Observation of electroweak production of same-sign W boson pairs in the two jet and two same-sign lepton final state in proton-proton collisions at 13 TeV

The CMS Collaboration

Abstract

The observation of electroweak production of same-sign W boson pairs in proton-proton collisions at 13 TeV is reported. The data sample corresponds to an integrated luminosity of 35.9 fb$^{-1}$ collected with the CMS detector. Events are selected by requiring exactly two leptons of the same charge, moderate missing momentum, and two jets with large rapidity separation and large dijet mass. The observed significance is 5.5 standard deviations, where a significance of 5.7 standard deviations is expected based on the standard model. A cross section measurement in a fiducial region is reported. Bounds on the structure of quartic vector-boson interactions are given in the framework of dimension-eight effective field theory operators, together with upper limits on the production of doubly charged Higgs bosons.
The standard model (SM) of particles physics provides an exceptionally accurate description of observations from many accelerator and non-accelerator based experiments. The discovery of a Higgs boson [1–3] confirmed that W and Z gauge bosons acquire mass through the Higgs mechanism. This prediction motivates further study of the mechanism of electroweak symmetry breaking through measurements of vector boson scattering (VBS) processes. Physics models beyond the SM predict enhancements for VBS through modifications to the Higgs sector or the presence of additional resonances [4, 5].

This document presents a study of VBS in pp collisions at \( \sqrt{s} = 13 \) TeV. The data sample corresponds to an integrated luminosity of \( 35.9 \pm 0.9 \) fb\(^{-1}\) collected with the CMS detector [6] at the LHC in 2016.

The main goal of the analysis is to identify same-sign W boson pair events produced purely via the electroweak interaction. Candidate events have exactly two identified leptons of the same charge, moderate missing transverse energy, and two jets with large rapidity separation and dijet mass. Requiring same-sign lepton events, reduces the contribution from the strong production of W boson pair events, making the experimental signature an ideal topology for VBS studies.

Figure 1 shows representative Feynman diagrams for the electroweak and QCD induced production.

Figure 1: Representative Feynman diagrams for the electroweak and QCD induced same-sign W boson pair production.

An excess of events could signal the presence of anomalous quartic gauge couplings (aQGC) [7] or the existence of a new resonance, such as a doubly charged Higgs boson. These particles are predicted in Higgs sectors beyond the SM where weak isotriplet scalars are included [8, 9]. They can be produced via weak vector-boson fusion (VBF) and decay to pairs of same-sign W bosons [10].

First experimental results were reported by the ATLAS and CMS Collaborations based on approximately 20 fb\(^{-1}\) of data collected at a center-of-mass energy of 8 TeV [11, 12]. The reported observed significance is 3.6 (2.0) standard deviations for the ATLAS (CMS) study, where a significance of 2.8 (3.1) standard deviations was expected based on the SM prediction.

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 tungsten crystal electromagnetic calorimeter, and a brass and scintillator hadron calorimeter, each composed of a barrel and two endcap sections. Forward calorimeters extend the pseudorapidity coverage provided by the barrel and endcap detectors. Muons are measured in gas-ionization detectors 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, can be found in Ref. [6].
The signal and background processes are generated using several Monte Carlo (MC) tools. The detector response is simulated by the GEANT4 package [13] using a detailed description of the CMS detector. Simultaneous proton-proton interactions overlapping with the event of interest are included in the simulated samples. The amount of additional interactions per event in the simulation corresponds to the conditions observed in the 13 TeV data collected in 2016, with a mean of approximately 20.

The leading-order (LO) event generator MADGRAPH 5.2 [14] is used to produce samples of diboson production via diagrams with two or fewer quantum chromodynamics (QCD) and up to six electroweak vertices. This includes two categories of diagrams: those with exactly two QCD vertices which we refer to as QCD production and those with no QCD vertices, which we refer to as electroweak (EW) production. We only consider EW production as the signal in the analysis, while QCD production is considered as background. This background from QCD production is small and differs in kinematics. The interference between the EW and QCD processes is found to be a few percent in the signal region and considered as systematic uncertainty. The EW category includes diagrams with WWW quartic interactions and diagrams where two same-sign W bosons scatter through the exchange of a Higgs boson, a Z boson, or a photon.

The WZ and ZZ production, via qQA annihilation, and the Wγ process are generated at LO with MADGRAPH. The gg → ZZ process is simulated with MCFM [15]. The Z+jets, Zτ, tQ, tW, tZ, WWZ, WWW, and ZZZ samples are generated with MG_AMC@NLO 2.3 [14]. The PYTHIA 8.205 [16, 17] package is used for parton showering, hadronization, and the underlying event simulation, with tune CUETP8M1 [18, 19]. The NNPDF 3.0 [20] set is used as the default set of parton distribution functions.

The final states considered are μ⁺μ⁺νμνμjj, e⁺e⁺νeνejj, e⁺μ⁺νeνμjj, and their charge conjugates. The τ-lepton decays to electrons and muons are included. A suite of signal and control triggers are designed for this analysis. The logical OR of single and double lepton triggers has a high efficiency, larger than 99.8% after all other selection requirements are applied.

A particle-flow (PF) algorithm [21, 22] is used to reconstruct all observable particles in the event. It combines all subdetector information to reconstruct individual particles and identify them as charged hadrons, neutral hadrons, photons, and leptons. The missing transverse momentum \( p_T^{\text{miss}} \) is defined as the magnitude of the negative vector sum of the transverse momenta of all reconstructed particles (charged and neutral) in the event.

The selection of events aims to single out same-sign lepton events with the VBS topology while reducing the top quark, Drell–Yan, and WZ background processes. Two same-sign lepton candidates, muons or electrons, with \( p_T > 25(20) \) GeV for the leading (trailing) lepton and \( |\eta| < 2.4 (2.5) \) for muons (electrons) are required. Electrons and muons are required to be isolated from other charged and neutral particles in the event. Jets are reconstructed using the anti-\( k_t \) clustering algorithm [23] with a distance parameter \( R = 0.4 \), as implemented in the FASTJET package [24, 25]. Events are required to have at least two selected jets with \( E_T > 30 \) GeV and \( |\eta| < 5.0 \). The VBS topology is targeted by requiring that the two jets leading in \( p_T \) have large dijet mass, \( m_{jj} > 500 \) GeV, large pseudorapidity separation, \( |\Delta\eta_{jj}| > 2.5 \), and \( \max(z^*_1) < 0.75 \), where \( z^*_1 = |\eta_1 - (\eta_{j1} + \eta_{j2})/2|/|\Delta\eta_{jj}| \) is the Zeppenfeld variable [26].

Identification techniques of decays of the bottom quark are used to veto top-quark backgrounds (tQ and tW). They are based on bottom quark jet tagging criteria that combine the information of displaced tracks with the information of secondary vertices associated to the jet using a multivariate technique, and on the presence of a soft muon in the event from the semileptonic
decay of the bottom-quark \[27\]. A minimum dilepton mass, \(m_{\ell\ell} > 20\) GeV, is required to reduce the \(W +\) jets and top-quark background processes. To reduce the background from WZ production, events with a third, loosely identified lepton with \(p_T > 10\) GeV or an identified tau hadronic decay with \(p_T > 18\) GeV are rejected. Drell–Yan events can be selected if the charge of one lepton is measured incorrectly. To reduce this background, \(|m_{\ell\ell} - m_Z| > 15\) GeV is required for \(e^+e^-\) events. The charge confusion in dimuon events is negligible. The Drell–Yan background is further reduced by requiring \(p_T^{\text{miss}} > 40\) GeV.

The non-prompt lepton background originating from leptonic decays of heavy quarks, hadrons misidentified as leptons, and electrons from photon conversions, is suppressed by the identification and isolation requirements imposed on muons and electrons. The remaining contribution from the non-prompt lepton background is estimated directly from data as follows. A control sample is defined by one lepton candidate that passes the standard lepton selection criteria, and another lepton candidate that fails the criteria but passes a looser selection, resulting in a sample of “pass-fail” lepton pairs. The efficiency \(\epsilon_{\text{pass}}\) for an object that satisfies the loose lepton requirements to pass the standard selection is determined from an independent sample dominated by events with non-prompt leptons from QCD multijet processes. This efficiency, parametrized as a function of \(p_T\) and \(\eta\) of the lepton, is then used to weight the events in the pass-fail sample by \(\epsilon_{\text{pass}}/(1 - \epsilon_{\text{pass}})\) to obtain the estimated contribution from the non-prompt lepton background in the signal region. The uncertainties from the determination of \(\epsilon_{\text{pass}}\) dominate the overall uncertainty of this method, arising from the statistical uncertainty in the measurement of the tight-to-loose ratios, from systematical uncertainties derived by comparing alternative methods, and from testing the closure of the method in simulated background events.

A WZ → 3\(\ell\nu\) control region is defined by requiring a third fully identified lepton with with \(p_T > 10\) GeV and an opposite-sign same-flavor lepton pair with a mass consistent with a Z boson decay. The contribution of opposite-sign (wrong-sign) lepton events to the signal region due to charged misidentification is estimated by applying data-to-simulation scale factors to charge-misidentified electrons in bins of \(\eta\). The charge-misidentification rates and the scale factors are estimated using Z boson events. The charge-misidentification rate is found to be between about 0.01% in the barrel region and about 0.3% in the endcap regions for electrons, while it is negligible for muons.

The signal efficiencies are estimated using simulated samples. In the statistical analysis, shape and normalization uncertainties are considered. The shape uncertainties are estimated by remaking the distribution of a given observable after considering the systematic variations for each source of uncertainty. The lepton trigger, reconstruction, and selection efficiencies are measured using \(Z/\gamma^* \rightarrow \ell^+\ell^-\) events that provide an unbiased sample with high purity. The estimated uncertainty is within 2% per lepton. The uncertainties due to the momentum scale for electrons and muons are also taken into account and contribute about 1%. The jet energy scale and resolution uncertainties give rise to an uncertainty in the yields of up to 7%. The uncertainty in the event selection efficiency for events with neutrinos yielding genuine \(p_T^{\text{miss}}\) in the final state is assessed and leads to an uncertainty of about 1%. The uncertainty in the estimated event yields, which is related to the top-quark veto, is evaluated by using a \(Z/\gamma^* \rightarrow \ell^+\ell^-\) sample with at least two reconstructed jets and is found to be up to 3%. The statistical uncertainty in the yield of each bin and for each process is also taken into account. The uncertainty of 2.5% in the integrated luminosity is considered for all processes estimated from simulation. The normalization of the processes with misidentified leptons has a 30% systematic uncertainty. The WZ normalization uncertainty is 20-40%, dominated by the small number of events in the trilepton control region. Theoretical uncertainties are estimated by varying the renormaliza-
tion and factorization scales up and down by a factor of two from their nominal value in the event, and found to be 12% for the signal normalization and 20% for the triboson background normalization. The interference between the EW signal and the QCD background processes are expected to be small and considered with a systematic uncertainty of up to 4.5% in the statistical analysis, estimated using the PHANTOM 1.2.8 generator [28]. A PDF uncertainty of 5% in the normalization of the signal is included. The systematic uncertainties of the background normalizations are taken into account.

The estimated signal and background yields, as well as the observed data yields, are shown in Table 1 for all six channels separately and their sum. The two dominating sources of background events are non-prompt leptons and the WZ → 3ℓν process. The distributions of mjj and mℓℓ in the signal region are shown in Fig. 2. An excess of events with respect to the background-only hypothesis is observed. In order to quantify the significance of the observation of the EW production of same-sign W boson pairs, a statistical analysis of the event yields is performed with a 2-dimensional fit of mjj and mℓℓ variables. The background contributions are allowed to vary within the estimated uncertainties. The WZ background contribution in the signal region is constrained as function of mjj using the control region. The observed (expected) significance was found to be 5.5 (5.7) standard deviations. The best-fit signal strength for the signal hypothesis is 0.90 ± 0.22 with respect to the SM expectation.

The cross section is extracted for a fiducial signal region. This region is defined by requiring two same-sign leptons from W boson leptonic decays with pT > 20 GeV and |ηℓ| < 2.5, two jets with pT > 30 GeV and |ηj| < 5.0, mjj > 500 GeV, and |Δηjj| > 2.5. In this definition, W → τν → ℓννν decays are excluded. The measured cross section is corrected for the acceptance in this region using the MadGraph MC generator, which is also used to estimate the theoretical cross section. The predicted theoretical cross section at leading order is 4.25 ± 0.21 fb, where 5% uncertainty is taken from the QCD scale variations. The fiducial cross section is measured to σfid(W ± W ± jj) = 3.83 ± 0.66 (stat) ± 0.35 (syst) fb, in agreement with the expectation. The overall efficiency within the fiducial region is 34.8 ± 0.3 (stat)%, while the fraction of events outside the fiducial region and selected at the reconstruction level is 20.6 ± 0.3 (stat)%.

Various extensions to the SM alter the couplings between vector bosons. Reference [7] proposes nine independent C- and P-conserving dimension-eight effective operators to modify the quartic couplings. In this case, the mℓℓ distributions both in the signal and WZ regions are used to extract the results. The observed and expected 95% CL limits on the nine coefficients are
shown in Table 2, where all the results are obtained by varying the effective operators one by one. The table also shows CMS limits from the LHC Run-I. The effect of possible aQGCs on the WZ process in the signal region is negligible because the background is normalized using data.

<table>
<thead>
<tr>
<th>Observed limits (TeV$^{-4}$)</th>
<th>Expected limits (TeV$^{-4}$)</th>
<th>Run-I limits (TeV$^{-4}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_{S0}/\Lambda$</td>
<td>[-7.7, 7.7]</td>
<td>[-7.0, 7.2]</td>
</tr>
<tr>
<td>$f_{S1}/\Lambda$</td>
<td>[-21.6, 21.8]</td>
<td>[-19.9, 20.2]</td>
</tr>
<tr>
<td>$f_{M0}/\Lambda$</td>
<td>[-6.0, 5.9]</td>
<td>[-5.6, 5.5]</td>
</tr>
<tr>
<td>$f_{M1}/\Lambda$</td>
<td>[-8.7, 9.1]</td>
<td>[-7.9, 8.5]</td>
</tr>
<tr>
<td>$f_{M6}/\Lambda$</td>
<td>[-11.9, 11.8]</td>
<td>[-11.1, 11.0]</td>
</tr>
<tr>
<td>$f_{M7}/\Lambda$</td>
<td>[-13.3, 12.9]</td>
<td>[-12.4, 11.8]</td>
</tr>
<tr>
<td>$f_{T0}/\Lambda$</td>
<td>[-0.62, 0.65]</td>
<td>[-0.58, 0.61]</td>
</tr>
<tr>
<td>$f_{T1}/\Lambda$</td>
<td>[-0.28, 0.31]</td>
<td>[-0.26, 0.29]</td>
</tr>
<tr>
<td>$f_{T2}/\Lambda$</td>
<td>[-0.89, 1.02]</td>
<td>[-0.80, 0.95]</td>
</tr>
</tbody>
</table>

Table 2: Observed and expected 95% CL limits on the coefficients for BSM higher order (dimension-eight) operators in the EFT Lagrangian. The last column is summarizing the LHC Run-I observed limits obtained by CMS.

Doubly charged Higgs bosons are predicted in models that contain a Higgs triplet field. Some of these scenarios predict same-sign lepton events from $W^\pm W^\pm$ decays with a VBF topology and the couplings depend on $m(H^\pm)$ and the parameter $\sin\theta_{H^\pm}$ or $s_{H^\pm}$, where $s_{H^\pm}^2$ denotes the fraction of the W boson mass squared generated by the vacuum expectation value (vev) of the triplets. The cross section for VBF production of $H^{\pm\pm}$ and decay to $W^\pm W^\pm$ is directly proportional to $s_{H^\pm}^2$. The remaining five parameters in the model are adjusted to get the given $m_{H^{\pm\pm}}$ hypothesis while requiring one of the scalar singlets to have a mass of 125 GeV. The Georgi–Machacek model of Higgs triplets [31] is considered. By using the $(m_{jj}, m_{\ell\ell})$ two-dimensional distributions to discriminate between signal and background processes, 95% CL upper limits on $\sigma_{VBF}(H^{\pm\pm}) \times B(H^{\pm\pm} \rightarrow W^\pm W^\pm)$ are derived as shown in Fig. 3 (left). The WZ background

Figure 2: Distributions of $m_{jj}$ (left) and $m_{\ell\ell}$ (right) in the signal region. The normalization of the predicted signal and background distributions corresponds to the result of the fit. The hatched bars include statistical and systematic uncertainties. For illustration, the doubly charged Higgs boson signal normalized to a cross section of 0.1 pb (left) and the distribution with aQGCs are shown. The histograms for other backgrounds include the contributions from QCD WW, $W\gamma$, wrong-sign events, DPS, and VVV processes.
contribution in the signal region is constrained using the control region. The excluded $s_H$ values as a function of $m(H^{±±})$ are shown in Fig. 3 (right). As discussed before, the WZ background contribution in the signal region is constrained using the control region. The blue region shows the parameter space for which the $H^{±±}$ total width exceeds 10% of $m(H^{±±})$, where the model is not applicable [32]. The observed limit excludes $s_H$ values greater than 0.18 and 0.44 at $m(H^{±±}) = 200$ and 1000 GeV, respectively. Results on this model have also been reported by the CMS collaboration in a search for VBF $H^{±} → W^{±}Z$ production [33].

In summary, we present a first observation of electroweak production of same-sign W boson pairs in proton-proton collisions at 13 TeV. The data sample corresponds to an integrated luminosity of 35.9 fb$^{-1}$ collected with the CMS detector. Events are selected by requiring exactly two leptons of the same charge, moderate missing transverse energy, and two jets with large rapidity separation and large dijet mass. The two dominating sources of background events after the event selection are non-prompt leptons and the WZ → 3ℓν. The event yield of the signal process is extracted using a 2-dimensional fit of $m_{jj}$ and $m_{ℓℓ}$ variables. The observed significance is 5.5 standard deviations, where a significance of 5.7 standard deviations is expected based on the standard model. A cross section measurement in a fiducial region is reported. No evidence for anomalous quartic gauge couplings is observed, and stringent bounds on the structure of quartic vector-boson interactions are given in the framework of dimension-eight effective field theory operators, together with upper limits on the production of doubly charged Higgs bosons.

References


