FERMILAB-08-415

The first measurement of the top quark mass at CDF II in the lepton+jets and dilepton channels simultaneously

T. Aaltonen,[24] J. Adelman,[14] T. Akimoto,[56] M.G. Albrow,[18] B. Álvarez González,[12] S. Amerio[x],[44] D. Amidei,[35] A. Anastassov,[39] A. Annovi,[20] J. Antos,[15] G. Apollinari,[18] A. Apresyan,[49] T. Arisawa,[58] A. Artikov,[16] W. Ashmanskas,[18] A. Attal,[4] A. Aurisano,[54] F. Azfar,[43] P. Azzurri[aa],[47] W. Badgett,[18] A. Barbaro-Galtieri,[29] V.E. Barnes,[49] B.A. Barnett,[26] V. Bartsch,[31] G. Bauer,[33] P.-H. Beauchemin,[34] F. Bedeschi,[47] D. Beecher,[31] S. Behari,[26] G. Bellettini[y],[47] J. Bellinger,[60] D. Benjamin,[17] A. Beretvas,[18] J. Beringer,[29] A. Bhatti,[51] M. Binkley,[18] D. Bisello[x],[44] I. Bizjak[dd],[31] R.E. Blair,[2] C. Blocker,[7] B. Blumenfeld,[26] A. Bocci,[17] A. Bodek,[50] V. Boisvert,[50] G. Bolla,[49] D. Bortoletto,[49] J. Boudreau,[48] A. Boveia,[11] B. Brau[a],[11] A. Bridgeman,[25] L. Brigliadori,[44] C. Bromberg,[36] E. Brubaker,[14] J. Budagov,[16] H.S. Budd,[50] S. Budd,[25] S. Burke,[18] K. Burkett,[18] G. Busetto[x],[44] P. Bussey[k],[22] A. Buzatu,[34] K. L. Byrum,[2] S. Cabrera[u],[17] C. Calancha,[32] M. Campanelli,[36] M. Campbell,[35] F. Canelli,[18] A. Canepa,[46] B. Carls,[25] D. Carlsmith,[60] R. Carosi,[47] S. Carrillo[m],[19] S. Carron,[34] B. Casal,[12] M. Casarsa,[18] A. Castro[w],[6] P. Catastini[z],[47] D. Cauz[cc],[55] V. Cavaliere[z],[47] M. Cavalli-Sforza,[4] A. Cerri,[29] L. Cerrito[n],[31] S.H. Chang,[28] Y.C. Chen,[1] M. Chertok,[8] G. Chiarelli,[47] G. Chlachidze,[18] F. Chlebana,[18] K. Cho,[28] D. Chokheli,[16] J.P. Chou,[23] G. Choudalakis,[33] S.H. Chuang,[53] K. Chung,[13] W.H. Chung,[60] Y.S. Chung,[50] T. Chwalek,[27] C.I. Ciobanu,[45] M.A. Ciocci[z],[47] A. Clark,[21] D. Clark,[7] G. Compostella,[44] M.E. Convery,[18] J. Conway,[8] M. Cordelli,[20] G. Cortiana[x],[44] C.A. Cox,[8] D.J. Cox,[8] F. Crescioli[y],[47] C. Cuenca Almenar[u],[8] J. Cuevas[r],[12] R. Culbertson,[18] J.C. Cully,[35] D. Dagenhart,[18] M. Datta,[18] T. Davies,[22] P. de Barbaro,[50] S. De Cecco,[52] A. Deisher,[29] G. De Lorenzo,[4] M. Dell'Orso[y],[47] C. Deluca,[4] L. Demortier,[51] J. Deng,[17] M. Deninno,[6] P.F. Derwent,[18] G.P. di Giovanni,[45] C. Dionisi[bb],[52] B. Di Ruzza[cc],[55] J.R. Dittmann,[5] M. D'Onofrio,[4] S. Donati[y],[47] P. Dong,[9] J. Donini,[44] T. Dorigo,[44] S. Dube,[53] J. Efron,[40] A. Elagin,[54] R. Erbacher,[8] D. Errede,[25] S. Errede,[25] R. Eusebi,[18] H.C. Fang,[29] S. Farrington,[43] W.T. Fedorko,[14] R.G. Feild,[61] M. Feindt,[27] J.P. Fernandez,[32] C. Ferrazza[aa],[47] R. Field,[19] G. Flanagan,[49] R. Forrest,[8] M.J. Frank,[5] M. Franklin,[23] J.C. Freeman,[18] I. Furic,[19] M. Gallinaro,[52] J. Galyardt,[13] F. Garberson,[11] J.E. Garcia,[21] A.F. Garfinkel,[49] K. Genser,[18] H. Gerberich,[25] D. Gerdes,[35] A. Gessler,[27] S. Giagu[bb],[52] V. Giakoumopoulou,[3] P. Giannetti,[47] K. Gibson,[48] J.L. Gimmell,[50] C.M. Ginsburg,[18] N. Giokaris,[3] M. Giordani[cc],[55] P. Giromini,[20] M. Giunta[y],[47] G. Giurgiu,[26] V. Glagolev,[16] D. Glenzinski,[18] M. Gold,[38] N. Goldschmidt,[19] A. Golossanov,[18] G. Gomez,[12] G. Gomez-Ceballos,[33] M. Goncharov,[54] O. González,[32] I. Gorelov,[38] A.T. Goshaw,[17] K. Goulianos,[51] A. Gresele[x],[44] S. Grinstein,[23] C. Grosso-Pilcher,[14] R.C. Group,[18] U. Grundler,[25] J. Guimaraes da Costa,[23] Z. Gunay-Unalan,[36] C. Haber,[29] K. Hahn,[33] S.R. Hahn,[18] E. Halkiadakis,[53] B.-Y. Han,[50] J.Y. Han,[50] F. Happacher,[20] K. Hara,[56] D. Hare,[53] M. Hare,[57] S. Harper,[43] R.F. Harr,[59] R.M. Harris,[18] M. Hartz,[48] K. Hatakeyama,[51] C. Hays,[43] M. Heck,[27] A. Heijboer,[46] J. Heinrich,[46] C. Henderson,[33] M. Herndon,[60] J. Heuser,[27] S. Hewamanage,[5] D. Hidas,[17] C.S. Hill[c],[11] D. Hirschbuehl,[27] A. Hocker,[18] S. Hou,[1] M. Houlden,[30] B.T. Huffman,[43] R.E. Hughes,[40] U. Husemann,[61] J. Huston,[36] J. Incandela,[11] G. Introzzi,[47] M. Iori[bb],[52] A. Ivanov,[8] E. James,[18] B. Jayatilaka,[17] E.J. Jeon,[28] M.K. Jha,[6] S. Jindariani,[18] W. Johnson,[8] M. Jones,[49] K.K. Joo,[28] S.Y. Jun,[13] J.E. Jung,[28] T.R. Junk,[18] T. Kamon,[54] D. Kar,[19] P.E. Karchin,[59] Y. Kato,[42] R. Kephart,[18] J. Keung,[46] V. Khotilovich,[54] B. Kilminster,[18] D.H. Kim,[28] H.S. Kim,[28] H.W. Kim,[28] J.E. Kim,[28] M.J. Kim,[20] S.B. Kim,[28] S.H. Kim,[56] Y.K. Kim,[14] N. Kimura,[56] L. Kirsch,[7] S. Klimenko,[19] B. Knuteson,[33] B.R. Ko,[17] K. Kondo,[58] D.J. Kong,[28] J. Konigsberg,[19] A. Korytov,[19] A.V. Kotwal,[17] M. Kreps,[27] J. Kroll,[46] D. Krop,[14] N. Krumnack,[5] M. Kruse,[17] V. Krutelyov,[11] T. Kubo,[56] T. Kuhr,[27] N.P. Kulkarni,[59] M. Kurata,[56] Y. Kusakabe,[58] S. Kwang,[14] A.T. Laasanen,[49] S. Lami,[47] S. Lammel,[18] M. Lancaster,[31] R.L. Lander,[8] K. Lannon[q],[40] A. Lath,[53] G. Latino[z],[47] I. Lazzizzera[x],[44] T. LeCompte,[2] E. Lee,[54] H.S. Lee,[14] S.W. Lee[t],[54] S. Leone,[47] J.D. Lewis,[18] C.S. Lin,[29] J. Linacre,[43] M. Lindgren,[18] E. Lipeles,[46] A. Lister,[8] D.O. Litvintsev,[18] C. Liu,[48] T. Liu,[18] N.S. Lockyer,[46] A. Loginov,[61] M. Loreti[x],[44] L. Lovas,[15] D. Lucchesi[x],[44] C. Luci[bb],[52] J. Lueck,[27] P. Lujan,[29] P. Lukens,[18] G. Lungu,[51] L. Lyons,[43] J. Lys,[29] R. Lysak,[15] D. MacQueen,[34] R. Madrak,[18] K. Maeshima,[18] K. Makhoul,[33] T. Maki,[24] P. Maksimovic,[26] S. Malde,[43] S. Malik,[31] G. Manca[e],[30] A. Manousakis-Katsikakis,[3] F. Margaroli,[49] C. Marino,[27] C.P. Marino,[25] A. Martin,[61] V. Martin[l],[22] M. Martínez,[4] R. Martínez-Ballarín,[32] T. Maruyama,[56] P. Mastrandrea,[52] T. Masubuchi,[56] M. Mathis,[26] M.E. Mattson,[59] P. Mazzanti,[6] K.S. McFarland,[50] P. McIntyre,[54] R. McNulty[j],[30] A. Mehta,[30] P. Mehtala,[24] A. Menzione,[47] P. Merkel,[49] C. Mesropian,[51] T. Miao,[18] N. Miladinovic,[7] R. Miller,[36] C. Mills,[23] M. Milnik,[27] A. Mitra,[1] G. Mitselmakher,[19] H. Miyake,[56] N. Moggi,[6] C.S. Moon,[28] R. Moore,[18] M.J. Morello[y],[47] J. Morlok,[27] P. Movilla Fernandez,[18] J. Mülmenstädt,[29] A. Mukherjee,[18] Th. Muller,[27] R. Mumford,[26] P. Murat,[18] M. Mussini[w],[6] J. Nachtman,[18] Y. Nagai,[56] A. Nagano,[56] J. Naganoma,[56] K. Nakamura,[56] I. Nakano,[41] A. Napier,[57]

V. Necula,[17] J. Nett,[60] C. Neu[v],[46] M.S. Neubauer,[25] S. Neubauer,[27] J. Nielsen[g],[29] L. Nodulman,[2] M. Norman,[10] O. Norniella,[25] E. Nurse,[31] L. Oakes,[43] S.H. Oh,[17] Y.D. Oh,[28] I. Oksuzian,[19] T. Okusawa,[42] R. Orava,[24] S. Pagan Griso[x],[44] E. Palencia,[18] V. Papadimitriou,[18] A. Papaikonomou,[27] A.A. Paramonov,[14] B. Parks,[40] S. Pashapour,[34] J. Patrick,[18] G. Pauletta[cc],[55] M. Paulini,[13] C. Paus,[33] T. Peiffer,[27] D.E. Pellett,[8] A. Penzo,[55] T.J. Phillips,[17] G. Piacentino,[47] E. Pianori,[46] L. Pinera,[19] K. Pitts,[25] C. Plager,[9] L. Pondrom,[60] O. Poukhov*,[16] N. Pounder,[43] F. Prakoshyn,[16] A. Pronko,[18] J. Proudfoot,[2] F. Ptohos[i],[18] E. Pueschel,[13] G. Punzi[y],[47] J. Pursley,[60] J. Rademacker[c],[43] A. Rahaman,[48] V. Ramakrishnan,[60] N. Ranjan,[49] I. Redondo,[32] V. Rekovic,[38] P. Renton,[43] M. Renz,[27] M. Rescigno,[52] S. Richter,[27] F. Rimondi[w],[6] L. Ristori,[47] A. Robson,[22] T. Rodrigo,[12] T. Rodriguez,[46] E. Rogers,[25] S. Rolli,[57] R. Roser,[18] M. Rossi,[55] R. Rossin,[11] P. Roy,[34] A. Ruiz,[12] J. Russ,[13] V. Rusu,[18] A. Safonov,[54] W.K. Sakumoto,[50] O. Saltó,[4] L. Santi[cc],[55] S. Sarkar[bb],[52] L. Sartori,[47] K. Sato,[18] A. Savoy-Navarro,[45] P. Schlabach,[18] A. Schmidt,[27] E.E. Schmidt,[18] M.A. Schmidt,[14] M.P. Schmidt*,[61] M. Schmitt,[39] T. Schwarz,[8] L. Scodellaro,[12] A. Scribano[z],[47] F. Scuri,[47] A. Sedov,[49] S. Seidel,[38] Y. Seiya,[42] A. Semenov,[16] L. Sexton-Kennedy,[18] F. Sforza,[47] A. Sfyrla,[25] S.Z. Shalhout,[59] T. Shears,[30] P.F. Shepard,[48] M. Shimojima[p],[56] S. Shiraishi,[14] M. Shochet,[14] Y. Shon,[60] I. Shreyber,[37] A. Sidoti,[47] P. Sinervo,[34] A. Sisakyan,[16] A.J. Slaughter,[18] J. Slaunwhite,[40] K. Sliwa,[57] J.R. Smith,[8] F.D. Snider,[18] R. Snihur,[34] A. Soha,[8] S. Somalwar,[53] V. Sorin,[36] J. Spalding,[18] T. Spreitzer,[34] P. Squillacioti[z],[47] M. Stanitzki,[61] R. St. Denis,[22] B. Stelzer[s],[9] O. Stelzer-Chilton,[17] D. Stentz,[39] J. Strologas,[38] G.L. Strycker,[35] D. Stuart,[11] J.S. Suh,[28] A. Sukhanov,[19] I. Suslov,[16] T. Suzuki,[56] A. Taffard[f],[25] R. Takashima,[41] Y. Takeuchi,[56] R. Tanaka,[41] M. Tecchio,[35] P.K. Teng,[1] K. Terashi,[51] J. Thom[h],[18] A.S. Thompson,[22] G.A. Thompson,[25] E. Thomson,[46] P. Tipton,[61] P. Ttito-Guzmán,[32] S. Tkaczyk,[18] D. Toback,[54] S. Tokar,[15] K. Tollefson,[36] T. Tomura,[56] D. Tonelli,[18] S. Torre,[20] D. Torretta,[18] P. Totaro[cc],[55] S. Tourneur,[45] M. Trovato,[47] S.-Y. Tsai,[1] Y. Tu,[46] N. Turini[z],[47] F. Ukegawa,[56] S. Vallecorsa,[21] N. van Remortel[b],[24] A. Varganov,[35] E. Vataga[aa],[47] F. Vázquez[m],[19] G. Velev,[18] C. Vellidis,[3] V. Veszpremi,[49] M. Vidal,[32] R. Vidal,[18] I. Vila,[12] R. Vilar,[12] T. Vine,[31] M. Vogel,[38] I. Volobouev[t],[29] G. Volpi[y],[47] P. Wagner,[46] R.G. Wagner,[2] R.L. Wagner,[18] W. Wagner,[27] J. Wagner-Kuhr,[27] T. Wakisaka,[42] R. Wallny,[9] S.M. Wang,[1] A. Warburton,[34] D. Waters,[31] M. Weinberger,[54] J. Weinelt,[27] W.C. Wester III,[18] B. Whitehouse,[57] D. Whiteson[f],[46] A.B. Wicklund,[2] E. Wicklund,[18] S. Wilbur,[14] G. Williams,[34] H.H. Williams,[46] P. Wilson,[18] B.L. Winer,[40] P. Wittich[h],[18] S. Wolbers,[18] C. Wolfe,[14] T. Wright,[35] X. Wu,[21] F. Würthwein,[10] S.M. Wynne,[30] S. Xie,[33] A. Yagil,[10] K. Yamamoto,[42] J. Yamaoka,[53] U.K. Yang[o],[14] Y.C. Yang,[28] W.M. Yao,[29] G.P. Yeh,[18] J. Yoh,[18] K. Yorita,[14] T. Yoshida,[42] G.B. Yu,[50] I. Yu,[28] S.S. Yu,[18] J.C. Yun,[18] L. Zanello[bb],[52] A. Zanetti,[55] X. Zhang,[25] Y. Zheng[d],[9] and S. Zucchelli[w],[6]

(CDF Collaboration[†])

[1]*Institute of Physics, Academia Sinica, Taipei, Taiwan 11529, Republic of China*
[2]*Argonne National Laboratory, Argonne, Illinois 60439*
[3]*University of Athens, 157 71 Athens, Greece*
[4]*Institut de Fisica d'Altes Energies, Universitat Autonoma de Barcelona, E-08193, Bellaterra (Barcelona), Spain*
[5]*Baylor University, Waco, Texas 76798*
[6]*Istituto Nazionale di Fisica Nucleare Bologna,* [w]*University of Bologna, I-40127 Bologna, Italy*
[7]*Brandeis University, Waltham, Massachusetts 02254*
[8]*University of California, Davis, Davis, California 95616*
[9]*University of California, Los Angeles, Los Angeles, California 90024*
[10]*University of California, San Diego, La Jolla, California 92093*
[11]*University of California, Santa Barbara, Santa Barbara, California 93106*
[12]*Instituto de Fisica de Cantabria, CSIC-University of Cantabria, 39005 Santander, Spain*
[13]*Carnegie Mellon University, Pittsburgh, PA 15213*
[14]*Enrico Fermi Institute, University of Chicago, Chicago, Illinois 60637*
[15]*Comenius University, 842 48 Bratislava, Slovakia; Institute of Experimental Physics, 040 01 Kosice, Slovakia*
[16]*Joint Institute for Nuclear Research, RU-141980 Dubna, Russia*
[17]*Duke University, Durham, North Carolina 27708*
[18]*Fermi National Accelerator Laboratory, Batavia, Illinois 60510*
[19]*University of Florida, Gainesville, Florida 32611*
[20]*Laboratori Nazionali di Frascati, Istituto Nazionale di Fisica Nucleare, I-00044 Frascati, Italy*
[21]*University of Geneva, CH-1211 Geneva 4, Switzerland*
[22]*Glasgow University, Glasgow G12 8QQ, United Kingdom*
[23]*Harvard University, Cambridge, Massachusetts 02138*
[24]*Division of High Energy Physics, Department of Physics, University of Helsinki and Helsinki Institute of Physics, FIN-00014, Helsinki, Finland*
[25]*University of Illinois, Urbana, Illinois 61801*
[26]*The Johns Hopkins University, Baltimore, Maryland 21218*
[27]*Institut für Experimentelle Kernphysik, Universität Karlsruhe, 76128 Karlsruhe, Germany*

[28]Center for High Energy Physics: Kyungpook National University, Daegu 702-701, Korea; Seoul National University, Seoul 151-742, Korea; Sungkyunkwan University, Suwon 440-746, Korea; Korea Institute of Science and Technology Information, Daejeon, 305-806, Korea; Chonnam National University, Gwangju, 500-757, Korea
[29]Ernest Orlando Lawrence Berkeley National Laboratory, Berkeley, California 94720
[30]University of Liverpool, Liverpool L69 7ZE, United Kingdom
[31]University College London, London WC1E 6BT, United Kingdom
[32]Centro de Investigaciones Energeticas Medioambientales y Tecnologicas, E-28040 Madrid, Spain
[33]Massachusetts Institute of Technology, Cambridge, Massachusetts 02139
[34]Institute of Particle Physics: McGill University, Montréal, Canada H3A 2T8; and University of Toronto, Toronto, Canada M5S 1A7
[35]University of Michigan, Ann Arbor, Michigan 48109
[36]Michigan State University, East Lansing, Michigan 48824
[37]Institution for Theoretical and Experimental Physics, ITEP, Moscow 117259, Russia
[38]University of New Mexico, Albuquerque, New Mexico 87131
[39]Northwestern University, Evanston, Illinois 60208
[40]The Ohio State University, Columbus, Ohio 43210
[41]Okayama University, Okayama 700-8530, Japan
[42]Osaka City University, Osaka 588, Japan
[43]University of Oxford, Oxford OX1 3RH, United Kingdom
[44]Istituto Nazionale di Fisica Nucleare, Sezione di Padova-Trento, [x]University of Padova, I-35131 Padova, Italy
[45]LPNHE, Universite Pierre et Marie Curie/IN2P3-CNRS, UMR7585, Paris, F-75252 France
[46]University of Pennsylvania, Philadelphia, Pennsylvania 19104
[47]Istituto Nazionale di Fisica Nucleare Pisa, [y]University of Pisa, [z]University of Siena and [aa]Scuola Normale Superiore, I-56127 Pisa, Italy
[48]University of Pittsburgh, Pittsburgh, Pennsylvania 15260
[49]Purdue University, West Lafayette, Indiana 47907
[50]University of Rochester, Rochester, New York 14627
[51]The Rockefeller University, New York, New York 10021
[52]Istituto Nazionale di Fisica Nucleare, Sezione di Roma 1, [bb]Sapienza Università di Roma, I-00185 Roma, Italy
[53]Rutgers University, Piscataway, New Jersey 08855
[54]Texas A&M University, College Station, Texas 77843
[55]Istituto Nazionale di Fisica Nucleare Trieste/Udine, [cc]University of Trieste/Udine, Italy
[56]University of Tsukuba, Tsukuba, Ibaraki 305, Japan
[57]Tufts University, Medford, Massachusetts 02155
[58]Waseda University, Tokyo 169, Japan
[59]Wayne State University, Detroit, Michigan 48201
[60]University of Wisconsin, Madison, Wisconsin 53706
[61]Yale University, New Haven, Connecticut 06520

We present a measurement of the mass of the top quark using data corresponding to an integrated luminosity of 1.9 fb^{-1} of $p\bar{p}$ collisions collected at $\sqrt{s}$ = 1.96 TeV with the CDF II detector at Fermilab's Tevatron. This is the first measurement of the top quark mass using top-antitop pair candidate events in the lepton + jets and dilepton decay channels simultaneously. We reconstruct two observables in each channel and use a non-parametric kernel density estimation technique to derive two-dimensional probability density functions from simulated signal and background samples. The observables are the top quark mass and the invariant mass of two jets from the W decay in the lepton + jets channel, and the top quark mass and the scalar sum of transverse energy of the event in the dilepton channel. We perform a simultaneous fit for the top quark mass and the jet energy scale, which is constrained *in situ* by the hadronic W boson mass. Using 332 lepton + jets candidate events and 144 dilepton candidate events, we measure the top quark mass to be $\mathrm{M_{top}} = 171.9 \pm 1.7$ (stat. + JES) $\pm$ 1.1 (other syst.) GeV/c^2 = 171.9 ± 2.0 GeV/c^2.

PACS numbers: pacs 14.65.Ha, 13.85.Qk, 12.15.Ff

*Deceased
†With visitors from [a]University of Massachusetts Amherst, Amherst, Massachusetts 01003, [b]Universiteit Antwerpen, B-2610 Antwerp, Belgium, [c]University of Bristol, Bristol BS8 1TL, United Kingdom, [d]Chinese Academy of Sciences, Beijing 100864, China, [e]Istituto Nazionale di Fisica Nucleare, Sezione di Cagliari,

I. INTRODUCTION

With a mass of approximately 172 GeV/c^2 [1], the top quark (t) is by far the most massive fundamental object observed to date in nature, some 40 times as massive as its isospin partner, the bottom quark (b). This large mass leads to an important role for the top quark in theoretical predictions from the standard model (SM) of particle physics. In particular, electroweak radiative corrections to the W boson mass, due to loops containing top quarks, play an important role in constraining the mass of the Higgs boson, which also contributes to radiative corrections. If the Higgs boson is discovered, a precise measurement of the mass of the top quark will help provide an important test of the SM, and would confirm that the newly observed object is the SM Higgs boson and not some other scalar particle or source of new physics. Independent of the Higgs boson, the large mass of the top quark may make precision measurements throughout the top quark sector necessary to help disentangle models of new physics [2].

The CDF and DØ collaborations jointly announced discovery of the top quark in 1995 [3, 4], but it was not until the availability of large datasets from Run II at the Tevatron that precision measurements of the top quark mass have been possible. The 2007 world average of published top quark mass measurements, $\mathrm{M_{top}} = 172.5 \pm 1.5$ (stat.) $\pm$ 2.3 (syst.) GeV/c^2 [1], compares to the CDF measurement of $\mathrm{M_{top}} = 174 \pm 10$ (stat.) $^{+13}_{-12}$ (syst.) GeV/c^2 upon finding first evidence for the top quark [5]. According to the SM, top quarks at the Tevatron are produced mainly in $t\bar{t}$ pairs resulting from $q\bar{q}$ annihilation (85%) and gluon-gluon fusion (15%). A top quark decays more than 99% of the time to a W boson and a b quark. The topology of $t\bar{t}$ events depends on the subsequent decay of each of the two W bosons. Each W boson can decay hadronically, to a pair of quarks, or leptonically, to a charged lepton and a neutrino. Due to the difficulty of reconstructing τ leptons when they decay, for the purposes of this analysis only electrons (e) and muons (μ) are considered as charged lepton candidates. The semileptonically decaying τ leptons can enter the data set when they are reconstructed as electrons or muons.

Dilepton $t\bar{t}$ events are those in which both W bosons decay leptonically. The dilepton channel has a small branching ratio of $\sim$ 5%, and suffers from underconstrained kinematics resulting from the presence of two neutrinos in each event that escape undetected. The advantages of the dilepton channel are a low background rate and simple combinatorics with only two quarks in the final state. Lepton + jets events are the roughly 30% of $t\bar{t}$ events in which one W boson decays hadronically and the other decays leptonically. The lepton + jets channel has four quarks, one lepton, and one neutrino in the final state, and sufficient amount of information measured in the detector to constrain the kinematics of the $t\bar{t}$ decay.

This analysis describes a measurement of the top quark mass in both the lepton + jets and dilepton decay channels using data collected at the Tevatron with the CDF II detector in a 1.9 fb^{-1} integrated luminosity run. This is the first analysis to combine likelihoods from multiple measurements of the top quark mass in different decay topologies into a single joint likelihood with a robust treatment of the correlations in systematic uncertainties between the two channels channels.

In the analysis described in this article we follow the template strategy [6]. In the lepton + jets channel we determine the kinematics of the decay by fitting for a reconstructed top quark mass m_t^{reco} [7]. In the dilepton channel, due to the two undetected neutrinos, there is not enough information to constrain the four-vectors of the top quarks. Instead we use the neutrino weighting algorithm (NWA), in which we scan over top quark masses, performing an integration over the polar angles of the neutrinos at each mass to calculate a weight based on the agreement with the measured momentum imbalance in the event [8, 9]. We select as the observable in an event the top quark mass (m_t^{NWA}) that yields the highest weight.

Uncertainties in jet modeling and in the calorimeter response result in a systematic uncertainty in the jet energy calibration, which in turn induces the largest systematic uncertainty in top quark mass measurements. In the lepton + jets channel, the hadronically decaying W boson provides an *in situ* calibration sample for these effects. The invariant mass of the hadronically decaying W bosons (m_{jj}) is used in the likelihood fit to measure and constrain the jet energy calibration. This procedure reduces the combined statistical and systematic uncertainty on the top quark mass measurement. The calibration obtained from the lepton + jets channel is for the first time applied to both channels by performing a simultaneous likelihood fit. To improve the precision of the dilepton measurement we also use a second observable, H_T [10], which is a scalar sum of the $t\bar{t}$ decay product

09042 Monserrato (Cagliari), Italy, [f]University of California Irvine, Irvine, CA 92697, [g]University of California Santa Cruz, Santa Cruz, CA 95064, [h]Cornell University, Ithaca, NY 14853, [i]University of Cyprus, Nicosia CY-1678, Cyprus, [j]University College Dublin, Dublin 4, Ireland, [k]Royal Society of Edinburgh/Scottish Executive Support Research Fellow, [l]University of Edinburgh, Edinburgh EH9 3JZ, United Kingdom, [m]Universidad Iberoamericana, Mexico D.F., Mexico, [n]Queen Mary, University of London, London, E1 4NS, England, [o]University of Manchester, Manchester M13 9PL, England, [p]Nagasaki Institute of Applied Science, Nagasaki, Japan, [q]University of Notre Dame, Notre Dame, IN 46556, [r]University de Oviedo, E-33007 Oviedo, Spain, [s]Simon Fraser University, Vancouver, British Columbia, Canada V6B 5K3, [t]Texas Tech University, Lubbock, TX 79409, [u]IFIC(CSIC-Universitat de Valencia), 46071 Valencia, Spain, [v]University of Virginia, Charlottesville, VA 22904, [dd]On leave from J. Stefan Institute, Ljubljana, Slovenia,

transverse energies.

Conventional procedures for combining values of the top quark mass measured in different decay channels require correlations in the systematic effects between the channels as inputs [11]. In addition, combinations of existing top quark mass results must also assume the functional form of the likelihood in each measurement. Typically, the likelihoods are assumed to be Gaussian, and likelihood asymmetries or other departures from Gaussian behavior are not taken into account. The analysis we describe here is free of these two assumptions. The systematic uncertainties are evaluated using the simultaneous fit, and the measurement uses one likelihood function that depends on the data in the two channels. As a cross-check, we also measure the top quark mass separately in the lepton + jets and dilepton channels, including a full evaluation of systematic uncertainties.

II. EXPERIMENTAL SETUP

A. CDF II Detector

The Collider Detector at Fermilab is located at one of two collision points along the ring of the Tevatron accelerator, which collides bunches of protons and anti-protons at a center-of-mass energy of 1.96 TeV. The detector has an approximate cylindrical geometry around the Tevatron beamline and is described in a Cartesian or in a polar coordinate system. In Cartesian coordinates the z axis is located along the beam axis with positive z in the direction of the proton beam, the x axis pointing outward in the plane of the Tevatron ring, and the y coordinate pointing up. It is often more convenient to use polar coordinates: the azimuthal angle ϕ is the angle from the x axis in the plane transverse to the beamline; the polar angle θ is the angle from the proton beam direction. The pseudorapidity $\eta \equiv -\ln(\tan\frac{\theta}{2})$ is a quantity numerically close to rapidity for highly relativistic particles; differences in pseudorapidity are therefore nearly invariant with respect to boosts along the z axis. Collisions occur along the beamline and are distributed about the center of detector with a spread of about 30 cm. We distinguish between η defined with respect to $z = 0$ (η_{det}) and η defined with respect to the event collision point. It is common to reference a region in a cone of ΔR around an object. This refers to the nearby region in η-ϕ space: $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$. The transverse momentum p_T refers to the momentum in the plane transverse to the beamline. The transverse energy of an object is defined as $E_T \equiv E \sin\theta$.

A detailed description of the CDF II detector is provided in Ref. [12]. In this section we briefly introduce the detector subsystems relevant to this analysis, starting with the detectors closest to the interaction region.

Charged particles are observed in the silicon tracking detectors. The innermost silicon detector, layer 00 [13], is a single-sided silicon strip detector mounted directly onto the beryllium beampipe, providing axial tracking information at a radius of 1.6 cm. The silicon vertex detector (SVX II) consists of five double-sided silicon strip detectors locating at radii from 2.5 cm to 10.6 cm from the beamline and 90 cm in length, providing axial and stereo information. Tracking of charged particles in the central region ($|\eta_{det}| < 1.0$) is provided by a 310 cm long cylindrical open cell drift chamber, the central outer tracker (COT), located at radii between 43 and 132 cm. The tracking detectors are immersed in a 1.4 T solenoidal magnetic field, allowing for charge determination and momentum measurements of charged particles [14].

The calorimeter system measures the energy and position of particles passing through and interacting with dense material. CDF uses lead-scintillator and steel-scintillator sampling devices for the electromagnetic and hadronic calorimetry, respectively. The calorimeter system is comprised of the central electromagnetic (CEM) [15], central (CHA) and wall (WHA) hadronic calorimeters [16] covering $|\eta_{det}| < 1.0$, and the plug electromagnetic (PEM) [17] and hadronic (PHA) calorimeters covering $1.1 < |\eta_{det}| < 3.6$. Shower maximum detectors are embedded in the central (CES) and plug (PES) electromagnetic calorimeters at approximately six radiation lengths from the collision point to provide the transverse shape of the shower [15, 18].

Muon detectors are located beyond the calorimeters. Directly outside the CHA is the central muon detector (CMU), which covers $|\eta_{det}| < 0.6$. Located behind further 60 cm of steel shielding is the central muon upgrade (CMP) detector. The central muon extension (CMX) covers the region $0.6 < |\eta_{det}| < 1.0$. Muons pass through the calorimeter and the shielding and leave behind a series of hits (stubs) in the muon detectors, which consist of four layers of single-wire drift cells.

CDF employs a three-level trigger system to select potentially interesting events, reducing the interaction rate from the 1.7 MHz average bunch crossing rate to a more manageable 75–150 Hz. The $t\bar{t}$ candidate events used in this analysis are collected by triggers that identify at least one high-p_T lepton candidate. For central electron events (CEM events), the first-level trigger requires a cluster in the electromagnetic calorimeter with $E_T > 8$ GeV, a matching track in the COT with $p_T > 8$ GeV/c, and a ratio of energy deposited in the hadronic to electromagnetic calorimeters less than 1:8. At the second trigger level, the cluster energy requirement is tightened to $E_T > 16$ GeV, and the third trigger level makes basic electron identification cuts and further tightens the energy requirement to $E_T > 18$ GeV. For muons in the central region, the first-level trigger requires stubs in both the CMU and CMP detectors (CMUP events) or the CMX detector (CMX events) and a matching track in the drift chamber with $p_T > 4$ GeV/c for CMUP events and with $p_T > 8$ GeV/c for CMX events. At the second trigger level, the p_T requirement is increased to 15 GeV/c, and at the third trigger level to 18 GeV/c.

III. EVENT SELECTION AND BACKGROUND ESTIMATION

The energetic, charged leptons and missing transverse energy from at least one leptonic W boson decay help distinguish lepton + jets and dilepton $t\bar{t}$ events from the QCD multijet background. Further rejection of background events containing real leptons is achieved by requiring high-p_T jets, and in some cases identifying one or more of those jets as arising from a b quark. We briefly describe here the reconstruction of physics objects in the detector, as well as event selection, background estimation, and Monte Carlo (MC) simulation.

A. Selection of leptons

Lepton identification is similar in the lepton + jets and dilepton channels. Events in both channels require a clean lepton in the central region of the detector. The second lepton in dilepton events can be from less pure categories such as forward or non-isolated leptons. The major distinctions between the two channels are noted in the following descriptions.

A small cluster of towers in the CEM containing $E_T >$ 20 GeV with a COT track that extrapolates to the face of a tower in the cluster is identified as an electron candidate. Electrons deposit most of their energy in the electromagnetic calorimeter; therefore we require the ratio of hadronic energy to electromagnetic energy to be less than $0.055 + 0.00045E$, where E is the total energy of the electron. To reject backgrounds with energetic π^0s plus a track, we require the ratio of energy in the cluster to the track momentum to be not more than 2.0 for electrons with $E_T < 100$ GeV. The lateral shower development measured in the calorimeter and the CES is required to match the electron shower shape as measured on a test beam. We also require that the COT track extrapolated to the depth of the shower maximum detector matches a CES cluster in the r-z and r-ϕ planes [19].

Full COT tracking information is available only for $|\eta_{det}| < 1.0$. To reconstruct forward (PHX) electron candidates with $|\eta_{det}| > 1.0$, we use clusters in the PES, the energy measurement in the PEM, and knowledge of the $p\bar{p}$ interaction vertex. Silicon detector hits are then added to form a track [20]. We require that forward electrons have $E_T > 20$ GeV and a ratio of hadronic to electromagnetic energy less than 0.05. We also require that shower profiles measured in the PEM and PES match the electron shower shape as measured on a test beam [19]. While PHX electrons are accepted in the dilepton sample, they are used in the lepton + jet sample only to veto dilepton events.

Muon candidates are required to have a track with $p_T > 20$ GeV/c matched to a calorimeter tower with electromagnetic energy less than 2 GeV and hadronic energy less than 6 GeV. The energy of the tower is required to be low since muons are minimum ionizing in the calorimeter material. Both requirements are slightly looser if the track momentum is greater than 100 GeV/c. If the track extrapolates to a fiducial region of the muon chambers, we require that it is matched to a stub inside these detectors. We categorize muons based on the detector chamber that the muon traverses. Muons fiducial to both the CMU and CMP detectors are called CMUP muons. If a track does not extrapolate to any muon chambers but all other quality criteria are satisfied, we accept this object as a "stubless" muon (CMIO). Only CMUP and CMX muons enter the lepton + jets dataset. All muon categories are allowed into the dilepton dataset for one of the two leptons to increase the statistics of the sample.

Electrons and muons produced in W boson decays will in general be well separated from other objects in the event. For electrons we define isolation as the ratio of energy deposited within $\Delta R < 0.4$ around and excluding the electron cluster to the electron cluster energy itself. For muons, isolation is defined as the ratio of the transverse energy within $\Delta R < 0.4$ from the tower crossed by the muon to the muon p_T. The tower traversed by the muon is excluded from transverse energy sum. We categorize leptons as isolated if their isolation variable is less than 0.1. The lepton + jets channel uses only isolated leptons. In the dilepton channel, to increase the sample statistics we allow one non-isolated lepton of any type except for PHX electrons and CMIO muons.

B. Jet corrections and systematic uncertainties

The property of quark confinement [21, 22] ensures that bare quarks are not directly observable after QCD processes take effect. Quarks and gluons ("partons") instead manifest themselves in the detector as jets of particles flowing in the direction of the original gluon or quark (parton). The sum of energies of particles within a cone around the direction of the fragmenting parton is strongly correlated with its energy. Jets for this analysis are reconstructed with the cone-based clustering algorithm JETCLU [23], using a cone in the azimuth-pseudorapidity space $\Delta R = \sqrt{\eta^2 + \phi^2} = 0.4$. The four-vector of a jet is constructed based on energies and locations of calorimeter towers belonging to it. A detailed explanation of CDF calibration of jet energy, the corrections applied to it and the associated systematic uncertainties can be found in Ref. [24]. We briefly summarize these corrections below.

From the raw jet energy several stages of corrections are applied that attempt to remove effects masking the initial parton energy. In the first stage, the relative corrections normalize the detector response as a function of η_{det} so that jets at all η_{det} have the same E_T response as jets in the well understood central region of the detector, $0.2 < |\eta_{det}| < 0.6$. The response varies across η_{det} due to uninstrumented regions of the detector and different amounts of material in front of the calorimeters. The

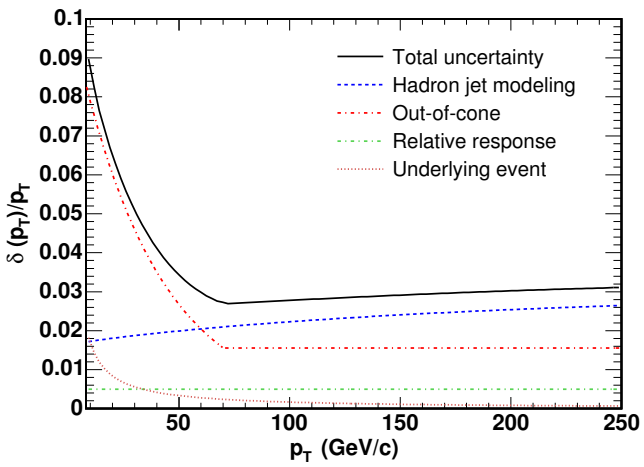


FIG. 1: Fractional uncertainty on corrected jet p_T as a function of corrected jet p_T.

next correction accounts for the average energy in a jet cone due to additional $p\bar{p}$ inelastic interactions occurring during the same bunch crossing. After accounting for energy due to extra $p\bar{p}$ interactions, corrections are applied for calorimeter non-linear response to hadron jets and energy loss in uninstrumented regions. At this stage, jet energies should be independent of the CDF detector, and correspond to the energies of "particle jets", which are defined as all long lived particles from the primary $p\bar{p}$ collision within the jet cone. Energy from spectator partons in the hard collision process that breaks up the proton and anti-proton to form the $t\bar{t}$ system is accounted for by the underlying energy correction. The final correction accounts for out-of-cone effects, in which some of the original parton energy lies outside the jet cone.

Modeling of each of the effects described above is a potential source of uncertainty on the measurement of jet energies. The combined fractional uncertainty on the jet energy calibration (σ_c) is shown in Fig. 1. The overall jet energy calibration is referred to as the jet energy scale (JES). We measure the difference $\Delta_{\rm JES}$ between the JES effects in simulation and data in units of σ_c. The *a priori* $\Delta_{\rm JES}$ estimate at CDF is thus by definition $0 \pm 1\ \sigma_c$. For the lepton + jets channel, jets with $|\eta_{det}| < 2.0$ and $E_T > 20$ GeV after applying the relative, multiple interactions, and hadron jet response linearity corrections are referred to as "tight jets." Jets not passing the tight cuts but having $E_T > 12$ GeV and $|\eta_{det}| < 2.4$ are referred to as "loose jets." The dilepton channel uses jets with $|\eta_{det}| < 2.5$ and $E_T > 15$ GeV after corrections to the particle jet level.

C. Neutrinos

Neutrinos from leptonic W decays escape the detector undetected, leading to an imbalance of energy in the transverse plane of the detector. This missing transverse energy ($\not\!\!E_T$) is defined with respect to the event's interaction vertex:

$$\not\!\!E_T = \left| \sum_i -\vec{n^i} E_T^i \right| \qquad (1)$$

where the sum on i runs over all calorimeter towers, $\vec{n^i}$ is a unit vector in the transverse plane pointing from the beamline to the face of the ith calorimeter tower, and E_T^i is the E_T in the ith tower, calculated using the reconstructed event vertex. In events with muons, the measured p_T of the muon track is used in the $\not\!\!E_T$ calculation, and not the energy deposited by the muon in the calorimeters. Since jets are measured with better energy resolution than the raw calorimeter towers, calorimeter tower energies are replaced with jet energies for the towers clustered into jets.

D. Identification of b jets

The B hadrons produced in the fragmentation of b quarks have an average lifetime on the order of 1.5 ps. Given typical boosts of b jets in $t\bar{t}$ events at the Tevatron, this implies that B hadrons on average travel several millimeters in the detector before decaying, leading to the identifiable signature of a displaced vertex. This analysis uses the SECVTX algorithm [25] to find jets containing secondary vertices consistent with the decay of a B hadron; when such a displaced vertex is found, the jet is identified ("tagged") as a b-jet. In the lepton + jets (dilepton) channel, only the four (two) jets with largest E_T are checked for b tags. Only tight jets are allowed to have b tags in the lepton + jets channel. The tagging efficiency of b quark jet depends on the E_T and η_{det} of the jet, but is roughly 40% within the fiducial acceptance of the silicon detector, with a roughly 1% per-jet probability to incorrectly tag a light flavor jet. Jets arising from charm quarks have finite lifetimes as well, and have a per-jet tagging efficiency of roughly 8%.

E. Lepton + jets selection and background estimation

The lepton + jets decay channel has four quarks in the final state, so we require at least four jets in the detector. Only events with at least one b tag are considered; events with no b tags have a substantial background from production of W bosons with four or more jets, and are also particularly sensitive to systematic uncertainties. We divide lepton + jets events into two exclusive subsamples based on the number of b tags. For events with a single b tag (1-tag), exactly four tight jets are required. Events with additional tight jets and events with only three tight jets, but one or more loose jets were predicted by MC simulation to contribute very little in terms of measurement

TABLE I: Lepton + jets event selection summary. Jets are corrected to the particle jet level.

	1-tag	2-tag
b-tags (Leading 4 jets)	$= 1$	> 1
Lepton p_T (GeV/c), E_T (GeV)	> 20	
$\not\!\!E_T$ (GeV)	> 20	
Leading 3 jets E_T (GeV)	> 20	
4th jet E_T (GeV)	>20	>12
Extra jets E_T (GeV)	<20	Any

precision. Events with two or more tags (2-tag) have extremely small background contamination, and also have fewer ways to assign jets to the quarks from the $t\bar{t}$ decay, so we allow events with three tight jets and one or more loose jets, as well as events with four or more tight jets. For each lepton + jets subsample, we also require a single, high-p_T, isolated CEM, CMUP or CMX lepton and $\not\!\!E_T$ higher than 20 GeV, indicating the presence of an escaping neutrino. Events with multiple leptons are vetoed. The lepton + jets event selection is summarized in Table I.

Background estimates for the lepton + jets events are derived from a hybrid of data- and MC-based measurements, similarly to previous dedicated analysis [26]. Data are used whenever possible, and MC information is used to fill in any remaining gaps in knowledge. In particular, the overall rate of events with real W bosons and additional jets (W + jets), which dominate the background sample, is determined using the data. The fractions of these events with one or two charm quarks (Wc and $Wc\bar{c}$ events) and two bottom quarks ($Wb\bar{b}$ events) are determined from MC samples. Overall normalizations of W + jets events come from the data after subtracting off an estimate for the fraction of W events coming from QCD processes (non-W events), and separating out a MC based estimate for other processes with real W's ($WW/WZ/ZZ$, $t\bar{t}$ and single-top production), where the lepton trigger efficiencies and MC-data differences have been taken into account. The number of background events also depends on the rate to mistag light-quark jets; this fake rate is determined using data samples triggered by presence of jets. Fake tags come mostly from mismeasured tracks, through interactions with material in the detector and real decays of long-lived light-flavor particles such as K_s and Λ also contribute.

Energetic charged leptons from W boson decay can be faked by QCD events via conversions (electrons) or misidentified pions and kaons (muons), as well as from semi-leptonic heavy-flavor decays. In such events, the $\not\!\!E_T$ requirement can also be passed when jets are mismeasured or fall into uninstrumented regions of the detector. The events which do not include a W boson, typically do fail the $\not\!\!E_T$ requirement, however, so the entire $\not\!\!E_T$ distribution, including the low-$\not\!\!E_T$ region, is used to fit for the number of events with fake W bosons and high $\not\!\!E_T$.

TABLE II: Expected event yield for the lepton + jets selection after all cuts. Uncertainties quoted are due to the uncertainty on the integrated luminosity, statistics of MC samples and uncertainty in QCD modeling.

	1-tag	2-tag
Wbb	9.1±3.7	2.1±0.9
$Wc\bar{c}$, Wc	8.3±3.4	0.5±0.3
W (mistags)	10.4±2.3	0.2±0.1
Single top	2.0±0.1	0.7±0.1
Diboson	2.4±0.2	0.21±0.02
QCD	10.4±8.7	0.3±1.6
Total Background	42.7 ± 12.5	4.2 ± 1.9
$t\bar{t}$ (6.7 pb)	156.7±21.1	76.6±12.0
Observed	233	99

The QCD background is modeled in the data by events passing all cuts where instead of a lepton an electron-like object is required. These pass all the kinematic cuts that are imposed on the electrons, but fail shower development or track quality cuts. In an alternative, high statistics model for the QCD background data events are used where the isolation cut on the charged lepton is removed and an anti-isolation cut is imposed, requiring the isolation variable defined earlier to be greater than 0.2. The non-isolated model is not used to obtain the QCD background normalization, however it is used at later stages in the analysis.

Table II shows the expected number of background events in the 1-tag and 2-tag samples after all cuts, the expected number of signal events based on the theoretical cross-section (6.7 pb) [27] at $\mathrm{M_{top}} = 175$ GeV/c^2, and the observed number of events. Uncertainties on the event yield expectations are due to integrated luminosity, statistics of the MC samples as well as uncertain $\not\!\!E_T$ spectrum in the QCD model.

F. Dilepton selection and background estimation

Signal events in the dilepton channel contain two b quarks, two neutrinos and two oppositely charged leptons. We require that at least one of the leptons is isolated and categorized as a CEM, CMUP or CMX lepton. The second lepton can belong to any category. If the second lepton is a PHX electron or a CMIO muon, it must be isolated. All other types of second leptons can be non-isolated to increase statistics of the sample. We require a minimum of two jets in each dilepton event.

To account for the two neutrinos, we require $\not\!\!E_T >$ 25 GeV. We increase the requirement to $\not\!\!E_T > 50$ GeV if the angle in the r-ϕ plane between the $\not\!\!E_T$ vector and any jet or lepton is less than 20° in order to reduce backgrounds such as Drell-Yan production of τ pairs and QCD events where jets fall into uninstrumented regions of the calorimeter.

We require that the H_T - the scalar sum of transverse

energies of jets, leptons and $\not\!\!E_T$ be greater than 200 GeV. This has a small effect on signal acceptance, as the sum of energies of the $t\bar{t}$ decay products must be equal to or greater than twice the top quark mass.

In events where the charged leptons are the same flavor, backgrounds with Z bosons are removed by requiring the invariant mass of the dilepton pair to be smaller than 76 GeV/c^2 or larger than 106 GeV/c^2. We impose this requirement only if the $\not\!\!E_T$ significance ($s_{\not\!\!E_T}$) is smaller than 4.0GeV$^{1/2}$:

$$s_{\not\!\!E_T} \equiv \frac{\not\!\!E_T}{\sqrt{E_T^{sum}}} < 4.0 \text{ GeV}^{1/2} \quad (2)$$

where E_T^{sum} is defined as the scalar sum of the transverse energies of all calorimeter towers, with muons and jets corrected as in the $\not\!\!E_T$ calculation.

Backgrounds for the dilepton channel include Drell-Yan processes, diboson production and QCD multijet production. We estimate the contribution from Drell-Yan production of τ pairs and diboson production using MC samples normalized to the theoretical cross sections. We apply trigger efficiencies as well as corrections accounting for differences between the data and MC simulations in lepton identification efficiencies and jet multiplicity distributions.

We employ a combined data-MC sample technique to estimate the contamination from Drell-Yan ee and $\mu\mu$. This background contains two components in the signal region: events outside the Z boson window (76 − 106 GeV/c^2), and events inside the window and passing the $\not\!\!E_T$ significance cut. We first count the number of events in data with a dilepton invariant mass within the Z window. After subtracting expected contributions from other sources, we multiply this estimate by the ratio of the number of events outside the window to the number of events inside the window, as measured in the MC samples. This gives the number of Drell-Yan events outside the Z window. We estimate the contribution from events inside the Z window with high $s_{\not\!\!E_T}$ by multiplying the number of events in data inside the window by the ratio of events passing and failing the $\not\!\!E_T$ significance cut, again obtained in MC samples.

The data are used to estimate the contribution of events where a real lepton is produced in association with multiple jets and one of the jets is misidentified as a second lepton. Data samples triggered on presence of jets are used to obtain the probability for a jet to fake a charged lepton. These probabilities depend on lepton category and the jet E_T. We apply these probabilities to the single-lepton data to obtain an estimate for the fake background contribution.

We divide the dilepton sample into non-tagged and tagged subsamples, which have very different purity. Since the fake background is modeled directly from data, the probabilities for fake leptons to be reconstructed are summed separately in events with and without a b tag. All other backgrounds are modeled using MC samples. For all MC events, we calculate the probability for each jet to be tagged, accounting for the probabilities for light flavor jets to be mistagged. Given the tag probabilities for the two leading jets, we calculate the probability for each event to enter the non-tagged and tagged subsamples. Signal and background estimates for the dilepton channel are summarized in Table III. Uncertainties are due to integrated luminosity, MC sample statistics and fake rates.

TABLE III: Expected event yield for the dilepton selection after all cuts. Uncertainties quoted capture the uncertainty on integrated luminosity, statistics of the MC samples and uncertainties on the fake rates.

	non-tagged	tagged
Diboson	9.1 ±2.2	0.3 ±0.1
Drell-Yan	16.0 ±2.5	0.9 ±0.1
Fakes	19.3 ±5.6	2.7 ±1.0
Total Background	44.3 ±7.0	3.9 ±1.0
$t\bar{t}$ (6.7 pb)	40.1 ±3.1	55.8 ±4.2
Observed	83	61

G. MC simulation

The signal ($t\bar{t}$) MC simulation is modeled by PYTHIA version 6.216 [28], with HERWIG version 6.510 [29] used as a cross-check. Most background kinematics are estimated from MC samples. The diboson backgrounds are modeled with PYTHIA version 6.216 and the W + jets and Drell-Yan+jets backgrounds are modeled by ALPGEN version 2.10′ [30], with jet fragmentation modeled by PYTHIA version 6.325 [28]. A matching scheme [31] is used to ensure that there is no double-counting of phase space in background events, as it is otherwise possible for events with hard hadronic shower evolution to give states already described by events at the matrix element level. The W + jets and Drell-Yan+jets background MC is divided into n-parton samples, where n refers to the total number of partons (quarks or gluons), including heavy flavor. The samples are combined according to the cross sections reported by ALPGEN, accounting for possibly different efficiencies for the samples to pass event selection. Similarly, events with heavy flavor after fragmentation are checked to ensure no double-counting of phase space across samples with different flavor types at the matrix element level. Double-counting can occur if samples generated at the matrix element level with light-flavor partons produce charm or bottom quark pairs in the parton shower. We remove such events unless both heavy flavor partons are within the same jet. In addition the heavy flavor quark pairs generated at the matrix element level can enter the same jet effectively reducing heavy jet multiplicity, therefore we also remove events of this type. Depending on the multiplicity of light and heavy flavor final state partons we remove a fraction of events between a few percent and approximately 20%.

Electroweak production of single top quarks in both the s- and t-channels contributes very few events to our sample. These events are treated as background, and are modeled using a fixed mass of $M_{top} = 175.0$ GeV/c^2. Single top quark events are generated by MADEVENT [32]; fragmentation is modeled with PYTHIA version 6.409 [33].

To model multiple proton-antiproton interactions occurring in a single bunch crossing, we add interactions where no partons with high transverse momenta are produced to the events simulated for each process. Those minimum bias collisions are simulated with PYTHIA version 6.216 . The number of minimum bias interactions added to a given event is equal to the expected number of $p\bar{p}$ interactions, which depends on the instantaneous luminosity profile of the data run of the event. The instantaneous luminosity profile is matched between MC samples and data only for the first 1.2 fb^{-1} of integrated luminosity. This incorrect model is a source of bias of 0.4 GeV/c^2 for the dilepton-only fit. No bias is present in the fitted top quark mass in the combined and lepton + jets only fits, however a bias of 0.04 σ_c is present in the fitted Δ_{JES} in both measurements. The bias in the dilepton measurement is higher due to the particular choice of observables used to make the measurement and also due to the fact that the *in situ* calibration absorbs the bias on mass and converts it into bias on Δ_{JES}. These biases were found in studies of MC samples with increased number of $p\bar{p}$ interactions and are corrected for in quoted results.

IV. EVENT RECONSTRUCTION

After selection and reconstruction of event parameters of physics interest, the data are processed to form estimators for the top quark mass. Simply forming invariant masses is not possible, as there are two top quarks per event. Events in the lepton+jets channel have many possible assignments of jets to the quarks, each of which give different reconstructed top quark masses. Each dilepton event has two undetected neutrinos, resulting in underconstrained kinematics. Both channels must account for the possibility that, due to radiation effects, jets in the detector may not correspond to quarks from the hard scattering. Both topologies also contain non-negligible backgrounds. We approach these problems by constructing quantities strongly correlated to the top quark mass and comparing the data to MC predictions that include all of the above effects. For the lepton+jets channel, we additionally account for the unique and dense environment of $t\bar{t}$ events by applying jet corrections specific to the $t\bar{t}$ events.

A set of generated distributions for a particular top quark mass is referred to as a template. A template is then a probability density function for a set of observables. Our measurement of the top quark mass is then a determination of the most likely parent template for the data. Further complicating the analysis, however, is the strong correlation between the Δ_{JES} in the detector and quantities sensitive to the top quark mass, including the top quark mass estimators. As explained in Sec. III B, scaling measured jet energies back to original parton energies is a difficult task, and any uncertainty on the JES directly translates to a systematic uncertainty on M_{top}. To reduce this effect, we introduce a second template in the lepton+jets channel that uses the hadronic decay of the W boson to make an *in situ* measurement of Δ_{JES} that can be applied to all jets in the event sample, including those from b quarks and those in the dilepton channel. The narrow width of the W makes its dijet mass (m_{jj}) a good estimator for Δ_{JES}.

A. Top-quark specific corrections

The jet corrections described in Sec. III B are generic algorithms derived for application in all high-energy CDF analyses. As such, they miss out on several key features of $t\bar{t}$ events in the lepton+jets channel. The generic jet corrections assume flat p_T spectra for all jets. The bias for the p_T spectra expected from the physics process under consideration may be corrected specifically. Top-antitop events have two different, non-flat p_T spectra for the W decay jets and the b jets. The generic jet corrections also do not account for differences between jets coming from b quarks and jets coming from light-flavor quarks. To account for all these effects, we derive jet corrections specific to the lepton+jets $t\bar{t}$ environment from MC simulations. After event selection, jets are corrected to the particle jet level. We separate the corrected jets by flavor (whether they came from a b quark or a light quark), and then into different η and p_T bins. The top-specific corrections are derived such that the p_T of a jet corresponds to the most probable value of the quark producing the jet.

The top-specific corrections can be over 50% for low-p_T jets in the central region, and slightly negative for high-p_T jets. We apply the corrections to the p_T of the jet. We assume the direction of the jet to be well measured, and do not apply correction to the jet angles (η and ϕ). To obtain the jet energy, the mass of jets assumed to come from b quarks is fixed to 5.0 GeV/c^2 and the mass of jets assumed to come from light quarks is fixed to 0.5 GeV/c^2, though the mass effects are small compared to typical jet energies in $t\bar{t}$ events. The top-specific corrections also provide the resolution on jet energy, once again separately for the two flavors and as a function of p_T and η. The resolution is worse than 20% for low-p_T jets, and better than 10% for high-p_T jets.

B. Lepton+jets reconstruction

The lepton+jets decay channel gives overconstrained kinematics for the $t\bar{t}$ system. Detailed information on the lepton+jets kinematic fitter can be found in Ref. [7].

The minimization package MINUIT [34] is used to minimize a χ^2-like function for the overconstrained kinematic system:

$$\chi^2 = \sum_{i=\ell,4\text{ jets}} \frac{(p_T^{i,\text{fit}} - p_T^{i,\text{meas}})^2}{\sigma_i^2} + \sum_{j=x,y} \frac{(U_j^{\text{fit}} - U_j^{\text{meas}})^2}{\sigma_j^2} + \frac{(M_{jj} - M_W)^2}{\Gamma_W^2} + \frac{(M_{\ell\nu} - M_W)^2}{\Gamma_W^2} + \frac{(M_{bjj} - m_t^{\text{reco}})^2}{\Gamma_t^2} + \frac{(M_{b\ell\nu} - m_t^{\text{reco}})^2}{\Gamma_t^2} \quad (3)$$

The first term constrains the p_T of the lepton and the 4 jets in the event to their measured values, within their uncertainties σ_i. The unclustered energy (U) is the energy in the calorimeter not associated with the primary lepton or one of the four leading jets. The second term constrains the x and y components of the unclustered energy ($U_x \equiv U \sin\theta \cos\phi$, $U_y \equiv U \sin\theta \sin\phi$) in the detector close to their measured values within uncertainties σ_j. The third term in the χ^2 expression constrains the dijet mass of the two jets assigned as W decay daughters to the well measured W mass within the W boson decay width. The fourth term similarly constrains the invariant mass of the leptonic W decay daughters. The last two terms constrain the invariant masses of the three-body top decay daughters to be consistent within the top quark decay width of 1.5 GeV. The value of m_t^{reco} is a free parameter in the fit, and is taken as the reconstructed mass used in the templates.

The neutrino transverse momentum is not a direct parameter in the χ^2 minimization, but is instead related to the unclustered energy, and is calculated at every stage of the minimization process:

$$p_{x,y}(\nu) = -\left(\sum_{\text{jets}} p_{x,y}(\text{jet}) + p_{x,y}(\text{lepton}) + U_{x,y}\right) \quad (4)$$

The longitudinal component of the neutrino momentum is a free parameter that is effectively determined by the constraint on the invariant mass of the leptonic W.

With the assumption that the leading (most energetic) four jets in the detector come from the four final quarks at the hard scatter level, there are 12 possible assignments of jets to quarks. The minimization is performed for each assignment, with m_t^{reco} taken from the assignment that yields the lowest χ^2. Events with the lowest $\chi^2 > 9.0$ are removed from the sample to reject poorly reconstructed events not fitting the $t\bar{t}$ hypothesis. The cut was optimized for expected statistical precision however we find that there is no strong dependence of the expected precision on the value of the cut. Identifying b jets reduces the number of combinations since tagged jets are assigned only to final state b quarks. In rare events with more than two tags among the leading four jets, only two tags of highest E_T are assigned to the b quarks and the additional tags are ignored.

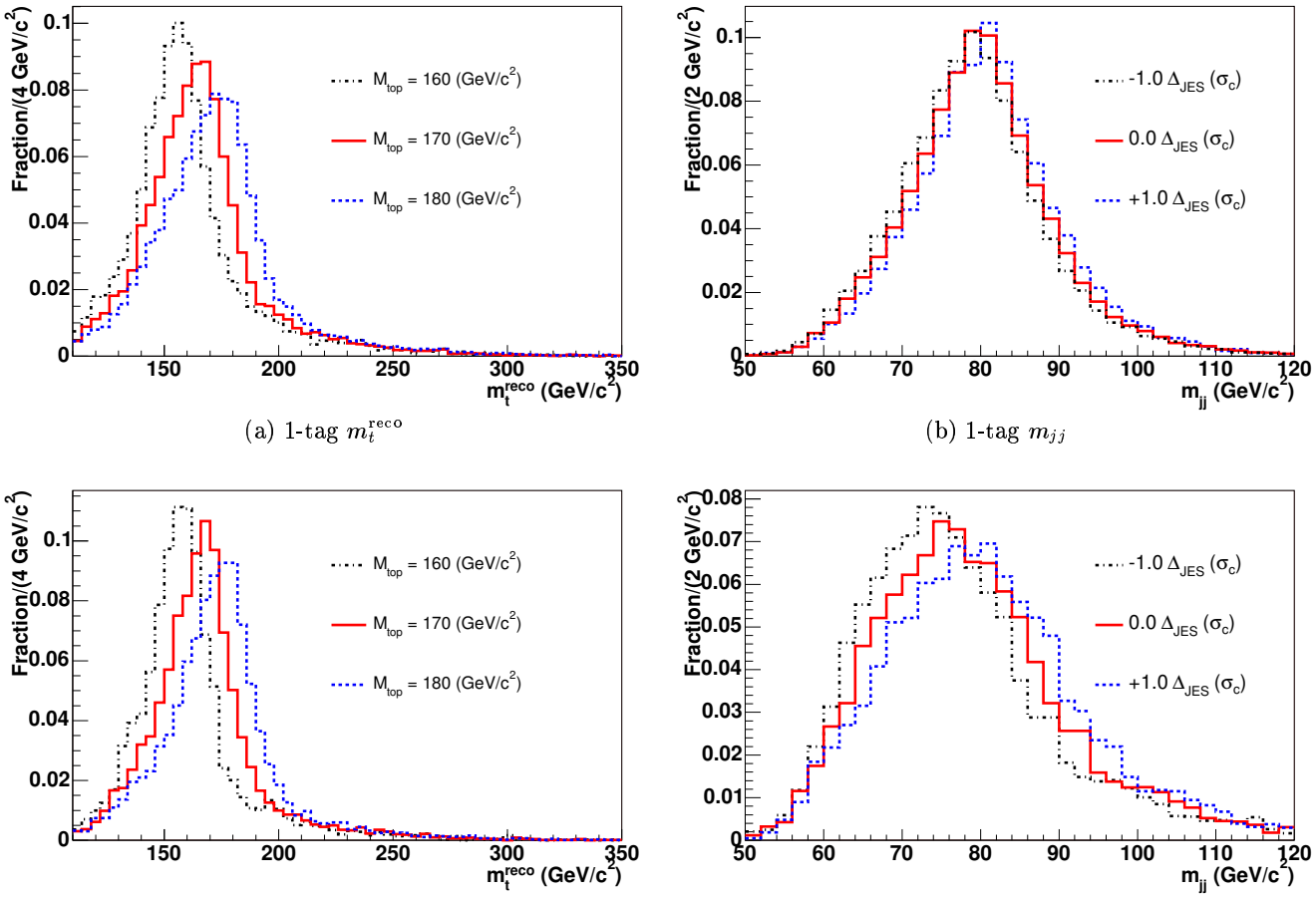
The calculation of the dijet mass m_{jj} is independent of the above minimization procedure to derive m_t^{reco}. Given a pair of jets, a simple invariant mass is calculated from the jet four-vectors; in particular, the W mass constraint of the kinematic fitter is not applied. There are multiple ways to choose two jets among the four or more jets in $t\bar{t}$ events. Tagged jets are assumed to come from final state b quarks. Additionally, the two jets from the hadronic W decay daughters are assumed to be among the leading four jets. For two-tag events, there is only one choice for the jet pair to be associated to the W boson. For 1-tag events, there are three possible dijet masses to be made from the 3 non-tagged leading jets; we pick the single dijet mass closest to the well known W mass. This sculpts the distribution, but is the choice most likely to be correct in selecting the two jets from the W decay daughters, and was found to give the best sensitivity to Δ_{JES}. We correct the jets using the light quark top-specific corrections.

C. Lepton+jets template results

We process MC samples with different values of M_{top} with full detector simulation [35] and event selection. The kinematic fitter is applied to each event, giving the m_t^{reco} templates shown in Fig. 2 (a) and (c). Though the peak of the templates depends strongly on M_{top}, the reconstruction is not perfect, and m_t^{reco} only gives an estimate for M_{top}. The large tails in the templates are a result of incorrect jet-quark assignments. The 2-tag subsample, with fewer jet-parton assignments, has narrower templates, and therefore has more sensitivity to the top quark mass. Templates for m_{jj} masses for three different values of Δ_{JES} are shown in Fig. 2 (b) and (d). The 1-tag m_{jj} templates are narrower than the 2-tag templates due to sculpting of the distributions. The sculpting also yields smaller shifts in the 1-tag template then in the 2-tag template, as the Δ_{JES} varies.

D. Dilepton templates

In the dilepton channel, the measured quantities and assumptions on the masses of particles in the decay cascade do not provide enough constraints to reconstruct the four-vectors of the top quarks. Instead, we form a reconstructed top quark mass (m_t^{NWA}) for each dilepton candidate using the Neutrino Weighting Algorithm [8, 9]. The algorithm assigns a weight to the event as a function of top quark mass. A top quark mass scan is performed in the range 80–380 GeV/c^2, and the value yielding the maximum weight is selected as m_t^{NWA}. The two most energetic jets in an event are considered to have originated from the b quarks, giving two possible jet-quark assignments. For each assumed top quark mass and jet-quark assignment, we integrate numerically over the possible pseudorapidities of the two neutrinos. The



(a) 1-tag $m_t^{\rm reco}$ (b) 1-tag m_{jj}

(c) 2-tag $m_t^{\rm reco}$ (d) 2-tag m_{jj}

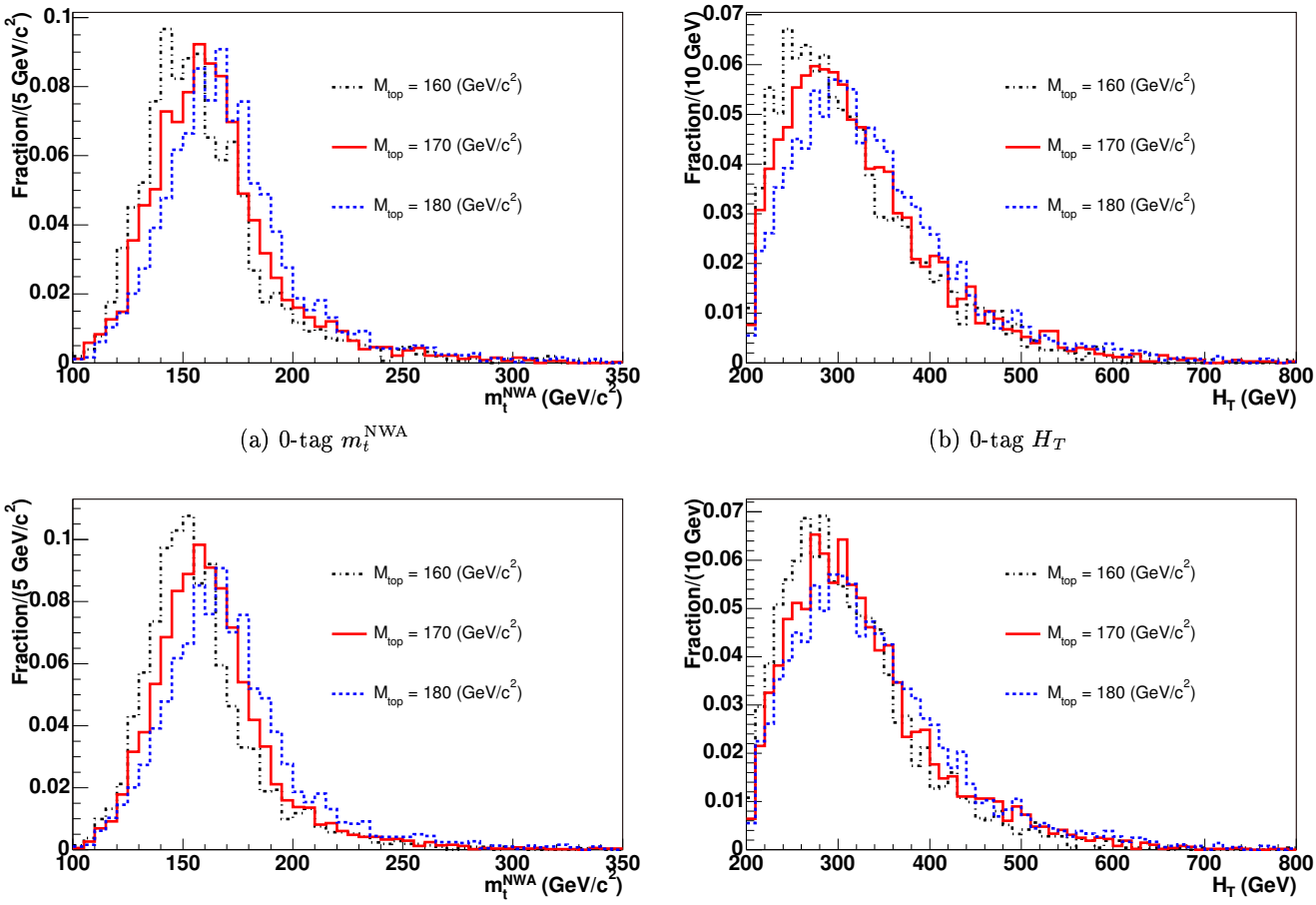
FIG. 2: Template distributions for MC events passing the lepton + jets selection. Shown are the 1-tag $m_t^{\rm reco}$ (a) and m_{jj} (b) distributions, and the 2-tag $m_t^{\rm reco}$ (c) and m_{jj} (d) distributions. The $m_t^{\rm reco}$ distributions are plotted using events with three values of $\mathrm{M_{top}}$ and with the nominal $\Delta_{\rm JES} = 0.0\ \sigma_{\rm c}$. The m_{jj} distributions are plotted using events with three values of $\Delta_{\rm JES}$ and with $\mathrm{M_{top}} = 170$ GeV/c^2.

distribution of the neutrino pseudorapidity is assumed to be Gaussian around zero, with a width of 1.0 obtained from PYTHIA. Given a neutrino pseudorapidity, we can solve for its transverse momentum. Up to two solutions are possible for each of the neutrino and antineutrino transverse momenta. For each of the four solutions, we compare the total momentum in the x and y directions carried by the neutrinos to the measured x and y components of the $\not{E}_T$. We calculate a weight that is the product of two Gaussians, one each for the x and y directions, of the difference between the measured $\not{E}_T$ in that direction and the sum of the momentum components of the two neutrinos. We use a Gaussian width of 19 GeV, which is optimized using MC $t\bar{t}$ events. The four weights are added to form the integrand. Note that we do not account for resolution effects in measurements of the jets and leptons. The two integrals corresponding to the two jet-b quark assignments are added to form the top-mass-dependent weight. In the calculation of the transverse momenta of the neutrinos, jets are corrected using the generic jet corrections; applying the top-specific corrections of Sec. IV A was not found to improve the resolution on the reconstructed top quark mass in this channel. Figure 3 (a) and (c) shows the output of the algorithm from fully simulated MC events with different input masses.

The momenta of the decay products of the $t\bar{t}$ pair are directly correlated to the invariant masses of the top quarks. We therefore use another variable, the H_T (*cf.* Section III F), to improve the precision of the measurement. quark mass. Figure 3 (b) and (d) shows H_T distributions for different top quark masses. The correlation coefficient between $m_t^{\rm NWA}$ and H_T is $\sim$ 40% in signal events and $\sim$ 60% in background events.

V. MASS FITTING

The distributions of the observables defined in Sec. IV are used to determine simultaneously the two parameters $\mathrm{M_{top}}$ and $\Delta_{\rm JES}$. For this, we need to know



(a) 0-tag m_t^{NWA}

(b) 0-tag H_T

(c) Tagged m_t^{NWA}

(d) Tagged H_T

FIG. 3: Template distributions for MC events passing the dilepton selection. Shown are the 0-tag m_t^{NWA} (a) and H_T (b) distributions, and the tagged m_t^{NWA} (c) and H_T (d) distributions. The distributions are plotted using events with three values of $\mathrm{M_{top}}$ and with the nominal $\Delta_{\mathrm{JES}} = 0.0\ \sigma_{\mathrm{c}}$.

$P(x, y; \mathrm{M_{top}}, \Delta_{\mathrm{JES}})$: the probability of observing a particular pair of values of the mass-sensitive parameters (x, y), given some $\mathrm{M_{top}}$ and Δ_{JES}. The observables x and y correspond to m_t^{reco} and m_{jj} for lepton + jets events and to m_t^{NWA} and H_T for dilepton events. When $\mathrm{M_{top}}$ and Δ_{JES} are fixed, the resulting $P(x, y)$ should be a normalized probability density function (PDF) over the two-dimensional space of the observables. The PDFs must be determined separately for signal and background events in each subsample (e.g. 1-tag lepton + jets events). The background probabilities do not depend on $\mathrm{M_{top}}$.

Inaccuracies in these families of PDFs lead to biases in the final measurement that can be difficult to uncover or to characterize. Therefore, in order to achieve a precision measurement of $\mathrm{M_{top}}$, it is essential to make a robust determination of P for each class of events. We accomplish this in two steps: First, at discrete values of $\mathrm{M_{top}}$ and Δ_{JES}, we estimate the two-dimensional PDFs for the observables from large samples of MC events using kernel density estimation (KDE), described in Sec. V A. Then we smooth and interpolate to find PDFs for arbitrary values of $\mathrm{M_{top}}$ and Δ_{JES} using local polynomial smoothing (LPS), described in Sec. V B.

The resulting probabilities are used in a combined likelihood fit (Sec. V C) to measure the top quark mass. We run a rigorous set of checks to validate the analysis machinery and calibrate the final result using events from MC simulation; these checks are described in Sec. V D.

A. Kernel Density Estimation

Previous template-based measurements of the top quark mass [7, 36] used arbitrary functional forms to fit parameterized PDFs from histograms of the observables. It is difficult to extend such parameterizations to two dimensions in observables while properly accounting for correlations between m_t^{reco} and m_{jj}, or m_t^{NWA} and H_T. In the lepton + jets channel, these correlations can lead to a bias of several hundred MeV, if not properly taken into account; in the dilepton channel, the correlations are larger and can make such a measurement with two

observables impossible. Using a functional form with a large number of parameters can also result in fits that are unstable with respect to small changes in the histograms or even in the parameter initialization.

This measurement takes a different approach based on KDE to form PDFs in two observables without any assumption about the functional form. Some useful introductions to KDE can be found in Refs. [37–39]. In KDE, the probability for an event with observable x is given by a linear sum of contributions from all entries in the MC sample. For a one-dimensional distribution, this probability is given by:

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right) \tag{5}$$

where $\hat{f}(x)$ is the probability to observe x given, as an example, a $t\bar{t}$ MC sample with known $\mathrm{M_{top}}$ and Δ_{JES}. The sample has n entries, with values of the observable given by x_i. The kernel function K is a normalized function that adds less probability to a measurement at x as its distance from x_i increases. The smoothing parameter h (sometimes called the bandwidth) is a number that determines the width of the kernel. Larger values of h smooth out the contribution to the kernel density estimate and give more weight at x farther from x_i. Smaller values of h provide less bias to the kernel density estimate, but are more sensitive to statistical fluctuations. We use an Epanechnikov kernel, defined as:

$$K(t) = \begin{cases} \frac{3}{4}(1 - t^2) & \text{for } |t| < 1, \\ K(t) = 0 & \text{otherwise.} \end{cases} \tag{6}$$

so that only events with $|x - x_i| < h$ contribute to $\hat{f}(x)$.

We use an adaptive KDE method in which the value of h is replaced by h_i so that the amount of smoothing applied to the ith event depends on the value of $\hat{f}(x_i)$ [40]. We run a first pass of kernel density estimation with constant h. This pilot kernel density estimate is then used in a second round of KDE to determine the individual h_i, with $h_i \propto \hat{f}(x_i)^{-0.5}$. In the peak of the distributions, where there are more events, we use small values of h_i to capture as much shape information as possible. In the tails of the distribution, where there are fewer events and the kernel density estimates are sensitive to statistical fluctuations, a larger value of h_i is used. The overall scale of h is set by the root mean square (RMS) of the distribution and by the number of entries in the MC sample; larger (smaller) smoothing is used when fewer (more) events are available [37, 38, 41]. If the smoothing parameters get too large in the tails of the distribution, the kernel density estimates can become non-local, and a point at x_i can contribute weight to an estimate at a distant x. Following Ref. [40], we guard against this by not allowing h_i to get too large:

$$h_i = \min(h_i, \sqrt{10} \cdot h_0) \tag{7}$$

where h_0 is the minimum adaptive bandwidth, which occurs in the peak of the distribution.

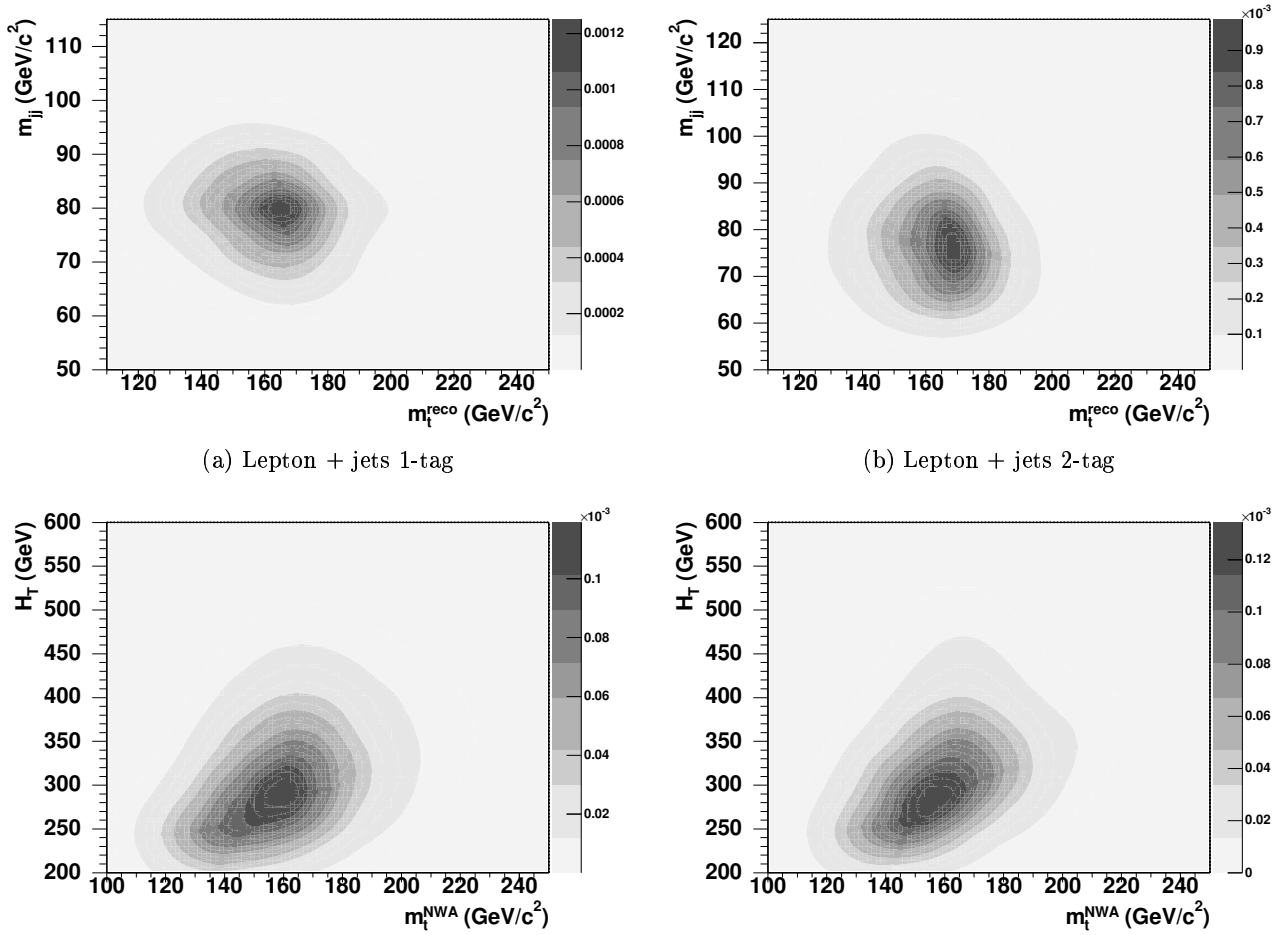
KDE is extended to two dimensions by multiplying two kernels together [42, 43]:

$$\hat{f}(x, y) = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{h_{x,i}h_{y,i}}\left[K\left(\frac{x - x_i}{h_{x,i}}\right) \times K\left(\frac{y - y_i}{h_{y,i}}\right)\right] \tag{8}$$

Note that the smoothing parameters for the two variables do not have to be identical. Typical values of h for kernel density estimates in the signal are 10–12 GeV/c^2 for m_t^{reco}, 3–6 GeV/c^2 for m_{jj}, 15–20 GeV/c^2 for m_t^{NWA} and 45–55 GeV for H_T. For background kernel density estimates, these numbers are slightly larger, as the number of events passing all the cuts is smaller and the templates are wider.

The kernels in Eq. (8) know nothing about the boundaries of the templates. Mathematically, the density functions can take on any real numbers given large enough smoothing, even though kinematic requirements and energy conservation limit possible values of the observables. When the probability density extends beyond such a limit, where the data are not found, the normalization condition of the kernels does not hold. To enforce unit normalization, we explicitly force hard boundaries and reject events in the tails of the distribution both from the MC templates and the data, typically removing 1-2% of signal events and a slightly larger fraction of background events. When kernel density estimates are calculated, we check that each of the individual kernels is normalized within the boundaries. If a kernel is not normalized and leaks probability outside the boundaries, it is renormalized such that it contributes unit weight inside the boundaries.

Using $\hat{f}(x, y)$ from (8) as $P(x, y)$, we can scan values of x and y to visualize the two-dimensional PDF for fixed $\mathrm{M_{top}}$ and Δ_{JES}. Figure 4 shows two-dimensional kernel density estimates for lepton + jets and dilepton signal events, given $\mathrm{M_{top}} = 170$ GeV/c^2 and $\Delta_{\mathrm{JES}} = 0.0$. Figure 5 shows the estimates for background events at $\Delta_{\mathrm{JES}} = 0.0$. The background kernel density estimates are derived separately for the individual contributions to the background model, taking into account the sample sizes and RMS values, and are then combined with the appropriate weights. Note that since jet energy scale shifts would affect all data events in a similar way, data-derived background templates do not depend on Δ_{JES}. The dilepton tagged background contains multiple peaks that come from the fake background. The data used to model the fakes, which comprise half of the total tagged background in the dilepton channel, contain very few events, and also peak at different locations than the other backgrounds.



(a) Lepton + jets 1-tag

(b) Lepton + jets 2-tag

(c) Dilepton 0-tag

(d) Dilepton tagged

FIG. 4: Kernel density estimates at $\mathrm{M_{top}} = 170$ GeV/c^2 and $\Delta_{\mathrm{JES}} = 0.0$ for lepton + jets 1-tag (a) and 2-tag (b) events, and for dilepton untagged (c) and tagged (d) events.

B. Local Polynomial Smoothing

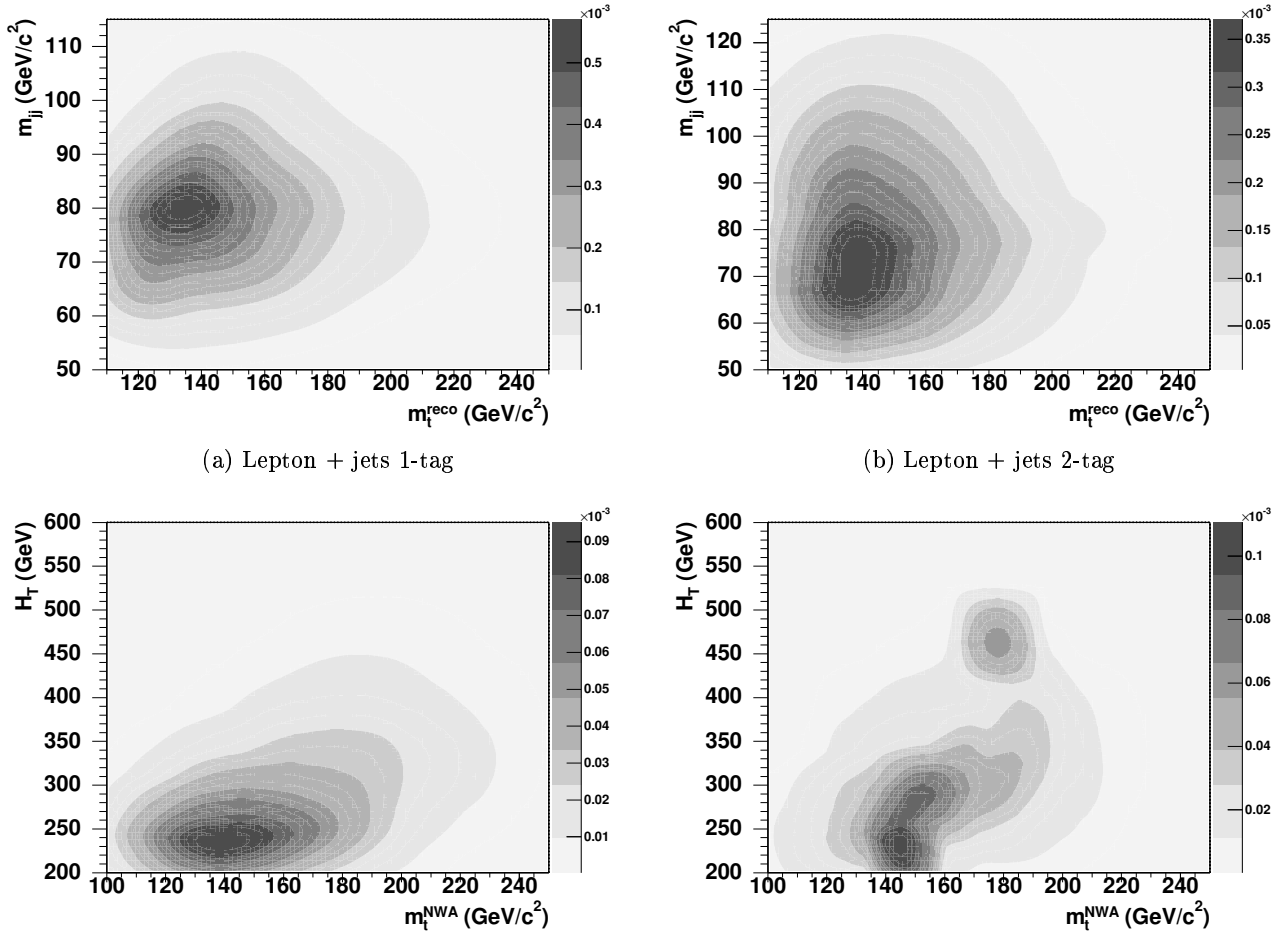
The PDF families $P(x, y; \mathrm{M_{top}}, \Delta_{\mathrm{JES}})$ are defined on the continuous parameters $\mathrm{M_{top}}$ and Δ_{JES}. But the signal MC samples are produced at discrete values of $\mathrm{M_{top}}$ and Δ_{JES}, and the background MC samples are produced at discrete values of Δ_{JES}, so KDE is capable of producing PDFs only at discrete points of the parameter space. In addition, statistical fluctuations in the kernel density estimates are correlated for events with similar observables, so it is useful to smooth out the PDFs before the likelihood fit. To obtain PDFs that are smoothly and continuously varying as a function of $\mathrm{M_{top}}$ and Δ_{JES} without assuming Gaussian likelihoods, we employ a technique known as local polynomial smoothing (LPS) [44], described briefly below.

LPS locally approximates the value of the PDF with a second-order polynomial. The expansion uses the estimates from KDE, but gives more weight to MC samples in a nearby region of Δ_{JES} (and $\mathrm{M_{top}}$, if we are smoothing out the signal probabilities). We look for an estimate $\hat{P}(\boldsymbol{\alpha})$ for the true value of the function $P(\boldsymbol{\alpha})$, where we have omitted the values of the observables (x, y) from the arguments of the function. The quantity $\boldsymbol{\alpha}$ is a two-dimensional vector $(\mathrm{M_{top}}, \Delta_{\mathrm{JES}})$ in the case of the signal probability function or the scalar Δ_{JES} for the background. Kernel density estimation provides estimates Y_k for the values of $P(\boldsymbol{\alpha_k})$ at a number of points $\boldsymbol{\alpha_k}$. We assume that Y_k are unbiased estimators of the true probability values $P(\boldsymbol{\alpha_k})$ with the same variance.

A second order expansion of the function P for points $\boldsymbol{t}$ in the neighborhood of $\boldsymbol{\alpha}$ can be written as:

$$P(\boldsymbol{t}) = \langle \boldsymbol{c}, \boldsymbol{F}(\boldsymbol{t} - \boldsymbol{\alpha}) \rangle \tag{9}$$

where the angle bracket denotes an inner product. The coefficients of the expansion are given by the components of the vector $\boldsymbol{c}$. The quantity $\boldsymbol{F}$ is a vector of basis functions for second order polynomials. For a two-



(a) Lepton + jets 1-tag

(b) Lepton + jets 2-tag

(c) Dilepton 0-tag

(d) Dilepton tagged

FIG. 5: Kernel density estimates at $\Delta_{\rm JES} = 0.0$ for lepton + jets 1-tag (a) and 2-tag (b) background events, and for dilepton untagged (c) and tagged (d) events.

dimensional $\boldsymbol{v}$, $\boldsymbol{F}(\boldsymbol{v})$ is:

$$\boldsymbol{F}(\boldsymbol{v}) = \begin{pmatrix} 1 \\ \boldsymbol{v}_0 \\ \boldsymbol{v}_1 \\ \boldsymbol{v}_0\boldsymbol{v}_1 \\ \frac{1}{2}\boldsymbol{v}_0^2 \\ \frac{1}{2}\boldsymbol{v}_1^2 \end{pmatrix} \tag{10}$$

If $\boldsymbol{v}$ is a scalar, $\boldsymbol{F}(\boldsymbol{v})$ reduces to a 3-component vector.

To evaluate $\hat{P}(\boldsymbol{\alpha})$, we minimize the criterion given by Eq. (11) with respect to $\boldsymbol{c}$. In other words, we find the second order expansion of P around $\boldsymbol{\alpha}$ that best matches the estimates Y_k at points $\boldsymbol{\alpha_k}$:

$$\sum_k w(\boldsymbol{\alpha_k})(Y_k - \langle \boldsymbol{c}, \boldsymbol{F}(\boldsymbol{\alpha_k} - \boldsymbol{\alpha})\rangle)^2 \tag{11}$$

where the weight of each estimate is given by the factor $w(\boldsymbol{\alpha_k}) = W(u_k)$, with:

$$u_k = \sqrt{\sum_{d=1}^{N_d} \left(\frac{\boldsymbol{\alpha}_d - \boldsymbol{\alpha}_{k,d}}{h_d}\right)^2} \tag{12}$$

The sum in Eq. (12) runs over the components of the vectors, and N_d is the dimensionality of the parameter space. We use $W = W(u) = (1 - |u|^3)^3$ for $|u| < 1$ and $W(u) = 0$ otherwise. This gives a smoothly decreasing weight to the estimates Y_k obtained at points far away from the evaluation point $\boldsymbol{\alpha}$. The constants h_d control the amount of smoothing, larger values of h_d give more weight to the estimates farther away from the point $\boldsymbol{\alpha}$.

For this analysis, signal MC samples are generated at 76 mass points with $\mathrm{M_{top}}$ ranging from 120 to 240 GeV/c^2. The spacing between mass points is small (0.5 GeV/c^2) in the region of interest (165–185 GeV/c^2), and gets larger in the tails of the grid. Each signal MC sample and MC-based background is processed using 29 different values of $\Delta_{\rm JES}$ from $-3.0\ \sigma_{\rm c}$ to $+3.0\ \sigma_{\rm c}$.

Near the range of interest of nominal $\Delta_{\rm JES}$, the spacing is 0.2 $\sigma_{\rm c}$. We set the constants h_d of Eq. (12) based on the performance of the analysis in terms of expected precision and biases. We choose $h_{\rm M_{top}} = 10.0$ GeV/c^2 for signal PDF smoothing in lepton + jets events and $h_{\rm M_{top}} = 15.0$ GeV/c^2 in the dilepton events. For both categories, $h_{\Delta_{\rm JES}} = 0.8\ \sigma_{\rm c}$. The background MC samples have smaller statistics and increased jitter, as many events are selected near jet energy thresholds and move in and out of the sample as $\Delta_{\rm JES}$ varies. To compensate for these effects, we choose the larger $h_{\Delta_{\rm JES}} = 3.0\ \sigma_{\rm c}$ for background smoothing.

C. Likelihood fit

We compare the two dimensional distributions of the observables in the data with the signal and background PDFs in an unbinned extended maximum likelihood fit [45]. The most important parameters of the fit are the mass of the top quark ($\rm M_{top}$) and the deviation from the nominal jet energy scale ($\Delta_{\rm JES}$). Each subsample gives two additional parameters: the expected number of signal events (n_s) and the expected number of background events (n_b) in the subsample. The likelihood form is given by:

$$\begin{aligned}\mathcal{L} &= e^{-\frac{\Delta_{\rm JES}^2}{2}} \\ &\times\ \mathcal{L}_{\rm lepton+jets,1\text{-}tag} \\ &\times\ \mathcal{L}_{\rm lepton+jets,2\text{-}tag} \\ &\times\ \mathcal{L}_{\rm dilepton,non\text{-}tagged} \\ &\times\ \mathcal{L}_{\rm dilepton,tagged}\end{aligned} \tag{13}$$

where the first term in the product constrains the measurement of the JES to its nominal value of 0 within the uncertainty of 1 $\sigma_{\rm c}$. Each of the subsequent terms corresponds to one subsample, and is given by:

$$\mathcal{L}_{\rm sample} = \mathcal{L}_{\rm shape} \times \mathcal{L}_{\rm bg} \tag{14}$$

The term in the likelihood most critical to the mass measurement is the extended maximum likelihood shape term:

$$\mathcal{L}_{\rm shape} = \frac{e^{-(n_s+n_b)}(n_s+n_b)^N}{N!} \times \prod_{i=1}^{N} \frac{n_s P_s(x_i,y_i;{\rm M_{top}},\Delta_{\rm JES}) + n_b P_b(x_i,y_i;\Delta_{\rm JES})}{n_s+n_b} \tag{15}$$

where the product runs over all events in a given subsample. The observables (x_i, y_i) in the ith event are $(m_t^{\rm reco}, m_{jj})$ in the lepton + jets channel and $(m_t^{\rm NWA}, H_T)$ in the dilepton channel. The quantities P_s and P_b designate the signal and background PDFs as determined by KDE and LPS. To improve the precision of the measurement, we apply a Gaussian constraint to the expected number of background events:

$$\mathcal{L}_{\rm bg} = e^{-\frac{(n_b - n_{b_0})^2}{2\sigma^2_{n_{b_0}}}} \tag{16}$$

where n_{b_0} is the *a priori* estimate for the expected number of background events, and $\sigma_{n_{b_0}}$ is the uncertainty on the estimate. Both sets of numbers are given in Table II and Table III.

We minimize the negative logarithm of the likelihood with respect to all 10 parameters using MINUIT. The uncertainty on $\rm M_{top}$ and $\Delta_{\rm JES}$ is found by searching for the points where the negative logarithm of the likelihood minimized with respect to all other parameters deviates by $\frac{1}{2}$ from the minimum. The uncertainty on the top quark measurement obtained this way includes the statistical uncertainty as well as the systematic uncertainty due to allowed variations in the jet energy scale and the background estimates. We scale the uncertainty on the top quark mass by the pull width as obtained in Sec. V D. For the dilepton-only cross-check, we fix the value of the $\Delta_{\rm JES}$ parameter to 0 $\sigma_{\rm c}$ and perform a fit only for $\rm M_{top}$, as the dilepton channel has no power to resolve $\rm M_{top}$ and $\Delta_{\rm JES}$ simultaneously.

D. Method check

We test the likelihood procedure using large numbers of MC simulated experiments, each of which is generated for a specific value of $\rm M_{top}$ and $\Delta_{\rm JES}$. In each experiment we select the number of background events from a Poisson distribution with a mean equal to the expected number of background events in the sample. The number of signal events is selected from a Poisson distribution with a mean equal to the expected number of signal events assuming a $t\bar{t}$ pair production cross section of 6.7 pb. The signal events are drawn at random from a MC sample generated at a given $\rm M_{top}$ and $\Delta_{\rm JES}$. The background events are drawn from the entire background sample with probabilities corresponding to the different background sources and the weights of individual events given by the model. Individual event rates can vary based for example on mistag probability of jets in a given event. An event (signal or background) can be drawn arbitrary number of times from a given MC sample so that different MC experiments can share simulated data. We refer to this method of drawing events from MC samples as "drawing with replacement". Once the simulated data is constructed we perform a maximum likelihood fit as described in the previous sections. We fluctuate the constraint on JES and the constraint on the expected number of background events in each subsample to reflect the possible deviation of those parameters from their *a priori* expected values. The fluctuations on the constraints are applied to estimate the effect of our limited knowledge about the nuisance parameters on the top quark

mass measurement. The jet energy scale constraint in Eq. (13) is replaced by:

$$e^{-\frac{(\Delta_{\rm JES}-\Delta_{{\rm JES}_f})^2}{2}} \tag{17}$$

where $\Delta_{{\rm JES}_f}$ is randomly selected in each MC simulated experiment from a Gaussian with mean corresponding to the $\Delta_{\rm JES}$ being tested and unit width. Similarly, the quantity n_{b_0} in Eq. (16) is replaced in each MC simulated experiment by a value drawn from a Gaussian with mean of n_{b_0} and width $\sigma_{n_{b_0}}$.

The likelihood fit should, on average, return the value of the top quark mass used to generate the MC simulated experiments. Figure 6 shows the average residual (deviation from expectation for an unbiased measurement) from 3000 MC simulated experiments for a range of $\mathrm{M_{top}}$. The fit to a constant shows no bias for the combined and lepton + jets-only fits, and a small positive bias for the dilepton-only fit. This bias does not warrant a correction since it is small in comparison to the expected uncertainty, and has a probability of 9% for a purely statistical fluctuation to generate the observed shift.

The statistical uncertainty on the measurement is extracted from the data. To test whether the error estimate is sound, for each MC simulated experiment we calculate the pull, defined as a ratio of the residual to the uncertainty reported by MINUIT. If the residual is positive (negative) we use the negative (positive) error in the ratio. The width of the pull distribution for $\mathrm{M_{top}}$ is shown in Fig. 7. The average pull width is larger than 1.0 for the combined and lepton + jets-only measurements due to the finite number of events in the two-dimensional fits. The pull width correction is 3% for these measurements, and thus we increase the reported uncertainty in the data by this amount.

The residual and pull width for the $\Delta_{\rm JES}$ parameter are also investigated using the MC simulated experiment ensembles. Both the combined fit and the lepton + jets-only fit show a negative bias of $\lesssim 0.02\ \sigma_{\rm c}$ (with marginal statistical significance) and a pull width of ~ 1.04. Since $\Delta_{\rm JES}$ is a nuisance parameter, whose precise value is not as important as its effect on the top quark mass measurement, we do not correct the $\Delta_{\rm JES}$ value measured in data for this bias.

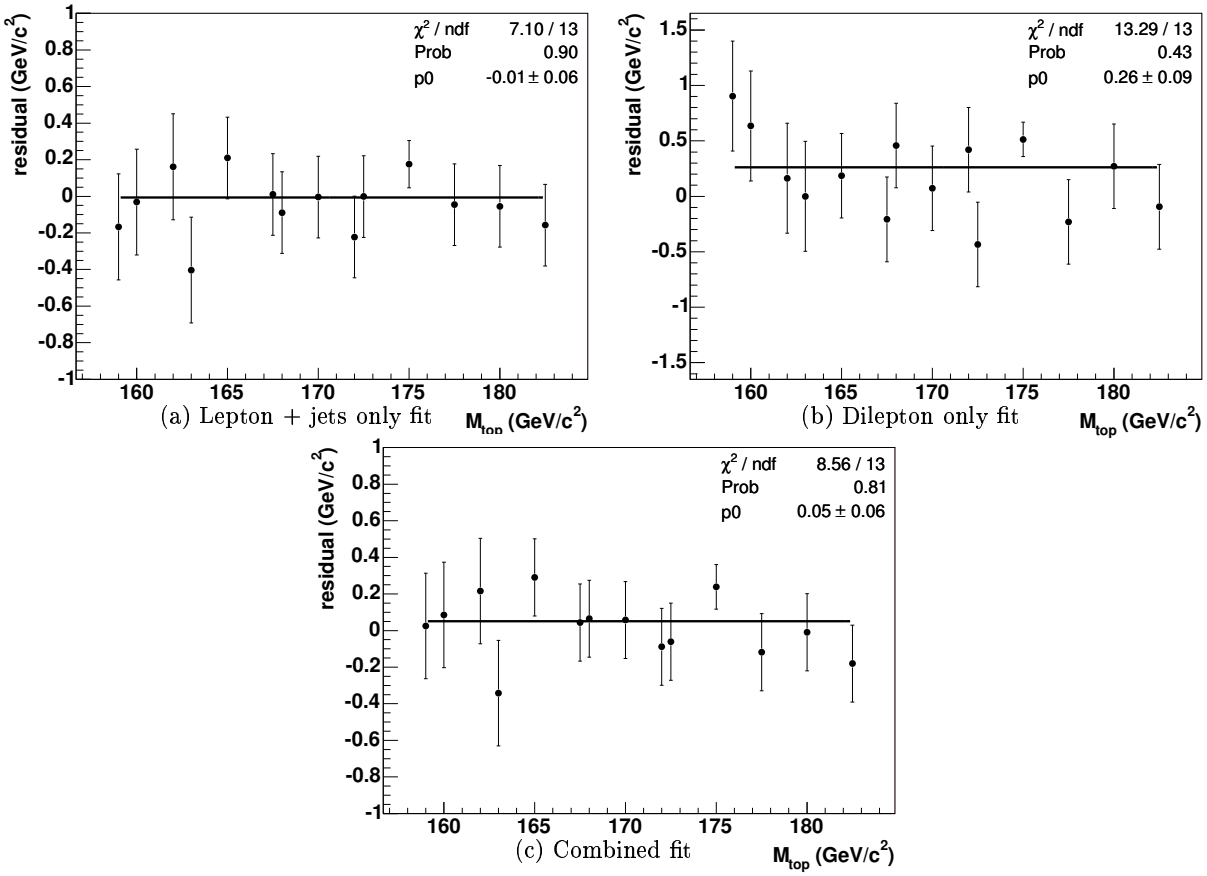
As noted above, when ensembles of simulated data are constructed the events are drawn with replacement. If we were to draw the events from MC samples without replacement such that no two MC simulated experiments were to share events, we would have only ~ 100 MC simulated experiments for each $t\bar{t}$ sample. Drawing events with replacement allows us to perform an arbitrarily large number of MC simulated experiments, fully exploring the possible combinations of events in order to check our machinery for possible biases. To evaluate the uncertainties on statistics such as the residuals and pull widths, we employ the bootstrap technique [46, 47]. In each bootstrap ensemble, we draw events from the signal MC sample with replacement until we reach the same number of events as in the original sample. We then run 3000 MC simulated experiments using this bootstrapped sample in place of the original sample. We repeat the above procedure 60 times. For each of the bootstrap ensembles, we calculate the desired statistic. The RMS of the statistic in question from the 60 bootstrap ensembles is taken as the uncertainty on the statistic. For example the uncertainty on the residual of fitted top quark mass in a typical MC sample is 0.4 GeV/c^2 for the dilepton fit and 0.2 GeV for the lepton + jets and combined fits

VI. RESULTS ON DATA

The likelihood fit when applied to the data yields $\mathrm{M_{top}} = 171.9 \pm 1.7$ (stat. + JES) GeV/c^2. The lepton + jets-only fit yields $\mathrm{M_{top}} = 171.8 \pm 1.9$ (stat. + JES) GeV/c^2. The dilepton-only fit, which does not include an *in situ* $\Delta_{\rm JES}$ measurement but instead fixes $\Delta_{\rm JES}$ to 0.0 $\sigma_{\rm c}$, yields $\mathrm{M_{top}} = 171.2\ ^{+3.6}_{-3.4}$ (stat.) GeV/c^2. The combined fit returns $\Delta_{\rm JES} = -0.17 \pm 0.35$ (stat. + $\mathrm{M_{top}}$) $\sigma_{\rm c}$, and the lepton + jets fit returns $\Delta_{\rm JES} = -0.12 \pm 0.37$ (stat. + $\mathrm{M_{top}}$) $\sigma_{\rm c}$. The results above have been corrected for the pull width and high instantaneous luminosity effects. Results from the combined fit, including fitted numbers of signal and background events for each subsample, are summarized in Table IV. The dilepton-only fit and lepton + jets-only fit both return $\mathrm{M_{top}}$ value lower than the $\mathrm{M_{top}}$ measured in both channels simultaneously. This is due to the *in situ* JES calibration extracted from the lepton + jets channel events being applied to the dilepton channel data in the combined fit.

The log-likelihood contours for the combined measurement are shown in Fig. 8. The one-dimensional log-likelihood for the dilepton-only measurement and the log-likelihood contours for the lepton + jets-only measurement are shown in Fig. 9 and Fig. 10, respectively. Figure 11 shows the one-dimensional lepton + jets data with the best-fit one-dimensional signal and background distributions overlaid on top. Figure 12 shows the distributions for the dilepton data. Using the observed number of events in data and the background expectations, 10% of MC experiments have a smaller error than the value measured in the combined fit. The p-value for the lepton + jets-only fit is 21%; the value for the dilepton-only fit is 14%.

We fit without the *a priori* JES and background constraints and measure the same $\mathrm{M_{top}} = 171.9 \pm 1.7$ (stat.+ JES) GeV/c^2, showing that these priors do not significantly affect our result. We also fit separately in the several individual subsamples: the first 1 fb^{-1} and last 0.9 fb^{-1} of data, electron and muon events in the lepton + jets-only fit, and different lepton pair type events in the dilepton-only fit. In addition we quote the top quark mass fitted in lepton + jets and dilepton subsamples separated by b tagging multiplicity. The results are consistent across cross-checks and are summarized in Table V.



(a) Lepton + jets only fit (b) Dilepton only fit (c) Combined fit

FIG. 6: Checks for bias in the fitted top quark mass for a) lepton + jets only fit, b) dilepton only fit and c) combined fit.

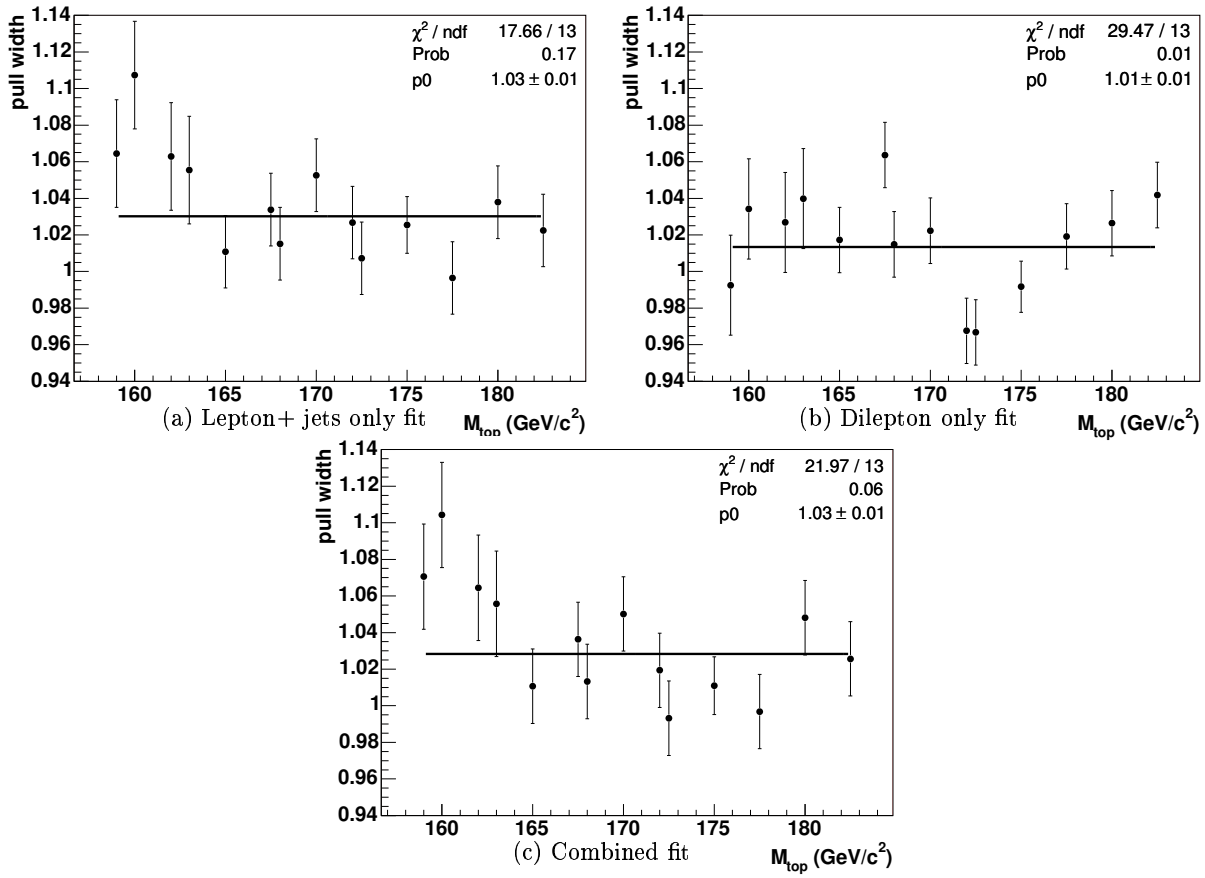
TABLE IV: The input constraints and fitted values are given for all free parameters in the combined likelihood fit. LJ refers to lepton + jets subsamples and DIL refers to dilepton subsamples. Con. refers to the constraint used in the likelihood.

Category		LJ 2-tag	LJ 1-tag	DIL tagged	DIL 0-tag
M_{top}	con.	None			
	fit	$\mathbf{171.9 \pm 1.7\ GeV/c^2}$			
JES	con.	$0.0 \pm 1.0\ \sigma_c$			
	fit	$-0.17 \pm 0.35\ \sigma_c$			
n_s	con.	None			
	fit	$96.4\ ^{+10.4}_{-9.7}$	$184.1\ ^{+17.7}_{-17.1}$	$56.9\ ^{+8.2}_{-7.5}$	$43.8\ ^{+10.4}_{-9.8}$
n_b	con.	4.2 ± 1.9	42.7 ± 12.5	3.9 ± 1.0	44.3 ± 7.0
	fit	3.4 ± 1.9	$47.4\ ^{+10.2}_{-10.0}$	3.9 ± 1.0	$41.5\ ^{+6.5}_{-6.4}$

The results are not corrected for pull width effects and bias due to incorrect instantaneous luminosity profile of the MC samples. With the exception of the fit without the *a priori* JES constraint, all cross-checks include the JES prior. The fits in the four subsamples separated by b tagging multiplicity include the *a priori* background constraints; all other cross-checks do not include background constraints.

VII. SYSTEMATIC UNCERTAINTIES

We examine a variety of effects that could affect our measurement by comparing MC simulated experiments in which we change systematic parameters within their uncertainties. As a single nuisance parameter, the measured $\Delta_{\rm JES}$ does not fully capture the complexities of jet energy scale uncertainties, particularly those with different η and p_T dependence. Fitting for the global JES removes most of these effects, but not all of them. In order to estimate the total residual JES uncertainty, we vary JES parameters within their uncertainties in both signal and background MC generated data and measure resulting shifts in M_{top}. We also conduct MC simulated experiments where we assume JES uncertainties are not fully correlated between jets of different momenta. So as not to bias the results, we remove the JES prior for these experiments. For the dilepton-only measurement, which has no *in situ* calibration, these systematics dominate. To form a b jet energy scale systematic we replace the default parameters of the Bowler fragmentation function [48] in the PYTHIA simulation with the parameters obtained by the DØ collaboration in a PYTHIA tune to the LEP and SLD data [49]. We also vary the semileptonic branching fractions of b and c quarks within their uncertainties given by [50] and [1]. The calorimeter response



(a) Lepton+ jets only fit (b) Dilepton only fit (c) Combined fit

FIG. 7: Width of the pull distribution for the fitted M_{top} for the a) lepton + jets only fit, b) dilepton only fit and c) combined fit.

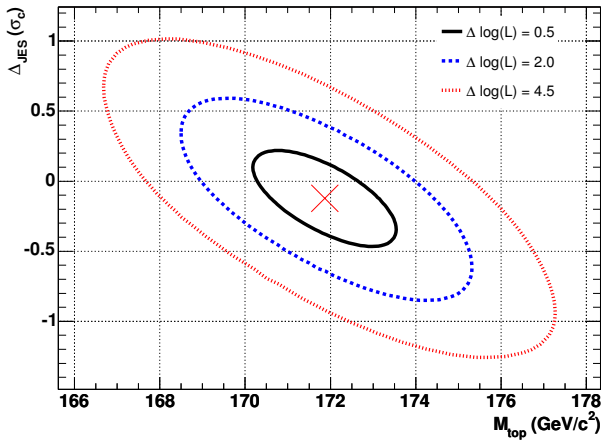


FIG. 8: Negative log-likelihood contours for the combined fit. The minimum is indicated by the 'x' and corresponds to the most probable top quark mass and $\Delta_{\rm JES}$, given the data. The contours are drawn at values of 0.5, 2.0 and 4.5 of the increase of the log-likelihood from the minimum value. These curves correspond to the 1,2 and 3 σ uncertainty on the measurement of the top quark mass.

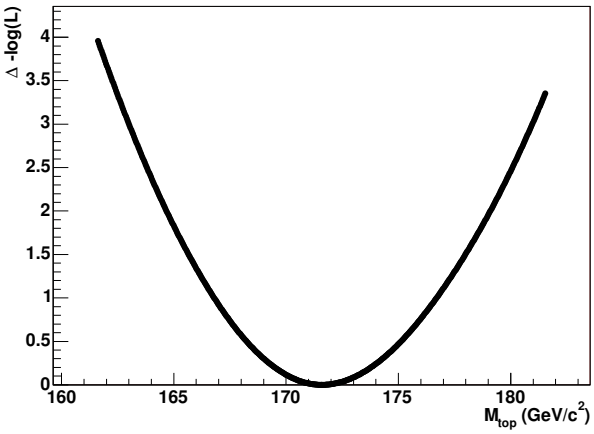


FIG. 9: One-dimensional log-likelihood for the dilepton-only fit.

to b jets is varied to capture differences in absolute jet energy scale uncertainties for light flavor and b quarks. Effects due to uncertain modeling of initial-state gluon radiation (ISR) and final-state gluon radiation (FSR) are studied by extrapolating uncertainties in the p_T of Drell-

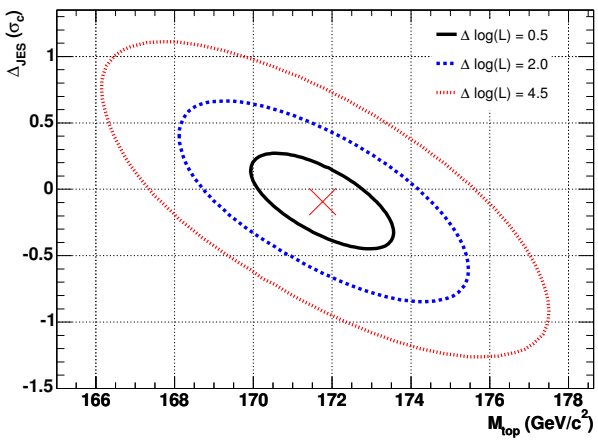


FIG. 10: Negative log-likelihood contours for the lepton + jets-only fit. The minimum is indicated by the 'x' and corresponds to the most probable top quark mass and $\Delta_{\rm JES}$, given the data. The contours are drawn at values of 0.5, 2.0 and 4.5 of the increase of the log-likelihood from the minimum value. These curves correspond to the 1,2 and 3 σ uncertainty on the measurement of the top quark mass.

TABLE V: Cross-checks on the data. LJ refers to the lepton + jets-only fit, DIL refers to the dilepton-only fit, and Combo refers to the combined fit. All numbers are uncorrected for pull width effects and bias due to incorrect instantaneous luminosity profile of the MC samples. For the dilepton-only fits, $\Delta_{\rm JES}$ is fixed to 0.0 σ_c.

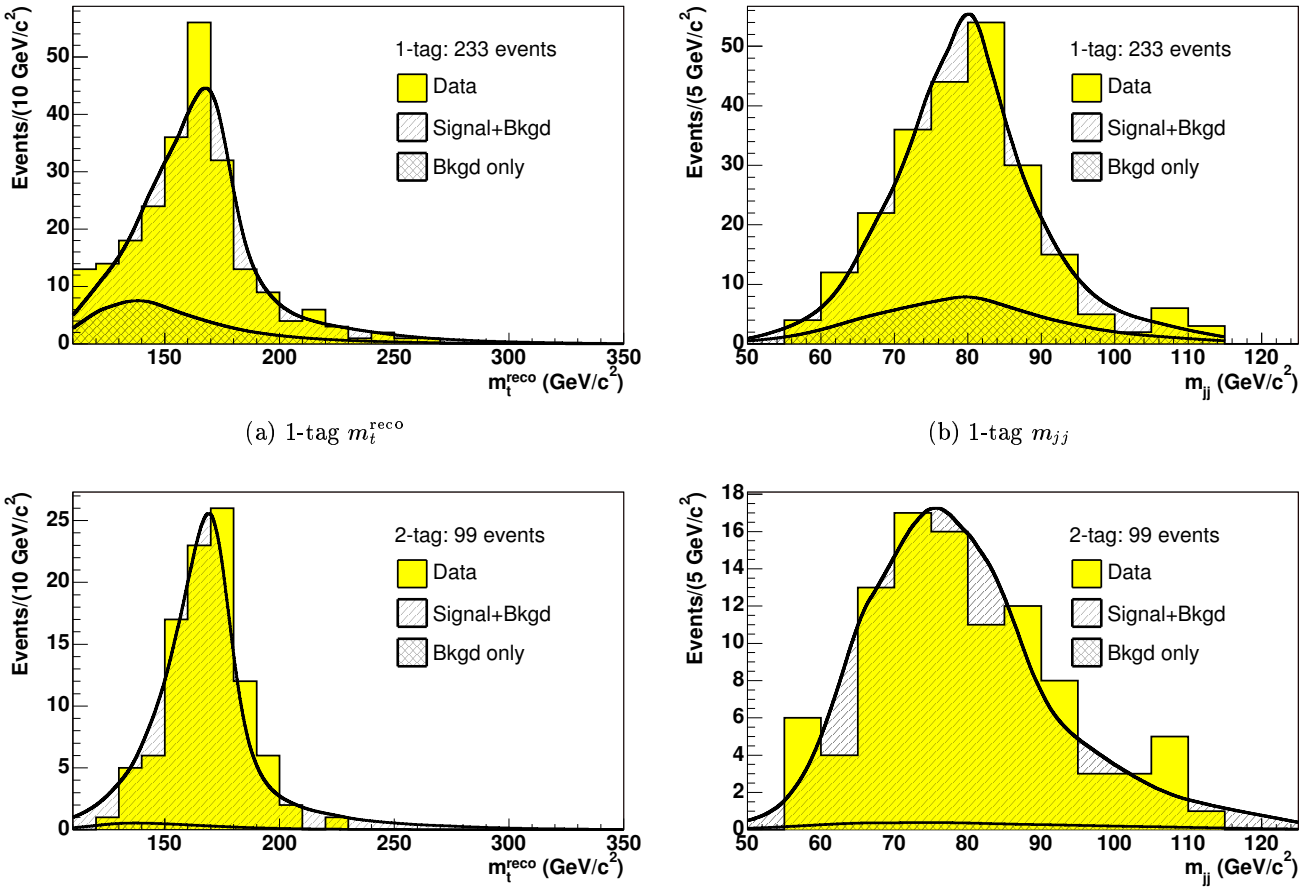
Fit type	Sample	$M_{\rm top}$ (GeV/c^2)	$\Delta_{\rm JES}$ (σ_c)
Nominal	Combo	171.9 ± 1.7	-0.12 ± 0.34
	LJ	171.8 ± 1.8	-0.09 ± 0.36
	DIL	$171.6\ ^{+3.5}_{-3.3}$	-
No JES prior	Combo	171.9 ± 1.7	$-0.14\ ^{+0.36}_{-0.37}$
	LJ	171.8 ± 1.9	$-0.11\ ^{+0.39}_{-0.38}$
	DIL	$171.6\ ^{+3.5}_{-3.3}$	-
No bkgd prior	Combo	171.9 ± 1.7	$-0.11\ ^{+0.35}_{-0.34}$
	LJ	171.8 ± 1.8	-0.06 ± 0.36
	DIL	171.5 ± 3.4	-
1-tag LJ	LJ	$169.1\ ^{+3.1}_{-2.6}$	$-0.17\ ^{+0.48}_{-0.57}$
2-tag LJ	LJ	$173.6\ ^{+2.6}_{-2.3}$	$0.20\ ^{+0.47}_{-0.50}$
0-tag DIL	DIL	$170.1\ ^{+6.4}_{-7.6}$	-
Tagged DIL	DIL	$172.2\ ^{+4.4}_{-4.0}$	-
e only LJ	LJ	172.2 ± 2.7	-0.09 ± 0.51
μ only LJ	LJ	$171.3\ ^{+2.4}_{-2.3}$	$-0.04\ ^{+0.46}_{-0.47}$
ee only DIL	DIL	169.0 ± 8.0	-
$e\mu$ only DIL	DIL	$173.6\ ^{+5.2}_{-4.0}$	-
$\mu\mu$ only DIL	DIL	$167.9\ ^{+7.5}_{-6.8}$	-
First 1 fb^{-1}	Combo	$171.7\ ^{+2.3}_{-2.4}$	$0.45\ ^{+0.55}_{-0.50}$
	LJ	$172.2\ ^{+2.5}_{-2.4}$	$0.59\ ^{+0.52}_{-0.55}$
	DIL	166.1 ± 5.0	-
Last 0.9 fb^{-1}	Combo	171.7 ± 2.7	$-0.70\ ^{+0.53}_{-0.59}$
	LJ	$170.2\ ^{+3.1}_{-3.0}$	$-0.61\ ^{+0.54}_{-0.64}$
	DIL	$175.2\ ^{+5.3}_{-4.7}$	-

Yan events to the $t\bar{t}$ mass region, resulting in a systematic on ISR-FSR modeling [7]. Note that unlike in Ref. [7], we coherently shift parameters affecting both ISR and FSR, as the uncertainties on the two effects should be correlated. We measure the uncertainty due to generator choice by comparing MC simulated experiments generated with HERWIG and PYTHIA. A systematic on different parton distribution functions is obtained by varying the independent eigenvectors of the CTEQ6M set [51], comparing parton distribution functions with different values of Λ_{QCD}, and comparing CTEQ5L [52] with MRST72 [53]. The gluon fusion fraction uncertainty is calculated by reweighting the MC samples to increase the fraction of $t\bar{t}$ events initiated by gluons instead of quarks from the 6% in the leading order MC samples used for the measurement to 20%, which is given as the 1σ upper bound on the gluon fusion fraction in [27]. Systematic uncertainties due to lepton energy scale are estimated by propagating shifts on electron and muon energies within their uncertainties. Background shape systematic uncertainties are obtained by varying the fraction of the different types of backgrounds in MC simulated experiments. For lepton+jets backgrounds, we generate further changes in the shapes by varying the Q^2 used in the calculation of hard scattering and shower evolution in the range $M_W^2/4$–$4M_W^2$. We also substitute the anti-electron QCD model for the non-isolated lepton model in MC simulated experiments. For dilepton backgrounds changing the shape of the Drell-Yan sample according to the difference in the missing energy distribution observed in data and simulation gives one systematic effect. We also shift the fake model in ways expected to maximally correlate with the reconstructed mass. The systematic uncertainty due to limited signal MC statistics is taken as the uncertainty on the fit to a constant of the residuals obtained in MC experiments (Fig. 6). We study the effects of limited background MC statistics using the bootstrap technique, where multiple background MC data ensembles are generated. The "pileup" systematic is induced by the possible mismodelling of the Minimum Bias events in the MC samples. These events are found to deposit more energy in jets in simulation than expected from data. It has been suggested that color effects may cause a systematic bias of order 0.5 GeV/c^2 which is not accounted in our studies [54].

The systematic uncertainties are summarized in Table VI. The total systematic uncertainty is 1.1 GeV/c^2 for both the combined and the lepton+jets measurement, and 3.8 GeV/c^2 for the dilepton-only measurement.

VIII. CONCLUSIONS

We present the first measurement of the top quark mass across multiple decay topologies using a joint likelihood fit. Our procedure includes a full treatment of correlations between systematics, and does not assume Gaussian likelihoods or symmetric errors in the channels



(a) 1-tag $m_t^{\rm reco}$ (b) 1-tag m_{jj}

(c) 2-tag $m_t^{\rm reco}$ (d) 2-tag m_{jj}

FIG. 11: One-dimensional lepton + jets data with density estimates overlaid using $\mathrm{M_{top}} = 172.0$ GeV/c^2, $\Delta_{\rm JES} = 0.0$, and a full background model. The expected numbers of events are set to the values from the constrained fit. Shown are the 1-tag $m_t^{\rm reco}$ (a) and m_{jj} (b) distributions, and the 2-tag $m_t^{\rm reco}$ (c) and m_{jj} (d) distributions.

being combined. In 1.9 fb^{-1} of data, we measure:

$$\begin{aligned}\mathrm{M_{top}} &= 171.9 \pm 1.7 \text{ (stat.} + \text{JES)} \\ &\quad \pm 1.1 \text{ (other syst.) GeV}/c^2 \\ &= 171.9 \pm 2.0 \text{ GeV}/c^2\end{aligned}$$

with cross-checks using events from the lepton+jets and dilepton channels separately:

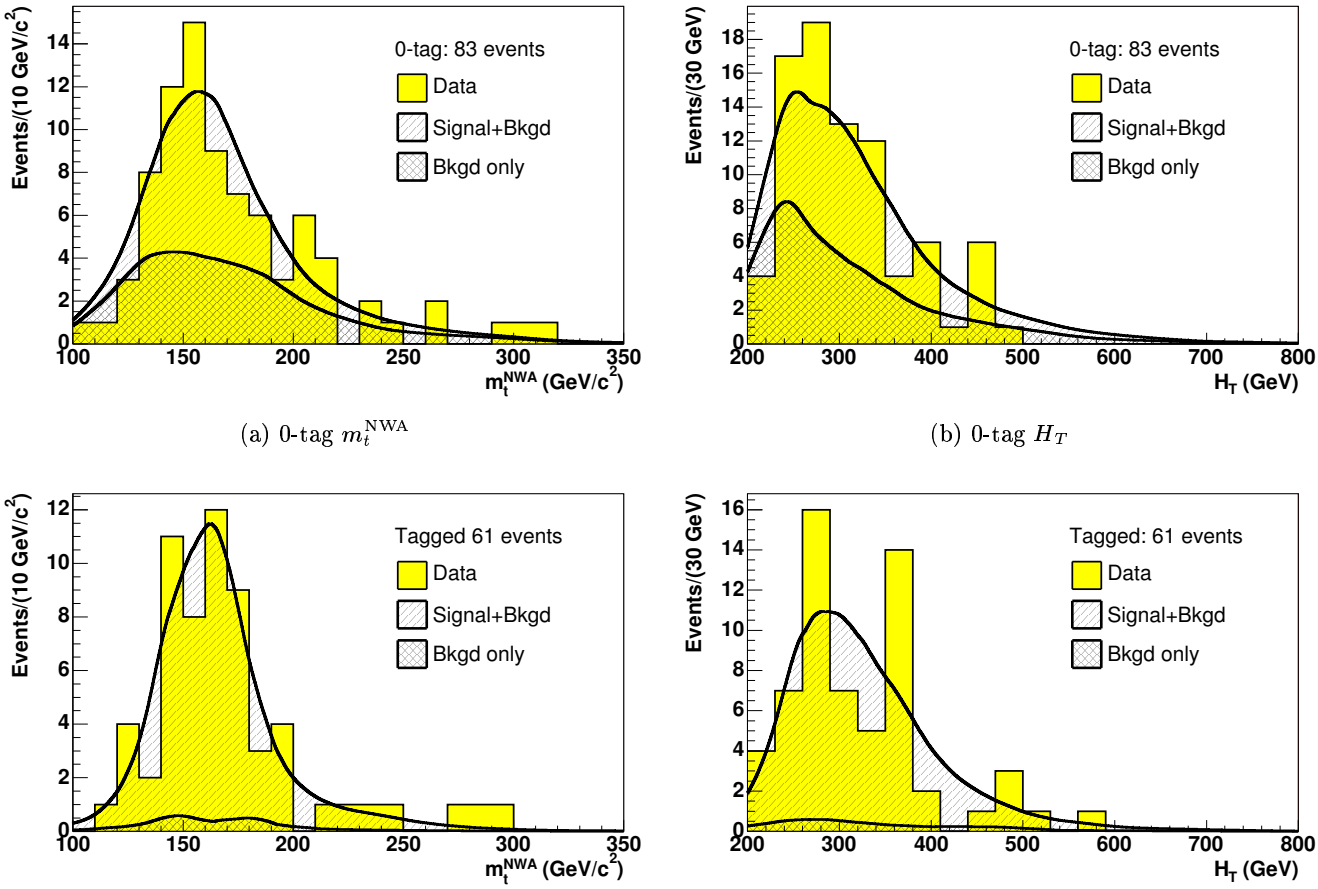
$$\begin{aligned}\mathrm{M_{top}} &= 171.8 \pm 1.9 \text{ (stat.} + \text{JES)} \\ &\quad \pm 1.1 \text{ (other syst.) GeV}/c^2 \\ &= 171.8 \pm 2.2 \text{ GeV}/c^2 \\ &\qquad \text{(lepton+jets only);} \\ \mathrm{M_{top}} &= 171.2\ ^{+3.6}_{-3.4} \text{ (stat.)} \\ &\quad \pm 3.8 \text{ (syst.) GeV}/c^2 \\ &= 171.2\ ^{+5.3}_{-5.1} \text{ GeV}/c^2 \\ &\qquad \text{(dilepton only).}\end{aligned}$$

This measurement increases our understanding of physics in the top quark sector, and contributes to tests of the mechanism of electroweak symmetry breaking. In addition, the analysis methods and tools described in this article will be applicable to other measurements at the Tevatron experiments, and soon at CERN's Large Hadron Collider.

The precision of top quark mass measurements at the Tevatron is approaching limits set by our understanding of non-perturbative QCD phenomena. Nevertheless, some further improvements are expected as CDF accumulates a factor of 3–4 times more data during run II and as studies of important systematic effects provide additional constraints on those uncertainties.

Acknowledgments

We thank the Fermilab staff and the technical staffs of the participating institutions for their vital contributions. This work was supported by the U.S. Department of En-



(a) 0-tag m_t^{NWA}

(b) 0-tag H_T

(c) Tagged m_t^{NWA}

(d) Tagged H_T

FIG. 12: One-dimensional dilepton data with density estimates overlaid using $\mathrm{M_{top}} = 172.0$ GeV/c^2, $\Delta_{\mathrm{JES}} = 0.0$, and full background model. The expected numbers of events are set to the values from the constrained fit. Shown are the 0-tag m_t^{NWA} (a) and H_T (b) distributions, and the tagged m_t^{NWA} (c) and H_T (d) distributions.

ergy and National Science Foundation; the Italian Istituto Nazionale di Fisica Nucleare; the Ministry of Education, Culture, Sports, Science and Technology of Japan; the Natural Sciences and Engineering Research Council of Canada; the National Science Council of the Republic of China; the Swiss National Science Foundation; the A.P. Sloan Foundation; the Bundesministerium für Bildung und Forschung, Germany; the Korean Science and Engineering Foundation and the Korean Research Foundation; the Science and Technology Facilities Council and the Royal Society, UK; the Institut National de Physique Nucleaire et Physique des Particules/CNRS; the Russian Foundation for Basic Research; the Ministerio de Ciencia e Innovación, Spain; the Slovak R&D Agency; and the Academy of Finland.

[1] W. M. Yao *et al.* (Particle Data Group), J. Phys. G **33**, 1 (2006), and 2007 partial update for the 2008 edition.
[2] C. T. Hill and S. J. Parke, Phys. Rev. D **49**, 4454 (1994).
[3] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **74**, 2626 (1995).
[4] S. Abachi *et al.* (D0 Collaboration), Phys. Rev. Lett. **74**, 2632 (1995).
[5] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **73**, 225 (1994).
[6] F. Abe *et al.* (CDF Collaboration), Phys. Rev. D **50**, 2966 (1994).
[7] A. Abulencia *et al.* (CDF Collaboration), Phys. Rev. D **73**, 032003 (2006).
[8] B. Abbott *et al.* (D0 Collaboration), Phys. Rev. D **60**, 052001 (1999).
[9] A. Abulencia *et al.* (CDF Collaboration), Phys. Rev. D **73**, 112006 (2006).
[10] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **75**, 3997 (1995).
[11] L. Lyons, D. Gibaut, and P. Clifford, Nucl. Instrum.

TABLE VI: Summary of systematics uncertainties. All numbers have units of GeV/c^2. Comb refers to the combined fit, LJ refers to the lepton+jets-only fit and DIL refers to the dilepton-only fit.

Systematic	Comb	LJ	DIL
Residual JES	0.7	0.7	3.5
Generator	0.7	0.8	1.3
Parton distribution functions	0.3	0.3	0.5
b jet energy	0.2	0.2	0.2
Background shape	0.2	0.2	0.3
Gluon fusion fraction	0.2	0.2	0.2
Initial and final state radiation	0.1	0.2	0.2
MC statistics	0.1	0.1	0.5
Lepton energy scale	0.1	0.1	0.3
Pileup	0.1	0.1	0.1
Combined	1.1	1.1	3.8

Methods A **270**, 110 (1988).

[12] D. Acosta *et al.* (CDF Collaboration), Phys. Rev. D **71**, 032001 (2005).

[13] A. Sill *et al.*, Nucl. Instrum. Methods A **447**, 1 (2000).

[14] T. Affolder *et al.*, Nucl. Instrum. Methods A **526**, 249 (2004).

[15] L. Balka *et al.*, Nucl. Instrum. Methods A **267**, 272 (1988).

[16] S. Bertolucci *et al.*, Nucl. Instrum. Methods A **267**, 301 (1988).

[17] M. Albrow *et al.*, Nucl. Instrum. Methods A **480**, 524 (2002).

[18] G. Apollinari *et al.*, Nucl. Instrum. Methods A **412**, 515 (1998).

[19] A. Abulencia *et al.* (CDF Collaboration), J. Phys. G **34**, 2457 (2007).

[20] C. Issever, AIP Conf. Proc. **670**, 371 (2003).

[21] D. Gross and F. Wilczek, Phys. Rev. Lett **30**, 1343 (1973).

[22] H. D. Politzer, Phys. Rev. Lett **30**, 1346 (1973).

[23] F. Abe *et al.* (CDF Collaboration), Phys. Rev. D **45**, 1448 (1992).

[24] A. Bhatti *et al.*, Nucl. Instrum. Methods A **566**, 375 (2006).

[25] T. Affolder *et al.* (CDF Collaboration), Phys. Rev. D **64**, 032002 (2001).

[26] D. Acosta *et al.* (CDF Collaboration), Phys. Rev. D **71**, 052003 (2005).

[27] M. Cacciari *et al.*, J. High Energy Phys. **04**, 068 (2004).

[28] T. Sjostrand *et al.*, Comput. Phys. Commun. **135**, 238 (2001).

[29] G. Corcella, I. G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M. H. Seymour, and B. R. Webber, J. High Energy Phys. **01**, 010 (2001).

[30] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, J. High Energy Phys. **07**, 001 (2003).

[31] M. L. Mangano, M. Moretti, F. Piccinini, and M. Treccani, J. High Energy Phys. **01**, 013 (2007).

[32] F. Maltoni and T. Stelzer, J. High Energy Phys. **0302**, 027 (2003).

[33] S. M. T. Sjostrand and P. Skands, J. High Energy Phys. **05**, 026 (2006).

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

[35] E. A. Gerchtein and M. Paulini, ECONF **C0303241**, TUMT005 (2003), physics/0306031.

[36] A. Abulencia *et al.* (CDF Collaboration), Phys. Rev. D **73**, 112006 (2006).

[37] D. W. Scott, *Multivariate Density Estimation: Theory, Practive and Visualization* (Wiley-Interscience, 1992).

[38] B. W. Silverman, *Density Estimation for Statistics and Data Analysis* (Chapman and Hall, 1986).

[39] K. Cranmer, Computer Physical Communcations **136**, 198 (2001).

[40] I. S. Abramson, The annals of Statistics **10**, 1217 (1982).

[41] G. R. Terrell, Journal of the American Statistical Association **85**, 470 (1990).

[42] M. P. Wand and M. C. Jones, Journal of the American Statistical Association **88**, 520 (1993).

[43] S. Sain, *Multivariate locally adaptive density estimation* (1999).

[44] C. Loader, *Local Regression and Likelihood* (Springer, 1999).

[45] R. J. Barlow, Nucl. Instrum. Methods A **297**, 496 (1990).

[46] B. Efron, The Annals of Statistics **7**, 1 (1979).

[47] B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap: Monographs on Statistics and Applied Probability 57* (Chapman and Hall, 1993).

[48] M. Bowler, Z. Phys. C **22**, 155 (1984).

[49] W. Fedorko, Ph.D. thesis, University of Chicago (2008), FERMILAB-THESIS-2008-42.

[50] Phys. Rep. **427**, 257 (2006).

[51] J. Pumplin *et al.*, J. High Energy Phys. **07**, 012 (2002).

[52] H. L. Lai *et al.* (CTEQ Collaboration), Eur. Phys. J. C **12**, 375 (2000).

[53] A. D. Martin *et al.*, Eur. Phys. J. C **14**, 133 (2000).

[54] D. Wicke and P. Z. Skands, arXiv:hep-ph/0807.3248 (2008), to appear in the proceedings of TOP2008: International Workshop on Top-Quark Physics, La Biodola, Isola d'Elba, Italy, 18-24 May 2008.