



Measurement of the WZ Production Cross Section in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV using 5.9 fb^{-1} of CDF Run II Data

The CDF Collaboration
URL <http://www-cdf.fnal.gov>
(Dated: June 6, 2010)

We report a new measurement of the WZ production cross section in the three charged lepton (e, μ) and one neutrino final state in $p\bar{p}$ collisions at a center of mass energy of 1.96 TeV. The data were collected with the CDF II detector at the Tevatron collider at Fermilab and correspond to an integrated luminosity of approximately 5.9 fb^{-1} . A NeuroBayes neural network is used to distinguish the WZ signal from backgrounds in the final selection region. The WZ cross section is then extracted using a binned maximum likelihood method which best fits the neural network output score's signal and background shapes to the data. The measured WZ cross section is $3.7 \pm 0.6(\text{stat})_{-0.4}^{+0.6}(\text{syst})(\text{pb})$.

I. INTRODUCTION

The direct production of WZ diboson pairs in proton-antiproton collisions is an important background in dilepton searches for a high mass Standard Model Higgs boson decaying to WW , and the primary background for the related $H \rightarrow WW$ trilepton search. Thus, understanding and verifying the modeling of the WZ process is essential to the $H \rightarrow WW$ search. Both Tevatron experiments have measured the WZ cross section with a three-lepton signature in the past [3] [5]. This note presents the most precise measurement to date of the WZ production cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV. We use approximately 5.9 fb^{-1} of integrated luminosity collected by the CDF II detector at the Fermilab Tevatron.

The WZ production cross section was first measured by CDF in Run II using 1.1 fb^{-1} of data. This was performed by making a fit to the missing transverse energy distribution to yield an excess above the background prediction equivalent to six standard deviations and yielding a measured cross section of $5.0_{-1.6}^{+1.8} \text{ pb}$ [3]. Later, the measurement was updated to 1.9 fb^{-1} of data yielding a measured cross section of $4.3_{-1.1}^{+1.4} \text{ pb}$.

In this note, we present a measurement of $p\bar{p} \rightarrow WZ \rightarrow l\nu l\nu l\nu$, where the lepton l indicates a reconstructed electron or muon, or also a tau if it then decays to an electron or muon. In 5.9 fb^{-1} of data, we find 53 events. From modeling, we expected about 40 WZ events and 8 events from Standard Model background processes. To measure the WZ cross section, we first use a NeuroBayes neural network to create a distribution separating background-like and signal-like events, then the measured cross section is extracted using a maximum likelihood method which best fits the neural network score distribution for the normalization of WZ events.

II. EVENT SELECTION

This analysis is based on an integrated luminosity of 5.9 fb^{-1} . The data are collected with inclusive high- p_T lepton (electron or muon) triggers. We use the same data sample and lepton selection as the search for a high mass Standard Model Higgs boson decaying to two W -bosons [4], updated to 5.9 fb^{-1} , which involves three categories of electrons, eight categories of muons, and one category that identifies a lepton track but cannot tell the flavor.

All leptons are required to be isolated such that the sum of the E_T for the calorimeter towers in a cone of $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.4$ around the lepton is less than 10% of the electron E_T or muon p_T . Although, one of the electron categories has replaced the set of hard cuts for identification with a likelihood function that uses the identification criteria as terms, so a small number of events may slightly exceed this isolation cut if the other electron identification requirements are strong. The transverse energy E_T of a shower or calorimeter tower is $E \sin \theta$, where E is the associated energy. Similarly, p_T is the component of track momentum transverse to the beam line.

Electron candidates are required to have a ratio of hadronic energy to electromagnetic energy consistent with originating from an electromagnetic shower and are further divided into central and forward categories. The central electron category requires a well-measured track satisfying $p_T > 10 \text{ GeV}/c$ that is fiducial to the central shower maximum detector (SMX) and matched to a central EM energy cluster. The candidate is also required to have a matching cluster in the SMX, minimal energy sharing between towers, and a ratio for shower energy E to track momentum p of less than $2.5 + 0.0015E_T$. A forward electron is required to be fiducial to the forward SMX detector and have an energy deposition in both the calorimeter towers and SMX detector consistent with an electron shower shape. One of the calorimeter seeded tracks is required to be matched with a silicon track to reduce the photon background.

Muons are identified by either a charged track matched to a reconstructed track segment (“stub”) in muon chambers or as a stubless minimum ionizing particle fiducial to the calorimeters. In addition, stubless muons are required to have at least 0.1 GeV in total calorimeter energy. For $|\eta_{\text{det}}| < 1.2$, strict requirements on the number of tracking chamber hits and the χ^2 of the track fit are placed on the muon tracks in order to suppress kaon decay-in-flight backgrounds. The category of stubless muons with $|\eta_{\text{det}}| > 1.2$ requires that at least 60% of the tracking chamber layers crossed by the track have hits. In order to suppress background from cosmic rays, the track’s point of closest approach to the beamline must be consistent with originating from the beam line.

The final category of leptons are constructed from tracks which are not fiducial to the SMX detectors nor identified as stubbed muons. The requirements for the tracks are the same as stubless muons with $|\eta_{\text{det}}| < 1.2$, but without any of the calorimeter requirements. Due to the lack of calorimeter information, electrons and muons cannot be reliably differentiated in this region, and are therefore treated as having either flavor in the candidate selection.

To identify the presence of neutrinos, we use the missing transverse energy $\cancel{E}_T = |\sum_i E_{T,i} \hat{n}_{T,i}|$, where the $\hat{n}_{T,i}$ is the transverse component of the unit vector pointing from the interaction point to calorimeter tower i . The \cancel{E}_T is corrected for muons which do not deposit all of their energy in the calorimeter.

Candidate events are required to pass one of five online trigger selections implemented in three successively more stringent levels. The final central electron requirement is an EM energy cluster with $E_T > 18 \text{ GeV}$ matched to a

CDF Run II	$\int \mathcal{L} = 5.9 \text{ fb}^{-1}$				
	<i>WZ</i> signal	Low \cancel{E}_T	Control	No- <i>Z</i>	Control
<i>ZZ</i>	4.97 ± 0.66	$6.97 \pm$	0.92	$1.70 \pm$	0.23
<i>Z</i> +Jets	2.41 ± 0.59	$41.4 \pm$	10.1	$14.2 \pm$	3.49
<i>Z</i> γ	0.77 ± 0.27	$71.4 \pm$	25.0	$80.3 \pm$	28.2
$t\bar{t}$	0.15 ± 0.04	$0.02 \pm$	0.005	$0.32 \pm$	0.10
Total Background	8.29 ± 0.97	$119.7 \pm$	27.2	$96.6 \pm$	28.5
<i>WZ</i>	40.2 ± 4.06	$6.25 \pm$	0.63	$3.52 \pm$	0.36
Sig.+Back.	48.5 ± 4.20	$126.0 \pm$	27.4	$100.1 \pm$	28.6
Data	53	118		104	

High Mass

TABLE I: Event count for the predicted signal and background contributions in the analysis signal region, as well as the two control regions defined for this analysis.

track with $p_T > 8 \text{ GeV}/c$. Muon triggers are based on information from muon chambers matched to a track with $p_T > 18 \text{ GeV}/c$. The trigger for forward electrons requires an $E_T > 20 \text{ GeV}$ EM energy cluster and an online measurement of the $\cancel{E}_T > 15 \text{ GeV}$ [6].

A. Signal Region

The trilepton candidates are selected from events with exactly three leptons, not all having the same sign. At least one of the leptons is required to satisfy the trigger criteria to be included in the datasets used and have $E_T > 20 \text{ GeV}$ ($p_T > 20 \text{ GeV}/c$) for electrons (muons). The other two leptons are required to have $> 10 \text{ GeV}$ (GeV/c). Further, all events in the signal region are required to have a missing transverse energy value of at least 25 GeV and must pass a *Z*-boson selection. With three leptons there are three possible pairings of leptons. Events that have at least one lepton pairing where the two leptons have opposite charge, same flavor, and a dilepton invariant mass of $m_{ll} \in [86.0, 106.0] \text{ GeV}/c^2$ pass the *Z*-boson selection.

The Standard Model background processes are $ZZ \rightarrow llll$ diboson where one lepton fails to be identified, $Z\gamma$ (i.e. Drell-Yan that emits a photon which undergoes conversion to an electron pair, mimicking an electron signature), *Z*+jets where a jet is misidentified as a lepton, and a very small amount of $t\bar{t}$ where a *b*-jet provides the third lepton signature. The \cancel{E}_T cut at 25.0 GeV drastically reduces the amount of $Z\gamma$ and *Z*+jets backgrounds since these processes do not tend to produce high p_T neutrinos.

The cross section of the *WZ* events from in monte carlo simulation is

$$\sigma_{WZ}^{NLO} = 3.46 \pm 0.21$$

B. Control Regions

Two control regions are defined for this analysis: one in the low \cancel{E}_T region and one with a *Z*-boson rejection. The signal region defined in section II A and both control regions are all mutually exclusive sets. Beginning with the signal region definition, the low \cancel{E}_T control region replaces the $\cancel{E}_T > 25.0 \text{ GeV}$ cut with a $\cancel{E}_T < 20.0 \text{ GeV}$ cut. As such, this region is heavy in $Z\gamma$ and *Z*+jets. The data count for this region is 118 events—in good agreement with the 126 ± 27 predicted by simulation. Similarly, the *Z*-removed control region reverses the *Z*-selection criteria of the signal region, except with a wider window around the *Z* mass at 91 GeV . In this case, all lepton pairings with opposite sign and same flavor must reject the *Z* mass with $m_{ll} \notin [66.0, 116.0] \text{ GeV}/c^2$. This region is largely dominated by $Z\gamma$ because if the photon is radiated from one of the leptons that decayed directly from the *Z*, then it is the three-body invariant mass—not the dilepton invariant mass—that sums back to the original *Z* mass. As such, the dilepton invariant mass tends to fall below the *Z*-selection mass window defined above. The 104 data events in this region are in excellent agreement with the 100 ± 29 predicted by simulation.

III. DATA MODELING

The geometric and kinematic acceptance for the *WZ*, *ZZ*, $Z\gamma$, and $t\bar{t}$ processes are determined using a Monte Carlo calculation of the collision followed by a GEANT3-based simulation of the CDF II detector [2] response. For

WZ , ZZ , DY , and $t\bar{t}$ the generator used is PYTHIA [8], and for $Z\gamma$ it is the generator described in [1]. We use the CTEQ5L parton distribution functions (PDFs) to model the momentum distribution of the initial-state partons [7].

A correction of up to 10% per lepton is applied to the simulation based on measurements of the lepton reconstruction and identification efficiencies in data using Z decays. Additional 5% and 10% corrections based on $Z \rightarrow \ell\ell$ cross section measurements are applied to stubless muons reconstructed from only central tracks and muons reconstructed from minimum ionizing energy deposits in the forward calorimeter, respectively, to account for the known poor modeling of these lepton types. Trigger efficiencies are determined from $W \rightarrow e\nu$ data for electrons and from $Z \rightarrow \mu^+\mu^-$ data for muons.

The background from Z +jets is estimated from a sample of events with an identified lepton and a jet that is required to pass loose isolation requirements and contain a track or energy cluster similar to those required in the lepton identification. The contribution of each event to the total yield is scaled by the probability that the jet is identified as a lepton. This probability is determined from multijet events collected with a set of jet-based triggers. A correction is applied for the small real lepton contribution using single W and Z boson Monte Carlo simulation.

IV. NEURAL NETWORK AND LIKELIHOOD RATIO FIT

The trilepton WZ analysis relies on the NeuroBayes neural network package to discriminate signal from background; we do not attempt the Matrix Element method in this study. Twelve input variables are used in this measurement (see figure 2). The neural net results can be seen in figure 1.

Because the interaction topology under consideration involves three leptons and also because we do not separate the analyses by jet bin, the signatures of the signal region under consideration involve many variables whose discriminatory power must be explored. As such, a large quantity of discriminating variables are used to train the NeuroBayes neural nets.

The discriminating variables are, in order of significance:

1. $\Delta\phi$ (W -Lep., \cancel{E}_T)
2. m_T (jets), 0-jet events just assigned a value of zero.
3. Lepton type combinations: discriminate by whether an event is eee (three electrons), $ee\mu$, $e\mu\mu$, etc. This variable is particularly good at discriminating $Z\gamma$ because of the photon conversion to electrons. For instance, eee is strong in $Z\gamma$, but $ee\mu$ is not because the same-flavor pair ee would be from the Z which means the μ would have to be from the conversion. However, photon conversion tends to be to electrons, not to muons.
4. \cancel{E}_T
5. m_T (W -Lep., \cancel{E}_T)
6. $\Delta\phi$ (2^{nd} Lep. by p_T ., \cancel{E}_T)
7. H_T
8. m_T (all Lep., \cancel{E}_T , Jets)
9. m_T (3^{rd} Lep. by p_T , \cancel{E}_T)
10. $\Delta\phi$ (vector sum of the three leptons, \cancel{E}_T)
11. m_T (three leptons)
12. NJet

A. Maximum Likelihood Method

A binned maximum likelihood method is used to extract the WZ cross section using the shape of the neural network score distributions (figure 1) from signal and background along with their estimated normalizations and systematic uncertainties. The best fit to these distributions, or the maximum likelihood, gives the best measure of the WZ cross section.

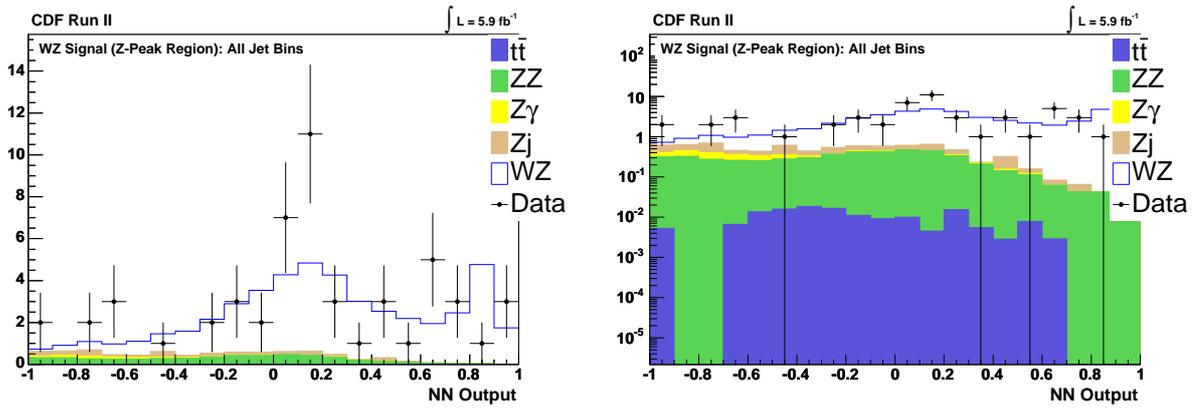


FIG. 1: Trilepton WZ NeuroBayes Neural Network output. Top two: signal region. Middle two: low \cancel{E}_T control region. Bottom two: Z -Peak removed control region.

The likelihood function is formed from a product of Poisson probabilities for each bin in the neural net score. Additionally, Gaussian constraints are applied corresponding to each systematic S_c (shown in Table II). The likelihood is given by

$$\mathcal{L} = \left(\prod_i \frac{\mu_i^{n_i} e^{-\mu_i}}{n_i!} \right) \cdot \prod_c e^{-\frac{S_c^2}{2}} \quad (1)$$

where μ_i is the total expectation in the i -th bin and n_i is the number of data events in the i -th bin. μ_i is given by

$$\mu_i = \sum_k \alpha_k \left[\prod_c (1 + f_k^c S_c) \right] (N_k^{Exp})_i \quad (2)$$

Here f_k^c is the fractional uncertainty associated with the systematic S_c and process k . This is constructed such that the systematics are properly correlated or uncorrelated between the different contributions. $(N_k^{Exp})_i$ is the expected number of events from process k in the i -th bin. α_k is the parameter which is used to measure the WZ cross section. It is a freely floating parameter for α_{WZ} and fixed to one for all other processes. The measured value of α_{WZ} multiplied by the input WZ cross section gives the measured value of the WZ cross section:

$$\sigma_{WZ}^{measured} = \alpha_{WZ} \cdot \sigma_{WZ}^{NLO}$$

V. SYSTEMATIC UNCERTAINTIES

Systematic uncertainties associated with the Monte Carlo simulation affect the WZ , ZZ , $Z\gamma$, and $t\bar{t}$ acceptances taken from the simulated event samples. Uncertainties originating from lepton selection and trigger efficiency measurements are propagated through the acceptance calculation, giving uncertainties typically around 2% for the different processes.

Although $t\bar{t}$ is a small contribution to the background for the WZ cross section in the trilepton case, we do have to account for the peculiar situation that our third lepton is faked from a b -jet and the rate at which a b -jet fakes a lepton—as opposed to a light jet—is not well-known. Further, as a background with two real leptons and one faked, we cannot ignore the possible coverage of $t\bar{t}$ in the data-based Fakes (Z +Jets) category. We know that the fake rates used in the Fakes category is based on jet samples populated mostly with light jets and presume that b -jets in particular are more likely than light jets to produce a signature that could fake a lepton. Hence, whatever $t\bar{t}$ contribution that exists in the Fakes category is scaled down by the light jet dominated fake rate, meaning it is scaled down too far. To make up for the difference we use an MC $t\bar{t}$ sample that allows reconstructed leptons to match to generator-level leptons, photons, or b -jets (typically, for these reconstructed MC leptons to be considered fully "found" they must pass a matching criterion to a generator-level lepton or photon only). Now, of course, we have the problem of possible

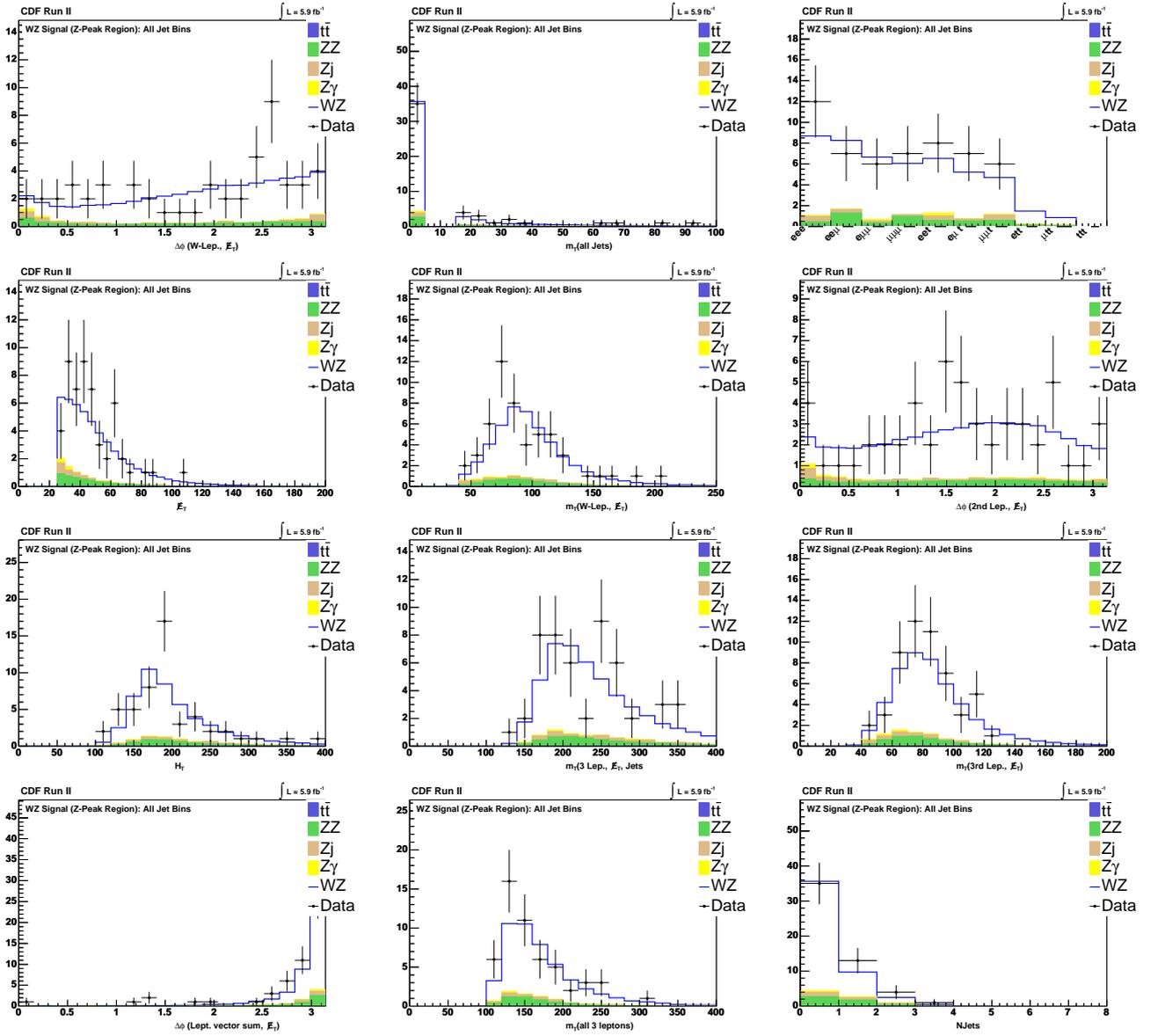


FIG. 2: The twelve neural network discriminating variable inputs.

Uncertainty Source	<i>WZ</i>	<i>ZZ</i>	<i>tt</i>	<i>Zγ</i>	<i>Z+jet</i>
Cross Section		6.0%	10.0%	5.0%	
PDF Model	2.7%	2.7%	2.1%	2.7%	
Higher-order Diagrams	10.0%	10.0%	10.0%	11.0%	
Conversion Modeling				20.0%	
Jet Fake Rates					24.5%
<i>b</i> -jet Fake Rate			0.23%		
Jet Energy Scaling				1.2%	
MC Run Dependence			5.0%		
Lepton ID Efficiencies	2.0%	2.0%	2.0%	2.0%	
Trigger Efficiencies	2.1%	2.1%	2.0%	2.1%	
Luminosity	5.9%	5.9%	5.9%	5.9%	

TABLE II: Summary of all systematics in this analysis. Systematics in italics are taken to be correlated across processes.

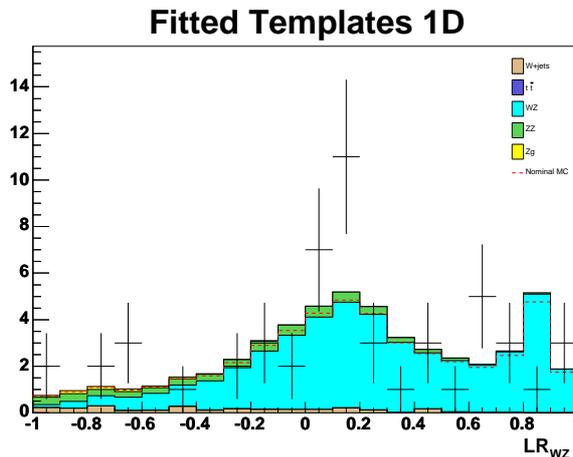


FIG. 3: Fitted signal and background templates where the backgrounds are Gaussianly constrained within their estimated uncertainties and the WZ normalization is allowed to float freely. The data is shown as the solid points. Additionally the sum of nominal predictions (before fitting) for all processes is shown as the red line superimposed.

double-counting of $t\bar{t}$ between the MC and what implicit $t\bar{t}$ contribution populates the Fakes category. To account for the double-counting possibility, we assign a systematic error defined to be one half the percentage difference between the MC $t\bar{t}$ sample that allows leptons to match to generator-level leptons, photons, and b -jets; and the MC $t\bar{t}$ sample that allows such matching to generator-level leptons and photons only.

For the $Z\gamma$ background contribution, there is an additional uncertainty of 20% from the detector material description and conversion veto efficiency. The systematic uncertainty on the Z +jets background contribution is determined from differences in the measured probability that a jet is identified as a lepton for jets collected using different jet E_T trigger thresholds. These variations correspond to changing the parton composition of the jets and the relative amount of contamination from real leptons.

VI. RESULTS

The binned maximum likelihood fit to data gives a measured value for the WZ cross section of

$$\sigma(p\bar{p} \rightarrow WZ) = 3.7 \pm 0.6(\text{stat})_{-0.4}^{+0.6}(\text{syst})(\text{pb})$$

where the uncertainty includes statistical, systematic, and luminosity uncertainties. The fitted templates are shown in Figure 3.

This is in good agreement with the theoretical expectation of $\sigma(p\bar{p} \rightarrow WZ) = 3.46 \pm 0.21$ pb, and is the world's best measurement to date of $\sigma(p\bar{p} \rightarrow WZ)$.

Acknowledgments

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 fuer Bildung und Forschung, Germany; the Korean Science and Engineering Foundation and the Korean Research Foundation; the Particle Physics and Astronomy Research Council and the Royal Society, UK; the Russian Foundation for Basic Research; the Comision Interministerial de Ciencia y Tecnologia, Spain; and in part by the European Community's

Human Potential Programme under contract HPRN-CT-20002, Probe for New Physics.

- [1] U. Baur and E. L. Berger. Probing the weak boson sector in z gamma production at hadron colliders. *Phys. Rev.*, D47:4889–4904, 1993.
- [2] R. Brun, R. Hagelberg, M. Hansroul, and J.C. Lassalle.
- [3] CDF Collaboration. Observation of wz production. *Phys. Rev. Lett.*, 98:161801, 2007.
- [4] CDF Collaboration. Search for $h\text{-}\cancel{L}ww$ higgs production at cdf using 5.3 fb^{-1} of data, 2010. See http://www-cdf.fnal.gov/physics/new/hdg//Results_files/results/hwwmenn_100305/.
- [5] D0 Collaboration. Measurement of wz diboson production cross section in trilepton final states at $\sqrt{s} = 1.96 \text{ tev}$, 2006. See <http://www-d0.fnal.gov/Run2Physics/WWW/results/prelim/EW/E15/E15.pdf>.
- [6] K. Hagiwara, S. Ishihara, R. Szalapski, and D. Zeppenfeld. Low energy effects of new interactions in the electroweak boson sector. *Phys. Rev. D*, 48(5):2182–2203, Sep 1993.
- [7] H. L. Lai *et al.* Global QCD analysis of parton structure of the nucleon: Cteq5 parton distributions. *Eur. Phys. J.*, C12:375–392, 2000.
- [8] Torbjorn Sjostrand, Stephen Mrenna, and Peter Skands. Pythia 6.4 physics and manual. *JHEP*, 05:026, 2006.