Measurement of the WW+WZ Production Cross Section Using the Lepton + Jets Final State at CDF II


(CDF Collaboration)

1 Institute of Physics, Academia Sinica, Taipei, Taiwan 11529, Republic of China
2 Argonne National Laboratory, Argonne, Illinois 60439
3 University of Athens, 157 71 Athens, Greece
4 Institut de Física d’Altes Energies, Universitat Autònoma de Barcelona, E-08193, Bellaterra (Barcelona), Spain
5 Baylor University, Waco, Texas 76798
6 Istituto Nazionale di Fisica Nucleare, c/o University of Bologna, I-40127 Bologna, Italy
7 Brandeis University, Waltham, Massachusetts 02254
8 University of California, Davis, Davis, California 95616
9 University of California, Los Angeles, Los Angeles, California 90024
10 University of California, San Diego, La Jolla, California 92093
11 University of California, Santa Barbara, Santa Barbara, California 93106
12 Instituto de Fisica de Cantabria, CSIC-University of Cantabria, 39005 Santander, Spain
13 Carnegie Mellon University, Pittsburgh, PA 15213
14 Enrico Fermi Institute, University of Chicago, Chicago, Illinois 60637
15 Comenius University, 842 48 Bratislava, Slovakia; Institute of Experimental Physics, 040 01 Kosice, Slovakia
16 Joint Institute for Nuclear Research, RU-141980 Dubna, Russia
17 Duke University, Durham, North Carolina 27708
18 Fermi National Accelerator Laboratory, Batavia, Illinois 60510
19 University of Florida, Gainesville, Florida 32611
20 Laboratori Nazionali di Frascati, Istituto Nazionale di Fisica Nucleare, I-00044 Frascati, Italy
21 University of Geneva, CH-1211 Geneva 4, Switzerland
22 Glasgow University, Glasgow G12 8QQ, United Kingdom
23 Harvard University, Cambridge, Massachusetts 02138
24 Division of High Energy Physics, Department of Physics, University of Helsinki and Helsinki Institute of Physics, FIN-00014, Helsinki, Finland
25 University of Illinois, Urbana, Illinois 61801
26 The Johns Hopkins University, Baltimore, Maryland 21218
We report two complementary measurements of the diboson (WW + WZ) cross section in the final state consisting of an electron or muon, missing transverse energy, and jets, performed using pp collision data at √s = 1.96 TeV collected by the Collider Detector at Fermilab. The first method uses the dijet invariant mass distribution while the second method uses more of the kinematic information in the event through matrix-element calculations of the signal and background processes and has a higher sensitivity. The result from the second method has a signal significance of 5.4 and is the first observation of WW + WZ production using this signature. Combining the results from both methods gives \( \sigma_{WW + WZ} = 16.0 \pm 3.3 \) pb, in agreement with the standard model prediction.

PACS numbers: 14.80.Bn, 14.70.Fm, 14.70.Hp, 12.15.Ji
Measurements involving heavy vector boson pairs (WW, WZ, and ZZ) are important tests of the electroweak sector of the standard model (SM). Deviations of the production cross section from predictions could arise from anomalous triple gauge boson interactions [1] or from new resonances decaying to vector bosons. Furthermore, the topology of diboson events is similar to that of events in which a Higgs boson is produced in association with a W or a Z, allowing diboson measurements to provide an important step towards future measurements of Higgs boson production.

Diboson production has been observed at the Tevatron in channels in which both bosons decay leptonically [2, 3]. Extraction of the diboson signal in hadronic channels is more challenging because of significantly larger backgrounds. In addition, due to limited detector resolution, it is difficult to distinguish hadronically decaying W bosons from Z bosons. We report on two measurements of the cross section, $\sigma(pp \rightarrow W + WZ)$, that use different techniques applied to the leptonic decay of one W and the hadronic decay of the associated W or Z ($W/WZ \rightarrow \ell\nu q q$, where $\ell$ represents an electron or muon). Our result represents the first observation of this signal in the lepton + jets channel. Evidence has previously been reported by the D0 collaboration [4], and the CDF collaboration set a limit on its cross section times branching ratio [5]. In addition the CDF collaboration has reported observation of $WW + WZ + ZZ$ in a different hadronic channel with large missing transverse energy and jets [6].

The first method uses the invariant mass of the two-jet system ($M_{jj}$) to extract a signal peak from data corresponding to $3.9$ fb$^{-1}$ of $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV. The second method takes advantage of more kinematic information in the event by constructing a discriminant based on calculations of the differential cross sections of the signal and background processes. This so-called matrix-element (ME) method has been employed in a search for a low-mass Higgs produced in association with a W boson [7] and in a measurement of single top production [8]. It is expected to achieve greater discriminating power and here uses data corresponding to an integrated luminosity of 2.7 fb$^{-1}$.

The aspects of the CDF II detector [9] relevant to these analyses are briefly described here. The tracking system is composed of silicon microstrip detectors and an open-cell drift chamber inside a 1.4 T solenoid. Electromagnetic lead-scintillator and hadronic iron-scintillator sampling calorimeters segmented in a projective geometry surround the tracking detectors. A central calorimeter covers a pseudorapidity range of $|\eta| < 1.1$ while plug calorimeters extend the acceptance into the region $1.1 < |\eta| < 3.6$. Outside of the calorimeters are muon detectors composed of scintillators and drift chambers. Cherenkov counters around the beam pipe and in the plug calorimeters count the inelastic collisions per bunch crossing and provide the luminosity measurement.

Data samples common to both analyses use trigger selections requiring a central electron (muon) with $E_T(p_T) > 18$ GeV. The ME method utilizes an additional sample derived from a trigger requiring two jets and large missing transverse energy ($E_T$) [10].

Offline we select events with electron (muon) candidates with $E_T(p_T) > 20$ GeV, and with $E_T$, jet, and other kinematic requirements chosen differently for the two methods. Jets are clustered using a fixed-cone algorithm with radius $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0.4$ and their energies are corrected for detector effects [11]. Cosmic ray and photon conversion candidates are identified and removed.

Further event selection requirements are made to reduce backgrounds and the sensitivity to systematic uncertainties. In the $M_{jj}$ method, we require events to have $E_T > 25$ GeV, at least two jets with $E_T > 15$ GeV and $|\eta| < 2.4$, and the dijet vector boson candidate to have $p_T > 40$ GeV/$c$. As a result of these selection criteria, the $M_{jj}$ distribution for background is smoothly falling in the region where the signal is expected to peak. The invariant mass of the dijet vector boson candidate, $M_{jj}$, is evaluated from the two most energetic jets. Additional requirements are made to reduce backgrounds and improve the Monte Carlo modeling of event kinematics: the transverse mass of the lepton and $E_T$ system ($M_T(W)$ [10]) must be greater than 30 GeV/$c^2$, and the two most energetic jets must be separated by $|\Delta\eta| < 2.5$.

In the ME method, we require events to have $E_T > 20$ GeV and exactly two jets with $E_T > 25$ GeV and $|\eta| < 2.0$. Additional selection criteria to reduce backgrounds and achieve good modeling of the quantities used.

---

Antwerp, Belgium, 4University of Bristol, Bristol BS8 1TL, United Kingdom, 5University of California Irvine, Irvine, CA 92697, 6University of California Santa Cruz, Santa Cruz, CA 95064, 7Cornell University, Ithaca, NY 14855, 8University of Cyprus, Nicosia CY-1678, Cyprus, 9University College Dublin, Dublin 4, Ireland, 10University of Edinburgh, Edinburgh EH9 3JZ, United Kingdom, 11University of Fukai, Fukui City, Fukui Prefecture, Japan 910-0017 12Kinki University, Higashi-Osaka City, Japan 577-8502 13Universidad Iberoamericana, Mexico D.F., Mexico, 14University of Iowa, Iowa City, IA 52242, 15Kansas State University, Manhattan, KS 66506 16Queen Mary, University of London, London, E1 4NS, England, 17University of Manchester, Manchester M13 9PL, England, 18Muons, Inc., Batavia, IL 60510, 19Nagasaki Institute of Applied Science, Nagasaki, Japan, 20University of Notre Dame, Notre Dame, IN 46556, 21University of Oviedo, E-33007 Oviedo, Spain, 22Texas Tech University, Lubbock, TX 79096, 23IFIC(CSIC-Universitat de Valencia), 56071 Valencia, Spain, 24Universidad Tecnica Federico Santa Maria, 110 Valparaiso, Chile, 25University of Virginia, Charlottesville, VA 22906 26Bergische Universitat Wuppertal, 42097 Wuppertal, Germany, 27Yarmouk University, Irbid 211-63, Jordan 28On leave from J. Stefan Institute, Ljubljana, Slovenia,
in the matrix element calculation include the rejection of events with either an additional jet of $E_T > 12$ GeV or a second charged lepton. The latter reduces $Z+$jets, $t\bar{t}$, and leptonic diboson backgrounds. For events with an electron candidate, there is a significant background from production of multiple jets (multi-jet in the following) by quantum chromodynamical (QCD) processes, where the electron is faked by a hadronic jet. The ME method deals with this background by applying stringent selection criteria, while the $M_{jj}$ method assigns a systematic uncertainty to the background shape. The reduction of the multi-jet QCD background in the ME analysis is achieved by raising the $E_T$ cut to 40 GeV, requiring $M_T(W) > 70$ GeV/$c^2$, and imposing additional cuts on the angles between the jets, the lepton, and the $E_T$ [12]. There is a less stringent requirement of $M_T(W) > 10$ GeV/$c^2$ imposed on muon events to reduce the QCD background in that channel.

After these selections for both methods, the dominant background to the diboson signal is a $W$ boson produced with accompanying jets ($W+$jets), where the $W$ decays leptonically. Smaller but non-negligible backgrounds come from QCD multi-jet (where one jet mimics a lepton signature), $Z+$jets, $t\bar{t}$, and single top production. QCD multi-jet events are modeled using data with loosened lepton selection criteria. All other signal and background processes are modeled using event generators and a GEANT-based CDF II detector simulation. The diboson signals as well as the $t\bar{t}$ and single top backgrounds are simulated using the PYTHIA event generator [14]. The $W+$jets and $Z+$jets backgrounds are simulated using the tree-level event generator ALPGEN [15], with an interface to PYTHIA providing parton showering and hadronization.

The normalization of the $Z+$jets background is based on the measured cross section while for $t\bar{t}$ and single top backgrounds the NLO predicted cross section is used [16]. The efficiencies for the $Z+$jets, $t\bar{t}$, and single top backgrounds are estimated from simulation. The normalization of the QCD background is estimated by fitting the $E_T$ spectrum in data to the sum of all contributing processes, where the QCD and $W+$jets normalizations float in the fit. In the final signal extractions from both methods, the multi-jet QCD background is Gaussian constrained to the result of this $E_T$ fit and the $W+$jets background is left unconstrained.

We now describe the methodology and results from each technique. In the $M_{jj}$ method we extract the signal fraction from the data by performing a $\chi^2$ fit to the dijet invariant mass spectrum, separately for electron and muon events. Templates of $M_{jj}$ distributions are constructed with the multi-jet QCD background, the signal $WW + WZ$ processes, and the sum of the electroweak backgrounds ($Z+$jets, $W+$jets, and $t\bar{t}$ production).

Figure 1 shows the fit results superimposed on data after the electron and muon samples are combined. Also shown is the data $M_{jj}$ distribution after having subtracted the estimated background, superimposed on the signal Monte Carlo shape extracted from the fit. Combining the two $\chi^2$ fit results we get a total of $1079 \pm 232$(stat) $\pm 86$(syst) $WW/WZ \rightarrow \ell\nu jj$ events, of which about 60% are muon events and 40% are electron events. The observed significance is $4.6\sigma$ where $4.9\sigma$ is expected, which is obtained by combining the separate results from the electron and muon channels. The resultant $WW + WZ$ production cross section measurement is $\sigma_{WW+WZ} = 14.4 \pm 3.1$(stat) $\pm 2.2$(syst) pb. The sources of systematic uncertainty in this measurement are discussed together with those from the ME method below.

**TABLE I:** Expected and observed event yields after the ME method selection in $2.7$ fb$^{-1}$ of data.

<table>
<thead>
<tr>
<th>Process</th>
<th>Predicted event yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>$WW$ signal</td>
<td>446 $\pm$ 29</td>
</tr>
<tr>
<td>$WZ$ signal</td>
<td>79 $\pm$ 6</td>
</tr>
<tr>
<td>$W+$jets</td>
<td>10175 $\pm$ 305</td>
</tr>
<tr>
<td>$Z+$jets</td>
<td>584 $\pm$ 88</td>
</tr>
<tr>
<td>QCD multi-jet</td>
<td>283 $\pm$ 113</td>
</tr>
<tr>
<td>$t\bar{t}$ + single top</td>
<td>241 $\pm$ 29</td>
</tr>
<tr>
<td><strong>Observed</strong></td>
<td>11812</td>
</tr>
</tbody>
</table>

In the ME method a probability density $P(x)$ that an event was produced by a given process is determined using the standard model differential cross section for that
process. For an event with measured quantities \( x \), we integrate the appropriate differential cross section \( \sigma(y) \) over the partonic quantities \( y \) convolved with the parton distribution functions (PDFs), \( f(y_1), f(y_2) \), and over transfer functions describing detector resolution effects, \( W(y, x) \):

\[
P(x) = \frac{1}{\sigma} \int \sigma(y) dq_1 dq_2 f(y_1) f(y_2) W(y, x).
\] (1)

We use the CTEQ5L PDF parameterization [17]. \( W(x, y) \) is a mapping of measured jet energy to partonic energy derived using the full detector simulation, while the lepton momenta and jet angles are assumed to be measured exactly. The integration is performed over the energy of the partons and the longitudinal momentum of the neutrino. The matrix element is calculated with tree-level diagrams from madgraph [18]. Event probability densities \( P_x \) are calculated for the signal processes \( WW \) and \( WZ \), as well as for \( W \) plus two parton and single top background processes. The event probabilities are combined into an event probability discriminant:

\[
EPD = P_{\text{signal}}/(P_{\text{signal}} + P_{\text{background}}),
\]

where \( P_{\text{signal}} = P_{\text{WW}} + P_{\text{WZ}} \) and \( P_{\text{background}} = P_{\text{W+jets}} + P_{\text{single top}} \). We make templates of the \( EPD \) for all signal and background processes and ultimately extract the signal using a fit of the observed \( EPD \) distribution to a sum of the signal and background templates. The expected event yields are as shown in Table I for the ME method’s event selection.

Figure 2 shows the dijet mass in bins of \( EPD \). Most of the background events have low \( EPD \). Events with \( EPD > 0.25 \) have a dijet mass peak close to the expected \( W/Z \) resonance, and the signal-to-background ratio improves with increasing \( EPD \).

![FIG. 2: \( M_{jj} \) for events with (a) \( EPD < 0.25 \), (b) \( 0.25 < EPD < 0.5 \), (c) \( 0.5 < EPD < 0.75 \), and (d) \( EPD > 0.75 \).](image)

Before comparing the observed \( EPD \) to the prediction, we validate the Monte Carlo modeling of the quantities that enter the matrix element calculation. We compare the observed distributions to the predicted ones in control regions with very little signal and also in the signal-rich region. The different regions are chosen according to the invariant mass of the two-jet system (\( M_{jj} \)): the signal-rich region has \( 55 < M_{jj} < 120 \text{ GeV} \) and the control regions cover the rest of the \( M_{jj} \) range. We also check the modeling of the properties (mass, \( p_T \), and \( \eta \)) of the leptonic \( W \) boson and the hadronic \( W \) or \( Z \) boson candidate. All of these quantities are well described by the simulation for our event selection. There is a small discrepancy in the description of \( M_{jj} \) in the control regions, as is visible in the low-\( EPD \) region of Figure 2. Associated with this discrepancy we assign a systematic mismodeling uncertainty which is derived in the control regions and extrapolated through the signal region. This uncertainty has a negligible effect on the results, because most background events lie in the first few bins of the \( EPD \) distribution. Small changes in modeling of those background events do not change the shape of the \( EPD \).

The observed and predicted \( EPDs \) are shown in Figure 3. We use a binned-likelihood fit of the observed \( EPD \) to a sum of templates, testing both a background-only hypothesis and a signal-plus-background \( (s + b) \) hypothesis. Systematic uncertainties, discussed further below, are included in the fit as constrained parameters. We perform pseudo experiments to calculate the probability (\( p \)-value) that the background-only discriminant fluctuates up to the observed result (observed \( p \)-value) and up to the median expected \( s + b \) result (expected \( p \)-value). We observe a \( p \)-value of \( 2.1 \times 10^{-7} \), corresponding to a signal significance of \( 5.4\sigma \), where \( 5.1\sigma \) is expected. The observed \( WW + WZ \) cross section is \( \sigma_{WW+WZ} = 17.7 \pm 3.1(\text{stat}) \pm 2.4(\text{syst}) \text{ pb} \).

![FIG. 3: Observed \( EPD \) distribution superimposed on distribution expected from simulated processes.](image)

We consider several sources of systematic uncertainty in both methods, taking into account their effect on both the signal acceptance and the shape of the background and signal templates. The uncertainty on the normaliza-
tion of the backgrounds is taken as part of the statistical uncertainty. In the $M_{jj}$ method the largest systematic uncertainties are due to the modeling of the electroweak and QCD shapes, about 8% and 6% respectively. In the ME method the uncertainty in the jet energy scale is the largest systematic uncertainty, at about 10%, which includes contributions both from the signal acceptance and from the shapes of the signal templates. In the $M_{jj}$ method this uncertainty is about 6%. Both methods include an uncertainty of about 5% due to initial and final state radiation and a 6% uncertainty on the integrated luminosity. Smaller contributions arise from PDFs, jet energy resolution, the factorization and renormalization scales used in the $W+$jets simulation, and trigger and lepton identification efficiencies.

One measure of how the two methods are correlated is the expected overlap of $WW + WZ$ signal. Accounting for the different integrated luminosities used, 15% of the signal in the $M_{jj}$ analysis is common to that in the EPD analysis. Conversely, 29% of the signal in the EPD analysis is common to that in the $M_{jj}$ analysis. This corresponds to a statistical correlation of about 21%. If we assume the systematic uncertainties are 100% correlated, then the total correlation between the two analyses is 49%, leading to a combined [19] result of $\sigma_{WW+WZ} = 16.0 \pm 3.3(\text{stat} + \text{syst})$ pb. Because the total uncertainties on the two input measurements are so similar, the combined central value does not depend significantly on the correlation assumed. The total uncertainty in the combined result increases with increasing correlation and we quote the value assuming maximum possible correlation. The signal overlap with the CDF $WW + WZ + ZZ$ observation in the $E_T$+jets channel [6] is also studied. While that analysis requires much larger $E_T$, it does not veto events with identified leptons. We found that about 15% of the $WW + WZ$ signal from the $E_T$+jets analysis appears in the analyses presented here.

In summary, we observe $WW + WZ$ production in the lepton plus jets plus $E_T$ final state. We perform two searches: one seeking a resonance on top of a smoothly falling dijet mass distribution, and another building a discriminant using a matrix element technique. The combined $WW + WZ$ cross section from these two methods is measured to be $\sigma_{WW+WZ} = 16.0 \pm 3.3(\text{stat} + \text{syst})$ pb, in good agreement with the SM prediction of 16.1 ± 0.9 pb [20]. Measurements of these diboson processes are a necessary step toward validating Higgs boson search techniques at the Tevatron.

We thank the Fermilab staff and the technical staffs of the participating institutions for their vital contributions. This work was supported by the U.S. Department of Energy and National Science Foundation; the Italian Istituto Nazionale di Fisica Nucleare; the Ministry of Education, Culture, Sports, Science and Technology of Japan; the Natural Sciences and Engineering Research Council of Canada; the National Science Council of the Republic of China; the Swiss National Science Foundation; the A.P. Sloan Foundation; the Bundesministerium für Bildung und Forschung, Germany; the World Class University Program, the National Research Foundation of Korea; the Science and Technology Facilities Council and the Royal Society, UK; the Institut National de Physique Nucleaire et Physique des Particules/CNRS; the Russian Foundation for Basic Research; the Ministerio de Ciencia e Innovación, and Programa Consolider-Ingenio 2010, Spain; the Slovak R&D Agency; and the Academy of Finland.

[10] We use a cylindrical coordinate system with its origin in the center of the detector, where $\theta$ and $\phi$ are the polar and azimuthal angles, respectively, and pseudorapidity is $\eta = -\ln \tan(\theta/2)$. The missing $E_T$ ($\vec{E}_T$) is defined by $\vec{E}_T = -\sum \vec{p}_i n_i$, where $\vec{n}_i$ is a unit vector perpendicular to the beam axis and pointing at the $i^{th}$ calorimeter tower. $\vec{E}_T$ is corrected for high-energy muons and also jet energy corrections. We define $E_T = |\vec{E}_T|$. The transverse momentum $p_T$ is defined to be $p\sin \theta$. The transverse mass of the W is defined as $M_T(W) = \sqrt{2p_T^2 E_T (1 - \cos(\Delta \phi_{iw}))}$.


