IL NUOVO CIMENTO **43 C** (2020) 38
DOI 10.1393/ncc/i2020-20038-y

COLLOQUIA: IFAE 2019

Recent results on single vector boson plus jets production at CMS

F. VAZZOLER(1)(2)

(1) *INFN, Sezione di Trieste - Trieste, Italy*
(2) *Dipartimento di Fisica, Università di Trieste - Trieste, Italy*

received 8 June 2020

Summary. — The most recent results on the hadroproduction of an electroweak boson in association with one or more jets within the CMS experiment are presented. The cross section is studied differentially with respect to the transverse momentum and the angular correlation of the final state particles. Attention is given to the associated production of a vector boson plus flavour-tagged (b, c) jets. Comparison between data and different state-of-the-art theoretical predictions including (next-to)-leading-order corrections are presented as well.

1. – Introduction

The high center-of-mass energy of collisions provided by the CERN Large Hadron Collider (LHC) allows to study with great precision the production of an electroweak (EW) boson in association with one or more jets of high transverse momentum p_T. In order to exploit the high statistics currently available, processes having large jet multiplicity are studied in many differential distributions with respect to the kinematic properties of the jets and vector boson. A precise measurement of the V+jets differential production cross section represents a rigorous test of the quantum chromodynamics (QCD) sector of the Standard Model (SM), providing valuable inputs for constraining the parton distribution functions (PDFs) in the high-Q^2 regime, thus reducing the systematic uncertainties on many relevant SM cross section measurements. Moreover, the comparison between data and predictions gives the opportunity to improve the Monte Carlo (MC) description of experimental data, while validating higher-order theoretical calculations. Finally, events with high boson p_T and jet multiplicity represent an important background for both SM and Beyond the Standard Model (BSM) processes, such as single top and $t\bar{t}$ production, vector boson fusion and scattering studies, Higgs boson production. Knowing with high precision the V+jets background is thus of fundamental importance to increase the overall sensitivity of many analyses.

2. – Monte Carlo predictions

Signal and background processes are mainly predicted using Monte Carlo simulations. Each of them includes the matrix element (ME), which describes the hard electroweak process, and the additional electroweak correction in perturbative Quantum Chromodynamic (pQCD) approach. The calculations are performed at leading order (LO) in pQCD for processes including up to four partons in the final state, while next-to-leading-order (NLO) predictions are available for processes up to two partons in the final state. The event generators used are MadGraph5 [1] or PYTHIA8 [2] for LO processes and MadGraph5_aMC@NLO [1] for NLO predictions. In the ME calculation the NNPDF 3.0 NNLO PDF set [3] is used. In order to include in the ME predictions the initial- and final-state gluon emission and the partonic cascade the event generator is interfaced with PYTHIA8. The settings of the parton shower are defined by the CUETP8M1 tune [4]. The matching between the number of partons generated at ME and the predictions of PYTHIA8 is performed using the k_T-MLM scheme [5] at LO while the FxFx one [6] is used at NLO. The measured differential cross section for the W/Z + jets processes are also compared to fixed-order calculations based on the N-jettines subtraction scheme (N_{jetti}) at NNLO for the W/Z + 1−jet production [7,8]. The NNPDF 3.0 NNLO PDF is used for the processes involving W production while the CT14 PDF [9] is used for the Z boson. Finally, the measured Z+jets cross section is also compared with a combination of next-to-next-to-leading order (NNLO) calculation for the Drell-Yan production combined with next-to-next-to-leading leading logarithm (NNLL) resummation, obtained with the GENEVA event generator [10]. Additional pp collisions must be included into the simulated samples. These events represent the effects of proton collisions that occur in the same or adjacent bunch crossing, commonly known as pileup (PU). The PU contribution is simulated as additional minimum bias events superimposed on the primary simulated events. Events are then passed to a detailed CMS detector simulation. The MC events are reconstructed with the same reconstruction software as real data.

3. – Vector boson plus jets production

3˙1. *W + jets at 13 TeV.* – The W + jets production is studied in the muon channel at a center-of-mass energy of 13 TeV corresponding to 2.2 fb^{-1} of data collected in 2015 [11]. The electron channel is not used due to a higher trigger threshold applied during data acquisition. Events are required to have exactly one muon with $p_T > 20\,\text{GeV}$ and $|\eta| < 2.4$ and one or more jets with $p_T > 30\,\text{GeV}$ and $|\eta| < 2.4$. The jets are reconstructed clustering particles with the anti-k_T algorithm requiring $\Delta R(\text{jet}, \mu) > 0.4$. Events are further required to be in the transverse mass peak region for the W boson $m_T > 50\,\text{GeV}$ while a b-tag veto is applied to reduce the $\bar{\text{t}}\text{t}$ contribution. The signal and background processes are estimated from MC simulation but the QCD multi-jet background is estimated from data. The obtained agreement between jet multiplicity distributions in data and simulation is on the 5% in almost every region. The W + jets fiducial cross section is obtained subtracting the backgrounds from data distributions and unfolding the obtained distribution back to its particle level. The unfolded distribution is then used to measure the W + jets cross section as a function of different kinematic observables. The measured cross section as a function of the inclusive jet multiplicities is compared with MG5_aMC predictions in fig. 1(left) showing good agreement within uncertainties for both LO and NLO predictions. The p_T of the most energetic jet is presented in fig. 1(center) and it is better described by the NLO MG5_aMC and the NNLO calculations, while the LO

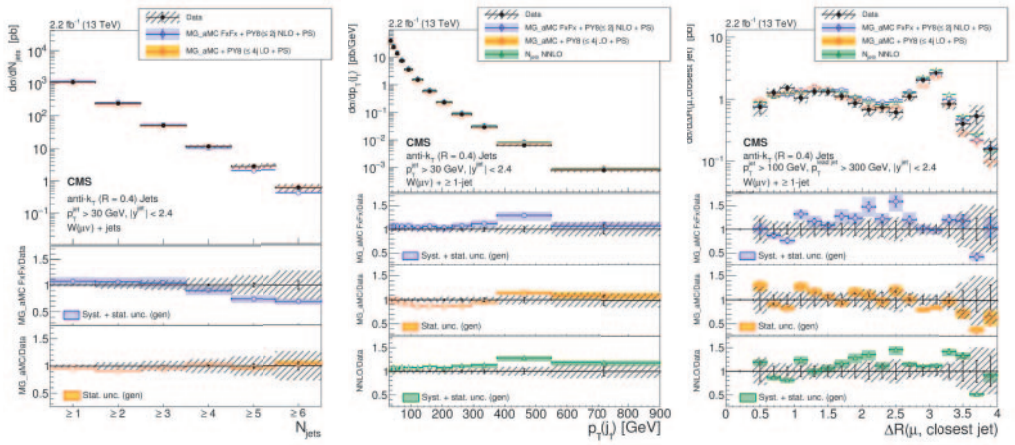

Fig. 1. – Measured cross section for W + *jets* as a function of the N_{jets} (left), the p_T of the most energetic jet (center) and of the $\Delta R(\mu, \text{jet})$ (right).

MG5_aMC prediction underestimate the signal at low jet p_T. The cross section is also measured as a function of different angular variables such as the angular distance between the muon and its closest jet in fig. 1(right), giving an insight on the EW radiative production of the W boson. Both LO (NLO) MG5_aMC and NNLO predictions show a decent modelling of the $\Delta R(\mu, \text{jet})$ distribution.

3˙2. γ + jets *at 13 TeV.* – The Z + jets production is studied in the electron and muon channel at a center-of-mass energy of 13 TeV corresponding to 2.2 fb^{-1} of data collected in 2015 [12]. The Z boson is defined as a pair of opposite charged leptons (e or μ) with $p_T > 20\,\text{GeV}$ and $|\eta| < 2.4$, corresponding to a reconstructed invariant mass range of $91 \pm 20\,\text{GeV}$. This definition provides a good balance between signal acceptance and background rejection. Jets are defined clustering all reconstructed particles with the anti-k_T algorithm requiring a $p_T > 30\,\text{GeV}$, $|\eta| < 2.4$ and $\Delta R(\text{jet}, \text{l}) > 0.4$. The contributions from signal and background events are estimated from MC simulation. The $\text{t}\bar{\text{t}}$ process, representing the dominant background, is also measured from data exploiting the fact that the $\sigma(\text{t}\bar{\text{t}} \to \text{e}^+\text{e}^-(\mu^+\mu^-))$ and $\sigma(\text{t}\bar{\text{t}} \to \text{e}^+\mu^-(\mu^+\text{e}^-))$ cross sections are expected to be equal. Events are then selected in a control sample using the same criteria as for the measurements but requiring the leptons not to have the same flavour. The obtained jet multiplicity distributions in data and simulation show a background contamination as low as 1% for the inclusive cross section, while representing a consistent contribution at higher jet multiplicities, mainly due to the $\text{t}\bar{\text{t}}$ process. The Z + jets fiducial cross section is then obtained subtracting the backgrounds from data distributions and unfolding the obtained distribution back to its particle level. The unfolded distribution is then used to measure the Z + jets cross section as a function of different kinematic observables. The agreement between LO MG5_aMC prediction and data tends to be lower than the NLO one for the jet multiplicity (N_{jets}), as can be seen in fig. 2(left). The GENEVA prediction describes well the measurements up to two jets, but fails to describe higher jet multiplicities where one or more jets arise from the parton shower. The p_T of the Z boson, directly related to the multiple (soft) gluon emission modelling, is shown in fig. 2(center), while the imbalance between the Z boson and the jet p_T, due to hadronic activity not included

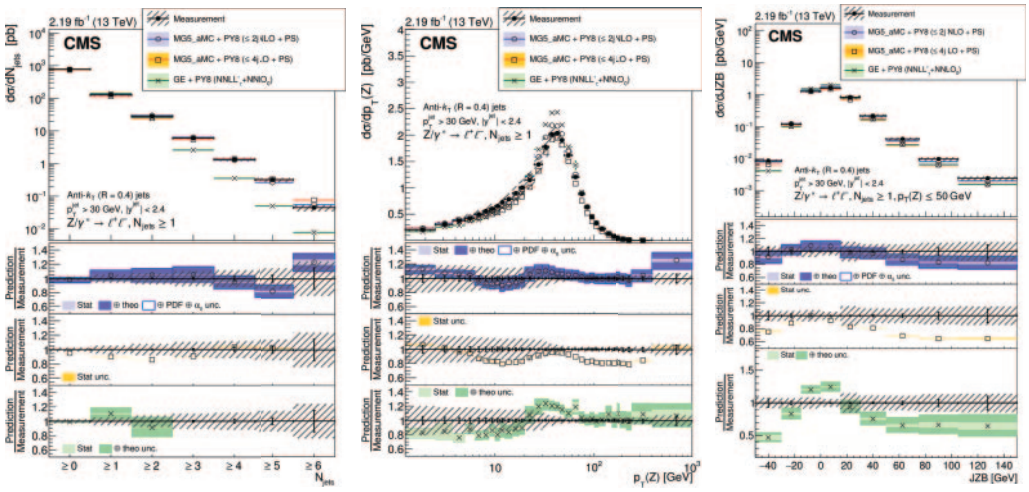

Fig. 2. – Measured cross section for Z + jets as a function of the N_{jets} (left), the p_T of the Z boson (center) and the jet-Z boson imbalance JZB $= |\sum_{\text{jets}} \vec{p}_T(\text{j}_i)| - |\vec{p}_T^{\,Z}|$ (right).

in the jets, is presented in fig. 2(right). The two distributions are well described by the NLO MG5_aMC prediction, while the LO MG5_aMC and GENEVA, both including LO accuracy for two partons in the final state, show consistent discrepancies with data. A NLO accuracy is then foreseen to correctly describe the hadronic activity in Z + jets events at 13 TeV.

3˙3. γ + jets *at 13 TeV*. – The γ + jets production is studied at a center-of-mass energy of 13 TeV corresponding to 2.26 fb^{-1} of data collected in 2015 [13]. The dominant production mechanisms of a photon plus jets in high-energy pp collisions are the quark-gluon Compton scattering $\text{qg} \to \text{q}\gamma$, annihilation $\text{q}\bar{\text{q}} \to \text{g}\gamma$ and the parton fragmentation $\text{q}\bar{\text{q}}(\text{gg}) \to \text{X}\gamma$. These measurements are directly sensitive to the gluon content of the proton over a wide range of x and Q^2 and were only recently included in the global PDF fits. Events are selected requiring jets clustered via the anti-k_T algorithm with $p_T > 30$ GeV and $|\eta| < 2.4$ and high-energy photon candidates with $E_T^\gamma > 190$ GeV and $|\eta| < 2.5$. An electron veto is applied to suppress electrons faking photons. To reject photons arising from electromagnetic decays of particles in the hadronic showers, further requirements are imposed on the ratio between energy measured in the hadronic to electromagnetic calorimeters, the isolation of the candidate photon with respect to nearby charged and neutral particles and the shape of the energy deposit in the electromagnetic calorimeter. A multivariate analysis is the implemented to further suppress the remaining jet-fake background and to extract the photon yield. Two boosted decision tree templates are defined for signal and background, starting from discriminating variables sensitive to the photon signal in the detectors, and then used to fit data in a binned likelihood fit in order to estimate the photon yield. The γ + jets cross section is then measured differentially with respect to the photon pseudo-rapidity and transverse energy, and compared to NLO QCD calculations provided by JETPHOX 1.3.1 [14] using NNPDF3.0 NLO PDFs in fig. 3(left). A good agreement is observed between theoretical predictions and experimental data, especially in the low-to-middle photon E_T range, where these measurement provide the potential to further constrain the proton gluon PDF. Possible disagreements

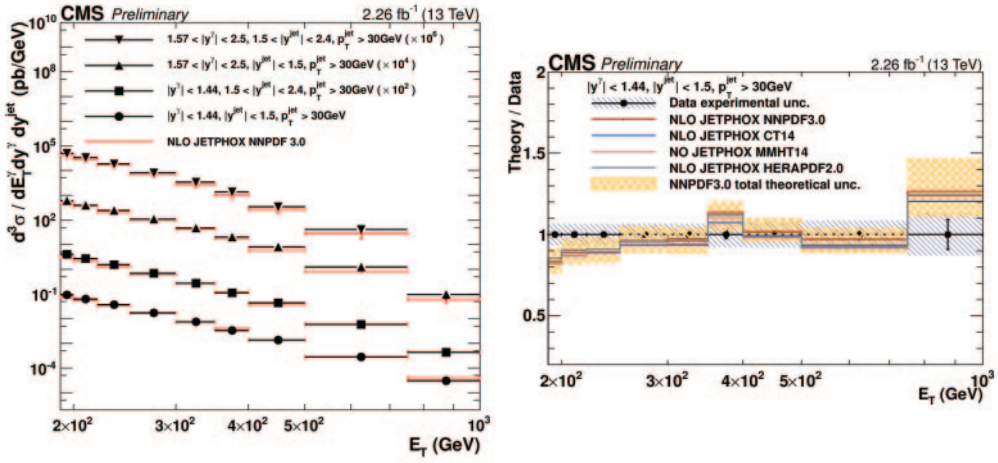

Fig. 3. – Measured cross section for $\gamma +$ jets as a function of the E_T and $|\eta|$ of the photon.

between JETPHOX predictions using different PDFs sets were also studied in fig. 3(right) showing small differences within the statistical and systematic uncertainties.

4. – Vector boson plus flavour-tagged jets production

4˙1. *$W + c$ at 13 TeV.* – The W + c production is studied at a center-of-mass energy of 13 TeV corresponding to 35.7 fb^{-1} of data collected in 2016 [15]. The dominant production mechanism of a W boson plus a c quark in high-energy pp collisions is the quark-gluon scattering $\text{gs}(\bar{\text{s}}) \rightarrow W^{\pm}\text{c}(\bar{\text{c}})$. These measurements, indeed, give an insight on the strange quark content of the proton, which is a valuable information for the measurement of the W boson mass at the LHC, while providing cross-checks with the results obtained in the global PDFs fits using DIS data. The W is reconstructed requiring a muon with $p_T > 26$ GeV and $|\eta| < 2.4$ and a transverse mass greater than 50 GeV. The charm quarks are tagged by fully reconstructing their strong decay $\text{c} \rightarrow \text{D}^{*\pm}(1020) \rightarrow D^0\pi^{\pm}_{\text{slow}} \rightarrow \text{K}^{\mp}\pi^{\pm}\pi^{\pm}_{\text{slow}}$. The signal consists of only opposite sign charge combinations of a W boson and a c quark while the dominant background is represented by $\text{W} + \text{c}\bar{\text{c}}$ production that contains equal number of opposite and same sign events. This background is thus removed subtracting events with the same sign combination from the opposite sign ones. The measured inclusive differential cross section is compared to predictions at NLO, obtained using MCFM 6.8 [16] and different PDFs sets, in fig. 4(left). The strangeness suppression distributions in all PDFs sets considered are in good agreement with each other and disagree with the ATLASepWZ16nnlo result [17]. The CMS measurements do not support the ATLAS hypothesis of an enhanced strange quark contribution in the proton sea.

4˙2. *$Z/W + b$ at 8 TeV.* – The production of a W or a Z boson plus b-jets is studied at a center-of-mass energy of 8 TeV corresponding to 19.8 fb^{-1} of data collected in 2012 [18, 19]. These measurements are of particular interest giving the opportunity to study the b content of the proton. A better understanding of the b hadron production mechanism and kinematic properties is indeed required to refine the background

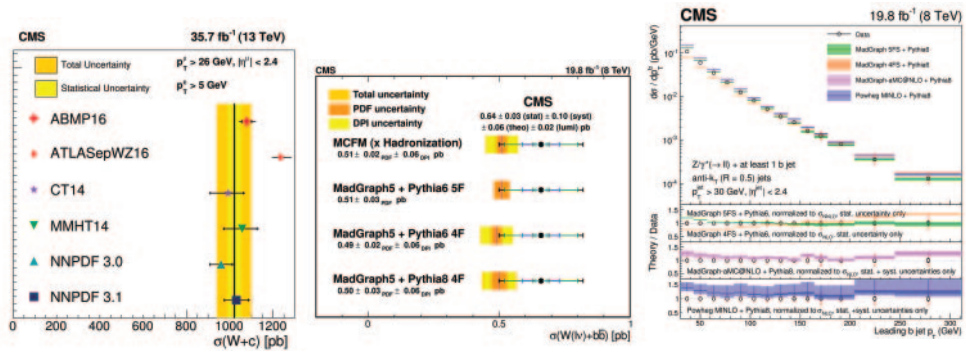

Fig. 4. – Measured cross section for W + c (left) and W + b (center) and differential cross section with respect to the p_T of the leading b-jet for Z + b events (right).

predictions and increase the sensitivity to many SM analyses BSM searches at the LHC. The W and Z bosons are reconstructed requiring electrons (muons) with $p_T >$ 30 (20) GeV and $|\eta| < 2.4$. Jets are clustered via the anti-k_T algorithm while the long lifetime and relatively large mass of b hadron is used to provide b-jet identification with a dedicated algorithm. The dominant background, represented by $t\bar{t}$ events, is estimated with a data-driven technique. The differential cross section is then measured as a function of different observables characterising the kinematic properties of the system. The results are compared with LO(NLO) predictions obtained using several flavour schemes in fig. 4(center) and (right). A good agreement is observed between data and theory.

5. – Summary

The CMS experiment has provided a broad range of V+jets measurements exploiting both 8 and 13 TeV proton proton collisions alongside LHC Run I and Run II data taking periods. All measurements share high precision in inclusive and differential cross sections, achieved with innovative experimental methods and larger available datasets. Measurements are generally in good agreement with predictions although it is clear that NLO calculations are needed to describe data. Future developments in this field are foreseen and include testing new generations of MC, able to push prediction beyond NLO, together with better constrain of systematic uncertainties, mainly coming from the jet energy scale corrections, and improvements of the unfolding and statistical techniques.

REFERENCES

[1] Alwall J. *et al.*, *JHEP*, **07** (2014) 79.
[2] Torbjorn S. *et al.*, *Comput. Phys. Commun.*, **191** (2015) 159.
[3] The NNPDF Collaboration, *JHEP*, **04** (2015) 40.
[4] Khachatryan V. *et al.*, *Eur. Phys. J. C*, **76** (2016) 155.
[5] Alwall J. *et al.*, *JHEP*, **02** (2009) 017.
[6] Frederix R. and Frixione S., *JHEP*, **12** (2012) 61.
[7] Boughezal R. *et al.*, *Phys. Rev. D*, **94** (2016) 11309.
[8] Boughezal R. *et al.*, *Phys. Rev. Lett.*, **116** (2016) 152001.
[9] Dulat S. *et al.*, *Phys. Rev. D*, **93** (2016) 033006

[10] Alioli S. *et al.*, *Phys. Rev. D*, **92** (2015) 094020.
[11] Sirunyan A. M. *et al.*, *Phys. Rev. D*, **96** (2017) 072005.
[12] Sirunyan A. M. *et al.*, *Eur. Phys. J. C*, **78** (2018) 965.
[13] Sirunyan A. M. *et al.*, *Eur. Phys. J. C*, **79** (2019) 20.
[14] Aurenche P. *et al.*, *Phys. Rev. D*, **73** (2006) 094007.
[15] Sirunyan A. M. *et al.*, *Eur. Phys. J. C*, **79** (2019) 269
[16] John M. Campbell and Ellis R. K., *Nucl. Phys. B*, **205–206** (2010) 10.
[17] Aaboud M. *et al.*, *Eur. Phys. J. C*, **77** (2017) 367.
[18] Khachatryan V. *et al.*, *Eur. Phys. J. C*, **77** (2017) 92.
[19] Khachatryan V. *et al.*, *Eur. Phys. J. C*, **77** (2017) 751.