Observation of the Production of a W Boson in Association with a Single Charm Quark at the Tevatron

Lucio Cerrito, John Paul Chou, Melissa Franklin, Tony Liss, Shulamit Moed Sher

October 3, 2011

We present a measurement of the production cross section of a W boson in association with a single charm quark in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV with the CDF detector. The analysis uses an integrated luminosity of 4.3 fb^{-1} and is based on the reconstruction of the final state with one high transverse-momentum electron or muon, missing transverse energy, and one hadronic jet. The signal is evinced by a charge asymmetry between the lepton from the W boson decay and a soft lepton from the semileptonic decay of the charm quark. We measure a production cross section times branching fraction of $13.3^{+3.3}_{-2.9}\,(\mathrm{stat}+\mathrm{syst})$ pb given a charm hadron with transverse momentum greater than 20 GeV and pseudorapidity within ± 1.5. This is consistent with the standard model expectation. Assuming a null hypothesis without the presence of the signal process, the number of events observed constitutes a $6.4\,\sigma$ excess.

Contents

1 Introduction

The production of a W boson in association with a single charm quark (W+c) proceeds at lowest order through sg and $\bar{s}g$ fusion. As can be seen Figs. 1 and 2, the electric charge of the lepton from the semileptonic decay of the charm quark and the electric charge of the lepton from the W boson decay are opposite in sign. We exploit this signature to distingush Wc events from the large background of other W+jet events, including W+c$\bar{\text{c}}$ and W+b$\bar{\text{b}}$. At the Tevatron, the W+c signal is approximately 5% of the inclusive W+1 jet cross section for jets with a transverse momentum greater than 10 GeV [1]. The production of W+c where the initial s quark is replaced by a d quark is suppressed by the CKM quark mixing matrix element V_{cd}. Given the larger d-quark partonic luminosity, this processes constitutes approximately 10% of the total production rate.

W+c production is a background process for several important physics signatures such as single-top production and associated W+Higgs production. It is also a significant component in the t$\bar{\text{t}}$ control region. Our motivation to perform this analysis is driven by the need to identify W+c events among W+heavy flavor events and to improve the understanding of backgrounds of the aforementioned processes. A CDF measurement of the W+c production cross section with soft muon tagging using ~ 1.8 fb^{-1} of data is reported in Ref. [2]. This note reports the measurements of the W+c production cross section using both soft electron and soft muon tagging, first separately, and then combined.

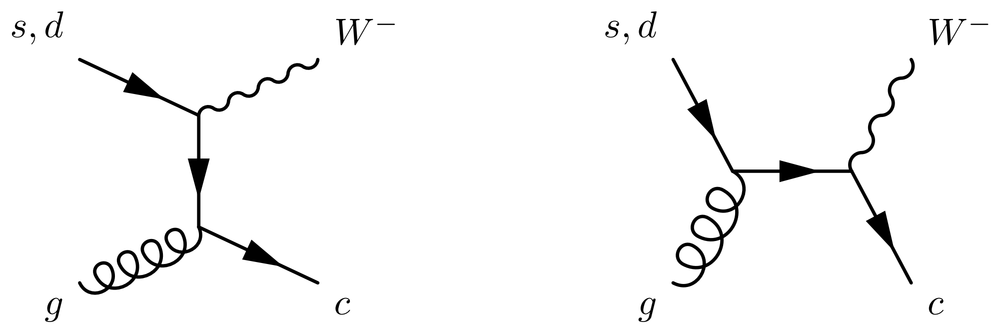

Figure 1: The leading order Feynman diagrams for W+charm production

2 Event Selection and Datasets

We apply the following criteria to our W+1$\tilde{\text{j}}$et event selection:

- one isolated > 20 GeV CEM/CMUP/CMX lepton (the lepton flavor must be consistent with the trigger path);
- one > 15 GeV jet corrected to level 5 with $|\eta_D| < 2.0$, reconstructed with the R=0.4 JETCLU algorithm;
- $\not{E}_{\text{T}} > 25$ GeV;

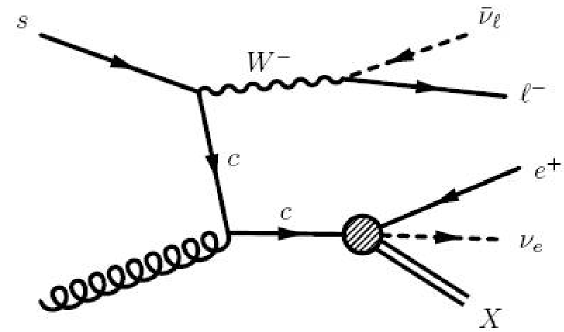

Figure 2: Feynman diagram of W plus single charm production with soft electron. When a soft muon is produced instead, the electron in the diagram final state is replaced by a muon.

- veto cosmic ray muons, conversion electrons and Z bosons.

We additionally require that

- transverse mass of the isolated lepton and the $\not\!\!E_{\mathrm{T}}$, M_{T}, is greater than 20 GeV;
- reject events when the invariant mass between the soft muon and tight muon is between 8–11 GeV or 70–110 GeV;
- reject events when the invariant mass between the soft electron and tight electron is greater than 45 GeV;
- reject events when the tight and soft leptons are both electrons and the difference in ϕ between the jet and $\not\!\!E_{\mathrm{T}}$ is less than 0.3.

These additional requirements are intended to further suppress multijet, Z+jet, and Drell–Yan events which are more prominent in this analysis than a typical top-quark analysis.

We tag the events by requiring the presence of a soft electron tag ($\mathrm{SLT_e}$) or a soft muon tag (SLT_μ) near the jet. The "nearness" requirement is that $\Delta R \leq 0.6$ for the SLT_μ and $\Delta R \leq 0.4$ for the $\mathrm{SLT_e}$, where ΔR is the distance between the track and the jet-axis in η–ϕ space.

We simulate the W+c signal with a combination of ALPGEN for the event generation and PYTHIA for the showering (stopwX). In these samples, the charm quark is required to have $p_{\mathrm{T}} > 8$ GeV and $|\eta| < 3.0$ at the generator level. In order to provide a more robust cross-section measurement claim, we further restrict the charm quark to have $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 1.5$. This also has the effect of minimizing the contribution from certain systematic uncertainties, such as the parton distribution functions.

After requiring that the charm has $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 1.5$, the inclusive cross section for the W($\rightarrow \ell\nu$)+c process is reduced from 21.1 pb to 7.5 pb.

Using MCFM, we expect that the inclusive NLO W+c cross section times W$\rightarrow \ell\nu$ branching fraction is 11.3 ± 2.2 pb for a charm quark with $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 1.5$.

3 Backgrounds

3.1 Monte Carlo Backgrounds

We estimate the backgrounds due to production of dibosons (WW, WZ, ZZ), single top, and $t\bar{t}$ from MC simulation. The contributions due to these processes are small, and the production cross sections are well-established. WW production contributes the most to the background among these processes and has a strong charge asymmetry

3.2 Drell–Yan and Z+jets events

Z+jets and Drell–Yan are distinguished in this analysis by the generator-level invariant-mass range of the dilepton pair. Events where $76 < M_{\ell\ell} < 106$ GeV are considered Z+jets, whereas events with dileptons pairs outside of this mass window are considered Drell–Yan.

3.3 Multijet QCD

Events due to multijet production can enter the event selection through hadronic mis-identification or heavy-flavor decay. Missing energy is the result of mis-measured jet energy, detector effects, as well as the occasional hard neutrino.

sample	$\mathrm{SLT_e}$	SLT_μ
pretagged	$(3.3 \pm 0.5) \times 10^4$	$(3.9 \pm 0.6) \times 10^4$
OS-only tagged	201 ± 6	150 ± 8
SS-only tagged	173 ± 6	158 ± 9
OS−SS difference	27 ± 12	-8 ± 17

Table 1: The multijet QCD fractions for SLT_e and SLT_μ in the different samples, tagged and pretag events, and their corresponding uncertainties.

We estimate this background by releasing the $\not\!\!E_T$ requirement on the events entirely and fitting the $\not\!\!E_T$ spectrum using a binned, negative log-likelihood minimization, floating only the templates for the W+jets and multijet backgrounds. Meanwhile, the MC background templates are fixed. The contribution of the W+c signal is so small and the $\not\!\!E_T$ distribution for W+c events is so similar to the W+jet events that we can confidently ignore the signal in the fit, allowing the W+jet background to cover the signal contribution.

Templates of the $\not\!\!E_T$ variable are constructed out of simulation for all of the backgrounds except for the multijet background. The multijet template is constructed in data from an "anti-electron" sample, in which electron-triggered events are required to have an electron-like object which fails at least two selection criteria.

We measure a charge asymmetry for the multijet background of -0.03 ± 0.06 (SLT_μ) and 0.07 ± 0.03 ($\mathrm{SLT_e}$). Figs. 3 and 4 show the fit results.

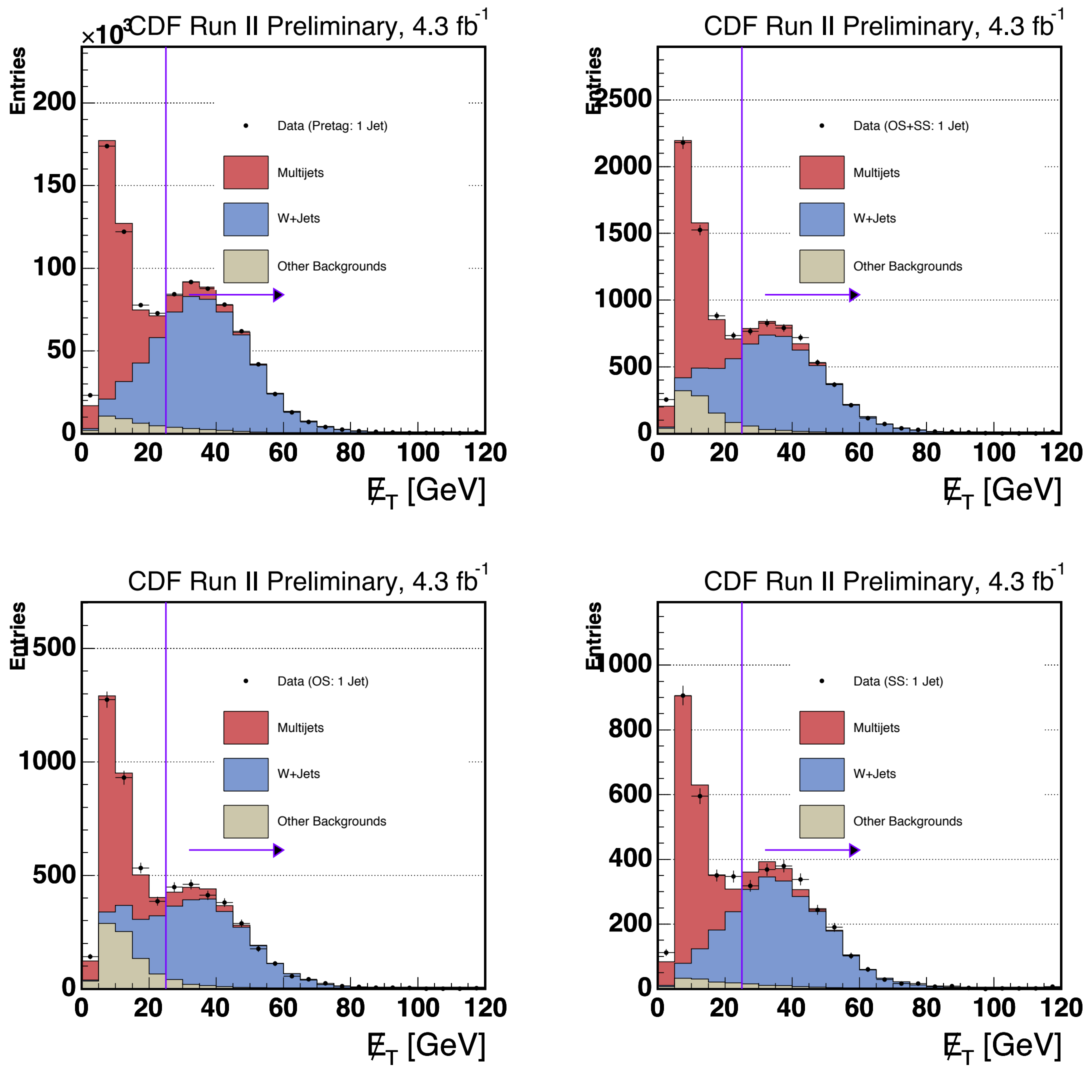

Figure 3: $\not\!\!E_T$ distribution in the anti-electron sample for SLT$_e$ events, upper left plot - for pretag events, upper right plot - OS+SS events, lower left - OS events, lower right - SS events. The fit for this template, shown above, is used to determine the QCD fraction in our selection.

3.4 W + Jets

The dominant background to W+c production is the production of a W boson associated with jets. Rather than rely on a theoretical prediction of the production cross section, we estimate the contribution by normalizing the pretag yield to the data. We rely on a combination of MC and data-driven techniques to estimate the contribution to the tagged sample.

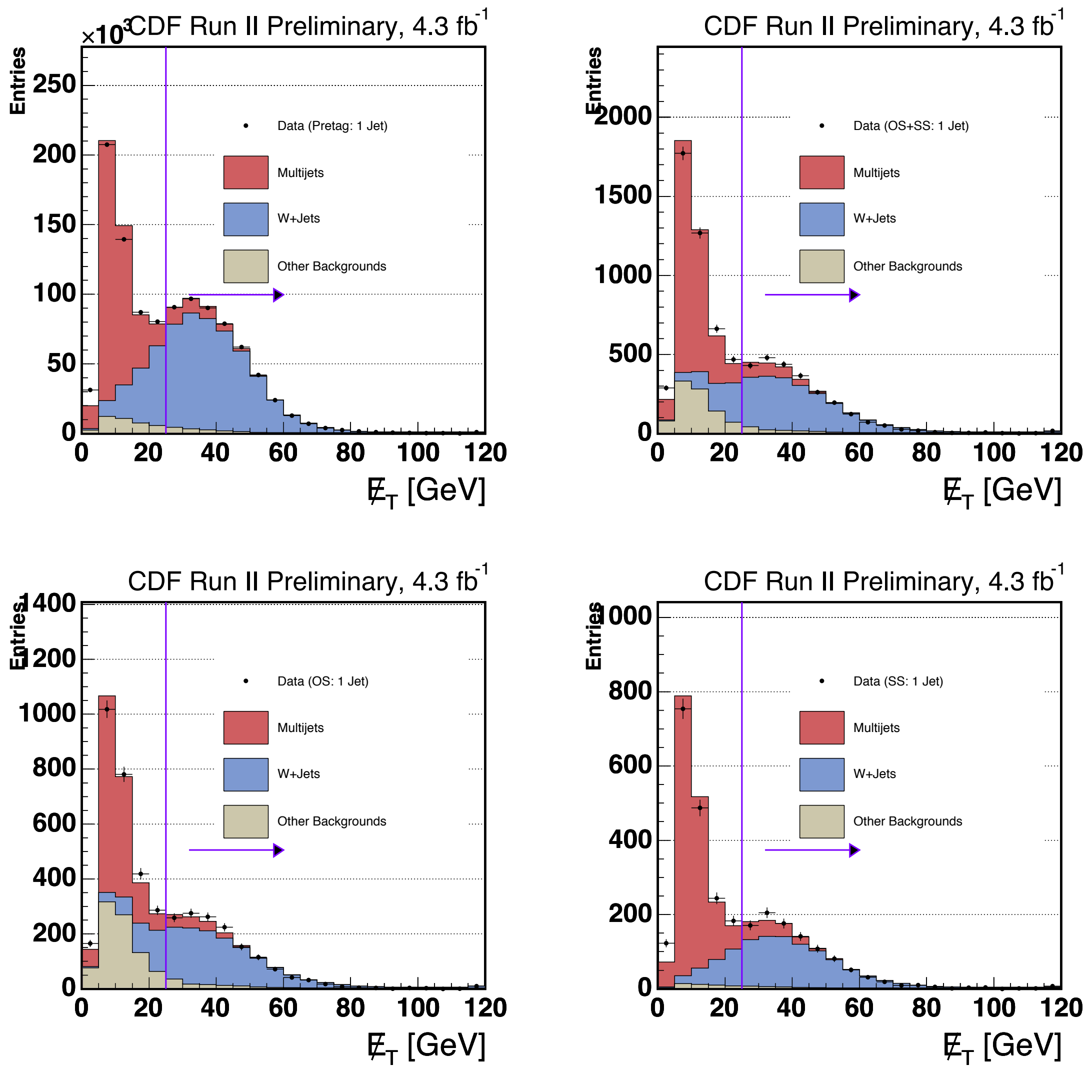

Figure 4: $\not\!\!E_{\rm T}$ distribution in the anti-electron sample for SLT$_\mu$ events, upper left plot - for pretag events, upper right plot - OS+SS events, lower left - OS events, lower right - SS events. The fit for this template, shown above, is used to determine the QCD fraction in our selection.

3.4.1 Pretag Estimate

The estimate of the number of pretag W+jet events is determined by subtracting off the pretag estimates of all other backgrounds and the signal from the total pretag event yield, that is:

$$N_{\rm pre}^{\rm W+jets} = N_{\rm pre}^{\rm data} - N_{\rm pre}^{\rm W+c} - N_{\rm pre}^{\rm MC} - N_{\rm pre}^{\rm Z+jets} - N_{\rm pre}^{\rm QCD}, \tag{1}$$

where the superscripts refer to the sample and the subscripts refer to the fact that these are the pretag expectations. In this case, W+jets does not include the W+c contribution typically associated with the process, but includes W+$b\bar{b}$, W+$c\bar{c}$, and W+light flavor. The superscript MC refers to all of the MC backgrounds (including WW production and single-top production), and Z+jets also includes Drell–Yan production.

process	# of events (SLT_μ)	# of events (SLT_e)	systematic source
diboson, single top, $t\bar{t}$	1449 ± 93	1421 ± 91	luminosity
Z+Jets, Drell–Yan	14744 ± 2990	13296 ± 2699	cross section
QCD multijet	39197 ± 5880	32827 ± 4925	QCD fit
W+c	12806 ± 794	12301 ± 763	luminosity
W+$b\bar{b}$, W+$c\bar{c}$, W+l.f.	448242 ± 6815	440773 ± 5847	QCD fit
data	516437	500618	—

Table 2: Summary of pretag contributions from each process, including the dominant source of systematic uncertainties. Also shown is the number of pretag events observed in the data. The W+c contribution in this table assumes the cross section predicted by ALPGEN, not the final measured result. The W+$b\bar{b}$, W+$c\bar{c}$, and W+light flavor pretag prediction is determined by normalizing to the data.

3.4.2 Tag Estimate

We estimate the number of SLT tags via the equation,

$$N_{\rm tag}^{\rm W+jets} = N_{\rm pre}^{\rm W+jets} \left(\sum_i \epsilon_i^{\rm HF} F_i^{\rm HF} + \epsilon^{\rm LF} \left(1 - \sum_i F_i^{\rm HF} \right) \right), \quad (2)$$

where the ϵ's refer to the SLT efficiency for a given sample, and the $F^{\rm HF}$'s refer to the heavy flavor fraction for a given sample. The index i runs over the different heavy flavor configurations, which indicate whether the jet is matched to a parton-level hadron of type b or c. The heavy flavor fractions are determined from MC simulation. The heavy-flavor fractions are also corrected by a K factor of 1.4 ± 0.4, although this has a negligible effect on the final result. The fractions are $F_{\rm b}^{\rm HF} = (0.94 \pm 0.27)\%$ and $F_{\rm c}^{\rm HF} = (2.5 \pm 0.7)\%$, so the light flavor fraction is $1 - F_{\rm b}^{\rm HF} - F_{\rm c}^{\rm HF} = (96.6 \pm 0.7)\%$.

Table 3 summarizes the total number of tags expected and measured in the data. We find good agreement in the number of predicted tags in both SLT channels. Since the primary component of the tagged sample is due to W+jet processes, this is a validation of the overall normalization due to this process. In the following section, we describe how we estimate the asymmetry due to W+jets.

Table 4 summarizes the asymmetry (OS-SS/OS+SS) measurements for each background.

4 Cross Section Measurement and Systematic Uncertainties

We interpret the excess of OS−SS events observed in the data over the background expectation as the production of W+c. Table 4 summarizes the observed number of events and expected number

process	# of events (SLT_μ)	# of events (SLT_e)	systematic source
diboson, single top, $t\bar{t}$	26 ± 3	35 ± 3	luminosity
Z+Jets, Drell–Yan	132 ± 30	138 ± 29	cross section
QCD multijet	308 ± 17	374 ± 12	QCD fit
W+c	214 ± 19	174 ± 16	luminosity
W+$b\bar{b}$, W+$c\bar{c}$, W+l.f.	1808 ± 271	4076 ± 305	fake rate
total expected	2488 ± 274	4797 ± 307	fake rate
data	2506	4582	—

Table 3: Summary of tag contributions from each process, including the dominant source of systematic uncertainties. Also shown is the number of tagged events observed in the data. The W+c contribution in this table assumes the cross section predicted by ALPGEN, not the final measured result. The overall agreement between predicted and observed tags is a consistency check of our method and demonstrates the reliability of our W+jet estimate in a control region.

process	SLT_μ		SLT_e	
diboson, single top, $t\bar{t}$	9 ± 1	0.33 ± 0.01	20 ± 2	0.58 ± 0.01
Z+Jets, Drell–Yan	84 ± 18	0.63 ± 0.02	36 ± 7	0.26 ± 0.01
QCD multijet	-8 ± 17	-0.03 ± 0.07	27 ± 12	0.07 ± 0.03
W+c	161 ± 13	0.75 ± 0.03	78 ± 7	0.45 ± 0.02
W+$b\bar{b}$, W+$c\bar{c}$, W+l.f.	86 ± 14	0.05 ± 0.01	174 ± 19	0.04 ± 0.01
total expected	331 ± 37	0.13 ± 0.02	336 ± 28	0.07 ± 0.01
data	458 ± 50	0.18 ± 0.02	406 ± 68	0.09 ± 0.01

Table 4: Summary of charge asymmety contributions from each process. The first column for each tagger consists of the expected number of OS−SS tags for that processes, and the second column consists of the expected asymmetry. Also shown is the observed result in the data, where the given errors reflect statistical uncertainties assuming an underlying Poisson process. The W+c contribution in this table assumes the cross section predicted by ALPGEN, not the final measured result. The difference between the observed and expected is interpreted as a underestimate of the W+c production cross section by the simulation.

of events for each process. We use Eq. **??** to measure the W+c production cross section. The statistical uncertainty on the cross section is proportional to the square root of the total number of tagged events—opposite-sign plus same-sign. We estimate systematic uncertainties due to the SLT taggers, the luminosity, the assumed theoretical cross sections, the W lepton reconstruction and trigger efficiencies, and the QCD multijet fit for each background individually and propagate them to the cross section uncertainty algebraically. Additional uncertainties due to the jet energy scale, initial- and final-state radiation, factorization and renormalization scales, hadronization modeling, and parton luminosity distributions are treated separately with a procedure to be discussed shortly.

We measure a production cross section for W+c where the charm parton has $p_T > 20$ GeV and

$|\eta| < 1.5$ of

$$\sigma = 13.4 \pm 2.3\,(\text{stat}) \pm 2.4\,(\text{syst}) \pm 1.1\,(\text{lumi})\,\text{pb}, \tag{3}$$

and

$$\sigma = 14.2 \pm 6.5\,(\text{stat}) \pm 3.4\,(\text{syst}) \pm 1.2\,(\text{lumi})\,\text{pb}, \tag{4}$$

for the SLT_μ and SLT_e measurements, respectively.

As a cross check, we measure the W+c cross section separately in the e+jet and μ+jet final states. This also acts as a cross check of the Z+jet background estimate, because this background is suppressed in the e+SLT_μ and μ+SLT_e channels. We find that

$$\sigma_{\text{e},\mu} = 14.6 \pm 3.0\,(\text{stat}) \pm 2.7\,(\text{syst}) \pm 0.9\,(\text{lumi})\,\text{pb} \tag{5}$$
$$\sigma_{\mu,\mu} = 11.6 \pm 3.7\,(\text{stat}) \pm 3.3\,(\text{syst}) \pm 1.2\,(\text{lumi})\,\text{pb} \tag{6}$$
$$\sigma_{\text{e},\text{e}} = 12.1 \pm 9.6\,(\text{stat}) \pm 4.0\,(\text{syst}) \pm 1.2\,(\text{lumi})\,\text{pb} \tag{7}$$
$$\sigma_{\mu,\text{e}} = 16.2 \pm 8.8\,(\text{stat}) \pm 4.1\,(\text{syst}) \pm 1.2\,(\text{lumi})\,\text{pb}, \tag{8}$$

where the first index refers to the lepton flavor from the W boson decay, and the second index refers to the lepton flavor of the SLT. As the statistical uncertainty in each of these measurements is independent, we note that they all give a consistent measurement of the W+c cross section. The same-flavor cross-section measurements are consistently lower than the opposite-flavor measurements. Although this is consistent with a slight over-estimate of the Z+jet and Drell–Yan backgrounds, there is little to suggest a faulty estimate.

We estimate the effect of the jet-energy scale on the acceptance calculation by varying the jet-energy correction $\pm 1\sigma$ about its uncertainties. We find that acceptance varies 2.0% as a result, which we apply as a systematic uncertainty on our cross section measurement. To measure the effect of initial-state and final-state radiation (ISR/FSR), we measure the W+jet acceptance in different samples with the ISR/FSR increased and decreased coherently. This has a 6% overall effect on the acceptance. Similarly, we vary the factorization and renormalization scales Q^2 up and down by a factor of 2.0 and 0.5, respectively, in the W+c simulation. We find a 1.3% overall effect on the acceptance.

We quote a systematic uncertainty of 4.6% on the simulation of the jet hadronization. This is done by swapping out the PYTHIA shower modeling for HERWIG in a sample similar to the W+c samples (see Ref [2] for details). PDF uncertainties are computed by remeasuring the acceptance with different PDFsets. We add all systematic uncertainties together in quadrature. The relative contribution from each is quoted in Table 5. The dominant systematics in this measurement are due to uncertainties surrounding the SLTs, the factorization and renormalization scales, the luminosity, and the uncertainty on the QCD multijet fit.

5 Combination

We combine the results from the two SLT taggers by performing a profile likelihood minimization. Systematic uncertainties are assumed to be either 100% correlated if they are shared between the two taggers or 0% correlated if not. We assume Wilks' theorem in establishing a 68% confidence interval by determining for which values of the cross section the negative log likelihood increases by half a unit. The combined measured value of the cross section times leptonic branching fraction is $13.3^{+3.3}_{-2.9}\,(\text{stat} + \text{syst})$ pb for a charm hadron $p_\text{T} > 20$ GeV and $|\eta| < 1.5$. Figure 5 shows twice the

source	uncertainty (%)	
	SLT_μ	SLT_e
SLT uncertainties	9.2	16.6
Factorization/Renormalization scales	1.3	
Luminosity	7.9	8.3
QCD multijet fit	6.3	9.9
ISR/FSR	6.0	
background cross sections	5.7	4.7
PDFs	3.6	
W-lepton ID	2.2	
Jet-energy scale	2.0	
Total	16.7	22.9

Table 5: Source of systematic uncertainties for the measurement ordered by relative size. The total systematic uncertainty is taken as the quadrature sum of the individual sources.

negative log likelihood ($-2\log\lambda(\sigma_{\mathrm{Wc}})$) as a function of the cross section. We note that the central value of the combined cross section is lower than either the SLT_e or SLT_μ measurements. This is due to the asymmetric error reporting. If we were to symmetrize the uncertainties, the central value of 13.5 pb would then lie between the two measurements.

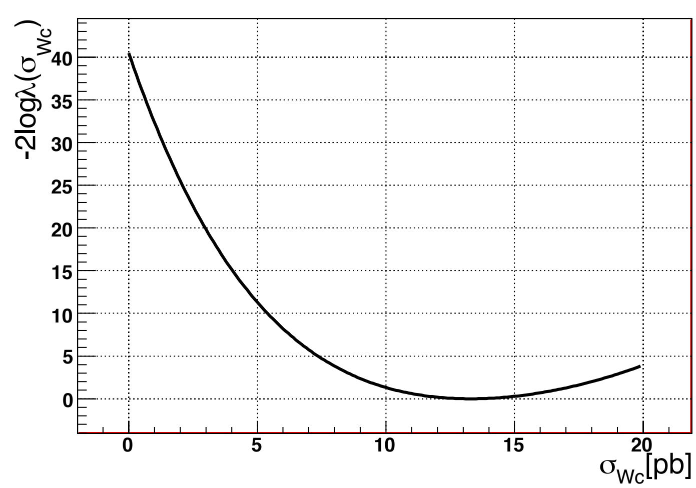

Figure 5: Twice the negative log likelihood as a function of the cross section. The minimum corresponds to the central value.

In the case of the null hypothesis where the signal cross section times branching fraction vanishes, the value of $\sqrt{-2\log\lambda(\sigma_{\mathrm{Wc}} = 0\ \mathrm{pb})}$ is $6.4\,\sigma$. We interpret this to be the first observation of W+c production at the Tevatron.

6 Conclusions

We have performed a measurement of the production cross section of a W boson in association with a single charm quark in $\mathrm{p\bar{p}}$ collisions at $\sqrt{s} = 1.96$ TeV with the CDF detector. The analysis uses an integrated luminosity of 4.3 fb^{-1} and is based on the reconstruction of the final state with one high transverse-momentum electron or muon, missing transverse energy, and one hadronic jet. We measure a production cross section times branching fraction of $13.3^{+3.3}_{-2.9}\,(\mathrm{stat} + \mathrm{syst})$ pb given a charm hadron with transverse momentum greater than 20 GeV and $|\eta| < 1.5$. This is consistent with the theoretical NLO production cross section times branching fraction of 11.3 ± 2.2 pb, but it is in tension with a LO production cross section times branching fraction of 7.5 ± 1.5 pb. Assuming a null hypothesis without the presence of the signal process, the number of events observed constitutes a $6.3\,\sigma$ excess.

References

[1] hep-ph/9308370; U. Baur, F. Halzen, S. Keller, M.L. Mangano, K. Riesselmann

[2] CDF note 8026, A. Abulencia, L. Cerrito, U. Grundler, T. Liss, X. Zhang, "First measurement of the Wc production cross-section with SLT flavor tagging".

[3] CDF note 9185, J. Adelman *et al*, "Method II For You".

[4] L. Cerrito, A. Taffard, CDF note 6305 "A Soft Muon Tagger for Run II", CDF note 7122 "A Soft Muon Tagger for Run II: Summer-04 version"

[5] John Paul Chou, CDF note 9603 "ttbar Production Cross Section Measurement using Soft Electron Tagging in ppbar Collisions", SLT_e CDF notes 8787-8790

[6] CDF note 9161, M. Franklin *et al*, "Updated Search for the Flavor Changing Neutral Current Decay t -¿ Zq with 1.9/fb of CDF-II Data"

[7] John Paul Chou Phd thesis, http://huhepl.harvard.edu/theses/chou.pdf

[8] PRL 105:101801, 2010 ; Phys.Rev.D 81:092002, 2010

[9] CDF note 6894. "Track Reconstruction Efficiency in Jets".

[10] A. D. Martin,R. G. Roberts, W. J. Stirling, and R. S. Thorne, Eur. Phys. J. C 4, 463 (1998)