Measurement of the $W + b$-jet and $W + c$-jet differential production cross sections in pp collisions at $\sqrt{s} = 1.96$ TeV

V.M. Abazov, B. Abbott, B.S. Acharya, M. Adams, J.P. Agnew, G.D. Alexeev, G. Alkhazov, A. Alton, A. Askew, S. Atkins, K. August, C. Avila, F. Badaud, L. Bagly, B. Baldin, D.V. Bandurin, S. Banerjee, E. Barberis, P. Baringer, J.F. Bartlett, U. Bassler, V. Bazterra, A. Bean, M. Begalli, L. Bellantoni, S.B. Beri, G. Bernardi, R. Bernhard, I. Bertram, M. Besançon, R. Beuselinck, P.C. Bhat, S. Bhatia, V. Bhatnagar, G. Blayze, S. Blessing, K. Bloom, A. Boehmlein, D. Boline, E.E. Boos, G. Borrisov, M. Borysoy, O. Borosy, A. Brandt, O. Brandt, R. Brock, A. Bross, D. Brown, X.B. Bu, M. Buehler, V. Buescher, V. Bunichev, S. Burdlin, C.P. Buszello, E. Camacho-Pérez, B.C.K. Casey, H. Castilla-Valdez, S. Caughron, S. Chakraborti, K.M. Chan, A. Chandra, E. Chapon, G. Chen, S.W. Cho, S. Choi, B. Choudhary, S. Chiang, D. Claes, J. Clutter, M. Cooke, W.E. Cooper, M. Corcoran, F. Cordier, M.-C. Cousinou, D. Cutts, D. Das, G. Davies, S.J. de Luca, E. Del Bene, D. De La Cruz-Burelo, F. Délion, R. Demina, D. Denisov, S.P. Denisov, S. Desai, C. Deterre, K. DeVauhn, H.T. Diehl, M. Diesburg, P.F. Ding, A. Dominguez, A. Dubey, L.V. Dudko, A. Duperrin, S. Dutt, M. Eads, D. Edmunds, J. Ellis, V.D. Elvira, Y. Enari, H. Evans, V.N. Evodikov, A. Fauré, L. Feng, T. Ferbel, F. Fiedler, F. Fillthorpe, W. Fisher, H.E. Fisk, M. Fortner, H. Fox, S. Fuess, P.H. Garbincius, A. García-Bellido, J.A. García-González, V. Gavrilov, W. Geng, C.E. Gerber, Y. Gershtein, G. Ginter, O. Gogota, G. Golovanov, P.D. Grannis, S. Greder, H. Greenlee, G. Grenier, Ph. Gris, J.-F. Grivaz, A. Grohsjean, S. Grünewald, M.W. Grünewald, T. Guillemin, G. Gutierrez, P. Gutierrez, J. Haley, L. Han, K. Harder, A. Harel, J.M. Hauptman, J. Hayas, T. Head, T. Hebbeker, D. Hedin, H. Hegab, A.P. Heinson, U. Heintz, C. Hensel, I. Hereida-De La Cruz, K. Herner, G. Hesketh, M.D. Hildreth, R. Hirosky, J.D. Hobbs, B. Hoeuwenisen, J. Hogan, M. Hohlfeld, J.L. Holzbauer, I. Howley, Z. Hubacek, V. Hynek, I. Iashvili, Y. Ilchenko, R. Illingworth, A.S. Ito, S. Jabeen, M. Jaffré, A. Jayasinghe, M.S. Jeong, R. Jesik, P. Jiang, K. Johns, E.Johnson, M. Johnson, A. Jonckheere, P. Jonsson, J. Joshi, A.W. Jung, A. Juste, E. Kajfasz, D. Karmanov, I. Katsanos, M. Kaur, R. Keheoe, S. Kerniche, N. Khalatyan, A. Khan, A. Kharchilava, Y.N. Kharaezmee, I. Kiselevich, J.M. Kohli, A.V. Kozelov, J. Kraus, A. Kumar, A. Kupco, T. Kurča, V.A. Kuzmin, S. Lammers, P. Lebrun, H.S. Lee, S.W. Lee, W.M. Lee, X. Lei, J. Lellouch, D. Li, L. Li, Q.Z. Li, J.K. Lim, D. Lincon, J. Linnemann, V.V. Lippert, R. Lipton, H. Liu, Y. Li, A. Lobkod, M. Lokeajcik, R. Lopes de Sa, R. Luna-Garcia, A.L. Lyon, A.K.A. Maciel, R. Madar, R. Magoña-Villalba, S. Malik, V.L. Malyshchev, M. Mansour, J. Martínez-Ortega, R. McCracken, C.L. Mcgivern, M.M. Meijer, A. Melnitchouk, D. Menezes, P.G. Mercadante, M. Merkin, A. Meyer, J. Meyer, F. Miconi, N.K. Mondal, M. Mulhearn, E. Nagy, M. Narain, R. Nayyar, H.A. Neal, J.P. Negret, P. Neustroev, H.T. Nguyen, T. Nunnenmann, J. Orduna, N. Osman, J. Osta, A. Pal, N. Parasch, V. Parihar, S.K. Park, R. Partridge, N. Parua, A. Patwa, B. Penning, M. Perilov, Y. Peters, K. Petridis, G. Petrillo, P. Pétroz, M.-A. Pleier, V.M. Podstavkov, A.V. Popov, M. Prewitt, D. Price, N. Prokopenko, J.Q. Qian, A. Quad, B. Quinn, P.N. Ratoff, I. Razumov, I. Rupp-Baudot, F. Rizatdinova, M. Rominsky, A. Ross, C. Royon, P. Rubinov, R. Ruchti, G. Sajot, A. Sánchez-Hernández, M.P. Sanders, A.S. Santos, G. Savage, M. Savitsky, L. Sawyer, T. Scanlon, R.D. Schamberger, Y. Scheglov, H. Schellman, C. Schwannberger, R. Schwienhorst, J. Sekaric, E. Severini, E. Shabalina, V. Shary, S. Shaw, A.A. Shchukin, V. Simak, P. Skubic, P. Slattery, D. Smirnov, G.R. Snow, J. Snow, S. Snyder, S. Söderlund-Rembold, L. Sonnenschein, K. Soustruník, J. Stark, D.A. Stoyanov, M. Strauss, L. Suter, P. Svoisky, M. Titov, V.V. Tokmenin, Y.-T. Tsai, D. Tsibychev, B. Tuchming, C. Tully, L. Uvarov, S. Uvarov, S. Uzunyan, R. Van Kooten, W.M. van Leeuwen, N. Varelas, E.W. Varnes, I.A. Vasilyev, A.Y. Verkheev, L.S. Vertogradov, M. Verzocchi, M. Vesterinen, D. Vilanova, P. Volck, H.D. Wahl, M.H.L.S. Wang, J. Warchol, G. Watts, M. Wayne, J. Weichert, L. Welty-Rieger,
M.R.J. Williams\textsuperscript{1,49}, G.W. Wilson\textsuperscript{53}, M. Wobisch\textsuperscript{54}, D.R. Wood\textsuperscript{55}, T.R. Wyatt\textsuperscript{41}, Y. Xie\textsuperscript{45}, R. Yamada\textsuperscript{45}, S. Yang\textsuperscript{4}, T. Yasuda\textsuperscript{45}, Y.A. Yatsunenko\textsuperscript{31}, W. Ye\textsuperscript{64}, Z. Ye\textsuperscript{45}, H. Yin\textsuperscript{45}, K. Yip\textsuperscript{65}, S.W. Youn\textsuperscript{45}, J.M. Yu\textsuperscript{56}, J. Zennamo\textsuperscript{62}, T.G. Zhao\textsuperscript{41}, B. Zhou\textsuperscript{56}, J. Zhu\textsuperscript{56}, M. Zielinski\textsuperscript{63}, D. Zieminska\textsuperscript{49}, and L. Zivkovic\textsuperscript{14}

(The D0 Collaboration\textsuperscript{*})

\textsuperscript{1}LAFFEX, Centro Brasileiro de Pesquisas Físicas, Rio de Janeiro, Brazil
\textsuperscript{2}Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil
\textsuperscript{3}Universidade Federal do ABC, Santo André, Brazil
\textsuperscript{4}University of Science and Technology of China, Hefei, People’s Republic of China
\textsuperscript{5}Universidad de los Andes, Bogotá, Colombia
\textsuperscript{6}Charles University, Faculty of Mathematics and Physics, Center for Particle Physics, Prague, Czech Republic
\textsuperscript{7}Czech Technical University in Prague, Prague, Czech Republic
\textsuperscript{8}Institute of Physics, Academy of Sciences of the Czech Republic, Prague, Czech Republic
\textsuperscript{9}Universidad San Francisco de Quito, Quito, Ecuador
\textsuperscript{10}LPSC, Université Joseph Fourier Grenoble 1, CNRS/IN2P3, Institut National Polytechnique de Grenoble, Grenoble, France
\textsuperscript{11}Universidad de los Andes, Bogotá, Colombia
\textsuperscript{12}CPPM, Aix-Marseille Université, CNRS/IN2P3, Marseille, France
\textsuperscript{13}LAL, Université Paris-Sud, CNRS/IN2P3, Orsay, France
\textsuperscript{14}LPNHE, Universités Paris VI and VII, CNRS/IN2P3, Paris, France
\textsuperscript{15}CEA, IRFU, SPP, Saclay, France
\textsuperscript{16}IPHC, Université de Strasbourg, CNRS/IN2P3, Strasbourg, France
\textsuperscript{17}IPNL, Université Lyon 1, CNRS/IN2P3, Villeurbanne, France and Université de Lyon, Lyon, France
\textsuperscript{18}IPNHE, Universités Paris VI and VII, CNRS/IN2P3, Paris, France
\textsuperscript{19}Physikalisches Institut A, RWTH Aachen University, Aachen, Germany
\textsuperscript{20}Physikalisches Institut, Universität Freiburg, Freiburg, Germany
\textsuperscript{21}Ludwig-Maximilians-Universität München, München, Germany
\textsuperscript{22}Panjab University, Chandigarh, India
\textsuperscript{23}Delhi University, Delhi, India
\textsuperscript{24}Korea Detector Laboratory, Korea University, Seoul, Korea
\textsuperscript{25}Tata Institute of Fundamental Research, Mumbai, India
\textsuperscript{26}University College Dublin, Dublin, Ireland
\textsuperscript{27}CINVESTAV, Mexico City, Mexico
\textsuperscript{28}Science Park, Amsterdam, the Netherlands
\textsuperscript{29}University of Nijmegen, Nijmegen, the Netherlands
\textsuperscript{30}Joint Institute for Nuclear Research, Dubna, Russia
\textsuperscript{31}Institut für Physik, Universität Mainz, Mainz, Germany
\textsuperscript{32}Ludwig-Maximilians-Universität München, München, Germany
\textsuperscript{33}II. Physikalisches Institut, Georg-August-Universität Göttingen, Göttingen, Germany
\textsuperscript{34}Physikalisches Institut, Universität Freiburg, Freiburg, Germany
\textsuperscript{35}Moscow State University, Moscow, Russia
\textsuperscript{36}Institut für Hoch Energien Physik, Protvino, Russia
\textsuperscript{37}Institució Catalana de Recerca i Estudis Avançats (ICREA) and Institut de Física d’Altes Energies (IFAE), Barcelona, Spain
\textsuperscript{38}University of Illinois at Chicago, Chicago, Illinois 60607, USA
\textsuperscript{39}Northern Illinois University, DeKalb, Illinois 60115, USA
\textsuperscript{40}University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA
\textsuperscript{41}University of Iowa, Iowa City, Iowa 52242, USA
\textsuperscript{42}University of Kansas, Lawrence, Kansas 66045, USA
\textsuperscript{43}Louisiana Tech University, Ruston, Louisiana 71272, USA
\textsuperscript{44}Northeastern University, Boston, Massachusetts 02115, USA
\textsuperscript{45}University of Michigan, Ann Arbor, Michigan 48109, USA

*Corresponding author.
We present a measurement of the cross sections for the associated production of a W boson with at least one heavy quark jet, b or c, in proton-antiproton collisions. Data corresponding to an integrated luminosity of 8.7 fb$^{-1}$ recorded with the D0 detector at the Fermilab Tevatron $\sqrt{s} = 1.96$ TeV are used to measure the cross sections differentially as a function of the jet transverse momenta in the range 20 to 150 GeV. These results are compared to calculations of perturbative QCD theory as well as predictions from Monte Carlo generators.

PACS numbers: 12.38.Qk, 13.85.Qk, 14.65.Fy, 14.70.Fm

Measurement of the production cross section of a W boson in association with a b or c-quark jet provides a stringent test of quantum chromodynamics (QCD). At hadron colliders, the associated production of a heavy quark with a W boson can also be a significant background to rare standard model (SM) processes, for example, production of top quark pairs [1], a single top quark with a W boson decaying to two b quarks [2], as well as for new physics processes, e.g., supersymmetric scalar top quark production [4].

The dominant processes contributing to $W + c$-jet production are $qg \rightarrow Wc$ and $qq' \rightarrow Wg$ followed by $g \rightarrow c\bar{c}$. The production cross section for the first process is sensitive to the quark and gluon parton density functions (PDFs). Since the c-b quark Cabibbo-Kobayashi-Maskawa matrix element is very small ($|V_{cb}|^2 \approx 0.0016$), the contribution of a b-quark initial state in the PDF is negligible. Comparing the d-quark and s-quark PDFs, the probability of interaction of a gluon with an d-quark is greater than that with a s-quark, but the CKM matrix element suppresses $d \rightarrow c$ transitions since $|V_{cd}|^2 \approx 0.04$. As a result, the expected contributions from s-quark and d-quark initial states for a jet transverse momentum $p_T > 20$ GeV at the Tevatron are around 85% and 15%, respectively [6]. According to the ALPGEN+PYTHIA [7,8] simulation for $W + c$-jet events, the contribution from $qg \rightarrow W + c$ dominates the entire $20 < p_T < 100$ GeV region with the contribution from $qq' \rightarrow W + c\bar{c}$ events increasing from about 25% to 45% as jet $p_T$ increases from 20 to 100 GeV. Measurement of the $p\bar{p} \rightarrow W + c$-jet differential cross section should provide information about the s-quark PDF. This PDF has been measured directly only in fixed target neutrino-nucleon deep inelastic scattering experiments at relatively low momentum transfer $Q \lesssim 15 – 20$ GeV [9–14]. A probe of the s-quark PDF at the Tevatron tests the universality of $s(x, Q^2)$, where x is the Feynman variable [15], and its QCD evolution up to $Q^2 \approx 10^4$ GeV$^2$.

There are only a few previous measurements of the $W + c$-jet cross section at hadron colliders, performed by the D0 [2], CDF [16,17], ATLAS [18], and CMS [19] Collaborations. The previous D0 and CDF measurements are inclusive; the CMS and ATLAS inclusive results were augmented by distributions in the pseudorapidity of the lepton from $W$ decay. It is important to note that the
measurements listed above are performed by requiring opposite electric charges of a soft lepton inside a jet from semileptonic charmed hadron decay with a lepton from $W$ decay elsewhere in the event, and measuring the cross section for “opposite-sign” (OS) versus “same-sign” (SS) events. A requirement of opposite signs for the leptons and subtraction of events with the same signs suppresses the sign-symmetric backgrounds as well as $W + c\bar{c}$ events due to gluon splitting, which become significant at high jet $p_T$. All measurements are in agreement with the perturbative next-to-leading order (NLO) QCD predictions that include contributions from gluon splitting within total theoretical uncertainties of 15–30%.

Measurements of the inclusive $W + b$-jet cross-sections have been reported by the CDF, D0, and ATLAS Collaborations. The CDF result is approximately $3\sigma$ higher than the NLO predictions while the D0 and ATLAS measurements agree with the theory within large (30–40%) theoretical uncertainties. A dominant ($\approx 85\%$) contribution to $W + b$-jet production at the Tevatron is due to the $q\bar{q}' \rightarrow W + g (g \rightarrow b\bar{b})$ process while the remaining contribution arises from the $b\bar{q} \rightarrow W b\bar{q}'$ process, with a negligible contribution from single top quark production.

We present, for the first time, measurements of $W + c$-jet and $W + b$-jet differential cross sections as a function of jet $p_T$, where no requirement of a soft lepton within a jet is made, and that are therefore sensitive to the gluon splitting contributions. The $W$ boson candidates are identified in the $\mu + \nu$ decay channel. The data used in this analysis were collected between July 2006 and September 2011 using the D0 detector at the Fermilab Tevatron Collider at $\sqrt{s} = 1.96$ TeV, and correspond to an integrated luminosity of 8.7 fb$^{-1}$. The D0 detector has a central tracking system consisting of a silicon microstrip tracker (SMT) and a central fiber tracker, both located within a 1.9 T superconducting solenoidal magnet, which are optimized for tracking and vertexing at pseudorapidities $|\eta| < 3$ and $|\eta| < 2.5$, respectively. A liquid argon and uranium calorimeter has a central section (CC) covering pseudorapidities $|\eta| \lesssim 1.1$, and two end calorimeters (EC) that extend coverage to $|\eta| \approx 4.2$, with all three housed in separate cryostats. An outer muon system covering $|\eta| < 2$ consists of a layer of tracking detectors and scintillation trigger counters in front of 1.8 T iron toroids, followed by two similar layers after the toroids. Luminosity is measured using plastic scintillator arrays located in front of the EC cryostats.

The $W + b/c$ candidate events are chosen by selecting single muon or muon-$i$-jet signatures with a three-level trigger system. The trigger efficiency has been estimated using $Z \rightarrow \mu^+\mu^-$ ($\pm$-jets) events in data. The trigger efficiency is parametrized as a function of muon $p_T$ and $\eta$ and is on average $\approx 70\%$.

Offline event selection requires a reconstructed $p\bar{p}$ interaction primary vertex (PV) that has at least three associated tracks and is located within 60 cm of the center of the detector along the beam direction. The vertex selection for $W + b/c$ events is approximately 90% efficient as measured in simulation.

We require a muon candidate to be reconstructed from hits in the muon system and matched to a reconstructed track in the central tracker. The transverse momentum of the muon must satisfy $p_T > 20$ GeV, with $|\eta| < 1.7$. Muons are required to be spatially isolated from other energetic particles using information from the central tracking detectors and calorimeter. Muons from cosmic rays are rejected by applying a timing criterion on the hits in the scintillator layers of the muon system and by applying restrictions on the displacement of the muon track with respect to the PV. The muon reconstruction efficiency is $\approx 90\%$.

Candidate $W + j$-jets events are selected by requiring at least one reconstructed jet with pseudorapidity $|\eta^{|\text{j}|}| < 1.5$ and $p_T^{|\text{j}|} > 20$ GeV. Jets are reconstructed from energy deposits in the calorimeter using the iterative midpoint cone algorithm with a cone of radius $R = \sqrt{\Delta \eta^2 + \Delta \phi^2} = 0.5$. The energies of jets are corrected for detector response, the presence of noise and multiple $p\bar{p}$ interactions. To enrich the sample with $W$ bosons, events are required to have missing transverse energy $E_T > 25$ GeV due to the neutrino escaping detection. We require that the $W$ boson candidates have a transverse mass $M_T > 40$ GeV.

Backgrounds for this analysis include events from the production of $W$ + light parton jets, $Z/\gamma^* +$ jets, $t\bar{t}$, single top quark, diboson $VV$ ($V = W, Z$) and QCD multijets in which a jet is misidentified as a muon. The $W + c$ and $W + b$ signal and the background processes excluding multijet are simulated using a combination of ALPGEN and PYTHIA MC event generators with PYTHIA providing parton showering and hadronization. We use PYTHIA with CTEQ6L1 PDFs. ALPGEN generates multi-parton final states using tree-level matrix elements (ME). When interfaced with PYTHIA, it employs the MLM scheme to treat ME partons produced from showering in PYTHIA. For the signal process, we also use the SHERPA MC generator that matches partons from the leading-order ME with up to two real parton emissions to the parton-shower jets according to the CKKW matching scheme. The generated events are processed through a GEANT-based simulation of the D0 detector geometry and response. To accurately model the effects of multiple $p\bar{p}$ interactions and detector noise, events from random $p\bar{p}$ crossings with a similar instantaneous luminosity spectrum as in data are overlaid on the MC events. These MC events are then processed using the same reconstruction code as for the data. The MC events are also weighted to take into account the trigger efficiency and small observed differences between MC and data in the distributions of the instantaneous luminosity.
and of the z coordinate of the \( p\bar{p} \) collision vertex.

The \( V + \text{jets} \) processes are normalized to the total inclusive \( W \) and \( Z \)-boson cross sections calculated at NNLO (next-to-next-to-leading order) \(^{38}\). The \( Z \)-boson \( p_T \) distribution is modeled to match the distribution observed in data \(^{39}\), taking into account the dependence on the number of reconstructed jets. To reproduce the \( W \)-boson \( p_T \) distribution in simulated events, we use the product of the measured \( Z \)-boson \( p_T \) spectrum times the ratio of \( W \) to \( Z \)-boson \( p_T \) distributions at NLO \(^{39,40}\). The NLO+NNLL (next-to-next-to-leading log) calculations are used to normalize \( t\bar{t} \) production \(^{41}\), while single top quark production is normalized to NNLO predictions \(^{42}\). The NLO \( WW\), \( WZ\), and \( ZZ \) production cross section values are obtained with the MCFM program \(^{13}\). The multijet background contribution is estimated from data using the “matrix method” \(^{30}\). For the final states studied here the multijet background is small (\( \lesssim 2\% \)) and arises mainly from the semileptonic decays of heavy quarks in which the muon satisfies the isolation requirements. To reduce the contribution from \( t\bar{t} \) production that increases with jet \( p_T \), we restrict the scalar sum of all the jet \( p_T \) values (\( H_T \)) to be less than 175 GeV. This requirement reduces the \( t\bar{t} \) fraction by a factor 1.5 to 2, depending on jet \( p_T \), and has signal efficiency greater than 95% except in the highest PT bin where it falls to 82%. The \( t\bar{t} \) fraction after the \( H_T \) cut varies between 5 and 20% with increasing jet \( p_T \).

Identification of \( b \) and \( c \) jets is crucial for this measurement. Once the inclusive \( W + \text{jets} \) sample is selected, at least one jet is required to be taggable, i.e. it must contain at least two tracks each with at least one hit in the SMT, \( p_T > 1 \) GeV for the highest-\( p_T \) track and \( p_T > 0.5 \) GeV for the next-to-highest \( p_T \) track. These criteria ensure sufficient information to classify the jet as a heavy-flavor candidate and have a typical efficiency of about 90%. Light parton jets (those resulting from light quarks or gluons) are suppressed using a dedicated artificial neural network (\( b\)-NN) \(^{43}\) that exploits the longer lifetimes of heavy-flavor hadrons relative to their lighter counterparts. The inputs to the \( b\)-NN include several characteristic quantities of the jet and associated tracks to provide a continuous output value that tends towards one for \( b \) jets and zero for the light jets. The \( b\)-NN input variables providing most of the discrimination are the number of reconstructed secondary vertices (SV) in the jet, the invariant mass of charged particle tracks associated with the SV (\( M_{SV} \)), the number of tracks used to reconstruct the SV, the two-dimensional decay length significance of the SV in the plane transverse to the beam, a weighted combination of the tracks’ transverse impact parameter significances, and the probability that the tracks associated with the jet originate from the \( p\bar{p} \) interaction vertex, which is referred to as the jet lifetime probability (JLIP). The jet is required to have a \( b\)-NN output greater than 0.5. For jet \( p_T \) in a range between 20 and 150 GeV this selection is (36 – 47)% efficient for \( b \)-jets and (8 – 11)% efficient for \( c \) jets with relative systematic uncertainties of (4.2 – 6.5)% for both the \( b \) jets and \( c \) jets. The systematic uncertainty is obtained from a comparison of the heavy flavor tagging efficiencies in data and MC as described in \(^{44}\). Only 0.2 – 0.4% of light jets are misidentified as heavy-flavor jets and comprise 7% to 15% of the final sample, with a larger fraction at lower jet \( p_T \). In addition to the \( b\)-NN output, we obtain further information by combining the \( M_{SV} \) and JLIP variables, which provide good discrimination between \( b \), \( c \), and light quark jets due to their different masses \(^{45,46}\). We form a single variable discriminant \( D_{MJL} = \frac{1}{2} (M_{SV}/(5 \text{ GeV}) - \text{ln}(\text{JLIP})/20) \) \(^{47}\) and require \( D_{MJL} > 0.1 \) to remove poorly reconstructed events and reduce the number of light-jet events. The efficiency for signal events to pass this selection is 98% for \( b \)-jets and 97% for \( c \)-jets.

After all selection requirements, 5260 events remain in the data sample. We measure the fraction of \( W + c \) and \( W + b \) events in the selected sample by performing a binned maximum likelihood fit to the observed data distribution of the \( D_{MJL} \) discriminant in bins of jet \( p_T \), as shown in Fig. 1 for the bin \( 30 < p_T < 40 \) GeV. The templates for \( W + b \) and \( W + c \) jets are taken from the simulation. Expected contributions from the background processes are subtracted from the \( D_{MJL} \) distribution in data before the fit. The ratio of the \( W \)-light parton jets to \( W + \) all jet flavors has been estimated using \( \text{ALPGEN}+\text{PYTHA} \) MC events taking into account the data-to-MC correction factors as described in Ref. \(^{44}\), and has been cross checked in data using looser cuts on \( b\)-NN output in the range from 0.15–0.3.

The fractions of \( W + b \) and \( W + c \) events after subtraction of background contribution are shown in Fig. 2 as a function of jet \( p_T \). The relative uncertainties on the fractions obtained from the fit range within (7–13)% for \( W + b \) and (6–11)% for \( W + c \). This includes the uncertainty due to \( W + b \) and \( W + c \) template shapes, studied in a previous analysis \(^{47}\). The contributions from the background events are varied within uncertainties on their predicted cross sections, and these uncertainties are propagated into the extracted signal fractions. The uncertainty due to the light parton jets template shape is taken from Ref. \(^{46}\). The overall relative uncertainties on the subtracted backgrounds range within (4–6)% for \( W + b \) and (3–4)% for \( W + c \).

We apply corrections to the measured number of signal events to account for the detector and kinematic acceptances and selection efficiencies using simulated samples of \( W + (c)\)-jet events. In these calculations, we apply the following selections at the particle level: at least one \( (c)\)-jet with \( p_T^{b(c)\text{-jet}} > 20 \) GeV, \(|p_T^{b(c)\text{-jet}}| < 1.5\), a muon with \( p_T^\mu > 20 \) GeV and \(|p_T^\mu| < 1.7\), and a neutrino with \( p_T^\nu > 25 \) GeV. In the following, we quote our cross section results for this restricted phase space as a fiducial cross
The distributions for the background events are subtracted from data before the fit. The distributions for the $c$-jet and $b$-jet templates (with statistical uncertainties) are shown normalized to their respective fitted fractions.

The acceptance is defined by the selection requirements in jet and muon transverse momenta and pseudorapidities. Correction factors to account for small differences between jet-$p_T$ and rapidity spectra in data and simulation are estimated, and used as weights to create a data-like MC sample. The differences between acceptance corrections obtained with standard and corrected MC samples are taken as a systematic uncertainty of up to 3% at low jet $p_T$. An additional systematic uncertainty of up to 4% is due to uncertainties in the jet energy correction and resolution. For $20 < p_T^{c-jet} < 150$ GeV the product of acceptance and muon selection efficiency varies within (50 – 65)% with a relative systematic uncertainty of (3 – 5)%. The systematic uncertainty on the muon selection efficiency is about 2% and is obtained from a comparison of the muon efficiencies in $Z \rightarrow \mu \mu$ events in data and MC. Uncertainties on $b$-jet identification are determined in simulations and data by using $b$-jet-enriched samples $^{[14]}$ and are about (2 – 5)% per jet. The integrated luminosity is known to a precision of 6.1% $^{[18]}$. By summing the uncertainties in quadrature we obtain a final total systematic uncertainty on the cross section measurements of (11 – 18)% depending on jet $p_T$ and final state.

To check the stability of the results, the $W + b$-jet and $W + c$-jet cross sections have been remeasured using a looser $b$-NN selection, $b$-NN output > 0.3, and with the light parton jet fraction included as an additional fit parameter, thus increasing the data statistics and the background fractions. The fractions of $b$- and $c$-jets are obtained from the maximum likelihood fit of the light, $b$- and $c$-jet templates to $D_{\text{MJL}}$ distribution in data as shown in Fig. 3. We also vary the default $H_T$ cut by ±15 GeV and remeasure the cross sections. In both cross-checks, the default and new cross sections are found to be in agreement within uncertainties that include the correlation between the two measurements.

In Figs. 3 and 5 and Tables 4 and 5 we present the $W + b$-jet and $W + c$-jet differential production cross sections times $W \rightarrow \mu \nu$ branching fraction for the fiducial phase space defined by $p_T^{\mu} > 20$ GeV, $|\eta^{\mu}| < 1.7$, $p_T^Z > 25$ GeV, and with at least one $b(c)$-jet with $p_T^{b(c)} > 20$ GeV and $|\eta^{p_{T}^{b(c)}}| < 1.5$. The cross sections are presented differentially in five $p_T^{c+b}$ bins in the region $20 – 150$ GeV. The data points are plotted at the value of $p_T^{c+b}$ for which the value of a smooth function describing the cross section equals the averaged cross section in the bin $^{[49]}$. 

FIG. 1: (Color online) Distribution of the $D_{\text{MJL}}$ discriminant after all selection criteria (including $b$-NN output > 0.5) for a representative bin of $30 < p_T^{c-jet} < 40$ GeV. The contributions from background events are subtracted from data before the fit. The distributions for the $c$-jet and $b$-jet templates (with statistical uncertainties) are shown normalized to their respective fitted fractions.

FIG. 2: (Color online) The $b$- and $c$-jet fractions (their total sum is normalized to 1.0) versus jet $p_T$ with total uncertainty from the $D_{\text{MJL}}$ fit.

FIG. 3: (Color online) Similar to Fig. 1 but for events selected with $b$-NN output > 0.3. This alternative selection is used as a cross-check of the main results.
The cross sections are compared to predictions from NLO QCD and two MC generators, SHERPA and ALPGEN. The NLO predictions are made using the MSTW2008 PDF set and CT10 PDF sets. We calculate the NLO QCD prediction using MC with central values of renormalization and fragmentation scales \( \mu_r = \mu_f = M_W \) and with the \( b \)-quark and \( c \)-quark masses \( m_b = 4.75 \) GeV and \( m_c = 1.5 \) GeV, respectively. Uncertainties are estimated by varying \( \mu_r \) and \( \mu_f \) independently by a factor of two in each direction.

The NLO predictions are corrected for non-perturbative effects such as parton-to-hadron fragmentation and multiple parton interactions. The latter are evaluated using SHERPA and PYTHIA, MC samples generated using their default settings. The overall corrections vary within a factor of 0.80–1.1 with an uncertainty of \( \lesssim 5\% \) assigned to account for the difference between the two MC generators. The ratios of data over the NLO QCD calculations and of the various theoretical predictions to the NLO QCD calculations are presented in Figs. 6 and 7. The measured \( W + b \)-jet cross sections are systematically above the NLO QCD predictions for all \( p_T \) bins. The \( W + c \)-jet data agree with the NLO QCD predictions at small \( p_T \) but disagree at higher \( p_T \) as the contribution from \( q\bar{q}' \to W + g \ (g \to c\bar{c}) \) events increases.

In addition to measuring the \( W + b \)-jet and \( W + c \)-jet cross-sections, we calculate the ratio \( \sigma(W+c)/\sigma(W+b) \) in \( p_T \) bins. In this ratio, many experimental systematic uncertainties cancel. Also, theory predictions of the ratio are less sensitive to the scale uncertainties and effects from missing higher-order terms that impact the normalizations of the cross sections. The remaining uncertainties are caused by largely anti-correlated uncertainties coming from the fitting of \( c \)-jet and \( b \)-jet \( D_M \) templates to data, and by other uncertainties on the \( b \)- and \( c \)-jet fractions discussed above. Experimental results as well as theoretical predictions for the ratios are presented in Table and Fig. The systematic uncertainties on the ratio vary within \((11–17)\%\). Theoretical scale uncertainties, estimated by varying the renormalization and factorization scales by a factor of two in the same way for the \( \sigma(W+b) \) and \( \sigma(W+c) \) predictions, are also significantly reduced. Specifically, residual scale uncertainties are typically \((0.5–4.6)\%\) for NLO QCD, which indicates a much smaller dependence of the ratio on the higher-order corrections. The ratio \( \sigma(W+c)/\sigma(W+b) \) for \( p_T^{jet} > 30 \) GeV is reasonably consistent with theoretical predictions except for SHERPA.

In summary, we have performed the first measurement of the differential cross section as a function of \( p_T^{jet} \) for the \( W + b \)-jet and \( W + c \)-jet final states with \( W \to \mu\nu \) decay at \( \sqrt{s} = 1.96 \) TeV, in a restricted phase space of \( p_T^{jet} > 20 \) GeV, \( |\eta^{\mu\nu}| < 1.7 \), \( p_T^{jet} > 25 \) GeV and with \( b(c) \) jets with the \( p_T \) range \( 20 < p_T^{jet} < 150 \) GeV and \( |\eta^{jet}| < 1.5 \).
TABLE I: The $W + b$-jet production cross sections times $W \rightarrow \mu \nu$ branching fraction, $d\sigma/dp_T^\mu$, together with statistical uncertainties ($\delta_{\text{stat}}$) and total systematic uncertainties ($\delta_{\text{syst}}$). The column $\delta_{\text{tot}}$ shows total experimental uncertainty obtained by adding $\delta_{\text{stat}}$ and $\delta_{\text{syst}}$ in quadrature. The last three columns show theoretical predictions obtained using NLO QCD with MSTW PDF set, and two MC event generators, SHERPA and ALPGEN.

| $p_T^\mu$ bin ($p_T^\mu$) (GeV) | $\langle p_T^\mu \rangle$ (GeV) | $d\sigma/dp_T^\mu$ (pb/GeV) | Data | $\delta_{\text{stat}}$(%) | $\delta_{\text{syst}}$(%) | $\delta_{\text{tot}}$(%) | NLO QCD | SHERPA | ALPGEN |
|-----------------------------|-----------------|----------------|---|----|----|----|---|---|---|
| 20–30 | 24.3 | $9.6 \times 10^{-2}$ | 2.4 | 17.8 | 18.0 | $6.5 \times 10^{-2}$ | $3.9 \times 10^{-2}$ | $3.9 \times 10^{-2}$ |
| 30–40 | 34.3 | $4.0 \times 10^{-2}$ | 2.9 | 13.6 | 13.9 | $3.0 \times 10^{-2}$ | $2.0 \times 10^{-2}$ | $2.0 \times 10^{-2}$ |
| 40–50 | 44.3 | $2.5 \times 10^{-2}$ | 3.6 | 14.4 | 14.8 | $1.6 \times 10^{-2}$ | $1.1 \times 10^{-2}$ | $1.1 \times 10^{-2}$ |
| 50–70 | 57.2 | $1.2 \times 10^{-2}$ | 3.4 | 15.2 | 15.6 | $7.4 \times 10^{-3}$ | $5.5 \times 10^{-3}$ | $5.2 \times 10^{-3}$ |
| 70–150 | 81.7 | $2.2 \times 10^{-3}$ | 4.5 | 17.7 | 18.3 | $1.4 \times 10^{-3}$ | $1.0 \times 10^{-3}$ | $9.3 \times 10^{-4}$ |

TABLE II: The $W + c$-jet production cross sections times $W \rightarrow \mu \nu$ branching fraction, $d\sigma/dp_T^c$, together with statistical uncertainties ($\delta_{\text{stat}}$) and total systematic uncertainties ($\delta_{\text{syst}}$). The column $\delta_{\text{tot}}$ shows total experimental uncertainty obtained by adding $\delta_{\text{stat}}$ and $\delta_{\text{syst}}$ in quadrature. The last three columns show theoretical predictions obtained using NLO QCD with MSTW PDF set, and two MC event generators, SHERPA and ALPGEN.

| $p_T^c$ bin ($p_T^c$) (GeV) | $\langle p_T^c \rangle$ (GeV) | $d\sigma/dp_T^c$ (pb/GeV) | Data | $\delta_{\text{stat}}$(%) | $\delta_{\text{syst}}$(%) | $\delta_{\text{tot}}$(%) | NLO QCD | SHERPA | ALPGEN |
|-----------------------------|-----------------|----------------|---|----|----|----|---|---|---|
| 20–30 | 24.2 | $4.1 \times 10^{-1}$ | 3.7 | 17.0 | 17.4 | $4.1 \times 10^{-1}$ | $2.1 \times 10^{-1}$ | $2.4 \times 10^{-1}$ |
| 30–40 | 34.2 | $2.6 \times 10^{-1}$ | 4.6 | 11.0 | 11.9 | $1.8 \times 10^{-1}$ | $9.2 \times 10^{-2}$ | $1.1 \times 10^{-1}$ |
| 40–50 | 44.2 | $1.5 \times 10^{-1}$ | 5.8 | 11.9 | 13.2 | $9.2 \times 10^{-2}$ | $4.6 \times 10^{-2}$ | $5.9 \times 10^{-2}$ |
| 50–70 | 57.0 | $8.4 \times 10^{-2}$ | 5.3 | 12.1 | 13.2 | $3.9 \times 10^{-2}$ | $2.0 \times 10^{-2}$ | $2.6 \times 10^{-2}$ |
| 70–150 | 80.7 | $1.3 \times 10^{-2}$ | 6.9 | 15.6 | 17.1 | $6.1 \times 10^{-3}$ | $3.1 \times 10^{-3}$ | $3.8 \times 10^{-3}$ |

FIG. 6: (Color online) The ratio of $W + b$-jet production cross sections to NLO predictions with the MSTW2008 PDF set [50] for data and theoretical predictions. The uncertainties on the data include both statistical (inner error bar) and total uncertainties (full error bar). Also shown are the uncertainties on the theoretical QCD scales. The ratio of NLO QCD predictions with CT10 [51] to those obtained with MSTW2008 as well as the predictions given by SHERPA [52] and ALPGEN [7] are also presented.

FIG. 7: (Color online) The ratio of $W + c$-jet production cross sections to NLO predictions with the MSTW2008 PDF set [50] for data and theoretical predictions. The uncertainties on the data include both statistical (inner error bar) and total uncertainties (full error bar). Also shown are the uncertainties on the theoretical QCD scales. The ratio of NLO QCD predictions with CT10 [51] to those obtained with MSTW2008 as well as the predictions given by SHERPA [52] and ALPGEN [7] are also presented.
TABLE III: The $\sigma(W+c)/\sigma(W+b)$ cross section ratio in bins of $c(b)$-jet $p_T$ together with statistical uncertainties ($\delta_{\text{stat}}$), total systematic uncertainties ($\delta_{\text{syst}}$). The column $\delta_{\text{tot}}$ shows total experimental uncertainty obtained by adding $\delta_{\text{stat}}$ and $\delta_{\text{syst}}$ in quadrature. The last three columns show theoretical predictions obtained with the NLO QCD using MSTW2008 PDF set, and two MC event generators, sherpa and alpgen.

| $p_T^{\text{c}}$ (GeV) | $\sigma (W + c)/ \sigma (W + b)$ | Data $\delta_{\text{stat}}$ (%) | Data $\delta_{\text{syst}}$ (%) | Data $\delta_{\text{tot}}$ (%) | NLO QCD | SHERPA | ALPGEN |
|-----------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|--------|--------|--------|
| 20–30                 | 24.3                          | 6.6                           | 13.6                          | 20.2                          | 6.2    | 5.4    | 6.2    |
| 30–40                 | 34.3                          | 6.1                           | 13.0                          | 21.2                          | 6.1    | 4.7    | 5.7    |
| 40–50                 | 44.3                          | 6.1                           | 13.9                          | 23.9                          | 5.8    | 4.2    | 5.4    |
| 50–70                 | 57.1                          | 7.2                           | 13.8                          | 24.7                          | 5.3    | 3.7    | 4.9    |
| 70–150                | 81.2                          | 5.7                           | 17.5                          | 23.2                          | 4.3    | 3.0    | 4.1    |

These are the first measurements of $W+b/c$ cross sections that are sensitive to the gluon splitting processes. The measured $W+b/c$-jet cross section is higher than the predictions in all $p_T$ bins and is suggestive of missing higher order corrections. The measured $W+c$-jet cross section agrees with NLO prediction for the low $p_T$ ($20–30$ GeV), but disagrees towards high $p_T$. The disagreement may be due to missing higher order corrections and an underestimated contribution from gluon splitting $g \rightarrow cc$ also observed earlier at LEP [52], LHCb [53], ATLAS [54] and D0 experiments [7], [55], and/or possible enhancement in the strange quark PDF as suggested by CHORUS [56], CMS [19] and ATLAS [18] data according to a recent PDF fit performed by ABKM group [57].

We thank John Campbell for useful discussions and predictions with MCFM. We thank the staffs at Fermilab and collaborating institutions, and acknowledge support from the DOE and NSF (USA); CEA and CNRS/IN2P3 (France); MON, NRC KI, and RFBR (Russia); CNPq and FAPERJ (Brazil); DAE and DST (India); COLCIENCIAS (Colombia); CONACYT (Mexico); NRF (Korea); FOM (The Netherlands); STFC and The Royal Society (UK); MSMT (Czech Republic); BMBF and DFG (Germany); SFI (Ireland); Swedish Research Council (Sweden); CAS and CNSF (China); and MESU (Ukraine).
The measured properties of W boson candidates are limited by their transverse energy and transverse mass, defined as 
$$M_T = \sqrt{(\hat{p}_T + p_T^l)^2 - (\hat{p}_x + p_x^l)^2 - (\hat{p}_y + p_y^l)^2},$$
where $\hat{p}_T$ is the magnitude of the missing transverse energy vector, $p_T^l$ is the transverse momentum of the lepton and $p_x^l$, $p_y^l$, and $p_z^l$ are the magnitude of the x and y components of the lepton’s momentum (missing transverse energy), respectively.

[17] T. Aaltonen et al. (CDF Collaboration), Phys. Rev. Lett. 110, 071801 (2013).
[18] G. Aad et al. (ATLAS Collaboration), J. High Energy Phys. 05, 068 (2014).
[19] S. Chatrchyan, et al. (CMS Collaboration), J. High Energy Phys. 02, 013 (2014).
[20] J. Campbell, R. K. Ellis, F. Maltoni, and S. Willenbrock, Phys. Rev. D 75, 054015 (2007).
[21] J. Campbell et al., Phys. Rev. D 79, 034023 (2009).
[22] T. Aaltonen et al. (CDF Collaboration), Phys. Rev. Lett. 104, 131801 (2010).
[23] V. M. Abazov et al. (D0 Collaboration), Phys. Lett. B 718, 1314 (2013).
[24] G. Aad et al. (ATLAS Collaboration), Phys. Lett. B 707, 418 (2012).
[25] V. M. Abazov et al. (D0 Collaboration), Nucl. Instrum. Methods in Phys. Res. A 565, 463 (2006).
[26] R. Angstadt et al., Nucl. Instrum. Methods in Phys. Res. Sect. A 622, 298 (2010).
[27] We use a standard right-handed coordinate system. The nominal collision point is the center of the detector with coordinates (0, 0, 0). The direction of the proton beam is the +z axis. The +x axis is horizontal, pointing away from the center of the Tevatron ring. The +y axis points vertically upwards. The polar angle, $\theta$, is defined such that $y = 0$ is the +z direction. The rapidity is defined as $y = -\ln [(E + p_y)/(E - p_y)]$, where $E$ is the energy and $p_y$ is the momentum component along the proton beam direction. Pseudorapidity is defined as $\eta = -\ln (\tan(\frac{\theta}{2}))$. $\phi$ is defined as the azimuthal angle in the plane transverse to the proton beam direction.
[28] S. Abachi et al. (D0 Collaboration), Nucl. Instrum. Methods in Phys. Res. A 324, 53 (1993).
[29] V. M. Abazov et al. (D0 Collaboration), Nucl. Instrum. Methods in Phys. Res. Sect. A 737, 281 (2014).
[30] V. M. Abazov et al. (D0 Collaboration), Phys. Rev. D. 86, 032005 (2012).
[31] G. C. Blazey et al., arXiv:hep-ex/0005012 (2000). We use the algorithm ILCA described in section 3.4.
[32] V. M. Abazov et al. (D0 Collaboration), Nucl. Instrum. Methods in Phys. Res. A 763, 442 (2014).
[33] The measured properties of W boson candidates are limited by their transverse energy and transverse mass, defined as $M_T = \sqrt{(\hat{p}_T + p_T^l)^2 - (\hat{p}_x + p_x^l)^2 - (\hat{p}_y + p_y^l)^2}$, where $\hat{p}_T$ is the magnitude of the missing transverse energy vector, $p_T^l$ is the transverse momentum of the lepton and $p_x^l$, $p_y^l$, and $p_z^l$ are the magnitude of the x and y components of the lepton’s momentum (missing transverse energy), respectively.