CMS Physics Analysis Summary

Contact: cms-pag-conveners-b2g@cern.ch 2024/03/26

Search for a heavy resonance decaying into ZH in events with an energetic jet and two electrons, two muons, or missing transverse momentum

The CMS Collaboration

Abstract

A search is presented for a heavy resonance decaying into a Z boson and a Higgs (H) boson. The analysis uses data from proton-proton collisions at a centre-of-mass energy of 13 TeV corresponding to 138 fb^{-1} of integrated luminosity, recorded with the CMS experiment in the years 2016 to 2018. Resonance masses between 1.4 and 5 TeV are considered, resulting in large transverse momenta of the Z and H bosons. The search targets the Z boson decay into two electrons, two muons, or two neutrinos. The H boson is reconstructed with a single large-radius jet, recoiling against the Z boson. The search is designed for the hadronic H boson decay modes $H \to c\bar{c}$ and $H \to VV^* \to 4$ quarks, where V denotes a W or Z boson. It achieves complementary sensitivity to previous searches targeting the $H \to b\bar{b}$ decays for high resonance masses.

1 Introduction

The standard model (SM) of particle physics is a powerful theory with remarkable predictive power. Nonetheless, there are indications that the SM is neither complete nor final. A plethora of theories extending the SM have been proposed and are currently being tested by experimental searches at the CERN LHC [1].

From an experimental perspective, direct searches for new elementary particles are typically not sensitive to all the free parameters of the underlying theory, but instead only to those that affect the production and decay rates of the new particles being searched for. As a consequence, simplified descriptions are commonly used in which the complete theory is approximated by an effective interaction that captures the dynamics of the targeted signal. One such simplified model is the Heavy Vector Triplet (HVT) model [2], which describes the production and decay of electroweak spin-1 resonances that arise from different theories such as weakly coupled [3–6], little Higgs models [7, 8], and composite Higgs scenarios [9–13]. Previous searches for a heavy resonance decaying to an SM Higgs (H) boson and an electroweak vector (V = W, Z) boson have already been carried out in semileptonic final states [14–20] and in fully hadronic final states [20–22] by the ATLAS and CMS Collaborations. The results of these searches are converted into upper limits on the production cross sections and lower limits on the heavy resonance masses in the HVT model. In this note, advanced jet identification algorithms using machine learning techniques are used to exploit the H boson decays to a pair of charm quarks ($c\bar{c}$) and four quarks through the intermediate $H \to VV^*$ decay. These identification techniques significantly improve the sensitivity at high resonance masses in the ZH channel of the spin-1 resonance decays, compared to a previous analysis by the CMS Collaboration [16].

This note presents a new model-independent search for a heavy resonance decaying into a Z and an H boson, where the H boson decays hadronically and the Z boson decays into a pair of oppositely charged leptons, e^+e^- or $\mu^+\mu^-$, or neutrinos. The H boson decay is identified with emphasis on the c flavour decay $H \to c\bar{c}$, and the four-prong decay $H \to VV^* \to 4$ quarks, which are explored for the first time in a search in the ZH channel. The results of this analysis represent a significant improvement in sensitivity compared to the previous CMS analysis [16].

The event reconstruction proceeds as follows. A pair of oppositely charged leptons or the missing transverse momentum is used to identify the Z boson decay. A large-radius jet recoiling against the reconstructed Z boson is used to identify the H boson decay. The substructure and the flavour of the H jet provide powerful tools to discriminate a potential signal from the SM background. Jet tagging plays an essential role in this analysis and state-of-the-art techniques are employed to improve the sensitivity for high resonance masses. The search is performed in a signal-enriched region by examining the distribution of the invariant mass or transverse mass of the reconstructed ZH system for a localised excess over a monotonically decreasing background distribution. The prediction of the SM background is obtained by fitting a one-dimensional function to the observed data and does not rely on simulation. The procedure is validated in a background-enriched validation region with kinematic properties similar to the signal region. The search uses proton-proton (pp) collision data at a centre-of-mass energy of 13 TeV, collected by the CMS experiment between 2016 and 2018, corresponding to a total integrated luminosity of $138\,\text{fb}^{-1}$.

2 The CMS detector and object reconstruction

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 tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter (HCAL), each composed of a barrel and two endcap sections. Forward calorimeters extend the pseudorapidity coverage provided by the barrel and endcap detectors. Muons are measured in gas-ionisation detectors embedded in the steel flux-return yoke outside the solenoid. More detailed descriptions of the CMS detector, together with a definition of the coordinate system used and the relevant kinematic variables, can be found in Refs. [23, 24].

Events of interest are selected using a two-tiered trigger system [25]. The first level, composed of custom hardware processors, uses information from the calorimeters and muon detectors to select events at a rate of around 100 kHz within a fixed latency of 4 μs [26]. The second level, known as the high-level trigger (HLT), consists of a farm of processors running a version of the full event reconstruction software optimised for fast processing, and reduces the event rate to around 1 kHz before data storage [25].

The particle-flow algorithm [27] aims to reconstruct and identify each particle in an event, with an optimised combination of information from the various elements of the CMS detector. The energy of photons is obtained from the ECAL measurement [28]. 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 [28]. The energy of muons is obtained from the curvature of the corresponding track [29]. 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 the response of the calorimeters to hadronic showers. Finally, the energy of neutral hadrons is obtained from the corresponding corrected ECAL and HCAL energies. Jet momentum is determined as the vectorial sum of all particle momenta in the jet, and is found from simulation to be, on average, within 5 to 10% of the true momentum over the whole p_{T} spectrum and detector acceptance. Additional pp interactions within the same or nearby bunch crossings (pileup) can contribute additional tracks and calorimetric energy depositions, increasing the apparent jet momentum.

The pileup per particle identification algorithm (PUPPI) [30, 31] is used to mitigate the effect of pileup at the reconstructed particle level, making use of local shape information, event pileup properties, and tracking information. A local shape variable is defined, which distinguishes between collinear and soft diffuse distributions of other particles surrounding the particle under consideration. The former is attributed to particles originating from the hard scatter and the latter to particles originating from pileup interactions. Charged particles identified to be originating from pileup vertices are discarded. For each neutral particle, a local shape variable is computed using the surrounding charged particles compatible with the primary vertex within the tracker acceptance of $|\eta| < 2.5$, where η denotes the pseudorapidity. Both charged and neutral particles are used in the region outside of the tracker coverage. The momenta of the neutral particles are then rescaled according to their probability of originating from the primary interaction vertex deduced from the local shape variable, superseding the need for jet-based pileup corrections [30].

The large-radius jets used in this analysis are clustered with the FASTJET package [32] using the anti-k_{T} algorithm [33] with a distance parameter of $R = 0.8$ (AK8 jets). The soft drop (SD) algorithm [34], which is a generalisation of the modified mass drop tagger algorithm [35], is used to identify the subjets of a boosted H boson candidate jet. This algorithm, with angular exponent $\beta = 0$ and soft cutoff threshold $z_{\mathrm{cut}} < 0.1$, is applied to AK8 jets reclustered using the Cambridge–Aachen algorithm [36, 37], and removes soft, wide-angle radiation from the jet.

Jet energy corrections are derived from simulation studies so that the average measured energy

of jets becomes identical to that of jets at the particle level. In situ measurements of the momentum balance in dijet, γ+jet, Z+jet, and multijet events are used to determine any residual differences between the jet energy scale in data and simulation, and appropriate corrections are made [38]. Additional selection criteria are applied to each jet to remove jets potentially dominated by instrumental effects or reconstruction failures [30].

The missing transverse momentum vector $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ is computed as the negative vector sum of the transverse momenta of all the PF candidates in an event and its magnitude is denoted as $p_{\mathrm{T}}^{\mathrm{miss}}$ [39]. The PUPPI algorithm is applied to reduce the pileup dependence of the $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ observable. The $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ is computed from the PF candidates weighted by their probability to originate from the primary interaction vertex [39]. The $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ is modified to account for corrections to the energy scale of the reconstructed jets in the event.

3 Data and simulated samples

Data events were collected with the CMS detector in pp collisions in the years 2016 to 2018 at $\sqrt{s} = 13\,\mathrm{TeV}$, corresponding to an integrated luminosity of $138\,\mathrm{fb}^{-1}$. The data analysed in this search were recorded by triggers requiring the presence of a single lepton or significant $p_{\mathrm{T}}^{\mathrm{miss}}$. To collect events where the Z boson decays into a pair of electrons, a combination of isolated and non-isolated electron triggers, and photon triggers is used to achieve optimal efficiency over the whole range of electron energies. Single muon triggers without isolation criteria are chosen to avoid losses in case of very collimated dimuon events.

The signal samples are generated using the HVT model [2]. Two free parameters are introduced to describe the coupling of the heavy spin-1 resonance to the H and V bosons (c_{H}), and to the fermions (c_{F}). The universality of lepton and quark couplings is assumed. These parameters are chosen to be dimensionless coefficients that control the relative contributions to the total interaction strength g_{V}. Two benchmark scenarios are considered in this analysis [2],

- Model A, with $c_{\mathrm{H}} = -0.556$, $c_{\mathrm{F}} = -1.316$ and $g_{\mathrm{V}} = 1$, and
- Model B, with $c_{\mathrm{H}} = -0.976$, $c_{\mathrm{F}} = 1.024$ and $g_{\mathrm{V}} = 3$.

We generate signal samples for different heavy resonance masses $m_{Z'}$ in the range of 1400 to 5000 GeV. The signal simulations use the narrow-width approximation, where we have verified that the natural widths of the signals are negligible compared to the detector resolution. The spin-1 resonance Z′ decays to a Z boson and an H boson in all simulated events. Only Z boson decays to leptons are simulated, whereas all possible decays of the H boson in the SM are considered.

The main backgrounds in this search originate from V boson production with additional jets (V+jets), which includes W+jets and Z+jets production. Both signal and V+jets events are generated with MADGRAPH5_aMC@NLO 2.6 [40, 41] at leading order (LO) in perturbative quantum chromodynamics (QCD). Subdominant background processes include diboson and top quark pair production, which are generated with PYTHIA 8.240 [42] at LO and POWHEG v2 [43–47] at next-to-LO (NLO), respectively. The cross section for the $\mathrm{t\bar{t}}$ background is adjusted to a prediction at next-to-NLO (NNLO) accuracy in perturbative QCD, using a next-to-next-to-leading-logarithmic soft-gluon approximation, obtained with the TOP++ 2.0 program [48].

For all simulated samples, the parton showering and hadronisation processes are simulated with PYTHIA 8, the underlying event simulation uses the CP5 tune [49], and the NNPDF 3.1 [50] NNLO parton distribution function (PDF) sets are employed.

All simulated samples are processed through a GEANT4-based [51] simulation of the CMS detector. To simulate the effect of pileup collisions, additional inelastic events are generated using PYTHIA 8 with a minimum bias cross section of 69.2 mb [52] and superimposed on the hard-scattering events. The simulation is corrected such that the distribution in the number of primary vertices matches that observed in the data.

4 Event selection and reconstruction

The final state targeted in this search consists of an energetic jet from the hadronic H boson decay, recoiling against a pair of charged leptons or $p_{\mathrm{T}}^{\mathrm{miss}}$ from the Z boson decay. Events are placed into mutually exclusive categories based on the flavour of the reconstructed leptons. Events in the charged lepton channels $\ell^+\ell^-$ ($\ell = e, \mu$) must have two leptons of the same flavour and opposite signs. Events with different-flavour leptons are discarded. In the $\ell^+\ell^-$ channels, events are considered where at least two leptons have $p_{\mathrm{T}} > 52\,\mathrm{GeV}$ and $|\eta| < 2.4$. The invariant mass of the dilepton pair must be within 81–101 GeV, consistent with the Z boson mass. To select Z bosons that possibly originate from the decay of a heavy resonance, we require the selected lepton pair to have $p_{\mathrm{T}} > 200\,\mathrm{GeV}$ and it must be isolated from the other activity in the event, except for the leptons from each other. In Z+jets events, the two leptons are expected to have a large angular separation compared to signal events, where the angular distance is defined as $\Delta R(\ell_1, \ell_2) = \sqrt{\Delta\eta(\ell_1, \ell_2)^2 + \Delta\phi(\ell_1, \ell_2)^2}$, with $\Delta\eta$ and $\Delta\phi$ denoting the differences in η and the azimuthal angle ϕ. We require that the angular separation between the two leptons must be less than 0.45. The neutrino channel requires the absence of a charged lepton and $p_{\mathrm{T}}^{\mathrm{miss}} > 250\,\mathrm{GeV}$, originating from the undetected neutrinos of the Z boson decay.

In all channels, events must contain at least one AK8 jet with $p_{\mathrm{T}} > 200\,\mathrm{GeV}$ and $|\eta| < 2.4$. The H boson candidate is reconstructed as a single AK8 jet, and must have a large azimuthal distance to the reconstructed Z boson system, $\Delta\phi(\mathrm{H}, \mathrm{Z}) > 2$.

The previous CMS result [16] has been obtained by using a combination of two mutually exclusive event categories. These are based on the number of b-tagged subjets of a boosted H boson candidate jet that includes all decay products of the H boson decay chain [53–55]. The two categories have been obtained by separating events into cases where the H jet contains two (2b category) or one or less ($\leq$1b category) b-tagged subjets. While the first category targets the $\mathrm{H} \to \mathrm{b\bar{b}}$ decays, the latter category includes all remaining hadronic H boson decays. In this work, the H boson candidates are rejected if all the SD subjets are b tagged. We introduce this requirement to exclude events that are part of the 2b category of the previous CMS analysis in the ZH channel [16]. We use the DEEPCSV algorithm [56] to identify subjets originating from the hadronisation of b quarks, identical to the previous analysis [16]. The chosen working point corresponds to a 10% misidentification rate for jets originating from light quarks or gluons, and an efficiency between 80 and 90% in selecting b-quark-initiated jets.

The primary source of background arises from V+jets production. For such events, the V boson is produced in association with a highly energetic jet, which is misidentified as an H boson jet. In the neutrino channel, the second largest source of background events are W+jets events where the charged lepton from the leptonic W boson decay is not reconstructed. The signal is searched for in the distributions of the reconstructed invariant mass of the ZH system, denoted by $m_{\mathrm{Z}'}^{\mathrm{rec}}$, in the charged lepton channels. In the neutrino channel, the sensitive distribution is the transverse mass,

$$m_{\mathrm{Z}'}^{\mathrm{T}} = \sqrt{2 p_{\mathrm{T}}^{\mathrm{H}} p_{\mathrm{T}}^{\mathrm{miss}} (1 - \cos\Delta\phi(\mathrm{H}, \vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}))}, \tag{1}$$

where $p_{\mathrm{T}}^{\mathrm{H}}$ is the transverse momentum of the H jet and $\Delta\phi(\mathrm{H}, \vec{p}_{\mathrm{T}}^{\,\mathrm{miss}})$ is the angular difference

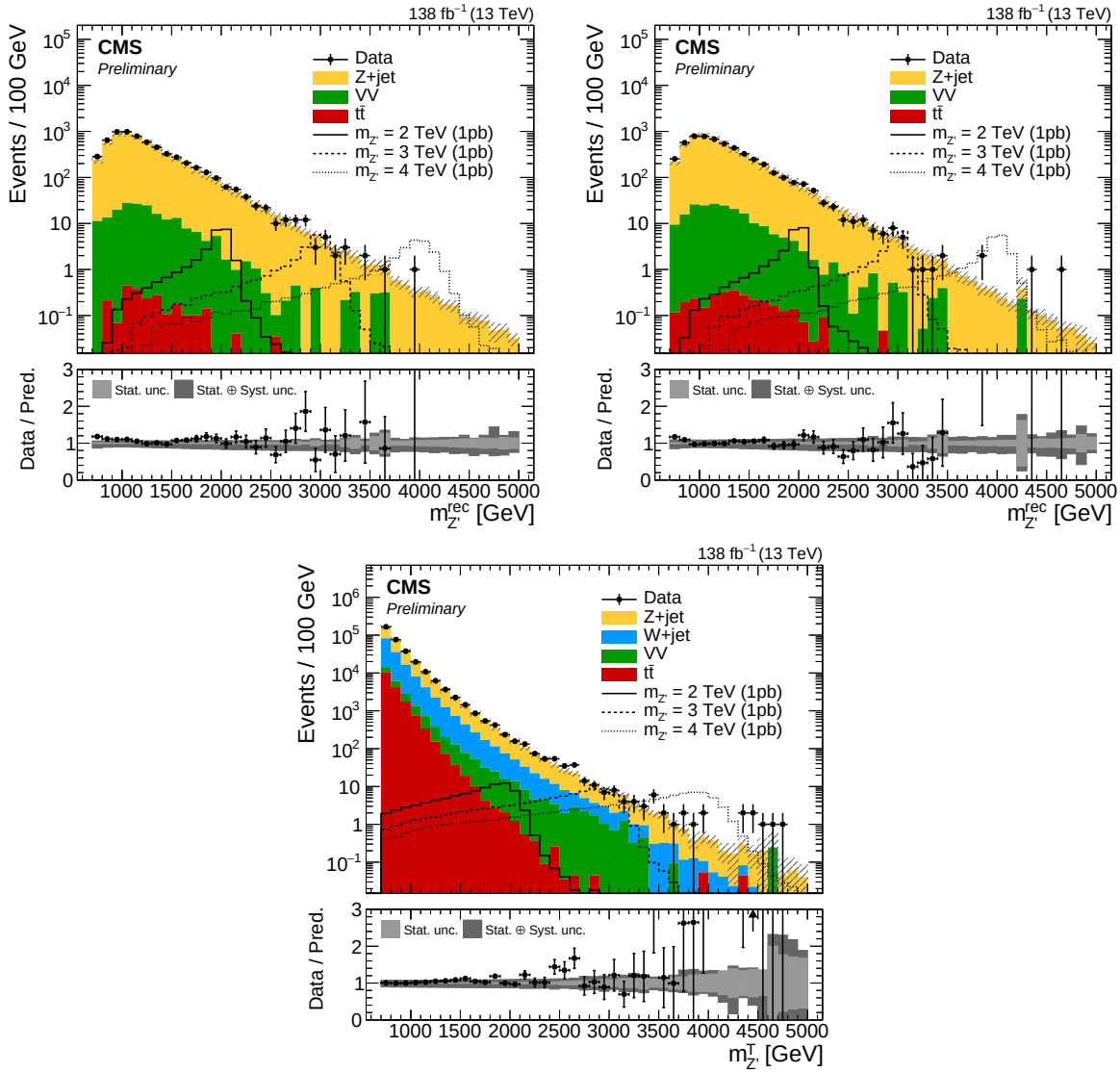

Figure 1: Distributions in $m_{Z'}^{\text{rec}}$ for the dimuon (upper left), dielectron (upper right), and in $m_{Z'}^{\text{T}}$ for the neutrino (lower) channels after the kinematic selections. The data are compared to simulation. The ratios to the total SM background are shown in the lower panels, where the statistical and total uncertainties are displayed as grey regions. The signal distributions are shown for an arbitrary cross section of 1 pb.

in azimuth between the H jet and $\vec{p}_{\text{T}}^{\text{miss}}$. Figure 1 shows the distributions in $m_{Z'}^{\text{rec}}$ for the charged lepton channels (upper) and the distribution in $m_{Z'}^{\text{T}}$ for the neutrino channel (lower).

The flavour content of the selected H jet provides discriminative power between signal and background processes. In V+jets events, jets are produced from initial or final state radiation and originate most commonly from the fragmentation of gluons or light-flavour quarks. Conversely, jets originating from decays of boosted H bosons are characterised by a more complex substructure and a larger component from the fragmentation of heavy-flavour quarks.

We improve the signal-to-background ratio by applying the ParticleNet [57] algorithm to the H jet. This algorithm is based on a deep neural network (DNN) with several classification nodes as output. A combination of the ParticleNet discriminants for the H boson decays to bb, cc, qq is used to discriminate H boson-initiated jets from light quark and gluon jets.

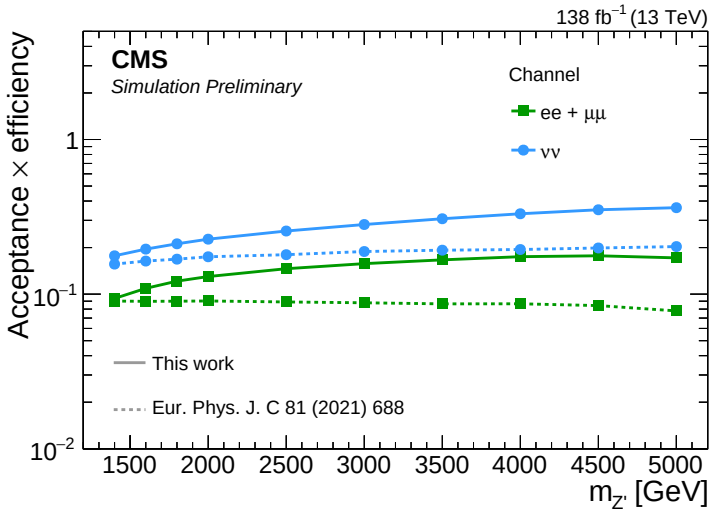

Figure 2: The product of signal acceptance and efficiency for signal events as a function of $m_{Z'}$ for the charged lepton and neutrino channels in the SR. The efficiency is calculated with respect to Z boson decays to neutrinos and to charged leptons for the neutrino and charged lepton channels, respectively. For comparison, the results from the $\leq$1b category of the previous CMS search in the ZH channel [16] are shown as dashed lines.

This combination of DNN scores is used to define our signal region (SR) by requiring that the combined score of the AK8 jet has to be greater than 0.95, corresponding to a background misidentification rate of 1%.

The product of the geometrical acceptance and the selection efficiency for different channels in the SR for all simulated signal samples is shown in Fig. 2. In the neutrino channel, the efficiency is calculated with respect to Z boson decays to neutrinos, and in the charged lepton channels with respect to Z boson decays to electrons, muons and τ leptons. We observe a gain in the product of signal acceptance and efficiency by about a factor of two compared to the $\leq$1b category of the previous CMS analysis.

5 Signal and background modelling

The search is performed by examining the distributions in $m_{Z'}^{\text{rec}}$ and $m_{Z'}^{\text{T}}$ in the SRs for a localised excess over a monotonically decreasing background distribution. The SM background prediction is obtained by fitting the free parameters of a one-dimensional function to the data. This reduces statistical fluctuations and systematic uncertainties associated with the simulation of background processes. The parametric background model is validated in a background-enriched validation region (VR), defined by inverting the selection on the DNN score. The VR has similar kinematic properties as the SR such that the functional form of the background prediction can be tested on data before examining the SRs.

The functional form used to fit the background component in the $m_{Z'}^{\text{rec}}$ distribution is given by

$$f_N(x) = \exp\left(\sum_{i=0}^{N} p_i x^i\right), \tag{2}$$

where N represents the degree of the polynomial function in the exponent and the coefficients p_i are free parameters. We test this function in fits to simulated events in the SR and VR, as well as in fits to data in the VR. Different values of N are tested, ranging from $N = 1$ to 3. A

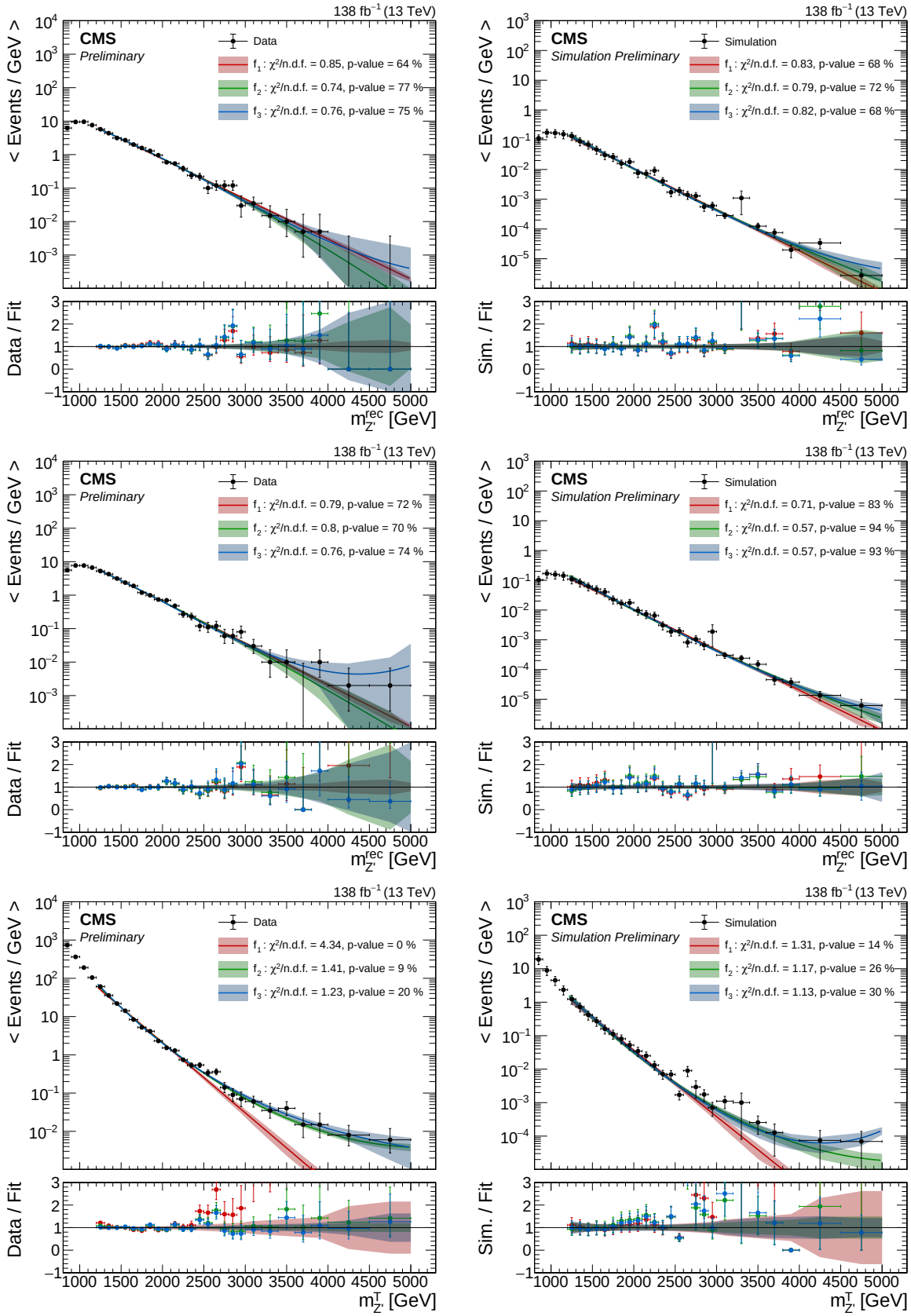

Figure 3: Fits of the background functions to the $m_{Z'}$ and $m_{Z'}^{T}$ distributions in data in the VRs (left) and simulation in the SRs (right) for the muon (upper), electron (middle), and neutrino (lower) channels. Each bin is divided by the bin width. The fit range excludes the kinematic turn-on, created by the selection criteria.

good description of the simulated backgrounds and the data in the VR is observed for $N > 1$ over the whole range up to 5 TeV. Fits of the background functions to the distributions in $m_{Z'}^{\text{rec}}$ and $m_{Z'}^{\text{T}}$ are shown in Fig. 3 for simulated events in the SR and collision data in the VR. An F-test [58] is used to determine the minimum number of free parameters needed to describe the distributions well. The result is $N = 2$, where functions with $N = 1$ are not able to describe the high-mass tail in the VR of the neutrino channel, and $N = 3$ does not lead to a statistically significant improvement of the fit quality. The simulation in the SR and the data in the VR have a higher statistical precision than the data in the SR, giving confidence that the chosen background function can reliably describe the background shape and normalisation in the SRs. The best-fit parameters and their uncertainties obtained from the simulation in the SR and the data in the VR are not used further in the analysis to not introduce a kinematic bias in the SR fits.

The line shape of the simulated signals in $m_{Z'}^{\text{rec}}$ consists of a Gaussian core centred around the generated $m_{Z'}$ and a long tail towards smaller masses. These asymmetric tails result from the off-shell production of the heavy resonance and from decay particles not reconstructed in the jet. The distributions in $m_{Z'}^{\text{rec}}$ of the signals are modelled with a Crystal-Ball function [59, 60], which describes well the line shape of the simulated signals.

6 Systematic uncertainties

The analysis is dominated by the statistical uncertainty from the limited size of the data set. This is reflected in the uncertainties in the fit parameters of the background functions from Eq. (2), which have the largest effect on the sensitivity of the analysis. Several additional sources of systematic uncertainties are considered, as these can affect the normalisation and lineshapes of the signal distributions. The effect of each source of systematic uncertainty on the signal normalisation is summarised in Tab. 1 and briefly described in the following.

Discrepancies in selection efficiencies between data and simulation are corrected with data-to-simulation scale factors (SFs). The experimental uncertainties are evaluated by varying the SFs up and down by one standard deviation for each uncertainty source. The resulting systematic uncertainties are treated as fully correlated across channels. The largest experimental systematic uncertainty is related to the H jet identification using the PARTICLENET algorithm and amounts to 2–5%. The uncertainty related to the rejection of double-b-tagged jets amounts to 0.4–1.0%. The jet energy scale and resolution uncertainties amount to up to 2% each. Uncertainties in the trigger efficiency, lepton identification and isolation are considered in the charged lepton channels, leading to uncertainties of 3–6% for the electron channel and up to 1% for the muon channel. Additional systematic uncertainties originate from estimations of the pileup contribution, the integrated luminosity [61–63], and trigger prefiring due to detector

Table 1: Sources of systematic uncertainties considered in this analysis, and their effect on the signal normalisation. The uncertainty ranges correspond to different signal masses.

Source	uncertainty	Source	uncertainty
H jet identification	2.0–5.0%	Trigger	0.9–1.5%
b tagging veto	0.4–1.0%	Muon identification	0.1–0.3%
Jet energy scale and resolution	0.2–2.0%	Electron identification	5.2–5.9%
Pileup	0.3–1.8%	Lepton reconstruction	0.9–1.7%
Luminosity	1.6%	PDF	0.3–13.4%
Prefiring	0.3–0.8%	QCD scales	6.6–17.2%

timing issues [26], each amounting up to 2%. While most of these uncertainties affect both, the normalisation and shape of the signal distributions, the effect on the normalisation is the dominating effect.

In addition, we consider theoretical uncertainties related to the production of the heavy resonance. The signal cross sections are affected by the choice of the QCD renormalisation and factorisation scales, as well as the uncertainties in the PDFs used to generate the signal samples. The effect of these is estimated following the PDF4LHC recommendations [64], and results in uncertainties of 6.6–17.2% and 0.3–13.4% for the scale variations and PDFs, respectively. The uncertainties increase towards higher signal masses. We include these theory uncertainties in the predictions of the production cross sections for Models A and B.

7 Results

The modified frequentist approach [65–67], known as the CL_s criterion with the profile likelihood ratio as the test statistic, is used in this search for setting limits on the possible presence of a signal. We use the $m_{Z'}$ and $m_{Z'}^{T}$ distributions measured in the three SRs for the statistical interpretation. The distributions in data are shown in Fig. 4. The background functions have been obtained from a fit with the background-only hypothesis, where the parameters of the fit functions have been left unconstrained to ensure an unbiased result. The data are in agreement with the predicted backgrounds from SM processes.

We set upper limits at the 95% confidence level (CL) on the product of the resonance production cross section $\sigma(pp \to Z')$ and the branching fraction $\mathcal{B}(Z' \to ZH)$ as a function of $m_{Z'}$. The expected limits obtained for the individual channels and their combination are shown in Fig. 5 (left). The neutrino channel is the most sensitive channel over the entire mass range because of the larger branching fraction and the higher selection efficiency. The final expected and observed exclusion limits resulting from the combination of the muon, electron and neutrino channels are shown in Fig. 5 (right). The observed upper limits agree with the expected limits from the background-only hypothesis within about one standard deviation over the full mass range considered.

We observe an improvement in the sensitivity and the mass exclusion limit compared to the the previous CMS analysis [16]. For $m_{Z'} > 2.4\,\text{TeV}$, the analysis has better sensitivity than the $\leq$1b category alone from the previous analysis by up to 50% on $\sigma(pp \to Z')\,\mathcal{B}(Z' \to ZH)$. The analysis has better sensitivity than the 2b category for $m_{Z'} > 3.4\,\text{TeV}$, where previously this crossing point was at 3.85 TeV. The lower Z' mass limits from this analysis alone are at 2.8 TeV and 3 TeV for Models A and B, respectively. This is the first time that identification of H boson decays to c quarks improves the sensitivity of a BSM search. These results demonstrate that the sensitivity for very high resonance masses can be improved by considering alternative decay channels than the $H \to b\bar{b}$ decay, despite its much larger branching fraction. The best possible constraints on the HVT model will be obtained from a combination with the 2b category of the previous analysis.

The upper limits on the cross sections are translated into two-dimensional upper limits on the coupling parameters for fermions, $g_F = g^2 c_F / g_V$, and bosons $g_H = c_H g_V$ in the HVT model, where g is the $SU(2)_L$ gauge coupling. The resulting exclusion contours are shown in Fig. 6.

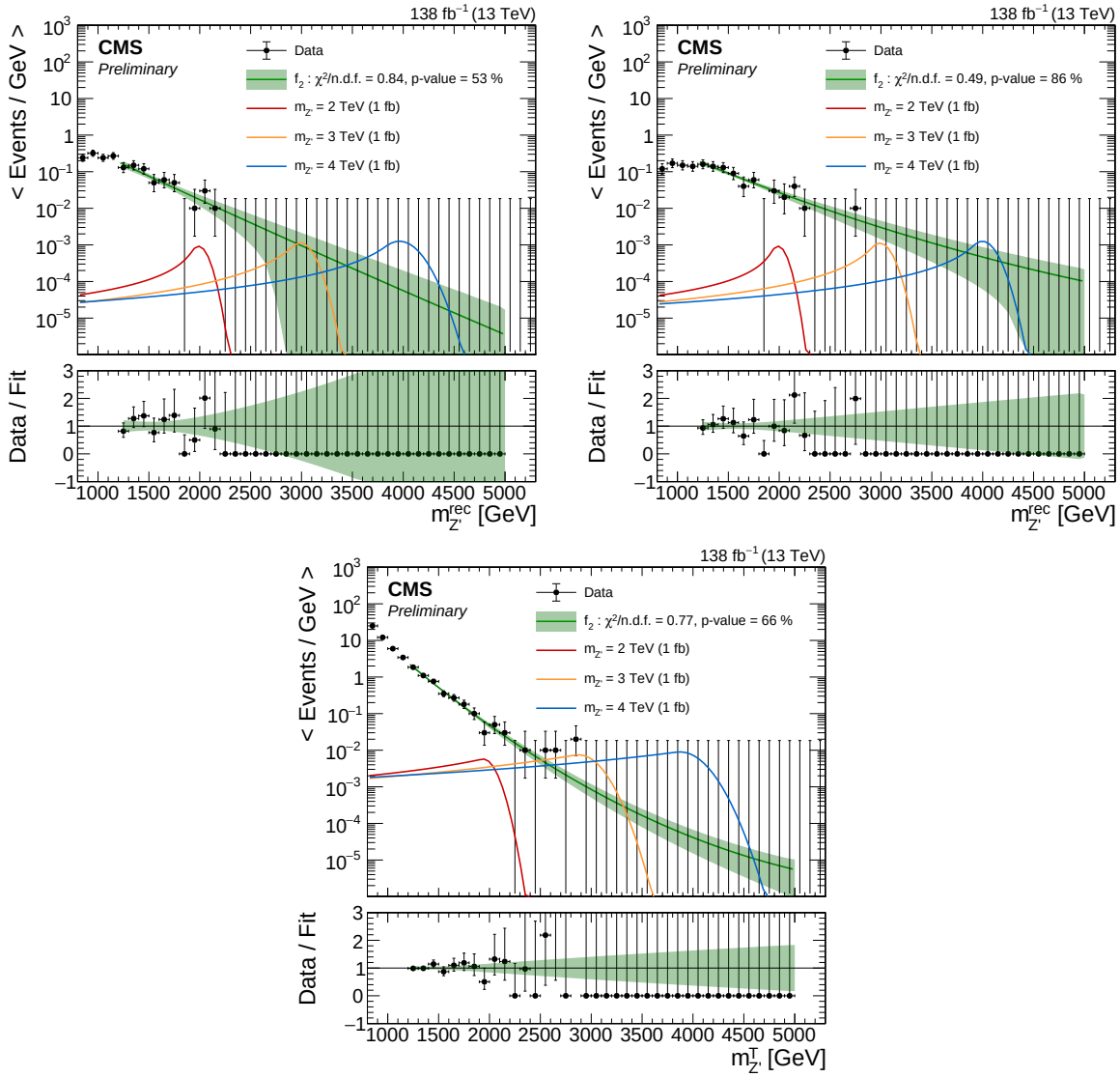

Figure 4: Distributions in $m^{\text{rec}}_{Z'}$ and $m^{\text{T}}_{Z'}$ for data in the SRs, together with fits of the background functions under the background-only hypothesis for the muon (upper left), electron (upper right), and neutrino (lower) channels. The number of observed events in each bin is divided by the bin width. The signal predictions are shown for different Z′ masses.

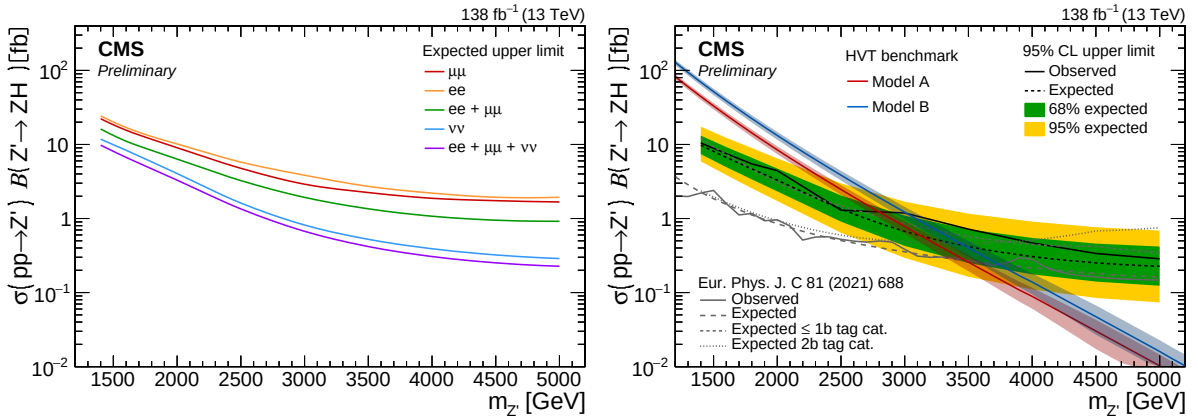

Figure 5: Expected and observed upper limits at 95% CL on the product of the production cross section $\sigma(\mathrm{pp} \to \mathrm{Z}')$ and the branching fraction $\mathcal{B}(\mathrm{Z}' \to \mathrm{ZH})$ as a function of the Z′ mass. Expected limits obtained from the three different final states are compared to the combined result (left). The expected and observed limits from the combination of all final states are compared to predictions from the HVT and limits from a previous analysis [16] (right).

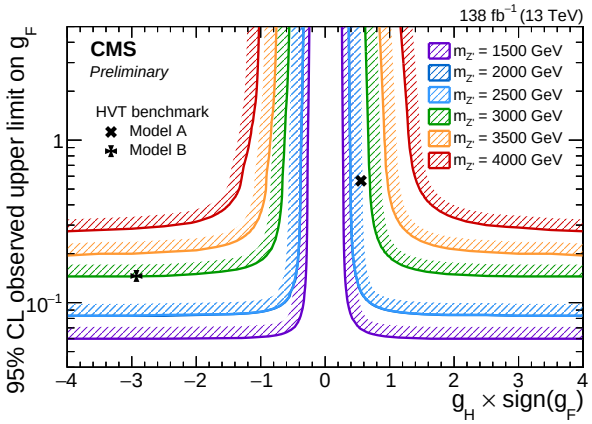

Figure 6: Observed upper limits at 95% CL on g_F for different Z′ masses as a function of the product of g_H with the sign of g_F. The two benckmark scenarios of the HVT model are shown by the black markers.

8 Summary

A search has been presented for the resonant production of a spin-1 particle with mass in the range of 1.4–5 TeV and the decay into a Z and a Higgs (H) boson. The analysis is performed using data recorded with the CMS detector at a centre-of-mass energy of 13 TeV, corresponding to an integrated luminosity of 138 fb^{-1}.

The final states explored include the Z boson decays into a pair of electrons, muons or neutrinos, and the hadronic decays of the H boson reconstructed as a single large-radius jet. A novel approach analysing the flavour content and substructure of the H boson jet was deployed to improve the sensitivity for high resonance masses. This analysis shows for the first time the benefit of including H boson decays into $\mathrm{c\bar{c}}$ and $\mathrm{VV}^* \to 4$ quarks, where V denotes a W or Z boson, besides the commonly used $\mathrm{H} \to \mathrm{b\bar{b}}$ decays in searches for new physics.

Exclusion limits at 95% confidence level are set on both the mass of a heavy resonance and the couplings to fermions and bosons in the HVT model. Resonances with masses below 3 TeV are excluded.

References

[1] S. Rappoccio, "The experimental status of direct searches for exotic physics beyond the standard model at the Large Hadron Collider", *Rev. Phys.* **4** (2019) 100027, doi:10.1016/j.revip.2018.100027, arXiv:1810.10579.

[2] 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.

[3] V. D. Barger, W. Keung, and E. Ma, "A gauge model with light W and Z bosons", *Phys. Rev. D* **22** (1980) 727, doi:10.1103/PhysRevD.22.727, arXiv:0801.1345.

[4] G. Altarelli, B. Mele, and M. Ruiz-Altaba, "Searching for new heavy vector bosons in $p\overline{p}$ colliders", *Z. Phys. C* **45** (1989) 109, doi:10.1007/BF01556677. [Erratum: doi:10.1007/BF01552335].

[5] E. Salvioni, G. Villadoro, and F. Zwirner, "Minimal Z′ models: Present bounds and early LHC reach", *JHEP* **11** (2009) 068, doi:10.1088/1126-6708/2009/11/068, arXiv:0909.1320.

[6] C. Grojean, E. Salvioni, and R. Torre, "A weakly constrained W′ at the early LHC", *JHEP* **07** (2011) 002, doi:10.1007/JHEP07(2011)002, arXiv:1103.2761.

[7] M. Schmaltz and D. Tucker-Smith, "Little Higgs review", *Ann. Rev. Nucl. Part. Sci.* **55** (2005) 229, doi:10.1146/annurev.nucl.55.090704.151502, arXiv:hep-ph/0502182.

[8] N. Arkani-Hamed, A. G. Cohen, E. Katz, and A. E. Nelson, "The littlest Higgs", *JHEP* **07** (2002) 034, doi:10.1088/1126-6708/2002/07/034, arXiv:hep-ph/0206021.

[9] R. Contino, D. Marzocca, D. Pappadopulo, and R. Rattazzi, "On the effect of resonances in composite Higgs phenomenology", *JHEP* **10** (2011) 081, doi:10.1007/JHEP10(2011)081, arXiv:1109.1570.

[10] D. Marzocca, M. Serone, and J. Shu, "General composite Higgs models", *JHEP* **08** (2012) 013, doi:10.1007/JHEP08(2012)013, arXiv:1205.0770.

[11] B. Bellazzini, C. Csáki, and J. Serra, "Composite Higgses", *Eur. Phys. J. C* **74** (2014) 2766, doi:10.1140/epjc/s10052-014-2766-x, arXiv:1401.2457.

[12] D. Greco and D. Liu, "Hunting composite vector resonances at the LHC: Naturalness facing data", *JHEP* **12** (2014) 126, doi:10.1007/JHEP12(2014)126, arXiv:1410.2883.

[13] K. Lane and L. Pritchett, "The light composite Higgs boson in strong extended technicolor", *JHEP* **06** (2017) 140, doi:10.1007/JHEP06(2017)140, arXiv:1604.07085.

[14] CMS Collaboration, "Search for heavy resonances decaying into a vector boson and a Higgs boson in final states with charged leptons, neutrinos and b quarks at $\sqrt{s} = 13\,\text{TeV}$", *JHEP* **11** (2018) 172, doi:10.1007/JHEP11(2018)172, arXiv:1807.02826.

[15] CMS Collaboration, "Search for heavy resonances decaying to ZZ or ZW and axion-like particles mediating nonresonant ZZ or ZH production at $\sqrt{s} = 13\,\text{TeV}$", *JHEP* **04** (2022) 087, doi:10.1007/JHEP04(2022)087, arXiv:2111.13669.

[16] CMS Collaboration, "Search for a heavy vector resonance decaying to a Z boson and a higgs boson in proton–proton collisions at $\sqrt{s} = 13\,\text{TeV}$", *Eur. Phys. J. C* **81** (2021) 688, doi:10.1140/epjc/s10052-021-09348-6, arXiv:2102.08198.

[17] CMS Collaboration, "Search for heavy resonances decaying to WW, WZ, or WH boson pairs in the lepton plus merged jet final state in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$", *Phys. Rev. D* **105** (2022) 032008, doi:10.1103/PhysRevD.105.032008, arXiv:2109.06055.

[18] CMS Collaboration, "Search for heavy resonances decaying to $\text{Z}(\nu\overline{\nu})\text{V}(\text{q}\overline{\text{q}}')$ in proton–proton collisions at $\sqrt{s} = 13\,\text{TeV}$", *Phys. Rev. D* **106** (2022) 012004, doi:10.1103/PhysRevD.106.012004, arXiv:2109.08268.

[19] ATLAS Collaboration, "Search for heavy diboson resonances in semileptonic final states in pp collisions at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector", *Eur. Phys. J. C* **80** (2020) 1165, doi:10.1140/epjc/s10052-020-08554-y, arXiv:2004.14636.

[20] ATLAS Collaboration, "Search for heavy resonances decaying into a Z or W boson and a Higgs boson in final states with leptons and b-jets in $139\,\text{fb}^{-1}$ of pp collisions at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector", *JHEP* **06** (2023) 016, doi:10.1007/JHEP06(2023)016, arXiv:2207.00230.

[21] ATLAS Collaboration, "Search for resonances decaying into a weak vector boson and a Higgs boson in the fully hadronic final state produced in proton–proton collisions at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector", *Phys. Rev. D* **102** (2020) 112008, doi:10.1103/PhysRevD.102.112008, arXiv:2007.05293.

[22] CMS Collaboration, "Search for new heavy resonances decaying to WW, WZ, ZZ, WH, or ZH boson pairs in the all-jets final state in proton–proton collisions at $\sqrt{s} = 13\,\text{TeV}$", *Phys. Lett. B* **844** (2023) 137813, doi:10.1016/j.physletb.2023.137813, arXiv:2210.00043.

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

[24] CMS Collaboration, "Development of the CMS detector for the CERN LHC Run 3", 2023. arXiv:2309.05466.

[25] CMS Collaboration, "The CMS trigger system", *JINST* **12** (2017) P01020, doi:10.1088/1748-0221/12/01/P01020, arXiv:1609.02366.

[26] CMS Collaboration, "Performance of the CMS Level-1 trigger in proton–proton collisions at $\sqrt{s} = 13\,\text{TeV}$", *JINST* **15** (2020) P10017, doi:10.1088/1748-0221/15/10/P10017, arXiv:2006.10165.

[27] CMS Collaboration, "Particle-flow reconstruction and global event description with the CMS detector", *JINST* **12** (2017) P10003, doi:10.1088/1748-0221/12/10/P10003, arXiv:1706.04965.

[28] CMS Collaboration, "Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC", *JINST* **16** (2021) P05014, doi:10.1088/1748-0221/16/05/P05014, arXiv:2012.06888.

[29] CMS Collaboration, "Performance of the CMS muon detector and muon reconstruction with proton–proton collisions at $\sqrt{s} = 13\,\text{TeV}$", *JINST* **13** (2018) P06015, doi:10.1088/1748-0221/13/06/P06015, arXiv:1804.04528.

[30] CMS Collaboration, "Pileup mitigation at CMS in 13 TeV data", *JINST* **15** (2020) P09018, doi:10.1088/1748-0221/15/09/p09018, arXiv:2003.00503.

[31] D. Bertolini, P. Harris, M. Low, and N. Tran, "Pileup per particle identification", *JHEP* **10** (2014) 059, doi:10.1007/JHEP10(2014)059, arXiv:1407.6013.

[32] 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.

[33] 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.

[34] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler, "Soft drop", *JHEP* **05** (2014) 146, doi:10.1007/JHEP05(2014)146, arXiv:1402.2657.

[35] M. Dasgupta, A. Fregoso, S. Marzani, and G. P. Salam, "Towards an understanding of jet substructure", *JHEP* **09** (2013) 029, doi:10.1007/JHEP09(2013)029, arXiv:1307.0007.

[36] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, "Better jet clustering algorithms", *JHEP* **08** (1997) 001, doi:10.1088/1126-6708/1997/08/001, arXiv:hep-ph/9707323.

[37] M. Wobisch and T. Wengler, "Hadronization corrections to jet cross-sections in deep inelastic scattering", in *Proceedings of the Workshop on Monte Carlo Generators for HERA Physics, Hamburg, Germany*, p. 270. 1998. arXiv:hep-ph/9907280.

[38] CMS Collaboration, "Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV", *JINST* **12** (2017) P02014, doi:10.1088/1748-0221/12/02/P02014, arXiv:1607.03663.

[39] CMS Collaboration, "Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$ using the CMS detector", *JINST* **14** (2019) P07004, doi:10.1088/1748-0221/14/07/P07004, arXiv:1903.06078.

[40] J. Alwall et al., "Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions", *Eur. Phys. J. C* **53** (2008) 473, doi:10.1140/epjc/s10052-007-0490-5, arXiv:0706.2569.

[41] 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.

[42] T. Sjöstrand et al., "An introduction to PYTHIA 8.2", *Comput. Phys. Commun.* **191** (2015) 159, doi:10.1016/j.cpc.2015.01.024, arXiv:1410.3012.

[43] 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.

[44] 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.

[45] S. Frixione, P. Nason, and G. Ridolfi, "A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction", *JHEP* **09** (2007) 126, doi:10.1088/1126-6708/2007/09/126, arXiv:0707.3088.

[46] 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.

[47] 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.

[48] 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.

[49] CMS Collaboration, "Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements", *Eur. Phys. J. C* **80** (2020) 4, doi:10.1140/epjc/s10052-019-7499-4, arXiv:1903.12179.

[50] NNPDF Collaboration, "Parton distributions from high-precision collider data", *Eur. Phys. J. C* **77** (2017) 663, doi:10.1140/epjc/s10052-017-5199-5, arXiv:1706.00428.

[51] GEANT4 Collaboration, "GEANT4—a simulation toolkit", *Nucl. Instrum. Meth. A* **506** (2003) 250, doi:10.1016/S0168-9002(03)01368-8.

[52] CMS Collaboration, "Measurement of the inelastic proton-proton cross section at $\sqrt{s} =$ 13 TeV", *JHEP* **07** (2018) 161, doi:10.1007/JHEP07(2018)161, arXiv:1802.02613.

[53] A. J. Larkoski, I. Moult, and B. Nachman, "Jet substructure at the Large Hadron Collider: A review of recent advances in theory and machine learning", *Phys. Rept.* **841** (2020) 1, doi:10.1016/j.physrep.2019.11.001, arXiv:1709.04464.

[54] R. Kogler, B. Nachman, A. Schmidt (editors) et al., "Jet substructure at the Large Hadron Collider", *Rev. Mod. Phys.* **91** (2019) 045003, doi:10.1103/RevModPhys.91.045003, arXiv:1803.06991.

[55] R. Kogler, "Advances in jet substructure at the LHC: Algorithms, measurements and searches for new physical phenomena", volume 284 of *Springer Tracts Mod. Phys.* Springer, 2021. doi:10.1007/978-3-030-72858-8, ISBN 978-3-030-72857-1, 978-3-030-72858-8.

[56] CMS Collaboration, "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV", *JINST* **13** (2018) P05011, doi:10.1088/1748-0221/13/05/P05011, arXiv:1712.07158.

[57] H. Qu and L. Gouskos, "ParticleNet: Jet tagging via particle clouds", *Phys. Rev. D* **101** (2020) 056019, doi:10.1103/PhysRevD.101.056019, arXiv:1902.08570.

[58] A. M. Mood, F. A. Graybill, and D. C. Boes, "Introduction to the theory of statistics". McGraw-Hill, 1973.

[59] M. Oreglia, "A Study of the Reactions $\psi' \to \gamma\gamma\psi$". PhD thesis, SLAC, 1980.

[60] T. Skwarnicki, "A study of the radiative CASCADE transitions between the Upsilon-Prime and Upsilon resonances". PhD thesis, Cracow, INP and DESY, 1986.

[61] CMS Collaboration, "Precision luminosity measurement in proton–proton collisions at $\sqrt{s}$ =13 TeV in 2015 and 2016 at CMS", *Eur. Phys. J. C* **81** (2021) 800, doi:10.1140/epjc/s10052-021-09538-2, arXiv:2104.01927.

[62] CMS Collaboration, "CMS luminosity measurement for the 2017 data-taking period at $\sqrt{s} = 13$ TeV", CMS Physics Analysis Summary CMS-PAS-LUM-17-004, 2018.

[63] CMS Collaboration, "CMS luminosity measurement for the 2018 data-taking period at $\sqrt{s} = 13$ TeV", CMS Physics Analysis Summary CMS-PAS-LUM-18-002, 2019.

[64] J. Butterworth et al., "PDF4LHC recommendations for LHC Run II", *J. Phys. G* **43** (2016) 023001, doi:10.1088/0954-3899/43/2/023001, arXiv:1510.03865.

[65] ATLAS and CMS Collaborations, and LHC Higgs Combination Group, "Procedure for the LHC Higgs boson search combination in Summer 2011", CMS Note CMS-NOTE-2011-005, ATL-PHYS-PUB-2011-11, 2011.

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

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