CMS Draft Analysis Note

2018/11/27
Head Id: 478965
Archive Id: 476886:479019M
Archive Date: 2018/10/23
Archive Tag: trunk

Search for high mass spin-0 resonances in the semileptonic WW $\to \ell\nu\,q\bar{q}$ final state at $\sqrt{s} = 13$ TeV with the 2016 dataset

D. Moran, M. Isabel Josa, and O. Gonzalez Lopez
CIEMAT, Madrid

Abstract

This analysis note describes the search for high mass spin-0 resonances in the semileptonic WW $\to \ell\nu\,q\bar{q}$ final state. The search is performed with an integrated luminosity of 35.9 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV in the mass range 200 GeV/c^2 $\leq m_X \leq$ 3000 GeV/c^2. In this mass regime the hadronically decaying boson may be sufficiently boosted that its decay products are contained in a single jet, thus merged jet reconstruction and substructure techniques are utilised. The case where the hadronic decay products are resolved is also considered.

Contents

1 Introduction

The ATLAS and CMS Collaborations have discovered a boson with mass close to 125 GeV and properties consistent with those expected for a scalar SM Higgs boson [1, 2]. To investigate further the SM Higgs sector, it is necessary to measure precisely the properties of the discovered Higgs boson and to perform searches for additional heavy scalar bosons that might form part of an extended Higgs sector. The existence of additional Higgs states, denoted X, are motivated in many BSM models including the electroweak singlet model [3] in which the SM Higgs boson mixes with a heavy EW singlet. A search for such a resonance in a number of diboson final states has previously been performed with the full Run 1 dataset at CMS [4]. The $WW \to \ell\nu\,q\bar{q}$ final state is the dominant decay channel of a SM-like Higgs boson for masses above 200 GeV. In this note we report the results of a search for high mass spin-0 resonances in the mass range 200 GeV/c^2 to 3000 GeV/c^2 in the semileptonic $WW \to \ell\nu\,q\bar{q}$ final state using the 2016 dataset.

In the mass regime of interest the hadronically decaying boson may be sufficiently boosted that its decay products are contained in a single merged jet. Jet substructure techniques are used to identify merged jets with two well defined subjets and to determine the merged jet mass, helping to discriminate vector boson decays from QCD jets coming from quarks and gluons. If the hadronic decay products are resolved then the boson decay may be reconstructed as two quark-jets (a dijet). In this analysis it is first attempted to reconstruct boson candidates using merged jets, if no boosted candidates are found then dijet reconstruction is attempted.

The leptonically decaying boson is reconstructed as a single isolated lepton and missing transverse energy ($\not\!\!E_T$) corresponding to the neutrino. An estimate of the neutrino longitudinal momentum is derived by imposing the constraint of the W mass on the invariant mass of the $\ell\nu$ system.

As the final state is fully reconstructable the analysis is performed by searching for a local enhancement in the diboson invariant mass spectrum. The dominant background processes are from W+jets and top production, with a smaller contribution from diboson, Z+Jets and QCD multijet events. Unlike for the signal the mass distribution of the background events is not resonant, providing a useful handle to isolate signal events. Data from signal-free control regions are used to normalize and tune the W+jets and top MC background prediction reducing the dependence on the simulation.

The kinematic information of the final state particles is fully exploited with the use of a matrix element based kinematic discriminant to enhance the sensitivity to signal. Identifying additional jets in the event also leads to increased sensitivity to signal produced through vector-boson fusion (VBF). Effects due to the interference between the signal, the SM Higgs boson and the WW continuum background are also considered.

2 Data and simulated samples

The analysis uses data from pp collisions produced in the LHC at a centre-of-mass energy $\sqrt{s} =$ 13 TeV with an integrated luminosity of 35.9 fb^{-1} [5].

The data collected by the CMS experiment during the 2016 B-H 25ns runs has been reprocessed in 2017 (Re-miniAOD data) using the official CMS software CMSSW, release CMSSW 8_0_X. The luminosity sections used are those certified in the official Golden JSON file: Cert_271036-284044_13TeV_23Sep2016ReReco_Collisions16_JSON.txt.

Official samples of simulated events (MC), produced in the RunIISummer16 simulation cam-

paign, are used in order to study the properties of the spin-0 resonance signal events and of the background processes. These samples are also reconstructed with the official software CMSSW, release 8_0_X. The global tags [6] used are :

- 80X_dataRun2_2016SeptRepro_v7 for the B-G era data
- 80X_dataRun2_Prompt_v16 for the H era data
- 80X_mcRun2_asymptotic_2016_TrancheIV_v8 for the simulation

A number of Higgs boson signal samples, with masses in the range from 200 GeV to 3 TeV, have been generated at NLO using POWHEG [7–9] and the subsequent decay of the Higgs bosons processed using JHUGEN [10–12]. Samples for both the gluon-gluon fusion and the vector boson fusion production processes have been generated. The signal lineshape for each mass point corresponds to the one expected for a SM Higgs boson at that mass. Samples corresponding to a Higgs boson with mass of 125 GeV have also been simulated and are treated as backgrounds in this analysis,

In this analysis the W+Jets background is characterised using a set of exclusive W+jets samples, binned in leptonic W Pt, produced at NLO with the MADGRAPH5_AMC@NLO event generator [13]. A scale factor is applied to each exclusive sample to match the normalization of the inclusive W+Jets sample rescaled to the NNLO cross section. Corrections to the W Pt spectrum coming from NLO electroweak contributions are applied on simulated W+Jets samples on a per-event basis depending on the generated W Pt [14].

$t\bar{t}$ events are characterised with fully leptonic and semileptonic samples produced with MADGRAPH5_AMC@NLO. Single top s-channel production is simulated with MADGRAPH5_AMC@NLO while single top t-channel and tW-channel production are simulated with POWHEG. For simplicity these backgrounds are collectively refered to as the top background in this analysis. The Pt spectra of top quarks in data is softer than that predicted by various MC simulations, as such the TOP-PAG periodically derives re-weighting functions based on the latest results and available theory predictions [15]. The application of the top-Pt reweighting has been found to improve the modelling and so is applied in this analysis.

Diboson events from standard model processes are simulated with WW, WZ and ZZ samples simulated using MADGRAPH5_AMC@NLO.

The background from DY Z+Jets is modelled using an inclusive sample produced with MADGRAPH5_AMC@NLO Corrections to the Z Pt spectrum coming from NLO electroweak contributions are applied on simulated Z+Jets samples on a per-event basis depending on the generated Z Pt [14].

QCD multi-jet production has been generated with PYTHIA8 [16]. The QCD samples are enriched in events with muons or electrons with dedicated filters. Samples are binned according to the momentum transfer in the matrix element.

All data and MC samples used in this analysis are displayed in appendix A, together with the background and signal MC cross sections [17].

In order to remove problematic or noise-dominated events the Moriond17 primary vertex, beam halo, HBHE noise, HBHE iso noise, ECAL TP, Bad PF muon, Bad charged hadron and ee badSC noise filters recommended for data by the JETMET POG [18] have been applied.

3 Event reconstruction

For this analysis the signature of a WW resonance signal event is a single isolated lepton and $\not\!\!E_T$ from a leptonic W decay and a dijet or a single merged jet from a hadronic W decay. The identification of $W \to \ell\nu$ and $W \to q\bar{q}$ is therefore a crucial step of the analysis. Events with leptons (electrons or muons), $\not\!\!E_T$ and jets are selected imposing quality criteria and kinematic constraints on the physics objects to ensure high efficiency in their reconstruction and identification, and good momentum and mass resolutions. The invariant mass of the $\ell\nu q\bar{q}$ system is consistent with the mass of a hypothetical resonance and is used as the main observable to discriminate signal events from background events.

3.1 Triggers

Data events used in the analysis belong to the SingleMuon and SingleElectron primary datasets, and are selected with the un-prescaled triggers :

- HLT_IsoMu24 || HLT_IsoTkMu24
- HLT_Ele25_eta2p1_WPTight

The muon trigger is seeded by the L1Mu22 bit while the electron trigger is seeded by an "OR" of the L1SingleEGX, L1SingleIsoEGXer and L1SingleIsoEGX bits. The event selection requirements of the analysis are consistent with those of the trigger. The trigger efficiency data–to–MC scale factors measured using tag-and-probe for muons [19, 20] and electrons [21] have been applied to the simulation.

3.2 Jets

Hadronic jets are reconstructed with the anti-k_T clustering algorithm [22] using two different radius parameters, $R = 0.8$ (AK8) and $R = 0.4$ (AK4). The AK8 algorithm is adopted for reconstructing the hadronic W decay in a single merged jet when the decay products are highly collimated, while the AK4 algorithm is used to reconstruct the hadronic W decay when the decay products are resolved. Both AK8 and AK4 jets are reconstructed from the list of particle candidates reconstructed with the particle flow (PF) algorithm [23]. Charged Particle Flow constituents not associated to the primary vertex are not used in the jet clustering procedure. The latest jet energy corrections [24] are applied to the jets. All jets are required to lie in the tracker acceptance $|\eta(j)| < 2.4$ and must pass the loose jet identification requirements defined by the JETMET POG for Run2 analyses [25]. To avoid double counting of the same object reconstructed in different collections, jets are required to be separated from the selected isolated leptons by $\Delta R(\ell j) > 0.8$ for AK8 jets by $\Delta R(\ell j) > 0.4$ for AK4 jets. For the selection of W candidates we require 2 AK4 jets both with $p_T(j) > 30$ GeV/c or an AK8 jet with $p_T(j) > 200$ GeV/c. The final hadronic W candidate requirements for the merged and resolved analyses will be discussed in detail in the following sections.

To reduce the top background we veto events that contain an additional jet with $p_T(j) > 20$ GeV/c which is b-tagged with the pfCombinedInclusiveSecondaryVertexV2BJetTags (CSV) algorithm [26] loose working point. We apply data–to–MC scale factors to correct the b-tagging efficiency and mistag probability in MC following the official Btag POG recommendations [27]. The jets are reweighted using the MC efficiencies and the scale factors, from the jet weights an overall event weight is derived following method 1a given in [28]. The CSV file used is :

BtagRecommendation80XReReco/subjet_CSVv2_Moriond17_B_H.csv

For AK8 jets the softdrop algorithm [29] is implemented to suppress the effects of pileup and

underlying event radiation. This technique declusters the jet and recursively removes soft wide-angle radiation, pushing QCD jets towards lower values while maintaining the jet mass for V-jets around the boson-mass. Each AK8 jet has two associated subjets corresponding to the protojets obtained by undoing the last iteration of the soft drop jet declustering procedure.

This analysis also makes use of the n-subjettiness variable [30] to quantify the compatibility of AK8 jets with a substructure hypothesis of n-subjets It is based on the distribution of the jet constituents with respect to the subjet axes. The ratio of 2-subjettiness and 1-subjettiness, τ_2/τ_1, allows the dipole structure of hadronic W decays to be distinguished from the monopole structure of QCD jets.

In order to mitigate the effect of pileup on the jet substructure observables used in this analysis (softdrop mass and n-subjettiness) the pileup per particle identification (PUPPI) [31] method is implemented. It has been observed that the PUPPI softdrop mass has a residual dependence on the jet Pt [32], this is accounted for in this analysis by applying dedicated Pt dependant mass corrections evaluated centrally by JMAR [33].

3.3 Missing transverse energy

This analysis uses type-1 corrected PFMET, where PFMET is defined as the magnitude of the negative vector sum of the transverse energy of all PF particles. Type-1 corrections involve propagation the jet energy corrections to the PFMET, which yields better performance. The $\not{E}_{\mathrm{T}}$ requirements are different for the merged and resolved analyses and will be discussed in the following sections.

3.4 Leptons

Electrons are reconstructed from energy deposits in the ECAL matched to tracks reconstructed in the silicon tracker. The electron trajectories are reconstructed using a dedicated modeling of the electron energy loss and fitted with a Gaussian sum filter. Electrons used in this analysis are required to pass the Particle Flow criteria, and to fall in the ECAL pseudorapidity fiducial range. To satisfy the electron trigger requirements the electrons must have $|\eta(\ell)| < 2.1$. The electron identification used in this analysis is based on the tight cut-based ID defined by the EGamma POG for Run 2 [21], Table 1 summarises the cuts used for this working point. An isolation cut is already applied within the cut-based ID and so no additional electron isolation requirements are made. In the isolation definition the effect of PU is corrected for by subtracting the median energy density in the event multiplied by the electron effective area. Muons are reconstructed by both the GlobalMuon and the Particle Flow muon reconstruction algorithms and are required to lie in the acceptance region $|\eta(\ell)| < 2.4$. The identification and isolation cuts used are the tight Muon ID and tight Particle Flow based isolation recommended by the Muon POG for Run 2 [34], Table 2 summarises the cuts used. The tracking, lepton reconstruction, identification and isolation efficiencies of the MC samples have been corrected using the scale factors (SF) provided by the Muon and EGamma POGs for 80X reconstruction [20, 35].

Table 1: Tight cut-based electron ID requirements.

cuts	barrel	endcap
full5x5_sigmaIetaIeta <	0.00998	0.0292
abs(dEtaInSeed) <	0.00308	0.00605
abs(dPhiIn) <	0.0816	0.0394
abs(dPhiIn) <	0.0816	0.0394
H/E <	0.0414	0.0641
Rel. comb. PF iso with EA corr <	0.0588	0.0571
abs(1/E-1/p) <	0.0129	0.0129
expected missing inner hits <=	1	1
pass conversion veto	yes	yes

Table 2: Tight muon ID and isolation requirements.

cuts	
isGlobalMuon	yes
isPFMuon	yes
Track $\chi^2/ndof$ <	10
Valid Muon hits >	0
Matched Stations >	1
d_{xy} <	0.2
d_z <	0.5
Valid Pixel hits >	0
Tracker Layers with hits >	5
PFIso <	0.15

Events with exactly one lepton satisfying the above criteria are selected in this analysis; those events containing additional loose electrons ($p_T(\ell) > 15$ GeV) or muons ($p_T(\ell) > 10$ GeV) are rejected. The W $\to e\nu$ and W $\to \mu\nu$ candidates are constructed by combining the $\not{E}_{\mathrm{T}}$ with a lepton which has $p_T(\ell) > 30$ GeV/c and $|\eta(\ell)|$ <2.4 (2.1) for muons (electrons). Assuming the W boson was on-shell, a quadratic equation for the z component of the neutrino momentum can be derived :

$$p_{z,\nu}^{\pm} = \frac{\mu p_{z,\ell}}{p_{T,\ell}^2} \pm \sqrt{\frac{\mu^2 p_{z,\ell}^2}{p_{T,\ell}^4} - \frac{E_\ell^2 p_{T,\nu}^2 - \mu^2}{p_{T,\ell}^2}} \tag{1}$$

where p_ℓ and p_ν are the four momenta of the charged lepton and the neutrino, respectively, and $\mu = \frac{M_W^2}{2} + p_{T,\ell} p_{T,\nu} cos\Delta\Phi$. The equation can give two solutions, in this case we select the one with the smaller magnitude. In case of no real solution the real part of the complex solution is used.

The baseline lepton and jet requirements for the reconstruction of the W candidates are summarized in Table 3.

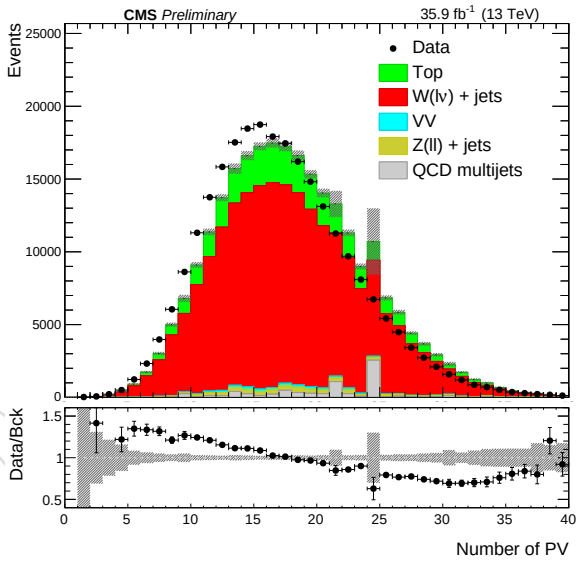

Figure 1: The number of primary vertices for events with a leptonic W candidate and 2 AK4 jets satisfying the resolved selection in data and simulation. The W+jets and top background distributions (red and green histograms) are normalized to data.

3.5 Primary Vertex

At least one good offline primary vertex (with longitudinal coordinate $|z| < 24$ mm, transverse position $\rho < 2$ mm, vertex fit variable $\chi^2 > 0$ and number of degrees of freedom larger than 4) is required. The MC description of the number of primary vertices (pileup) is known not to reproduce that of the data, therefore the simulated events are reweighted in order to improve the description. The standard CMS pileup reweighting technique [36] is implemented assuming a total inelastic cross section of 69.2 mb. The pileup distribution in data and MC after the pileup reweighting is applied is shown in Figure 1 for events with a leptonic W candidate and 2 AK4 jets satisfying the resolved selection.

Table 3: WW baseline lepton and jet requirements.

observable	selection criterion
lepton ID	see section 3
jet ID	see section 3
$p_T(\ell)$	> 30 GeV/c
$\lvert\eta(\ell)\rvert$	(e) < 2.1, (μ) < 2.4
$\lvert\eta(j)\rvert$	< 2.4
$p_T(j)$ (AK4)	> 30 GeV/c
$p_T(j)$ (AK8)	> 200 GeV/c
$\Delta R(\ell j)$ (AK8)	> 0.8
$\Delta R(\ell j)$ (AK4)	> 0.4

4 Boosted selection

For the boosted merged jet analysis an AK8 jet with corrected PUPPI softdrop mass $m_{\mathrm{J}} > 40$ GeV/c^2 is required. To suppress the background from QCD multijet events we require $\not{E}_{\mathrm{T}} > 40$ GeV/c^2. The $\mathrm{W} \to \ell\nu$ and $\mathrm{W} \to \mathrm{q\bar{q}}$ decay candidates satisfying the boosted analysis requirements are combined into $\ell\nu J$ resonance candidates. In order to suppress the dominant W+jets background, $\ell\nu J$ candidates are selected in the PUPPI softdrop mass region $65 \leq m_{\mathrm{J}} \leq 105$ GeV/c^2, called the signal region (SR). Outside of this signal region, candidates with PUPPI softdrop mass within $40 \leq m_{\mathrm{J}} \leq 250$ GeV/c^2 are used for background determination. This is the main control region of the analysis and is referred to as the sideband (SB). A top-enriched control region (Top CR) is also defined by reversing the b-veto and requiring at least one AK4 b-tagged jet in the event.

For heavy resonance decays the P_T of the W candidates are expected to be roughly half of the resonance mass. Therefore we require both the leptonic and hadronic W candidates to satisfy the requirement $Pt_{\mathrm{W}}/m_{\ell\nu J} > 0.4$. Figure 2 shows the $min(Pt_{\mathrm{W}_{Had}}, Pt_{\mathrm{W}_{Lep}})/m_{\ell\nu J}$ distributions for events in the SR and SB. For all the Figures in this section the normalisation of the W+jets and top backgrounds are taken from a fit to the data in the SR, SB and Top CR, the gray bands show the MC statistical uncertainties.

4.1 Boosted W-tagging

For the final selection a cut on the PUPPI n-subjettiness ratio τ_2/τ_1 is used to identify boosted hadronic boson candidates (boosted W-tagging). The PUPPI τ_2/τ_1 distribution of the boosted W candidates in the SR and SB are displayed in Figure 2. An optimal high purity cut of PUPPI $\tau_2/\tau_1 < 0.4$ [33] has been determined by previous analyses and is implemented here. The JetMET POG has measured a PUPPI τ_2/τ_1 high purity cut efficiency data-to-MC scale factor of 1.0 ± 0.06, a corresponding mass scale factor of $1.0 \pm .0094$ and a corresponding mass resolution scale factor of 1.0 ± 0.2 [33]. The PUPPI τ_2/τ_1 scale factor has been applied to the signal and diboson background MC. After applying the cut the events generally contain a single boosted W candidate; in the small number of events with greater than one candidate we choose the one with highest transverse momentum.

The entire selection procedure described above is referred to as the "boosted selection". The boosted selection criteria are summarized in Table 4. The groomed jet mass of the hadronic W candidates in data and simulation for events passing the boosted selection in the SR, SB and Top CR are shown in Figure 3. Figure 4 shows the lepton P_T and $\not{E}_{\mathrm{T}}$ for events passing the boosted selection in the SR and SB. The leptonic and hadronic W P_T for the same events are shown in Figure 5. Figure 6 displays the $m_{\ell\nu J}$ for events passing the boosted selection in the SR, SB and Top CR. Good agreement is observed between data and simulation for all significant variables.

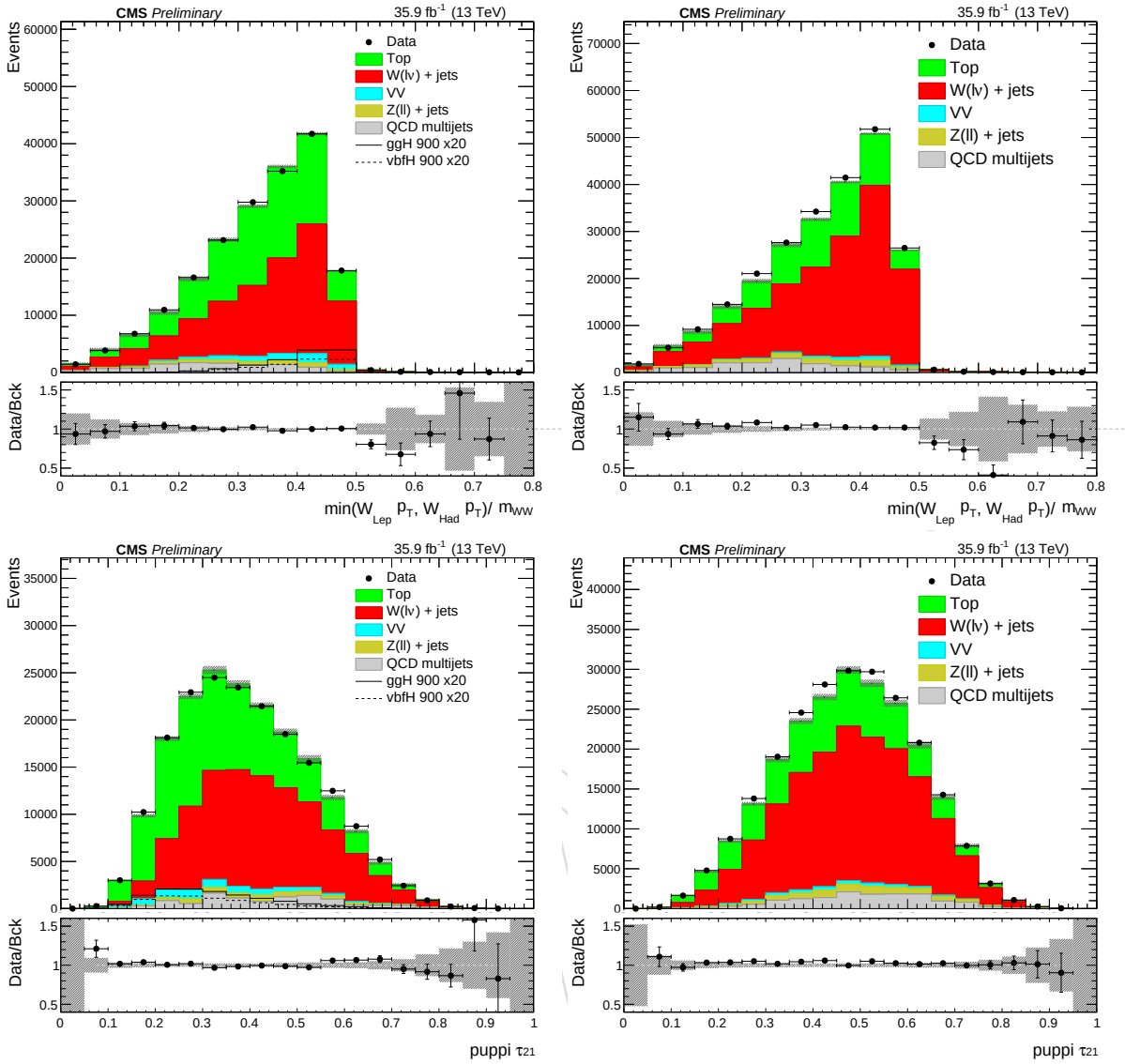

Figure 2: $min(Pt_{W_{Had}}, Pt_{W_{Lep}})/m_{\ell\nu J}$ (top) and PUPPI τ_2/τ_1 (bottom) in data and simulation for events with boosted W candidates in the SR (left) and SB (right). The W+jets and top background normalisations (red and green histograms) are taken from a fit to the data in the SR, SB and Top CR.

Table 4: Boosted selection.

observable	selection criterion
$\not{E}_T$	> 40 GeV/c^2
$p_T(J)$	> 200 GeV/c
$min(Pt_{W_{Had}}, Pt_{W_{Lep}})/m_{\ell\nu J}$	> 0.4
PUPPI τ_2/τ_1	< 0.4
m_J SR	[65,105] GeV/c^2
m_J SB	[40,65] $\cup$ [105,250] GeV/c^2

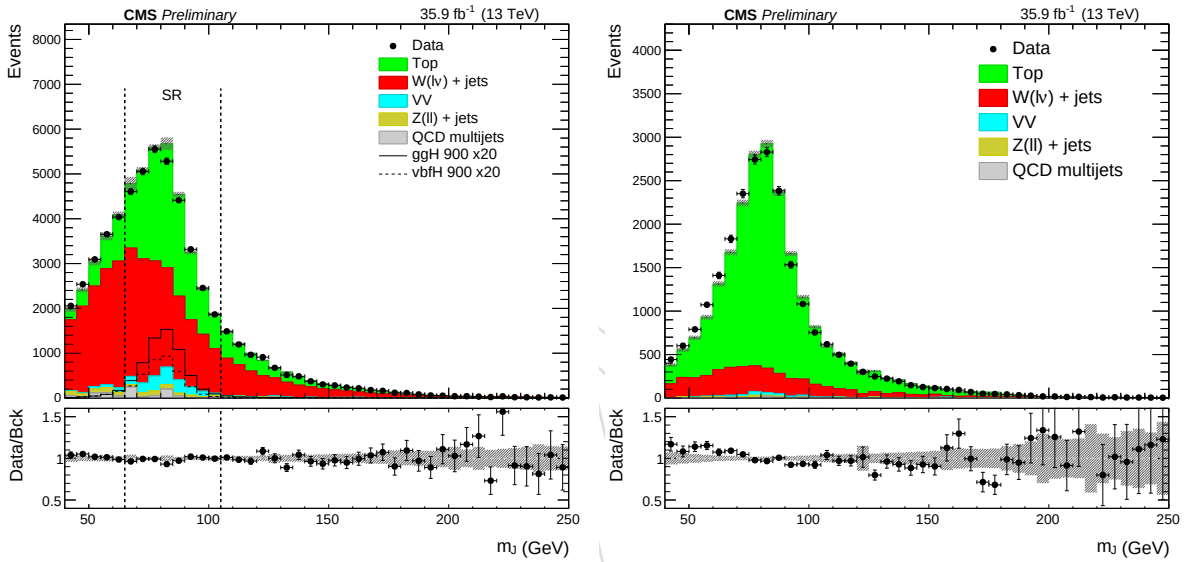

Figure 3: Hadronic W candidate PUPPI softdrop mass in data and simulation for events passing the boosted selection in the SR and SB (left), and the Top CR (right). The W+jets and top background normalisations (red and green histograms) are taken from a fit to the data in the SR, SB and Top CR.

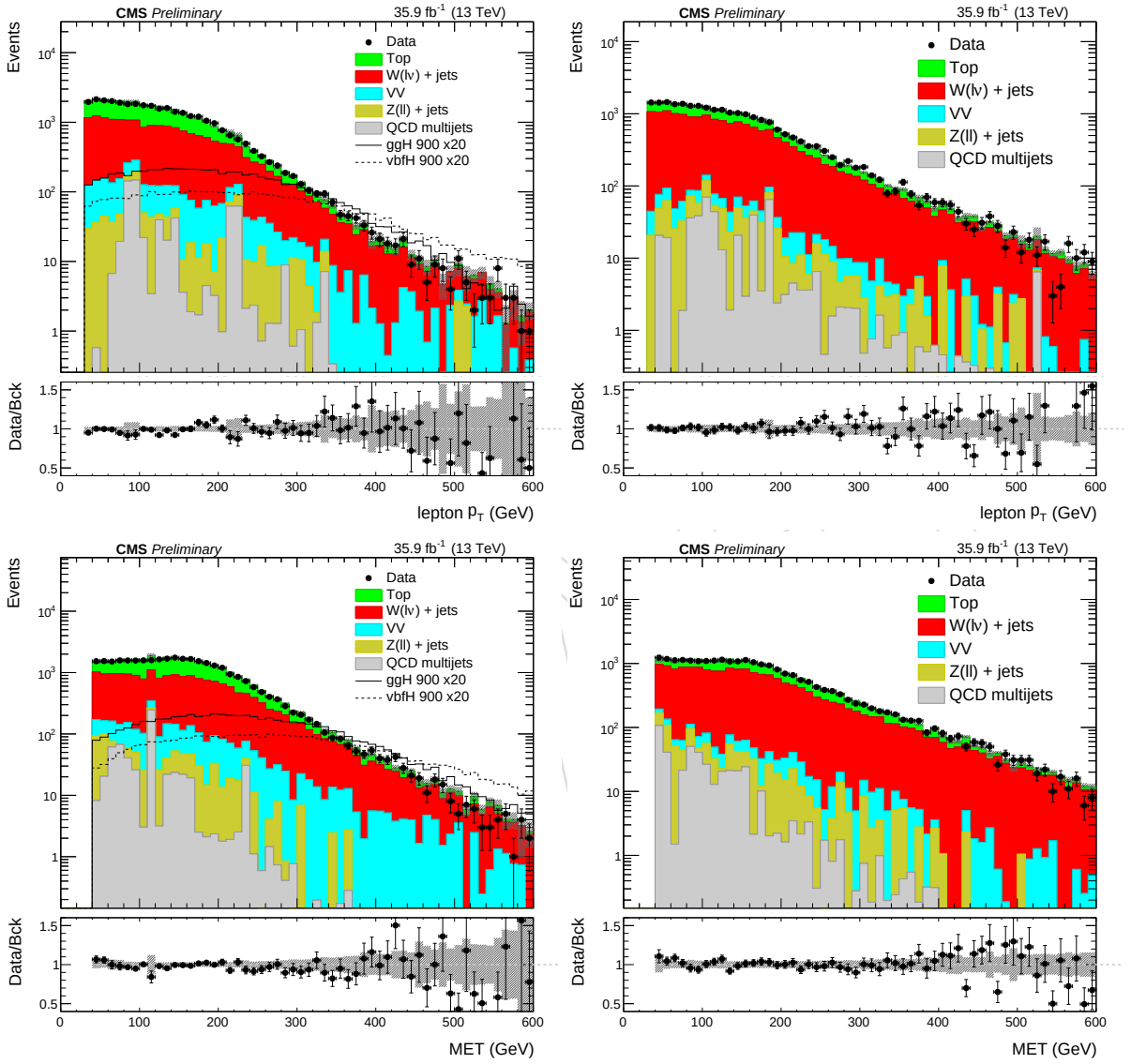

Figure 4: Lepton P_T (top) and $\not{E}_T$ (bottom) in data and simulation for events passing the boosted selection in the SR (left) and SB (right). The W+jets and top background normalisations (red and green histograms) are taken from a fit to the data in the SR, SB and Top CR.

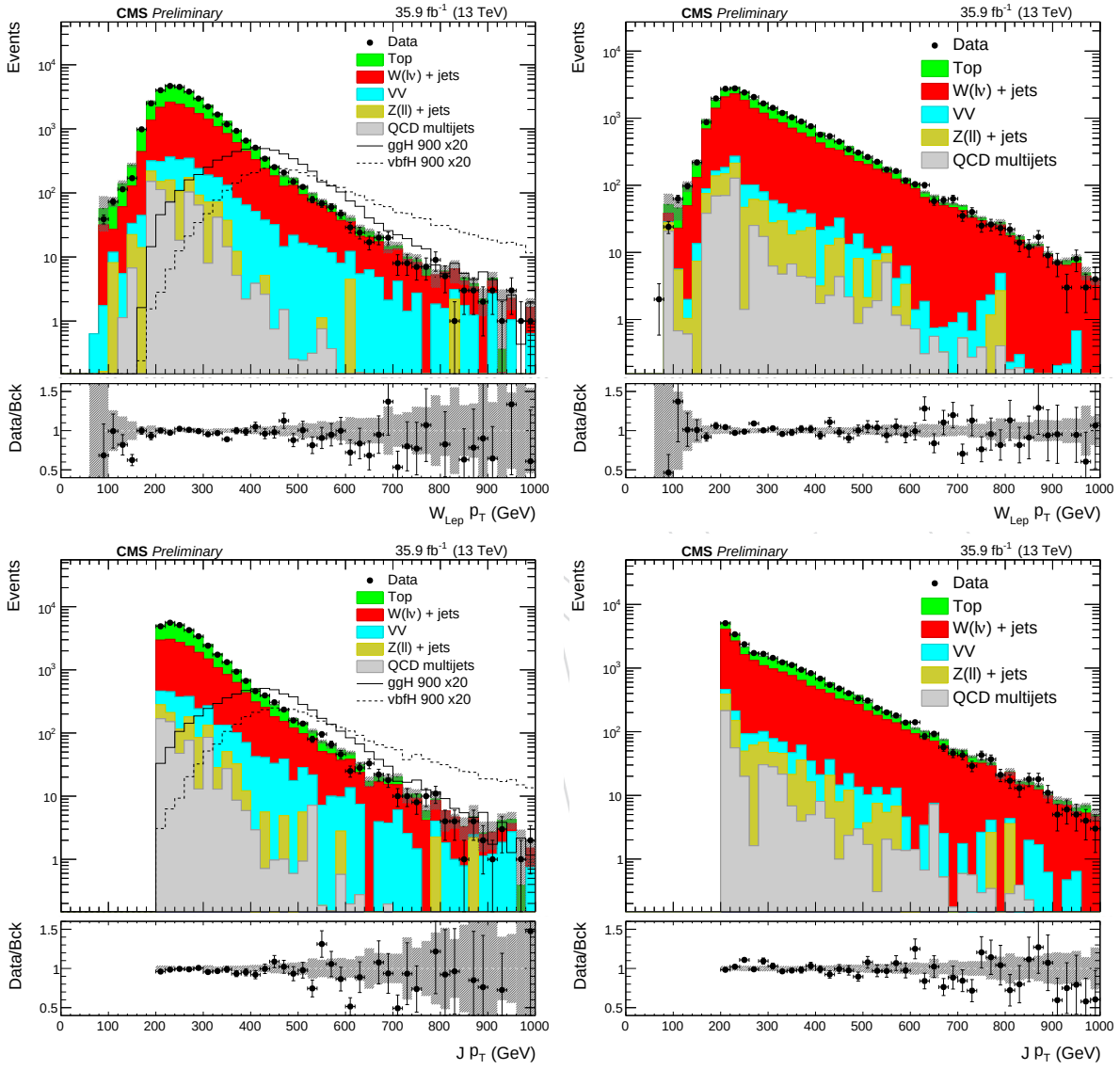

Figure 5: The leptonic W P_T (top) and hadronic W P_T (bottom) in data and simulation for events passing the boosted selection in the SR (left) and SB (right). The W+jets and top background normalisations (red and green histograms) are taken from a fit to the data in the SR, SB and Top CR.

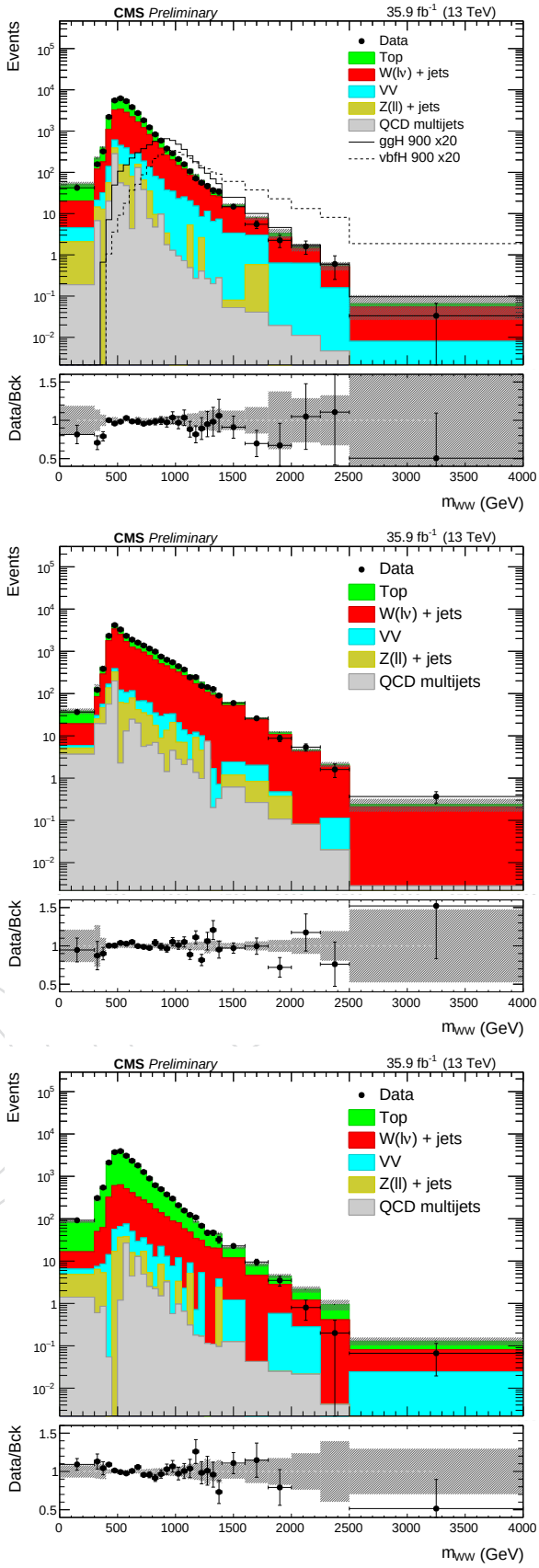

Figure 6: The $m_{\ell\nu J}$ in data and simulation for events passing the boosted selection in the SR (top), SB (middle) and Top CR (bottom). The W+jets and top background normalisations (red and green histograms) are taken from a fit to the data in the SR, SB and Top CR.

5 Resolved selection

For events which do not pass the boosted jet selection, it is attempted to reconstruct a resolved hadronic W decay using 2 AK4 jets. A kinematic fit is performed to the dijet system using the W mass constraint, in events with greater than 2 jets the dijet pair with the smallest χ^2 is chosen. To suppress the background from QCD multijet events we require $\not{E}_{\mathrm{T}} > 30$ GeV/c^2 and $\mathrm{W}_{m_T} > 50$ GeV/c^2, where W_{m_T} is the transverse mass of the leptonic W candidate. The $\mathrm{W} \to \ell\nu$ and $\mathrm{W} \to \mathrm{q\bar{q}}$ decay candidates satisfying these requirements are combined into $\ell\nu jj$ resonance candidates. To suppress background $\ell\nu jj$ candidates are selected in the dijet mass signal region $65 \leq m_{\mathrm{jj}} \leq 105$ GeV/c^2. Sideband and top control regions used for background estimation are defined as for the boosted selection.

As for the boosted category we require that the leptonic and hadronic W candidates satisfy the condition $Pt_{\mathrm{W}}/m_{\ell\nu jj} > 0.35$. Further reduction in the QCD multijet background is achieved by requiring $X_{m_T} > 60$ GeV/c^2, where X_{m_T} is the transverse mass of the $\ell\nu jj$ system. Figure 7 shows the $min(Pt_{\mathrm{W}_{Had}}, Pt_{\mathrm{W}_{Lep}})/m_{\ell\nu jj}$ and X_{m_T} distributions for events in the SR and SB. For all the Figures in this section the normalisation of the W+jets and top backgrounds are taken from a fit to the data in the SR, SB and Top CR, the gray bands show the MC statistical uncertainties.

The entire selection procedure described above is referred to as the "resolved selection" and is summarized in Table 5. The dijet mass of the hadronic W candidates in data and simulation for events passing the resolved selection in the SR, SB and Top CR are shown in Figure 8. Figure 9 shows the lepton P_T, $\not{E}_{\mathrm{T}}$ and the W_{m_T} for events passing the resolved selection in the SR and SB. The leptonic and hadronic W P_T for the same events are shown in Figure 10. Figure 11 displays the $m_{\ell\nu jj}$ for events passing the resolved selection in the SR, SB and Top CR. Good agreement is observed between data and simulation for all significant variables.

Table 5: resolved selection.

observable	selection criterion
$\not{E}_{\mathrm{T}}$	> 30 GeV/c^2
W_{m_T}	> 50 GeV/c^2
$min(Pt_{\mathrm{W}_{Had}}, Pt_{\mathrm{W}_{Lep}})/m_{\ell\nu jj}$	> 0.35
X_{m_T}	> 60 GeV/c^2
m_{jj} SR	[65,105] GeV/c^2
m_{jj} SB	[40,65] $\cup$ [105,250] GeV/c^2

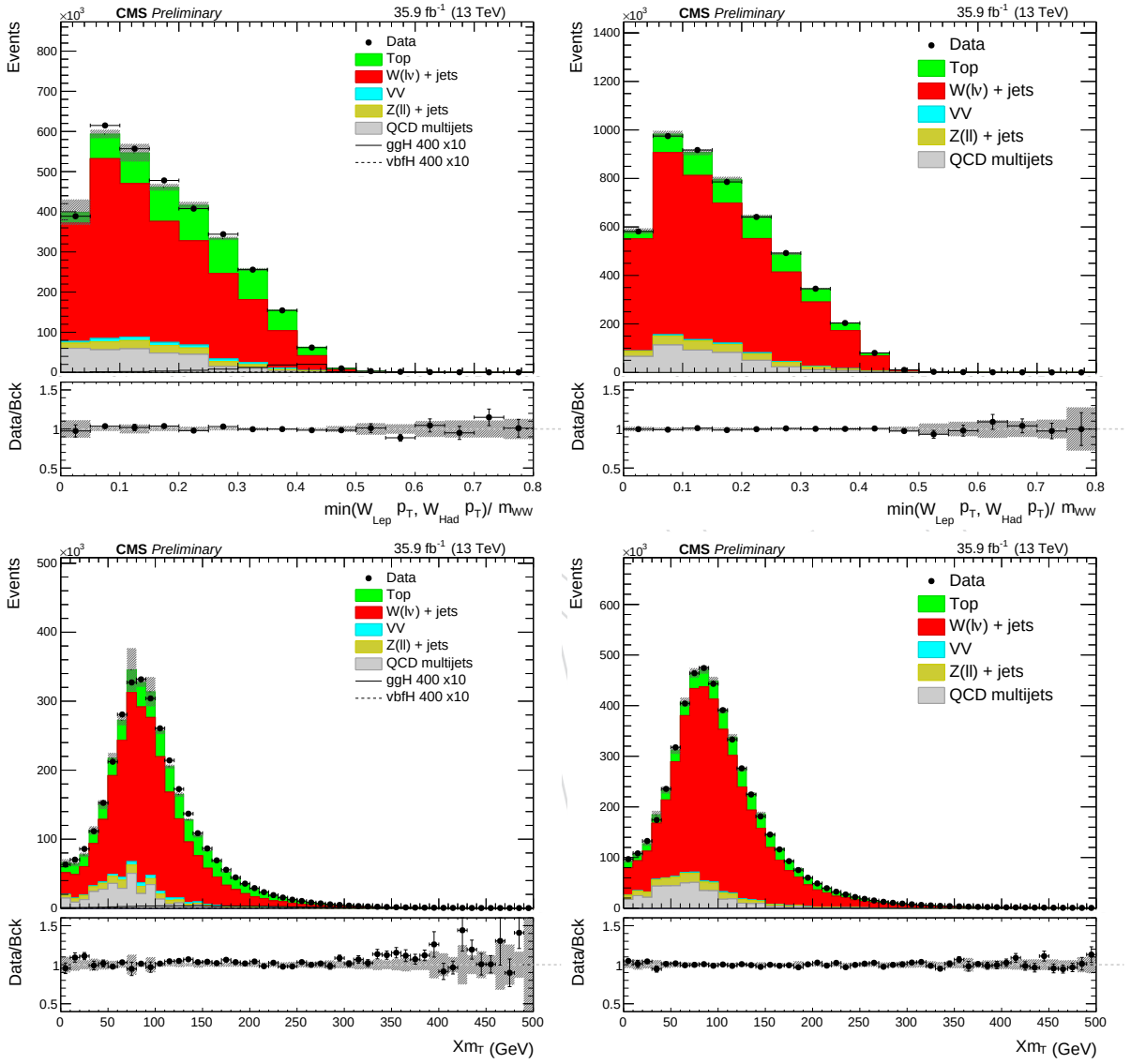

Figure 7: $min(Pt_{W_{Had}}, Pt_{W_{Lep}})/m_{\ell\nu jj}$ (top) and X_{m_T} (bottom) in data and simulation for events with resolved W candidates in the SR (left) and SB (right). The W+jets and top background normalisations (red and green histograms) are taken from a fit to the data in the SR, SB and Top CR.

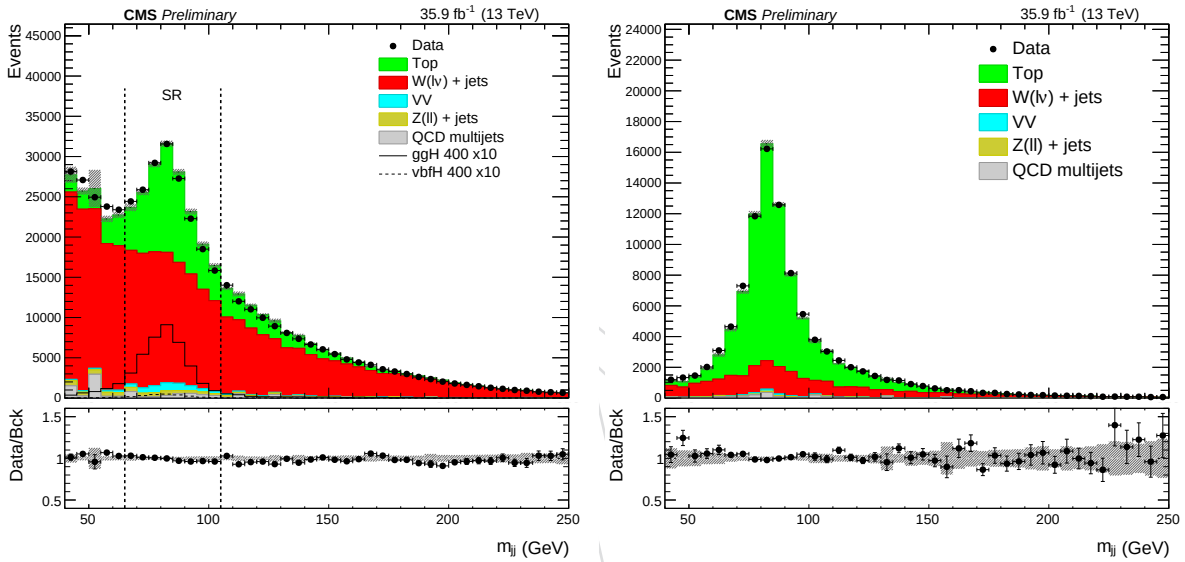

Figure 8: Hadronic W dijet mass in data and simulation for events passing the resolved selection in the SR and SB (left), and the Top CR (right). The W+jets and top background normalisations (red and green histograms) are taken from a fit to the data in the SR, SB and Top CR.

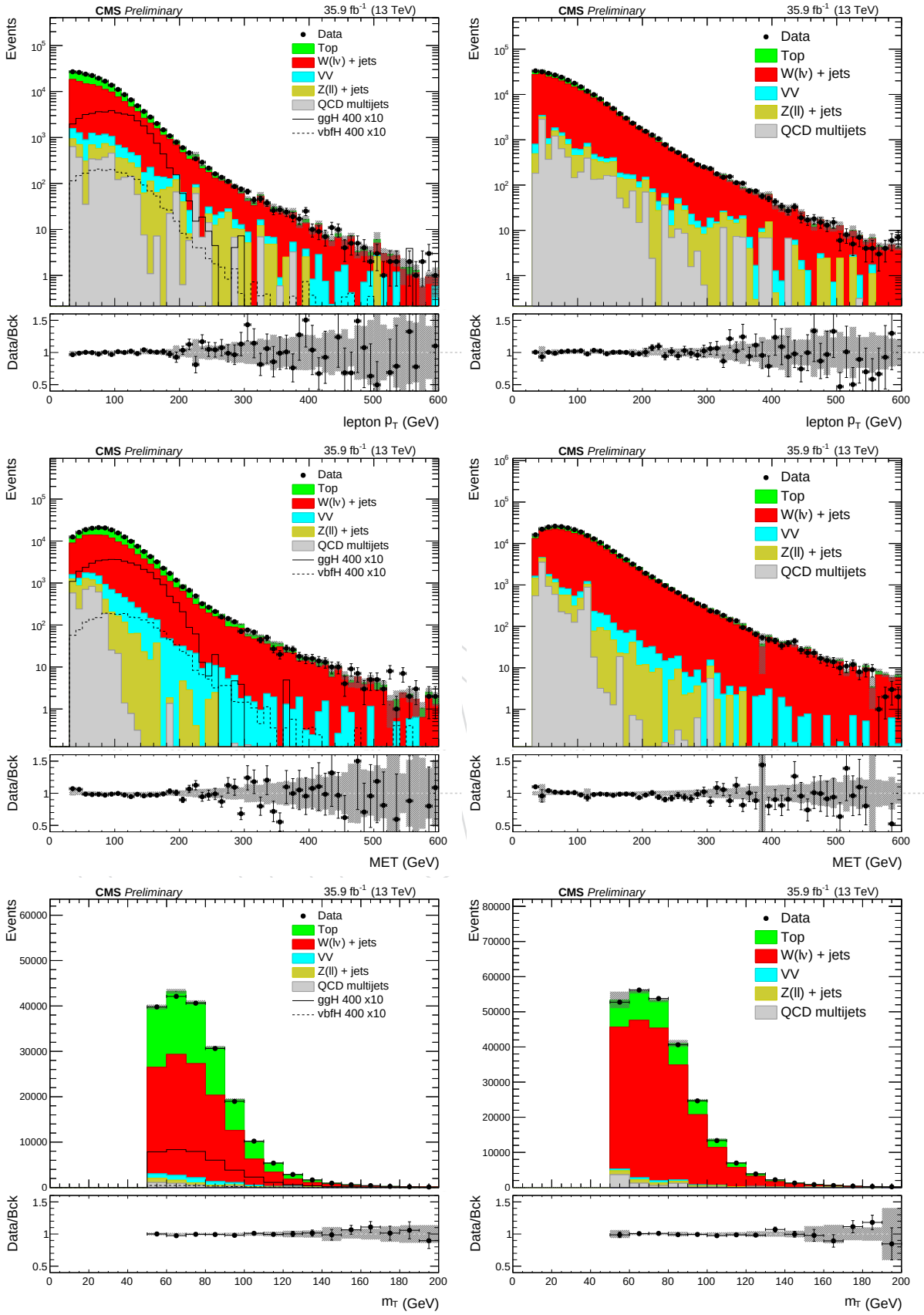

Figure 9: Lepton P_T (top), $\not{E}_T$ (middle) and W$_{m_T}$ (bottom) in data and simulation for events passing the resolved selection in the SR (left) and SB (right). The W+jets and top background normalisations (red and green histograms) are taken from a fit to the data in the SR, SB and Top CR.

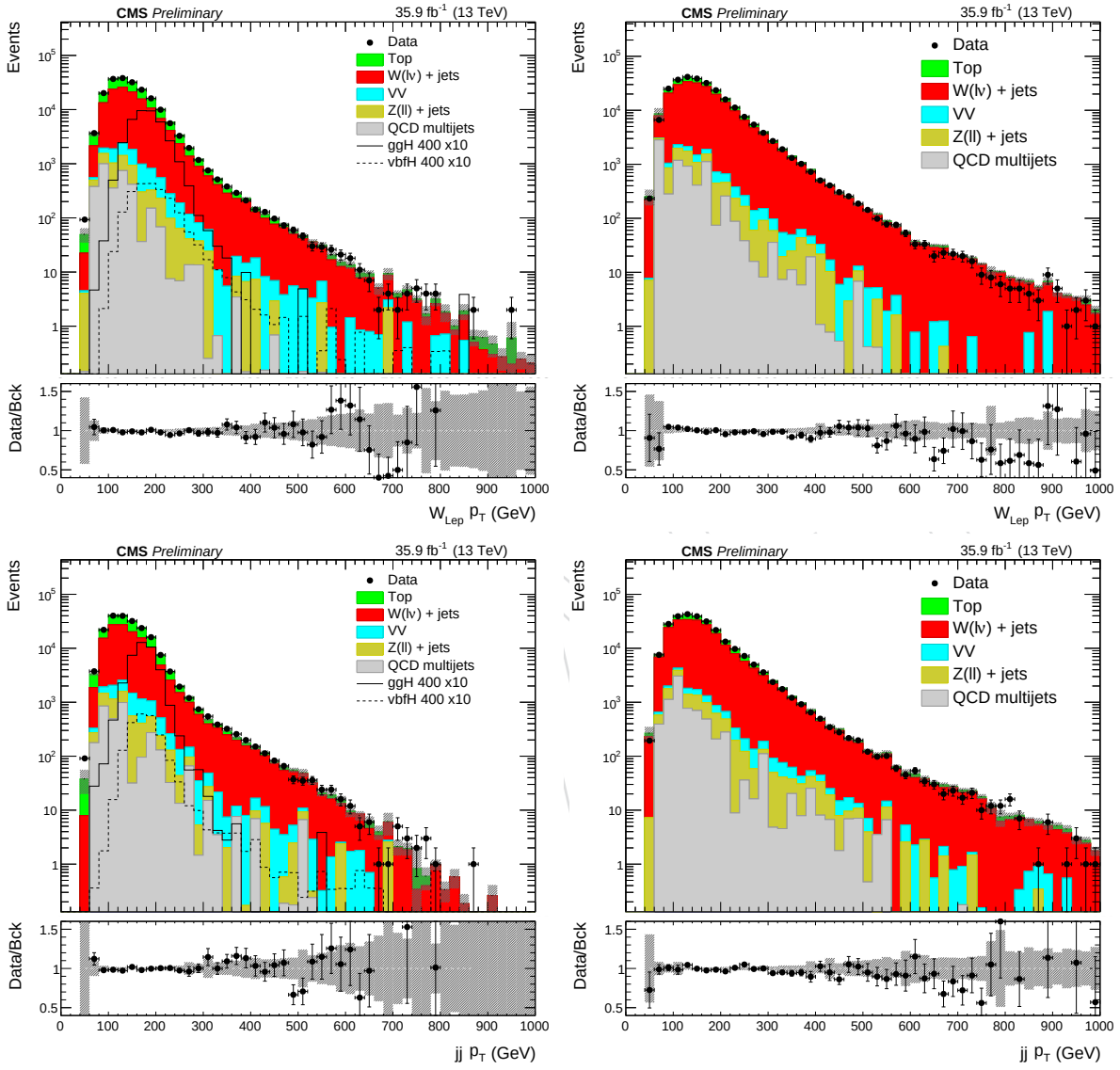

Figure 10: The leptonic W P_T (top) and hadronic W P_T (bottom) in data and simulation for events passing the resolved selection in the SR (left) and SB (right). The W+jets and top background normalisations (red and green histograms) are taken from a fit to the data in the SR, SB and Top CR.

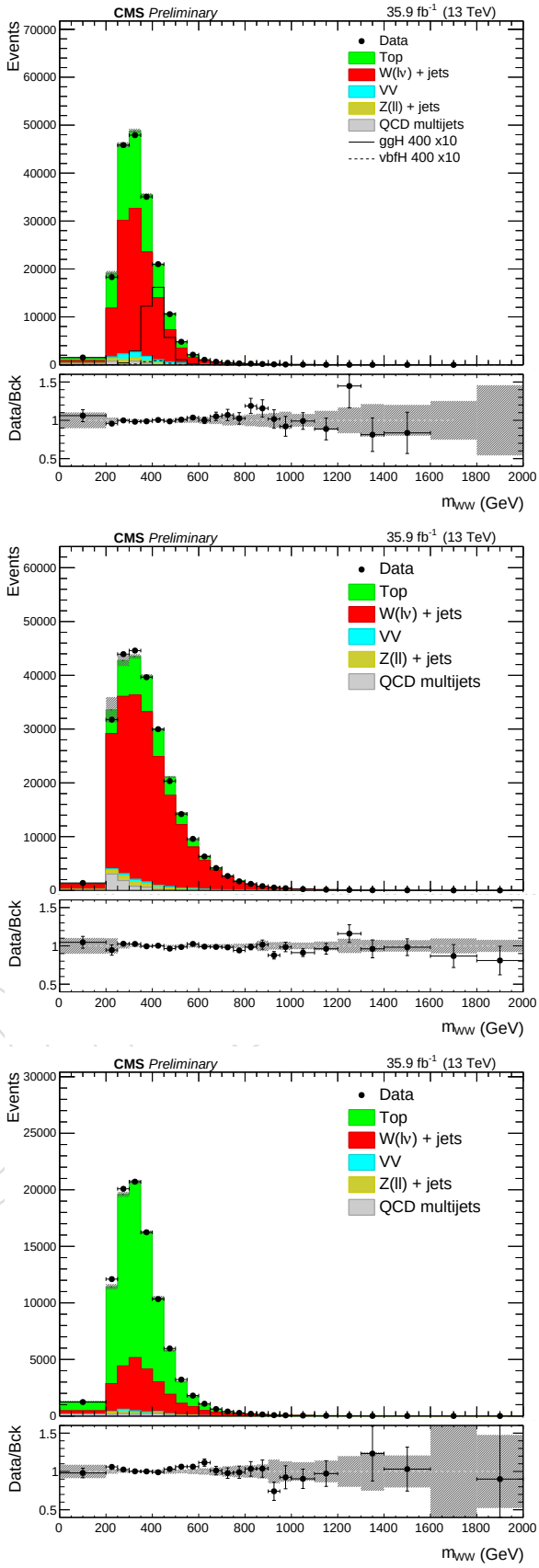

Figure 11: The $m_{\ell\nu jj}$ in data and simulation for events passing the resolved selection in the SR (top), SB (middle) and Top CR (bottom). The W+jets and top background normalisations (red and green histograms) are taken from a fit to the data in the SR, SB and Top CR.

6 Event categorisation

In this analysis the events are categorised by the flavour of the lepton (electron and muon) and type of hadronic W candidate (boosted and resolved). To increase the signal sensitivity further we also divide events into categories based on the tagging of VBF X production and gluon-gluon fusion X production.

6.1 VBF X category

For the identification of a VBF produced X candidate event we require 2 additional AK4 jets with $P_T > 30$ GeV/c^2 and η <4.7. If there are more then 2 additional jets then the pair with the highest invariant mass is chosen. An event enters the VBF category if the dijet pair has a separation in η, $\Delta\eta(jj)$, greater than 3.5 and an invariant mass, m_{jj}, greater than 500 GeV/c^2. The VBF $\Delta\eta(jj)$ and m_{jj} variables are shown for the boosted and resolved categories in Figure 13 and Figure 14 respectively.

6.2 gluon-gluon fusion X category

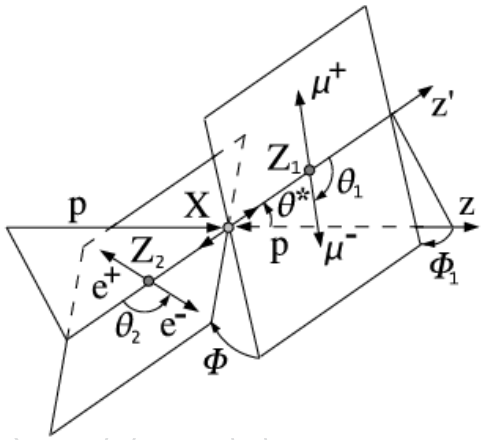

Figure 12: An illustration of the production and decay anlges for an $X \to VV$ decay with 4 final state particles.

Those events which are not tagged as VBF candidates are considered for the gluon-gluon fusion produced X category. The tagging of gluon-gluon fusion produced X candidates is achieved using a kinematic discriminant based on the angular distributions of the X candidate decays products. An illustration of the production and decay anlges for an $X \to VV$ decay with 4 final state particles is shown in Figure 12. The discriminant used in this analysis is implemented with MELA [10–12, 37] which uses JHUGEN and MCFM matrix elements to calculate probabilities, defined as the matrix elements squared, for an event to come from either signal or background respectively. The calculations are done at LO in QCD. For each event the reconstructed lepton, neutrino ($\not{E}_{\mathrm{T}}$ with P_z estimate from W mass constraint) and $q\bar{q}$ system (dijet or 2 subjets of merged jet W candidate) are used as inputs to MELA. A kinematic discriminant is then constructed as :

$$KD = (1 + \frac{c * P_{Bkg}}{P_{Sig}})^{-1} \tag{2}$$

where the constant c may be tuned to adjust the relative normalisations of probabilities. The discriminant is continuously distributed between 0 and 1, with signal being closer to 1 and

background closer to 0. In this analysis a cut at 0.5 is used to tag gluon-gluon fusion produced X candidates. The value of c used in the boosted category is then chosen to maximise the signal sensitivity for a 1500 GeV signal, likewise that for the resolved category is chosen to maximise the signal sensitivity for a 400 GeV signal. These signal mass values are chosen to represent the median of the signal masses under study for each analysis. The cosine of the production angle θ^* and the KD are shown for the boosted and resolved categories in Figure 15 and Figure 16 respectively.

Events failing the Mela KD requirement enter the untagged category resulting in three X-tagging categories in all (vbf$X$, gg$X$ and untagged).

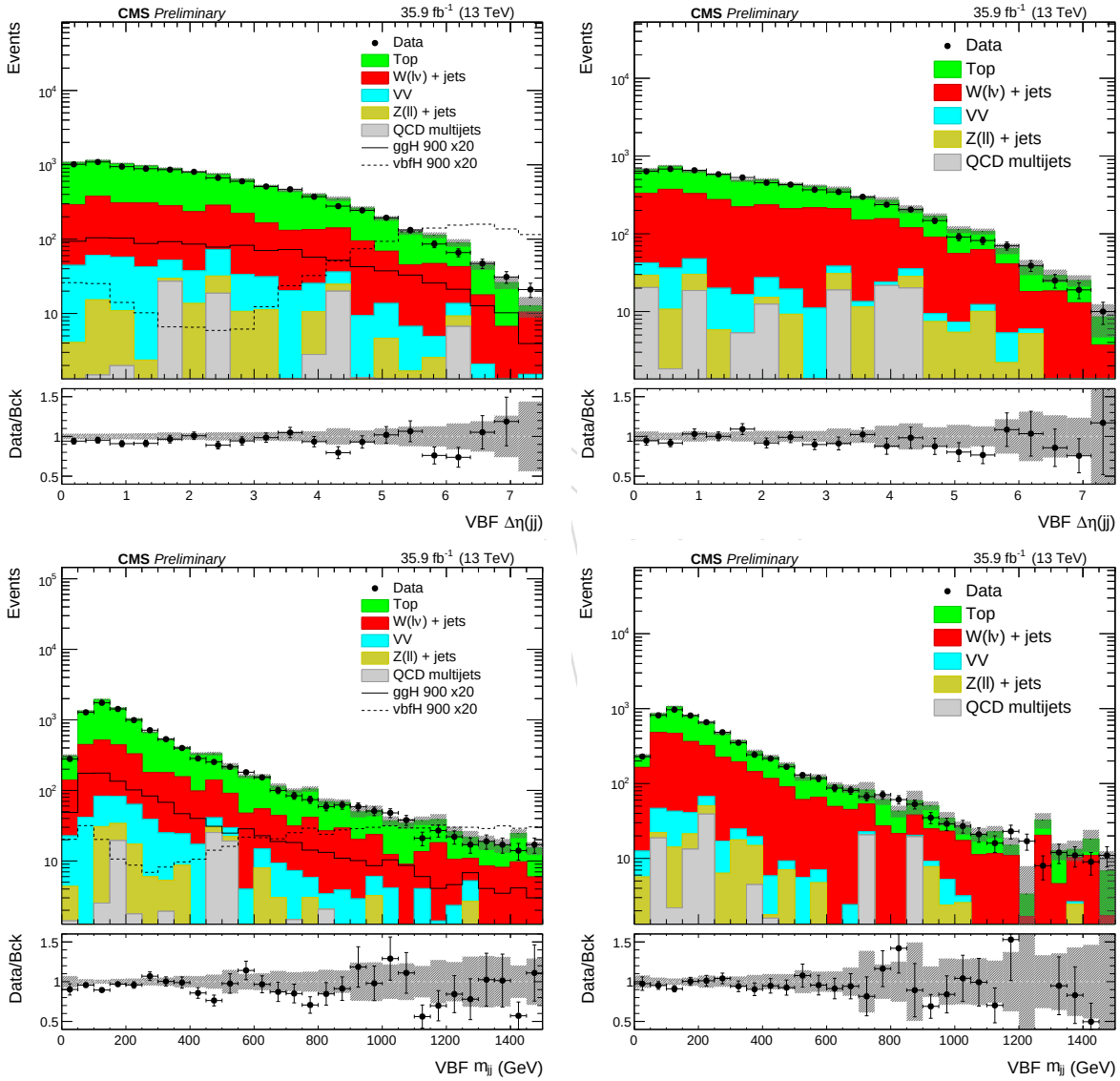

Figure 13: VBF $\Delta\eta(jj)$ (top) and m_{jj} (bottom) in data and simulation for events passing the boosted selection in the SR (left) and SB (right). The W+jets and top background normalisations (red and green histograms) are taken from a fit to the data in the SR, SB and Top CR.

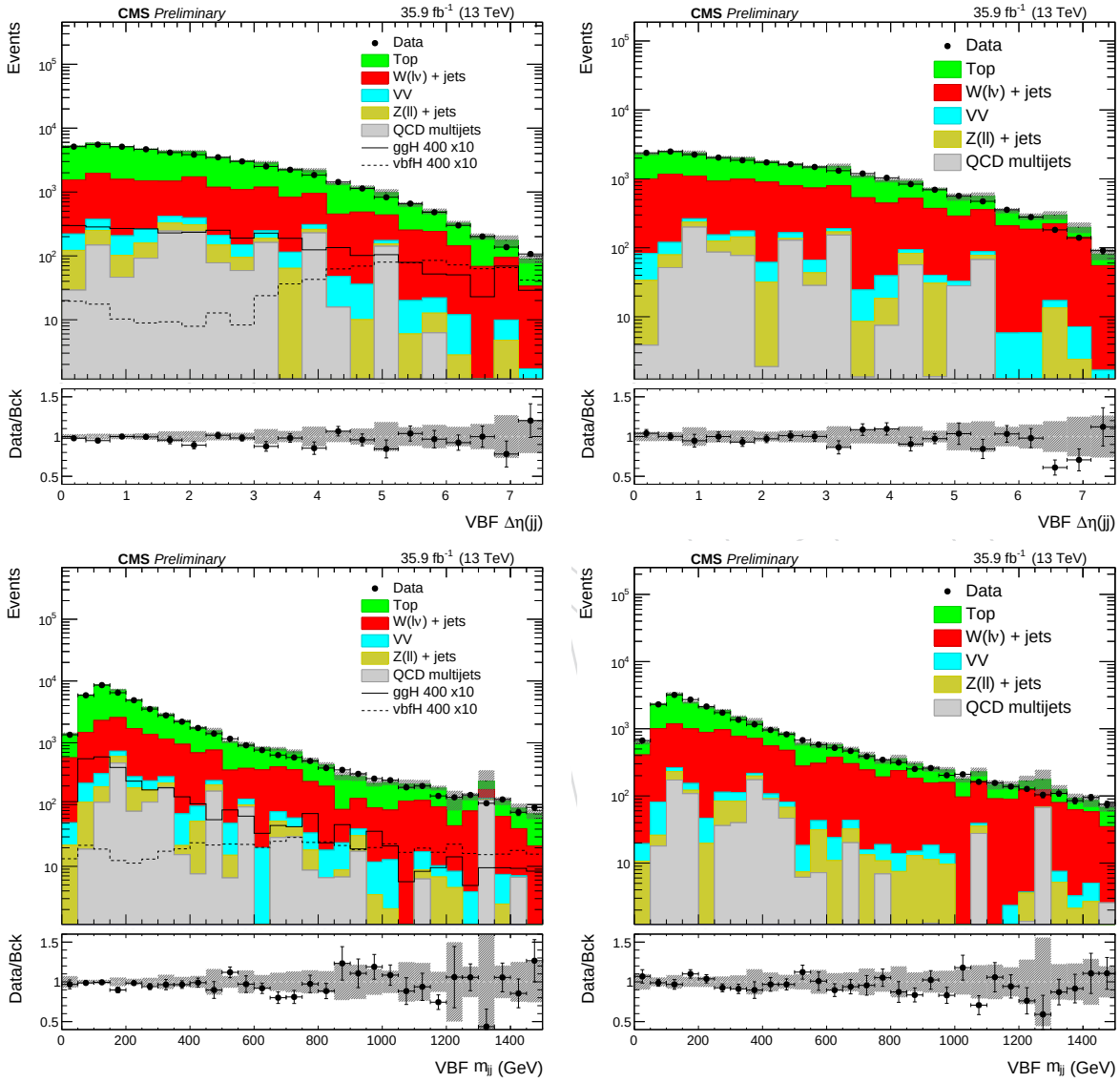

Figure 14: VBF $\Delta\eta(jj)$ (top) and m_{jj} (bottom) in data and simulation for events passing the resolved selection in the SR (left) and SB (right). The W+jets and top background normalisations (red and green histograms) are taken from a fit to the data in the SR, SB and Top CR.

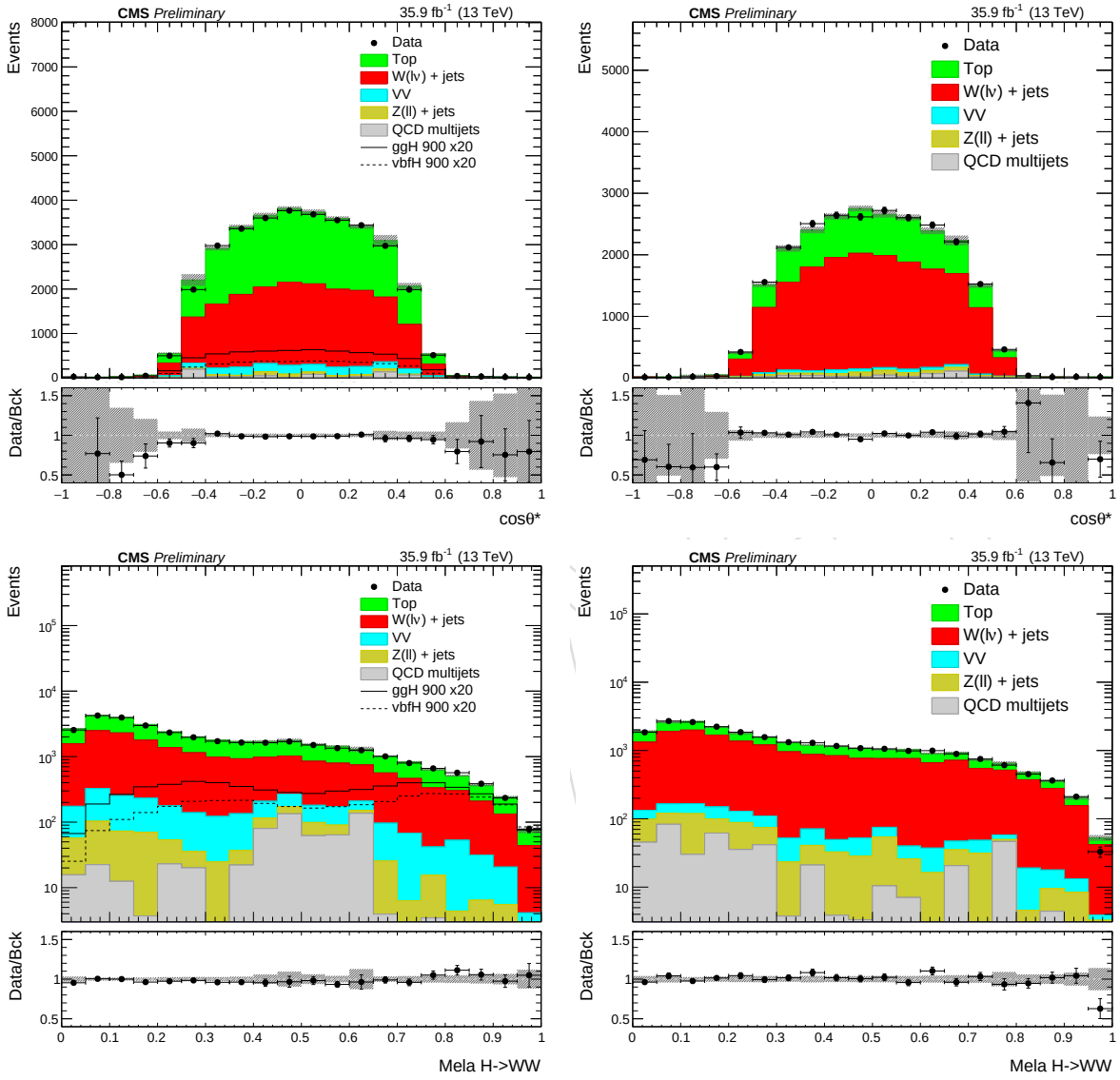

Figure 15: cosθ^* (top) and KD (bottom) in data and simulation for events passing the boosted selection in the SR (left) and SB (right). The W+jets and top background normalisations (red and green histograms) are taken from a fit to the data in the SR, SB and Top CR.

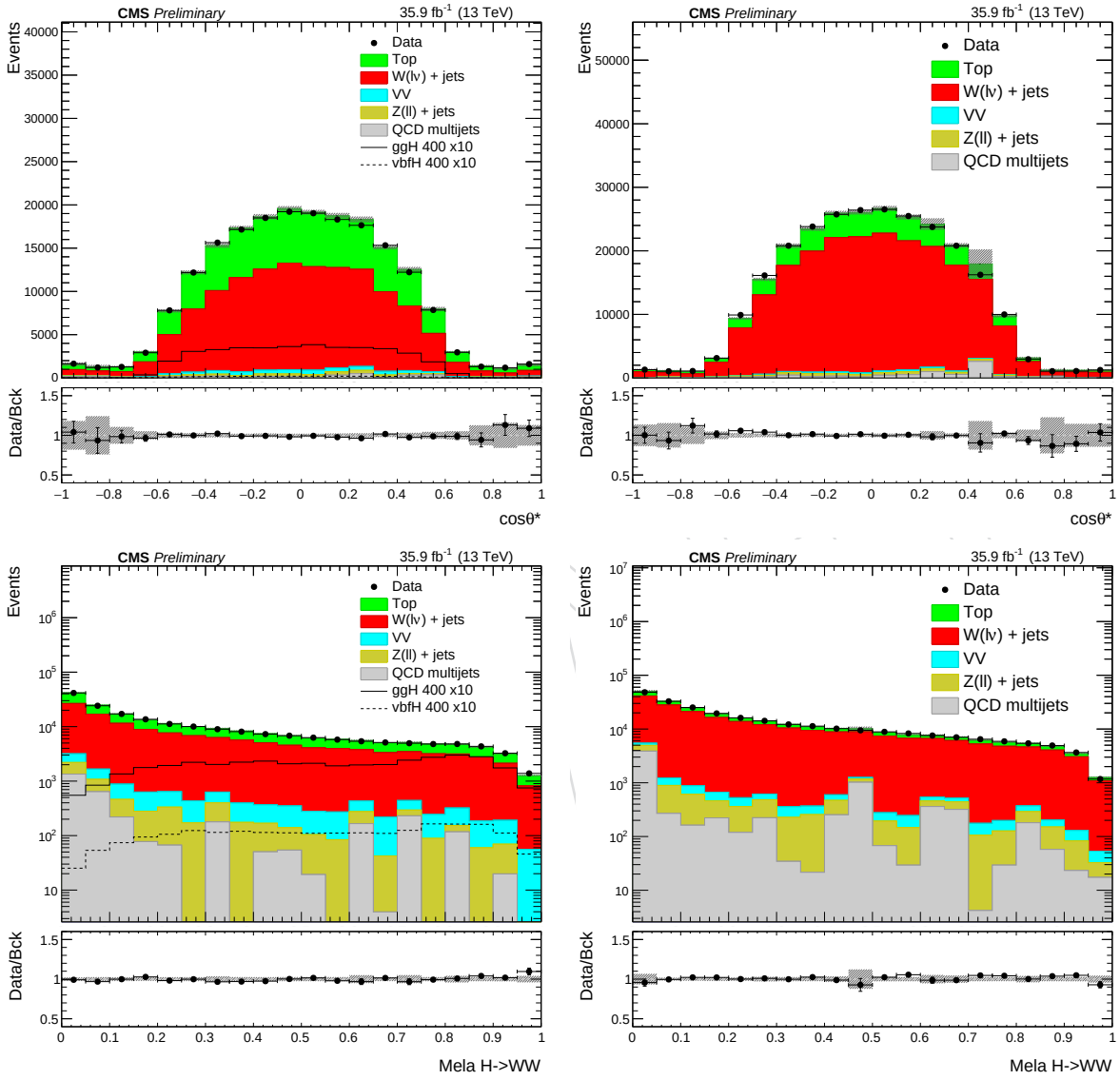

Figure 16: cosθ^* (top) and KD (bottom) in data and simulation for events passing the resolved selection in the SR (left) and SB (right). The W+jets and top background normalisations (red and green histograms) are taken from a fit to the data in the SR, SB and Top CR.

7 Interference

The high mass resonances considered in this analysis have a non negligible width and so it is important to take into account interference effects with the WW continuum and the H(125) off-shell tail. We use the MELA package to reweight the signal processes generated with POWHEG+JHUGen in order to obtain the interference terms between X, the WW continuum background and the H(125) off-shell tail. The generator level final decay products from the Higgs decay are used as input to the MELA $gg \to X$ weight calculations. Figure 17 displays the generator level Higgs mass for a number of $gg \to X$ signals, along with the interference contributions of the $gg \to$ WW continuum and the $gg \to H(125)$. To estimate the VBF $qq \to qqX$ interference it is also necessary to consider the incoming and outgoing quarks in the MELA weight calculation. For those signal events containing a hard radiation gluon we recast from NLO to LO topologies by merging the gluon with the closest quark. This approach insures that the input generator level particles are consistent with the matrix element calculation. Figure 18 displays the generator level Higgs mass for a number of $qq \to qqX$ signals, along with the interference contributions of the $qq \to qq$WW continuum and the $qq \to qqH(125)$. The continuum backgrounds, $gg \to$ WW and $qq \to qq$WW, are also estimated using the MELA reweighting scheme in this analysis.

It can be seen that the two sources of interference have opposite sign and partially cancel out. The size of the cancellation also depends on the signal mass. Generally the overall interference contribution is non negligible and so its effect is included in the final fit to the data. The appropiate physics model for the fit is

$$\text{Yield} = \mu S + \sqrt{\mu} I + B \quad (3)$$

where B is the background component, S is the signal component, I is the interference component and μ is the signal strength parameter. This is manipulated to avoid potenetial negative probability distribution functions due to the interference term giving the final model [38]

$$\text{Yield} = \mu|S + B + I| + (\mu - \sqrt{\mu})|S| + (1 - \sqrt{\mu})|B| \quad (4)$$

where the square brackets represent the input histograms to the fit.

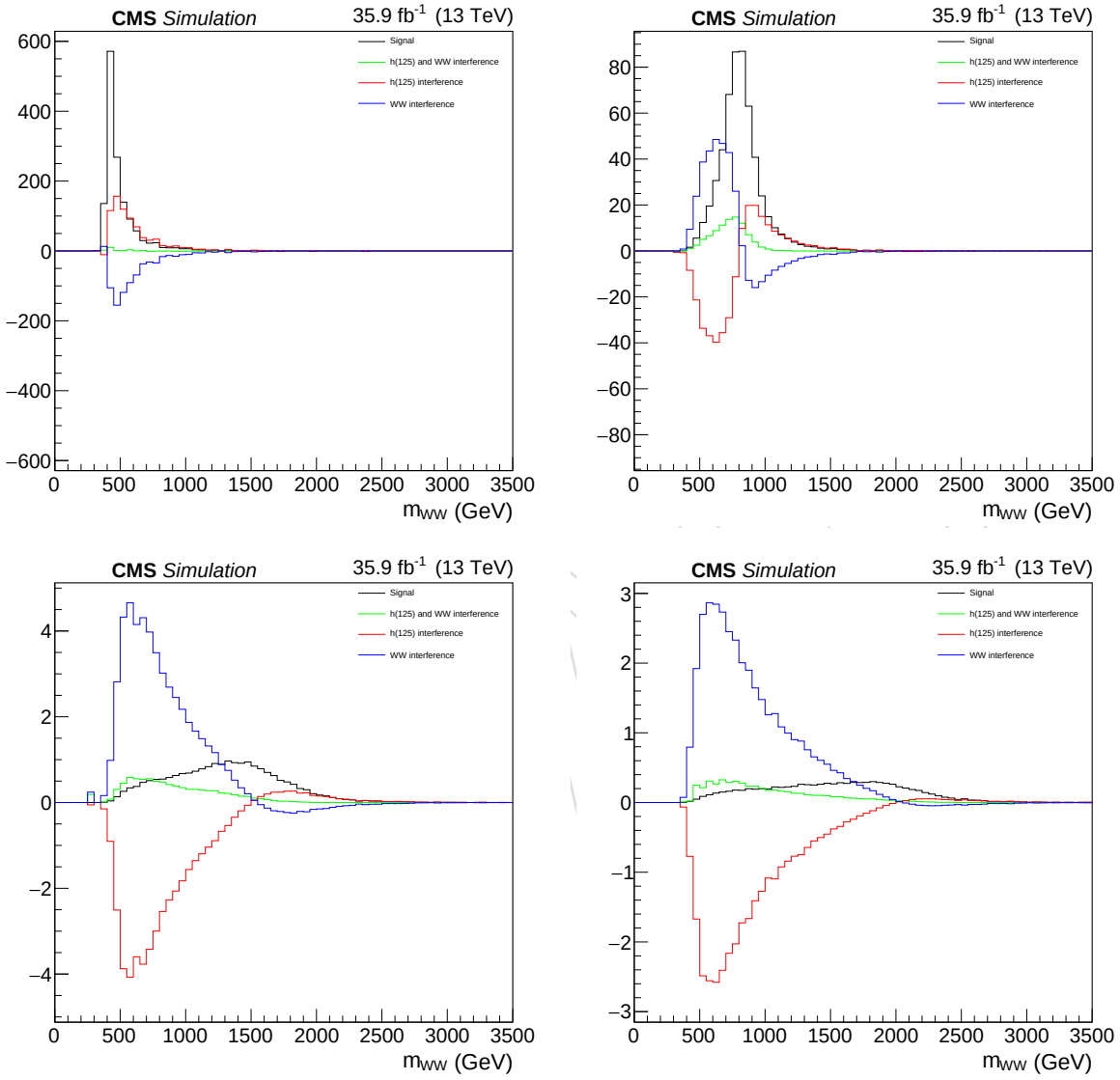

Figure 17: Generator level Higgs mass for a $gg \to X$ signal with a mass of 400 GeV (top left), 800 GeV (top right), 1500 GeV (bottom left) and 2000 GeV (bottom right). The contribution of the interference with the $gg \to \mathrm{WW}$ continuum and the $gg \to H(125)$ are shown in blue and red respectively. The overall interference contribution is shown in green.

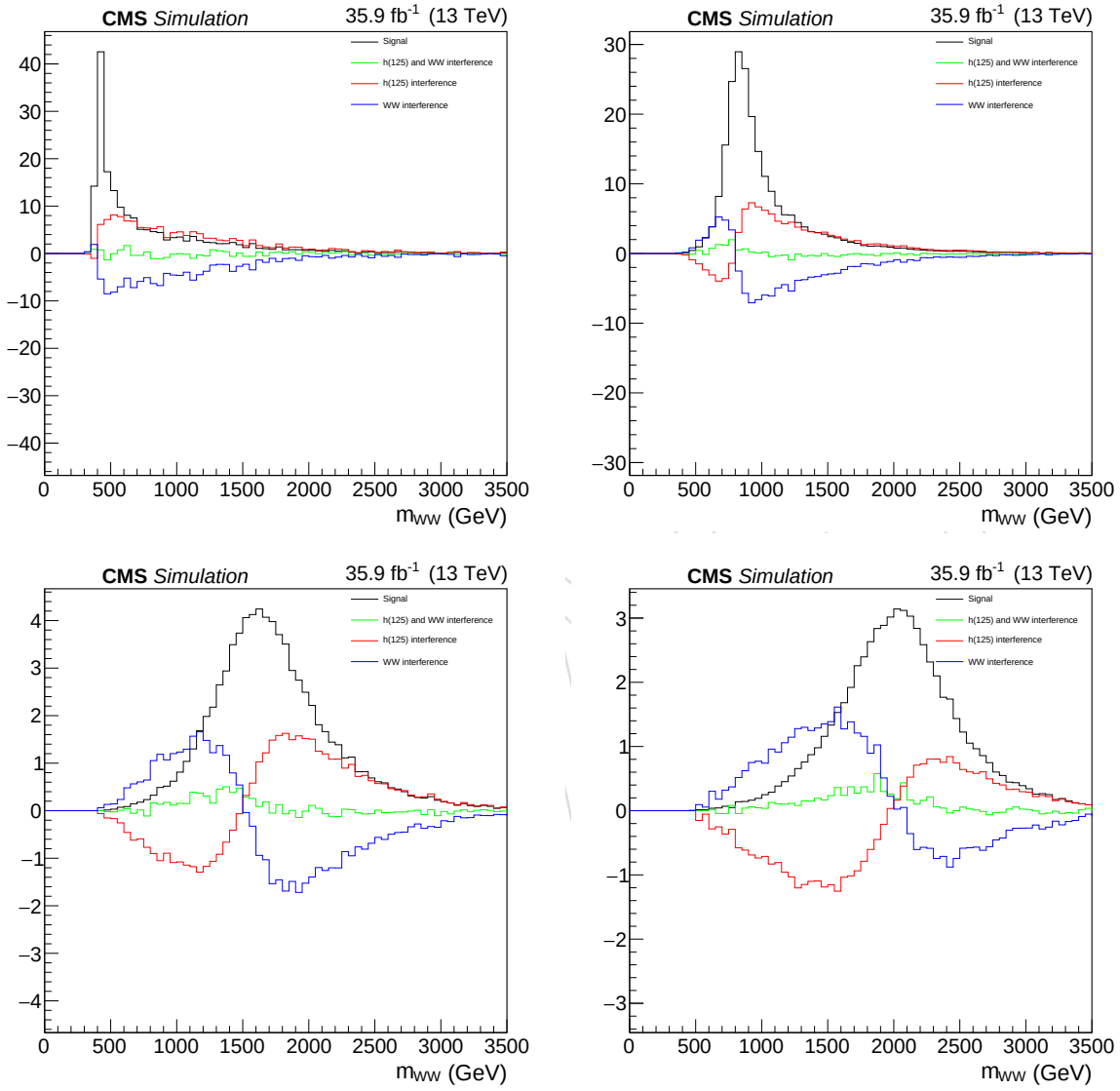

Figure 18: Generator level Higgs mass for a $vbf \rightarrow X$ signal with a mass of 400 GeV (top left), 800 GeV (top right), 1500 GeV (bottom left) and 2000 GeV (bottom right). The contribution of the interference with the $vbf \rightarrow$ WW continuum and the $vbf \rightarrow H(125)$ are shown in blue and red respectively. The overall interference contribution is shown in green.

8 Backgrounds

The main backgrounds in this analysis are from W+jets and top production. Sub-dominant backgrounds considered are SM diboson production, Z+Jets and QCD multijet production. Background samples are modeled by MC simulation that has been properly reweighted to account for known discrepancies between data and simulated events.

The diboson (WW, WZ and ZZ) and Z+jets production backgrounds are expected to be small and are estimated from NLO simulation. Those backgrounds which interfere with the signal($gg \to$ WW and $qq \to qq$WW) have been estmimated using the MELA reweighting scheme discussed in section 7. The boosted V tagging efficiency data-to-MC scale factor of 1.0 ± 0.06 taken from [33] is applied to diboson events satisfying the final selection cuts. The QCD multijet background is greatly suppressed in this analysis and is estimated with PYTHIA8.

Most events in the selected sample come from W+jets and top production. The MADGRAPH5_AMC@NLO MC is used to describe the contribution from W+jets, $t\bar{t}$ and single top s-channel production while POWHEG is used for single top t-channel and tW-channel production. A good level of agreement between data and simulation is demonstrated by the SR, SB and Top CR distributions shown in the previous sections.

For the final analysis, the W+jets and top X candidate mass distributions ($m_{\ell\nu J}$ and $m_{\ell\nu jj}$) are extracted from the simulation. However in each category an alternative normalization to that predicted by the MC is allowed. The normalization of each background is parametrized by a multiplicative factor that is allowed to float free in the fits to the data and is constrained by the observed yields in the SR, SB and Top CR.

9 Systematics

In this section the systematic uncertainties affecting the boosted and resolved categories are presented.

Effects related to the normalization of the background and signal $m_{\ell\nu J}$ and $m_{\ell\nu jj}$ distributions considered here are luminosity, subdominant background cross-sections and boosted V tagging efficiency. Effects related to the normalization and shape of the $m_{\ell\nu J}$ and $m_{\ell\nu jj}$ distributions are lepton selection efficiency, lepton energy scale, unclustered energy scale, jet energy scale and resolution, *b*-tagging, pileup reweighting, boosted V mass scale and resolution, PDF and QCD scale uncertainties.

Normalization systematic errors affecting the boosted category are summarized in Table 6. Table 7 summarizes the normalization systematic errors affecting the resolved category. Systematic uncertainties affecting the shape of the signal, W+Jets and top background $m_{\ell\nu J}$ and $m_{\ell\nu jj}$ distributions have been studied in all relevant cases. Those found to be non-negligible, and so considered in the final fit to the data, are highlighted in tables 6 and 7. Any correlations between the Top CR, SB and SR are fully taken into account in the final fit.

Luminosity uncertainty

The uncertainty on LHC luminosity is 2.5% [5].

Subdominant background normalization

We assign a 10% uncertainty on the normalization of the diboson, Z+jets and QCD multijets backgrounds.

Lepton trigger, identification and isolation

Lepton trigger, reconstruction, identification and isolation scale factors have been computed using tag-and-probe technique for both muons [19, 20] and electrons [35, 39, 40]. The scale factors are apllied in this analysis and the related uncertainties estimated by shifting up and down the scale factors by their uncertainty. Additionally for muons flat systematic uncertainties are assigned for the trigger (0.5%), identification (1%) and isolation (0.5%). For electrons with Pt $>$ 80 GeV a flat systematic uncertainty of 1% is also assigned to the reconstruction. The total uncertainty for muons is 1.5%, while for electrons it is 2.9%.

Lepton energy scale

Uncertainties related to the lepton energy scale are small [41, 42]. The effect on normalisation is found to be around 0.5%.

Jet energy scale and resolution

Uncertainties related to the jet energy scale are calculated by changing the jet energies by ± 1 sigma error of the corresponding jet energy corrections for AK8 and AK4 jets [24].

Uncertainties related to the jet energy resolution are accessed by applying the recommended resolution scale factors [43], which typically have a value of 1.1, and comparing the result with the central analysis. The variations in jet energy scale and resolution are also propagated to the $\not{E}_{\mathrm{T}}$ measurement.

$\not\!\!E_T$ unclustered energy uncertainty

The $\not\!\!E_T$ measurement is affected by uncertainties on the unclustered energy in the event. In order to evaluate the effect we used the JME POG $\not\!\!E_T$ uncertainty tool [44], which varies the scale and resolutions of the objects in the event according to the measured uncertainties for each object and then recalculates the $\not\!\!E_T$.

Pileup

The events with a given number of true interactions in the simulated samples were re-weighted using the instantaneous luminosity profile of the data and a minimum-bias cross section of 69.2 mb. The uncertainty is studied by re-estimating the number of events with a given number of true interactions using $\pm 5\%$ values of the minimum-bias cross section [36].

Heavy quark flavor tagging uncertainty

The uncertainties related to the b-tagging efficiency of b-jets and to the mistag probability of light jets has been estimated by independently varying the corresponding data-to-Monte Carlo scale factors (SF_b and SF_l) by one sigma, as a function of p_T and η of the jets [27].

Boosted V tagging efficiency uncertainty

The uncertainty related to the boosted V tagging efficiency for signal, top and diboson background events has been estimated using the results of [33]. Reconstructing merged jets produced by boosted W bosons in top decays, the study of [33] quotes a data-to-Monte Carlo scale factor uncertainty of 6% for our selection cuts. Since this is measured in the $t\bar{t}$ regime a V tagging p_T-extrapolation uncertainty is also considered for the high mass signals.

Boosted V mass scale and resolution uncertainties

The PUPPI softdrop mass scale (JMS) and resolution (JMR) have been measured using boosted W bosons from top decays. For signal and backgrounds containing hadronic W decays, uncertainties related to the JMS and JMR have been estimated by varying the mass and resolution calibrations within their measured uncertainties [33]. As there are no corresponding calibration measurements for QCD jets we treat the JMS and JMR measurements for boosted W events as a standard candle. For those backgounds containing QCD jets we apply JMS and JMR uncertainties based on varying the mass and resolution by the measured uncertainties for boosted W events. These uncertainties are found to have a negligible effect on the $m_{\ell\nu J}$ and $m_{\ell\nu jj}$ shapes and so are treated as normalisation systematics only. The boosted jet mass uncertainties for hadronic W decays and QCD jets are treated as uncorrelated.

PDF uncertainty

The acceptance uncertainties related to the PDF set NNPDF 3.0 have been determined according to the procedure described in [45], using the 100 PDF weights together in a single envelope around the central distribution.

QCD scale uncertainties

The acceptance uncertainties related to the renormalisation and factorisation scale choice have been determined according to the procedure described in [46]. The background and signal uncertainties from this source are treated as uncorrelated in the final fit to the data.

Table 6: Summary of systematic uncertainties on the normalization of signal and backgrounds in the boosted SR. Where appropiate the uncertainty in the signal, W+Jets and top background shapes are also considered in the final fit to the data (*).

Source	W+Jets	Top	Minor bkgs	ggX500	ggX2000	vbfX500	vbfX2000
Luminosity				2.5%			
Electron trigger & ID				2.9%			
Muon trigger & ID				1.5%			
Electron energy scale	0.5%	0.5%	0.5%	0.4%	0.08%	0.3%	0.05%
Muon momentum scale	0.4%	0.6%	0.5%	0.3%	0.1%	0.4%	0.07%
Unclustered energy scale	0.3%	0.2%	0.3%	0.7%	0.3%	0.4%	0.1%
Jet energy scale *	3.0%	1.0%	2.0%	0.9%	0.2%	0.4%	0.5%
Jet energy resolution *	1.0%	1.0%	1.0%	1.0%	0.3%	0.5%	0.3%
b-tagging efficiency SF	0.1%	1.0%	0.1%	0.1%	0.1%	0.1%	0.1%
mis-tag SF	0.4%	0.5%	0.4%	0.3%	0.6%	0.3%	0.6%
VV normalization			10%				
QCD normalization			10%				
Z+jets normalization			10%				
Boosted V tagging		6%	6% (VV)	6%	6%	6%	6%
Boosted V tagging extrapolation				1.5%	13%	1.5%	13%
Boosted V JMS	1%	0.1%	0.5% (VV)	0.6%	0.4%	0.6%	0.6%
Boosted V JMR	1.5%	4.5%	5.0% (VV)	5.1%	4.7%	5.0%	4.7%
Pileup	1.0%	0.5%	0.6%	0.2%	0.6%	0.7%	0.3%
PDFs *	1.5%	2.0%		1.0%	2.7%	6.0%	7.0%
QCD scale *	3.0%	18.0%		1.4%	1.6%	1.5%	0.5%

Table 7: Summary of systematic uncertainties on the normalization of signal and backgrounds in the resolved SR. Where appropiate the uncertainty in the signal, W+Jets and top background shapes are also considered in the final fit to the data (*).

Source	W+Jets	Top	Minor bkgs	ggX300	ggX600	vbfX300	vbfX600
Luminosity				2.5%			
Electron trigger & ID				2.9%			
Muon trigger & ID				1.5%			
Electron energy scale	0.6%	0.4%	0.7%	0.5%	0.2%	0.4%	0.2%
Muon momentum scale	0.3%	0.4%	0.1%	0.3%	0.7%	0.3%	0.2%
Unclustered energy scale	1.1%	0.9%	1.3%	3.0%	1.1%	3.0%	0.8%
Jet energy scale *	5.5%	2.5%	4.0%	2.0%	0.2%	1.5%	1.0%
Jet energy resolution *	0.6%	0.6%	0.9%	1.0%	1.3%	0.8%	1.7%
b-tagging efficiency SF	0.1%	1.0%	0.2%	0.1%	0.1%	0.1%	0.1%
mis-tag SF	0.3%	0.5%	0.4%	0.3%	0.4%	0.4%	0.3%
VV normalization			10%				
QCD normalization			10%				
Z+jets normalization			10%				
Pileup	0.7%	0.7%	1.0%	1.1%	0.4%	0.9%	0.9%
PDFs *	1.4%	1.8%		1.9%	2.0%	1.7%	4.0%
QCD scale *	11.0%	19.0%		1.1%	3.9%	3.1%	1.0%

10 Statistical analysis and results

Since the decay products of the spin-0 resonance can be fully reconstructed, it is possible to use the reconstructed resonance mass distribution to discriminate signal events from background events. The reconstructed mass will peak around the true resonance mass, m_X, for the signal, while for background processes it will have a broader distribution. The $m_{\ell\nu J}$ and $m_{\ell\nu jj}$ mass distributions in the sideband region, for the three X-tagging categories, are depicted in Figures 19 and 20. The mass distributions are compared to the background predictions after fitting to the data in the SB. In this fit the Top CR is used to constrain the background normalisations. The $m_{\ell\nu J}$ and $m_{\ell\nu jj}$ mass distributions in the signal region, for the three X-tagging categories, are depicted in Figures 21 and 22. The mass distributions are compared to the background predictions after fitting to the data in the SR with the SB and Top CR used to constrain the background normalisations. Figures 23 and 24 show the same events also split by lepton type. The total uncertainty on the background prediction is represented by the gray band in these Figures. Table 8 shows the predicted and observed number of events in the SR, SB and Top CR in the boosted and resolved production categories.

Table 8: Predicted yields in the boosted and resolved analysis categories with 35.9 fb^{-1}. The predicted numbers are derived from a fit to the data in the SR, SB and Top CR regions.

Boosted category	untagged	ggX	vbfX
SR predicted	23900 ± 300	7523 ± 110	1127 ± 43
SR observed	23864	7547	1133
SB predicted	17035 ± 230	6195 ± 100	766 ± 37
SB observed	17076	6190	759
Top CR predicted	16242 ± 162	5296 ± 77	1749 ± 54
Top CR observed	16229	5277	1751
Resolved category	**untagged**	**ggX**	**vbfX**
SR predicted	144369 ± 1066	45086 ± 349	5669 ± 153
SR observed	144227	45052	5695
SB predicted	195535 ± 1502	58609 ± 397	4270 ± 110
SB observed	195614	58619	4270
Top CR predicted	71936 ± 675	20139 ± 219	6322 ± 168
Top CR observed	71990	20161	6294

A shape-based treatment of the expected and observed distribution of the invariant mass of the resonance candidate increases the sensitivity of the analysis, compared to a simpler counting experiment. A binned histogram-based calculation has been used for the shape analysis.

The normalizations of the two main backgrounds, W+jets and top production, are allowed to float freely. A simultaneous fit to the number of events in the SB and Top CR serves to constrain the normalisations. The calculation to discriminate signal from background events is performed for the twelve individual channels (electrons and muons in the three X-tagging categories of the boosted and resolved analyses) and then combined taking into account the correlations among the systematic uncertainties. These uncertainties may affect either the shape of the distributions or their normalization. Systematics which affect the shape are taken into account using mass distributions with the systematic varied up and down by one sigma. The final analysis accounts for the finite statistics of the background simulations by including bin-by-bin MC statistical uncertainties.

Figure 25 displays the pull distributions for the physics systematics after fitting the data in the SR, SB and Top CR. The floating W+jets and top normalisations, shown in bins 7 to 12 for the

boosted category and bins 15 to 20 for the resolved category, are constrained by the fit and the predictions found to be in good agreement with the data. The goodness of fit is estimated using the saturated model algorithm [47]. Figure 26 shows the G.O.F estimator distribution for 250 background toy fits. The value obtained in data is shown by the red line, good compatibility with the toy distribution is observed.

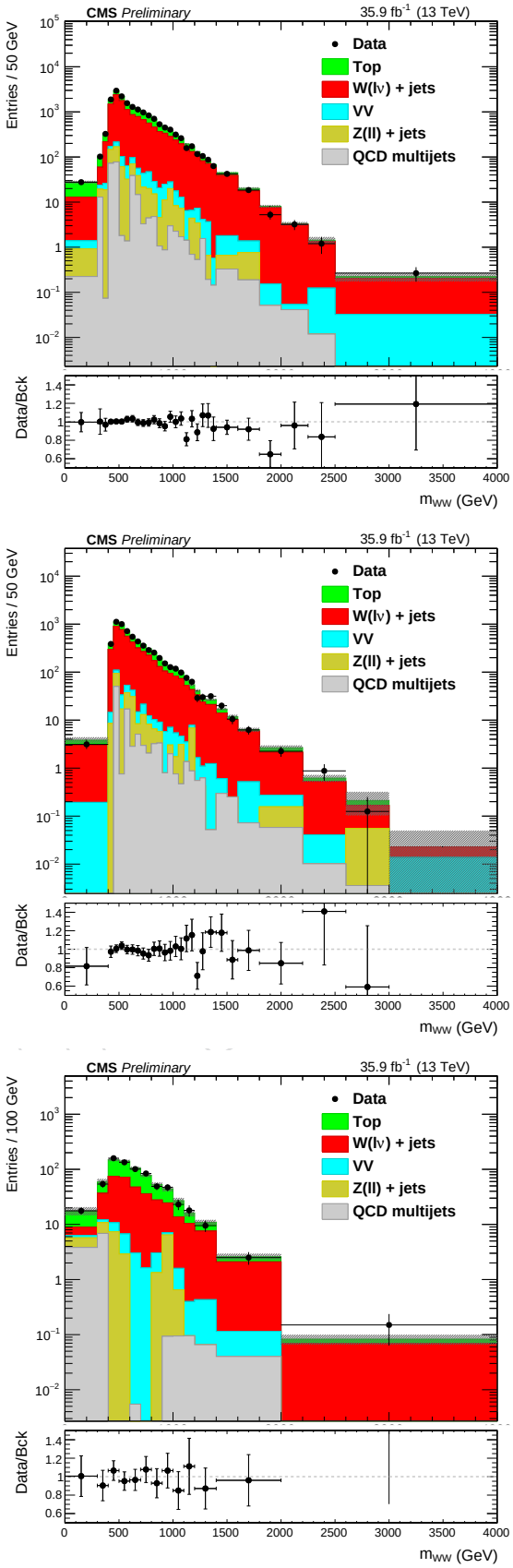

Figure 19: The $m_{\ell\nu J}$ for events passing the boosted SB selection in the untagged (top), gg$X$ (middle) and vbf$X$ categories (bottom) after fitting to the SB and Top CR

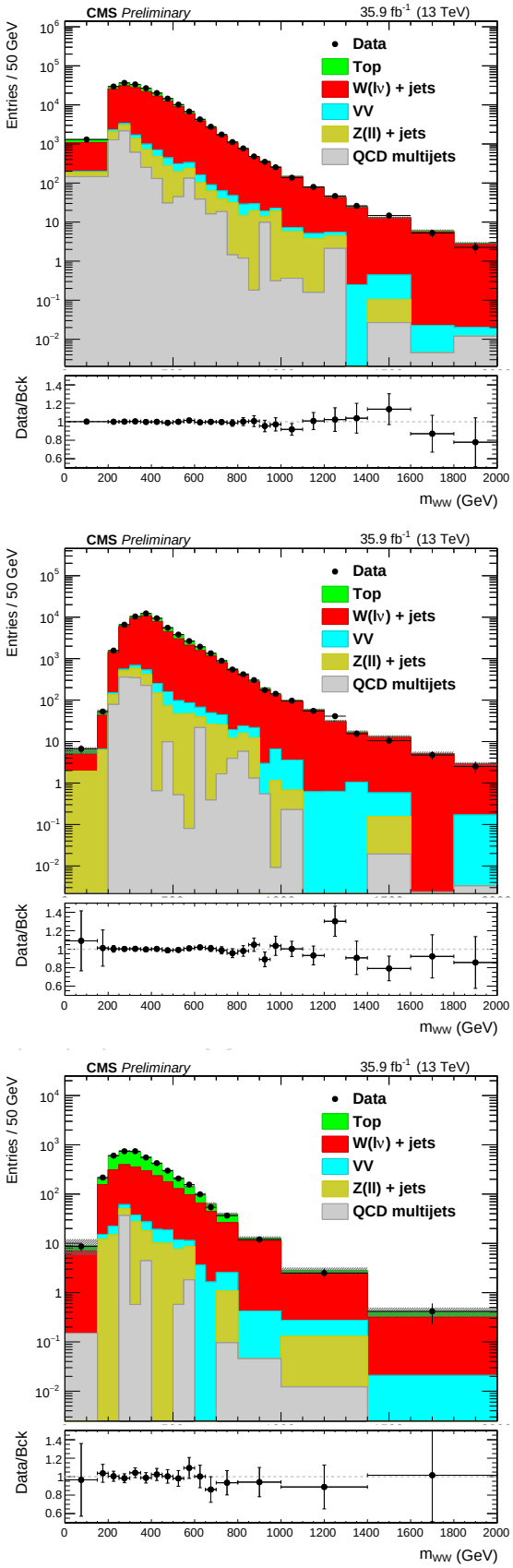

Figure 20: The $m_{\ell\nu jj}$ for events passing the resolved SB selection in the untagged (top), gg$X$ (middle) and vbf$X$ categories (bottom) after fitting to the SB and Top CR.

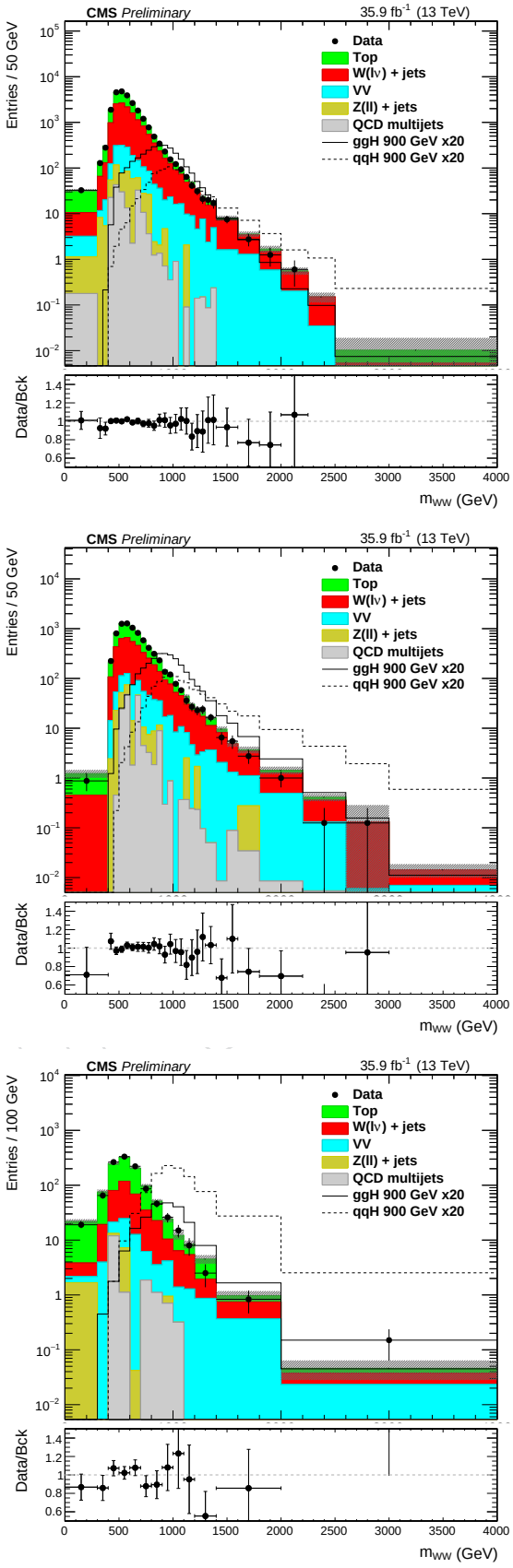

Figure 21: The $m_{\ell\nu J}$ in data and simulation for events passing the boosted SR selection in the untagged (top), gg$X$ (middle) and vbf$X$ categories (bottom) after fitting to the SR, SB and Top CR.

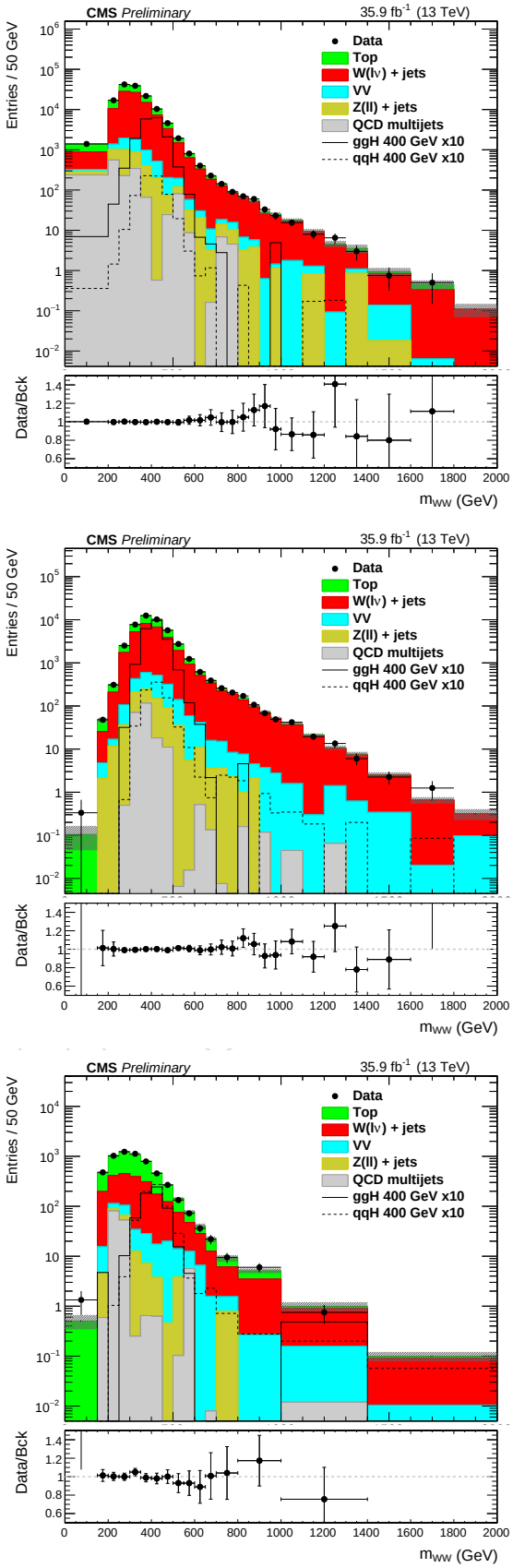

Figure 22: The $m_{\ell\nu jj}$ in data and simulation for events passing the resolved SR selection in the untagged (top), gg$X$ (middle) and vbf$X$ categories (bottom) after fitting to the SR, SB and Top CR.

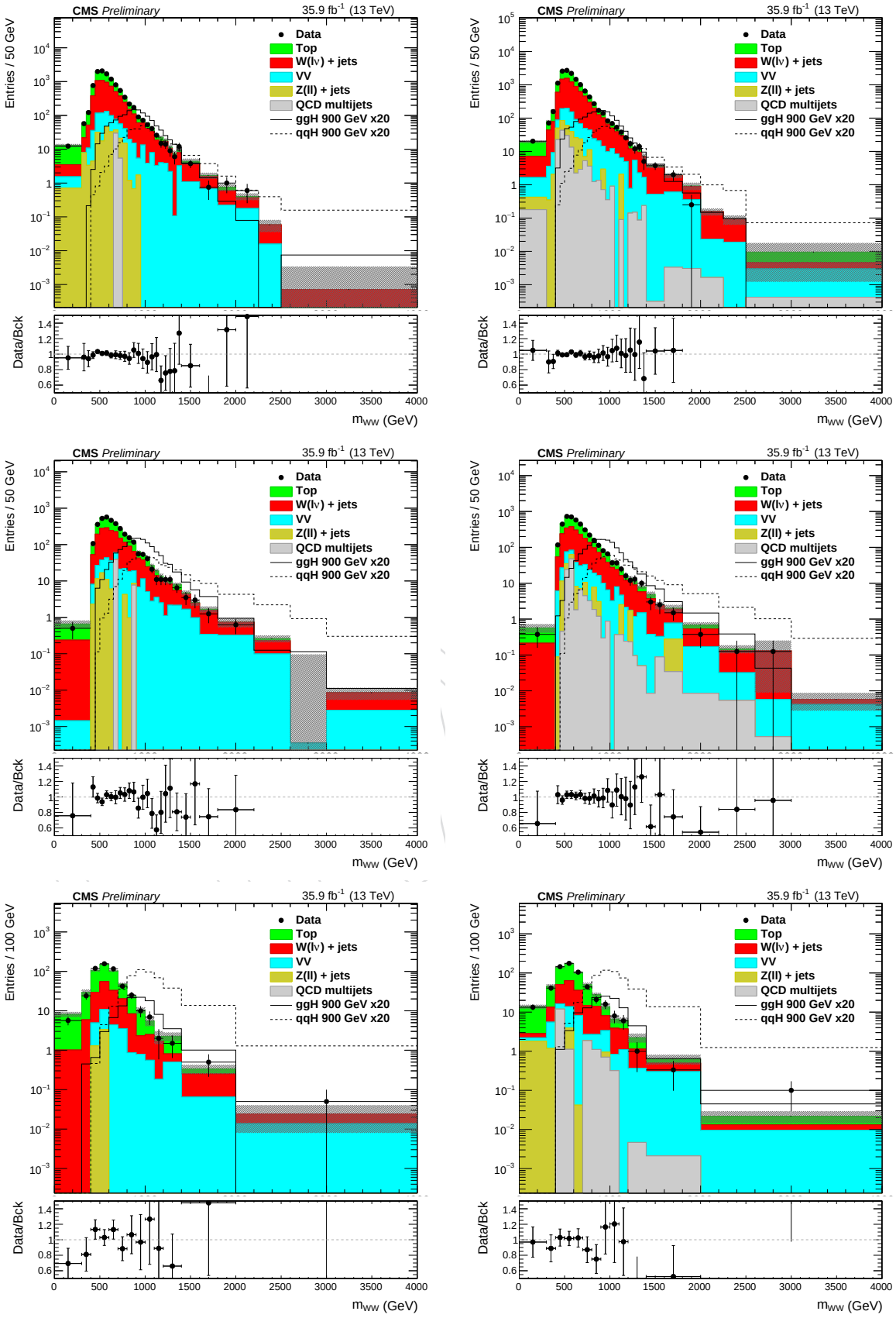

Figure 23: The $m_{\ell\nu J}$ in data and simulation for electron (left) and muon (right) reconstructed events passing the boosted SR selection in the untagged (top), gg$X$ (middle) and vbf$X$ categories (bottom) after fitting to the SR, SB and Top CR.

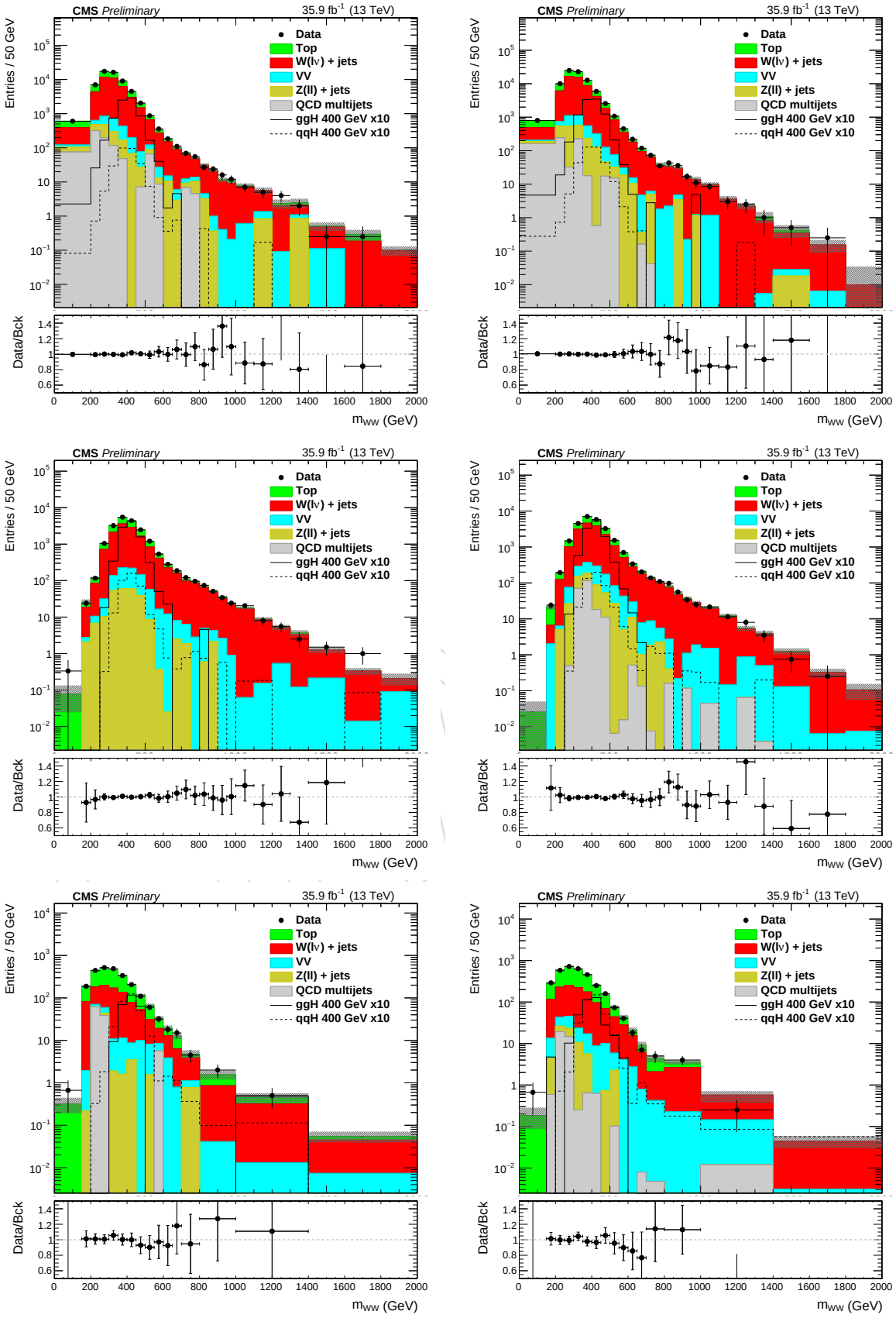

Figure 24: The $m_{\ell\nu jj}$ in data and simulation for electron (left) and muon (right) reconstructed events passing the resolved SR selection in the untagged (top), gg$X$ (middle) and vbf$X$ categories (bottom) after fitting to the SR, SB and Top CR.

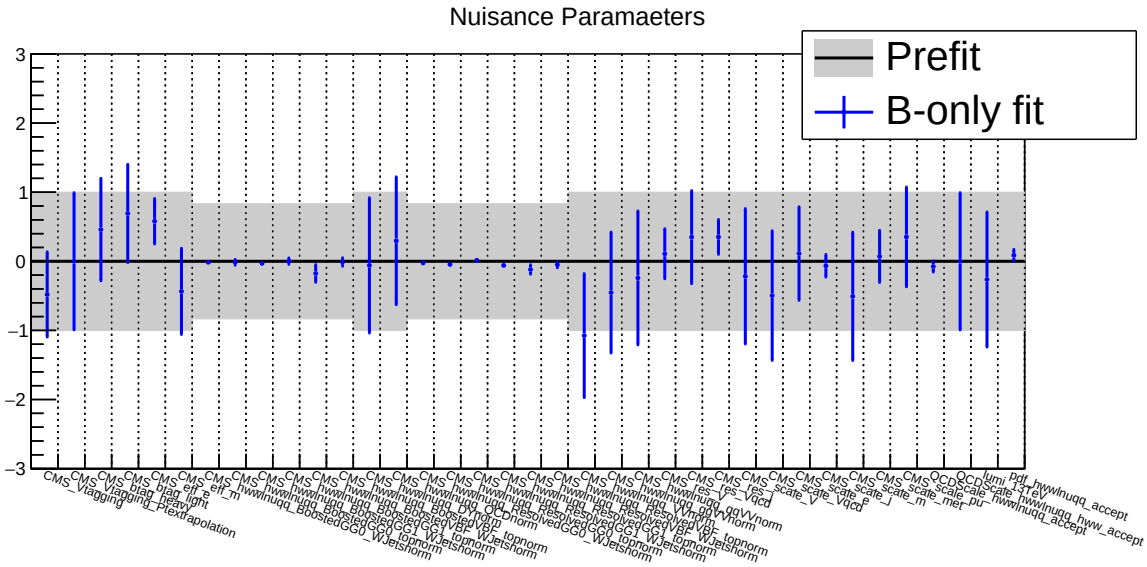

Figure 25: The pull distributions for the physics systematics obtained from a background only fit to the data in the SR, SB and Top CR.

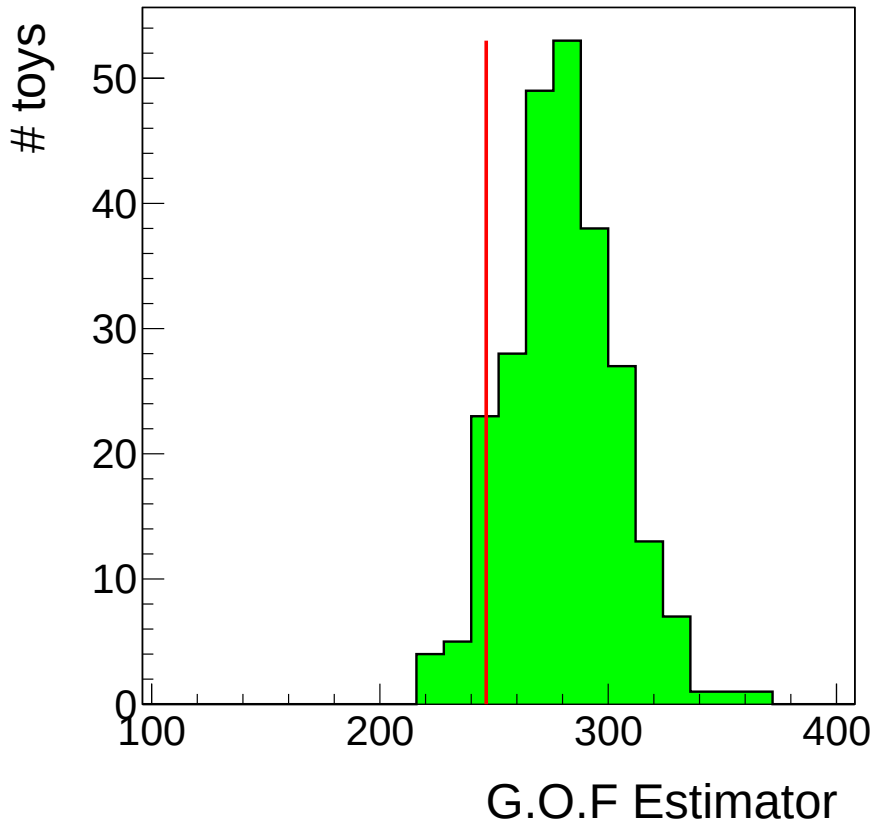

Figure 26: Distribution of the saturated model G.O.F estimator for 250 toy background fits. The red line indicates the value obtained in data.

Expected and observed upper limits on the X production cross-section time branching fraction to WW may be determined as function of the resonance mass hypothesis, taking as input the mass distribution for data, and for the background and signal expectations. The statistical procedure, based on the profile likelihood method, uses the asymptotic CL_s approach [48], implemented in the official tool developed by the CMS Higgs combination group [49]. Within this framework systematic uncertainties are treated as nuisance parameters and the dominant backgrounds allowed to float freely by implementing a large uniform a-priori uncertainty on the normalisation. The impact of each systematic on the signal strength, defined as the shift in the signal strength as a nuisance parameter is varied up and down by 1σ of its post-fit value, is shown in Figure 27. This is effectively a measure of the correlation between the systematic and the signal strength.

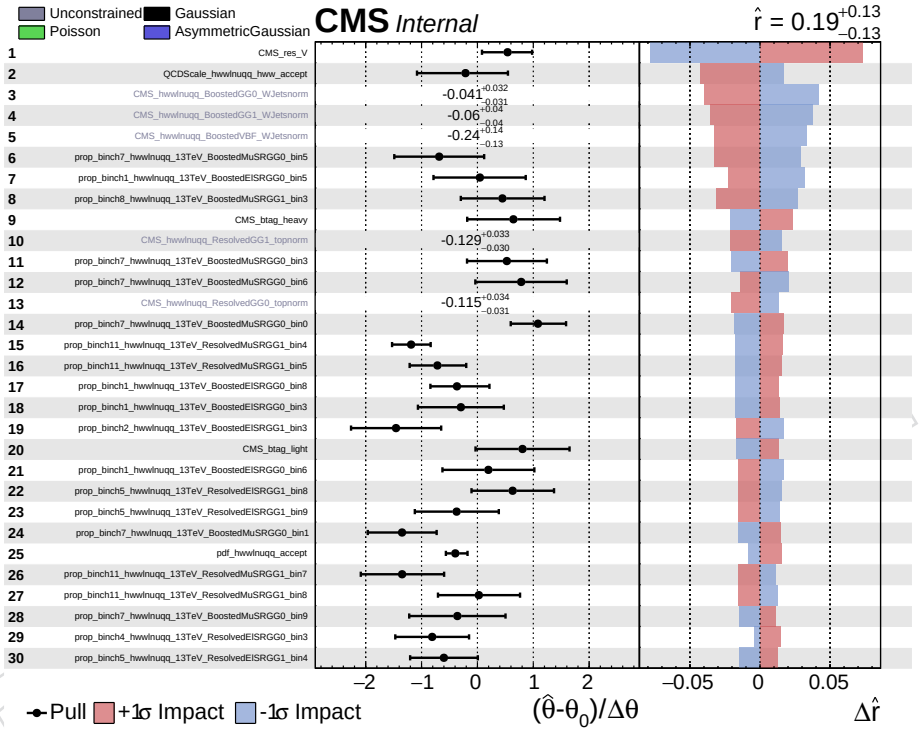

Figure 27: The impact of the systematics on the 500 GeV Higgs signal strength where the impact is defined as the shift in the signal strength as a nuisance paramater is varied up and down by 1σ of its post-fit value with all other parameters profiled as normal. The 30 most significant systematics, with respect to their impact on the signal strenght, are shown.

The results are expressed as upper limits on the cross section times branching fraction of the process $X \rightarrow$ WW for a SM like Higgs boson. The expected and observed limits on σ_X times branching fraction $X \rightarrow$ WW are determined for m_X hypotheses between 200 GeV/c^2 and 3000 GeV/c^2, in the range from 400 GeV/c^2 to 3000 GeV/c^2 for the boosted analysis and 200 GeV/c^2 to 600 GeV/c^2 for the resolved analysis. Figure 28 and Figure 29 show the expected and observed limits for the full dataset recorded during 2016 at 13 TeV, corresponding to a luminosity of 35.9 fb^{-1}.

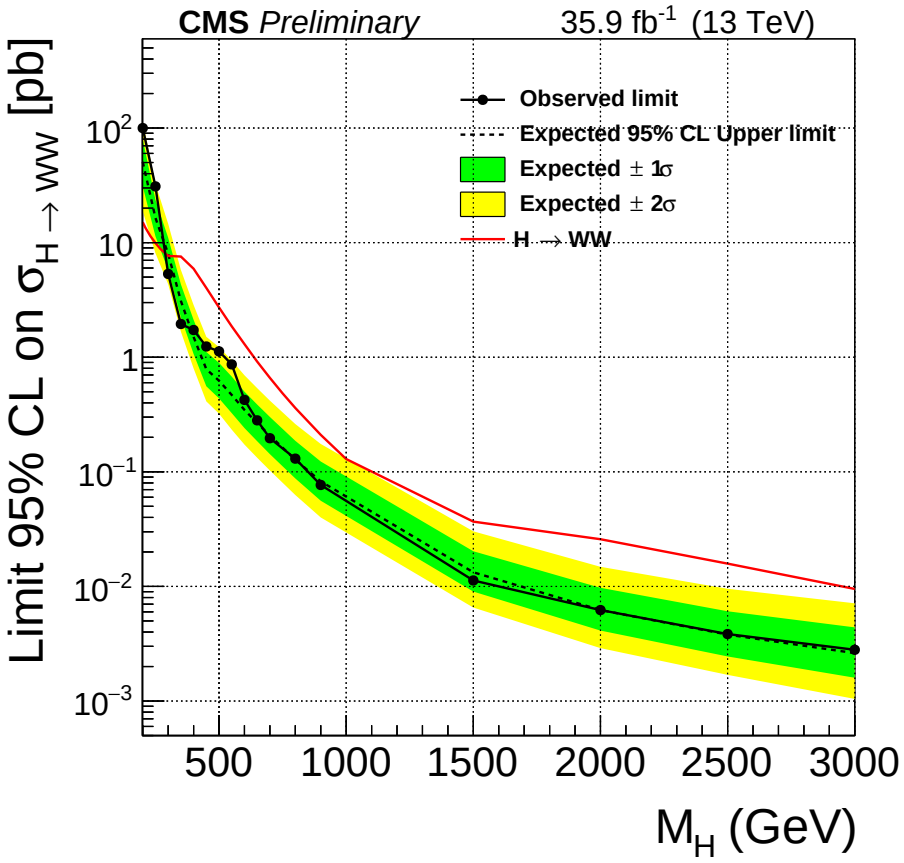

Figure 28: Expected and observed 95% CL upper limits on the X production cross section times branching fraction $X \rightarrow$ WW obtained using the CL_s technique. The 68% and 95% ranges of expectation for the background-only model are also shown with green and yellow bands, respectively.

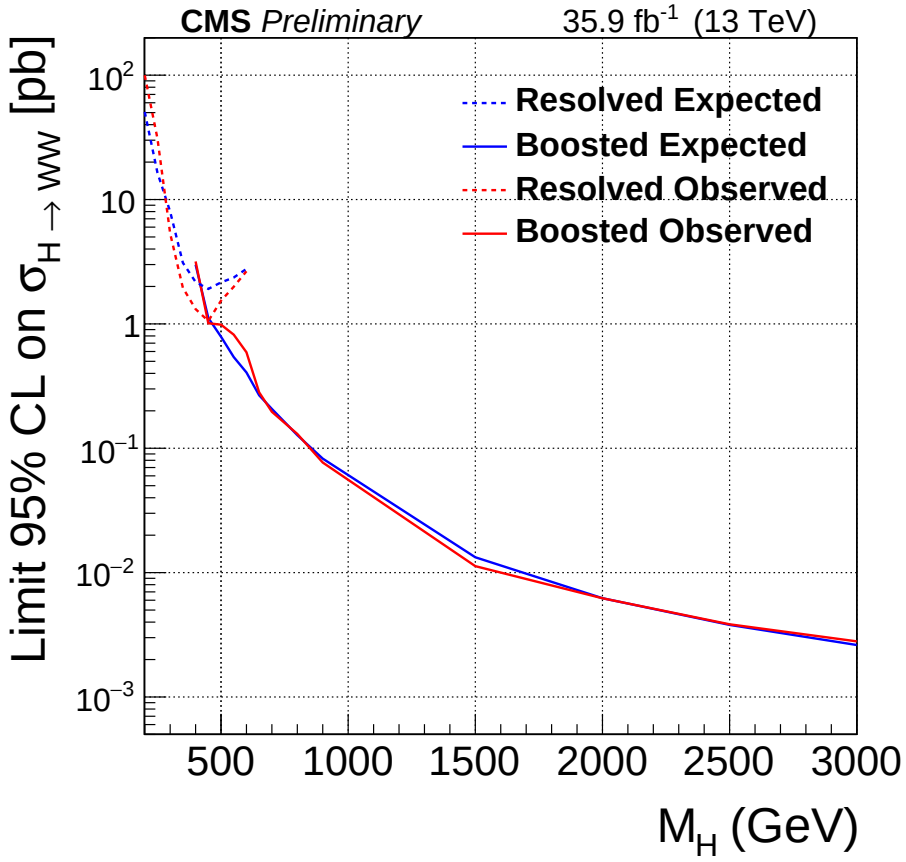

Figure 29: Expected and observed 95% CL upper limits on the X production cross section times branching fraction $X \rightarrow$ WW obtained using the CL_s technique. The resolved and boosted limits are represented by the dashed and full lines respectively.

11 Combination of $X \to WW \to \ell\nu\, q\bar{q}$ and the $X \to WW \to 2\ell 2\nu$

A search for high mass spin-0 resonances has also been conducted in the $X \to WW \to 2\ell 2\nu$ final state with 2016 data. To maximise sensitivity to a potential signal a full combination of the $X \to WW \to \ell\nu\, q\bar{q}$ and the $X \to WW \to 2\ell 2\nu$ analyses has been conducted. Figure 30 shows the expected and observed limits of the combined analyses. The relative rate of gluon gluon fusion and VBF production has also been investigated by introducing a parameter, f_{VBF}, which represents the fraction of the total cross section which is due to VBF production. Figure 31 shows the expected limits with f_{VBF} allowed to float, f_{VBF} set to 0 and f_{VBF} set to 1.

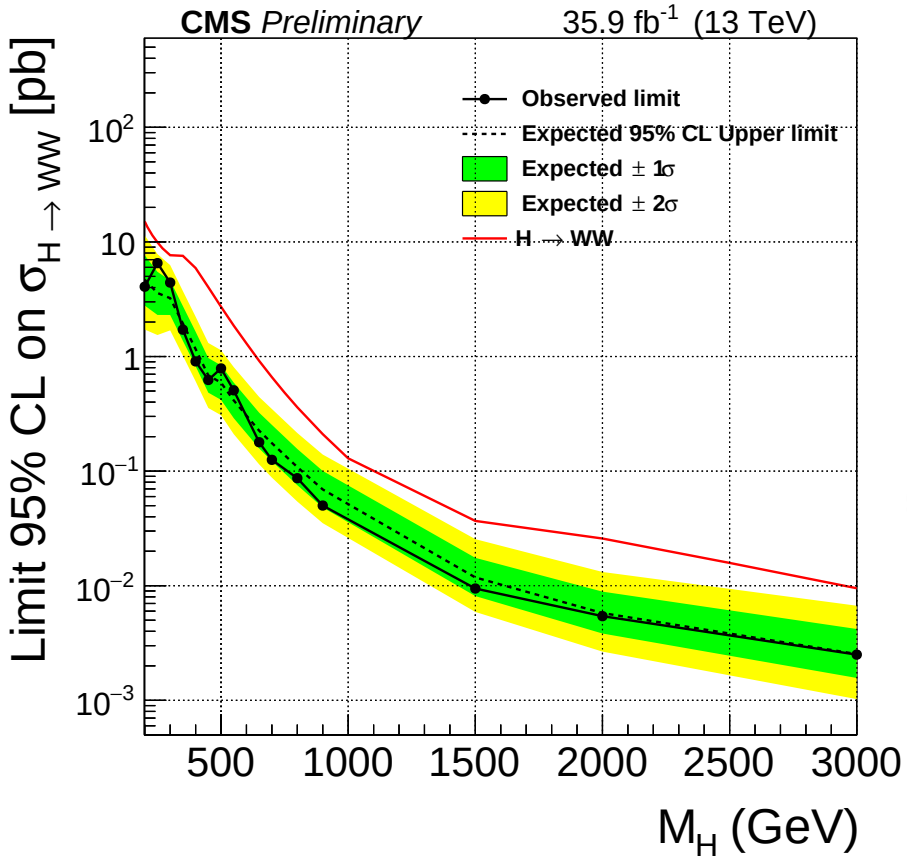

Figure 30: Expected and observed 95% CL upper limits on the X production cross section times branching fraction $X \to WW$ for the $X \to WW \to \ell\nu\, q\bar{q}$ and the $X \to WW \to 2\ell 2\nu$ analyses combined. The 68% and 95% ranges of expectation for the background-only model are also shown with green and yellow bands, respectively.

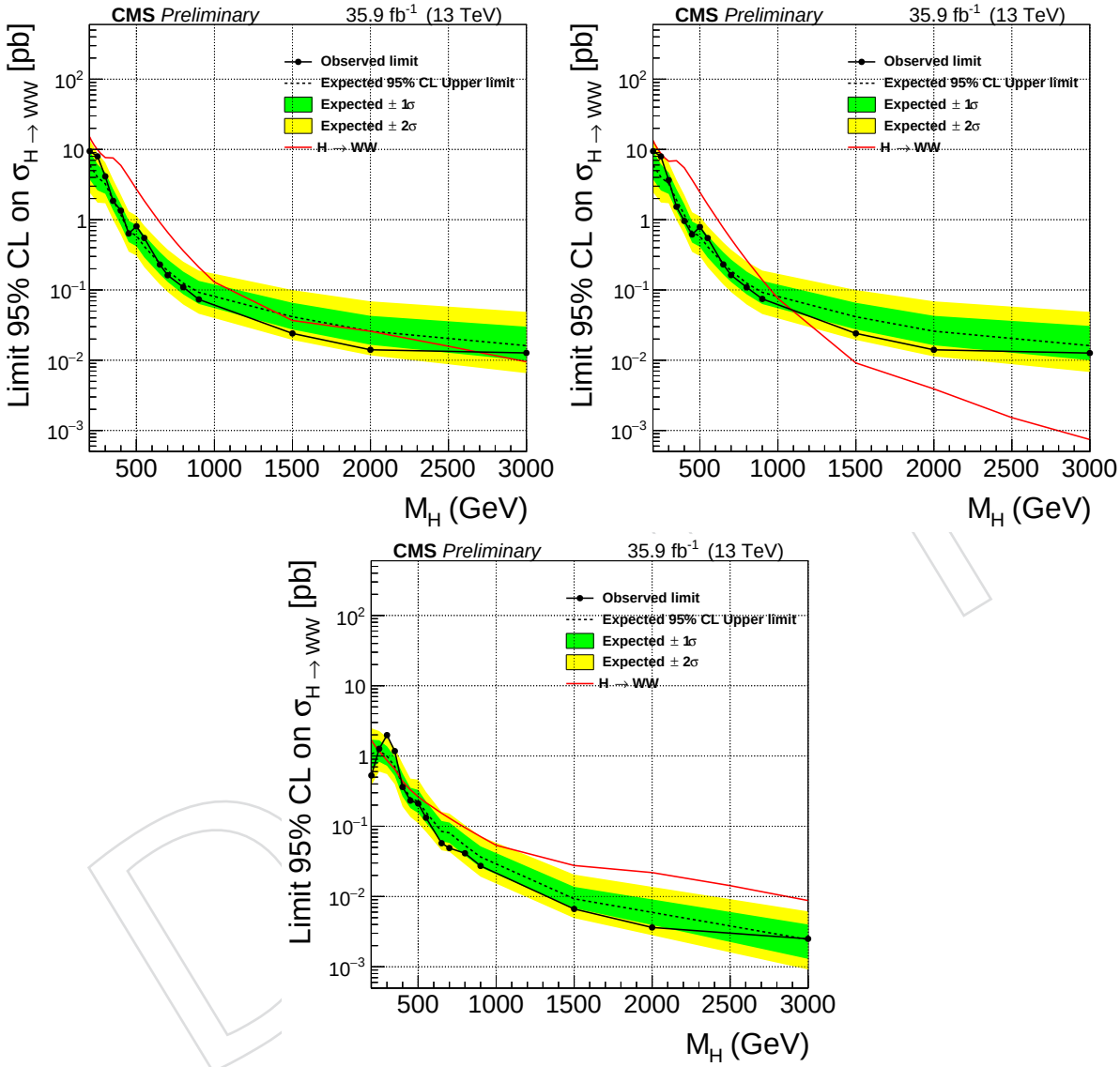

Figure 31: Expected and observed 95% CL upper limits on the X production cross section times branching fraction $X \to \mathrm{WW}$, for the $X \to \mathrm{WW} \to \ell\nu\,\mathrm{q\bar{q}}$ and the $X \to \mathrm{WW} \to 2\ell 2\nu$ analyses combined, with f_{VBF} allowed to float (top left), f_{VBF} set to 0 (top right) and f_{VBF} set to 1 (bottom). The 68% and 95% ranges of expectation for the background-only model are also shown with green and yellow bands, respectively.

References

[1] ATLAS Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", *Phys. Lett.* **B716** (2012) 1–29, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.

[2] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys. Lett.* **B716** (2012) 30–61, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.

[3] V. Barger et al., "LHC Phenomenology of an Extended Standard Model with a Real Scalar Singlet", *Phys. Rev.* **D77** (2008) 035005, doi:10.1103/PhysRevD.77.035005, arXiv:0706.4311.

[4] CMS Collaboration, M. Pelliccioni, "CMS High mass WW and ZZ Higgs search with the complete LHC Run1 statistics", in *Proceedings, 50th Rencontres de Moriond Electroweak Interactions and Unified Theories: La Thuile, Italy, March 14-21, 2015*, pp. 47–52. 2015. arXiv:1505.03831.

[5] CMS Collaboration, "CMS luminosity measurement for the 2016 data taking period", CMS PAS LUM-17-001, 2017.

[6] CMS Collaboration, "Global tags for condidtions data", twiki, CERN, 2018.

[7] 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.

[8] 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.

[9] 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.

[10] Y. Gao et al., "Spin determination of single-produced resonances at hadron colliders", *Phys. Rev.* **D81** (2010) 075022, doi:10.1103/PhysRevD.81.075022, arXiv:1001.3396.

[11] S. Bolognesi et al., "On the spin and parity of a single-produced resonance at the LHC", *Phys. Rev.* **D86** (2012) 095031, doi:10.1103/PhysRevD.86.095031, arXiv:1208.4018.

[12] A. V. Gritsan, R. Rntsch, M. Schulze, and M. Xiao, "Constraining anomalous Higgs boson couplings to the heavy flavor fermions using matrix element techniques", *Phys. Rev.* **D94** (2016), no. 5, 055023, doi:10.1103/PhysRevD.94.055023, arXiv:1606.03107.

[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] S. Kallweit et al., "NLO QCD+EW predictions for V + jets including off-shell vector-boson decays and multijet merging", *JHEP* **04** (2016) 021, doi:10.1007/JHEP04(2016)021, arXiv:1511.08692.

[15] CMS Collaboration, "pt based reweighting of ttbar mc", twiki, CERN, 2018.

[16] 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.

[17] CMS Collaboration, "Summary table of samples produced for the 1 billion campaign, with 25ns bunch-crossing - https://twiki.cern.ch/twiki/bin/viewauth/cms/summarytable1g25ns", twiki, CERN, 2017.

[18] CMS Collaboration, "Met filter recommendations for run 2 - https://twiki.cern.ch/twiki/bin/viewauth/cms/missingetoptionalfiltersrun2", twiki, CERN, 2017.

[19] CMS Collaboration, "Reference muon id, isolation and trigger efficiencies for run 2 - https://twiki.cern.ch/twiki/bin/viewauth/cms/muonreferenceeffsrun2", twiki, CERN, 2017.

[20] CMS Collaboration, "Muon work in progress and pag results - https://twiki.cern.ch/twiki/bin/view/cms/muonworkinprogressandpagresults", twiki, CERN, 2017.

[21] CMS Collaboration, "Egamma id recipes run2 - https://twiki.cern.ch/twiki/bin/view/cms/egammaidrecipesrun2", twiki, CERN, 2017.

[22] M. Cacciari, G. P. Salam, and G. Soyez, "The anti- k t jet clustering algorithm", *Journal of High Energy Physics* **2008** (2008), no. 04, 063.

[23] CMS Collaboration, "Particle-flow reconstruction and global event description with the CMS detector", *JINST* **12** (2017), no. 10, P10003, doi:10.1088/1748-0221/12/10/P10003, arXiv:1706.04965.

[24] CMS Collaboration, "Jet energy corrections in run2 - https://twiki.cern.ch/twiki/bin/view/cms/jecdatamc", twiki, CERN, 2017.

[25] CMS Collaboration, "Recommendations for the 13tev data analysis, run2016 (80x) - https://twiki.cern.ch/twiki/bin/view/cms/jetid13tevrun2016", twiki, CERN, 2017.

[26] CMS Collaboration, "Identification of b-quark jets with the CMS experiment", *JINST* **8** (2013) P04013, doi:10.1088/1748-0221/8/04/P04013, arXiv:1211.4462.

[27] CMS Collaboration, "Btag recommendation 80xrereco - https://twiki.cern.ch/twiki/bin/view/cms/btagrecommendation80xrereco", twiki, CERN, 2017.

[28] CMS Collaboration, "Methods to apply b-tagging efficiency scale factors", twiki, CERN, 2018.

[29] 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.

[30] J. Thaler and K. Van Tilburg, "Maximizing Boosted Top Identification by Minimizing N-subjettiness", *JHEP* **02** (2012) 093, doi:10.1007/JHEP02(2012)093, arXiv:1108.2701.

[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] CMS Collaboration, "Performance of grooming techniques for w-tagging", twiki, CERN, 2018.

[33] CMS Collaboration, "Wztagging of jets, working points and scale factors - https://twiki.cern.ch/twiki/bin/view/cms/jetwtagging", twiki, CERN, 2017.

[34] CMS Collaboration, "Baseline muon selections for run 2 - https://twiki.cern.ch/twiki/bin/view/cms/swguidemuonidrun2", twiki, CERN, 2017.

[35] CMS Collaboration, "Egamma efficiencies and scale factors - https://twiki.cern.ch/twiki/bin/viewauth/cms/egammaidrecipesrun2 efficiencies_and_scale_factors", twiki, CERN, 2017.

[36] CMS Collaboration, "Pileup json files for run 2 - https://twiki.cern.ch/twiki/bin/view/cms/pileupjsonfilefordata", twiki, CERN, 2017.

[37] J. M. Campbell and R. K. Ellis, "An Update on vector boson pair production at hadron colliders", *Phys. Rev.* **D60** (1999) 113006, doi:10.1103/PhysRevD.60.113006, arXiv:hep-ph/9905386.

[38] CMS Collaboration, "Combine physics models : Interference", twiki, CERN, 2018.

[39] V. Botta, "Lepton scale factors on rereco data and summer 16 mc", presentation, CERN, 2017.

[40] A. Kumar, "Electron trigger efficiencies with 2016 data", presentation, CERN, 2016.

[41] CMS Collaboration, "Reference guidelines and results for muon momentum scale and resolution in run 2 - https://twiki.cern.ch/twiki/bin/view/cms/muonreferencescaleresolrun2", twiki, CERN, 2016.

[42] CMS Collaboration, "Energy smearing and scale correction 80x - https://twiki.cern.ch/twiki/bin/view/cms/egmsmearer", twiki, CERN, 2017.

[43] CMS Collaboration, "Jer scaling factors and uncertainties for 13 tev (2015 and 2016) - https://twiki.cern.ch/twiki/bin/view/cms/jetresolution#jer_scaling_factors_and_uncertai", twiki, CERN, 2017.

[44] CMS Collaboration, "Workbookminiaod2016 etmiss", twiki, CERN, 2018.

[45] S. Alekhin et al., "The PDF4LHC Working Group Interim Report", arXiv:1101.0536.

[46] CMS Collaboration, "Scale choice and related uncertainty - https://twiki.cern.ch/twiki/bin/viewauth/cms/citationsforgenerators", twiki, CERN, 2017.

[47] R. D.Cousins, "Generalization of chisquare goodness-of-fit test for binned data using saturated models, with application to histograms", note, University of California, 2013.

[48] G. Cowan, K. Cranmer, E. Gross and O. Vitells, "Asymptotic formulae for likelihood based tests of new physics", *Eur. Phys. J.* **C71** (2011) 1554, doi:10.1140/epjc/s10052-011-1554-0.

[49] CMS Collaboration, "Swguide higgs analysis combined limit - https://twiki.cern.ch/twiki/bin/view/cms/higgswg", twiki, CERN, 2011.

[50] CMS Collaboration, "Sm higgs production cross sections at 13/14 tev (cern report 3)", twiki, CERN, 2016.

[51] LHC Higgs Cross Section Working Group Collaboration, "Handbook of LHC Higgs Cross Sections: 3. Higgs Properties: Report of the LHC Higgs Cross Section Working Group", Technical Report CERN-2013-004. CERN-2013-004, Jul, 2013.

A Data and Monte Carlo samples

Table 9 lists the re-miniAOD data samples used in the analysis and the corresponding luminosities. Table 10 provides the background datasets along with their cross sections. The signal samples, simulated with POWHEG, are

GluGluHToWWToLNuQQ_M*_13TeV_powheg_JHUgen*_pythia8
RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV
VBF_HToWWToLNuQQ_M*_13TeV_powheg_JHUgen*_pythia8
RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV

For the Higgs boson signals, the cross sections used are those reported by the LHC Higgs Cross Section Working Group [50], computed at NNLO and NNLL QCD and NLO EW for gluon fusion, and at NNLO QCD and NLO EW for VBF. The branching fractions are the ones reported in Ref. [51].

Table 9: Data Re-miniAOD samples used in the analysis.

Channel	Dataset	Luminosity [pb^{-1}]
$\mu\nu jj$	/SingleMuon/Run2016B03Feb2017_ver2v2 /SingleMuon/Run2016C03Feb2017v1 /SingleMuon/Run2016D03Feb2017v1 /SingleMuon/Run2016E03Feb2017v1 /SingleMuon/Run2016F03Feb2017v1 /SingleMuon/Run2016G03Feb2017v1 /SingleMuon/Run2016H03Feb2017_ver2v1 /SingleMuon/Run2016H03Feb2017_ver3v1	35900.
$e\nu jj$	/SingleElectron/Run2016B03Feb2017_ver2v2 /SingleElectron/Run2016C03Feb2017v1 /SingleElectron/Run2016D03Feb2017v1 /SingleElectron/Run2016E03Feb2017v1 /SingleElectron/Run2016F03Feb2017v1 /SingleElectron/Run2016G03Feb2017v1 /SingleElectron/Run2016H03Feb2017_ver2v1 /SingleElectron/Run2016H03Feb2017_ver3v1	35900.

Table 10: Background simulated MINIAODSIM samples from the RunIISummer16MiniAODv2 production used in the analysis.

Process	Dataset	σ [pb]
W+Jets	WJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8 RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV	61526.7
W+Jets Pt-0To50	WJetsToLNu_Wpt-0To50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8 RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV	57280.0
W+Jets Pt-50To100	WJetsToLNu_Wpt-50To100_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8 RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV	3258.0
W+Jets Pt-100To200	WJetsToLNu_Pt-100To250_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8 RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV	676.3
W+Jets Pt-250To400	WJetsToLNu_Pt-250To400_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8 RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV	23.94
W+Jets Pt-400To600	WJetsToLNu_Pt-400To600_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8 RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV	3.031
W+Jets Pt-600ToInf	WJetsToLNu_Pt-600ToInf_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8 RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV	0.4524
$t\bar{t}$ Dilep	TTJets_Dilept_TuneCUETP8M2T4_13TeV-amcatnloFXFX-pythia8 RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV	87.31
$t\bar{t}$ SingleLep T	TTJets_SingleLeptFromT_TuneCUETP8M2T4_13TeV-amcatnloFXFX-pythia8 RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV	182.17
$t\bar{t}$ SingleLep Tbar	TTJets_SingleLeptFromTbar_TuneCUETP8M2T4_13TeV-amcatnloFXFX-pythia8 RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV	182.17
t s-channel	ST_s-channel_4f_leptonDecays_13TeV-amcatnlo-pythia8_TuneCUETP8M1 RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV_v6	3.68
t tW antitop	ST_tW_antitop_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1 RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV	35.6
t tW top	ST_tW_top_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1 RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV	35.6
t t-channel antitop	ST_t-channel_antitop_4f_inclusiveDecays_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1 RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV	80.95
t t-channel top	ST_t-channel_top_4f_inclusiveDecays_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1 RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV	136.02
WZ 2L2Q	WZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8 RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV	5.5950
WZ 1L1Nu2Q	WZTo1L1Nu2Q_13TeV_amcatnloFXFX_madspin_pythia8 RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV	10.71
ZZ 2L2Q	ZZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8 RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV	3.221
WW 1L1Nu2Q	WWTo1L1Nu2Q_13TeV_amcatnloFXFX_madspin_pythia8 RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV	49.997
Z+Jets	DYJetsToLL_M-50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8 RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV	1921.8*3
QCD multijet μ EnrichedPt5	QCD_Pt-*_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8 RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV	LO $\sigma^*\epsilon_{filter}$ [17]
QCD multijet EM Enriched	QCD_Pt-*_EMEnriched_TuneCUETP8M1_13TeV_pythia8 RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV	LO $\sigma^*\epsilon_{filter}$ [17]