Measurement of associated Z + charm production in proton–proton collisions at $\sqrt{s} = 8$ TeV

CMS Collaboration*
CERN, 1211 Geneva 23, Switzerland

Received: 6 November 2017 / Accepted: 21 March 2018 / Published online: 9 April 2018
© CERN for the benefit of the CMS collaboration 2018

Abstract A study of the associated production of a $Z$ boson and a charm quark jet ($Z + c$), and a comparison to pro-
duction with a $b$ quark jet ($Z + b$), in $pp$ collisions at a
centre-of-mass energy of 8 TeV are presented. The analy-
sis uses a data sample corresponding to an integrated lumi-
nosty of 19.7 fb$^{-1}$, collected with the CMS detector at the
CERN LHC. The $Z$ boson candidates are identified through
developments into pairs of electrons or muons. Jets originat-
ing from heavy flavour quarks are identified using semi-
luminous decays of $c$ or $b$ flavoured hadrons and hadronic decays
d of charm hadrons. The measurements are performed in the
kinematic region with two leptons with $p_T > 20$ GeV,
$|\eta| < 2.1$, $71 < m_{\ell\ell} < 111$ GeV, and heavy flavour jets
with $p_T^{\text{jet}} > 25$ GeV and $|\eta^{\text{jet}}| < 2.5$. The $Z + c$ production cross
cross section is measured to be $\sigma(pp \to Z + c + X) B(Z \to \ell^+\ell^-) = 8.8 \pm 0.5 \text{(stat)} \pm 0.6 \text{(syst)}$ pb. The ratio of the $Z + c$
and $Z + b$ production cross sections is measured to be $\sigma(pp \to Z + c + X)/\sigma(pp \to Z + b + X) = 2.0 \pm 0.2 \text{(stat)} \pm 0.2 \text{(syst)}$.
The $Z + c$ production cross section and the cross section ratio
are also measured as a function of the transverse momentum
of the $Z$ boson and of the heavy flavour jet. The measurements
are compared with theoretical predictions.

1 Introduction

The CERN Large Hadron Collider (LHC) has delivered a
large sample of $pp$ collisions containing events with a vector
boson ($V$) accompanied by one or more jets ($V+\text{jets}$). Some
of these events involve the production of a vector boson in
association with jets originating from heavy flavour (HF)
quarks and can be used to study specific predictions of the
standard model (SM).

These $V+\text{jets}$ events constitute an important background
to many ongoing searches for new physics beyond the SM. A
proper characterization of these processes and validation of
their theoretical description is important to provide a reliable
estimate of their specific backgrounds to the various searches.
For example, third-generation scalar quarks ($\text{squarks}$) that
are predicted by supersymmetric theories to decay via charm
quarks have been searched for in final states with a charm
quark jet ($c$-jet) and a large transverse momentum imbalance
[1–3]. A dominant background to this process is the
associated production of a $c$-jet and a $Z$ boson that decays invisibly into neutrinos. An improved description of this
background can be obtained from a measurement of the same
process with the $Z$ boson decaying into charged leptons.

Similarly, the associated production of a $Z$ boson and HF
jets is a significant background to the production of the Higgs
boson in association with a $Z$ boson ($pp \to Z + H + X$; $H \to q\bar{q}$). Experimental studies of this process in the context of the
SM focus on an analysis with $b$-quarks in the final state [4–7],
although some models beyond the SM also predict enhanced
decay rates in the $c\bar{c}$ final state [8]. In either case, it is impor-
tant to understand the relative contribution of the different
flavours to the $Z+\text{HF}$ jets background to minimize the associ-
at-ed systematic uncertainties.

The possibility of observing evidence of an intrinsic charm
(IC) quark component in the nucleon has recently received
renewed interest [9]. The associated production of neutral
vector bosons and $c$ jets ($V+c$) has been identified [10–13] as
a suitable process to investigate this physics topic. One of the
main effects of an IC component would be an enhancement
of $Z+c$ production, mainly at large values of the transverse
momentum of the $Z$ boson and of the $c$ jet.

Production of a $Z$ boson and a $c$-jet has been studied in
high-energy hadron collisions by the D0 [14] and CDF [15]
experiments at the Tevatron $p\bar{p}$ collider. More recently, the
LHCb Collaboration has measured the associated production
of a $Z$ boson and a $D$ meson in the forward region in $pp$
collisions at $\sqrt{s} = 7$ TeV [16].

In this paper we present a measurement of the production
cross section at $\sqrt{s} = 8$ TeV of a $Z$ boson and at least one
jet from a $c$ quark. In addition, the relative production of a $Z$

* e-mail: cms-publication-committee-chair@cern.ch
boson and a jet from heavy quarks of different flavours (c or b) is quantified by the ratio of their production cross sections. The associated production of a Z boson and at least one or two b jets using an inclusive b tagging technique to identify Z + b events has been studied with the same dataset and the results are reported in Ref. [17]. To reduce the uncertainties in the ratio, the production cross section of a Z boson and a jet from a b quark is remeasured in this analysis using exactly the same methodology as for the Z + c cross section. The remeasured Z + b cross section agrees with the published value within one standard deviation and is used in the ratio measurement.

The Z boson is identified through its decay into a pair of electrons or muons. Jets with HF quark content are identified through (1) the semileptonic decay of c or b flavoured hadrons with a muon in the final state, and (2) using exclusive hadronic decays of charm hadrons. The cross section and cross section ratio are measured at the level of stable particles, which are defined prior to the emission of any electroweak radiation. To minimize acceptance corrections, the measurements are restricted to a phase space that is close to the experimental fiducial volume with optimized sensitivity for the investigated processes: two leptons with transverse momentum \( p_T > 20 \text{ GeV} \), pseudorapidity \( |\eta| < 2.1 \), and dilepton invariant mass consistent with the mass of the Z boson, \( 71 < m_{\ell\ell} < 111 \text{ GeV} \), together with a c (b) jet with \( p_T^{\text{jet}} > 25 \text{ GeV} \), \( |\eta^{\text{jet}}| < 2.5 \). The jet should be separated from the leptons of the Z boson candidate by a distance \( \Delta R(\text{jet}, \ell) = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2} > 0.5 \). The cross section \( \sigma(pp \to Z+c+X)/\sigma(pp \to Z+b+X) \) (abbreviated as \( \sigma \)) and the cross section ratio \( \sigma(pp \to Z+c+X)/\sigma(pp \to Z+b+X) \) (abbreviated as \( \sigma \)) are determined both inclusively and differentially as a function of the transverse momentum of the Z boson, \( p_T^Z \), and the \( p_T \) of the jet with heavy flavour content, \( p_T^{\text{jet}} \).

The paper is structured as follows. The CMS detector is briefly described in Sect. 2, and the data and simulated samples used are presented in Sect. 3. Section 4 deals with the selection of the Z + HF jets signal sample, the auxiliary samples of events from the associated production of W+c, and top quark-antiquark (t\(^\bar{\text{t}}\)) production. The determination of the c tagging efficiency is the subject of Sect. 5. The analysis strategy devised to separate the two contributions, Z + c and Z + b, in the sample of Z + HF jets is detailed in Sect. 6. Section 7 reviews the most important sources of systematic uncertainties and their impact on the measurements. Finally, the measurements of the inclusive Z + c cross section and the \( (Z+c)/(Z+b) \) cross section ratio are presented in Sect. 8, and the differential measurements are reported in Sect. 9. The main results of the paper are summarized in Sect. 10.

### 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 solenoid volume are a silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter, each composed of a barrel and two endcap sections. Extensive forward calorimetry complements the coverage provided by the barrel and endcap detectors. The silicon tracker measures charged particles within the pseudorapidity range \( |\eta| < 2.5 \). It consists of 1440 silicon pixel and 15 148 silicon strip detector modules. For nonisolated particles of \( p_T < 10 \text{ GeV} \) and \( |\eta| < 1.4 \), the track resolutions are typically 1.5% in \( p_T \) and 25–90 (45–150) \( \mu \text{m} \) in the transverse (longitudinal) impact parameter [18]. The electron momentum is estimated by combining the energy measurement in the ECAL with the momentum measurement in the tracker. The momentum resolution for electrons with \( p_T \approx 45 \text{ GeV} \) from \( Z \to e^+e^- \) decays ranges from 1.7% for nonshowering electrons in the barrel region to 4.5% for showering electrons in the endcaps [19]. Muons are measured in the pseudorapidity range \( |\eta| < 2.4 \), using three technologies: drift tubes, cathode strip chambers, and resistive plate chambers. Matching muons to tracks measured in the silicon tracker results in a relative transverse momentum resolution for muons with \( p_T < 100 \text{ GeV} \) of 1.3–2.0% in the barrel and better than 6% in the endcaps. The \( p_T \) resolution in the barrel is better than 10% for muons with \( p_T \) up to 1 TeV [20]. For nonisolated muons with \( p_T < 25 \text{ GeV} \), the relative transverse momentum resolution is 1.2–1.7% in the barrel and 2.5–4.0% in the endcaps [18]. The first level of the CMS trigger system [21], composed of custom hardware processors, uses information from the calorimeters and muon detectors to select events of interest in a fixed time interval of less than 4 \( \mu \text{s} \). The high-level trigger processor farm further decreases the event rate from around 100 kHz to less than 1 kHz, before data storage. A more detailed description of the CMS detector, together with a definition of the coordinate system used and the basic kinematic variables, can be found in Ref. [22].

### 3 Data and simulated samples

The data were collected by the CMS experiment during 2012 at the pp centre-of-mass energy of 8 TeV and correspond to an integrated luminosity of \( L = 19.7 \pm 0.5 \text{ fb}^{-1} \).

Samples of simulated events are produced with Monte Carlo (MC) event generators, both for the signal process and for the main backgrounds. A sample of signal Z boson events is generated with MADGRAPH v5.1.3.30 [23], interfaced with PYTHIA v6.4.26 [24] for parton showering and hadronizati-
tion using the MLM [25,26] matching scheme. The MADGRAPH generator produces parton-level events with a vector boson and up to four partons at leading order (LO) on the basis of a matrix-element calculation. The generation uses the parton distribution functions (PDF) set CTEQ6L [27]. The matching scale between jets from matrix element calculations and those produced via parton showers is 10 GeV, and the factorization and renormalization scales are set to \( q^2 = M_Z^2 + (p_T^Z)^2 \).

Other physics processes produce events with the same final state topology as the signal. The main background is the production of \( t\bar{t} \) events. Smaller contributions are expected from direct production of a pair of vector bosons: WW, WZ, and ZZ.

A sample of \( t\bar{t} \) events is generated with \textsc{powheg v1.0} [28–31], interfaced with \textsc{pythia} and using the CT10 [32] PDF set. The WW, WZ, and ZZ processes are modelled with samples of events generated with \textsc{pythia}6 and the CTEQ6L1 PDF set.

A sample of W boson events is generated with MADGRAPH interfaced with \textsc{pythia}6. It is used in the determination of the c tagging efficiency and to validate the modelling of relevant distributions with a data sample of W+jets events.

The matching scale between jets from matrix element calculations and those produced via parton showers is 10 GeV, and the factorization and renormalization scales are set to \( q^2 = M_W^2 + (p_T^W)^2 \). For all event generation the \textsc{pythia}6 parameters for the underlying event modelling are set to the MSTW2008NNLO [36] PDF set.

Generated events are processed through a full \textsc{geant4}-based [34] CMS detector simulation and trigger emulation. Simulated events are then reconstructed using the same algorithms as used to reconstruct collision data and are normalized to the integrated luminosity of the data sample using their respective cross sections. For electroweak processes the cross sections are evaluated to next-to-leading order (NNLO) with \textsc{mcfm} 6.6 [37] and using the MSTW2008NLO [36] PDF set. The \( \tau \tau \) cross section is taken at NNLO from Ref. [38].

The simulated samples incorporate additional pp interactions in the same or neighbouring bunch crossings (pileup). Simulated events are weighted so that the pileup distribution matches the measured one, with an average of about 21 pileup interactions per bunch crossing.

Simulated samples are corrected for differences between data and MC descriptions of lepton trigger, reconstruction, and selection efficiencies (\( \epsilon_\ell \)). Lepton efficiencies are evaluated with samples of dilepton events in the Z mass peak with the “tag-and-probe” method [39], and correction factors \( \epsilon_\ell^{\text{data}} / \epsilon_\ell^{\text{MC}} \), binned in terms of \( p_T \) and \( \eta \) of the leptons, are computed. These correction factors, based on the kinematics of each lepton in an event, are multiplied and used as an event weight.

The simulated signal sample includes Z boson events accompanied by jets originating from quarks of all flavours (b, c, and light). Events are classified as Z+b, Z+c, or Z+light flavour according to the flavour of the generator-level jets built from all showered particles after fragmentation and hadronization (all stable particles except neutrinos) and clustered with the same algorithm that is used to reconstruct data jets. A generator-level jet is defined to be b flavoured if \( p_T^{\text{gen jet}} > 15 \text{ GeV} \) and there is a b hadron among the particles generated in the event within a cone of radius \( \Delta R = 0.5 \) around the jet axis. Similarly, a generator-level jet is considered to be c flavoured if \( p_T^{\text{gen jet}} > 15 \text{ GeV} \) and there is a c hadron and no b hadrons within a cone of \( \Delta R = 0.5 \) around the jet axis. A Z+jets event is assigned as a Z+b event if there is at least a generator-level jet identified as a b flavoured jet regardless of the number of c flavoured or light jets. Z+c if there is at least a c flavoured jet at the generator-level and no b flavoured generator-level jets, and Z+light flavour otherwise.

4 Event reconstruction and selection

Electron and muon candidates are reconstructed following standard CMS procedures [19,20]. Jets, missing transverse energy, and related quantities are determined using the CMS particle-flow (PF) reconstruction algorithm [40], which identifies and reconstructs stable particle candidates arising from a collision with an optimized combination of the signals measured from all subdetectors.

Jets are built from PF candidates using the anti-\( k_T \) clustering algorithm [41] with a distance parameter of \( R = 0.5 \). The energy and momentum of the jets are corrected as a function of the \( p_T \) and \( \eta \) to account for the nonlinear response of the calorimeters and for the presence of pileup interactions [42,43]. Jet energy corrections are derived using samples of simulated events and further adjusted using dijet, photon+jet and Z+jet events in data.

The missing transverse momentum vector \( \mathbf{p}_T^{\text{miss}} \) is the projection on the plane perpendicular to the beams of the negative vector sum of the momenta of all particles that are reconstructed with the PF algorithm. The missing transverse energy variable, \( E_T^{\text{miss}} \), is defined as the magnitude of the \( \mathbf{p}_T^{\text{miss}} \) vector, and it is a measure of the transverse energy of particles leaving the detector undetected [44].

The primary vertex of the event, representing the hard interaction, is selected among the reconstructed vertices as the one with the highest sum of the transverse momenta squared of the tracks associated to it.
4.1 Selection of Z+HF jet events

Events with a pair of leptons are selected online by a trigger system that requires the presence of two lepton candidates of the same flavour with $p_T^{\ell} > 17$ and 8 GeV for the leading-$p_T$ and subleading-$p_T$ lepton candidates, respectively. The analysis follows the offline selections as used in the CMS $Z \rightarrow e^+e^-$ and $Z \rightarrow \mu^+\mu^-$ inclusive analyses [39] and requires the presence of two high-$p_T$ reconstructed leptons with opposite charges in the pseudorapidity region $|\eta^{\ell}| < 2.1$. The transverse momentum of the leptons has to be greater than 20 GeV.

The leptons are required to be isolated. The combined isolation $I_{\text{comb}}$ is used to quantify the additional hadronic activity around the selected leptons. It is defined as the sum of the transverse energy of neutral hadrons and photons and the transverse momentum of charged particles in a cone with $R < 0.3$ (0.4) around the electron (muon) candidate, excluding the contribution from the lepton itself. Only charged particles originating from the primary vertex are considered in the sum to minimize the contribution from pileup interactions. The contribution of neutral particles from pileup vertices is estimated and subtracted from $I_{\text{comb}}$. For electrons, this contribution is evaluated with the jet area method described in Ref. [45]; for muons, it is taken to be half the sum of the $p_T$ of all charged particles in the cone originating from pileup vertices. The factor one-half accounts for the expected ratio of charged to neutral particle energy in hadronic interactions. The electron (muon) candidate is considered to be isolated when $I_{\text{comb}}/p_T^{\ell} < 0.15$ (0.20). Finally, the analysis is restricted to events with a dilepton invariant mass, $m_{\ell\ell}$, in the range $91 \pm 20$ GeV in accordance with previous Z+jets measurements [17,46].

A $Z+\text{jets}$ sample is selected by requiring the presence of at least one jet with $p_T^{\text{jet}} > 25$ GeV and $|\eta^{\text{jet}}| < 2.5$. Jets with an angular separation between the jet axis and any of the selected leptons less than $\Delta R(\text{jet}, \ell) = 0.5$ are not considered. To reduce the contribution from $t\bar{t}$ events, we require $E_T^{\text{miss}}$ to be smaller than 40 GeV.

Hadrons with c or b quark content decay weakly with lifetimes of the order of $10^{-12}$ s and mean decay lengths larger than 100 µm at the LHC energies. Secondary vertices well separated from the primary vertex can be reconstructed from the tracks of their charged decay products. We focus on the following three signatures to identify jets originating from a heavy flavour quark:

- **Semileptonic mode** – A semileptonic decay of a heavy flavour hadron leading to a well-identified muon associated to a displaced secondary vertex.
- **D$^\pm$ mode** – A displaced secondary vertex with three tracks consistent with a $D^\pm \rightarrow K^\mp \pi^\pm \pi^\pm$ decay.
- **D*(2010)$^\pm$ mode** – A displaced secondary vertex with two tracks consistent with a $D^0 \rightarrow K^-\pi^+$ ($\bar{D}^0 \rightarrow K^+\pi^-$) decay and associated with a $D^{*+}(2010) \rightarrow D^0\pi^+$ ($D^{*-}(2010) \rightarrow \bar{D}^0\pi^-$) decay at the primary vertex.

Displaced secondary vertices for the first two categories are formed with either the Simple Secondary Vertex (SSV) [47] or the Inclusive Vertex Finder (IVF) [48,49] CMS vertex reconstruction algorithms. Both algorithms follow the adaptive vertex fitter technique [50] to construct a secondary vertex, but differ in the tracks used. The SSV algorithm takes as input the tracks constituting the jet; the IVF algorithm starts from a displaced track with respect to the primary vertex (seed track) and searches for nearby tracks, in terms of their separation distance in three dimensions and their angular separation around this seed, to build the vertex. Tracks used in a secondary vertex reconstruction must have $p_T > 1$ GeV. Vertices reconstructed with the IVF algorithm are considered first because of the higher efficiency of the algorithm. If no IVF vertex is found, SSV vertices are searched for, thus providing additional event candidates. We employ a different technique for the third (D*(2010)$^\pm$ mode) category, as described below in the text. The typical mass resolution in the D$^\pm$ and D*(2010)$^\pm$ reconstruction is $\approx$17 MeV in the decay modes analyzed here.

4.1.1 Selection in the semileptonic mode

The $Z+c$ ($Z+b$) events with a semileptonic c (b) quark decay are selected by looking for a reconstructed muon (muon-inside-a-jet) among the constituents of any of the selected jets. This muon-inside-a-jet candidate has to satisfy the same quality criteria as those imposed on the muons from the Z boson decay. The muon has to be reconstructed in the region $|\eta^{\mu}| < 2.4$, with $p_T^{\mu} < 25$ GeV, $p_T^{\mu}/p_T^{\text{jet}} < 0.6$, and it should not be isolated from hadron activity. The combined isolation has to be large, $I_{\text{comb}}/p_T^{\mu} > 0.2$. Furthermore, the muon-inside-a-jet is required to be associated to a secondary vertex, reconstructed either with the IVF or SSV algorithm. No minimum $p_T$ is required for the muon beyond the general $p_T > 1$ GeV requirement for the tracks used in the reconstruction of the secondary vertices. Muon reconstruction sets a natural threshold of $p_T \gtrsim 3$ GeV in the barrel region and $p_T \gtrsim 2$ GeV in the endcaps to ensure the muon passes the material in front of the muon detector and travels deep enough into the muon system to be reconstructed and satisfy the identification criteria [39]. The above selection results in 4145 events in the $Z \rightarrow e^+e^-$ channel and 5258 events in the $Z \rightarrow \mu^+\mu^-$ channel.

Figure 1 shows the transverse momentum distribution of the selected muon-inside-a-jet for $Z \rightarrow e^+e^-$ (left) and $Z \rightarrow$
The data are compared with the predictions of the MC simulations, which are composed of $Z + b$ events ($\approx 65\%$), $Z + c$ events ($\approx 25\%$), $Z +$ light flavour ($\lesssim 5\%$), and other backgrounds, such as $t\bar{t}$ and diboson production ($\approx 5\%$).

### 4.1.2 Selection in the $D^{\pm}$ mode

Event candidates in the $D^{\pm}$ mode are selected by looking for secondary vertices made of three tracks and with a reconstructed invariant mass consistent with the $D^{\pm}$ mass: $1869.5 \pm 0.4$ MeV [51]. The sum of the charges of the tracks participating in the secondary vertex must be $\pm 1$. The kaon mass is assigned to the track with opposite sign to the total charge of the three-prong vertex, and 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 [51].

The distribution of the reconstructed invariant mass for $D^{\pm}$ candidates associated with $Z \to e^+e^-$ (left) and $Z \to \mu^+\mu^-$ (right) is presented in Fig. 2. The signal and background contributions shown in the figure are estimated with the simulated samples. The charm fraction $B(c \to D^{\pm})$ in the PYTHIA simulation ($19.44 \pm 0.02\%$) is lower than the value ($22.7 \pm 0.9 \pm 0.5\%$) obtained from a combination [52] of published measurements performed at LEP [53–55] and the branching fraction of the decay $D^{\pm} \to K^\mp\pi^\pm\pi^\pm$ ($7.96 \pm 0.03\%$), is also lower than the PDG value ($10.13 \pm 0.19\%$) [51]; predicted event rates from the MC simulation are scaled in order to match the experimental charm fractions.

The signal region is defined by the constraint $|\Delta m(D^{\pm})| \equiv |m^{\text{rec}}(D^{\pm}) - 1.87\text{ GeV}| < 0.05\text{ GeV}$, where $m^{\text{rec}}(D^{\pm})$ is the reconstructed mass of the $D^{\pm}$ meson candidate. The mass range of the signal region is indicated in Fig. 2 as two dashed, vertical lines. The width of the signal region approximately corresponds to three times the measured mass resolution. The nonresonant background is subtracted from the event count in the signal window, and is estimated using the number of events selected in a control region away from the resonance, extending up to a window of 0.1 GeV width, $N[0.05 < \Delta m(D^{\pm}) < 0.10\text{ GeV}]$, as also shown in Fig. 2.

The number of selected events in data after background subtraction is $375 \pm 44$ in the $Z \to e^+e^-$ channel and $490 \pm 48$ in the $Z \to \mu^+\mu^-$ channel. Based on the simulation, the selected sample is enriched in $Z + c$ events ($\approx 60\%$), while the fraction of $Z + b$ events is $\approx 35\%$. The contribution from $Z +$ light flavour events is negligible, and the contribution of $t\bar{t}$ and diboson events is smaller than 5%.

### 4.1.3 Selection in the $D^*(2010)^{\pm}$ mode

Events with $Z +$ jets candidates in the $D^*(2010)^{\pm}$ mode are selected by requiring a displaced vertex with two oppositely charged tracks among the tracks constituting the jet. These tracks are assumed to be the decay products of a $D^0$ meson. The candidate is combined with a third track from the jet constituents that should represent the soft pion, emitted in the strong decay $D^*(2010)^+ \to D^0\pi^+$. To be a soft pion candidate, the track must have a transverse momentum larger than 0.5 GeV and lie in a cone of radius $\Delta R(D^0, \pi) = 0.1$ around the line of flight of the $D^0$ meson candidate.

The track of the $D^0$ meson candidate with a charge opposite to the charge of the soft pion is taken to be the kaon from the $D^0$ meson decay and is required to have $p_T > 1.75\text{ GeV}$. The other track is assigned to be the pion and is required...
to have $p_T > 0.75$ GeV. Two-track combinations with an invariant mass different from the nominal $D^0$ meson mass ($1864.86 \pm 0.13$ MeV) by less than 100 MeV are selected, and a secondary vertex is constructed using the two tracks and the CMS Kalman vertex fitter algorithm [56]. The two-track system is kept as a valid $D^0$ meson candidate if the probability for the vertex fit is greater than 0.05. To ensure a clean separation between the secondary and primary vertices, the 2D-distance in the transverse plane between them, divided by the uncertainty in the distance measurement (defined as decay length significance) has to be larger than 3. Furthermore, to guarantee that the reconstruction of the angle between the two directions has to be larger than $3$. The cosine measurement (defined as decay length significance) has to exceed a threshold.

The signal region is defined by the constraint $\Delta m(D^*(2010)^{\pm}) \equiv |m_{rec}(D^*(2010)^{\pm}) - 2.01$ GeV $| < 0.04$ GeV, where $m_{rec}(D^*(2010)^{\pm})$ is the reconstructed mass of the $D^*(2010)^{\pm}$ candidate, and which corresponds to slightly more than twice the measured mass resolution. The two dashed, vertical lines present in Fig. 3 indicate the mass range of the signal region. The nonresonant background contribution to the signal region is subtracted using the number of events selected in a control region away from the resonance. We use a window of 0.12 $(2 \times 0.06)$ GeV width, $N[0.04 < \Delta m(D^*(2010)^{\pm}) < 0.10$ GeV $]$, also shown in Fig. 3, and apply the proper weight to account for the different width of the signal and control regions $(8/12)$. The number of data selected events after background subtraction is $234 \pm 22$ in the $Z \rightarrow e^+e^-$ channel and $308 \pm 24$ in the $Z \rightarrow \mu^+\mu^-$ channel. According to the predictions obtained from the simulated samples, the fraction of $Z+c$ events in the selected sample is high ($\approx 65\%$) and the contribution of $Z+b$ events is $\approx 30\%$. No contribution is expected from $Z+\text{light flavour}$ events. Less than 5% of the selected events arise from $t\bar{t}$ and diboson production.

Systematic biases due to the background subtraction are expected to be negligible compared to the statistical uncertainty, because of the approximate agreement observed between data and simulation as shown in Figs. 2 and 3.

4.2 Selection of $W$+charm jet events ($c$ jet control sample)

Additional data and simulated samples consist of events from associated production of a $W$ boson and a jet originating from a $c$ quark ($W+c$). They are used to model characteristic distributions of jets with $c$ quark content and to measure the $c$ quark.
tagging efficiency in a large, independent sample. Jet flavour assignment in the simulated W + jets events follows the criteria presented in Sect. 3 for Z + jets events.

The production of a W boson in association with a c quark proceeds at LO via the processes qg → W−+c and qg → W+ (q = s, d). A key property of the qg → W+c reaction is the presence of a charm quark and a W boson with opposite-sign (OS) charges. Background processes deliver OS and same-sign (SS) events in equal proportions, whereas qg → W+c is always OS. Therefore, distributions obtained after OS—SS subtraction are representative of the W+c component, allowing for detailed studies of c jets.

We select W+c events following the criteria of the analysis reported in Ref. [57]. Candidate events are selected online using single-lepton triggers, which require at least one isolated electron (muon) with pT > 27 (24) GeV and |η| < 2.1. The lepton identification and isolation criteria are very similar to those used for the Z+jets selection. The offline pT threshold is increased to 30 (25) GeV for electrons (muons) because of the higher thresholds of the single-lepton triggers. The transverse invariant mass of the lepton and pTmiss system is defined as M_{T} = \sqrt{2 p_{T}^{\ell} E_{T}^{\text{miss}} [1 - \cos(\phi^{\ell} - \phi^{E_{T}^{\text{miss}}})]}, where \phi^{\ell} and \phi^{E_{T}^{\text{miss}}} are the azimuthal angles of the lepton momentum and pTmiss. The M_{T} must be larger than 55 (50) GeV for events in the W → eν (W → μν) channel.

Identification of jets originating from c quarks proceeds exactly as described in Sect. 4.1. In all cases the charge of the c quark is unequivocally known. In the semileptonic mode the charge of the muon determines the charge of the c quark. In the D± and D*(2010)± modes the charge of the D candidates defines the charge of the c quark. OS events are events when the muon, D±, or D*(2010)± candidate has a charge opposite to the lepton from the W boson decay, and SS events when the charge is the same.

Based on the simulations, after subtracting the SS from the OS samples, W + c events are the dominant contributor to the distributions; ≈ 90% in the semileptonic decay modes and larger than 98% in the D± and D*(2010)± exclusive channels. The remaining backgrounds, mainly from top quark production, are subtracted using the simulation.

4.3 Selection of t\bar{t} samples

A sample of t\bar{t} events (eμ-\text{t\bar{t}} sample) is selected using the leptonic decay modes of the W bosons from the t\bar{t} pair when they decay into leptons of different flavour. The t\bar{t} production is a natural source of b flavoured jets and enables tests of the MC description of the relevant distributions for b jets as well as of the performance of the secondary vertexing method. This sample is also used to model the t\bar{t} background in the discriminant variables used to extract the signal yields.

An e\mu-\text{t\bar{t}} sample is selected online by a trigger path based on the presence of an electron-muon pair. The offline selection proceeds as for the Z+HF jet events, but the two leptons must be different flavours. After the selection, contributions from processes other than t\bar{t} production are negligible.

An additional t\bar{t} enriched sample is used to estimate the normalization of the remaining t\bar{t} background. The same selection used for the Z+HF jet signal is applied: two leptons of the same flavour, e\mu or μμ, and E_T^{\text{miss}} > 80 GeV, instead...
of $E_T^{\text{miss}} < 40\text{ GeV}$. The small contribution from $Z + \text{jets}$
events in these samples ($\lesssim 3\%$) is subtracted according to its MC expectation.

5 Measurement of the c and b quark tagging efficiencies

The accuracy of the description in the MC simulations of the secondary vertex reconstruction part of the c tagging method is evaluated with a control sample of $W + c$ events with a well-identified $\mu + \text{jet}$ decay mode except for Drell–
Yan events. The contamination from the Drell–Yan process, which yields genuine OS dimuon events may reach 25%. The $W + c$ sample in the $W \rightarrow \mu \nu$ decay mode, with the lepton from the W decay of different flavour from the $\muon-inside-a-jet$, is not affected by this background and is employed for the c tagging study.

A $W + c$ event is “SV-tagged” if there is a reconstructed secondary vertex in the jet and the $\muon-inside-a-jet$ is one of the tracks used to form the vertex. The c jet tagging efficiency is the fraction of “SV-tagged” $W + c$ events, over all $W + c$ events, after OS – SS subtraction:

$$
\epsilon_c = \frac{N(W + c)_{\text{OS–SS}}(\text{SV-tagged})}{N(W + c)_{\text{OS–SS}}}.
$$

Efficiencies are obtained independently with the data and with the W+jets simulated samples. Data-to-simulation scale factors, $S\epsilon_c$, are then computed as the ratio between the c jet tagging efficiencies in data and simulation,

$$
S\epsilon_c = \frac{\epsilon_{c,\text{data}}}{\epsilon_{c,\text{MC}}}.
$$

They are used to correct the simulation efficiency.

The c jet tagging efficiencies and the scale factors are computed both inclusively and as a function of the jet $p_T$. The expected average c tagging efficiency is $\approx 33\%$ for the IVF algorithm and $\approx 21\%$ for the SSV algorithm. The c tagging efficiency ranges from 24% for the IVF algorithm (15% for the SSV algorithm) for $p_T^{\text{jets}}$ of 25–30 GeV and up to 37% (26%) for $p_T^{\text{jets}}$ of $\approx 100\text{ GeV}$. The $S\epsilon_c$ for jets with a $p_T$ larger than 25 GeV is found to be 0.93 ± 0.03 (stat) ± 0.02 (syst) for IVF vertices. It is 0.92 ± 0.03 (stat) ± 0.02 (syst) for SSV vertices. The systematic uncertainty accounts for inaccuracies in pileup description, jet energy scale and resolution, lepton efficiencies, background subtraction, and modelling of charm production and decay fractions in the simulation.

Detailed studies of the behaviour of the b tagging methods developed in CMS are available in Ref. [58]. Following the same procedure, we have used the $e\mu$-tt sample to investigate the data-to-MC agreement for the b tagging methods in this analysis. The b tagging efficiencies in data and simulated events are computed as the fraction of $e\mu$-tt events with a muon-inside-a-jet participating in a secondary vertex with respect to the number of events when the secondary vertex condition is released. The $S\epsilon_b = \epsilon_{b,\text{data}} / \epsilon_{b,\text{MC}}$ is measured to be $0.96 \pm 0.03$ for both IVF and SSV vertices, where the uncertainty includes statistical and systematic effects due to the jet energy scale and resolution and the pileup.

6 Analysis strategy

The extraction of $Z + c$ and $Z + b$ event yields is based on template fits to distributions of variables sensitive to the jet flavour. In the semileptonic mode we use the corrected invariant mass, $M_{\text{corr}}^{\text{vertex}}$ (corrected secondary-vertex mass), of the charged particles attached to the secondary vertex (the muon-inside-a-jet included). All charged particles are assigned the mass of the pion, except for the identified muon. A correction is included to account for additional particles, either charged or neutral, that may have been produced in the semileptonic decay but were not reconstructed [59],

$$
M_{\text{vertex}}^{\text{corr}} = \sqrt{M_{\text{vertex}}^2 + p_{\text{vertex}}^2 \sin^2 \theta + p_{\text{vertex}} \sin \theta},
$$

where $M_{\text{vertex}}$ and $p_{\text{vertex}}$ are the invariant mass and modulus of the vectorial sum of the momenta of all reconstructed particles associated to the secondary vertex, and $\theta$ is the angle between the momentum vector sum and the vector from the primary to the secondary vertex.

In the $D^{\pm}$ and $D^*(2010)^{\pm}$ modes a likelihood estimate of the probability that the jet tracks come from the primary vertex, called jet probability (JP) discriminant [47], is used.

The shapes of the Z+c discriminant distributions are modelled in data using OS $W + c$ events, after subtraction of the SS $W + c$ distributions. It is checked using simulated events that the corresponding distributions obtained from the $W + c$ samples accurately describe the $Z + c$ distributions. The main features of the jets, such as $p_T$, $\eta$, jet charged multiplicity, and the number of secondary vertices are found to be consistent between $Z + c$ and $W + c$ simulated samples and are in agreement with the observed distributions in the sample of $W + c$ events in data. Figure 4 (left) shows the simulated $p_T^{\text{jets}}$ distributions of $W + c$ and $Z + c$ events compared to $W + c$ data after OS – SS subtraction. The number of secondary vertices, identified with the IVF algorithm, is shown in Fig. 4 (right). Events with no reconstructed IVF vertices

Springer
have at least one reconstructed vertex with the SSV vertex algorithm. All distributions in Fig. 4 are normalized to unity.

The corrected secondary-vertex mass and JP discriminant distributions, normalized to unity, are presented in Fig. 5 for the three analysis categories. The simulated W+c and Z+c distributions are compared to W+c data. In general, the simulated W+c and Z+c distributions agree with the W+c data in all categories. A noticeable discrepancy is observed between the simulated and measured distributions of the corrected secondary-vertex mass in W+c events as shown in Fig. 5 (left). This difference is due to a different fraction of events with two- and three-track vertices in data and in the simulation. Studies with simulated events demonstrate that the fraction of events with two- and three-track vertices for W+c and Z+c production is the same. Therefore, we assume that the W+c corrected secondary-vertex mass distribution measured in data properly reproduces the same distribution for the Z+c measured events. The distributions obtained in the electron and muon decay channels are consistent and are averaged to obtain the final templates, thereby decreasing the associated statistical uncertainty.

The shape of the discriminant variables for Z+b events is modelled with the simulated samples. The simulated distribution of the corrected secondary-vertex mass is validated with the sample of $e\mu$-tt events as shown in Fig. 6. The simulation describes the data well, apart from the mass regions 3–4 GeV
and above 7.5 GeV. The observed differences, ≈ 13% in the 3–4 GeV mass region and ≈ 50% above 7.5 GeV, are used to correct the simulated Z+b distribution. However, the number of events in the eμ-τf sample does not allow a validation of the shape of JP discriminant distributions for Z+b events in the exclusive channels.

The distributions of the discriminant variables obtained in data are corrected by subtracting the contributions from the various background processes. They are estimated in the following way:

- The shapes of the discriminant distributions for t̅¯t production are evaluated with the eμ-τf sample. The normalization difference between same and different flavour combinations, \( N_{ee}^{Z+b} / N_{\mu\mu}^{Z+b} (N_{\mu\mu}^{Z+b} / N_{e\mu}^{Z+b}) \) is estimated from the sideband region \( E_{T}^{miss} > 80 \text{ GeV} \), and applied to the signal region \( E_{T}^{miss} < 40 \text{ GeV} \).
- The shape and normalization of the corrected secondary-vertex mass distribution for the Z+ light flavoured quark background in the semileptonic channel are evaluated with the simulated samples. Discrepancies between data and simulation in the rate of Z+ light flavour jet misidentification are corrected by applying the appropriate scale factors to the simulation [58]. No background from the Z+ light flavour quark process is expected in the exclusive channels.
- The shapes and normalization of the discriminant distributions for the remaining background from diboson production are taken from simulation.

The yields of Z+c and Z+b events in data are estimated by performing least squares fits between the background-subtracted data and template distributions. Independent fits are performed in the dielectron and dimuon channels and in the three analysis categories. The expected Z+c and Z+b distributions are fitted to data with scaling factors \( \mu_{Z+c} \) and \( \mu_{Z+b} \) defined with respect to the initial normalization predicted from simulation as free parameters of the fit. Typical values of the scaling factors are in the range 0.95–1.05 with a correlation coefficient between \( \mu_{Z+c} \) and \( \mu_{Z+b} \) of the order of −0.4. The scaling factor obtained for the Z+b component is consistent with that reported in Ref. [17] for a similar fiducial region. The fitted \( \mu_{Z+c} \) and \( \mu_{Z+b} \) are applied to the expected yields to obtain the measured ones in the data. The measured yields are summarized in Table 1.

Figure 7 shows the background-subtracted distributions of the corrected secondary-vertex mass for the Z+jets events with a muon-inside-a-jet associated with a secondary vertex. The corrected secondary-vertex mass tends to be larger for Z+b than for Z+c events because the larger mass of the b quark gives rise to heavier hadrons \( (m_{hadrons} \approx 5 \text{ GeV}, m_{b} \approx 2 \text{ GeV}) \).

The JP discriminant takes lower values for Z+c events than for Z+b events. The \( D^\pm \) or \( D^{*}(2010)^\pm \) mesons in Z+b events are “secondary” particles, i.e. they do not originate from the hadronization of a c quark produced at the primary vertex, but are decay products of previous b hadron decays at unobserved secondary vertices. Figure 8 shows the background-subtracted distribution of the JP discriminant for the Z+jets events with a \( D^\pm \rightarrow K^\pm \pi^\pm \pi^\pm \) candidate. Two bins are used to model the JP discriminant in this channel; as a result, the determination of the scaling factors \( \mu_{Z+c} \) and \( \mu_{Z+b} \) is reduced to solving a system of two equations with two unknowns.

Figure 9 presents the background-subtracted distribution of the JP discriminant for the Z+jets events with a \( D^{*}(2010)^\pm \) candidate. In this latter channel the particle identified as the soft pion in the \( D^{*}(2010)^\pm \rightarrow D^0 \pi^\pm \) decay is a true primary particle in the case of Z+c events, whereas it arises from a secondary decay (b hadron \( \rightarrow D^{*}(2010)^\pm + X \rightarrow D^0 \pi^\pm + X \)) for Z+b events. This “secondary” origin of the soft pion generates a distinctive dip in the first bin of the JP discriminant distribution for Z+b events.

7 Systematic uncertainties

Several sources of systematic uncertainties are identified, and their impact on the measurements is estimated by performing the signal extraction fit with the relevant parameters in the simulation varied up and down by their uncertainties. The effects are summarized in Fig. 10. The contributions from the various sources are combined into fewer categories for presentation in Fig. 10.
channels. The \( N_{Z+c}^{\text{signal}} \) and \( N_{Z+b}^{\text{signal}} \) are the yields of \( Z+c \) and \( Z+b \) events, respectively, extracted from the fit to the corrected secondary-vertex mass distributions. The factors \( C_i \) that correct the selection efficiencies are also given. They include the relevant branching fraction for the corresponding channel. All uncertainties quoted in the table are statistical, except for those of the measured cross sections and cross section ratios where the first uncertainty is statistical and the second is the estimated systematic uncertainty from the sources discussed in the text.

<table>
<thead>
<tr>
<th>Channel</th>
<th>( N_{Z+c}^{\text{signal}} )</th>
<th>( C_{Z+c} ) (%)</th>
<th>( \sigma(Z+c) B [\text{pb}] )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Semileptonic mode</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( Z \rightarrow e^+e^- )</td>
<td>1070 ± 100</td>
<td>0.63 ± 0.03</td>
<td>8.6 ± 0.8 ± 1.0</td>
</tr>
<tr>
<td>( Z \rightarrow \mu^+\mu^- )</td>
<td>1450 ± 140</td>
<td>0.81 ± 0.03</td>
<td>9.1 ± 0.9 ± 1.0</td>
</tr>
<tr>
<td>( Z \rightarrow \ell^+\ell^- )</td>
<td>( \sigma(Z+c) B = 8.8 \pm 0.6 \text{(stat)} \pm 1.0 \text{(syst)} \text{pb} )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Channel</th>
<th>( N_{Z+b}^{\text{signal}} )</th>
<th>( C_{Z+b} ) (%)</th>
<th>( \sigma(Z+c)/\sigma(Z+b) )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>D(\pm) mode</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( Z \rightarrow e^+e^- )</td>
<td>2610 ± 110</td>
<td>2.90 ± 0.08</td>
<td>1.9 ± 0.2 ± 0.2</td>
</tr>
<tr>
<td>( Z \rightarrow \mu^+\mu^- )</td>
<td>3240 ± 150</td>
<td>3.93 ± 0.10</td>
<td>2.2 ± 0.3 ± 0.2</td>
</tr>
<tr>
<td>( Z \rightarrow \ell^+\ell^- )</td>
<td>( \sigma(Z+c)/\sigma(Z+b) = 2.0 \pm 0.2 \text{(stat)} \pm 0.2 \text{(syst)} )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Channel</th>
<th>( N_{Z+c}^{\text{signal}} )</th>
<th>( C_{Z+c} ) (%)</th>
<th>( \sigma(Z+c) B [\text{pb}] )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>D(\ast(2010)\pm) mode</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( Z \rightarrow e^+e^- )</td>
<td>280 ± 60</td>
<td>0.13 ± 0.02</td>
<td>10.9 ± 2.2 ± 0.9</td>
</tr>
<tr>
<td>( Z \rightarrow \mu^+\mu^- )</td>
<td>320 ± 80</td>
<td>0.18 ± 0.02</td>
<td>8.8 ± 2.0 ± 0.8</td>
</tr>
<tr>
<td>( Z \rightarrow \ell^+\ell^- )</td>
<td>( \sigma(Z+c) B = 9.7 \pm 1.5 \text{(stat)} \pm 0.8 \text{(syst)} \text{pb} )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Channel</th>
<th>( N_{Z+c}^{\text{signal}} )</th>
<th>( C_{Z+c} ) (%)</th>
<th>( \sigma(Z+c) B [\text{pb}] )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Combination</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( Z \rightarrow \ell^+\ell^- )</td>
<td>( \sigma(Z+c) B = 8.8 \pm 0.5 \text{(stat)} \pm 0.6 \text{(syst)} \text{pb} )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 7 Corrected secondary-vertex mass distributions, after background subtraction, in the dielectron (left) and dimuon (right) channels for events selected in the semileptonic mode. Events with \( M_{\text{corr}} > 8 \text{GeV} \) are included in the last bin of the distribution. The shape of the \( Z+c \) and \( Z+b \) contributions is estimated as explained in the text. Their normalization is adjusted to the result of the signal extraction fit. Vertical bars on data points represent the statistical uncertainty in the data. The hatched areas represent the sum in quadrature of the statistical uncertainties of the templates describing the two contributions (\( Z+c \) from W+c data events and \( Z+b \) from simulation).

One of the main uncertainties is related to the charm fractions for the production and decay of c hadrons in the simulated samples and to the determination of the c tagging efficiency. The average of the inclusive charm quark semileptonic branching fractions is \( B(c \rightarrow \ell) = 0.096 \pm 0.004 \) [51], and the exclusive sum of the individual contributions from
Fig. 8 Background-subtracted distributions of the JP discriminant in the dielectron (left) and dimuon (right) channels for $Z + \text{jets}$ events with a $D^\pm \rightarrow K\pi\pi\pi^\pm$ candidate. The shape of the $Z + c$ and $Z + b$ contributions is estimated as explained in the text. Their normalization is adjusted to the result of the signal extraction fit. Vertical bars on data points represent the statistical uncertainty in the data. The hatched areas represent the sum in quadrature of the statistical uncertainties of the templates describing the two contributions ($Z + c$ from $W + c$ data events and $Z + b$ from simulation).

Fig. 9 Background-subtracted distributions of the JP discriminant in the dielectron (left) and dimuon (right) channels for $Z + \text{jets}$ events with a $D^{*+}(2010)\pm \rightarrow D^0\pi^\pm \rightarrow K\pi\pi\pi^\pm$ candidate. The shape of the $Z + c$ and $Z + b$ contributions is estimated as explained in the text. Their normalization is adjusted to the result of the signal extraction fit. Vertical bars on data points represent the statistical uncertainty in the data. The hatched areas represent the sum in quadrature of the statistical uncertainties of the templates describing the two contributions ($Z + c$ from $W + c$ data events and $Z + b$ from simulation).

all weakly decaying charm hadrons is $0.086 \pm 0.004$ [51,52]. The average of these two values, $B(c \rightarrow \ell) = 0.091 \pm 0.003$, is consistent with the PYTHIA value used in our simulations (9.3%). We assign a 5% uncertainty in order to cover both central values within one standard deviation. The average of the inclusive b quark semileptonic branching fractions is $B(b \rightarrow \ell) = 0.1069 \pm 0.0022$ [51], which is consistent with the PYTHIA value used in our simulations (10.5%). The corresponding uncertainty of 2% is propagated. The 5% systematic uncertainty in $B(c \rightarrow \ell)$ is further propagated for the fraction of $Z + b$ events with a lepton in the final state through the decay chain $b \rightarrow c \rightarrow \ell$. Uncertainties in the branching ratios of other $b$ hadron decay modes with a lepton in the final state, such as $b$ hadron $\rightarrow \tau (\rightarrow \ell + X) + X'$, $b$ hadron $\rightarrow J/\psi (\rightarrow \ell^+\ell^-) + X$, are not included since the expected contribution to the selected sample is negligible.

Since the simulation in the $D^\pm$ and $D^{*+}(2010)^\pm$ modes is reweighted to match the experimental values [52], the uncer-
PF objects to the reconstructed $E_\text{m}\text{estimated by propagating the jet energy scale uncertainties and}
ment of the missing transverse energy in the event is esti-
CMS studies [42,43]. The uncertainty from a mismeasure-
within their uncertainties according to the results of dedicated
jet energy scale (jet energy resolution) correction factors
jet energy resolution are assessed by varying the correspond-

Fig. 10 Contributions to the systematic uncertainty in the measured
$Z+c$ cross section and in the $(Z+c)/(Z+b)$ cross section ratio. The
first three blocks in the graph show the uncertainties in the $Z+c$ cross
section in the three decay modes, semileptonic (SL), $D^*$, and $D^0(2010)^\pm$,
calculated from the combination of the dimuon and dielectron $Z$ boson
decay channels. The fourth block shows the systematic uncertainties in the
combined (Comb.) $Z+c$ cross section. The last block presents the sys-
tematic uncertainty in the $(Z+c)/(Z+b)$ cross section ratio measured in
the semileptonic mode. For every block, the height of the hatched
bars indicates the contribution from the different sources of systematic
uncertainty. The last, solid bar shows their sum in quadrature
difference
in the reweighting factors (5% for $D^\pm$ and 3.2% for
$D^*(2010)^\pm$) is propagated to the cross section.

The contribution from gluon splitting processes to $Z+c$
production in the phase space of the measurement is small,
and its possible mismodelling has little impact on the
measurements. Its effect is evaluated with the simulated sample
by independently increasing the weight of the events with at
least two $c$ ($b$) quarks in the list of generated particles close
to the selected jet ($\Delta R(\text{jet, } c(b)) < 0.5$) by three times the
experimental uncertainty in the gluon splitting rate into $c\bar{c}$,
$b\bar{b}$ quark pairs [60,61].

The effects of the uncertainty in the jet energy scale and
jet energy resolution are assessed by varying the correspon-
ding jet energy scale (jet energy resolution) correction factors
within their uncertainties according to the results of dedicated
CMS studies [42,43]. The uncertainty from a mismeasure-
ment of the missing transverse energy in the event is esti-
mated by propagating the jet energy scale uncertainties and
by adding 10% of the energy unassociated with reconstructed
PF objects to the reconstructed $E_T^{\text{miss}}$

The uncertainty in the $c$ tagging scale factors is in the range
3.5–4%, and it is around 2.5% for the $b$ tagging efficiency. In
the $D^*(2010)^\pm$ mode, the candidate reconstruction procedure
is repeated by independently changing by one standard devia-
tion, in terms of the $p_T$ resolution, the different $p_T$-thresholds
imposed and the decay length significance requirement. We
assume the uncertainty is the quadratic sum of the respective
differences between data and simulation in the change of the
number of $D^*(2010)^\pm$ candidates (2.8%).

The uncertainty in the lepton efficiency correction factors
is 4% in the $Z \rightarrow e^+e^-$ and 2% in the $Z \rightarrow \mu^+\mu^-$ chan-
nels. The uncertainty in the efficiency for the identification
of muons inside jets is approximately 3%, according to ded-
icated studies in multijet events [20].

An additional systematic uncertainty is assigned to account
for a possible mismodelling of the subtracted backgrounds.
For the $t\bar{t}$ background the uncertainty is taken as the differ-
ence between the estimate based on data, as described in
Sect. 6, and the one based on simulation. For $Z+$light flavour
events, the systematic uncertainty is evaluated by using the
MC correction factors associated with different misidentifi-
cation probabilities. Finally, the diboson contribution is var-
ied by the difference between the theoretical cross sections
calculated at NNLO and NLO ($\approx 15\%$) [62–64].

The reference signal simulated sample is generated with
MadGraph $+\text{Pythia6}$ using the PDF CTEQ6L1 and
rewighted to NNLO PDF set MSTW2008NNLO. The differ-
ence resulting from using other NNLO PDF sets is small
($\lesssim 1\%$). Following the prescription of the PDF groups, the
PDF uncertainty is of the same order.

The shapes of the discriminant distributions obtained from
the $W+c$ event sample are observed to be very stable. Changes
in the jet energy scale and variations in the $p_T$ threshold
imposed to select $W$ boson candidates do not affect the shape
of the templates. The correction factors applied in certain
regions to the corrected secondary-vertex mass template for
$Z+b$ events are varied within their uncertainties.

Uncertainties due to the pileup modelling are calculated
using a modified pileup profile obtained with a pp inelastic
cross section changed by its estimated uncertainty, 6%. The
uncertainty in the determination of the integrated luminosity
of the data sample is 2.6% [65].

Systematic uncertainties in the differential $Z+c$ cross
section and in the $(Z+c)/(Z+b)$ cross section ratio are in
the range 11–15%. The main sources of systematic uncer-
tainty in the differential distributions are due to the jet energy
scale determination, the charm fractions for $c$ hadron produc-
tion and decay in simulation, and the efficiencies of heavy
flavour tagging. The uncertainty in the binned $c$ tagging effi-
ciency scaling factors is 7–8%. Uncertainties in the $b$ tagging
efficiencies are 3–5%. An additional source of systematic
uncertainty in the differential measurement as a function of
the transverse momentum of the jet arises from the statisti-
cal uncertainty in the determination of the response matrix
used to correct for migration of events across $p_T^{\text{jet}}$ bins, as
described in Sect. 9. Its impact is evaluated by repeating
the correction procedure using a large number of response mat-
trices, built from the nominal one by varying its components
according to their statistical uncertainties. The effect is in the
range 4–6% for the $Z + c$ cross section and 4.5–7% for the $(Z + c)/(Z + b)$ cross section ratio.

8 Inclusive $Z + c$ cross section and $(Z + c)/(Z + b)$ cross section ratio

For all channels under study, the $Z + c$ cross section is determined in the fiducial region $p_T^j > 20$ GeV, $|\eta^j| < 2.1$, $71 < m_{\ell\ell} < 111$ GeV, $p_T^{\text{jet}} > 25$ GeV, $|\eta^{\text{jet}}| < 2.5$, and $\Delta R(\text{jet}, \ell) > 0.5$, using the following expression:

$$\sigma(Z + c) B = \frac{N_{Zc}^{\text{signal}}}{C \mathcal{L}},$$

where $N_{Zc}^{\text{signal}}$ is the fitted yield of $Z + c$ events and $\mathcal{L}$ is the integrated luminosity. The factor $C$ corrects for event losses in the selection process and is estimated using simulated events. The $C$ factors also include the relevant branching fraction for the corresponding channel.

Similarly, the ratio of cross sections $\frac{\sigma(Z + c)}{\sigma(Z + b)}$ is calculated in the same fiducial region applying the previous expression also for the $Z + b$ contribution:

$$\frac{\sigma(Z + c)}{\sigma(Z + b)} = \frac{N_{Zc}^{\text{signal}}}{N_{Zb}^{\text{signal}}} \frac{C(Z + b)}{C(Z + c)},$$

Table 1 shows the $Z + c$ production cross section obtained in the three modes and the $(Z + c)/(Z + b)$ cross section ratio (semileptonic mode only).

For the three categories of this analysis the $Z + c$ cross sections obtained in the dielectron and dimuon $Z$ boson decay channels are consistent. The results obtained in the three analysis categories are also consistent. Several combinations are performed to improve the precision of the measurement taking into account statistical and systematic uncertainties of the individual measurements. Systematic uncertainties arising from a common source and affecting several measurements are considered as fully correlated. In particular, all systematic uncertainties are assumed fully correlated between the electron and muon channels, except those related to lepton reconstruction. The average $Z + c$ cross sections obtained in the three categories, together with the combination of the six measurements, are also presented in Table 1. The combination is dominated by the result in the semileptonic mode. The contribution of the $D^0(2010)^{+}\bar{\tau}$ mode to the average is also significant despite the limited size of the selected samples.

The cross section ratio $\frac{\sigma(Z + c)}{\sigma(Z + b)}$ has been measured in the semileptonic mode, in the two $Z$ boson decay channels, and the results among them are consistent. Both cross section ratios are combined taking into account the statistical and systematic uncertainties in the two channels, and the correlations among them. The combination is given in Table 1.

The measured $Z + c$ cross section and the $(Z + c)/(Z + b)$ cross section ratio are compared to theoretical predictions obtained using two MC event generators and the MCFM program.

A prediction of the $Z + c$ fiducial cross section is obtained with the MADGRAPH sample. It is estimated by applying the phase space definition requirements to generator level quantities: two leptons from the $Z$ boson decay with $p_T^\ell > 20$ GeV, $|\eta^\ell| < 2.1$, and dilepton invariant mass in the range $71 < m_{\ell\ell} < 111$ GeV; a generator-level $c$ jet with $p_T^{c,\text{jet}} > 25$ GeV, $|\eta^{c,\text{jet}}| < 2.5$ and separated from the leptons by a distance $\Delta R(c\text{ jet}, \ell) > 0.5$. A prediction of the $Z + b$ cross section, and hence of the $(Z + c)/(Z + b)$ cross section ratio, is similarly derived applying the relevant phase space definition requirements to $b$ flavoured generator-level jets.

The MADGRAPH prediction, $\sigma(Z + c) B = 8.14 \pm 0.03$ (stat) $\pm 0.25$ (PDF) pb, is in agreement with the measured value. The quoted PDF uncertainty corresponds to the largest difference in the predictions obtained using the central members of two different PDF sets (MSTW2008 vs NNPDF2.3); uncertainties computed using their respective PDF error sets are about half this value.

We have also compared the measurements with predictions obtained with a sample of events generated with MADGRAPH5_aMC@NLO v2.2.1 [66] (hereafter denoted as MG5_aMC) generator interfaced with PYTHIA v8.212 [67] using the CUETP8M1 tune [68] for parton showering and hadronization. The matrix element calculation includes the $Z$ boson production process with 0, 1, and 2 partons at NLO. The FxFx [69] merging scheme between jets from matrix element and parton showers is implemented with a merging scale parameter set to 20 GeV. The NNPDF3.0 PDF set [70] is used for the matrix element calculation, while the NNPDF2.3 LO is used for the showering and hadronization.

The MG5_aMC prediction of the $Z + c$ cross section is slightly higher, $\sigma(Z + c) B = 9.46 \pm 0.04$ (stat) $\pm 0.15$ (PDF) $\pm 0.50$ (scales) pb, but still in agreement with the measurement. Uncertainties in the prediction are evaluated using the reweighting features implemented in the generator [71]. The quoted PDF uncertainty corresponds to the standard deviation of the predictions obtained using the one hundred replicas in the NNPDF3.0 PDF set. The scale uncertainty is the envelope of the predictions when the factorization and renormalization scales are varied by a factor of two or one half independently, always keeping the ratio between them less than or equal to two.

Theoretical predictions in perturbative quantum chromodynamics at NLO for the associated production of a $Z$ boson and at least one $c$ quark are obtained with the MCFM 7.0 program [72]. Several sets of NLO PDF sets are used, accessed
through the LHAPDF6 [73] library interface. Partons are clustered into jets using the anti-\(k_T\) algorithm with a distance parameter of 0.5. The kinematic requirements follow the experimental selection: the two leptons from the Z boson decay with \(p_T > 20\text{GeV}, |\eta| < 2.1, 71 < m_{\ell\ell} < 111\text{GeV}\) and a c parton jet with \(p_T^{\text{parton jet}} > 25\text{GeV}, |\eta^{\text{parton jet}}| < 2.5\), and separated from the leptons by \(\Delta R(\text{parton jet}, \ell) > 0.5\). The factorization and renormalization scales are set to the mass of the Z boson. The PDF uncertainty in the predictions is evaluated following the prescription recommended by the individual PDF groups; the scale uncertainty is estimated as the envelope of the results with (twice, half) factorization and renormalization scales variations.

The prediction computed with MCFM follows the calculation reported in Refs. [72,74]. The leading contribution gc \(\rightarrow Zc\) is evaluated at NLO including virtual and real corrections. Some of these corrections feature two jets in the final state, one of them with heavy flavour quark content. The calculation also includes the process q\overline{q} \(\rightarrow Zc\) evaluated at LO, where either one of the heavy flavour quarks escapes detection or the two of them coalesce into a single jet.

The MCFM prediction, which is a parton-level calculation, is corrected for hadronization effects so it can be compared with the particle-level measurements reported in this paper. The correction factor is computed with the MadGraph simulated sample comparing the predicted cross section using generator-level jets and parton jets. Parton jets are defined using the same anti-\(k_T\) clustering algorithm with a distance parameter of 0.5, applied to all quarks and gluons after showering, but before hadronization. The flavour assignment for parton jets follows similar criteria as for generator-level jets: a parton jet is labelled as a b jet if there is at least a b quark among its constituents, regardless of the presence of any c or light quarks. It is classified as c jet if there is at least a c quark, and no b quark, among the constituents, and light otherwise. The size of the correction is \(\approx 10\%\) for Z+c and \(\approx 15\%\) for Z+b cross sections, in good agreement with the estimation in Ref. [75].

After the hadronization correction the MCFM prediction still misses contributions from the parton shower evolution, underlying event, and multiple parton interactions. An approximate value of the total correction due to these processes and hadronization is estimated using MadGraph and amounts to \(\approx 30\%\). This correction is not applied to MCFM predictions, but can explain the observed differences between MCFM and the predictions of other generators.

Predictions are produced using MSTW08 and CT10 PDF sets and a recent PDF set from the NNPDF Collaboration, NNPDF3IC [76], where the charm quark PDF is no longer assumed to be perturbatively generated through pair production from gluons and light quarks, but is parameterized and determined along with the light quark and gluon PDFs. The PDF set where the charm quark PDF is generated perturbatively, NNPDF3nIC [76], is also used.

No differences in the predictions are observed using either NNPDF3IC or NNPDF3nIC PDF sets. Differences among them start to be sizeable when the transverse momentum of the Z boson is \(\gtrsim 100\text{GeV}\) [76]. The largest prediction is obtained using the MSTW08 PDF set, \(\sigma(\text{Z} + c)B = 5.32 \pm 0.01\text{ (stat)} \pm 0.06\text{ (PDF)} +0.34 -0.35\text{ (scales)}\) pb. Predictions obtained using CT10 and NNPDF3IC are \(5\%\) smaller than with MSTW08. The uncertainties in all the calculations are of the same order.

The MadGraph prediction for the \(\sigma(\text{Z} + c)/(\text{Z} + b)\) cross section ratio is \(1.781 \pm 0.006\text{ (stat)} \pm 0.004\text{ (PDF)}\), where the PDF uncertainty reflects the largest variation using the various PDF sets. The expectation from MG5_aMC is \(1.84 \pm 0.01\text{ (stat)} \pm 0.07\text{ (scales)}\). The uncertainties from the several members within one PDF set essentially vanish in the ratio. Both predictions agree with the measured ratio.

A prediction for the cross section ratio is also obtained with MCFM, as the ratio of the predictions for \(\sigma(\text{Z} + c)\) and \(\sigma(\text{Z} + b)\), using the same parameters emulating the experimental scenario for both processes. The calculation of the \(\sigma(\text{Z} + b)\) cross section follows the same reference as \(\sigma(\text{Z} + c)\) [72,74]. The highest predicted value is \(\sigma(\text{Z} + c)/\sigma(\text{Z} + b) = 1.58 \pm 0.07\text{ (stat+PDF syst)}\) obtained when the CT10 PDF set is used. The prediction from NNPDF3IC is about \(10\%\) lower, mainly because the predicted \(Z + b\) cross section using this PDF is the highest one.

9 Differential Z + c cross section and (Z + c)/(Z + b) cross section ratio

The \(Z + c\) production cross section and the \(\sigma(\text{Z} + c)/(\text{Z} + b)\) cross section ratio are measured differentially as a function of the transverse momentum of the Z boson, \(p_T^Z\), and of the transverse momentum of the HF jet with the sample selected in the semileptonic mode described in Sect. 4.1. The transverse momentum of the Z boson is reconstructed from the momenta of the two selected leptons. The sample is divided into three different subsamples according to the value of the variable of interest, \(p_T^Z\) or \(p_T^{HF}\), and the fit procedure is performed independently for each of them and for each Z boson decay mode. The number and size of the bins is chosen such that the corrected secondary-vertex mass distribution for each bin is sufficiently populated to perform the signal extraction fit.

Potential effects of event migration between neighbouring bins and inside/outside the acceptance due to the detector resolution are studied using simulated samples. A detector
response matrix is built with those events fulfilling the selection criteria both with generated and reconstructed variables. The element \((i, j)\) in the matrix determines the probability that an event with generated \(p_T^{\text{jet}}\) \((p_T^{\text{jet}})\) in bin \(i\) ends up reconstructed in bin \(j\) of the distribution.

Migration effects in \(p_T\) are found to be negligible and no correction is applied. An uncertainty of 1\%, which corresponds to the difference between the cross sections with and without corrections, is included in the systematic uncertainties.

Some migration of events between neighbouring bins in \(p_T^{\text{jet}}\) is expected because of the energy resolution, mainly between the first and second bins (<30\%), while migrations between the second and third bins are less than 10\%. Migration effects are expected to be the same in the two Z boson decay modes. The response matrix is used to unfold the fitted signal yields to actual signal yields at particle level. Events with a generated \(p_T^{\text{jet}}\) outside the fiducial region and reconstructed inside it because of resolution effects are subtracted prior to the unfolding procedure. Corrections are made for acceptance losses at the border of the kinematical region because of the detector resolution and reconstruction inefficiencies. The unfolding is performed with an analytical inversion of the matrix defining the event migrations. Statistical and systematic uncertainties are propagated through the unfolding procedure.

Tables 2 and 3 summarize the fitted \(Z+c\) and \(Z+b\) signal yields, the \(Z+c\) cross section, and the \((Z+c)/(Z+b)\) cross section ratio in the three \(p_T^{Z}\) and \(p_T^{\text{jet}}\) bins and in the two Z boson decay channels. The differential cross section and cross section ratio measured in the two Z boson decay channels are consistent and are combined to obtain the final results, taking into account the statistical and systematic uncertainties in the two channels and the correlations among them. The combined cross section and cross section ratio are presented in Table 4. They are also shown graphically in Fig. 11 in bins of \(p_T^{Z}\) (top) and \(p_T^{\text{jet}}\) (bottom).

Theoretical predictions for the differential cross section and cross section ratio are also obtained with the two MC generator programs and with MCFM. They are shown in Fig. 11 for comparison with the measured values. The uncertainties in the MadGraph predictions include the statistical and PDF uncertainties. Scale variations are additionally included in the uncertainties of MG5_aMC and MCFM. Predictions from MG5_aMC are higher than the predictions from MadGraph in the three bins of the \(Z+c\) differential distributions. A higher \((Z+c)/(Z+b)\) cross section ratio is predicted up to 60 GeV, although consistent within uncertainties. The predictions from MadGraph and MG5_aMC successfully reproduce the measurements. The level of agreement is similar in terms of the \(Z+c\) cross section and the \((Z+c)/(Z+b)\) cross section ratio.

A similar ordering appears in the differential cross sections and the inclusive cross sections for theoretical predictions calculated with MCFM and the various PDF sets. The highest \(Z+c\) cross section is predicted using the MSTW08 PDF set, the largest differential \((Z+c)/(Z+b)\) cross section ratio in the two variables is obtained with the CT10 PDF set. All MCFM predictions are lower than the differential cross section measurements as a function of \(p_T^{Z}\). This discrepancy is most pronounced in the first bin in \(p_T^{\text{jet}}\). Differences between predictions and data are reduced in the \((Z+c)/(Z+b)\) cross section ratio comparison.

The fitted charm PDF in NNPDF3IC [76] set is consistent with having an intrinsic component. The fitted fraction of the proton momentum that the charm quark component carries is \((0.7 \pm 0.3)\%\) if EMC data [77] is included in the fit and \((1.6 \pm 1.2)\%\) without it. After subtraction of the perturbative component, the momentum fraction of the proton carried by the IC component is \((0.5 \pm 0.3)\%\) if EMC data is included in the fit, or \((1.4 \pm 1.2)\%\) if not. Upper limits from the CTEQ-TEA Collaboration are also available [78, 79]. Quoted limits on the proton momentum fraction carried by the IC component vary between 1.5 and 2.5\% at 90\% confidence level depending on the parameterization used.

If the proton momentum fraction taken by the charm quark component (intrinsic + perturbative) is of order \(\approx 2\%\), an increase in the production of \(Z+c\) events with a \(p_T^{Z} \approx 100\) GeV of at least 20–25\% would be expected [76]. Should it be smaller than 1\%, the cross section increase would be limited in the \(p_T^{Z}\) region around 100–200 GeV and only become visible at significantly higher \(p_T^{Z}\) (>500 GeV). The measured cross section in the \(p_T^{Z}\) bin [60, 200] GeV is in agreement with predictions from MadGraph and MG5_aMC using a perturbative charm quark PDF. This measurement is in agreement with no increase in the production rate or with a very modest one, as expected from current upper limits on the IC component. No increase in the production rate in the highest \(p_T^{\text{jet}}\) bin is observed, either.

### 10 Summary

The associated production of a Z boson with at least one charm quark jet in proton-proton collisions at a centre-of-mass energy of 8 TeV was studied with a data sample corresponding to an integrated luminosity of \(19.7 \pm 0.5 \) fb\(^{-1}\). It was compared to the production of a Z boson with at least one b quark jet. Selection of event candidates relies on the identification of semileptonic decays of c or b hadrons with a muon in the final state and through the reconstruction of exclusive decay channels of D\(^+\) and D\(^0\) (2010)\(^\pm\) mesons. The Z boson is identified through its decay into an e\(^+\)e\(^-\) or \(\mu^+\mu^-\) pair.
Table 2 Differential cross section $d\sigma(Z+c)/dp_T^Z$ and cross section ratio $<d\sigma(Z+c)/dp_T^Z>/<d\sigma(Z+b)/dp_T^Z>$ in the semileptonic mode and in the two $Z$ boson decay channels. The $N_{Z+c}^\text{signal}$ and $N_{Z+b}^\text{signal}$ are the yields of $Z+c$ and $Z+b$ events, respectively, extracted from the fit. All uncertainties quoted in the table are statistical, except for those of the measured cross sections and cross section ratios, where the first uncertainty is statistical and the second is the estimated systematic uncertainty from the sources discussed in the text.

<table>
<thead>
<tr>
<th>Channel</th>
<th>$N_{Z+c}^\text{signal}$</th>
<th>$d\sigma(Z+c)/dp_T^Z$ B [pb]</th>
<th>$N_{Z+b}^\text{signal}$</th>
<th>$d\sigma(Z+b)/dp_T^Z$ B [pb]</th>
</tr>
</thead>
<tbody>
<tr>
<td>$0 &lt; p_T^Z &lt; 30$ GeV</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Z \rightarrow e^+ e^-$</td>
<td>$212 \pm 44$</td>
<td>$0.067 \pm 0.014 \pm 0.010$</td>
<td>$578 \pm 52$</td>
<td>$1.5 \pm 0.4 \pm 0.2$</td>
</tr>
<tr>
<td>$Z \rightarrow \mu^+ \mu^-$</td>
<td>$380 \pm 61$</td>
<td>$0.102 \pm 0.016 \pm 0.017$</td>
<td>$693 \pm 68$</td>
<td>$2.7 \pm 0.6 \pm 0.4$</td>
</tr>
<tr>
<td>$30 &lt; p_T^Z &lt; 60$ GeV</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Z \rightarrow e^+ e^-$</td>
<td>$501 \pm 60$</td>
<td>$0.144 \pm 0.017 \pm 0.019$</td>
<td>$1035 \pm 66$</td>
<td>$2.4 \pm 0.4 \pm 0.3$</td>
</tr>
<tr>
<td>$Z \rightarrow \mu^+ \mu^-$</td>
<td>$586 \pm 92$</td>
<td>$0.123 \pm 0.019 \pm 0.017$</td>
<td>$1422 \pm 87$</td>
<td>$1.9 \pm 0.4 \pm 0.3$</td>
</tr>
<tr>
<td>$60 &lt; p_T^Z &lt; 200$ GeV</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Z \rightarrow e^+ e^-$</td>
<td>$363 \pm 53$</td>
<td>$0.017 \pm 0.002 \pm 0.002$</td>
<td>$913 \pm 67$</td>
<td>$1.7 \pm 0.3 \pm 0.2$</td>
</tr>
<tr>
<td>$Z \rightarrow \mu^+ \mu^-$</td>
<td>$474 \pm 73$</td>
<td>$0.017 \pm 0.003 \pm 0.002$</td>
<td>$1056 \pm 81$</td>
<td>$2.0 \pm 0.4 \pm 0.3$</td>
</tr>
</tbody>
</table>

Table 3 Differential cross section $d\sigma(Z+c)/dp_T^{\text{jet}}$ and cross section ratio $(d\sigma(Z+c)/dp_T^{\text{jet}})/(d\sigma(Z+b)/dp_T^{\text{jet}})$ in the semileptonic mode and in the two $Z$ boson decay channels. The $N_{Z+c}^\text{signal}$ and $N_{Z+b}^\text{signal}$ are the yields of $Z+c$ and $Z+b$ events, respectively, extracted from the fit. All uncertainties quoted in the table are statistical, except for those of the measured cross sections and cross section ratios, where the first uncertainty is statistical and the second is the estimated systematic uncertainty from the sources discussed in the text.

<table>
<thead>
<tr>
<th>Channel</th>
<th>$N_{Z+c}^\text{signal}$</th>
<th>$d\sigma(Z+c)/dp_T^{\text{jet}}$ B [pb]</th>
<th>$N_{Z+b}^\text{signal}$</th>
<th>$d\sigma(Z+b)/dp_T^{\text{jet}}$ B [pb]</th>
</tr>
</thead>
<tbody>
<tr>
<td>$25 &lt; p_T^{\text{jet}} &lt; 40$ GeV</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Z \rightarrow e^+ e^-$</td>
<td>$476 \pm 58$</td>
<td>$0.342 \pm 0.048 \pm 0.041$</td>
<td>$1022 \pm 67$</td>
<td>$2.7 \pm 0.6 \pm 0.3$</td>
</tr>
<tr>
<td>$Z \rightarrow \mu^+ \mu^-$</td>
<td>$583 \pm 91$</td>
<td>$0.337 \pm 0.059 \pm 0.055$</td>
<td>$1393 \pm 90$</td>
<td>$2.4 \pm 0.6 \pm 0.4$</td>
</tr>
<tr>
<td>$40 &lt; p_T^{\text{jet}} &lt; 60$ GeV</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Z \rightarrow e^+ e^-$</td>
<td>$289 \pm 47$</td>
<td>$0.090 \pm 0.027 \pm 0.018$</td>
<td>$843 \pm 59$</td>
<td>$1.3 \pm 0.4 \pm 0.2$</td>
</tr>
<tr>
<td>$Z \rightarrow \mu^+ \mu^-$</td>
<td>$456 \pm 66$</td>
<td>$0.104 \pm 0.027 \pm 0.014$</td>
<td>$1044 \pm 75$</td>
<td>$1.9 \pm 0.5 \pm 0.3$</td>
</tr>
<tr>
<td>$60 &lt; p_T^{\text{jet}} &lt; 200$ GeV</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Z \rightarrow e^+ e^-$</td>
<td>$311 \pm 56$</td>
<td>$0.012 \pm 0.003 \pm 0.008$</td>
<td>$686 \pm 64$</td>
<td>$1.7 \pm 0.5 \pm 0.3$</td>
</tr>
<tr>
<td>$Z \rightarrow \mu^+ \mu^-$</td>
<td>$369 \pm 63$</td>
<td>$0.013 \pm 0.003 \pm 0.007$</td>
<td>$800 \pm 75$</td>
<td>$1.9 \pm 0.5 \pm 0.3$</td>
</tr>
</tbody>
</table>

Table 4 Differential $Z+c$ cross section and $(Z+c)/(Z+b)$ cross section ratio. The first column presents the $p_T$ range for each bin. Column 2 presents the cross section and column 3 the ratio. The differential measurements as a function of the transverse momentum of the $Z$ boson (jet with heavy flavour content) are given in the upper (lower) part of the table. The first uncertainty is statistical and the second is the systematic uncertainty arising from the sources discussed in the text.

<table>
<thead>
<tr>
<th>$[p_T^Z, min, p_T^Z, max]$</th>
<th>$d\sigma(Z+c)/dp_T^Z$ B [pb]</th>
<th>$d\sigma(Z+c)/dp_T^Z / d\sigma(Z+b)/dp_T^Z$</th>
</tr>
</thead>
<tbody>
<tr>
<td>[0, 30]</td>
<td>$0.077 \pm 0.011 \pm 0.011$</td>
<td>$1.7 \pm 0.3 \pm 0.2$</td>
</tr>
<tr>
<td>[30, 60]</td>
<td>$0.133 \pm 0.013 \pm 0.017$</td>
<td>$2.1 \pm 0.3 \pm 0.3$</td>
</tr>
<tr>
<td>[60, 200]</td>
<td>$0.017 \pm 0.002 \pm 0.002$</td>
<td>$1.8 \pm 0.3 \pm 0.2$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$[p_T^{\text{jet}}, min, p_T^{\text{jet}}, max]$</th>
<th>$d\sigma(Z+c)/dp_T^{\text{jet}}$ B [pb]</th>
<th>$d\sigma(Z+c)/dp_T^{\text{jet}} / d\sigma(Z+b)/dp_T^{\text{jet}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>[25, 40]</td>
<td>$0.341 \pm 0.037 \pm 0.040$</td>
<td>$2.5 \pm 0.4 \pm 0.3$</td>
</tr>
<tr>
<td>[40, 60]</td>
<td>$0.097 \pm 0.019 \pm 0.012$</td>
<td>$1.5 \pm 0.3 \pm 0.2$</td>
</tr>
<tr>
<td>[60, 200]</td>
<td>$0.013 \pm 0.002 \pm 0.002$</td>
<td>$1.8 \pm 0.4 \pm 0.2$</td>
</tr>
</tbody>
</table>
The cross section for the production of a $Z$ boson associated with at least one $c$ quark jet is measured. The measurement is performed in the kinematic region with two leptons with transverse momentum $p_T^\ell > 20 \text{ GeV}$, pseudorapidity $|\eta^\ell| < 2.1$, dilepton invariant mass $71 < m_{\ell\ell} < 111 \text{ GeV}$ and a jet with $p_T^{\text{jet}} > 25 \text{ GeV}$, $|\eta^{\text{jet}}| < 2.5$, separated from the leptons of the $Z$ boson candidate by a distance $\Delta R(\ell, \text{jet}) > 0.5$.

The $Z + c$ production cross sections measured in all the analysis categories are fully consistent, and the combined value is $\sigma(pp \rightarrow Z + c + X)B(Z \rightarrow \ell^+\ell^-) = 8.8 \pm 0.5 \text{ (stat) } \pm 0.6 \text{ (syst)} \text{ pb}$. This is the first measurement at the LHC of $Z + c$ production in the central pseudorapidity region.

The cross section ratio for the production of a $Z$ boson and at least one $c$ and at least one $b$ quark jet is measured in the same kinematic region and is $\sigma(pp \rightarrow Z + c + X)/\sigma(pp \rightarrow Z + b + X) = 2.0 \pm 0.2 \text{ (stat) } \pm 0.2 \text{ (syst)}$.

The size of the sample selected in the semileptonic channel allows for the first differential measurements of the $Z + c$ cross section at the LHC. The $Z + c$ cross section and $(Z + c)/(Z + b)$ cross section ratio are measured as a function of the transverse momentum of the $Z$ boson and of the heavy flavour jet.

The measurements are in agreement with the leading order predictions from MadGraph and next-to-leading-order predictions from MadGraph$_5$ aMC@NLO. Predictions from the MCFM program are lower than the measured $Z + c$ cross...
section and \((Z+c)/(Z+b)\) cross section ratio, both inclusively and differentially. This difference can be explained by the absence of parton shower development and nonperturbative effects in the MCFM calculation.

Measurements in the highest \(p_T^{Z} (p_T^{Z,jet})\) region analyzed, \(60 < p_T^{Z} (p_T^{Z,jet}) < 200\) GeV, would be sensitive to the existence of an intrinsic charm component inside the proton if this IC component were large enough to induce a significant enhancement in the \(Z+c\) production cross section. However, our measurements of the \(Z+c\) cross section and \((Z+c)/(Z+b)\) cross section ratio are consistent with predictions using PDF sets with no IC component.

Acknowledgements We congratulate our colleagues in the CERN accelerator departments for the excellent performance of the LHC and thank the technical and administrative staff of CERN and at other CMS institutes for their contributions to the success of the CMS effort. In addition, we gratefully acknowledge the computing centres and personnel of the Worldwide LHC Computing Grid for delivering so effectively the computing infrastructure essential to our analyses. Finally, we acknowledge the enduring support for the construction and operation of the LHC and the CMS detector provided by the following funding agencies: BMWFW and FWF (Austria); FNRS and FWO (Belgium); CNPq, CAPES, FAPERJ, and FAPESP (Brazil); MES (Bulgaria); CERN; CAS, MoST, and NSFC (China); COCENCIAS (Colombia); MSES and CSF (Croatia); RPF (Cyprus); SENESCYT (Ecuador); MoER, ERC IUT, and ERDF (Estonia); Academy of Finland, MEC, and HIP (Finland); CEA and CNRS/IN2P3 (France); BMBF, DFG, and HGF (Germany); GSRT (Greece); OTKA and NIH (Hungary); DAE and DST (India); IPM (Iran); SFI (Ireland); INFN (Italy); MSIP and NRF (Republic of Korea); LAS (Lithuania); MEYET, and MST (Malaysia); BUAP, CINVESTAV, CONACYT, LNS, SEP, and UASLP-FAI (Mexico); MBIE (New Zealand); PAEC (Pakistan); MSHE and NSC (Poland); FCT (Portugal); JINR (Dubna); MON, RosAtom, RAS, RFBR and RAEP (Russia); MESTD (Serbia); SEIDI, CPAN, PCTI and FEDER (Spain); Swiss Funding Agencies (Switzerland); MST (Taipei); ThEPCenter, IPST, STAR, and NSTDA (Thailand); TUBITAK and TAEK (Turkey); NASU and SFFR (Ukraine); STFC (UK); DOE and NSF (USA).

Individuals have received support from the Marie-Curie programme and the European Research Council and Horizon 2020 Grant, contract No. 675440 (European Union); the Leventis Foundation; the A. P. Sloan Foundation; the Alexander von Humboldt Foundation; the Royal Society; the Packard Foundation; the US Department of Energy; the National Science Foundation; the National Research Foundation; the Rachadapisesk Somport Fund for Postdoctoral Fellowship, Chulalongkorn University and the Chulalongkorn Academic into Its 2nd Century Project Advancement Project (Thailand); the Welch Foundation, contract C-1845; and the Westen Havens Foundation (USA).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. Funded by SCOAP3.

References


79. T.-J. Hou et al., Heavy flavors on CT14. in 23rd International Workshop on Deep-Inelastic Scattering and Related Subjects (DIS 2015), volume DIS2015 (Dallas, Texas, USA, 2015), p. 166

CMS Collaboration

Yerevan Physics Institute, Yerevan, Armenia
A. M. Sirunyan, A. Tumasyan

Institut für Hochenergiephysik, Wien, Austria

Institute for Nuclear Problems, Minsk, Belarus
V. Chekovsky, V. Mossolov, J. Suarez Gonzalez

Page 21 of 34 287
National Centre for Particle and High Energy Physics, Minsk, Belarus
N. Shumeiko

Universiteit Antwerpen, Antwerpen, Belgium

Vrije Universiteit Brussel, Brussel, Belgium

Université Libre de Bruxelles, Brussels, Belgium

Ghent University, Ghent, Belgium

Université Catholique de Louvain, Louvain-la-Neuve, Belgium

Université de Mons, Mons, Belgium
N. Beliy

Centro Brasileiro de Pesquisas Fisicas, Rio de Janeiro, Brazil

Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil

Universidade Estadual Paulista, Universidade Federal do ABC, São Paulo, Brazil

Institute for Nuclear Research and Nuclear Energy, Bulgarian Academy of Sciences, Sofia, Bulgaria
A. Aleksandrov, R. Hadjiiska, P. Iaydjiev, M. Misheva, M. Rodozov, S. Stoykova, G. Sultanov, M. Vutova

University of Sofia, Sofia, Bulgaria
A. Dimitrov, I. Glushkov, L. Litov, B. Pavlov, P. Petkov

Beihang University, Beijing, China
W. Fang, X. Gao

Institute of High Energy Physics, Beijing, China

State Key Laboratory of Nuclear Physics and Technology, Peking University, Beijing, China
Y. Ban, G. Chen, Q. Li, S. Liu, Y. Mao, S. J. Qian, D. Wang, Z. Xu

Universidad de Los Andes, Bogota, Colombia
C. Avila, A. Cabrera, L. F. Chaparro Sierra, C. Flores, C. F. González Hernández, J. D. Ruiz Alvarez

Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture, University of Split, Split, Croatia
N. Godinovic, D. Lelas, I. Puljak, P. M. Ribeiro Cipriano, T. Sculac
Faculty of Science, University of Split, Split, Croatia
Z. Antunovic, M. Kovac

Institute Rudjer Boskovic, Zagreb, Croatia
V. Brigljevic, D. Ferencek, K. Kadija, B. Mesic, T. Susa

University of Cyprus, Nicosia, Cyprus

Charles University, Prague, Czech Republic
M. Finger, M. Finger Jr.

Universidad San Francisco de Quito, Quito, Ecuador
E. Carrera Jarrin

Academy of Scientific Research and Technology of the Arab Republic of Egypt, Egyptian Network of High Energy Physics, Cairo, Egypt
E. El-khateeb, S. Elgammal, A. Ellithi Kamel

National Institute of Chemical Physics and Biophysics, Tallinn, Estonia
R. K. Dewanjee, M. Kadastik, L. Perrini, M. Raidal, A. Tiko, C. Veelken

Department of Physics, University of Helsinki, Helsinki, Finland
P. Eerola, J. Pekkanen, M. Voutilainen

Helsinki Institute of Physics, Helsinki, Finland

Lappeenranta University of Technology, Lappeenranta, Finland
J. Talvitie, T. Tuuva

IRFU, CEA, Université Paris-Saclay, Gif-sur-Yvette, France

Laboratoire Leprince-Ringuet, Ecole polytechnique, CNRS/IN2P3-Université Paris-Saclay, Palaiseau, France

CNRS-IN2P3 IPHC UMR 7178, Université de Strasbourg, 67000 Strasbourg, France

CNRS/IN2P3, Centre de Calcul de l’Institut National de Physique Nucleaire et de Physique des Particules, Villeurbanne, France
S. Gadrat

CNRS-IN2P3, Institut de Physique Nucléaire de Lyon, Université de Lyon, Université Claude Bernard Lyon 1, Villeurbanne, France

Georgian Technical University, Tbilisi, Georgia
A. Khvedelidze

Tbilisi State University, Tbilisi, Georgia
Z. Tsamalaidze
I. Physikalisches Institut, RWTH Aachen University, Aachen, Germany

III. Physikalisches Institut A, RWTH Aachen University, Aachen, Germany

III. Physikalisches Institut B, RWTH Aachen University, Aachen, Germany

Deutsches Elektronen-Synchrotron, Hamburg, Germany

University of Hamburg, Hamburg, Germany

Institut für Experimentelle Kernphysik, Karlsruhe, Germany

Institute of Nuclear and Particle Physics (INPP), NCSR Demokritos, Aghia Paraskevi, Greece
G. Anagnostou, G. Daskalakis, T. Geralis, V. A. Giakoumopoulou, A. Kyriakis, D. Loukas, I. Topsis-Giotis

National and Kapodistrian University of Athens, Athens, Greece
S. Kesisoglou, A. Panagiotou, N. Saoulidou

University of Ioánnina, Ioánnina, Greece
I. Evangelou, C. Foudas, T. Geralis, V. A. Giakoumopoulou, A. Kyriakis, D. Loukas, I. Topsis-Giotis

MTA-ELTE Lendület CMS Particle and Nuclear Physics Group, Eötvös Loránd University, Budapest, Hungary
M. Csanad, N. Filipovic, G. Pasztor

Wigner Research Centre for Physics, Budapest, Hungary

Institute of Nuclear Research ATOMKI, Debrecen, Hungary
N. Beni, S. Czellar, J. Karancsi, A. Makovec, J. Molnar, Z. Szillasi

Institute of Physics, University of Debrecen, Debrecen, Hungary
M. Bartók, P. Raics, Z. L. Trocsanyi, B. Ujvari

Indian Institute of Science (IISc), Bangalore, India
S. Choudhury, J. R. Komaragiri
Institute for Theoretical and Experimental Physics, Moscow, Russia
V. Epshteyn, V. Gavrilov, N. Lychkovskaya, V. Popov, I. Pozdnyakov, G. Safronov, A. Spiridonov, M. Toms, E. Vlasov, A. Zhokin

Moscow Institute of Physics and Technology, Moscow, Russia
T. Aushev, A. Bylinkin

National Research Nuclear University ‘Moscow Engineering Physics Institute’ (MEPhI), Moscow, Russia
M. Chadeeva, R. Chistov, E. Tarkovskii

P.N. Lebedev Physical Institute, Moscow, Russia
V. Andreev, M. Azarkin, I. Dremin, M. Kirakosyan, A. Terkulov

Skobeltsyn Institute of Nuclear Physics, Lomonosov Moscow State University, Moscow, Russia
A. Baskakov, A. Belyaev, E. Boos, M. Dubinin, L. Dudko, A. Ershov, A. Gribushin, V. Klyukhin, O. Kodolova, I. Lokhtin, I. Miagkov, S. Obraztsov, S. Petrushanko, V. Savrin, A. Snigirev

Novosibirsk State University (NSU), Novosibirsk, Russia
V. Blinov, Y. Skovpen, D. Shtol

State Research Center of Russian Federation, Institute for High Energy Physics, Protvino, Russia

Faculty of Physics and Vinca Institute of Nuclear Sciences, University of Belgrade, Belgrade, Serbia
P. Adzic, P. Cirkovic, D. Devetak, M. Dordevic, J. Milosevic, V. Rekovic

Centro de Investigaciones Energéticas Medioambientales y Tecnológicas (CIEMAT), Madrid, Spain

Universidad Autónoma de Madrid, Madrid, Spain
C. Albajar, J. F. de Trocóniz, M. Missiroli, D. Moran

Universidad de Oviedo, Oviedo, Spain
J. Cuevas, C. Erice, I. Fernandez Menendez, J. Gonzalez Caballero, J. R. González Fernández, E. Palencia Cortezon, S. Sanchez Cruz, I. Suárez Andrés, P. Vischia, J. M. Vizan Garcia

Instituto de Física de Cantabria (IFCA), CSIC-Universidad de Cantabria, Santander, Spain

CERN, European Organization for Nuclear Research, Geneva, Switzerland
Brunel University, Uxbridge, UK
J. E. Cole, P. R. Hobson, A. Khan, P. Kyberd, I. D. Reid, P. Symonds, L. Teodorescu, M. Turner

Baylor University, Waco, USA
A. Borzou, K. Call, J. Dittmann, K. Hatakeyama, H. Liu, N. Pastika

Catholic University of America, Washington, DC, USA
R. Bartek, A. Dominguez

The University of Alabama, Tuscaloosa, USA
A. Buccilli, S. I. Cooper, C. Henderson, P. Rumerio, C. West

Boston University, Boston, USA
D. Arcaro, A. Avetisyan, T. Bose, D. Gastler, D. Rankin, C. Richardson, J. Rohlf, L. Sulak, D. Zou

Brown University, Providence, USA

University of California, Davis, Davis, USA

University of California, Los Angeles, USA

University of California, Riverside, Riverside, USA

University of California, San Diego, La Jolla, USA

Department of Physics, University of California, Santa Barbara, Santa Barbara, USA

California Institute of Technology, Pasadena, USA

Carnegie Mellon University, Pittsburgh, USA

University of Colorado Boulder, Boulder, USA
J. P. Cumalat, W. T. Ford, F. Jensen, A. Johnson, M. Krohn, S. Leontsinis, T. Mulholland, K. Stenson, S. R. Wagner

Cornell University, Ithaca, USA

Fairfield University, Fairfield, USA
D. Winn

Fermi National Accelerator Laboratory, Batavia, USA

University of Florida, Gainesville, USA

Florida International University, Miami, USA
S. Linn, P. Markowitz, G. Martinez, J. L. Rodriguez

Florida State University, Tallahassee, USA

Florida Institute of Technology, Melbourne, USA

University of Illinois at Chicago (UIC), Chicago, USA

The University of Iowa, Iowa City, USA

Johns Hopkins University, Baltimore, USA

The University of Kansas, Lawrence, USA

Kansas State University, Manhattan, USA
A. Ivanov, K. Kaadze, Y. Maravin, A. Mohammadi, L. K. Saini, N. Skhirtladze, S. Toda

Lawrence Livermore National Laboratory, Livermore, USA
F. Rebassoo, D. Wright

University of Maryland, College Park, USA

Massachusetts Institute of Technology, Cambridge, USA

University of Minnesota, Minneapolis, USA
35: Now at National Research Nuclear University ’Moscow Engineering Physics Institute’ (MEPhI), Moscow, Russia
36: Also at Institute of Nuclear Physics of the Uzbekistan Academy of Sciences, Tashkent, Uzbekistan
37: Also at St. Petersburg State Polytechnical University, St. Petersburg, Russia
38: Also at University of Florida, Gainesville, USA
39: Also at P.N. Lebedev Physical Institute, Moscow, Russia
40: Also at California Institute of Technology, Pasadena, USA
41: Also at Budker Institute of Nuclear Physics, Novosibirsk, Russia
42: Also at Faculty of Physics, University of Belgrade, Belgrade, Serbia
43: Also at INFN Sezione di Roma; Sapienza Università di Roma, Rome, Italy
44: Also at University of Belgrade, Faculty of Physics and Vinca Institute of Nuclear Sciences, Belgrade, Serbia
45: Also at Scuola Normale e Sezione dell’INFN, Pisa, Italy
46: Also at National and Kapodistrian University of Athens, Athens, Greece
47: Also at Riga Technical University, Riga, Latvia
48: Also at Institute for Theoretical and Experimental Physics, Moscow, Russia
49: Also at Albert Einstein Center for Fundamental Physics, Bern, Switzerland
50: Also at Istanbul University, Faculty of Science, Istanbul, Turkey
51: Also at Adiyaman University, Adiyaman, Turkey
52: Also at Istanbul Aydin University, Istanbul, Turkey
53: Also at Mersin University, Mersin, Turkey
54: Also at Cag University, Mersin, Turkey
55: Also at Piri Reis University, Istanbul, Turkey
56: Also at Gaziosmanpasa University, Tokat, Turkey
57: Also at Izmir Institute of Technology, Izmir, Turkey
58: Also at Necmettin Erbakan University, Konya, Turkey
59: Also at Marmara University, Istanbul, Turkey
60: Also at Kafkas University, Kars, Turkey
61: Also at Istanbul Bilgi University, Istanbul, Turkey
62: Also at Rutherford Appleton Laboratory, Didcot, UK
63: Also at School of Physics and Astronomy, University of Southampton, Southampton, UK
64: Also at Instituto de Astrofísica de Canarias, La Laguna, Spain
65: Also at Utah Valley University, Orem, USA
66: Also at Beykent University, Istanbul, Turkey
67: Also at Bingol University, Bingol, Turkey
68: Also at Erzincan University, Erzincan, Turkey
69: Also at Sinop University, Sinop, Turkey
70: Also at Mimar Sinan University, Istanbul, Turkey
71: Also at Texas A&M University at Qatar, Doha, Qatar
72: Also at Kyungpook National University, Daegu, Korea