Measurements of $WW$ and $WZ$ production in $W + \text{jets}$ final states in $pp$ collisions

(The D0 Collaboration*)

1Universidad de Buenos Aires, Buenos Aires, Argentina
2LAFEX, Centro Brasileiro de Pesquisas Fisicas, Rio de Janeiro, Brazil
3Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil
4Universidade Federal do ABC, Santo André, Brazil
5Instituto de Fisica Teorica, Universidade Estadual Paulista, São Paulo, Brazil
6University of Science and Technology of China, Hefei, People’s Republic of China
7Universidad de los Andes, Bogotá, Colombia
8Charles University, Faculty of Mathematics and Physics, Center for Particle Physics, Prague, Czech Republic
9Czech Technical University in Prague, Prague, Czech Republic
10Center for Particle Physics, Institute of Physics, Academy of Sciences of the Czech Republic, Prague, Czech Republic
11Universidad San Francisco de Quito, Quito, Ecuador
12LPC, Université Blaise Pascal, CNRS/IN2P3, Clermont, France
13LPSC, Université Joseph Fourier Grenoble 1, CNRS/IN2P3, Institut National Polytechnique de Grenoble, Grenoble, France
14CPPM, Aix-Marseille Université, CNRS/IN2P3, Marseille, France
15LAL, Université Paris-Sud, CNRS/IN2P3, Orsay, France
16LPNHE, Universités Paris VI and VII, CNRS/IN2P3, Paris, France
17CEA, Ifis, SPP, Saclay, France
18IPHC, Université de Strasbourg, CNRS/IN2P3, Strasbourg, France
19IPNL, Université Lyon 1, CNRS/IN2P3, Villeurbanne, France and Université de Lyon, Lyon, France
20Physikalisches Institut A, RWTH Aachen University, Aachen, Germany
21Physikalisches Institut, Universität Freiburg, Freiburg, Germany
22II. Physikalisches Institut, Georg-August-Universität Göttingen, Göttingen, Germany
23Institut für Physik, Universität Mainz, Mainz, Germany
24Ludwig-Maximilians-Universität München, München, Germany
25Fachbereich Physik, Bergische Universität Wuppertal, Wuppertal, Germany
26Panjab University, Chandigarh, India
27Delhi University, Delhi, India
28Tata Institute of Fundamental Research, Mumbai, India
29University College Dublin, Dublin, Ireland
30Korea Detector Laboratory, Korea University, Seoul, Korea
31CINVESTAV, Mezco City, Mexico
32Nikhef, Science Park, Amsterdam, the Netherlands
33Radboud University Nijmegen, Nijmegen, the Netherlands and Nikhef, Science Park, Amsterdam, the Netherlands
34Joint Institute for Nuclear Research, Dubna, Russia
35Institute for Theoretical and Experimental Physics, Moscow, Russia
36Moscow State University, Moscow, Russia
37Institute for High Energy Physics, Protvino, Russia
38Petersburg Nuclear Physics Institute, St. Petersburg, Russia
39Institució Catalana de Recerca i Estudis Avançats (ICREA) and Institut de Fisica d’Altes Energies (IFAE), Barcelona, Spain
40Stockholm University, Stockholm and Uppsala University, Uppsala, Sweden
41Lancaster University, Lancaster LA1 4YB, United Kingdom
42Imperial College London, London SW7 2AZ, United Kingdom
43The University of Manchester, Manchester M13 9PL, United Kingdom
44University of Arizona, Tucson, Arizona 85721, USA
45University of California Riverside, Riverside, California 92521, USA
46Florida State University, Tallahassee, Florida 32306, USA
47Fermi National Accelerator Laboratory, Batavia, Illinois 60510, USA
48University of Illinois at Chicago, Chicago, Illinois 60607, USA
We study WW and WZ production with $\ell qq$ ($\ell = e, \mu$) final states using data collected by the D0 detector at the Fermilab Tevatron Collider corresponding to 4.3 fb$^{-1}$ of integrated luminosity from $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV. Assuming the ratio between the production cross sections $\sigma(WW)$ and $\sigma(WZ)$ as predicted by the standard model, we measure the total $(WV(W = W, Z))$ cross section to be $\sigma(WV) = 19.6^{+3.2}_{-3.0}$ pb, and reject the background-only hypothesis at a level of 7.9 standard deviations. We also use $b$-jet discrimination to separate the WZ component from the dominant WW component. Simultaneously fitting WW and WZ contributions, we measure $\sigma(WW) = 15.9^{+3.7}_{-3.2}$ pb and $\sigma(WZ) = 3.3^{+0.3}_{-0.3}$ pb, which is consistent with the standard model predictions.

PACS numbers: 14.70.Fm, 14.70.Hp, 13.85.Ni, 13.85.Qk

The study of the production of $VV (V = W, Z)$ boson pairs provides an important test of the electroweak sector of the standard model (SM). In $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV, the next-to-leading order (NLO) SM cross sections for these processes are $\sigma(WV) = 11.7 \pm 0.8$ pb, $\sigma(WZ) = 3.5\pm0.3$ pb and $\sigma(ZZ) = 1.4\pm0.1$ pb $^3$. Measuring a significant departure in cross section or deviations in the predicted kinematic distributions would indicate the presence of anomalous gauge boson couplings $^2$ or new particles in extensions of the SM $^3$. This analysis also provides a proving ground for the advanced analysis techniques used in low mass Higgs searches $^4$. The production of VV in $p\bar{p}$ collisions at the Fermilab Tevatron Collider has been observed in fully leptonic decay modes $^5$ and more recently, in leptons+jets decay modes $^6$, where the combined $WW+ZZ$ cross section was measured.

In this Letter, we report observation of the production of a W boson that decays leptonically in associated production with a second vector boson that decays hadronically ($WV \rightarrow \ell qq$; $\ell = e^\pm$ or $\mu^\pm$, and $\nu$ and $q$ denote matter or anti-matter as appropriate). The data used for this analysis correspond to 4.3 fb$^{-1}$ of integrated lumi-
nosity collected between 2006 and 2009 by the D0 detector \cite{7} at the Fermilab Tevatron Collider. The D0 detector dijet mass resolution for $W/Z$ decays of $\approx 18\%$ results in significant overlap of $W \to q\bar{q}$ and $Z \to q\bar{q}$ dijet mass peaks. Therefore, we first consider $WW$ and $WZ$ simultaneously and measure the total $WV$ cross section assuming the ratio of $WW$ to $WZ$ cross sections as predicted by the SM. We then apply $b$-jet identification to separate the $WZ$ contribution, where the $Z$ boson decays into $b\bar{b}$ pairs, from the dominant $WW$ production.

Candidate events in the electron channel are required to satisfy a single electron trigger or a trigger requiring electrons and jets, which results in a combined trigger efficiency of $(98.3 \pm 2.3)\%$ for the $e\nu qq$ event selection described below. A comprehensive suite of triggers in the muon channel, based on leptons, jets and their combination, achieves a trigger efficiency of $(95 \pm 5)\%$ for the $\mu qq$ event selection.

To select $WW \to e\nu qq$ candidates, we require a single reconstructed electron (muon) with transverse momentum $p_T > 15$ GeV (20 GeV) and pseudorapidity $|\eta| < 1.1$ (2.0) \cite{8}, missing transverse energy $E_T > 20$ GeV, and two or three jets reconstructed using a cone algorithm \cite{9}. The jets must have $p_T > 20$ GeV, $|\eta| < 2.5$, and at least two tracks within the jet cone \cite{9} originating from the $p\bar{p}$ interaction vertex. Lepton candidates must be spatially matched to a track that originates from the primary $p\bar{p}$ interaction vertex and they must be isolated from energy depositions in the calorimeter and other tracks in the central tracking detector. To reduce background from processes that do not contain $W \to \ell \nu$, we require that the transverse mass \cite{10} is $M_T^\ell (\text{GeV}) > 40 - 0.5 E_T$. In addition, we restrict $M_T^\mu < 200$ GeV to suppress muon candidates with poorly measured momenta.

Signal and most of the background processes are modeled with Monte Carlo (MC) simulation. The signal events are generated with PYTHIA \cite{11} using CTEQ6L1 parton distribution functions (PDFs) \cite{12} and include all SM decays. The fixed-order matrix element (FOME) generator ALPGEN \cite{13} with CTEQ6L1 PDF is used to generate $W+\text{jets}, \ Z+\text{jets}$, and $t\bar{t}$ events. The FOME generator COMPHEN \cite{14} is used to produce single top-quark MC samples with CTEQ6M PDF \cite{12}. Both ALPGEN and COMPHEN are interfaced to PYTHIA for parton showering and hadronization. The MC events undergo a GEANT-based \cite{15} detector simulation and are reconstructed using the same algorithms as used for D0 data. The effect of multiple $p\bar{p}$ interactions is included by overlaying data events from random beam crossings on simulated events. The next-to-NLO (NNLO) cross section is used to normalize the $Z+\text{jets}$ (light and heavy-flavor jets) \cite{16}. The approximate NNLO cross section \cite{17} is used to normalize the $t\bar{t}$ samples, while the single top-quark MC samples are normalized to the approximate next-to-NNLO cross section \cite{18}. The normalization of the $W+\text{jets}$ MC sample (for all flavor contributions) is determined from data. Additional NLO heavy-flavor corrections are calculated with MCFM \cite{19} and applied to $Z/W+\text{heavy-flavor jets}$ MC samples.

The multijet background in which a jet is misidentified as a prompt lepton is determined from data. For the muon channel, the multijet background is modeled with data that fail the muon isolation requirements, but pass all other selections. For the electron channel, the multijet background is estimated using a data sample containing events that pass less restrictive electron quality requirements. Both multijet samples are corrected for contributions from processes modeled by MC. The multijet normalizations are determined from fits to the $M_T^\mu \nu$ distributions and assigned uncertainties of 20%.

To identify $b$-quark jets, in particular those originating from $Z \to b\bar{b}$ decays, we use the D0 neural network (NN) $b$-tagging algorithm \cite{20}. The NN is trained to separate light-flavor jets from heavy-flavor jets based on a combination of variables sensitive to the presence of tracks and vertices displaced from the primary $p\bar{p}$ interaction vertex. The NN outputs for the two highest $p_T$ jets are then used as inputs to the final multivariate discriminant. We define non-overlapping 0, 1, and 2-tag sub-channels based on whether neither, only one, or both of the two highest $p_T$ jets pass the least restrictive NN operating point, for which the $b$-jet identification efficiency and the light-flavor jet misidentification rate are approximately 80% and 10%. Scale factors are applied to the MC events to account for any difference in efficiency or misidentification rate between data and simulation.

The dominant background is $W+\text{jets}$ and therefore the modeling of this process in ALPGEN and the corresponding sources of uncertainties were studied in detail. Comparison of ALPGEN with other generators \cite{21} and with data shows discrepancies in jet $\eta$, dijet angular separation and the transverse momentum of the $W$ boson candidate. Thus, data are used to correct these quantities in the ALPGEN $W+\text{jets}$ and $Z+\text{jets}$ samples before $b$-tagging is performed \cite{22}. The possible bias in this procedure from the presence of the diboson signal in data is small, but is taken into account as a systematic uncertainty.

As the diboson events are generated with a LO generator, changes to the event kinematics and the acceptance due to a NLO and resummation effects are studied using events from the MC@NLO \cite{23} interfaced to HERWIG \cite{24} for parton showering and hadronization and using the CTEQ6M PDF set. Comparing kinematics at the generator level after final state radiation, we parameterize a two-dimensional correction matrix in the $p_T$ of the diboson system and of the highest $p_T$ boson. After applying this correction to our PYTHIA sample, we find good agreement with MC@NLO for all distributions studied. Half of the difference between the PYTHIA and MC@NLO predictions is used as systematic uncertainty on the diboson production model, accounting for the possible effects of higher order corrections beyond NLO and of different
TABLE I: Number of events for signal and each background after the combined fit of $WV$ using the RF output distribution (with total uncertainties determined from the fit) and the number of events observed in data.

<table>
<thead>
<tr>
<th>Bkgd</th>
<th>Electron channel</th>
<th>Muon channel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diboson signal</td>
<td>1725 ± 84</td>
<td>1465 ± 67</td>
</tr>
<tr>
<td>$W/Z$ + light-flavor jets</td>
<td>37232 ± 1033</td>
<td>33516 ± 709</td>
</tr>
<tr>
<td>$W/Z$ + heavy-flavor jets</td>
<td>5371 ± 608</td>
<td>4854 ± 490</td>
</tr>
<tr>
<td>$t\bar{t}$ and single top</td>
<td>1746 ± 127</td>
<td>1214 ± 86</td>
</tr>
<tr>
<td>Multijet</td>
<td>10630 ± 1007</td>
<td>1982 ± 384</td>
</tr>
<tr>
<td>Total predicted</td>
<td>56704 ± 635</td>
<td>43031 ± 531</td>
</tr>
<tr>
<td>Data</td>
<td>56698</td>
<td>43044</td>
</tr>
</tbody>
</table>

The signal and the backgrounds are further separated using a multivariate classifier to combine information from several variables. This analysis uses a random forest (RF) classifier [25, 26], from which the output distribution is used as a final variable to measure the production cross sections by performing a template fit. Fifteen well-modeled variables [27] that demonstrate a difference in probability density between signal and at least one of the backgrounds are used as inputs to the RF. The RF is trained using a fraction of each MC sample. The remainder of each MC sample, along with the multijet background samples, is then evaluated by the RF and used in the measurement.

Depending on the source, we consider the effect of systematic uncertainty on the normalization and/or on the shape of differential distributions for signal and backgrounds [27]. Systematic effects on the differential distributions of the ALPGEN $W$+jets and $Z$+jets MC events from changes of the renormalization and factorization scales and of the parameters used in the MLM parton-jet matching algorithm [28] are also considered. Uncertainties on PDFs [29], as well as uncertainties from object reconstruction and identification, are evaluated for all MC samples.

The total $WV$ cross section is determined from a fit to the data of the signal and background RF output distributions. This fit is performed simultaneously on the distributions in the electron and muon channels, and in the 0, 1, and 2-tag sub-channels, by minimizing a Poisson $\chi^2$ function with respect to variations in the systematic uncertainties [30]. The magnitude of the systematic uncertainties is effectively constrained by the regions of the RF output distribution with low signal over background ratio. A Gaussian prior is used for each systematic uncertainty. The effects on separate samples or sub-channels due to the same uncertainty are assumed to be 100% correlated. However, different uncertainties are assumed to be mutually independent.

The fit simultaneously varies the signal and $W$+jets contributions, thereby also determining the normalization factor for the $W$+jets MC sample. This obviates the need for using the predicted ALPGEN cross section, and provides a more rigorous approach that incorporates an unbiased uncertainty from $W$+jets when extracting the signal cross section. The $W$+jets normalization factor from the fit is consistent with the theoretical NNLO prediction [31]. The yields for signal and each background are given in Table I. Though the total diboson yield includes a small contribution from $ZZ \rightarrow \ell\ell qq$ events (1.5%), in which one of the charged leptons escapes detection, the cross sections presented here are corrected for this contribution assuming that the ratios between $WW$, $WZ$ and $ZZ$ cross sections are given by the SM.
FIG. 3: (color online) A comparison of the signal+background prediction to data in the RF output distribution (summed over electron and muon channels) for 0, 1, and 2-tag sub-channels after the combined fit to data using the dijet mass distribution. Also shown is the RF output distribution (points) in the dijet mass distribution (summed over electron and muon channels) for 0, 1, and 2-tag sub-channels after the combined fit to data using the RF output distribution. The fit is then performed with the signal divided into the separate \( W W \), \( W Z \), and \( W Z + \) LP components, which are allowed to float independently. The result of this simultaneous fit of \( \sigma(WW) \) and \( \sigma(WZ) \) using the RF output distributions is shown in Fig. 2. It yields \( \sigma(WW) = 15.9^{+1.9}_{-1.5} (\text{stat})^{+3.2}_{-2.9} (\text{syst}) \) pb and \( \sigma(WZ) = 3.3^{+3.4}_{-2.7} (\text{stat})^{+2.2}_{-1.8} (\text{syst}) \) pb. The RF output distributions for the 0, 1, and 2-tag sub-channels from this fit are shown in Fig. 3. This measurement is also verified fitting the dijet mass distribution, which yields \( \sigma(WW) = 13.3^{+2.2}_{-2.0} (\text{stat})^{+3.6}_{-2.7} (\text{syst}) \) pb and \( \sigma(WZ) = 5.4^{+4.5}_{-2.6} (\text{stat})^{+4.3}_{-2.6} (\text{syst}) \) pb. Figure 4 shows plots for the background-subtracted dijet mass after the dijet mass fit.

We also perform a fit in which we constrain the \( WW \) cross section to its SM prediction with a Gaussian prior equal to the theoretical uncertainty of 7\% [1]. The fit of the RF output distribution yields a \( W Z \) cross section of \( \sigma(WZ) = 6.5 \pm 0.9 (\text{stat}) \pm 3.0 (\text{syst}) \) pb with observed (expected) significance of 2.2 (1.2) s.d., and the dijet mass fit yields \( \sigma(WZ) = 6.7 \pm 1.0 (\text{stat}) \pm 3.9 (\text{syst}) \) pb with observed (expected) significance of 1.7 (0.9) s.d. As expected, now that \( \sigma(WW) \) is constrained to the SM prediction, the fit requires a higher rate for \( W Z \) in order to account for the excess of signal-like events.

In summary, we have measured the cross section for total \( WV \) production to be \( \sigma(WV) = 19.6^{+3.2}_{-3.0} \) pb \((V = W \ or \ Z)\) with a significance of 7.9 s.d. above the background-only hypothesis. This result demonstrates the ability of the D0 experiment to measure a dijet signal in a background-dominated final state directly relevant to low mass Higgs searches. Furthermore, we have used b-jet tagging to measure the contributions from \( W W \) and \( W Z \) and measured the cross sections for the separate processes to be \( \sigma(WW) = 15.9^{+3.2}_{-3.0} \) pb and \( \sigma(WZ) = 3.3^{+4.3}_{-3.3} \) pb. Although we cannot yet claim 3 s.d. evidence of a \( W Z \) signal in the \( \ell\nu jj \) final states, the extracted \( WV \) and \( W Z \) cross sections are in agreement with the SM prediction and their precise measurement represents an independent test to new physics which
could manifest itself differently in different final states.

We thank the staffs at Fermilab and collaborating institutions, and acknowledge support from the DOE and NSF (USA); CEA and CNRS/IN2P3 (France); FASI, Rosatom and RFBR (Russia); CNFq, FAPERJ, FAPESP and FUNDUNESP (Brazil); DAE and DST (India); Colciencias (Colombia); CONACyT (Mexico); KRF and KOSEF (Korea); CONICET and UBACyT (Argentina); Colciencias (Colombia); CONACyT (Mexico); KRF and KOSEF (Korea); CONICET and UBACyT (Argentina); FOM (The Netherlands); STFC and the Royal Society (United Kingdom); MSMT and GACR (Czech Republic); CRC Program and NSERC (Canada); BMBF (Germany); SFI (Ireland); The Swedish Research Council (Sweden); and CAS and CNSF (China).


[8] D0 uses a spherical coordinate system with the z axis running along the proton beam axis. The angles \( \theta \) and \( \phi \) are the polar and azimuthal angles, respectively. Pseudorapidity is defined as \( \eta = -\ln [\tan(\theta/2)] \), in which \( \theta \) is measured with respect to the proton beam direction.

[9] G. C. Blazey et al., arXiv:hep-ex/0005012 (2000). The seeded cone algorithm with radius \( \Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2} = 0.5 \) is used.


[27] See Appendix.


APPENDIX

INPUT VARIABLES TO THE RANDOM FOREST CLASSIFIER

Here we define the fifteen variables used as inputs to the RF classifier. The observed distribution for each variable is shown in Figs. 5[4] and 7 along with the predicted distribution after the fit of the total WW (V = W, Z) cross section using the RF output distribution.

The RF input variables can be classified into three categories: (i) b-jet identification variables, (ii) kinematics of individual final state particles, and (iii) kinematics of multiple final state particles. Several variables are calculated using the four-momentum of the dijet system, which we define as the sum of the four-momenta of the two highest p_T jets. We also reconstruct a W → ℓv candidate W^{ℓν}, from the charged lepton and the E_T. The neutrino from the W → ℓv decay is assigned the transverse momentum defined by E_T and a longitudinal momentum that is calculated assuming the mass of the ℓν system is 80.4 GeV. Of the two possible solutions, we choose the real component that provides the smaller total invariant mass of all objects in the event.

- **b-jet Identification Variables:**

  The NN b-tagger has 12 operating points characterized by different purities. Each jet is assigned an integer b-tag value based on the highest purity operating point that it passes.

  1. Max b-tag Value: The greater b-tag value of the two highest p_T jets. The neural network b-tagger has 12 operating points of increasing purity and the b-tag value corresponds to the highest operating point satisfied by the jet (or zero if the jet did not satisfy any of the operating points).

  2. Min b-tag Value: The lesser b-tag value of the two highest p_T jets.

  These variables are shown in Fig. 5 both in logarithmic and linear scales.

- **Kinematics of Individual Final State Particles:**

  1. \( p_T(ℓ) \): The \( p_T \) of the charged lepton.

  2. \( p_T(\text{jet}_1) \): The highest jet \( p_T \).

  3. \( p_T(\text{jet}_2) \): The second highest jet \( p_T \).

  4. \( E_T \): The imbalance in transverse energy determined from the energy measured in each calorimeter cell and then corrected for reconstructed muons, jets, and electrons/photons.

- **Kinematics of Multiple Final State Particles:**

  1. \( M_{jj} \): The invariant mass of the dijet system reconstructed from the two highest \( p_T \) jets.

  2. \( M_T^\nu = \sqrt{2 p_T^\nu E_T (1 - \cos(\Delta \phi(ℓ, E_T)))} \): The transverse W mass reconstructed from the charged lepton and the \( E_T \).

  3. \( H_T = p_T(\text{jet}_1) + p_T(\text{jet}_2) \): The scalar sum of the two highest jet \( p_T \)s.

  4. \( p_T^E(\text{dijet, jet}_1)^W = |p_T(\text{jet}_1) \times p_T(\text{jet}_1 + \text{jet}_2)| \): The magnitude of the leading jet transverse momentum perpendicular to the dijet system in the rest frame of the \( W^{ℓν} \) candidate;

  5. \( p_T^E(\text{dijet, jet}_2) = |p_T(\text{jet}_2) \times p_T(\text{jet}_1 + \text{jet}_2)| \): The magnitude of the second-leading jet transverse momentum perpendicular to the dijet system in the laboratory frame.

  6. \( k_T^{min} = \Delta R(\text{jet}_1, \text{jet}_2) \frac{p_T(\text{jet}_1)}{p_T(\text{jet}_1 + \text{jet}_2)} \): The angular separation between the two jets of highest \( p_T \), weighted by the ratio of the transverse momentum of the second-leading jet and a scalar sum of the transverse momenta of the \( W^{ℓν} \) constituents.

  7. \( \cos(\angle(\text{dijet, jet}_1)) \): The cosine of the angle between the momentum vectors of the dijet system and the highest \( p_T \) jet in the laboratory frame.

  8. \( \cos(\angle(W^{ℓν}, \text{jet}_1)^{jj}) \): Cosine of the angle between the momentum vectors of the leading jet and the \( W^{ℓν} \) candidate, evaluated in the rest frame of the dijet system.

  9. Centrality: The scalar sum of transverse momenta of the charged lepton and all jets in the event divided by the sum of their energies.

SYSTEMATIC UNCERTAINTIES

Table II gives the size of the systematic uncertainties for Monte Carlo simulations and multijet estimates. We consider the effect of systematic uncertainties both on the normalization and on the shape of differential distributions for signal and backgrounds. Although Table II lists uncertainties for the diboson and W+jets simulation, these uncertainties are not used when measuring the diboson signal cross section, for which the diboson and W+jets normalizations are free parameters. However, the size of the uncertainty must be specified when estimating the significance and when we constrain the cross section for WW production to its SM prediction in the fit.
FIG. 5: (color online) Distributions of the $b$-jet identification variables used as inputs to the RF classifier (first two of fifteen) for electron and muon channels combined, with logarithmic ((a) and (b)) and linear ((c) and (d)) scales. To better show the $WW$ and $WZ$ signals the lowest bin is cut off in the distributions with a linear scale. The signal and background predictions and the systematic uncertainty band are evaluated after the fit of the total $WV$ cross section in the RF output distribution. Definitions for each variable are provided in the text (LP denotes light partons such as $u$, $d$, $s$ or gluon, and HF denotes heavy-flavor such as $c\bar{c}$ or $b\bar{b}$).

$\Delta \chi^2$ FOR W V MEASUREMENT

The statistical significance of the diboson signal yield from the fit to the data is estimated via analysis of the $\Delta \chi^2$ curve obtained by fitting the data to the sum of background and signal templates as a function of the signal rate. The results of this analysis are given in Table III. Figure 8 shows how the $\chi^2$ of the fit changes as a function of the signal cross section when using either the dijet mass or the RF output distribution to measure the total $WV$ cross section.
FIG. 6: (color online) Distributions of the variables (next eight of fifteen) used as inputs to the RF classifier for electron and muon channels combined, and before b-tagging. The signal and background predictions and the systematic uncertainty band are evaluated after the fit of the total WV cross section in the RF output distribution. Definitions for each variable are provided in the text (LP denotes light partons such as $u$, $d$, $s$ or gluon, and HF denotes heavy-flavor such as $c\bar{c}$ or $b\bar{b}$).
FIG. 7: (color online) Distributions of the variables (remaining five of fifteen) used as inputs to the RF classifier for electron and muon channels combined, and before $b$-tagging. The signal and background predictions and the systematic uncertainty band are evaluated after the fit of the total $WV$ cross section in the RF output distribution. Definitions for each variable are provided in the text (LP denotes light partons such as $u$, $d$, $s$ or gluon, and HF denotes heavy-flavor such as $c\bar{c}$ or $b\bar{b}$).
TABLE II: The RMS amplitude (in percent) of each systematic uncertainty in the RF output distributions for the signal and background predictions. The RMS amplitude is defined as: \( \sqrt{\sum_{i=0}^{n} p_i \Delta_i^2 / \sum_{i=0}^{n} p_i} \); where \( p_i \) is the predicted number of events in bin \( i \), \( \Delta_i \) is the percent change in bin \( i \) when the uncertainty is varied by 1 s.d., and \( n \) is the number of bins. In cases where the amplitude is different for different subchannels, the range of amplitudes is given. The rightmost column indicates whether the uncertainty only affects the normalization (N) or if it has also a differential dependence (D).

<table>
<thead>
<tr>
<th>Source of systematic uncertainty</th>
<th>Diboson signal</th>
<th>W+jets</th>
<th>Z+jets</th>
<th>Top</th>
<th>Multijet</th>
<th>Nature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electron trigger/ID efficiency</td>
<td>±5</td>
<td>±5</td>
<td>±5</td>
<td>±5</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>Muon trigger/ID efficiency</td>
<td>±5</td>
<td>±5</td>
<td>±5</td>
<td>±5</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>Jet identification</td>
<td>±1</td>
<td>±1-2</td>
<td>±1-2</td>
<td>±1-2</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>Jet energy scale</td>
<td>±2-4</td>
<td>±6-8</td>
<td>±4-12</td>
<td>±2-3</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>Jet energy resolution</td>
<td>±2-3</td>
<td>±3-12</td>
<td>±4-10</td>
<td>±1-2</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>Jet vertex confirmation</td>
<td>±2-3</td>
<td>±3-4</td>
<td>±3-5</td>
<td>±1-3</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>Taggability correction</td>
<td>±1-5</td>
<td>±1-4</td>
<td>±1-5</td>
<td>±1-5</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>b-tagging</td>
<td>±1-5</td>
<td>±1-4</td>
<td>±1-5</td>
<td>±1-5</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>Luminosity</td>
<td>±6.1</td>
<td>±6.1</td>
<td>±6.1</td>
<td>±6.1</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>Cross section</td>
<td>±7</td>
<td>±6.3</td>
<td>±6.3</td>
<td>±10</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>V+heavy-flavor cross section</td>
<td>±20</td>
<td>±20</td>
<td>±20</td>
<td></td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>V+2 jets/V+3 jets cross section</td>
<td>±10</td>
<td>±10</td>
<td>±10</td>
<td></td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>Multijet normalization</td>
<td>±20</td>
<td>±20</td>
<td>±20</td>
<td></td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>Multijet shape, electron channel</td>
<td>± &lt;1</td>
<td>± &lt;1</td>
<td>± &lt;1</td>
<td>± &lt;1</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>Multijet shape, muon channel</td>
<td>± &lt;1</td>
<td>± &lt;1</td>
<td>± &lt;1</td>
<td>± &lt;1</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>Diboson modeling</td>
<td>±2-3</td>
<td>±2</td>
<td>±1-3</td>
<td>±2-4</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>Parton distribution function</td>
<td>±1</td>
<td>±1</td>
<td>±1</td>
<td>±1</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>Unclassified Energy correction</td>
<td>± &lt;1</td>
<td>± &lt;1</td>
<td>± &lt;1</td>
<td>± &lt;1</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>ALPGEN jet ( \eta ) corrections</td>
<td>± &lt;1</td>
<td>± &lt;1</td>
<td>± &lt;1</td>
<td>± &lt;1</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>ALPGEN ( \Delta R(jj) ) and ( p_T(W) ) corrections</td>
<td>± &lt;1</td>
<td>± &lt;1</td>
<td>± &lt;1</td>
<td>± &lt;1</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>Re-weighting diboson bias</td>
<td>± &lt;1</td>
<td>± &lt;1</td>
<td>± &lt;1</td>
<td>± &lt;1</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>Renormalization and factorization scales</td>
<td>± &lt;1</td>
<td>± &lt;1</td>
<td>± &lt;1</td>
<td>± &lt;1</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>Underlying event model</td>
<td>±1</td>
<td>±1</td>
<td>±1</td>
<td></td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>ALPGEN parton-jet matching parameters</td>
<td>± &lt;1</td>
<td>± &lt;1</td>
<td>± &lt;1</td>
<td>± &lt;1</td>
<td>D</td>
<td></td>
</tr>
</tbody>
</table>

TABLE III: Results from fitting the total \( WV \) cross section and the uncertainties resulting from limited data statistics (stat), and possible systematic biases (syst). Also, the expected and observed \( \Delta \chi^2 \) obtained by fitting the data with and without the specified signal process and the corresponding significance in number of standard deviations (s.d.) for a one-sided Gaussian integral.

<table>
<thead>
<tr>
<th></th>
<th>Measured ( \sigma(WV) ) [pb]</th>
<th>Expected</th>
<th>( \Delta \chi^2 ) (significance)</th>
<th>Observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>RF Output</td>
<td>19.6 ± 1.4 (stat) ±2.7 (syst)</td>
<td>35.8 (5.9 s.d.)</td>
<td>63.5 (7.9 s.d.)</td>
<td></td>
</tr>
<tr>
<td>Dijet Mass</td>
<td>18.3 ± 1.5 (stat) ±3.5 (syst)</td>
<td>22.1 (4.6 s.d.)</td>
<td>33.0 (5.6 s.d.)</td>
<td></td>
</tr>
</tbody>
</table>
FIG. 8: (color online) The change in $\chi^2$ relative to the best fit value when fitting the total $WV$ cross section using either the dijet mass or the RF output distribution.