signal to background ratio is significantly higher for the $e\mu$ channel than for the ee or $\mu\mu$ channels, it is important to treat the three channels separately when extracting the top quark mass.

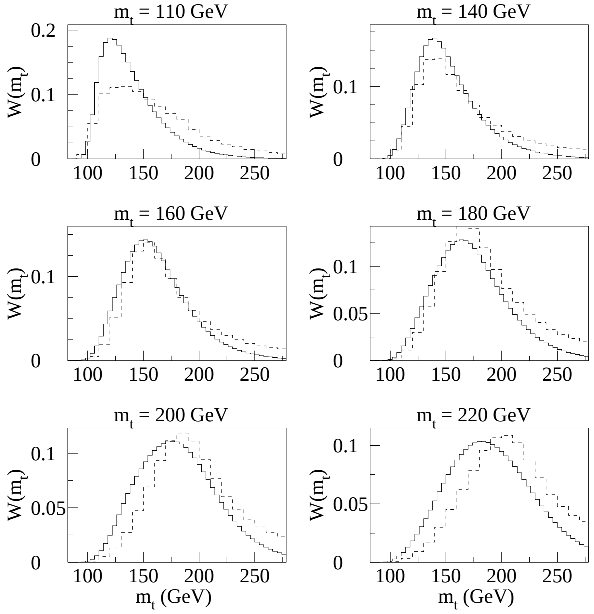


FIG. 12. Average weight functions for fully simulated $t\bar{t}$ decays in the ee channel from the $\mathcal{M}$WT analysis (solid line) and the νWT analysis (dashed line).

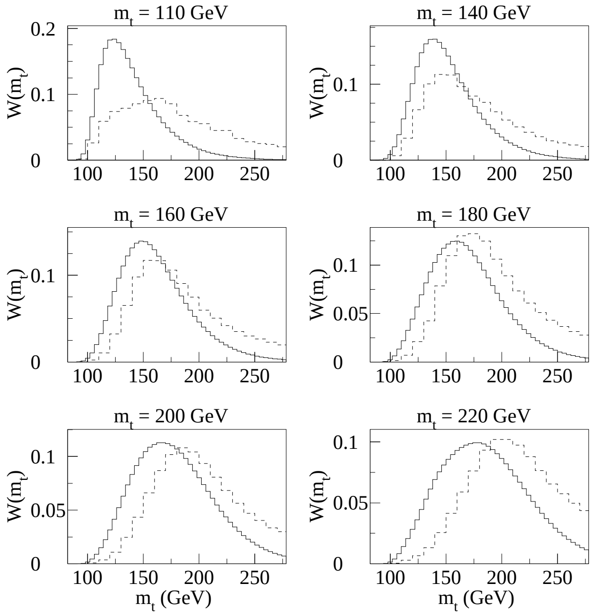


FIG. 13. Average weight functions for fully simulated $t\bar{t}$ decays events in the $\mu\mu$ channel from the $\mathcal{M}$WT analysis (solid line) and the νWT analysis (dashed line).

VII. MASS FITS

A. General Procedure

We estimate the top quark mass by comparing weight functions from Monte Carlo $t\bar{t}$ samples, generated at different values of the top quark mass, with the weight functions for the collider data. We use a maximum likelihood fit to find the value of the top quark mass for which the Monte Carlo predictions agree best with the data.

For each dilepton event, we compute the weights $W(m_t)$ at 50 values of the top quark mass between 80 and 280 GeV. To fit these 50 values directly we would need the probability density as a function of 50 arguments, which is impractical. We can, however, reduce the number of quantities without losing too much information. The individual weight functions are much broader than the size of the steps for which the weights are computed. As shown in Figures 12–13, their rms is 35–40 GeV. Therefore, we integrate the weights over five bins 40 GeV wide, as shown in Fig. 14. Since we need information only about the shape of the weight function, we normalize the area under the function to unity, such that the integrals over four of the bins are independent quantities. We thereby reduce the weight function for each event to the four-dimensional vector

$$\vec{W} = (W_1, W_2, W_3, W_4)\,, \tag{22}$$

where

$$W_1 = \int_{80\ \mathrm{GeV}}^{120\ \mathrm{GeV}} W(m)\,dm \tag{23}$$

and W_2, W_3, and W_4 are computed analogously.

We now maximize the joint likelihood

$$L = \frac{1}{\sqrt{2\pi}\sigma_b} e^{-\frac{(n_b-\overline{n}_b)^2}{2\sigma_b^2}} \frac{(n_s+n_b)^N\, e^{-(n_s+n_b)}}{N!} \times \prod_i^N \frac{n_s f_s(\vec{W}_i|m_t) + n_b f_b(\vec{W}_i)}{n_s+n_b} \tag{24}$$

with respect to the parameters n_s (the expected number of signal events), n_b (the expected number of background events), and m_t (the top quark mass). The product is taken over all events. The first term in the likelihood is a Gaussian constraint that forces the expected number of background events to agree with the background estimate $\overline{n}_b$ within its uncertainty σ_b. The second is a Poisson constraint that forces the expected number of events to be consistent with the observed number of dilepton events N. The remaining part is the probability density for the vector $\vec{W}_i$ for the collider data for n_s signal and n_b background events. Here f_s is the probability density function for signal and f_b for background events. We maximize L with respect to n_s and n_b at each value of m_t using the MINUIT program [22] to eliminate the nuisance parameters n_s and n_b. We are left with L at the discrete values of m_t for which we have Monte Carlo samples. Each dilepton channel is treated separately in this fit and the final likelihood L is the product of the likelihoods from each channel. We fit a polynomial to $-\ln L$, the minimum of which gives the measured value of the top quark mass.

The following sections describe the derivation of the probability density function for $\vec{W}$, the parametrization of the likelihood functions, and the fit results.

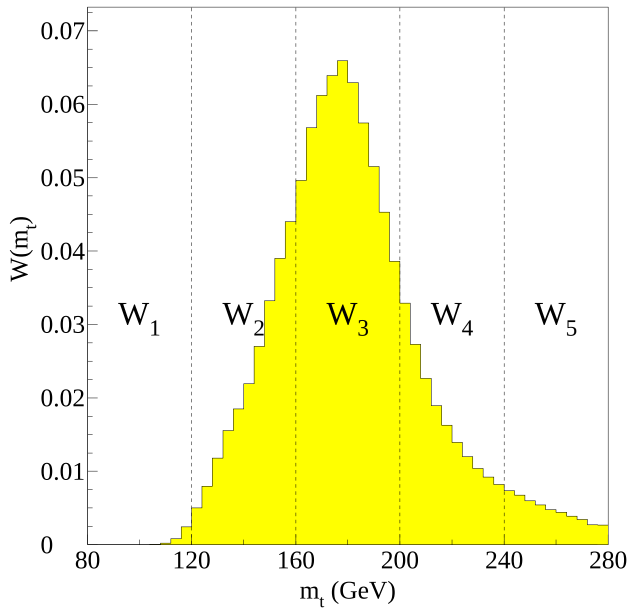


FIG. 14. The weight function for a typical Monte Carlo event, normalized to unity. The vertical lines show the five intervals over which the weight function is integrated.

B. Probability Density Estimation

To estimate the continuous functions f_s and f_b from the discrete sample of Monte Carlo points available for each value of m_t would require a prohibitively large number of Monte Carlo events to populate the four dimensional parameter space. We therefore use a probability density estimation (PDE) technique employing continuous kernels [23].

Consider that each event in the sample is characterized by a set of d uncorrelated values, which are grouped into the d-dimensional vector $\vec{\zeta}$. Then the probability density f for any $\vec{\zeta}$ can be estimated based on a sample of N^{MC} Monte Carlo events as

$$f(\vec{\zeta}) = \frac{1}{N^{\mathrm{MC}} h^d} \sum_{i=1}^{N^{\mathrm{MC}}} K\left(\frac{\vec{\zeta} - \vec{\zeta}_i}{h}, \mathbf{C}\right), \tag{25}$$

where $\mathbf{C}$ is the covariance matrix for the components of $\vec{\zeta}$, h is a free parameter, and K is the kernel function.

Any function which is maximal at zero and asymptotically approaches zero as the absolute value of its argument becomes large would be an acceptable choice for K. For simplicity, we choose a multidimensional Gaussian. In our application, the results of applying either the $\mathcal{M}$WT or νWT techniques to an event is the 4-dimensional vector $\vec{W}$. The elements of $\vec{W}$ are highly correlated, and so a linear transformation must be applied to the data to remove the correlations before using Eq. 25:

$$\vec{W}' = \mathbf{A}\vec{W}. \tag{26}$$

The transformation matrix $\mathbf{A}$ is chosen so that the covariance matrix $\mathbf{C}$ of the transformed variables is diagonal. It can be shown that for two distinct sources of events (signal and background in our case), there exists a unique matrix **A** which results in the covariance matrix for one source to be the identity matrix **I** and that from the other source to be a general diagonal matrix **D** [23]. We choose to have **C** be the identity matrix for background. The matrix **A** is computed only once, using the distribution of Monte Carlo $t\bar{t}$ events generated at all top quark masses. After transformation, the kernel function has the form:

$$K\left(\frac{\vec{W}'-\vec{W}'_i}{h},\mathbf{C}\right)=\prod_{j=1}^{d}\frac{1}{\sqrt{2\pi c_j}}\exp\left(-\frac{\left((\vec{W}'-\vec{W}'_i)_j/h\right)^2}{2c_j}\right) \tag{27}$$

where the c_j are the diagonal elements of **C**.

One minor extension of this method is needed to properly model the background. As described in Sec. V, the backgrounds in the dilepton channel arise from a variety of sources. We assign weight factors b_j such that their contribution to the probability density corresponds to the relative strengths of the n background sources:

$$\frac{b_j N_j^{\mathrm{MC}}}{\sum_{i=1}^{n} b_i N_i^{\mathrm{MC}}}=\frac{\overline{n}_j}{\overline{n}_b}, \tag{28}$$

where N_j^{MC} is the number of Monte Carlo events and $\overline{n}_j$ is the number of events expected from the j^{th} background source. The estimate for the probability density for an event weight vector $\vec{W}$ is then given by:

$$f_s(\vec{W}|m_t)=\frac{1}{Nh^4}\sum_{i=1}^{N}K\left(\frac{\vec{W}'-\vec{W}'_i}{h},\mathbf{D}\right) \tag{29}$$

for signal and

$$f_b(\vec{W})=\frac{1}{\left(\sum_{j=1}^{n} b_j N_j^{\mathrm{MC}}\right)h^4}\sum_{j=1}^{n} b_j\sum_{i=1}^{N_j^{\mathrm{MC}}}K\left(\frac{\vec{W}'-\vec{W}'_i}{h},\mathbf{I}\right) \tag{30}$$

for background.

The remaining step is to fix the value of the free parameter h to maximize the expected resolution of the measurement. Using the ensemble test method described below, we find that values of h in the range $0.1-0.4$ are preferred, and we choose $h=0.3$.

C. Ensemble Tests

Ensemble tests are mock experiments in which the dilepton events are simulated using a Monte Carlo program with a known top quark mass (m_t^{MC}) and processed in exactly the same manner as the collider data. The procedure is as follows: if there are N_j events in the j^{th} decay channel, we draw N_j events from the MC samples for this decay channel. We then select a random number between 0 and 1 for each event. If the random number is greater than $\overline{n}_j/N_j$, we take an event from the signal sample. Otherwise we select an event from the background sample. If there are multiple sources of background, another random number is

TABLE VI. Results of ensemble tests using the νWT algorithm showing the effect of different parametrizations of the $-\ln L$ function. The fits are polynomials of degree m to n points.

Fit		$m_t^{\rm MC} = 150$ GeV			$m_t^{\rm MC} = 200$ GeV		
n	m	Median GeV	Mean GeV	R^{68} GeV	Median GeV	Mean GeV	R^{68} GeV
5	2	152.2	154.1	13.4	198.1	197.8	18.6
7	2	151.6	154.0	13.0	198.2	198.1	19.0
9	2	151.9	154.5	13.6	198.8	199.4	18.9
9	3	151.6	151.8	13.3	196.0	190.0	19.6
11	3	151.9	152.5	13.8	193.4	196.3	19.3

selected in order to decide the source of background from which to draw the event. We then fit the ensemble using the maximum likelihood procedure described above. We repeat this procedure for a large number of ensembles (typically 1000). In this manner we can gauge the statistical properties of the maximum likelihood estimate of the top quark mass, $\widehat{m_t}$.

We characterize the width of the (in general not Gaussian) distribution of fit results by half the length of the shortest interval in m_t that contains 68.3% of the ensembles, R^{68}.

D. Parametrization of the Likelihood Function

We fit a polynomial to the values of $-\ln L$ computed for different top quark masses. The fitted top quark mass is the value of m_t for which the polynomial assumes its minimum $-\ln L_0$. The statistical uncertainty δm_t due to the finite size of the event sample is given by half of the interval in m_t for which $-\ln L < -\ln L_0 + \frac{1}{2}$.

We have a choice of what order polynomial, and how many points around L_0, to include in the fit. The values of $\widehat{m_t}$ and $\delta\widehat{m_t}$ returned by the fit depend on these choices. We therefore perform ensemble tests to select the choice that gives the most accurate values. For the fitted top quark mass this means agreement with the input mass used to generate the ensembles. For the uncertainty it means agreement with the observed scatter of ensemble results.

We fit quadratic and cubic polynomials to five to eleven points, centered on the point of maximum likelihood. Table VI gives the results of ensemble tests using these fitting options. The cubic does not improve the accuracy of the fitted mass and we therefore choose to fit the $-\ln L$ points with a quadratic polynomial.

The width of the fitted quadratic polynomial increases with the number of points included in the fit. We choose the number of points that results in pull distributions of unit widths. If $\widehat{m_t}$ is an unbiased estimate of $m_t^{\rm MC}$ with a Gaussian resolution of width $\delta\widehat{m_t}$, then the pull

$$s = \frac{\widehat{m_t} - m_t^{\rm MC}}{\delta\widehat{m_t}} \tag{31}$$

is normally distributed around zero with unit width. We fit Gaussians to histograms of the

TABLE VII. Pull means and widths from ensemble tests of the $\mathcal{M}$WT algorithm.

$m_t^{\rm MC}$	$n=5$	$n=7$	$n=9$	
GeV	Width	Width	Width	Mean
130	1.16	0.90	0.79	0.65
140	1.01	0.90	0.81	0.38
150	1.12	0.95	0.87	0.13
160	1.34	1.12	1.03	0.12
170	1.26	1.08	0.99	0.11
180	1.24	1.08	0.98	0.00
190	1.12	1.02	1.03	−0.06
200	1.17	1.10	1.06	−0.11
210	1.09	1.04	1.04	−0.09

TABLE VIII. Pull means and widths from ensemble tests of the νWT algorithm.

$m_t^{\rm MC}$	$n=5$	$n=7$	$n=9$	
GeV	Width	Width	Width	Mean
130	1.22	1.04	1.04	0.58
140	1.09	0.97	0.88	0.40
150	1.03	0.92	0.86	0.16
160	1.18	0.99	0.96	0.17
170	1.17	1.06	0.98	0.08
180	1.27	1.11	1.03	0.03
190	1.16	1.05	0.99	−0.07
200	1.07	1.10	1.02	−0.08
210	1.08	1.01	1.03	−0.08

pulls for all ensembles generated with the same $m_t^{\rm MC}$. The pull widths are tabulated in Table VII for the $\mathcal{M}$WT algorithm and in Table VIII for the νWT algorithm.

The fits that include only five points underestimate $\delta\widehat{m}_t$. The nine point fits give pull widths closest to unity over the whole range of m_t. Therefore we choose to fit the quadratic polynomial to nine points for the final results. The pull distributions for ensemble tests at a variety of top quark masses are shown in Fig. 15 for the $\mathcal{M}$WT algorithm and in Fig. 16 for the νWT algorithm.

Tables IX and X list the median and mean fitted top quark masses from ensemble tests using a quadratic fit to nine points. The differences between $\widehat{m}_t$ and $m_t^{\rm MC}$ at masses below 150 GeV can be traced to the small number of events available to model some of the backgrounds ($Z \to \ell\ell$, WW). For these background processes the selection efficiency is so low that a significant increase in the number of Monte Carlo events that satisfy the selection criteria is not possible due to limited computing resources. When we replace these small samples

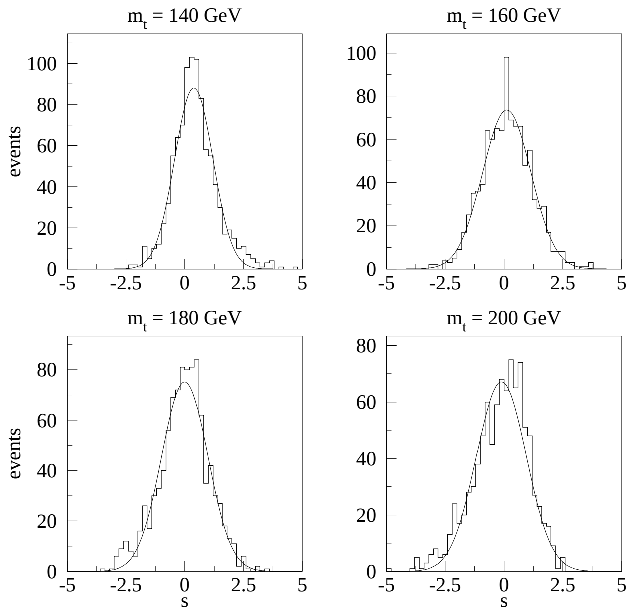


FIG. 15. Pull distributions for the $\mathcal{M}$WT algorithm. The smooth curves are fits to Gaussians.

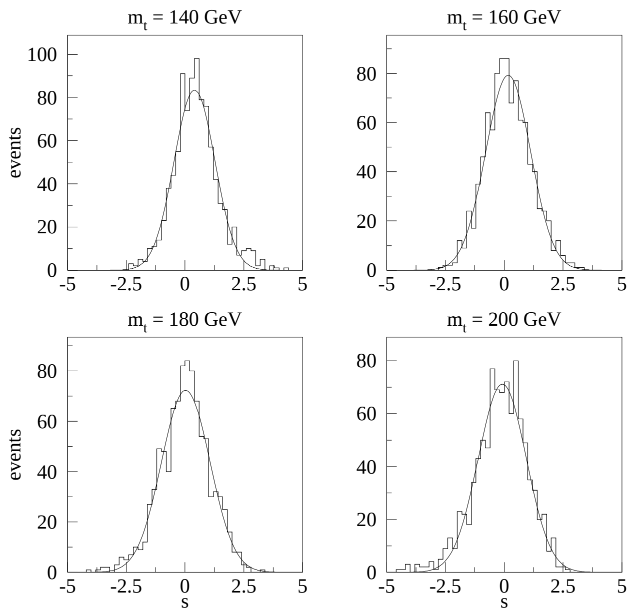


FIG. 16. Pull distributions for the νWT algorithm. The smooth curves are fits to Gaussians.

TABLE IX. Median and mean of the fitted top quark masses and 68% confidence intervals from ensemble tests of the $\mathcal{M}$WT algorithm.

m_t^{MC} GeV	Median GeV	Mean GeV	R^{68} GeV
130	138.1	138.3	13.6
140	144.6	147.1	12.7
150	151.6	153.4	12.8
160	161.6	163.9	15.8
170	172.2	173.7	16.7
180	180.5	181.0	17.3
190	189.5	190.5	17.8
200	200.3	200.1	19.5
210	210.0	210.9	21.4

TABLE X. Median and mean of the fitted top quark masses and 68% confidence intervals from ensemble tests of the νWT algorithm.

m_t^{MC} GeV	Median GeV	Mean GeV	R^{68} GeV
130	138.2	139.8	18.1
140	145.9	147.5	13.9
150	151.9	154.5	13.6
160	161.5	163.5	14.4
170	172.2	173.0	16.2
180	180.5	181.3	18.1
190	188.7	189.6	17.7
200	198.8	199.4	18.9
210	210.1	210.0	20.2

with large samples picked randomly from a smooth distribution these differences vanish. For fitted masses above about 150 GeV, these differences become small. We choose not to correct the results for this effect. It is included in the uncertainty assigned to the fit procedure in Sect. VIII F. Figures 17 and 18 show that for the two algorithms, the peak of the $\widehat{m_t}$ distribution is consistent with $m_t^{\rm MC}$.

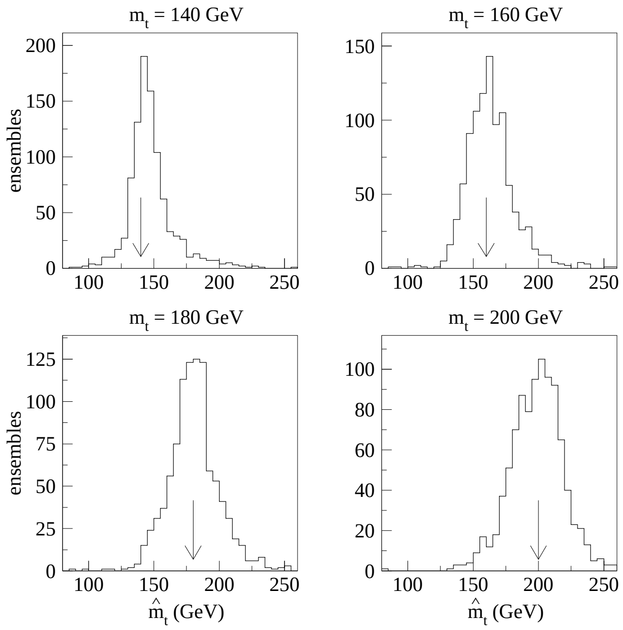


FIG. 17. Distribution of $\widehat{m_t}$ from ensemble tests of the $\mathcal{M}$WT algorithm. The arrows point to the input mass.

E. Results

Applying the procedure outlined above to the dilepton event sample, we find

$$m_t = 168.2 \pm 12.4 \text{ (stat) GeV} \tag{32}$$

for the $\mathcal{M}$WT algorithm and

$$m_t = 170.0 \pm 14.8 \text{ (stat) GeV} \tag{33}$$

for the νWT algorithm. Figures 19 and 20 compare $\sum_i \vec{W}_i$ for collider data to the fitted signal plus background shapes. The insets show the corresponding fits to $-\ln L$.

In Figures 21(a) and 22(a) we compare the statistical uncertainties for the $\mathcal{M}$WT and νWT analyses with the distribution of R^{68} observed in ensemble tests with $m_t^{\rm MC} = 170$ GeV. For the $\mathcal{M}$WT analysis there is a 21% probability to obtain a smaller statistical uncertainty than 12.4 GeV and for the νWT analysis there is a 47% probability to obtain a smaller statistical uncertainty than 14.8 GeV. The pull distributions indicate that $\delta\widehat{m_t}$ is a good estimate of the statistical uncertainty. We verify this by considering the subset of ensembles

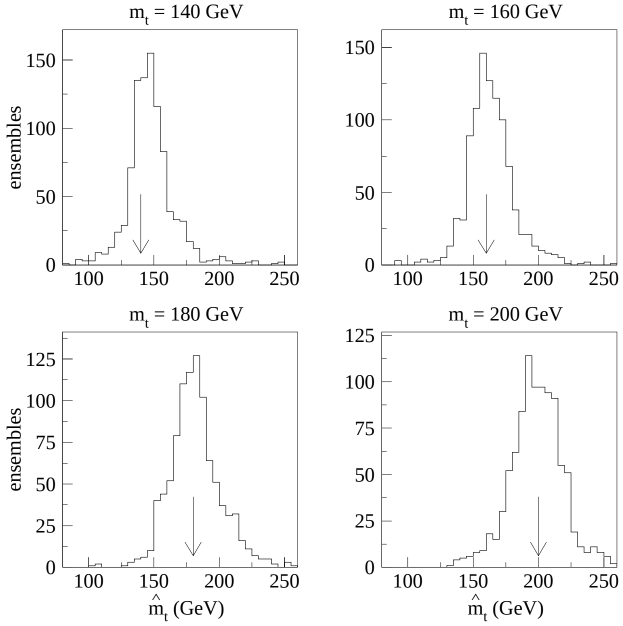


FIG. 18. Distribution of $\widehat{m_t}$ from ensemble tests of the νWT algorithm. The arrows point to the input mass.

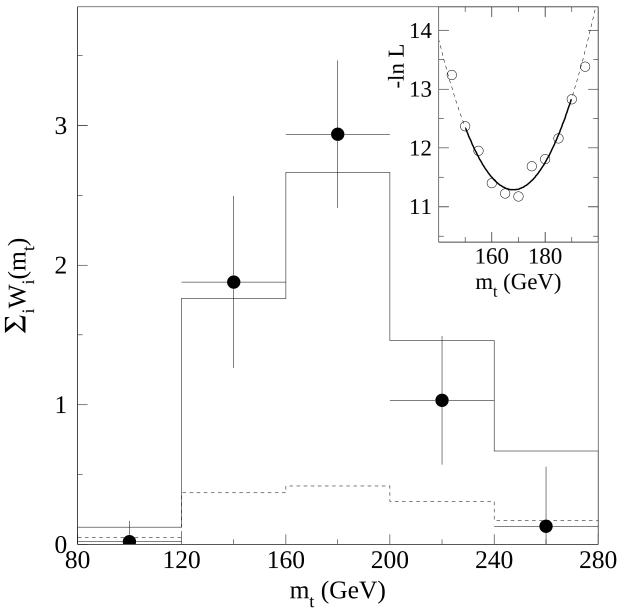


FIG. 19. Summed event weight function $\sum_i \vec{W}_i$ for the data sample (points), the fitted signal plus background (solid), and the background alone (dashed) for the $\mathcal{M}$WT algorithm. The error bars indicate the rms observed for five event samples in ensemble tests. The inset shows the corresponding fit to $-\ln L$, drawn as a solid line in the region considered in the fit.

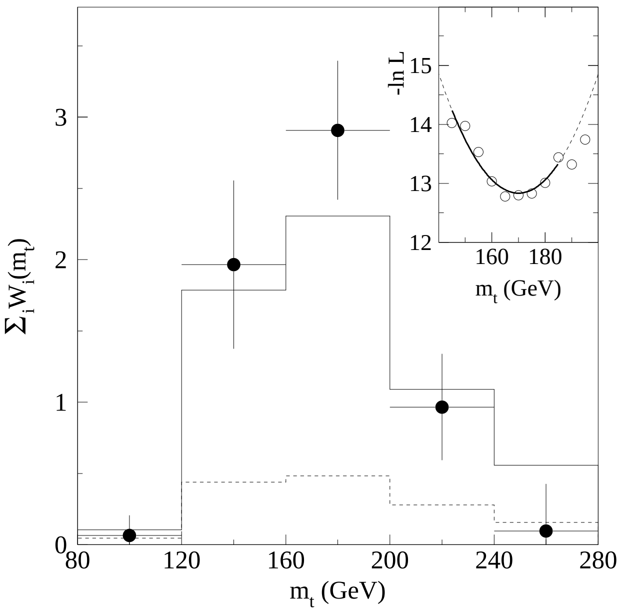


FIG. 20. Summed event weight function $\sum_i \vec{W}_i$ for the data sample (points), the fitted signal plus background (solid), and the background alone (dashed) for the νWT algorithm. The error bars indicate the rms observed for five event samples in ensemble tests. The inset shows the corresponding fit to $-\ln L$, drawn as a solid line in the region considered in the fit.

with $\delta\widehat{m_t}$ consistent with the observed value. Figures 21(b) and 22(b) show the distribution of mass estimates $\widehat{m_t}$ for the ensembles with $\delta\widehat{m_t}$ between the dashed lines in (a). The widths R^{68} of all such ensembles are consistent with the observed values of $\delta\widehat{m_t}$.

The $e\mu$ channel, with the largest number of events and smallest background, should dominate the result of the fit, while the $\mu\mu$ channel with only one event and a sizeable background should have the least effect. We therefore also fit separately the five events from the ee and $e\mu$ samples and the three $e\mu$ events. Table XI lists the results. This table also shows the effect of varying the degree of the polynomial used to fit $-\ln L$ and the number of points included in the fit. No excursions comparable to the statistical uncertainty of the measurement are seen in the results of any of these variations.

VIII. SYSTEMATIC ERRORS

A. Estimation of Systematic Uncertainties

Systematic uncertainties give rise to biases in the result of the analysis no matter how many events are analyzed. They are due to differences between the collider data and our signal or background models. Variation in the event selection or the fit procedure, which in general also result in a change in the final result when applied to a small sample of events, do not represent systematic uncertainties. Rather, these are statistical effects and are properly accounted for by our use of a maximum likelihood fit to define the statistical uncertainty.

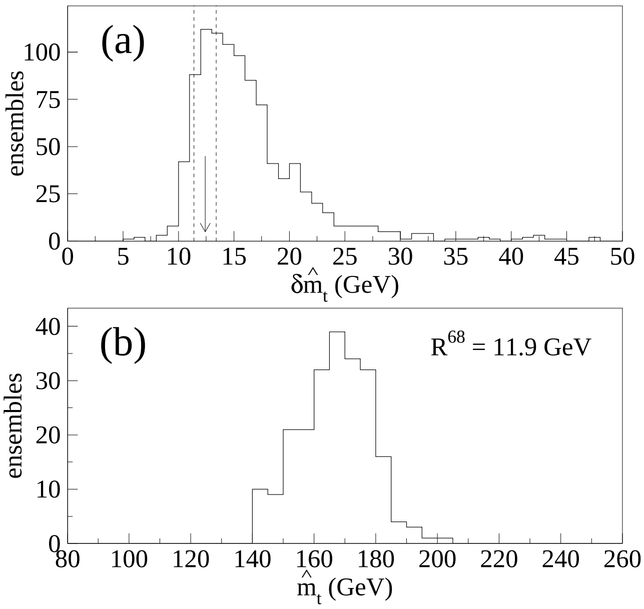


FIG. 21. (a) Distribution of uncertainties $\delta\widehat{m_t}$ obtained from ensemble tests for the $\mathcal{M}$WT algorithm with $m_t^{\rm MC} = 170$ GeV. The arrow marks the value returned by the fit to the data (12.4 GeV). (b) Distribution of $\widehat{m_t}$ for the ensembles with $\delta\widehat{m_t}$ between the dashed lines in (a).

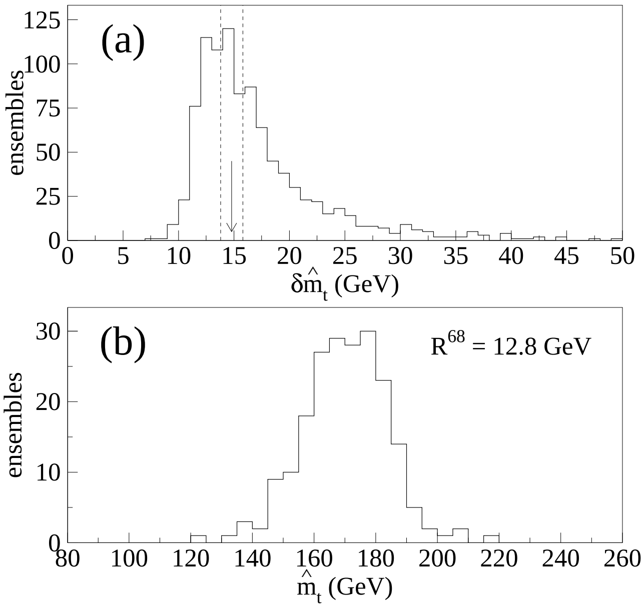


FIG. 22. (a) Distribution of uncertainties $\delta\widehat{m_t}$ obtained from ensemble tests of the νWT algorithm with $m_t^{\rm MC} = 170$ GeV. The arrow marks the value returned by the fit to the data (14.8 GeV). (b) Distribution of $\widehat{m_t}$ for the ensembles with $\delta\widehat{m_t}$ between the dashed lines in (a).

TABLE XI. Results of several variations of the maximum likelihood fit to the data. The fits are polynomials of degree m to n points.

Channels	Fit		Fitted Mass (GeV)	
	n	m	$\mathcal{M}$WT	νWT
$ee, e\mu, \mu\mu$	5	2	166±12	169±11
	7	2	168±12	170±13
	9	2	168±12	170±15
	11	3	167^{+11}_{-13}	171±16
$ee, e\mu$	5	2	166±13	173±12
	7	2	167±12	172±15
	9	2	168±13	173±14
	11	3	166^{+11}_{-13}	172^{+13}_{-15}
$e\mu$	5	2	173±15	169±14
	7	2	173±13	169±13
	9	2	173±13	170±15
	11	3	172^{+13}_{-15}	170^{+15}_{-16}

Systematic uncertainties can, in general, be estimated using ensemble tests in which a mismatch is introduced between the conditions under which the ensembles are created, and the assumptions used in the probability density estimation. In most cases we vary conditions in the ensembles and then analyse them with the same probability density functions used for the collider data, *i.e.*, assuming the nominal conditions. Any deviation of the fitted mass values from the mass used when generating the ensembles indicates a systematic effect. Due to the finite number of Monte Carlo events available, these systematic effects can be estimated with an uncertainty of about 1 GeV. Table XII summarizes the sources of systematic uncertainties and their estimated magnitudes. The estimated uncertainties differ insignificantly between the two algorithms so that we use the average of the uncertainties from both analyses, weighted by the respective statistical uncertainty in the measured top quark mass, as an estimate for both algorithms. The following sections describe the individual uncertainties in more detail.

B. Jet Energy Scale

To propagate the jet energy scale uncertainty (section IV C) to the top mass measurement, we generate signal Monte Carlo samples ($m_t = 170$ GeV) and background samples with jet energy responses one standard deviation higher and lower than the nominal response. We also scale the energy in the calorimeter that is not included in any jet by the same factor as the jets, and the $\not{p}_T$ is recomputed to reflect the scale change. We then create Monte Carlo ensembles from the scaled samples and fit them using the probability density functions generated with the nominal jet energy response. Table XIII shows the results of this mismatch in jet energy scale. Averaging the upward and downward excursions of the

TABLE XII. Summary of systematic uncertainties for the dilepton mass fits.

Source	Uncertainty (GeV)		
	$\mathcal{M}$WT	νWT	average
Jet Energy Scale	2.0	2.9	2.4
Multiple Interactions	1.4	1.2	1.3
Background Model	0.9	1.5	1.1
Signal Generator	2.3	1.1	1.8
Monte Carlo Sample Size	0.3	0.3	0.3
Likelihood Fit	0.9	1.3	1.1
Total	3.5	3.9	3.6

TABLE XIII. Effect of varying the jet energy response in ensemble tests with $m_t = 170$ GeV.

Jet Scale	Median $\widehat{m_t}$ (GeV)	
	$\mathcal{M}$WT	νWT
$+2.5\% + 0.5$ GeV	172.9	174.0
Nominal	172.2	172.2
$-2.5\% - 0.5$ GeV	168.9	168.3

median results in a systematic uncertainty of 2.0 GeV for the $\mathcal{M}$WT algorithm and 2.9 GeV for the νWT algorithm.

C. Signal Monte Carlo Generator

The accurate determination of the top quark mass depends on the signal Monte Carlo providing a faithful description of $t\bar{t}$ events. Some features, in particular gluon radiation and parton fragmentation, are only modeled approximately by HERWIG and other reasonable approximations exist. In the absence of large samples of $t\bar{t}$ events, none of them can be directly excluded. To test the sensitivity of the result to the Monte Carlo generator, we generate ensembles of events with the ISAJET event generator. We simulate the detector response using GEANT and analyse them in the standard way. We then fit the weight functions of ensembles of these events with the probability density functions obtained from Monte Carlo events generated by the HERWIG program. Tables XIV and XV list the results. For a given top quark mass, we take the difference ΔMedian between the medians of the results from the ISAJET samples (Tables XIV and XV) and the HERWIG samples (Tables IX and X). We compute the average of the magnitude of these differences for all top quark masses, 2.3 GeV for the $\mathcal{M}$WT algorithm and 1.1 GeV for the νWT algorithm, and assign these values as the systematic uncertainty in the top quark mass measurement.

In addition, we have performed studies to directly assess the impact of gluon radiation by varying the fraction of events with gluon radiation in a HERWIG Monte Carlo sample by

TABLE XIV. Results of analyzing ensembles of events generated by ISAJET with the $\mathcal{M}$WT algorithm.

m_t^{MC} GeV	Median GeV	Mean GeV	R^{68} GeV	ΔMedian GeV	ΔMean GeV
140	143.6	145.0	14.4	−1.0	−2.1
150	151.0	151.6	14.3	−0.6	−1.8
160	160.0	161.4	16.4	−1.6	−2.5
170	169.0	168.6	17.3	−3.2	−5.1
180	178.0	178.4	18.0	−2.5	−2.6
190	186.2	186.9	19.8	−3.3	−3.6
200	197.2	196.1	20.2	−3.1	−4.0
210	206.7	206.1	22.1	−3.3	−4.8

TABLE XV. Results of analyzing ensembles of events generated by ISAJET with the νWT algorithm.

m_t^{MC} GeV	Median GeV	Mean GeV	R^{68} GeV	ΔMedian GeV	ΔMean GeV
140	145.9	147.8	15.6	0.0	0.3
150	152.6	154.4	15.4	0.7	−0.1
160	160.1	161.6	15.8	−1.4	−1.9
170	170.8	171.6	17.6	−1.4	−1.4
180	179.1	179.5	18.2	−1.4	−1.8
190	189.4	188.7	18.5	0.7	−0.9
200	198.6	198.3	19.5	−0.2	−1.1
210	206.8	205.6	20.3	−3.3	−4.4

50%. This results in a change of 1.3 GeV in the measured top quark mass, which is quite consistent with the uncertainties quoted above based on HERWIG-ISAJET differences.

We studied the sensitivity of the results to variations in our choice of parton distribution functions. We expect the sensitivity to parton distribution functions to be larger for the $\mathcal{M}$WT analysis because it uses them explicitly in the mass reconstruction. Our default choice is the CTEQ3M set of parton distribution functions [24]. We also perform ensemble tests with weight functions derived using MRSA′ parton distribution functions [25] with three different values of Λ_{QCD}. The Monte Carlo events for the ensembles were generated with an input mass of 170 GeV and CTEQ3M parton distribution functions in the generation and the top mass reconstruction. The results are summarized in Table XVI. The variation in the median of the ensemble tests is 20 MeV. We conclude that any sensitivity to parton distribution functions is negligible compared to other systematic effects in the generation of the Monte Carlo samples.

TABLE XVI. Results of varying the choice of parton distribution functions (pdf) in the $\mathcal{M}$WT analysis.

pdf	Median GeV	Mean GeV
CTEQ3M	172.25	173.67
MRSA$'$ ($\Lambda_{\rm QCD} = 266$ MeV)	172.27	173.66
MRSA$'$ ($\Lambda_{\rm QCD} = 344$ MeV)	172.27	173.51
MRSA$'$ ($\Lambda_{\rm QCD} = 435$ MeV)	172.26	173.38

TABLE XVII. Effect of introducing dummy models for the poorly modeled portion of the background.

Background Model	Median $\widehat{m_t}$ (GeV)	
	$\mathcal{M}$WT	νWT
Low Mass	172.9	172.7
Nominal	172.2	172.2
High Mass	172.0	171.2

D. Background Shape

The modeling of the background also depends on a Monte Carlo simulation. In addition, for some sources of background ($Z \to \ell\ell$, WW) very few Monte Carlo events satisfy the selection criteria. To estimate how sensitive the result is to the poorly constrained distribution of these events, we use dummy models instead of the Monte Carlo samples. These models assume that the $W(m_t)$ distributions for these backgrounds are Gaussian, with a width chosen randomly between 20 and 60 GeV. In one of the models ("low mass"), the mean of the Gaussian was randomly selected between 120 and 160 GeV, and in the other ("high mass") between 180 and 220 GeV. We then perform ensemble tests using the known background components plus the dummies to estimate the background probability densities, with events drawn from the standard signal and background models. The results are listed in Table XVII. Based on the observed shifts in the median $\widehat{m_t}$ the uncertainties are 0.9 GeV and 1.5 GeV for the $\mathcal{M}$WT and νWT analyses, respectively.

E. Multiple Interactions

The beams in the Tevatron are structured into six proton and six antiproton bunches. Proton and antiproton bunches collide every 3.5 μs in the center of the detector. More than one $p\overline{p}$ interaction can take place during a crossing and the detector sees the superposition of all these interactions. At the mean luminosity at which the data were taken ($7.5 \times 10^{30}/\mathrm{cm}^2/\mathrm{s}$) on average 1.3 interactions occur per crossing. Since the cross section for the production of high-p_T secondaries is small, it is very unlikely that more than one of these interactions produces high-p_T particles or jets. However, the Monte Carlo models do not include the effect of the additional low-p_T particles due to multiple interactions during the same crossing.

There are two ways in which these additional interactions may affect the reconstructed event. First, the additional particles deposit energy in the calorimeter, some of which falls into the jet cones. Second, the additional tracks may confuse the algorithm that determines the z-position of the interaction vertex, leading to mismeasurement of the jet directions. The jet energy scale calibration accounts for the former effect on average. To study the latter effect, we add particles from one or two simulated additional $p\overline{p}$ interactions to a sample of 5000 Monte Carlo $t\overline{t}$ decays with $m_t = 170$ GeV. The signatures of the resulting events in the detector are simulated by the GEANT program. The events are reconstructed by the same programs as the collider data. For this study ensemble tests are of little help, since the small sample sizes prohibit the generation of a large number of independent ensembles. We estimate the size of the systematic effect by comparing the $W(m_t)$ distributions in the samples with zero, one, and two additional interactions. Although the resolution of the z vertex degrades with the additional interaction, the effect on the $W(m_t)$ distribution is modest. The difference in mean between a sample without additional interactions and the sample in which 33% of the events have one and 36% two additional interactions, approximating the conditions at which the data were taken, is only 0.6 GeV for the νWT analysis. A change of this magnitude is roughly equivalent to a change of 1.2 GeV in the top quark mass. For the $\mathcal{M}$WT analysis we get a similar value, 1.4 GeV.

F. Likelihood Fit and Monte Carlo Statistics

There are systematic uncertainties in the value of the top quark mass that minimizes $-\ln L$. These arise both from the finite number of Monte Carlo events used in determining the $-\ln L$ points and the choice of function to fit these points.

To estimate the effect of the Monte Carlo sample size, we split the signal Monte Carlo samples into five subsets and repeat the fit to the data using each subset as the signal model. The rms variation observed in the central value is then divided by $\sqrt{5}$, yielding a systematic uncertainty of 0.3 GeV for either algorithm.

To estimate the uncertainty arising from the choice of the parabolic fit to nine likelihood points, we fit Monte Carlo ensembles with $m_t = 170$ GeV using a variety of parametrizations and observe the resulting changes in the median of $\widehat{m}_t$. We fit quadratic polynomials to five and seven points and cubic polynomials to nine and eleven points. The largest variations of 0.9 GeV ($\mathcal{M}$WT) and 1.3 GeV (νWT) give estimates of the systematic uncertainties.

IX. RESULTS

A. Combination of the $\mathcal{M}$WT and νWT Measurements

The two algorithms we use give consistent results. The weights computed by the $\mathcal{M}$WT and νWT algorithms are based on different aspects of $t\overline{t}$ production and decay and are therefore not completely correlated. To gauge the degree of correlation, we fit ensembles of $t\bar{t}$ Monte Carlo events for a top quark mass of 170 GeV using both algorithms. We then select the subset of these ensembles with likelihood functions of similar widths as observed in the data (*i.e.* those for which the $\mathcal{M}$WT analysis yields $11.4 < \delta m_t < 13.4$ GeV and the νWT analysis yields $13.8 < \delta m_t < 15.8$ GeV). Based on these tests we find that the correlation coefficient between the $\mathcal{M}$WT and νWT algorithms is 0.77. A statistical combination of the results from the two algorithms then yields

$$m_t = 168.4 \pm 12.3\ (\mathrm{stat}) \pm 3.6\ (\mathrm{syst})\ \mathrm{GeV}. \tag{34}$$

The systematic uncertainties are taken as completely correlated between the two algorithms. Since they differ insignificantly between the two algorithms we quote the mean from Table XII.

B. Combination of the Dilepton and Lepton+Jets Measurements

The value of the top quark mass obtained from the dilepton channel is in good agreement with that found by fitting $t\bar{t} \to \ell$+jets events [7], supporting the hypothesis that both are due to the decays of the same pair-produced particles. We obtain our best measurement of the mass of the top quark by combining the results of the analyses in the two channels. Since the two measurements are statistically independent the combination is straight forward. The systematic uncertainties in the combined measurement are evaluated by propagating the uncertainties in each channel with correlation coefficients of either 0 (for MC statistics, likelihood fit, and background model) or 1 (for jet energy scale, multiple interactions, and HERWIG–ISAJET differences). We obtain

$$m_t = 172.1 \pm 5.2\ (\mathrm{stat}) \pm 4.9\ (\mathrm{syst})\ \mathrm{GeV}. \tag{35}$$

The effective correlation coefficient between the two measurements is 0.15. If we neglected all correlations the result would change by less than 200 MeV.

C. Conclusions

We have reported the measurement of the top quark mass using six dilepton events. We use maximum likelihood fits to the dynamics of the decays to achieve maximum sensitivity to the mass of the top quark. We developed two algorithms for the computation of the likelihood that exploit complementary features of $t\bar{t}$ production and decay. Both result in very similar measurements of the top quark mass. They also agree well with the mass measured from fits to $t\bar{t} \to \ell +$jets events, supporting the hypothesis that both channels correspond to decays of the same particle. We combine the mass measurements from both channels to obtain

$$m_t = 172.1 \pm 7.1\ \mathrm{GeV}. \tag{36}$$

ACKNOWLEDGEMENTS

We thank the staffs at Fermilab and collaborating institutions for their contributions to this work, and acknowledge support from the Department of Energy and National Science Foundation (U.S.A.), Commissariat à L'Energie Atomique (France), Ministry for Science and Technology and Ministry for Atomic Energy (Russia), CAPES and CNPq (Brazil), Departments of Atomic Energy and Science and Education (India), Colciencias (Colombia), CONACyT (Mexico), Ministry of Education and KOSEF (Korea), and CONICET and UBACyT (Argentina).

REFERENCES

* Visitor from Universidad San Francisco de Quito, Quito, Ecuador.

† Visitor from IHEP, Beijing, China.

[1] S.L. Glashow, Nucl. Phys. **22**, 579 (1961); S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967); A. Salam, *Proceedings of the 8^{th} Nobel Symposium*, ed. N. Svartholm (Almqvist and Wiksells, Stockholm 1968), p. 367.

[2] G. Degrassi *et al.* Phys. Lett. B **418**, 209 (1998); G. Degrassi, P. Gambino, and A. Sirlin, Phys. Lett. B **394**, 188 (1997).

[3] The LEP Collaborations, the LEP Electroweak Working Group, and the SLD Heavy Flavour Group, CERN-PPE/97-154 (unpublished).

[4] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **74**, 2626 (1995); S. Abachi *et al.* (DØ Collaboration), Phys. Rev. Lett. **74**, 2632 (1995).

[5] B. Abbott *et al.* (DØ Collaboration), Phys. Rev. Lett. **80**, 2063 (1998).

[6] E.W. Varnes, Ph.D. thesis, University of California, Berkeley (1997) (unpublished), `http://www-d0. fnal.gov/publications_talks/thesis/thesis. html`.

[7] S. Abachi *et al.* (DØ Collaboration), Phys. Rev. Lett. **79**, 1197 (1997); B. Abbott *et al.* (DØ Collaboration), to be published in Phys. Rev. D, FERMILAB-PUB-98/031-E.

[8] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **80**, 2767 (1998).

[9] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **80**, 2779 (1998).

[10] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **79**, 1992 (1997).

[11] S. Abachi *et al.* (DØ Collaboration), Nucl. Instrum. Methods in Phys. Res. **A338**, 185 (1994).

[12] S. Abachi *et al.* (DØ Collaboration), Phys. Rev. Lett. **79**, 1203 (1997).

[13] S. Abachi *et al.* (DØ Collaboration), Phys. Rev. D**52**, 4877 (1995).

[14] B. Abbott *et al.* (DØ Collaboration), Phys. Rev. Lett. **80**, 3008 (1998); *ibid.* to be published in Phys. Rev. D, FERMILAB-Pub-97/424-E.

[15] B. Abbott *et al.* (DØ Collaboration), submitted to Nucl. Instrum. Methods in Phys. Res. A, Fermilab-Pub-97/330-E. This work uses a slightly different version of these corrections, described in R. Kehoe (DØ Collaboration), Proc. 6^{th} Inernational Conf. on Calorimetry in High Energy Physics, Frascati (1996).

[16] G. Marchesini *et al.*, Comput. Phys. Commun. **67** 465 (1992).

[17] F. Carminati *et al.*, GEANT *Users Guide*, CERN Program Library W5013, 1991 (unpublished).

[18] S. Abachi *et al.* (DØ Collaboration), Phys. Rev. Lett. **75**, 1034 (1995).

[19] K. Kondo, J. Phys. Soc. Jpn. **57**, 4126 (1988); *ibid.* **60**, 836 (1991).

[20] R.H. Dalitz and G.R. Goldstein, Phys. Rev. D**45**, 1531 (1992).

[21] F. Paige and S. Protopopescu, BNL Report No. BNL38034, 1986 (unpublished).

[22] F. James, CERN Program Library D506, 1978 (unpublished).

[23] L. Holmström, R. Sain, and H.E. Miettinen, Comp. Phys. Comm. **88**, 195 (1995).

[24] H.L. Lai *et al.*, Phys. Rev. D **51**, 4763 (1995).

[25] A.D. Martin, W.J. Stirling, and R.G. Roberts, Phys. Rev. D **50**, 6734 (1994) and Phys. Rev. D **51**, 4756 (1995).