Study of the dijet invariant mass distribution in $p\bar{p} \to W(\to \ell\nu) + jj$ final states at $\sqrt{s} = 1.96$ TeV

We present a study of the dijet invariant mass spectrum in events with two jets produced in association with a W boson in data corresponding to an integrated luminosity of 4.3 fb$^{-1}$ collected with the D0 detector at $\sqrt{s} = 1.96$ TeV. We find no evidence for anomalous resonant dijet production and derive upper limits on the production cross section of an anomalous dijet resonance recently reported by the CDF Collaboration, investigating the range of dijet invariant mass from 110 to 170 GeV/$c^2$. The probability of the D0 data being consistent with the presence of a dijet resonance with 4 pb production cross section at 145 GeV/$c^2$ is $8 \times 10^{-6}$.

PACS numbers: 12.15.Ji, 12.38.Qk, 13.85.Rm, 14.80.-j

The CDF Collaboration at the Fermilab Tevatron $p\bar{p}$ collider recently reported a study of the dijet invariant mass ($M_{jj}$) spectrum in associated production with $W \rightarrow \ell \nu$ ($\ell = e$ or $\mu$) at $\sqrt{s} = 1.96$ TeV with an integrated luminosity of 4.3 fb$^{-1}$ [1]. In that paper they present evidence for an excess of events corresponding to 3.2 standard deviations (s.d.) above the background expectation, centered at $M_{jj} = 144 \pm 5$ GeV/$c^2$ [1]. The CDF authors model this excess using a Gaussian peak with a width corresponding to an expected experimental $M_{jj}$ resolution for the CDF detector [2] of 14.3 GeV/$c^2$ and further estimate the acceptance and selection efficiencies by simulating associated $W +$ Higgs boson ($H$) production in the decay mode $H \rightarrow b\bar{b}$ and with a mass $M_H = 150$ GeV/$c^2$. Assuming the excess is caused by a particle $X$ with $B(X \rightarrow jj) = 1$, the CDF Collaboration reports an estimated production cross section of $\sigma(p\bar{p} \rightarrow WX) \approx 4$ pb.

Using 5.3 fb$^{-1}$ of integrated luminosity, the D0 Collaboration has previously set limits on resonant $b\bar{b}$ pro-
duction in association with a $W$ boson in dedicated searches for standard model (SM) Higgs bosons in the $WH \to \ell\nu b\bar{b}$ channel [3]. The D0 Collaboration reported upper limits on $\sigma(p\bar{p} \to WH) \times B(H \to b\bar{b})$ ranging from approximately 0.62 pb for $M_H = 100$ GeV/$c^2$ to 0.33 pb for $M_H = 150$ GeV/$c^2$. The CDF Collaboration has performed a similar search using 2.7 fb$^{-1}$ of integrated luminosity and reported no excess of events [4]. Furthermore, the D0 Collaboration has not observed a significant excess of associated $W$ boson and dijet production in analyses of either $WW/WZ \to \ell\nu jj$ [5] or $H \to WW \to \ell\nu jj$ [6] using 1.1 fb$^{-1}$ and 5.4 fb$^{-1}$ of integrated luminosity, respectively.

In this Letter we report a study of associated $W(\to \ell\nu)$ and dijet production using data corresponding to 4.3 fb$^{-1}$ of integrated luminosity collected with the D0 detector [7] at $\sqrt{s} = 1.96$ TeV at the Fermilab Tevatron $p\bar{p}$ Collider. The CDF study of this production process uses the same integrated luminosity. We investigate the dijet invariant mass range from 110 to 170 GeV/$c^2$ for evidence of anomalous dijet production.

To select $W(\to \ell\nu) +jj$ candidate events, we impose similar selection criteria to those used in the CDF analysis: a single reconstructed lepton (electron or muon) with transverse momentum $p_T > 20$ GeV/$c$ and pseudorapidity $|\eta| < 1.0$; missing transverse energy $E_T > 25$ GeV; two jets reconstructed using a jet cone algorithm [9] with a cone of radius $\Delta R = 0.5$ that satisfy $p_T > 30$ GeV/$c$ and $|\eta| < 2.5$, while vetoing events with additional jets with $p_T > 30$ GeV/$c$. The separation between the two jets must be $|\Delta\eta(jet_1,jet_2)| < 2.5$, and the azimuthal separation between the most energetic jet and the direction of the $E_T$ must satisfy $\Delta\phi(jet,E_T) > 0.4$. The transverse momentum of the dijet system is required to be $p_T(jj) > 40$ GeV/$c$. To reduce the background from processes that do not contain $W \to \ell\nu$ decays, we require a transverse mass [10] of $M^{\ell\nu}_T > 30$ GeV/$c^2$. In addition, we restrict $M^{\ell\nu}_T < 200$ GeV/$c^2$ to suppress muon candidates with poorly measured momenta. 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 for the $e\nu jj$ selection of $(98.2\pm1.7)\%$. A suite of triggers in the muon channel achieves a trigger efficiency of $(95.0\pm5.1)\%$ for the $\mu\nu jj$ selection. Lepton candidates must be spatially matched to a track that originates from the $p\bar{p}$ interaction vertex and they must be isolated from other energy depositions in the calorimeter and other tracks in the central tracking detector.

Most background processes are modeled using Monte Carlo (MC) simulation as in the CDF analysis. Diboson contributions ($WW, WZ, ZZ$) are generated with PYTHIA [11] using CTEQ6L1 parton distribution functions (PDF) [12]. The fixed-order matrix element (FOME) generator ALPGEN [13] with CTEQ6L1 PDF is used to generate $W$+jets, $Z$+jets, and $t\bar{t}$ events. The FOME generator COMPHEP [14] is used to produce single top-quark MC samples with CTEQ6M PDF. Both ALPGEN and COMPHEP are interfaced to PYTHIA for subsequent parton showering and hadronization. The MC events undergo a GEANT-based [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. All MC samples except the $W$+jets are normalized to next-to-leading order (NLO) or next-to-NLO (NNLO) predictions for SM cross sections; the $t\bar{t}$, single $t$, and diboson cross sections are taken from Ref. [16], Ref. [17], and the MCFM program [18], respectively. The $Z$+jets sample is normalized to the NNLO cross section [19]. The multijet background, in which a jet misidentified as an isolated lepton passes all selection requirements, is determined from data. In the muon channel, the multijet background is modeled with data events that fail the muon isolation requirements, but pass all other selections. In the electron channel, the multijet background is estimated using a data sample containing events that pass loosened electron quality requirements, but fail the tight electron quality criteria. All multijet samples are corrected for contributions from processes modeled by MC. The multijet normalizations in the two lepton channels are determined from fits to the $M^{\ell\nu}_T$ distributions, in which the multijet and $W$+jets relative normalizations are allowed to float. The expected rate of multijet background is determined by this normalization, with an assigned uncertainty of 20%.

Corrections are applied to the MC to account for differences from data in reconstruction and identification efficiencies of leptons and jets. Also, trigger efficiencies measured in data are applied to MC. The instantaneous luminosity profile and $z$ position of the $p\bar{p}$ interaction vertex of each MC sample are adjusted to match those in data. The $p_T$ distribution of $Z$ bosons is corrected at the generator level to reproduce dedicated measurements [20].

Other D0 analyses of this final state apply additional corrections to improve the modeling of the $W$+jets and $Z$+jets production in the MC [3]. For the results presented in this Letter, we choose not to apply those corrections in order to parallel the CDF analysis. We did, however, study the effects of applying such corrections [21] and find they do not alter our conclusions.

We consider the effect of systematic uncertainties on both the normalization and the shape of dijet invariant mass distributions. Systematic effects are considered from a range of sources: the choice of renormalization and factorization scales, the ALPGEN parton-jet matching algorithm [22], jet energy resolution, jet energy scale, and modeling of the underlying event and parton showering. Uncertainties on the choice of PDF, as well as uncertainties from object reconstruction and identification, are evaluated for all MC samples.
form fits to electron and muon selections simultaneously and then sum them to obtain the dijet invariant mass distributions shown in Fig. 1. The measured yields after the fit are given in Table I.

To probe for an excess similar to that observed by the CDF Collaboration [1], we model a possible signal as a Gaussian resonance in the dijet invariant mass with an observed width corresponding to the expected resolution of the D0 detector given by

$$\text{observed width} = \frac{\text{mean dijet invariant mass}}{\text{expected resolution}}.$$ 

Here, $\lambda_{W}$ and $\lambda_{W'}$ are the width and mass of the $W \rightarrow jj$ resonance, determined to be $\lambda_{W} = 11.7\,\text{GeV/c}^2$ and $\lambda_{W'} = 81\,\text{GeV/c}^2$ from a simulation of $WW \rightarrow \ell\nu jj$ production. For a dijet invariant mass resonance at $M_{jj} = 145\,\text{GeV/c}^2$, the expected width is $\sigma_{jj} = 15.7\,\text{GeV/c}^2$.

We normalize the Gaussian model in the same way as reported in the CDF Letter [1]. We assume that any such excess comes from a particle $X$ that decays to jets with 100% branching fraction. The acceptance for this hypothetical process ($WX \rightarrow \ell\nu jj$) is estimated from a MC simulation of $WH \rightarrow \ell\nu b\bar{b}$ production. When testing the Gaussian signal with a mean of $M_{jj} = 145\,\text{GeV/c}^2$, the acceptance is taken from the $WH \rightarrow \ell\nu b\bar{b}$ simulation with $M_H = 150\,\text{GeV/c}^2$. This prescription is chosen to be consistent with the CDF analysis, which used a simulation of $WH \rightarrow \ell\nu b\bar{b}$ production with $M_H = 150\,\text{GeV/c}^2$ to estimate the acceptance for the excess that they observe at $M_{jj} = 144\,\text{GeV/c}^2$. When probing other values of $M_{jj}$, we use the acceptance obtained for $WH \rightarrow \ell\nu b\bar{b}$ MC events with $M_H = M_{jj} + 5\,\text{GeV/c}^2$.

We use this Gaussian model to derive upper limits on the cross section for a possible dijet resonance as a function of dijet invariant mass using the $CL_s$ method with a negative log-likelihood ratio (LLR) test statistic [24] that is summed over all bins in the dijet invariant mass spectrum. Upper limits on cross section are calculated at the 95% confidence level (C.L.) for Gaussian signals with mean dijet invariant mass in the range $110 < M_{jj} < 170\,\text{GeV/c}^2$, in steps of 5 GeV/c², allowing the cross sections for $W+$jets production to float with no constraint. Other contributions are constrained by the a priori uncertainties on their rate, either derived from

In Fig. 1 we present the dijet invariant mass distribution after a fit of the sum of SM contributions to data. Other distributions are available in the supplementary material [21]. The fit minimizes a Poisson $\chi^2$ function with respect to variations in the rates of individual background sources and systematic uncertainties that may modify the predicted dijet invariant mass distribution [23]. A Gaussian prior is used for each systematic uncertainty, including those on the normalization of each sample, but the cross sections for diboson and $W+$jets production in the MC are floated with no constraint. The fit computes the optimal values of the systematic uncertainties, accounting for departures from the nominal predictions by including a term in the fit function that sums the squared deviation of each systematic in units normalized by its ±1 s.d. Different uncertainties are assumed to be mutually independent, but those common to both lepton channels are treated as fully correlated. We per-

![FIG. 1: (color online) Dijet invariant mass summed over electron and muon channels after the fit without (a) and with (b) subtraction of SM contributions other than that from the SM diboson processes, along with the ±1 s.d. systematic uncertainty on all SM predictions. The $\chi^2$ fit probability, $P(\chi^2)$, is based on the residuals using data and MC statistical uncertainties. Also shown is the relative size and shape for a model with a Gaussian resonance with a production cross section of 4 pb at $M_{jj} = 145\,\text{GeV/c}^2$.](image-url)
theory or subsidiary measurements.

The Gaussian model is assigned systematic uncertainties affecting both the normalization and shape of the distribution derived from the systematic uncertainties on the diboson simulation. A fit [23] of both the signal+background and background-only hypotheses is performed for an ensemble of pseudo-experiments as well as for the data distribution. The results of the cross section upper limit calculation are shown in Fig. 2 and are summarized in Table II.

In a further effort to evaluate the sensitivity for any excess of events of the type reported by the CDF collaboration, we perform a signal-injection test. We repeat the statistical analysis after injecting a Gaussian signal model, normalized to a cross section of 4 pb, into the D0 data sample, thereby creating a mock “data” sample modeling the expected outcome with a signal present. The size and shape of the injected Gaussian model for $M_{jj} = 145 \text{ GeV}/c^2$ relative to other data components is shown in Fig. 1.

The LLR metric provides a sensitive measure of model compatibility, providing information on both the rate and mass of any signal-like excess. We therefore study the LLR distributions obtained with actual data as well as the signal-injected mock data sample. The results of the LLR test in Fig. 3 show a striking difference between the two hypotheses, demonstrating that this analysis is sensitive to the purported excess. In the actual data, however, no significant evidence for an excess is observed.

In Fig. 4, we show as a function of cross section the $p$-value obtained by integrating the LLR distribution populated from pseudo-experiments drawn from the signal+background hypothesis above the observed LLR, assuming a Gaussian invariant mass distribution with a mean of $M_{jj} = 145 \text{ GeV}/c^2$. The $p$-value for a Gaussian signal with cross section of 4 pb is $8.0 \times 10^{-6}$, corresponding to a rejection of this signal cross section at a Gaussian equivalent of 4.3 s.d. We set a 95% C.L. upper limit of 1.9 pb on the production cross section of such a resonance.

In summary, we have used $4.3 \text{ fb}^{-1}$ of integrated luminosity collected with the D0 detector to study the dijet invariant mass spectrum in events containing one $W \rightarrow \ell \nu$ ($\ell = e$ or $\mu$) boson decay and two high-$p_T$ jets. Utilizing a similar data selection as the CDF collaboration we find no evidence for anomalous, resonant
TABLE II: Expected and observed upper limits on the cross section (in pb) at the 95% C.L. for a dijet invariant mass resonance.

<table>
<thead>
<tr>
<th>$M_{jj}$ (GeV)</th>
<th>110</th>
<th>115</th>
<th>120</th>
<th>125</th>
<th>130</th>
<th>135</th>
<th>140</th>
<th>145</th>
<th>150</th>
<th>155</th>
<th>160</th>
<th>165</th>
<th>170</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Expected:</strong></td>
<td>2.20</td>
<td>2.01</td>
<td>1.90</td>
<td>1.78</td>
<td>1.71</td>
<td>1.64</td>
<td>1.58</td>
<td>1.52</td>
<td>1.47</td>
<td>1.40</td>
<td>1.37</td>
<td>1.31</td>
<td>1.24</td>
</tr>
<tr>
<td><strong>Observed:</strong></td>
<td>2.57</td>
<td>2.44</td>
<td>2.35</td>
<td>2.27</td>
<td>2.19</td>
<td>2.09</td>
<td>2.00</td>
<td>1.85</td>
<td>1.69</td>
<td>1.58</td>
<td>1.46</td>
<td>1.36</td>
<td>1.28</td>
</tr>
</tbody>
</table>

production of dijets in the mass range $110-170 \text{ GeV}/c^2$. Using a simulation of $WH \rightarrow t\bar{t}bb$ production to model acceptance and efficiency, we derive upper limits on the cross section for anomalous resonant dijet production. For $M_{jj} = 145 \text{ GeV}/c^2$, we set a 95% C.L. upper limit of 1.9 pb on the cross section and we reject the hypothesis of a production cross section of 4 pb at the level of 4.3 s.d.

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); CNPq, FAPERJ, FAPESP and FUNDUNESP (Brazil); DAE and DST (India); Colciencias (Colombia); CONACyT (Mexico); KRF and KOSEF (Korea); CONACET and UBACYT (Argentina); FOM (The Netherlands); MSMT and GACR (Czech Republic); CRC Program and NSERC (Canada); BMBF and DFG (Germany); SFI (Ireland); 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 iterative mid-point cone algorithm with radius $\Delta R = \sqrt{(\Delta \phi)^2 + (\Delta y)^2} = 0.5$ was used.
APPENDIX

Fit of a 145 GeV/c² Dijet Resonance

The D0 data do not indicate the presence of a non-SM dijet resonance such as indicated by the CDF Collaboration. In the Letter we showed the fit of the SM predictions to the data. However, fitting only the SM contributions could hide an excess if the systematic uncertainties allowed the SM contributions to be distorted in such a way that they filled in the excess. To study this question, we present a fit to the data of the SM predictions plus the Gaussian signal template with $M_{jj} = 145 \text{ GeV/c}^2$. The resulting dijet mass distribution is shown in Fig. 5. The fit is performed in the same way as described in the Letter (no constraint on diboson or $W$+jets normalizations), except that it now also includes the Gaussian signal template for $M_{jj} = 145 \text{ GeV/c}^2$ with a freely floating normalization. The Gaussian model includes systematic uncertainties affecting both the normalization and shape of the template, analogous to the systematic uncertainties for the diboson prediction. The best fit value for the Gaussian template yields a cross section of \( \sigma(p\bar{p} \rightarrow WX) = 0.82^{+0.83}_{-0.82} \text{ pb} \), consistent with no excess. When we fix the diboson cross section to the SM prediction with a Gaussian prior of 7\% on the rate, the best fit \( \chi^2 = 0.464 \). The Gaussian model includes systematic uncertainties, greatly reducing the need for the kinematic corrections of the simulation. Thus, these analyses use parameterized functions to correct the pseudorapidities of the two highest \( p_T \) jets and the \( \Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2} \) separation between those two jets in $W$+jets and $Z$+jets samples, and the transverse momentum of the $W$ boson candidate, \( p_T(W) \), in the $W$+jets samples, to better model the data.

The tight kinematic selection criteria employed in this analysis (e.g., $p_T(jj) > 40 \text{ GeV/c}^2$) remove much of the phase space in which the MC generators have difficulty modeling data (e.g., low $p_T(W)$), greatly reducing the need for the kinematic corrections of the simulation. Therefore, the plots and results in the Letter do not use any of these kinematic corrections, which is consistent with the CDF analysis.

Although kinematic corrections are not required to achieve adequate modeling when applying the tight selection criteria of this analysis, modeling issues are probably still present. In this section we present the results obtained when the kinematic corrections (derived from a selection similar to the search for $WH$ production [3]) are applied to this analysis.

The following figures are analogous to those in the Letter, except that the above mentioned kinematic corrections have been applied to the simulation. Figure 6 shows the dijet invariant mass distribution after the fit of the sum of SM predictions to data. The change relative to Fig. 1 in the Letter is not large, but improved modeling is evident in the higher \( \chi^2 \) probability. The resulting upper limits on the cross section for production of a dijet invariant mass resonance are presented in Table III and shown in Fig. 7. They are consistent with those in Fig. 2 from the Letter. Figure 8 shows the LLR distributions analogous to Fig. 3 from the Letter.

From this study we conclude that the kinematic cor-
FIG. 6: (color online) Dijet invariant mass summed over lepton channels after the fit (a) without and (b) with SM contributions subtraction other than from the SM diboson processes, along with the ±1 s.d. systematic uncertainty on all SM predictions. These distributions have the additional kinematic corrections applied to the MC.

TABLE III: Expected and observed upper limits on the cross section (in pb) at the 95% C.L. for a dijet invariant mass resonance. These limits are derived with the additional kinematic corrections applied to the MC.

<table>
<thead>
<tr>
<th>$M_{jj}$ (GeV)</th>
<th>110</th>
<th>115</th>
<th>120</th>
<th>125</th>
<th>130</th>
<th>135</th>
<th>140</th>
<th>145</th>
<th>150</th>
<th>155</th>
<th>160</th>
<th>165</th>
<th>170</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected</td>
<td>2.35</td>
<td>2.16</td>
<td>2.05</td>
<td>1.97</td>
<td>1.88</td>
<td>1.81</td>
<td>1.73</td>
<td>1.68</td>
<td>1.65</td>
<td>1.56</td>
<td>1.52</td>
<td>1.45</td>
<td>1.42</td>
</tr>
<tr>
<td>Observed</td>
<td>2.26</td>
<td>2.02</td>
<td>1.93</td>
<td>1.83</td>
<td>1.74</td>
<td>1.64</td>
<td>1.55</td>
<td>1.48</td>
<td>1.37</td>
<td>1.27</td>
<td>1.18</td>
<td>1.09</td>
<td>1.06</td>
</tr>
</tbody>
</table>

Kinematic distributions presented in this Section are modeled without the additional corrections applied to the MC. Figure 9 shows the dijet invariant mass distribution for the separate lepton channels after the simultaneous fit in these two distributions of the SM predictions to data. Figure 1(a) in the Letter is the combination of these two plots. Figure 10 shows comparisons between data and simulation for other kinematic variables after the fit.

Additional Data-MC Comparisons

Rections to the simulation do not substantially change the result and we reach the same conclusion that there is no excess of dijet events in the D0 data similar to that reported by the CDF collaboration.
FIG. 7: (color online) Upper limits on the cross section (in pb) at the 95% C.L. for a Gaussian signal in dijet invariant mass. These results are derived with the additional kinematic corrections applied to the MC. Shown in the figure are the limit expected using the background prediction, the observed data limit, and the regions corresponding to a 1 s.d. and 2 s.d. fluctuation of the backgrounds.

FIG. 8: (color online) Log-likelihood ratio test statistic as a function of probed dijet invariant mass. These results are derived with the additional kinematic corrections applied to the MC. Shown are the expected LLR for the background prediction (dashed black) with regions corresponding to a 1 s.d. and 2 s.d. fluctuation of the backgrounds, for the signal+background prediction (dashed red), for the data (solid black), and for data with a dijet mass resonance at 145 GeV/c$^2$ injected using a cross section of 4 pb (solid red).

FIG. 9: (color online) Dijet invariant mass distributions separately for the (a) electron and (b) muon channel after the simultaneous fit of these two distributions.
FIG. 10: (color online) Distributions of kinematic variables (combined electron and muon channels) evaluated using the results of a $\chi^2$ fit of SM predictions to data for the dijet invariant mass distribution: (a) transverse W mass, (b) $\Delta R$ separation between the two selected jets, (c) lepton $p_T$, (d) missing transverse energy, (e) highest jet $p_T$, (f) second highest jet $p_T$. The $\pm 1$ s.d. systematic uncertainty on all SM predictions is presented by the cross-hatched area.