PROCEEDINGS OF SCIENCE

Measurement of same sign WW VBS processes at CMS with one hadronic tau in the final state

Andrea Piccinelli[a]

[a] *On behalf of the CMS Collaboration*
University of Notre Dame, South Bend, Indiana, USA

E-mail: andrea.piccinelli@cern.ch

This contribution presents an investigation into the electroweak production of same-sign W boson pairs (ssWW) with a hadronically decaying tau lepton (τ_h) in the final state. This study is based on proton-proton collisions at a center-of-mass energy of 13 TeV, as recorded by the CMS experiment at the CERN LHC, with an integrated luminosity of 138 fb^{-1}. This work delves deeper into the electroweak spontaneous symmetry breaking (EWSB) sector of the standard model (SM) of particle physics, focusing on Vector Boson Scattering (VBS) processes, which play a pivotal role in understanding the non-abelian structure of the SM. The article introduces a Machine Learning approach to enhance the identification of the final state in ssWW VBS processes. The results show an observed electroweak signal strength of $1.44^{+0.63}_{-0.56}$ corresponding to a signal significance of 2.74 standard deviations. The study concludes with the intention to explore the sensitivity to indirect New Physics effects using the Standard Model Effective Field Theory (SMEFT) framework.

The European Physical Society Conference on High Energy Physics (EPS-HEP2023)
21-25 August 2023
Hamburg, Germany

https://pos.sissa.it/

PoS(EPS-HEP2023)322

1. Introduction

The discovery of the Higgs Boson [1] provided evidence that fermions and bosons acquire their masses through the Brout-Englert-Higgs mechanism [2, 3], completing the standard model (SM) picture of particle physics. However, the electroweak spontaneous symmetry breaking (EWSB) sector is yet particularly interesting to investigate, allowing to probe the non-abelian structure of the SM. In this framework, Vector Boson Scattering (VBS) processes play a special role: the unitarity of the scattering of the longitudinal polarization of vector bosons is granted by a delicate cancellation of diagrams involving the mediation of Higgs boson. Therefore, for any deviation from the SM couplings of the Higgs with the vector bosons, the VBS cross-section would diverge increasing energy, leaving us with an incredible probe to investigate the Higgs sector and particularly the EWSB process. From a theoretical standpoint, the VBS process is defined, at Born level, as the purely EW processes of the order $O(\alpha^6)$, where α is the electroweak constant, and are therefore characterized by very low cross-section. In addition, VBS signatures also feature irreducible contribution to order $O(\alpha_s^2\alpha^4)$, usually referred to as QCD background. Typical Feynman diagrams for these processes are shown in Fig. 1. Among all the VBS processes, the scattering of a W pair of the same sign is one of the most favored channels, distinguished for the highest cross-section for the EW-mediated process and the large cross-section ratio between EW and QCD production modes. The present search [4] investigates a new and so far unexplored final state characterized by the decay of one of the scattered W bosons into a τ lepton, with the latter ultimately decaying into hadrons (hadronic tau, τ_h). The final state thus consists of a charged light lepton $\ell = e, \mu$, the corresponding neutrino ν_ℓ, a τ_h, the corresponding ν_τ, and two jets produced by the quarks that emitted the W boson pair. The τ lepton is included for the first time in the study of VBS processes and represents an indirect probe of possible deviation from SM due to its large mass and a preferential coupling to the Higgs boson. In this manuscript, we introduce a Machine Learning approach for the identification of the $\ell\tau_h\nu\nu jj$ final state in ssWW VBS processes and perform a measurement of its EW (EW+QCD) signal strength. The data collected by the CMS experiment from 2016 to 2018 are analyzed, corresponding to an integrated luminosity of 138 fb^{-1}.

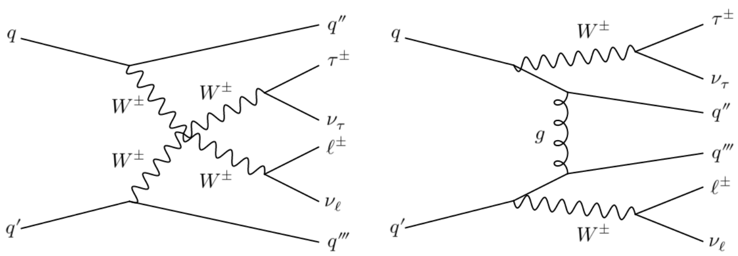

Figure 1: Representative Born level Feynman diagrams contributing to the process $pp \to \tau^\pm\nu_\tau\ell^\pm\nu_\ell jj$, $\ell = e,\ \mu$, with $O(\alpha^6)$ (left) and $O(\alpha_s^2\alpha^4)$ (right) couplings.

2. The CMS experiment

The central feature of the CMS apparatus is a superconducting solenoid of $6m$ internal diameter, providing a magnetic field of 3.8 T. Within the solenoid volume are a silicon pixel and silicon strip tracker, a lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass-and-scintillator hadron calorimeter (HCAL), each composed of a barrel and two endcap sections. Forward calorimeters extend the coverage in pseudorapidity η provided by the barrel and endcap detectors. Muons are detected in gas-ionization chambers embedded in the steel flux-return yoke outside the solenoid. A more detailed description of the CMS detector, together with a definition of the coordinate system used and the relevant kinematic variables [5].

3. Event reconstruction and selection

Collision events are collected using single-lepton triggers that require the presence of an isolated electron or muon with p_T of at least 30 or 35 GeV. Jet-based triggers are used to estimate from data the fake lepton background, i.e. jets that are misreconstructed as leptons (e, μ, τ_h). Events are reconstructed using the CMS particle-flow (PF) algorithm [7, 8] that reconstructs and identifies each individual particle with an optimized combination of all subdetector information. The missing transverse momentum vector $\vec{p}_T^{miss}$ is defined as the projection onto the plane perpendicular to the beam axis of the negative vector sum of the momenta of all reconstructed PF objects in an event. Its magnitude is referred to as p_T^{miss}. Jets are reconstructed by clustering PF candidates using the anti-kT algorithm [6] with a distance parameter of 0.4. Jets are calibrated in the simulation, and separately in data, accounting for energy deposits of neutral particles from pileup and any nonlinear detector response. The primary vertex (PV) is defined as the vertex with the largest value of the sum of the track p_T squared. The physics objects are derived only from the tracks assigned to the vertex as inputs by clustering them into jets, including leptons. The $\vec{p}_T^{miss}$ is recalculated only from those jets by summing their negative $\vec{p}_T$ vectors. Electron or muon candidates are required to originate from the PV, pass quality selection criteria, and be isolated relative to other activities in the event. These objects are reconstructed by associating a track reconstructed in the tracking detectors with either a cluster of energy in the ECAL or a track in the muon system. Electron and muon candidates must pass certain identification criteria to be further selected in the analysis. For the "loose" identification, they must satisfy $p_T > 10$ GeV and $|\eta| < 2.5(2.4)$ for electrons (muons). At the final stage of the lepton selection the "tight" working points criteria are chosen, including requirements on the impact parameter of the candidates with respect to the PV and their isolation with respect to other particles in the event, together with higher thresholds for their p_T. Hadronic τ lepton decays (τ_h) are reconstructed with the CMS hadrons-plus-strips (HPS) algorithm [9] and are identified with a DNN-based algorithm, called DeepTau, capable to discriminate taus against jets, electrons, and muons, using three different classifiers [10]. For the latter, the tightest working points (WPs) that return the best signal sensitivity against the background are implemented. An event is considered consistent with the topology of the signal if it presents only one "tight" τ_h, one "tight" electron/muon, and at least two jets with a pseudorapidity separation $|\Delta\eta| > 2.5$. Among all the possible jet pairs that satisfy the latter request, the pair with the highest invariant mass m_{jj} is chosen.

4. Analysis strategy

The analysis targets the VBS production of pairs of W bosons with the same sign, with one decaying to a hadronic τ lepton, in association with two jets originating from the scattered incoming partons. Events are first selected requiring one electron or muon, $\ell = e$, μ, one τ_h, and vetoing additional "loose" leptons (e, μ, τ_h). In addition,the signal region (SR) and a series of control regions (CRs) are defined to estimate and validate the background predictions, as specified in the following lines. The signal region (SR) is designed to enhance the VBS signal events against all backgrounds. Events with same-sign $\ell\tau_h$ pair, $p_T^{miss} > 50$ GeV and $M_{jj} > 500$ GeV are selected. In this region, almost 95% of the background is due to the non-prompt lepton contribution, about 2% to Z/γ^*+jets background, 1% to dileptonic $t\bar{t}$ production, and the remaining processes. The largest background contribution arises from events containing hadronic jets mistakenly reconstructed as leptons (e, μ, τ_h) and is referred to as *non-prompt* lepton background. Non-prompt leptons are produced mainly by QCD-mediated multijet, associated $W+$ jets, and hadronic and semileptonic $t\bar{t}$ production and they are estimated with the data-driven method described in detail in Ref. [**?**]. The non-prompt CR is intended to validate the data-driven fake leptons background contribution, and to contain events that include misreconstructed leptons, mainly from W+jets and QCD multijet events. Finally, a $t\bar{t}$ and an opposite-sign control region ($t\bar{t}$ and OS CR) are defined to constrain MC simulations of these background sources. Because of the large background and complex signal topology, nine significant features to separate signals and backgrounds, chosen from a larger set of kinematic observables, are condensed in a single Machine Learning discriminator (a feed-forward Deep Neural Network, DNN). The DNN implemented is optimized to discriminate signals against the main sources of background. The statistical analysis is implemented with a maximum likelihood (ML) fit to extract the signal strength, the ratio of the signal yield observed to that predicted by the model. The inputs to the fit are the distributions in the DNN output of the data, the signal, and the backgrounds estimated according to the description in the next Section.

5. Results

From the ML fit, we extract values of the signal strength and the statistical significance in two separate approaches: in one case, we measure the purely EW signal strength keeping the QCD ssWW production contribution fixed to the SM prediction; in a second case, we measure the signal strength considering as signal the EW and QCD ssWW processes together. For the primary result of the EW signal strength measurement, the DNN output distributions of the SR and the $t\bar{t}$ and OS CRs are shown in Figure 2 for both the electron and muon flavors. In the figure, the data are compared with the background estimated before (pre-fit) and after (post-fit) the simultaneous fit of the signal and control regions. The pulls shown in the lower panels are defined as the difference between the data and the post-fit background prediction divided by the quadratic difference between the uncertainties in the data and the post-fit yields. The quadratic difference of the uncertainties is taken to account for the correlation between the data and the post-fit prediction. The observed (expected) EW signal strength is $1.44^{+0.63}_{-0.56}$ ($1.00^{+0.60}_{-0.53}$), corresponding to a signal significance of 2.74 standard deviations with 1.94 expected. The simultaneous measurement of the EW and QCD-associated diboson production is measured with an observed (expected) significance of 2.87 (2.04) standard deviations.

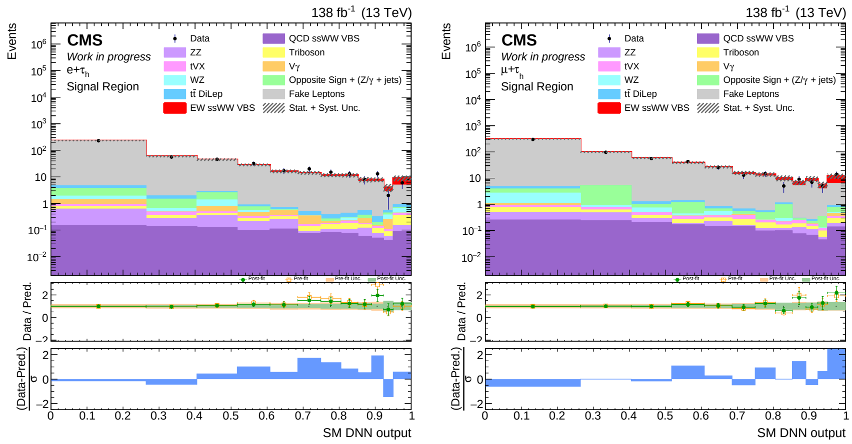

Figure 2: Distribution of SM DNN output for the $e + \tau_h$ (left) and $\mu + \tau_h$ (right) channels for the full data sample in SR. Data points are overlaid on the post-fit background (stacked histograms). The overflow is included in the last bin. The middle panels show ratios of the data to the pre-fit background prediction and post-fit background yield in yellow and green, respectively. The yellow (green) bands in the middle panels indicate the systematic component of the pre-fit (post-fit) uncertainty.

6. Conclusions

Electroweak production of same-sign W boson pairs (ssWW) with a hadronically decaying τ (τ_h) in the final state is investigated for the first time. The analysis is performed using a sample of proton-proton collisions at $\sqrt{s} = 13$ TeV recorded by the CMS experiment at the CERN LHC, corresponding to an integrated luminosity of 138 fb^{-1}. The sensitivity to ssWW VBS processes with a τ_h in the final state is estimated by implementing ML-based algorithms to improve the discrimination of signal events against the main backgrounds. The observed EW signal strength is $1.44^{+0.63}_{-0.56}$, corresponding to a signal significance of 2.74 standard deviations with 1.94 expected, and it is measured keeping the QCD associated diboson production fixed to the standard model prediction. The simultaneous measurement of the EW and QCD associated diboson production is measured with an observed (expected) significance of 2.87 (2.04) standard deviations. In the near future, this study aims at studying the sensitivity to indirect New Physics effects by implementing the Standard Model Effective Field Theory (SMEFT) framework. Assuming that New Physics with energy scale $\Lambda_{NP} \gg \Lambda_{SM}$ induces only perturbative effects in VBS processes, SMEFT can be properly modeled by introducing the following effective Lagrangian:

$$\mathcal{L}_{eff} = \mathcal{L}_{SM} + \sum_{D>4} \sum_{i} \frac{c_i^{(D)}}{\Lambda_{NP}^{D-4}} O_i^{(D)},$$

Effects due to relevant operators $O_i^{(D)}$ with $D = 6, 8$ are going to be investigated with the analysis workflow illustrated above.

References

[1] Aad, Georges et al. (ATLAS, CMS Collaboration), "Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s}$ = 7 and 8 TeV," *JHEP* **08** (2016) 045. [hep-ex/1606.02266] doi:10.1007/JHEP08(2016)045

[2] Higgs, Peter W., "Spontaneous Symmetry Breakdown without Massless Bosons," *Phys. Rev.* **145**, 1156–1163 (1966). doi:10.1103/PhysRev.145.1156

[3] Englert, F. and Brout, R., "Broken Symmetry and the Mass of Gauge Vector Mesons," *Phys. Rev. Lett.* **13**, 321–323 (1964). doi:10.1103/PhysRevLett.13.321

[4] CMS Collaboration, "Measurement of $W^{\pm}W^{\pm}$ scattering in proton-proton collisions at $\sqrt{s}$ = 13 TeV in final states with one tau lepton," CMS PAS SMP-22-008, CERN, Geneva, 2023. http://cds.cern.ch/record/2867989

[5] The CMS Collaboration, "The CMS experiment at the CERN LHC," *Journal of Instrumentation* **3**, S08004 (2008). doi:10.1088/1748-0221/3/08/s08004

[6] Cacciari, Matteo and Salam, Gavin P. and Soyez, Gregory, "The anti-k_t jet clustering algorithm," *JHEP* **04** (2008) 063. [hep-ex/0802.1189] doi:10.1088/1126-6708/2008/04/063

[7] Sirunyan, A. M. et al. (CMS Collaboration), "Particle-flow reconstruction and global event description with the CMS detector," *JINST* **12** (2017) P10003. [physics.ins-det/1706.04965] doi:10.1088/1748-0221/12/10/P10003

[8] CMS Collaboration, "Commissioning of the Particle-Flow Event Reconstruction with the first LHC collisions recorded in the CMS detector," CMS PAS PFT-10-001, 2010.

[9] Sirunyan, A. M. et al. (CMS Collaboration), "Performance of reconstruction and identification of τ leptons decaying to hadrons and ν_τ in pp collisions at $\sqrt{s}$ = 13 TeV," *JINST* **13** (2018) P10005. [hep-ex/1809.02816] doi:10.1088/1748-0221/13/10/P10005

[10] Sirunyan, A. M. et al. (CMS Collaboration), "Performance of reconstruction and identification of τ leptons decaying to hadrons and ν_τ in pp collisions at $\sqrt{s}$ = 13 TeV," *JINST* **13** (2018) P10005. [hep-ex/1809.02816] doi:10.1088/1748-0221/13/10/P10005

PoS(EPS-HEP2023)322