,pdftitle=Measurement of associated charm production in W final states at sqrt(s) = 7 TeV,pdfsubject=CMS,pdf
physics, EWK, Standard Model Physics

Available on the CERN CDS information server **CMS PAS SMP-12-002**

CMS Physics Analysis Summary

Contact: cms-pag-conveners-smp@cern.ch 2013/03/10

Measurement of associated charm production in W final states at $\sqrt{s} = 7$ TeV

The CMS Collaboration

Abstract

We present a detailed study of the associated production of a W boson and one c-quark jet in pp collisions at a center-of-mass energy of 7 TeV. The analysis has been conducted with a data sample corresponding to a total integrated luminosity of 5 fb^{-1}, collected by the CMS detector at the LHC.
Candidates of W bosons are identified through their decay into a charged lepton (muon or electron) and a neutrino, in the pseudorapidity range $|\eta^{\ell}| < 2.1$. The W + c measurements are performed for charm jets in the kinematic region $p_T^{jet} > 25$ GeV, $|\eta^{jet}| < 2.5$ and for two different thresholds of transverse momentum of the lepton from the W decay. Using hadronic and inclusive semileptonic decays of charm hadrons, we measure the following total cross sections:

$$\sigma(\mathrm{pp} \to \mathrm{W} + \mathrm{c} + \mathrm{X}) \times \mathcal{B}(\mathrm{W} \to \ell\nu) = 107.7 \pm 3.3\,(\mathrm{stat.}) \pm 6.9\,(\mathrm{syst.})\ \mathrm{pb}\ (\mathrm{p_T^{\ell}} > 25\,\mathrm{GeV})$$

and

$$\sigma(\mathrm{pp} \to \mathrm{W} + \mathrm{c} + \mathrm{X}) \times \mathcal{B}(\mathrm{W} \to \ell\nu) = \ \ 84.1 \pm 2.0\,(\mathrm{stat.}) \pm 4.9\,(\mathrm{syst.})\ \mathrm{pb}\ (\mathrm{p_T^{\ell}} > 35\,\mathrm{GeV}),$$

and the cross section ratios:

$$\frac{\sigma(\mathrm{pp} \to \mathrm{W}^+ + \bar{\mathrm{c}} + \mathrm{X})}{\sigma(\mathrm{pp} \to \mathrm{W}^- + \mathrm{c} + \mathrm{X})} = 0.954 \pm 0.025\,(\mathrm{stat.}) \pm 0.001\,(\mathrm{syst.})\ (\mathrm{p_T^{\ell}} > 25\,\mathrm{GeV})$$

and

$$\frac{\sigma(\mathrm{pp} \to \mathrm{W}^+ + \bar{\mathrm{c}} + \mathrm{X})}{\sigma(\mathrm{pp} \to \mathrm{W}^- + \mathrm{c} + \mathrm{X})} = 0.939 \pm 0.019\,(\mathrm{stat.}) \pm 0.005\,(\mathrm{syst.})\ (\mathrm{p_T^{\ell}} > 35\,\mathrm{GeV}).$$

These cross sections and cross section ratios provide important information on the strange and anti-strange quark parton density functions of the proton at the electroweak scale. Cross sections and cross section ratios are also measured differentially with respect the pseudorapidity of the lepton from the W decay. Results are compared with theoretical predictions at next-to-leading order using different parton distribution function sets.
Whereas current theoretical predictions are in agreement with the measured $\mathrm{W} + \mathrm{c}$ cross sections at the 10% level, an analysis of data distributions sensitive to $\mathrm{W} + \mathrm{c\bar{c}}$ and $\mathrm{W} + \mathrm{b\bar{b}}$ backgrounds reveals large differences between Monte Carlo generators. In particular, simulations based on PYTHIA show a large deficit when compared with experimental observations in some specific kinematic configurations.

1 Introduction

The study of associated production of a W boson and a charm (c) quark at hadron colliders (hereafter referred as $W + \mathrm{c}$ production) provides direct access to the strange quark content of the proton at the electroweak scale [1–3]. This sensitivity is due to the dominance of $\bar{\mathrm{s}}\mathrm{g} \to W^+ + \bar{\mathrm{c}}$ and $\mathrm{sg} \to W^- + \mathrm{c}$ contributions at the hard scattering level (Figure 1). Recent work [4] indicates that precise measurements of this process at the LHC may significantly reduce the uncertainties on the strange parton distribution function (PDF) and help resolving existing ambiguities and limitations of low energy neutrino Deep Inelastic Scattering (DIS) data. Improving the PDF knowledge is essential for many present and future precision analyses, like the measurement of the W mass [5]. An asymmetry between the strange and anti-strange PDFs has also been proposed as an explanation of the so called NuTeV anomaly [6]. The $\mathrm{pp} \to W + \mathrm{c} + \mathrm{X}$ process is a sizable background for signals involving heavy quarks and missing transverse energy in the final state. Particularly relevant cases are top-quark studies and third-generation squark searches. Measurements of the $\mathrm{p\bar{p}} \to W + \mathrm{c} + X$ cross section and of the cross section ratio $\mathrm{p\bar{p}} \to W + \mathrm{c} + X / \mathrm{p\bar{p}} \to W + \mathrm{jets} + X$ have been performed with a relative precision of about 20-30% at the Tevatron hadron collider [7–9], using semileptonic charm hadron decays.

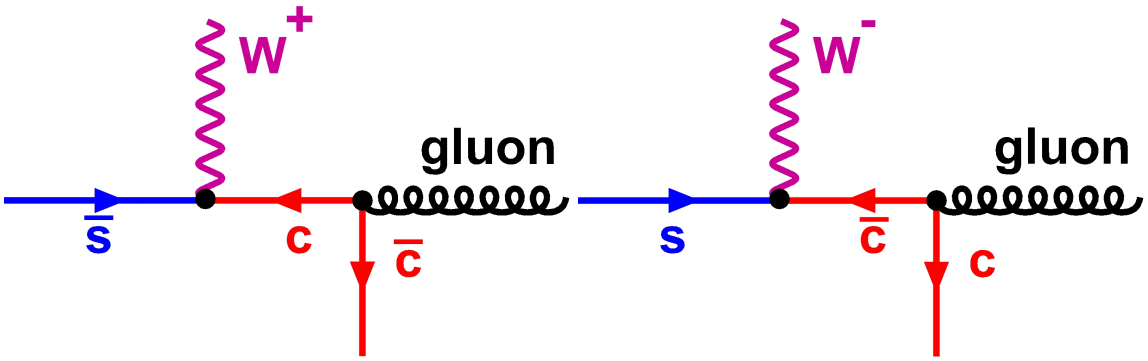

Figure 1: Main diagrams at the hard scattering level for associated W - charm production at the LHC.

$W + \mathrm{c}$ production receives contributions at the few percent level from the processes $\bar{\mathrm{d}}\mathrm{g} \to W^+ + \bar{\mathrm{c}}$ and $\mathrm{dg} \to W^- + \mathrm{c}$, which are Cabibbo suppressed. Overall, the $W^- + \mathrm{c}$ yield is expected to be slightly larger than the $W^+ + \bar{\mathrm{c}}$ yield at the LHC, due to the participation of down valence quarks in the initial state. A key property of the $\mathrm{qg} \to W + \mathrm{c}$ reaction is the presence of a charm quark and a W boson with opposite sign charges.

This document presents a detailed study of the $\mathrm{pp} \to W + \mathrm{c} + \mathrm{X}$ process with the CMS detector, using a data sample corresponding to a total integrated luminosity of $5\,\mathrm{fb}^{-1}$ collected in 2011, at a center-of-mass energy of $\sqrt{s} = 7$ TeV. We measure the total cross section and the charge ratio $R_c^{\pm} = \sigma(W^+ + \bar{\mathrm{c}})/\sigma(W^- + \mathrm{c})$ using muonic and electronic decays of the W boson. Charm jets are identified within the fiducial cuts $p_{\mathrm{T}}^{jet} > 25$ GeV, $|\eta^{jet}| < 2.5$ using exclusive hadronic and inclusive semileptonic decays of charm hadrons. Furthermore, the cross section and the $R_c^{\pm}$ ratio are measured differentially as a function of the pseudorapidity of the lepton from W decay.

The structure of this note is as follows: the CMS detector is briefly described in section 2 and the general analysis strategy is outlined in section 3. The samples used to carry out the measurement and the event selection criteria are presented in sections 4 and 5. Section 6 details the

measurement of the total cross section and sections 7 and 8 are devoted to studies of the differential cross section and the $R_c^{\pm}$ ratio. Results and comparisons with theoretical predictions are discussed in section 9. Section 10 contains a specific study of distributions sensitive to $W + c\bar{c}$ and $W + b\bar{b}$ backgrounds. Finally, we summarize the results of this study in section 11.

2 The CMS detector

The central feature of the CMS apparatus is a superconducting solenoid of 6 m internal diameter, providing a magnetic field of 3.8 T. Within the field volume are a silicon pixel and strip tracker, an electromagnetic calorimeter (ECAL) and a brass/scintillator hadron calorimeter (HCAL). Muons are detected in gas–ionization detectors embedded in the steel return yoke. In addition to the barrel and endcap detectors, CMS has extensive forward calorimetry.

CMS uses a right-handed coordinate system, with the origin at the nominal interaction point, the x-axis pointing to the center of the LHC ring, the y-axis pointing up (perpendicular to the LHC plane), and the z-axis along the anticlockwise-beam direction. The polar angle θ is measured from the positive z-axis and the azimuthal angle ϕ is measured in radians in the xy-plane. The pseudorapidity is given by $\eta = -\ln(\tan(\theta/2))$.

The inner tracker measures charged particle trajectories in the pseudorapidity range $|\eta| \leq 2.5$. It consists of 1440 silicon pixel and 15148 silicon strip detector modules. It provides an impact parameter resolution of 15 μm and a transverse momentum (p_T) resolution of about 1% for charged particles with p_T around 40 GeV. The electromagnetic calorimeter consists of nearly 76000 lead tungstate crystals which provide coverage in pseudorapidity $|\eta| \leq 1.479$ in a cylindrical barrel region (EB) and $1.479 \leq |\eta| \leq 3.0$ in two endcap regions (EE). A preshower detector consisting of two planes of silicon sensors interleaved with a total of three radiation lengths of lead is located in front of the EE. The ECAL has an ultimate energy resolution of better than 0.5% for unconverted photons with transverse energies above 100 GeV. The energy resolution is 3% or better for the range of electron energies relevant for this analysis. The hadronic calorimeter is a sampling device with brass as passive material and scintillator as active material. The combined calorimeter cells are grouped in projective towers of granularity $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$ at central rapidities and 0.175×0.175 at forward rapidities. Muons are detected in the pseudorapidity range $|\eta| \leq 2.4$, with detection planes based on three technologies: drift tubes, cathode strip chambers, and resistive plate chambers. A high-p_T muon originating from the interaction point produces track segments in typically three or four muon stations. Matching these segments to tracks measured in the inner tracker results in a p_T resolution between 1% and 2% for p_T values up to 100 GeV. The first level (L1) of the CMS trigger system, composed of custom hardware processors, is designed to select the most interesting events in less than 1 ms using information from the calorimeters and muon detectors. The High Level Trigger (HLT) processor farm further decreases the event rate to a few hundred hertz, before data storage. A more detailed description of CMS can be found elsewhere [10].

3 Analysis Strategy

We study $W + c$ associated production in final states containing a $W \to \ell\nu$ decay plus a leading jet with charm content. Jets originating from a c ($\bar{c}$) parton are identified using one of the three following signatures: a) a displaced secondary vertex with 3 tracks and an invariant mass consistent with a $D^+ \to K^-\pi^+\pi^+$ ($D^- \to K^+\pi^-\pi^-$) decay, b) a displaced secondary vertex with 2 tracks consistent with a $D^0 \to K^-\pi^+$ ($\bar{D^0} \to K^+\pi^-$) decay and associated to a previous $D^{*+}(2010) \to D^0\pi^+$ ($D^{*-}(2010) \to \bar{D^0}\pi^-$) decay at the primary vertex, c) a semileptonic decay

leading to a well identified muon. In total, and taking into account that both electron and muon channels are considered in the W decay, six different final states are explored.

$D^{\pm}$, $D^{*\pm}(2010)$ and $c \to \ell\nu$ decays provide a direct measurement of the charm-jet charge, a powerful tool to disentangle the $W + c$ signal component from most of the background processes. We define two types of distributions: opposite-sign distributions, denoted by OS in the following, are built on samples containing a W and a charm jet with opposite charge sign; same-sign distributions, denoted by SS, are built from samples where the W and the charm jet have the same charge sign. A dedicated study of SS distributions is presented in section 10. The final distributions used in the analysis are obtained by subtracting the OS and SS distributions for any given variable. This subtraction has no effect on the signal at leading order. In contrary, $W + c\bar{c}$ and $W + b\bar{b}$ events provide the same OS and SS contributions and are suppressed in OS-SS subtracted distributions. Moreover, any OS-SS asymmetry present in $t\bar{t}$, single-top or $W +$ light jet backgrounds is found to be negligible for practical purposes according to simulations. As a consequence OS-SS subtracted distributions are largely dominated by the $W + c$ component, allowing many detailed studies of the $pp \to W + c + X$ process at current LHC integrated luminosities.

Using displaced vertices and exclusive final states from $D^{\pm}$ and $D^{*\pm}(2010)$ mesons constitutes a simple way of suppressing a priori several backgrounds, like Drell-Yan, $W +$ light jet and multijet final states with no heavy flavor content. It also reduces backgrounds containing b-hadron decays, which often lead to more charged tracks than a typical D-meson decay. The sample containing semileptonic charm decays is complementary: it has a larger statistics than the D-meson exclusive channels, but it is more affected by backgrounds, in particular Drell-Yan.

Cross section measurements demand a precise accounting of systematic uncertainties on charm branching fractions and acceptance. In this context exclusive identification of $D^{\pm}$ and $D^{*\pm}(2010)$ becomes mandatory. However, only charge identification is strictly required for studies that are independent of the overall $W + c$ normalization, like relative differential measurements or measurements of the $(W^+ + \bar{c})/(W^- + c)$ ratio. We employ semiinclusive selections in these cases in order to increase the statistics. The semiinclusive sample associated to the $D^{\pm}$ channel consists of all secondary vertices with three tracks. The semiinclusive equivalent sample of the $D^{*\pm}(2010)$ channel is obtained by significantly relaxing the criteria on the mass difference between $D^{*\pm}(2010)$ and D^0 candidates in order to include decays with additional neutral pions. Even with these relaxed criteria simulations predict that the background contributions to the OS-SS subtracted distributions are small compared with the signal yield.

4 Data streams, Monte Carlo samples and signal definition

The analysis reported in this document has been performed with the data sample of proton-proton collisions at $\sqrt{s} = 7\,\text{TeV}$ collected with the CMS detector in 2011. A detailed data validation and certification process guarantees that the total integrated luminosity available for analysis, $\mathcal{L} = 5.0 \pm 0.1\,\text{fb}^{-1}$, fulfills all quality requirements for all detectors used in this analysis. Candidate events in the muon decay channel of the W boson are selected online by a single muon trigger that requires a reconstructed muon with a $p_T > 24\,\text{GeV}$. Candidate events in the electron channel are selected by a variety of electron triggers. Trigger conditions were tightened for the different periods of the data taking to cope with the increasing instantaneous luminosity of the machine. Most of the data used in this analysis are selected requiring an electron candidate with $E_T > 32\,\text{GeV}$.

Muons and electron candidates are reconstructed following CMS standard algorithms [11], [12].

Jets and missing transverse energy related quantities are computed using particle-flow techniques [13] where a full reconstruction of the event in terms of the individual particles is performed from the hits and energy depositions in the different subdetectors.

Large samples of events simulated with Monte Carlo (MC) techniques are used to evaluate signal and background efficiencies. W signal ($W \to \mu\nu$ and $W \to e\nu$) as well as other electroweak (EWK) processes (such as Z$\to \ell\ell$, W$\to \tau\nu$ and Z$\to \tau\tau$ production) are generated with the MADGRAPH [14] (v5.1.1) event generator, interfaced to the PYTHIA [15] (v6.4.24) program for parton shower simulation. The MADGRAPH generator produces parton-level events with a vector boson and up to four partons on the basis of a matrix-element calculation. It has shown to reproduce successfully the observed jet multiplicity and kinematic properties of W + jets final states at the LHC energy regime [16]. The matching scale is $m^2 = (10\,\text{GeV})^2$ and the factorization and renormalization scales are set at $q^2 = M^2_{W/Z} + p^2_{T,W/Z}$. No cuts were applied in phase space at the generator level, except for the $Z(\gamma^*)$ case ($M_{\ell\ell} > 10$ GeV).

Potential backgrounds in this analysis are $t\bar{t}$ and single-top. A sample of $t\bar{t}$ events is generated also with the MADGRAPH generator interfaced with PYTHIA. Single top events are generated in the t-channel, s-channel and tW associated modes with the next-to-leading order (NLO) generator POWHEG [17] (v1.0), interfaced with PYTHIA. The parton distribution function set used in these POWHEG productions is CT10 [18]. We also consider the small contributions from diboson (WW, WZ, ZZ) events and QCD multijet events using PYTHIA. All leading-order (LO) generations use CTEQ6L1 PDF set [19] and parameters set according to the Z2 tune [20].

Cross sections of the EWK processes are normalized to the cross section predictions from FEWZ [21] evaluated at next-to-next-to-leading Order (NNLO) using MSTW08NNLO [22] PDFs. For the rest of the processes, cross sections are normalized to the next-to-leading order (NLO) cross section predictions from MCFM [23] using the MSTW08NLO PDF set, except for the $t\bar{t}$ cross section that is taken at next-to-next-to-leading-log order (NNLL) from [24]. QCD multijets cross section is evaluated at LO.

Several minimum bias interactions are superimposed on the hard reaction to simulate the real experimental conditions of several pp collisions occurring simultaneously (PileUp events, PU). To reach an optimal agreement with the experimental data, the simulated distributions are reweighted according to the actual number of interactions occurring given the instantaneous luminosity for each bunch crossing. Generated events are processed through the full GEANT4 [25] detector simulation, trigger emulation and event reconstruction chain of the CMS experiment.

At the hard-scattering level we identify W + c signal events as those containing an odd number of charm partons in the final state. This choice provides a simple operational definition of the process and ensures that pure QCD splittings of the $g \to c\bar{c}$ type are associated to the background. The W + c signal reference is defined at the hard scattering level of MADGRAPH, which provides an implicit parton-jet matching for a jet separation parameter of $R = 1$, suitable for comparisons with the NLO theoretical predictions of MCFM at the $\lesssim 1\%$ level. The phase space cuts at the generator level are chosen in order to match approximately the experimental selection cuts used in the analysis. For charm partons we require $p^c_T > 25\,\text{GeV}$, $|\eta^c| < 2.5$. Two different transverse momentum thresholds are considered for the charged leptons from the W decay: $p^\ell_T > 25$ GeV and $p^\ell_T > 35$ GeV. Differential measurements are performed as a function of the lepton pseudorapidity η^ℓ, whereas total cross sections and average ratios require $|\eta^\ell| < 2.1$.

5 Event Selection

Selection of W candidates closely follows the criteria used in the $W \to \mu\nu$ and $W \to e\nu$ inclusive analysis [26]. The leptonic decay of a W boson into a muon or an electron and a neutrino is characterized by the presence of a high momentum isolated lepton. The neutrino escapes detection causing an apparent imbalance in the transverse energy of the event. Experimentally, the negative value of the modulus of the vector sum of all the transverse energy and momentum of the particles reconstructed by the Particle Flow algorithm defines the missing transverse energy of an event, $E_{\mathrm{T}}^{\mathrm{miss}}$. In W boson events, this variable is an estimator of the transverse energy of the undetected neutrino.

Muon tracks are required to have a transverse momentum $p_{\mathrm{T}}^{\mu} > 25\,\mathrm{GeV}$ and be measured in the pseudorapidity range $|\eta^{\mu}| < 2.1$. A muon isolation variable, I^{rel}, is built from the sum of transverse energies or momenta of reconstructed particles in a $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.4$ cone around the direction of the muon, and normalized to the muon momentum. The muon is required to be isolated from any other detector activity according to the criterion $I^{\mathrm{rel}} < 0.12$.

Electrons candidates with $p_{\mathrm{T}}^{\mathrm{e}} > 35\,\mathrm{GeV}$ measured in the pseudorapidity range $|\eta^{\mathrm{e}}| < 2.1$, excluding the region $1.44 < |\eta| < 1.57$ where services and cables exit are considered for the analysis. The isolation variable is defined from the sum of the transverse component of ECAL and HCAL energy deposits and transverse momenta of tracks reconstructed in the inner tracker, in a $\Delta R < 0.3$ cone around the electron direction, $I_{\mathrm{comb}} = \sum(E_T(\mathrm{ECAL}) + E_T(\mathrm{HCAL}) + p_{\mathrm{T}}(\mathrm{tracks}))$ and normalized to the electron momentum. An isolated electron must satisfy $I_{\mathrm{comb}}/p_{\mathrm{T}}^{\mathrm{e}} < 0.05$.

Contamination arising from Drell-Yan processes is reduced by removing events containing two or more muons (electrons) with $p_{\mathrm{T}}^{\ell} > 25(20)\,\mathrm{GeV}$ in the pseudorapidity region $|\eta^{\mu}| < 2.4$ ($|\eta^{\mathrm{e}}| < 2.5$). Finally, the reconstructed transverse mass, M_T, built from the transverse momentum of the isolated lepton, p_{T}^{ℓ}, and the missing transverse energy in the event, $E_{\mathrm{T}}^{\mathrm{miss}}$, $M_T \equiv \sqrt{2\,p_T^{\ell}\,E_{\mathrm{T}}^{\mathrm{miss}}\,(1-\cos(\phi_{\ell}-\phi_{E_{\mathrm{T}}^{\mathrm{miss}}}))}$, must be large. In the muon channel the M_T variable has to be larger than 40 GeV. A higher threshold is set in the electron channel, $M_T > 55\,\mathrm{GeV}$, as a condition on this variable ($M_T > 50\,\mathrm{GeV}$) is already included in the online selection. This requirement reduces the QCD multijet background to a negligible level in the muon channel. Remnant QCD background in the electron channel will be subtracted from the experimental $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution.

A W + jets sample is selected by demanding the presence of at least one jet with $p_T^{jet} > 25\,\mathrm{GeV}$ in the pseudorapidity range $|\eta^{jet}| < 2.5$, thus ensuring that the jet is contained in the tracker volume, and hence, achieving an optimal jet resolution. Jets are reconstructed from the Particle Flow candidates collection using an anti-k_T clustering algorithm [27] with a size parameter of $R = 0.5$. Charged particles with tracks not originated at the primary vertex are not considered for the jet clustering. Jets within a cone of radius $\Delta R < 0.3$ around a lepton candidate are disregarded for the jet counting. Energy not related with the hard interaction is subtracted from the jet energy [28], [29]. Finally, energy corrections are applied to flatten the jet response as a function of η and p_{T} [30].

A W + c candidate sample is further selected searching for a distinct signature of a charmed particle decay among the constituents of the leading jet associated to the W boson, as introduced in section 3. For that purpose, events with a secondary vertex arising from the decay of a relatively long-lived quark are kept. Secondary vertices are reconstructed using an adaptive vertex finder [31] algorithm with well understood performance [32]. This algorithm is rather

stable with respect to alignment uncertainties and is an essential component of the vertex-based b-tagging algorithms used in the CMS experiment. In its default implementation, also used in this analysis, tracks within a $\Delta R < 0.3$ cone around the jet axis, with a transverse momentum larger than 1 GeV and with a probability of consistency with the primary vertex below 50%, are considered to build a secondary vertex. Finally, only secondary vertices with a transverse decay length significance larger than 3 with respect to the primary vertex position are kept.

A search for $D^{\pm}$ and D^0 exclusive and semiinclusive charm meson decays is carried out in those events with secondary vertices with three and two charged prongs, respectively. In addition, a $W + c$ candidate sample with the charm quark decaying semileptonically is selected from the events with an identified muon among the particles constituting the jet. The several samples selected are described in the following in more detail.

5.1 Selection of exclusive $D^{\pm}$ decays

We identify $D^{\pm} \to K^{\mp}\pi^{\pm}\pi^{\pm}$ decays in the selected W+jets sample using secondary vertices with three tracks and a reconstructed invariant mass consistent with the $D^{\pm}$ mass: 1869.5 ± 0.4 MeV [33]. The Kaon mass is assigned to the track that has opposite sign to the total charge of the three-prong vertex, the remaining tracks are assumed to have the mass of a charged pion. This assignment is correct in more than 99% of the cases, since the fraction of double Cabibbo-suppressed decays is extremely small: $\mathcal{B}(D^+ \to K^+\pi^+\pi^-)/\mathcal{B}(D^+ \to K^-\pi^+\pi^+) = 0.00577 \pm 0.00022$ [33].

The branching fraction of the decay chain $\mathcal{B}(c \to D^{\pm} \to K^{\mp}\pi^{\pm}\pi^{\pm})$ in the PYTHIA simulation ($(1.528 \pm 0.008)\%$) is found to be about 40% smaller than the experimental measurement, $\mathcal{B}(c \to D^{\pm} \to K^{\mp}\pi^{\pm}\pi^{\pm}) = (2.08 \pm 0.10)\%$, as extracted from our combination of published measurements performed at LEP [34–36] of this branching fraction times $R_c = \Gamma(Z \to c\bar{c})/\Gamma(Z \to \text{had})$, the partial decay width of the Z into charm quark pairs. The measurements are rescaled to take into account the latest experimental value from the PDG [33] of $R_c = 0.1721 \pm 0.0030$. Experimental systematic uncertainties are assumed to be uncorrelated between the different measurements, given the substantially different sources of uncertainty considered in each case, whereas the experimental uncertainty on R_c is propagated in a correlated way. In all plots comparing data and MC the predicted rate is scaled in order to match the experimental central value.

Figure 2 shows the OS-SS subtracted distributions of the reconstructed invariant mass for $D^{\pm}$ candidates associated to $W \to \mu\nu$ and $W \to e\nu$ decays. They are almost exclusively populated by $W + c$ events. Data are fitted to the sum of a Gaussian for the signal plus a second-order polynomial for the non-resonant component. The PDG value of the $D^{\pm}$ mass is reproduced rather precisely in all cases.

5.2 Selection of exclusive $D^{*\pm}(2010)$ decays

The first step in the identification of $D^{*+}(2010) \to D^0\pi^+$ ($D^{*-}(2010) \to \bar{D}^0\pi^-$) decays is the selection of a secondary vertex with two tracks of opposite charge, as expected from a $D^0 \to K^-\pi^+$ ($\bar{D}^0 \to K^+\pi^-$) decay. This two-track system is combined with a primary track with $p_T > 0.3$ GeV found in a cone of $\Delta R < 0.1$ around the direction of the D^0 candidate momentum. The secondary track with charge opposite to the charge of the primary track is assumed to be the Kaon in the D^0 decay. Only combinations with a reconstructed mass differing from the D^0 mass (1864.86 ± 0.13 MeV [33]) by less than 70 MeV are kept. The $D^{*\pm}(2010)$ signal is identified as a clean peak around the expected value of the mass between the $D^{*\pm}(2010)$ and D^0 reconstructed masses, $m^{rec}(D^{*\pm}(2010)) - m^{rec}(D^0) = 145.421 \pm 0.010$ MeV [33].

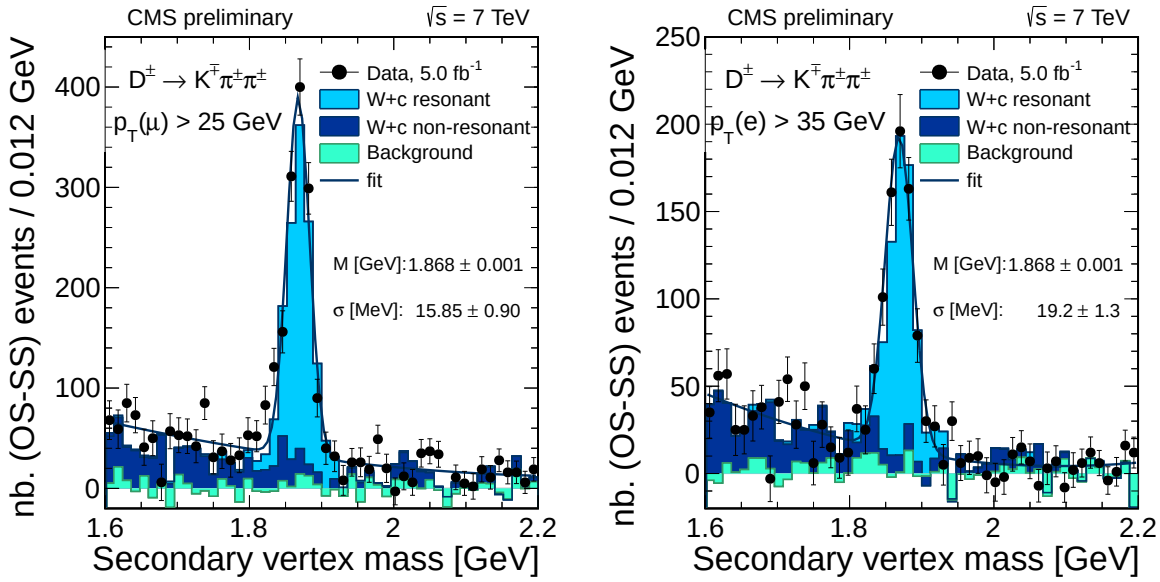

Figure 2: Fit to the invariant mass distribution of three-prong secondary vertices in data, after subtraction of the SS component. The position and width of the resonance peak are in reasonable agreement with the MC expectations (only statistical uncertainties are quoted). The channels shown correspond to muonic and electronic decays of the W with cuts of $p_{\mathrm{T}}^{\mu} > 25$ GeV (left) and $p_{\mathrm{T}}^{\mathrm{e}} > 35$ GeV cuts (right). Note that the amount of non-W + c background predicted by the simulation is almost negligible.

The product of the branching fractions $\mathcal{B}(\mathrm{c} \to \mathrm{D}^{*+}(2010)) \times \mathcal{B}(\mathrm{D}^{*+}(2010) \to \mathrm{D}^0\pi^+) \times \mathcal{B}(\mathrm{D}^0 \to \mathrm{K}^-\pi^+)$ in the PYTHIA simulation is $(0.743 \pm 0.005)\%$, which is about 20% larger than our estimation of the experimental value, $(0.622 \pm 0.020)\%$. The latter number is a weighted average that uses as inputs the dedicated measurements of this product by OPAL [37] and ALEPH [35], as well as the measurement of $\mathcal{B}(\mathrm{c} \to \mathrm{D}^{*+}(2010)) \times \mathcal{B}(\mathrm{D}^{*+}(2010) \to \mathrm{D}^0\pi^+)$ by DELPHI [38]. The measurements are rescaled to take into account the latest $\mathcal{B}(\mathrm{D}^0 \to \mathrm{K}^-\pi^+) = 0.0388 \pm 0.0005$ and R_{c} experimental values from the PDG [33]. Also in this case, experimental systematic uncertainties are assumed to be uncorrelated between the different measurements, and the experimental uncertainty on R_{c} is propagated in a correlated way. The MC predictions for the full $\mathrm{D}^{*\pm}(2010)$ decay chain are scaled in order to match the correct experimental values.

Figure 3 shows the OS-SS subtracted distribution of the reconstructed mass difference $m(\mathrm{D}^{*\pm}(2010)) - m(\mathrm{D}^0)$, where MC predictions are normalized to the recorded luminosity in data. $W \to \mu\nu$ and $W \to \mathrm{e}\nu$ decays are considered, with transverse momentum cuts of $p_{\mathrm{T}}^{\mu} > 25$ GeV and $p_{\mathrm{T}}^{\mathrm{e}} > 35$ GeV. Note that the amount of background predicted by the simulation and also observed in data is extremely small.

5.3 Selection of semileptonic charm decays

In addition to the previous exclusive channels, we also consider the identification of charm jets via semileptonic decays of the c quark. Only semileptonic decays into muons are considered. Muons in jets are identified with the same criteria used for muon identification in W decays, with the exception of the isolation requirements. Since the OS-SS subtraction strategy effectively suppresses all backgrounds but Drell-Yan, additional cuts are applied in order to reduce the Drell-Yan contamination to manageable levels without affecting the signal in an appreciable way. We require: $p_{\mathrm{T}}^{\mu} < 25$ GeV, $p_{\mathrm{T}}^{\mu}/p_{\mathrm{T}}^{jet} < 0.6$ and $p_{\mathrm{T}}^{rel} < 2.5$ GeV, where p_{T}^{μ} denotes here the

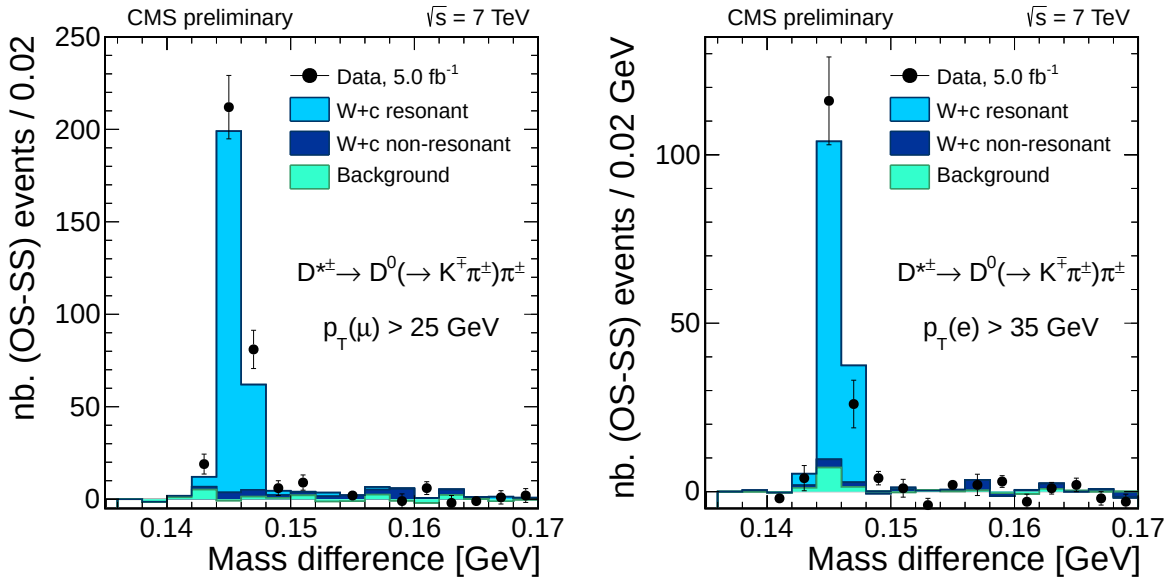

Figure 3: Distributions of the reconstructed mass difference between $D^{*\pm}(2010)$ and D^0 candidates in the selected $W + c$ sample in CMS, after subtraction of the SS component. The position and width of the peak, near 145 MeV is in agreement with the MC expectations. The channels shown correspond to muonic and electronic decays of the W with cuts of $p_T^{\mu} > 25$ GeV (left) and $p_T^{e} > 35$ GeV (right).

transverse momentum of the muon identified inside the jet and p_T^{rel} is its transverse momentum with respect to the jet direction. We also require the invariant mass of the dilepton system to be above 12 GeV, in order to avoid the region of low mass QCD resonances. Finally, dimuon events with a mass above 85 GeV are rejected. The latter cut is not applied to the sample with W boson decays into electrons, which is minimally affected by high-mass dilepton contamination.

Regarding the input semileptonic branching fraction of charm jets, we employ in this analysis the value $\mathcal{B}(c \to \ell) = 0.0911 \pm 0.0049$, which is the average of the inclusive value 0.096 ± 0.004 [33] and of the exclusive sum of the individual contributions from all weakly decaying charmed hadrons: 0.086 ± 0.004 [33, 39]. The uncertainty is increased in order to cover both central values within one standard deviation. This value is also rather consistent with the PYTHIA value present in our simulations (9.3%).

Figure 4 shows the resulting transverse momentum distribution of the selected muons inside the leading jet, OS-SS subtracted. Data are compared with the MC predictions normalized to the total integrated luminosity. Again both $W \to \mu\nu$ and $W \to e\nu$ decays are considered, with transverse momentum cuts of $p_T^{\mu} > 25$ GeV and $p_T^{e} > 35$ GeV for the leptons from the W decay. Note that the background predicted by the simulation is rather small in the electron channel but has a substantial Drell-Yan component in the muon channel. A small signal excess over the predictions (of about 10%) is visible. This excess is consistent with the observations in the $D^{\pm}$ and $D^{*\pm}(2010)$ channels.

5.4 Selection of semiinclusive $D^{\pm}$ and $D^{*\pm}(2010)$ decays

Enlarged samples of $W + c$ candidates are also selected through the identification of semiinclusive $D^{\pm}$ and $D^{*\pm}(2010)$ charm decays in order to increase the size of the samples available for the differential measurements.

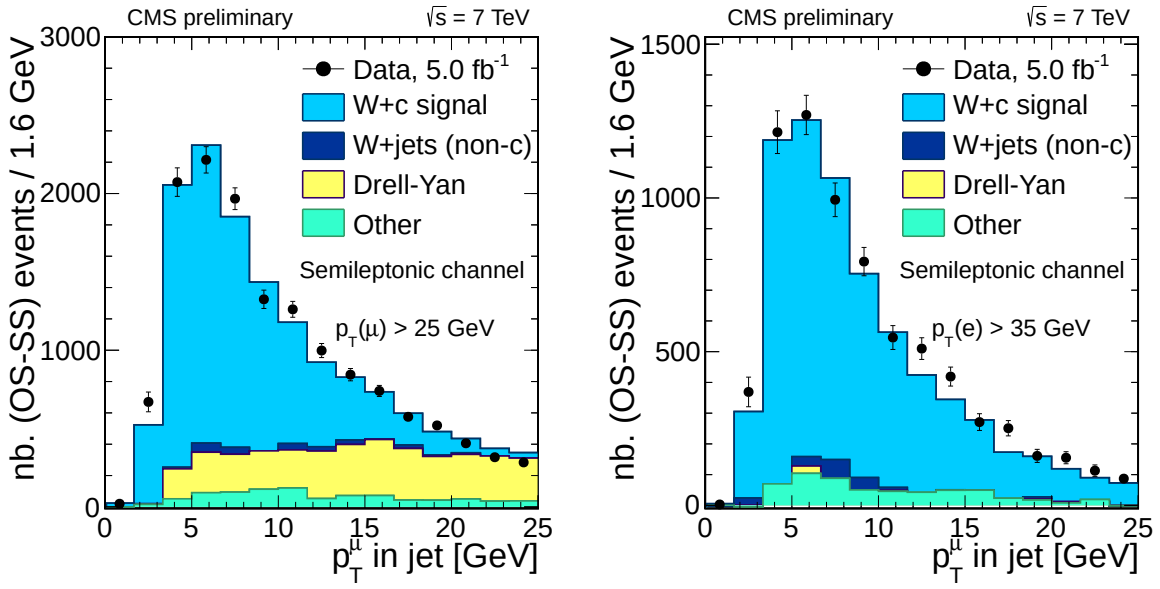

Figure 4: Distributions of the transverse momentum of the muon inside the leading jet of the event, after subtraction of the SS component. The channels shown correspond to muonic and electronic decays of the W with cuts of $p_{\mathrm{T}}^{\mu} > 25$ GeV (left) and $p_{\mathrm{T}}^{\mathrm{e}} > 35$ GeV (right).

Semiinclusive $\mathrm{D}^{\pm} \to \mathrm{K}^{\mp}\pi^{\pm}\pi^{\pm}$ charm decay candidates are selected among the events with a secondary vertex with three charged tracks and with a vertex charge, computed as the sum of the charges of the tracks constituting the vertex, equal to ± 1. Mass assignment for the secondary tracks follows what is described in section 5.1. The OS-SS subtracted distribution of the reconstructed invariant mass in events with three prongs is presented in Figure 5. In addition to the resonant peak at the $\mathrm{D}^{\pm}$ mass, there is a non-resonant spectrum with lower values of the invariant mass and corresponding mainly to $\mathrm{D}^{\pm}$ decays with one or several unaccounted neutral particles in the final state. For the differential cross section measurement we consider the region of the invariant mass spectrum $m(\mathrm{K}^{\mp}\pi^{\pm}\pi^{\pm}) < 2.5\,\mathrm{GeV}$.

Semiinclusive D^0 charm meson decays are reconstructed in the W+jets events with a displaced secondary vertex built from two charged tracks of opposite sign. The two tracks are assumed to correspond to the decay products of a D^0. The decay chain $\mathrm{D}^{*\pm}(2010) \to \mathrm{D}^0\pi^{\pm}$, $\mathrm{D}^0 \to \mathrm{K}^{\mp}\pi^{\pm}$ is searched for, according to the procedure described in section 5.2 but releasing the D^0 mass constraint $|m(\mathrm{K}^{\mp}\pi^{\pm}) - 1864.86| < 70\,\mathrm{MeV}$. Figure 6 shows the (OS-SS) subtracted distributions of the mass difference $m(\mathrm{K}^{\mp}\pi^{\pm}\pi^{\pm}) - m(\mathrm{K}^{\mp}\pi^{\pm})$. The peak at $m(\mathrm{K}^{\mp}\pi^{\pm}\pi^{\pm}) - m(\mathrm{K}^{\mp}\pi^{\pm}) \sim 145$ MeV corresponds to the nominal $\mathrm{D}^{*\pm}(2010) - \mathrm{D}^0$ mass difference. $W + \mathrm{c}$ events are still the dominant contribution at larger values of the mass difference. Remnant background is small and it is mainly due to residual W+light jets, $W + \mathrm{c\bar{c}}$ and $\mathrm{t\bar{t}}$ production. We select the events with an invariant mass difference $m(\mathrm{K}^{\mp}\pi^{\pm}\pi^{\pm}) - m(\mathrm{K}^{\mp}\pi^{\pm}) < 0.7\,\mathrm{GeV}$.

6 Measurement of the $W + \mathrm{c}$ cross section

The measurement of the total $W + \mathrm{c}$ cross section within the fiducial cuts discussed in the initial sections is performed on several different final states containing a well identified $W \to \ell\nu$ decay plus a leading jet with charm content. Only leptonic decays where ℓ is a muon or an electron are considered. The charm jet is identified in the three different modes presented in section 3 and described in section 5. We provide two sets of measurements: one with $p_{\mathrm{T}}^{\ell} > 25$ GeV using

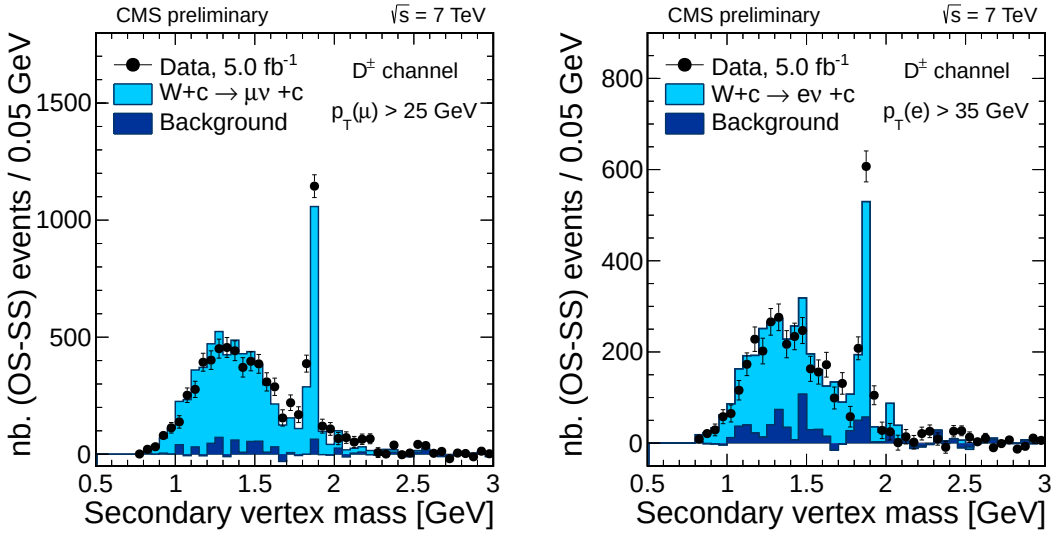

Figure 5: Invariant mass distribution of the three prongs composing a secondary vertex assuming a $D^{\pm} \to K^{\mp}\pi^{\pm}\pi^{\pm}$ hypothesis. The left plot is for $W \to \mu\nu$ events, with $p_T^{\mu} > 25\,\text{GeV}$. The right plot is for $W \to e\nu$ events, with $p_T^{e} > 35\,\text{GeV}$. Distributions are presented after subtraction of the SS component.

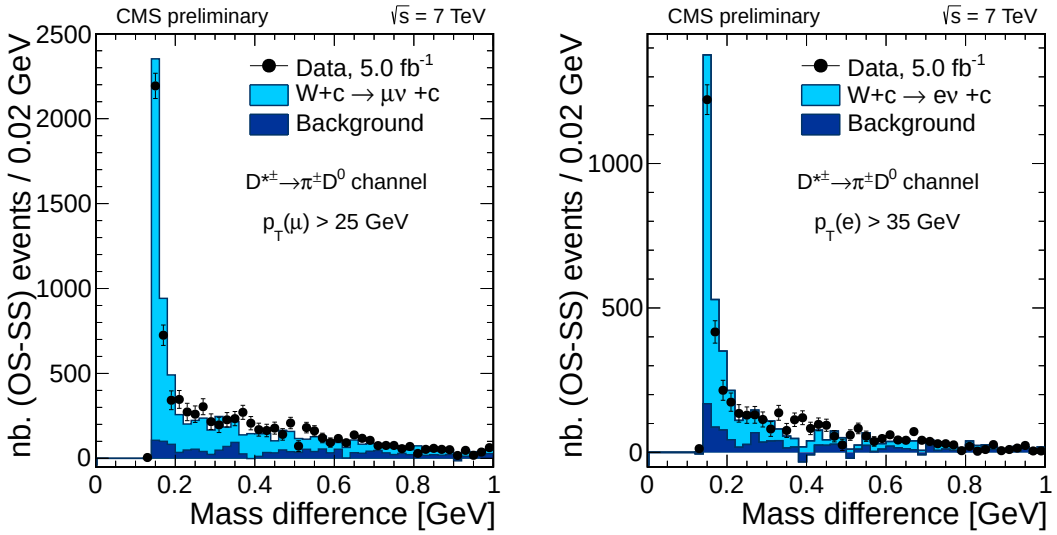

Figure 6: W+jets candidate events with a two-prong secondary vertex. Difference of the invariant masses of the two-prong system plus the closest track from the primary pp interaction vertex and the invariant mass of the two secondary vertex tracks ($m(K^{\mp}\pi^{\pm}\pi^{\pm}) - m(K^{\mp}\pi^{\pm})$), assuming the decay chain $D^{*\pm}(2010) \to D^0\pi^{\pm} \to K^{\mp}\pi^{\pm}\pi^{\pm}$ decay. The sharp peak at 145 MeV reflects the nominal $m_{D^{*\pm}(2010)} - m_{D^0}$ mass difference for the decay chain $D^{*\pm}(2010) \to D^0\pi^{\pm}$. The left plot is for $W \to \mu\nu$ events, with $p_T^{\mu} > 25\,\text{GeV}$. The right plot is for $W \to e\nu$ events, with $p_T^{e} > 35\,\text{GeV}$. Distributions are presented after subtraction of the SS component.

only $W \to \mu\nu$ decays, and a second one, using both $W \to \mu\nu$ and $W \to e\nu$ decays with $p_T^\ell > 35$ GeV.

6.1 Cross section measurements

For all channels under study the $W + c$ cross section within the reference fiducial cuts $p_T^\ell > 25(35)$ GeV, $|\eta^\ell| < 2.1$, $p_T^{jet} > 25$ GeV, $|\eta^{jet}| < 2.5$ is determined using the following Equation:

$$\sigma(W + c) = \frac{N_{sel} - N_{bkg}}{\mathcal{L}_{\rm int}\, \mathcal{B}\, \mathcal{A}\, \epsilon} \tag{1}$$

where N_{sel} is the number of selected signal events, N_{bkg} is the number of selected background events and $\mathcal{L}_{\rm int}$ is the integrated luminosity. $\mathcal{B}$ is the relevant charm branching fraction for the channel under study, i.e. $\mathcal{B} \equiv \mathcal{B}(c \to D^+ \to K^-\pi^+\pi^+) = (2.08 \pm 0.10)\%$ in the case of the $D^\pm$ channel, $\mathcal{B} \equiv \mathcal{B}(c \to D^{*+}(2010); D^{*+}(2010) \to D^0\pi^+; D^0 \to K^-\pi^+) = (0.622 \pm 0.020)\%$ for the $D^{*\pm}(2010)$ channel and $\mathcal{B} \equiv \mathcal{B}(c \to \ell) = (9.11 \pm 0.49)\%$ for the dilepton channel. The acceptance times efficiency factor $\mathcal{A}\,\epsilon$ depends on the specific decay channel considered. It is largely independent of the detector acceptance because it is determined for the reference fiducial cuts. On the other hand, it accounts for unavoidable losses due to selection criteria, which may be significant. In $W + c$ events, less than 20% of the events have a well identified secondary vertex, while less than 50% of the semimuonic charm decays lead to a muon with enough energy and quality to be identified in the muon spectrometer.

The signal enriched region for the $D^\pm$ channel is defined by the cut $|m_D^{rec} - 1.87\,\text{GeV}| < 0.05\,\text{GeV}$, where m_D^{rec} is the reconstructed mass of the $D^\pm$ candidate (Figure 2). The same cut is applied to the MC predictions in order to determine the acceptance times efficiency factor for the measurement, $\mathcal{A}\,\epsilon$. We estimate values of $\mathcal{A}\,\epsilon = 0.1114 \pm 0.0033$ ($p_T^\mu > 25$ GeV) and $\mathcal{A}\,\epsilon = 0.0834 \pm 0.0032$ ($p_T^e > 35$ GeV), where the quoted uncertainties are statistical only. The background is fully dominated by the non-resonant $W + c$ component. It is subtracted from the selected number of events in the data window by using the number of events selected in a control region away from the resonance, extending up to a window of 200 MeV width, $N(0.05\ \text{GeV} < |m_D^{rec} - 1.87\ \text{GeV}| < 0.20\ \text{GeV})$. This number is scaled by the ratio $N(|m_D^{rec} - 1.87\ \text{GeV}| < 0.05\ \text{GeV})/N(0.05\ \text{GeV} < |m_D^{rec} - 1.87\ \text{GeV}| < 0.20\ \text{GeV})$ observed in MC in order to obtain the number of background events expected in the reference window. The procedure is largely independent of uncertainties on the charm fractions present in PYTHIA. Systematic biases due to the assumed non-resonant background subtraction are expected to be negligible compared to the statistical uncertainty, given the approximate agreement between data and MC distributions.

The signal enriched region for the $D^{*\pm}(2010)$ channel is defined by the events selected in the region $|m^{rec}(D^{*\pm}(2010)) - m^{rec}(D^0) - 145\,\text{MeV}| < 5\,\text{MeV}$, where $m^{rec}(D^{*\pm}(2010)) - m^{rec}(D^0)$ is the reconstructed mass difference between the D mesons (Figure 3). The same procedure is applied to the MC predictions in order to determine the acceptance times efficiency factor for the measurement, $\mathcal{A}\,\epsilon$. We estimate values of $\mathcal{A}\,\epsilon = 0.0849 \pm 0.0040$ ($p_T^\mu > 25$ GeV) and $\mathcal{A}\,\epsilon = 0.0559 \pm 0.0036$ ($p_T^e > 35$ GeV), where the quoted uncertainties are statistical only. Like in the $D^\pm$ case, the background in data is subtracted from the selected number of data events in a sideband sample, defined in this case as: $5\ \text{MeV} < |m^{rec}(D^{*\pm}(2010)) - m^{rec}(D^0) - 145\ \text{MeV}| < 20\ \text{MeV}$. This number is scaled by the ratio $N(|m^{rec}(D^{*\pm}(2010)) - m^{rec}(D^0) - 145\ \text{MeV}| < 5\ \text{MeV})/N(5\ \text{MeV} < |m^{rec}(D^{*\pm}(2010)) - m^{rec}(D^0) - 145\ \text{MeV}| < 20\ \text{MeV})$ observed in MC.

For the semileptonic channels the $\sigma(W + c)$ cross section is simply obtained by counting of the total number of dilepton selected events (Figure 4). The acceptance times efficiency factors for the different lepton cuts are estimated in the simulation as: $\mathcal{A}\,\epsilon = 0.2035 \pm 0.0021$ ($p_T^\mu > 25$ GeV) and $\mathcal{A}\,\epsilon = 0.1706 \pm 0.0021$ ($p_T^e > 35$ GeV), where the quoted uncertainties are statistical only. In the dimuon channel, the accepted Drell-Yan background corresponds to a region of significant missing transverse energy, with one low-p_T lepton inside a jet. Potential discrepancies between data and MC are corrected by analyzing the Drell-Yan dominated control region with dimuon invariant mass above 85 GeV. A correction factor of 1.2 ± 0.1 is found to provide agreement between data and simulation in this region. The uncertainty, propagated as a systematic uncertainty, takes conservatively into account possible differences in the description of events below and around the Z pole.

We perform two combinations of cross section measurements. A first combination uses all channels associated to $W \to \mu\nu$ decays with a muon cut of $p_T^\mu > 25$ GeV. A second combination considers all channels associated to W decays into muons and electrons, for a lepton cut of $p_T^\ell > 35$ GeV. Tables 1 and 2 contain all the relevant inputs used in the measurements, as well as the total cross sections obtained for each final state and their average, including systematic uncertainties. In all presented averages, correlations are taken into account. Full correlation is taken between channels sharing a given source of systematic uncertainty.

	$W \to \mu\nu$, $p_T^\mu > 25\,\text{GeV}$			
Final state	$(S+B)_{\text{data}}$	S_{data}	$\mathcal{A}\,\epsilon$ [%]	$\sigma(W + c)$ [pb]
$D^\pm$	1502 ± 62	1203 ± 91	11.1 ± 0.3	103.6 ± 7.8 (stat.) $\pm$ 8.1 (syst.)
$D^{*\pm}(2010)$	318 ± 21	309 ± 23	8.5 ± 0.4	116.9 ± 8.7 (stat.) $\pm$ 10.0 (syst.)
$c \to \mu$	14215 ± 196	9867 ± 237	20.4 ± 0.2	106.5 ± 2.6 (stat.) $\pm$ 9.6 (syst.)
Average				107.7 ± 3.3 (stat.) $\pm$ 6.9 (syst.)

Table 1: Cross section results for a transverse momentum lepton cut of $p_T^\mu > 25$ GeV in the different final states considered. $(S+B)_{\text{data}}$ is the number of selected events in the signal region (around the resonance in the case of $D^\pm$ and $D^{*\pm}(2010)$ final states). S_{data} is the corrected number of signal events after background subtraction using the method described in the text. $\sigma(W + c)$ is the measured $W + c$ cross section after correction for the charm fractions as discussed in the text, namely: $\mathcal{B}(c \to D^\pm \to K^\mp\pi^\pm\pi^\pm) = (2.085 \pm 0.100)\%$, $\mathcal{B}(c \to D^{*\pm}(2010); D^{*\pm}(2010) \to D^0\pi^\pm; D^0 \to K^\mp\pi^\pm) = (0.622 \pm 0.020)\%$ and $\mathcal{B}(c \to \ell\nu) = (9.11 \pm 0.49)\%$. All uncertainties quoted in the Table are just statistical, except for the measured cross section.

Several sources of systematic uncertainties are considered. Their contribution to the total uncertainty is presented in Table 3. The limited precision on the branching fractions of the charm decays is one of the dominant sources of systematics. Tracking reconstruction inefficiencies are intrinsically small ($< 1\%$ [40]). Given the nature of the method used to build secondary vertices, tracks are assigned to either the primary vertex or the secondary vertex in a way that may be different in data and in MC. In order to estimate the size of a potential discrepancy, the set of secondary tracks is either increased by adding a nearby primary track or decreased by dropping one of the original secondary tracks. The observed differences at the resonant D^0 and $D^\pm$ peaks between data and simulation suggest a possible maximum bias of 3.3%. This uncertainty is assigned to the $D^\pm$ and $D^{*\pm}(2010)$ channels and assumed to be fully correlated between them. Pileup uncertainties are accounted for by increasing the assumed minimum bias cross section by its estimated uncertainty ($\approx 6\%$). Jet enegy scale uncertainties are extracted from dedicated CMS studies [41], which also take into account possible variations in the jet

		$W \to \mu\nu$, $p_{\mathrm{T}}^{\mu} > 35\,\mathrm{GeV}$		
Final state	$(S+B)_{\mathrm{data}}$	S_{data}	$\mathcal{A}\,\epsilon$ [%]	$\sigma(W+\mathrm{c})$ [pb]
$\mathrm{D}^{\pm}$	1209 ± 55	981 ± 79	11.4 ± 0.4	$82.9 \pm 6.7\,(\mathrm{stat.}) \pm 6.4\,(\mathrm{syst.})$
$\mathrm{D}^{*\pm}(2010)$	260 ± 19	248 ± 21	8.6 ± 0.5	$92.3 \pm 7.8\,(\mathrm{stat.}) \pm 8.2\,(\mathrm{syst.})$
$\mathrm{c} \to \mu$	11462 ± 172	7875 ± 207	21.6 ± 0.2	$79.9 \pm 2.1\,(\mathrm{stat.}) \pm 6.9\,(\mathrm{syst.})$
Average (μ)				$82.9 \pm 2.6\,(\mathrm{stat.}) \pm 5.1\,(\mathrm{syst.})$
		$W \to \mathrm{e}\nu$, $p_{\mathrm{T}}^{\mathrm{e}} > 35\,\mathrm{GeV}$		
$\mathrm{D}^{\pm}$	838 ± 47	726 ± 55	8.3 ± 0.3	$83.5 \pm 6.3\,(\mathrm{stat.}) \pm 7.1\,(\mathrm{syst.})$
$\mathrm{D}^{*\pm}(2010)$	148 ± 15	145 ± 18	5.6 ± 0.4	$83.3 \pm 10.4\,(\mathrm{stat.}) \pm 8.5\,(\mathrm{syst.})$
$\mathrm{c} \to \mu$	7156 ± 151	6701 ± 175	17.1 ± 0.2	$86.5 \pm 2.2\,(\mathrm{stat.}) \pm 6.9\,(\mathrm{syst.})$
Average (e)				$85.3 \pm 2.5\,(\mathrm{stat.}) \pm 5.7\,(\mathrm{syst.})$
Final Average				$84.1 \pm 2.0\,(\mathrm{stat.}) \pm 4.9\,(\mathrm{syst.})$

Table 2: Cross section results for a transverse momentum lepton cut of $p_{\mathrm{T}}^{\ell} > 35$ GeV in the different final states considered. $(S+B)_{\mathrm{data}}$ is the number of selected events in the signal region (around the resonance in the case of $\mathrm{D}^{\pm}$ and $\mathrm{D}^{*\pm}(2010)$ final states). S_{data} is the corrected number of signal events after background subtraction using the method described in the text. $\sigma(W+\mathrm{c})$ is the measured $W+\mathrm{c}$ cross section after correction for the charm fractions as discussed in the text, namely: $\mathcal{B}(\mathrm{c} \to \mathrm{D}^{\pm} \to \mathrm{K}^{\mp}\pi^{\pm}\pi^{\pm}) = (2.085 \pm 0.100)\%$, $\mathcal{B}(\mathrm{c} \to \mathrm{D}^{*\pm}(2010); \mathrm{D}^{*\pm}(2010) \to \mathrm{D}^{0}\pi^{\pm}; \mathrm{D}^{0} \to \mathrm{K}^{\mp}\pi^{\pm}) = (0.622 \pm 0.020)\%$ and $\mathcal{B}(\mathrm{c} \to \ell\nu) = (9.11 \pm 0.49)\%$. All uncertainties quoted in the Table are just statistical, except for the measured cross section.

flavor composition. Additional $\not{E}_{\mathrm{T}}$ effects ($\approx 2\%$) are estimated by smearing the M_T Monte Carlo distribution in order to match the M_T shape observed in data. Muon trigger and selection inefficiencies are included in the simulation by applying the data/MC factors determined in dedicated tag-and-probe studies as a function of the muon pseudorapidity. For muons we estimate a 0.7% uncertainty according to CMS studies on dimuon events at the Z pole. The effect of momentum resolution corrections determined at the Z pole is also propagated as an additional uncertainty. In the electron case we consider the difference between switching on and off the efficiency corrections, due to the presence of missing transverse requirements at the trigger level that may affect differently the W + jet sample selected in this analysis. The lepton efficiency uncertainty for leptons inside jets is taken to be 3.0% according to dedicated studies in multijet events. The ISR related uncertainties are estimating by recalculating the acceptance on samples generated with a different scale ($Q = 0.5, 2$ times the reference scale used in the generation). The average value of the meson energy fraction in charm decays is varied by 4%, which is about twice the uncertainty on the $\mathrm{D}^{*\pm}(2010)$ fragmentation determined at LEP, in order to cover possible uncertainties on the assumed shape. Other theory uncertainties on $\mathcal{A}\,\epsilon$ include remnant PDF effects and potential biases due to the adoption of the MadGraph jet-parton matching scheme as the reference to be compared with MCFM calculations ($\approx 1\%$). The integrated luminosity measurement has a 2.2% uncertainty [42]. Physics backgrounds, including the gluon-splitting $W + \mathrm{c\bar{c}}$ component, have a negligible contribution to the systematics compared with the statistical uncertainties in the background subtraction.

For $p_{\mathrm{T}}^{\mu} > 25$ GeV the average $W+\mathrm{c}$ cross section is measured to be:

$$\sigma(\mathrm{pp} \to \mathrm{W} + \mathrm{c} + \mathrm{X}) \times \mathcal{B}(\mathrm{W} \to \mu\nu, p_{\mathrm{T}}^{\mu} > 25\ \mathrm{GeV}) = 107.7 \pm 3.3\,(\mathrm{stat.}) \pm 6.9\,(\mathrm{syst.})\ \mathrm{pb} \qquad (2)$$

For $p_{\mathrm{T}}^{\ell} > 35$ GeV we obtain:

$$\sigma(\mathrm{pp} \to \mathrm{W} + \mathrm{c} + \mathrm{X}) \times \mathcal{B}(\mathrm{W} \to \ell\nu, p_{\mathrm{T}}^{\ell} > 35 \text{ GeV}) = 84.1 \pm 2.0 \text{ (stat.)} \pm 4.9 \text{ (syst.) pb} \quad (3)$$

These measurements are largely background-free. The overall relative uncertainty, $6 - 7\%$, is dominated by systematic uncertainties on the theoretical modeling of the signal and by experimental uncertainties on the acceptance of the selection criteria. A detailed comparison with theoretical predictions is provided in section 9.

Source	$p_{\mathrm{T}}^{\mu} > 25$ GeV $\Delta_{\mathrm{syst}}[\%]$	$p_{\mathrm{T}}^{\ell} > 35$ GeV $\Delta_{\mathrm{syst}}[\%]$
MC statistics	1.6	1.3
Lepton efficiency, resolution	0.8	1.5
Muon efficiency in charm decay	1.4	1.5
Vertex reconstruction	1.8	1.7
Pileup	0.9	0.8
Jet energy scale	3.0	1.7
$\not{E}_{\mathrm{T}}$	2.0	2.0
$\mathcal{B}(\mathrm{c} \to \mathrm{D}^{\pm} \to \mathrm{K}^{\mp}\pi^{\pm}\pi^{\pm})$	1.5	1.5
$\mathcal{B}(\mathrm{c} \to \mathrm{D}^{*\pm}(2010) \to \mathrm{D}^{0} \to \mathrm{K}^{\mp}\pi^{\pm})$	0.7	0.6
$\mathcal{B}(\mathrm{c} \to \ell)$	2.6	2.7
ISR and Q^2-matching	0.2	0.2
Fragmentation function	0.8	0.6
Other theory uncertainties on $\mathcal{A}\,\epsilon$	0.8	0.7
DY background	1.4	0.9
Luminosity	2.2	2.2
Total	6.3	5.7

Table 3: Systematic contributions (Δ_{syst}) to the combined $\sigma(W + \mathrm{c})$ measurements in the reference fiducial volume given by $p_{\mathrm{T}}^{jet} > 25$ GeV, $|\eta^{jet}| < 2.5$, $|\eta^{\ell}| < 2.1$ and two different cuts on the transverse momentum of the lepton in the W decay: $p_{\mathrm{T}}^{\ell} > 25$ GeV (muon channel only) and $p_{\mathrm{T}}^{\ell} > 35$ GeV (muon and electron channels combined).

The high signal purity of the selected samples allows a deeper study of the properties of $W + \mathrm{c}$ events. Figure 7 shows the distributions of the jet pseudorapidity and the jet momentum fraction carried out by the D meson or the muon for $\mathrm{D}^{\pm}$, $\mathrm{D}^{*\pm}(2010)$ and semileptonic candidates in the enriched signal region. The latter observable is directly related with the charm fragmentation function. The normalization of the $W + \mathrm{c}$ component in the Monte Carlo has been scaled by 1.1 in order to approximately match the experimental rate measured in data. Electron and muon channels are added in order to enhance the significance of the comparison. All distributions show reasonable agreement with the predictions of MADGRAPH + PYTHIA.

7 Measurement of the normalized differential cross section as a function of the lepton pseudorapidity

The $W + \mathrm{c}$ cross section is also measured differentially as a function of the pseudorapidity of the lepton from the W decay. We will first determine the normalized differential cross section,

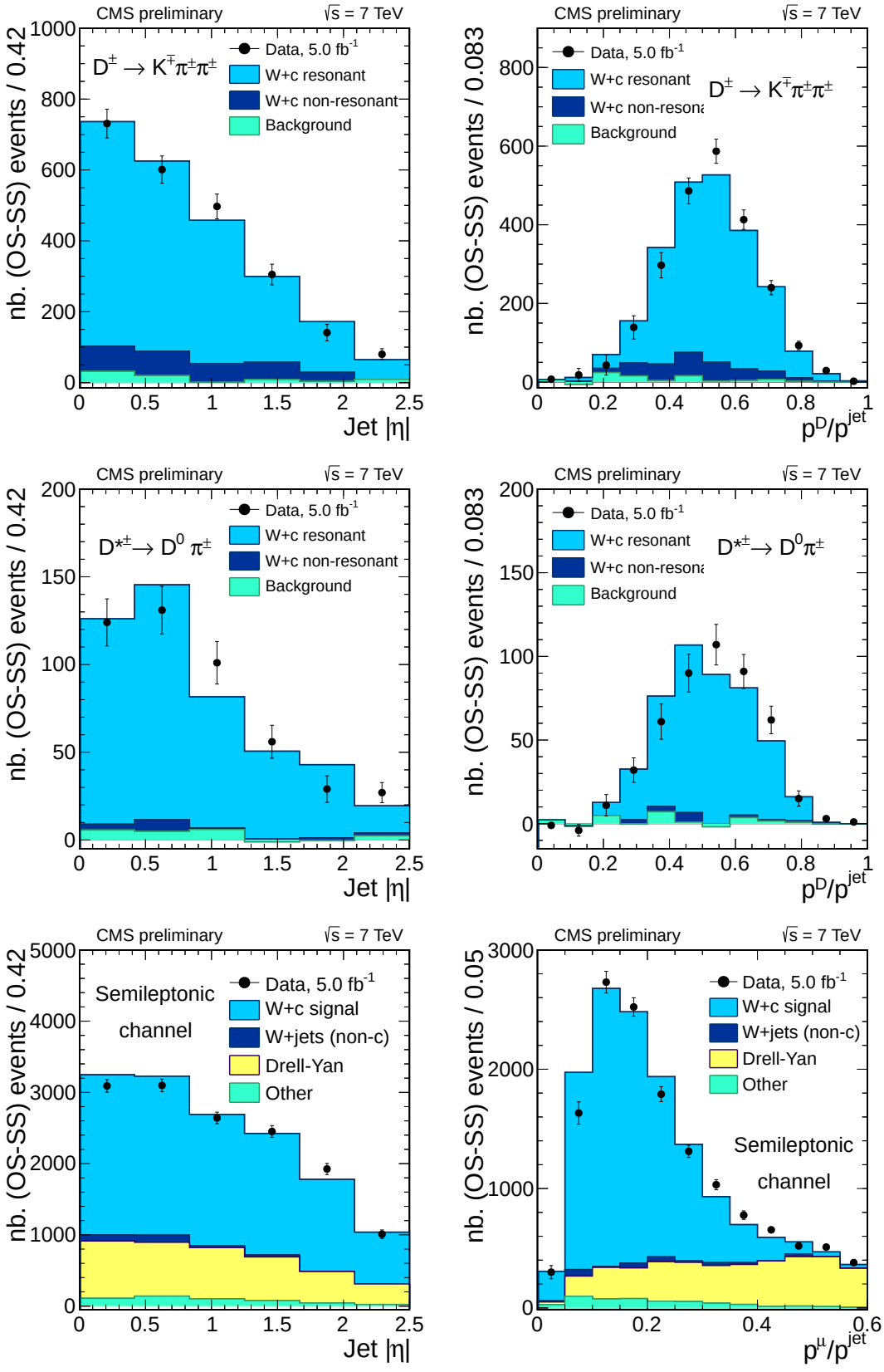

Figure 7: Distributions of $W + c$ selected events in the different charm decay channels as a function of the jet pseudorapidity (left) and the jet momentum fraction carried out by the D meson/muon (right). Top histograms correspond to the $D^{\pm}$ decay channel, middle histograms correspond to the $D^{*\pm}(2010)$ decay channel and bottom histograms to semileptonic charm decays into muons. Only events in the enriched signal region used to determine the cross section are used. The Monte Carlo predictions have been scaled by a factor of 1.1 in order to approximately match the $W + c$ yield measured in data.

$\frac{1}{\sigma(W+c)}\frac{d\sigma(W+c)}{d\eta}$, as the fraction of $W+c$ events having the lepton from the W emitted in a given pseudorapidity interval and divided by the size of the interval:

$$\frac{1}{\sigma(W+c)}\frac{d\sigma(W+c)}{d\eta} = \frac{[N_{OS,i} - N_{SS,i} - (N^{bck}_{OS,i} - N^{bck}_{SS,i})]/(\mathcal{A}\,\epsilon)_i}{\sum_{i=1}^{i=5}[N_{OS,i} - N_{SS,i} - (N^{bck}_{OS,i} - N^{bck}_{SS,i})]/(\mathcal{A}\,\epsilon)_i} \times \frac{1.}{\Delta\eta_i} \quad (4)$$

where $N_{OS,i} - N_{SS,i}$ is the difference between Opposite Sign and Same Sign events with the lepton in the i^{th} pseudorapidity bin.

For this measurement, identification of $W+c$ final states with a defined charm content follows the semiinclusive charm meson decays selection described in section 5.4. In addition to them, the sample of $W+c$ candidates with an identified muon inside the jet is also employed for the differential measurement. Residual background after the SS subtraction ($N^{bck}_{OS,i} - N^{bck}_{SS,i}$) is evaluated with the simulated samples and subtracted from the OS-SS data sample.

Relative normalization among the different η_ℓ bins is accounted through acceptance times efficiency correcting factors, $(\mathcal{A}\,\epsilon)_i$, referring the phase space defined as $W+c$ events with $p_T^{jet} > 25$ GeV and $|\eta^{jet}| < 2.5$, with the lepton from the W decay with a minimum transverse momentum defined in the analysis and pseudorapidity in a particular η^ℓ interval. The resulting distribution is then normalized to the total cross section measured. Five bins in the absolute value of the lepton pseudorapidity are considered: [(0, 0.35), (0.35, 0.7), (0.7, 1.1), (1.1, 1.6), (1.6, 2.1)]; this binning is chosen in order to have a uniform share-out of the events among the five bins.

Normalized differential cross sections are evaluated following expression 4 from the number of SS subtracted events in the five pseudorapidity bins, together with the expected backgrounds and the $\mathcal{A}\,\epsilon$ correcting factors. The normalized differential cross section distributions are presented in Figure 8.

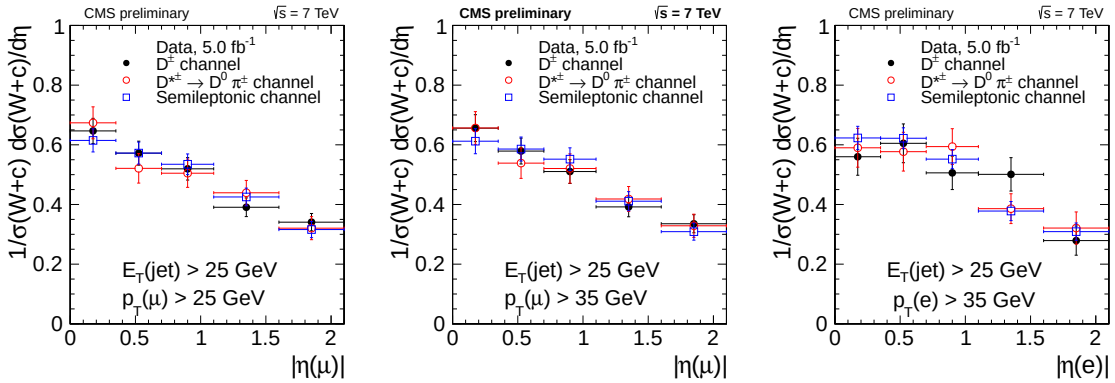

Figure 8: Normalized differential cross section distribution of $W+c$ ($W \to \ell\nu$) events as a function of the pseudorapidity of the lepton from the W decay. The two first plots show the results from the $W \to \mu\nu$ sample, with $p_T^\mu > 25\,\text{GeV}$ (left plot) and $p_T^\mu > 35\,\text{GeV}$ (central plot). The right plot shows the results from the $W \to e\nu$ sample, with $p_T^e > 35\,\text{GeV}$. The results obtained with the sample of events with a three-prong secondary vertex are shown as black solid points. Open red points represent the results obtained with the sample of events having a two-prong secondary vertex and the blue open squares give the result from the sample of events with an identified muon inside the jet.

There is a good agreement among the normalized differential cross section measured with the different $W+c$ subsamples (three-prong secondary vertices, two-prong secondary vertices and

muon-in-a-jet samples) and between the two $W \to \ell\nu$ decay channels considered. In order to compare the experimental results with the theoretical predictions, results obtained in the $W \to \mu\nu$ channel with $p_{\mathrm{T}}^{\mu} > 25\,\mathrm{GeV}$ are averaged. Results from the $W \to \mu\nu$ sample and with $p_{\mathrm{T}}^{\mu} > 35\,\mathrm{GeV}$ and in the $W \to \mathrm{e}\nu$ channel ($p_{\mathrm{T}}^{\mathrm{e}} > 35\,\mathrm{GeV}$) are combined together taking into account both the statistical and the systematic uncertainties affecting the different channels. The normalized differential cross sections obtained for $p_{\mathrm{T}}^{\mu} > 25\,\mathrm{GeV}$ and $p_{\mathrm{T}}^{l} > 35\,\mathrm{GeV}$ are combined with the total $W + \mathrm{c}$ cross section presented in section 6 to give an absolute measurement of the differential cross-section ($\frac{d\sigma(W+\mathrm{c})}{d\eta}$) in each pseudorapidity bin.

7.1 Systematic uncertainties in the normalized differential cross section

Most of the sources that have been discussed in section 6 have a negligible impact in the differential distributions as their effect largely cancel out in the ratios.

The main source of systematic uncertainties in the normalized differential distributions are those related with lepton reconstruction and identification. Differential distributions are largely independent of jet energy scale effects as they are measured as a function of the pseudorapidity of the lepton from the W decay and the jet kinematic region spanned is in all cases the same, independently of the pseudorapidity of the lepton. Possible effects due to jet energy scale uncertainties are evaluated changing the jet energy scale in the simulated $W + \mathrm{c}$ sample according to the results of CMS dedicated studies. Observed variations in the resulting differential distribution can be largely accounted by mere statistical fluctuations in the MC sample.

Calibration factors for lepton momentum scale and resolution have been derived in CMS from detailed studies of the position and width of the Z boson peak. Systematic uncertainty in the normalized differential cross section is estimated in the $W \to \mathrm{e}\nu$ channel comparing the resulting distributions applying or not calibration corrections. Variations are smaller than 1% in the barrel, and of the order of 1.5% in the endcap region. For the $W \to \mu\nu$ channel the measurement is repeated one thousand times varying in each of them the muon calibration factors within their uncertainties and comparing with the values obtained applying the central value of the correcting factors. The width of the resulting distribution is taken as the systematic uncertainty due to the knowledge of the muon momentum scale and resolution. Uncertainties between from 0.2% and 0.4% in the normalized differential distributions were obtained, depending on the particular muon pseudorapidity bin considered, the sample selection and the p_{T}^{μ} threshold applied.

Correction factors to the lepton reconstruction efficiencies predicted by the simulated samples are determined in CMS with clean samples of Z ($\to \ell^{+}\ell^{-}$) events. An invariant mass ($m_{\ell^{+}\ell^{-}}$) constraint and tight quality requirements set to one of the leptons allows the other lepton to be used as a probe to test the different steps in lepton reconstruction and identification. Efficiencies are calculated for data and simulated samples and correcting factors $\mathrm{efficiency}_{\mathrm{data}}/\mathrm{efficiency}_{\mathrm{MC}}$ are computed. Available samples to cover certain muon pseudorapidity regions have a smaller statistics. We take half of the total systematic uncertainty in the efficiency correcting factors, $\sim 0.35\%$) as uncorrelated between the different pseudorapidity bins. For the $W \to \mathrm{e}\nu$ channel, the effect of the efficiency corrections in the measured ratios ($\sim 0.25\%$) is computed and taken as an estimation of the systematic uncertainty.

The only physical process with a significant contribution to the final selected sample is Drell-Yan production affecting the candidate $W \to \mu\nu$ with a second muon identified inside the jet. The correction factor (1.2 ± 0.1) applied to DY prediction is varied by one sigma and the differential distribution is reevaluated. Variations smaller than 0.3% are observed and taken as the associated systematic uncertainty. Top contributions have also been varied by 6% for $\mathrm{t\bar{t}}$

production and by a 15% for single top production. Variations in the differential distributions are smaller 0.2%. A total systematic uncertainty due to the background subtraction of 0.3% is taken.

Concerning uncertainties arising from an imperfect knowledge of the PDFs, it is observed that the uncertainties related with strange PDFs within the same PDF set are smaller or equal than the differences between the central values obtained with MSTW08 [22], CT10 [18] and NNPDF23 [43]. No variation in the $\mathcal{A}\,\epsilon$ correction factors computed with these sets of PDF is observed and thus, no change is expected in the final result.

Systematic uncertainties arising from other sources, such as the knowledge of the Pileup profile or the average energy fraction in charm fragmentation have been evaluated in the $W + c$ MC sample and found to be negligible.

8 Measurement of the cross sections ratio $\sigma(W^+ + \bar{c})/\sigma(W^- + c)$

Total and differential cross section ratios of W^+ to W^- bosons associated production together with a charm jet ($\sigma(W^+ + \bar{c})/\sigma(W^- + c)$) are also measured. They are determined as the ratios of the differences between the number of OS and SS events in which the lepton from the W decay is positively charged and the OS-SS difference in which the lepton from the W decay is negatively charged:

$$R_c^{\pm} = \frac{\sigma(W^+ + \bar{c})}{\sigma(W^- + c)} = \frac{(N^+{}_{OS} - N^+{}_{SS})}{(N^-{}_{OS} - N^-{}_{SS})}$$

The differential measurement is done as a function of the pseudorapidity of the lepton from the W decay.

The number of $(N^+{}_{OS} - N^+{}_{SS})$ and $(N^-{}_{OS} - N^-{}_{SS})$ are extracted from the same subsamples used for the differential cross section measurement, presented in the previous section, further dividing them according to the sign of the lepton from the W decay. Background contributions to the W^+ and W^- samples are expected to be equal. Predictions from the simulated samples for the main sources of background (top and DY production) are very similar. Expectations for the residual contributions arising from $W +$ light and $W + c\bar{c}$ production in the positive and negative lepton samples are also compatible within their statistical uncertainties. In order to diminish statistical fluctuations in the measurement arising from the background estimation, no background subtraction is performed. To reduce the DY contribution to the $W \to \mu\nu$ candidate sample in the category with a muon inside the jet, its transverse momentum has to be smaller than 12 GeV. No $\mathcal{A}\,\epsilon$ correction is applied as they affect by the same amount to the positively and negatively charged samples and correction cancels in the ratio.

Figure 9 presents the cross section ratios $R_c^{\pm}(\eta_\ell)$ obtained from the three subsamples: events with a displaced secondary vertex with three tracks (semiinclusive $D^{\pm}$ states), events with a displaced secondary vertex with two tracks (semiinclusive $D^{*\pm}(2010)$ states) and events with an identified muon inside the jet corresponding to a semileptonic charm decay.

There is a good agreement among the cross section ratios measured with the different $W + c$ subsamples (three-prong secondary vertices, two-prong secondary vertices and muon-in-a-jet samples) and between the two $W \to \ell\nu$ decay channels considered. In order to compare the experimental results with the theoretical predictions, results obtained in the $W \to \mu\nu$ channel with $p_T^{\mu} > 25$ GeV are averaged. Results from the $W \to \mu\nu$ sample and with $p_T^{\mu} > 35$ GeV and in the $W \to e\nu$ channel ($p_T^{e} > 35$ GeV) are combined together taking into account both the statistical and the systematic uncertainties affecting the different channels.

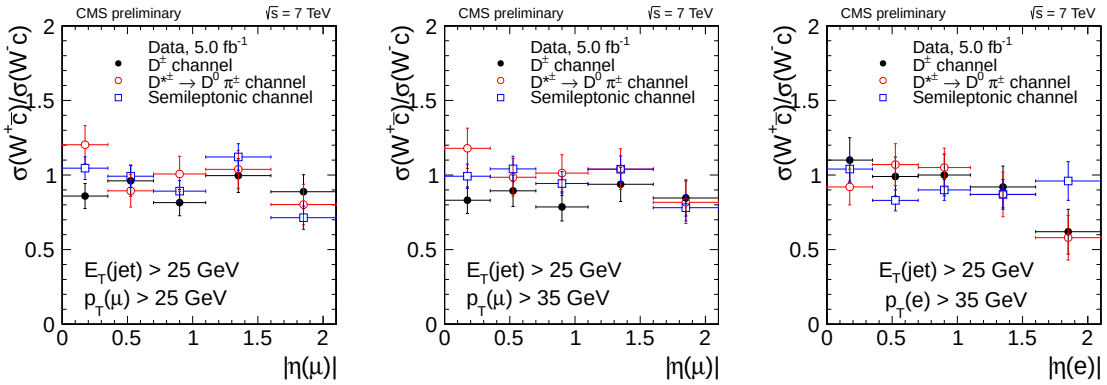

Figure 9: Differential measurement of $R_c^{\pm} = \sigma(W^+ + \bar{c})/\sigma(W^- + c)$ as a function of the lepton pseudorapidity from the W decay. The two first plots show the results from the $W \to \mu\nu$ sample, with $p_T^{\mu} > 25\,\text{GeV}$ (left plot) and $p_T^{\mu} > 35\,\text{GeV}$ (central plot). The right plot shows the results from the $W \to e\nu$ sample, with $p_T^{e} > 35\,\text{GeV}$. The results obtained with the sample of events with a three-prong secondary vertex are shown as black solid points. Open red points represent the results obtained with the sample of events having a two-prong secondary vertex and the blue open squares give the result from the sample of events with an identified muon inside the jet.

The main source of systematic uncertainties in the cross section ratios are those related with lepton reconstruction and identification, and in particular any lepton-charge dependent effect that may affect in a different amount to the W^+ and W^- candidate samples.

Systematic uncertainty in the cross section ratio due to lepton momentum scale and resolution is estimated following the same approach than for the normalized differential cross section. Uncertainties in the $W \to e\nu$ channel are smaller than 1% in the barrel, and of the order of 1.5% in the endcap region. They vary in the range $0.4 - 0.8\%$ in the $W \to \mu\nu$ channel, depending again on the muon pseudorapidity bin, the selected sample and the muon p_T threshold. They reduce to $\sim 0.2 - 0.3\%$ for the inclusive cross section ratios as the effect of muon momentum corrections among the different muon pseudorapidity bins cancel to a large extent, thus decreasing the final uncertainty.

Correction factors to the lepton reconstruction efficiencies $\text{efficiency}_{\text{data}}/\text{efficiency}_{\text{MC}}$ do not exhibit any dependence on the lepton charge. Apparent differences in efficiency between positive and negative leptons may arise from a different behaviour in lepton momentum scale and resolution. To this extent, systematic uncertainties quoted above for this effect also cover differences in efficiencies between ℓ^+ and ℓ^-.

Charge lepton misidentification in CMS is $< 0.3\%$ for electrons [44] and $\sim 10^{-4}$ for muons [45]. Associated systematic uncertainty in the positive to negative cross section ratio is proportional to the relative difference between $W^+ + \bar{c}$ and $W^- + c$ production, and this is small (measured cross section ratios are close to 1) and thus the total effect is neglected.

The following averaged ratios for the full lepton pseudorapidity range are derived:

$$\frac{\sigma(\mathrm{pp} \to \mathrm{W}^+ + \bar{\mathrm{c}} + \mathrm{X})}{\sigma(\mathrm{pp} \to \mathrm{W}^- + \mathrm{c} + \mathrm{X})}(p_{\mathrm{T}}^{\mu} > 25 \text{ GeV}) = 0.954 \pm 0.025 \text{ (stat.)} \pm 0.001 \text{ (syst.)},$$

$$\frac{\sigma(\mathrm{pp} \to \mathrm{W}^+ + \bar{\mathrm{c}} + \mathrm{X})}{\sigma(\mathrm{pp} \to \mathrm{W}^- + \mathrm{c} + \mathrm{X})}(p_{\mathrm{T}}^{\mu} > 35 \text{ GeV}) = 0.947 \pm 0.026 \text{ (stat.)} \pm 0.002 \text{ (syst.)},$$

$$\frac{\sigma(\mathrm{pp} \to \mathrm{W}^+ + \bar{\mathrm{c}} + \mathrm{X})}{\sigma(\mathrm{pp} \to \mathrm{W}^- + \mathrm{c} + \mathrm{X})}(p_{\mathrm{T}}^{\mathrm{e}} > 35 \text{ GeV}) = 0.93 \pm 0.03 \text{ (stat.)} \pm 0.01 \text{ (syst.)},$$

$$\frac{\sigma(\mathrm{pp} \to \mathrm{W}^+ + \bar{\mathrm{c}} + \mathrm{X})}{\sigma(\mathrm{pp} \to \mathrm{W}^- + \mathrm{c} + \mathrm{X})}(p_{\mathrm{T}}^{\ell} > 35 \text{ GeV}) = 0.939 \pm 0.019 \text{ (stat.)} \pm 0.005 \text{ (syst.)}.$$

9 Results and comparisons with theoretical predictions

Measured total and differential cross sections and cross section ratios can be compared with analytical calculations from the MCFM program. The $W + \mathrm{c}$ process [46] is available in MCFM up to $\mathcal{O}({\alpha_s}^2)$ with a massive charm quark. MCFM does not include contributions from gluon splitting into a $\mathrm{c\bar{c}}$ pair, but only contributions where the strange (or the down) quark couples to the W boson.

We show predictions for three global NNLO PDFs sets: MSTW2008, CT10 and NNPDF2.3. These three PDF sets have in common the use of neutrino charm production data to provide information on the proton's strangeness. In addition, we also compare with predictions using the NNPDF2.3$_{\mathrm{coll}}$ NNLO set [47]. This set of PDFs is based on high energy collider data only, and thus does not rely on the neutrino charm information. In particular, it includes W and Z production data from ATLAS, CMS and LHCb, and leads to a larger strangeness than that of gobal PDF sets.

The parameters of the calculation have been adjusted to match the experimental measurement and the generator level conditions: $E_{\mathrm{T}}^{jet} > 25$ GeV and $|\eta^{jet}| < 2.5$. Two sets of predictions are computed, to reproduce the different lepton p_{T} thresholds used in the analysis: $p_{\mathrm{T}}^{\ell} > 25$ GeV used in the $W \to \mu\nu$ channel and $p_{\mathrm{T}}^{\ell} > 35$ GeV in the $W \to \mu\nu$ and in the $W \to \mathrm{e}\nu$ channel.

Both the factorization and the renormalization scales are set to the value of the W boson mass. To estimate the uncertainty from missing higher perturbative orders, cross section predictions have also been computed varying independently the factorization and renomalization scales to twice and half the nominal value (with the constraint that the ratio of scales is never larger than two). The envelope of the cross sections with scale variation defines the theoretical scale uncertainty.

9.1 Total cross section

The measured total cross sections are consistent with theoretical expectations, although there are significant variations depending on the PDF set used in the prediction. The detailed theoretical predictions are compared graphically with the experimental measurement in Figure 10 where the central value of the prediction is shown, together with the uncertainty due to the PDF variations within each set. Scale uncertainties in the total cross section are of the order of $\pm 5\%$.

From Figure 10 we see that measured $W + c$ cross sections favour the PDF sets which include neutrino charm data, and in particular, are consistent with the result that the strange sea is suppresed with respect the up and down quark sea by a certain factor. NNPDF2.3$_{coll}$ PDF set instead, predicts a symmetric light quark sea, though larger PDF uncertainties were found. The hypothesis of a symmetric strange sea, as implemented in NNPDF2.3$_{coll}$ is disfavored by the results presented in this note, though measured cross sections are still consistent with the theoretical predictions within theoretical and experimental uncertainties.

9.2 Normalized differential cross section

Predictions for the normalized differential distribution are also obtained from analytical calculations from MCFM. Cross section predictions are obtained differentially in the lepton pseudorapidity using the same binning as in the data analysis: [0, 0.35], [0.35, 0.7], [0.7, 1.1], [1.1, 1.6], [1.6, 2.1] and the normalized distribution is computed from these cross section values as:

$$\frac{1}{\sigma(W+c)}\frac{d\sigma(W+c)}{d\eta} = \frac{\sigma(W+c)(\eta_{i,min} < \eta_\ell < \eta_{i,max})}{\sum_{i=1}^{i=5}(\sigma(W+c)(\eta_{i,min} < \eta_\ell < \eta_{i,max})} \times \frac{1.}{\Delta\eta_i} \quad (5)$$

The theoretical predictions are shown graphically in Figure 11 together with the average of the experimental measurements. There is a good agreement between the measured distributions and the theoretical predictions. Differences among the predictions obtained with the various PDF sets are of the same order than the expected uncertainties (at 68% CL). They increase when the transverse momentum of the lepton is larger than 35 GeV with respect to those for $p_T^\ell > 25$ GeV, but the associated uncertainties are also larger. It has to be noted as well the very different size of the associated uncertainties. This fact is mostly due to the different assumptions of the several PDF groups about the strange and anti-strange quarks content of the proton and to the different experimental inputs used [3]. Systematic uncertainties due to the scale variations are smaller than 1% for all muon pseudorapidity bins.

These differential distributions are normalized to the total cross sections obtained in section 6. The $\frac{d\sigma(W+c)}{d\eta}$ is presented in Figure 12. The theoretical predictions are also shown.

9.3 Charged ratios

Theoretical predictions for the cross section ratio $\sigma(W^+ + \bar{c})/\sigma(W^- + c)$ are derived in the same conditions explained before. Predictions for $\sigma(W^+ + \bar{c})$ and $\sigma(W^- + c)$ production are computed independently in the same lepton pseudorapidity intervals considered in the analysis. We note that this ratio is sensitive to the strangeness asymmetry in the proton, but also to the asymmetries of up down quarks from the Cabibbo suppressed $d \to W^- c$ ($\bar{d} \to W^+ \bar{c}$) process.

Both the central value and the associated PDF uncertainties for the various sets of predictions are quite different. These differences arise from the assumptions underlying each global fit. For instance, the CT10 set assumes equal content of strange and anti-strange quarks in the proton, leading to a very small PDF uncertainty in the charged ratio. On the other hand, both NNPDF2.3 and MSTW08 provide independent parametrizations of the strangeness asymmetry, and lead to similar PDF errors in this case.

PDF uncertainties are larger when the transverse momentum of the lepton from the W decay increases, they are larger for $p_T^\ell > 35$ GeV than for $p_T^\ell > 25$ GeV. Differences among the predictions are relatively large for some of the lepton pseudorapidity bins, $\sim 4-5\%$, although of the same order of the associated uncertainty. As expected, PDF uncertainties increase at forward rapidities, where the range of Bjorken-x is outside that covered by available data sensitive to

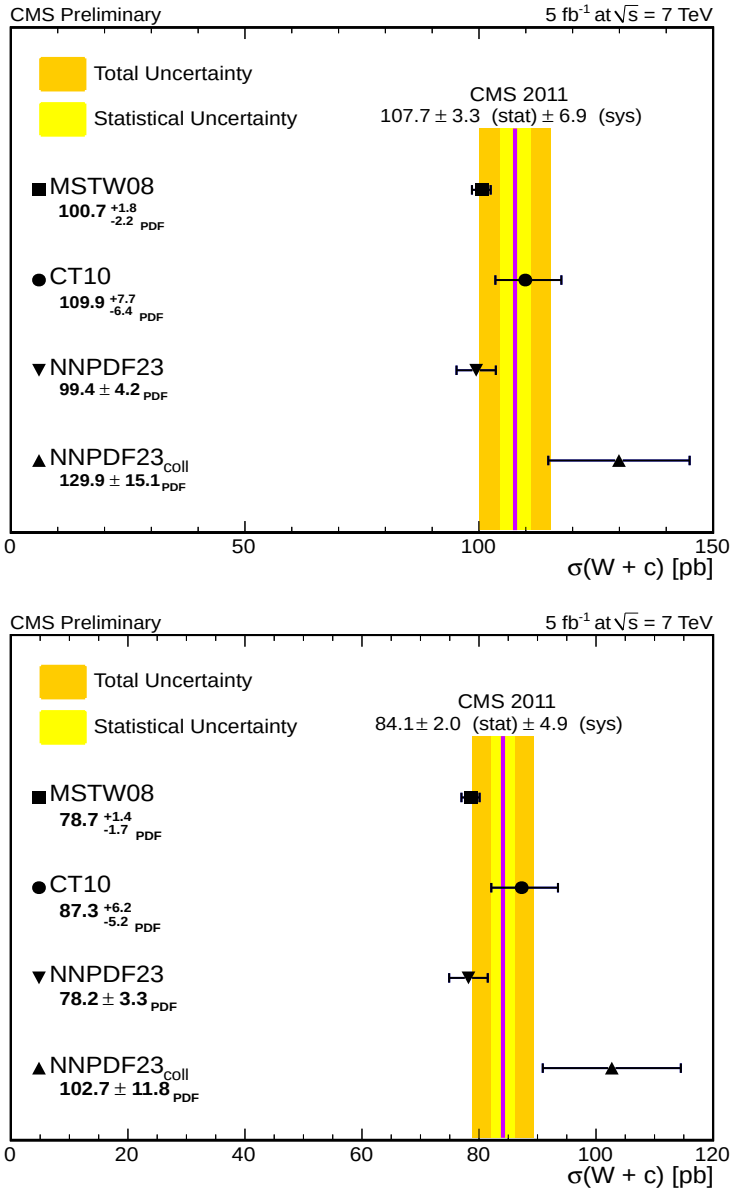

Figure 10: Comparison of the theoretical predictions for $\sigma(W + \mathrm{c})$ computed with MCFM and several sets of PDFs with the average of the experimental measurements. The top plot shows the predictions for a p_{T} threshold of the lepton from the W decay of $p_{\mathrm{T}}^{\ell} > 25\,\mathrm{GeV}$ and the bottom plot presents the predictions for $p_{\mathrm{T}}^{\ell} > 35\,\mathrm{GeV}$.

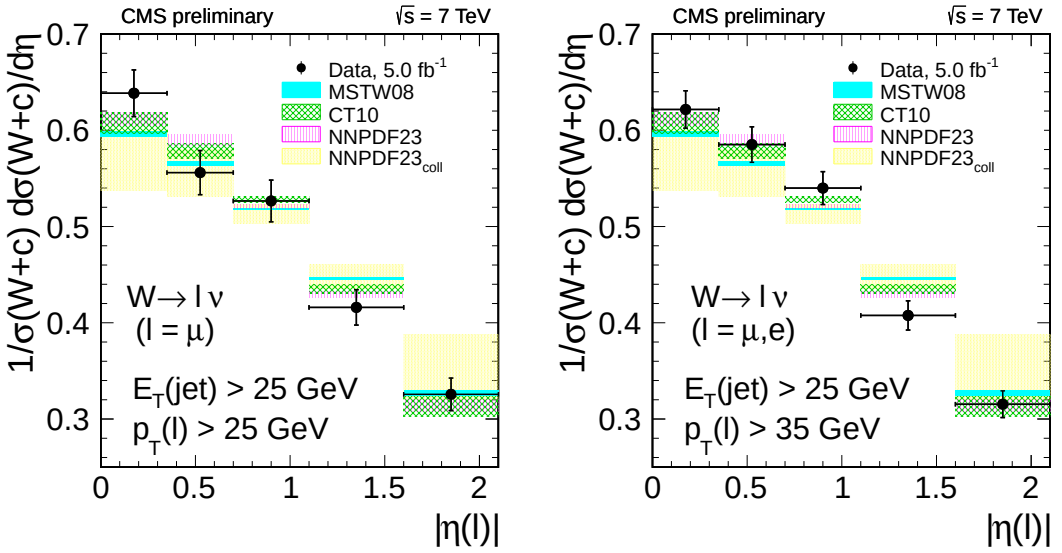

Figure 11: Normalized differential cross section, $\frac{1}{\sigma(W+c)}\frac{d\sigma(W+c)}{d\eta}$, as a function of the pseudorapidity of the lepton from the W decay. The transverse momentum of the lepton is larger than 25 GeV in the left plot (muon channel only) and larger than 35 GeV in the right plot. The black data points are the average of the results presented before with the three different subsamples: $D^{\pm}$, $D^{*\pm}(2010)$ and dileptons. In the right plot the results obtained with the $W \to \mu\nu$ samples and $W \to e\nu$ samples are combined. Theoretical predictions at NLO computed with MCFM and four different sets of PDFs are also shown.

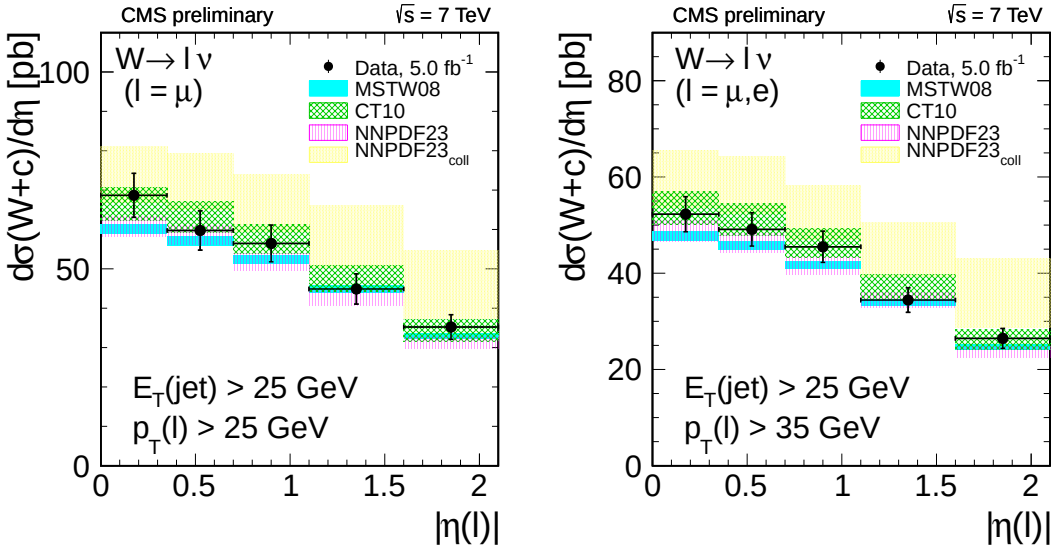

Figure 12: $W + c$ differential cross section, $\frac{d\sigma(W+c)}{d\eta}$, as a function of the pseudorapidity of the lepton from the W decay. The transverse momentum of the lepton is larger than 25 GeV in left plot (muon channel only) and larger than 35 GeV in the right plot. The black data points are the average of the results obtained with the four different subsamples: $D^{\pm}$, $D^{*\pm}(2010)$ and dileptons. In the right plot the results achieved with the $W \to \mu\nu$ samples and $W \to e\nu$ samples are combined. Theoretical predictions at NLO computed with MCFM and four different sets of PDFs are also shown.

strangeness. Systematic uncertainties in the charge ratio due to the scale variations are smaller than 1% for the full muon pseudorapidity range studied $[0., 2.1]$ and $1 - 2\%$ for the smaller pseudorapidity bins used for the differential measurement. The theoretical predictions are shown graphically in Figure 13 together with the average of the experimental measurements. There is a good agreement between the measured distributions and the theoretical predictions. Finally, the averaged inclusive cross section ratios are compared with their theoretical predictions in Figure 14.

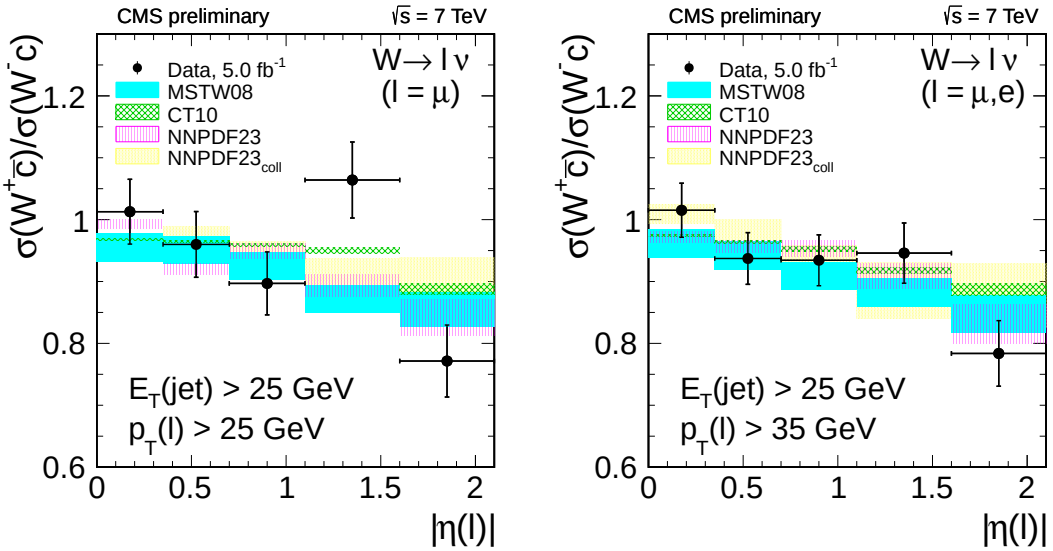

Figure 13: Differential cross section ratios, $R_c^{\pm}(\eta_\ell)$, as a function of the pseudorapidity of the lepton from the W decay. Results for the $p_{\mathrm{T}}^{\ell} > 25\,\mathrm{GeV}$ case are shown in the left plot (mon channel only). In the right plot, the transverse momentum of the lepton is larger than 35 GeV. The black data points are the average of the results presented before with the three different subsamples: $\mathrm{D}^{\pm}$, $\mathrm{D}^{*\pm}(2010)$ and dileptons. In the right plot the results obtained with the $W \to \mu\nu$ samples and $W \to \mathrm{e}\nu$ samples are combined. Theoretical predictions at NLO computed with MCFM and four different sets of PDFs are also shown.

10 Study of $W + c\bar{c}$ and $W + b\bar{b}$ background contributions

In this section we analyze in more detail events where the selected charm meson ($\mathrm{D}^{*\pm}(2010)$ or $\mathrm{D}^{\pm}$) and the associated W boson have the same sign of the charge. These events are part of the backgrounds that are suppressed via OS-SS subtraction, and therefore have no influence on the results discussed in the previous sections. However, they also contain important experimental information, since they are sensitive to $W + c\bar{c}$ and $W + b\bar{b}$ final states. One advantage of D-meson control samples is the a priori reduction of contributions containing light-jets. In addition, the $\mathrm{D}^{*\pm}(2010) \to \pi^{\pm}\mathrm{D}^0$ channel is less affected by contamination from b-quark decays, due to the primary vertex constraint imposed on the charged pion.

In order to increase the statistical significance of the study, $W \to \mu\nu$ and $W \to \mathrm{e}\nu$ decays are combined. Since the strong background suppression via OS-SS subtraction is not operational in this case, the Drell-Yan plus jets contamination is reduced by rejecting events containing two muons or two electrons with $p_{\mathrm{T}}^{\ell} > 20$ GeV. No specific cuts are used to reduce $t\bar{t}$ backgrounds. However, their contribution is controlled by studying the observed yields as a function of the jet multiplicity.

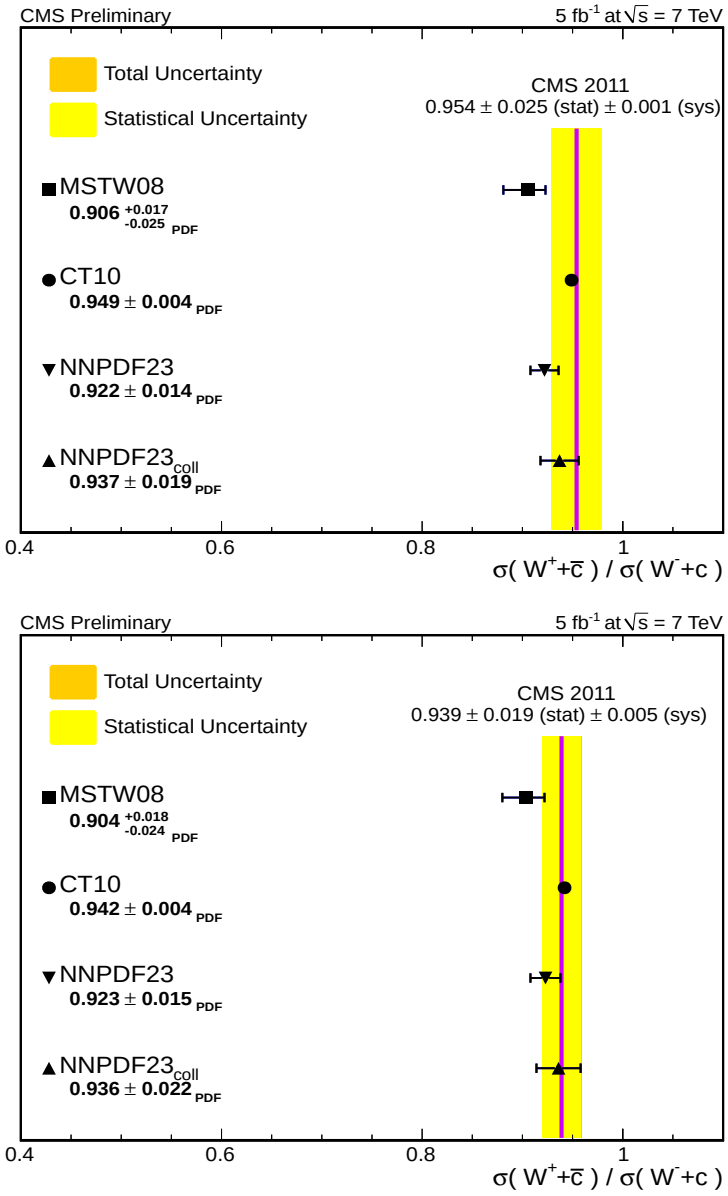

Figure 14: Comparison of the theoretical predictions for $R_c^{\pm}$ computed with MCFM and several sets of PDFs with the average of the experimental measurements. The top plot compares the average of the measurements done in the muon channel for a p_{T} threshold of the lepton from the W decay of $p_{\mathrm{T}}^{\ell} > 25\,\mathrm{GeV}$. The bottom plot presents the average of the measurements in the muon and in the electron channel with the predictions for $p_{\mathrm{T}}^{\ell} > 35\,\mathrm{GeV}$.

$D^{\pm}$ and $D^{*\pm}(2010)$ decays are identified following exactly the same procedure as followed in the measurement of the total cross section. The simulated rates for the resonant decay chains are reweighted in order to match the reference experimental values, and the $W + c$ contribution in the continuum is scaled up by 10%, in order to better match the central value of the measured cross section. Figure 15-left shows the invariant mass distribution for the case where W and $D^{\pm}$ have same-sign charges. The different Monte Carlo sub-contributions in the W component are also detailed. There is an overall significant excess in data with respect to the predictions of MADGRAPH + PYTHIA, which is more evident at the resonant peak. A fit to a Gaussian peak plus a second order polynomial gives $537 \pm 51(stat.)$ events for the resonant signal, whereas a similar fit in Monte Carlo gives $187 \pm 43(stat.)$. Figure 15-right shows the dependence of the excess at the $D^{\pm}$ resonance peak as a function of the number of jets with $E_T^{jet} > 25$ GeV, $|\eta^{jet}| < 2.5$. The excess is concentrated in the region of low multiplicities, and it is much more pronounced in events with only one jet, which excludes an interpretation in terms of a $t\bar{t}$ excess. The single-top component is predicted to be small, but it is also studied on the same-sign subsample containing only one jet plus an additional forward jet in the region $|\eta^{jet}| > 3.5$, which is dominated by t-channel single-top production. We find that the POWHEG-based single-top simulation describes adequately the data. A possible explanation of the observed deficit in the $W +$ jets Monte Carlo is a significant deficit in the PYTHIA description of the kinematic region dominated by hard-collinear gluon-splitting into heavy quarks. A global rescaling of these components leads to approximate agreement for most of the kinematic distributions studied. Figure 16-right shows the equivalent distributions for the $D^{*\pm}(2010)$ decay channel. Qualitatively, the conclusions are similar to the $D^{\pm}$ case, with the excess more concentrated in events with only one jet. Within the limited statistics, we do not observe significant discrepancies in events containing two D-meson tagged jets, suggesting that the matrix-element generation part governed by MadGraph is adequate.

In order to further test the dependence on the parton-jet matching approach and on the details of the parton shower generator, we perform an alternative study of the $D^{\pm}$ channel using SHERPA (v1.3.1)[48] as $W +$ jets Monte Carlo reference. SHERPA employs a matching method different from MadGraph and has its own fragmentation and hadronization steps, different from PYTHIA. Indeed SHERPA provides a better description of the same-sign distributions measured in data, as illustrated in Figure 17.

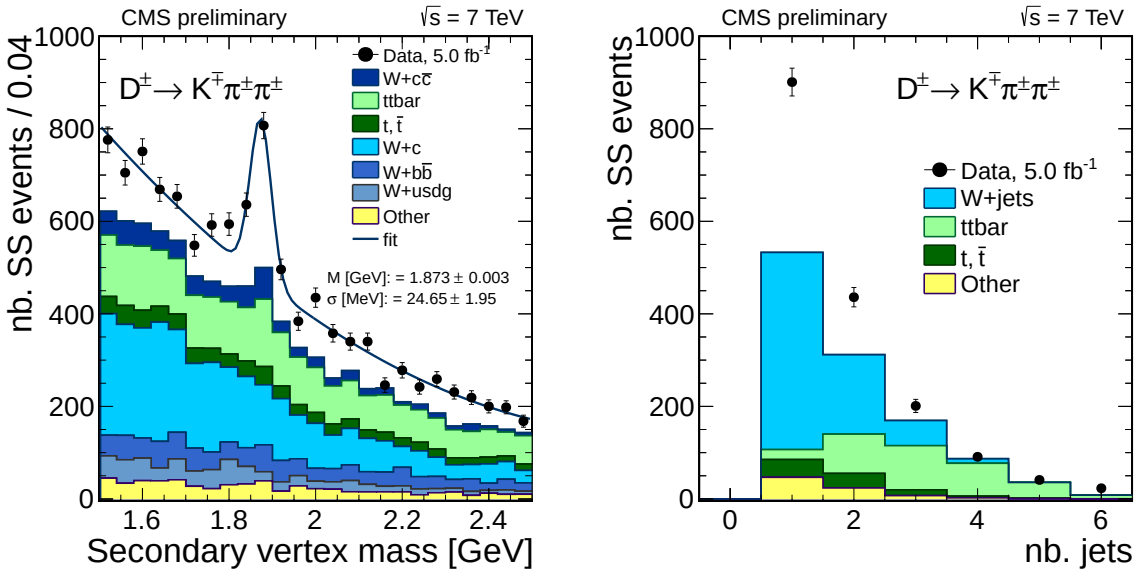

Figure 15: Left: invariant mass distribution of three-prong secondary vertices in W + jets events under the assumption of a $K^{\mp}\pi^{+}\pi^{+}$ final state. The figure shows those events where the W and the charmed meson have same charge sign (SS). A clear $D^{\pm}$ data excess is observed. The excess in the SS can not be attributed to a genuine W + c signal. Right: dependence of the signal in the peak region with respect to the number of jets. The excess is concentrated in the region of very low jet multiplicity, dominated by W + jets events.

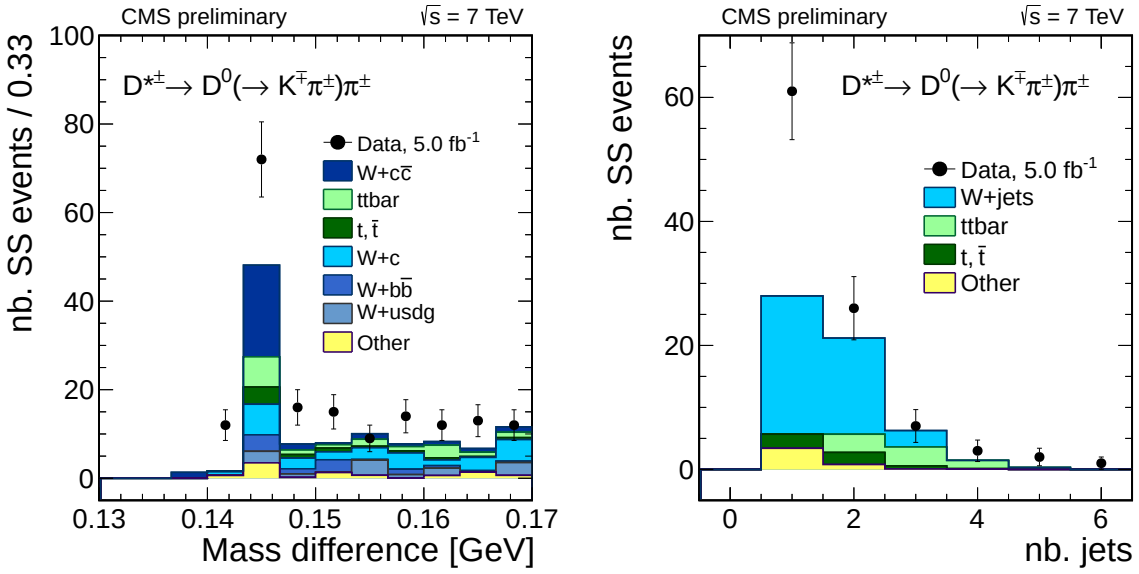

Figure 16: Left: invariant mass difference between $D^{*\pm}(2010)$ and D^{0} candidates in W + jets events. The figure shows those events where the W and the charmed meson have same charge sign (SS). A clear $D^{*\pm}(2010)$ data excess is observed. The excess in the SS can not be attributed to a genuine W + c signal. Right: dependence of the signal in the peak region with respect to the number of jets. The excess is concentrated in the region of very low jet multiplicity, dominated by W + jets events.

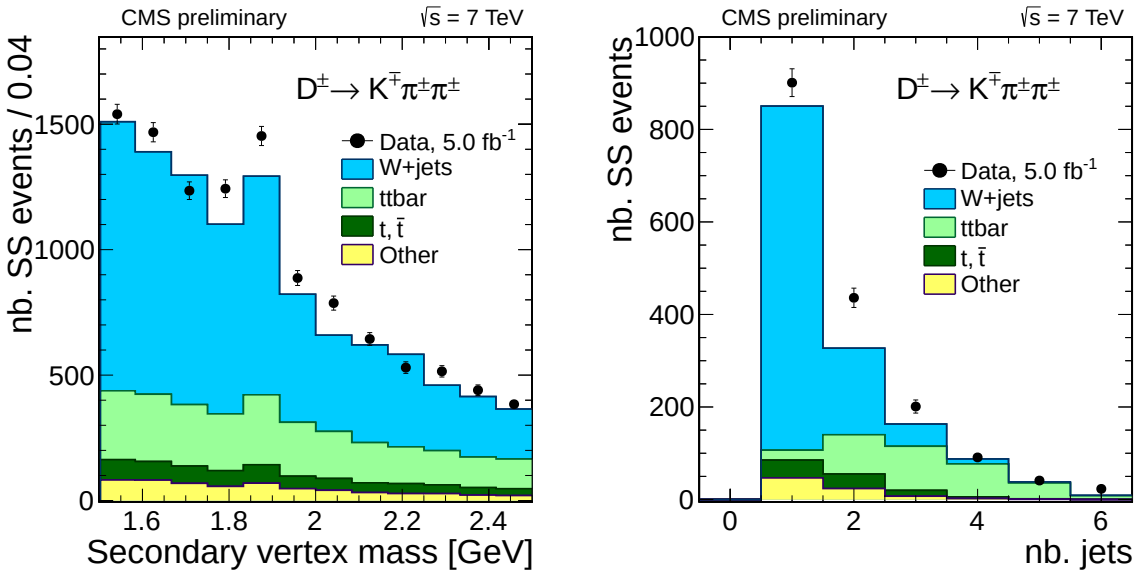

Figure 17: Left: invariant mass distribution of three-prong secondary vertices in W + jets events under the assumption of a $K^{\mp}\pi^{+}\pi^{+}$ final state. The figure shows those events where the W and the charmed meson have same charge sign (SS), but the Monte Carlo W + jets reference in this case is SHERPA. Right: distribution of the number of jets for events in the region of the resonance peak. Compared with the MADGRAPH − PYTHIA case, a better data-MC agreement is observed in this case.

11 Summary and Conclusions

Associated production of a W boson with a charm jet in pp collisions at a $\sqrt{s} = 7\,\text{TeV}$ is experimentally established for the first time, in the data sample collected by the CMS experiment during the 2011 LHC run, with an integrated luminosity of $5\,\text{fb}^{-1}$. The signal of W boson production together with a charm jet is evinced through the identification of the leptonic decay of the W in its muonic and electronic decay modes and the reconstruction of exclusive and semiinclusive final states from the decay of charm hadrons.

CMS heavy flavour tagging algorithms allow to select efficiently candidate samples with a displaced secondary vertex with two and three charged tracks corresponding to the decay products of charm mesons. Clear signals of $\text{D}^{\pm}$ mesons are observed through the reconstruction of the decay mode $\text{D}^{\pm} \to \text{K}^{\mp}\pi^{\pm}\pi^{\pm}$ in events with three-prongs secondary vertices. A clean signal from D^0 production in the decay chain $\text{D}^{*\pm}(2010) \to \text{D}^0\pi^{\pm}$, with the subsequent decay $\text{D}^0 \to \text{K}^{\mp}\pi^{\pm}$ is extracted in events with two-prongs secondary vertices. In addition, an independent $W + \text{c}$ sample with an identified muon from the semileptonic decay of the charm quark is selected.

Analysis of these samples has allowed to perform a detailed study of the $W + \text{c}$ process at the LHC energy. Total and differential production cross sections have been measured for the kinematic region $p_T^{jet} > 25$ GeV, $|\eta^{jet}| < 2.5$ in the three charm particles decay modes. Two different regimes for the transverse momentum of the lepton from the W decay are explored: $p_{\text{T}}^{\ell} > 25\,\text{GeV}$ in the W muon decay channel only, and $p_{\text{T}}^{\ell} > 35\,\text{GeV}$ in both the muon and the electron W decay channels. Differential measurements are produced as a function of the pseudorapidity of the lepton from the W decay.

The following total cross sections are measured:

$$\sigma(\text{pp} \to \text{W} + \text{c} + \text{X}) \times \mathcal{B}(\text{W} \to \mu\nu, p_{\text{T}}^{\mu} > 25\ \text{GeV}) = 107.7 \pm 3.3\ (\text{stat.}) \pm 6.9\ (\text{syst.})\ \text{pb}$$

$$\sigma(\text{pp} \to \text{W} + \text{c} + \text{X}) \times \mathcal{B}(\text{W} \to \mu\nu, p_{\text{T}}^{\mu} > 35\ \text{GeV}) = 82.9 \pm 2.6\ (\text{stat.}) \pm 5.1(\text{syst.})\ \text{pb}$$

$$\sigma(\text{pp} \to \text{W} + \text{c} + \text{X}) \times \mathcal{B}(\text{W} \to \text{e}\nu, p_{\text{T}}^{\text{e}} > 35\ \text{GeV}) = 85.3 \pm 2.5\ (\text{stat.}) \pm 5.7\ (\text{syst.})\ \text{pb}$$

$$\sigma(\text{pp} \to \text{W} + \text{c} + \text{X}) \times \mathcal{B}(\text{W} \to \ell\nu, p_{\text{T}}^{\ell} > 35\ \text{GeV}) = 84.1 \pm 2.0\ (\text{stat.}) \pm 4.9\ (\text{syst.})\ \text{pb}$$

Total and differential cross sections are compared with theoretical predictions calculated using four different sets of Parton Distribution Functions. A good agreement with the total cross sections measured is observed. Theoretical calculations also predict differential cross section shapes in agreement with the measured ones.

Total and differential cross section ratios of W^+ to W^- bosons associated production together with a charm jet ($\sigma(W^+ + \bar{\text{c}})/\sigma(W^- + \text{c})$) are also measured and compared with the corresponding theoretical predictions. Predictions for these ratios are specially sensitive to the assumptions applied by the several PDF groups in the global fits about the strange and antistrange quark content. The following total cross section ratios are obtained:

$$\frac{\sigma(pp \to W^+ + \bar{c} + X)}{\sigma(pp \to W^- + c + X)}(p_T^{\mu} > 25\ \text{GeV}) = 0.954 \pm 0.025\ (\text{stat.}) \pm 0.001\ (\text{syst.}),$$

$$\frac{\sigma(pp \to W^+ + \bar{c} + X)}{\sigma(pp \to W^- + c + X)}(p_T^{\mu} > 35\ \text{GeV}) = 0.947 \pm 0.026\ (\text{stat.}) \pm 0.002\ (\text{syst.}),$$

$$\frac{\sigma(pp \to W^+ + \bar{c} + X)}{\sigma(pp \to W^- + c + X)}(p_T^{e} > 35\ \text{GeV}) = 0.93 \pm 0.03\ (\text{stat.}) \pm 0.01\ (\text{syst.}),$$

$$\frac{\sigma(pp \to W^+ + \bar{c} + X)}{\sigma(pp \to W^- + c + X)}(p_T^{\ell} > 35\ \text{GeV}) = 0.939 \pm 0.019\ (\text{stat.}) \pm 0.005\ (\text{syst.}).$$

Theoretical predictions are in good agreement with these measurements.

Whereas reference simulations reproduce the measured total $W + c$ cross sections at the 10% level, an analysis of data distributions that are sensitive to $W + c\bar{c}$ and $W + b\bar{b}$ backgrounds shows a large excess with respect to MADGRAPH + PYTHIA predictions in the kinematic region where the heavy quark and the heavy antiquark are contained in the same jet. This strong Monte Carlo deficit with respect to data is not confirmed by other parton shower approaches, like the one implemented in SHERPA.

References

[1] U. Baur et al., "The Charm content of W + 1 jet events as a probe of the strange quark distribution function", *Phys. Lett. B* **318** (1993) 544–548.

[2] A. Kusina et al., "Strange Quark PDFs and Implications for Drell-Yan Boson Production at the LHC", *Phys. Rev. D* **85** (2012) 094028, doi:10.1103/PhysRevD.85.094028.

[3] W. Stirling and E. Vryonidou, "Charm production in association with an electroweak gauge boson at the LHC", *Phys. Rev. Lett.* **109** (2012) 082002, doi:10.1103/PhysRevLett.109.082002.

[4] J. Rojo, "Parton Distributions and LHC Data", *Nuovo Cim.* **C035N1** (2012) 179–186, doi:10.1393/ncc/i2012-11118-8.

[5] G. Bozzi, J. Rojo, and A. Vicini, "The Impact of PDF uncertainties on the measurement of the W boson mass at the Tevatron and the LHC", *Phys. Rev. D* **83** (2011) 113008, doi:10.1103/PhysRevD.83.113008.

[6] NNPDF Collaboration, "Precision determination of electroweak parameters and the strange content of the proton from neutrino deep-inelastic scattering", *Nucl. Phys. B* **823** (2009) 195–233, doi:10.1016/j.nuclphysb.2009.08.003.

[7] CDF Collaboration, "First measurement of the production of a W boson in association with a single charm quark in p anti-p collisions at $\sqrt{s} = 1.96$ TeV", *Phys. Rev. Lett.* **100** (2008) 091803, doi:10.1103/PhysRevLett.100.091803.

[8] CDF Collaboration, "Observation of the Production of a W Boson in Association with a Single Charm Quark", *Phys. Rev. Lett.* **110** (2013) 071801, doi:10.1103/PhysRevLett.110.071801.

[9] D0 Collaboration, "Measurement of the ratio of the $p\bar{p} \to W+$ c-jet cross section to the inclusive $p\bar{p} \to W+$ jets cross section", *Phys.Lett. B* **666** (2008) 23–30, doi:10.1016/j.physletb.2008.06.067.

[10] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.

[11] CMS Collaboration, "Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV", *JINST* **7** (2012) P10002, doi:10.1088/1748-0221/7/10/P10002.

[12] CMS Collaboration, "Electron reconstruction and identification at $\sqrt{s} = 7$ TeV", *CMS PAS* **EGM-10-004** (2010).

[13] CMS Collaboration, "Particle-flow commissioning with muons and electrons from J/Psi, and W events at 7 TeV", *CMS PAS* **PFT-2010-003** (2010).

[14] J. Alwall et al., "MadGraph5: Going Beyond", *JHEP* **06** (2011) 128, doi:10.1007/JHEP06(2011)128.

[15] T. Sjostrand, S. Mrenna, and P. Skands, "PYTHIA 6.4 Physics and Manual", *JHEP* **05** (2006) 026, doi:10.1088/1126-6708/2006/05/026.

[16] CMS Collaboration, "Jet Production Rates in Association with W and Z Bosons in pp Collisions at $\sqrt{s} = 7$ TeV", *JHEP* **01** (2012) 010, doi:10.1007/JHEP01(2012)010.

[17] S. Alioli et al., "NLO vector-boson production matched with shower in POWHEG", *JHEP* **07** (2008) 060, doi:10.1088/1126-6708/2008/07/060.

[18] H.-L. Lai et al., "New parton distributions for collider physics", *Phys. Rev. D* **82** (2010) 074024, doi:10.1103/PhysRevD.82.074024.

[19] J. Pumplin et al., "New generation of parton distributions with uncertainties from global QCD analysis", *JHEP* **07** (2002) 012, doi:10.1088/1126-6708/2002/07/012.

[20] CMS Collaboration, "Measurement of the Underlying Event Activity at the LHC with $\sqrt{s} = 7$ TeV and Comparison with $\sqrt{s} = 0.9$ TeV", *JHEP* **09** (2011) 109, doi:10.1007/JHEP09(2011)109.

[21] K. Melnikov and F. Petriello, "Electroweak gauge boson production at hadron colliders through $\mathcal{O}(\alpha_s{}^2)$", *Phys. Rev. D* **74** (2006) 114017, doi:10.1103/PhysRevD.74.114017.

[22] A. Martin et al., "Parton distributions for the LHC", *Eur. Phys. J. C* **63** (2009) 189–285, doi:10.1140/epjc/s10052-009-1072-5.

[23] J. M. Campbell and R. Ellis, "MCFM for the Tevatron and the LHC", *Nucl. Phys. Proc.* **215 Suppl. 205** (2010) 10, doi:10.1016/j.nuclphysbps.2010.08.011.

[24] N. Kidonakis, "Next-to-next-to-leading soft-gluon corrections for the top quark cross section and transverse momentum distribution", *Phys. Rev. D* **82** (2010) 114030, doi:10.1103/PhysRevD.82.114030.

[25] GEANT4 Collaboration, "GEANT4: A simulation toolkit", *Nucl. Instrum. and Methods* **A506** (2003) 250–303, doi:10.1016/S0168-9002(03)01368-8.

[26] CMS Collaboration, "Measurements of Inclusive W and Z Cross Sections in pp Collisions at $\sqrt{s} = 7$ TeV with the CMS experiment", *JHEP* **010** (2011) 132, doi:10.1007/JHEP10(2012)132.

[27] M. Cacciari, G. Salam, and G. Soyez, "The anti-kt jet clustering algorithm", *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063.

[28] M. Cacciari and G. Salam, "Pileup subtraction using jet areas", *Phys. Lett. B* **659** (2008) 119–126, doi:10.1016/j.physletb.2007.09.077.

[29] M. Cacciari, G. Salam, and G. Soyez, "The catchment area of jets", *JHEP* **04** (2008) 005, doi:10.1088/1126-6708/2008/04/005.

[30] CMS Collaboration, "Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS", *J. Instrum.* **6** (2011) P11002, doi:10.1088/1748-0221/6/11/P11002.

[31] W. Waltenberger, R. Fruhwirth, and P. Vanlaer, "Adaptive vertex fitting", *J. Phys. G* **34** (2007) N343, doi:10.1088/0954-3899/34/12/N01.

[32] CMS Collaboration, "Identification of b-quark jets with the CMS experiment", *CMS PAS* **BTV-12-001** (2012) arXiv:1211.4462. Comments: Submitted to the Journal of Instrumentation.

[33] Particle Data Group and J. Beringer and others, "Review of Particle Physics", *Phys. Rev. D* **86** (2012) 010001, doi:10.1103/PhysRevD.86.010001.

[34] OPAL Collaboration, "A study of charm hadron production in $Z \to c\bar{c}$ and $Z \to b\bar{b}$ decays at LEP", *Z. Phys. C* **72** (1996) 1–16, doi:10.1007/s002880050218.

[35] ALEPH Collaboration, "Study of charm production in Z decays", *Eur. Phys. J. C* **16** (2000) 597–611, doi:10.1007/s100520000421.

[36] DELPHI Collaboration, "Measurements of the Z partial decay width into $c\bar{c}$ and multiplicity of charm quarks per b decay", *Eur. Phys. J. C* **12** (2000) 225–241, doi:10.1007/s1005299000228.

[37] OPAL Collaboration, "Measurement of $f(c \to D^* + X)$, $f(b \to D^* + X)$ and $\Gamma(c\bar{c})/\Gamma(\textit{hadronic})$ using $D^{*\pm}$ mesons", *Eur. Phys. J. C* **1** (1998) 439–459, doi:10.1007/s100520050095.

[38] DELPHI Collaboration, "Determination of $P(c \to D^{*+})$ and $BR(c \to \ell^+)$ at LEP 1", *Eur. Phys. J. C* **12** (2000) 209–224, doi:10.1007/s100529900227.

[39] L. Gladilin, "Charm Hadron Production Fractions", arXiv:0603175.

[40] CMS Collaboration, "Measurement of Tracking Efficiency", *CMS PAS* **TRK-10-002** (2010).

[41] R. Eusebi, on behalf of the CMS Collaboration, "Jet energy corrections and uncertainties in CMS: Reducing their impact on physics measurements", *J.Phys.Conf.Ser.* **404** (2012) 012014, doi:10.1088/1742-6596/404/1/012014.

[42] CMS Collaboration, "Absolute Calibration of the Luminosity Measurement at CMS: Winter 2012 Update", *CMS PAS* **SMP-12-008** (2012).

[43] NNPDF Collaboration, "Unbiased global determination of parton distributions and their uncertainties at NNLO and at LO", *Nucl.Phys. B* **855** (2012) 153–221, doi:10.1016/j.nuclphysb.2011.09.024.

[44] CMS Collaboration, "Measurement of the Electron Charge Asymmetry in Inclusive W Production in pp Collisions at $\sqrt{s} = 7$ TeV", *Phys. Rev. Lett.* **109** (2012) 111806, doi:10.1103/PhysRevLett.109.111806.

[45] CMS Collaboration, "Measurement of the charge ratio of atmospheric muons with the CMS detector", *Phys. Lett. B* **692** (2010) 83–104, doi:10.1016/j.physletb.2010.07.033.

[46] W. Giele, S. Keller, and E. Laenen, "QCD corrections to W boson plus heavy quark production at the Tevatron", *Phys. Lett. B* **372** (1996) 141–149, doi:10.1016/0370-2693(96)00078-0.

[47] NNPDF Collaboration, "Parton distributions with LHC data", *Nucl.Phys. B* **867** (2013) 244–289, doi:10.1016/j.nuclphysb.2012.10.003.

[48] T. Gleisberg et al., "Event Generation with Sherpa 1.1", *JHEP* **02** (2009) 007, doi:10.1088/1126-6708/2009/02/007.