CMS Physics Analysis Summary

Contact: cms-pag-conveners-exotica@cern.ch
2016/06/08

Search for heavy neutrinos and third-generation leptoquarks in final states with two hadronically decaying \( \tau \) leptons and two jets in proton-proton collisions at \( \sqrt{s} = 13 \) TeV

The CMS Collaboration

Abstract

A search for heavy, right-handed neutrinos, \( N_l \), and right-handed \( W_R \) bosons, which arise in the left-right symmetric extensions of the standard model, has been performed. The search focuses on the scenario where the \( W_R \) and \( N_l \) decay chains result in a pair of high-\( p_T \) \( \tau \) leptons which decay hadronically, in addition to two high-\( p_T \) jets and missing transverse energy from the \( \tau \) lepton decays. The analysis is performed using 2.1 fb\(^{-1}\) of data collected by the CMS experiment in 2015 at \( \sqrt{s} = 13 \) TeV. For models with strict left-right symmetry, and assuming only \( N_l \) flavor contributes significantly to the \( W_R \) decay width, \( W_R \) masses below 2.35 (1.63) TeV are excluded at a 95% confidence level, assuming the \( N_l \) mass is 0.8 (0.2) times the mass of \( W_R \) boson. To illustrate the sensitivity of this analysis to other new physics models, focus is also placed on pair production of third-generation scalar leptoquarks with decay into \( \tau \tau bb \). Third-generation scalar leptoquarks with masses below 740 GeV are excluded, assuming a 100% branching fraction for the leptoquark decay to a \( \tau \) lepton and a bottom quark.
1 Introduction

In the standard model (SM) the neutrinos of the three generations are considered to be massless. However, the observation of neutrino oscillations implies a non-zero mass and points to new physics models. Results from neutrino oscillation experiments together with cosmological constraints imply very small neutrino masses [1–3]. The leading model which generates light neutrino masses is the “seesaw” mechanism, which is realized in various schemes [1–3]. In the simplest case, the small observed neutrino masses are generated as a result of a heavy neutrino state $N$. In this model, the SM neutrino mass is given by $m_{\nu} \sim y^2_{\nu} v^2 / m_N$, where $y_{\nu}$ is a Yukawa coupling, $v$ the Higgs vacuum expectation value in the SM, and $m_N$ the mass of the heavy neutrino state. If the seesaw mechanism is to explain the masses of the known neutrinos, the light and heavy neutrinos must be Majorana particles, so processes that violate lepton number conservation by two units would be possible. Therefore, searches for heavy Majorana neutrinos using hadron colliders are important in resolving the nature of neutrinos and the origin of neutrino masses.

One way to confer mass to neutrinos, in the context of the see-saw mechanism, is provided by the left-right symmetry extension (LRSM), in which the SM group SU(2)$_L$ has a right-handed counterpart, originally introduced to explain the non-conservation of parity in weak interactions. The new SU(2)$_R$ group, similar to the SU(2)$_L$, predicts the existence of three new gauge bosons, $W^+_R$ and $Z'$, and three heavy right-handed neutrino states $N_l (l = e, \mu, \tau)$, partners of the light neutrinos states $\nu_l$. A reference process allowed by this model is the production of a $W_R$ that decays in a heavy neutrino $N_\tau$ and a lepton of the same generation. Searches for heavy neutrinos have been performed in the $\mu\mu jj$ and $ee jj$ channels assuming $N_\tau$ is too heavy to play a role in the decay of $W_R$ [4]. Of particular interest here is the scenario where the above decay chain results in a pair of high-$p_T \tau$ leptons and two energetic jets produced by the decay $W_R \rightarrow \tau + N_\tau \rightarrow \tau + \tau qq'$. A similar $\tau\tau jj$ final state can be realized in other extensions of the SM. For example, many extensions of the SM predict a new scalar or vector boson, called leptoquark (LQ), which carries non-zero lepton and baryon numbers, as well as color and fractional electric charge [5, 6]. Such particles are motivated by a unified description of quarks and leptons [7]. The combination of baryon and lepton numbers implies that pair production of third-generation LQs can mediate quark-lepton transitions and decay into $\tau\tau bb$.

The analysis presented is a general search for physics beyond the SM (BSM) in events with two energetic $\tau$ leptons and two energetic jets in the detector. To illustrate the sensitivity of this search for BSM processes, two benchmark new physics scenarios are considered: (1) LRSM with decay chain $pp \rightarrow W_R \rightarrow \tau + N_\tau \rightarrow \tau + \tau qq'$; (2) pair production of third-generation scalar leptoquarks with decay into $\tau\tau bb$.

A $\tau$ lepton is the heaviest known lepton with a mass of 1.777 GeV and a lifetime of $2.9 \times 10^{-13}$ seconds. Around one third of all $\tau$ leptons decay to $e/\mu$ and two neutrinos, and the remainder decay into hadronic jets and one neutrino ($\tau_h$). In the latter case, a $\tau_h$ consists of one, three, or (rarely) five charged mesons usually accompanied by one or more neutral pions. The channel in which the pair of $\tau$ leptons decays to $\tau_h \tau_h$ is considered. Because the hadronic decay of the $\tau\tau$ system has two associated neutrinos, missing transverse energy ($E_T^{\text{miss}}$) is present. Unlike heavy neutrino searches in the $eejj$ or $\mu\mu jj$ [8, 9] final states, due to the presence of neutrinos from the $\tau$-lepton decays, the $W_R$ resonance mass in the $\tau_h \tau_h$ channel cannot be fully reconstructed. To successfully distinguish between signal and backgrounds, the visible $\tau$-lepton decay products,
two jets, and the $E_T^{\text{miss}}$ are used to reconstruct the partial mass:

$$m(\tau_{h,1}, \tau_{h,2}, j, j, E_T^{\text{miss}}) = \sqrt{(E_{\tau_1} + E_{\tau_2} + E_j + E_j + E_T^{\text{miss}})^2 - (p_{\tau_1}^\gamma + p_{\tau_2}^\gamma + p_j^\gamma + p_j^\gamma + E_T^{\text{miss}})^2}. \quad (1)$$

The partial mass is expected to be large in the heavy neutrino case, $\langle m(\tau_{h,1}, \tau_{h,2}, j, j, E_T^{\text{miss}}) \rangle \approx m(W_R)$. The heavy neutrino search strategy is to look for a broad enhancement in the mass distribution consistent with new physics. For the pair production of leptoquarks, the scalar sum of the transverse momenta ($p_T$) of the decay products, $S_T = p_T^{\tau_{h,1}} + p_T^{\tau_{h,2}} + p_T^j + p_T^j$, is expected to be large ($\langle S_T \rangle \approx m_{LQ}$). In this case the strategy is similar to other leptoquark analyses and involves searching for a broad enhancement in the high $S_T$ part of the spectrum.

In hadronic $\tau$-lepton decays, there is only one $\tau$ neutrino (anti-neutrino) present, leading to a higher visible momentum of $\tau$ decay products compared to leptonic decays of $\tau$ leptons ($\tau_l$) on average. Therefore, $m(\tau_{h}, j, j, E_T^{\text{miss}})$ and $S_T$ is typically higher than in channels containing $\tau_l$. This characteristic combined with the $\approx 42\%$ branching ratio of $\tau\tau \rightarrow h\tau_h$ makes this analysis a promising channel in the search for new physics. Because a $\tau_h$ resembles QCD jets, the typical probability of misidentifying a QCD jet as a $\tau_h$ is at least an order of magnitude higher than that for a QCD jet to be misidentified as an electron or muon. As a result the QCD multijet background in the $h\tau_h$ channel is larger than in $\tau\tau \rightarrow \tau_1\tau_2\tau_3\tau_4$ channels. However, the multijet QCD contribution at high mass and $S_T$ is strongly reduced owing to its fast falling production cross section.

The overall strategy of the analysis is similar to other heavy neutrino and leptoquark searches. Upon selecting two high quality $\tau_h$ candidates and two additional jet candidates, the data distribution of $m(\tau_{h}, j, j, E_T^{\text{miss}})$ (in the heavy neutrino scenario) or $S_T$ (in the leptoquark scenario) is used to fit for a potential signal that would appear as an excess of events over the SM expectation in the high parts of the distributions. The selections defining the signal region (SR), described in Section 5, allow for a reduction of the background contribution in the high mass or $S_T$ part of the spectrum to a reasonable level. A main challenge of this analysis is to ensure high and well-understood signal selection and trigger efficiency with SM signatures containing real $\tau_h$ candidates. The strategy is described in Section 6 and relies on the selection of $Z \rightarrow \ell\ell + \text{jets}$ events. A number of additional background enriched control regions are described in Section 6. The control samples are defined to ensure a good understanding of the background contributions as well as to cross-check the accuracy of our efficiency measurements and assign appropriate systematic uncertainties (Section 7). The background contributions in the SR are derived from data wherever possible using samples enriched with background events. These control regions are used to measure the mass shapes, $S_T$ shapes, and selection efficiencies in order to extrapolate to the region where the signal is expected. In cases where the background contributions are small ($< 10\%$) or the above approach is not feasible, data-to-simulation scale factors, defined as a ratio between observed data events and expected simulated yields in background enhanced regions, are used to validate or correct the expected contributions obtained from the simulation samples.

## 2 CMS Detector

The central feature of the CMS apparatus is a superconducting solenoid of 6 m inner diameter, providing a field of 3.8 T. Within the field volume are the silicon pixel and strip tracker, the crystal electromagnetic calorimeter (ECAL), which includes a silicon sensor preshower detector in front of the ECAL endcaps, and the brass/scintillator hadron calorimeter (HCAL). Muons
are measured in gas-ionization detectors embedded in the steel return yoke. In addition to the barrel and endcap detectors, CMS has extensive forward calorimetry.

The inner tracker measures charged particles within $|\eta| < 2.5$ and provides an impact parameter resolution of $\sim 15 \mu m$ and a transverse momentum resolution of about 1.5% for 100 GeV particles. Collision events are selected by a first level trigger made of a system of fast electronics and a higher level trigger that consists of a farm of commercial CPUs running a version of the offline reconstruction optimized for fast processing. A more detailed description of the CMS detector can be found elsewhere [10].

### 3 Object Reconstruction and Identification

The jets are reconstructed using the Particle Flow (PF) algorithm [11]. In the PF approach, information from all subdetectors is combined to reconstruct and identify final-state particles (muons, electrons, photons, and charged and neutral hadrons) produced in the collision. The anti-$k_T$ clustering algorithm [12] with a distance parameter $R = 0.4$ is used for jet clustering. Jets are required to pass identification criteria designed to reject particles from pileup interactions, anomalous behavior from the calorimeters, and be fairly well separated from any identified leptons. For jets with $p_T > 30$ GeV and $|\eta| < 2.4$, the identification efficiency is $\approx 99\%$, while 90–95% of pileup jets are rejected [13]. The jet energy scale and resolution is calibrated through correction factors that depend on the $p_T$ and $\eta$ of the jet [14]. Jets originating from the hadronization of bottom quarks are identified using the loose working point of the combined secondary vertex (CSV) algorithm [15], which exploits observables related to the long lifetime of b hadrons. For b-quark jets with $p_T > 30$ GeV and $|\eta| < 2.4$, the identification efficiency is $\approx 85\%$ with a $\approx 10\%$ fake rate for light quark and gluon jets [16]. The b-quark jets are used to obtain $t\bar{t}$ enriched control samples used to estimate the background rate in the SR.

Although muons are not used to define the SR, they are utilized to obtain control samples for the background estimations. Muons are reconstructed using the tracker and muon chambers. Quality cuts based on the minimum number of hits in the silicon tracker, pixel detector and muon chambers are applied to suppress backgrounds from decays in flight and hadron shower remnants that reach the muon system [17]. Muon candidates are required to pass isolation requirements. Isolation is defined as the sum of the $p_T$ of the reconstructed PF charged and neutral particles, within an isolation cone of radius $\Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2} = 0.4$ centered around the muon track. In both cases the contribution from the muon candidate is removed from the sum and corrections are applied to remove the contribution from particles produced in pileup interactions. The muon identification efficiency for the quality requirements and kinematic range used in this analysis is $\approx 98\%$.

Hadronic decays of the $\tau$ lepton are reconstructed and identified using the hadrons plus strips (HPS) algorithm [18] designed to optimize the performance of $\tau$ reconstruction by considering specific $\tau_h$ decay modes. To suppress backgrounds from light-quark or gluon jets, a $\tau_h$ is required to be isolated from other energy in the event. The isolation variable is calculated using a cone of radius $\Delta R = 0.5$ in the vicinity of the identified $\tau_h$ and considering the energy deposits of particles not considered in the reconstruction of the $\tau_h$ decay mode. Additionally, $\tau_h$ candidates are required to be distinguishable from electrons and muons by using dedicated discriminators in the event. The algorithm to discriminate a $\tau_h$ from an electron utilizes observables that quantify the compactness and shape of energy deposits in the ECAL, to distinguish electromagnetic from hadronic showers, in combination with observables that are sensitive to the amount of bremsstrahlung emitted along the leading track and observables that are sensitive to the overall particle multiplicity. The discriminator against muons is based on the
presence of hits in the muon system associated with the track of the $\tau_h$ candidate. The “tight” isolation working point is used to define the SR, which results in a $\tau_h$ identification efficiency for the kinematic range used in this analysis of $\approx 55\%$.

The presence of neutrinos in the $\tau\tau$ decays must be inferred from the imbalance of total momentum in the detector. The magnitude of the negative vector sum of the transverse momentum of visible PF objects is known as missing transverse energy, denoted $E_T^{\text{miss}}$. Information from the forward calorimeter is included in the calculation of $E_T^{\text{miss}}$ and the jet corrections described above are propagated as corrections to $E_T^{\text{miss}}$. Missing transverse energy is one of the most important observables for discriminating the signal events from background events which do not contain neutrinos, such as QCD multijet events.

## 4 Signal and Background Samples

QCD multijet processes are the prevailing background in the SR. QCD multijet events are characterized by jets with a high-multiplicity of particles, which can be misidentified as $\tau_h$. Apart from QCD multijets, the other much smaller backgrounds are Drell-Yan (DY) processes giving rise to $\tau$ leptons and top pair production ($t\bar{t}$). The DY + jets events are characterized by two isolated $\tau$ leptons from the decay of a $Z$ boson and uncorrelated jets from initial state radiation. Background from $t\bar{t}$ events is accompanied by two b-quark jets, in addition to similar contributions from genuine isolated $\tau_h$ leptons and mis-identified $\tau_h$ candidates.

Collision data is compared to samples of Monte-Carlo (MC) simulated events and data-driven techniques are employed wherever needed. The MADGRAPH program (v5.1.5)\cite{19} is used for DY + jets, W + jets, and $t\bar{t}$ + jets production. The MADGRAPH generator is interfaced with PYTHIA 6.4.22 \cite{20} for parton shower and fragmentation. The PYTHIA generator is used to model the signal and QCD multijet processes. The heavy neutrino signal event samples are generated with $W_R$ masses ranging from 1 to 3 TeV. The $N_\tau$ mass varies between 0.1 and 0.8 times the $W_R$ mass. The leptoquark signal event samples are generated with masses ranging from 200 to 1000 GeV. The $\tau$-lepton decays have been performed with TAUOLA \cite{21}. MC generated events have been processed with a detailed simulation of the CMS apparatus using the GEANT4 package \cite{22}. The MC background and signal yields are normalized to integrated luminosity using next-to-next-to-leading order (NNLO) or next-to-leading order (NLO) cross-sections \cite{23}. The mean number of interactions in a single bunch crossing in the analysed dataset is 21. In MC events, multiple interactions are superimposed on the primary collision, and each MC event is reweighted such that the distribution of the number of true interactions matches that in data.

## 5 Event Selection

Candidate signal events were collected using a trigger requiring the presence of at least two $\tau_h$ trigger objects with $p_T(\tau_h) > 35$ GeV and $|\eta(\tau_h)| < 2.1$ \cite{24}. Additional kinematic requirements on $p_T$ and $\eta$ are imposed on the reconstructed $\tau_h$ candidates used in the SR to achieve a trigger efficiency greater than $\approx 90\%$ per $\tau_h$ candidate. Pre-selected events are required to have at least two $\tau_h$ candidates with $p_T$ greater than 70 GeV. The $\tau_h \tau_h$ pairs are required to be separated by $\Delta R > 0.4$. A $\tau_h$ candidate is required to have pseudorapidity $|\eta| < 2.1$ in order to ensure that it is reconstructed fully within the acceptance of the tracking system. Candidates are also required to satisfy the reconstruction and identification criteria described in section 3. Unlike other $\tau\tau$ analyses where an opposite-sign requirement can help discriminate against backgrounds from mis-identified $\tau_h$ candidates, our LRSM benchmark scenario can produce
both oppositely-charged and same-sign charged $\tau_h \tau_h$ candidates. Therefore, no charge requirement is imposed in this analysis.

In addition to the pre-selection described above, the final selection is defined by requiring at least two jets with $p_T$ greater than 50 GeV and pseudorapidity $|\eta| < 2.4$. Only jets separated from the leptons in the $\tau_h \tau_h$ pair by $\Delta R > 0.4$ are considered. Because there are neutrinos in the $\tau \tau$ system decay, we require $E_T^{miss} > 50$ GeV in order to control the level of QCD multijet background. Further, to reduce the contribution from $Z + jets$, the invariant mass of the $\tau_h \tau_h$ pair is required to be greater than 100 GeV.

The set of events satisfying the selections described above define the SR. According to simulation, the total background yield in the SR is $\approx 20$ events, with QCD multijet, $t\bar{t}$, and $Z + jets$ composing 76.6%, 12.6%, and 6.6% of the rate respectively. Finally, the $m(\tau_h \tau_h, j, j, E_T^{miss})$ and $S_T$ shapes normalized to the values obtained from the background estimation methods (section 6) are used to search for a broad enhancement above the SM background prediction.

The signal selection efficiency for $W_R \rightarrow \tau + N_t \rightarrow \tau + \tau q q'$ events depends on the $W_R$ and $N_t$ masses. The total signal acceptance, assuming the $N_t$ mass is half the $W_R$ mass, is $1.65%$ for $m(W_R) = 1.0$ TeV and $5.15%$ for $m(W_R) = 2.7$ TeV. The signal selection efficiency for $LQ \rightarrow \tau b$ events is $4.14%$ for $m(LQ) = 0.6$ TeV and $6.68%$ for $m(LQ) = 1.0$ TeV. These efficiencies include the $\approx 42%$ branching fraction of $\tau \tau$ to $\tau_h \tau_h$.

6 Background Estimation

As discussed above, $E_T^{miss}$ and $\tau_h$ isolation are the main discriminating variables against QCD multijet events. Thus, the QCD multijet background estimation methodology utilizes control samples obtained by inverting these requirements. In the remainder of this section, events obtained by inverting the isolation requirement on both $\tau_h$ candidates will be referred to as non-isolated $\tau_h \tau_h$ samples. The QCD multijet background is estimated using a completely data-driven approach which relies on the classic ABCD method. The regions $ABCD$ are defined as follows:

- $A$: fail the $E_T^{miss} > 50$ GeV cut; non-isolated $\tau_h \tau_h$
- $B$: fail the $E_T^{miss} > 50$ GeV cut; pass nominal isolation
- $C$: pass the $E_T^{miss} > 50$ GeV cut; non-isolated $\tau_h \tau_h$
- $D$: pass the $E_T^{miss} > 50$ GeV cut; pass nominal isolation (signal region)

Region $D$ is the nominal SR. The QCD component $N^i_{QCD}$ in regions $A, B, C$ is predicted by subtracting MC non-QCD backgrounds from data ($N^i_{QCD} = N^i_{Data} - N^i_{\neq QCD}$). The signal contamination in control regions $A, B, C$ is negligible. The contribution of QCD events in the SR ($N^D_{QCD}$) is estimated using the predicted rate of QCD events in region $C (N^C_{QCD})$, weighted by a scale factor used to extrapolate from the non-isolated to the isolated $\tau_h$ region. The extrapolation factor is obtained by dividing the expected number of QCD events in region $B (N^B_{QCD})$ by the expected number of QCD events in region $A (N^A_{QCD})$. The shapes for the variables of interest, $m(\tau_h \tau_h, j, j, E_T^{miss})$ and $S_T$, are obtained from region $C$.

Closure and validation tests for the background estimation method outlined above are performed with data. Two aspects are simultaneously tested: (1) closure on the normalization $N^D_{QCD} = N^C_{QCD} \cdot (N^B_{QCD} / N^A_{QCD})$; (2) correct determination of the $m(\tau_h \tau_h, j, j, E_T^{miss})$ and $S_T$ shapes. The first set of closure tests in data are performed using the same method and event selection criteria described above for the different regions, except with an inverted jet multi-
plicity requirement, $N_j < 2$, in order to provide an exclusive set of samples, $A'B'C'D'$. The purity of QCD events in these control samples ranges approximately from 96% to 99%. Agreement is observed between the observed number of QCD events in region $D'$, 123, and the predicted yield using the estimation method, $N_{QCD}^{B'} = N_{QCD}^C \cdot (N_{QCD}^{B'}/N_{QCD}^A) = 122.2 \pm 10.3$. Figure 1(top) shows the $m(\tau_h, \tau_h, j, E_T^{miss})$ and $S_T$ distributions in region $D'$, where the shapes of QCD events were obtained from region $C'$ and normalized to the expected yield of QCD events in the region $D'$. Note there is good agreement across the $m(\tau_h, \tau_h, j, E_T^{miss})$ and $S_T$ spectrum, showing that $\tau_h$ isolation does not bias the shape. An additional test on the extraction of the shape from the non-isolated $\tau_h$ regions, with $N_j \geq 2$, is performed using the shape from QCD events falling in region $A$, to estimate the shape of QCD events in region $B$. Figure 1 (bottom) shows the $m(\tau_h, \tau_h, j, E_T^{miss})$ and $S_T$ distributions in region $B$, using the shape for QCD events from region $A$, which provides further confidence on the method. The procedure outlined in this section yields a QCD estimate of $N_{QCD}^{Signal} = 15.1 \pm 4.1$. The systematic uncertainty is based on the statistics of the control samples.

The measurement of the $Z \rightarrow \tau\tau +$ jets contribution to the SR is based on both simulation and data. The efficiency for the trigger and requirement of at least two high quality $\tau_h$ leptons is expected to be well modeled by simulation. Mismodeling of the $Z \rightarrow \tau\tau +$ jets background rate and shapes in the SR can come from the requirement of two additional jets. Therefore, its contribution is determined from two control samples. The first control sample is used to validate the correct modeling of the requirement of at least two high quality $\tau_h$ leptons. The second control sample is used to measure a correction factor for the correct modeling of two additional jets.

The first control sample used to validate the correct modeling of the trigger and requirement of at least two high quality $\tau_h$ leptons is obtained by using the pre-selection cuts defined previously, but requiring $\tau_h \tau_h$ pairs to have an invariant mass less than 100 GeV, resulting in a sample composed of $\approx 90\%$ of $Z \rightarrow \tau\tau$, according to simulation. The rates and shapes in data and MC are consistent, with a measured data-to-MC scale factor of 0.97 ± 0.19.

The second control sample, used to measure a correction factor for the efficiency of the dijet selections, is obtained by applying selection criteria similar to those used in the final analysis to obtain a sample of $Z \rightarrow \mu\mu +$ jets events with two muons (instead of $\tau_h \tau_h$) having invariant mass compatible with the $Z$-mass hypothesis ($60 < m_{\mu\mu} < 120$ GeV). Candidate events for the $Z \rightarrow \mu\mu +$ jets control sample were collected using a trigger that requires the presence of at least one muon trigger object with $p_T(\mu) > 18$ GeV. Due to lepton universality, muons are produced in $Z$-decays as often as $\tau$ leptons. This is used to properly model the extra hadronic activity in $Z \rightarrow \tau\tau +$ jets. Therefore, a scale factor for $Z \rightarrow \tau\tau +$ jets can be measured from this resultant dimuon sample. The predicted rate for $Z \rightarrow \tau\tau$ in simulation, after final selections, can be corrected for the mis-modeled dijet selection efficiency, to determine the expected contribution of $Z \rightarrow \tau\tau +$ jets in the SR. The measured correction factor is 1.20 ± 0.01.

Similar to the estimation of $Z \rightarrow \tau\tau +$ jets, the measurement of the $t\bar{t}$ contribution to the SR is based on both data and simulation. A $t\bar{t}$ enriched control sample is obtained by applying all the signal selection criteria with at least one $b$-tagged jet and two isolated muons, as opposed to $\tau_h \tau_h$, and additionally requiring a $Z$-mass veto cut ($m_{\mu\mu}$ outside the region between 80 and 110 GeV) to suppress $Z \rightarrow \mu\mu +$ jets. The $t\bar{t}$ prediction from simulation agrees with the observed yield and shape in the control sample, and thus the $t\bar{t}$ prediction in the SR is based on simulation without further corrections. The $t\bar{t}$ estimate in the SR is 2.5 ± 0.9.
7 Systematic Uncertainties

Various imperfectly known or simulated effects can alter the shape and normalization of the $m(\tau_h,\tau_h,j,j,E_T^{\text{miss}})$ and $S_T$ spectrum. Since the estimation of the background contributions in the SR is partly based on simulation, the signal and certain backgrounds are affected by similar sources of systematic uncertainties. For example, the uncertainty in the luminosity measurement is 2.7% [25] and affects the signal, DY + jets, and $t\bar{t}$ background. The dominant source of systematic uncertainties on the signal, DY + jets, and $t\bar{t}$ predictions are due to uncertainties in the $\tau_h$ identification and trigger efficiency. The $\tau_h$ trigger efficiency per object are measured from $Z \rightarrow \tau\tau \rightarrow \mu \tau_h$ events, selected by single muon triggers and which satisfy the same $\tau_h$ identification criteria used in the SR, by determining the fraction of $\tau_h$ candidates which additionally pass the $\tau_h$ trigger requirements. This leads to a relative uncertainty of 5.0% per $\tau_h$ candidate. Systematic effects associated with $\tau_h$ identification are extracted from a fit to the $Z \rightarrow \tau\tau$ visible mass distribution, $m(\tau_1,\tau_2)$. In order to extract the uncertainty on the $\tau_h$ identi-
8 Results

Figure 2 shows the background predictions as well as the observed \(m(\tau_h, \tau_h, j, j, E_T^{\text{miss}})\) and \(S_T\) spectrum. The last bin in the mass plot represents the yield for \(m(\tau_h, \tau_h, j, j, E_T^{\text{miss}}) > 2.25\) TeV, while the last bin in the \(S_T\) plot represents the yield for \(S_T > 1\) TeV (i.e. includes the overflow). The observed yield is 14 events, while the total predicted background yield is 19.8 ± 4.2 events, with QCD multijet, \(t\bar{t}\), and \(Z \rightarrow \tau\tau\) composing 76.3%, 12.6%, and 6.6% of the rate respectively (see Table 1). The distributions from the \(m(W_R) = 2700\) GeV and \(m(LQ) = 1\) TeV signal hypotheses are plotted with the background prediction in Figure 2 to illustrate how a hypothetical signal would appear above the SM background prediction. Since the observed \(m(\tau_h, \tau_h, j, j, E_T^{\text{miss}})\) and \(S_T\) distributions do not reveal evidence for \(W_R \rightarrow \tau N_f \rightarrow \tau\tau jj\) and \(LQ \rightarrow \tau b\) production, an upper bound at 95% confidence level (CL) is set on \(\sigma\), where \(\sigma\) is the signal production cross-section.

The calculation of the exclusion limit is obtained by using each bin of the \(m(\tau_h, \tau_h, j, j, E_T^{\text{miss}})\) distribution (or \(S_T\) for the LQ interpretation) to construct one bin of the poisson likelihood and computing the 95% CL upper limit on the signal cross-section using the modified frequentist construction CL\textsubscript{s} method. Systematic uncertainties are represented by nuisance parameters, assuming a gamma or log normal prior for normalization parameters, and Gaussian priors for shape uncertainties.

Figure 3 shows the expected and observed limits as well as the theoretical cross-section as functions of \(m(W_R)\) and \(m(LQ)\). For heavy neutrino models with strict left-right symmetry, and assuming only \(N_f\) flavor contributes significantly to the \(W_R\) decay width, \(W_R\) masses below 2.31 TeV are excluded at a 95% CL, assuming the \(N_f\) mass is 0.5 · \(m(W_R)\). The heavy neutrino limits depend on the \(N_f\) mass. For example, the \(x = m(N_f) / m(W_R) = 0.1 (0.25)\) scenario yields significantly lower average jet and sub-leading \(\tau_h p_T\) than the \(x = 0.5\) mass assumption, and the acceptance is lower by a factor of \(\approx 16\) (3) for \(m(W_R) = 1\) TeV and \(\approx 6\) (1.9) for \(m(W_R) = 2.7\) TeV. On the other hand, the \(x = 0.75\) scenario produces similar or larger average jet and \(\tau_h p_T\) than the \(x = 0.5\) mass assumption, yielding an event acceptance that is up to 10% larger. Figure 4 shows the 95% CL upper limits on the production cross section, as a function of \(m(W_R)\) and
\[ x = m(N_\tau)/m(W_R). \]
The signal acceptance and mass shape is evaluated for each \( \{m(W_R), x\} \) combination in Figure 4 and used in the limit calculation procedure described above. Masses below \( m(W_R) = 2.35 \) (1.63) TeV are excluded at a 95% confidence level, assuming the \( N_\tau \) mass is 0.8 (0.2) times the mass of \( W_R \) boson.

This is the first LHC result for \( N_\tau \) searches with \( \tau \) leptons. Other searches for heavy neutrinos have been performed in the \( \mu\mu jj \) and \( ee jj \) channels assuming \( N_\tau \) is too heavy to play a role in the decay of \( W_R \) (and thus free of \( \tau \) leptons). In those searches, an excess of 2.4-2.8\( \sigma \) has been observed at \( \approx 2.2 \) TeV in the \( eejj \) channel, while the \( \mu\mu jj \) channel excludes \( m(W_R) \) less than \( \approx 3 \) TeV under the assumption of very heavy \( N_\tau \) [4]. For the leptoquark interpretation using \( S_T \) as the final fit variable, the expected 95% CL exclusion is LQ masses below 790 GeV, while the observed exclusion is approximately 740 GeV, resulting in the most stringent limit to date.

Figure 2: Left: \( m(\tau_h, \tau_h, j, j, E_{T}^{\text{miss}}) \) distribution in the SR. Right: \( S_T \) distribution in the SR.

Figure 3: Left: Expected and observed limits, at 95% confidence level, as functions of \( m(W_R) \) mass. Right: Expected and observed limits, at 95% confidence level, as functions of LQ mass. The bands on the expected limits represent the one and two standard deviations obtained using a large sample of pseudo-experiments based on the background-only hypothesis for each bin of the mass and \( S_T \) distributions.
Table 1: Number of observed events in data and estimated background and signal rates in the SR.

<table>
<thead>
<tr>
<th>Process</th>
<th>Prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>DY + jets</td>
<td>1.3 ± 0.5</td>
</tr>
<tr>
<td>W + jets</td>
<td>0.9 ± 0.4</td>
</tr>
<tr>
<td>t\bar{t}</td>
<td>2.5 ± 0.9</td>
</tr>
<tr>
<td>QCD</td>
<td>15.1 ± 4.1</td>
</tr>
<tr>
<td>Total</td>
<td>19.8 ± 4.2</td>
</tr>
<tr>
<td>Observed</td>
<td>14</td>
</tr>
<tr>
<td>m(W_R) = 1.0 TeV</td>
<td>61.1 ± 1.5</td>
</tr>
<tr>
<td>m(W_R) = 2.7 TeV</td>
<td>1.6 ± 0.02</td>
</tr>
<tr>
<td>m(LQ) = 0.6 TeV</td>
<td>14.7 ± 0.3</td>
</tr>
<tr>
<td>m(LQ) = 1.0 TeV</td>
<td>0.8 ± 0.01</td>
</tr>
</tbody>
</table>

Figure 4: Observed and expected 95% CL upper limits on the production cross section for the ratio between m(N_t) and m(W_R) as a function of m(W_R).

9 Summary

A search is performed for physics beyond the SM in events with two energetic \( \tau \) leptons, two energetic jets, and large momentum imbalance, using data corresponding to an integrated luminosity of 2.1 fb\(^{-1}\) collected by the CMS detector in proton-proton collisions at \( \sqrt{s} = 13 \) TeV. The search focuses on two benchmark new physics scenarios: (1) production of heavy right-handed neutrinos, \( N_l \), and right-handed \( W_R \) bosons which arise in the left-right symmetric extensions of the SM and where the \( W_R \) and \( N_l \) decay chains result in a pair of high-\( p_T \) \( \tau \) leptons; (2) pair production of third-generation scalar leptoquarks which decay to \( \tau \bar{\tau} b\bar{b} \). The observed \( m(\tau_b, \tau_b, j, j, E_T^{\text{miss}}) \) and \( S_T \) distributions do not reveal any evidence for new physics. Assuming only \( N_t \) flavor contributes significantly to the \( W_R \) decay width, \( W_R \) masses below 2.35 (1.63) TeV are excluded at a 95% confidence level, assuming the \( N_t \) mass is 0.8 (0.2) times the mass of \( W_R \) boson. This is the first LHC result for \( N_l \) searches with \( \tau \) leptons. This analysis is the first to focus on searches for pair production of third-generation scalar leptoquarks using the \( \tau_b \tau_b jj \).
final state, resulting in an expected 95% CL exclusion of $m(LQ) < 790$ GeV and an observed exclusion of approximately 740 GeV.

References


