Search for the standard model Higgs boson in $\ell\nu$ + jets final states in 9.7 fb^{-1} of $p\bar{p}$ collisions with the D0 detector

V. M. Abazov,[32] A. Abbinante,[45] B. Abbott,[67] B. S. Acharya,[26] M. Adams,[46] T. Adams,[44] G. D. Alexeev,[32] G. Alkhazov,[36] A. Alton,[56,*] A. Askew,[44] S. Atkins,[54] K. Augsten,[7] C. Avila,[5] F. Badaud,[10] L. Bagby,[45] B. Baldin,[45] D. V. Bandurin,[44] S. Banerjee,[26] E. Barberis,[55] P. Baringer,[53] J. F. Bartlett,[45] U. Bassler,[15] V. Bazterra,[46] A. Bean,[53] M. Begalli,[2] L. Bellantoni,[45] S. B. Beri,[24] G. Bernardi,[14] R. Bernhard,[19] I. Bertram,[39] M. Besançon,[15] R. Beuselinck,[40] P. C. Bhat,[45] S. Bhatia,[58] V. Bhatnagar,[24] G. Blazey,[47] S. Blessing,[44] K. Bloom,[59] A. Boehnlein,[45] D. Boline,[64] E. E. Boos,[34] G. Borissov,[39] A. Brandt,[70] O. Brandt,[20] R. Brock,[57] A. Bross,[45] D. Brown,[14] X. B. Bu,[45] M. Buehler,[45] V. Buescher,[21] V. Bunichev,[34] S. Burdin,[39,†] C. P. Buszello,[38] E. Camacho-Pérez,[29] B. C. K. Casey,[45] H. Castilla-Valdez,[29] S. Caughron,[57] S. Chakrabarti,[64] D. Chakraborty,[47] K. M. Chan,[51] A. Chandra,[72] E. Chapon,[15] G. Chen,[53] S. W. Cho,[28] S. Choi,[28] B. Choudhary,[25] S. Cihangir,[45] D. Claes,[59] J. Clutter,[53] M. Cooke,[45] W. E. Cooper,[45] M. Corcoran,[72] F. Couderc,[15] M.-C. Cousinou,[12] D. Cutts,[69] A. Das,[42] G. Davies,[40] S. J. de Jong,[30,31] E. De La Cruz-Burelo,[29] F. Déliot,[15] R. Demina,[63] D. Denisov,[45] S. P. Denisov,[35] S. Desai,[45] C. Deterre,[20,§] K. DeVaughan,[59] H. T. Diehl,[45] M. Diesburg,[45] P. F. Ding,[41] A. Dominguez,[59] A. Dubey,[25] L. V. Dudko,[34] A. Duperrin,[12] S. Dutt,[24] A. Dyshkant,[47] M. Eads,[47] D. Edmunds,[57] J. Ellison,[43] V. D. Elvira,[45] Y. Enari,[14] H. Evans,[49] V. N. Evdokimov,[35] L. Feng,[47] T. Ferbel,[63] F. Fiedler,[21] F. Filthaut,[30,31] W. Fisher,[57] H. E. Fisk,[45] M. Fortner,[47] H. Fox,[39] S. Fuess,[45] A. Garcia-Bellido,[63] J. A. García-González,[29] G. A. García-Guerra,[29,‡] V. Gavrilov,[33] W. Geng,[12,57] C. E. Gerber,[46] Y. Gershtein,[60] G. Ginther,[45,63] G. Golovanov,[32] P. D. Grannis,[64] S. Greder,[16] H. Greenlee,[45] G. Grenier,[17] Ph. Gris,[10] J.-F. Grivaz,[13] A. Grohsjean,[15,§] S. Grünendahl,[45] M. W. Grünewald,[27] T. Guillemin,[13] G. Gutierrez,[45] P. Gutierrez,[67] J. Haley,[55] L. Han,[4] K. Harder,[41] A. Harel,[63] J. M. Hauptman,[52] J. Hays,[40] T. Head,[41] T. Hebbeker,[18] D. Hedin,[47] H. Hegab,[68] A. P. Heinson,[43] U. Heintz,[69] C. Hensel,[20] I. Heredia-De La Cruz,[29] K. Herner,[45] G. Hesketh,[41,¶] M. D. Hildreth,[51] R. Hirosky,[73] T. Hoang,[44] J. D. Hobbs,[64] B. Hoeneisen,[9] J. Hogan,[72] M. Hohlfeld,[21] I. Howley,[70] Z. Hubacek,[7,15] V. Hynek,[7] I. Iashvili,[62] Y. Ilchenko,[71] R. Illingworth,[45] A. S. Ito,[45] S. Jabeen,[69] M. Jaffré,[13] A. Jayasinghe,[67] M. S. Jeong,[28] R. Jesik,[40] P. Jiang,[4] K. Johns,[42] E. Johnson,[57] M. Johnson,[45] A. Jonckheere,[45] P. Jonsson,[40] J. Joshi,[43] A. W. Jung,[45] A. Juste,[37] E. Kajfasz,[12] D. Karmanov,[34] I. Katsanos,[59] R. Kehoe,[71] S. Kermiche,[12] N. Khalatyan,[45] A. Khanov,[68] A. Kharchilava,[62] Y. N. Kharzheev,[32] I. Kiselevich,[33] J. M. Kohli,[24] A. V. Kozelov,[35] J. Kraus,[58] A. Kumar,[62] A. Kupco,[8] T. Kurča,[17] V. A. Kuzmin,[34] S. Lammers,[49] P. Lebrun,[17] H. S. Lee,[28] S. W. Lee,[52] W. M. Lee,[44] X. Lei,[42] J. Lellouch,[14] D. Li,[14] H. Li,[73] L. Li,[43] Q. Z. Li,[45] J. K. Lim,[28] D. Lincoln,[45] J. Linnemann,[57] V. V. Lipaev,[35] R. Lipton,[45] H. Liu,[71] Y. Liu,[4] A. Lobodenko,[36] M. Lokajicek,[8] R. Lopes de Sa,[64] R. Luna-Garcia,[29,**] A. L. Lyon,[45] A. K. A. Maciel,[1] R. Magaña-Villalba,[29] S. Malik,[59] V. L. Malyshev,[32] J. Mansour,[20] J. Martínez-Ortega,[29] R. McCarthy,[64] C. L. McGivern,[41] M. M. Meijer,[30,31] A. Melnitchouk,[45] D. Menezes,[47] P. G. Mercadante,[3] M. Merkin,[34] A. Meyer,[18] J. Meyer,[20,§§] F. Miconi,[16] N. K. Mondal,[26] M. Mulhearn,[73] E. Nagy,[12] M. Naimuddin,[25] M. Narain,[69] R. Nayyar,[42] H. A. Neal,[56] J. P. Negret,[5] P. Neustroev,[36] H. T. Nguyen,[73] T. Nunnemann,[22] J. Orduna,[72] N. Osman,[12] J. Osta,[51] M. Padilla,[43] A. Pal,[70] N. Parashar,[50] V. Parihar,[69] S. K. Park,[28] R. Partridge,[69,∥] N. Parua,[49] A. Patwa,[65,∥∥] B. Penning,[45] M. Perfilov,[34] Y. Peters,[20] K. Petridis,[41] G. Petrillo,[63] P. Pétroff,[13] M.-A. Pleier,[65] P. L. M. Podesta-Lerma,[29,††] A. Podkowa,[45,¶¶] V. M. Podstavkov,[45] A. V. Popov,[35] M. Prewitt,[72] D. Price,[49] N. Prokopenko,[35] J. Qian,[56] A. Quadt,[20] B. Quinn,[58] M. S. Rangel,[1] P. N. Ratoff,[39] I. Razumov,[35] I. Ripp-Baudot,[16] F. Rizatdinova,[68] M. Rominsky,[45] A. Ross,[39] C. Royon,[15] P. Rubinov,[45] R. Ruchti,[51] G. Sajot,[11] P. Salcido,[47] A. Sánchez-Hernández,[29] M. P. Sanders,[22] A. S. Santos,[1,‡‡] G. Savage,[45] L. Sawyer,[54] T. Scanlon,[40] R. D. Schamberger,[64] Y. Scheglov,[36] H. Schellman,[48] C. Schwanenberger,[41] R. Schwienhorst,[57] J. Sekaric,[53] H. Severini,[67] E. Shabalina,[20] V. Shary,[15] S. Shaw,[57] A. A. Shchukin,[35] R. K. Shivpuri,[25] V. Simak,[7] P. Skubic,[67] P. Slattery,[63] D. Smirnov,[51] K. J. Smith,[62] G. R. Snow,[59] J. Snow,[66] S. Snyder,[65] S. Söldner-Rembold,[41] L. Sonnenschein,[18] K. Soustruznik,[6] J. Stark,[11] D. A. Stoyanova,[35] M. Strauss,[67] L. Suter,[41] P. Svoisky,[67] M. Titov,[15] V. V. Tokmenin,[32] Y.-T. Tsai,[63] D. Tsybychev,[64] B. Tuchming,[15] C. Tully,[61] L. Uvarov,[36] S. Uvarov,[36] S. Uzunyan,[47] R. Van Kooten,[49] W. M. van Leeuwen,[30] N. Varelas,[46] E. W. Varnes,[42] I. A. Vasilyev,[35] A. Y. Verkheev,[32] L. S. Vertogradov,[32] M. Verzocchi,[45] M. Vesterinen,[41] D. Vilanova,[15] P. Vokac,[7] H. D. Wahl,[44] M. H. L. S. Wang,[45] J. Warchol,[51] G. Watts,[74] M. Wayne,[51] J. Weichert,[21] L. Welty-Rieger,[48] A. White,[70] D. Wicke,[23] M. R. J. Williams,[39] G. W. Wilson,[53] M. Wobisch,[54] D. R. Wood,[55] T. R. Wyatt,[41] Y. Xie,[45] R. Yamada,[45] S. Yang,[4] T. Yasuda,[45] Y. A. Yatsunenko,[32] W. Ye,[64] Z. Ye,[45] H. Yin,[45] K. Yip,[65] S. W. Youn,[45] J. M. Yu,[56] J. Zennamo,[62] T. G. Zhao,[41] B. Zhou,[56] J. Zhu,[56] M. Zielinski,[63] D. Zieminska,[49] and L. Zivkovic[14]

(D0 Collaboration)

[1]*LAFEX, Centro Brasileiro de Pesquisas Físicas, Rio de Janeiro, Brazil*
[2]*Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil*
[3]*Universidade Federal do ABC, Santo André, Brazil*
[4]*University of Science and Technology of China, Hefei, People's Republic of China*
[5]*Universidad de los Andes, Bogotá, Colombia*
[6]*Charles University, Faculty of Mathematics and Physics, Center for Particle Physics, Prague, Czech Republic*
[7]*Czech Technical University in Prague, Prague, Czech Republic*
[8]*Center for Particle Physics, Institute of Physics, Academy of Sciences of the Czech Republic, Prague, Czech Republic*
[9]*Universidad San Francisco de Quito, Quito, Ecuador*
[10]*LPC, Université Blaise Pascal, CNRS/IN2P3, Clermont, France*
[11]*LPSC, Université Joseph Fourier Grenoble 1, CNRS/IN2P3, Institut National Polytechnique de Grenoble, Grenoble, France*
[12]*CPPM, Aix-Marseille Université, CNRS/IN2P3, Marseille, France*
[13]*LAL, Université Paris-Sud, CNRS/IN2P3, Orsay, France*
[14]*LPNHE, Universités Paris VI and VII, CNRS/IN2P3, Paris, France*
[15]*CEA, Irfu, SPP, Saclay, France*
[16]*IPHC, Université de Strasbourg, CNRS/IN2P3, Strasbourg, France*
[17]*IPNL, Université Lyon 1, CNRS/IN2P3, Villeurbanne, France and Université de Lyon, Lyon, France*
[18]*III. Physikalisches Institut A, RWTH Aachen University, Aachen, Germany*
[19]*Physikalisches Institut, Universität Freiburg, Freiburg, Germany*
[20]*II. Physikalisches Institut, Georg-August-Universität Göttingen, Göttingen, Germany*
[21]*Institut für Physik, Universität Mainz, Mainz, Germany*
[22]*Ludwig-Maximilians-Universität München, München, Germany*
[23]*Fachbereich Physik, Bergische Universität Wuppertal, Wuppertal, Germany*
[24]*Panjab University, Chandigarh, India*
[25]*Delhi University, Delhi, India*
[26]*Tata Institute of Fundamental Research, Mumbai, India*
[27]*University College Dublin, Dublin, Ireland*
[28]*Korea Detector Laboratory, Korea University, Seoul, Korea*
[29]*CINVESTAV, Mexico City, Mexico*
[30]*Nikhef, Science Park, Amsterdam, The Netherlands*
[31]*Radboud University Nijmegen, Nijmegen, The Netherlands*
[32]*Joint Institute for Nuclear Research, Dubna, Russia*
[33]*Institute for Theoretical and Experimental Physics, Moscow, Russia*
[34]*Moscow State University, Moscow, Russia*
[35]*Institute for High Energy Physics, Protvino, Russia*
[36]*Petersburg Nuclear Physics Institute, St. Petersburg, Russia*
[37]*Institució Catalana de Recerca i Estudis Avançats (ICREA) and Institut de Física d'Altes Energies (IFAE), Barcelona, Spain*
[38]*Uppsala University, Uppsala, Sweden*
[39]*Lancaster University, Lancaster LA1 4YB, United Kingdom*
[40]*Imperial College London, London SW7 2AZ, United Kingdom*
[41]*The University of Manchester, Manchester M13 9PL, United Kingdom*
[42]*University of Arizona, Tucson, Arizona 85721, USA*
[43]*University of California Riverside, Riverside, California 92521, USA*
[44]*Florida State University, Tallahassee, Florida 32306, USA*
[45]*Fermi National Accelerator Laboratory, Batavia, Illinois 60510, USA*
[46]*University of Illinois at Chicago, Chicago, Illinois 60607, USA*
[47]*Northern Illinois University, DeKalb, Illinois 60115, USA*
[48]*Northwestern University, Evanston, Illinois 60208, USA*
[49]*Indiana University, Bloomington, Indiana 47405, USA*
[50]*Purdue University Calumet, Hammond, Indiana 46323, USA*
[51]*University of Notre Dame, Notre Dame, Indiana 46556, USA*
[52]*Iowa State University, Ames, Iowa 50011, USA*
[53]*University of Kansas, Lawrence, Kansas 66045, USA*
[54]*Louisiana Tech University, Ruston, Louisiana 71272, USA*
[55]*Northeastern University, Boston, Massachusetts 02115, USA*
[56]*University of Michigan, Ann Arbor, Michigan 48109, USA*
[57]*Michigan State University, East Lansing, Michigan 48824, USA*

^{58}*University of Mississippi, University, Mississippi 38677, USA*
^{59}*University of Nebraska, Lincoln, Nebraska 68588, USA*
^{60}*Rutgers University, Piscataway, New Jersey 08855, USA*
^{61}*Princeton University, Princeton, New Jersey 08544, USA*
^{62}*State University of New York, Buffalo, New York 14260, USA*
^{63}*University of Rochester, Rochester, New York 14627, USA*
^{64}*State University of New York, Stony Brook, New York 11794, USA*
^{65}*Brookhaven National Laboratory, Upton, New York 11973, USA*
^{66}*Langston University, Langston, Oklahoma 73050, USA*
^{67}*University of Oklahoma, Norman, Oklahoma 73019, USA*
^{68}*Oklahoma State University, Stillwater, Oklahoma 74078, USA*
^{69}*Brown University, Providence, Rhode Island 02912, USA*
^{70}*University of Texas, Arlington, Texas 76019, USA*
^{71}*Southern Methodist University, Dallas, Texas 75275, USA*
^{72}*Rice University, Houston, Texas 77005, USA*
^{73}*University of Virginia, Charlottesville, Virginia 22904, USA*
^{74}*University of Washington, Seattle, Washington 98195, USA*

We present, in detail, a search for the standard model Higgs boson, H, in final states with a charged lepton (electron or muon), missing energy, and two or more jets in data corresponding to 9.7 fb^{-1} of integrated luminosity collected at a center-of-mass energy of $\sqrt{s} = 1.96$ TeV with the D0 detector at the Fermilab Tevatron $p\bar{p}$ Collider. The search uses b-jet identification to categorize events for improved signal versus background separation and is sensitive to associated production of the H with a W boson, $WH \to \ell\nu b\bar{b}$; gluon fusion with the Higgs decaying to W-boson pairs, $H \to WW \to \ell\nu jj$; and associated production with a vector boson where the Higgs decays to W-boson pairs, $VH \to VWW \to \ell\nu jjjj$ production (where $V = W$ or Z). We observe good agreement between data and expected background. We test our method by measuring WZ and ZZ production with $Z \to b\bar{b}$ and find production rates consistent with the standard model prediction. For a Higgs boson mass of 125 GeV, we set a 95% C.L. upper limit on the production of a standard model Higgs boson of $5.8 \times \sigma_{SM}$, where σ_{SM} is the standard model Higgs boson production cross section, while the expected limit is $4.7 \times \sigma_{SM}$. We also interpret the data considering models with fourth generation fermions, or a fermiophobic Higgs boson.

I. INTRODUCTION

The Higgs boson is the massive physical state that emerges from electroweak symmetry breaking in the Higgs mechanism [1–3]. This mechanism generates the masses of the weak gauge bosons and explains the fermion masses through their Yukawa couplings to the Higgs boson

*Visitor from Augustana College, Sioux Falls, SD, USA.
†Visitor from The University of Liverpool, Liverpool, UK.
‡Visitor from UPIITA-IPN, Mexico City, Mexico.
§Visitor from DESY, Hamburg, Germany.
$^{\parallel}$Visitors from SLAC, Menlo Park, CA, USA.
¶Visitor from University College London, London, UK.
**Visitor from Centro de Investigacion en Computacion—IPN, Mexico City, Mexico.
††Visitor from ECFM, Universidad Autonoma de Sinaloa, Culiacán, Mexico.
‡‡Visitor from Universidade Estadual Paulista, São Paulo, Brazil.
§§Visitor from Karlsruher Institut für Technologie (KIT)—Steinbuch Centre for Computing (SCC.)
$^{\parallel\parallel}$Visitor from Office of Science, U.S. Department of Energy, Washington, D.C. 20585, USA.
¶¶Visitor from Bradley University, Peoria, IL, USA.

field. The mass of the Higgs boson (M_H) is a free parameter in the standard model (SM). Precision measurements of various SM electroweak parameters constrain M_H to be less than 152 GeV at the 95% C.L. [4–6]. Direct searches at the CERN e^+e^- Collider (LEP) [7] exclude $M_H < 114.4$ GeV at the 95% C.L. The ATLAS and CMS Collaborations, using pp collisions at the CERN LHC, exclude masses from $110 < M_H < 600$ GeV, except for a narrow region between 122 and 127 GeV [8,9]. Both experiments observe a resonance at a mass of ≈ 125 GeV, primarily in the $\gamma\gamma$ and ZZ final states, with a significance greater than 5 standard deviations (s.d.) that is consistent with SM Higgs boson production [10,11]. The CDF and D0 Collaborations at the Fermilab Tevatron Collider report a combined analysis that excludes the region $147 < M_H < 179$ GeV [12] and shows evidence at the 3 s.d. level for a particle decaying to $b\bar{b}$, produced in association with a W or Z boson, consistent with SM WH/ZH production [13]. Demonstrating that the observed resonance is the SM Higgs boson requires also observing it at the predicted rate in the $b\bar{b}$ final state, which is the dominant decay mode for masses below $M_H \lesssim 135$ GeV.

The dominant production process for the Higgs boson at the Tevatron Collider is gluon fusion ($gg \rightarrow H$), followed by the associated production of a Higgs boson with a vector boson (VH), then via vector-boson fusion ($VVqq' \rightarrow Hqq'$). At masses below $M_H \approx 135$ GeV, the Higgs boson mainly decays to a pair of b quarks, while for larger masses, the dominant decay is to a pair of W bosons. Because the $H \rightarrow b\bar{b}$ process is difficult to distinguish from background at hadron colliders, it is more effective to search for the Higgs boson produced in association with a vector boson for this decay channel.

This article presents a search by the D0 collaboration for the SM Higgs boson using events containing one isolated charged lepton ($\ell = e$ or μ), a significant imbalance in transverse energy ($\not{E}_T$), and two or more jets. It includes a detailed description of the $WH \rightarrow \ell\nu b\bar{b}$ search, initially presented in Ref. [14] and used as an input to the result presented in Ref. [13], differing from and superseding that result due to an updated treatment of some systematic uncertainties as described in Sec. X below. The complete analysis comprises searches for the production and decay channels $WH \rightarrow \ell\nu b\bar{b}$, $H \rightarrow WW^* \rightarrow \ell\nu jj$ (where $j = u$, d, s, c), and $VH \rightarrow VWW^* \rightarrow \ell\nu jjjj$ (where $V = W$ or Z). This search also considers contributions from ZH production and from the decay $H \rightarrow ZZ$ when one of the charged leptons from $Z \rightarrow \ell\ell$ decay is not identified in the detector. We optimize the analysis by subdividing data into mutually exclusive subchannels based on charged lepton flavor, jet multiplicity, and the number and quality of candidate b-quark jets. This search also extends the most recent D0 $WH \rightarrow \ell\nu b\bar{b}$ search [14] by adding subchannels with looser b-quark jet identification requirements and subchannels with four or more jets. These additional subchannels are primarily sensitive to $H \rightarrow WW^* \rightarrow \ell\nu jj$ and $VH \rightarrow VWW^* \rightarrow \ell\nu jjjj$ production and extend the reach of our search to $M_H = 200$ GeV. We present a measurement of VZ production with $Z \rightarrow b\bar{b}$ as a cross-check on our methodology in Sec. XI. In addition to our standard model interpretation, we consider interpretations of our result in models with a fourth generation of fermions, and models with a fermiophobic Higgs as described in Sec. XIII.

Several other searches for $WH \rightarrow \ell\nu b\bar{b}$ production have been reported at a $p\bar{p}$ center-of-mass energy of $\sqrt{s} = 1.96$ TeV, most recently by the CDF Collaboration [15]. The results presented here supersede previous searches by the D0 Collaboration, presented in Refs. [16–20], which used subsamples of the data presented in this article. They also supersede a previous search for Higgs boson production in the $\ell\nu jj$ final state by the D0 Collaboration [21].

II. THE D0 DETECTOR

This analysis relies on all major components of the D0 detector: tracking detectors, calorimeters, and the muon identification system. These systems are described in detail in Refs. [22–25].

Closest to the interaction point is the silicon microstrip tracker (SMT) followed by the central scintillating fiber tracker. These detector subsystems are located inside a 2 T magnetic field provided by a superconducting solenoid. They track charged particles and are used to reconstruct primary and secondary vertices for pseudorapidities [26] of $|\eta| < 3$. Outside the solenoid is the liquid argon/uranium calorimeter consisting of one central calorimeter (CC) covering $|\eta| \lesssim 1$ and two end calorimeters (EC) extending coverage to $|\eta| \approx 4$. Each calorimeter contains an innermost finely segmented electromagnetic layer followed by two hadronic layers, with fine and coarse segmentation, respectively. The main functions of the calorimeters are to measure energies and help identify electrons, photons, and jets using coordinate information of significant energy clusters. They also give a measure of the $\not{E}_T$. A preshower detector between the solenoidal magnet and central calorimeter consists of a cylindrical radiator and three layers of scintillator strips covering the region $|\eta| < 1.3$. The outermost system provides muon identification. It is divided into a central section that covers $|\eta| < 1$ and forward sections that extend coverage out to $|\eta| \approx 2$. The muon system is composed of three layers of drift tubes and scintillation counters, one layer before and two layers after a 1.8 T toroidal magnet.

III. EVENT TRIGGER

Events in the electron channel are triggered by a logical OR of triggers that require an electromagnetic object and jets, as described in Ref. [20]. Trigger efficiencies are modeled in the Monte Carlo (MC) simulation by applying the trigger efficiency, measured in data, as an event weight. This efficiency is parametrized as a function of electron η, azimuthal angle ϕ [27], and transverse momentum p_T. For the events selected in our analysis, these triggers have an efficiency of (90–100)% depending on the trigger and the region of the detector.

The muon channel uses an inclusive trigger approach, based on the logical OR of all available triggers, except those containing lifetime-based requirements that can bias the performance of b-jet identification. To determine the trigger efficiency, we compare data events selected with a well-modeled logical OR of the single muon and muon + jets triggers ($T_{\mu\text{OR}}$), which are about 70% efficient, to events selected using all triggers. The increase in event yield in the inclusive trigger sample is used to determine an inclusive trigger correction for the MC trigger efficiency, P_{corr}, relative to the $T_{\mu\text{OR}}$ trigger ensemble,

$$P_{\text{corr}} = \frac{(N_{\text{Data}} - N_{\text{MJ}})_{\text{incl}} - (N_{\text{Data}} - N_{\text{MJ}})_{T_{\mu\text{OR}}}}{N_{\text{MC}}}, \quad (1)$$

where the numerator is the difference between the number of data events in the inclusive trigger sample and the $T_{\mu\text{OR}}$

trigger sample, after subtracting off instrumental multijet (MJ) backgrounds, and the denominator is the number of MC events (after the event selection and normalization to data described in Sec. VIII and the MC corrections are applied as described in Sec. VI A) with the trigger efficiency set to 1. The total trigger efficiency estimate for events in the muon channel is $T_{\mu\mathrm{OR}} + P_{\mathrm{corr}}$, limited to be ≤ 1.

Triggers based on jets and $\not{E}_T$ make the most significant contributions to the inclusive set of triggers beyond those included in the well-modeled $T_{\mu\mathrm{OR}}$ trigger set. To account for these contributions, the correction from $T_{\mu\mathrm{OR}}$ triggers to the inclusive set of triggers is parametrized as a function of the scalar sum of the transverse momenta of all jets, H_T, and the $\not{E}_T$, and is derived for separate regions in muon η.

For $|\eta| < 1.0$, events are dominantly triggered by single muon triggers, while for $|\eta| > 1.6$, triggers based on the logical OR of muon + jets prevail. The third region, $1.0 < |\eta| < 1.6$, is a mixture of single muon and muon + jets triggers. In the $|\eta| < 1.0$ and $1.0 < |\eta| < 1.6$ regions the detector support structure allows only partial coverage by the muon system. This impacts the muon trigger efficiency

in the region $-2 < \phi < -1.2$. In these regions, we therefore derive separate corrections. The inclusive trigger approach results in a gain of about 30% in efficiency over using only muon and the muon + jets triggers. Examples of these corrections, P_{corr}, are shown in Fig. 1.

IV. IDENTIFICATION OF LEPTONS, JETS, AND $\not{E}_T$

To reconstruct the candidate $W(\rightarrow \ell\nu)$ boson, our selected events are required to contain a single identified electron or muon together with significant $\not{E}_T$. To ensure statistical independence with channels that contain more than one lepton, we do not consider events with more than one electron or muon. Two or more jets are also required in order to study $WH \rightarrow \ell\nu b\bar{b}$, $H \rightarrow WW \rightarrow \ell\nu jj$, and $VH \rightarrow VWW \rightarrow \ell\nu jjjj$ production. Two sets of lepton identification criteria are applied for each lepton channel in order to form a "loose" sample, used to estimate the multijet background from data as described in Sec. VII, and a "tight" sample used to perform the search. The event selection procedure, prior to b-jet categorization, is similar to that described in Ref. [20] and described in more detail below.

Electrons with $p_T > 15$ GeV are selected in the pseudorapidity regions $|\eta| < 1.1$ and $1.5 < |\eta| < 2.5$, corresponding to the CC and EC, respectively. Multivariate discriminants are used to identify electrons, with a separate discriminant trained for the CC and EC regions. The discriminants are based on boosted decision trees [28–32] (BDTs) as implemented in the TMVA package [33] with input variables that are listed below. The BDTs are discussed in more detail in Sec. IX. The loose and tight electron samples are defined by different requirements on the response of these multivariate discriminants that are chosen to retain high electron selection efficiencies while suppressing backgrounds at differing rates.

Leptons coming from the leptonic decays of W bosons tend to be isolated from jets. Isolated electromagnetic showers are identified within a cone in η-ϕ space of $\Delta\mathcal{R} = \sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.4$ [34]. In the CC (EC), an electromagnetic shower is required to deposit 97% (90%) of its total energy within a cone of radius $\Delta\mathcal{R} = 0.2$ in the electromagnetic calorimeter. The showers must have transverse and longitudinal distributions that are consistent with those expected from electrons. In the CC region, a reconstructed track, isolated from other tracks, is required to have a trajectory that extrapolates to the electromagnetic (EM) shower. The isolation criteria restrict the sum of the scalar p_T of tracks with $p_T > 0.5$ GeV within a hollow cone of radius $0.05 < \Delta\mathcal{R} < 0.4$ surrounding the electron candidate to be less than 2.5 GeV. The BDTs are constructed using additional information, such as the number and scalar p_T sum of tracks in the cone of radius $\Delta\mathcal{R} < 0.4$ surrounding the candidate cluster, track-to-cluster-matching probability, the ratio of the transverse energy of the cluster to the transverse momentum of the track associated with the shower, the EM energy fraction,
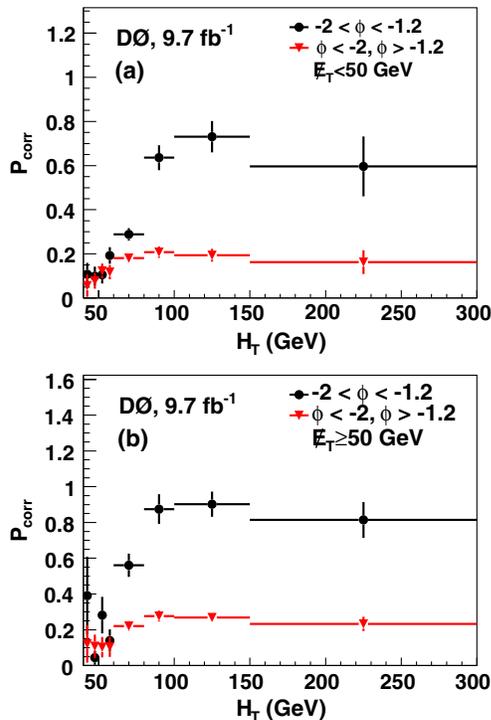

FIG. 1 (color online). Data-derived muon trigger correction to account for the resulting efficiency gain in moving from single muon and muon + jets triggers to inclusive triggers as a function of H_T for $|\eta| < 1.0$, shown (a) for events with $\not{E}_T < 50$ GeV and (b) for events with $\not{E}_T \geq 50$ GeV. The black circles show the correction when the muon is in the region of ϕ ($-2 < \phi < -1.2$) where there is a gap in the muon coverage for detector supports, and the red triangles show the correction elsewhere in ϕ.

lateral and longitudinal shower-shape characteristics, as well as the number of hits in the various layers of the tracking detector, and information from the central pre-shower detector. The discriminants are trained using $Z/\gamma^* \to ee$ data events.

We select muons with $p_T > 15$ GeV and $|\eta| < 2.0$. They are required to have reconstructed track segments in layers of the muon system both before and after the toroidal magnet, except where detector support structure limits muon system coverage, for which the presence of track segments in any layer is sufficient. The local muon system track must be spatially matched to a track in the central tracker.

Muons originating from semileptonic decays of heavy flavored hadrons are typically not isolated due to jet fragmentation and secondary particles from the partial hadronic decays. We employ a loose muon definition, requiring a minimal separation of $\Delta \mathcal{R}(\mu, j) > 0.5$ between the muon and any jet, while the tight identification has additional isolation requirements. For tight muons, the scalar sum of the p_T of tracks with $\Delta \mathcal{R} < 0.5$ around the muon candidate is required to be less than $0.4 \times p_T^\mu$. Furthermore, the transverse energy deposits in the calorimeter in a hollow cone of $0.1 < \Delta \mathcal{R} < 0.4$ around the muon must be less than $0.12 \times p_T^\mu$. To suppress cosmic ray muons, scintillator timing information is used to require hits in the detector to coincide with a beam crossing.

To reduce backgrounds from $Z/\gamma^* \to \ell\ell +$ jets and $t\bar{t}$ production, we reject events containing more than one tight-isolated charged lepton. Jets are reconstructed in the calorimeters in the region $|\eta| < 2.5$ using an iterative midpoint cone algorithm, with a cone size of $\Delta \mathcal{R} = 0.5$ [35]. To minimize the possibility that jets are caused by noise or spurious energy deposits, the fraction of the total jet energy contained in the electromagnetic layers of the calorimeter is required to be between 5% and 95%, and the energy fraction in the coarse hadronic layers of the calorimeter is required to be less than 40%. To suppress noise, different energy thresholds are also applied to clustered and to isolated cells [36]. The energy of the jets is scaled by applying a correction determined from $\gamma +$ jet events using the same jet-finding algorithm. This scale correction accounts for additional energy (e.g., residual energy from previous bunch crossings and energy from multiple $p\bar{p}$ interactions) that is sampled within the finite cone size, the calorimeter energy response to particles produced within the jet cone, and energy flowing outside the cone or moving into the cone via detector effects [36]. We also apply an additional correction that accounts for the flavor composition of jets [37].

Jet energy calibration and resolution are adjusted in simulated events to match those measured in data. This correction is derived from $Z(\to ee) +$ jet events from the p_T imbalance between the Z boson and the recoiling jet in the MC simulation when compared to that observed in

data, and applied to jet samples in MC events. Differences in reconstruction thresholds in simulation and data are also taken into account, and the jet identification efficiency and jet resolution are adjusted in the simulation to match those measured in data. All selected jets are required to have $p_T > 20$ GeV and $|\eta| < 2.5$. We require that jets originate from the primary $p\bar{p}$ vertex (PV), such that each selected jet is matched to at least two tracks with $p_T > 0.5$ GeV that have at least one hit in the SMT detector and a distance of closest approach with respect to the PV of less than 0.5 cm in the transverse plane and less than 1 cm along the beam axis (z). Interaction vertices are reconstructed from tracks that have $p_T > 0.5$ GeV with at least two hits in the SMT. The primary vertex is the reconstructed vertex with the highest average p_T of its tracks. Vertex reconstruction is described in more detail in Ref. [38]. We also require that the PV be reconstructed within $z_{PV} = \pm 60$ cm of the center of the detector.

The $\not{E}_T$ is calculated from individual calorimeter cell energies in the electromagnetic and fine hadronic sections of the calorimeter and is required to satisfy $\not{E}_T > 15$ GeV for the electron channel and $\not{E}_T > 20$ GeV for the muon channel. Energy from the coarse hadronic layers that is contained within a jet is also included in the $\not{E}_T$ calculation. A correction for the presence of any muons and all energy corrections applied to electrons and jets are propagated to the value of $\not{E}_T$.

V. TAGGING OF *b*-QUARK JETS

The *b*-tagging algorithm for identifying jets originating from *b* quarks is based on a multivariate discriminant using a combination of variables sensitive to the presence of tracks or secondary vertices displaced significantly in the *x*-*y* plane from the $p\bar{p}$ interaction vertex. This algorithm provides improved performance over the neural network algorithm described in Ref. [38].

Jets considered by the *b*-tagging algorithm are required to be "taggable," i.e., contain at least two tracks with $p_T > 0.5$ GeV, including at least one with $p_T > 1$ GeV, that each have at least one hit in the SMT and have a distance of closest approach with respect to the PV of less than 0.15 (0.4) cm in the transverse plane (in z). The efficiency of this requirement accounts for variations in detector acceptance and track reconstruction efficiencies at different locations of the PV prior to the application of the *b*-tagging algorithm, and depends on the z position of the PV and the p_T and η of the jet. For jets that pass through the geometrical acceptance of the tracking system, this efficiency is typically about 97%. The efficiency for *b*-tagging is determined with respect to taggable jets. The correction for taggability is measured in the selected data sample, while the corrections for *b*-tagging are determined in an independent heavy-flavor jet enriched sample of events that include a jet containing a muon, as described in Ref. [38]. The efficiency for jets to be taggable and to satisfy
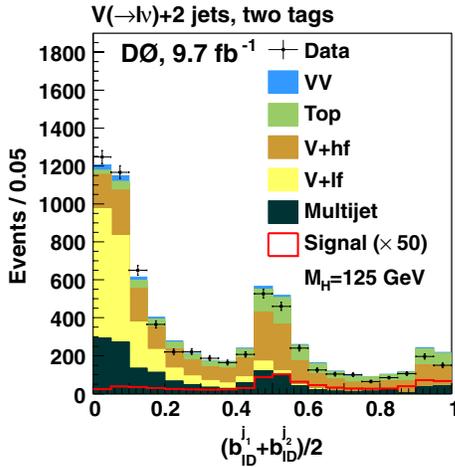
FIG. 2 (color online).   Average of the b-identification discriminant outputs of each jet in events with two jets. Discrepancies in data-MC agreement are within the systematic uncertainties described in Sec. X.

b-tagging requirements in the simulation is corrected to reproduce the respective efficiencies in the data.

We define six independent tagging samples with zero, one loose, one tight, two loose, two medium, or two tight b-tagged jets. An inclusive "pretag" sample is also considered for parts of this analysis. Events with no jets satisfying the taggability or b-tagging criteria define the zero b-tag sample. If exactly one jet is b-tagged, and the b-identification discriminant output for that jet, $b_{\text{ID}}^{j_i}$, satisfies the tight selection threshold ($b_{\text{ID}}^{j_i} > 0.15$), that event is considered part of the one tight b-tag sample. Events with exactly one b-tagged jet that fails the tight selection threshold, but passes the loose selection threshold ($b_{\text{ID}}^{j_i} > 0.02$) are included in the one loose b-tag sample. Events with two or more b-tagged jets are assigned to either the two loose b-tags, two medium b-tags, or two tight b-tags category, depending on the value of the average b-identification discriminant of the two jets with the highest discriminant values, i.e., the double tight category is required to satisfy $(b_{\text{ID}}^{j_1} + b_{\text{ID}}^{j_2})/2 > 0.55$; the medium category is $0.35 < (b_{\text{ID}}^{j_1} + b_{\text{ID}}^{j_2})/2 \leq 0.55$; and the loose category is $0.02 < (b_{\text{ID}}^{j_1} + b_{\text{ID}}^{j_2})/2 \leq 0.35$ (see Fig. 2). The operating point for the loose (medium, tight) threshold has an identification efficiency of 79% (57%, 47%) for individual b jets, averaged over selected jet p_{T} and η distributions, with a b-tagging misidentification rate of 11% (0.6%, 0.15%) for light-quark jets (lf), calculated by the method described in Ref. [38].

VI. MONTE CARLO SIMULATION

We account for all Higgs boson production and decay processes that can lead to a final state containing exactly one charged well-isolated lepton, $\not{E}_{\text{T}}$, and two or more jets. The signal processes considered are the following.

(i) Associated production of a Higgs boson with a vector boson where the Higgs boson decays to $b\bar{b}$, $c\bar{c}$, $\tau\tau$, or VV. The associated weak vector boson decays leptonically in the case of $H \rightarrow b\bar{b}$ and either leptonically or hadronically in the case of $H \rightarrow WW$. Contributions from $Z(\rightarrow \ell\ell)H(\rightarrow b\bar{b})$ production arise from identifying only one charged lepton in the detector, with the other contributing to the $\not{E}_{\text{T}}$.

(ii) Higgs boson production via gluon fusion with the subsequent decay $H \rightarrow VV$, where one weak vector boson decays leptonically (with exactly one identified lepton).

(iii) Higgs boson production via vector-boson fusion with the subsequent decay $H \rightarrow VV$, where one weak vector boson decays leptonically (with exactly one identified lepton).

Various SM processes can mimic expected signal signatures, including $V + $ jets, diboson (VV), MJ, $t\bar{t}$, and single top-quark production.

All signal processes and most of the background processes are estimated from MC simulation, while the MJ background is evaluated from data, as described in Sec. VII. We use PYTHIA [39] to simulate all signal processes and diboson processes. The $V + $ jets and $t\bar{t}$ samples are simulated with the ALPGEN [40] MC generator interfaced to PYTHIA for parton showering and hadronization, while the SINGLETOP event generator [41,42] interfaced to PYTHIA is used for single top-quark events. To avoid overestimating the probability of further partonic emissions in PYTHIA, the MLM factorization ("matching") scheme [43] is used. All of these simulations use CTEQ6L1 [44,45] parton distribution functions (PDFs).

A full GEANT-based [46] detector simulation is used to process signal and background events. To account for residual activity from previous beam crossings and contributions from the presence of additional $p\bar{p}$ interactions, events from randomly selected beam crossings with the same instantaneous luminosity profile as the data are overlaid on the simulated events. All events are then reconstructed using the same software as used for data.

The signal cross sections and branching fractions are normalized to the SM predictions [12]. The WH and ZH cross sections are calculated at next-to-next-to-leading order (NNLO) [47], with MSTW2008 NNLO PDFs [48]. The gluon fusion process uses the NNLO plus next-to-next-to-leading-log (NNLL) calculation [49], and the vector-boson fusion process is calculated at NNLO in QCD [50]. The Higgs boson decay branching fractions are obtained with HDECAY [51,52]. We use next-to-leading-order cross sections to normalize single top-quark [53] and diboson [54,55] production, while we use an approximate NNLO calculation for $t\bar{t}$ production [56]. The p_{T} of the Z boson in $Z + $ jets events is corrected to match that observed in the data [57]. The p_{T} of the W boson in $W + $ jets events is corrected using the same

dependence but taking into account the differences between the p_T spectra of the Z and W bosons in NNLO QCD [58]. Additional scale factors to account for higher-order terms in the ALPGEN MC for the $V +$ heavy flavor jets, $V + hf$, are obtained from MCFM [55,59]. The $V +$ jets processes are then normalized to the data for each lepton flavor and jet multiplicity separately as described in Sec. VIII.

A. MC reweighting

Motivated by previous comparisons of ALPGEN with data [60] and with other event generators [43], we develop corrections to $W +$ jets and $Z +$ jets MC samples to correct for the shape discrepancies in kinematic distributions between data and simulation. The corrections are derived based on the direct comparison between data and MC samples prior to the application of b-tagging, where any contamination from signal is very small.

To improve the description of jet directions, we correct the η distributions of the leading and second leading jets in $W/Z +$ jets events. The correction function is a fourth-order polynomial determined from the ratio of the $V +$ jets events in MC and data minus non-$V +$ jets backgrounds. The modeling of the lepton η in $W +$ jets events is adjusted by applying a second-order polynomial correction. Correlated discrepancies observed in the leptonically decaying W-boson transverse momentum, p_T^W, and the jet angular separation, $\Delta\mathcal{R}(j_1, j_2)$, are corrected through two reweighting functions in the two-dimensional $\Delta\mathcal{R}$-p_T^W plane [20]. The p_T^W reweighting is applied only to $W +$ jets events, while the $\Delta\mathcal{R}$ reweighting is applied to both $W +$ jets and $Z +$ jets events. Each of these corrections is designed to change differential distributions, but to preserve normalization. Corrections are on the order of a few percent in the highly populated region of each distribution and may exceed 10% for extreme values of each distribution.

All corrections are derived in events selected with muon $+$ jets triggers to minimize uncertainties due to contamination from MJ events, and are applied to both the electron and muon channels. Additional p_T^W, $\Delta\mathcal{R}$, and lepton η corrections and corresponding systematic uncertainties are determined from events selected with inclusive muon triggers and are applied to events containing muons, accounting for variations in modeling distributions of the inclusively triggered events.

VII. MULTIJET BACKGROUND

The MJ background—events where a jet is misidentified as a lepton—is determined from the data prior to the application of b-tagging, using a method similar to the one used in Ref. [20]. This method involves applying event weights that depend on the relative efficiency ε_{LT}^{ℓ} of a lepton passing loose requirements to subsequently pass the tight requirements and a similar relative probability,

P_{LT}^{MJ}, for an MJ event to pass these sequential selections. An MJ template is constructed by selecting events from data in which the lepton passes the loose isolation requirement, but fails the tight requirement, as described in Sec. IV. Each event in the MJ template is weighted by
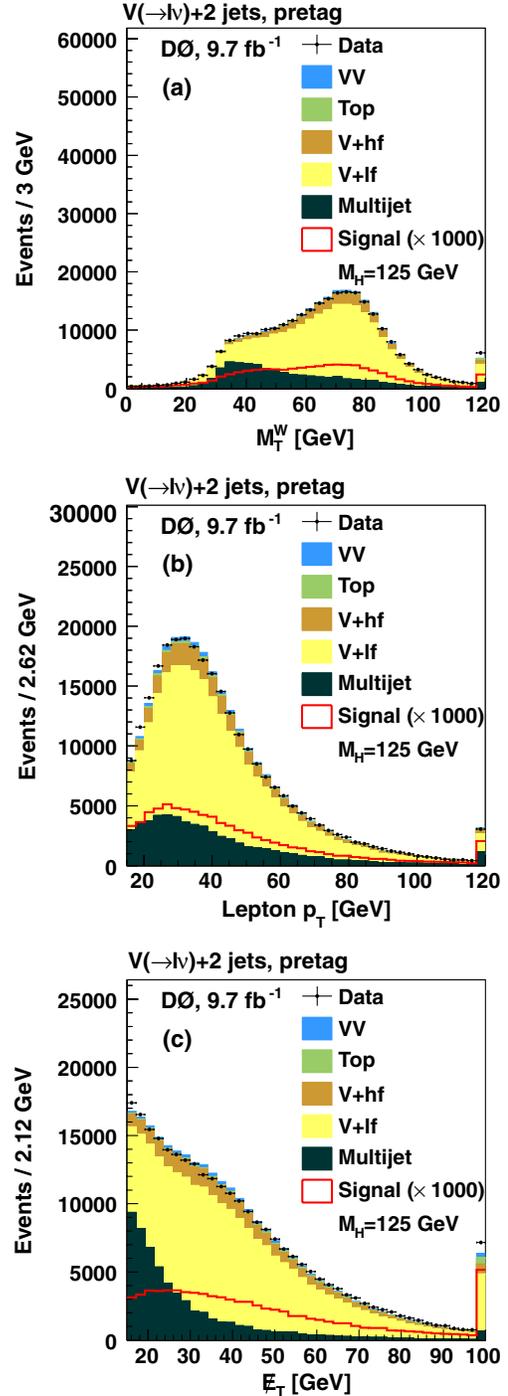

FIG. 3 (color online).   Distributions for all selected events with two jets of (a) transverse mass of the lepton-$\slashed{E}_T$ system, (b) charged lepton p_T, and (c) $\slashed{E}_T$. The signal is multiplied by 1000. Overflow events are added to the last bin.

$$w_{\mathrm{MJ}} = \frac{P_{\mathrm{LT}}^{\mathrm{MJ}}}{1 - P_{\mathrm{LT}}^{\mathrm{MJ}}}, \qquad (2)$$

where $P_{\mathrm{LT}}^{\mathrm{MJ}}$ is a function of the event kinematics. Since the MJ template contains a contribution from events with real leptons originating from leptonic decays of W/Z bosons, we correct the normalization of the $V +$ jets MC using the event weight

$$w_{\mathrm{VJ}} = 1 - \frac{P_{\mathrm{LT}}^{\mathrm{MJ}}(1 - \epsilon_{\mathrm{LT}}^{\ell})}{\epsilon_{\mathrm{LT}}^{\ell}(1 - P_{\mathrm{LT}}^{\mathrm{MJ}})}, \qquad (3)$$

where $\epsilon_{\mathrm{LT}}^{\ell}$ and $P_{\mathrm{LT}}^{\mathrm{MJ}}$ are functions of event kinematics. The efficiencies $\epsilon_{\mathrm{LT}}^{\ell}$ are functions of lepton p_{T}, and they are determined from $Z/\gamma^* \to \ell\ell$ events. The probabilities $P_{\mathrm{LT}}^{\mathrm{MJ}}$ are determined in the region $5 < \not{E}_{\mathrm{T}} < 15$ GeV from the measured ratio of the number of events with tight leptons and those with loose leptons after correcting each sample for the expected MC contribution from real leptons in the

specific kinematic interval. Electron channel probabilities are parametrized in p_{T}, calorimeter detector η, and $\min \Delta\phi(\not{E}_{\mathrm{T}}, j)$, while probabilities in the muon channel are parametrized in p_{T} for different regions in muon detector η and $\Delta\phi(\not{E}_{\mathrm{T}}, \mu)$.

VIII. EVENT SELECTION

Events are required to have one isolated charged lepton, large $\not{E}_{\mathrm{T}}$, and two or more jets, as described in Sec. IV. To suppress MJ backgrounds, events must satisfy the additional requirement that $M_T^W > 40$ GeV $- 0.5 \times \not{E}_{\mathrm{T}}$, where M_T^W is the transverse mass [61] of the W boson. We then perform the final normalization of the $V +$ jets and MJ backgrounds via a simultaneous fit to data in the M_T^W distribution after subtracting the other SM background predictions from the data as described in Refs. [14,19,20]. The distribution of M_T^W after this normalization procedure is shown in Fig. 3(a). We perform separate fits for each

TABLE I. Observed number of events in data and expected number of events from each signal and background source (where $V = W, Z$) for events with exactly two jets. The expected signal is quoted at $M_H = 125$ GeV. The total background uncertainty includes all sources of systematic uncertainty added in quadrature.

	Pretag	Zero b tags	One loose b tag	One tight b tag	Two loose b tags	Two med. b tags	Two tight b tags
$VH \to \ell\nu b\bar{b}$	37.3	6.4	4.0	11.6	3.2	4.6	7.7
$H \to VV \to \ell\nu jj$	24.7	18.8	3.9	1.8	0.3	0.07	0
$VH \to VVV \to \ell\nu jjjj$	13.0	9.3	2.3	1.2	0.3	0.04	0.01
Diboson	5686	4035	968	535	109	42	38
$V + (g, u, d, s)$-jets	182 271	148 686	26 421	6174	1762	132	13
$V + (b\bar{b}/c\bar{c})$	27 443	15 089	4872	5236	978	691	691
Top ($t\bar{t}$ + single top)	3528	758	455	1289	247	333	462
Multijet	58 002	43 546	9316	3700	946	298	195
Total expectation	276 930	212 114	42 032	16 935	4043	1496	1400
Total uncertainty	±14 998	±11 352	±2438	±1696	±362	±117	±175
Observed events	276 929	211 169	42 774	16 406	4057	1358	1165

TABLE II. Observed number of events in data and expected number of events from each signal and background source (where $V = W, Z$) for events with exactly three jets. The expected signal is quoted at $M_H = 125$ GeV. The total background uncertainty includes all sources of systematic uncertainty added in quadrature.

	Pretag	Zero b tags	One loose b tag	One tight b tag	Two loose b tags	Two med. b tags	Two tight b tags
$VH \to \ell\nu b\bar{b}$	8.6	1.3	1.0	2.4	0.9	1.1	1.7
$H \to VV \to \ell\nu jj$	8.8	6.0	1.7	0.8	0.3	0.07	0.01
$VH \to VVV \to \ell\nu jjjj$	7.3	4.5	1.6	0.9	0.3	0.05	0.01
Diboson	1138	727	238	113	42	14	10
$V + (g, u, d, s)$-jets	24 086	18 078	4577	976	582	34	3
$V + (b\bar{b}/c\bar{c})$	6625	3213	1349	1250	411	228	164
Top ($t\bar{t}$ + single top)	3695	563	419	1123	365	460	570
Multijet	10 364	6629	2162	933	367	130	82
Total expectation	45 908	29 209	8746	4395	1768	867	830
Total uncertainty	±2582	±1619	±587	±528	±209	±118	±113
Observed events	45 907	28 924	8814	4278	1815	879	797

TABLE III. Observed number of events in data and expected number of events from each signal and background source (where $V = W, Z$) for events with four or more jets. The expected signal is quoted at $M_H = 125$ GeV. The total background uncertainty includes all sources of systematic uncertainty added in quadrature.

	Pretag	Zero *b* tags	One loose *b* tag
$VH \rightarrow \ell\nu b\bar{b}$	1.4	0.2	0.2
$H \rightarrow VV \rightarrow \ell\nu jj$	2.4	1.4	0.6
$VH \rightarrow VVV \rightarrow \ell\nu jjjj$	3.6	2.0	0.8
Diboson	199	112	46
$V + (g, u, d, s)$-jets	3055	2143	679
$V + (b\bar{b}/c\bar{c})$	1280	542	286
Top ($t\bar{t}$ + single top)	2889	311	268
Multijet	2092	1110	450
Total expectation	9516	4217	1729
Total uncertainty	±530	±231	±144
Observed events	9685	3915	1786

lepton flavor and jet multiplicity category before dividing events into categories based on the number and quality of identified *b* jets, as described in Sec. V. All events passing these selection criteria constitute the pretag sample, and each pretag event also belongs to exactly one of the six independent *b*-tag categories. Only the zero and one-loose *b*-tag categories are considered when searching for the signal in events with four or more jets because $t\bar{t}$ production dominates the small amount of signal present in higher *b*-tag categories.

The expected number of events from each signal and background category is compared to the observed data for each *b*-jet identification category for events with two jets, three jets, and four or more jets in Tables I, II, and III, respectively. Selected kinematic distributions are shown for all selected events in Figs. 3 and 4, and the dijet invariant mass for events with two jets is shown for all *b*-tag categories in Figs. 5 and 6. In all plots, data points are shown with error bars that reflect the statistical uncertainty only. Discrepancies in data-MC agreement are within our systematic uncertainties described in Sec. X.

IX. MULTIVARIATE SIGNAL DISCRIMINANTS

We employ multivariate analysis (MVA) techniques to separate signal from background events. To separate signal from the MJ events, we use a boosted decision tree implemented with the TMVA package [33]. This multivariate analysis is described in Sec. IX A. A BDT is also used to separate signal from other specific background sources in events with four or more jets (see Sec. IX D). For the final multivariate analysis, we use a BDT in the one tight *b*-tag channel and all three two-*b*-tag channels, and we use a random forest decision tree (RF) [62] implemented in the
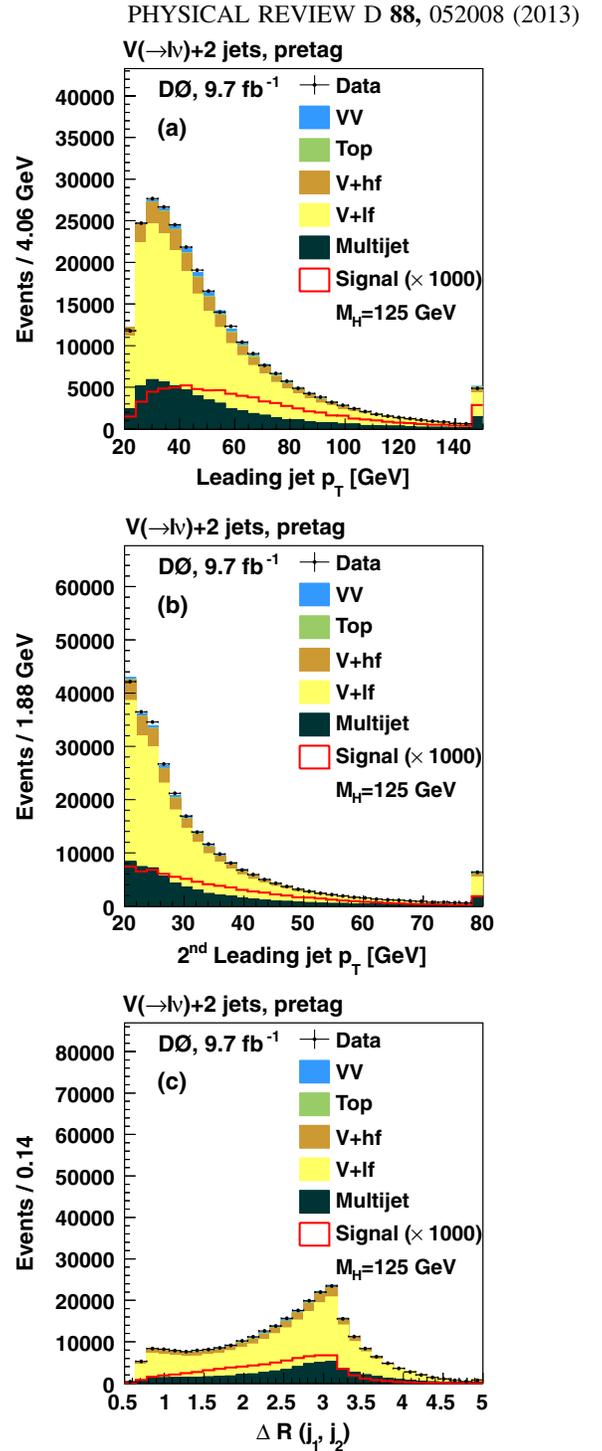

FIG. 4 (color online). Distributions for all selected events with two jets of (a) leading jet p_T, (b) second-leading jet p_T, and (c) $\Delta\mathcal{R}$ between the leading and second-leading jets. The signal is multiplied by 1000. Overflow events are added to the last bin.

STATPATTERNRECOGNITION package [28,63] for events in the zero and one loose *b*-tag channels.

The BDT and the RF are forms of machine learning techniques known as decision trees. Decision trees operate on a series of yes/no splits on events that are known to
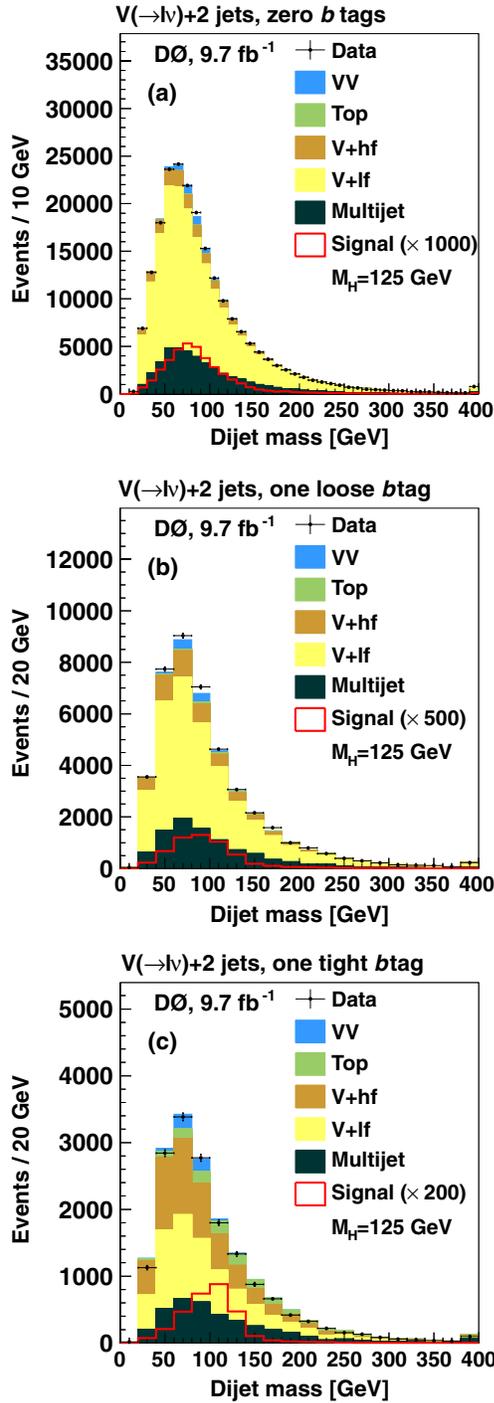
FIG. 5 (color online). Invariant mass of the leading and second-leading jets in events with two jets and (a) zero b tags, (b) one loose b tag, and (c) one tight b tag. The signal is multiplied by 1000, 500, and 200, respectively. Overflow events are added to the last bin.
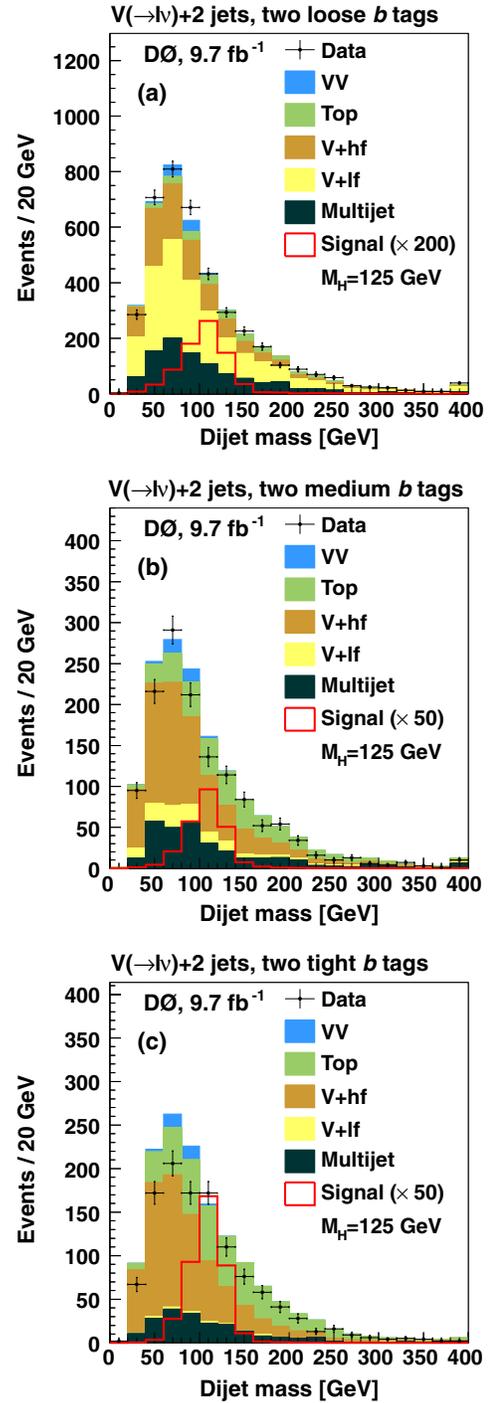
FIG. 6 (color online). Invariant mass of the leading and second-leading jets in events with two jets and (a) two loose b tags, (b) two medium b tags, and (c) two tight b tags. The signal is multiplied by 200, 50, and 50, respectively. Overflow events are added to the last bin.

be classified as either signal or background. The splitting is done to maximally separate signal from background. The resulting nodes are continually split to optimally separate signal from background until either a minimum number of events in a node is reached or the events in a

node are pure signal or pure background. The technique of boosting in the BDT builds up a series of trees where each tree is retrained, boosting the weights for events that are misclassified in the previous training. The RF technique creates a collection of decision trees where each
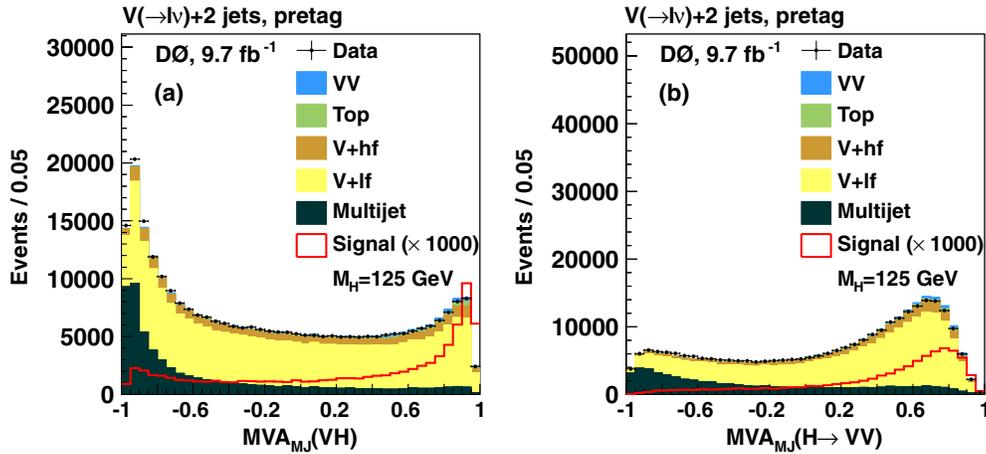
FIG. 7 (color online).   The multivariate discriminant output for (a) $\mathrm{MVA}_{\mathrm{MJ}}(VH)$ and (b) $\mathrm{MVA}_{\mathrm{MJ}}(H \to VV)$, for all events. The signal is multiplied by 1000.

tree is trained on a subset of the training data that is randomly sampled.

We train separate BDTs and RFs for each lepton flavor, jet multiplicity, and tagging category, and for each hypothesized Higgs boson mass in steps of 5 GeV. Since the branching fraction for the Higgs decay to *b* quarks is only significant over the mass range 90–150 GeV, we restrict the search in the one tight and two-*b*-tag channels to this range of M_H. In the zero and one loose *b*-tag channels, the primary signal contribution is from Higgs decays to vector bosons, the search is performed over the mass range of 100–200 GeV.

Each of the final BDTs and RFs are trained to distinguish the signal from all of the backgrounds. We choose variables to train the BDTs and RFs that have good agreement between data and background simulation (since the expected contribution from signal events is small), and so that there is a good separation between signal and at least one background. Background and signal samples are each split into three independent samples for use in training, testing, and performing the final statistical analysis with each multivariate discriminant. We ensure that the discriminant is not biased towards statistical fluctuations in the training sample by comparing the training output to the testing sample. The independent sample used for the limit-setting procedure ensures that any optimizations performed based on the output of the training and testing samples do not bias the final limits.

A. Multivariate multijet discriminators

We train two separate BDTs to separate the MJ background from signal events: one for $VH(\to b\bar{b}, c\bar{c}, \tau\tau)$ signals, $\mathrm{MVA}_{\mathrm{MJ}}(VH)$, and one for $H \to VV$ signals, $\mathrm{MVA}_{\mathrm{MJ}}(H\to VV)$. The variables used in training these BDTs are chosen to exploit kinematic differences between the MJ and signal events, and are documented in Ref. [64].

To improve the training statistics, we combine signal events for $M_H = 120$, 125, and 130 GeV in training. We find that a BDT trained on this combination of Higgs boson masses has a similar performance when applied to other masses, eliminating the need for a mass-dependent MJ discriminant. The BDT outputs $\mathrm{MVA}_{\mathrm{MJ}}(VH)$ and $\mathrm{MVA}_{\mathrm{MJ}}(H\to VV)$ are shown in Fig. 7. The $\mathrm{MVA}_{\mathrm{MJ}}(VH)$ and $\mathrm{MVA}_{\mathrm{MJ}}(H \to VV)$ discriminant outputs are used as input variables to the final MVAs, as detailed in Ref. [64].

B. Final $WH \to \ell\nu b\bar{b}$ MVA analysis

In events with two or three jets and one tight *b* tag or two *b* tags, the $WH \to \ell\nu b\bar{b}$ process provides the dominant signal contribution. To separate signal from background, we train a BDT on the $WH \to \ell\nu b\bar{b}$ signal and all backgrounds. The lists of input variables to the MVA and their descriptions are included in Ref. [64]. Figures 8 and 9 show examples of some of the most effective discriminating variables used in our BDTs for the two-jet and three-jet channels, respectively, in the one tight *b*-tag and all two-*b*-tag channels. Figures 10 and 11 show the BDT output for the two- and three-jet channels, respectively, in the one tight *b*-tag and all the two-*b*-tag channels.

C. Final $H \to WW \to \ell\nu jj$ MVA analysis

The $H \to WW \to \ell\nu jj$ process provides the dominant signal in events with two or three jets and zero *b* tags or one loose *b* tag, since the *W*-boson decays producing a *b* quark are rare. For signal searches in these channels, we apply a multivariate technique based on the RF discriminant. Events in the above tagging categories are examined for $100 \leq M_H \leq 150$ GeV. Since we do not perform the search in the one tight and two-*b*-tag channels for $M_H > 150$ GeV, events having exactly two or three jets in all *b*-tagging categories (i.e. pretag events) are used in the search for $155 \leq M_H \leq 200$ GeV.
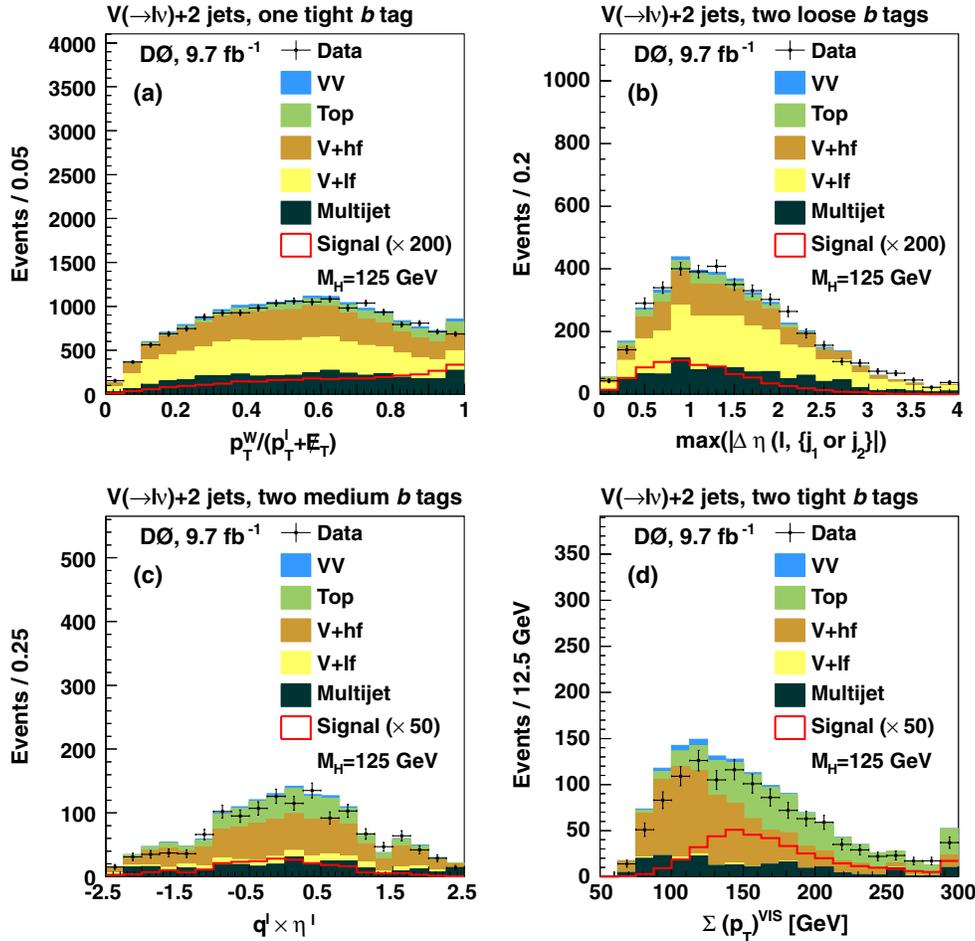
FIG. 8 (color online). Distributions of some of the most significant inputs to the final discriminant in events with exactly two jets and either one tight b tag, two loose b tags, two medium b tags, or two tight b tags: (a) $p_T^W/(p_T^\ell + \not{E}_T)$, shown for events with one tight b tag; (b) $\max|\Delta\eta(\ell, \{j_1 \text{ or } j_2\})|$ [85], shown for events with two loose b tags; (c) $q^\ell \times \eta^\ell$ [86], shown for events with two medium b tags; (d) $\sum(p_T)^{VIS}$ [87], shown for events with two tight b tags. The signal is multiplied by 200, 200, 50, and 50, respectively. Overflow events are added to the last bin.

To suppress MJ background in the electron channel in these subchannels, we select events with $\text{MVA}_{MJ}(H \to VV) > -0.4$ for $M_H \le 150$ GeV in events with zero or one loose tag, and $\text{MVA}_{MJ}(VH) > -0.5$ for $M_H \ge 155$ GeV in all events. These requirements were optimized to maximize the ratio of the number of signal events to the square root of the number of background events. The MJ component in the zero or one loose b-tag muon channel is small, so there is no cut applied to the MJ MVA outputs.

We train an RF on the total signal and background from all considered physics processes. We optimize the RF independently in the electron and muon channels for each b-tag and jet multiplicity category. As the signal shape is strongly driven by the signal mass hypothesis, we optimize the MVA variable list at two different mass points: at $M_H = 125$ GeV for masses below 150 GeV and at $M_H = 165$ GeV for masses above 150 GeV. Because the resolution of the reconstructed Higgs boson mass is about 20 GeV for channels presented in this article, optimizing

the input variable list at only these mass points is sufficient. Each RF is trained using between 14 and 30 well-modeled discriminating variables formed from kinematic properties of either elementary objects like jets or leptons, or composite objects, such as reconstructed W-boson candidates (see Figs. 12 and 13). The lists of input variables and their descriptions are included in Ref. [64]. The final RF discriminants for the electron and muon channels are shown in Figs. 14 and 15.

D. Final $VH \to VWW \to \ell\nu jjjj$ MVA analysis

The majority of signal events with four or more jets and zero b tags or one loose b tag are from the $VH \to VWW \to \ell\nu jjjj$ process, but there are significant contributions from direct production via gluon fusion and vector-boson fusion. Identification of the Higgs boson decay products in $VH \to VWW$ events is complicated by the combinatorics of pairing four jets into two hadronically decaying vector-boson candidates and then two of the three
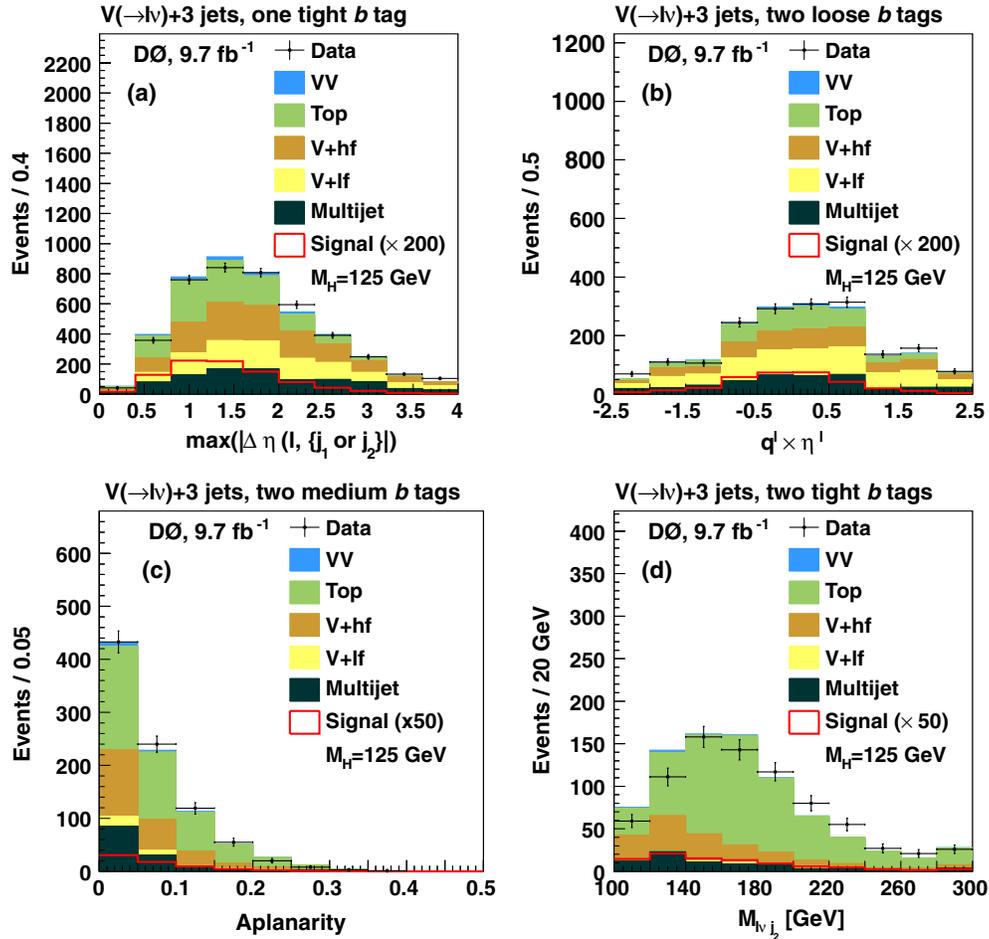
FIG. 9 (color online).   Distributions of some of the most significant inputs to the final discriminant in events with exactly three jets and either one tight *b* tag, two loose *b* tags, two medium *b* tags, or two tight *b* tags: (a) max $|\Delta\eta(\ell, \{j_1 \text{ or } j_2\})|$ [85], shown for events with one tight *b* tag; (b) $q^\ell \times \eta^\ell$ [86], shown for events with two loose *b* tags; (c) aplanarity [88], shown for events with two medium *b* tags; (d) $m_{\ell\nu j_2}$ [89,90], shown for events with two tight *b* tags. The signal is multiplied by 200, 200, 50, and 50, respectively. Overflow events are added to the last bin.

total vector-boson candidates into the Higgs boson candidate. The discriminating variables are different for fully hadronic and semileptonic Higgs boson decays, and determining the Higgs boson candidate for an event also determines which of these two decay scenarios is considered. Variables unique to a particular decay scenario are set to a default value outside of the physical range of that variable in events reconstructed under the alternate decay scenario. To reconstruct the two hadronically decaying vector-boson candidates, we examine the leading four jets in an event and choose the jet pairings that minimize

$$E_{ab,cd} = |m_{ab} - M_W| + |m_{cd} - M_W|, \qquad (4)$$

where m_{ab} (m_{cd}) is the invariant mass of the *a*th and *b*th (*c*th and *d*th) jets, and $M_W = 80.4$ GeV [65]. The Higgs boson candidate is then determined by considering the semileptonically decaying *W* boson and the two hadronically decaying vector bosons and selecting the

vector-boson candidate pair with the minimum $\Delta\mathcal{R}$ separation in an event, out of the three possible pairings.

Diverse signal processes contribute to the inclusive four-jet channel with relative contributions varying with M_H. To help mitigate the effect of having many signal and background contributions to this search channel, we use two layers of multivariate discriminants to improve the separation of signal from background. The first layer of training focuses on separating the sum of all signal processes from specific sets of backgrounds. Input variables for each background-specific discriminant are selected based on the separation power between the total signal and the backgrounds being considered. Background-specific discriminants are trained to separate the sum of all Higgs boson signal processes from three specific background categories: $t\bar{t}$ and single top-quark production, $V +$ jets production, and diboson production. The input variables and their descriptions are listed in Ref. [64]. Separate background-specific discriminants are trained for each
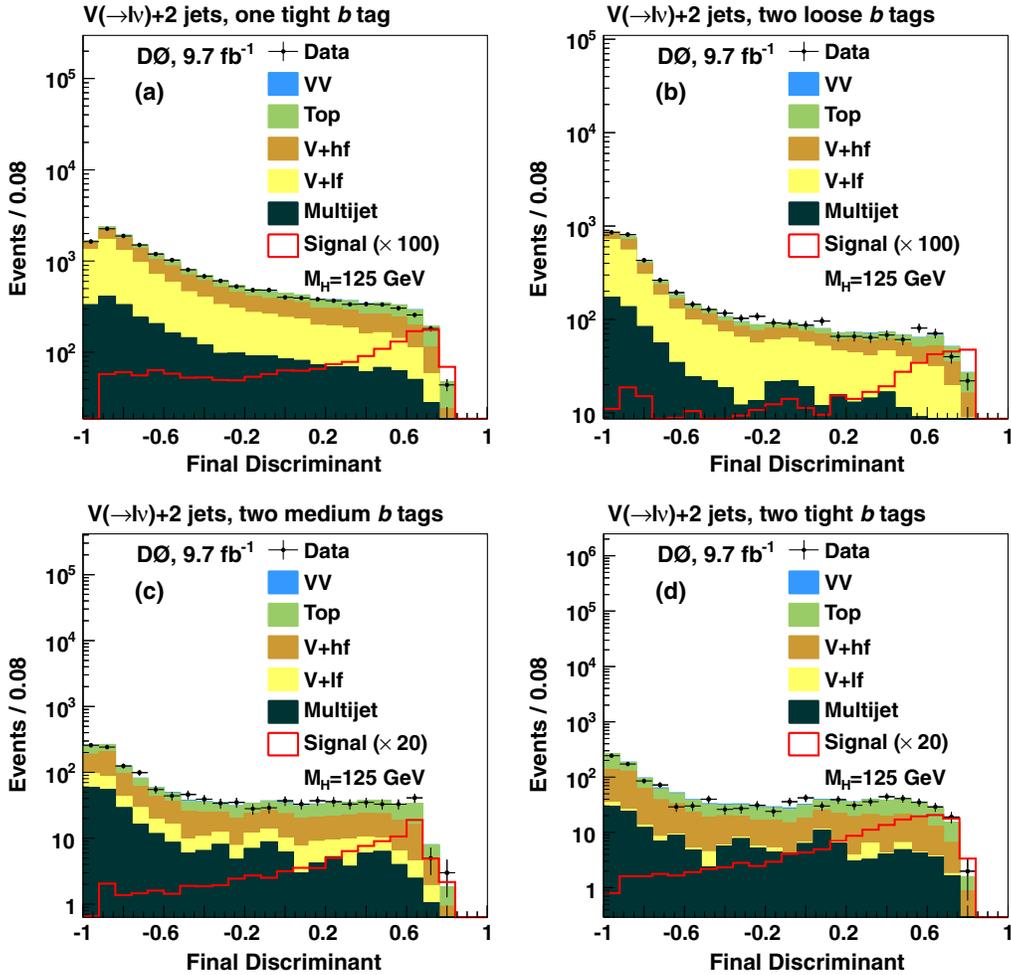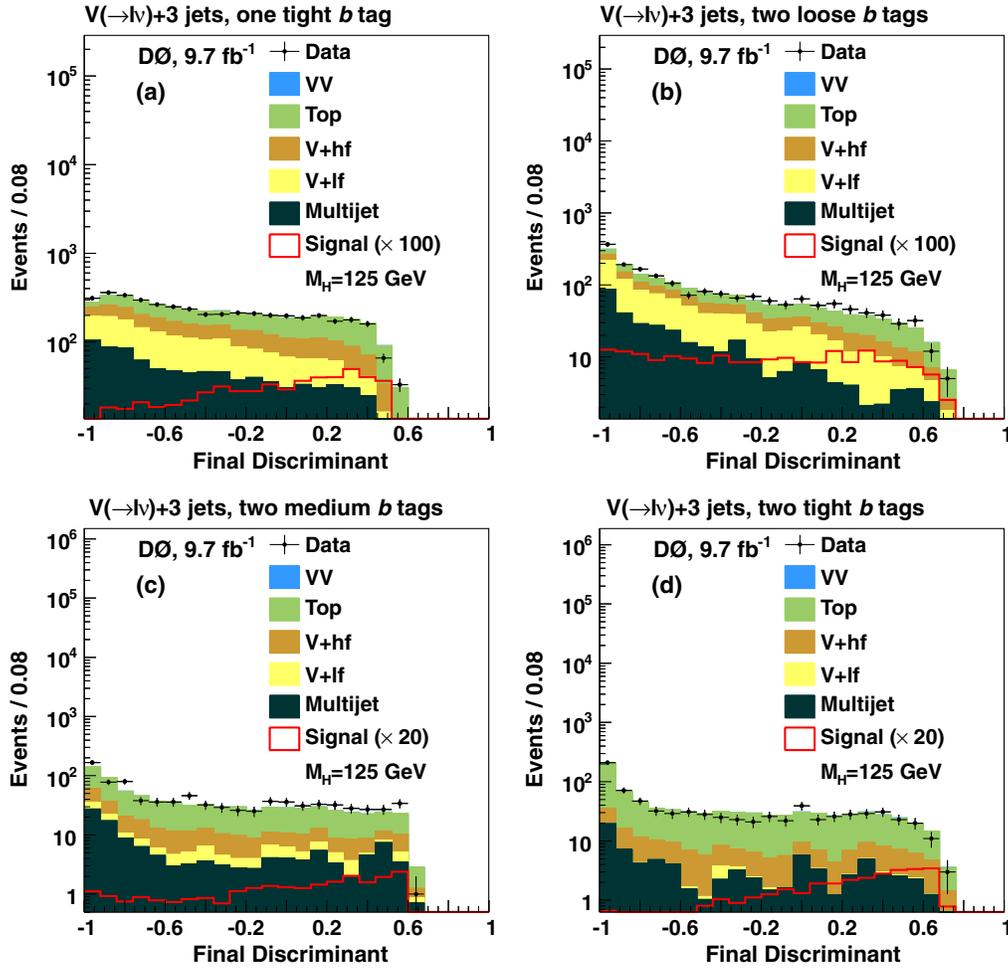
FIG. 10 (color online).   Distributions of the final discriminant output, after the maximum likelihood fit (described in Sec. XII), in events with exactly two jets and (a) one tight b tag, (b) two loose b tags, (c) two medium b tags, and (d) two tight b tags. The signal is multiplied by 100, 100, 20, and 20, respectively.

Higgs boson mass point considered. Sample inputs and output responses of the background-specific discriminants are shown in Figs. 16 and 17, respectively, for $M_H = 125$ GeV.

The background-specific discriminants are used as inputs to the final RF discriminant that is trained to discriminate all signal processes from the total background contributions. Additional input variables for the final discriminant are selected based on their separation power between the total signal and the total background, and are required to be well modeled. The input variables for each lepton and b-tag category are listed in Ref. [64]. Sample inputs and output responses of the final discriminants are shown in Figs. 18 and 19, respectively, for $M_H = 125$ GeV.

X. SYSTEMATIC UNCERTAINTIES

We assess systematic uncertainties on signals and backgrounds for each of the jet multiplicity and b-tag channels by repeating the full analysis after varying each source of uncertainty by ± 1 s.d. We consider uncertainties that affect both the normalizations and the shapes of our MVA outputs.

We include theoretical uncertainties on the $t\bar{t}$ and single top-quark production cross sections (7% [53,56]), the diboson production cross section (6% [54]), $V + lf$ production (6%), and $V + hf$ production (20%, estimated from MCFM [55,59]). Since the $V +$ jets experimental scaling factors for the three- and four-jet channels are different from unity, we apply an additional systematic uncertainty on the $V +$ jets samples that is uncorrelated across jet multiplicity and lepton flavor bins. The size of this uncertainty is taken as the uncertainty from the $V +$ jets fit to data, described in Sec. VII.

An uncertainty on the integrated luminosity (6.1% [66]) affects the normalization of the expected signal and simulated backgrounds. Uncertainties that affect the final MVA distribution shapes include jet taggability (3% per jet), b-tagging efficiency (2.5%–3% per heavy-quark jet), the light-quark jet misidentification rate (10% per jet), jet

FIG. 11 (color online). Distributions of the final discriminant output, after the maximum likelihood fit (described in Sec. XII), in events with exactly three jets and (a) one tight *b* tag, (b) two loose *b* tags, (c) two medium *b* tags, and (d) two tight *b* tags. The signal is multiplied by 100, 100, 20, and 20, respectively.

identification efficiency (5%), and jet energy calibration and resolution (varying between 5% and 15%, depending on the process and channel), as described in Ref. [20]. We also include uncertainties from modeling that affect both the shapes and normalizations of the final MVA distributions. These include an uncertainty on the trigger efficiency in the muon channel as derived from the data (3%–5%), lepton identification and reconstruction efficiency (5%–6%), the MLM matching [40] applied to V + light-flavor events ($\approx 0.5\%$), the ALPGEN renormalization and factorization scales, and the choice of parton distribution functions (2%) as described in Ref. [20]. The trigger uncertainty in the muon channel is calculated as the difference between applying a trigger correction calculated using the ALPGEN reweightings derived on the $T_{\mu\mathrm{OR}}$ trigger sample and applying the nominal trigger correction. Since we reweight our ALPGEN samples, we include separate uncertainties on each of the five functions used to apply the reweighting. The adjusted functions are calculated by shifting the parameter responsible for the largest shape

variation of the fit by ± 1 s.d. then calculating the remaining parameters for the function using the covariance matrix obtained from the functional fit.

We determine the uncertainty on the MJ background shape by relaxing the requirement from Sec. VIII on M_T^W to $M_T^W > 30$ GeV $- 0.5 \times \not{E}_T$ and repeating the analysis with this selection in place. The positive and negative variations are taken to be symmetric. The uncertainty in the MJ rate is 15% (20%) for the electron (muon) channel. Since our MJ sample is statistically limited, we do not correlate the uncertainties in the rate and shape across the subchannels. Since we simultaneously fit MJ and V + jets to match data, we apply a normalization uncertainty to the V + jets samples that is anticorrelated with the MJ normalization systematics and scales as the relative MJ to V + jets normalization.

XI. *WZ* AND *ZZ* PRODUCTION WITH $Z \to b\bar{b}$

The SM processes $W(\to \ell\nu)Z(\to b\bar{b})$ and $Z(\to \ell\ell) \times Z(\to b\bar{b})$ where one of the leptons from the $Z \to \ell\ell$ decay
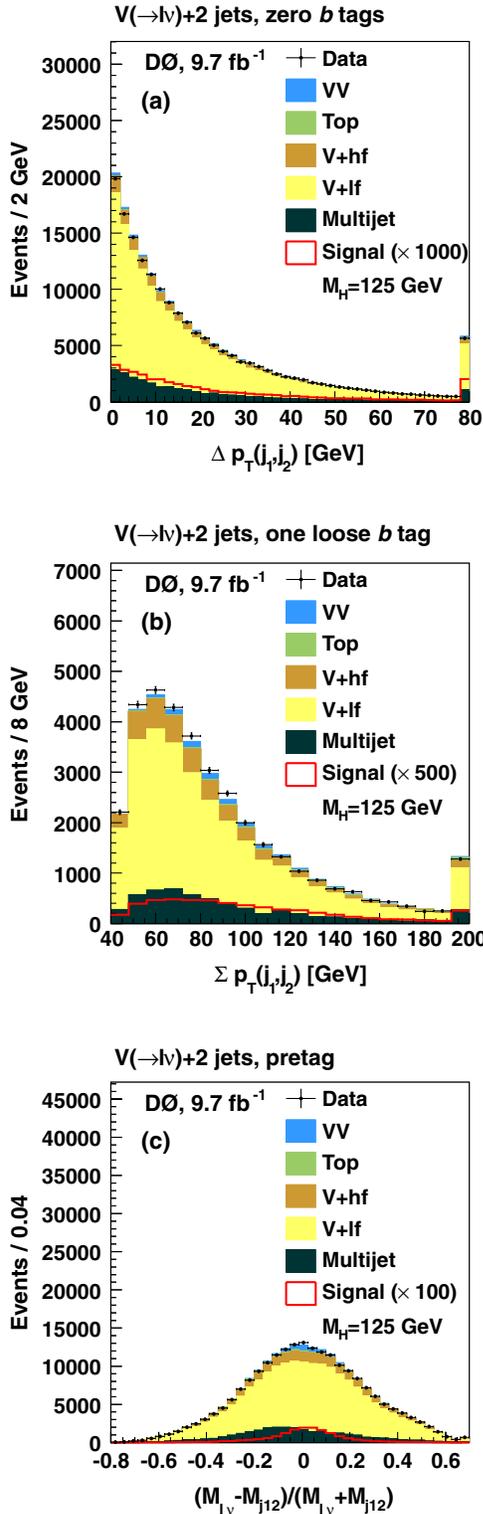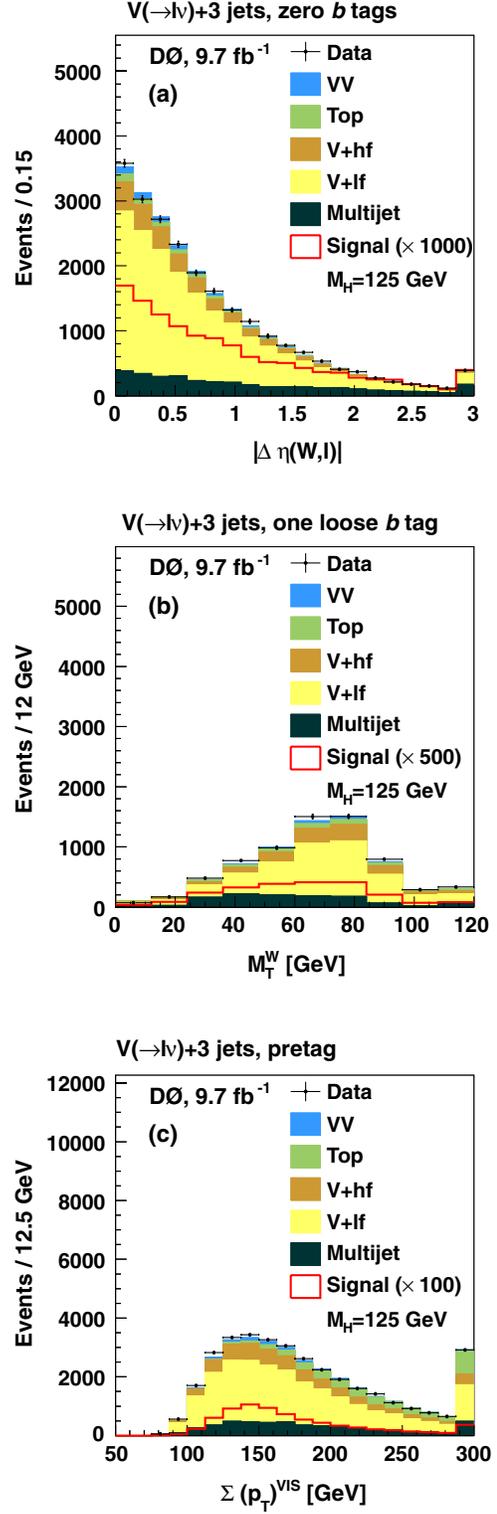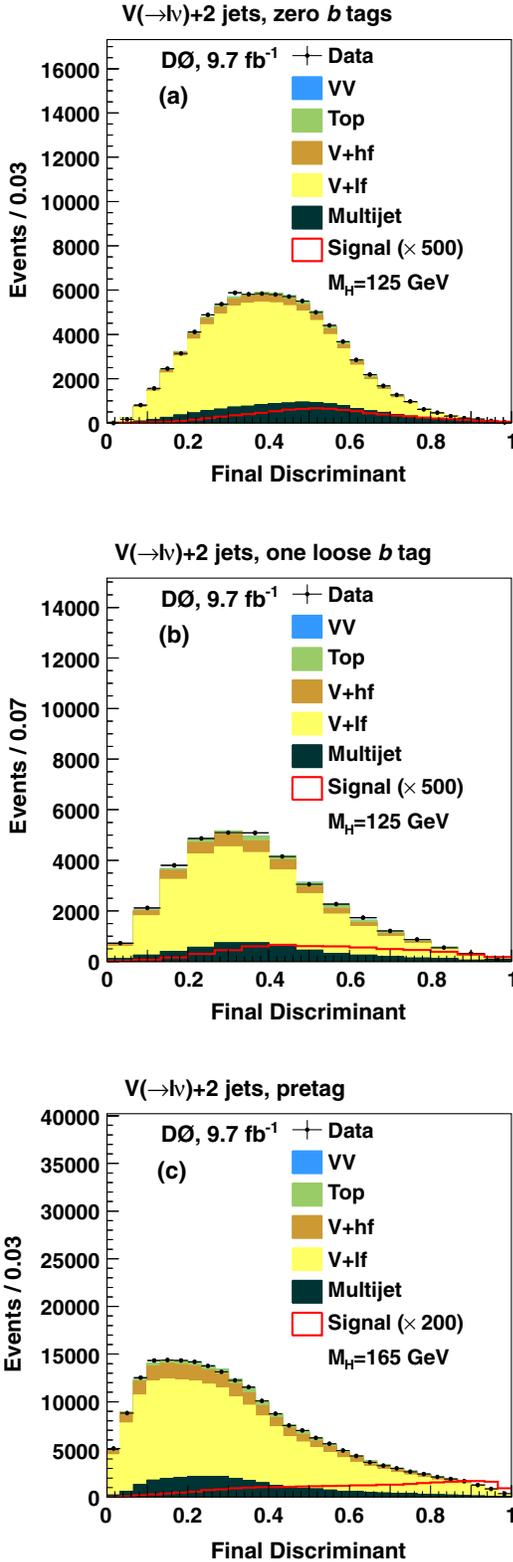
FIG. 12 (color online). Distributions of the most significant inputs to the final multivariate discriminants for the two-jets zero and one loose b-tag channels: (a) $\Delta p_T(j_1, j_2)$, shown for events with zero b tags for $M_H = 125$ GeV; (b) $\sum_{i=1}^{2} p_T^{j_i}$, shown for events with one loose b tag for $M_H = 125$ GeV; (c) $(M_{\ell\nu} - M_{j_{12}})/(M_{\ell\nu} + M_{j_{12}})$, shown for all tags for $M_H = 165$ GeV. The signal is multiplied by 1000, 500, and 100, respectively. Overflow events are added to the last bin.

FIG. 13 (color online). Distributions of the most significant inputs to the final multivariate discriminants for the three-jets zero and one loose b-tag channels: (a) $|\Delta\eta(W, \ell)|$ [90,91], shown for events with zero b tags for $M_H = 125$ GeV; (b) M_T^W, shown for events with one loose b tag for $M_H = 125$ GeV; (c) $\sum(p_T)^{\mathrm{VIS}}$ [87], shown for all tags for $M_H = 165$ GeV. The signal is multiplied by 1000, 500, and 100, respectively. Overflow events are added to the last bin.
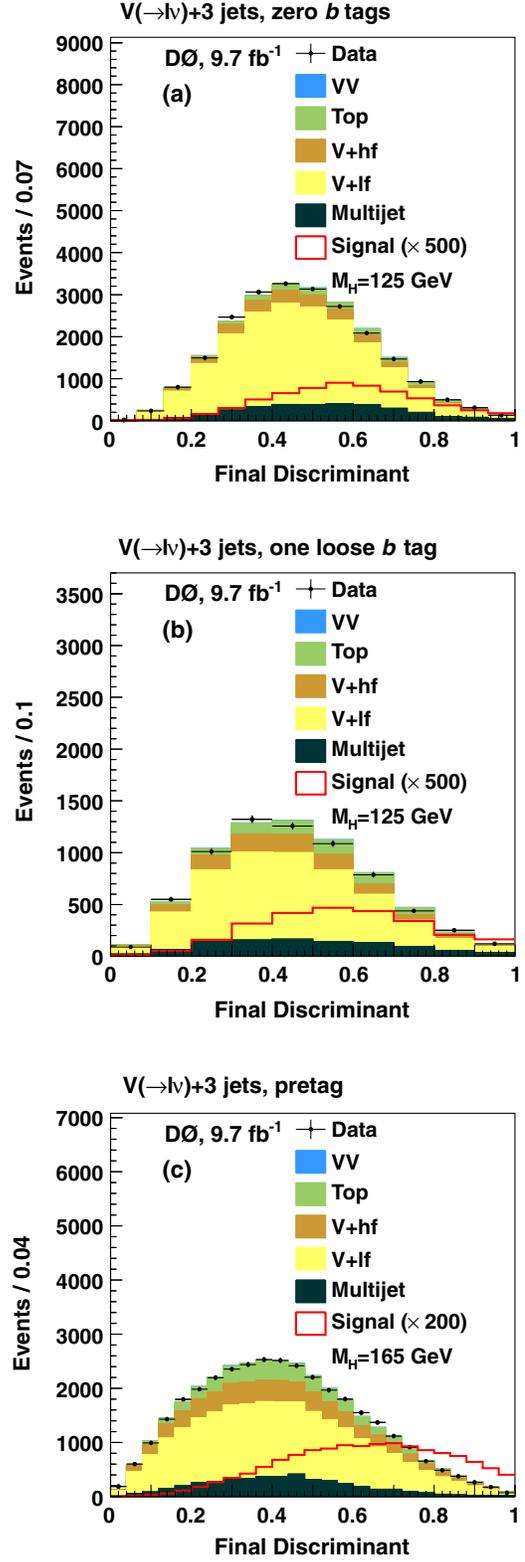
FIG. 14 (color online). Distributions of the final discriminant output, after the maximum likelihood fit (described in Sec. XII), for events in the following channels: (a) two jets, zero *b* tags for $M_H = 125$ GeV, (b) two jets, one loose *b* tag for $M_H = 125$ GeV, and (c) two jets, all tags $M_H = 165$ GeV. The signal is multiplied by 500, 500, and 200, respectively.
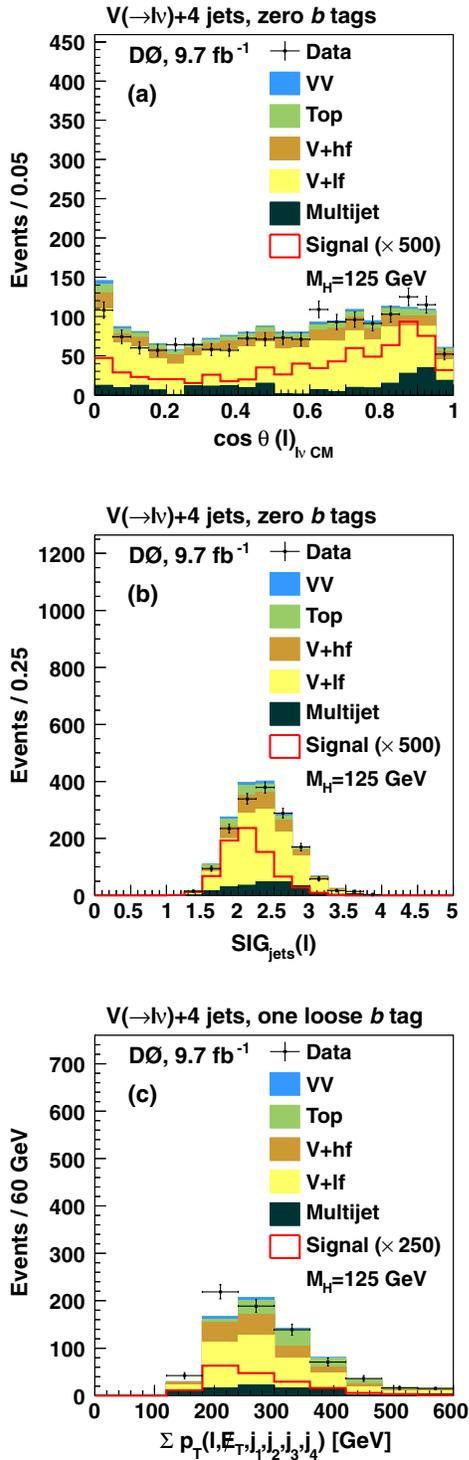
FIG. 15 (color online). Distributions of the final discriminant output, after the maximum likelihood fit (described in Sec. XII), for events in the following channels: (a) three jets, zero *b* tags for $M_H = 125$ GeV, (b) three jets, one loose *b* tag for $M_H = 125$ GeV, and (c) three jets, all tags $M_H = 165$ GeV. The signal is multiplied by 500, 500, and 200, respectively.
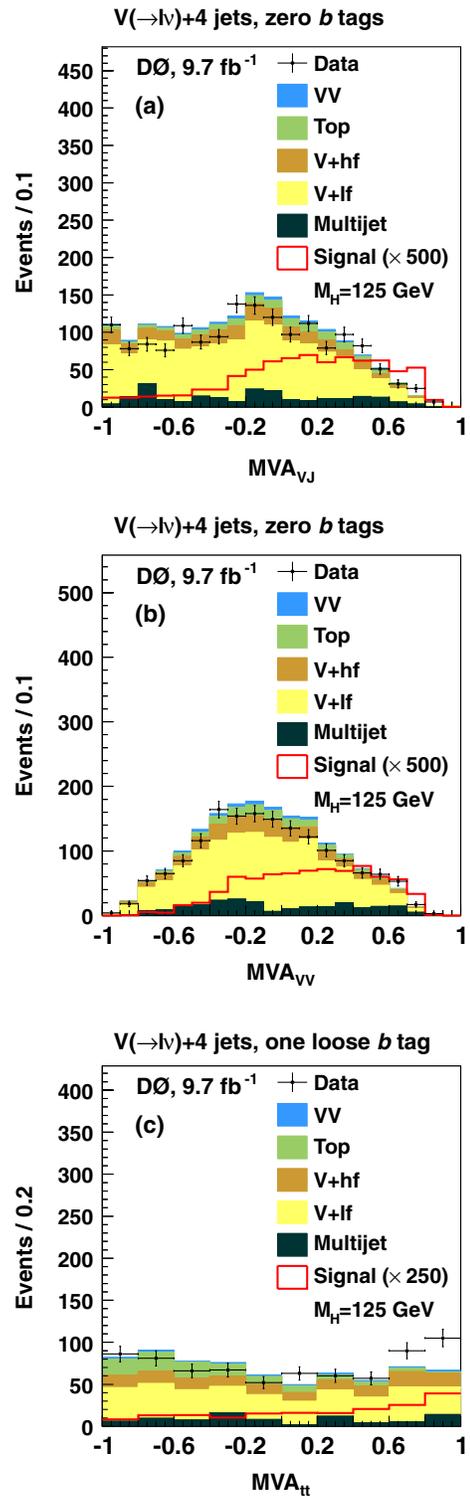
FIG. 16 (color online). Distributions of the most significant inputs to background-specific multivariate discriminants for the ≥ four-jet subchannels: (a) $\cos\theta(\ell)_{\ell\nu\,\mathrm{CM}}$ [90,92], input to discriminant against $V +$ jets backgrounds, shown for events with zero b tags; (b) $\mathcal{SIG}_{\mathrm{jets}}(\ell)$ [93], input to discriminant against diboson backgrounds, shown for events with zero b tags; (c) $\sum p_T(\ell, \not{E}_T, j_1, j_2, j_3, j_4)$, input to discriminant against top-quark backgrounds, shown for events with one loose b tag. The $M_H = 125$ GeV signal is multiplied by 250 in (c) and by 500 in (a) and (b). Overflow events are added to the last bin.

FIG. 17 (color online). Distributions of the output of background-specific multivariate discriminants for the ≥ four-jet subchannels: (a) discriminant against $V +$ jets backgrounds, shown for events with zero b tags; (b) discriminant against diboson backgrounds, shown for events with zero b tags; (c) discriminant against top-quark backgrounds, shown for events with one loose b tag. The $M_H = 125$ GeV signal is multiplied by 250 in (c) and by 500 in (a) and (b).
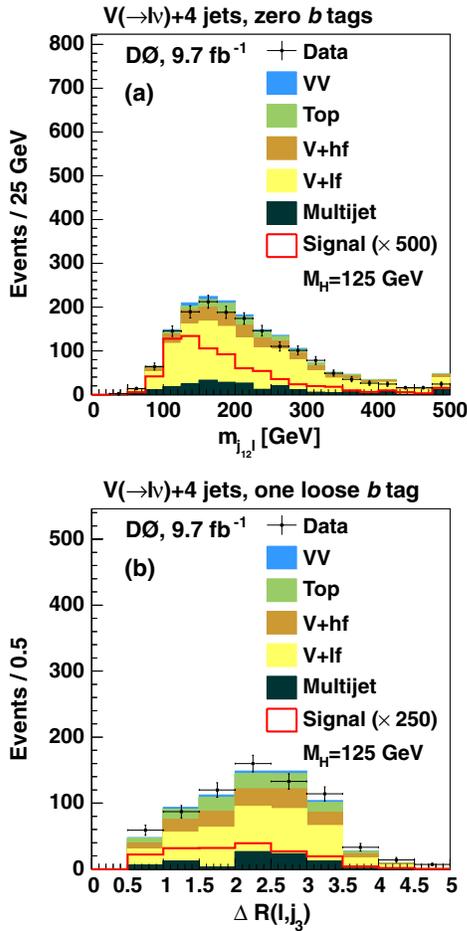
FIG. 18 (color online). Distributions of the most significant inputs, other than background-specific multivariate discriminants, to the final multivariate discriminants for the $\geq$ four-jet subchannels: (a) $m_{j_{12}\ell}$, shown for events with zero b tags; (b) $\Delta\mathcal{R}(\ell, j_3)$, shown for events with one loose b tag. The $M_H =$ 125 GeV signal is multiplied by 500 in (a) and 250 in (b). Overflow events are added to the last bin.
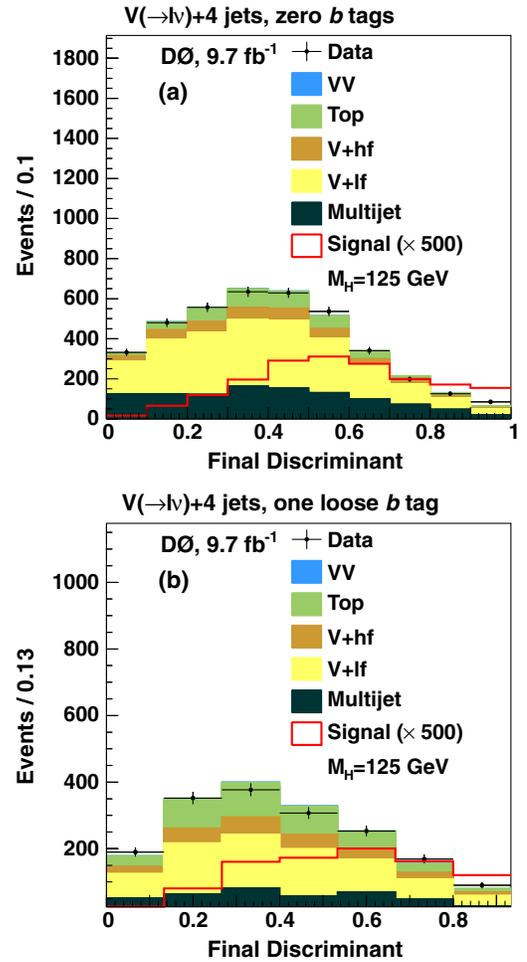

FIG. 19 (color online). Distributions of the final discriminant output, after the maximum likelihood fit (described in Sec. XII), at $M_H = 125$ GeV for the four-or-more-jets channels with (a) zero b tags and (b) one loose b tag. The $M_H = 125$ GeV signal is multiplied by 500.

is not reconstructed result in the same final-state signature as the Higgs boson in this search. Therefore, we search for these processes to validate our analysis methodology. The only change in the analysis is in the training of the final discriminant in events with two or three jets with one tight b tag or two b tags. We train using the WZ and ZZ diboson processes as signal while leaving the WW process as a background. The output of this discriminant is used to measure the combined WZ and ZZ cross section by performing a maximum likelihood fit to data using signal-plus-background models, with maximization over the systematic uncertainties as described in detail in Sec. XII. The expected significance of the measurement using the MVA output is 1.8 s.d. We measure a cross section of 0.50 ± 0.34(stat) $\pm$ 0.36(syst) times the expected SM cross section of $4.4 \pm$ 0.3 pb. Figure 20 shows the MVA discriminant output for the diboson cross section ($WZ + ZZ$) with background-subtracted data and signal scaled to the best-fit value.

XII. UPPER LIMITS ON THE HIGGS BOSON PRODUCTION CROSS SECTION

We derive upper limits on the Higgs boson production cross section multiplied by the corresponding branching fraction in units of the SM prediction. The limits are calculated using the modified frequentist CL_s approach [67–69], and the procedure is repeated for each assumed value of M_H.

Two hypotheses are considered: the background-only hypothesis (B), in which only background contributions are present, and the signal-plus-background (S + B) hypothesis, in which both signal and background contributions are present.

The limits are determined using the MVA output distributions, together with their associated uncertainties, as inputs to the limit-setting procedure. To preserve the stability of the limit derivation procedure in regions of small background statistics in the one tight b-tag and all
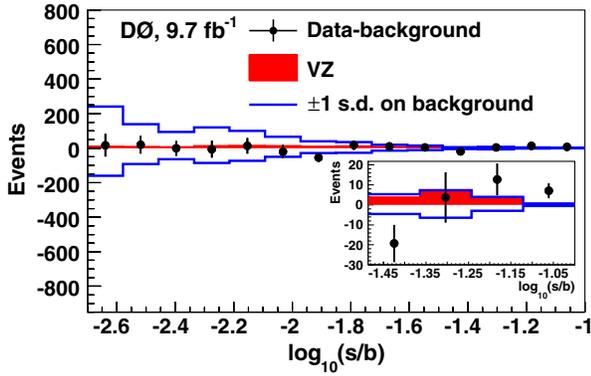
FIG. 20 (color online). Final MVA discriminant output shown for the expected diboson signal and background-subtracted data rebinned as a function of $\log(S/B)$, after the maximum likelihood fit, summed over b-tag channels. The error bars on data points represent the statistical uncertainty only. The post-fit systematic uncertainties are represented by the solid lines. The signal expectation is shown scaled to the best fit value. The inset gives an expanded view of the high-$\log(S/B)$ region.

two-b-tag categories, the width of the bin at the largest MVA output value is adjusted by comparing the total background and signal + background expectations until the statistical significances for B and S + B are, respectively,

greater than 3.6 and 5.0 s.d. from zero. The remaining part of the distribution is then divided into equally sized bins. In the zero b-tags and one loose b-tag categories, the width of the bin at largest MVA output is set such that the relative statistical uncertainty on the signals-plus-background entries is less than 0.15. The remaining bins are distributed uniformly. The rebinning procedure is checked for potential biases in the determination of the final limits, and no such bias is observed.

We evaluate the compatibility of the data with the background-only and signal + background hypotheses. This is done using the log-likelihood ratio (LLR), which is twice the negative logarithm of the ratio of the Poisson likelihoods, L, of the signal + background hypothesis to the background only hypothesis, $LLR = -2\ln(L_{S+B}/L_B)$.

Systematic uncertainties are included through nuisance parameters that are assigned Gaussian probability distributions (priors). The signal and background predictions are functions of the nuisance parameters. Each common source of systematic uncertainty (such as the uncertainties on predicted SM cross sections, identification efficiencies, and energy calibration, as described in Sec. X) is taken to be correlated across all channels except as otherwise noted in Sec. X.
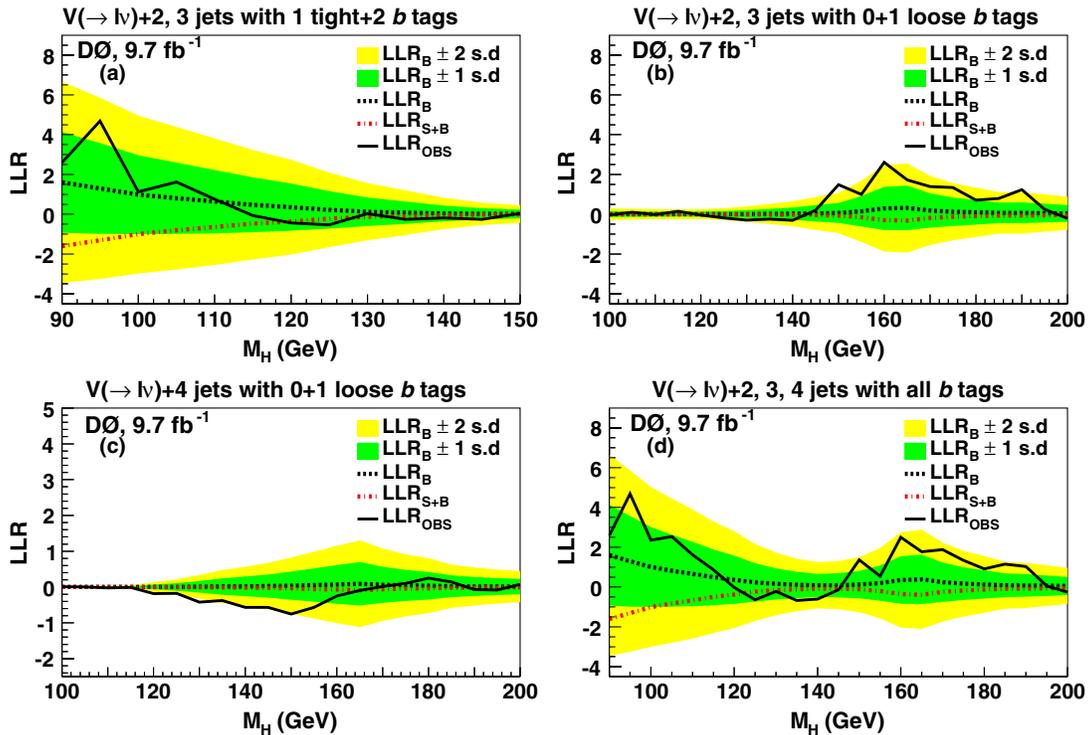

FIG. 21 (color online). The expected and observed log-likelihood ratios as functions of the hypothesized Higgs boson mass M_H for the (a) electron and muon, two- and three-jet, one tight and two-b-tag channels; (b) electron and muon, two- and three-jet, zero and one loose b-tag channels; (c) electron and muon, four-or-more-jets, zero and one loose b-tag channels; (d) combination of all channels. The dashed red and black lines correspond to the median LLR of the signal + background and background-only hypotheses, respectively. The solid line corresponds to the LLR obtained from the data, and the shaded regions are the ±1 s.d. and ±2 s.d. values for the background-only hypothesis.
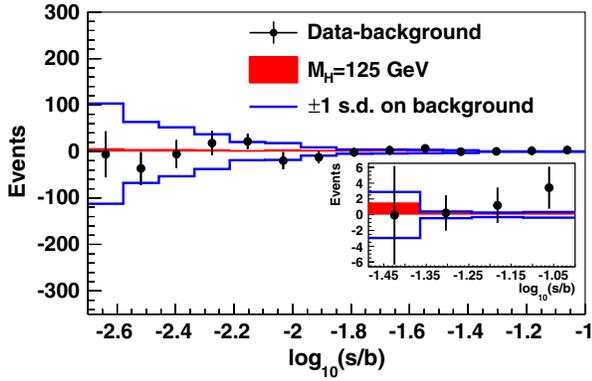
FIG. 22 (color online). The MVA discriminant output distribution minus the total background expectation for $M_H = 125$ GeV rebinned as a function of $\log(S/B)$. The post-fit uncertainties are represented by the solid lines. The signal expectation is shown scaled to the best-fit value. The inset gives an expanded view of the high-$\log(S/B)$ region.

The inclusion of systematic uncertainties in the generation of pseudoexperiments has the effect of broadening the expected LLR distributions and, thus, reducing the ability to resolve signal-like excesses. This degradation can be partially reduced by performing a maximum likelihood fit to each pseudoexperiment (and data), once for the B hypothesis and once for the S + B hypothesis. The maximization is performed over the systematic uncertainties. The LLR is evaluated for each outcome using the ratio of maximum likelihoods for the fit to each hypothesis. The resulting degradation of the limits due to systematic uncertainties is ~30% for searches in the vicinity of $M_H = 125$ GeV.

The medians of the obtained LLR distributions for the B and S + B hypotheses for each tested mass are presented in Fig. 21. The corresponding ± 1 s.d. and ± 2 s.d. values for the background-only hypothesis at each mass point are represented by the shaded regions in the figure. The LLR values obtained from the data are also presented in the figure.

The MVA discriminant distributions, for the Higgs boson mass point $M_H = 125$ GeV, after subtracting the total posterior background expectation are shown in Fig. 22. The signal expectation is shown scaled to the observed upper limit (described later) and the uncertainties in the background after the constrained fit are shown by the solid lines.

Upper limits are calculated at 23 discrete values of the Higgs boson mass, spanning the range 90–200 GeV and spaced in increments of 5 GeV, by scaling the expected signal contribution to the value at which it can be excluded
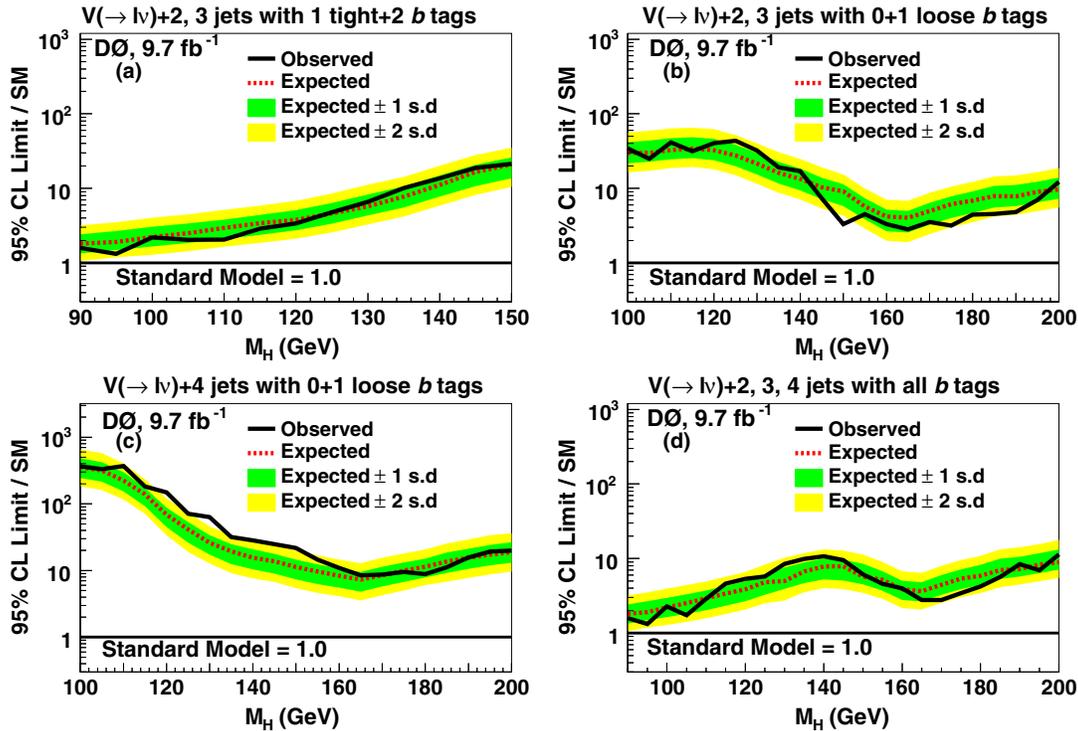

FIG. 23 (color online). The expected and observed 95% C.L. upper limits on SM Higgs boson production for the (a) electron and muon, two- and three-jet, one tight and two-*b*-tag channels; (b) electron and muon, two- and three-jet, zero and one loose *b*-tag channels ($M_H \leq 150$ GeV) and pretag channels ($M_H \geq 155$ GeV); (c) electron and muon, four-or-more-jets, zero and one loose *b*-tag channels; (d) combination of all channels. The limits are presented as ratios to the expected SM prediction. The dashed line corresponds to the expected limit, and the solid line corresponds to the limit observed in data. The shaded regions are the ± 1 s.d. and ± 2 s.d. values for the expected limit.

TABLE IV. The expected and observed 95% C.L. limits, as a function of the Higgs boson mass M_H, presented as ratios of production cross section times branching fraction to the expected SM prediction.

M_H (GeV)	Combined 95%C.L./$\sigma_{\rm SM}$																						
	90	95	100	105	110	115	120	125	130	135	140	145	150	155	160	165	170	175	180	185	190	195	200
Two or three jets with one tight b tag or two b tags																							
Expected	1.8	1.9	2.2	2.5	2.9	3.4	3.8	4.7	5.8	7.9	11.1	16.7	20.8	···	···	···	···	···	···	···	···	···	···
Observed	1.6	1.3	2.2	2.0	2.1	2.9	3.4	4.8	6.6	10.1	13.6	18.8	18.5	···	···	···	···	···	···	···	···	···	···
Two or three jets with zero b tags or one loose b tag																							
Expected	···	···	29.8	30.0	32.6	34.0	32.5	27.5	21.6	16.2	13.3	10.3	9.1	5.7	4.2	4.0	5.0	6.1	6.8	7.9	7.8	9.0	9.7
Observed	···	···	34.4	24.9	41.4	31.4	40.3	43.5	32.3	19.1	17.0	7.3	3.3	4.5	3.3	2.8	3.5	3.2	4.4	4.5	4.8	7.0	12.2
Four or more jets with zero b tags or one loose b tag																							
Expected	···	···	357	316	224	139	68.6	41.2	26.2	19.4	15.5	13.7	11.3	9.7	8.3	7.3	8.5	10.0	11.4	13.7	15.6	17.3	18.8
Observed	···	···	365	331	369	182	149	71.2	63.4	31.8	28.3	24.9	21.9	14.6	10.9	8.5	8.7	9.5	8.8	11.2	15.7	19.2	19.8
All channels combined																							
Expected	1.8	1.9	2.2	2.5	2.9	3.4	3.8	4.7	5.0	6.7	7.8	7.9	5.7	5.2	3.8	3.7	4.4	5.4	5.9	7.0	7.2	8.3	8.9
Observed	1.6	1.3	2.3	1.7	2.9	4.6	5.3	5.8	8.5	9.9	10.7	9.6	6.1	4.6	4.0	2.8	2.8	3.4	4.2	5.7	8.4	6.9	11.4

at the 95% C.L. The expected limits are calculated from the background-only LLR distribution whereas the observed limits are quoted with respect to the LLR values measured in data. The expected and observed 95% C.L. upper limits results for the Higgs boson production cross section multiplied by the decay branching fraction are shown, as a function of the Higgs boson mass M_H, in units of the SM prediction in Fig. 23. The values obtained for the expected and observed limit to SM ratios at each mass point are listed in Table IV for all one-tight, two-loose, two-medium, and two-tight b-tag subchannels together, for the two-jet and three-jet, zero and one loose b-tag subchannels (all b-tag categories for $M_H > 150$ GeV) together, the $\geq$ four-jet subchannels, and the combination of all subchannels.

XIII. INTERPRETATIONS IN FOURTH GENERATION AND FERMIOPHOBIC HIGGS MODELS

Extensions of the minimal electroweak symmetry-breaking mechanism of the SM may be allowed, including models with a fourth generation of fermions or with a Higgs boson that has modified couplings to fermions, as in fermiophobic Higgs models (FHM). We interpret our results in these scenarios using the subchannels that are sensitive to $H \rightarrow WW$ decays: events with two or more jets and zero or one loose b tag for $M_H \leq 150$ GeV, extended to include pretag two- and three-jet events for $M_H \geq 155$ GeV. These are the first results for these models in the $\ell\nu +$ jets final state.

Previous results from the Tevatron Collider experiments in the context of a fourth generation of fermions set a limit on the M_H of $131 < M_H < 207$ GeV [70]. The ATLAS [71] and CMS [72] collaborations exclude $140 < M_H <$ 185 GeV and $144 < M_H < 207$ GeV, respectively. Previous searches for the fermiophobic Higgs boson in $H \rightarrow \gamma\gamma$ and $H \rightarrow VV$ channels, with two leptons in the final state, were

carried out at the LEP e^+e^- Collider [73–76], by the CDF [77] and D0 [78] Collaborations, and by the ATLAS [79] and CMS [80] Collaborations, with the most stringent limits being set by the CMS experiment where the excluded range is $110 < M_H < 194$ GeV.

The Hgg coupling is enhanced in fourth generation models, which leads to a higher rate of $gg \rightarrow H$ production and a larger decay width of $H \rightarrow gg$ than in the SM [81–84]. However, since $H \rightarrow gg$ is loop mediated, the $H \rightarrow WW^*$ decay mode dominates for $M_H > 135$ GeV, as in the SM. We consider two scenarios for the presence of a fourth generation. In the "low-mass" scenario, we assume a fourth generation neutrino mass of $m_{\nu 4} = 80$ GeV and a value for the fourth generation charged lepton mass of $m_{\ell 4} = 100$ GeV, while in the "high-mass" scenario, we assume values for the fourth generation neutrino and lepton masses of $m_{\nu 4} = m_{\ell 4} = 1$ TeV. Both scenarios set the fourth generation quark masses to the values in Ref. [84]. After applying our selection criteria, the total expected signal for $gg \rightarrow H$ production in the low-mass (high-mass) fourth generation model is enhanced by a factor of 7.2 (7.5) over the SM production rate for
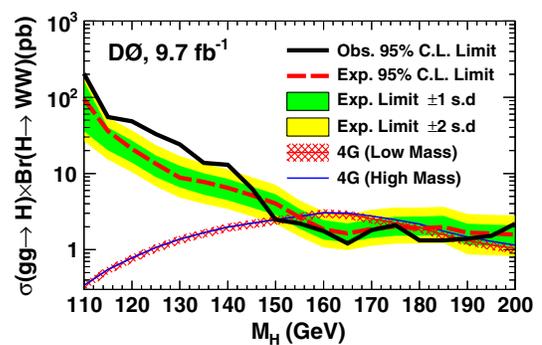

FIG. 24 (color online). The expected and observed 95% C.L. upper limits on $\sigma(gg \rightarrow H) \times \mathcal{B}(H \rightarrow WW)$ compared to the prediction from the fourth generation fermion model.
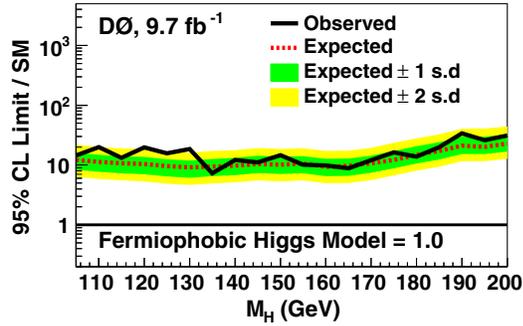
FIG. 25 (color online).   The expected and observed 95% C.L. upper limits on fermiophobic Higgs boson production.

$M_H = 125$ GeV. We only consider gluon-fusion Higgs boson production, and we set limits on $\sigma(gg \to H) \times \mathcal{B}(H \to WW^*)$. These limits are compared with the predicted $gg \to H$ production cross section results from HDECAY [51], as shown in Fig. 24. We exclude the "low-mass" scenario for $150 < M_H < 188$ GeV, and the "high-mass" scenario for $150 < M_H < 190$ GeV.

In the FHM, the Higgs boson does not couple to fermions at tree level but is otherwise SM-like. This suppresses production via gluon fusion to a negligible rate and forbids direct decay to fermions. Production in association with a vector boson or via vector-boson fusion is allowed. For this interpretation, we set the contribution from $gg \to H$ production to zero and scale the contributions from other production and decay mechanisms to reflect the predicted rate in the FHM. After applying our selection criteria, the total expected signal for vector-boson fusion and $VH \to VWW$ production in the FHM is enhanced by a factor of 4.2 over the SM production rate for $M_H = 125$ GeV. The expected and observed cross section times branching fraction limits are compared to the FHM predictions in Fig. 25.

XIV. SUMMARY

We have presented a search for SM Higgs boson production in lepton $+ \not{E}_T +$ jets final states with a data set corresponding to 9.7 fb^{-1} of integrated luminosity collected with the D0 detector. The search is sensitive to $VH \to Vb\bar{b}$, $H \to WW^* \to \ell\nu jj$, and $WH \to WWW^* \to \ell\nu jjjj$ production and decay, and supersedes previous $VH \to Vb\bar{b}$ and $H \to WW^* \to \ell\nu jj$ searches published by D0. To maximize our signal sensitivity, we subdivided the data set into 36 independent subchannels according to lepton flavor, jet multiplicity, and the number and quality of b-tagged jets and applied multivariate analysis techniques to further discriminate between signal and background. We tested our method by examining SM WZ and ZZ production with $Z \to b\bar{b}$ decay and found production rates consistent with the SM prediction. We observed no significant excess over the background prediction as expected from the amplitude of a 125 GeV SM Higgs boson signal, given the sensitivity of this single channel. Significance is achieved by combining this channel with the other low-mass channels analyzed at the Tevatron [13], while here we set 95% C.L. upper limits on the Higgs boson production cross section for masses between 90 and 200 GeV. For $M_H = 125$ GeV, the observed (expected) upper limit is 5.8 (4.7) times the SM prediction. We also interpreted the data in models with fourth generation fermions, or a fermiophobic Higgs boson. In these interpretations, we excluded $150 < M_H < 188(190)$ GeV in the "low-mass" ("high-mass") fourth generation fermion scenario, and provided 95% C.L. limits on the production cross section in the fermiophobic model.

[1] F. Englert and R. Brout, Phys. Rev. Lett. **13**, 321 (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] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **108**, 151803 (2012).
[5] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. Lett. **108**, 151804 (2012).
[6] LEP Electroweak Working Group, http://lepewwg.web.cern.ch/LEPEWWG/.
[7] R. Barate *et al.* (LEP Working Group for Higgs boson searches), Phys. Lett. B **565**, 61 (2003).
[8] G. Aad *et al.* (ATLAS Collaboration), Phys. Rev. D **86**, 032003 (2012).
[9] S. Chatrchyan *et al.* (CMS Collaboration), Phys. Lett. B **710**, 26 (2012).
[10] G. Aad *et al.* (ATLAS Collaboration), Phys. Lett. B **716**, 1 (2012).
[11] S. Chatrchyan *et al.* (CMS Collaboration), Phys. Lett. B **716**, 30 (2012).

[12] The TEVNPH (Tevatron New Phenomena and Higgs) Working Group, arXiv:1203.3774.

[13] T. Aaltonen *et al.* (CDF and D0 Collaborations), Phys. Rev. Lett. **109**, 071804 (2012).

[14] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. Lett. **109**, 121804 (2012).

[15] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **109**, 111804 (2012).

[16] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. Lett. **94**, 091802 (2005).

[17] V. M. Abazov *et al.* (D0 Collaboration), Phys. Lett. B **663**, 26 (2008).

[18] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. Lett. **102**, 051803 (2009).

[19] V. M. Abazov *et al.* (D0 Collaboration), Phys. Lett. B **698**, 6 (2011).

[20] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. D **86**, 032005 (2012).

[21] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. Lett. **106**, 171802 (2011).

[22] S. Abachi *et al.* (D0 Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **338**, 185 (1994).

[23] V. M. Abazov *et al.* (D0 Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **565**, 463 (2006).

[24] M. Abolins *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **584**, 75 (2008).

[25] R. Angstadt *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **622**, 298 (2010).

[26] The pseudorapidity $\eta = -\ln\left[\tan\frac{\theta}{2}\right]$, where θ is the polar angle as measured from the proton-beam axis.

[27] The azimuthal angle, ϕ, is defined as the opening angle with respect to the $+x$ direction in a right-handed coordinate system defined by $+y$ as up and $+z$ as the proton-beam direction.

[28] I. Narsky, arXiv:physics/0507157.

[29] L. Breiman, J. Friedman, R. Olshen, and C. Stone, *Classification and Regression Trees* (Wadsworth & Brooks/Cole Advanced Books and Software, Pacific Grove, CA, 1984).

[30] R. E. Schapire, "The Boosting Approach to Machine Learning: An Overview," in Proceedings of the MSRI Workshop on Nonlinear Estimation and Classification, Berkeley, CA, USA, 2001 (unpublished).

[31] Y. Freund and R. E. Schapire, Journal of Japanese Society for Artificial Intelligence **14**, 771 (1999) (in Japanese, translation by Naoki Abe).

[32] J. H. Friedman, *PHYSTAT 2003, Stanford, USA, 2003*, eConf C030908, WEAT003 (2003).

[33] A. Hoecker *et al.*, Proc. Sci., ACAT (2007) 040. We use version 4.1.0.

[34] $\Delta\mathcal{R} = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ is the separation between two objects in (η, ϕ) space, where ϕ is the azimuthal angle.

[35] G. C. Blazey *et al.*, arXiv:hep-ex/0005012.

[36] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. D **85**, 052006 (2012).

[37] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. D **84**, 032004 (2011).

[38] V. M. Abazov *et al.* (D0 Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **620**, 490 (2010).

[39] T. Sjöstrand, S. Mrenna, and P. Z. Skands, J. High Energy Phys. 05 (2006) 026.

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

[41] E. Boos, V. Bunichev, M. Dubinin, L. Dudko, V. Edneral, V. Ilyin, A. Kryukov, V. Savrin, A. Semenov, and A. Sherstnev, Nucl. Instrum. Methods Phys. Res., Sect. A **534**, 250 (2004).

[42] E. Boos, V. Bunichev, L. Dudko, V. Savrin, and V. Sherstnev, Phys. At. Nucl. **69**, 1317 (2006).

[43] J. Alwall *et al.*, Eur. Phys. J. C **53**, 473 (2008).

[44] H. L. Lai, J. Huston, S. Kuhlmann, F. Olness, J. Owens, D. Soper, W. K. Tung, and H. Weerts, Phys. Rev. D **55**, 1280 (1997).

[45] J. Pumplin, D. Stump, J. Huston, N. P. Lai, H.-L. and W.-K. Tung, J. High Energy Phys. 07 (2002) 012.

[46] R. Brun and F. Carminati, GEANT Detector Description and Simulation Tool, CERN Program Library Long Writeup W5013, 1993 (unpublished).

[47] J. Baglio and A. Djouadi, J. High Energy Phys. 10 (2010) 064.

[48] A. Martin, W. Stirling, R. Thorne, and G. Watt, Eur. Phys. J. C **63**, 189 (2009).

[49] D. de Florian and M. Grazzini, Phys. Lett. B **674**, 291 (2009).

[50] P. Bolzoni, F. Maltoni, S.-O. Moch, and M. Zaro, Phys. Rev. D **85**, 035002 (2012).

[51] A. Djouadi, J. Kalinowski, and M. Spira, Comput. Phys. Commun. **108**, 56 (1998).

[52] J. Butterworth *et al.*, arXiv:1003.1643.

[53] N. Kidonakis, Phys. Rev. D **74**, 114012 (2006).

[54] J. M. Campbell and R. K. Ellis, Phys. Rev. D **60**, 113006 (1999).

[55] J. M. Campbell, R. K. Ellis, and C. Williams, MCFM—Monte Carlo for FeMtobarn processes, http://mcfm.fnal.gov/.

[56] U. Langenfeld, S. Moch, and P. Uwer, Phys. Rev. D **80**, 054009 (2009).

[57] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. D **76**, 012003 (2007).

[58] K. Melnikov and F. Petriello, Phys. Rev. D **74**, 114017 (2006).

[59] J. M. Campbell, arXiv:hep-ph/0105226.

[60] V. M. Abazov *et al.* (D0 Collaboration), Phys. Lett. B **669**, 278 (2008).

[61] Transverse mass of a leptonically decaying W is defined as $M_T^2 \equiv 2p_T^\ell \not{E}_T[1 - \cos\Delta\phi(\ell, \not{E}_T)]$.

[62] L. Breiman, Mach. Learn. **45**, 5 (2001).

[63] I. Narsky, arXiv:physics/0507143.

[64] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevD.88.052008 for descriptions of MVA input variables.

[65] J. Beringer *et al.* (Particle Data Group), Phys. Rev. D **86**, 010001 (2012).

[66] T. Andeen *et al.*, Report No. FERMILAB-TM-2365, 2007.

[67] T. Junk, Nucl. Instrum. Methods Phys. Res., Sect. A **434**, 435 (1999).

[68] A. L. Read, J. Phys. G **28**, 2693 (2002).

[69] W. Fisher, Report No. FERMILAB-TM-2386-E, 2007.

[70] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. D **82**, 011102 (2010).

[71] G. Aad *et al.* (ATLAS Collaboration), Eur. Phys. J. C **71**, 1728 (2011).

[72] S. Chatrchyan *et al.* (CMS Collaboration), Phys. Lett. B **699**, 25 (2011).

[73] A. Heister *et al.* (ALEPH Collaboration), Phys. Lett. B **544**, 16 (2002).

[74] J. Abdallah *et al.* (DELPHI Collaboration), Eur. Phys. J. C **35**, 313 (2004).

[75] P. Achard *et al.* (L3 Collaboration), Phys. Lett. B **568**, 191 (2003).

[76] G. Abbiendi *et al.* (OPAL Collaboration), Phys. Lett. B **544**, 259 (2002).

[77] T. Aaltonen *et al.* (CDF Collaboration), Phys. Lett. B **717**, 173 (2012).

[78] V. Abazov *et al.* (D0 Collaboration), Phys. Rev. Lett. **107**, 151801 (2011).

[79] G. Aad *et al.* (ATLAS Collaboration), Eur. Phys. J. C **72**, 2157 (2012).

[80] S. Chatrchyan *et al.* (CMS Collaboration), J. High Energy Phys. 09 (2012) 111.

[81] B. Holdom, W. S. Hou, T. Hurth, M. L. Mangano, S. Sultansoy, and G. Ünel, PMC Phys. A **3**, 4 (2009).

[82] G. D. Kribs, T. Plehn, M. Spannowsky, and T. M. P. Tait, Phys. Rev. D **76**, 075016 (2007).

[83] E. Arik, O. Cakir, S. A. Cetin, and S. Sultansoy, Acta Phys. Pol. B **37**, 2839 (2006).

[84] C. Anastasiou, R. Boughezal, and E. Furlan, J. High Energy Phys. 06 (2010) 101.

[85] Maximum $\Delta\eta$ between the charged lepton and the leading or second leading jet.

[86] Product of the lepton charge and its pseudorapidity.

[87] Scalar sum of the p_{T} of the visible particles, including the charged lepton and jets.

[88] Aplanarity is defined as $3\lambda_3/2$, where λ_3 is the smallest eigenvalue of the normalized momentum tensor $S^{\alpha\beta} = (\sum_i p_i^\alpha p_i^\beta)/(\sum_i |\vec{p}_i|^2)$, where α, $\beta = 1, 2, 3$ correspond to the x, y, z momentum components, and i runs over all visible objects.

[89] Invariant mass of the system consisting of the charged lepton, reconstructed neutrino, and second leading jet.

[90] The p_Z of the neutrino candidate is estimated by constraining the charged lepton and the neutrino system to the mass of the W boson and choosing the lowest magnitude solution.

[91] $\Delta\eta$ between the $\ell\nu$ system and the charged lepton.

[92] Cosine of the angle between the charged lepton and the proton beam axis in the center of mass of the $\ell\nu$ system.

[93] $\mathcal{SIG}_{\mathrm{jets}}$ for ℓ is defined with respect to all jets in an event as $\sum_{\mathrm{jets}}[p_{\mathrm{T}}^{\mathrm{jet}} \times \Delta\mathcal{R}(\ell, \mathrm{jet})]/\sum_{\mathrm{jets}}(p_{\mathrm{T}}^{\mathrm{jet}})$.