First Observation of Vector Boson Pairs in a Hadronic Final State at the Tevatron Collider

We present the first observation in hadronic collisions of the electroweak production of vector boson pairs (VV, \( V= W, Z \)) where one boson decays to a dijet final state. The data correspond to 3.5 fb\(^{-1}\) of integrated luminosity of \( p\bar{p} \) collisions at \( \sqrt{s} = 1.96 \) TeV collected by the CDF II detector at the Fermilab Tevatron. We observe 1516 \( \pm 239 \) (stat) \( \pm 144 \) (syst) diboson candidate events and measure a cross section \( \sigma(p\bar{p} \rightarrow VV + X) \) of 18.0 \( \pm 2.8 \) (stat) \( \pm 2.4 \) (syst) \( \pm 1.1 \) (lumi) pb, in agreement with the expectations of the standard model.
The production of heavy gauge boson pairs (WW, WZ, or ZZ) in $p\bar{p}$ collisions has been observed in the fully leptonic final states at the Fermilab Tevatron collider [1, 2]. Diboson production has not yet been conclusively observed in $p\bar{p}$ collisions in decay channels involving hadrons [3]; however, evidence for diboson decays into an $\ell\nu q\bar{q}^{(l)}$ final state ($l=e, \mu, \tau; q=u, d, s, c, b$) has been recently presented by DØ collaboration [4].

Measurements of diboson production cross sections provide tests of the self-interactions of the gauge bosons. Deviations from the standard model (SM) prediction for the production rates could indicate new physics [5]. Furthermore, given that diboson production is topologically similar to associated Higgs boson production, $pp\to VH+X$ ($V=W, Z$), the analysis techniques described in this Letter are important for Higgs boson searches.

Here we present the first observation at a hadron collider of diboson production with one boson decaying into leptons and the other into hadrons. The analysis is performed on a sample of events with large transverse momentum imbalance ($E_T$) [6, 7] and two jets whose invariant mass can be reconstructed. This signature is sensitive to not only $\ell\nu q\bar{q}^{(l)}$, but also to $\nu\bar{\nu} q\bar{q}^{(l)}$ decays because we do not explicitly require presence of identified charged leptons. The limited dijet mass resolution results in a significant overlap of the $W\to q\bar{q}'$ and the $Z\to q\bar{q}$ dijet mass peaks, and therefore the combination of the three diboson signals, WW, WZ, and ZZ, is considered.

We analyze a dataset of $p\bar{p}$ collisions corresponding to an integrated luminosity of $3.5\text{fb}^{-1}$ collected with the CDF II detector at the center-of-mass energy of 1.96 TeV. The CDF II detector is described in detail elsewhere [8]. Here the components that are relevant to this search are briefly discussed. Surrounding the beam pipe, there is a tracking system consisting of a silicon microstrip detector, a cylindrical drift chamber, and a solenoid that provides a 1.4 T magnetic field along the beam axis. The central and plug calorimeters, which respectively cover the pseudorapidity regions of $|\eta|<1.1$ and $1.1<|\eta|<3.6$ [6], surround the tracking system with a projective tower geometry. The calorimeters are composed of inner electromagnetic and outer hadronic sections that consist of lead-scintillator and iron-scintillator, respectively. In the central region, the calorimeter consists of 48 modules, segmented into towers of granularity $\Delta\eta \times \Delta\phi \approx 0.1 \times 0.26$. The energy resolution of the central electromagnetic calorimeter for electrons is $\sigma(E_T)/E_T = 13.5\%/\sqrt{E_T(\text{GeV})} \mp 1.5\%$ [9], while the energy resolution of the central hadron calorimeter for charged pions that do not interact in the electromagnetic section is $\sigma(E_T)/E_T = 50\%/\sqrt{E_T(\text{GeV})} \mp 3\%$ [10], where $E_T$ is the transverse energy [6]. The plug electromagnetic (PEM) and plug hadron (PHA) calorimeters are identically segmented into 480 towers whose sizes are $\Delta\eta \times \Delta\phi \approx 0.1 \times 0.13$ for $|\eta|<1.8$ to $\Delta\eta \times \Delta\phi \approx 0.6 \times 0.26$ for the towers centered at $|\eta|=3.3$. The corresponding PEM and PHA energy resolutions are $\sigma(E_T)/E_T = 14.4\%/\sqrt{E_T(\text{GeV})} \mp 0.7\%$ and $\sigma(E_T)/E_T = 74\%/\sqrt{E_T(\text{GeV})} \mp 4\%$ [10] for charged pions that do not interact in the electromagnetic section. A system of Cherenkov counters [12], located around the beam pipe and inside the plug calorimeters, is used to measure the number of inelastic $p\bar{p}$ collisions per bunch crossing and thereby the luminosity.

The diboson signal (WW, WZ, and ZZ) is simulated using the PYTHIA 6.2 Monte Carlo generator [13]. The most significant backgrounds to the diboson signal are $W(\ell\nu)+\text{jets}$, $Z(\nu\bar{\nu})+\text{jets}$, and QCD multijet production (in the following referred to as multijet background or MJB). Other less significant backgrounds include $Z(l\bar{l})+\text{jets}$, $t\bar{t}$, and single $t$-quark production. The $W+jets$ backgrounds are simulated using the fixed-order matrix element generator ALPGEN v2.1 [14] which is interfaced with PYTHIA v6.3 to simulate parton showering and fragmentation, the underlying event, and additional $p\bar{p}$ interactions in the same bunch crossing. The $Z$+jets and $t$-quark production processes are simulated with
PYTHIA V6.2. The detector response in all Monte Carlo samples is modeled by a GEANT-based CDF II detector simulation [15]. The MJB does not typically result in signatures of large intrinsic \( E_T \). However, when jet energy is not measured accurately an event may be reconstructed with large \( E_T \) and pass the analysis selection criteria. Because of the large multijet production rate, this can still be a significant background in a \( E_T \)+jets based analysis. The MJB is determined from the data. All other background predictions are normalized using next-to-leading-order (NLO) calculations for SM cross sections.

The selection of signal events proceeds as follows: a set of \( E_T \)-based triggers select events with a variety of \( E_T \) and jet requirements. All these triggers have benefited significantly from the calorimeter trigger upgrade completed in 2007 [16]. The majority (94\%) of events satisfy the inclusive \( E_T \) trigger which requires \( E_T > 45 \text{ GeV} \). Jets are reconstructed in the calorimeter using the JETCLU cone algorithm [17] with a cone radius of 0.4 in (\( \eta, \phi \)) space. The energy measured by the calorimeter is corrected for effects that distort the true jet energy [18]. Such effects include the non-linear response of the calorimeter to particle energy, loss of energy in uninstrumented regions of the detector, spectator interactions, and energy radiated outside of the jet cone. We select events that have \( E_T > 60 \text{ GeV} \) [19] and exactly two jets with \( E_T > 25 \text{ GeV} \) and \( |\eta| < 2.0 \). This ensures a trigger efficiency of 96\%±2\% on signal. We search for diboson production in the dijet mass range \( 40 < M_{jj} < 160 \text{ GeV}/c^2 \). The lower edge of the mass range is chosen to ensure that events are on the trigger efficiency plateau with respect to dijet mass. In addition to the requirements discussed above, the electromagnetic fraction of the total energy for each of the two jets is required to be less than 90\% to ensure that electrons and photons are not counted as jets.

In order to suppress the MJB we use a \( E_T \) resolution model to distinguish true \( E_T \) originating from undetected neutrinos from fake \( E_T \) due to jets that are not measured accurately. The \( E_T \) significance is a dimensionless quantity based on the energy resolution of the jets, on soft unclustered particles, and on the event topology; details of its definition can be found in Ref. [20]. The \( E_T \) significance is typically low when \( E_T \) arises from mismeasurement. In addition to having a small significance, the \( \vec{E}_T \) will often be aligned with a jet. We select events with \( E_T \) significance larger than 4 and azimuthal angle between \( \vec{E}_T \) and nearest jet (\( \Delta \phi_{\vec{E}_T}^{\text{jet}} \)) greater than 0.4 radians.

Finally, we apply several requirements that suppress contamination due to cosmic-ray, beam-related, and other non-collision backgrounds. Events are required to have at least one reconstructed vertex formed by charged particle tracks. The transverse energies of all calorimeter towers are calculated with respect to the \( z \) position of the primary vertex with the largest \( \sum p_T \) of associated tracks. The electromagnetic fraction of the total event energy has to be larger than 45\% in order to reduce beam-related backgrounds. The arrival time of both leading jets as measured by the electromagnetic shower timing system [21] has to be consistent with the pp collision time. The remaining non-collision background has a smooth \( M_{jj} \) distribution and accounts for less than 90\% of the 44,910 selected events.

The shape and normalization of the MJB are determined from the data. A vector, \( \vec{p}_T \), analogous to the calorimeter-based \( \vec{E}_T \), is constructed from the vector sum of the transverse momenta of particles measured in the tracking system, and is largely uncorrelated to \( \vec{E}_T \) for events where jets are not reconstructed accurately. In the absence of \( \vec{E}_T \) arising from mismeasurement in the calorimeter, the \( \vec{E}_T \) and \( \vec{p}_T \) will be aligned in most events. The MJB is expected to be the dominant background component at larger values of \( \Delta \phi(\vec{E}_T, \vec{p}_T) \). The dijet mass shape and normalization for the remaining MJB contribution in the sample is found by selecting events with \( \Delta \phi(\vec{E}_T, \vec{p}_T) > 1.0 \) and subtracting out the non-MJB backgrounds. The normalization is scaled up to account for the MJB contamination in the region \( \Delta \phi(\vec{E}_T, \vec{p}_T) < 1.0 \). The shape of the MJB is fit to an exponential in \( M_{jj} \) to derive a dijet mass template. The MJB shapes of \( M_{jj} \) and \( \Delta \phi(\vec{E}_T, \vec{p}_T) \) distributions are verified with a large statistics MC sample.

The signal extraction is performed using a minimization of the unbinned extended negative log likelihood with RooFt program [22]. Three \( M_{jj} \) template distributions are used in the fit: the first is \( V+\)jets and \( t\)-quark production (in the following referred to as “electroweak” (EWK) backgrounds) and is taken from Monte Carlo simulation; the second is the MJB template, where the slope and normalization are Gaussian constrained to their previously measured values; the third template describes the signal. The signal shape is comprised of the \( WW, WZ, \) and \( ZZ \) distributions. This template is obtained from a Gaussian+polynomial fit to the signal Monte Carlo simulation where the mean and the width of the Gaussian distribution are linearly dependent on the jet energy scale (JES).

To assess the effect of systematic uncertainties on the measurement, we address separately two classes of sources: those that affect the signal extraction procedure, and those that affect the signal acceptance in the cross section calculation. The signal extraction systematic uncertainties come from uncertainties in signal and background shapes. The shape uncertainties take into account the effect of jet energy resolution (JER), JES, MJB shape, and the shape of the EWK background. The jet energy scale, and the shape and the normalization of MJB, are treated as nuisance parameters in the fit and Gaussian constrained to their independently measured values. These uncertainties are therefore accounted for.
in the statistical uncertainty of the extraction.

The shape uncertainty for the EWK background is determined by using $\gamma+$jets data [23] as an alternative background model in the $M_{jj}$ fit. All major non-MJB backgrounds include a gauge boson accompanied by jets. There are similarities between the $\gamma+$jets and $V+$jets production; however, due largely to the mass difference between the $\gamma$ and the $W/Z$, the kinematics is not identical. To take this into account the $\gamma+$jets data are weighted by the ratio of the dijet mass distributions of the EWK background MC samples to $\gamma+$jets PYTHIA sample. We use these adjusted $\gamma+$jets data to determine a systematic uncertainty on the EWK $M_{jj}$ template. Selection cuts applied to $\gamma+$jets events are not identical to those applied to the $E_T^{\gamma}+$jets sample. For example, the $Z$ decay into neutrinos will register as $E_T^{\gamma}$ in the detector, while the photon $E_T$ will be measured in the calorimeter. For this reason we cut on the vector sum of the photon $E_T$ and any $E_T^{\gamma}$ present in $\gamma+$jets events at 60 GeV, treating this sum as analogous to $E_T^{\gamma}$ in $V+$jets events. A further consideration in the construction of the $\gamma+$jets template is the effect of $\gamma+$V events, as these events will cause a peak in the $\gamma+$jets dijet mass distribution. We subtract this contribution using the $\gamma+$V PYTHIA sample. Finally, we perform two signal extraction fits using the default EWK and $\gamma+$jets templates, respectively. The uncertainty due to the shape of the EWK background is then estimated as the difference in the results obtained from these two fits. The described method accounts for a combined effect of JES, JER, and modeling of jets in MC on the EWK $M_{jj}$ template.

The uncertainty associated with the JES is the dominant source of systematic uncertainty on the acceptance, and, therefore, the cross section. The JES affects several of the variables used in the event selection. The effect of the JES systematic uncertainty is quantified by varying the jet energies in the signal Monte Carlo simulation to account for the $\pm 1\sigma$ variations of the JES. Other less significant sources of systematic uncertainty that affect the measured cross section are jet energy resolution, initial and final state radiation (ISR/FSR), and parton distribution functions (PDF). A summary of all sources of systematic uncertainty is presented in Table I.

The measured yields for signals and backgrounds are given in Table II. We extract $1516\pm239(stat)\pm144(syst)$ signal events which correspond to a cross section of $18.0\pm2.8(stat)\pm2.4(syst)\pm1.1(lumi)$ pb, in agreement with the SM prediction of $16.8\pm0.5$ pb obtained using the mcfm v5.4 program [24] with CTEQ6.1M PDFs [25]. Based on the MC simulation, the acceptances for the $WW$, $WZ$, and $ZZ$ production are 2.5%, 2.6%, and 2.9%, respectively. In the calculation of the combined diboson cross section, we assume that each signal process contributes proportionally to its predicted SM cross section: 11.7 pb for $WW$, 3.6 pb for $WZ$, and 1.5 pb for $ZZ$.

Figure 1 shows a comparison between the observed $\Delta \phi^{jet}$ distribution and the MJB and EWK (signal+background) components. This distribution provides a strong consistency check on our MJB model. Figure 2 shows the fit result and a comparison between the expected signal and data after background subtraction. We bin the data as in Fig. 2 and obtain a $\chi^2$ of 9.4 for 9 degrees of freedom corresponding to a p-value of 40%.

In summary, we use the $E_T^{\gamma}+$jets final state to measure the $W^+W^-+jets$ cross section in $pp$ collisions at $\sqrt{s} = 1.96$ TeV to be $18.0\pm2.8(stat)\pm2.4(syst)\pm1.1(lumi)$ pb. This is consistent with the SM prediction of $16.8\pm0.5$ pb.

To assess the strength of the observed signal, the effect of parameter variations due to all relevant sources of uncertainty are studied by comparing the likelihood of background-only fit with the full fit result, and converting the difference into significance numbers. We thus measure that the signal corresponds to a significance of at least 5.3 standard deviations from the background-only hypothesis.

### Table I: The systematic uncertainties and their effect on the number of extracted signal events, the acceptance, and the cross section. All systematics are added in quadrature.

<table>
<thead>
<tr>
<th>Systematic</th>
<th>% uncert.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extraction</td>
<td></td>
</tr>
<tr>
<td>EWK shape</td>
<td>7.7</td>
</tr>
<tr>
<td>Resolution</td>
<td>5.6</td>
</tr>
<tr>
<td>Total</td>
<td>9.5</td>
</tr>
<tr>
<td>Acceptance</td>
<td></td>
</tr>
<tr>
<td>$E_T$ resolution model</td>
<td>1.0</td>
</tr>
<tr>
<td>Trigger inefficiency</td>
<td>2.2</td>
</tr>
<tr>
<td>ISR/FSR</td>
<td>2.5</td>
</tr>
<tr>
<td>PDF</td>
<td>2.0</td>
</tr>
<tr>
<td>Total</td>
<td>9.0</td>
</tr>
<tr>
<td>Luminosity</td>
<td>5.9</td>
</tr>
<tr>
<td>Total</td>
<td>14.4</td>
</tr>
</tbody>
</table>

### Table II: Value of parameters in the model used to fit the $M_{jj}$ distribution. $1516\pm239(stat)$ signal events are extracted from the 44,910 data events which pass our selection cuts. The jet energy scale is also extracted from the fit and agrees well with the default value (1.0) as measured from calibrations [18].

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Fitted value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jet energy scale, JES</td>
<td>$0.985 \pm 0.019$</td>
</tr>
<tr>
<td>Yield of EWK background events</td>
<td>$36,140 \pm 1230$</td>
</tr>
<tr>
<td>Yield of MJB events</td>
<td>$7249 \pm 1130$</td>
</tr>
<tr>
<td>Yield of diboson candidates</td>
<td>$1516 \pm 239$</td>
</tr>
</tbody>
</table>

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 Korean Science and Engineering Foundation and the Korean Research Foundation; 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.

[6] The detector is cylindrically symmetric around the proton beam axis which is oriented in the positive z direction. The polar angle, θ, is measured from the origin of the coordinate system at the center of the detector with respect to the z axis. The pseudorapidity, transverse energy, and transverse momentum are defined as y = −ln tan(θ/2), ET = E sin(θ), and pT = E sin(θ), respectively.
[7] The missing ET is defined by \( ET = |\vec{E}_T| = \sum_i E_T^i \hat{\mathbf{n}}_i \), where \( \hat{\mathbf{n}}_i \) is a unit vector perpendicular to the beam axis and pointing at the \( i^{th} \) calorimeter tower. The sum \( E_T \) is defined by \( \sum E_T = \sum_i E_T^i \). Both sums are over all calorimeter towers with \( E_T > 100 \text{MeV} \).
[19] In the offline event selection, we use \( ET \) after correcting for jet energy losses in cracks between detector components and for non-linear calorimeter response [18].
The $E_T$ significance is calculated as follows:

$$E_T \text{ significance} = - \log_{10}(\tilde{P}),$$

$$\tilde{P} = \prod \int_{-1}^{y_i} P_i(x)dx, \quad \text{if} \quad y_i < 0,$$

or

$$\tilde{P} = \prod \left(1 - \int_{-1}^{y_i} P_i(x)dx\right), \quad \text{if} \quad y_i > 0,$$

where $P_i(x)$ is the jet energy resolution function obtained from MC and validated with data, $E_T$ is the energy of the $i$-th jet, and $\Delta \phi_i$ is the azimuthal angle between the $i$-th jet and $E_T$.