This is the accepted manuscript made available via CHORUS. The article has been published as:

Search for the standard model Higgs boson produced in association with a W^{±} boson with 7.5 fb^{-1} integrated luminosity at CDF

T. Aaltonen *et al.* (CDF Collaboration)

Phys. Rev. D **86**, 032011 — Published 20 August 2012

DOI: 10.1103/PhysRevD.86.032011

Search for the standard model Higgs boson produced in association with a $W^{\pm}$ boson with 7.5 fb^{-1} integrated luminosity at CDF

T. Aaltonen,[21] B. Álvarez González[z],[9] S. Amerio,[40] D. Amidei,[32] A. Anastassov[x],[15] A. Annovi,[17] J. Antos,[12] G. Apollinari,[15] J.A. Appel,[15] T. Arisawa,[54] A. Artikov,[13] J. Asaadi,[49] W. Ashmanskas,[15] B. Auerbach,[57] A. Aurisano,[49] F. Azfar,[39] W. Badgett,[15] T. Bae,[25] A. Barbaro-Galtieri,[26] V.E. Barnes,[44] B.A. Barnett,[23] P. Barria[hh],[42] P. Bartos,[12] M. Bauce[ff],[40] F. Bedeschi,[42] S. Behari,[23] G. Bellettini[gg],[42] J. Bellinger,[56] D. Benjamin,[14] A. Beretvas,[15] A. Bhatti,[46] D. Bisello[ff],[40] I. Bizjak,[28] K.R. Bland,[5] B. Blumenfeld,[23] A. Bocci,[14] A. Bodek,[45] D. Bortoletto,[44] J. Boudreau,[43] A. Boveia,[11] L. Brigliadori[ee],[6] C. Bromberg,[33] E. Brucken,[21] J. Budagov,[13] H.S. Budd,[45] K. Burkett,[15] G. Busetto[ff],[40] P. Bussey,[19] A. Buzatu,[31] A. Calamba,[10] C. Calancha,[29] S. Camarda,[4] M. Campanelli,[28] M. Campbell,[32] F. Canelli,[11,15] B. Carls,[22] D. Carlsmith,[56] R. Carosi,[42] S. Carrillo[m],[16] S. Carron,[15] B. Casal[k],[9] M. Casarsa,[50] A. Castro[ee],[6] P. Catastini,[20] D. Cauz,[50] V. Cavaliere,[22] M. Cavalli-Sforza,[4] A. Cerri[f],[26] L. Cerrito[s],[28] Y.C. Chen,[1] M. Chertok,[7] G. Chiarelli,[42] G. Chlachidze,[15] F. Chlebana,[15] K. Cho,[25] D. Chokheli,[13] W.H. Chung,[56] Y.S. Chung,[45] M.A. Ciocci[hh],[42] A. Clark,[18] C. Clarke,[55] G. Compostella[ff],[40] M.E. Convery,[15] J. Conway,[7] M.Corbo,[15] M. Cordelli,[17] C.A. Cox,[7] D.J. Cox,[7] F. Crescioli[gg],[42] J. Cuevas[z],[9] R. Culbertson,[15] D. Dagenhart,[15] N. d'Ascenzo[w],[15] M. Datta,[15] P. de Barbaro,[45] M. Dell'Orso[gg],[42] L. Demortier,[46] M. Deninno,[6] F. Devoto,[21] M. d'Errico[ff],[40] A. Di Canto[gg],[42] B. Di Ruzza,[15] J.R. Dittmann,[5] M. D'Onofrio,[27] S. Donati[gg],[42] P. Dong,[15] M. Dorigo,[50] T. Dorigo,[40] K. Ebina,[54] A. Elagin,[49] A. Eppig,[32] R. Erbacher,[7] S. Errede,[22] N. Ershaidat[dd],[15] R. Eusebi,[49] S. Farrington,[39] M. Feindt,[24] J.P. Fernandez,[29] R. Field,[16] G. Flanagan[u],[15] R. Forrest,[7] M.J. Frank,[5] M. Franklin,[20] J.C. Freeman,[15] Y. Funakoshi,[54] I. Furic,[16] M. Gallinaro,[46] J.E. Garcia,[18] A.F. Garfinkel,[44] P. Garosi[hh],[42] H. Gerberich,[22] E. Gerchtein,[15] S. Giagu,[47] V. Giakoumopoulou,[3] P. Giannetti,[42] K. Gibson,[43] C.M. Ginsburg,[15] N. Giokaris,[3] P. Giromini,[17] G. Giurgiu,[23] V. Glagolev,[13] D. Glenzinski,[15] M. Gold,[35] D. Goldin,[49] N. Goldschmidt,[16] A. Golossanov,[15] G. Gomez,[9] G. Gomez-Ceballos,[30] M. Goncharov,[30] O. González,[29] I. Gorelov,[35] A.T. Goshaw,[14] K. Goulianos,[46] S. Grinstein,[4] C. Grosso-Pilcher,[11] R.C. Group[53],[15] J. Guimaraes da Costa,[20] S.R. Hahn,[15] E. Halkiadakis,[48] A. Hamaguchi,[38] J.Y. Han,[45] F. Happacher,[17] K. Hara,[51] D. Hare,[48] M. Hare,[52] R.F. Harr,[55] K. Hatakeyama,[5] C. Hays,[39] M. Heck,[24] J. Heinrich,[41] M. Herndon,[56] S. Hewamanage,[5] A. Hocker,[15] W. Hopkins[g],[15] D. Horn,[24] S. Hou,[1] R.E. Hughes,[36] M. Hurwitz,[11] U. Husemann,[57] N. Hussain,[31] M. Hussein,[33] J. Huston,[33] G. Introzzi,[42] M. Iori[jj],[47] A. Ivanov[p],[7] E. James,[15] D. Jang,[10] B. Jayatilaka,[14] E.J. Jeon,[25] S. Jindariani,[15] M. Jones,[44] K.K. Joo,[25] S.Y. Jun,[10] T.R. Junk,[15] T. Kamon[25],[49] P.E. Karchin,[55] A. Kasmi,[5] Y. Kato[o],[38] W. Ketchum,[11] J. Keung,[41] V. Khotilovich,[49] B. Kilminster,[15] D.H. Kim,[25] H.S. Kim,[25] J.E. Kim,[25] M.J. Kim,[17] S.B. Kim,[25] S.H. Kim,[51] Y.K. Kim,[11] Y.J. Kim,[25] N. Kimura,[54] M. Kirby,[15] S. Klimenko,[16] K. Knoepfel,[15] K. Kondo*,[54] D.J. Kong,[25] J. Konigsberg,[16] A.V. Kotwal,[14] M. Kreps,[24] J. Kroll,[41] D. Krop,[11] M. Kruse,[14] V. Krutelyov[c],[49] T. Kuhr,[24] M. Kurata,[51] S. Kwang,[11] A.T. Laasanen,[44] S. Lami,[42] S. Lammel,[15] M. Lancaster,[28] R.L. Lander,[7] K. Lannon[y],[36] A. Lath,[48] G. Latino[hh],[42] T. LeCompte,[2] E. Lee,[49] H.S. Lee[q],[11] J.S. Lee,[25] S.W. Lee[bb],[49] S. Leo[gg],[42] S. Leone,[42] J.D. Lewis,[15] A. Limosani[t],[14] C.-J. Lin,[26] M. Lindgren,[15] E. Lipeles,[41] A. Lister,[18] D.O. Litvintsev,[15] C. Liu,[43] H. Liu,[53] Q. Liu,[44] T. Liu,[15] S. Lockwitz,[57] A. Loginov,[57] D. Lucchesi[ff],[40] J. Lueck,[24] P. Lujan,[26] P. Lukens,[15] G. Lungu,[46] J. Lys,[26] R. Lysak[e],[12] R. Madrak,[15] K. Maeshima,[15] P. Maestro[hh],[42] S. Malik,[46] G. Manca[a],[27] A. Manousakis-Katsikakis,[3] F. Margaroli,[47] C. Marino,[24] M. Martínez,[4] P. Mastrandrea,[47] K. Matera,[22] M.E. Mattson,[55] A. Mazzacane,[15] P. Mazzanti,[6] K.S. McFarland,[45] P. McIntyre,[49] R. McNulty[j],[27] A. Mehta,[27] P. Mehtala,[21] C. Mesropian,[46] T. Miao,[15] D. Mietlicki,[32] A. Mitra,[1] H. Miyake,[51] S. Moed,[15] N. Moggi,[6] M.N. Mondragon[m],[15] C.S. Moon,[25] R. Moore,[15] M.J. Morello[ii],[42] J. Morlock,[24] P. Movilla Fernandez,[15] A. Mukherjee,[15] Th. Muller,[24] P. Murat,[15] M. Mussini[ee],[6] J. Nachtman[n],[15] Y. Nagai,[51] J. Naganoma,[54] I. Nakano,[37] A. Napier,[52] J. Nett,[49] C. Neu,[53] M.S. Neubauer,[22] J. Nielsen[d],[26] L. Nodulman,[2] S.Y. Noh,[25] O. Norniella,[22] L. Oakes,[39] S.H. Oh,[14] Y.D. Oh,[25] I. Oksuzian,[53] T. Okusawa,[38] R. Orava,[21] L. Ortolan,[4] S. Pagan Griso[ff],[40] C. Pagliarone,[50] E. Palencia[f],[9] V. Papadimitriou,[15] A.A. Paramonov,[2] J. Patrick,[15] G. Pauletta[kk],[50] M. Paulini,[10] C. Paus,[30] D.E. Pellett,[7] A. Penzo,[50] T.J. Phillips,[14] G. Piacentino,[42] E. Pianori,[41] J. Pilot,[36] K. Pitts,[22] C. Plager,[8] L. Pondrom,[56] S. Poprocki[g],[15] K. Potamianos,[44] F. Prokoshin[cc],[13] A. Pranko,[26] F. Ptohos[h],[17] G. Punzi[gg],[42] A. Rahaman,[43] V. Ramakrishnan,[56] N. Ranjan,[44] I. Redondo,[29] P. Renton,[39] M. Rescigno,[47] T. Riddick,[28] F. Rimondi[ee],[6] L. Ristori[42],[15] A. Robson,[19] T. Rodrigo,[9] T. Rodriguez,[41] E. Rogers,[22] S. Rolli[i],[52] R. Roser,[15] F. Ruffini[hh],[42] A. Ruiz,[9] J. Russ,[10] V. Rusu,[15] A. Safonov,[49] W.K. Sakumoto,[45] Y. Sakurai,[54] L. Santi[kk],[50] K. Sato,[51] V. Saveliev[w],[15] A. Savoy-Navarro[aa],[15] P. Schlabach,[15] A. Schmidt,[24] E.E. Schmidt,[15] T. Schwarz,[15] L. Scodellaro,[9] A. Scribano[hh],[42] F. Scuri,[42] S. Seidel,[35] Y. Seiya,[38] A. Semenov,[13] F. Sforza[hh],[42] S.Z. Shalhout,[7] T. Shears,[27]

P.F. Shepard,[43] M. Shimojima[v],[51] M. Shochet,[11] I. Shreyber-Tecker,[34] A. Simonenko,[13] P. Sinervo,[31] K. Sliwa,[52] J.R. Smith,[7] F.D. Snider,[15] A. Soha,[15] V. Sorin,[4] H. Song,[43] P. Squillacioti[hh],[42] M. Stancari,[15] R. St. Denis,[19] B. Stelzer,[31] O. Stelzer-Chilton,[31] D. Stentz[x],[15] J. Strologas,[35] G.L. Strycker,[32] Y. Sudo,[51] A. Sukhanov,[15] I. Suslov,[13] K. Takemasa,[51] Y. Takeuchi,[51] J. Tang,[11] M. Tecchio,[32] P.K. Teng,[1] J. Thom[g],[15] J. Thome,[10] G.A. Thompson,[22] E. Thomson,[41] D. Toback,[49] S. Tokar,[12] K. Tollefson,[33] T. Tomura,[51] D. Tonelli,[15] S. Torre,[17] D. Torretta,[15] P. Totaro,[40] M. Trovato[ii],[42] F. Ukegawa,[51] S. Uozumi,[25] A. Varganov,[32] F. Vázquez[m],[16] G. Velev,[15] C. Vellidis,[15] M. Vidal,[44] I. Vila,[9] R. Vilar,[9] J. Vizán,[9] M. Vogel,[35] G. Volpi,[17] P. Wagner,[41] R.L. Wagner,[15] T. Wakisaka,[38] R. Wallny,[8] S.M. Wang,[1] A. Warburton,[31] D. Waters,[28] W.C. Wester III,[15] D. Whiteson[b],[41] A.B. Wicklund,[2] E. Wicklund,[15] S. Wilbur,[11] F. Wick,[24] H.H. Williams,[41] J.S. Wilson,[36] P. Wilson,[15] B.L. Winer,[36] P. Wittich[g],[15] S. Wolbers,[15] H. Wolfe,[36] T. Wright,[32] X. Wu,[18] Z. Wu,[5] K. Yamamoto,[38] D. Yamato,[38] T. Yang,[15] U.K. Yang[r],[11] Y.C. Yang,[25] W.-M. Yao,[26] G.P. Yeh,[15] K. Yi[n],[15] J. Yoh,[15] K. Yorita,[54] T. Yoshida[l],[38] G.B. Yu,[14] I. Yu,[25] S.S. Yu,[15] J.C. Yun,[15] A. Zanetti,[50] Y. Zeng,[14] C. Zhou,[14] and S. Zucchelli[ee]6

(CDF Collaboration[†])

[1]*Institute of Physics, Academia Sinica, Taipei, Taiwan 11529, Republic of China*
[2]*Argonne National Laboratory, Argonne, Illinois 60439, USA*
[3]*University of Athens, 157 71 Athens, Greece*
[4]*Institut de Fisica d'Altes Energies, ICREA, Universitat Autonoma de Barcelona, E-08193, Bellaterra (Barcelona), Spain*
[5]*Baylor University, Waco, Texas 76798, USA*
[6]*Istituto Nazionale di Fisica Nucleare Bologna, [ee] University of Bologna, I-40127 Bologna, Italy*
[7]*University of California, Davis, Davis, California 95616, USA*
[8]*University of California, Los Angeles, Los Angeles, California 90024, USA*
[9]*Instituto de Fisica de Cantabria, CSIC-University of Cantabria, 39005 Santander, Spain*
[10]*Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA*
[11]*Enrico Fermi Institute, University of Chicago, Chicago, Illinois 60637, USA*
[12]*Comenius University, 842 48 Bratislava, Slovakia; Institute of Experimental Physics, 040 01 Kosice, Slovakia*
[13]*Joint Institute for Nuclear Research, RU-141980 Dubna, Russia*
[14]*Duke University, Durham, North Carolina 27708, USA*
[15]*Fermi National Accelerator Laboratory, Batavia, Illinois 60510, USA*
[16]*University of Florida, Gainesville, Florida 32611, USA*
[17]*Laboratori Nazionali di Frascati, Istituto Nazionale di Fisica Nucleare, I-00044 Frascati, Italy*
[18]*University of Geneva, CH-1211 Geneva 4, Switzerland*
[19]*Glasgow University, Glasgow G12 8QQ, United Kingdom*
[20]*Harvard University, Cambridge, Massachusetts 02138, USA*
[21]*Division of High Energy Physics, Department of Physics,*
University of Helsinki and Helsinki Institute of Physics, FIN-00014, Helsinki, Finland
[22]*University of Illinois, Urbana, Illinois 61801, USA*
[23]*The Johns Hopkins University, Baltimore, Maryland 21218, USA*
[24]*Institut für Experimentelle Kernphysik, Karlsruhe Institute of Technology, D-76131 Karlsruhe, Germany*
[25]*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; Chonbuk National University, Jeonju 561-756, Korea
[26]*Ernest Orlando Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*
[27]*University of Liverpool, Liverpool L69 7ZE, United Kingdom*
[28]*University College London, London WC1E 6BT, United Kingdom*
[29]*Centro de Investigaciones Energeticas Medioambientales y Tecnologicas, E-28040 Madrid, Spain*
[30]*Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
[31]*Institute of Particle Physics: McGill University, Montréal, Québec,*
Canada H3A 2T8; Simon Fraser University, Burnaby, British Columbia,
Canada V5A 1S6; University of Toronto, Toronto, Ontario,
Canada M5S 1A7; and TRIUMF, Vancouver, British Columbia, Canada V6T 2A3
[32]*University of Michigan, Ann Arbor, Michigan 48109, USA*
[33]*Michigan State University, East Lansing, Michigan 48824, USA*
[34]*Institution for Theoretical and Experimental Physics, ITEP, Moscow 117259, Russia*
[35]*University of New Mexico, Albuquerque, New Mexico 87131, USA*
[36]*The Ohio State University, Columbus, Ohio 43210, USA*
[37]*Okayama University, Okayama 700-8530, Japan*
[38]*Osaka City University, Osaka 588, Japan*
[39]*University of Oxford, Oxford OX1 3RH, United Kingdom*

[40]Istituto Nazionale di Fisica Nucleare, Sezione di Padova-Trento, [ff]University of Padova, I-35131 Padova, Italy
[41]University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA
[42]Istituto Nazionale di Fisica Nucleare Pisa, [gg]University of Pisa,
[hh]University of Siena and [ii]Scuola Normale Superiore, I-56127 Pisa, Italy
[43]University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA
[44]Purdue University, West Lafayette, Indiana 47907, USA
[45]University of Rochester, Rochester, New York 14627, USA
[46]The Rockefeller University, New York, New York 10065, USA
[47]Istituto Nazionale di Fisica Nucleare, Sezione di Roma 1,
[jj]Sapienza Università di Roma, I-00185 Roma, Italy
[48]Rutgers University, Piscataway, New Jersey 08855, USA
[49]Texas A&M University, College Station, Texas 77843, USA
[50]Istituto Nazionale di Fisica Nucleare Trieste/Udine,
I-34100 Trieste, [kk]University of Udine, I-33100 Udine, Italy
[51]University of Tsukuba, Tsukuba, Ibaraki 305, Japan
[52]Tufts University, Medford, Massachusetts 02155, USA
[53]University of Virginia, Charlottesville, Virginia 22906, USA
[54]Waseda University, Tokyo 169, Japan
[55]Wayne State University, Detroit, Michigan 48201, USA
[56]University of Wisconsin, Madison, Wisconsin 53706, USA
[57]Yale University, New Haven, Connecticut 06520, USA

(Dated: August 1, 2012)

We present a search for the standard model Higgs boson produced in association with a $W^{\pm}$ boson. This search uses data corresponding to an integrated luminosity of 7.5 fb^{-1} collected by the CDF detector at the Tevatron. We select $WH \to \ell\nu b\bar{b}$ candidate events with two jets, large missing transverse energy, and exactly one charged lepton. We further require that at least one jet be identified to originate from a bottom quark. Discrimination between the signal and the large background is achieved through the use of a Bayesian artificial neural network. The number of tagged events and their distributions are consistent with the standard model expectations. We observe no evidence for a Higgs boson signal and set 95% C.L. upper limits on the WH production cross section times the branching ratio to decay to $b\bar{b}$ pairs, $\sigma(p\bar{p} \to W^{\pm}H) \times \mathcal{B}(H \to b\bar{b})$, relative to the rate predicted by the standard model. For the Higgs boson mass range of 100 GeV/c^2 to 150 GeV/c^2 we set observed (expected) upper limits from 1.34 (1.83) to 38.8 (23.4). For 115 GeV/c^2 the upper limit is 3.64 (2.78). The combination of the present search with an independent analysis that selects events with three jets yields more stringent limits ranging from 1.12 (1.79) to 34.4 (21.6) in the same mass range. For 115 and 125 GeV/c^2 the upper limits are 2.65 (2.60) and 4.36 (3.69), respectively.

PACS numbers: 13.85.Rm, 14.80.Bn

I. INTRODUCTION

The standard model (SM) describes not only the fundamental particles of quarks and leptons and their interactions, but also predicts the existence of a single scalar particle, the Higgs boson, which arises as a result of spontaneous electroweak symmetry breaking [1–4]. The Higgs boson remains the only fundamental SM particle that has not been observed by experiment. Direct searches at LEP2 [5], the Tevatron [7], and recently LHC experiments [8, 9] have constrained the Higgs boson mass to lie

*Deceased
†With visitors from [a]Istituto Nazionale di Fisica Nucleare, Sezione di Cagliari, 09042 Monserrato (Cagliari), Italy, [b]University of CA Irvine, Irvine, CA 92697, USA, [c]University of CA Santa Barbara, Santa Barbara, CA 93106, USA, [d]University of CA Santa Cruz, Santa Cruz, CA 95064, USA, [e]Institute of Physics, Academy of Sciences of the Czech Republic, Czech Republic, [f]CERN, CH-1211 Geneva, Switzerland, [g]Cornell University, Ithaca, NY 14853, USA, [h]University of Cyprus, Nicosia CY-1678, Cyprus, [i]Office of Science, U.S. Department of Energy, Washington, DC 20585, USA, [j]University College Dublin, Dublin 4, Ireland, [k]ETH, 8092 Zurich, Switzerland, [l]University of Fukui, Fukui City, Fukui Prefecture, Japan 910-0017, [m]Universidad Iberoamericana, Mexico D.F., Mexico, [n]University of Iowa, Iowa City, IA 52242, USA, [o]Kinki University, Higashi-Osaka City, Japan 577-8502, [p]Kansas State University, Manhattan, KS 66506, USA, [q]Ewha Womans University, Seoul, 120-750, Korea, [r]University of Manchester, Manchester M13 9PL, United Kingdom, [s]Queen Mary, University of London, London, E1 4NS, United Kingdom, [t]University of Melbourne, Victoria 3010, Australia, [u]Muons, Inc., Batavia, IL 60510, USA, [v]Nagasaki Institute of Applied Science, Nagasaki, Japan, [w]National Research Nuclear University, Moscow, Russia, [x]Northwestern University, Evanston, IL 60208, USA, [y]University of Notre Dame, Notre Dame, IN 46556, USA, [z]Universidad de Oviedo, E-33007 Oviedo, Spain, [aa]CNRS-IN2P3, Paris, F-75205 France, [bb]Texas Tech University, Lubbock, TX 79609, USA, [cc]Universidad Tecnica Federico Santa Maria, 110v Valparaiso, Chile, [dd]Yarmouk University, Irbid 211-63, Jordan,

in the range between 115.5 and 127 GeV/c^2 at 95% C.L., which is consistent with the 95% C.L. upper limit of 152 GeV/c^2 obtained from global fits to precision electroweak data [10].

In $\sqrt{s} = 1.96\,\text{TeV}$ proton-antiproton collisions, the Higgs boson is expected to be produced mainly through gluon fusion ($gg \to H$) and in association with a W or Z boson [11]. The cross section for WH production is twice that of ZH and is about a factor of 10 smaller than $gg \to H$. The Higgs boson decay branching fraction is dominated by $H \to b\bar{b}$ for the Higgs boson mass $m_H < 135$ GeV/c^2 and by $H \to W^+W^-$ for $m_H > 135$ GeV/c^2 [12]. A search for a low-mass Higgs ($m_H < 135$ GeV/c^2) in the $gg \to H \to b\bar{b}$ channel is extremely challenging because the $b\bar{b}$ QCD production rate is many orders of magnitude larger than the Higgs boson production rate. Requiring the leptonic decay of the associated W boson improves greatly the expected signal over background ratio in this channel. As a result, $WH \to l\nu b\bar{b}$ [6] is one of the most promising channels for the low-mass Higgs boson searches, and it significantly contributes to the combined search for the Higgs boson at the Tevatron [7].

This paper presents a search for Higgs boson production in proton-antiproton collisions using the $WH \to \ell\nu b\bar{b}$ channel at $\sqrt{s} = 1.96\,\text{TeV}$ using data collected between February 2002 and March 2011 with the CDF detector. The acquired data correspond to an integrated luminosity of approximately 7.5 fb^{-1}. Searches for the standard model Higgs boson using the same final state have been reported before by CDF [13, 14] and D0 [15] with data corresponding to an integrated luminosity of 5.6 fb^{-1} and 5.3 fb^{-1}, respectively. Compared to the previously reported analysis, we have employed a Bayesian artificial neural network (BNN) discriminant [16, 17] to improve discrimination between signal and background. The signal acceptance is improved by using additional triggers based on jets and missing transverse energy, as well as a novel method to combine them into a single analysis stream in order to maximize the event yield while properly accounting for correlations between triggers. The signal acceptance is also increased by using several different lepton reconstruction algorithms, for muon and electron candidates. We have optimized b-tagging algorithms used in the analysis to increase signal acceptance. We also employed multivariate methods to improve the rejection of multi-jet QCD background, as well as to improve di-jet invariant mass resolution.

Recently, the experiments at the Large Hadron Collider (LHC) have obtained enough data to set limits on the Higgs boson mass exceeding the sensitivity of the Tevatron experiments [8, 9]. However, at the LHC the most sensitive low-mass search is in the diphoton final state and searches for $H \to b\bar{b}$ will take more data before the Tevatron sensitivity is reached in this channel. In this sense, the Tevatron and LHC are complementary and both will provide important information in the search for a low-mass Higgs boson.

This paper is organized as follows. Section II describes the experimental apparatus, the Collider Detector at Fermilab (CDF). Section III presents the data samples and the event selection used to identify the $WH \to \ell\nu b\bar{b}$ candidate events. Section IV presents the background modeling and its estimation. Section V discusses the signal acceptance and its systematic uncertainty. Section VI introduces advanced techniques to improve the analysis sensitivity further. The final results and conclusions are presented in Sec. VII and Sec. VIII.

II. THE CDF II DETECTOR

The CDF II detector [18] geometry is described using a cylindrical coordinate system. The z-axis follows the proton direction, and the polar angle θ is usually expressed through the pseudorapidity $\eta = -\ln(\tan(\theta/2))$. The detector is approximately symmetric around $\eta = 0$ and in the azimuthal angle ϕ. The energy transverse to the beam is defined as $E_T = E\sin\theta$, and the momentum transverse to the beam is $p_T = p\sin\theta$.

Charged particles are tracked by a system of silicon microstrip detectors [19] and a large open cell drift chamber [20] in the region $|\eta| \leq 2.0$ and $|\eta| \leq 1.0$, respectively. The tracking detectors are immersed in a 1.4 T solenoidal magnetic field aligned with the incoming beams, allowing measurement of charged particle p_T.

The transverse momentum resolution is measured to be $\delta p_T/p_T \approx 0.07\% \cdot p_T(\text{GeV/c})$ for the combined tracking system [18]. The resolution on the track impact parameter (d_0), the distance from the beam-line axis to the track at the track's closest approach in the transverse plane, is $\sigma(d_0) \approx 40\,\mu$m, of which about $30\,\mu$m is due to the transverse size of the Tevatron beam itself [19].

Outside of the tracking systems and the solenoid, segmented calorimeters with projective tower geometry are used to reconstruct electromagnetic showers and hadronic jets [21–23] over the pseudorapidity range $|\eta| < 3.6$. The transverse energy is measured in each calorimeter tower where the polar angle (θ) is calculated using the measured z position of the event vertex and the tower location.

Contiguous groups of calorimeter towers with signals are identified and summed together into an energy cluster. Electron candidates are identified in the central electromagnetic calorimeter (CEM) or in the forward, known as the plug, electromagnetic calorimeter (PEM) as isolated, mostly electromagnetic, clusters that match a reconstructed silicon track in the pseudorapidity range $|\eta| < 1.1$ and $1.1 < |\eta| < 2.0$, respectively. The electron transverse energy is reconstructed from the electromagnetic cluster with a precision $\sigma(E_T)/E_T \approx 13.5\%/\sqrt{E_T(\text{GeV})} \oplus 2\%$ for central electrons [21] and $\sigma(E_T)/E_T = 16.0\%/\sqrt{E_T(\text{GeV})} \oplus 2\%$ for plug electrons [24]. Jets are identified as a group of electromagnetic calorimeter energy (E_{EM}) and hadronic calorimeter energy (E_{HAD}) clusters populating a cone of radius

$\Delta R \approx \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} \leq 0.4$ units around a high-E_T seed cluster [25]. Jet energies are corrected for calorimeter nonlinearity, losses in the gaps between towers, and multiple primary interactions. The jet energy resolution is approximately $\sigma(E_T) \approx [0.1E_T + 1.0 \text{ GeV}]$ [26].

Muon candidates are detected in three separate subdetectors. After at least five interaction lengths in the calorimeter, central muons first encounter four layers of planar drift chambers (CMU), capable of detecting muons with $p_T > 1.4\,\text{GeV}/c$ [27]. Four additional layers of planar drift chambers (CMP) behind another 60 cm of steel detect muons with $p_T > 2.8$ GeV/c [28]. These two systems cover the same central pseudorapidity region with $|\eta| \leq 0.6$. A track that is linked to both CMU and CMP stubs is called a CMUP muon. Muons that exit the calorimeters at $0.6 \leq |\eta| \leq 1.0$ are detected by the CMX system of four drift layers. Muon candidates are then identified as isolated tracks that extrapolate to line segments or "stubs" in the muon subdetectors.

Missing transverse energy ($\not\!\!E_T$) is defined as the opposite of the vector sum of all calorimeter tower energy depositions projected on the transverse plane. It is used as a measure of the sum of the transverse momenta of the particles that escape detection, most notably neutrinos. The corrected energies are used for jets in the vector sum defining $\not\!\!E_T$. The muon momentum is also added for any minimum ionizing high-p_T muon found in the event.

Muon and electron candidates used in this analysis are identified during data taking with the CDF trigger system, a three-level filter with tracking information available at the first level [29]. The first stage of the central electron trigger (CEM) requires a track with $p_T > 8$ GeV/c pointing to a tower with $E_T > 8$ GeV and $E_{\text{HAD}}/E_{\text{EM}} < 0.125$. As appropriate for selecting W-decay electrons, the plug electron trigger (MET+PEM) requires a tower with $E_T > 8$ GeV, $E_{\text{HAD}}/E_{\text{EM}} < 0.125$ and the missing transverse energy $\not\!\!E_T > 15$ GeV. The first stage of the muon trigger requires a track with $p_T > 4$ GeV/c (CMUP) or 8 GeV/c (CMX) pointing to a muon stub. A complete lepton reconstruction is performed online in the final trigger stage, where we require $E_T > 18\,\text{GeV}$ for central electrons (CEM), $E_T > 18\,\text{GeV}$ and $\not\!\!E_T > 20\,\text{GeV}$ for plug electron (MET+PEM) and $p_T > 18\,\text{GeV}/c$ for muons (CMUP, CMX).

The $\not\!\!E_T$ + 2 jet trigger has been previously used in the WH analysis [13], which complements the high-p_T lepton triggers by identifying a lepton from WH decay as a high-p_T track isolated from other tracks which has failed the lepton triggers mentioned above. At high instantaneous Tevatron luminosity, the accept rate of this trigger is reduced (pre-scaled) by randomly sampling a luminosity-dependent fraction of events. This trigger also requires two jets with $E_T > 10$ GeV, one of them central ($|\eta| < 1.1$), and $\not\!\!E_T > 35$ GeV. We also include a second $\not\!\!E_T$ and two-jets trigger, which was introduced only in the second part of the data and requires two jets with $E_T > 10$ GeV and $\not\!\!E_T > 30$ GeV. We also include a third trigger based on $\not\!\!E_T$ only, and $\not\!\!E_T > 45$ GeV for the first part of the data, while the selection criteria is relaxed to 40 GeV for the second part of the data.

The efficiency of the different triggers is measured using the lepton triggered data and is parametrized using sigmoid turn-on curves as a function of $\not\!\!E_T$, without correcting for the muon momenta. The novel method exploited to combine and parametrize all the three trigger paths is described in [30], which can be generalized to any combination of different trigger paths, allowing optimal performance.

III. DATA SAMPLES AND EVENT SELECTION

The data collected using the lepton-based (CEM, CMUP, CMX and MET+PEM) triggers correspond to 7.5 ± 0.4 fb^{-1} of integrated luminosity, while the data from the $\not\!\!E_T$-based triggers correspond to 7.3 ± 0.4 fb^{-1}.

The $WH \to \ell\nu b\bar{b}$ signal consists of two b jets, a high-p_T lepton, and large missing energy. This section provides an overview of the signal reconstruction with a focus on the improvements of this analysis over a previous WH search [13].

A. Improving Lepton Identification

We use several different lepton identification algorithms in order to include events from multiple trigger paths. Each algorithm requires a single high-p_T (> 20 GeV/c), isolated charged lepton consistent with leptonic W boson decay. Because the lepton from a leptonic W decay is well-isolated from the rest of the event, the additional energy in the cone of $\Delta R = 0.4$ surrounding the lepton is required to have less than 10% of the lepton energy. We employ the same lepton identification algorithms as the prior CDF WH search [13]. The tight lepton is required to be identified as either an electron (CEM, PEM), a muon (CMUP, CMX), or an isolated track from the data collected with $\not\!\!E_T$ triggers.

We further improve the lepton acceptance by about 10% by including two additional lepton identification algorithms. One lepton type is selected from CEM-triggered events using a multivariate likelihood method to select electron candidates that fail the standard electron requirements. Another lepton type is selected from $\not\!\!E_T$-triggered events by requiring an isolated track with significant deposits of energy in the calorimeter. Such tracks primarily originate from the leptonic decay of the W boson, where the electrons fail the standard identification, or from τ leptons that decay into single charged hadrons.

The efficiency of lepton identification is measured using $Z \to e^+e^-$ and $Z \to \mu^+\mu^-$ samples. A pure sample of leptons is obtained by selecting events where the invariant mass of two high-p_T tracks is near the mass of the Z boson and one track passed the trigger and tight lepton selection. The efficiency is then measured using

the other unbiased track. The same procedure is applied to simulated events events and a correction factor is applied to correct the difference due to imperfect detector modeling.

B. b-jet Identification

Multijet final states have dominant contributions from QCD light-flavor jet production. The low-mass standard model Higgs boson decays predominantly to b-quark pairs. Jets from b quarks can be distinguished from light-flavor jets by looking for the decay of long-lived B hadrons within the jet cone. We employ three b-identification algorithms to optimize the selection of b-quark jets. The secondary vertex tagging algorithm [31] (SECVTX) attempts to reconstruct a secondary vertex using tracks found within a jet. If a vertex is found and it is significantly displaced from the $p\bar{p}$ interaction point (primary vertex), the jet is identified as a b-jet("b-tagged"). The Jet Probability algorithm [32] (JP) uses tracking information from tracks inside a jet to identify B decays. The algorithm looks at the distribution of impact parameters for tracks inside a jet to form a probability that the jet originated from the primary vertex. Light jets yield a probability distribution approximately constant between 0 and 1, while b jets preferentially populate low values of probability. A jet is considered as b-tagged if the jet probability value is less than 5%. The neural network tagging algorithm [33] (NN) combines the strengths of existing b-tagging information more efficiently using a multivariate technique exploiting variables such as displaced vertices, displaced tracks, and low-p_T muons from b-quark decay. The NN provides an output value ranging from -1 (light-jet-like) to 1 (b-jet-like). The cut on this continuous output has been tuned to provide maximum sensitivity: a jet is considered as b-tagged if the jet's NN output is positive (> 0).

To increase the signal to background ratio for WH events, at least one jet must be b-tagged by the SECVTX algorithm. We then divide our sample into four exclusive categories in a preferential order based on the purity of b-tagged jets. The first category (ST+ST) comprises events where there are two SECVTX b-tagged jets. The second category (ST+JP) consists of events where only one of the jets is b-tagged by SECVTX and the second jet is b-tagged only by JP. The third category (ST+NN) is similar to the second, but the second jet is b-tagged only by NN. The fourth category (ST) contains events where only one of the jets is b-tagged by SECVTX and the second jet is not b-tagged.

C. Lepton + Jets Selection

After identifying the final state in the event, we require that the events contain one high-p_T lepton (> 20 GeV/c), corrected $\not\!\!E_T >$ 20 GeV (25 GeV in the case of forward electrons), and two jets with corrected $E_T > 20$ GeV and $|\eta| < 2.0$. The event's primary vertex is calculated by fitting a subset of well-measured tracks coming from the beam line and is required to be within 60 cm of the center of the CDF II detector [18]. The longitudinal coordinate z_0 of the lepton track at point of closest approach to the beam line must be within 5 cm of the primary vertex to ensure that the lepton and the jets come from the same hard interaction. In order to reduce the Z + jets and WW/WZ background rates, events with more than one lepton are rejected. Events from $Z \to l^+l^-$ decays in which one lepton is not identified are removed by vetoing events where the invariant mass of the lepton and any track in the event is within the Z mass window between 76 and 106 GeV/c^2.

Before applying any b-tagging algorithm, the sample (pretag sample) has dominant contributions from W + jets and QCD multijet production. We use the b-tagging strategies outlined above to increase the signal purity of the W + 2 jet events. We further purify the sample with exactly one secondary vertex tagged jet (ST) by applying additional kinematic and angular cuts to reduce QCD multijet events that mimic the W-boson signature. The rejection is based on a support vector machine multivariate discriminant that was optimized to identify the W + jets events against the QCD events [34].

IV. BACKGROUND ESTIMATION

The final state signature of $WH \to l\nu b\bar{b}$ production can be mimicked by a number of processes. The dominant backgrounds are W + jets production, $t\bar{t}$ production, single top production, and QCD multijet production. Several electroweak production processes (diboson or Z + jets) also contribute with smaller rates. We estimate the background rates based on the same strategies used in the previous top cross section measurement [31], single top searches [35], and WH analysis [13]. We provide an overview of each background estimate below.

A. Top and Electroweak Backgrounds

Production of both top-quark pairs and single top quarks contributes to the tagged W+jets sample. Several electroweak boson production processes also contribute. Pairs of WW can decay to a lepton, a neutrino (seen as missing energy), and two jets, one of which may originate from a charm quark. Pairs of WZ events can decay to the signal $\ell\nu b\bar{b}$ or $\ell\nu c\bar{c}$ final state. Finally, $Z \to \tau^+\tau^-$ events with one leptonic τ decay and one hadronic decay contribute, yielding a lepton, missing traverse energy, and a narrow jet displaced from the primary interaction point.

The normalizations of the diboson and top production backgrounds are based on the theoretical cross sections [36–38] listed in Table I, the time-integrated luminosity, and the acceptance and b-tagging efficiency de-

TABLE I: Theoretical cross sections and uncertainties for the electroweak and single top backgrounds, along with the theoretical cross section for $t\bar{t}$ at $m_t = 172.5\,\mathrm{GeV}/c^2$.

Background	Theoretical cross sections [pb]
WW	11.66 ± 0.70
WZ	3.46 ± 0.30
ZZ	1.51 ± 0.20
single-top s-channel	1.05 ± 0.07
single-top t-channel	2.10 ± 0.19
$t\bar{t}$	7.04 ± 0.44

rived from Monte Carlo events. The acceptance is corrected based on measurements using data for lepton identification, trigger efficiencies, b-tagging efficiencies, and the z vertex cut. The total top and electroweak contributions in each tagging category are shown in Table II. We use the measured inclusive cross section (787.4±85.0 pb) for Z + jets [39].

B. W + heavy flavor

The $Wb\bar{b}$, $Wc\bar{c}$, and Wc processes (W + heavy flavor) are major background sources after the b-tagging requirement. Large theoretical uncertainties exist for the overall normalization because current Monte Carlo event generators can generate W+heavy-flavor events only to tree-level. Consequently, the rates for these processes are normalized to data. The contribution from true heavy-flavor production in W+jets events is determined from measurements of the heavy-flavor event fraction in W+jets events and the b-tagging efficiency for those events.

The fraction of W+jets events produced with heavy-flavor jets has been studied extensively using a combination of ALPGEN + PYTHIA Monte Carlo generators [40–42]. Calculations of the heavy-flavor fraction in ALPGEN have been calibrated using a jet data sample, and a scaling factor of 1.4 ± 0.4 is necessary to make the heavy-flavor production in Monte Carlo match the production in W+1 jet events.

For the tagged W+heavy flavor (HF) background estimate, the heavy-flavor fractions and tagging rates are multiplied by the number of pretag W+jets candidate events (N_{pretag}) in data, after correction for the contribution of non-W ($f_{\mathrm{non-}W}$) as determined from the fits described in Section IV C , $t\bar{t}$, and other background events to the pretag sample. The W+heavy flavor background contribution is obtained by the following relation:

$$N_{W+\mathrm{HF}} = f_{\mathrm{HF}}\epsilon_{\mathrm{tag}}\left[N_{\mathrm{pretag}}(1 - f_{\mathrm{non-}W}) - N_{\mathrm{TOP}} - N_{\mathrm{EWK}}\right], \quad (1)$$

where f_{HF} is the heavy-flavor fraction, ϵ_{tag} is the tagging efficiency, N_{TOP} is the expected number of $t\bar{t}$ and single top events, and N_{EWK} is the expected background contribution from WW, WZ, ZZ and Z boson events, as described in Section IV A.

The total W + heavy flavor contributions in each tagging category are shown in Table II.

C. Non-W QCD Multijet

Events from QCD multijet production may mimic the W-boson signature due to instrumental background. When a jet passes the charged lepton selection criteria or a heavy-flavor jet produces a charged leptons via semileptonic decay, the jet is reconstructed incorrectly as a charged lepton, which is denoted as a non-W lepton. Non-W $\not\!\!E_T$ can result from mismeasurements of energy or semileptonic decays of heavy-flavor quarks. Since the $\not\!\!E_T$ mismeasurement is usually not well modeled in the detector simulation, we use several different samples of observed events to model the non-W multijet contribution. One sample is based on events that fired the central electron trigger but failed at least two of the five electron selection identification requirements that do not depend on the kinematic properties of the event, such as the fraction of energy in the hadronic calorimeter. This sample is used to estimate the non-W contribution from CEM, CMUP, and CMX events. A second sample is formed from events that pass a generic jet trigger with transverse energy $E_T > 20$ GeV to model PEM events. These jets are additionally required to have a fraction of energy deposited in the electromagnetic calorimeter between 80% and 95%, and fewer than four tracks, to mimic electrons. A third sample, used to model the non-W background in isolated track events, consists of events that are required to pass the $\not\!\!E_T$ triggers and contain a muon that passes all identification requirements but fails the isolation requirement.

To estimate the non-W fraction in both the pretag and tagged sample, the $\not\!\!E_T$ spectrum is fit to a sum of the predicted background shapes. The fit has one fixed component and two templates whose normalization can float. The fixed component is obtained by adding the contributions of the simulated processes based on theoretical cross sections. The two floating templates are a Monte Carlo W + jets template and a non-W template. The non-W template is different depending on the lepton category, as explained above. The total non-W contribution for each tagging category is also shown in Table II.

D. Mistagged Jets

Events with W + light-flavor jets containing no b or c quark with a fake b tag (mistags) can contribute to our tagged signal sample. We estimate the amount of mistags using the number of pretag W + light flavor events and the event mistag probability. The amount of pretag W + light flavor is determined from the pretag sample by subtracting the events from non-W, top and electroweak,

TABLE II: Background summary table for each b-tagging category after all lepton categories combined. As a reference, the expected signal for $m_H = 115$ GeV/c^2 is also shown.

	ST+ST	ST+JP	ST+NN	1-ST
Pretag events	184050			
$t\bar{t}$	142 ± 22	114 ± 12	62.8 ± 6.4	479 ± 49
Single top (s-ch)	45.0 ± 6.7	35.1 ± 3.4	18.9 ± 1.8	106 ± 10
Single top (t-ch)	13.9 ± 2.4	13.3 ± 2.0	8.7 ± 1.2	191 ± 23
WW	1.67 ± 0.42	6.23 ± 2.08	5.14 ± 1.35	186 ± 25
WZ	12.9 ± 2.0	10.7 ± 1.2	5.84 ± 0.62	53.3 ± 6.2
ZZ	0.62 ± 0.09	0.49 ± 0.06	0.29 ± 0.03	2.05 ± 0.23
Z + jets	9.64 ± 1.40	11.9 ± 1.7	8.75 ± 1.30	182 ± 25
$Wb\bar{b}$	257 ± 104	228 ± 91	125 ± 50	1450 ± 580
$Wc\bar{c}/c$	31.0 ± 12.6	98.3 ± 40.5	63.8 ± 26.0	1761 ± 708
Mistag	12.1 ± 2.9	52.8 ± 15.2	57.0 ± 14.3	1646 ± 220
Non-W QCD	57.9 ± 23.6	85.3 ± 34.1	74.9 ± 29.9	747 ± 299
Total background	584 ± 169	656 ± 194	432 ± 126	6802 ± 1822
Observed events	519	568	402	6482
WH and ZH signal (115 GeV/c^2)	7.28	5.34	2.80	16.0

and W + heavy flavor contributions. The event mistag probability is based on the per-jet mistag matrix that is derived from inclusive jet data by counting the number of false tags per jet for each b tagger and is parametrized as a function of jet E_T, η, number of vertices, track multiplicity, and the scalar sum of jet E_T in the event. For each event in our W + light flavor Monte Carlo samples, we apply the per-jet mistag matrix to each jet and combine the probability to get an event mistag probability. The total mistag contribution for each tagging category is also shown in Table II.

E. Summary of Background Estimation

The summary of the background and signal ($m_H = 115$ GeV/c^2) estimates and the number of observed events are shown in Table II for each tagging category. In this table, all lepton types are combined. In general, the numbers of expected and observed events are in good agreement.

V. SIGNAL ACCEPTANCE

In this section, the number of expected Higgs events and systematic uncertainties on the signal acceptance are discussed. We consider the signal acceptance for the $WH \to \ell\nu b\bar{b}$ process and the residual contribution of $ZH \to \ell\ell\, b\bar{b}$ where one of the leptons fails the Z removal cut. We generated $WH \to l\nu b\bar{b}$ and $ZH \to l^+l^- b\bar{b}$ samples using the Pythia Monte Carlo program [42] for 11 values of the SM Higgs mass sampled between 100 and 150 GeV/c^2. The number of expected $WH \to l\nu b\bar{b}$ events (N) is given by

$$N = \epsilon \mathcal{L} \sigma(p\bar{p} \to WH) \mathcal{B}(H \to b\bar{b}), \quad (2)$$

where ϵ, $\sigma(p\bar{p} \to WH)$, and $\mathcal{B}(H \to b\bar{b})$ are the event detection efficiency, production cross section, and branching ratio, respectively, and $\mathcal{L}$ is the integrated luminosity of the data-taking period. The production cross section and branching ratio are calculated to next-to-leading order (NLO) precision [11].

The total event detection efficiency is the product of several efficiencies: the trigger efficiency, the primary vertex reconstruction efficiency, the lepton identification efficiency, the b-tagging efficiency, and the event kinematic selection efficiency. The lepton trigger efficiency is measured using a clean $W \to l\nu$ data sample, obtained from other triggers after applying more stringent offline cuts. The $\not\!\!E_T$ trigger efficiency is obtained using a trigger combination method [30]. The primary vertex efficiency is obtained using the vertex distribution from the minimum bias data. The lepton identification efficiency is calculated using $Z \to l^+l^-$ data and Monte Carlo samples. The b-tag efficiency is measured in a b-enriched sample from semileptonic heavy flavor decay.

The expected number of signal events is estimated for each of the probed values of the Higgs boson mass. Table II shows the number of expected WH and ZH events for $M_H = 115$ GeV/c^2 in 7.5 fb^{-1}.

The total systematic uncertainty on the acceptance comes from several sources, including trigger efficiencies, the jet energy scale, initial and final state radiation, lepton identification, luminosity, and b-tagging efficiencies. The lepton trigger uncertainties are measured using Z boson decays. The acceptance uncertainty due to the jet energy scale (JES) [26] is calculated by shifting jet energies in WH Monte Carlo samples by ± one standard

deviation. The deviation from the nominal acceptance is taken as the systematic uncertainty. We estimate the impact of changes in initial state radiation (ISR) and final state radiation (FSR) by halving and doubling the parameters related to initial and final state radiation in the Monte Carlo event generation [43]. The difference from the nominal acceptance is taken as the systematic uncertainty. The uncertainty in the incoming partons' energies relies on the parton distribution function (PDF) fits. A NLO version of the PDFs, CTEQ6M, provides a 90% confidence interval for each of the eigenvector input parameters [44]. The nominal PDF value is reweighted to have a 90% confidence level value, and the corresponding reweighted acceptance is computed. The differences between the nominal and the reweighted acceptance are added in quadrature, and the total is assigned as the systematic uncertainty [31].

The lepton identification uncertainties are estimated based on studies comparing $Z \to l^+l^-$ events in data and Monte Carlo.

The systematic uncertainty of 6% in the CDF luminosity measurement is treated as fully correlated between the signal and all Monte Carlo based background samples.

The systematic uncertainty on the event tagging efficiency is estimated by varying the b-tagging efficiency and mistag prediction by $\pm$ one standard deviation and calculating the difference between the shifted acceptance and the default one.

Total systematic uncertainties are summarized in Tables III, IV, and V.

VI. ANALYSIS OPTIMIZATION

In this section we discuss the analysis optimization procedure after the event selection.

A. b-jet Energy Correction

The dijet invariant mass provides discrimination between signal and background and is a critical variable used in the multivariate analysis as described below. Improvement of the dijet mass resolution directly results in an improvement of the WH signal sensitivity. To improve dijet invariant mass resolution, we developed a neural network b-jet energy correction method. The neural network was trained using a sample of Monte Carlo simulated $WH \to l\nu b\bar{b}$ events. During training, jet observables were used as input values, and the energy of the corresponding b quark was used as the target value.

For each jet, we studied 40 variables related to the calorimeter energy, the charged tracks, and the displaced vertices within the jet cone of 0.4 and converged on nine well-modeled input variables most optimal for the jet-energy correction. The four calorimeter variables chosen are the jet E_T before and after the standard jet correction, jet p_T, and jet transverse mass. The tracking variables chosen are the sum p_T and the maximum p_T of the set of tracks within the jet cone. For the jet tagged by SECVTX we also include the vertex variables such as the secondary vertex transverse decay length, its uncertainty, and fitted secondary vertex p_T. Further details can be found in Ref. [45]. Without (with) applying NN corrections to b-jets in the Higgs decays, the dijet mass resolution is $\sim$15% ($\sim$11%) for double-tagged events, and $\sim$17% ($\sim$13%) for single-tagged events.

B. Bayesian Neural Network Discriminant

To improve the signal-to-background discrimination further, we employed a Bayesian neural network (BNN) trained on a variety of kinematic variables to distinguish WH events from the background [16, 17]. For this analysis, we employ distinct BNN discriminant functions that were optimized separately for the different tagging categories and each Higgs boson mass in order to maximize the sensitivity.

The BNN configuration has N input variables, 2N hidden nodes, and one output node. The input variables were selected by an iterative BNN optimization procedure from a large number of possible variables. The optimization procedure identified the most sensitive one-variable BNN, then looped over all remaining variables and found the most sensitive two-variable BNN. The process continued until adding a new variable no longer improved the sensitivity. The discriminant then is used to do hypothesis testing of a WH signal in the simulated data as a function of Higgs mass, which improves the background rejection with a sensitivity gain of 25% compared to the most sensitive variable alone.

The discriminant used for the ST+ST tag category is trained using N=7 input variables. The most sensitive variable is M_{jj}, the invariant mass calculated from the two tight jets after using the neural-network-based jet energy correction as described in Sec. VI A. The second input variable is the p_T imbalance, which is the difference between the scalar sum of the p_T of all measured objects and the $\not\!\!E_T$, $p_T(\text{jet1}) + p_T(\text{jet2}) + p_T(\text{lep}) - \not\!\!E_T$. The third variable, $M_{l\nu j}^{max}$, is the invariant mass of the lepton, $\not\!\!E_T$, and one of the two jets, where the jet is chosen to give the maximum invariant mass. The fourth variable is $Q_{lep} \times \eta_{lep}$, the signed product of the electric charge times the η of the charged lepton. The fifth variable is $\Sigma E_T(\text{loose jets})$, which is the scalar sum of loose jets transverse energy. A loose jet is defined as a jet having $|\eta| < 2.4$, $E_T > 12$ GeV, but failing the tight-jet requirement ($E_T > 20$ GeV and $|\eta| < 2.0$). The sixth variable is the p_T of the reconstructed W. The last variable is H_T, the scalar sum of the event transverse energies $H_T = \Sigma E_T$ (jets) + p_T(lepton) + $\not\!\!E_T$.

The discriminant used for both the ST+JP and ST+NN tag categories is trained with the same input variables as the ST+ST category, except that the variable $M_{l\nu j}^{max}$ is replaced with $M_{l\nu j}^{min}$ and the p_T imbal-

TABLE III: Systematic uncertainties on the acceptance for central leptons (in percent).

Category	JES	ISR/FSR/PDF	Lepton ID	Trigger	b-tag	Total
ST+ST	2.0	4.9	2	< 1	8.6	10.3
ST+JP	2.8	4.9	2	< 1	8.1	10.1
ST+NN	2.2	7.7	2	< 1	13.6	15.9
1-ST	2.3	3.0	2	< 1	4.3	6.1

TABLE IV: Systematic uncertainties on the acceptance for forward electrons (in percent).

Category	JES	ISR/FSR/PDF	Lepton ID	Trigger	b-tag	Total
ST+ST	2.4	7.7	2	< 1	8.6	12.0
ST+JP	3.9	4.5	2	< 1	8.1	10.3
ST+NN	6.7	12.9	2	< 1	13.6	20.0
1-ST	2.9	5.7	2	< 1	4.3	8.0

ance is replaced with the $\not\!\!E_T$. The discriminant used for the single ST tag category is trained with the same input variables as the ST+ST category with the exception that $M_{l\nu j}^{max}$ is replaced by $\not\!\!E_T$ and an extra variable is added. The new variable is the output of an artificial-neural-network-based heavy flavor separator trained to distinguish b-quark jets from the charm and light flavor jets after SECVTX tagging [35]. Distributions of all these variables are checked for both the pretag and tagged sample to ensure that they are described well by the background model.

The training is defined such that the neural network attempts to produce an output as close to 1.0 as possible for the Higgs boson signal events and as close to 0.0 as possible for background events. Figure 1 shows a shape comparison of the BNN output between signal and background events for the ST+ST, ST+JP, ST+NN, and ST sample, respectively.

VII. RESULTS

We perform a direct search for an excess in the signal region of the BNN output distribution from double-tagged and single-tagged W + 2 jets events. Figure 2 shows the BNN output distributions for each b-tagging category. The data and background predictions are in good agreement.

We use a binned likelihood fit [13, 46] to the observed BNN output distributions to test the presence of a WH signal. For optimal sensitivity, we perform a simultaneous search in each b-tag and lepton category. The total likelihood is the product of the single Poisson likelihoods used in each independent sample. The likelihood fit accommodates the uncertainties on our background estimate by letting the overall background prediction float within Gaussian constraints. The systematic uncertainties associated with the shape of the BNN output due to JES uncertainty are also included for both signal and background. We use a different set of background and signal BNN template shapes for each combination of lepton type and tag category. We correlate the systematic uncertainties appropriately across different lepton types and tag categories. We find no evidence for a Higgs boson signal in our sample. We use Bayesian limits with a positive flat prior and set 95% C.L. upper limits on the WH cross section times branching ratio, $\sigma(p\bar{p} \to W^{\pm}H) \cdot \mathcal{B}(H \to b\bar{b})$, relative to the rate predicted by the standard model.

We compare our observed limits to our expected sensitivity by generating statistical trials according to the background-only model and analyzing them as our data. The combined expected and observed limits for all the lepton types are shown in Figure 3 and Table VI. Limits are also determined for the combination of this analysis with the independent WH search using a matrix element technique for events with three jets [14]. The luminosity used in the three jet analysis is 5.6 fb^{-1}. The combination improves the expected WH sensitivity by about 5% over the W + two jet result alone. The observed limits in the two jet channel in the mass range above $m_H > 110$ GeV/c^2 are one standard deviation higher than expected. After combining with three jet bin, our limits become closer to the expectation.

This $WH(ZH) \to \ell\nu(\ell\ell)b\bar{b}$ analysis represents a substantial improvement in sensitivity over the prior analysis using a neural network [13]. The increase in sensitivity is 25% at $m_H = 115$ GeV/c^2 in addition to the improvement from a larger sample size, and is mainly due to the improvement of analysis techniques that include the BNN discriminant, the b-jet energy correction, the additional $\not\!\!E_T$ triggers, the loose leptons, and the optimized b-tagging strategies.

TABLE V: Systematic uncertainties on the acceptance for additional leptons (in percent).

Category	JES	ISR/FSR/PDF	Lepton ID	Trigger	b-tag	Total
ST+ST	1.7	7.1	4.5	3.0	8.6	12.5
ST+JP	2.4	6.4	4.5	3.0	8.1	11.9
ST+NN	1.9	19.5	4.5	3.0	13.6	24.5
1-ST	4.7	8.4	4.5	3.0	4.3	11.8

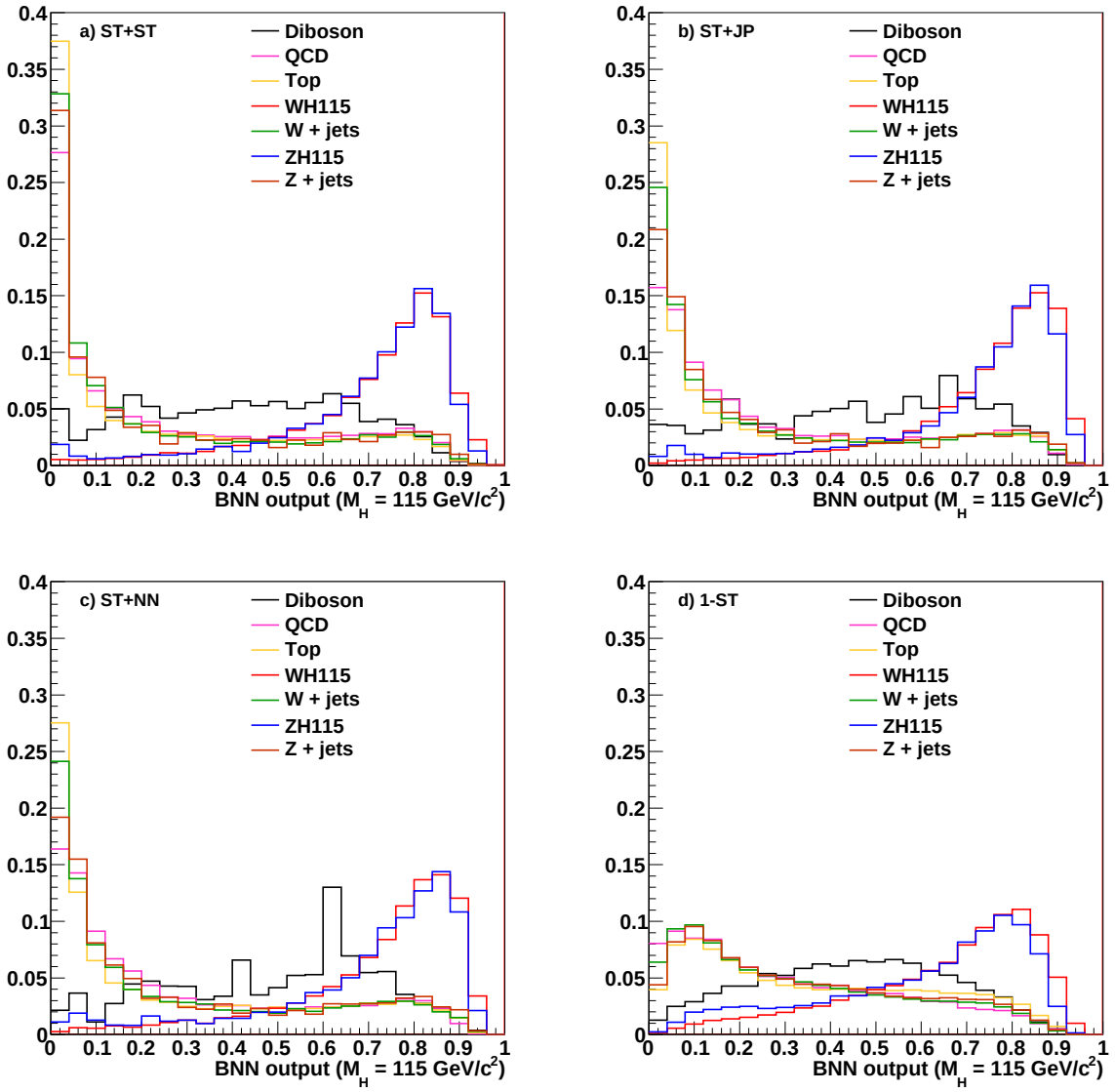

FIG. 1: Comparison of the BNN output for signal ($M_H = 115$ GeV/c^2) and background events with all lepton types included. From (a) - (d) the b-tag categories are ST+ST, ST+JP, ST+NN, and ST, respectively. Signal and background histograms are each normalized to unit area. The WH and ZH signals peak near the 1.0 value. The QCD multijet, top quark, W + jets and Z + jets peak near the 0.0 value. The diboson background has a broad peak in the middle region as its kinematics is very close to the signal ones. The diboson spike in figure (c) is a statistical fluctuation.

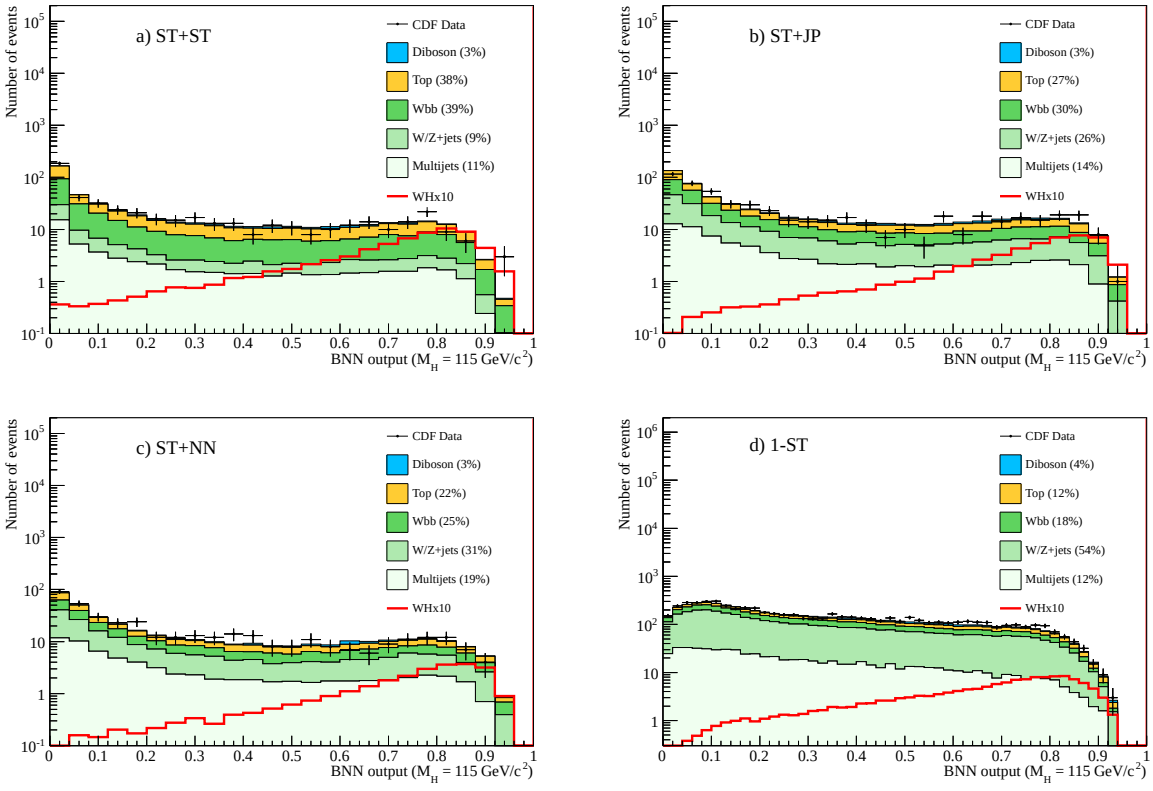

FIG. 2: The observed data and predicted BNN output for signal ($M_H = 115$ GeV/c^2) and background events with all leptons included. From (a) - (d) the b-tag categories are ST+ST, ST+JP, ST+NN, and ST, respectively.

TABLE VI: Observed and expected upper limits at 95% C.L. normalized to the SM expectation on $\sigma(p\bar{p} \to WH) \times \mathcal{B}(H \to b\bar{b})$ as a function of Higgs mass, including all lepton and tag categories, in the presented analysis and after combination with an independent search using a matrix element analysis for events with three jets.

Upper Limits/SM for Combined Lepton and Tag Categories				
	W + 2 jets		W + 2, 3 jets	
m_H (GeV/c^2)	Observed	Expected	Observed	Expected
100	1.34	1.83	1.12	1.79
105	2.10	2.08	2.06	1.98
110	3.42	2.26	2.78	2.17
115	3.64	2.78	2.65	2.60
120	4.68	3.22	3.40	3.06
125	5.84	4.01	4.36	3.69
130	8.65	5.13	6.09	4.80
135	10.2	7.02	7.71	6.40
140	16.4	9.39	12.3	8.84
145	24.7	15.3	18.9	14.2
150	38.8	23.4	34.4	21.6

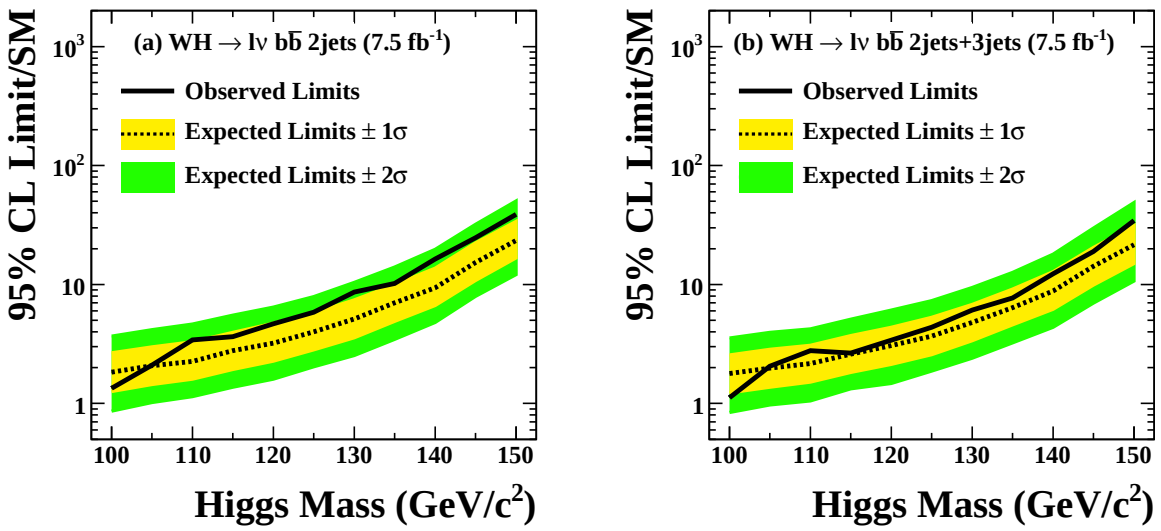

FIG. 3: Observed and expected upper limits at 95% C.L. on Higgs boson production times branching ratio with respect to the SM expectation for all lepton and tag categories combined as a function of the Higgs boson mass for the present analysis (a) and after combination with the independent three jet analysis with a matrix element (b).

VIII. CONCLUSIONS

We have presented the results of a CDF search for the standard model Higgs boson decaying to $b\bar{b}$ final states, produced in association with a W boson decaying into a charged lepton and neutrino. We find that for the dataset corresponding to an integrated luminosity of 7.5 fb^{-1}, the data agree with the SM background predictions. We therefore set upper limits on the Higgs boson production cross section times the $H \to b\bar{b}$ branching ratio with respect to the standard model prediction. For the mass range of 100 GeV/c^2 through 150 GeV/c^2 we set observed (expected) upper limits at 95% C.L. from 1.34 (1.83) to 38.8 (23.4). For 115 GeV/c^2 the upper limit is 3.64 (2.78). When we combine this search with an independent search using events with three jets [14], we set more stringent limits in the same mass range from 1.12 (1.79) to 34.4 (21.6). For 115 and 125 GeV/c^2 the upper limits are 2.65 (2.60) and 4.36 (3.69), respectively. Improved analysis techniques have resulted in an increase in sensitivity over the previous 2.7 fb^{-1} analysis [13] by 25% more than the expectation from simple luminosity scaling.

The search results in this channel at the CDF experiment are the most sensitive low-mass Higgs boson search at the Tevatron. While the LHC experiments will continue to improve their sensitivity to the low-mass Higgs boson, which is obtained primarily from searches in the diphoton final state, we expect that the searches in the $H \to b\bar{b}$ channel at the Tevatron will provide a crucial test on the existence and nature of the low-mass Higgs boson.

Acknowledgments

We thank the Fermilab staff and the technical staffs of the participating institutions for their vital contributions. This work was supported by the U.S. Department of Energy and National Science Foundation; the Italian Istituto 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 World Class University Program, the National Research Foundation of Korea; the Science and Technology Facilities Council and the Royal Society, UK; the Russian Foundation for Basic Research; the Ministerio de Ciencia e Innovación, and Programa Consolider-Ingenio 2010, Spain; the Slovak R&D Agency; the Academy of Finland; and the Australian Research Council (ARC).

[1] P. W. Higgs, Phys. Lett. **12**, 132 (1964).
[2] P. W. Higgs, Phys. Rev. Lett. **13**, 508 (1964).
[3] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, Phys. Rev. Lett. **13**, 585 (1964).
[4] F. Englert and R. Brout, Phys. Rev. Lett. **13**, 321 (1964).
[5] R. Barate *et al.* (LEP Working Group for Higgs boson

searches), Phys. Lett. B **565**, 61 (2003).
[6] In this paper, lepton (ℓ) denotes electron ($e^{\pm}$), muon ($\mu^{\pm}$), or small contribution from one-prong tau ($\tau^{\pm}$) decay, and ν denotes each corresponding neutrinos.
[7] The Tevatron New Phenomena and Higgs Working Group (2011), arXiv:1107.5518.
[8] G. Aad *et al.* (ATLAS Collaboration), Phys. Lett. B **710**, 49 (2012).
[9] S. Chatrchyan *et al.* (CMS Collaboration), Phys. Lett. B **710**, 26 (2012).
[10] The LEP Electroweak Working Group, http://lepewwg.web.cern.ch/LEPEWWG/,(arXiv:0911.2604).
[11] Tev4LHC Higgs Working Group, *Standard Model Higgs cross sections at hadron colliders*, http://maltoni.home.cern.ch/maltoni/TeV4LHC/SM.html.
[12] A. Djouadi, J. Kalinowski, and M. Spira, Comput. Phys. Commun. **108**, 56 (1998).
[13] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. D **85**, 052002 (2012).
[14] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. D **85**, 072001 (2012).
[15] V. M. Abazov *et al.* (D0 Collaboration), Phys. Lett. B **698**, 6 (2011).
[16] R. M. Neal, http://www.cs.toronto.edu/ř̃adford/fbm.software.html.
[17] R. M. Neal, *Bayesian Learning for Neural Networks* (Springer, 1996).
[18] D. Acosta *et al.* (CDF Collaboration), Phys. Rev. D **71**, 032001 (2005).
[19] A. Sill (CDF Collaboration), Nucl. Instrum. Methods A **447**, 1 (2000).
[20] T. Affolder *et al.*, Nucl. Instrum. Methods A **526**, 249 (2004).
[21] L. Balka *et al.*, Nucl. Instrum. Methods A **267**, 272 (1988).
[22] S. Bertolucci *et al.*, Nucl. Instrum. Methods A **267**, 301 (1988).
[23] M. G. Albrow *et al.*, Nucl. Instrum. Methods A **480**, 524 (2002).
[24] L. Breccia *et al.*, Nucl. Instrum. Methods A **532**, 575 (2004).
[25] F. Abe *et al.* (CDF Collaboration), Phys. Rev. D **45**, 1448 (1992).
[26] A. Bhatti *et al.*, Nucl. Instrum. Methods A **566**, 375 (2006).
[27] G. Ascoli *et al.*, Nucl. Instrum. Methods A **268**, 33 (1988).
[28] T. Dorigo (CDF Collaboration), Nucl. Instrum. Methods A **461**, 560 (2001).
[29] E. J. Thomson *et al.*, IEEE Trans. Nucl. Sci. **49**, 1063 (2002).
[30] A. Buzatu, A. Warburton, N. Krumnack, and W.-M. Yao, submitted to Nucl. Instrum. Methods A, arXiv:1206.4813.
[31] D. E. Acosta *et al.* (CDF Collaboration), Phys. Rev. D **71**, 052003 (2005).
[32] A. Abulencia *et al.* (CDF Collaboration), Phys. Rev. D **74**, 072006 (2006).
[33] J. Keung , Ph.D. Thesis, University of Pennsylvania, FERMILAB-THESIS-2010-32, (2010).
[34] F. Sforza *et al.* J. Phys. Conf. Ser., **331**, 032045 (2011).
[35] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. D **82**, 112005 (2010).
[36] J. Campbell, R. K. Ellis, Phys. Rev. D **65**, 113007 (2002).
[37] M. Cacciari, S. Frixione, M. L. Mangano, P. Nason, and G. Ridolfi, J. High Energy Phys. 04 (2004) 068.
[38] B. W. Harris, E. Laenen, L. Phaf, Z. Sullivan, and S. Weinzierl, Phys. Rev. D **66**, 054024 (2002).
[39] D. Acosta *et al.* (CDF Collaboration), Phys. Rev. Lett. **94**, 091803 (2005).
[40] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, J. High Energy Phys. 07 (2003) 001.
[41] G. Corcella *et al.*, arXiv:hep-ph/0201201.
[42] T. Sjostrand *et al.*, Comput. Phys. Commun. **135**, 238 (2001).
[43] A. Abulencia *et al.* (CDF Collaboration), Phys. Rev. D **73**, 032003 (2006).
[44] J. Pumplin *et al.*, J. High Energy Phys. 07 (2002) 012.
[45] T. Aaltonen, A. Buzatu, B. Kilminster, Y. Nagai, and W.-M. Yao, submitted to Nucl. Instrum. Methods A, arXiv:1107.3026.
[46] T. Junk, Nucl. Instrum. Methods A **434**, 435 (1999).