Measurement of the production and differential cross sections of $W^+W^-$ bosons in association with jets in $p\bar{p}$ collisions at $s = 1.96$ TeV.
Measurement of the production and differential cross sections of $W^+W^-$ bosons in association with jets in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV

We present a measurement of the $W^+W^-$-boson-pair production cross section in $p\bar{p}$ collisions at 1.96 TeV center-of-mass energy and the first measurement of the differential cross section as a function of jet multiplicity and leading-jet energy. The $W^+W^-$ cross section is measured in the final state comprising two charged leptons and neutrinos, where either charged lepton can be an electron or a muon. Using data collected by the CDF experiment corresponding to 9.7 fb$^{-1}$ of integrated luminosity, a total of 3027 collision events consistent with $W^+W^-$ production are observed with an estimated background...
contribution of $1790 \pm 190$ events. The measured total cross section is $\sigma(p \bar{p} \to W^+W^-) = 14.0 \pm 0.6_{\text{stat}}^{+1.3}_{-1.0}{\text{syst}} \pm 0.8_{\text{lumi}}$ pb, consistent with the standard model prediction.

DOI: 10.1103/PhysRevD.91.111101 PACS numbers: 14.70.Fm, 12.38.Qk, 12.15.Ji

The measurement of $W$-boson-pair production is an important test of the electroweak gauge sector of the standard model (SM) of particle physics [1,2]. Extending the measurement to include the properties of jets produced in association with the boson pair is an interesting test of quantum chromodynamics (QCD) that has not been performed before in events with multiple heavy gauge bosons. This process is a significant background for Higgs-boson studies in the $W^+W^-(WW)$ decay mode at particle colliders, where the use of jet vetoes and jet counting is an essential element of the analysis technique [3–6]. Furthermore, a final state consisting of two massive gauge bosons and two jets is typical of vector-boson scattering, a process sensitive to many non-SM contributions to the electroweak-symmetry-breaking sector [7]. With the first evidence of vector-boson scattering at the Large Hadron Collider (LHC) [8,9], it is essential to verify and tune the simulation tools used to describe electroweak processes produced in association with jets.

This article reports measurements of the $W^+W^-$ production and differential cross sections as a function of jet multiplicity and jet energy in a final state consisting of events with two oppositely charged leptons, where each lepton is identified as either an electron or a muon, and an imbalance in the total event transverse momentum (transverse energy), due to the presence of neutrinos. The differential measurements of jet multiplicity and jet energy are unfolded to hadronic-jet level to account for the detector response to jets of hadrons and compared directly to predictions from two Monte Carlo (MC) simulations. The simulations considered are the ALPGEN generator [10], which calculates the lowest perturbative order of the strong interaction that produces a $W^+W^-$ pair of vector bosons and a fixed number of partons in the final state (typically referred to as fixed-order MC), and the MC@NLO next-to-leading order (NLO) generator [11] both interfaced to parton-shower generators. Fixed-order and NLO simulation are widely used methods for modeling electroweak processes with multiple jets. The measurement of differential cross sections as a function of the number of associated jets is particularly suited to the Tevatron because the large rate of top-quark-pair production at the LHC makes this analysis much more difficult. The production of $W$-boson pairs was first observed by the CDF experiment using Tevatron Run I data [12]. This process has since been measured by the CDF [13], D0 [14], ATLAS [15], and CMS [16,17] experiments.

This measurement uses an integrated luminosity of $9.7 \text{ fb}^{-1}$ of $p \bar{p}$ collision data collected by the CDF experiment which comprises the full Tevatron Run II data set. A 12% increase in signal acceptance is obtained by using events in which one or both $W$ bosons decay to a tau that subsequently decays to an electron or muon. Events are classified based on the multiplicity and the transverse energy of jets, where jets are the products of the parton showering and hadronization of high energy partons produced in the initial scattering process. The cross section measurement uses an artificial neural network (NN) method to distinguish signal and background events. The jet multiplicity classifications are zero, one, and two or more jets and separate NNs are used for each multiplicity class. Events with a single jet are further classified according to the energy of the jet transverse to the beam line into bins of 15–25 GeV, 25–45 GeV, and more than 45 GeV. The cross sections are extracted via a maximum-likelihood fit to the data of a weighted sum of the normalized binned NN distributions for signal and backgrounds, simultaneously over the five event classes.

The CDF experiment consists of a solenoidal spectrometer with a silicon tracker and an open-cell drift chamber surrounded by calorimeters and muon detectors [18]. The kinematic properties of particles and jets are characterized using the azimuthal angle $\phi$ and the pseudorapidity $\eta = -\ln (\tan(\theta/2))$, where $\theta$ is the polar angle relative to the nominal proton beam axis. Transverse energy, $E_T$, is defined to be $E \sin \theta$, where $E$ is the energy deposited in pointing-tower geometry electromagnetic and hadronic calorimeters. Transverse momentum, $p_T$, is the momentum component of a charged particle transverse to the beam line.

This analysis uses jets, electrons, muons and missing transverse energy. Electron and muon candidates (hereafter electrons and muons) are typically identified using the drift chamber and electromagnetic calorimeter or muon chambers, respectively. Jet candidates (jets) are measured using an iterative cone algorithm that clusters signals from calorimeter towers within a cone of $\Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2} = 0.4$. Corrections are applied to the calorimeter cluster-energy sums to account for the average calorimeter response to jets [19]. The missing transverse energy vector, $\vec{E}_T$, is defined as the opposite of the vector sum of the $E_T$ of all calorimeter towers, corrected to account for the calorimeter response to jets and to muons [3].

The experimental signature of the signal events is two leptons with opposite charge and large $E_T$. A number of SM processes result in this observed final state and are thus backgrounds. The WZ and ZZ processes can yield the same final state if one boson decays hadronically, or some decay products are not detected. Top quarks decay predominantly to a $b$ quark and a $W$ boson, which can decay leptonically,
making $t\bar{t}$ production a dominant background for events with two jets. Incorrect measurements of the energy or momentum of leptons or jets result in apparent $E_T$. This allows the large-production-rate Drell-Yan (DY) process, $Z^+\gamma^* \rightarrow \ell^+\ell^-$, where $\ell$ refers to any charged lepton, to be included in the candidate sample. A third source of background arises from the misidentification of a final-state particle, such as in $W + j$ and $W\gamma$ processes, where the photon or jet is incorrectly reconstructed as a lepton.

Events are selected from a sample containing at least one electron (muon) of $E_T(p_T) > 20$ GeV and $|\eta| < 2.0$ that satisfy an online single-lepton selection requirement in the trigger [20]. A second lepton is required to have opposite charge and $E_T(p_T) > 10$ GeV. Jets are required to have $E_T > 15$ GeV and $|\eta| < 2.5$. In order to reduce the contribution of DY events with mismeasured $E_T(p_T)$, the $E_T^{rel}$ is defined as $E_T^{rel} = E_T \sin \Delta \phi(E_T, \ell/j)$ if the azimuthal separation between the $E_T$ and the momentum vector of the nearest lepton or jet is less than $\pi$. Otherwise $E_T^{rel} = E_T$. $E_T^{rel}$ is required to be greater than 25 GeV for same-flavor lepton pairs, or 15 GeV for electron-muon events from the leptonic decays of $\tau^+\tau^-$ DY events, where the DY contribution is small. Requiring the invariant mass of the same-flavor lepton pairs ($m_{\ell\ell}$) to be outside the $Z$ mass range (between 80 and 90 GeV/$c^2$) further suppresses $Z^+\gamma^* \rightarrow \ell^+\ell^-$ events, while requiring one radial of azimuthal separation between the $E_T$ and the dilepton momentum suppresses $Z^+\gamma^* \rightarrow \ell^+\ell^-\gamma$ events. Leptons are required to be isolated, by applying the criteria that the sum of $E_T$ for calorimeter towers (or the sum of the $p_T$ of charged particles for central muons) in a cone of $\Delta R$ around the lepton is less than 10% of the electron $E_T$ (muon $p_T$). This helps to purify the lepton sample and reduce backgrounds due to misidentified particles, particularly $W + j$ and $W\gamma$. The $W\gamma$ background is further reduced by requiring $m_{\ell\ell}$ greater than 16 GeV/$c^2$. To suppress $t\bar{t}$ background we reject events with two or more jets in which any jet passes a $b$-flavored-quark identification algorithm [21].

The geometric and kinematic acceptance for the WW signal and the yields of the DY, $WZ$, $ZZ$, $W\gamma$, and $t\bar{t}$ backgrounds is determined by simulation. Backgrounds consisting of DY leptons associated with the production of zero or one jet are simulated using the PYTHIA generator [22]. DY background with two or more jets and WW samples are simulated with up to three partons with the ALPGEN generator. The contributions to the WW production from the next-to-next-to-leading-order (NNLO) process $gg \rightarrow W^+W^-$ are accounted for by reweighting the WW simulation as a function of the $W p_T$ distribution to incorporate the extra contribution predicted in Ref. [23]. The DY background is studied in a sample of low $E_T^{rel}$ events and a correction is applied to account for the mismodeling of $E_T$ in the simulation of zero-jet and one-jet events [3]. The $WZ$, $ZZ$, and $t\bar{t}$ samples are simulated with the PYTHIA generator [22]. The expected yields for simulated processes are normalized using cross section calculations performed at NNLO for $t\bar{t}$ [24], NLO for $WZ$, $ZZ$ [25] and DY [26], and fixed order for WW [10]. The $W\gamma$ background is simulated using the Baur MC generator [27], but is normalized according to a study of same-charge lepton with low $m_{\ell\ell}$ [3]. Samples are generated using the CTEQ5L [28] parton distribution functions (PDFs), and interfaced to PYTHIA for parton showering, fragmentation, and hadronization. The detector response for these processes is modeled with a GEANT3-based simulation [29], and further corrected for trigger, reconstruction, and identification efficiencies using samples of $W \rightarrow ev$ and $Z \rightarrow \ell\ell$ events [20]. The $W + j$ background is determined using collision data. The probability for a jet to be reconstructed as a lepton is measured in jet-triggered data samples and applied to a $W + j$ data sample to estimate the contribution from $W + j$ events. The estimated and observed yields for signal and background in each jet region are shown in Table I.

We use a NN method to further discriminate signal from background. Separate NNs are trained using the NEUROBAYES [30] program with simulated signal and background events for events with zero, one, and two or more jets. Inputs to the NNs are selected to exploit features of the signal and background processes. The inputs include the total energy of the event, which tends to be modest in WW events compared to $t\bar{t}$ events; the $p_T$ of the second highest $p_T$ lepton, which tends to be low in events where a jet or photon is misidentified as a lepton; missing $E_T$ that is not aligned with a lepton or jet, suggesting that it is the result of neutrino production; and, in events with two or more jets, the vector sum of the $p_T$ of the leading two jets, which tends to be greater if the jets are $t\bar{t}$ decay products. For events without jets, we use a likelihood ratio of the signal probability density over the sum of signal and background events. The estimated and observed yields for signal and background in each jet region are shown in Table I.

### Table I. Estimated and observed event yields.

<table>
<thead>
<tr>
<th>Process</th>
<th>Zero jets</th>
<th>One jet</th>
<th>Two or more jets</th>
</tr>
</thead>
<tbody>
<tr>
<td>$WZ$</td>
<td>19.5 ± 3.0</td>
<td>16.7 ± 2.3</td>
<td>4.26 ± 0.81</td>
</tr>
<tr>
<td>$ZZ$</td>
<td>13.2 ± 1.9</td>
<td>4.25 ± 0.61</td>
<td>1.33 ± 0.26</td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>3.7 ± 1.0</td>
<td>76 ± 12</td>
<td>158 ± 16</td>
</tr>
<tr>
<td>$DY$</td>
<td>150 ± 34</td>
<td>83 ± 21</td>
<td>20.2 ± 8.6</td>
</tr>
<tr>
<td>$W\gamma$</td>
<td>214 ± 27</td>
<td>44.0 ± 6.4</td>
<td>7.5 ± 1.9</td>
</tr>
<tr>
<td>$W + j$</td>
<td>685 ± 118</td>
<td>250 ± 46</td>
<td>81 ± 15</td>
</tr>
<tr>
<td>Total bk</td>
<td>1086 ± 124</td>
<td>474 ± 57</td>
<td>272 ± 26</td>
</tr>
<tr>
<td>WW</td>
<td>963 ± 108</td>
<td>224 ± 29</td>
<td>73 ± 20</td>
</tr>
<tr>
<td>WW + bk</td>
<td>2049 ± 177</td>
<td>698 ± 73</td>
<td>345 ± 39</td>
</tr>
<tr>
<td>Data</td>
<td>2090</td>
<td>682</td>
<td>331</td>
</tr>
</tbody>
</table>
background probability densities as an input to the NN [3]. The probability densities are calculated on an event-by-event basis for the signal and each background hypothesis using the observed kinematic properties of the events in leading-order (LO) matrix-element calculations [25]. The sum of final NN outputs for all event classes is shown in Fig. 1. Systematic uncertainties affecting both the normalization and shape of the signal and background NN distributions are assessed for the following sources. Uncertainties on acceptance originating from lepton-selection and trigger-efficiency measurements contribute a 4.3% uncertainty on all event yields. Acceptance uncertainties due to potential contributions from higher-order effects are evaluated as follows. For the WW contribution, the ALPGEN sample is reweighted by the $p_T$ of the WW system according to samples generated with different choices of PDF, renormalization, and factorization scales. The PDF uncertainty is evaluated with the CTEQ61 PDF error set, and ranges from 1.2% to 1.8%, increasing as a function of jet multiplicity and jet energy. The scale uncertainty is evaluated by doubling or halving the renormalization and factorization scales simultaneously and ranges from 0.5% to 24%, increasing as a function of jet multiplicity and decreasing as a function of jet energy. The scale uncertainties are anticorrelated between adjacent bins of jet energy and jet multiplicity. Between the zero-jet and one-jet bins the anticorrelation is specifically between the bin with no jets and one jet but having the lowest $E_T$. The effect of the scale uncertainty on the shape of the WW NN distributions is also evaluated. For $t\bar{t}$ contribution an uncertainty of 2.7% is assigned due to QCD effects [31]. For $WZ$ and ZZ backgrounds, which are simulated at LO, a 10% uncertainty is assigned, arising from the difference in the observed acceptance of WW events generated at LO and NLO with PYTHIA and MC@NLO, respectively. Because modeling of higher-order amplitudes can affect the extrapolation of the normalization to the predicted $W\gamma$ event yield, this uncertainty is also applied to the $W\gamma$ background. A 6.8% uncertainty is also assigned to the $W\gamma$ background due to the photon-conversion modeling. An uncertainty of 19% to 26%, increasing as both a function of jet multiplicity and jet energy, is assigned to the $t\bar{t}$ DY background to account for mismodeling of $E_T$ based on the differences in acceptance observed when varying the $E_T$ template in the study of low $E_{T,rel}$ DY events described above. Uncertainty on jet modeling for simulated backgrounds varies from 1.0% to 29%, and is anticorrelated between jet multiplicity and jet energy and generally increases as a function of jet multiplicity. For WW and DY processes, the effect on the shape of the distribution is also evaluated. For the $W + $ jets background, systematic uncertainties of 20% to 30%, depending on lepton types, are determined by calculating the misidentification probabilities at different jet-energy thresholds. Theoretical uncertainties of 6.0% are assigned to WZ and ZZ cross sections [25] and 4.3% to the $t\bar{t}$ cross section [24]. The measured luminosity, which is used to normalize the yield of all signal and background processes modeled using simulations, has an uncertainty of 5.9% [32]. Tables including the systematic uncertainties used in each jet classification and all correlations and anticorrelations are given in Ref. [3]. The sum of uncertainties from statistical and systematic sources is included for each signal and background for each jet multiplicity in Table I. The uncertainties on the background predictions are those after the fit described below.

The differential WW cross section is extracted from the NN output shapes incorporating the normalizations and systematic uncertainties of signal and background in each signal region via a binned-maximum-likelihood method. The likelihood is formed from the Poisson probabilities of observing $n_i$ events in the $i$th bin, in which $\mu_i$ are expected. Systematic uncertainties are included as normalization parameters on signal and background and subject to constraint by Gaussian terms. The likelihood is given by

$$\mathcal{L} = \prod_i \mu_i^{n_i} e^{-\mu_i} \prod_c e^{-\tilde{\lambda}_c}$$

(1)

where

$$\mu_i = \sum_k \alpha_k \prod_c \left(1 + f^c_k S_c \right) (N_k)_i,$$

(2)

For a process $k$ and a systematic effect $c$, $f^c_k$ is the fractional uncertainty assigned, and $S_c$ is a Gaussian-constrained floating parameter associated with $c$. $(N_k)_i$ is the expected number of events from process $k$ in bin $i$. Systematic uncertainties that affect the shape of the distribution are treated as correlated with the appropriate rate uncertainties. The parameter $\alpha_k$ is an overall normalization...
that is fixed to one for all processes except $WW$, for which it is determined by the fit independently for each analysis region. The likelihood function is maximized with respect to the systematic parameters $S_x$ and cross section normalizations $\alpha_{WW}$ simultaneously in all regions. The cross section in each region is calculated by multiplying the value of $\alpha_{WW}$ by the predicted cross sections calculated by ALPGEN. The total cross section is determined to be $\sigma(p\bar{p} \to W^+W^- + X) = 14.0 \pm 0.6(\text{stat})^{+1.2}_{-1.0}(\text{syst}) \pm 0.8(\text{lumi})$ pb which is consistent within one $\sigma$ with the inclusive NLO cross section prediction of $11 \pm 0.7$ pb as calculated by the MC@NLO program and the total prediction of the fixed-order program ALPGEN. The result is unfolded to the hadronic-jet level based on the results of a study of the bin-to-bin migration of events due to jet reconstruction, jet-energy scale, and jet-resolution effects as determined in simulated events. The final result is iteratively corrected to account for differences in acceptance between the reconstructed and true distributions using a Bayesian [33,34] technique. Migrations between jet-multiplicity and jet-energy bins are typically of order 10% or less. An independent training sample of simulated events is used to test the unfolding process. The correct differential cross sections are reproduced stably with a minimal number of iterations. The unfolded results are compared to ALPGEN, using the CTEQ5L PDFs and interfaced to PYTHIA for parton showering with the MLM matching algorithm [35], and MC@NLO, using the CTEQ5M PDFs and interfaced to HERWIG [36,37] for parton showering. The measured and predicted differential cross sections are shown in Table II and Fig. 2. The differential cross section measurements are consistent with both simulation predictions. The largest deviation occurs in the two-jet sample being less that two standard deviations. NNLO contributions to the $q\bar{q} \to W^+W^-$ cross section are not accounted for in the calculations and are expected to increase the predicted cross section.

In summary, the $WW$ cross section is measured in the dilepton channel both inclusively and differentially in jet multiplicity and $E_T$ using a neutral-net discriminant and binned-maximum-likelihood fit. This is the first measurement of the differential cross section for pair production of massive-vector bosons. The measured cross section, $14.0 \pm 0.6(\text{stat})^{+1.2}_{-1.0}(\text{syst}) \pm 0.8(\text{lumi})$ pb, and the differential cross sections are consistent with both the NLO and fixed-order predictions. This result indicates the suitability of using either of these theoretical techniques to study processes with multiple gauge bosons and jets. Processes of this type will be used extensively at the LHC to perform searches for non-SM physics and to investigate the nature of electroweak-symmetry breaking by studying the process of vector-boson scattering.

We thank the Fermilab staff and the technical staffs of the participating institutions for their vital contributions. This work was supported by the U.S. Department of Energy and National Science Foundation; the Italian Istituto Nazionale di Fisica Nucleare; the Ministry of Education, Culture, Sports, Science and Technology of Japan; the Natural Sciences and Engineering Research Council of Canada; the National Science Council of the Republic of China; the Swiss National Science Foundation; the A.P. Sloan Foundation; the

![FIG. 2 (color online). Measurement and predictions of $\sigma(p\bar{p} \to W^+W^- + n \text{jets})$. Values are given inclusively and differentially as functions of jet multiplicity and jet-transverse energy. Transverse energy ranges are (a) $15 < E_T < 25$ GeV, (b) $25 < E_T < 45$ GeV, and (c) $E_T > 45$ GeV.](image-url)
Bundesministerium für Bildung und Forschung, Germany; the Korean World Class University Program, the National Research Foundation of Korea; the Science and Technology Facilities Council and the Royal Society, United Kingdom; the Russian Foundation for Basic Research; the Ministerio de Ciencia e Innovación, and Programa Consolider-Ingenio 2010, Spain; the Slovak R&D Agency; the Academy of Finland; the Australian Research Council (ARC); and the EU community Marie Curie Fellowship Grant No. 302103.