PROCEEDINGS OF SCIENCE

# Electroweak production of Z bosons with forward-backward jets at CMS

**Pedro Ferreira da Silva***, *for the CMS Collaboration*
CERN, *Geneva, Switzerland and* LIP, *Lisbon, Portugal*
*E-mail:* psilva@cern.ch

The first measurement of the electroweak production cross section of a Z boson with two jets (EWK Zjj) in proton-proton collisions at $\sqrt{s} = 7$ TeV is discussed. Proton-proton data, recorded by the CMS experiment at the LHC and corresponding to an integrated luminosity of 5 fb$^{-1}$, is used to measure $\sigma(\text{EWK Zjj}) = 154 \pm 24_{\text{stat}} \pm 46_{\text{exp}} \pm 27_{\text{th}} \pm 3_{\text{lumi}}$ fb in agreement with the standard model prediction. The result establishes the basis for the more general study of vector boson fusion processes: Higgs boson searches, anomalous electroweak gauge couplings and the scattering of vector bosons.

*The European Physical Society Conference on High Energy Physics -EPS-HEP2013*
*18-24 July 2013*
*Stockholm, Sweden*

*Speaker.



PoS(EPS-HEP 2013)466

## 1. Introduction

Establishing the properties of the production of vector bosons (V=W, Z) at the LHC is crucial not only as a test of perturbative QCD (pQCD) and tuning of the Monte-Carlo (MC) generators but also as window to probe the content of the proton offering constraints on the modelling of the parton density functions (PDF). In the context of different searches for new physics these processes often constitute important backgrounds.

The production of vector bosons in association with two jets (V+2 jets) is dominated by radiative processes which involve QCD vertices. However, the production of V+2 jets at $\mathscr{O}(\alpha_{EW}^4)$ has distinct properties with respect to the QCD production as it is dominated by processes in which the two extra jets are produced after the scattering of quarks through color singlet exchanges and where the vector bosons are exchanged via a t-channel. These characteristics lead to small transferred momentum, i.e. $Q^2 \sim 0$. As a consequence the two jets are expected to scatter with a large rapidity interval in which the production of additional hadronic activity is suppressed. The invariant mass of the two jets, referred to as tag jets, is also expected to be large, reaching the TeV scale.

Pure electroweak production of a Z+2 jets (EWK Zjj) includes the so-called vector boson fusion production of a Z (VBF Z). Negative interference terms suppress however the pure VBF Z production by a factor close to 2.5. Using VBFNLO [1], a production cross section of $\sigma(\text{EWK Zjj}) =$ 166 fb is predicted for proton-proton collisions at 7 TeV, after requiring $p_T^j > 25$ GeV, $M_{\text{jj}} >$ 120 GeV, $|\eta|^j < 4$ and $M_{ll} > 50$ GeV at generator level and using $\mu_R = \mu_F = m_Z$ and the CT10 PDF [2]. The expected cross section is therefore large enough to provide a high statistics sample which can be used to probe the properties of a color singlet exchange at the LHC [3]. The study of these properties is of obvious interest in view of a better understanding of the rarer VBF Higgs and vector boson scattering productions [4].

The first measurement of this process at the LHC is reported in this manuscript. The results, obtained by the CMS experiment [5], have published recently in [6].

## 2. Event selection and signal discrimination strategy

Events with at least two charged leptons (electrons or muons) are selected in data using dilepton triggers. Each lepton is required to have a transverse momentum $p_{\text{T}} > 20$ GeV and to be reconstructed within $|\eta| < 2.4$, where $\eta$ is its pseudo-rapidity. Isolation criteria is furthermore imposed by using the tracks which are reconstructed in a cone of radius $R = \sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.3$. The sum of the transverse momenta of the tracks is required not to exceed more than 10% of the $p_{\text{T}}$ of each lepton candidate. In each event the leptons are required to have opposite charge and the invariant mass of the system must be compatible with a Z boson, i.e. $|M_{\ell\ell} - M_Z| < 15$ GeV.

The two tag jets are selected within $|\eta| < 3.6$ and required to have $p_{\text{T}} > 40$ GeV. The leading-$p_{\text{T}}$ jet is required to have $p_{\text{T}} > 65$ GeV. In the tag jet selection two alternative approaches are considered: particle-flow-based jets [7] and jet-plus-tracks [8]. The first approach combines the relevant information from the different sub-detectors to identify and reconstruct particle candidates in the event which are subsequently used in the jet clustering algorithm. The second approach uses calorimeter jets whose energy response and resolution is improved by incorporating track-

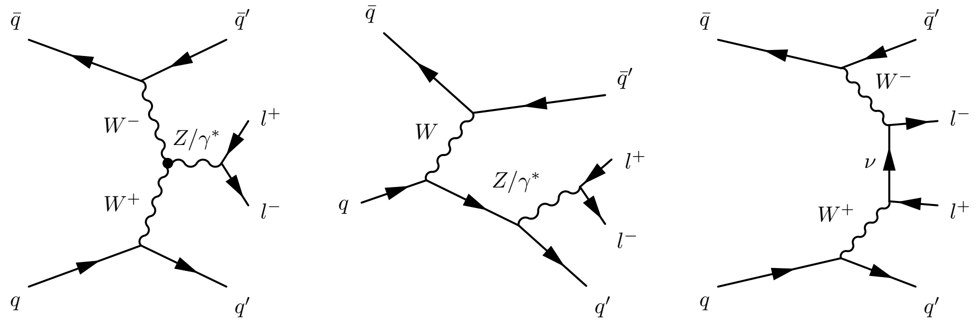


**Figure 1:** Representative diagrams for dilepton production in association with two jets from pure electroweak processes. Vector boson fusion (left), Bremmstrahlung-like (center), and multiperipheral (right) productions.

based corrections. In both cases jets are reconstructed using the anti-$k_{\mathrm{T}}$ algorithm with a distance parameter of 0.5 [9].

Two final selection criteria are applied: the invariant mass of the dijet system is required to be $M_{\mathrm{jj}} > 600$ GeV; the Z boson candidate is required to have been centrally produced in the dijet rest frame, i.e. $|y^{*}| = |y_{Z} - 0.5 \cdot (y_{j1} + y_{j2})| < 1.2$ where $y$ ($y^{*}$) is the rapidity in the laboratory (dijet rest) frame.

Overall the selection criteria for jets is found to be stricter when compared with the one usually applied for Higgs boson searches in the VBF channel [10]. Higher $p_{\mathrm{T}}$ thresholds are used and no central jet veto is applied. The kinematics for the EWK Zjj and the VBF Higgs boson production are expected to differ given that the first includes the contributions from bremmstrahlung and multiperipheral processes, as shown in Fig. 1. Both the additional processes and the interference between them leads to higher jet transverse momenta with respect to the VBF production alone. This fact is a consequence of the polarisation in the different processes: both longitudinal and transversely polarised W bosons participate in EWK Zjj processes.

After the full selection a 6% efficiency is expected for both $ee$ and $\mu\mu$. The efficiency is estimated using a signal sample simulated with the MADGRAPH generator (version 5) [11] interfaced with PYTHIA 6.4.25 [12] for showering and hadronization. The ratio of signal to background events is expected to be $\approx$11% with the main background contribution being due to QCD production of a Z boson with at least 2 jets ($\approx$97% of the total expected backgrounds). The contributions from each process at each selection step, obtained in simulation, are summarised in Table 1. The dijet invariant mass distribution obtained after selection (except the $M_{\mathrm{jj}}$ requirement) is shown in Fig. 2 (*left*). After the fit for the contribution of signal and backgrounds is performed, a good agreement is observed across the full $M_{\mathrm{jj}}$ spectrum.

The dijet invariant mass has good discrimination power between signal and background but it is also expected to be prone to large uncertainties stemming from jet energy scale and resolution in the forward pseudo-rapidity region where the tag jets tend to be reconstructed. A more sensitive approach is used based on a multivariate analysis. A boosted decision tree with decorrelation (BDTD) is trained to give optimal separation between signal-like and background-like events. Besides making use of the $p_{\mathrm{T}}$ of the tag jets, $M_{\mathrm{jj}}$ and $y^{*}$ the following variables are included in the training: difference ($\Delta\eta_{jj}$) and sum ($\eta_{\mathrm{j1}} + \eta_{\mathrm{j2}}$) of the tag jets' pseudo-rapidities; azimuthal sepa-

**Table 1:** Event yields after each selection step for the data, for the signal Monte Carlo and the backgrounds when PF jets are used. The expected contributions from the signal and background processes are evaluated from simulation, for 5 fb$^{-1}$ of integrated luminosity. VV is used an abbreviation for the sum of WW, WZ and ZZ processes.

| Selection step | | Data | EWK Zjj | QCD Zjj | $t\bar{t}$ | VV |
|---|---|---|---|---|---|---|
| $Z \to \ell\ell$ | $e^+e^-$ | $1.5\times10^6$ | 410 | $1.5\times10^6$ | 1600 | 2160 |
| | $\mu^+\mu^-$ | $1.7\times10^6$ | 460 | $1.7\times10^6$ | 1400 | 2450 |
| Tag jets | $e^+e^-$ | 24000 | 270 | 23000 | 880 | 253 |
| | $\mu^+\mu^-$ | 26000 | 280 | 26000 | 680 | 285 |
| $\lvert y^*\rvert < 1.2$ | $e^+e^-$ | 15000 | 200 | 15000 | 760 | 162 |
| | $\mu^+\mu^-$ | 16000 | 200 | 16000 | 580 | 172 |
| $M_{\mathrm{jj}} > 600$ GeV | $e^+e^-$ | 560 | 67 | 550 | 17 | 4 |
| | $\mu^+\mu^-$ | 640 | 72 | 610 | 14 | 4 |

ration of the two tagging jets ($\Delta\phi_{jj}$); transverse momentum ($p_{\mathrm{T}}^{\ell\ell}$) and rapidity ($y^{\ell\ell}$) of the dilepton system; azimuthal separations between the dilepton system and each tagging jet ($\Delta\phi(\ell\ell,\mathrm{j1})$ and $\Delta\phi(\ell\ell,\mathrm{j2})$). In the *ee* channel the analysis makes also use of a quark-gluon discriminator. The BDTD has been trained using a MadGraph-based simulation of signal and Drell-Yan and top pair events. The discriminator output is shown in Fig. 2 (*right*).

## 3. Results

The dijet invariant mass and the BDTD spectra are used to extract the signal strength, i.e. $\mu = \sigma/\sigma_{\mathrm{th}}$, using a fraction fit to the data with signal and background shapes. To evaluate the uncertainties, alternative shapes are used for each process after shifting the sources by $\pm1\sigma$. If the source affects solely the rate of a given process, the initial prediction is altered accordingly. In each case the fit is repeated and the difference to the central value is used as an estimator of the uncertainty. Table 2 summarises the uncertainties affecting the measurement with the two techniques as well as the result obtained from the fit for the signal and background strength.

The dominant source of uncertainty is observed to be the jet energy scale and resolution. The impact is observed to be larger when a simple fit to the dijet invariant mass is performed.

The second source of systematic uncertainty is due to the modelling of the background. This uncertainty is derived by comparing the MadGraph-based prediction with a next-to-leading order (NLO) computation performed with MCFM [13]. The NLO calculation is performed using a dynamical QCD scale set to $\mu_0 = \sum_{i=1}^{n} p_{\mathrm{T}}^{i}$, and it is used to correct the MadGraph-predicted $M_{\mathrm{jj}}$ shape using parton jets. After propagating to the reconstructed dijet mass (or BDT discriminator) the difference induced in the measurement by the altered shape is assigned as systematic uncertainty. The variation of the QCD scale in the NLO calculation, to $2\mu_0$ or $\mu_0/2$, and the variation by the PDF uncertainties, are found to have negligible impact on the modelling of the background.

The results of the fits are found to be in agreement between the different channels and methods. Using the most precise method (BDT) and combining the di-muon and di-electron analyses with PF jets, the average cross section obtained is $\sigma(\mathrm{EWK}\ \ell\ell\mathrm{jj})_{\ell=e,\mu} = 154 \pm 24_{\mathrm{stat}} \pm 46_{\mathrm{syst}} \pm 27_{\mathrm{th}} \pm 3_{\mathrm{lum}}$ pb, in agreement with the NLO standard model prediction.



PoS(EPS-HEP 2013)466

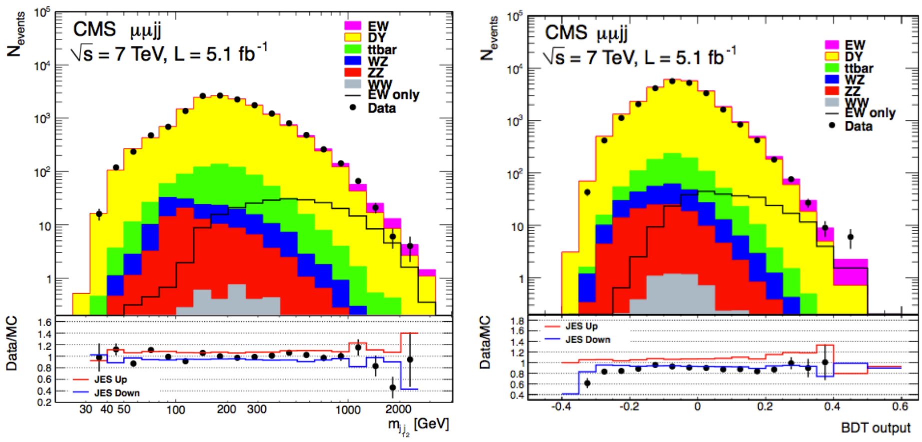


**Figure 2:** Dijet invariant mass *left* and BDTD output (*right*) for the dimuon channel after selecting the Z candidate and the tag jets. The expected contributions from the signal and background processes are evaluated from simulation. The solid line with the label "EWK only" shows the distribution for the signal alone. The bottom panels show the ratio of data over the expected contribution of the signal plus background. The region between the two lines, with the labels JES Up and JES Down, shows the $1\sigma$ uncertainty of the simulation prediction due to the jet energy scale uncertainty. The data points are shown with the statistical uncertainties.

**Table 2:** Uncertainties affecting the determination of the signal strength in the different analysis.

| Discriminator | $M_{jj}$ | | BDT | | |
|---|---|---|---|---|---|
| Channel | $\mu^+\mu^-$ | $\mu^+\mu^-$ | $\mu^+\mu^-$ | $\mu^+\mu^-$ | $e^+e^-$ |
| Jet algorithm | JPT | PF | JPT | PF | PF |
| $\mu = \sigma/\sigma_{th}$ | 1.14 | 1.14 | 0.90 | 0.85 | 1.17 |
| Statistical uncertainty | 0.28 | 0.30 | 0.19 | 0.18 | 0.27 |
| Systematic uncertainty | 0.47 | | 0.26 | | 0.35 |
| Background modelling | 0.20 | | 0.16 | | 0.16 |
| Signal modelling | 0.05 | | 0.05 | | 0.05 |
| $t\bar{t}$ cross section | 0.02 | | 0.03 | | 0.03 |
| Diboson cross sections | 0.01 | | 0.02 | | 0.02 |
| Jet energy scale/resolution | 0.44 | | 0.22 | | 0.29 |
| Pileup modelling | 0.05 | | 0.03 | | 0.03 |
| MC statistics | 0.14 | | 0.13 | | 0.19 |
| Gluon-quark discriminator | - | | - | | 0.02 |
| Dilepton selection | 0.02 | | 0.02 | | 0.02 |
| Luminosity | 0.02 | | 0.02 | | 0.03 |

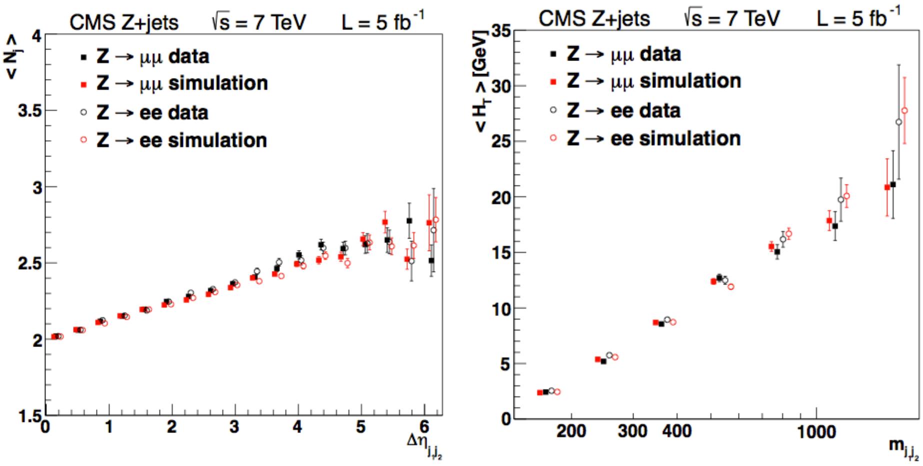


**Figure 3:** *Left*: average number of jets with $p_T > 40$ GeV as a function of the $\Delta\eta_{j1j2}$ separation. *Right*: average $H_T$ of the three leading soft track jets in the pseudo-rapidity interval between the tagging jets for $p_T^{j1,j2} > 65, 40$GeV as a function of the dijet invariant mass. The data points and the points from simulation are shown with the statistical uncertainties.

## 4. Jet activity profiles in Z+jets events

As previously referred, the main background is modelled from simulation and it is largely affected by the theory uncertainty which is assigned from the difference between the Born-level MADGRAPH predictions and the NLO MCFM computation. To gain further insight into the adequacy of the MADGRAPH prediction we have studied the radiation patterns of multijet events in association with a Z boson candidate in both control and signal regions. The variables chosen for this study follow the prescriptions and suggestions from [14] in which the model dependence is estimated by comparing different generators. Our studies have been performed either using PF jets or track-based jets which can be unambiguously associated to the primary vertex of the event and and are therefore expected to be pileup-resilient. The anti-$k_T$ algorithm with a distance parameter of 0.5 is used to cluster the track jets collection built using tracks which are found to be within 2 mm of the main primary vertex [1] in the longitudinal direction and with a relative error >1/3. Overall a fair agreement between data and simulation is observed for the different variables studied and the MADGRAPH+PYTHIA-based simulation. Figure 3 (*left*) shows the results obtained for the average number of PF jets with $p_T > 40$ GeV as function of the pseudo-rapidity separation between the tag jets. No requirement is made on the relative position of the extra jet with respect to the tag jets. Figure 3 (*right*) shows the evolution of the $H_T$ variable, computed from the scalar sum of the three leading $p_T$ track jets reconstructed in the $\eta_{min}^{tagjet} + 0.5 < \eta < \eta_{max}^{tagjet}$ range. Complementary V+jets studies have been reported in this conference as well, see [15].

## 5. Summary

A summary of the first measurement of EWK Zjj production at the LHC, which comprises

[1]The main primary vertex in the event is chosen to be that with the largest $\sum p_T^2$ where the sum is made over all the tracks which have been used in the vertex fit.



PoS(EPS-HEP 2013)466

the vector boson fusion production of a Z, is presented in this manuscript. The production cross section for this process is overwhelmed by one of the QCD production of a Z boson with two associated jets. A multivariate analysis is used to gain in significance and robustness against systematic uncertainties when separating the pure EWK production of Z+2 jets. The measured cross section is observed to be in agreement with the NLO prediction corresponding to a 2.6$\sigma$ significance. The high statistics sample selected in data opens the door to probe the properties of colour-singlet exchange at the LHC with higher integrated luminosity. For the selected events the measurement of the hadronic activity in events with Drell-Yan production in association with two jets is in good agreement with Born-level MADGRAPH+PYTHIA simulation. In the future it is expected that the EWK Zjj production can be more accurately probed by making use of higher integrated luminosity.

## References


[1] K. Arnold *et al.*, "VBFNLO: A Parton level Monte Carlo for processes with electroweak bosons", Comput. Phys. Commun. **180** (2009) 1661

[2] H. -L. Lai *et al.*, "New parton distributions for collider physics," Phys. Rev. D **82** (2010) 074024

[3] D. L. Rainwater, R. Szalapski and D. Zeppenfeld, "Probing color singlet exchange in $Z$ + two jet events at the CERN LHC", Phys. Rev. D **54** (1996) 6680

[4] V. A. Khoze *et al*, "A Z monitor to calibrate Higgs production via vector boson fusion with rapidity gaps at the LHC", Eur. Phys. J. C **26** (2003) 429

[5] CMS Collaboration, "The CMS experiment at the CERN LHC", JINST **3** (2008) S08004

[6] CMS Collaboration, "Measurement of the hadronic activity in events with a Z and two jets and extraction of the cross section for the electroweak production of a Z with two jets in pp collisions at $\sqrt{s}$ = 7 TeV", JHEP **10** (2013) 062

[7] CMS Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and $E_{\mathrm{T}}^{\mathrm{miss}}$", CMS-PAS-PFT-09-001 (2009)

[8] CMS Collaboration, "Determination of jet energy calibration and transverse momentum resolution in CMS", JINST **6** (2011) 11002

[9] M. Cacciari, G. P. Salam and G. Soyez, "The Anti-k(t) jet clustering algorithm", JHEP **0804** (2008) 063

[10] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", Phys. Lett. B **716** (2012) 30

[11] J. Alwall *et al.*, "MadGraph 5 : Going Beyond", JHEP **1106** (2011) 128

[12] T. Sjostrand, S. Mrenna and P. Z. Skands, "PYTHIA 6.4 Physics and Manual", JHEP **0605** (2006) 026

[13] J. M. Campbell and R. K. Ellis, "MCFM for the Tevatron and the LHC", Nucl. Phys. Proc. Suppl. **205-206** (2010) 10

[14] J. R. Andersen *et al.* [SM and NLO Multileg Working Group Collaboration], "The SM and NLO Multileg Working Group: Summary report", [arXiv:hep-ph/1003.1241]

[15] M.A. Weber, "Vector boson plus jets measurements at CMS", in proceedings of EPS-HEP 2013, PoS(EPS-HEP 2013)469