IL NUOVO CIMENTO **48 C** (2025) 81
DOI 10.1393/ncc/i2025-25081-6

Colloquia: IFAE 2024

Recent results in VBS and VBF measurement performed with the CMS experiment(*)

A. C. M. Bulla(1)(2) on behalf of the CMS Collaboration

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

received 2 December 2024

Summary. — The study of vector boson scattering (VBS) and vector boson fusion (VBF) processes is crucial for testing the Standard Model (SM). This report focuses on two analyses: the electroweak production of opposite-sign $W^{\pm}W^{\mp}$ bosons in the fully leptonic decay channel and the electroweak WV production in the semileptonic decay channel. Additionally, it briefly covers the ongoing single W boson electroweak production analysis. The data sample analyzed corresponds to an integrated luminosity of 138 fb^{-1} of proton-proton collisions at $\sqrt{s} = 13$ TeV collected by the CMS detector. The VBS $W^{\pm}W^{\mp}$ signal is observed with a significance of 5.6 standard deviations (5.2 expected) and a measured fiducial cross-section of 10.2 ± 2.0 fb, consistent with the SM prediction of 9.1 ± 0.6 fb. The VBS WV process shows the first evidence with a significance of 4.4 standard deviations observed (5.1 expected) and a measured fiducial cross-section of 1.9 ± 0.5 fb, matching the SM prediction of 2.2 ± 0.1 fb. Finally, the VBF W analysis is expected to have a 15% uncertainty on the fiducial cross-section measurement.

1. – Introduction

Vector Boson Scattering (VBS) and Vector Boson Fusion (VBF) processes occur when the initial quarks radiate vector bosons that interact with each other. These rare processes, predicted by the Standard Model, have cross sections in the range of 10-1000 fb. They are deeply connected to the electroweak spontaneous symmetry-breaking mechanism and exhibit rich phenomenology, including triple and quartic gauge couplings, interactions with the Higgs boson, and Effective Field Theory (EFT) applications. VBS and VBF processes are characterized by distinct kinematic features, such as large pseudorapidity separation ($|\Delta\eta_{jj}|$), high invariant mass (m_{jj}), and low values of the Zeppenfeld variable [6] $|Z_X| = |\eta_X - \frac{1}{2}(\eta_{j_1} + \eta_{j_2})|$ with η_X being the pseudorapity of the X object. The accurate measurement of VBS and VBF cross-sections can provide insights into the Higgs and Gauge sectors of the standard model (SM) lagrangian and potentially reveal new physics beyond the Standard Model (BSM).

(*) IFAE 2024 - "Energy Frontier" session

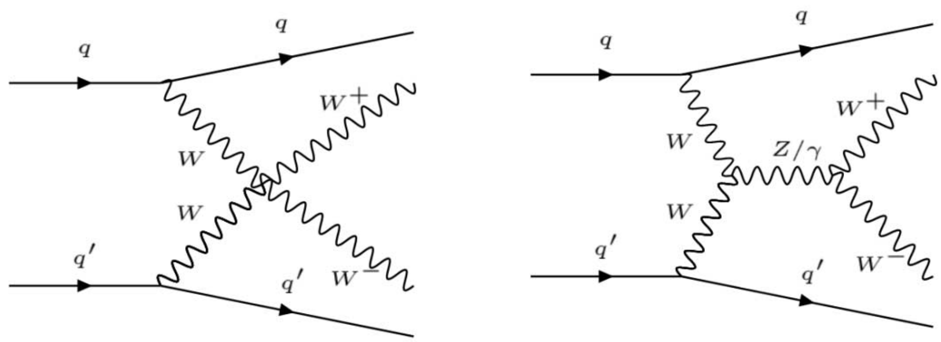

Fig. 1. – Examples of Feynman diagrams α_{EW}^6 for the EW production of $W^{\pm}W^{\mp}$ at LO.

2. – VBS opposite-sign W^+W^-

The first CMS observation of a purely electroweak production of a pair of opposite-sign ($W^{\pm}W^{\mp}$) bosons that decay in a fully leptonic final state is presented [1]. Figure 1 shows some examples of Feynman's diagrams at the order α_{EW}^6 that contribute to the signal.

Pre-selected events require two opposite-sign leptons, moderate missing transverse energy (MET), and at least two jets with high invariant mass and large pseudorapidity separation. Events with an additional lepton are rejected to suppress multiboson backgrounds. The analysis strategy varies significantly with the lepton flavour due to background contamination. The Drell-Yan processes heavily contaminate the same flavour (SF) channel, so the opposite flavour (OF) channel has a better signal-to-noise ratio.

The signal extraction is carried out through a binned maximum likelihood fit on the most discriminating variable distribution with signal and background templates. The fit is performed simultaneously in both the signal regions (SRs), defined by requiring no b-jet (using the DeepJet algorithm [3]) and split according to the $Z_{ll} \gtrless 1$ value with $Z_{ll} = \eta_{ll} - \frac{1}{2}(\eta_{j_1} + \eta_{j_2})$ being the Zeppenfeld variable for the di-lepton system. Background normalization is estimated using control regions (CRs) as single-bin templates and systematic uncertainties are treated as nuisance parameters in the fit.

In the SF category SRs, events are divided into three $m_{jj} - |\Delta\eta_{jj}|$ regions. For $m_{jj} > 300$ GeV and $|\Delta\eta_{jj}| > 3.5$, the m_{jj} distribution is used in the fit (fig. 2). In other cases, a single bin number of events is used. In both SRs in the OF category, two identical feed-forward deep neural networks (DNNs) are trained to separate the signal from $t\bar{t}$ and QCD-induced WW backgrounds. The DNNs are trained with 9 variables, including jet and lepton kinematic and angular properties. The DNN outputs are used as discriminating variables for the fit.

The signal's observed (expected) significance is 5.6 (5.2) standard deviations from the background-only hypothesis. The cross-section is measured in two fiducial volumes; in the inclusive one, the cross-section is 99 ± 20 fb (89 ± 5 fb expected), whereas in the one closer to the signal region the measured cross-section is 10.2 ± 2.0 fb (9.1 ± 0.6 fb expected), confirming the SM expectations. The statistical uncertainty is responsible for 50% of the relative error. In contrast, the most impactful systematic uncertainties are the normalization of the main background and the QCD scale on the $t\bar{t}$ processes.

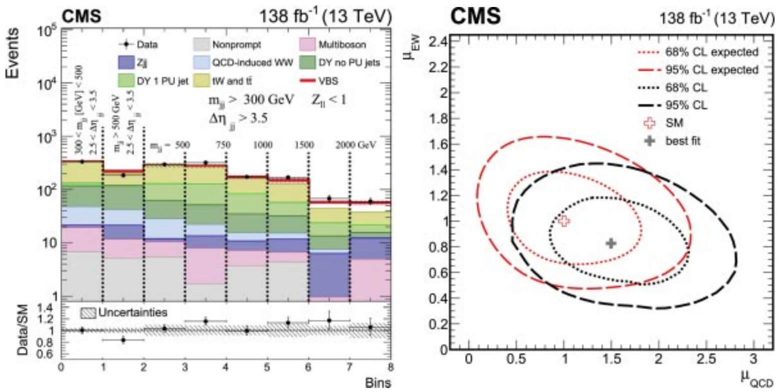

Fig. 2. – Left: OSWW distribution used in the fit for the $Z_{ll} > 1$ SR in SF category. Right: Simultaneous expected and observed fits for the electroweak and QCD-induced WV productions.

3. – VBS WV semileptonic

The first CMS evidence for the purely electroweak production of a WV pair of bosons in a semileptonic channel is presented [2]. The data sample analyzed corresponds to an integrated luminosity of 138 fb^{-1} of proton-proton collisions at $\sqrt{s} = 13$ TeV collected by the CMS detector.

Since the W boson decays leptonically, events are required to have one isolated lepton and a moderate amount of MET. The analysis strategy strongly depends on the topology of the hadronic decay of the V boson, which can either be a W or a Z boson. If the V is highly boosted, its decay products are reconstructed as a single AK8 jet [5]. In cases where the V boson is resolved, at least four AK4 jets are required in the final state. The signal regions (SRs) are defined by selecting events where the V boson decay is on-shell, and events with at least one b-jet, identified using the DeepJet algorithm, are vetoed.

To improve the sensitivity of the analysis, two distinct deep neural network (DNN) models are implemented for the different SRs. Both models are trained to distinguish signal events from all backgrounds. In the resolved topology, a more complex DNN with several input variables is used due to the additional jets, while in the boosted topology, a slightly simpler DNN is employed to account for the lower statistical availability of events. The most important variables for both models are the invariant mass of the tagging jets (m_{jj}), followed by the Zeppenfeld variable for the lepton.

Through a maximum likelihood fit, two different measurements are performed in an inclusive fiducial phase space. In the first case, only the cross-section of the pure electroweak production is measured, obtaining 1.9 ± 0.5 pb (2.23 ± 0.10 pb expected) with a significance of 4.4 (5.1) standard deviations to the background-only hypothesis. In the second case, also the QCD-induced cross-section is measured simultaneously in the same phase-space, obtaining 16.3 ± 3.2 pb (16.9 ± 2.5 pb expected) matching the expectation from the SM. Figure 2 shows the simultaneous expected and observed fits. Even in this case, the statistical uncertainty is responsible for 50% of the relative error.

4. – VBF W

For the pure electroweak production of a W boson in association with two jets, the analysis targets both inclusive and differential cross-section measurements. It also investigates the indirect effects of new physics through an EFT approach, studying the action of 9(+2) CP-even(odd) dimension-6 operators. The analyzed data sample corresponds to an integrated luminosity of 138 fb^{-1} of proton-proton collisions at $\sqrt{s} = 13$ TeV collected by the CMS detector.

Candidate signal events are preselected by requiring one isolated lepton and vetoing additional leptons. Two jets are also required, largely separated in $\Delta\eta_{jj}$ and with high invariant mass m_{jj} to mimic the typical VBS/VBF topology. The SR is then defined by requiring high m_T(W) and vetoing events with at least one b-jet, identified using the DeepJet algorithm.

A unique 64-64-64 DNN model is trained in the joint electron + muon SR with 8 input variables. Among the most important are the Zeppenfeld variable and m_{jj}, which is expected due to their intrinsic physical meaning. The model is trained to distinguish the signal from the main background, which is the mixed QCD-induced production of a W boson with two jets, representing 75% of the background in the SR. The DNN output is then used in the maximum likelihood fit, being the most discriminating variable. The normalizations of the mixed QCD-induced W production and the top-related backgrounds enter the fit as single bin constraints from their respective CR. Systematic uncertainties are treated as nuisance parameters in the fit. Using a MC Asimov toy [4], 15% uncertainty on the cross-section measurement is expected, with the greatest contribution coming from the QCD scale variation on the signal.

The goal of the unfolding measurement is to perform a differential cross-section measurement in generator-level bins of some interesting variables. The DNN is again used to define 2D variables. This technique helps in reducing the various backgrounds and provides sensitivity to the measurement. Additionally, the chosen reconstructed-level variable gives sensitivity to the generator-level bins in which the cross-section measurement is performed. The expected sensitivity strongly depends on the number of required generator-level bins, with the measurement being statistically limited.

Finally, the introduction of dimension-6 operators in the SM lagrangian allows to look for any deviation from SM predictions. Having trilinear gauge couplings in the LO description of the process, VBF analyses are sensitive to bosonic operators that mainly modify the coupling between bosons. An @LHE study has been conducted to select the operators to which the analysis is sensitive. Out of 59 operators from the Warsaw Basis, 25 have been found to affect the processes under study. Only 11 are then selected and will be studied at the reconstructed level. The fitting strategy again involves the DNN. The SR is split according to the DNN score output $\gtrless 0.5$. The low-DNN score region is used as a single-bin region while in the high-DNN region, different variables are considered for the fit, with $\Delta\phi_{jj}$ being one of the most sensitive to the dimension-6 operators effect. The channel is found to be competitive with other analyses and will contribute to setting limits on the most sensitive operators.

REFERENCES

[1] The CMS Collaboration, *Phys. Lett. B*, **841** (2023) 137495.
[2] The CMS Collaboration, *Phys. Lett. B*, **834** (2022) 137438.
[3] Bols E., Kieseler J., Verzetti M., Stoye M. and Stakia A., *J. Instrum.*, **15** (2020) P12012.
[4] Cowan G., Cranmer K., Gross E. and Vitells O., *Eur. Phys. J. C*, **71** (2011) 1554.
[5] Cacciari M., Salam G. and Soyez G., *JHEP*, **04** (2008) 063.
[6] Rainwater D., Szalapski R. and Zeppenfeld D., *Phys. Rev. D*, **54** (1996) 6680.