CMS Physics Analysis Summary

Contact: cms-pag-conveners-b2g@cern.ch 2016/08/08

Search for new resonances decaying to WW/WZ $\to \ell\nu$qq

The CMS Collaboration

Abstract

A search for massive narrow resonances decaying to pairs of W and Z bosons in the $\ell\nu$qq final state is presented, based on 12.9 fb^{-1} of pp collision data collected in 2016 by the CMS experiment at the CERN LHC with a center-of-mass energy of 13 TeV. Spin-1 and spin-2 resonances corresponding to masses in the range $600 - 4500$ GeV and decaying to WW/WZ are probed using the $\ell\nu$qq final state. Cross section and resonance mass exclusion limits are set for models that predict gravitons and heavy spin-1 bosons.

1 Introduction

There are several theory models that motivate the existence of heavy particles that decay to pairs of bosons. These models usually aim to answer open questions of the Standard Model (SM) such as the integration of gravity into the SM using extra dimensions. Popular examples of such models include the bulk scenario [1–3] of the Randall-Sundrum Warped Extra Dimensions model [4, 5], and the composite heavy vector triplet (HVT) model [6]. The composite HVT generalises a large number of explicit models predicting spin-1 resonances, which can be described by a rather small set of parameters.

Extensive studies have been performed over recent years from the ATLAS and CMS experiments on heavy resonance searches decaying into a pair of vector bosons [7–9], covering a mass range of 600-4000 GeV.

In this physics analysis summary we perform the study with the first 2016 proton-proton collision data at $\sqrt{s} = 13$ TeV collected by the CMS experiment. We search for the production of a narrow-width resonance within the mass range 600-4500 GeV, decaying into WV (V denotes hadronically decaying W or Z boson). Narrow refers to the assumption that the natural width of the resonance is much smaller than the experimental resolution. This analysis looks exclusively at semileptonic final states from a W boson that decays to a lepton (μ or e) and a neutrino while the other W or Z boson decays hadronically, leading to a merged jet. Final states with $W \to \tau\nu$ candidates, where the τ decays to one electron or muon and two neutrinos, are also considered.

In this physics analysis summary, Section 2 briefly describes the CMS detector. Section 3 gives an overview of the simulations used in this analysis. Section 4 provides a detailed description of the reconstruction and event selection. In Section 5 we study the performance of the so-called "V tagging", i.e. the reconstruction of the hadronically decaying vector boson using information from jet substructures. This is done by studying a sample of events enriched in top quarks. Section 6 describes the background estimation and the signal modeling. Systematic uncertainties are discussed in Sec. 7. The results of the search for a bulk graviton are presented in Sec. 8.

2 The CMS Detector

The central feature of the CMS apparatus is a superconducting solenoid of 6 m internal diameter, providing a magnetic field of 3.8 T. Contained within the superconducting solenoid volume are a silicon pixel and 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. Muons are measured in gas-ionization detectors embedded in the steel flux-return yoke outside the solenoid. Extensive forward calorimetry complements the coverage provided by the barrel and endcap detectors.

The particle-flow event algorithm [10, 11] reconstructs and identifies each individual particle with an optimized combination of information from the various elements of the CMS detector. The energy of photons is directly obtained from the ECAL measurement, corrected for zero-suppression effects. The energy of electrons is determined from a combination of the electron momentum at the primary interaction vertex as determined by the tracker, the energy of the corresponding ECAL cluster, and the energy sum of all bremsstrahlung photons spatially compatible with originating from the electron track. The energy of muons is obtained from the curvature of the corresponding track. The energy of charged hadrons is determined from a

combination of their momentum measured in the tracker and the matching ECAL and HCAL energy deposits, corrected for zero-suppression effects and for the response function of the calorimeters to hadronic showers. Finally, the energy of neutral hadrons is obtained from the corresponding corrected ECAL and HCAL energy.

Jet energy-momentum is determined as the vectorial sum of all particle four-momenta in the jet. An offset correction is applied to jet energy-momenta to take into account the contribution from additional proton-proton interactions within the same bunch crossing. These corrections are derived from simulation, and are confirmed with in situ measurements of the energy-momentum balance in dijet and photon+jet events. Additional selection criteria are applied to each event to remove spurious jet-like features originating from isolated noise patterns in certain HCAL regions.

Muons are measured in the pseudorapidity range $|\eta| < 2.4$, with detection planes made using three technologies: drift tubes, cathode strip chambers, and resistive plate chambers. Matching muons to tracks measured in the silicon tracker results in a relative transverse momentum resolution for muons with $20 < p_T < 100\,\text{GeV}$ of 1.3–2.0% in the barrel and better than 6% in the endcaps, the p_T resolution in the barrel is better than 10% for muons with p_T up to 1 TeV.

Electrons are measured in the electromagnetic calorimeter, in the pseudorapidity range $|\eta| < 2.5$. The electron momentum is estimated by combining the energy measurement in the ECAL with the momentum measurement in the tracker. The momentum resolution for electrons with $p_T \approx 45$ GeV from $Z \to ee$ decays ranges from 1.7% for nonshowering electrons in the barrel region to 4.5% for showering electrons in the endcaps.

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. [12].

3 Simulated samples

The bulk graviton model and the HVT model (W') are used as benchmark signal processes. In these models, the vector gauge bosons are produced with a longitudinal polarisation (V_L) in more than 99% of events. For the HVT model, a scenario (model A) is chosen where the W' predominantly couples to bosons; that scenario is the most representative of a composite Higgs model. Different mass hypotheses are considered for the posited resonance, in the range 600 to 4500 GeV. Simulated signal samples are generated with the leading-order (LO) mode of MADGRAPH5_AMC@NLO v5.2.2.2 [13] with a relative resonance width of 0.1%. We restrict the analysis to scenarios where the natural width of the resonance is sufficiently small to be neglected when compared to the detector resolution. This makes our modeling of the detector effects on the signal shape independent of the actual model used for generating the events.

Monte Carlo (MC) simulated samples of the SM background processes are used to optimise the analyses. The W+jets and Z+jets SM processes are simulated with MADGRAPH5_AMC@NLO, $t\bar{t}$ and single top quark events are generated with POWHEG v2 [14–19], while diboson (WW, WZ, and ZZ) processes are produced with PYTHIA v8.205 [20, 21]. The parton showering and hadronisation are performed with PYTHIA using the CUETP8M1 tune [22, 23]. The NNPDF 3.0 [24] parton distribution functions (PDF) are used in all generated samples. All generated samples are processed through a GEANT4-based [25] simulation of the CMS detector. The simulated background samples are normalized using inclusive cross sections calculated at next-to-leading order (NLO), or next-to-next-to-leading order (NNLO) where available, calculated with MCFM v6.6 [26–29], FEWZ v3.1 [30] and TOP++ v2.0 [31].

Minimum bias interactions are added to the generated events in order to match the additional particle production observed in data from the large number of proton-proton interactions occurring per LHC bunch crossing (pileup). The simulated samples are corrected for observed differences between data and simulation in the efficiencies of lepton trigger selection, lepton identification/isolation, and selection criteria identifying jets originating from hadronisation of b quarks (b jets).

4 Reconstruction and selection of events

4.1 Trigger and preliminary offline selection

Events are selected online with a trigger requiring either one muon or one electron. The muon selected online must have p_T higher than 45 GeV, $|\eta| < 2.1$, and satisfy isolation requirements. The electron selected online must have E_T higher than 45 GeV, and satisfy identification criteria.

Offline, all events are required to have at least one primary vertex reconstructed within a 24 cm window along the beam axis, with a transverse distance from the nominal pp interaction region of less than 2 cm [32]. In the presence of more than one vertex passing these requirements, the primary-event vertex is chosen to be the one with the highest total p_T^2, summed over all the associated tracks.

4.2 Jet reconstruction

Hadronic jets are clustered from the four-momenta of the particles reconstructed by the CMS particle-flow (PF) algorithm [10, 11], using the FASTJET software package [33]. The reconstructed PF constituents are assigned to one of the five candidate categories (electrons, muons, photons, charged hadrons, and neutral hadrons). Jets used for identifying the hadronically decaying W bosons are clustered using the anti-k_T algorithm [34] with a distance parameter $R = 0.8$ ("AK8 jets"). In order to identify b jets, the anti-k_T jet clustering algorithm is used [35] with a distance parameter $R = 0.4$ ("AK4 jets") and the inclusive combined secondary vertex b tagging algorithm [36] is applied to the reconstructed AK4 jets. The ratio of the b-tagging efficiency between data and simulation is used as a scale factor to correct the simulated events. A correction based on the projected area of the jet on the front face of the calorimeter is used to take into account the extra energy clustered in jets due to neutral particles coming from pileup. Jet energy corrections are derived from simulation and from dijet and photon+jet events in data as in Ref. [37]. Additional quality criteria are applied to the jets in order to remove spurious jet-like features originating from isolated noise patterns in the calorimeters or the tracker. The efficiency of these jet quality requirements for signal events is above 99%. The AK8 (AK4) jets are required to be separated from any well-identified electron or muon by $\Delta R > 0.8$ (0.3). All jets must have $p_T > 30$ GeV and $|\eta| < 2.4$ in order to be considered in the subsequent steps of the analysis.

4.3 $W \to q\bar{q}'$ and $Z \to q\bar{q}$ identification using jet substructure

AK8 jets are used to reconstruct the W-jet and Z-jet candidates from hadronic decays of boosted W and Z bosons, respectively. In order to discriminate against multijet backgrounds we exploit both the reconstructed jet mass, which is required to be close to the W- or Z-boson mass, and the two-prong jet substructure produced by the particle cascades of two high-p_T quarks merging into one jet. Jets that are identified as coming from the merged decay products of a single V boson are hereafter referred to as "V jets".

4.3.1 Jet mass reconstruction

As the first step in exploring potential substructure, the jet constituents are subjected to a jet grooming algorithm that improves the resolution on the jet mass and reduces the effect of pileup [38]. The goal of jet grooming is to re-cluster the jet constituents while applying additional requirements that eliminate soft, large-angle QCD radiation that increases the V-jet mass compared to the initial V boson mass. Different jet grooming algorithms have been explored at CMS and their performance on jets in multijet processes has been studied in detail [38]. In this analysis, we use the *jet pruning* [39, 40] algorithm. The jet pruned mass, m_{jet}, is computed from the sum of the four-momenta of the constituents surviving the pruning and it is then corrected similarly to what is done on the jet p_{T}. The jet is considered as a W(Z) -jet candidate if m_{jet} falls in the range $65(75) < m_{\text{jet}} < 95(105)$ GeV, which we define as the signal jet mass window.

4.3.2 N-subjettiness reconstruction

Further discrimination against jets from gluon and single-quark hadronization is obtained from the quantity called *N-subjettiness* [41]. The constituents of the jet before the pruning procedure are re-clustered with the k_{T} algorithm [42, 43], until N joint objects (*subjets*) remain in the iterative combination procedure of the k_{T} algorithm. The N-subjettiness, τ_N, is then defined as

$$\tau_N = \frac{1}{d_0} \sum_k p_{\text{T},k} \min(\Delta R_{1,k}, \Delta R_{2,k}, \ldots, \Delta R_{N,k}), \tag{1}$$

here the index k runs over the PF constituents of the jet and the distances $\Delta R_{n,k}$ are calculated with respect to the axis of the nth subjet. We use a one step optimization of the exclusive k_{T} axes as a definition for the subjet axes. The normalization factor d_0 is calculated as $d_0 = \sum_k p_{\text{T},k} R_0$, setting R_0 to the jet radius of the original jet. The variable τ_N quantifies the capability of clustering the jet constituents in exactly N subjets, with small values representing configurations more compatible with the N-subjets hypothesis. The ratio between 2-subjettiness and 1-subjettiness, $\tau_{21} = \tau_2/\tau_1$, is found to be a powerful discriminant between jets originating from hadronic V-boson decays and from gluon and single-quark hadronization. We reject V-jet candidates with $\tau_{21} > 0.45$ (0.6), for low (high) mass searches. The remaining events are further categorised according to their value of τ_{21} in order to enhance the sensitivity of the analysis. Jets coming from hadronic V boson decays in signal events are characterised by lower values of τ_{21} compared to the SM backgrounds.

4.4 Muon reconstruction and selection

Muons are reconstructed with a fit using both the inner tracking system and the muon spectrometer [44]. Muons used in the semileptonic analysis must satisfy identification requirements on the impact parameters of the track, the number of hits reconstructed on both the silicon tracker and the muon detectors, and the relative error on the p_{T} measurement. These quality selections ensure a precise measurement of the four-momentum and reject misreconstructed muons.

An isolation requirement is applied to suppress the background from multijet events where jet constituents are identified as muons. A cone of radius $\Delta R = 0.3$ is constructed around the muon direction, and the isolation parameter is defined as the scalar sum of the transverse momenta of all the additional reconstructed tracks within the cone, divided by the muon p_{T}. We then require that the value of this isolation parameter is smaller than 0.1. The efficiency of this muon selection has been measured with a tag-and-probe method using Z bosons [45], and it has a negligible dependence on the number of reconstructed primary vertices in the event.

In the muon channel analysis, events must have exactly one isolated muon with $p_T > 50\,\text{GeV}$ and $|\eta| < 2.1$.

4.5 Electron reconstruction and selection

Electron candidates are reconstructed by matching energy deposits in the ECAL with reconstructed tracks [46]. In order to suppress multijet QCD background, electron candidates must pass stringent identification and isolation criteria optimised for high-p_T electrons [47]. Those criteria include requirements on the geometrical matching between the ECAL deposit and the reconstructed track, the ratio of the energies deposited in hadronic and electromagnetic calorimeters, the shape of the ECAL deposit, the impact parameters of the track, and the number of reconstructed hits on the silicon tracker.

In the electron channel analysis, we require then exactly one electron with $p_T > 55\,\text{GeV}$ and $|\eta| < 2.5$. Reconstructed electrons must be outside of the overlap region between the ECAL barrel and endcaps ($1.44 < |\eta| < 1.57$).

4.6 Missing transverse energy reconstruction and selection

The missing transverse energy E_T^{miss} is defined as the magnitude of the vector sum of the transverse momenta of the reconstructed PF objects. Corrections to the raw E_T^{miss} value to account for the jet energy scale of AK4 jets are applied. More details on the E_T^{miss} performance in CMS can be found in Refs. [48, 49].

A requirement $E_T^{miss} > 40\ (80)\,\text{GeV}$ is applied for the muon (electron) channel. The threshold is higher in the electron channel to further suppress the larger background from multijet processes.

4.7 $W \to \ell\nu$ reconstruction and identification

Identified electrons or muons are associated with the $W \to \ell\nu$ candidate. The transverse momentum of the undetected neutrino is assumed to be equal to the E_T^{miss}. The longitudinal momentum of the neutrino is obtained by solving a second-order equation that sets the $\ell\nu$ invariant mass to be equal to the known W boson mass [50]. In the case of two real solutions, the smaller one is chosen; in the case of two complex solutions, their real part is used. The four-momentum of the neutrino is used to build the four-momentum of the $W \to \ell\nu$ candidate. The same procedure is applied for $W \to \tau\nu$ candidates, where the τ decays to one electron or muon and two neutrinos. In this case, the E_T^{miss} represents the transverse momentum of the three-neutrino system. The contribution from $W \to \tau\nu$ events to the analysis is around 9% of the total number of signal events, measured on the two extreme mass points (600 and 1000 GeV). Finally, the invariant mass of the WV system (M_{WV}) is computed using the AK8 jet identified as the hadronic V boson, the selected lepton and the reconstructed neutrino.

4.8 Final event selection

After reconstructing the two vector bosons, we apply the final selections used for the search. For all channels, any V-boson candidate must have a p_T greater than 200 GeV. In addition, there are specific topological selections. We require that the two V bosons from the decay of a massive resonance are approximately back-to-back: the ΔR distance between the lepton and the V-jet is greater than $\pi/2$; the azimuthal angular separation between the missing transverse energy vector and the V-jet is greater than 2.0 radians; and the azimuthal angular separation between the $W \to \ell\nu$ and V-jet candidates is greater than 2.0 radians. To further reduce the level of the $t\bar{t}$ background, events are rejected if there are one or more b tagged AK4 jets in the

event, using a working point of the b-tagging algorithm tuned to provide a misidentification rate of ∼1% and efficiency of ∼70%. This veto preserves about 90% of signal events, while rejecting ∼55% of $t\bar{t}$ events.

Events having m_{jet} in the range 105–135 GeV after the final selection are not considered in this analysis, in order to avoid bias in future searches for WH, ZH and HH resonances with the same dataset.

The pruned mass and τ_{21} distributions for the different signal samples and background events are taken from simulation, and shown along with data in Fig. 1 for both the muon and electron channels. The τ_{21} distribution shows some disagreement between data and simulation. Previous studies suggest that part of this discrepancy can be attributed to a mismodeling of the parton showering in simulation [51]. The analysis is designed to be robust against differences between data and simulation independent of their specific sources, as described in the next sections.

In this iteration of the heavy resonance search we optimize separately the selection for the low mass region (600-1000 GeV) and high mass region (1000-4500 GeV). We briefly summarize here the main differences from previous analyses [8, 52]:

- we used triggers with lower p_T tresholds for both low and high mass searches;
- the offline p_T cut is 50 (55) GeV for muon (electron) categories;
- the ID selections for the selected lepton, which have been loosened from [52];
- to suppress possible QCD contaminations, W transverse mass in muon channel is required to be larger than 40 GeV.
- the jet pruned mass window for the signal region, changed to 65-95 GeV for W-jet and 75-105 GeV for Z-jet.
- the N-subjettiness cut: we used the 0.45 working point for low mass searches, and 0.60 for high mass cases.

5 W tagging in a top-quark enriched control sample

The data/simulation discrepancy observed in the key variable τ_{21} (Fig. 1) could bias the signal efficiency estimated from the simulated samples. It is therefore important to study the mismodeling in a signal-free sample with jets having characteristics that are similar to those expected for a genuine signal. In this way one can extract correction factors to apply to the signal efficiency suggested by the simulation and obtain a small systematic uncertainty related to this effect. A sample of high-p_T W bosons, which decay hadronically and are reconstructed as a single AK8 jet, can be cleanly studied in $t\bar{t}$ and single top-quark events. The control sample is selected by applying all analysis requirements but inverting the b-jet veto. The signal contamination (from a Bulk Graviton of 750 GeV) is less than 0.1%. The data are compared with the predictions from simulation. Discrepancies between data and simulation are corrected in the analysis using the scale factors for top-quark background normalization, W tagging efficiency, as well as the peak and resolution of the V-jet mass distribution derived in this section. Since the jet substructure produced in simulation depends on the modeling of the parton shower, PYTHIA v8.2 is used for this part of the event simulation. In this way the results of this study can be consistently applied to the simulated signal samples, that are also generated with the same parton showering. The distributions of the τ_{21} and the pruned jet mass in the top-quark enriched control sample are shown in [52].

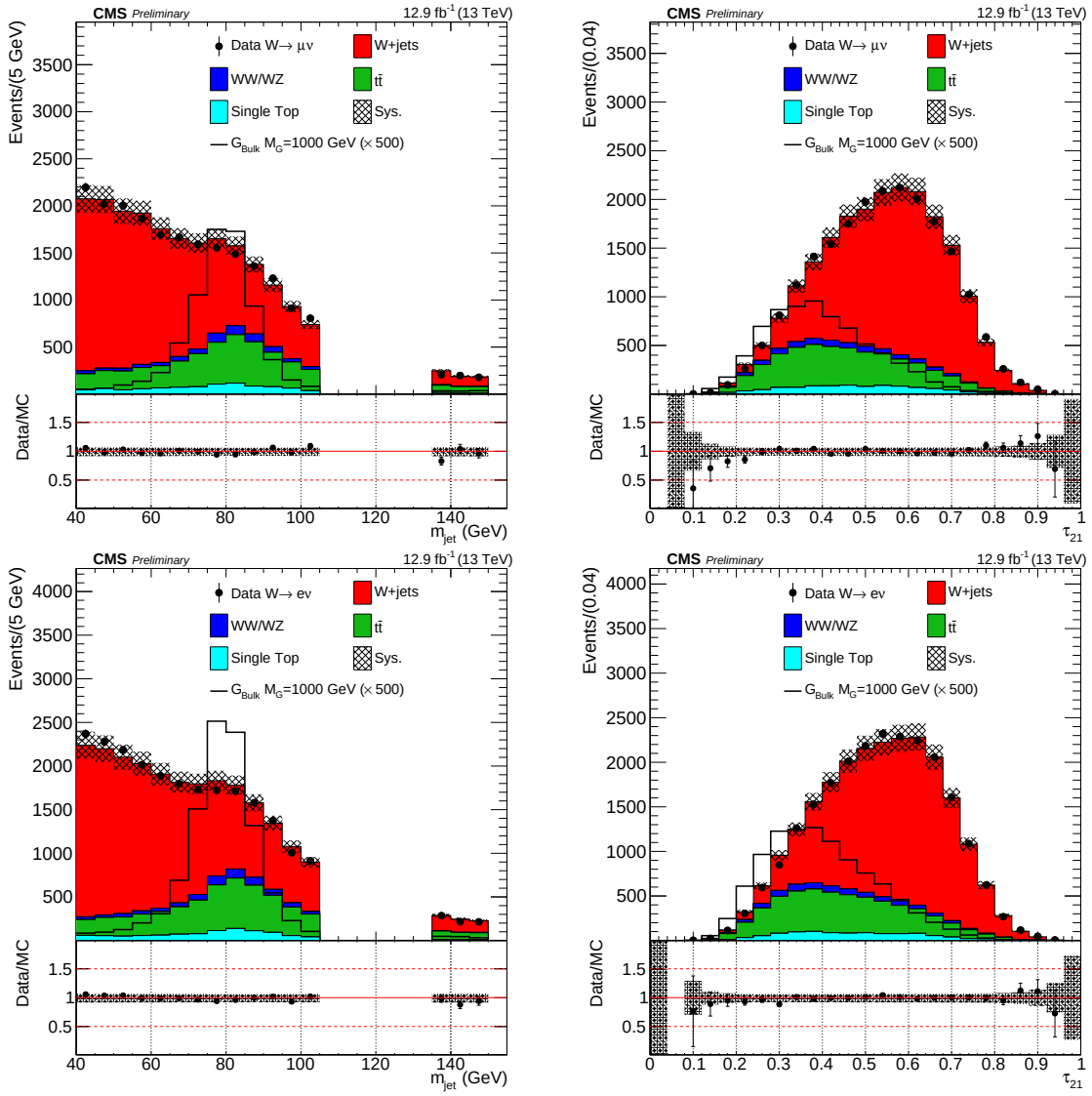

Figure 1: Pruned jet mass and N-subjettiness ratio τ_{21} distributions in the m_{jet} signal region and sideband. The signal distributions are scaled to an arbitrary cross section for better visualisation. The W+jets background from simulation is rescaled such that the total number of background events matches the number of events in data. The grey band represents the statistical uncertainty. Top row: muon channel, bottom row: electron channel.

For the $t\bar{t}$ and single top contributions, we compute a scale factor based on the difference between data and simulation expectation in the signal region of the control sample. We apply all the cuts used to define the $t\bar{t}$ control region; then, looking only inside the jet mass signal region and applying the N-subjettiness τ_{21} selection, we account for the difference after the τ_{21} cut, to get the data/simulation scale factor. The final computed scale factor is 0.81 ± 0.02 and 0.83 ± 0.01 for $\tau_{21} < 0.45$ and 0.6, respectively, combining muon and electron channels.

For the WW/WZ and signal contributions, we are only concerned with the efficiency for pure W-jet signal, so it is necessary to subtract background contributions to get the correct signal efficiency scale factor. A simultaneous fit to the jet mass distributions is performed in the top-quark enriched sample, in both data and simulation, to extract the efficiency of the τ_{21} cut. Differences in the resulting W-tagging efficiencies will be driven by the discrepancy between data and simulation in the τ_{21} distribution. The ratio of the efficiency in data and simulation yields W-tagging scale factors. These are used to correct the total signal efficiency predicted by the simulation. The scale factor for W tagging is 0.976 ± 0.048 and 1.002 ± 0.018 for $\tau_{21} < 0.45$ and 0.6, respectively, combining the muon and electron channels. A more detailed explanation of the procedure for the extraction of the W-tagging scale factor is given in [52].

In addition, the W-jet mass peak position and resolution are extracted from the same fit. The extracted jet mass peak positions are 84.9 ± 0.2 and 83.8 ± 0.2 in data and MC respectively, while the extracted jet mass resolution is 7.9 ± 0.2 in data and 7.5 ± 0.2 in MC, where the quoted uncertainties are statistical only. The mass peak position is slightly shifted with respect to the W boson mass because of the presence of extra energy deposited in the jet cone coming from pileup, underlying event, and initial-state radiation not completely removed by the jet pruning procedure. For events with top quarks, additional energy contributions come also from the possible presence of a b jet close to the W-jet candidate.

6 Modelling of background and signal

The m_{WV} distribution observed in data is largely dominated by two SM background processes: inclusive W-boson production where quark (q) or gluon (g) jets are falsely identified as a Vjet and two real W bosons from $t\bar{t}$ production, single top quark production and Standard Model non-resonant diboson processes.

6.1 Background estimation

Estimation of the $t\bar{t}$ background uses the control sample defined in the previous section, Section 5. A numerical factor is derived to scale the simulated $t\bar{t}$ contribution to the data in the top-enriched control region. This factor is then used in the signal region as well, modulo corrections for b tagging efficiencies derived between data and MC. This factor is 0.81 ± 0.02 and 0.83 ± 0.01 for $\tau_{21} < 0.45$ and 0.6, respectivey, for muon and electron channels. Uncertainties in the scale factors are treated as systematic effects in the final fits. The single top and non-resonant SM diboson contributions are sub-dominant and taken from simulation with corrections to the W peak shape from data.

To estimate the background from W+jets events, we use the data in pruned jet mass sidebands to fit for the contribution in the signal region. Further, we assume that there is a correlation between the m_{jet} and m_{WV} variables for the dominant W+jets background, which can be adequately modelled by simulation. A signal-depleted control region (sideband) is defined by requiring the mass of the V-jet to be lower than that in the nominal selection; the m_{WV} distribution observed in this region is then extrapolated to the nominal region by a transfer function

estimated from simulation. The contamination from the signal (a Bulk Graviton of 750 GeV) in the sideband region is at the order of 0.2%. Other minor sources of background, such as $t\bar{t}$, single top-quark, and Standard Model diboson production, are estimated using the simulated samples after applying correction factors based on control samples in data, as described in Sec. 5. The signal-depleted control ("sideband") region is defined around the jet mass window described in Sec. 4. The *lower* and *upper* sidebands are defined in the m_{jet} ranges 40–65 GeV and 135–150 GeV, respectively.

The overall normalization of the W+jets background in the signal region is determined from a fit to the m_{jet} distribution in the lower and upper sidebands of the observed data. The analytical form of the fit function is chosen from simulation studies and the minor backgrounds are taken from the simulation. Figure 2 shows the result of this fit procedure.

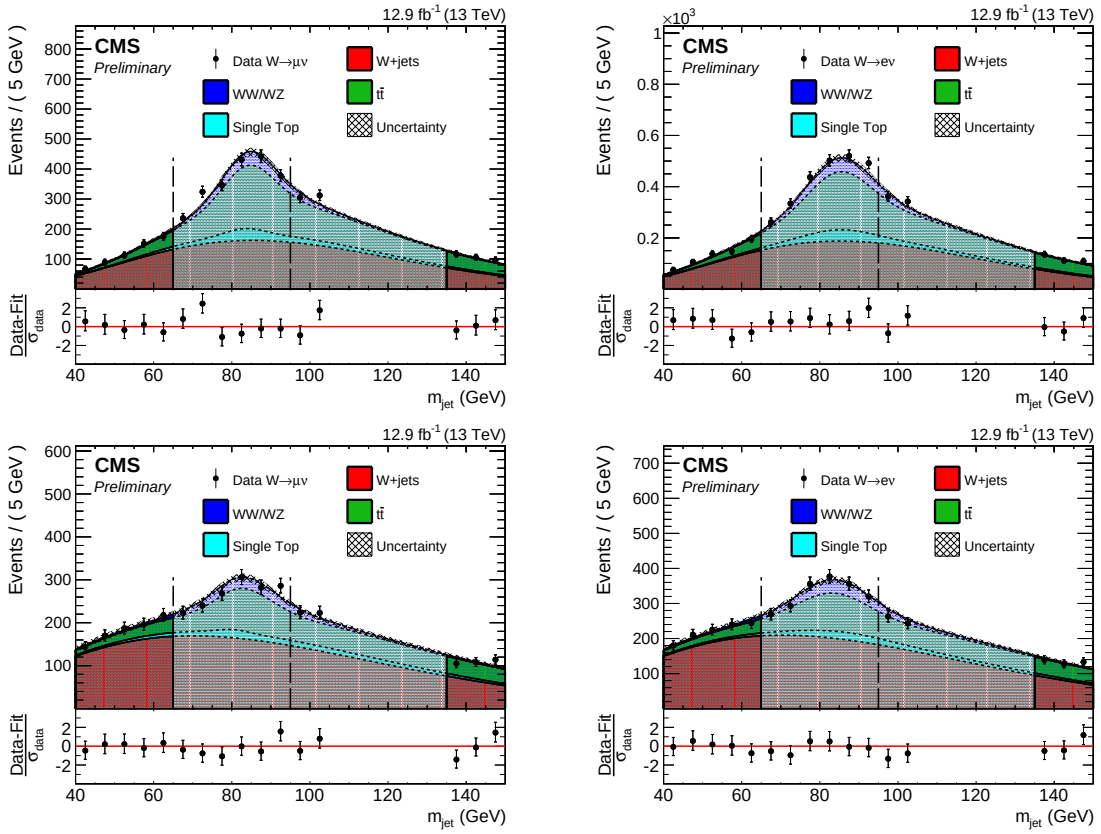

Figure 2: Distributions of the pruned jet mass m_{jet} in the muon (left) and electron (right) channels, for low mass (top) and high mass (bottom) region. The uncertainty band represents the statistical uncertainty on the background prediction. All selections are applied except the final m_{jet} signal window requirement. Data are shown as black markers. The contribution of the W+jets background is extrapolated from the sideband to the signal region 65-95 GeV, which is represented with a grey shaded area and marked with a label. The grey shaded area between 95-135 GeV is not used in the analysis in order to avoid bias in future searches for WH, ZH and HH resonances.

The shape of the m_{WV} distribution of the W+jets background in the signal region is determined from the lower m_{jet} sideband only, through the transfer function $\alpha_{\text{MC}}(m_{\text{WV}})$ derived from the

W+jets simulation, defined as:

$$\alpha_{\mathrm{MC}}(m_{\mathrm{WV}}) = \frac{F^{\mathrm{W+jets}}_{\mathrm{MC,SR}}(m_{\mathrm{WV}})}{F^{\mathrm{W+jets}}_{\mathrm{MC,SB}}(m_{\mathrm{WV}})}, \tag{2}$$

where $F(m_{\mathrm{WV}})$ is the probability density function used to describe the m_{WV} spectrum in the different regions, and has the functional form of a levelled exponential

$$F(m_{\mathrm{WV}}) = N_0 \cdot \exp\left(am_{\mathrm{WV}} + \frac{b}{m_{\mathrm{WV}}}\right), \tag{3}$$

where N_0 is a normalization parameter obtained as described above, and a and b control the shape of the function tail.

The m_{WV} distribution observed in the lower sideband region is corrected for the presence of minor backgrounds in order to have an estimation of the W+jets contribution in the control region of the data, $F^{\mathrm{W+jets}}_{\mathrm{DATA,SB}}(m_{\mathrm{WV}})$. The shape of the W+jets background distribution in the signal region is then obtained by rescaling $F^{\mathrm{W+jets}}_{\mathrm{DATA,SB}}(m_{\mathrm{WV}})$ by $\alpha_{\mathrm{MC}}(m_{\mathrm{WV}})$. The minor backgrounds are then added to the W+jets background to obtain the total SM prediction in the signal region.

Figure 3 and figure 4 show the final observed spectrum in m_{WW} and m_{WZ} of the selected data events in muon and electron channel.The observed data and the predicted background agree with each other.

6.2 Signal modeling

Figure 5 shows the simulated m_{WW} distributions for resonance masses from 600 to 4500 GeV. We adopt an analytical description of the signal shape, choosing a double-sided Crystal-Ball function [53] (i.e., a Gaussian core with power law tails on both sides) to describe the simulated resonance distributions.

7 Systematic uncertainties

The systematic uncertainties for the low mass search are largely the same as those described in [52]. We verified further that the jet energy scale and resolution uncertainties on signal yields and shapes are also conservative for low mass searches.

7.1 Systematic uncertainties in the background estimation

Uncertainties in both the shape and the normalization of the background prediction are relevant. The uncertainty of the shape is dominated by the effect of the covariance matrices of the simultaneous fits to the data sideband region and the simulation signal and sideband regions. An effect of almost equal magnitude is due to the uncertainties in the modeling transfer function $\alpha(m_{\mathrm{WV}})$ between the sideband region and the signal region. The uncertainty on the normalization of the background prediction is composed of three different effects: the W+jets component, dominated by the statistical uncertainty of the number of events in data in the pruned jet mass sideband, at the order of 3–5%; the $\mathrm{t\bar{t}}$/single top component, dominated by the b-tag scale factor uncertainties, amounting to about 5–6%; and the diboson component, dominated by the theoretical uncertainties, which is at the order of 20%.

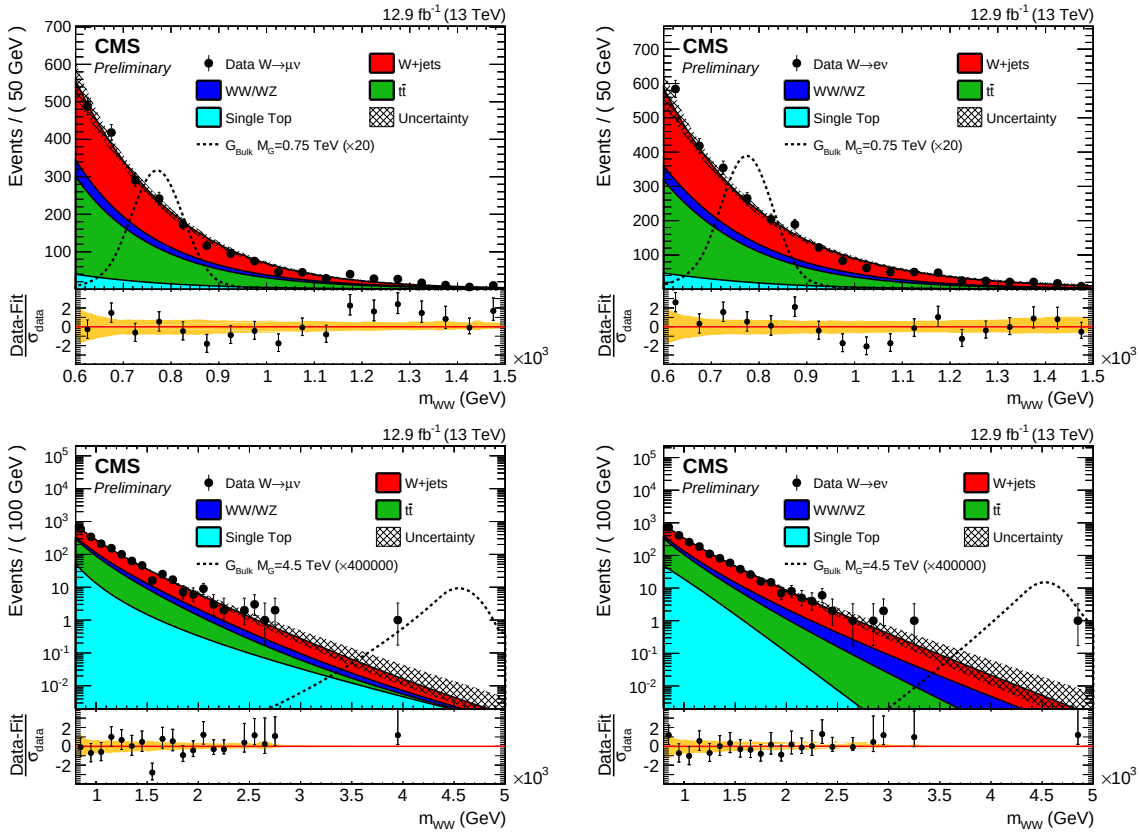

Figure 3: Final observed m_{WW} distributions for the WW analyses in the signal region. Left(right) shows muon(electron) channel. Top (bottom) shows low(high) mass. The solid curve represents the background estimation provided by the data-driven method as discussed in Sec. 6.1. The hatched band includes both statistical and systematic (of normalization and of shape) uncertainties. The data are shown as black markers. The bottom panels show the corresponding pull distributions, quantifying the agreement between the background-only hypothesis and the data. The pull distribution is defined as the difference between the data and the background prediction, divided by the error on data. The error bars on the points represent the statistical uncertainty of the data, while the uncertainty band (statistics+systematics) on the background prediction is shown with a yellow band.

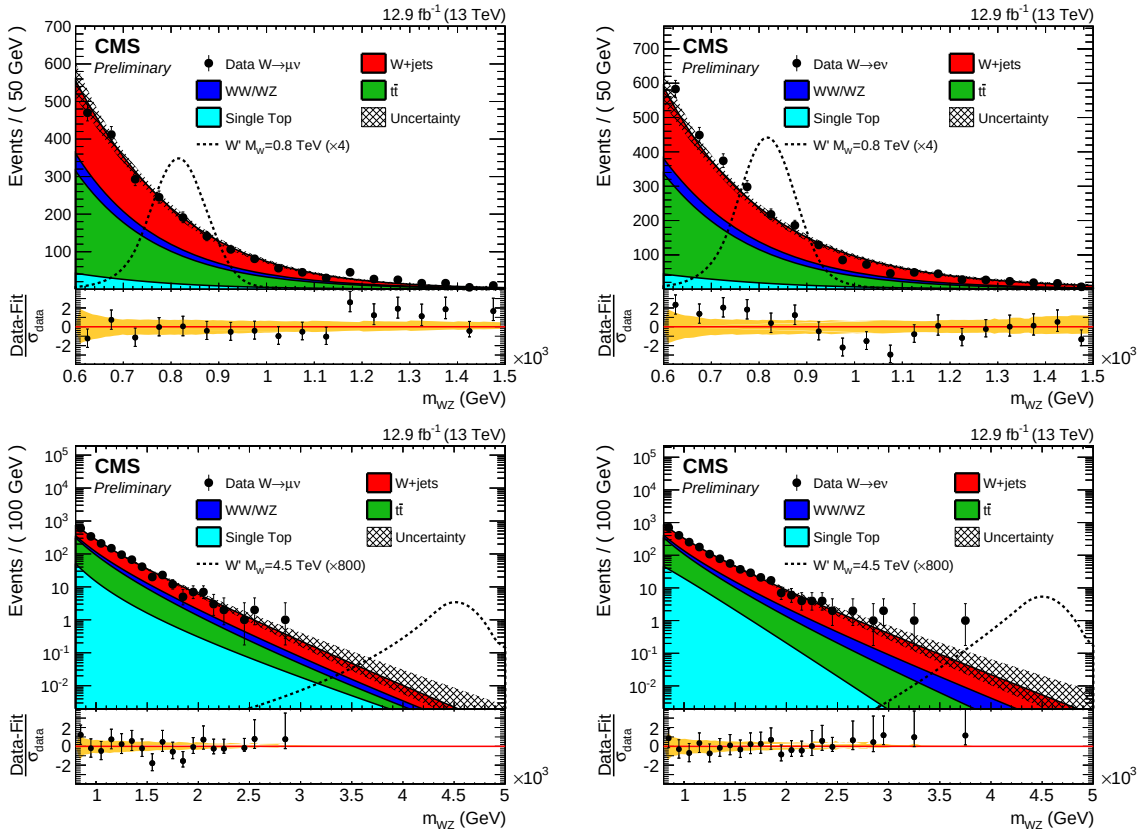

Figure 4: Final observed m_{WZ} distributions for the WZ analyses in the signal region. Left(right) shows muon(electron) channel. Top (bottom) shows low(high) mass. The solid curve represents the background estimation provided by the data-driven method as discussed in Sec. 6.1. The hatched band includes both statistical and systematic (of normalization and of shape) uncertainties. The data are shown as black markers. The bottom panels show the corresponding pull distributions, quantifying the agreement between the background-only hypothesis and the data. The pull distribution is defined as the difference between the data and the background prediction, divided by the error on data. The error bars on the points represent the statistical uncertainty of the data, while the uncertainty band (statistics+systematics) on the background prediction is shown with a yellow band.

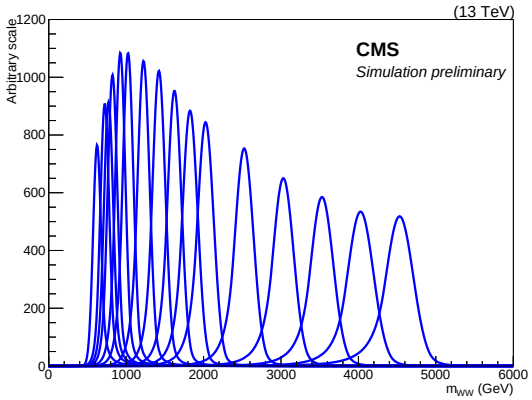

Figure 5: m_{WW} distribution for different signal mass points. The area of each shape is proportional to the total signal efficiency of the corresponding mass point.

7.2 Systematic uncertainties in the signal prediction

The dominant uncertainty on the signal selection efficiency arises from uncertainties in the lepton trigger and identification, and the efficiency of tagging the hadronic jet as a V boson (V tagging). As described in Sec. 5, the efficiency of the V tagging selection is measured in data using a sample enriched in semi-leptonic $t\bar{t}$ events. A simultaneous fit to that data sample and to a corresponding simulated sample of a suitable mixture of $t\bar{t}$, single top and W+jets yields both a correction factor to the V tagging efficiency on signal samples as well as the systematic uncertainty on that efficiency. The normalization uncertainties are summarized in Tab. 1.

Uncertainties on the reconstruction of jets affect both the signal efficiency and the shape of the reconstructed resonance in the signal samples. The four-momenta of the reconstructed jets are rescaled (smeared) according to the uncertainties on the jet energy-momentum scale (jet energy-momentum resolution). The selection efficiencies are recalculated on these modified samples, with the resulting changes taken as a systematic uncertainty depending on the resonance mass. Additionally, the induced changes on the shape of the reconstructed resonances are propagated as uncertainties on signal lineshape.

Changes in the lepton energy/momentum are propagated to the reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$, and propagated through the entire analysis. The relative variation in the number of the selected signal events is taken as a systematic uncertainty on signal normalization; for both lepton flavors, those uncertainties are smaller than 1%. These uncertainties are uncorrelated for the different lepton flavors but correlated with the obtained signal efficiency. In addition, when fitting the nominal signal lineshape and the scaled lineshapes, the observed variation of the peak position (mean of the m_{WV} distribution) and its width (RMS of the m_{WV} distribution) are added as a systematic uncertainty on the fitted signal shape. Again, for both lepton flavors, those uncertainties are smaller than 1%.

The systematic uncertainties in the lepton trigger, identification, and isolation efficiencies are derived using a tag-and-probe analysis in $\mathrm{Z} \to \ell\ell$ events. An uncertainty of 5% is assigned to the trigger efficiency for both lepton flavors. For the lepton identification and isolation efficiency, the systematic uncertainty is estimated to be 5% for the muon and electron flavors.

The uncertainty on the knowledge of the integrated luminosity [54] of the data sample (6.2%) introduces an uncertainty on the number of signal events passing the final selection. This uncertainty is fully correlated in all channels. The uncertainty due to parton distribution functions varies from 5 (10)% to 30 (100)% for W′ (bulk graviton) mass increased from 0.6 TeV to 4.5TeV,

while the scale uncertainties varies from 1 (10)% to 15 (25)%.

Tables 1 and 2 summarize the systematic uncertainties for this analysis.

Table 1: Summary of the systematic uncertainties and their impact on the signal yield and reconstructed m_{WV} shape (mean and width) for both muon and electron channels. Ranges are denoted in square brackets.

Source	Effect	$\mu\nu$+jet uncertainty	eν+jet uncertainty
Lepton trigger	Yield	5%	5%
Lepton identification	Yield	5%	5%
b tag	Yield	0.6%	
Jet energy scale	Yield	[1-2]%	
Jet energy scale	Shape (mean)	1.3%	
Jet energy scale	Shape (width)	[2%–3%]	
Jet mass scale	Yield	[1-4]%	
Jet mass resolution	Yield	[0.1-2]%	
Jet energy resolution	Yield	$<$0.1%	
Jet energy resolution	Shape (mean)	0.1%	
Jet energy resolution	Shape (width)	4%	
Integrated luminosity	Yield	6.2%	
W′ PDF	Yield	[5-30]%	
W′ Scale	Yield	[1-15]%	
Bulk graviton PDF	Yield	[10-100]%	
Bulk graviton Scale	Yield	[10-25]%	
V tagging τ_{21}	Yield	5%	

Table 2: Systematic uncertanties affecting the background normalization. Ranges are denoted in square brackets.

Source	$\mu\nu$+jet uncertainty	eν+jet uncertainty
Lepton trigger	5%	5%
Lepton identification	5%	5%
Integrated luminosity	6.2%	
Diboson cross section	20%	
$t\bar{t}$ and single top normalization	2%	
W+jets normalization, low (high) mass	5%(4%)	5%(3.5%)
Single top b tag	5%	
$t\bar{t}$ b tag	6%	
Diboson b tag	0.6%	
V tagging τ_{21}	5%	

8 Statistical interpretation

The comparison between the m_{WV} distribution observed in data and the standard model background prediction is used to test the standard model against the presence of a new resonance decaying to vector bosons. We set upper limits on the production cross section of such a resonance by combining the electron and muon event categories. We follow the modified frequentist prescription described in Refs. [55, 56] (asymptotic CL_S method). The limits are computed using a shape analysis. Systematic uncertainties are treated as nuisance parameters and profiled in the statistical interpretation using log-normal priors.

Exclusion limits are set in the context of the bulk graviton model, under the assumption of a natural width negligible with respect to the experimental resolution (narrow-width approximation). Figure 6 and figure 7 show the 95% confidence level (CL) expected and observed exclusion limits as a function of the resonance mass for each channel individually and also for their combination. The limits are compared with the cross section times the branching fraction to WW for a bulk graviton with $k/\overline{M}_{\mathrm{Pl}} = 0.5$.

Exclusion limits are also set in the context of HVT model A. Figure 8 and figure 9 show the 95% confidence level (CL) expected and observed exclusion limits as a function of the resonance mass for each channel individually and also for their combination. The limits are compared with the cross section times the branching fraction to WZ for a W′ boson from HVT model A.

An uncertainty of about $5-20\%$ on the observed limit is included due to the choice of the background functional form. This value has been estimated through several bias tests, performed evaluating the results using different parameterizations of the background.

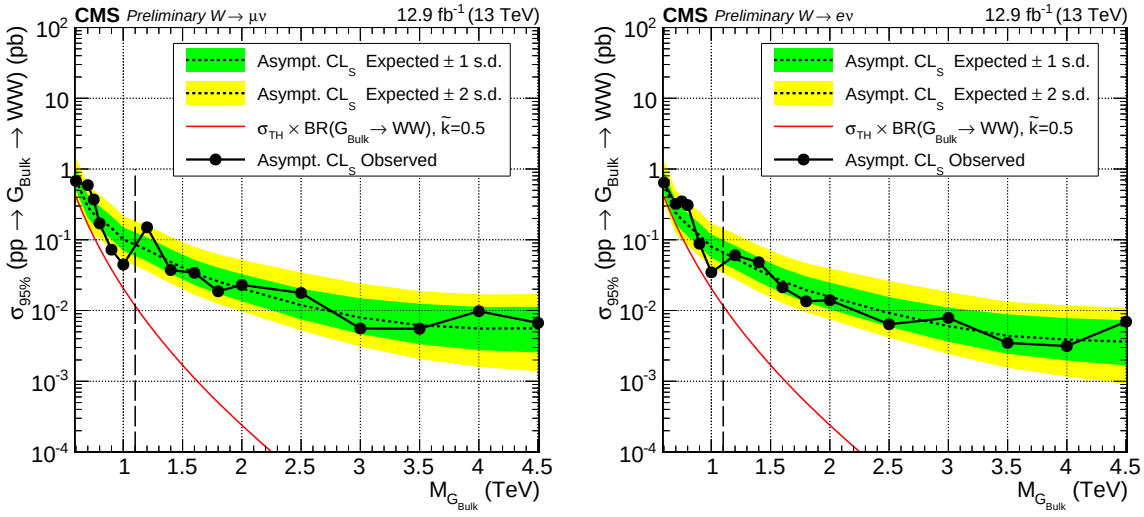

Figure 6: Observed (black solid) and expected (black dashed) 95% CL upper limits on the product of the graviton production cross section and the branching fraction of $G_{\mathrm{bulk}} \rightarrow WW$ in the muon(left) and electron(right) channel. The theoretical cross section multiplied by the relevant branching ratio is shown as a red solid line. The dashed vertical line delineates the transition between the low and high mass searches.

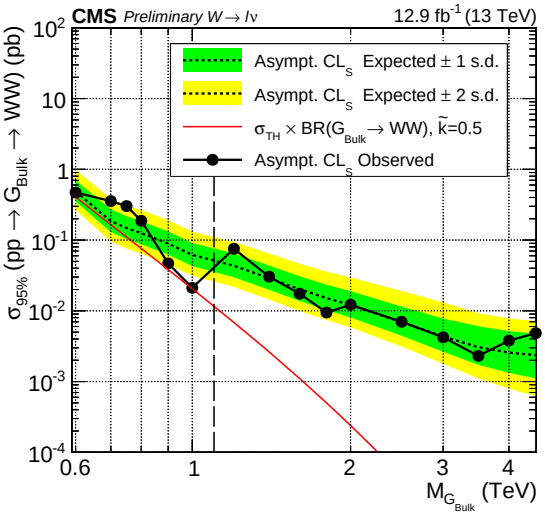

Figure 7: Observed (black solid) and expected (black dashed) 95% CL upper limits on the product of the graviton production cross section and the branching fraction of $G_{bulk} \rightarrow WW$ for the statistical combination of electron and muon channels. The theoretical cross section multiplied by the relevant branching ratio is shown as a red solid line. The dashed vertical line delineates the transition between the low and high mass searches.

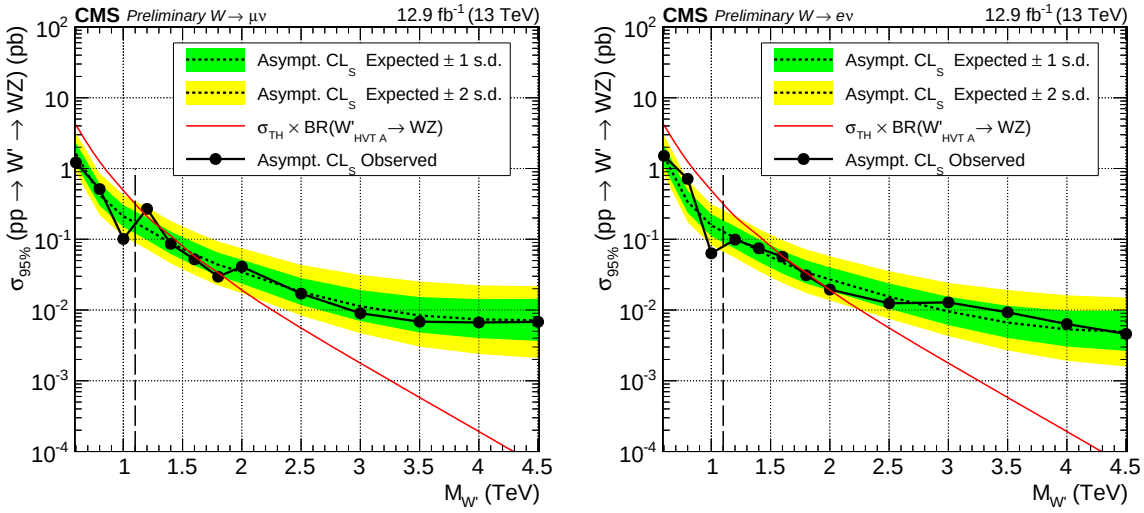

Figure 8: Observed (black solid) and expected (black dashed) 95% CL upper limits on the product of the W′ production cross section and the branching fraction of $W' \rightarrow WZ$ in the muon(left) and electron(right) channel. The theoretical cross section multiplied by the relevant branching ratio is shown as a red solid line. The dashed vertical line delineates the transition between the low and high mass searches.

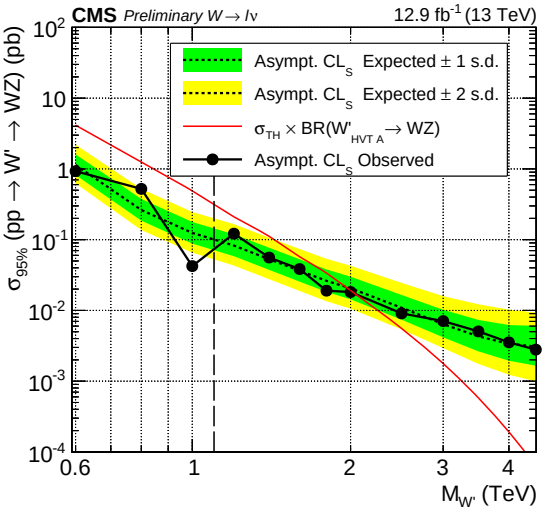

Figure 9: Observed (black solid) and expected (black dashed) 95% CL upper limits on the product of the W′ production cross section and the branching fraction of W′ $\rightarrow$ WZ for the statistical combination of electron and muon channels. The theoretical cross section multiplied by the relevant branching ratio is shown as a red solid line. The dashed vertical line delineates the transition between the low and high mass searches.

9 Summary

We have presented a search for new resonances decaying to WW and WZ in which one of the bosons decays hadronically. The final state considered is $\ell\nu q\bar{q}^{(\prime)}$ with $\ell = \mu$ or e. The results include the case in which $W \to \tau\nu$ where τ decays leptonically. W or Z bosons that decay to quarks are identified by requiring a jet with mass compatible with the W or Z noson mass. Additional information from jet substructure is used to reduce the background from W+jets and multijet processes. No evidence for a signal is found, and upper limits at 95% CL are set for resonance masses between 600 and 4500 GeV on the bulk graviton $\to$ WW and HVT $W' \to$ WZ production cross section, in the range of 400 to 4 fb and 1000 to 3 fb, respectively. These cross section limits are the most stringent to date in these final states.

References

[1] K. Agashe, H. Davoudiasl, G. Perez, and A. Soni, "Warped Gravitons at the LHC and Beyond", *Phys. Rev. D* **76** (2007) 036006, doi:10.1103/PhysRevD.76.036006, arXiv:hep-ph/0701186.

[2] A. L. Fitzpatrick, J. Kaplan, L. Randall, and L.-T. Wang, "Searching for the Kaluza-Klein Graviton in Bulk RS Models", *JHEP* **09** (2007) 013, doi:10.1088/1126-6708/2007/09/013, arXiv:hep-ph/0701150.

[3] O. Antipin, D. Atwood, and A. Soni, "Search for RS gravitons via W(L)W(L) decays", *Phys. Lett. B* **666** (2008) 155–161, doi:10.1016/j.physletb.2008.07.009, arXiv:0711.3175.

[4] L. Randall and R. Sundrum, "A Large mass hierarchy from a small extra dimension", *Phys. Rev. Lett.* **83** (1999) 3370–3373, doi:10.1103/PhysRevLett.83.3370, arXiv:hep-ph/9905221.

[5] L. Randall and R. Sundrum, "An Alternative to compactification", *Phys. Rev. Lett.* **83** (1999) 4690–4693, doi:10.1103/PhysRevLett.83.4690, arXiv:hep-th/9906064.

[6] D. Pappadopulo, A. Thamm, R. Torre, and A. Wulzer, "Heavy Vector Triplets: Bridging Theory and Data", *JHEP* **09** (2014) 060, doi:10.1007/JHEP09(2014)060, arXiv:1402.4431.

[7] ATLAS Collaboration, "Searches for heavy diboson resonances in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector", arXiv:1606.04833.

[8] CMS Collaboration, "Search for WW in semileptonic final states: low mass extension", CMS Physics Analysis Summary CMS-PAS-B2G-16-004, 2016.

[9] CMS Collaboration, "Combination of diboson resonance searches at 8 and 13 TeV", CMS Physics Analysis Summary CMS-PAS-B2G-16-007, 2016.

[10] CMS Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET", CMS Physics Analysis Summary CMS-PAS-PFT-09-001, 2009.

[11] CMS Collaboration, "Commissioning of the Particle-flow Event Reconstruction with the first LHC collisions recorded in the CMS detector", CMS Physics Analysis Summary CMS-PAS-PFT-10-001, 2010.

[12] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.

[13] J. Alwall et al., "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations", *JHEP* **07** (2014) 079, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.

[14] P. Nason, "A New method for combining NLO QCD with shower Monte Carlo algorithms", *JHEP* **11** (2004) 040, doi:10.1088/1126-6708/2004/11/040, arXiv:hep-ph/0409146.

[15] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method", *JHEP* **11** (2007) 070, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.

[16] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX", *JHEP* **06** (2010) 043, doi:10.1007/JHEP06(2010)043, arXiv:1002.2581.

[17] S. Alioli, P. Nason, C. Oleari, and E. Re, "NLO single-top production matched with shower in POWHEG: s- and t-channel contributions", *JHEP* **09** (2009) 111, doi:10.1007/JHEP02(2010)011,10.1088/1126-6708/2009/09/111, arXiv:0907.4076. [Erratum: JHEP02,011(2010)].

[18] E. Re, "Single-top Wt-channel production matched with parton showers using the POWHEG method", *Eur. Phys. J. C* **71** (2011) 1547, doi:10.1140/epjc/s10052-011-1547-z, arXiv:1009.2450.

[19] S. Alioli, S.-O. Moch, and P. Uwer, "Hadronic top-quark pair-production with one jet and parton showering", *JHEP* **01** (2012) 137, doi:10.1007/JHEP01(2012)137, arXiv:1110.5251.

[20] T. Sjostrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 Physics and Manual", *JHEP* **05** (2006) 026, doi:10.1088/1126-6708/2006/05/026, arXiv:hep-ph/0603175.

[21] T. Sjostrand, S. Mrenna, and P. Z. Skands, "A Brief Introduction to PYTHIA 8.1", *Comput. Phys. Commun.* **178** (2008) 852–867, doi:10.1016/j.cpc.2008.01.036, arXiv:0710.3820.

[22] P. Skands, S. Carrazza, and J. Rojo, "Tuning PYTHIA 8.1: the Monash 2013 Tune", *Eur. Phys. J. C* **74** (2014), no. 8, 3024, doi:10.1140/epjc/s10052-014-3024-y, arXiv:1404.5630.

[23] CMS Collaboration, "Underlying Event Tunes and Double Parton Scattering", CMS Physics Analysis Summary CMS-PAS-GEN-14-001, 2014.

[24] R. D. Ball et al., "Impact of Heavy Quark Masses on Parton Distributions and LHC Phenomenology", *Nucl. Phys. B* **849** (2011) 296–363, doi:10.1016/j.nuclphysb.2011.03.021, arXiv:1101.1300.

[25] GEANT4 Collaboration, "GEANT4: A Simulation toolkit", *Nucl. Instrum. Meth. A* **506** (2003) 250–303, doi:10.1016/S0168-9002(03)01368-8.

[26] J. M. Campbell, R. K. Ellis, and D. L. Rainwater, "Next-to-leading order QCD predictions for W + 2 jet and Z + 2 jet production at the CERN LHC", *Phys. Rev. D* **68** (2003) 094021, doi:10.1103/PhysRevD.68.094021, arXiv:hep-ph/0308195.

[27] J. M. Campbell, R. K. Ellis, and C. Williams, "Vector boson pair production at the LHC", *JHEP* **07** (2011) 018, doi:10.1007/JHEP07(2011)018, arXiv:1105.0020.

[28] J. M. Campbell and R. K. Ellis, "Top-quark processes at NLO in production and decay", *J. Phys. G* **42** (2015), no. 1, 015005, doi:10.1088/0954-3899/42/1/015005, arXiv:1204.1513.

[29] J. M. Campbell, R. K. Ellis, and F. Tramontano, "Single top production and decay at next-to-leading order", *Phys. Rev. D* **70** (2004) 094012, doi:10.1103/PhysRevD.70.094012, arXiv:hep-ph/0408158.

[30] Y. Li and F. Petriello, "Combining QCD and electroweak corrections to dilepton production in FEWZ", *Phys. Rev. D* **86** (2012) 094034, doi:10.1103/PhysRevD.86.094034, arXiv:1208.5967.

[31] M. Czakon and A. Mitov, "Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders", *Comput. Phys. Commun.* **185** (2014) 2930, doi:10.1016/j.cpc.2014.06.021, arXiv:1112.5675.

[32] CMS Collaboration, "Tracking and Primary Vertex Results in First 7 TeV Collisions", CMS Physics Analysis Summary CMS-PAS-TRK-10-005, 2010.

[33] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet User Manual", *Eur. Phys. J. C* **72** (2012) 1896, doi:10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097.

[34] M. Wobisch and T. Wengler, "Hadronization corrections to jet cross-sections in deep inelastic scattering", in *Monte Carlo generators for HERA physics. Proceedings, Workshop, Hamburg, Germany, 1998-1999*. 1998. arXiv:hep-ph/9907280.

[35] M. Cacciari, G. P. Salam, and G. Soyez, "The Anti-k(t) jet clustering algorithm", *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.

[36] CMS Collaboration, "Identification of b quark jets at the CMS Experiment in the LHC Run 2", CMS Physics Analysis Summary CMS-PAS-BTV-15-001, CERN, 2016.

[37] CMS Collaboration, "Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS", *JINST* **6** (2011) P11002, doi:10.1088/1748-0221/6/11/P11002, arXiv:1107.4277.

[38] CMS Collaboration, "Studies of jet mass in dijet and W/Z + jet events", *JHEP* **05** (2013) 090, doi:10.1007/JHEP05(2013)090, arXiv:1303.4811.

[39] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, "Techniques for improved heavy particle searches with jet substructure", *Phys. Rev. D* **80** (2009) 051501, doi:10.1103/PhysRevD.80.051501, arXiv:0903.5081.

[40] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, "Recombination Algorithms and Jet Substructure: Pruning as a Tool for Heavy Particle Searches", *Phys. Rev. D* **81** (2010) 094023, doi:10.1103/PhysRevD.81.094023, arXiv:0912.0033.

[41] J. Thaler and K. Van Tilburg, "Identifying Boosted Objects with N-subjettiness", *JHEP* **03** (2011) 015, doi:10.1007/JHEP03(2011)015, arXiv:1011.2268.

[42] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber, "Longitudinally invariant K_t clustering algorithms for hadron hadron collisions", *Nucl. Phys. B* **406** (1993) 187–224, doi:10.1016/0550-3213(93)90166-M.

[43] S. D. Ellis and D. E. Soper, "Successive combination jet algorithm for hadron collisions", *Phys. Rev. D* **48** (1993) 3160–3166, doi:10.1103/PhysRevD.48.3160, arXiv:hep-ph/9305266.

[44] CMS Collaboration, "Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV", *JINST* **7** (2012) P10002, doi:10.1088/1748-0221/7/10/P10002, arXiv:1206.4071.

[45] CMS Collaboration, "Measurements of Inclusive W and Z Cross Sections in pp Collisions at $\sqrt{s} = 7$ TeV", *JHEP* **01** (2011) 080, doi:10.1007/JHEP01(2011)080, arXiv:1012.2466.

[46] CMS Collaboration, "Energy Calibration and Resolution of the CMS Electromagnetic Calorimeter in pp Collisions at $\sqrt{s} = 7$ TeV", *JINST* **8** (2013) P09009, doi:10.1088/1748-0221/8/09/P09009, arXiv:1306.2016.

[47] CMS Collaboration, "Search for leptonic decays of W' bosons in pp collisions at $\sqrt{s} = 7$ TeV", *JHEP* **08** (2012) 023, doi:10.1007/JHEP08(2012)023, arXiv:1204.4764.

[48] CMS Collaboration, "Missing transverse energy performance of the CMS detector", *JINST* **6** (2011) P09001, doi:10.1088/1748-0221/6/09/P09001, arXiv:1106.5048.

[49] CMS Collaboration, "Performance of Missing Transverse Momentum Reconstruction Algorithms in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV with the CMS Detector", CMS Physics Analysis Summary CMS-PAS-JME-12-002, 2013.

[50] Particle Data Group Collaboration, "Review of Particle Physics (RPP)", *Phys. Rev. D* **86** (2012) 010001, doi:10.1103/PhysRevD.86.010001.

[51] CMS Collaboration, "Identification techniques for highly boosted W bosons that decay into hadrons", *JHEP* **12** (2014) 017, doi:10.1007/JHEP12(2014)017, arXiv:1410.4227.

[52] CMS Collaboration, "Search for massive resonances decaying into pairs of boosted W and Z bosons at $\sqrt{s} = 13$ TeV", CMS Physics Analysis Summary CMS-PAS-EXO-15-002, 2015.

[53] M. J. Oreglia, "A study of the reactions $\psi' \to \gamma\gamma\psi$". PhD thesis, Stanford University, 1980. SLAC Report SLAC-R-236.

[54] CMS Collaboration, "Preliminary CMS Luminosity Measurement for the 2015 Data Taking Period", CMS Physics Analysis Summary CMS-PAS-LUM-15-001, 2015.

[55] A. L. Read, "Presentation of search results: The CL(s) technique", *J. Phys. G* **28** (2002) 2693–2704, doi:10.1088/0954-3899/28/10/313.

[56] T. Junk, "Confidence level computation for combining searches with small statistics", *Nucl. Instrum. Meth. A* **434** (1999) 435–443, doi:10.1016/S0168-9002(99)00498-2, arXiv:hep-ex/9902006.