Improved Search for a Higgs Boson Produced in Association with $Z \rightarrow \ell^+\ell^-$ in $pp$ Collisions at $\sqrt{s} = 1.96$ TeV

T. Aaltonen, B. Álvarez González, S. Amerio, D. Amidei, A. Anastassov, A. Anno, J. Antos, G. Apollinari, J.A. Appel, A. Apresyan, T. Arisawa, A. Artikov, J. Asaadli, W. Ashmanekas, B. Auerbach, A. Aurisano, F. Azzar, W. Badgett, A. Barbaro-Galtieri, V.E. Barnes, B.A. Barnett, P. Barria, P. Bartos, M. Bause, G. Bauer, F. Bedeschi, D. Beecher, S. Behari, G. Bellettini, J. Bellinger, D. Benjammin, A. Beretvas, A. Bhatti, G. Binkley, D. Bisello, I. Bizjak, K.R. Bland, C. Blocker, B. Blumenfeld, A. Bocci, A. Bodek, D. Bortoletto, J. Boudreau, A. Boevel, B. Braun, L. Brigliadori, A. Brisuda, C. Bromberg, E. Brucken, M. Bussco, J. Budagov, H.S. Budh, S. Budd, K. Burket, G. Busetto, P. Bussey, A. Buzatu, S. Cabrera, C. Calancha, S. Camarda, M. Campanelli, C. Canner, F. Canavero, L. Canepa, B. Carlso, D. Carlsmith, R. Carosi, S. Carrillo, S. Carron, B. Casal, M. Casars, A. Castro, P. Catastini, D. Cauz, V. Cavaliere, M. Cavallì-Sforza, A. Cerri, L. Cerrito, Y.C. Chen, M. Cherot, G. Chiarelli, G. Chlachidze, F. Chlebana, K. Cho, D. Chokheli, J.P. Chou, W.H. Chung, Y.S. Chung, C.I. Ciobanu, M.A. Ciocci, A. Clark, D. Clark, G. Compostella, M.E. Convery, J. Conway, M. Corbo, M. Cordelli, C.A. Cox, D.J. Cox, F. Crescioli, C. Cuenca Almenar, J. Cues, R. Culbertson, D. Dagenhart, N. D'Ascanzo, M. Datta, P. de Barbaro, G. De Cecco, G. De Lorenzo, M. Dell'Orso, C. Deluca, L. Demortier, J. Deng, M. Denino, F. Devoto, M. D'Errico, A. Di Canto, B. Di Ruzza, J.R. Dittmann, M. D'Onofrio, S. Donati, P. Dong, T. Dorigo, K. Ebina, A. Elagin, A. Eppig, R. Erbacher, D. Errede, S. Errede, E. Eshardt, R. Eusebi, H.C. Fang, S. Farrington, M. Feindt, J.P. Fernandez, C. Ferrazza, R. Field, G. Flanagan, R. Forrest, M.J. Frank, M. Franklin, J.C. Freeman, I. Furic, M. Galarino, J. Galyndt, J.E. Garcia, A.F. Garfinkl, P. Garosi, H. Gerberich, E. Gerchtein, S. Giagni, V. Giakoumpoulou, P. Giannetti, K. Gibson, C.M. Ginsburg, N. Giokaris, P. Giornini, M. Giunta, G. Giurgiu, V. Glagolev, D. Glenzinski, M. Gold, D. Goldin, N. Goldschmidt, A. Golossan, G. Gomez, G. Gomez-Ceballos, M. Goncharov, O. González, I. Gorelov, A.T. Goshaw, K. Gouliamos, A. Gresele, S. Grinstein, C. Gross-Pichler, R.C. Group, J. Guimarães da Costa, Z. Gunay-Unalan, P. Haber, S.R. Hahn, E. Halkiadakis, A. Hamaguchi, J.Y. Han, F. Happacher, K. Hara, D. Hare, M. Hare, R.F. Harr, K. Hatakeyama, M. Hays, M. Heck, J. Heinrich, M. Herndon, S. Hewamane, D. Hidas, A. Hocker, W. Hopkins, D. Horn, S. Hou, R.E. Hughes, M. Hurwitz, U. Husemann, N. Hussain, M. Huston, G. Intorzi, M. Iori, A. Ivanov, E. James, D. Kang, B. Jayatilaka, E.J. Jeon, M.K. Jha, S. Jindariani, W. Johnson, M. Jones, K.K. Joo, S.Y. Jum, T.R. Junk, T. Kamon, P.E. Kardim, Y. Kato, W. Ketchum, J. Keung, V. Khotilovich, B. Kilminster, D.H. Kim, H.S. Kim, H.W. Kim, J.E. Kim, M.J. Kim, S.B. Kim, S.H. Kim, Y.K. Kim, N. Kimura, S. Klimenko, K. Kondo, D.J. Kong, J. Konig, A. Korytov, A.V. Kotwal, M. Kreps, J. Kroll, D. Krop, N. Krummack, M. Kruse, V. Krutelyov, T. Kuhr, K. Kurata, S. Kwang, A.T. Laasanen, S. Lami, S. Lammel, L. Lancaster, R.L. Lander, K. Lannon, A. Lath, G. Latino, I. Lazzizzera, T. LeCompte, E. Lee, H.S. Lee, J.S. Lee, S.W. Lee, S. Leo, S. Leone, J.D. Lewis, C.-J. Lin, S. Linacre, M. Lindgren, E. Lipiello, A. Lister, D.O. Litvintsev, C. Liu, Q. Liu, T. Liu, S. Lockhart, N.S. Lockyer, A. Logino, D. Lucchesi, J. Lueck, P. Lujan, P. Lukens, G. Lungu, J. Lys, R. Lysak, R. Madrak, K. Maeshima, K. Makhoul, P. Maksimonov, S. Malik, G. Manca, A. Manousakis-Katsikakis, F. Margaroli, C. Marino, M. Martínez, R. Martínez-Ballarin, M. Mathis, M.E. Mattson, P. Mazzanti, K.S. McFarland, P. McIntyre, R. McNulty, A. Mehta, P. Mehtala, A. Menzione, C. Mesropian, T. Miao, D. Mietlicki, A. Mitra, H. Miyake, S. Moed, N. Moggi, M.N. Mondragon, C.S. Moon, R. Moore, M.R. Morello, J. Morlock, P. Movilla Fernandez, A. Mukherjee, Th. Muller, P. Murat, M. Mussini, J. Nachtman, Y. Nagai, J. Naganoma, I. Nakano, A. Napier, J. Nett, C. Neu, M.S. Neubauer, J. Nielsen, L. Nodulman, O. Norniella, E. Nurse, L. Oakes, S.H. Oh, J.D. Oh, I. Okusuzian, T. Okusawa, R. Orava, L. Ortolan, S. Pagan Griso, C. Pagliarone, E. Palenciac, V. Papadimitriou, A.A. Paramonov, J. Patrick, G. Pauletta, M. Paulini, C. Paus, D.E. Pellet, C. Paus, D.E. Pellet.
We present a search for the standard model Higgs boson produced with a $Z$ boson in 4.1 fb$^{-1}$ of data collected with the CDF II detector at the Tevatron. In events consistent with the decay of the Higgs boson to a bottom-quark pair and the $Z$ boson to electrons or muons, we set 95% credibility level upper limits on the $ZH$ production cross section times the branching ratio. Improved analysis methods enhance signal sensitivity by 20\% relative to previous searches beyond the gain due to the larger data sample. At a Higgs boson mass of 115 GeV we set a limit of 5 \times 10^{-9}$ times the standard model value.

PACS numbers: 14.80.Bn,13.85.Rm

Establishing the mechanism of electroweak symmetry breaking (EWSB) is one of the outstanding issues of particle physics. In the standard model (SM), EWSB is
mediated by a Higgs field that manifests a particle, the as-yet-unobserved Higgs boson. A SM Higgs boson with mass \( M_H \) below 114.4 GeV/c\(^2\) or with \( M_H \) between 162 and 166 GeV/c\(^2\) has been excluded at 95\% confidence level in direct searches at LEP [1] and the Tevatron [2]. A recently reported [3] preliminary update of the Tevatron result has expanded this 95\% exclusion range to values of \( M_H \) between 158 and 175 GeV/c\(^2\).

At the Tevatron, production of the Higgs boson is dominated by the direct production process \( gg \rightarrow H \), and for \( M_H < 135 \) GeV/c\(^2\) the Higgs boson decays primarily to a pair of \( b \) quarks, \( H \rightarrow b\bar{b} \) [4]. The process \( gg \rightarrow H \rightarrow b\bar{b} \) is overwhelmed by multi-jet background. Associated production of a Higgs boson with a leptonically decaying \( W \) or \( Z \) boson yields a distinct signature for efficient selection and in turn greater sensitivity despite a significantly smaller cross section than direct production [5] [6]. This Letter presents an improved search for the SM Higgs boson produced in association with a \( Z \) or \( W \) boson, \( ZH \) or \( WH \) (for the SM Higgs boson produced in association with a \( \ell^+\ell^- \) or at least one central track of magnetic energy (EM) cluster of \( E \)). The remaining fraction of events is either an electron or a forward EM cluster of \( E \).

Inclusion of the low-\( S/B \) subsample increases the total signal yield in the PreTag sample by 16\%.

The PreTag sample consists mainly of \( Z \)+light flavor (l.f.) jet \((u,d,s,g)\) events, with smaller contributions from \( Z \)+heavy flavor (h.f.) jet \((c,b)\), \( \ell \ell \), and diboson processes. To reduce the \( Z \)+l.f. background, we look for \( b \) jets in the event. We use two algorithms to identify (tag) \( b \) jets: one based on evidence for a decay displaced spatially from the \( pp \) interaction point (SV) [7] and one based on track impact parameters with respect to the \( pp \) interaction point (JP) [19]. For the SV algorithm, there are two operating points: tight and loose [20]. The tight operating point has better l.f.-jet rejection (smaller mistag probability) at the expense of reduced \( b \)-jet identification efficiency.

We select events in the PreTag sample using the \( b \) tagging algorithms on the jet pairs forming Higgs candidates. We require the jet pairs to satisfy one of the following classifications, in order of precedence from highest to lowest ratio of \( ZH \) signal to background: a pair...
containing two SV-tight-tagged jets is classified as tight-double-tagged (TDT); a pair consisting of one SV-loose-tagged jet and a second JP-tagged jet is classified as loose-double-tagged (LDT); and a pair where only one jet has a SV-tight-tag is classified as single-tagged (ST). When there are multiple $H \rightarrow b\bar{b}$ candidates in an event, the jet pair with the highest $b$ tag class is selected (TDT > LDT > ST). This is the first time the JP algorithm has been used in the $ZH \rightarrow ℓ⁺ℓ⁻ b\bar{b}$ channel. While the $b$ tag selection matches the signal selection efficiency (approximately 60%) and background rejection rate (96%) of previous efforts, the addition of the new LDT class increases sensitivity to a $ZH$ signal by 6%. Due to differing background composition, the events are divided into independent subsamples based on $b$-tagging class. With two $Z$ boson $S/B$ categories and three $b$-tagging classes, we form a total of six subsamples that we analyze for $ZH$ content.

We compare the $b$-tag data to a model of signal and backgrounds to estimate the signal content. Signal, $E_T$, and diboson events are modeled with the PYTHIA [21] event generator. Backgrounds from $Z$+h.f. processes are simulated at the quark level using ALPGEN [22], then passed to PYTHIA for hadronization. The detector response is modeled with a detailed detector simulation [7]. We estimate the background from $Z$+l.f. mistags using re-weighted PreTag data with weights reflecting the probability for a l.f. jet to be erroneously $b$ tagged. A small fraction, much less than 1%, of jets can be erroneously identified as electrons, resulting in a background of misidentified (misID) $Z \rightarrow ee$ candidates. The likelihood of jet-electron misidentification is measured in generic jet data and applied to all electron-jet and jet-jet pairings in the electron triggered samples to generate a model of misID $Z$ events. The misID $Z \rightarrow μμ$ background contribution is estimated from the events passing all selection requirements but containing like-charge muon pairs. Predicted and observed event totals are listed in Table I for the $b$-tagged subsamples.

To improve the separation of $ZH$ from background, we utilize several multivariate techniques that use kinematic quantities as inputs. The dijet mass $M_{jj}$ is one of the most useful quantities, with its separating power mainly by the jet-energy resolution. In $ZH$ signal events with $Z \rightarrow ℓ⁺ℓ⁻$, incorrect measurement of jet energies results in apparent missing transverse energy $E_T$ [22]. We correct jet energies, based upon the $E_T$, and thereby improve the resolution on the dijet invariant mass. Jet-energy correction factors are computed by a NN trained to match measured jet energies to parton-level energies in $Z$+jets and signal events. In this analysis, this neural network is improved compared to the previous analysis [10] by utilizing additional input variables describing the recoil of the $Z$ boson. The corrected jet energies are used to recompute the Higgs candidate mass $M_H$, the $p_T$ of the jets, the $p_T$ of the Higgs candidate, the projection of $E_T$ onto the lower-$E_T$ Higgs jet, and the sphericity. The effect of the NN corrections on the reconstructed $H \rightarrow b\bar{b}$ mass is shown in Fig. 1. In signal the resolution [24] on $M_H$ is improved from about 18% to 12%.

To exploit the combined signal-to-background discrimination power of event quantities and their correlations, we employ a neural network discriminant (NN$_D$) trained to simultaneously separate $ZH$, $t\bar{t}$, and $Z$+jets events. The NN$_D$ is configured to return values of $(x, y) = (1, 0)$ for $ZH$ events, $(0, 0)$ for $Z$+jets, and $(1, 1)$ for $t\bar{t}$. A separate NN$_D$ is formed for each of the three $b$-tag classes. In addition to the quantities recomputed with corrected jet energies, the NN$_D$ inputs include: $E_T$; MEPs for $ZH$, $t\bar{t}$, and $Z$+jets processes; the number of jets in the event; and the output of a $b$ jet identifying artificial neural network (NN$_b$) [12]. The $ZH$ process MEP is computed [11] by convolving the theoretical matrix element for $ZH$ production with detector resolution functions. The resulting MEP reflects the likelihood that an event is signal. We calculate two additional MEPs under the hypotheses that an event is $Z$+jets or $t\bar{t}$ to aid background rejection. The NN$_b$ augments the performance of the SV algorithm by isolating incorrectly $b$-tagged l.f. jets. The addition of NN$_b$ as an input enhances the ability of the NN$_D$ to distinguish $ZH$ from $Z$+l.f., which constitutes 40% of the total background in the ST class. One-dimensional projections of NN$_D$ output are shown in Fig. 2 for each $b$-tag class.

The effect of systematic uncertainties on the determination of signal content requires propagating uncertainties on NN$_D$ input quantities to the output distributions. The dominant effect on the result comes from uncertainties on cross sections for background processes — a 40% uncertainty is assumed on the normalization of $Z$+h.f. samples [25], 11.5% for the diboson samples [26], 20% to account for the theoretical and experimental uncertainty on the $t\bar{t}$ cross section [27], and 5% for $ZH$ signal [28]. Uncertainties of 4% (ST), 8% (TDT), and 11%
TABLE I: Comparison of the expected mean event totals for background and ZH signal with the observed number of data events. The totals are for full event selection, and systematic and statistical uncertainties are combined in quadrature. Systematic uncertainties dominate. The background composition is provided in the first five rows. Each of the six ZH subsamples is presented in a separate column.

| Process          | High S/B | Low S/B |
|------------------|----------|---------|
|                  | TDT      | LDT     | ST      | TDT      | LDT     | ST      |
| tt               | 7.0±1.5  | 8±2     | 17±4    | 2.9±0.6  | 3.2±0.8 | 8.9±1.9 |
| Diboson          | 2.9±0.4  | 4±1     | 16±2    | 0.5±0.1  | 0.6±0.1 | 3.3±0.5 |
| Z + h.f.         | 18±7     | 30±13   | 159±67  | 3.5±1.5  | 5.6±2.4 | 32±14   |
| Z + l.f.         | 0.9±0.3  | 9±3     | 152±23  | 0.4±0.1  | 3.8±1.3 | 50±7.6  |
| misID Z          | 0.7±0.3  | 2±1     | 22±11   | 1.4±0.7  | 1.1±0.5 | 23±12   |
| Total Bkg.       | 29±8     | 53±14   | 366±72  | 9±2      | 14±3    | 117±20  |

ZH(115 GeV/c²) 0.7±0.1 0.7±0.1 1.7±0.2 0.11±0.01 0.11±0.03 0.28±0.05

Data 23 56 406 12 14 116

FIG. 2: Projections of the two-dimensional neural network (NN_D) output onto the x-axis (x and y are defined in the text) for events in the b-tag categories ST, LDT, and TDT. Events with a NN_D score of y ≥ 0.1 are omitted to highlight the signal region. The ZH contribution is shown, multiplied by a factor of 25, for MH = 115 GeV/c².

(LDT) on the normalization of b-tagged samples are applied to account for different b-tag efficiencies in data and simulation. Other uncertainties affecting sample normalizations include: 6% on the integrated luminosity, 1% on the trigger and lepton reconstruction efficiencies [29], 1.5% on the measurement of lepton energies, and a 50% uncertainty on the total lepton misID estimate. We include uncertainties on jet energies, the modeling of initial and final state radiation, and mistag event weighting as variations on the shape and normalization of the NN_D output.

We calculate a limit on the associated production of a SM Higgs boson and Z boson based on the comparison of the full NN_D output of the b-tagged data to expectations for signal and background. We consider eleven Higgs mass hypotheses between 100 and 150 GeV/c². We use a Bayesian algorithm [30] with a flat prior in the production cross section, integrating over the priors for the systematic uncertainties, incorporating correlated rate and shape uncertainties, and uncorrelated bin-by-bin statistical uncertainties [31]. Systematic uncertainties reduce the sensitivity of this search by 16%. The expected 95% credibility level (C.L.) limits are calculated assuming no signal, based on expected backgrounds only, as the median of 1000 simulated experiments. The ±1σ and ±2σ expected limits are derived from the distribution of the simulation limits at the 16th, 84th, 2nd, and 98th percentiles of the distribution, respectively. The observed 95% C.L. on the ZH cross section are calculated from the b-tagged data, and displayed in Fig. 3 and summarized in Table 11.

In conclusion, we have searched for the SM Higgs boson produced in association with a Z boson, where Z → ℓ⁺ ℓ⁻ and H → bb, finding no significant evidence for the process. We set 95% C.L. upper limits on the ZH produc-
TABLE II: Expected (Exp.) and observed (Obs.) 95% C.L. upper limits on the $ZH$ production cross section times the branching ratio for $H \to b \bar{b}$ normalized to the SM value for Higgs bosons between 100 and 150 GeV/c$^2$. The assumed $ZH$ cross section and branching fraction for $H \to b \bar{b}$ are 0.11 pb [6] and 0.73 [4] for a 115 GeV/c$^2$ Higgs boson.

| $M_H$ (GeV) | 100 | 105 | 110 | 115 | 120 | 125 | 130 | 135 | 140 | 145 | 150 |
|------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Exp.       | 6.7 | 6.4 | 6.3 | 6.8 | 8.5 | 10. | 13  | 19  | 29  | 45  | 74  |
| Obs.       | 4.5 | 4.6 | 5.3 | 5.9 | 7.9 | 8.1 | 10. | 14  | 19  | 24  | 43  |

...tion cross section times the $H \to b \bar{b}$ branching ratio for Higgs boson masses between 100 and 150 GeV/c$^2$. For a Higgs boson mass of 115 GeV/c$^2$ we set (expect) a 95% C.L. upper limit of 5.9 (6.8) times the standard model prediction. This result is an important step forward in the search for the Higgs boson and the source of EWSB, improving upon the previous observed (expected) limits in this channel by factors of 2.2 to 3.7 (1.9 to 2.4).

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 World Class University Program, the National Research Foundation of Korea; the Science and Technology Facilities Council and the Royal Society, UK; the Institut National de Physique Nucléaire et Physique des Particules/CNRS; the Russian Foundation for Basic Research; the Ministerio de Ciencia e Innovación, and Programa Consolider-Ingenio 2010, Spain; the Slovak R&D Agency; and the Academy of Finland.

[1] R. Barate et al. (LEP Working Group), Phys. Lett. B 565, 61 (2003).
[2] T. Aaltonen et al. (CDF and D0 Collaborations), Phys. Rev. Lett. 104, 061802 (2010).
[3] The TEVNPH Working Group (CDF and D0 Collaborations) (2010), arXiv:1007.4587.
[4] A. Djouadi, J. Kalinowski, and M. Spira, Comput. Phys. Commun. 108, 56 (1998).
[5] O. Brein, A. Djouadi, and R. Harlander, Phys. Lett. B579, 149 (2004), hep-ph/0307206.
[6] M. L. Ciccolini, S. Dittmaier, and M. Kramer, Phys. Rev. D68, 073003 (2003), hep-ph/0306234.
[7] D. Acosta et al. (CDF Collaboration), Phys. Rev. D 71, 052003 (2005).
[8] T. Aaltonen et al. (CDF Collaboration), Phys. Rev. Lett. 103, 101802 (2009).
[9] T. Aaltonen et al. (CDF Collaboration), Phys. Rev. Lett. 104, 141801 (2010).
[10] T. Aaltonen et al. (CDF Collaboration), Phys. Rev. Lett. 101, 251803 (2008).
[11] T. Aaltonen et al. (CDF Collaboration), Phys. Rev. D 80, 071101 (2009).
[12] S. Richter (2007), Ph.D. Thesis, Karlsruhe U., EKP, FERMILAB-THESIS-2007-35.
[13] T. Aaltonen et al. (CDF) (2010), submitted to Phys. Rev. D, arXiv:1004.1181.
[14] S. Z. Shalhout (2010), Ph.D. Thesis, Wayne State Univ., FERMILAB-THESIS-2010-33.
[15] We use a cylindrical coordinate system with $z$ along the proton beam direction, $r$ the perpendicular radius from the central axis of the detector, and $\phi$ the azimuthal angle. For $\theta$ the polar angle from the proton beam, we define $\eta = -\ln \tan(\theta/2)$, transverse momentum $p_T = p \sin \theta$ and transverse energy $E_T = E \sin \theta$.
[16] G. Blazey and B. Flaugher, Ann. Rev. Nucl. Part. Sci. 49, 633 (1999).
[17] The $\Delta R$ between two objects is given by $\sqrt{\Delta \eta^2 + \Delta \phi^2}$.
[18] A. Bhatti et al., Nucl. Instrum. Methods Phys. Res., Sect. A 566, 375 (2006).
[19] A. Abulencia et al. (CDF Collaboration), Phys. Rev. D 74, 072006 (2006).
[20] C. Neu for the CDF Collaboration, presented at TOP 2006, Coimbra, Portugal, 12-15 Jan 2006.
[21] T. Sjöstrand et al., Comput. Phys. Comm. 135, 238 (2001). We use “PYTHIA Tune A”, R. Field and R. C. Group, hep-ph/0510198v1.
[22] M. L. Mangano et al., J. High Energy Phys. 07 , 001 (2003).
[23] The missing $E_T$ ($\tilde{E}_T$) is defined by the sum over calorimeter towers: $\tilde{E}_T = -\sum_i E_T^{i}\hat{n}_i$, where $i =$ calorimeter tower number with $|\eta| < 3.6$, $\hat{n}_i$ is a unit vector perpendicular to the beam axis and pointing in the $i$th calorimeter tower. We also define $E_T = |\tilde{E}_T|$.
[24] We define dijet mass resolution as the standard deviation of the dijet mass distribution divided by its mean.
[25] J. M. Campbell and R. K. Ellis, Phys. Rev. D 62, 114012 (2000).
[26] J. M. Campbell and R. K. Ellis, Phys. Rev. D 60, 113006 (1999).
[27] U. Langenfeld, S. Moch, and P. Uwer, Phys. Rev. D 80, 054009 (2009).
[28] J. Baglio and A. Djouadi (2010), arXiv:1003.4266.
[29] D. Acosta et al. (CDF Collaboration), Phys. Rev. D 74, 071101 (2009).
[30] F. Fabbri et al. (Particle Data Group), J. Phys. G 37, 075021 (2010).
[31] T. Aaltonen et al. (CDF Collaboration), Phys. Rev. D 80, 012002 (2009).