

# CMS Draft Analysis Note

*The content of this note is intended for CMS internal use and distribution only*

2015/02/25

Head Id: 271137

Archive Id: 278646:278657M

Archive Date: 2014/12/10

Archive Tag: trunk

## Search for a BSM resonance decaying to W and Higgs bosons in the $lvb\bar{b}$ final state

Andreas Hinzmann<sup>1</sup>, Benjamin Kilminster<sup>1</sup>, Clemens Lange<sup>1</sup>, Shuai Liu<sup>2</sup>, Qiang Li<sup>2</sup>, Jennifer Ngadiuba<sup>1</sup>, Mengmeng Wang<sup>2</sup>, and Zijun Xu<sup>2</sup>

<sup>1</sup> University of Zurich

<sup>2</sup> Peking University

### Abstract

A search for a massive resonance decaying into W and H bosons in the  $lvb\bar{b}$  final state is presented. Results are based on data corresponding to an integrate luminosity of  $19.7 \text{ fb}^{-1}$  of proton-proton collisions at  $\sqrt{s} = 8 \text{ TeV}$  and collected with the CMS detector at the CERN LHC. For a high-mass ( $\gtrsim 1 \text{ TeV}$ ) resonance, the two quarks coming from Higgs give rise to the presence of one single “merged” jet, which can be tagged through a study of its substructure and flavor. A model-independent statistical interpretation is applied to the selected dataset, setting limits on the cross section of a narrow resonance decaying to WH as a function of its mass.

This box is only visible in draft mode. Please make sure the values below make sense.

PDFAuthor: EXO-WH team

PDFTitle: Search for a BSM resonance decaying to W and Higgs bosons in the  $lvb\bar{b}$  final state

PDFSubject: CMS

PDFKeywords: CMS, physics, software, computing

Please also verify that the abstract does not use any user defined symbols



# Contents

1	1	Introduction . . . . .	2
2	2	Data Sets and Monte Carlo Samples . . . . .	5
3	2.1	Data sets . . . . .	5
4	2.2	Signal Samples . . . . .	5
5	2.3	Background Samples . . . . .	8
6	3	Physics Objects . . . . .	10
7	3.1	Events Cleanup . . . . .	10
8	3.2	Trigger and skim requirements . . . . .	11
9	3.3	Lepton selection . . . . .	12
10	3.4	Jet selection . . . . .	14
11	3.5	$E_T^{\text{miss}}$ selection . . . . .	15
12	4	Reconstruction of the WH resonance . . . . .	16
13	4.1	Leptonic W reconstruction . . . . .	16
14	4.2	Hadronic H reconstruction . . . . .	16
15	5	Event Selection . . . . .	19
16	5.1	b-tagging veto optimization . . . . .	19
17	5.2	Higgs b-tagging optimization . . . . .	19
18	5.3	$t\bar{t}$ background rejection . . . . .	22
19	5.4	Final Selection and Control plots . . . . .	26
20	6	Signal Modeling . . . . .	35
21	6.1	Parametrization of the $m_{WH}$ signal shape . . . . .	40
22	6.2	Signal Efficiency . . . . .	40
23	7	Background estimate . . . . .	44
24	7.1	Strategy . . . . .	44
25	7.2	Alpha Method . . . . .	44
26	7.3	SF Method . . . . .	48
27	7.4	$t\bar{t}$ control sample . . . . .	50
28	8	Systematic Uncertainties . . . . .	54
29	8.1	Background Normalization and Shape . . . . .	54
30	8.2	Signal Normalization . . . . .	54
31	8.3	Signal Shape . . . . .	57
32	9	Results . . . . .	60
33	9.1	Tools for statistical analysis . . . . .	60
34	9.2	Limits - Method 1 . . . . .	61
35	10	Conclusions . . . . .	66
36	A	$\tau_{21}$ optimization . . . . .	67
37	B	$t\bar{t}$ rejection optimization . . . . .	72
38	C	Closure test for alpha method . . . . .	77
39	D	Studies on bad event filters . . . . .	80
40	E	Event displays . . . . .	84
41	F	Asymptotic Limits . . . . .	92

## 1 Introduction

The Standard Model (SM) of particle physics has been very successful in describing high-energy physics phenomena investigated so far [1]. The discovery of the Higgs has completed the Standard Model (SM) as we know it [2, 3]. On the other hand, the SM is expected to be just an effective theory valid at energy scales up to  $\sim 1$  TeV. The light Higgs mass leaves open the questions about the naturalness of the SM and leaves the door open to the presence of new physics to stabilize the Higgs vacuum. For larger scales, the theory predictions violate unitarity unless the constants of the SM are extremely fine tuned. This problem can be reformulated in terms of the large difference between the mass of the Higgs boson ( $M_H \lesssim 1$  TeV/ $c^2$ ) and the Planck scale,  $M_{Pl}$ , where the gravitational force is expected to be of the same size of the other fundamental forces ( $M_{Pl} \sim 10^{16}$  TeV/ $c^2$ ). This is usually called the *hierarchy problem*.

Composite-Higgs models [4–6], provide a direct solution to the hierarchy problem. These models consider the Higgs boson as a pseudo-Nambu-Goldstone boson of a global symmetry and other composite bound states beyond the Higgs are expected to exist and can be experimentally observed. If deviations from the SM are observed for the Higgs couplings in future, precision LHC data, it has been shown that the masses of these resonances can be in the TeV range.

Composite-Higgs models predict many new particles, including additional gauge bosons, e.g. new, heavy  $W'$  boson. For  $W'$  with the SM coupling, the most stringent limits are reported in searches with leptonic final states [7, 8], and the current lower limit on the  $W'$  mass is 2.9 TeV/ $c^2$ . The limit varies by 0.1 TeV/ $c^2$ , depending on the chirality of the  $W'$  couplings. Specific searches in the WZ final state have also been reported [9–11] setting a lower limit of 1.1 TeV/ $c^2$ .

In this analysis, we are interested in  $W' \rightarrow WH$ , which can be crucial for  $W'$  with non-SM coupling. Considering the following rescaling on  $W'$  couplings:

$$W'WH = x \times WWH_{SM}, \quad W'F\bar{F} = y \times W\bar{F}F_{SM}, \quad (1)$$

which is typically true in models like Composite-Higgs, up to  $\mathcal{O}(v^2/f^2)$ , with  $v$  as the Higgs field vacuum expectation value and  $f$  as the composite scale.  $x$  and  $y$  are the rescaling coefficients: in the Little Higgs model [12, 13], one has

$$x = -\cot 2\theta, \quad y = -\cot \theta, \quad (2)$$

while in the Heavy Vector Triplet (HVT) Model [14], one has

$$x = C_H \frac{g_v}{g_w}, \quad y = C_q \frac{g_w}{g_v}, \quad (3)$$

where  $g_w = e/s_W \sim 0.651$ .

We note in scenario B of the HVT model in Ref. [14]:

$$|C_H| \sim |C_q| \sim 1, \quad g_v \geq 3, \quad (4)$$

the dominant BRs are into di-bosons and the fermionic decays are extremely suppressed. In Figure 2 the BRs for the two body decays of the charged,  $W'$ , and neutral resonance,  $Z'$ , are shown as a function of the resonance mass. As shown in the figure, the BRs of  $W'$  and  $Z'$  are expected to be of the same order of magnitude, since the charged and the neutral resonances are practically degenerate. Furthermore,  $W' \rightarrow WH$  becomes dominant and almost degenerate with  $W' \rightarrow WZ$  (as the longitudinal modes of the SM light gauge bosons is essentially from the Higgs doublet [13]), and these parameter points are not constrained so far from experiments (see Fig 3.2 in Ref. [14]).

The signal studied in this analysis is the production of a resonance with mass above 0.8 TeV, decaying to WH where Higgs decays to two b-quarks and the W-boson decays to lepton and neutrino (see Fig. 1), with no assumption on additional particles produced in the final state. It is assumed that the resonance is narrow, i.e. that its natural width is much smaller than the experimental resolution.

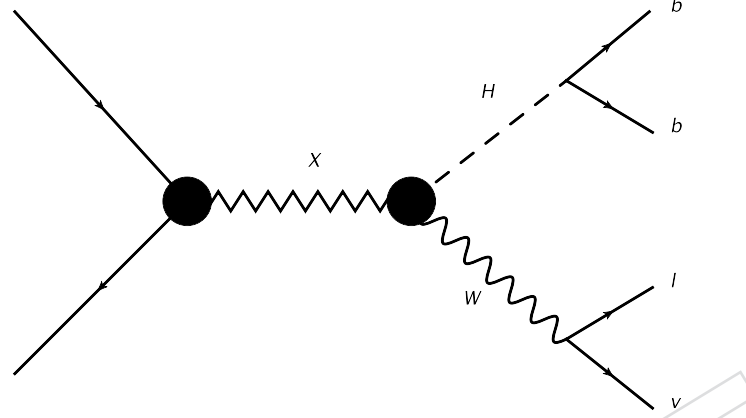


Figure 1: Production of a resonance decaying to WH

In the high mass regime, the two quarks are separated by a small angle in space, resulting in the presence of one single merged Jet after hadronization. This Jet that can be tagged as coming from one massive boson, through the study of its substructure and subjet b-tag, using techniques already used in other CMS analyses ([15–17], [18, 19]).

This note is organized as follows: Section 2 shows the data and MC samples, together with the triggers used in the analysis; in Section 3 the objects reconstruction with the boosted H-tag and subjet b-tag techniques are shown, further with the reconstruction of WH resonance mass in Section 4; in Section 5, event selection criteria and data/MC comparison plots in signal free region are shown; then we describe the background estimation methods in Section 7 and the systematic errors in Section 8. Finally the results are shown in Section 9, with brief discussions in Section 10.

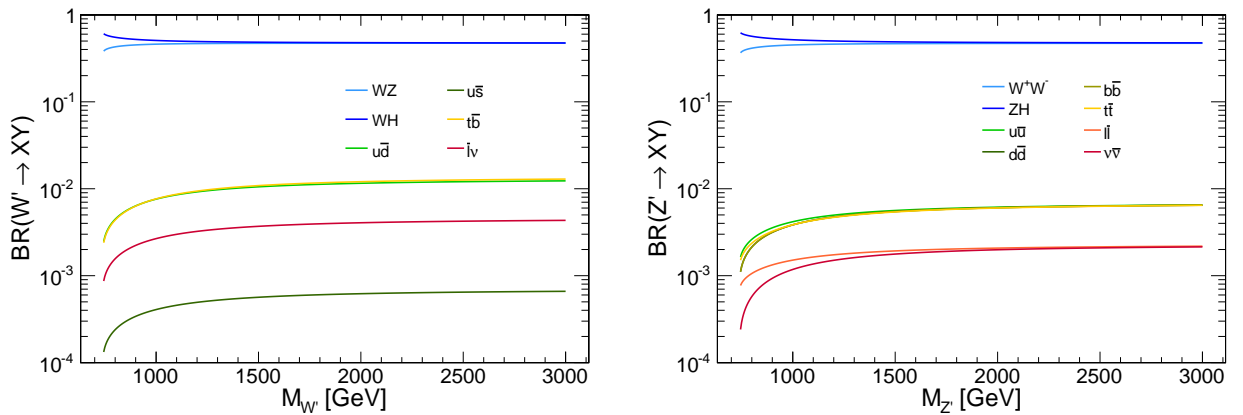


Figure 2: Branching Ratios for the two body decays of the charged  $W'$  (left) and of the neutral resonance  $Z'$  (right). The charged and the neutral resonances are practically degenerate.

## 2 Data Sets and Monte Carlo Samples

### 2.1 Data sets

We use about  $19.7 \text{ fb}^{-1}$  of data collected in 2012 and re-reco'ed with CMSSW\_5.3.X. The data samples are listed in Table 1. We use only lumisections that have been declared good for analysis by the central certification team, listed in the following official JSON file:

- Cert\_190456–193621.8TeV.22Jan2013ReReco\_Collisions12.JSON.txt – for Run2012A re-reconstruction.
- Cert\_193833–196531.8TeV.22Jan2013ReReco\_Collisions12.JSON.txt – for Run2012B re-reconstruction.
- Cert\_198022–203742.8TeV.22Jan2013ReReco\_Collisions12.JSON.txt – for Run2012C re-reconstruction.
- Cert\_203777–208686.8TeV.22Jan2013ReReco\_Collisions12.JSON.txt – for Run2012D re-reconstruction.

Table 1: Data samples used in the analysis.

Sample	Run Range	$\mathcal{L} [\text{pb}]^{-1}$
/SingleMu/Run2012A-22Jan2013-v1/AOD	190456–193621	889.362
/SingleMu/Run2012B-22Jan2013-v1/AOD	193833–196531	4424
/SingleMu/Run2012C-22Jan2013-v1/AOD	198022–203742	7144
/SingleMu/Run2012D-22Jan2013-v1/AOD	203777–208686	7307
TOTAL SingleMu	190456–208686	19764
/SingleElectron/Run2012A-22Jan2013-v1/AOD	190456–193621	889.362
/SingleElectron/Run2012B-22Jan2013-v1/AOD	193833–196531	4422
/SingleElectron/Run2012C-22Jan2013-v1/AOD	198022–203742	7080
/SingleElectron/Run2012D-22Jan2013-v1/AOD	203777–208686	7314
TOTAL SingleElectron	190456–208686	19705

### 2.2 Signal Samples

The signal hypothesis has been simulated with a  $W'$  boson produced via quark-antiquark interaction and decaying into  $W$  and Higgs bosons in the  $lvq\bar{q}$  final state with  $q = b, c, g$  and  $l = e, \mu, \tau$ . Les Houches Event (LHE) files were produced at parton level with MadGraph version 5.1.5.12 [20] considering a representative set of couplings and masses as described in [21]. A set of simulated samples were privately produced using CMSSW version 5.3.2\_patch4, following the conditions used for the official Summer12 campaign. Showering and hadronization were performed using PYTHIA 6.426 with the tune Z2star configuration. The full simulation was performed in three-steps, GEN-SIM, DIGI-RAW and RECO, and GEANT4 [22] was used for the simulation of the CMS detector. In the DIGI-RAW step, simulated minimum bias samples from the dataset /MinBias\_TuneZ2star\_8TeV-pythia6/Summer12-START50\_V13-v3/GEN-SIM have been mixed to signal events, in order to simulate effects of pile-up (PU), according to the distribution of number of PU events defined in the 2012\_Summer\_50ns\_PoissonOOTPU configuration. This emulation of the official full simulation was validated comparing the private

117 production with samples from the official Summer12 campaign in two previously approved  
118 analyses [23, 24]. The generated samples are summarized in Table 2.

DRAFT

Table 2: Signal samples used in the analysis.

Sample	Number of events	$M_{W'}$ [TeV]
/EXOWH.Wprime.M800_GENSIM.V2/mwang-EXOWH.Wprime.M800 _AODSIM.V2-2c74483358b1f8805e5601fc325d256c/USER	19600	0.8
/EXOWH.Wprime.M900_GENSIM.V2/mwang-EXOWH.Wprime.M900 _AODSIM.V2-2c74483358b1f8805e5601fc325d256c/USER	19999	0.9
/EXOWH.Wprime.M1000_GENSIM.V2/mwang-EXOWH.Wprime.M1000 _AODSIM.V2-2c74483358b1f8805e5601fc325d256c/USER	12687	1.0
/EXOWH.Wprime.M1100_GENSIM.V2/mwang-EXOWH.Wprime.M1100 _AODSIM.V2-2c74483358b1f8805e5601fc325d256c/USER	18815	1.1
/EXOWH.Wprime.M1200_GENSIM.V2/mwang-EXOWH.Wprime.M1200 _AODSIM.V2-2c74483358b1f8805e5601fc325d256c/USER	19998	1.2
/EXOWH.Wprime.M1300_GENSIM.V2/mwang-EXOWH.Wprime.M1300 _AODSIM.V2-2c74483358b1f8805e5601fc325d256c/USER	19997	1.3
/EXOWH.Wprime.M1400_GENSIM.V2/mwang-EXOWH.Wprime.M1400 _AODSIM.V2-2c74483358b1f8805e5601fc325d256c/USER	19999	1.4
/EXOWH.Wprime.M1500_GENSIM.V2/mwang-EXOWH.Wprime.M1500 _AODSIM.V2-2c74483358b1f8805e5601fc325d256c/USER	19600	1.5
/EXOWH.Wprime.M1600_GENSIM.V2/mwang-EXOWH.Wprime.M1600 _AODSIM.V2-2c74483358b1f8805e5601fc325d256c/USER	19997	1.6
/EXOWH.Wprime.M1700_GENSIM.V2/mwang-EXOWH.Wprime.M1700 _AODSIM.V2-2c74483358b1f8805e5601fc325d256c/USER	19201	1.7
/EXOWH.Wprime.M1800_GENSIM.V2/mwang-EXOWH.Wprime.M1800 _AODSIM.V2-2c74483358b1f8805e5601fc325d256c/USER	14516	1.8
/EXOWH.Wprime.M1900_GENSIM.V2/mwang-EXOWH.Wprime.M1900 _AODSIM.V2-2c74483358b1f8805e5601fc325d256c/USER	19995	1.9
/EXOWH.Wprime.M2000_GENSIM.V2/mwang-EXOWH.Wprime.M2000 _AODSIM.V2-2c74483358b1f8805e5601fc325d256c/USER	19152	2.0
/EXOWH.Wprime.M2100_GENSIM.V2/mwang-EXOWH.Wprime.M2100 _AODSIM.V2-2c74483358b1f8805e5601fc325d256c/USER	19999	2.1
/EXOWH.Wprime.M2200_GENSIM.V2/mwang-EXOWH.Wprime.M2200 _AODSIM.V2-2c74483358b1f8805e5601fc325d256c/USER	16803	2.2
/EXOWH.Wprime.M2300_GENSIM.V2/mwang-EXOWH.Wprime.M2300 _AODSIM.V2-2c74483358b1f8805e5601fc325d256c/USER	17200	2.3
/EXOWH.Wprime.M2400_GENSIM.V2/mwang-EXOWH.Wprime.M2400 _AODSIM.V2-2c74483358b1f8805e5601fc325d256c/USER	19600	2.4
/EXOWH.Wprime.M2500_GENSIM.V2/mwang-EXOWH.Wprime.M2500 _AODSIM.V2-2c74483358b1f8805e5601fc325d256c/USER	19998	2.5
/EXOWH.Wprime.M2600_GENSIM.V2/mwang-EXOWH.Wprime.M2600 _AODSIM.V2-2c74483358b1f8805e5601fc325d256c/USER	14412	2.6
/EXOWH.Wprime.M2700_GENSIM.V2/mwang-EXOWH.Wprime.M2700 _AODSIM.V2-2c74483358b1f8805e5601fc325d256c/USER	19998	2.7
/EXOWH.Wprime.M2800_GENSIM.V2/mwang-EXOWH.Wprime.M2800 _AODSIM.V2-2c74483358b1f8805e5601fc325d256c/USER	18401	2.8
/EXOWH.Wprime.M2900_GENSIM.V2/mwang-EXOWH.Wprime.M2900 _AODSIM.V2-2c74483358b1f8805e5601fc325d256c/USER	19201	2.9
/EXOWH.Wprime.M3000_GENSIM.V2/mwang-EXOWH.Wprime.M3000 _AODSIM.V2-2c74483358b1f8805e5601fc325d256c/USER	13600	3.0



## 2.3 Background Samples

The dominant background in the  $W' \rightarrow WH \rightarrow \ell\nu 2b$  analysis is the inclusive  $W$  production with jets ( $W$ +jets). Other backgrounds considered are inclusive  $t\bar{t}$  production, standard model diboson production ( $WW$ ,  $WZ$  and  $ZZ$ ), single- $t$  production, and the inclusive Drell-Yan production with jets ( $Z/\gamma/Z^*$ +jets). All the background samples considered were produced in the frame of the Summer12 production.

The  $W$ +jets and DY+jets samples that we use have been produced with MadGraph. The  $t\bar{t}$ +jets and single- $t$  samples have been produced with POWHEG. The diboson samples were produced with Pythia6. A full list of the samples is found in Table 3. The cross sections are taken from <https://twiki.cern.ch/twiki/bin/view/>

CMS/StandardModelCrossSectionsat8TeV: the cross sections for  $W$ +jets and DY+jets samples (generated in  $p_T$  bins) are obtained by rescaling the corresponding LO cross sections by a flat NNLO/LO  $k$ -factor derived from the inclusive production; the cross sections for the remaining processes are either NLO or approx. NNLO calculations.

DRAFT

Table 3: Background samples used in the analysis.

Sample	Number of events	$\sigma$ [pb]
MadGraph samples		
/WJetsToLNu_PtW-100_TuneZ2star_8TeV-madgraph/ Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	12.1M	282.5
/WJetsToLNu_PtW-180_TuneZ2star_8TeV-madgraph-tarball/ Summer12_DR53X-PU_S10_START53_V7C-v1/AODSIM	9.73M	29.00
/DYJetsToLL_PtZ-100_TuneZ2star_8TeV_ext-madgraph-tarball/ Summer12_DR53X-PU_S10_START53_V7C-v1/AODSIM	2.5M	39.1
Powheg samples		
/TT_CT10_TuneZ2star_8TeV-powheg-tauola/ Summer12_DR53X-PU_S10_START53_V7A-v2/AODSIM	21.7M	225.2
/T_s-channel_TuneZ2star_8TeV-powheg-tauola/ Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	0.209M	3.79
/T_t-channel_TuneZ2star_8TeV-powheg-tauola/ Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	3.75M	56.4
/T_tW-channel-DR_TuneZ2star_8TeV-powheg-tauola/ Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	0.498M	11.1
/Tbar_s-channel_TuneZ2star_8TeV-powheg-tauola/ Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	0.140M	1.76
/Tbar_t-channel_TuneZ2star_8TeV-powheg-tauola/ Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	1.90M	30.7
/Tbar_tW-channel-DR_TuneZ2star_8TeV-powheg-tauola/ Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	0.493M	11.1
Pythia samples		
/WW_TuneZ2star_8TeV_pythia6_tauola/ Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	3.87M	57.1
/WZ_TuneZ2star_8TeV_pythia6_tauola/ Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	1.91M	33.2
/ZZ_TuneZ2star_8TeV_pythia6_tauola/ Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	0.486M	8.06

### 3 Physics Objects

Events are processed with PAT in CMSSW\_5\_3\_9 and CMGTools version V5\_10\_0.

The analysis relies on the standard reconstruction algorithms produced by the CMS community. Event data is reconstructed using the particle-flow (PF) reconstruction technique [25, 26]. Particle flow attempts to reconstruct all stable particles in an event by combining information from all sub-detectors. The algorithm categorizes all particles into the following five types: muons, electrons, photons, charged and neutral hadrons. The list of reconstructed particles is used as the set of inputs for a jet clustering algorithm to create particle-flow jets.

We describe in the next sections the reconstruction of muons, electrons, jets,  $E_T^{\text{miss}}$  resonances. The general strategy is to keep the selection requirements as loose as possible, in order not to bias the selected sample towards a specific BSM model.

#### 3.1 Events Cleanup

##### 3.1.1 Noise Filters

Events are required to pass the following set of standard noise filters:

- `primaryVertexFilterPath`
- `noscrapingFilterPath`
- `hcalLaserEventFilterPath`
- `HBHENoiseFilterPath`
- `trackingFailureFilterPath`
- `CSCTightHaloFilterPath`
- `eeBadScFilterPath`
- `EcalDeadCellTriggerPrimitiveFilterPath`
- `EcalLaserCorrectionFilterPath`
- `tobtecfalesfilterPath`

The `primaryVertexFilterPath` filter requires the presence of at least one reconstructed vertex, passing the following requirements:

- obtained from a fit of a suitable set of tracks, i.e., `!isFake`
- number of degrees of freedom larger than 4
- absolute value of the longitudinal coordinate `abs(z)` smaller than 24 cm
- transversal distance to the longitudinal axis `ρ` smaller than 2 cm.

In case more than one suitable primary vertex is found, the vertex with the highest tracks sum- $p_T^2$  is taken as the primary vertex of the event. This vertex serves as the reference to identify the tracks associated to the event that fired the trigger and it is used for PU removal in jet clustering.

The `tobtecfalesfilterPath` removes events characterized by large amounts of fake tracks from iteration 6 (TOB-TEC seeding) in the transition region between the barrel and end-cap regions of the detector ( $0.9 < |\eta| < 1.6$ ). In particular, it rejects the event if the ratio between the TOBTEC-seeded tracks and pixel-seeded tracks is larger than 3. The efficiency of this filter as a function of the H-bb jet  $p_T$  and  $\eta$  has been studied in details as reported in Appendix D.

### 3.1.2 Pileup removal

The presence of additional interactions with respect to the primary one, known as Pile-up (PU), is expected to affect this analysis in the following ways:

- additional energy from PU gets added to the jets from the main interaction, modifying the energy and the position of the reconstructed jet;
- additional low  $p_T$  jets composed of PU energy get added to the event.
- tracks and calorimetric towers from PU energy deposits get added to the isolation energy sum of the leptons, thus making isolation cuts less efficient.

We borrow Fig. 3 from [27] to show the inconsistency of PU interaction between data and MC samples. In the following this distribution in MC has been re-weighted to match the data. In Fig. 4 we show consistency in the number of reconstructed vertex between data and MC after MC samples have been re-weighted.

Various algorithms are available to correct for PU effects. PF algorithms can be used to decrease the effects of pileup. The charged particles coming from PU can also be removed before the jet clustering requiring that all the tracks come from the primary vertex with an algorithm called “Charged Hadron Subtraction” (CHS). The so-called Fastjet and L1-offset corrections remove the additional energy released in the event from PU interactions.

In the rest of the analysis the effect of the PU on the jets is corrected using both the CHS and the Fastjet algorithms. Following the recommendations of the MUO and EGM POGs, no attempt is done to correct the electrons and muons for PU for the dedicated high- $p_T$  selection employed in this search. The final systematics due to PU on the signal acceptance is discussed in Section 8.

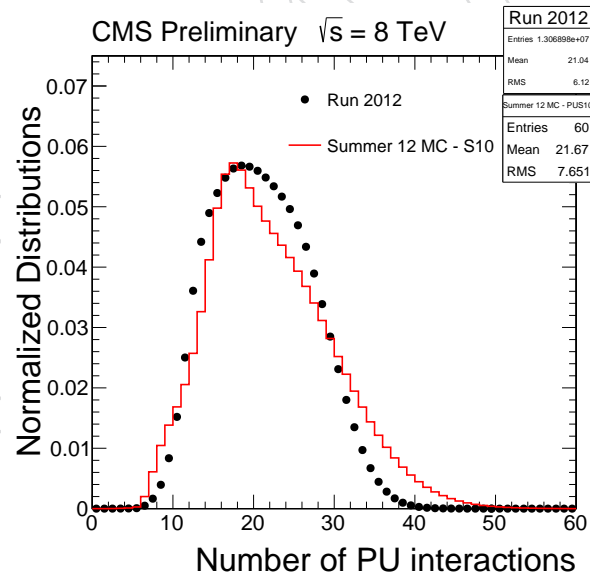


Figure 3: Number of PU interactions in data and MC. In the following the MC sample has been reweighted to match the data.

### 3.2 Trigger and skim requirements

The results presented in this note are based on the SingleMu, and SingleElectron datasets. Each of these datasets contain at least one un-prescaled single lepton trigger with looser  $p_T$  requirements than our offline selections. The lowest threshold un-prescaled, non-isolated triggers are used:

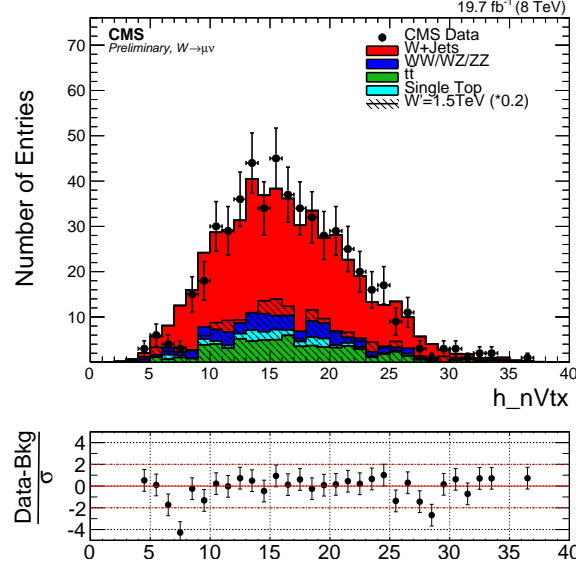


Figure 4: Number of reconstructed vertices after the reweighting of the MC sample by the pile-up corrections. Points with error bars show data, histograms show contribution of dominant background channels, both after the selection.

• HLT\_Mu40\_eta2p1\_\*

for the SingleMu dataset, and

• HLT\_Ele80\_CaloIdVT\_TrkIdT\_\* (active for 5% of the data period)

• HLT\_Ele80\_CaloIdVT\_GsfTrkIdT\_\* (active for 95% of the data period)

for the SingleElectron dataset.

The trigger selection is not applied in the MC. Instead the trigger efficiency is computed from data with a Tag-and-Probe technique and the Monte Carlo samples are corrected accordingly by using the trigger efficiency as an event weight. The efficiency for the single muon trigger with respect of the offline lepton selection employed in this analysis has been taken from the muon POG twiki page <https://twiki.cern.ch/twiki/bin/viewauth/CMS/MuonReferenceEffs> using the file `SingleMuonTriggerEfficiencies_eta2p1_Run2012ABCD_v5trees.root`:  $0.9290 \pm 0.0002$  ( $0.0 < |\eta| < 0.9$ ),  $0.8314 \pm 0.0006$  ( $0.9 < |\eta| < 1.2$ ), and  $0.8027 \pm 0.0005$  ( $1.2 < |\eta| < 2.1$ ) depending on the muon  $\eta$ . While the efficiency of the single electron trigger has been studied in detail in the context of the search for  $W'$  in single lepton + MET final state[28]:  $0.991$  ( $0.0 < |\eta| < 1.4442$ ) and  $0.976$  ( $1.566 < |\eta| < 2.5$ )<sup>1</sup> depending on the electron  $\eta$ .

### 3.3 Lepton selection

#### 3.3.1 Muon Selection

For muon identification and their momentum assignment we follow the recommendation of the MUON POG at

<https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideMuonId>.

The new versions (after CMSSW\_5.3.6\_patch1) of the HighPT muon ID and the TuneP momentum assignment are used, which translate into the following requirements:

<sup>1</sup>The EB-EE transition region  $1.4442 < |\eta| < 1.566$  is excluded from the analysis

- muon identified as `GlobalMuon`;
- at least one muon chamber hit included in the global-muon track fit;
- muon segments in at least two muon stations;
- matched tracker track has transverse impact parameter  $d_{xy} < 2$  mm w.r.t. the primary vertex (calculated using dB method);
- analogously, longitudinal impact parameter  $d_z < 5$  mm;
- number of pixel hits larger than zero;
- number of tracker layers with hits  $> 5$ .

The track used to obtain the muon momentum needs to pass  $dp_T/p_T < 0.3$ .

On top of the identification cuts, muon candidates are requested to pass a relative isolation cut,  $\frac{\sum p_T^{\text{tracks}}}{p_T^{\text{muon}}} < 0.1$ , where  $\sum p_T^{\text{tracks}}$  is the scalar sum of tracks'  $p_T$  in a cone of  $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.3$  around the muon direction, and  $p_T^{\text{muon}}$  is the transverse momentum of the muon itself.

Tight muon candidates have to pass the HighPT muon ID and isolation described above and need to have  $p_T > 50$  GeV and  $|\eta| < 2.1$  (to be in the plateau of the trigger efficiency). Loose muon candidates (used for the lepton veto selection) are required to pass the same ID and isolation with less stringent kinematic cuts of  $p_T > 20$  GeV and  $|\eta| < 2.4$ .

The muon ID+ISO scale factors have been taken from the muon POG twiki page <https://twiki.cern.ch/twiki/bin/viewauth/CMS/MuonReferenceEffs> using the files `MuonEfficiencies.Run2012ReReco_53X.root` and `MuonEfficiencies.ISO.Run2012ReReco_53X.root`:

- the scale factor for the muon ID is  $0.98996 \pm 0.00032$  ( $0.0 < |\eta| < 0.9$ ),  $0.9923 \pm 0.0006$  ( $0.9 < |\eta| < 1.2$ ) and  $0.9949 \pm 0.0004$  ( $1.2 < |\eta| < 2.1$ ) depending on the muon  $\eta$ ;
- the scale factor for the muon ISO is  $0.99965 \pm 0.00008$  ( $0.0 < |\eta| < 0.9$ ),  $0.9994 \pm 0.0001$  ( $0.9 < |\eta| < 1.2$ ) and  $0.99968 \pm 0.00009$  ( $1.2 < |\eta| < 2.1$ ) depending on the muon  $\eta$ ;

### 3.3.2 Electron Selection

For the electron identification and reconstruction of electron  $p_T$  assignment we follow the recommendation of the EGM POG at <https://twiki.cern.ch/twiki/bin/viewauth/CMS/HEEPElectronID>.

The latest version of the HEEP selection (v4.1) is used:

- supercluster pseudorapidity ( $\eta_{SC}$ ) smaller than 1.442 (for barrel electrons) or in the range 1.56–2.5 (for endcap electrons);
- driven by calorimetric deposits (`isEcalDriven`);
- have  $|\Delta\eta_{in}|$  smaller than 0.005 (0.007) for barrel electrons (endcap electrons);
- have  $|\Delta\phi_{in}|$  smaller than 0.06 (both for barrel and endcap electrons);
- have the ratio of hadronic to electromagnetic energy smaller than 0.05;
- have  $\sigma_{i\eta i\eta}$  smaller than 0.03 (for endcap electrons only);
- have either  $E_{2\times 5}/E_{5\times 5}$  larger than 0.94 or  $E_{1\times 5}/E_{5\times 5}$  larger than 0.83 (for barrel electrons only);
- have the number of inner layer lost hits smaller than 2;

- have  $|d_{xy}|$  smaller than 0.02 (0.05) cm for barrel electrons (endcap electrons).

Further, the electron candidates are requested to pass the following detector-based isolation cuts (isolation cone  $\Delta R = 0.3$ ):

- EM + Had Depth 1 Isolation  $< 2 + 0.03 \cdot E_t + 0.28 \cdot \rho$  for the barrel electrons;
- EM + Had Depth 1 Isolation  $< 2.5 + 0.28 \cdot \rho$  for  $E_t < 50$ , else  $< 2.5 + 0.03 \cdot (E_t - 50) + 0.28 \cdot \rho$  for the endcap electrons;
- Trk Isolation (sum of tracks  $p_T$ )  $< 5$  GeV for both barrel and endcap electrons.

We select `Tight` electron candidates passing the HEEP 4.1 electron ID+ISO with  $p_T > 90$  GeV and  $0.0 < |\eta| < 1.442$  OR  $1.56 < |\eta| < 2.5$  (to be in the plateau of the trigger efficiency). `Loose` electron candidates (used for the lepton veto selection) are required to pass the HEEP 4.1 electron ID+ISO with  $p_T > 35$  GeV (this  $p_T$  cut is included by default in the HEEP 4.1 selection) and the same  $\eta$  range of the tight electrons.

The electron ID+ISO efficiencies have been measured in data and MC [28]. Data/MC scale factors of  $\approx 0.98$  are used as event weights to correct the MC.

### 3.4 Jet selection

Jets are reconstructed from the list of Particle Flow (PF) candidates [26] reconstructed in the event. Charged hadrons originating from vertices other than the primary vertex were not used in the jet clustering with the Charged Hadron Subtraction (CHS) procedure. We use two jet collections:

- Cambridge-Aachen algorithm with distance parameter  $R = 0.8$  (CA8) to select Higgs candidates decaying to jets in boosted regime;
- *anti* -  $k_T$  algorithm with distance parameter  $R = 0.5$  (AK5) to require or veto the presence of b-tagged jets in the event.

#### 3.4.1 CA8 Jet Selection

The selected PF candidates are clustered using a Cambridge-Aachen algorithm with distance parameter  $R = 0.8$ . The CA8 jet energy is corrected applying the `L1L2L3` (`L1L2L3+Residual`) corrections for MC (data), derived for AK7 jets by the JetMET group [29]. This procedure was already validated in other analysis using wide jets [30, 31]. The global tag used for data (MC) is `FT_53_V21_AN4` (`START53_V23`).

We select jet candidates with  $p_T > 80$  GeV and  $|\eta| < 2.4$ . In addition, jets are subject to the following loose identification criteria ( $> 99\%$  efficiency), `JetID Loose`:

- muon energy fraction smaller than 0.99
- photon energy fraction smaller than 0.99
- charged electromagnetic energy fraction smaller than 0.99
- neutral hadron energy fraction smaller than 0.99
- charged hadron energy fraction larger than 0
- number of constituent particles larger than 1

If a CA8 jet is within  $\Delta R < 1.0$  of any tight electron or tight muon defined in the previous sections, the jet is not used in the analysis.

Furthermore jets are rejected if their pseudorapidity falls in the region  $1 < |\eta| < 1.8$ , corresponding to the barrel-endcap transition region of the silicon tracker. The goal of this require-



ment is to remove jets where the tracking software behaves in an anomalous way, and many fake tracks are associated to the jet. The choice for this fiducial cut is explained in details in Appendix D.

### 3.4.2 AK5 Jet Selection

The selected PF candidates are clustered using an *anti* -  $k_T$  algorithm with distance parameter  $R = 0.5$ . The jet energy is corrected applying the L1L2L3 (L1L2L3+Residual) corrections for MC (data), derived for the same algorithm by the JetMET group [29]. The global tag used for data (MC) is FT\_53\_V21\_AN4 (START53\_V23). We select jet candidates with  $p_T > 30$  GeV and  $|\eta| < 2.4$ , passing the JetID Loose criteria described above. If an AK5 jet is within  $\Delta R < 0.3$  of any tight electron or tight muon defined in the previous sections, the jet is not used in the analysis.

### 3.5 $E_T^{\text{miss}}$ selection

The missing transverse energy,  $E_T^{\text{miss}}$ , of the event is computed as the negative vector sum of all particle flow objects' transverse momenta.

We follow the recommendations of the JetMET POG reported at <https://twiki.cern.ch/twiki/bin/view/CMS/EXOMissingET2012> and <https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookMetAnalysis>.

Corrections are applied to the  $E_T^{\text{miss}}$  to account for jet energy corrections (Type-1). The minimum  $p_T$  of the jets used in the correction is 10 GeV; jets with EM fraction greater than 90% or identified as muons are not used.

A  $\phi$  modulation of reconstructed  $E_T^{\text{miss}}$  is observed in Data as well as in Monte Carlo simulated events. The modulation is due to a systematic shift of the  $E_T^{\text{miss}}$  x/y components which increases linearly with the number of reconstructed vertices. The magnitude of the shift is different in Data and MC though. A correction for this  $\phi$  modulation has been developed and it is applied in this analysis (MET x/y Shift Correction)

An initial requirement  $E_T^{\text{miss}} > 40$  GeV is applied in the analysis.



## 4 Reconstruction of the WH resonance

### 4.1 Leptonic W reconstruction

$W \rightarrow \ell \nu$  candidates are formed from all the charged leptons passing the selection requirements (electrons or muons) and the  $E_T^{\text{miss}}$  reconstructed in the event. We assume that the  $x$  and  $y$  components of the corrected missing transverse energy are entirely due to the escaping neutrino. The longitudinal component of the neutrino momentum,  $p_{z,\nu}$ , is calculated following the method used previously for the reconstruction of the invariant mass of the top quark as described in [32]. The method consist of solving the second-order equation that makes use of the known W mass:

$$(E_\ell + \sqrt{E_T^{\text{miss}2} + p_{z,\nu}^2})^2 - (\mathbf{p}_{T,\ell} + \mathbf{E}_T^{\text{miss}})^2 - (p_{z,\ell} + p_{z,\nu})^2 = M_W^2 = (80.4)^2 \quad (5)$$

This equation has in general two solutions:

$$p_{z,\nu}^{A,B} = \frac{\mu \cdot p_{z,\ell}}{\mathbf{p}_{T,\ell}^2} \pm \sqrt{\frac{\mu^2 \cdot p_{z,\ell}^2}{\mathbf{p}_{T,\ell}^4} - \frac{E_\ell^2 \cdot E_T^{\text{miss}2} - \mu^2}{\mathbf{p}_{T,\ell}^2}} \quad (6)$$

with

$$\mu = \frac{M_W^2}{2} + \mathbf{p}_{T,\ell} \cdot \mathbf{E}_T^{\text{miss}} \quad (7)$$

If the discriminant of equation 6 is positive, two real solutions are obtained. In this case, the solution with the smallest value of  $|p_{z,\nu}|$  is chosen. If the discriminant in equation 6 becomes negative, or equivalently the transverse mass  $M_T$  is larger than the W-boson pole mass of  $80.4 \text{ GeV}/c^2$  used in the constraint, the solutions have an imaginary part. The method used here eliminates the imaginary component by modifying the components of the missing transverse energy such to give  $M_T = M_W$ , still respecting equation 5.

### 4.2 Hadronic H reconstruction

In the  $X \rightarrow WH$  events, the Lorentz factor  $\gamma$  (boost factor) of the H boson is approximately  $M_X/2M_H$  [33]. Due to the kinematics of the event, a large boost factor implies that the decay products will be predominantly emitted with a small  $\Delta R \approx 2M_H/p_T^H$  [34].

This analysis searches for X resonances with mass greater than  $\approx 0.8 \text{ TeV}$ , for which the boosted topology is by far the dominant one. The  $H \rightarrow b\bar{b}$  candidate is therefore reconstructed as a single massive CA8 jet. The pruned jet mass (jet substructure reconstructed quantity) is examined to discriminate against standard QCD jets.

Besides pruned jet mass, the presence of b-quarks in the final state is exploited in order to further suppress the background from light quark jets, by means of fat jet or subjet b-tagging (see Section 4.2.3).

In addition to the pruned jet mass, the N-subjettiness, showed in past analysis good performance for V-tagging. An optimization of this variable was performed for this case as well, showing that a cut on the  $\tau_{21}$  does not result in any additional sensitivity after b-tagging. Because of this consideration we decided to make use of pruned jet mass and b-tagging only. For more details about N-subjettiness and the optimization see Appendix A.

### 4.2.1 Jet grooming: pruning

*Jet grooming* techniques improve the resolution on the jet mass and help in reducing the effect of pile-up. The goal of jet grooming is to define a procedure for reclustering the jet constituents eliminating those originating from soft, large-angle QCD radiation or from other QCD processes different from the hard one where the H boson was created. Different jet grooming algorithms have been implemented in the CMSSW software and their performance on jets in QCD processes has been studied in detail in [35].

In this analysis, we chose to use the *jet pruning* algorithm because of its good capability in improving the jet mass resolution, the stability in presence of pileup interactions, and because of the better operational experience, thanks to the many analysis that already used it at CMS [30, 31, 35, 36].

The jet pruning algorithm [37] reclusters the jet constituents removing at the same time soft or large angle recombinations. The jet pruning algorithm that we use is based on the CA clustering algorithm with distance parameter,  $R = 0.8$ . In the reclustering procedure, recombinations of jet constituents characterized by soft energies are vetoed through a cut on the quantity  $z_{ij}$

$$z_{ij} = \frac{\min(p_{T,i}, p_{T,j})}{p_{T,(i+j)}} < z_{cut} \quad (8)$$

where the indexes  $i$  and  $j$  refer to the two jet constituents considered in the recombination step and  $z_{cut}$  is a parameter of the pruning set to  $z_{cut} = 0.1$ . A second requirement is applied on the  $\Delta R$  between the two constituents,  $\Delta R_{ij}$

$$\Delta R_{i,j} > D_{cut} = \kappa_{pr} \times \frac{m}{p_T} \quad (9)$$

In this case,  $m$  and  $p_T$  refer to the original ungroomed jet and the parameter  $\kappa_{pr} = 0.5$ . If any of the conditions  $z_{ij} < z_{cut}$  or  $\Delta R_{i,j} > D_{cut}$ , is satisfied, the softest of the two branches in the recombination is taken out of the list of the jet constituents. Otherwise, the two constituents are reHiggs according to the normal CA procedure.

The resulting pruned jet mass ( $m_{jet}^{pruned}$ ) is used for H-tagging, since it exhibits better resolution and smaller biases with respect to the ungroomed (standard) jet mass.

The modelling of the signal efficiency for the pruned jet mass cut is cross-checked by measuring the pruned jet mass scale and resolution using merged  $W \rightarrow \bar{q}q'$  decays in  $\ell$ +jets  $t\bar{t}$  events [38]. The data events are compared to simulated  $t\bar{t}$  events, generated with MADGRAPH, interfaced to PYTHIA for parton showering. The pruned jet mass scale and resolution agree within 1.7% and 11% respectively, which lead to an uncertainty of 1.9% in the efficiency of the pruned jet mass cut. In addition, the impact of the difference between b-fragmentation in H to  $b\bar{b}$  decays and light-quark-fragmentation in W to  $q\bar{q}$  events was checked by comparing the boosted top mass scale, which contains one  $q$  quark in AN-2013-029.v9 Fig 31c and agrees within uncertainties. The difference between light quark and bottom quark fragmentation is taken into account through the jet energy scale uncertainty on the pruned jet mass cut efficiency of 1-5%, amounting to a total of 2.1-5.3% uncertainty on the pruned jet mass cut efficiency.

For reconstructing the kinematics of the  $X \rightarrow WH$  resonance we use still the kinematics of the original jet, however. In this way we still use properly calibrated jets for reconstructing the

398 kinematics of the event <sup>2</sup>.

#### 399 4.2.2 Fat jet and subjet b-tagging

400 The final state includes two b (anti-)quarks in this analysis, while because of boosted reason the  
 401 two b quarks merging into one fat jet. That means we can apply a cut on the b-tagging of the  
 402 subjets to reduce the background. Considering the jet track sharing, we used a Higgs b-tagging  
 403 cut in this analysis. More details are provided in Sec. 5.2.

#### 404 4.2.3 Selection criteria

405 At this stage of the analysis,  $H \rightarrow b\bar{b}$  candidates are simply identified by CA8 jets passing the se-  
 406 lection criteria described in Section 3.4.1. No requirements on  $m_{jet}^{pruned}$ , or b-tagging are applied  
 407 here. The final cuts on the pruned jet mass, and  $b - tagging$  are optimized at the stage of the  
 408 final event selection. The pruned jet mass is also used to define control regions i) to cross check  
 409 the MC description of the data and ii) for the data-driven background estimation as described  
 410 later.

DRAFT

---

<sup>2</sup>Jet energy scale corrections for pruned jets are not yet available in CMS.

## 5 Event Selection

### 5.1 b-tagging veto optimization

Each event passing the final selection is required to have no AK5 jets far from the Higgs hadronic candidate in the event by  $\Delta R = 0.8$  tagged as b-jet accordingly with the CSV algorithm, medium working point (CSVM). The reason for this choice is discussed below.

We considered different options for the b-tagging veto using the CSV algorithm:

- different working points (loose L, medium M, tight T);
- different jet selection: either all the selected AK5, or only those that are far from the Higgs hadronic candidate in the event by  $\Delta R = 0.8$  (the latter option is identified by the label “clean” in the following plots).

We choose as a figure of merit the quantity suggested in Ref. [39] (“Punzi’s significance”):

$$P = \frac{\epsilon_S}{1 + \sqrt{B}} \quad (10)$$

where  $\epsilon_S$  is the signal selection efficiency and  $B$  is the remaining background after all the cuts have been applied. The signal and background are evaluated in a region of  $\pm 15\%$  around the target resonance mass. Since the presence of bjets in the event is independent on the  $\tau_{21}$  value of the leading jet, we don’t need to consider  $\tau_{21}$  cut in this study. In fact, from appendix A we can see that  $\tau_{21}$  cut is not a key variable anymore for this analysis and it does not show any degree of separation in between signal and background.

The significance calculated for the different b-tagging veto options and for different resonance masses (1, 1.5, 2 and 2.5 TeV) is reported in Figures 5, from which we can see that for low mass point the CSV Loose working point gives the best Punzi’s significance, while with the increase of resonance mass (at least from 1.5 TeV point) the CSV Medium working point gives the best significance. Considering our signal samples cover from 0.8 TeV to 3 TeV, we choose CSV Medium as the b-tagging veto cut, which gives the best significance performance for most of signal samples.

### 5.2 Higgs b-tagging optimization

The final state includes two b (anti-)quarks in this analysis, while because of boosted reason the two b quarks merging into one fat jet. That means we can apply a cut on the b-tagging of the subjets to reduce the background. Considering the jet track sharing, we used a Higgs b-tagging cut in this analysis. Following the recommendations of the Btag POG reported at <https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagPOG>, in *Boosted event topologies* bullet: using subjet b-tagging cut when the  $\Delta R$  bigger than 0.3 and using fat jet b-tagging cut when the  $\Delta R$  smaller than 0.3. In this analysis, we choose  $\Delta R = 0.3$  of the subjets as the demarcation point.

In addition, from <https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagPOG>, we know Loose or Medium working point is the recommendation of Btag POG. Considering both efficiency and significance of b-tagging cut (in Figures 6, 7, 8 and 9), we choose the following cut as our final Higgs b-tagging cut:

- using 2 CSVL subjet b-tagging cut if  $\Delta R > 0.3$  of the subjets;
- changing to use 1 CSVL fat jet b-tagging cut if  $\Delta R < 0.3$  of the subjets

which is also adopted in Ref. [40].

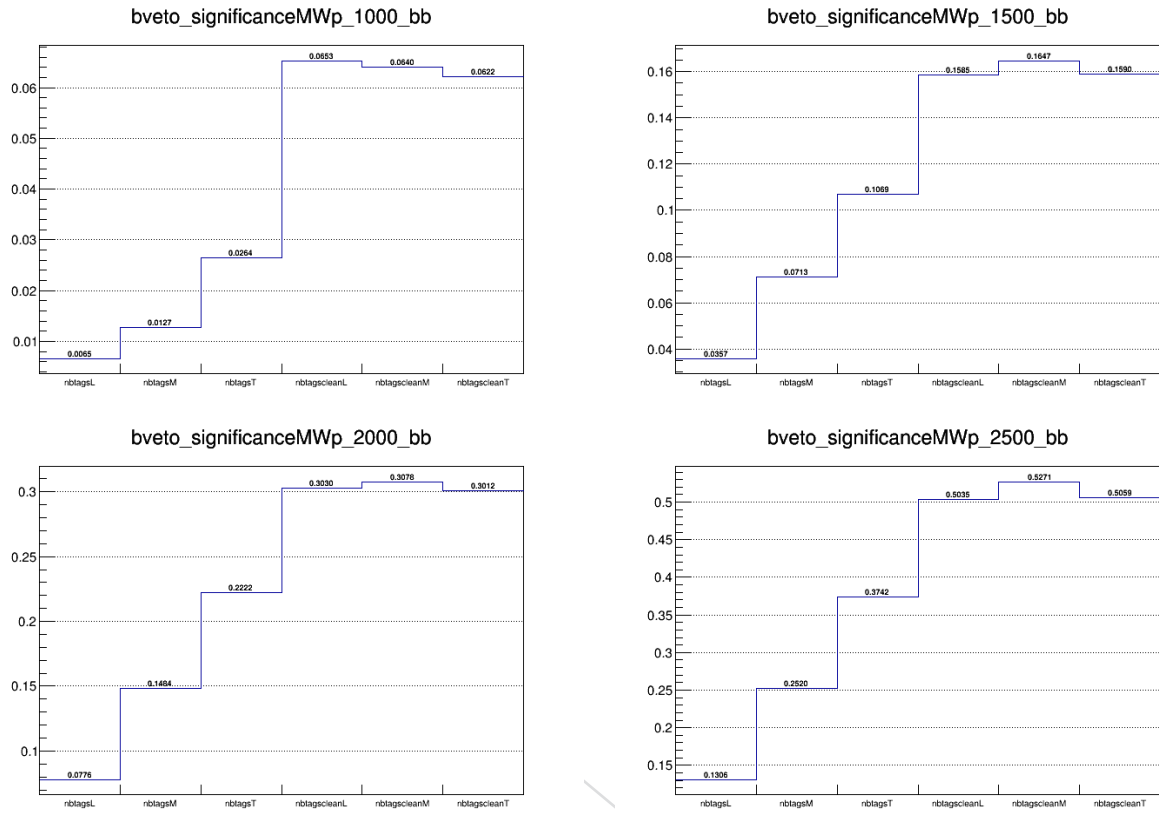


Figure 5: Punzi significance for various choices of b-tagging veto and for different resonance masses (1 TeV (top left), 1.5 TeV (top right), 2 TeV (bottom left) and 2.5 TeV (bottom right)).

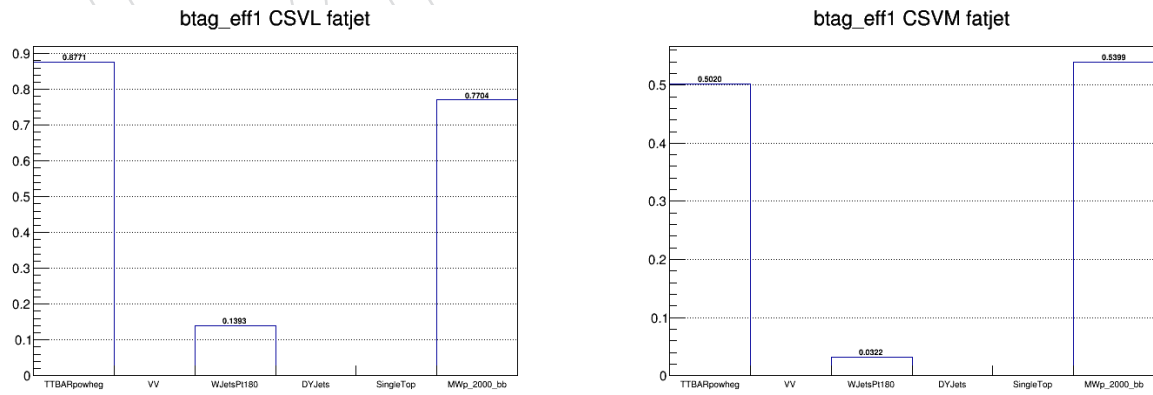


Figure 6: Efficiency for CSV b-tagging fat jet cut (1 CSVL fat jet b-tagging cut left, 1 CSVM fat jet b-tagging cut right).

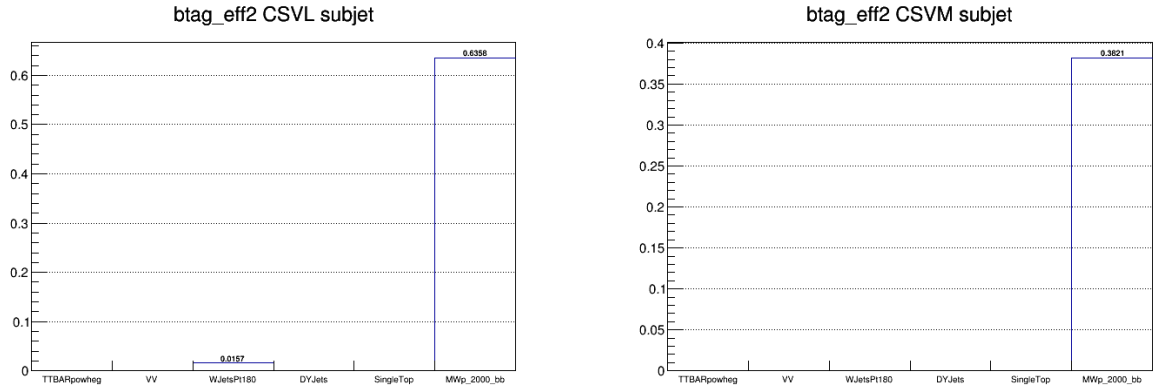


Figure 7: Efficiency for CSV b-tagging subjet cut (2 CSVL subjet b-tagging cut left, 2 CSVM subjet b-tagging cut right).

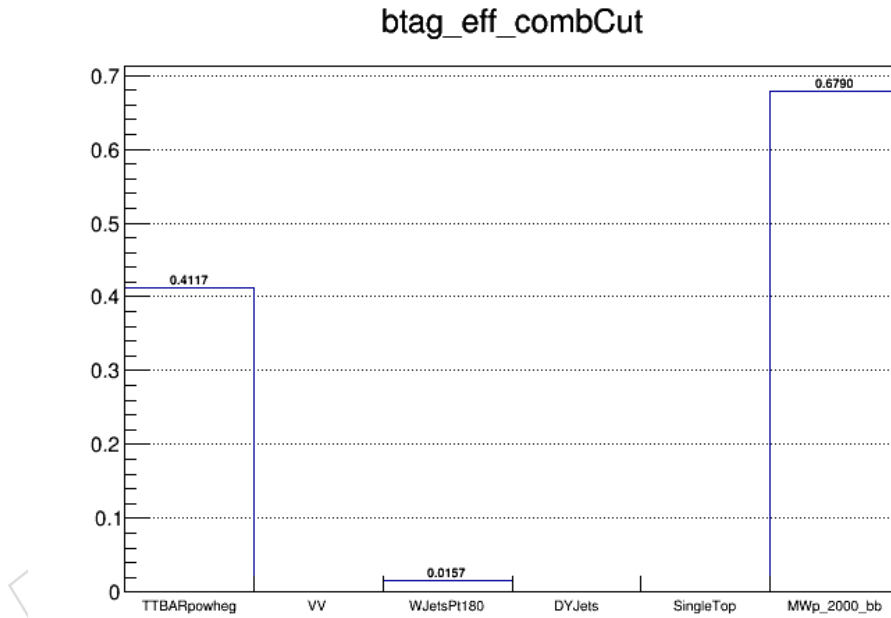


Figure 8: Efficiency for Higgs b-tagging cut.

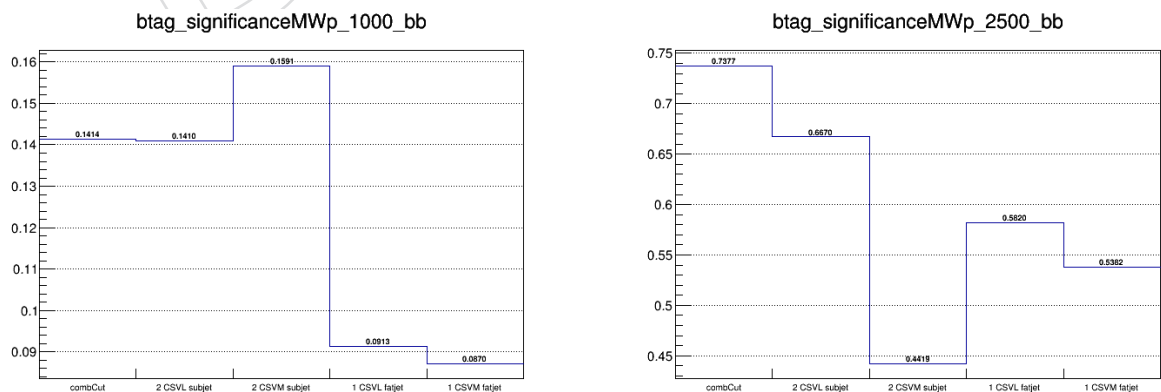


Figure 9: Punzi significance for various choices of b-tagging cut and Higgs b-tagging cut.

### 5.3 $t\bar{t}$ background rejection

The  $t\bar{t}$  process can have two real Ws one decaying in leptons and the other in hadrons. In addition this background have b quarks produced from the top decays, thus generating the same final state of the  $X \rightarrow WH$  signal. In  $t\bar{t}$  process, the b quarks are not overlapping with CA8 jet, which is opposite in signal samples. In the analysis, we use a veto on the number of b-tag jets (see Section 5.1) to suppress this contribution. Still some residual background is present after final selection consisting of  $\sim 85\%$  of the all background. Nevertheless, given the particular and well known signature of the top quark, more efforts can be spent to further reduce its contamination and gain additional sensitivity.

The analysis of the events left after the full set of selections described in the previous sections, have been studied in order to first understand the topology of the left  $t\bar{t}$  events. The distribution of the  $\Delta R(b_{gen}, jet^{CA8})$  in Figure 10 (*left*) shows that in  $\sim 96\%$  of the  $t\bar{t}$  events, the generated b-quark is inside the “fake” Higgs-jet candidate cone of 0.8. When considering in addition the mis-tagging of the second b-quark, the event is likely to pass the b-tag veto, leading to an event topology like the one shown in the scheme of Figure 11. When applying subjet (fatjet) b-tagging, one of the subjets is then b-tagged while the other subjet is likely to be due to one of the quarks coming from the hadronic W boson. Furthermore, the distribution of the PDG id’s [41] of the generated W daughters in Figure 10 (*right*) shows that, in the  $\sim 70\%$  of the events the W boson decays in a c-quark, leading to higher mis-tagging probability when applying subjet/fatjet b-tagging.

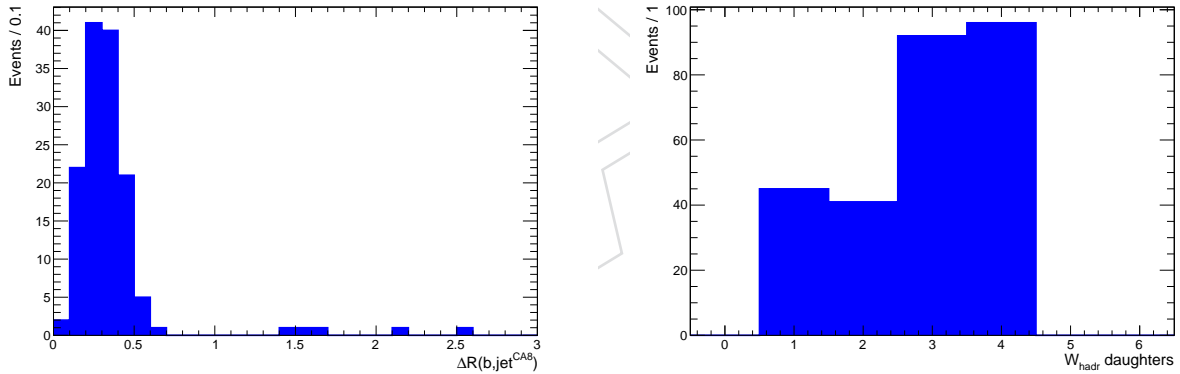


Figure 10: *Left*:  $\Delta R$  between the generated b-quark and the “fake” Higgs-jet in  $t\bar{t}$  events after full selections. *Right*: PDG id’s of the generated daughters of the hadronic W in  $t\bar{t}$  events after full selections.

Given the topology described above, the hadronic and leptonic top quarks are easily identified exploiting the presence of one or more extra not b-tagged jet in most of the  $t\bar{t}$  events (see Figure 12). In fact, given the selected “boosted” event topology, we expect the extra jet in the event to be close in the  $\phi - \eta$  space to the selected fat jet (but still outside the fat jet cone of 0.8 to avoid overlap in the jet reconstruction) or to the selected lepton. In the first case we expect the hadronic top quark to be correctly reconstructed, while in the latter we expect the leptonic one to be the candidate. If more than one extra jet is present, the closest to the lepton is chosen to reconstruct the leptonic top quark, while the closest to the fat jet is chosen to reconstruct the hadronic one.

Once the top quarks are identified the invariant mass of the leptonic and/or hadronic top quark can be computed (see Section 4.1 for the reconstruction of the leptonic W), only when at least

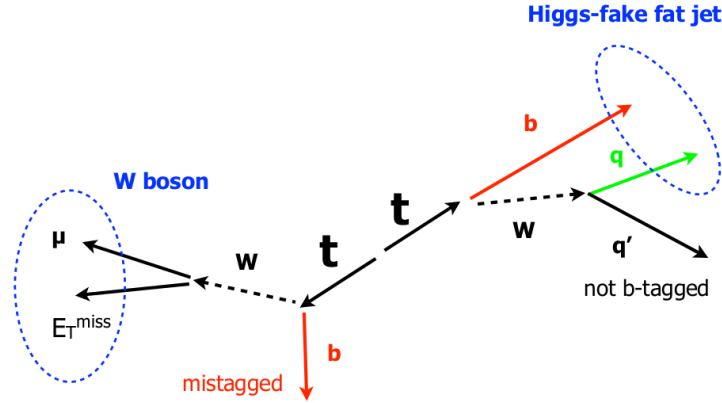


Figure 11: Topology of a  $t\bar{t}$  event where a jet is tagged as Higgs-jet passing the full selections.

one extra jet is present in the event. Figure 13 and Figure 14 show the distributions of the invariant mass of the hadronic (left) and leptonic (right) top mass, when the top reconstruction algorithm described above is applied to both signal and  $t\bar{t}$  events.

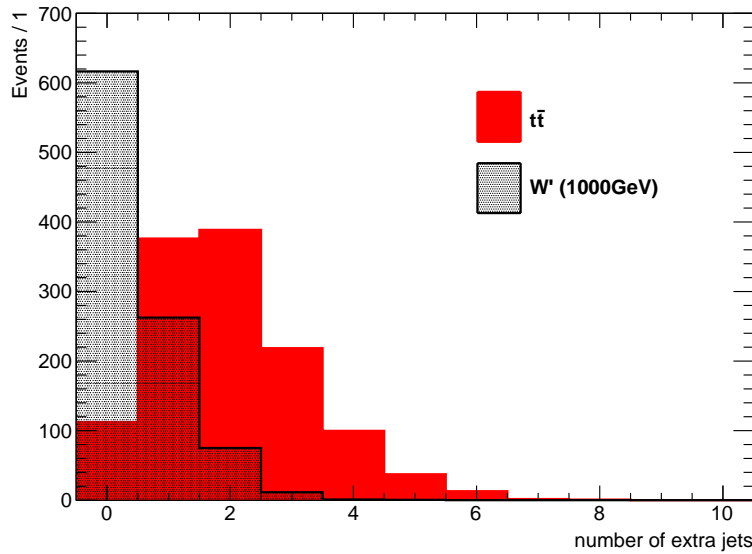


Figure 12: Distribution of the number of extra jets in  $t\bar{t}$  and signal events.

Once the topology has been fully understood, several strategies have been tested in order to reject these events:

- BDT with different sets of input variables
- combination of b-tag veto with BDT
- combination of b-tag veto with a veto on the reconstructed hadronic and/or leptonic top mass
- veto on the total number of extra jets

The different options have been compared in terms of the “Punzi’s significance”. For more details on the optimization see Appendix B, where the strategies are described and the results summarized.

Since the sensitivity reached with the BDT is comparable with the one reached with the veto on the reconstructed invariant mass of the top, the choice was made to include the top mass veto



in the selections. If no extra jets are present the veto does not have any effect.

Since for each event at least one of the two top quarks is correctly reconstructed we decided to apply the veto on the invariant mass in a way that it combines the information from both the top quarks as follows

$$120 < m_{top}^{leptonic} < 240 \parallel 160 < m_{top}^{hadronic} < 280$$

Furthermore, the lower and upper limits of the hadronic and leptonic top mass for the veto have been optimized on the three mass points of 1, 1.5 and 2 TeV giving the optimum choice above.

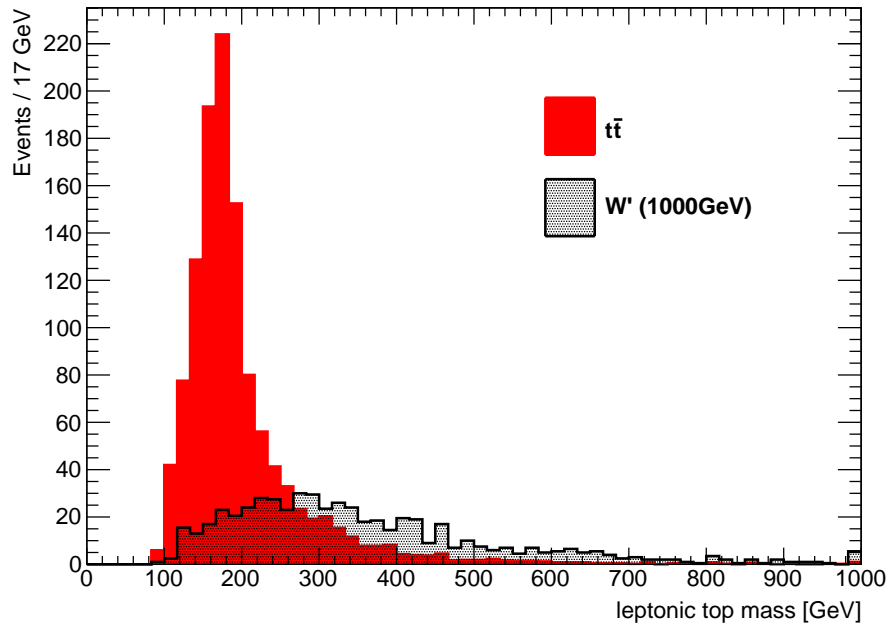


Figure 13: leptonic top invariant mass.

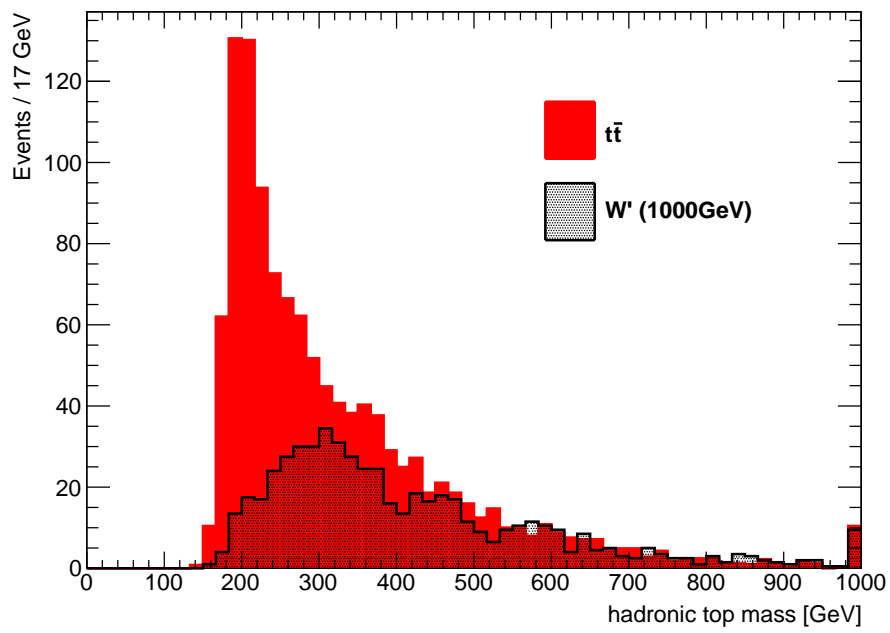


Figure 14: *Left*: hadronic top invariant mass

## 5.4 Final Selection and Control plots

A summary of the final kinematic and jet substructure selection criteria is presented in Table 4.

We use the low sideband region (events with  $40 < m_{jet}^{pruned} < 110$  GeV) to form a sample of data events with small signal contamination from a potential new  $X \rightarrow WH$  resonance. In this control sample we compare data and MC simulation. All the backgrounds (except W+jets) are normalized to the cross sections described in Section 2.3. Instead, the W+jets cross section described in Section 2.3 is multiplied by a factor 2.02 (2.08) for the muon (electron) channel, such that the total integral of the MC matches the number of events in data in the low sideband region. For W+jets, the WJetsToLNu\_PtW-100\_TuneZ2star\_8TeV-madgraph MC sample is used. Various distributions are shown for both electron and muon channels:

- Figure 15: lepton  $p_T$  and  $\eta$ ;
- Figure 16:  $E_T^{miss}$  and  $p_{z\nu}$  (defined in Section 4.1);
- Figure 17: leptonic W  $p_T$  and ;
- Figure 18: hadronic H  $p_T$  and  $\eta$ ;
- Figure 19:  $m_{jet}^{pruned}$ ;
- Figure 20:  $m_{WH}$  in sideband region (using the  $p_{z\nu}$  defined in Section 4.1);
- Figure 21:  $m_{WH}$  with only simulation in signal region (using the  $p_{z\nu}$  defined in Section 4.1);

From these figures, we can see that the agreement between data and simulation is quite good, which is a different case with [27]. That is why we use a relative simple way to estimate the background, especially for  $t\bar{t}$  and VV backgrounds (do not derive scale factors using  $t\bar{t}$  enriched sample, see Section 7.4).

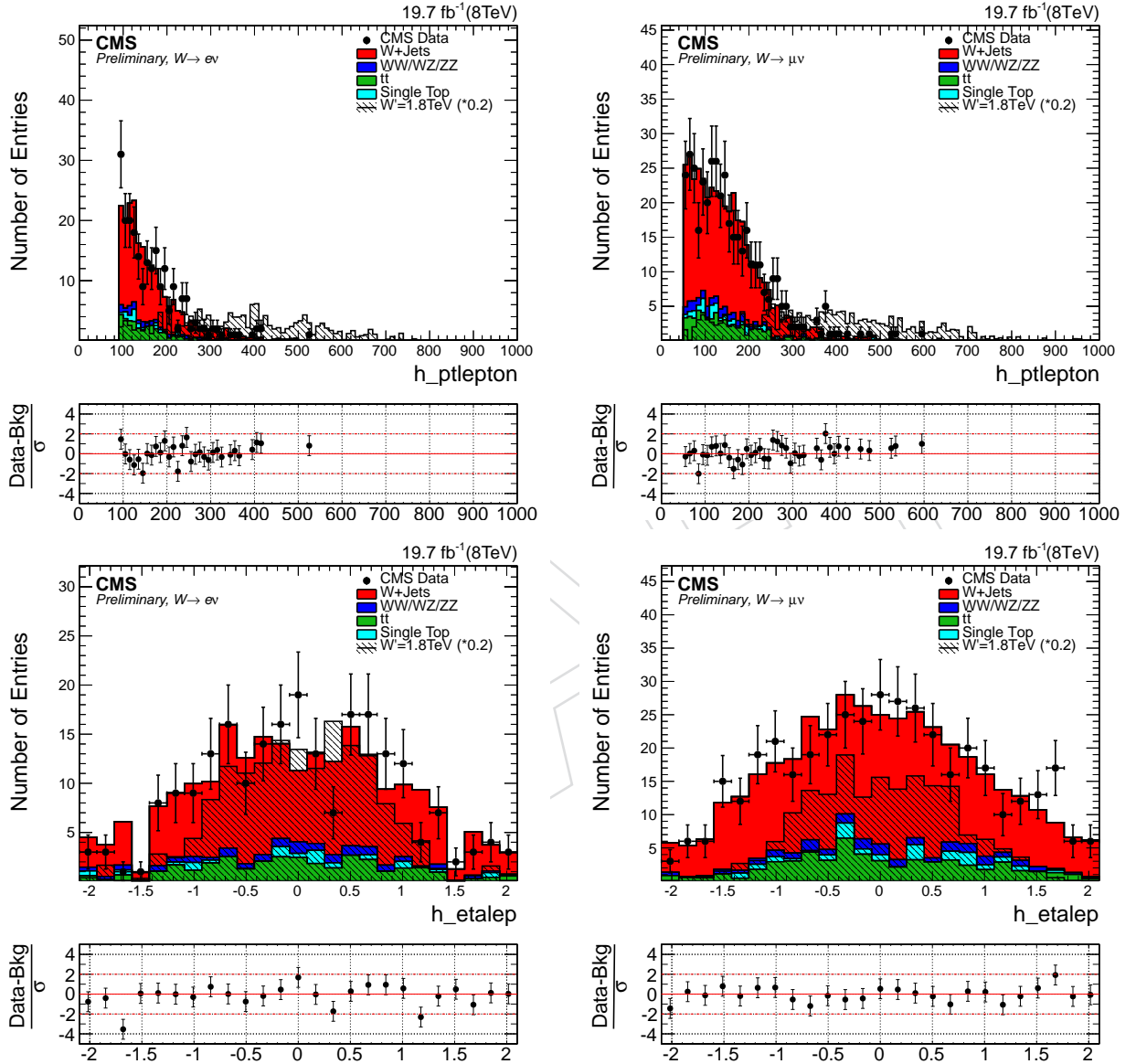


Figure 15: Lepton  $p_T$  and  $\eta$  for electron channel (left) and muon channel (right) for events with  $40 < m_{jet}^{pruned} < 110$  GeV. The signal is normalized to 0.1 times of the selected events for the model considered ( $W'$ ,  $M = 1$  TeV).

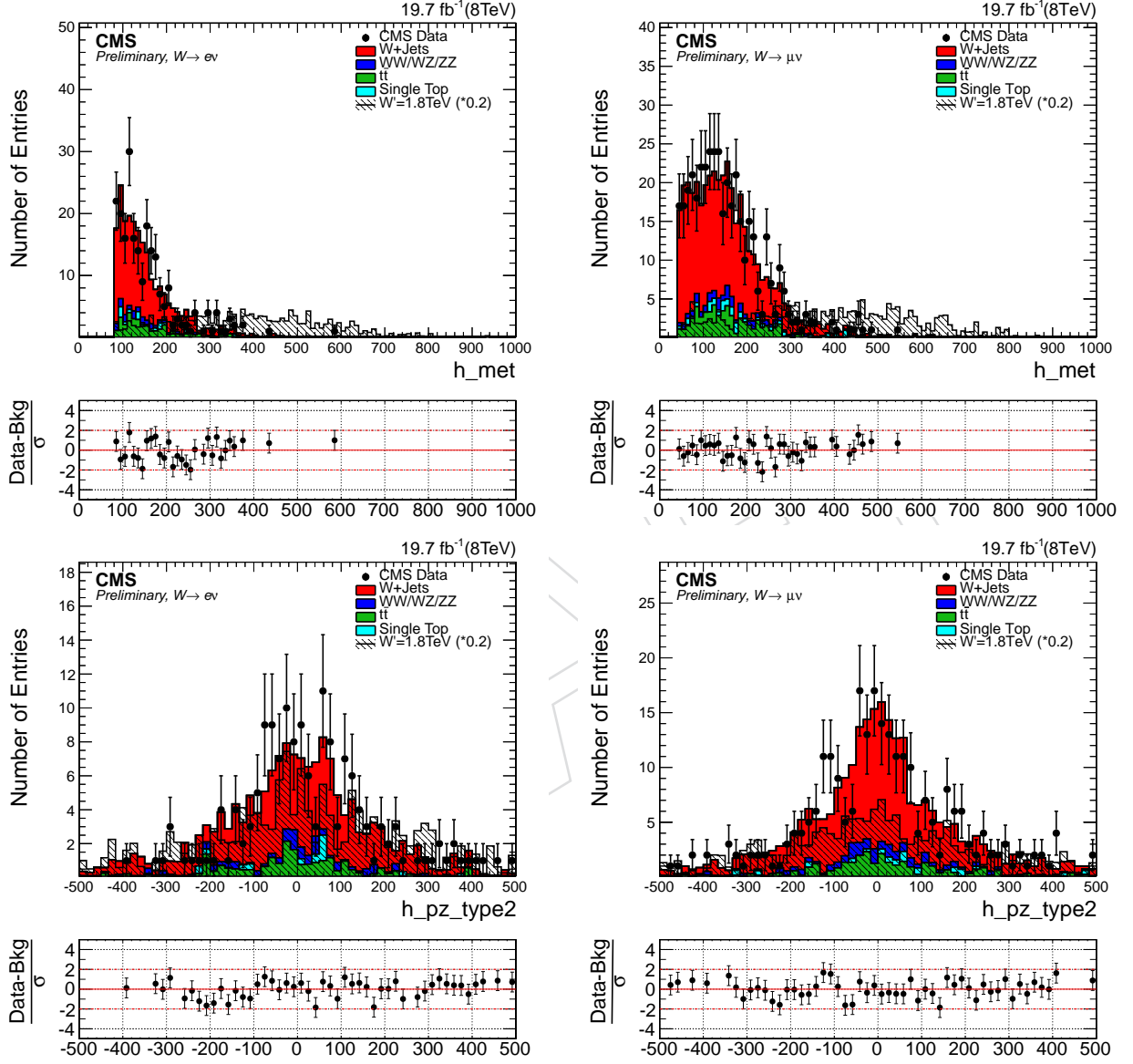


Figure 16:  $E_T^{\text{miss}}$  and  $p_{z\nu}$  (defined in Section 4.1) for electron channel (left) and muon channel (right) for events with  $40 < m_{\text{jet}}^{\text{pruned}} < 110\text{ GeV}$ . The signal is normalized to 0.1 times of the selected events for the model considered ( $W'$ ,  $M = 1\text{TeV}$ ).

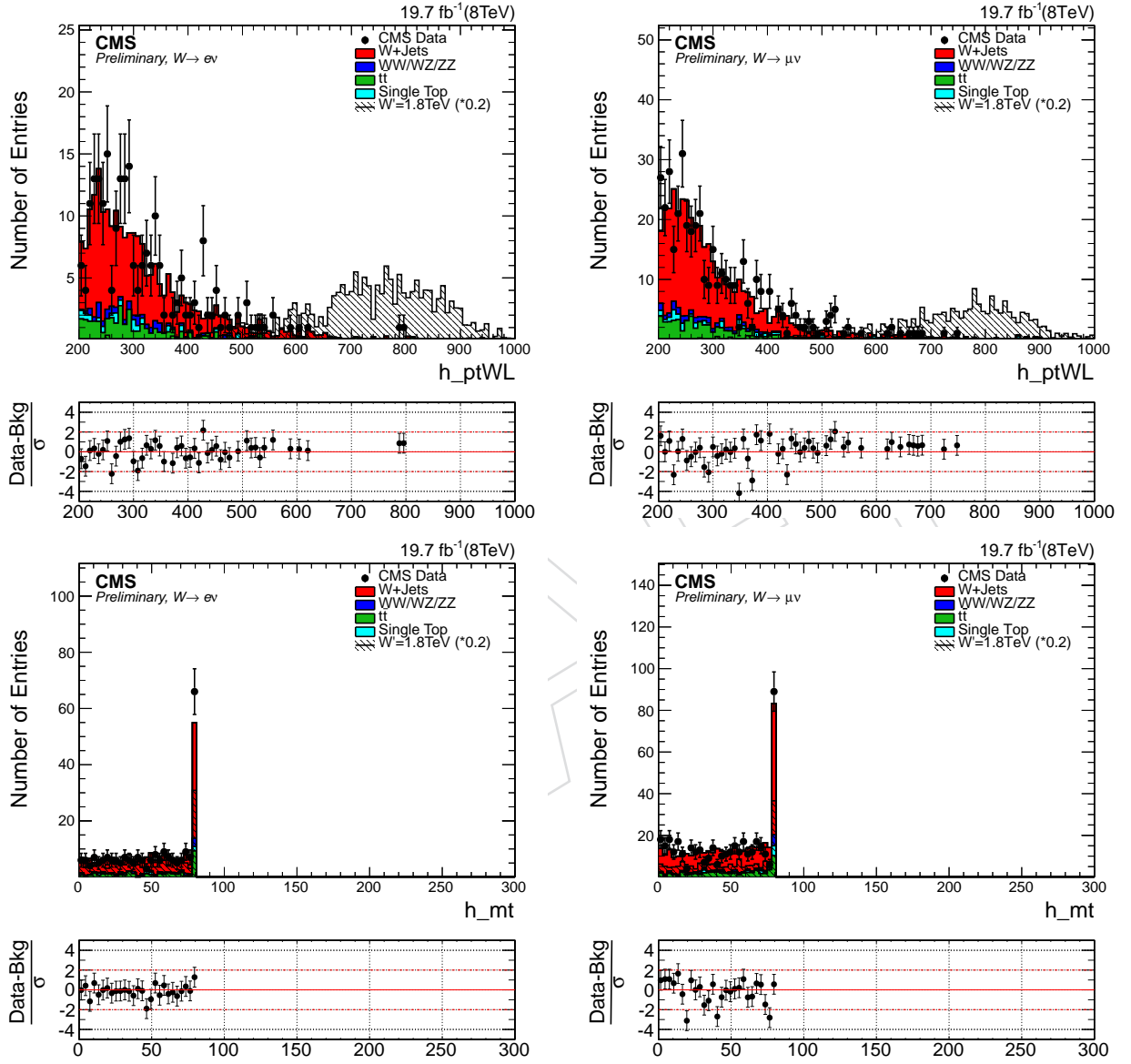


Figure 17: Leptonic  $W$   $p_T$  and for electron channel (left) and muon channel (right) for events with  $40 < m_{jet}^{pruned} < 110$  GeV. The signal is normalized to 0.1 times of the selected events for the model considered ( $W'$ ,  $M = 1$  TeV).

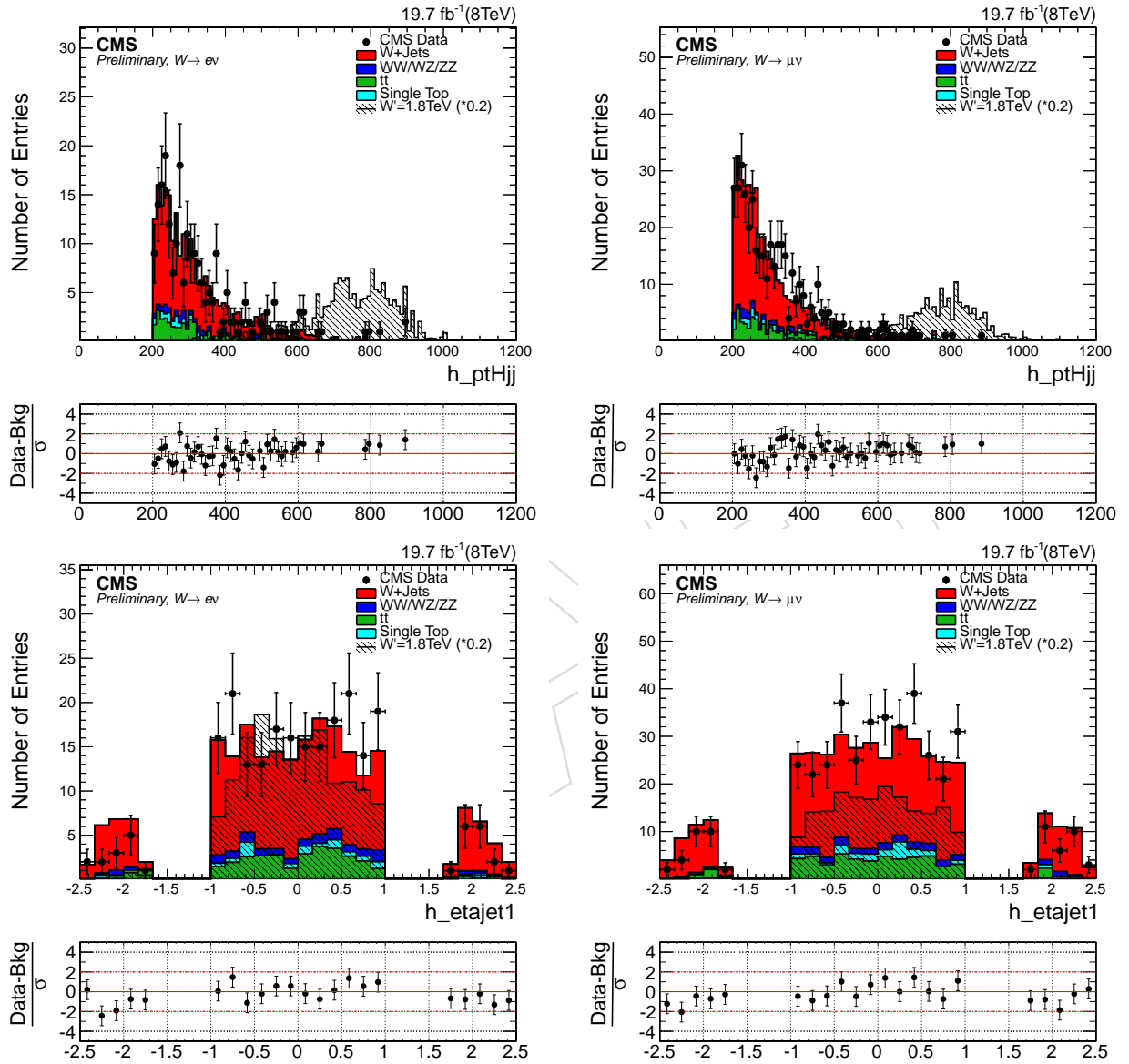


Figure 18: Hadronic W  $p_T$  and  $\eta$  for electron channel (left) and muon channel (right) for events with  $40 < m_{jet}^{pruned} < 110$  GeV. The signal is normalized to 0.1 times of the selected events for the model considered (W', M = 1TeV).

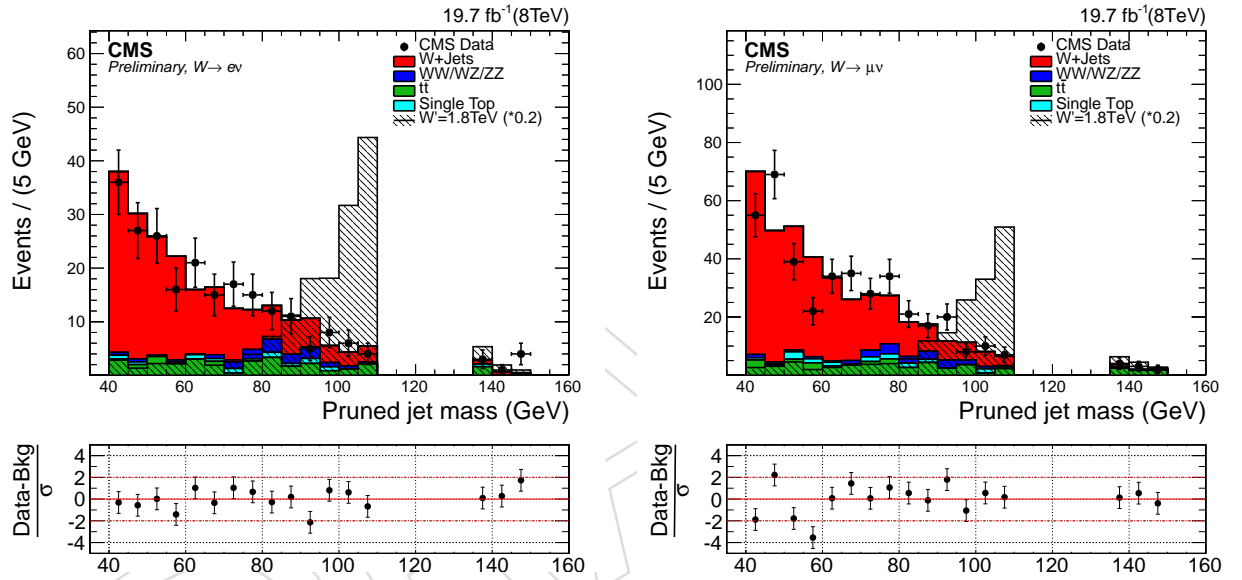


Figure 19:  $m_{jet}^{pruned}$  for electron channel (left) and muon channel (right) for events with  $40 < m_{jet}^{pruned} < 110$  GeV. The signal is normalized to 0.1 times of the selected events for the model considered ( $W', M = 1$  TeV).



Table 4: Summary of final selection.

Selection	Value	Comments
<b>Tight Lepton selection</b>		
Electron $p_T$	$p_T > 90 \text{ GeV}$	
Muon $p_T$	$p_T > 50 \text{ GeV}$	
Electron $\eta$	$ \eta _{SC} < 2.5$ except $[1.4442, 1.566]$ range	Avoid the ECAL gap.
Muon $\eta$	$ \eta  < 2.1$	
<b>Loose Lepton selection</b>		
Electron $p_T$	$p_T > 35 \text{ GeV}$	
Muon $p_T$	$p_T > 20 \text{ GeV}$	
Electron $\eta$	$ \eta _{SC} < 2.5$ except $[1.4442, 1.566]$ range	Avoid the ECAL gap.
Muon $\eta$	$ \eta  < 2.4$	
<b>CA8 jet selections</b>		
Jet $p_T$	$p_T > 80 \text{ GeV}$	Used for hadronic H reconstruction
Jet $\eta$	$ \eta  < 2.4$ except $[1, 1.5]$ range	
<b>AK5 jet selections</b>		
Jet $p_T$	$p_T > 30 \text{ GeV}$	Used for b-veto
Jet $\eta$	$ \eta  < 2.4$	jet selection
<b><math>E_T^{\text{miss}}</math> selections</b>		
$E_T^{\text{miss}}$ (electron ch.)	$E_T^{\text{miss}} > 80 \text{ GeV}$	
$E_T^{\text{miss}}$ (muon ch.)	$E_T^{\text{miss}} > 40 \text{ GeV}$	
<b>Boson selections</b>		
Pruned jet mass (signal)	$110 < m_{jet}^{\text{pruned}} < 135 \text{ GeV}$	
Pruned jet mass (low-mass sideband)	$40 < m_{jet}^{\text{pruned}} < 110 \text{ GeV}$	
Pruned jet mass (high-mass sideband)	$135 < m_{jet}^{\text{pruned}} < 150 \text{ GeV}$	
Leptonic W $p_T$	$p_T > 200 \text{ GeV}$	
Hadronic H $p_T$	$p_T > 200 \text{ GeV}$	
Back-to-back topology	$\Delta R(\ell, H_{had}) > \pi/2$ , $\Delta\phi(H_{had}, E_T^{\text{miss}}) > 2$ $\Delta\phi(H_{had}, W_{lep}) > 2$	
B-tag veto	no CSV Medium AK5 jet within $\Delta R(ca8, ak5) = 0.8$	
Combined b-tagging cut	$\geq 2$ CSVL subjets if $\Delta R > 0.3$ of subjets $\geq 1$ CSVL CA8 jet if $\Delta R < 0.3$ of subjets	
Top mass veto	$m_{top}^{\text{leptonic}} < 120 \parallel m_{top}^{\text{leptonic}} > 240$ $m_{top}^{\text{hadronic}} < 160 \parallel m_{top}^{\text{hadronic}} > 280$	
<b>Veto</b>		
Number of loose electrons	0	in addition to tight lepton
Number of loose muons	0	in addition to tight lepton

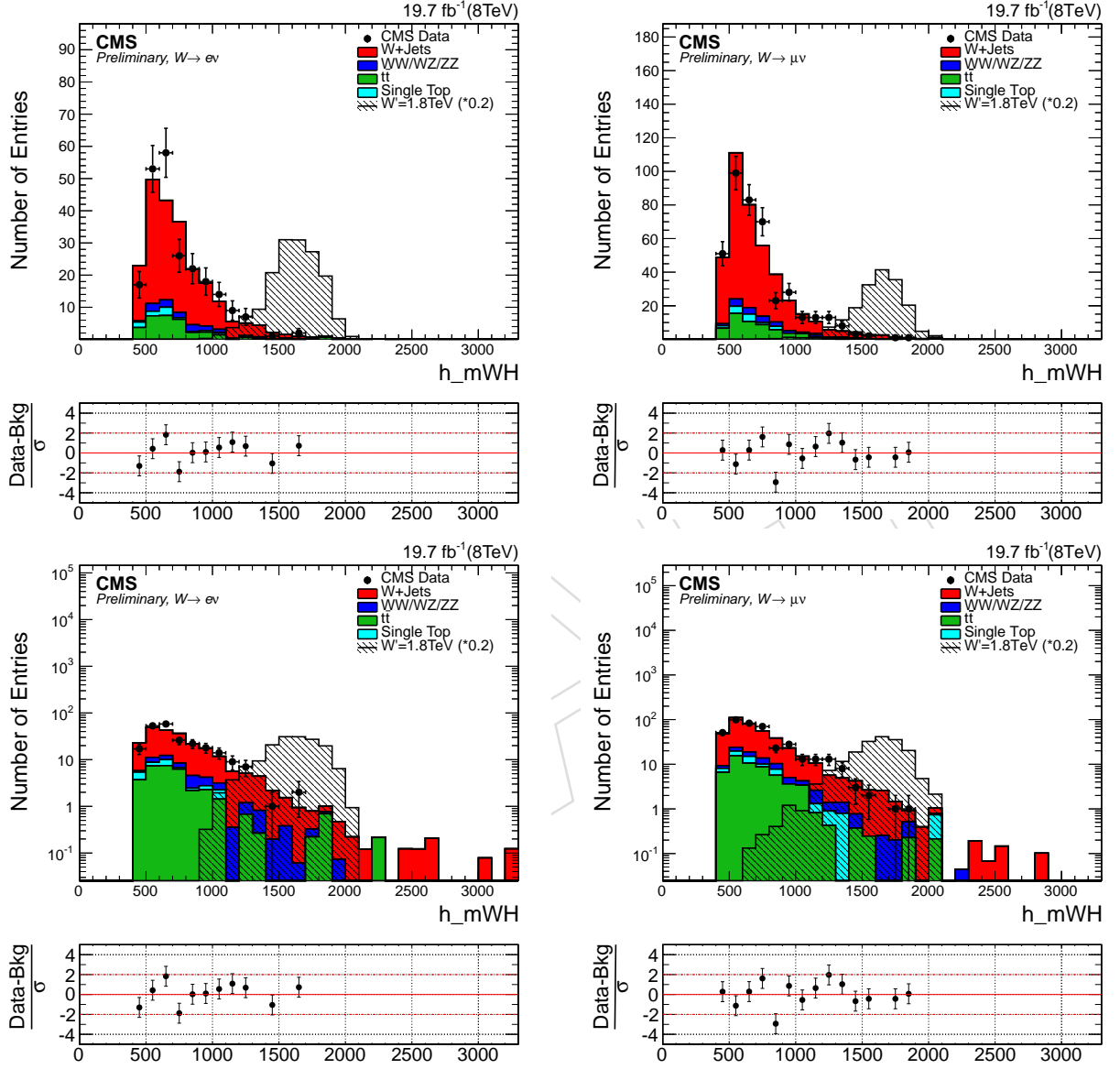


Figure 20:  $m_{WH}$ (using the  $p_{z\nu}$  defined in Section 4.1) in linear and log scale for electron channel (left) and muon channel (right) for events with  $40 < m_{jet}^{pruned} < 110$  GeV (sideband region). The signal is normalized to 0.1 times of the selected events for the model considered ( $W'$ ,  $M = 1$  TeV).

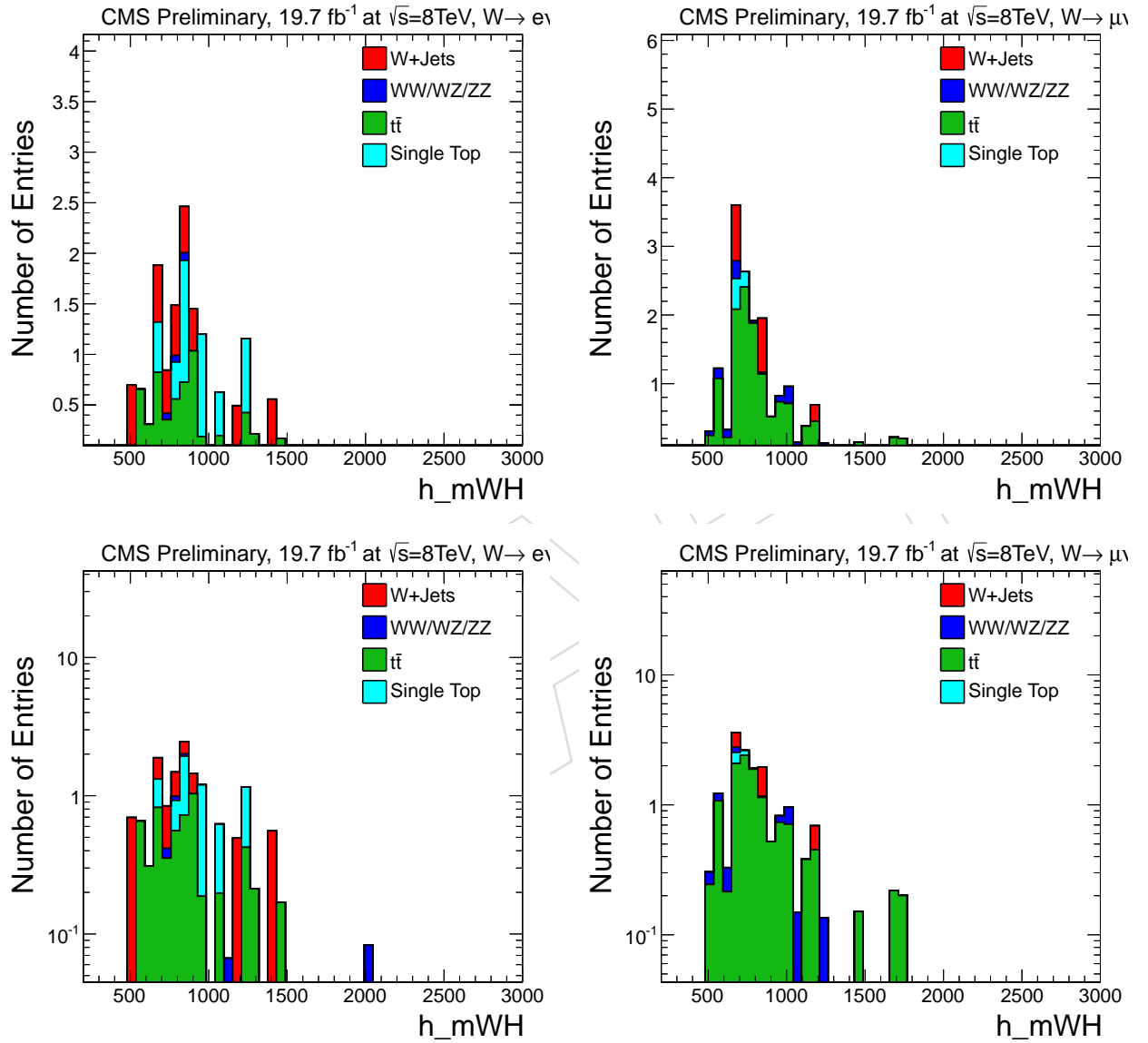


Figure 21:  $m_{WH}$ (using the  $p_{z\nu}$  defined in Section 4.1) in linear and log scale for electron channel (left) and muon channel (right) for events with  $110 < m_{jet}^{pruned} < 135$  GeV (signal region).

## 6 Signal Modeling

We used the  $W'$  samples listed in Table 2 for modeling the shape of the reconstructed  $m_{WH}$  distribution of the signal. The toy model used to generate the signal events has been compared at generator level for one mass point (1.5 TeV) with the two known benchmark models used for the interpretation of the results: the Little Higgs and the Heavy Vector Triplet. For the latter, the scenario B is considered as pointed out in Section 1. In addition to the signal samples mentioned above, two LHE files were produced with MadGraph for these two models. The model parameters have been chosen as it follows

HVT	Little Higgs
$g_V = 3$	
$C_H = -1$	$\cot 2\theta = 2.3$
$C_q = 1$	$\cot \theta = -0.20799$

where  $C_H$  describes interactions involving the Higgs boson or longitudinally polarized SM vector bosons,  $C_q$  describes the direct interactions of the  $W'$  boson with fermions and  $g_V$  is the typical strength of the new interaction.

The natural width and the cross section for both the models are listed in Table 5 for several resonance masses. The widths for the HVT have been computed by means of Equations 2.25 and 2.31 in Ref. [14], while the cross sections by means of the online tools provided by the authors of [14] and available at a dedicated webpage [42]. The widths for the Little Higgs have been computed by means of Equation 15 in [43].

Fig. 22 shows the width over mass ratio for  $W'$  in the Little Higgs and the HVT model. For wide range of  $\cot \theta$  or  $g_V$  in the two models, the ratio can be lower than 10%.

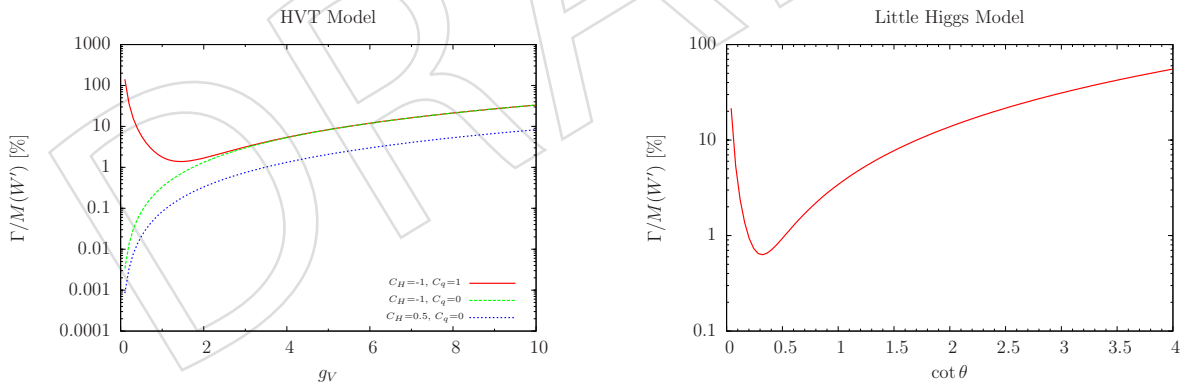


Figure 22: Ratio of width over mass for  $W'$  in the Little Higgs and the HVT<sub>B</sub>.

The comparison between the toy model, the Little Higgs and the HVT<sub>B</sub> for several kinematical distributions are represented in Figures 23-27 showing a good agreement between each others and validating the interpretation of the final results. Figure 28 shows the dependence of the width on the resonance mass for different values of  $g_V$  in the scenarios B of the HVT model (taken from [14]). The natural width of the resonance obtained using the chosen value of  $g_V$  is sufficiently small to be neglected when compared to the detector resolution. This makes our modeling of the detector effects on the signal shape independent of the actual model used for generating the events and allows us to apply it to a model-independent search for narrow-resonances. Figure 29 shows the fit of the signal resonance mass performed on the events

Table 5: Natural total widths and cross sections for the Little Higgs and  $HVT_B$  models for several resonance masses. The branching ratio of the  $WH \rightarrow l\nu b\bar{b}$  is not included in calculation.

Resonance mass [GeV]	Little Higgs		$HVT_B$	
	$\Gamma$ [GeV]	$\sigma$ [pb]	$\Gamma$ [GeV]	$\sigma$ [pb]
800	7.22	5.09E-01	24.08	3.37E-01
900	8.12	3.03E-01	27.10	2.48E-01
1000	9.02	1.87E-01	30.11	1.71E-01
1100	9.92	1.18E-01	33.12	1.16E-01
1200	10.83	7.65E-02	36.13	8.05E-02
1300	11.73	5.06E-02	39.14	5.59E-02
1400	12.63	3.39E-02	42.15	3.88E-02
1500	13.53	2.29E-02	45.16	2.51E-02
1600	14.44	1.56E-02	48.17	1.87E-02
1700	15.34	1.08E-02	51.18	1.30E-02
1800	16.24	7.43E-03	54.19	9.03E-03
1900	17.14	5.17E-03	57.20	6.27E-03
2000	18.05	3.61E-03	60.21	4.25E-03
2100	18.95	2.53E-03	63.22	3.02E-03
2200	19.85	1.76E-03	66.23	2.10E-03
2300	20.75	1.24E-03	69.24	1.46E-03
2400	21.65	8.67E-04	72.25	1.01E-03
2500	22.56	6.07E-04	75.27	7.31E-04
2600	23.46	4.25E-04	78.28	4.89E-04
2700	24.36	2.96E-04	81.29	3.39E-04
2800	25.26	2.07E-04	84.30	2.36E-04
2900	26.17	1.44E-04	87.31	1.64E-04
3000	27.07	9.97E-05	90.32	1.21E-04

554 generated with the  $HVT_B$  model. The distribution is parametrized with a Breit-Wigner function  
555 and the fit results are shown on the plot.

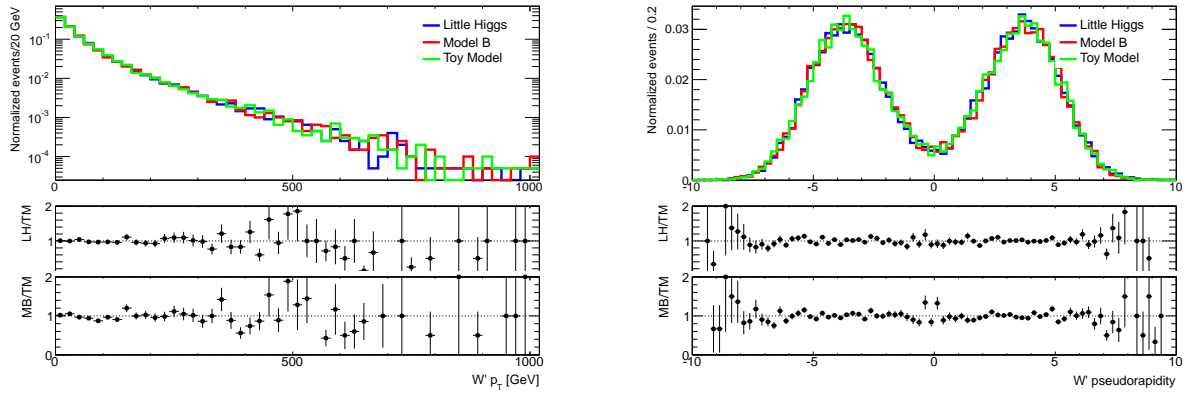


Figure 23: Comparison of the generated  $W'$   $\eta$  and  $p_T$  between the toy model, the Little Higgs and the  $HVT_B$ .

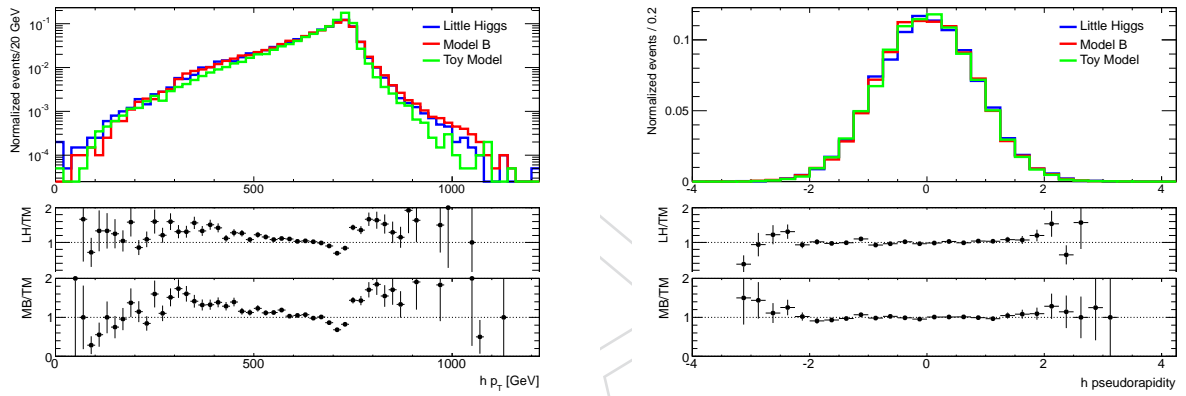


Figure 24: Comparison of the generated Higgs  $\eta$  and  $p_T$  between the toy model, the Little Higgs and the  $HVT_B$ .

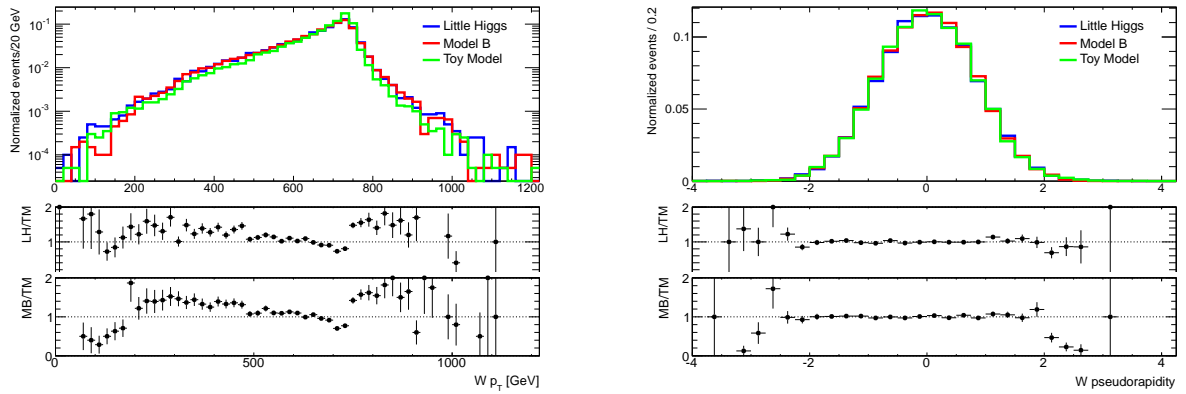


Figure 25: Comparison of the generated  $W$   $\eta$  and  $p_T$  between the toy model, the Little Higgs and the  $HVT_B$ .

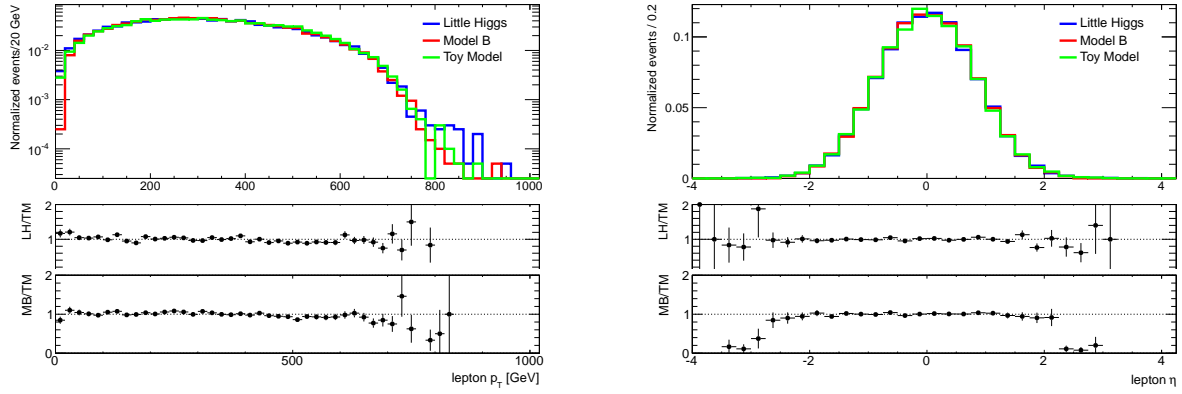


Figure 26: Comparison of the generated lepton  $\eta$  and  $p_T$  between the toy model, the Little Higgs and the HVT<sub>B</sub>.

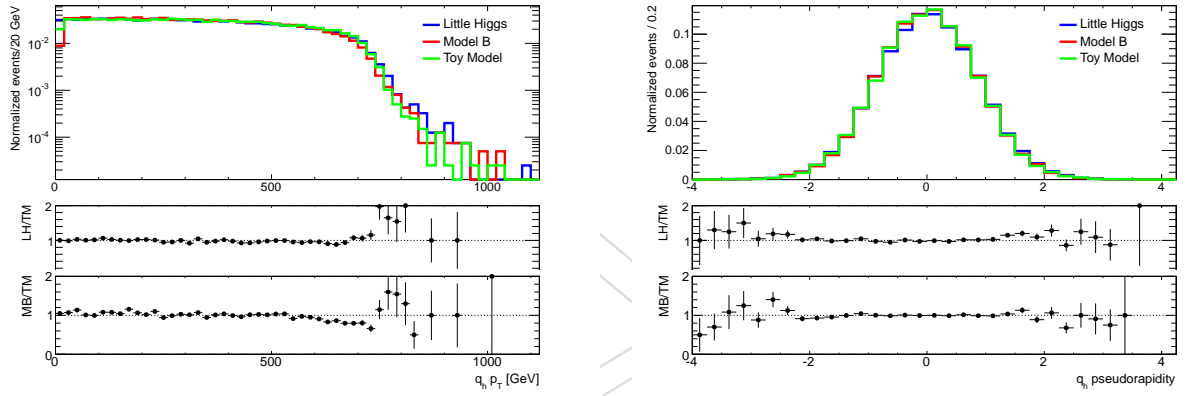


Figure 27: Comparison of the generated b-quarks  $\eta$  and  $p_T$  between the toy model, the Little Higgs and the HVT<sub>B</sub>.

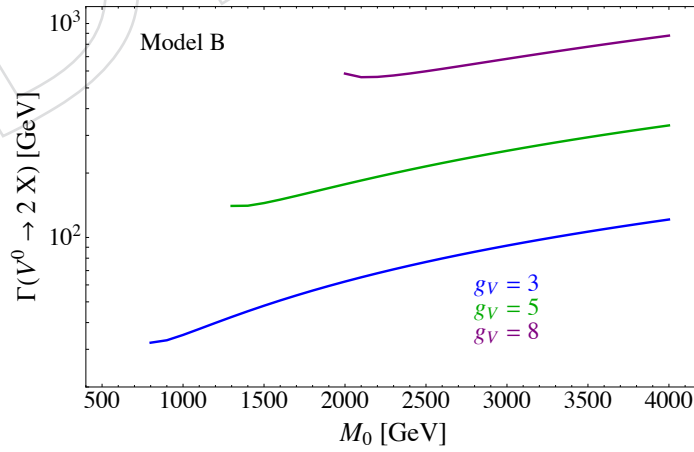


Figure 28: Total widths corresponding to different values of the coupling  $g_V$  in the scenario B of the HVT.

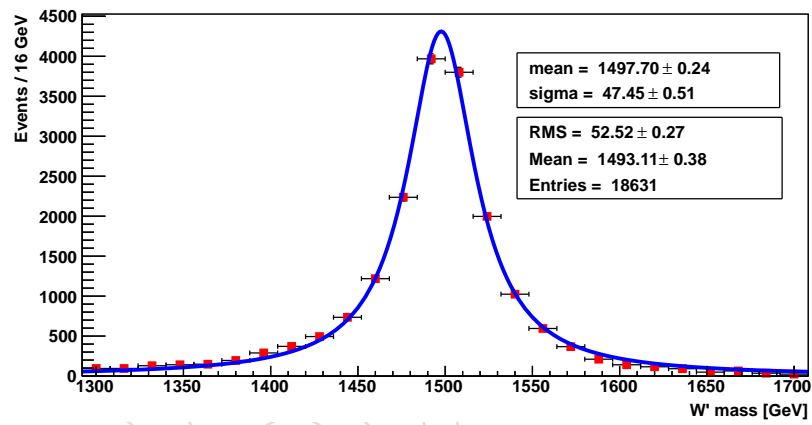


Figure 29: Fit of the resonance mass performed on the events generated with the  $HVT_B$  model.



## 6.1 Parametrization of the $m_{WH}$ signal shape

The reconstructed  $m_{WH}$  spectrum of the signal is modeled with a Crystal-Ball function (i.e. a Gaussian core with powerlaw tails on both sides, RooDoubleCB implemented in RooFit). The  $m_{WH}$  signal distribution varies depending on the resonance mass hypothesis. Several examples are shown in Figure 30 for electron and muon channels. A summary of the standard deviation of the Gaussian core is presented for several mass points in Table 6 for the electron and muon channels. The electron and muon shapes are very similar, even at high  $m_{WH}$ , since the mass resolution is dominated anyway by the  $E_T^{miss}$  contribution.

## 6.2 Signal Efficiency

The signal reconstruction efficiency in each of the two channels (electron and muon) is defined as the ratio:

$$\varepsilon_{SIG} = \frac{\text{Nr. of events passing the full selection in that channel}}{\text{Total nr. of events generated}} \quad (11)$$

where the total number of events generated includes the decay of the W boson in electron, muon and tau and the decay of the Higgs boson in b quarks. The signal efficiencies (as defined above) for the final selection criteria reported in Section 5.4 (signal region, i.e.  $110 < m_{jet}^{pruned} < 135$  GeV) are summarized in Table 7. The efficiency as a function of the nominal  $W'$  mass is shown in Fig. 31 for the muon (left) and the electron (right) channels.

Table 6: The main parameters (mass, mean and  $\sigma$ ) of the signal shape for different mass points at reconstructed level for the electron and muon channels.

Resonance mass [GeV]	Mean of CB function [GeV]	$\sigma$ of CB function [GeV]
electron channel		
800	797.8	46.45
900	889.85	56.04
1000	997.52	58.71
1100	1094.8	56.69
1200	1200	64.29
1300	1300	62.59
1400	1400	74.16
1500	1500	76.10
1600	1600	70
1700	1700	84.99
1800	1800	75.72
1900	1900	83.06
2000	2000	81.05
2100	2100	85
2200	2200	77.08
2300	2300	95
2400	2400	100
2500	2500	90
2600	2600	100
2700	2700	104.37
2800	2800	105
2900	2900	102.49
3000	3000	100
muon channel		
800	799.10	40
900	888.25	53.59
1000	993.45	59.34
1100	1089.20	62.66
1200	1200	56.24
1300	1300	60
1400	1400	68.60
1500	1500	71
1600	1600	72.05
1700	1700	75
1800	1800	75
1900	1900	80
2000	2000	80
2100	2100	85
2200	2200	79.51
2300	2300	95
2400	2400	100
2500	2500	90
2600	2600	100
2700	2700	100
2800	2800	105
2900	2900	116.98
3000	3000	100

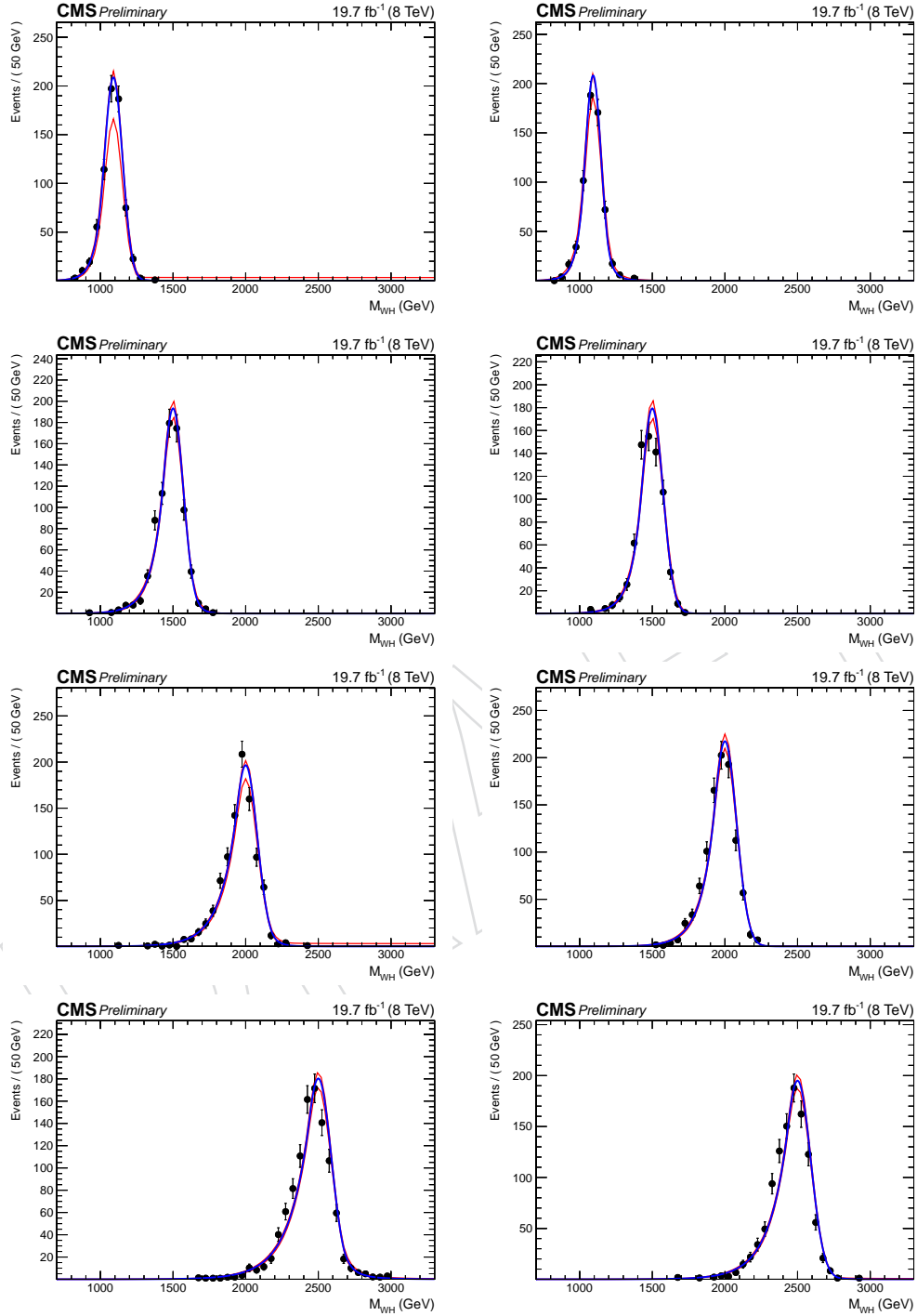


Figure 30: Distribution of the reconstructed invariant mass,  $m_{WH}$ , and shape parameterization for resonance masses of 1.1, 1.5, 2, and 2.5 TeV. Events reconstructed in muon (electron) channel are shown in the left (right) column.

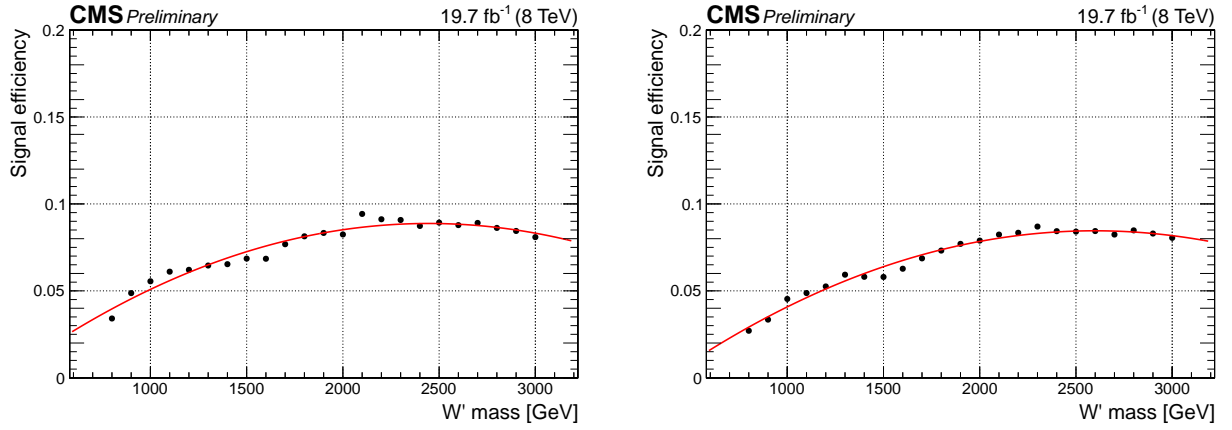


Figure 31: Signal efficiency for the final selection criteria as a function of the  $W'$  mass hypothesis in the muon (left) and electron (right) channel.

Table 7: Expected signal efficiency for the electron and muon channel separately, for several resonance mass hypothesis.

$M_{W'}$ [GeV]	Signal efficiency	
	electron channel	muon channel
800	0.027	0.034
900	0.033	0.049
1000	0.045	0.056
1100	0.049	0.061
1200	0.053	0.062
1300	0.059	0.065
1400	0.058	0.065
1500	0.058	0.069
1600	0.063	0.068
1700	0.069	0.077
1800	0.073	0.081
1900	0.077	0.083
2000	0.079	0.082
2100	0.082	0.094
2200	0.083	0.091
2300	0.087	0.091
2400	0.084	0.087
2500	0.084	0.089
2600	0.084	0.088
2700	0.082	0.089
2800	0.085	0.086
2900	0.083	0.084
3000	0.080	0.081

## 7 Background estimate

### 7.1 Strategy

We search for bumps in the  $m_{WH}$  spectrum due to new, heavy  $W'$  resonances. This section describes the methods used to derive both the normalization and the shape of the  $m_{WH}$  distribution for SM background events.

The background in the signal region is estimated from data (with the aid of MC for cross-checks) using two different methods:

- In the first method ( $\alpha$  method), the W+jets background prediction in the signal region is obtained from the data in the low-mass sideband region ( $40 < m_{jet}^{pruned} < 110$  GeV), rescaled for an appropriate transfer function from the sideband to the signal region derived from simulation.
- In the second method (SF method), we employ two scale factors to rescale W+jets and  $t\bar{t}$  backgrounds in signal region ( $110 < m_{jet}^{pruned} < 135$  GeV) and then perform a simple and direct “cut-and-count” procedure.

In this analysis, we mainly depend on the first  $\alpha$  method (Section 7.2), and the second method (Section 7.3) is only used to cross check.

A final cross check in a  $t\bar{t}$  enriched control sample is then performed and an uncertainty on its normalization derived (Section 7.4).

### 7.2 Alpha Method

The method presented in this section is similar to the one already used in previous CMS analysis to search for an BSM to WW [27] and [44] and X to ZZ [24], in the semi-leptonic channels.

As shown in Fig. 20, the main backgrounds after final selection are W+jets,  $t\bar{t}$ , SingleTop and SM VV processes. For non-W+jets backgrounds, we estimate them from MC for their shape and normalization, and they are initially rescaled to the integrated luminosity of the data using their corresponding NLO cross sections listed in Section 2.3.

The total background prediction (normalization and shape) as a function of the reconstructed  $m_{WH}$  is obtained by this formula, separately for each analysis category (ele LP, ele HP, mu LP, mu HP):

$$\begin{aligned}
 N_{tot}^{signal}(m_{WH}) &= N_{W+jets}^{signal}(m_{WH}) + N_{OB}^{signal}(m_{WH}) \\
 &= N_{data}^{sideband}(m_{WH}) \times (1 - R_0(m_{WH})) \times \alpha^{MC}(m_{WH}) \times F_{W+jets} \\
 &\quad + N_{OB}^{signal}(m_{WH})
 \end{aligned} \tag{12}$$

where

- $N_{tot}^{signal}(m_{WH})$  is the total background prediction in the signal region;
- $N_{W+jets}^{signal}(m_{WH})$  is the W+jets (plus the small Z+jets) background prediction in the signal region;
- $N_{OB}^{signal}(m_{WH}) = N_{VV}^{signal}(m_{WH}) + N_{t\bar{t}}^{signal}(m_{WH}) + N_{singletop}^{signal}(m_{WH})$  is the background prediction in the signal region for the sum of the non-W+jets backgrounds (OB = Other Backgrounds = VV,  $t\bar{t}$ , single top);

- $N_{data}^{sideband}(m_{WH})$  is the  $m_{WH}$  distribution in data for the low-mass sideband region ( $40 < m_{jet}^{pruned} < 110$  GeV);
- $R_0(m_{WH}) = N_{OB}^{sideband}(m_{WH}) / N_{data}^{sideband}(m_{WH})$  is the fraction of non-W+jets backgrounds (VV,  $t\bar{t}$ , single top) in the sideband region with respect to the total background in the sideband region (i.e. data). The  $1 - R_0(m_{WH})$  multiplicative term represents the subtraction of the non-W+jets contribution from the data in the sideband region.
- $\alpha^{MC}(m_{WH}) = N_{W+jets}^{signal}(m_{WH}) / N_{W+jets}^{sideband}(m_{WH})$  is the W+jets background prediction in the signal region divided by the one in the sideband region, calculated from MC. It represents the transport function (from sideband to signal region) used to correct the data in the sideband region and extract the W+jets  $m_{WH}$  shape.
- $F_{W+jets}$  is an overall scale factor used to set the normalization of the W+jets background prediction.

The terms reported in the formula above are discussed in details in the following sections.

### 7.2.1 W+jets background

The W+jets background is an important (reducible) background in this analysis. The real  $W \rightarrow \ell \nu$  is reconstructed, while the hadronic H candidate is a b-quark initiated jet which gets a large mass due to QCD radiation in the fragmentation process. This choice is motivated by the disagreement in the jet mass distribution between the data and simulation, as proved in [45].

- $F_{W+jets}$  - The normalization of the W+jets background is obtained using the same fit procedure employed in Ref. [44] (the actual software implementation is also exactly the same<sup>3</sup>). We select events passing the full selection and additional requirement on the minimum reconstructed  $m_{WH} > 800$  GeV (corresponding to the region where we perform the search for  $X \rightarrow WH$  resonances). We then take the pruned jet mass distribution of the data in the range  $40 < m_{jet}^{pruned} < 150$  GeV, as shown in Figure 32. The background contributions for VV,  $t\bar{t}$ , and single top are obtained by fitting the individual MC predictions with various functions describing the combinatorial background and the Gaussian core of these distributions (more details can be found in Section 5.4 of [44]); the fit parameters (both normalization and shape) for these backgrounds are fixed by the MC prediction. The W+jets component is also parametrized with a combinatorial-like function which is found to well describe the MC (more details can be found in Section 5.4 of [44]). We finally perform a fit to the data in the *low-mass plus high-mass sideband regions* ( $40 < m_{jet}^{pruned} < 110$  GeV plus  $135 < m_{jet}^{pruned} < 150$  GeV) using the background parametrization introduced above for each component. In this fit, the only free parameters are associated to the normalization and shape of the W+jets background, while the other background components are fixed from the simulation. The fit result for the W+jets background is extrapolated in the signal region ( $110 < m_{jet}^{pruned} < 135$  GeV) to provide the normalization of the W+jets background (i.e. the scale factors  $F_{W+jets}$ ).
- $\alpha^{MC}(m_{WH}) = N_{W+jets}^{signal}(m_{WH}) / N_{W+jets}^{sideband}(m_{WH})$  - The shape of the W+jets: is the W+jets background prediction in the signal region divided by the one in the sideband region, and it is calculated from MC. It represents the transport function (from sideband to signal region) used to correct the  $m_{WH}$  spectrum of the data in the sideband

<sup>3</sup>Thanks to the authors of [44] for providing their code.

region. This method only provides the W+jets shape, while the normalization has been discussed before in this Section. Most of the systematic uncertainties associated to the MC cancels in this ratio. The method relies on the assumption that the correlation between  $m_{WH}$  and the pruned jet mass for W+jets events in data is reasonably well reproduced by the MC. The procedure has been successfully validated performing a closure test with data in the sideband region, as described in Appendix C.

A number of different functions are used both for the continuous  $m_j$  and  $m_{lvj}$  shapes. A default shape is chosen for each fit and the other shapes, which also describe well the fit sample, are used as alternative to estimate systematic uncertainties:

- Functions used for fitting  $m_j$  spectrum:

$$F_{\text{Exp}}(x) = e^{ax}$$

$$F_{\text{ErfExp}}(x) = e^{ax} \cdot \frac{1 + \text{Erf}((x - b)/w)}{2}$$

$$F_{\text{Pow2}}(x) = \frac{((1 - c_0x)^a}{c_1x^b}$$

$$F_{\text{ExpGaus}}(x) = e^{ax} \cdot e^{-(x-b)^2/2s^2}$$

$$F_{\text{Gaus}}(x) = e^{-(x-a)^2/2s^2}$$

$$F_{\text{GausErfExp}}(x) = f_0 \cdot F_{\text{Gaus}}(x) + (1 - f_0) \cdot F_{\text{ErfExp}}(x)$$

- Functions used for fitting  $m_{lvj}$  spectrum:

$$F_{\text{ExpN}}(x) = e^{ax+b/x} \quad (13)$$

$$F_{\text{ExpTail}}(x) = e^{-x/(a+bx)}$$

$$F_{\text{Exp}}(x) = e^{ax}$$

where  $c_{0,or1}$  is constant, selected with different initial values depending if we fit the  $m_j$  or the  $m_{lvj}$  distribution, while the other parameter (including  $a, b, f_0$  and  $s, w$ ) are floating.

All the chosen shapes used to fit  $m_j$  and  $m_{lvj}$  are listed in Table. 8 and Table. 9. In addition, for the fit function, the  $F_{\text{ExpN}}(m_{lvj})$ , defined in Eq. 13, is taken as default shape when W+jets spectrum is fitted in the range between [0.7-3.0] TeV, while the  $F_{\text{ExpTail}}(m_{lvj})$  is used as alternative function to evaluate shape systematics. The reason of this choice reported at [44].

Mass (GeV)	W+Jets	$t\bar{t}$	Single Top	VV
40-110	$F_{\text{Pow2}}$	$F_{\text{GausErfExp}}$	$F_{\text{Exp}}$	$F_{\text{ExpGaus}}$

Table 8: Summary of all the shapes used for fit the  $m_j$  spectrum of each background component

Mass (TeV)	$m_j$ Region	W+Jets	$t\bar{t}$	Single Top	VV
[0.8-3]	SB	$F_{\text{ExpN}}$	$F_{\text{Exp}}$	$F_{\text{Exp}}$	$F_{\text{Exp}}$
[0.8-3]	SR	$F_{\text{ExpN}}$	$F_{\text{Exp}}$	$F_{\text{Exp}}$	$F_{\text{Exp}}$

Table 9: Summary of all the shapes used for fit the  $m_{lvj}$  spectrum of each background component

Figure 33 shows the  $\alpha^{MC}(m_{WH})$  functions for the different channels.

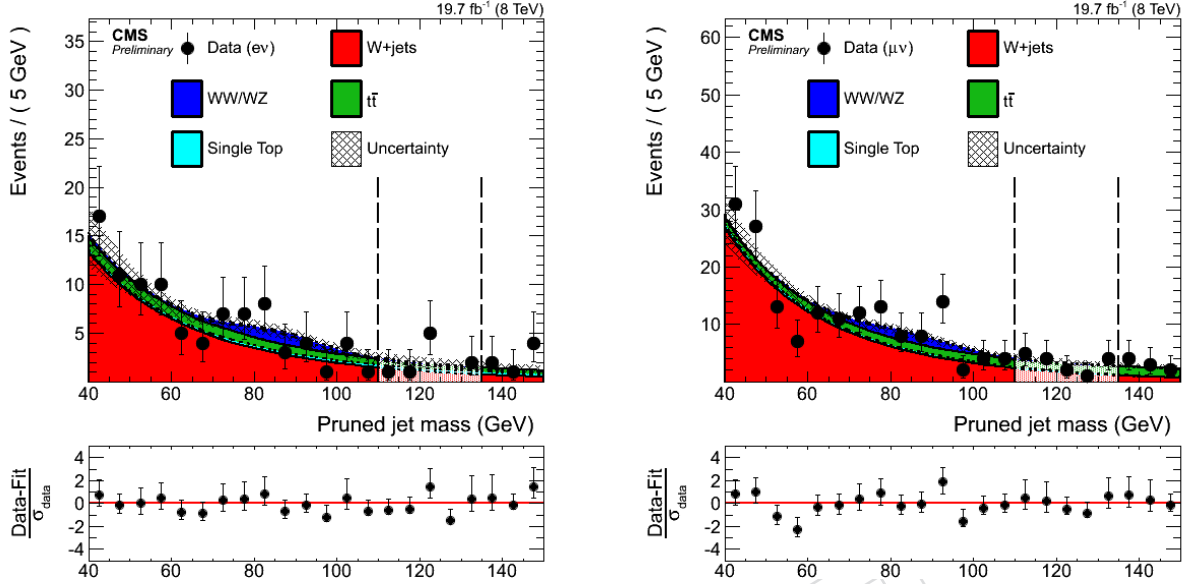


Figure 32: Pruned jet mass distributions for the electron (left), and muon (right) categories for events with  $40 < m_{jet}^{pruned} < 150$  GeV passing the full selection plus the  $m_{WH} > 800$  GeV requirement. The different background components, results of the fit to the low-mass and high-mass sideband regions in data, are shown.

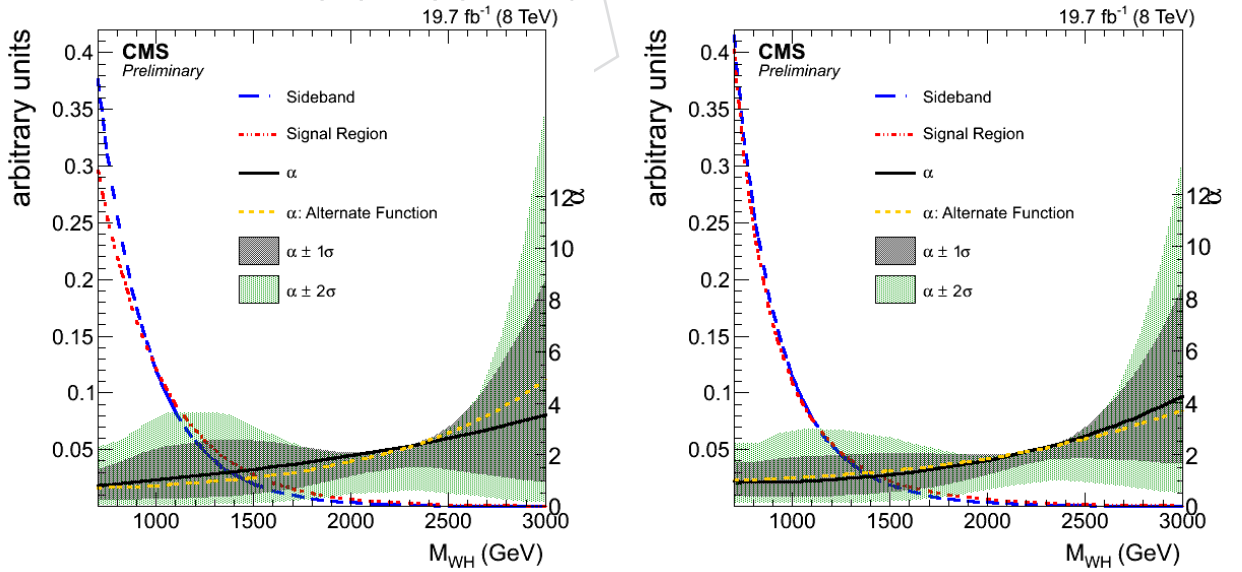


Figure 33:  $\alpha^{MC}(m_{WH})$  functions vs  $m_{WH}$  for the electron(left), muon(right) channels.



### 7.2.2 Background uncertainty

We consider two type of uncertainties for the total background prediction:

- **Uncertainty on background fit parameters (shape and normalization)**

We use a leveled exponential fit function with 2 shape parameters and 1 normalization parameter. Each fit returns uncertainties on all the fitted parameter, which are in general correlated. We want to include systematic uncertainties on the fit parameters in the limit/p-value calculation. However, from a technical point of view, the statistical tool that we use for the limit calculation (see Sec. 9) is not able to properly deal with partially correlated parameters. For this reason we redefine the parameters by diagonalizing the covariance matrix in order to decorrelate them. In this procedure, the new parameters are defined in such a way to be centered at zero and with error equal to unity<sup>4</sup>. The background fit parameterization is then redefined as a function of these new, uncorrelated parameters. This new fit function (and its related uncertainties on fit parameters) is used to describe the background in the limit calculation of Sec. 9.

- **Uncertainty on  $\alpha^{MC}(m_{WH})$**

In addition to the statistical uncertainties from the number of events in the sidebands of data, there are statistical uncertainties on the extrapolation factors  $\alpha^{MC}(m_{WH})$  due to the limited size of the W+jets MC sample. These uncertainties are propagated to the final background parameterization in the following way: using two different functions to fit W+jets sample to get W+jets shape, and the uncertainty coming from the W+jets shape can be converting into difference between the fit parameters.

Studies reported at [44] have shown that  $\alpha^{MC}(m_{WH})$  can change due to different MC with different parton showering algorithm considered. Herwig++ and Pythia were compared in this study, showing that the variation on  $\alpha^{MC}(m_{WH})$  is of the same size of the statistical uncertainties on the MC. For this reason the uncertainties on  $\alpha^{MC}(m_{WH})$  have been inflated by a factor  $\sqrt{2}$  to take this effect into account.

Figure 34 gives the final estimated background distribution.

## 7.3 SF Method

This method is a simple cut-and-count method to calculate limit.

- First, in signal region ( $110 < m_{jet}^{pruned} < 135$ ) of  $t\bar{t}$  control samples, we perform a cut-and-count flow to get a scale factor (yields of  $t\bar{t}$  in data over yields of  $t\bar{t}$  in MC) for  $t\bar{t}$  background.
- Second, we put the SF into the sideband region ( $40 < m_{jet}^{pruned} < 110$ ) of normal sample and perform again the cut-and-count procedure to calculate the W+jets scale factor (also get the yields of data and MC individually).
- Third, we put the two SFs into signal region of normal samples to correct the  $t\bar{t}$  and W+jets backgrounds.

Note, both of the two SFs are get using the whole  $m_{WH}$  region ( $m_{WH} > 800$  GeV). While, estimating the backgrounds rate for each of  $W'$  mass point, we only count the yields of up and down 100GeV of that mass (e.g. count yields  $m_{WH}$  belong to [900,1100] for  $W'$  1TeV point).

<sup>4</sup>the shift and renormalization of the parameter is a somewhat arbitrary convention present in the diagonalization tool that we use.

709  
710

From Figure 35, we can see that the two methods of backgrounds estimation is consistent with each other.

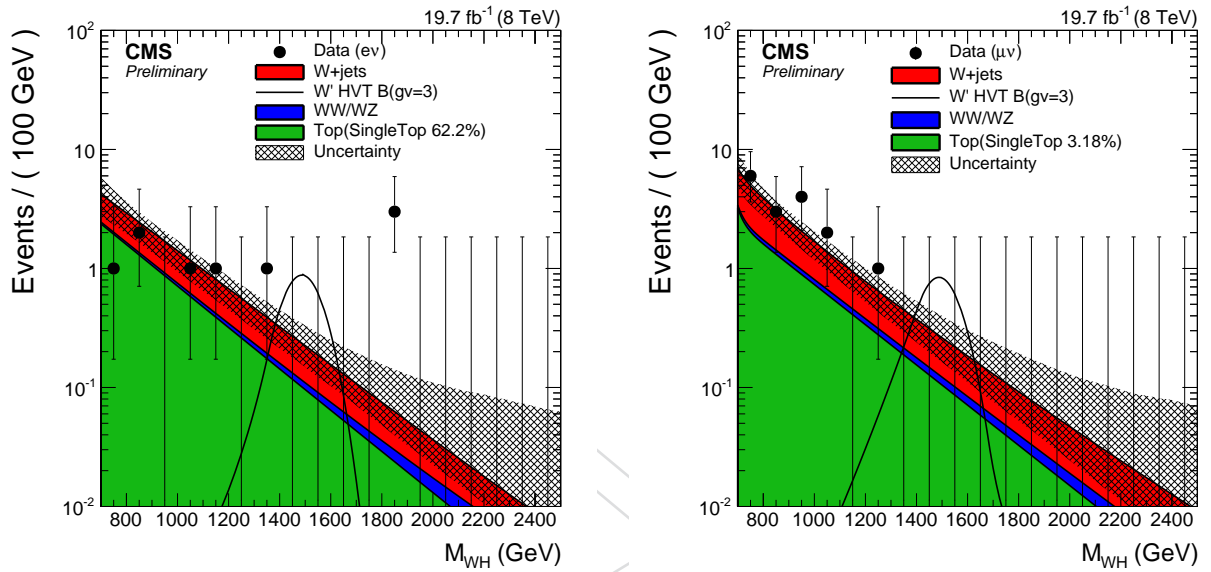


Figure 34: Final estimated background distributions for the electron (left), and muon (right) categories for events in signal region, i.e. in  $110 < m_{jet}^{pruned} < 135$  GeV passing the full selection plus the  $m_{WH} > 800$  GeV requirement. The different background components are shown.

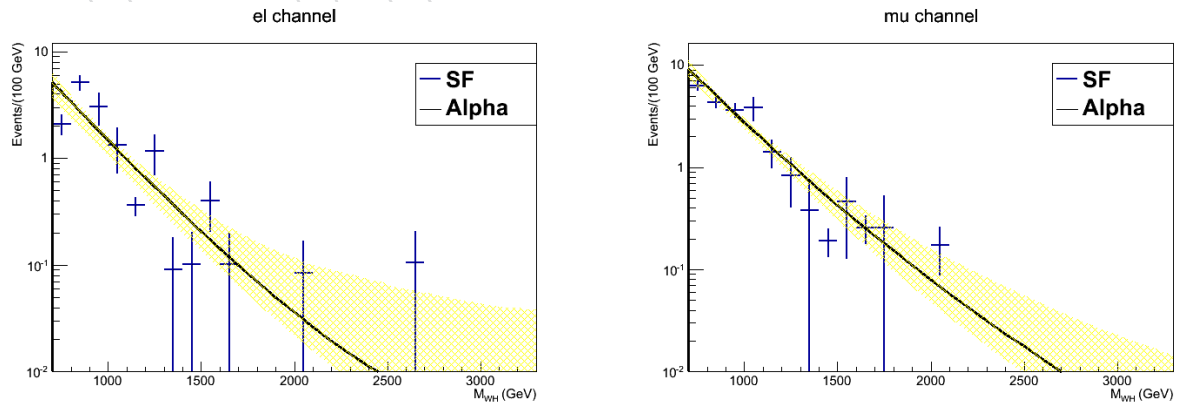


Figure 35: Comparison of the two methods of backgrounds estimation for the electron (left), and muon (right) categories for events with  $110 < m_{jet}^{pruned} < 135$  GeV passing the full selection plus the  $m_{WH} > 800$  GeV requirement. The black line is the background estimated by alpha method and the light yellow shadow is the corresponding uncertainty of this method. The blue crossing line is from SF method estimation.

## 7.4 $t\bar{t}$ control sample

To select a  $t\bar{t}$  enriched control sample we use the standard kinematic pre-selections described in Section 5.4, but changing the requirement on the number of b-tagged AK5 jets: looking at the AK5 reconstructed jets, outside of the CA8 jet cone, it is required at least one CSV “medium” AK5 b-jet. No selection is applied on the pruned jet mass and the whole region [40-150] GeV is considered.

After these selections, data are represented by a pure sample of  $t\bar{t}$  events, with a small contamination from single top, W+jets and VV events. The pruned jet mass, jet  $p_T$ ,  $m_{WH}$ ,  $m_{top}^{leptonic}$  and  $m_{top}^{hadronic}$  variables are shown in Figures 36-38 for both muon and electron channels. In these plots the MC samples are rescaled to the integrated luminosity of the data using the cross sections listed in Table 3. The minor W+jets contribution has been normalized using the same scale factors employed in Section 5.4 (2.02 and 2.08 for muon and electron channel respectively).

A good agreement between the data and the MC is here observed and a data-driven estimation of the  $t\bar{t}$  normalization is considered not necessary. Nevertheless, a background estimation cross check from  $t\bar{t}$  control sample is considered useful to verify the stability of the whole procedure and to derive an uncertainty on the  $t\bar{t}$  normalization.

We derive here W+jets and  $t\bar{t}$  scale factors performing a simultaneous fit of four regions: pruned jet mass low sideband and  $t\bar{t}$  control region in the muon and electron channels separately. In order to simplify, the shapes are not taken into account, but a simple counting is instead performed leaving all the other backgrounds fixed. The obtained scale factors with this method are  $1.18 \pm 0.02$  and  $1.71 \pm 0.02$  for  $t\bar{t}$  and W+jets respectively.

The result is used to assign a systematic uncertainty on the  $t\bar{t}$  cross section of 18% as reported in Section 8.

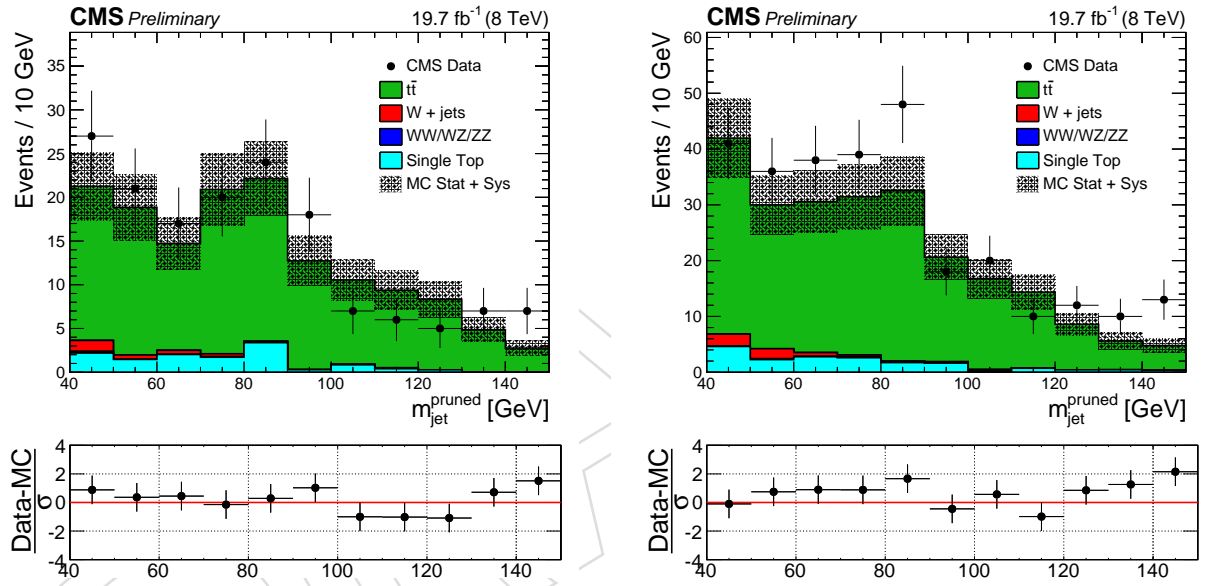


Figure 36: Pruned jet mass for electron channel (left) and muon channel (right) for events with  $40 < m_{jet}^{pruned} < 150$  GeV in the  $t\bar{t}$  control sample.

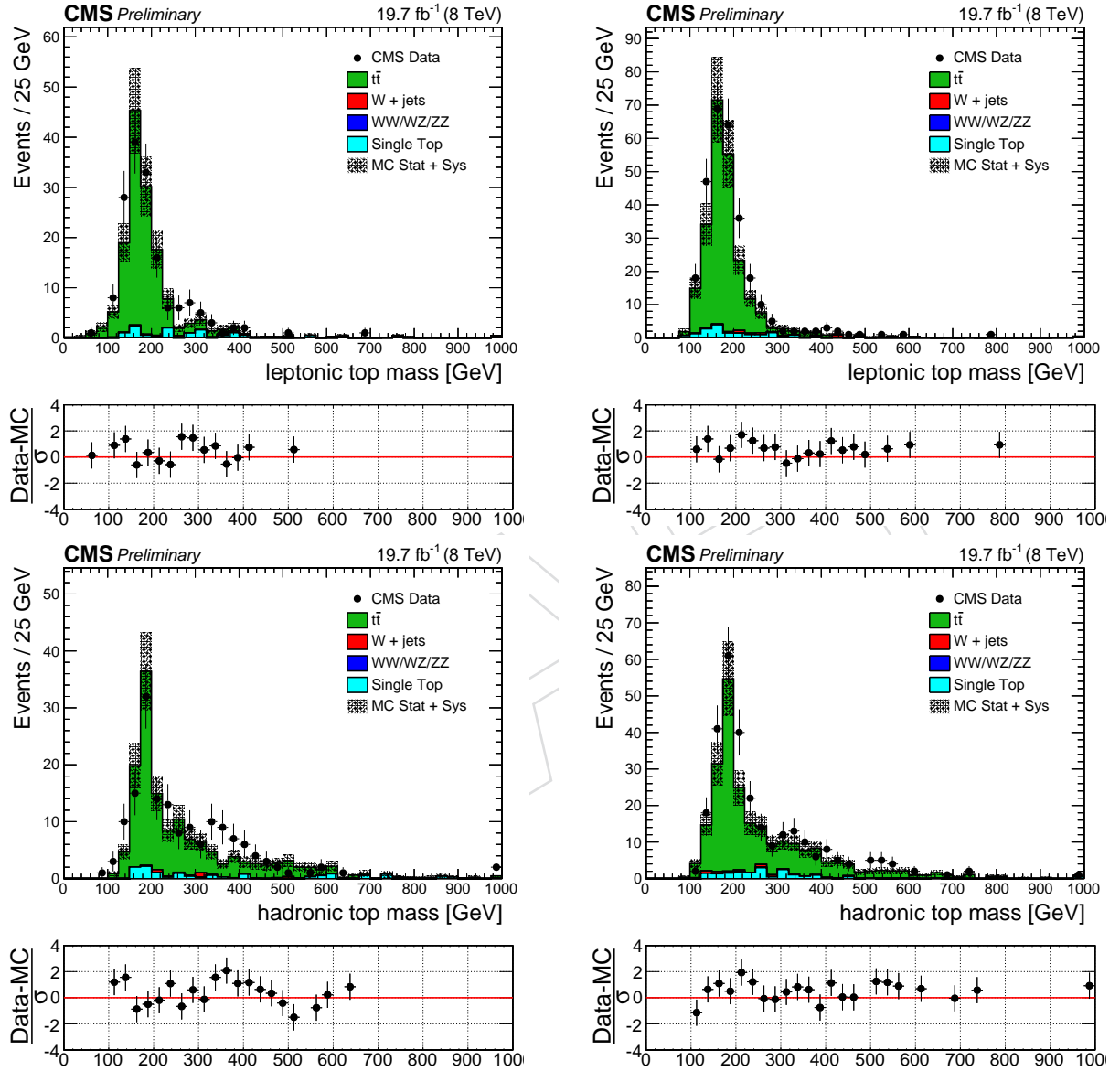


Figure 37:  $m_{top}^{leptonic}$  (top) and  $m_{top}^{hadronic}$  (bottom) for electron channel (left) and muon channel (right) for events with  $40 < m_{jet}^{pruned} < 150$  GeV in the  $t\bar{t}$  control sample.

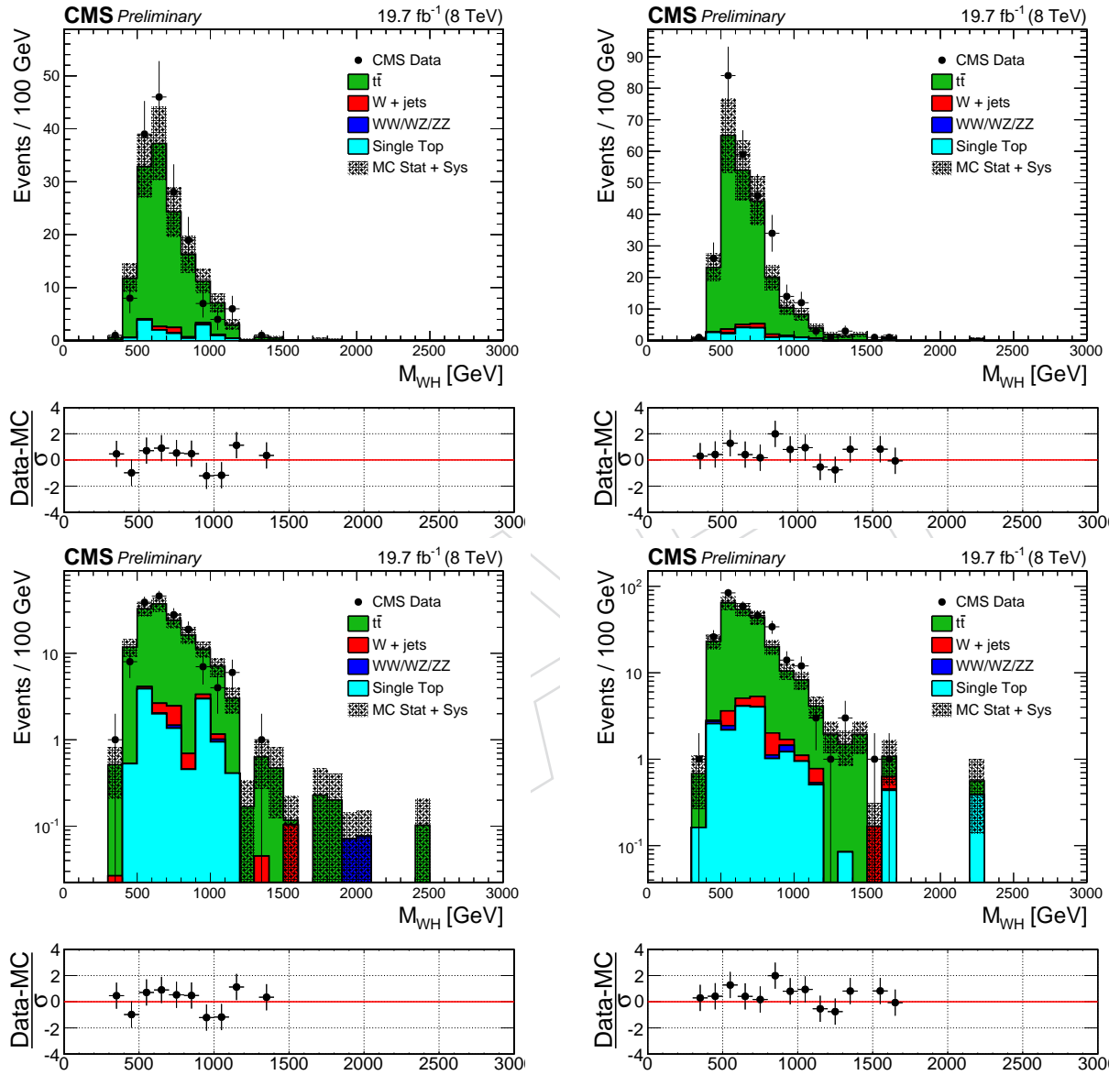


Figure 38:  $m_{WH}$  in linear (top) and log (bottom) scale for electron channel (left) and muon channel (right) for events with  $40 < m_{jet}^{pruned} < 150$  GeV in the  $t\bar{t}$  control sample.

## 8 Systematic Uncertainties

This section describes the systematic uncertainties on signal and background used for the calculation of the limits on cross section and the significance/p-value of potential excesses of events observed in the data.

The uncertainties described below are all included as nuisance parameters in the statistic tool described in Sec. 9.

### 8.1 Background Normalization and Shape

Uncertainties in the estimation of the background affect both the normalization and shape of the  $m_{WH}$  distribution. The systematics uncertainties on the parameters modeling the background shape and background normalization for  $\alpha$  method are described in Sec. 7.2.2.

Uncertainties in the background normalization are mainly statistical in nature and scale with the amount of data in the sideband regions and the number of events in the simulated samples.

#### 8.1.1 Background uncertainties - Method 1

The main uncertainties on the W+jets background modeling (both shape and normalization) come from the number of data events in the  $m_{jet}^{pruned}$  sideband region and the size of the MC sample used to estimate the transport function  $\alpha^{MC}(m_{WH})$ . More details on how these uncertainties are evaluated are discussed in Sec. 7.2.2.

The systematic uncertainty in the  $t\bar{t}$  normalization comes from the data-to-simulation scale factor of 11% derived in the simultaneous fit of a top-quark enriched control sample together with the pruned jet mass low sideband (see Section 7.4). The same uncertainty of 11% is assigned to single top normalization as well. The systematic uncertainty in the VV inclusive cross section is assigned to be 20%, taken from the relative difference in the mean value between the published CMS cross section measurement at  $\sqrt{s} = 8$  TeV and the SM expectation [46].

In the statistical analysis described in Sec. 9, the background parameters are nuisance parameters floating within the range specified by their uncertainties using a Gaussian constraint ("param" option in Higgs combination tool).

### 8.2 Signal Normalization

The following sources of systematic uncertainties affect the number of signal events expected after the final selection (either due to signal efficiency or integrated luminosity uncertainties). The signal efficiency and the integrated luminosity are used as input in the limit and significance calculation.

The uncertainties are calculated separately for electron and muon channel.

#### 8.2.1 Integrated luminosity

The uncertainty on the knowledge of the integrated luminosity of the data sample (2.6%) [47] introduce an uncertainty on the number of signal events passing the final selection.

This uncertainty is fully correlated in all channels.

### 8.2.2 Trigger

The uncertainty on the single lepton trigger efficiency introduces an uncertainty in the signal efficiency. Uncertainties of 1% flat for the electron channel taken from Ref. [28] and  $\eta$ -dependent based on <https://twiki.cern.ch/twiki/bin/viewauth/CMS/MuonReferenceEffs> ranging up to 3.3% are used for the muon channel.

This uncertainty is uncorrelated for different lepton flavors.

### 8.2.3 Lepton energy/momentum scale

The uncertainty on the lepton energy/momentum scale introduces an uncertainty in the signal efficiency. We scale up and down the electron energy by 0.6% for barrel and 1.5% for endcap, taken from Ref. [48], propagate the change to the  $E_T^{\text{miss}}$  reconstructed in the event, and re-run the entire analysis for the signal samples. We check the variation in the number of selected signal events from the nominal value and take the largest relative difference as systematic uncertainty. The same procedure is used also for the muon channel, by scaling up and down the muon momentum by 0.2% (or 5% if the muon  $p_T$  is greater than 200 GeV). A flat value of 1 (2)% for electron (muon) channel is used for all the resonance mass hypotheses.

This uncertainty is uncorrelated for different lepton flavors.

### 8.2.4 Lepton energy/momentum resolution

The uncertainty on the lepton energy/momentum resolution introduces an uncertainty in the signal efficiency. We smear the electron energy momentum 1.3% for barrel and 2.8% for endcap, taken from Ref. [49]. Similarly, for muons we use an uncertainty of 0.2% for momentum scale and 0.6% on  $p_T$  resolution for  $p_T < 200$  GeV and a  $p_T$ -dependent smearing of  $0.05p_T/\text{TeV}$  (5% uncertainty on momentum scale at 1 TeV) beyond (see <https://twiki.cern.ch/twiki/bin/view/CMS/MuonReferenceResolution>). We propagate the change to the  $E_T^{\text{miss}}$  reconstructed in the event, and re-run the entire analysis for the signal samples. We check the variation in the number of selected signal events from the nominal value and take the largest relative difference as systematic uncertainty.

The variation is found to be of the order of 0.4% (0.25%) for electron (muon) channel for all mass hypotheses and therefore negligible for this analysis.

This uncertainty is uncorrelated for different lepton flavors.

### 8.2.5 Lepton identification/isolation efficiency

The uncertainty on the lepton identification and isolation efficiency introduces an uncertainty in the signal efficiency. Based on Ref. [28], a flat 3% uncertainty is used for electron channel. For the muon channel,  $\eta$ -dependent uncertainties based on <https://twiki.cern.ch/twiki/bin/viewauth/CMS/MuonReferenceEffs> are evaluated.

This uncertainty is uncorrelated for different lepton flavors.

### 8.2.6 Jet energy scale

The uncertainty on the jet energy scale introduces an uncertainty in the signal efficiency. We scale up and down the CA8 jet energy accordingly with the uncertainties in the  $p_T$ - and  $\eta$ -dependent jet energy corrections (accessible via the global tag) [50, 51]. At the same time, we scale up and down (in the same direction of the CA8 jet case) also the energy of all the jets in the AK5 collection with  $p_T > 10$  GeV,  $EMEnergyFraction < 0.9$ , and  $MuonEnergyFraction < 0.9$ .



(the latter two cuts to clean the collection from electrons and jets) and propagate the change to the  $E_T^{\text{miss}}$  (the change on the  $E_T^{\text{miss}}$  is calculated only using variations of AK5 jets; this is done in order not to double-count the contribution from CA8 jets matching with AK5 jets). We then re-run the entire analysis for the signal samples using the new jet  $p_T$ 's and new  $E_T^{\text{miss}}$ . We check the variation in the number of selected signal events from the nominal value and take the largest relative difference as systematic uncertainty. The resulting systematic uncertainty depends on the resonance mass ( $M_{\text{res}}$ ), from 1–5%.

This uncertainty is fully correlated in all channels.

### 8.2.7 Jet energy resolution

The uncertainty on the jet energy resolution introduces an uncertainty in the signal efficiency. We smear the CA8 jet energy by the data/MC scale factors for jet energy resolution as a function of jet  $\eta$  following the recommendations reported at <https://twiki.cern.ch/twiki/bin/view/CMS/JetResolution>. At the same time, we scale up and down (in the same direction of the CA8 jet case) also the energy of all the jets in the AK5 collection with  $p_T > 10$  GeV, electromagnetic energy fraction  $< 0.9$ , and muon energy fraction  $< 0.9$  (the latter two cuts to clean the collection from electrons and jets) and propagate the change to the  $E_T^{\text{miss}}$  (the change on the  $E_T^{\text{miss}}$  is calculated only using variations of AK5 jets; this is done in order not to double-count the contribution from CA8 jets matching with AK5 jets). We then re-run the entire analysis for the signal samples using the new jet  $p_T$ 's and new  $E_T^{\text{miss}}$ . We check the variation in the number of selected signal events between the nominal value and the value obtained after the correction for the different jet energy resolution in data and MC. The relative variation is taken as systematic uncertainty. A flat value of 1.5% for both electron and muon channel is used for all the resonance mass hypotheses.

This uncertainty is fully correlated in all channels.

### 8.2.8 Higgs mass tagging efficiency

The Higgs mass tagging uncertainty was obtained from the measurement of the W jet mass [38] and amounts to 1.9%. The difference between light quark and bottom quark fragmentation is already taken into account by the jet energy scale uncertainty discussed above.

### 8.2.9 $E_T^{\text{miss}}$ scale and resolution

Uncertainties in the energy/momentum scale and resolution of leptons, jets in the event are propagated to an uncertainty on the  $E_T^{\text{miss}}$  determination, and therefore introduce an uncertainty in the signal efficiency. The procedure used to include the  $E_T^{\text{miss}}$  uncertainties is described in the sub-sections above for each source of systematic uncertainty. The generic formula for the  $E_T^{\text{miss}}$  after variations is:

$$\vec{E}_T^{\text{miss}}(\text{mod.}) = \vec{E}_T^{\text{miss}} + \sum_{\text{recoobjects}} [\vec{p}_T - \vec{p}_T(\text{mod.})] \quad (14)$$

where “mod.” indicates the new vector after the variations have been applied.

### 8.2.10 b-tagging veto efficiency

For b-tagging veto, we use the Combined Secondary Vertex algorithm with loose working point. We apply b-tagging scale factors using the method 1a) (with event reweighting) described at <https://twiki.cern.ch/twiki/bin/viewauth/CMS/BTagSFMethods>. All MC sam-

ples used in the analysis are re-weighted accordingly with these factors. The MC-based b-tagging efficiencies (for the calculation of the event weights) are obtained from the  $t\bar{t}$  sample used in this analysis as a function of the jet  $p_T$ , jet  $\eta$ , and separately for light and heavy flavour jets. The same efficiencies are used also for the calculation of the weights in the signal samples. The uncertainty in the data/MC scale factors (described at <https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagPOG>, "2012 Data and MC (Moriond13 prescription)" bullet) for light-flavour (u,d,s) and heavy-flavour (c and b) quarks introduces an uncertainty in the weight. We scale up and down the b-tagging weights accordingly with the uncertainties on the scale factors reported at <https://twiki.cern.ch/twiki/bin/viewauth/CMS/BTagSFMETHODS>. We then re-run the entire analysis for the signal samples using the new b-tagging weights. We check the variation in the number of selected signal events between the nominal value and the value obtained with the new weights. The relative variation is taken as systematic uncertainty. The variation is found to be 0.4% for both electron and muon channels, and for all mass hypotheses.

### 8.2.11 Higgs b-tagging identification efficiency

We adopt the same method described in [18]. Two efficiency maps in  $p_T$ - $\eta$  bins are required in order to estimate this systematic: one is created out of the subjects with  $\Delta R$  distance  $> 0.3$ , while the second is created out of the fatjet using only those events where the  $\Delta R$  between the subjects is  $< 0.3$ . The maps are obtained from the signal samples merging all the available mass points and separately for light and heavy flavour subjects/fatjets.

All MC samples used in the analysis are re-weighted using these maps and following the same procedure described in the section above. The same method to compute the b-tagging systematic uncertainty is described above and holds here as well.

The variation is found to be in the range 2-8% depending on the mass hypotheses, while similar numbers are found for the muon and electron channel at the same mass point.

### 8.2.12 Pileup modeling

The MC signal events are reweighted such that the number of pileup interactions matches with what present in the data. The number of pileup interactions in data can be estimated using measurements of instantaneous luminosity and knowing the total inelastic pp cross section at 8 TeV. The uncertainties on these two quantities introduce an uncertainty on the number of pileup interactions considered in the simulation, and therefore on the signal efficiency. We borrow this uncertainty of 0.5% from [27] for this analysis.

## 8.3 Signal Shape

The following sources of systematic uncertainties affect the shape of the  $m_{WH}$  distribution for signal events. The  $m_{WH}$  signal shapes are used in the limit and significance calculation.

The uncertainties on the  $m_{WH}$  signal shape are evaluated at the same time of the uncertainties on the signal normalization described above (including the propagation of the effects to the  $E_T^{\text{miss}}$ ). In all cases, for each variation (lepton or jet energy/resolution) we create the histograms of  $m_{WH}$  signal shapes for all the resonance mass hypotheses considered after significant variation in the  $m_{WH}$  shapes between the two cases.

### 8.3.1 Lepton energy/momentum scale

The uncertainty on the lepton energy/momentum scale introduces an uncertainty in the position and the width of the  $m_{WH}$  signal shape. A systematic uncertainty of  $< 0.3\%$  (1%) on the

mean (width) of the DoubleCB function describing the  $m_{WH}$  signal shape are used for the electron channel. A systematic uncertainty of  $< 0.7\%$  ( $2.5\%$ ) on the mean (width) of the DoubleCB function describing the  $m_{WH}$  signal shape are used for the muon channel.

### 8.3.2 Lepton energy/momentum resolution

The uncertainty on the lepton energy/momentum resolution introduces an uncertainty in the position and the width of the  $m_{WH}$  signal shape. A systematic uncertainty of  $< 0.1\%$  ( $1.2\%$ ) on the mean (width) of the DoubleCB function describing the  $m_{WH}$  signal shape are used for the electron channel. A systematic uncertainty of  $< 0.1\%$  ( $0.8\%$ ) on the mean (width) of the DoubleCB function describing the  $m_{WH}$  signal shape are used for the muon channel.

### 8.3.3 Jet energy scale

The uncertainty on the jet energy scale introduces an uncertainty in the position and the width of the  $m_{WH}$  signal shape. A systematic uncertainty of  $0.5\%$  ( $4\%$ ) on the mean (width) of the DoubleCB function describing the  $m_{WH}$  signal shape are used for the electron channel. A systematic uncertainty of  $0.5\%$  ( $3.5\%$ ) on the mean (width) of the DoubleCB function describing the  $m_{WH}$  signal shape are used for the muon channel.

### 8.3.4 Jet energy resolution

The uncertainty on the jet energy resolution introduces an uncertainty in the position and the width of the  $m_{WH}$  signal shape. A systematic uncertainty of  $< 0.01\%$  ( $0.01\%$ ) on the mean (width) of the DoubleCB function describing the  $m_{WH}$  signal shape are used for both the electron and muon channels. Therefore it is considered negligible for this analysis.

### 8.3.5 Unclustered energy scale

We have evaluated this uncertainty with the following procedure: select all the jets (NOT used for the JES and JER systematics) with  $p_T < 10$  GeV and  $\text{MuFraction} < 0.9$ ; for each event perform the vectorial sum of these jets; the new created jet is assumed to be the contribution from the unclustered energy and a JEC uncertainty of  $10\%$  is assigned to it; the energy of this jet is varied accordingly with its uncertainty and the MET is recalculated; the variation with respect to the default value is used as systematic uncertainty. A systematic uncertainty of  $< 0.01\%$  ( $0.01\%$ ) on the mean (width) of the DoubleCB function describing the  $m_{WH}$  signal shape are used for both the electron and muon channels. Therefore it is considered negligible for this analysis.

### 8.3.6 $E_T^{\text{miss}}$ scale and resolution

Uncertainties in the energy/momentum scale and resolution of leptons, jets in the event are propagated to an uncertainty on the  $E_T^{\text{miss}}$  determination, and therefore introduce an uncertainty in the  $m_{WH}$  signal shape. The procedure used to include the  $E_T^{\text{miss}}$  uncertainties is described in the sub-sections above for each source of systematic uncertainty.

Table 10 summarize the impact of each source of systematic uncertainty on both the signal yields after full selection (both due to uncertainty on signal efficiency or integrated luminosity of the data sample) and on the reconstructed  $m_{WH}$  signal shape. The numbers indicated in this table are included in the limit and significance calculation for all the resonance mass hypotheses and analysis categories considered, unless otherwise noted by a variable range.

Table 10: Summary of sources of systematic uncertainties and their impact on the number of signal events expected after full selection and on the reconstructed  $m_{WH}$  signal shape (peak position and width). The total uncertainty is reported as the sum in quadrature of the different terms. See text for more details.

Source of Syst. Uncert.	Number of Signal Events [%]	$m_{WH}$ Signal Shape	
		Mean [%]	Width [%]
Electron energy (Muon momentum) scale	1 (2)	< 0.3(0.7)	1 (2.5)
Electron energy (Muon momentum) resolution	< 0.4(0.25)	< 0.1	< 1.2(0.8)
Jet energy scale	1–5	0.5	< 4(3.5)
Jet energy resolution	1.5	< 0.15(0.1)	< 3.5(3)
Unclustered energy scale	< 0.01	< 0.01	< 0.01
$E_T^{\text{miss}}$ scale and resolution	Included in lepton/jet/unclust.en. uncert.		
Trigger	1	-	-
Electron (Muon) identification/isolation efficiency	3 (1)	-	-
Higgs mass tagging efficiency	1.9	-	-
b-tagging veto efficiency	0.4	-	-
Higgs b-tagging identification efficiency	2–8	-	-
Pileup modeling	0.5	-	-
Integrated luminosity	2.6	-	-
Total Syst. Uncert.	22–26	1.3(1.6)	9.8

## 9 Results

In this section, we use the signal  $m_{WH}$  shapes presented in Section 6.1, the signal efficiency presented in Section 6.2, the background parameterization introduced in Section 7.2 and the systematic uncertainties discussed in Section 8 to perform a search for new resonances in the  $m_{WH}$  spectrum of data selected accordingly with the criteria summarized in Section 5.

We present the upper limits on  $X \rightarrow WH$  resonance cross section and  $X \rightarrow WH \rightarrow l\nu b\bar{b}$  cross section, using Alpha Method.

### 9.1 Tools for statistical analysis

The limits showed in Section 9.2 are calculated with the modified frequentist paradigm (full  $CL_s$ ) as described in [52, 53].

Systematic uncertainties are treated as nuisance parameters and profiled in the statistical interpretation. The shape analysis that we pursued uses simultaneously the two categories (electron and muon channel separately) of the events, using shapes and normalization predicted for the signal. We tested 13 different mass hypotheses of the HVT B model and LH model separated scanning the mass range  $m_{WH} \in [800, 2000]$  GeV at regular steps of 100 GeV. This choice covers the mass range where previous CMS analysis stopped focusing on the interesting region above the TeV scale where the boosted topology is by far the dominant one. For the implementation of the statistical analysis, we use the Higgs combination tool [54] by calling the following example command:

- **Observed and expected limits:**

```
combine -M HybridNew --frequentist --testStat LHC -n ${label} -m $mass -s $
-d ${datacard}.txt -H ProfileLikelihood
--rMax $maxBoundary --rMin $minBoundary
```

We also show the same final upper limit calculated with the asymptotic approximation [55] of the full  $CL_s$  method in Appendix F.

## 9.2 Limits - Method 1

The background estimation Method 1 is described in Section 7.2.

Figure 39 shows the 95% CL upper limits on the cross section of the process  $pp \rightarrow X \rightarrow WH$  and on the cross section of the process  $pp \rightarrow X \rightarrow WH \rightarrow l\nu b\bar{b}$  combining the two event categories as a function of the resonance mass, while Fig. 42 shows the same limits for the two separate categories. For comparison, the production cross sections times branching ratio of  $X \rightarrow WH$  (and times branching ratio of  $X \rightarrow WH \rightarrow l\nu b\bar{b}$ ) are presented for the benchmark HVT and LH models, [14] and [12, 13].

We also evaluate the local significance of the data under the assumption of a narrow resonance decaying to the  $WH$  final state and lepton universality for the SM  $W$  boson decay, by combining the two event categories. The result is shown in Fig. 40, while the local significances for the two separate categories is shown in Fig. 43. The highest local significance of 2.2 standard deviations is found for a resonance mass of 1.8 TeV. The local significance for a resonance of 1.8 TeV in the electron and muon channels separately are 2.9 and 0.5 standard deviations, respectively.

An estimate of the look-else-where effect by means of pseudo-experiments is shown in Fig. 41. Taking this effect into account, a local significance of 2.9 standard deviations for an excess at a specific resonance mass in a specific channel translates into a global significance of about 1.9 standard deviations searching for resonances over the full mass range 0.8-2.5 TeV in two channels. The result is thus statistically compatible with the standard model expectation within 2 standard deviations.

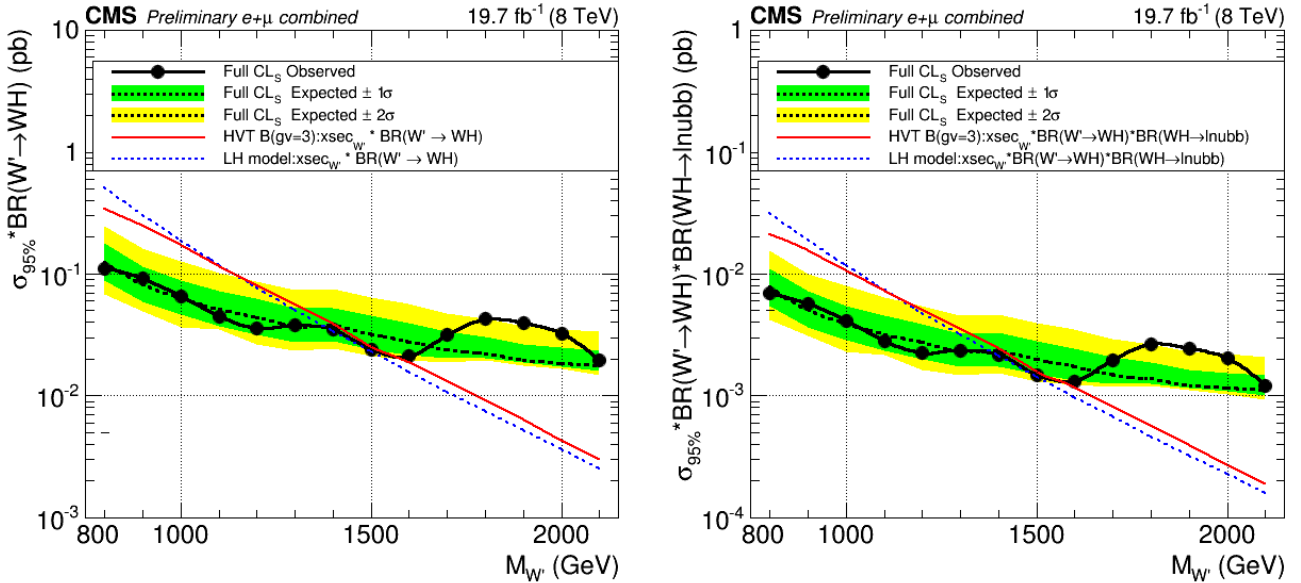


Figure 39: Method 1 - All analysis categories combined - Observed (dashed) and expected (solid) 95% CL upper limit on the product of the  $W'$  production cross section times the branching fraction of  $X \rightarrow WH$  (left), or times the branching fraction of  $X \rightarrow WH \rightarrow l\nu b\bar{b}$  (right) using  $19.7 \text{ fb}^{-1}$  of data. The 68% and 95% ranges of expectation are also shown with green and yellow bands. The theoretical cross section times the corresponding BR for  $X \rightarrow WH$  or  $X \rightarrow WH \rightarrow l\nu b\bar{b}$  are shown as a red solid (blue dashed) curve for HVT (LH) model.

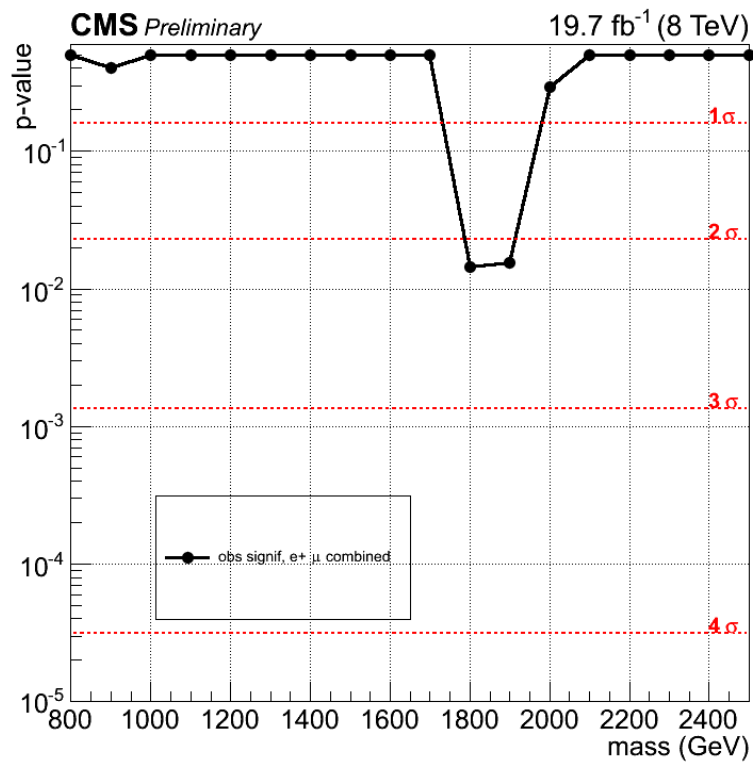


Figure 40: Method 1 - All analysis categories combined p-value performance, based on Profile-Likelihood.

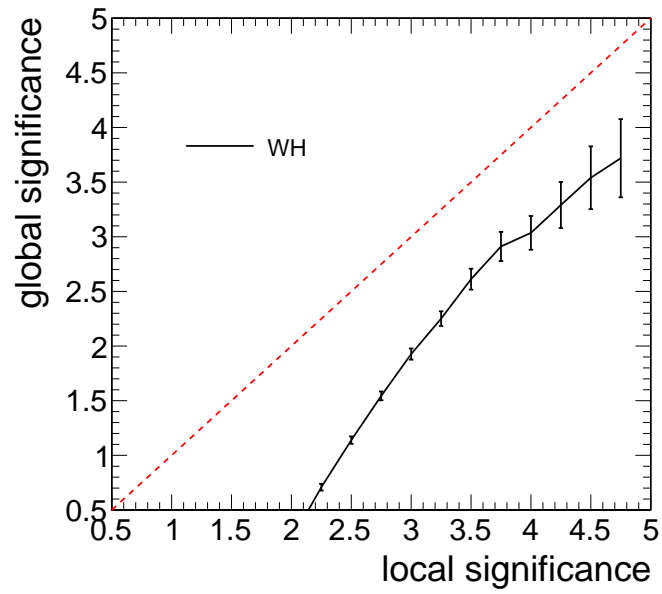


Figure 41: Estimate of the look-else-where effect for this search. Shown is the global significance as a function of the local significance which corresponds to the maximal significance in the  $M_{WH}$  distribution in the mass ranges 0.8-2.5 TeV in the two categories. The global significance is estimated with a frequentist approach using background-only toys and corresponds to the fraction of toys (translated from a p-value to significance) with at least a certain local significance in the  $M_{WH}$  distribution in the two categories.



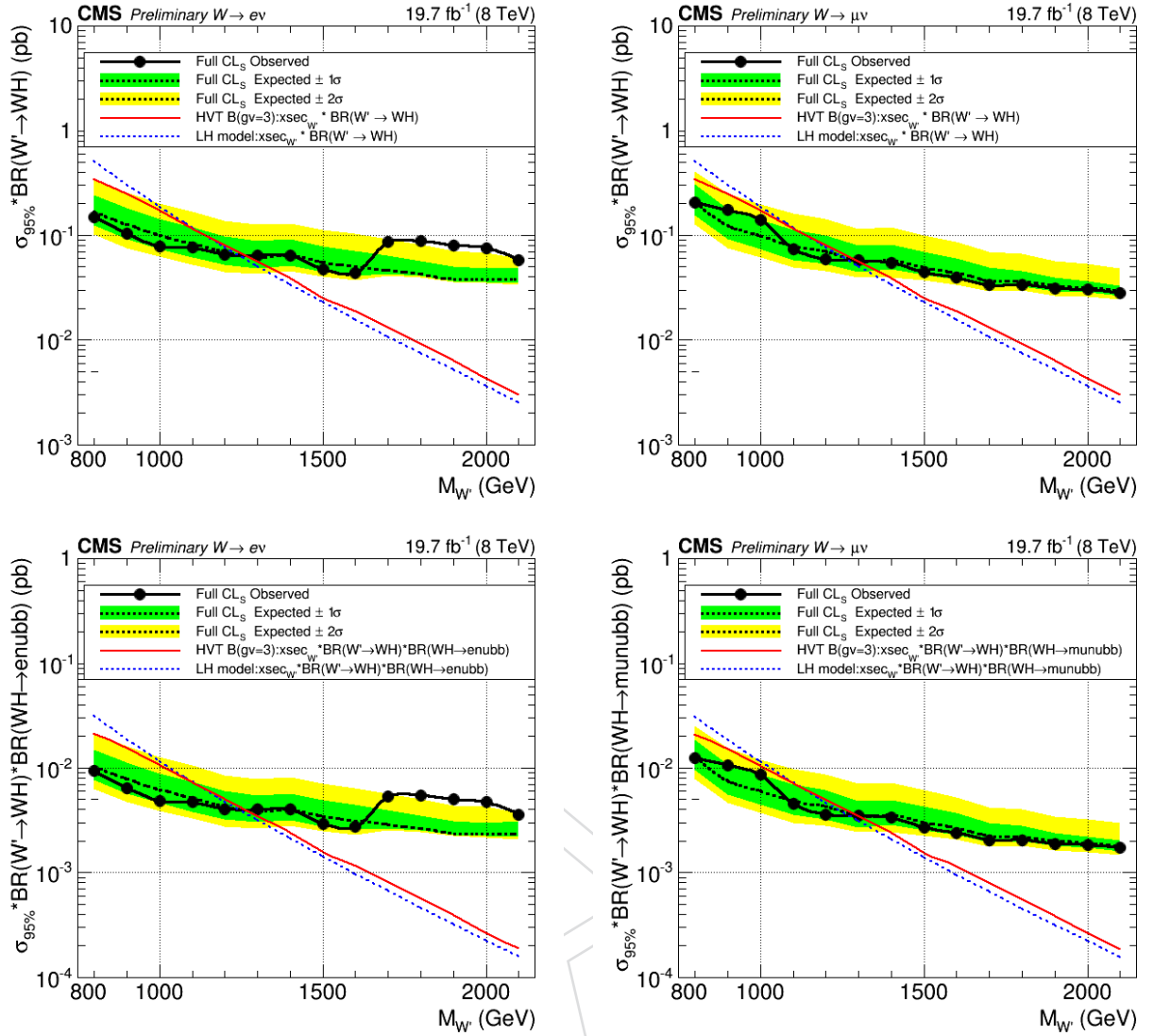


Figure 42: Method 1 - Ele (left), Mu (right) - Observed (dashed) and expected (solid) 95% CL upper limit on the product of the  $W'$  production cross section and the branching fraction of  $X \rightarrow WH$  (top) or the branching fraction of  $X \rightarrow WH \rightarrow l\nu_{bb}$  (bottom) using 19.7  $\text{fb}^{-1}$  of data. The 68% and 95% ranges of expectation are also shown with green and yellow bands. The theoretical cross section times BR for  $X \rightarrow WH$  is shown as a red solid (blue dashed) curve for HVT (LH) model.

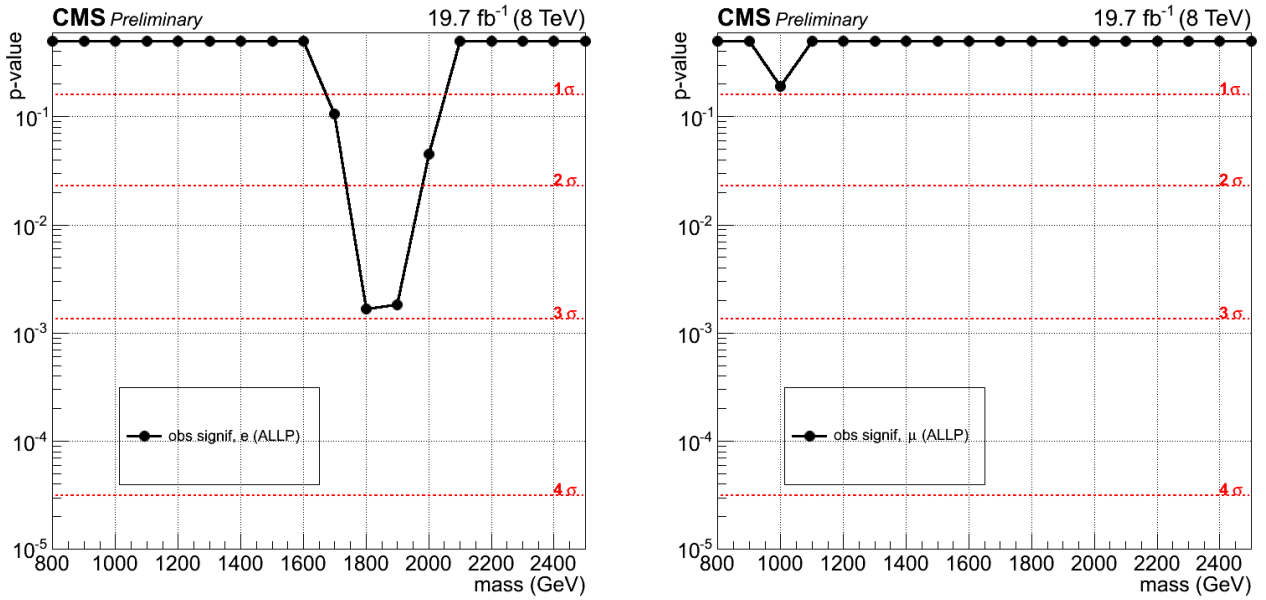


Figure 43: p-value performance for electron (left) and muon (right) channels.

## 10 Conclusions

A search for new particles decaying to W boson and Higgs boson with subsequent decay to a final state containing one charge lepton, one neutrino, and two b quarks,  $X \rightarrow WW \rightarrow b\bar{b}\ell\nu$ , is presented. Results are based on data corresponding to an integrated luminosity of  $19.7 \text{ fb}^{-1}$  of proton-proton collisions at  $\sqrt{s} = 8 \text{ TeV}$  and collected with the CMS detector at the CERN LHC.

Jet sub-structure and subjet b-tagging techniques are exploited for separating the signal from the SM background when the boost of the H causes the two b quarks to merge into the same jet reconstructed in the detector.

We set upper limits at 95% CL on the  $X \rightarrow WH$  cross section, which can be interpreted as generic limits on narrow resonances X decaying into the same final state.

DRAFT

## A $\tau_{21}$ optimization

From [27] we can see the 2- to 1-subjettiness ratio,  $\tau_{21}$ , is a key variable that it can discriminate between jets coming from hadronic decays of boosted boson and QCD. But in this analysis, we find that its performance is not important anymore, and make no sense for the final limit. The details are in the following.

In this section we use same strategy in [27] to study  $\tau_{21}$  performance. And all the results are based on top of the selection criteria listed in Sec. 5. For convenience, here we write again the whole procedure. we choose as a figure of merit the quantity suggested in Ref. [39] ("Punzi's significance" with  $a = b = 2$ , from Eq. 7 of the reference cited):

$$P = \frac{\epsilon_S}{1 + \sqrt{B}} \quad (15)$$

where  $\epsilon_S$  is the signal selection efficiency and  $B$  is the remaining background.

- merge the electron and muon channels together to do this optimization study;
- set a window of  $\pm 15\%$  around the target resonance mass;
- histogram the  $\tau_{21}$  variable for signal events (for the chosen mass) and for background events which pass all the final selection requisites including the signal region  $m_{jet}^{pruned}$  cut and also have a reconstructed resonances mass  $m_{WW}$  in the window defined above.
- set an upper threshold  $\tau_{21}^{\max}$  and integrate the  $\tau_{21}$  distributions of signal and background up to the threshold. The value obtained for the signal sample is proportional to the signal efficiency  $\epsilon_S$ , while the value obtained for the background sample is an estimate of  $B$ ;
- calculate the significance  $P$ .

In order to increase the MC statistics at high  $m_{WH}$  values, the W+jets sample `WJetsToLNu_PtW-180_TuneZ2star_8TeV-madgraph-tarball` is used. The results are shown in Figure 44, 45, 46 and 47 for target resonance masses of  $M'_W = 1000, 1500, 2000$  and  $2500$  GeV.

From Figure 44, 45, 46 and 47, we can see that because of limited statistics, especially at high mass point,  $\tau_{21}$  distribution does not show obvious tendency for signal samples and background samples. Figure 48 can further show that  $\tau_{21}$  cut make no sense for the final limit in this analysis. Thus, in this analysis we give up this cut.

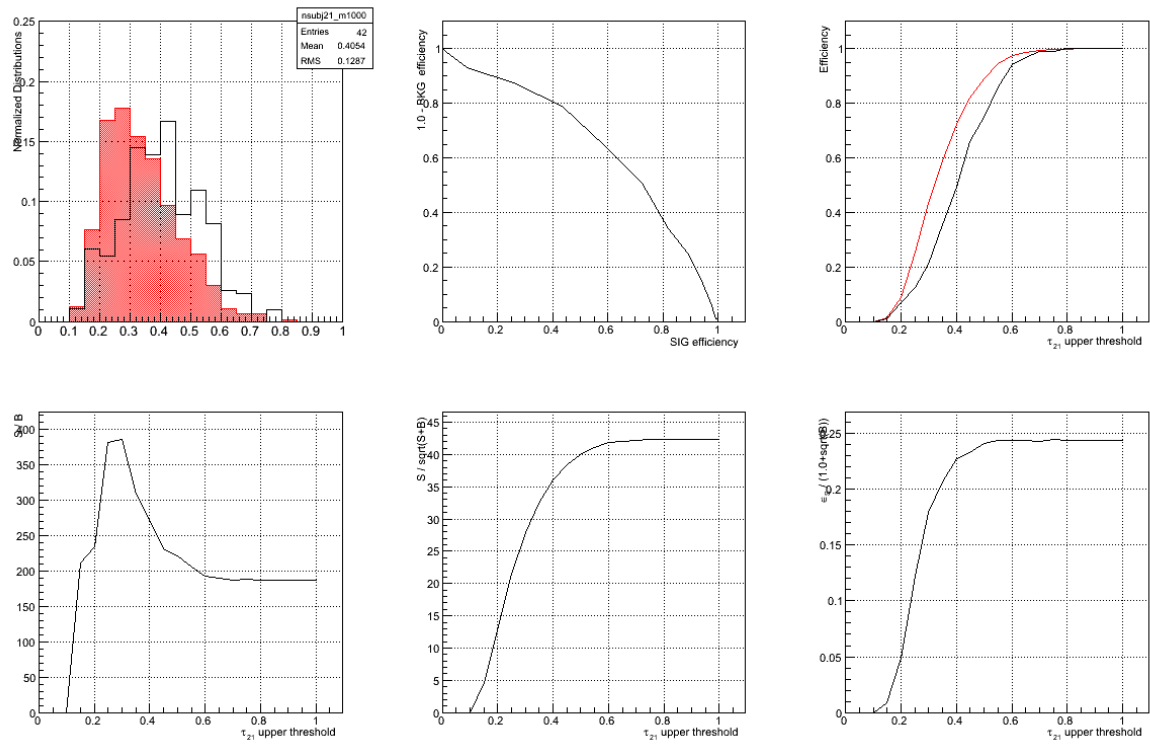


Figure 44: Steps of optimization for  $M'_W = 1000$  GeV. Top left: distributions of  $\tau_{21}$  for signal and background (the red filled is signal and the black line is background). Top middle: signal efficiency  $\times$  background rejection curve (ROC curve). Top right: signal (red) and background (black) efficiencies as function of  $\tau_{21}$  cut. Bottom left: signal  $S$  over background  $B$  as function of  $\tau_{21}$  cut. Bottom middle:  $S/\sqrt{S+B}$  as function of  $\tau_{21}$  cut. Bottom right: Punzi's significance (as defined in Eq.15) as function of  $\tau_{21}$  cut.

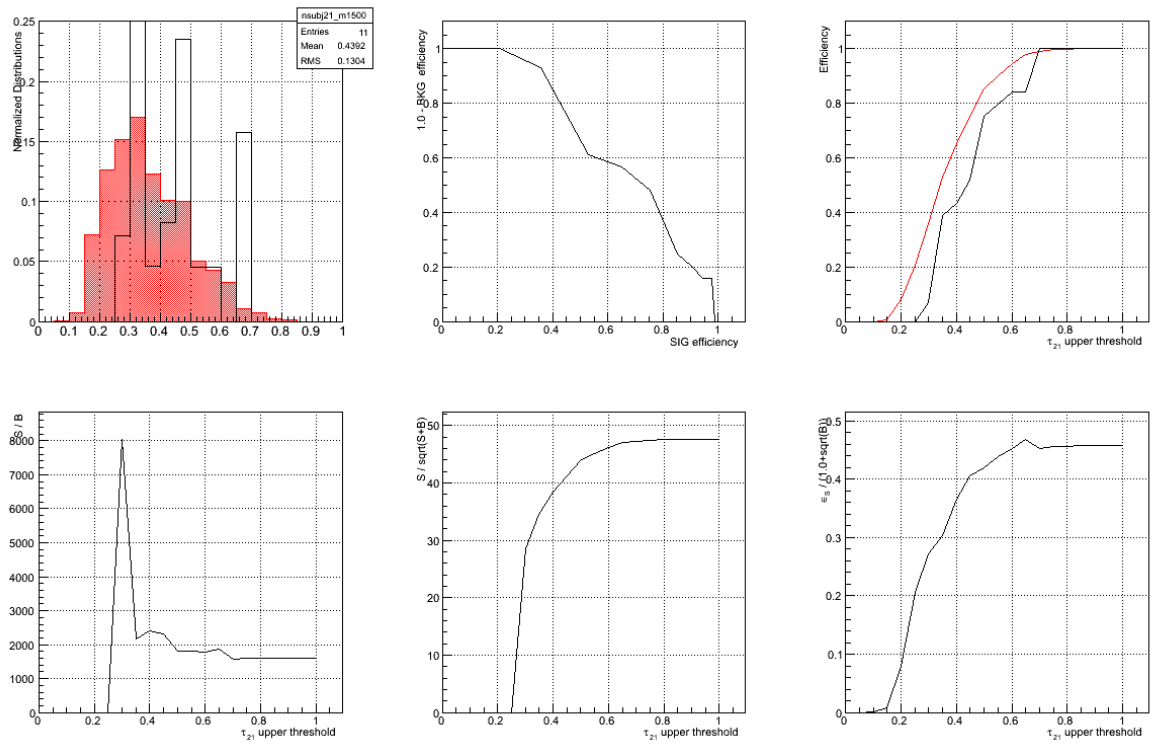


Figure 45: Steps of optimization for  $M'_W = 1500$  GeV. Top left: distributions of  $\tau_{21}$  for signal and background (the red filled is signal and the black line is background). Top middle: signal efficiency  $\times$  background rejection curve (ROC curve). Top right: signal (red) and background (black) efficiencies as function of  $\tau_{21}$  cut. Bottom left: signal  $S$  over background  $B$  as function of  $\tau_{21}$  cut. Bottom middle:  $S/\sqrt{S+B}$  as function of  $\tau_{21}$  cut. Bottom right: Punzi's significance (as defined in Eq.15) as function of  $\tau_{21}$  cut.

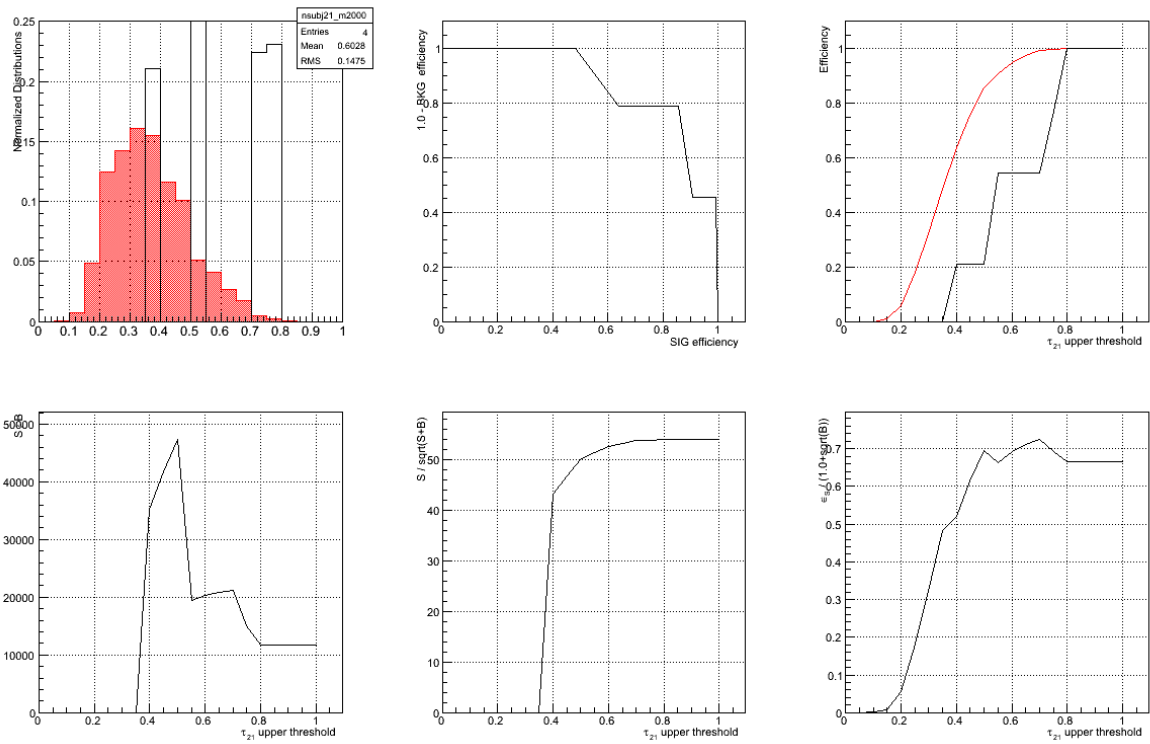


Figure 46: Steps of optimization for  $M'_W = 2000$  GeV. Top left: distributions of  $\tau_{21}$  for signal and background (the red filled is signal and the black line is background). Top middle: signal efficiency  $\times$  background rejection curve (ROC curve). Top right: signal (red) and background (black) efficiencies as function of  $\tau_{21}$  cut. Bottom left: signal  $S$  over background  $B$  as function of  $\tau_{21}$  cut. Bottom middle:  $S/\sqrt{S+B}$  as function of  $\tau_{21}$  cut. Bottom right: Punzi's significance (as defined in Eq.15) as function of  $\tau_{21}$  cut.

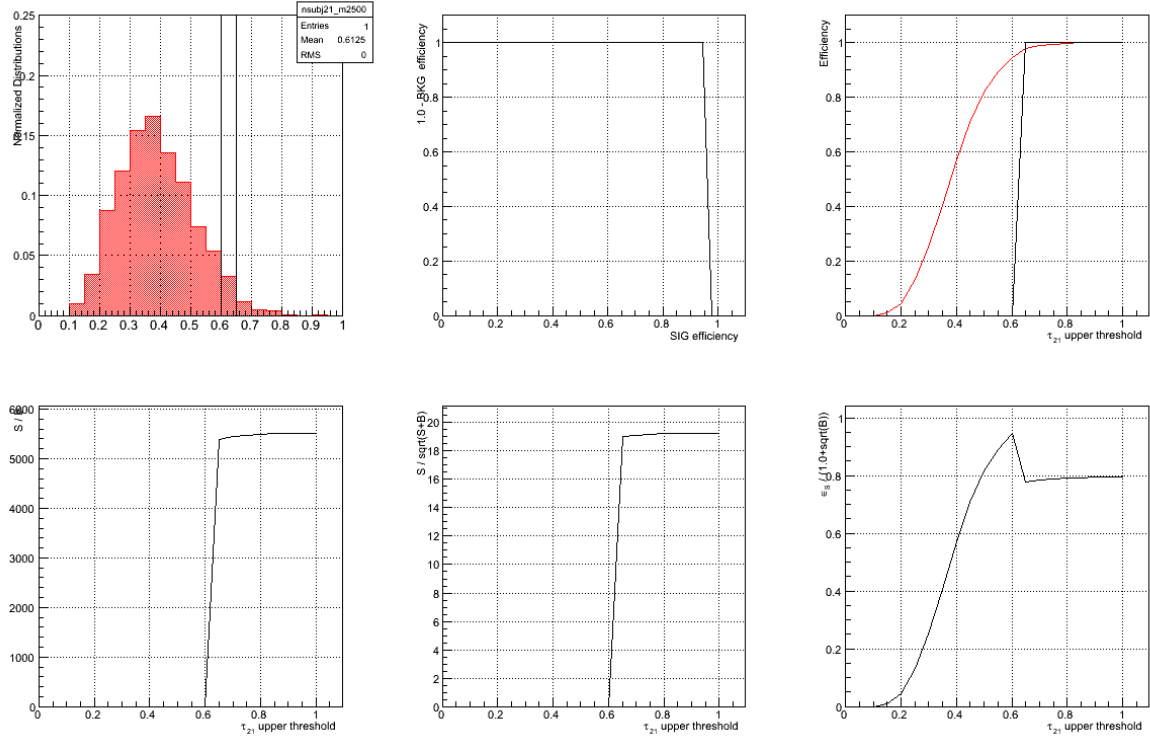


Figure 47: Steps of optimization for  $M'_W = 2500$  GeV. Top left: distributions of  $\tau_{21}$  for signal and background (the red filled is signal and the black line is background). Top middle: signal efficiency  $\times$  background rejection curve (ROC curve). Top right: signal (red) and background (black) efficiencies as function of  $\tau_{21}$  cut. Bottom left: signal  $S$  over background  $B$  as function of  $\tau_{21}$  cut. Bottom middle:  $S/\sqrt{S+B}$  as function of  $\tau_{21}$  cut. Bottom right: Punzi's significance (as defined in Eq.15) as function of  $\tau_{21}$  cut.

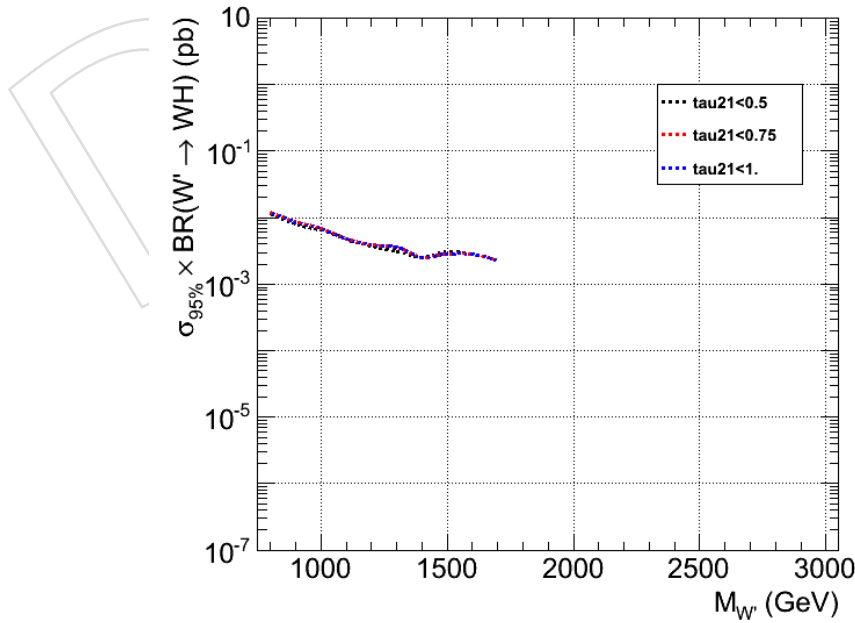


Figure 48: Final limit plot comparison with  $\tau_{21} < 0.5$  (black line),  $\tau_{21} < 0.75$  (red line) and  $\tau_{21} < 1$  (blue line), using SF method introduced in Section 7.3, from 800GeV to 1700GeV.



## B $t\bar{t}$ rejection optimization

In this section the different strategies tested in order to reject  $t\bar{t}$  events are described, following the discussion made on the event topology in Section 5.3. The options considered are the following

- BDT with different sets of input variables
- combination of b-tag veto with BDT
- combination of b-tag veto with a veto on the reconstructed hadronic and/or leptonic top mass
- veto on the total number of extra jets

In order to estimate the potential separation power of several variables and the correlation between them 100 Boosted Decision Trees have been trained on the  $t\bar{t}$  and on the 1 TeV signal samples. Both the samples include only the events that pass the full selection excluding the b-tag veto and the Higgs b-tagging, and selecting the pruned jet mass region [110,150] GeV, in order to gain enough statistics for the training (the shape of the input variables considered for the training do not exhibit any bias due to the excluded selections). Furthermore, only events with at least one extra AK5 jet (as defined in 3.4.2) outside of the fat jet cone of 0.8 are considered. This condition ensures the validity of the considered input variables.

The input variables considered for the training are listed in Table 11 and shown in Figures 49 to 52 for the signal and for the background. The invariant mass of the leptonic and hadronic top quarks has been reconstructed using the algorithm described in Section 5.3.

Table 11: Input variables considered for the BDT training.

$p_T^{AK5}$	$p_T$ of the leading AK5 jet
$CSV^{AK5}$	CSV of the leading AK5 jet
$M_{ht}$	invariant mass of the hadronic top
$M_{lt}$	invariant mass of the leptonic top
$\Delta\phi(jet^{AK5}, jet^{CA8})$	$\Delta\phi$ between the CA8 jet and the closest AK5 jet
$\Delta R(jet^{AK5}, jet^{CA8})$	$\Delta R$ between the CA8 jet and the closest AK5 jet
$\Delta\phi(l, jet^{AK5})$	$\Delta\phi$ between the lepton and the closest AK5 jet
$\Delta R(l, jet^{AK5})$	$\Delta R$ between the lepton and the closest AK5 jet

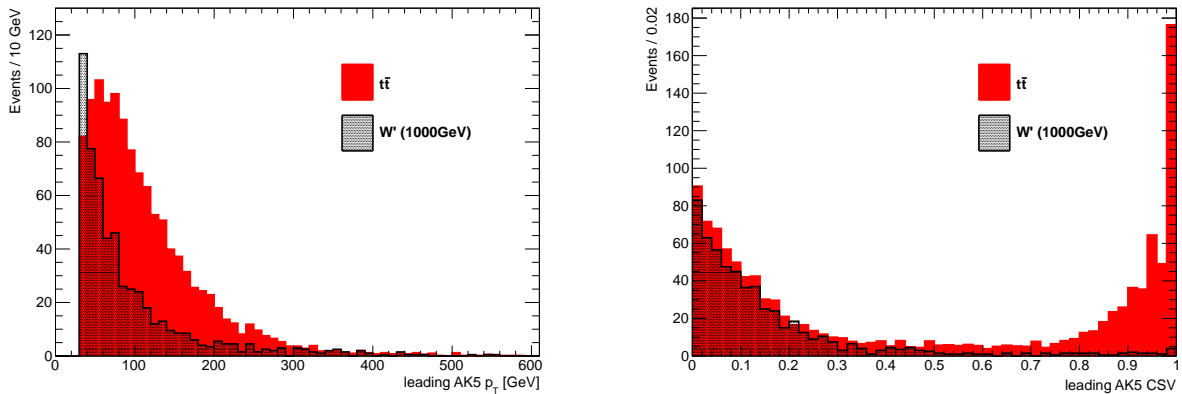


Figure 49: Left:  $p_T$  of the leading AK5 jet. Right: CSV of the leading AK5 jet.

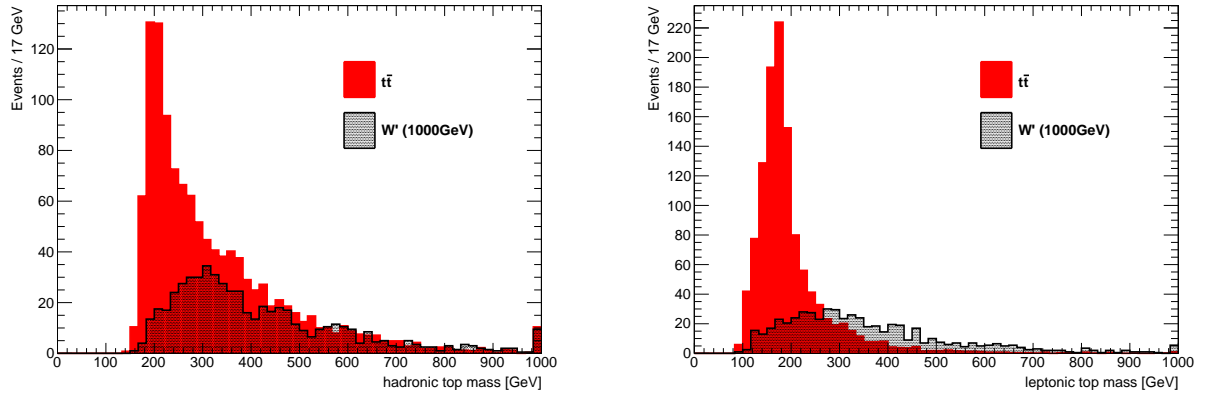


Figure 50: *Left*: hadronic top invariant mass. *Right*: leptonic top invariant mass.

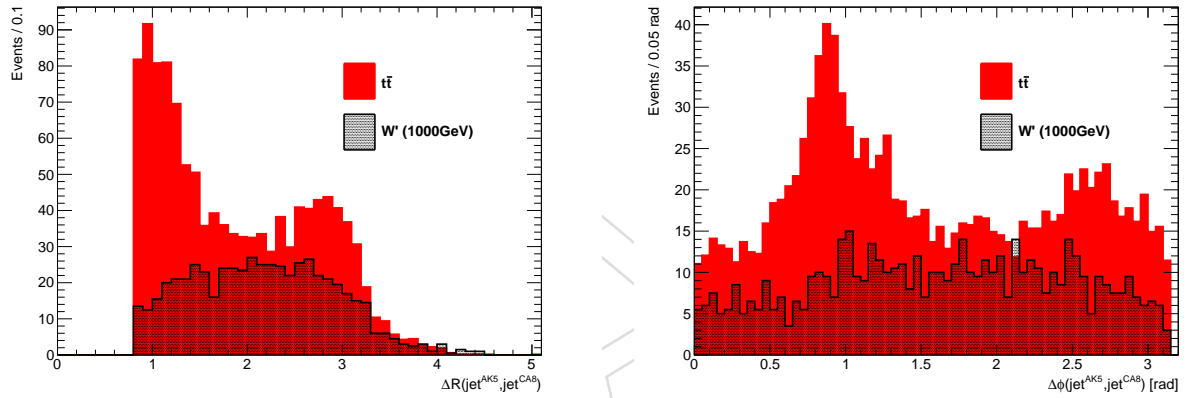


Figure 51: *Left*:  $\Delta R$  between the CA8 jet and the closest AK5 jet. *Right*:  $\Delta\phi$  between the CA8 jet and the closest AK5 jet.

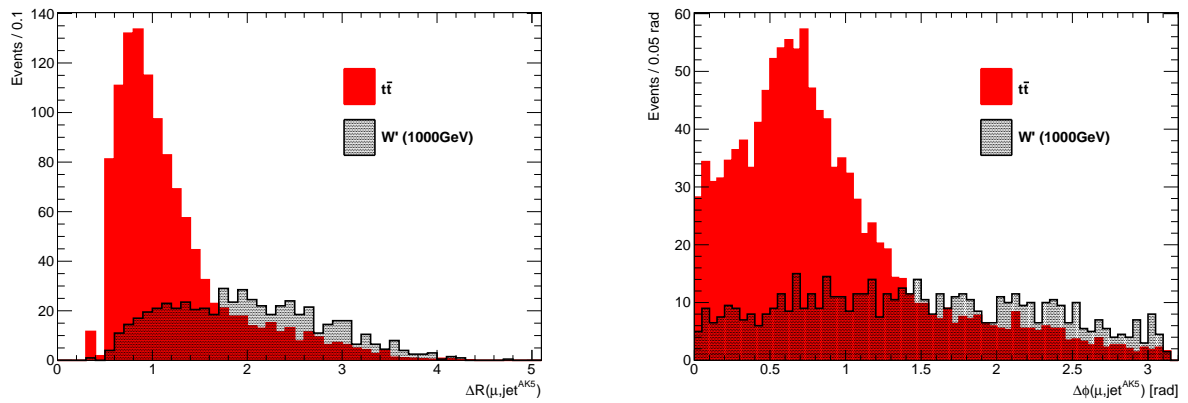


Figure 52: *Left*:  $\Delta R$  between the lepton and the closest AK5 jet. *Right*:  $\Delta\phi$  between the lepton and the closest AK5 jet.

The final set of input variables have been optimized after studying the separation power of each tested set, as represented by the ROC curves in Figure 53. Table 12 summarizes the  $t\bar{t}$  rejection at the signal efficiency of 80% for different sets of BDT input variables.

The BDT has been tested and trained on two independent subset of events and the response is shown in Figure 54. A veto is then applied on different values of the BDT outputs until an optimum working point is found, that maximizes the “Punzi’s significance”  $P$ . Other alternative options to the BDT for  $t\bar{t}$  rejection have been tested, the significance computed and compared to the BDT veto performance. The final results are listed in Table 13 where the optimization is performed on the 1 TeV sample only.

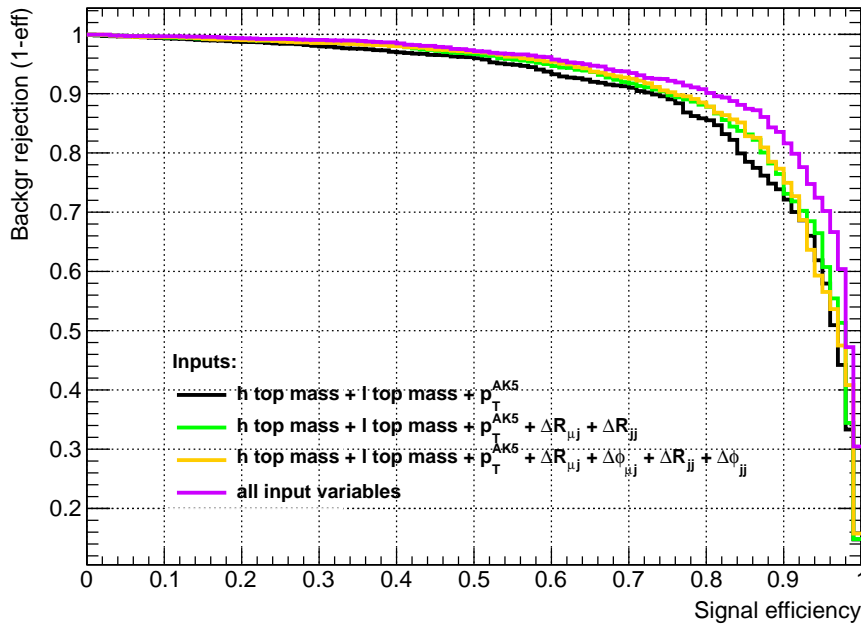


Figure 53: ROC curves showing the separation power between signal and  $t\bar{t}$  background of the BDT after training on different sets of input variables.

Table 12:  $t\bar{t}$  background rejection at at the signal efficiency of 80% for different sets of BDT input variables.

	Input variables	$t\bar{t}$ rejection
Set 1	$M_{ht}, M_{lt}, p_T^{AK5}, CSV^{AK5}$ $\Delta\phi(jet^{AK5}, jet^{CA8}), \Delta R(jet^{AK5}, jet^{CA8}),$ $\Delta\phi(l, jet^{AK5}), \Delta R(l, jet^{AK5})$	90%
Set 2	$M_{ht}, M_{lt}, p_T^{AK5},$ $\Delta\phi(jet^{AK5}, jet^{CA8}), \Delta R(jet^{AK5}, jet^{CA8}),$ $\Delta\phi(l, jet^{AK5}), \Delta R(l, jet^{AK5})$	88%
Set 3	$M_{ht}, M_{lt}, p_T^{AK5},$ $\Delta R(jet^{AK5}, jet^{CA8}), \Delta R(l, jet^{AK5})$	88%
Set 4	$M_{ht}, M_{lt}, p_T^{AK5}$	86%

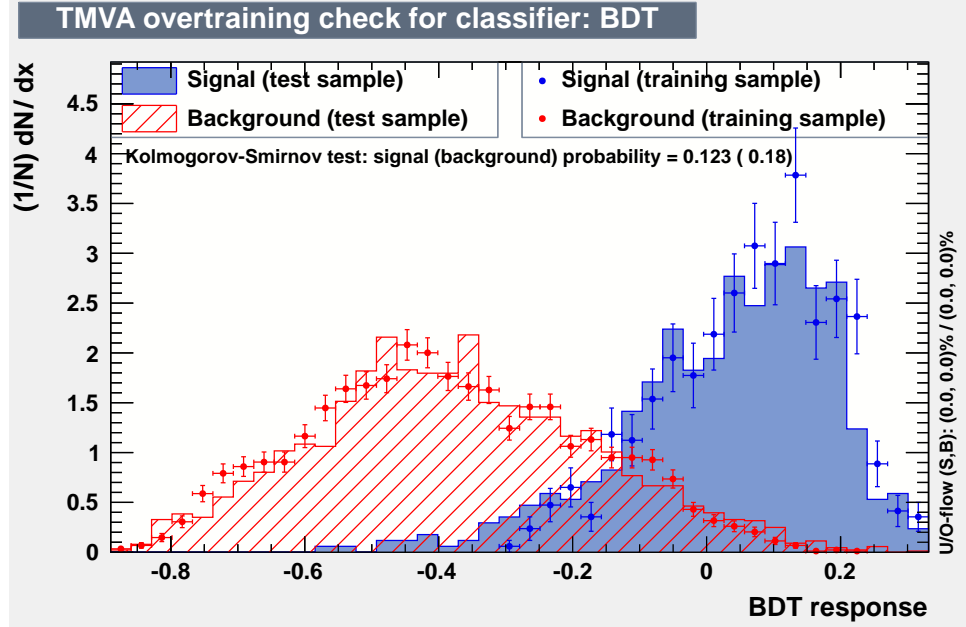


Figure 54: BDT response on training and test samples.

Table 13: Comparison between BDT veto performances and alternative options for  $t\bar{t}$  rejection in terms of “Punzi’s significance”.

Veto	BDT input variables	$P$
b-tag veto	-	$0.030 \pm 0.003$
b-tag veto + ( $150 < M_{lt} < 220 \text{ GeV} \parallel$ $150 < M_{ht} < 300 \text{ GeV}$ )	-	$0.038 \pm 0.004$
BDT	Set 1	$0.043 \pm 0.004$
b-tag veto + BDT	Set 3	$0.044 \pm 0.004$
b-tag veto + BDT	Set 4	$0.042 \pm 0.004$
0 extra jets	-	$0.036 \pm 0.003$

Since the improvement obtained adding the  $\Delta R$ 's and the CSV to the set of BDT input variables is not particularly sensitive we decided to keep the minimal set (Set 4) and to combine the veto on the "kinematics-only" BDT output with the b-tag veto. With this set of variables a new BDT has been trained and tested on the  $t\bar{t}$  sample and on several signal samples at different mass points (1, 1.5, 2, 2.5 TeV). This choice is made to avoid a bias of the BDT response toward a particular mass. The BDT has been then evaluated on each of these mass points and a common optimum working point has been found. Table 14 reports the significances for each mass point comparing four different options: b-tag veto only, b-tag + "kinematics-only" BDT veto, b-tag + top mass veto. In the table the significances are computed in a region of  $\pm 15\%$  around the target resonance mass. The results for 2.5 TeV mass point are not shown because no  $t\bar{t}$  events are present in that region.

Since the sensitivity reached with the BDT veto is comparable with the one reached with the top mass veto the choice was made to include the top mass veto in the selections. The lower and upper limits of the hadronic and leptonic top mass for the veto have been optimized on the three mass points of 1, 1.5 and 2 TeV giving the following optimum choice

$$m_{top}^{leptonic} < 120 \parallel m_{top}^{leptonic} > 240$$

$$m_{top}^{hadronic} < 160 \parallel m_{top}^{hadronic} > 280$$

Table 14: Comparison between BDT veto performances and alternative options for  $t\bar{t}$  rejection in terms of "Punzi's significance".

Veto	1 TeV	1.5 TeV	2 TeV
b-tag veto	$0.039 \pm 0.002$	$0.099 \pm 0.008$	$0.1532 \pm 0.019$
b-tag veto + ( $150 < M_{lt} < 220$ GeV $\parallel$ $150 < M_{ht} < 300$ GeV)	$0.047 \pm 0.003$	$0.102 \pm 0.009$	$0.158 \pm 0.020$
b-tag veto + BDT	$0.050 \pm 0.003$	$0.115 \pm 0.011$	$0.169 \pm 0.024$

## C Closure test for alpha method

We perform a validation of the background estimation method described in Section 7.2, in particular in Section 7.2.1 for the W+jets background estimation.

In the main analysis, we use the data in the low-mass sideband region ( $40 < m_{jet}^{pruned} < 110$  GeV) to predict the total background in the signal region ( $110 < m_{jet}^{pruned} < 135$  GeV).

$$\begin{aligned} N_{tot}^{signal}(m_{WH}) &= N_{W+jets}^{signal}(m_{WH}) + N_{OB}^{signal}(m_{WH}) \\ &= N_{data}^{sideband}(m_{WH}) \times (1 - R_0(m_{WH})) \times \alpha^{MC}(m_{WH}) \times F_{W+jets} \\ &\quad + N_{OB}^{signal}(m_{WH}) \end{aligned} \quad (16)$$

This method relies on the assumption that the correlation between the reconstructed  $m_{WH}$  and  $m_{jet}^{pruned}$  are reasonably well modeled by the simulation. The validity of the method assumptions can be tested using real data in a control region dominated by W+jets events, with a small expected signal contamination.

For this closure test, we define new sideband and signal regions:

- low-mass sideband region ( $40 < m_{jet}^{pruned} < 80$  GeV);
- high-mass sideband region ( $135 < m_{jet}^{pruned} < 150$  GeV);
- “signal region” ( $80 < m_{jet}^{pruned} < 110$  GeV).

The possible contamination from a  $W' \rightarrow WZ \rightarrow l\nu bb$  signal in the closure test signal region has been evaluated, using the efficiency on a simulated sample and the current limits set by other analyses. In particular, we used the sample /WprimeToWZ\_M-1500\_TuneZ2star\_8TeV-pythia6/Summer12\_DR53X-PU\_S10\_START53\_V7A-v1/AODSIM available on the DAS, and the limits set by EXO-13-009[56] for a neutral resonance of a mass of 1.5 TeV decaying into two W bosons (we assume here degeneracy between the neutral and the charged resonances as pointed out in the Section 1). The expected signal efficiency in the closure test signal region is then expected to be about 0.9%, corresponding to a contamination of less than 2 events, thus considered negligible for the closure test.

The goal of the closure test is to predict the number of data events in the signal region using the data in the low-mass sideband region, following the same procedure employed in the main analysis (Section 7.2).

- The W+jets normalization ( $F_{W+jets}$ ) is obtained from fits to the  $m_{jet}^{pruned}$  distribution as described in Section 7.2.1. The fit is performed using both low- and high- mass sideband regions defined above, and the number of W+jets in the “signal” region is obtained. Figure 55 show the  $m_{jet}^{pruned}$  distributions for the 2 categories with the fit results.
- The  $\alpha^{MC}(m_{WH})$  functions are shown in Figure 56 for the 2 categories.
- The total background prediction is shown for the 2 analysis categories in Figures 57 (log scale), together with the actual data in the signal region. A good agreement between the data and the background prediction is observed, both in the shape and the normalization of the distributions.

This cross check confirms that the alpha method to extract the W+jets background is reliable and can be used to search for new  $X \rightarrow WH$  resonances.

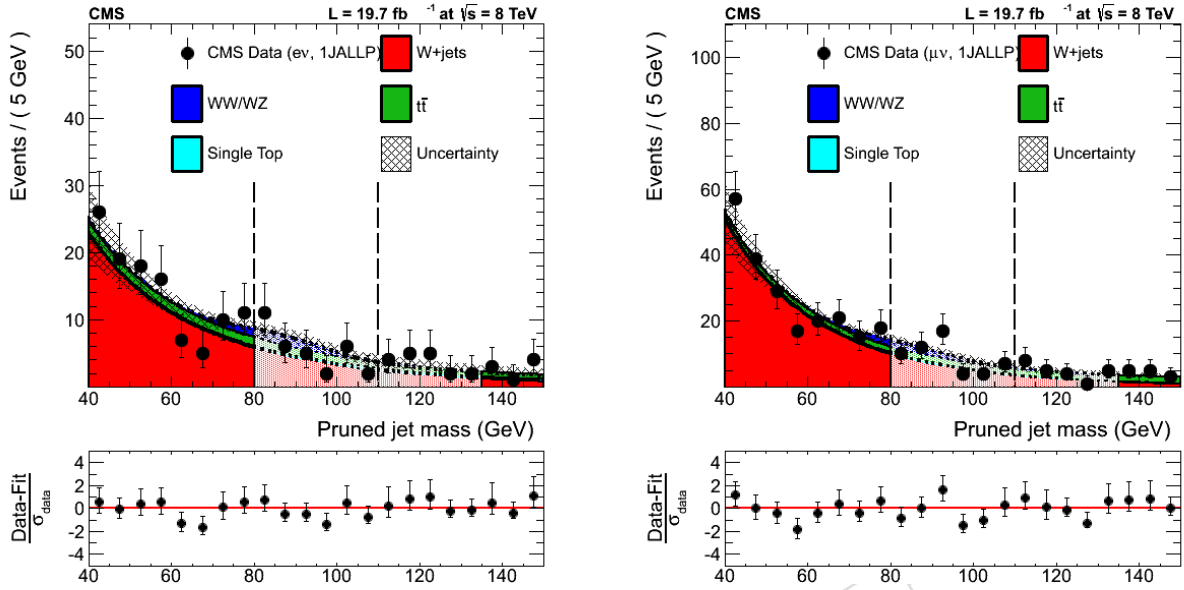


Figure 55: Closure test - Pruned jet mass distributions for the electron (left), and muon (right) categories for events with  $40 < m_{jet}^{pruned} < 80$  GeV passing the full selection plus the  $m_{WH} > 800$  GeV requirement. The different background components, results of the fit to the low-mass and high-mass sideband regions in data, are shown. In the closure test, the “signal region” is identified by the events in the range  $80 < m_{jet}^{pruned} < 110$  GeV.

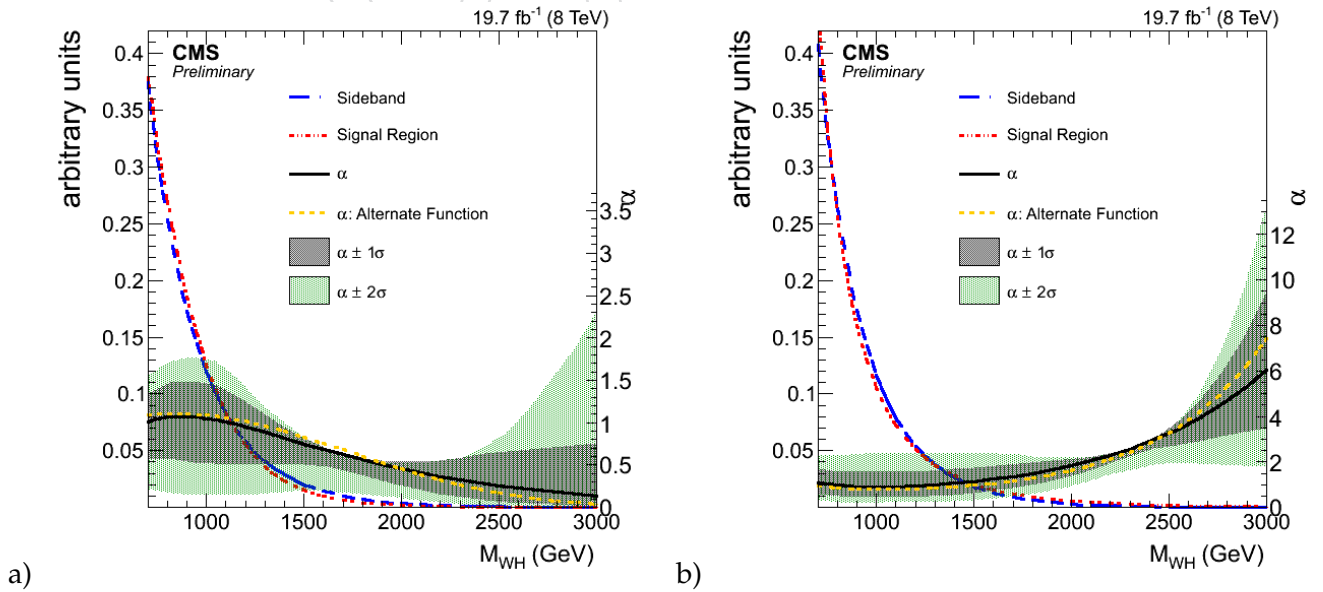


Figure 56: Closure test -  $\alpha^{MC}(m_{WH})$  functions vs  $m_{WH}$  for the electron (a) and muon (d) categories.

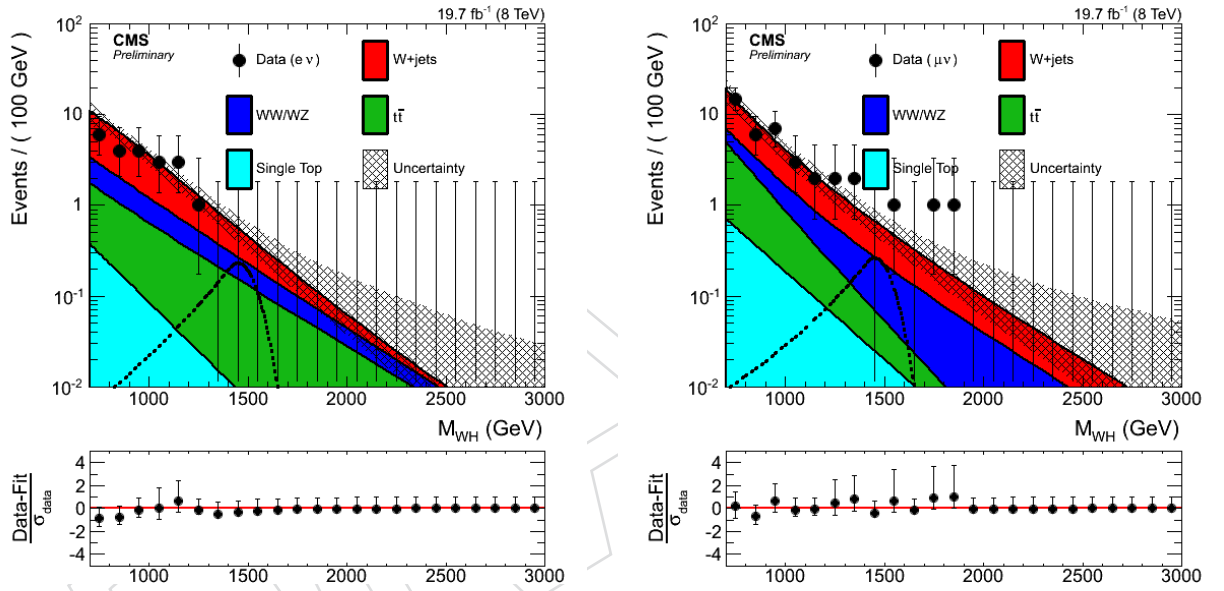


Figure 57: Closure test -  $m_{WH}$  distributions for data (black points) in the signal region after full selection for the two event categories used in this analysis: electron (left), muon (right).



## D Studies on bad event filters

A scan of the event displays of all the events with  $M_{WH} > 1.6$  TeV passing all the selection criteria, showed that two of them are due to anomalous behavior of the tracking software during jet reconstruction. The anomalous events are characterized by large amounts of fake tracks due to the failure of the tracking algorithm iteration 6 (TOB-TEC seeding) in the transition region between the barrel and end-cap regions of the detector ( $1 < |\eta| < 1.5$ ). Figure 58 shows the event display of one of the two events affected by this problem, while Figure 59 shows the same feature in simulation. In order to reject this type of noise we apply the `tobtecfakesfilter` (see Section 3.1) following the recommendation of the tracking POG. The efficiency of the `tobtecfakesfilter` on signal MC samples is about 97% for each of the mass points.

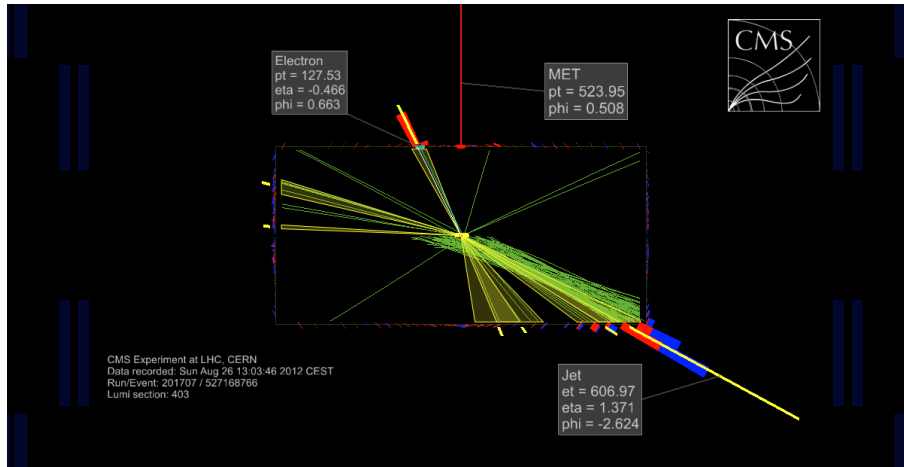


Figure 58: Event display of one typical anomalous event in the SingleElectron trigger dataset. A large number of fake tracks are reconstructed creating a bias in the jet reconstruction. Only tracks with  $p_T$  larger than 2 GeV are shown.

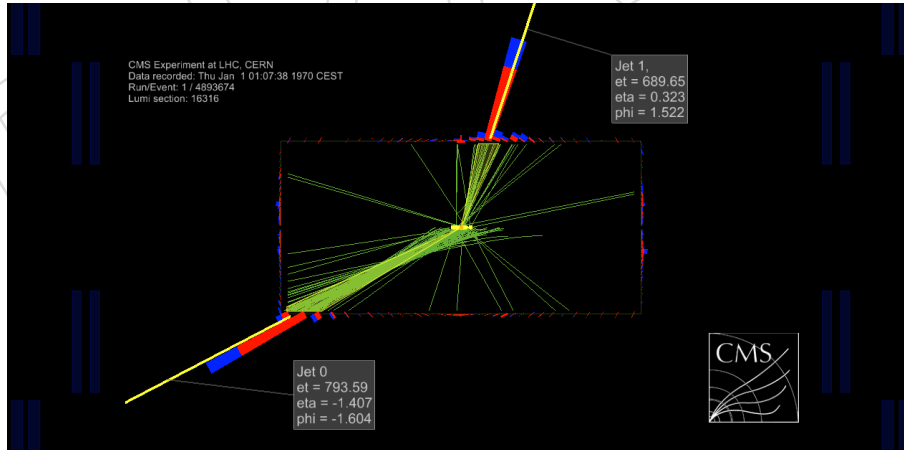


Figure 59: Event display of one anomalous event in simulation. The event is taken from a QCD MC sample. Only tracks with  $p_T$  larger than 2 GeV are shown.

The checks performed on the anomalous events showed that, after applying the `tobtecfakesfilter`, residual noise can still be identified in the problematic  $\eta$  region. Therefore we decide to apply an additional requirement on the selected jet  $\eta$ : as described in Section 3.4.1, jets are rejected if their pseudorapidity falls in the problematic region  $1 < |\eta| < 1.8$ . The choice for this fiducial cut is motivated by the different rate with which anomalous events appear in data and

MC as described in what follows.

We study the efficiency of the `tobtecfakesfilter` as a function of the H-bb jet  $p_T$  and  $\eta$  in a dijet sample with high statistics. The datasets used for these studies are summarized in Table 15. Events are selected if one of the following trigger has fired: HLT\_PFJet320, HLT\_PFHT650, HLT\_PFNPUHT650. The performance of the filters in simulation has been studied using the QCD MC samples listed in Table 16. We require at least two jets, with  $p_T > 400$  GeV for the leading jet and  $p_T > 80$  for the sub-leading jet. We require at least one jet to be b-tagged using the same combined b-tagging algorithm as used for the signal region. The jet that fails the b-tagging is required to have low pruned mass ( $< 40$  GeV).

Table 15: Summary of 8 TeV collision data used to study the efficiencies of the filters in data. The certification file used for these data is `Cert_190456-208686_8TeV_22Jan2013ReReco_Collisions12_JSON.txt`.

Dataset
/Jet/Run2012A-22Jan2013-v1/AOD
/JetHT/Run2012B-22Jan2013-v1/AOD
/JetHT/Run2012C-22Jan2013-v1/AOD
/JetHT/Run2012D-22Jan2013-v1/AOD

Table 16: Summary of the samples used to study the efficiencies of the two filters in simulation.

Sample
/QCD_HT-250To500_TuneZ2star_8TeV-madgraph-pythia6/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM
/QCD_HT-500To1000_TuneZ2star_8TeV-madgraph-pythia6/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM
/QCD_HT-1000ToInf_TuneZ2star_8TeV-madgraph-pythia6/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM

Figure 60 shows the effect of the filter on the jet  $\eta$  distribution comparing data, QCD background and signal MC: the signal shape is rather unaffected while data and MC background shapes become smoother after the application of the filter. Figure 61 shows the filter efficiency on data and simulated signal and QCD background as a function of the H-bb jet  $p_T$  for different H-bb jet  $\eta$  regions. A little dependence of the filter efficiency with the jet  $p_T$  is observed in the central  $\eta$  regions 0-1 and 1-1.5, while in the forward  $\eta$  region 1.5-2.4 we observe a drop of the efficiency with the jet  $p_T$ . The performance of the filter in the different  $\eta$  regions is summarized in Figure 62. A large discrepancy between data and MC is found in the  $\eta$  region 1-1.8, probably due to a poor material description in this region in MC.

We also perform the same efficiency studies removing the b-tagging requirement. The filter efficiency as a function of the leading jet  $\eta$  for this case is shown in Fig. 63, for both data and MC. This shows that the b-tagging requirement enriches the samples with TOBTEC anomalous events, hence the analysis is particularly sensitive to it.

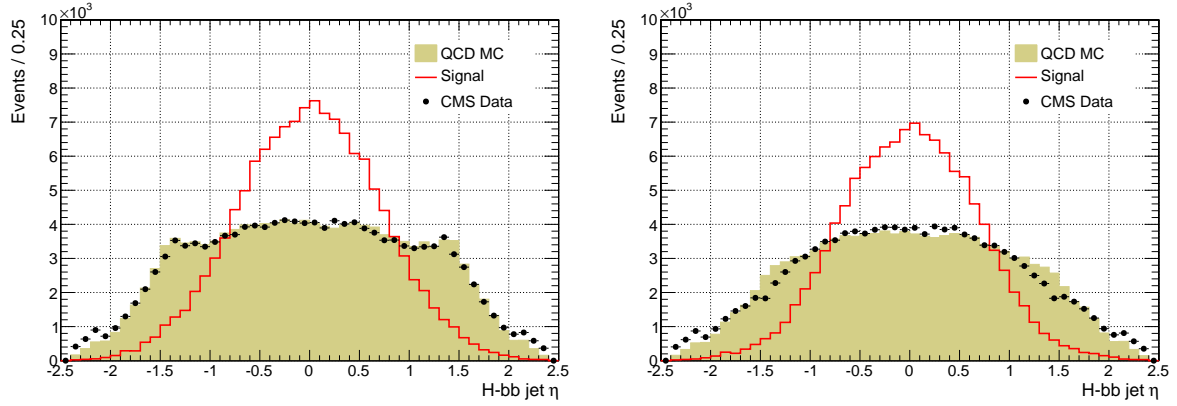


Figure 60: Comparison of the H-bb jet  $\eta$  distributions for data, QCD and signal MC before (left) and after (right) applying the `tobtecfakesfilter`. Data events are selected from the Jet trigger datasets. Signal jets are mostly central in the detector.

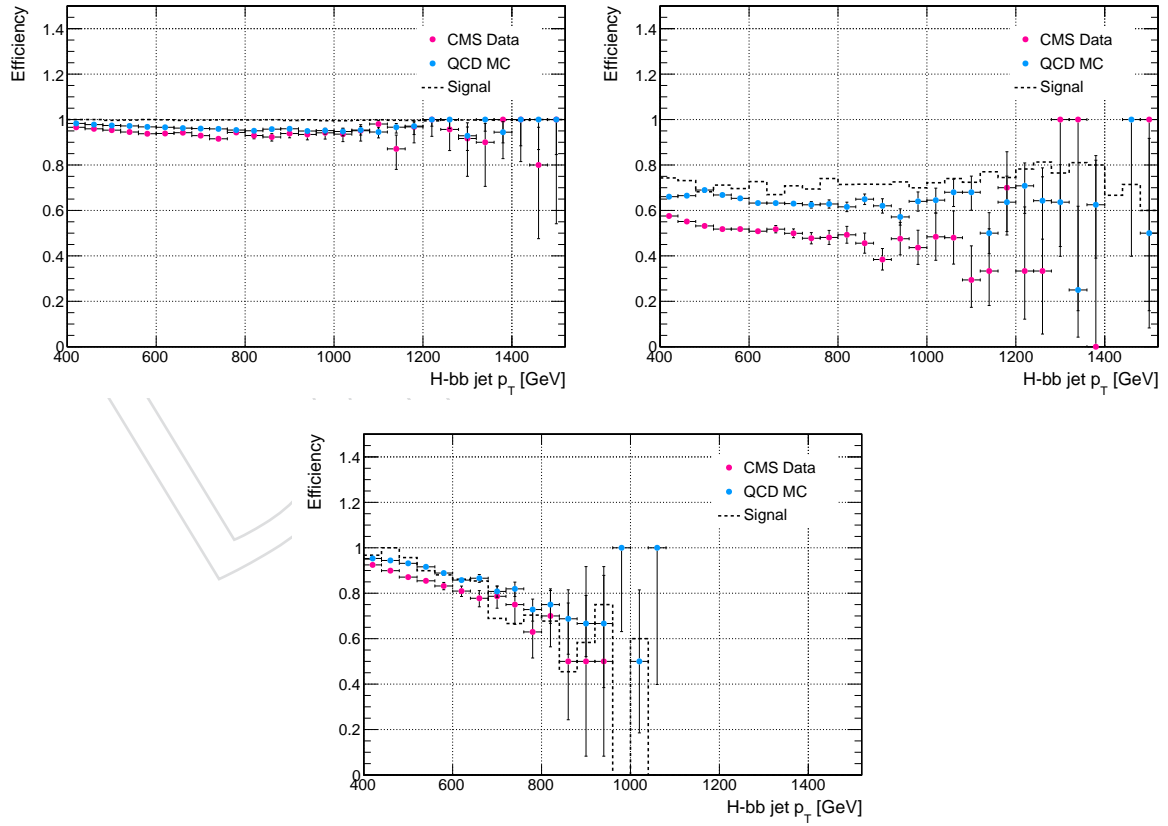


Figure 61: Comparison of the `tobtecfakesfilter` efficiency as a function of the H-bb jet  $p_T$  for data (magenta), QCD MC (azure) and signal MC (dashed) in the jet  $\eta$  region 0-1 (top left), 1-1.5 (top right) and 1.5-2.4 (bottom).

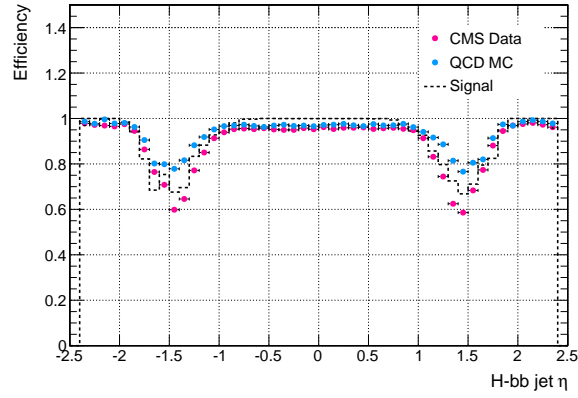


Figure 62: Efficiency of the `tobtecfakesfilter` as a function of the H-bb jet  $\eta$  for data (magenta), QCD MC (azure) and signal MC (dashed).

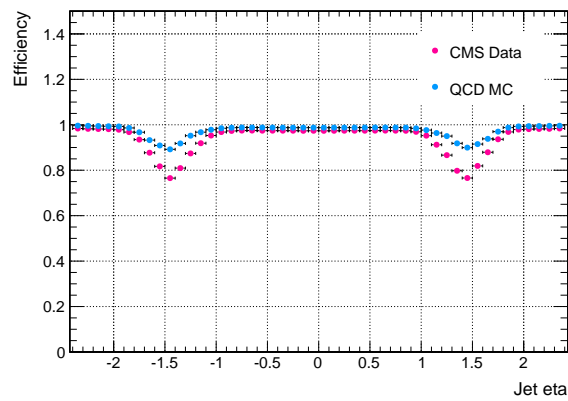


Figure 63: Efficiency of the `tobtecfakesfilter` as a function of the leading jet  $\eta$  for data (magenta), QCD MC (azure) and signal MC (dashed). B-tagging is not applied to the leading jet.

## E Event displays

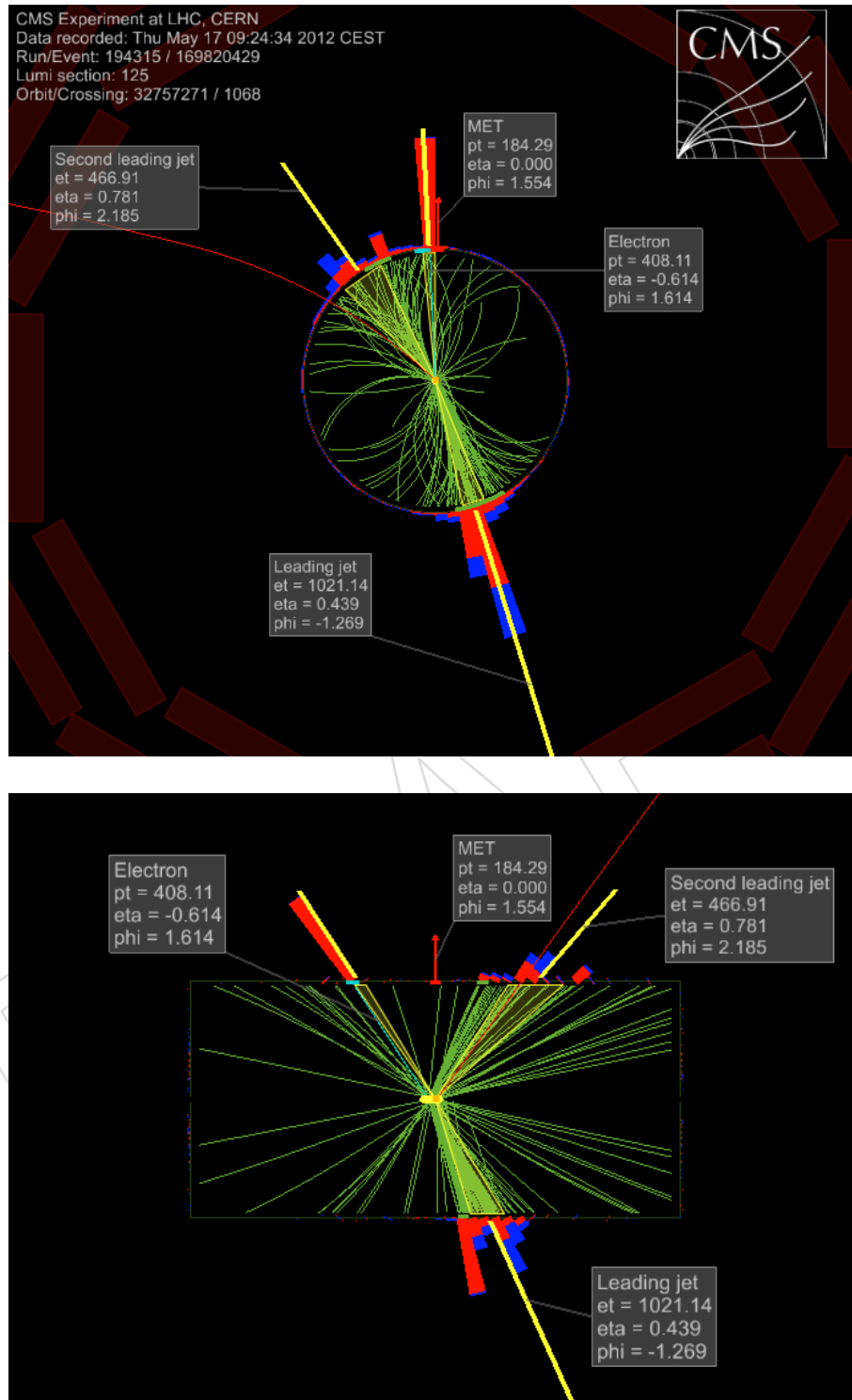


Figure 64: Event display of leptonic W plus Higgs bb-jet event. The leptonic W has a transverse momentum of 0.61 TeV, while the Higgs bb-jet has a transverse momentum of 1.1 TeV. The invariant mass of the WH system is 1.81 TeV.

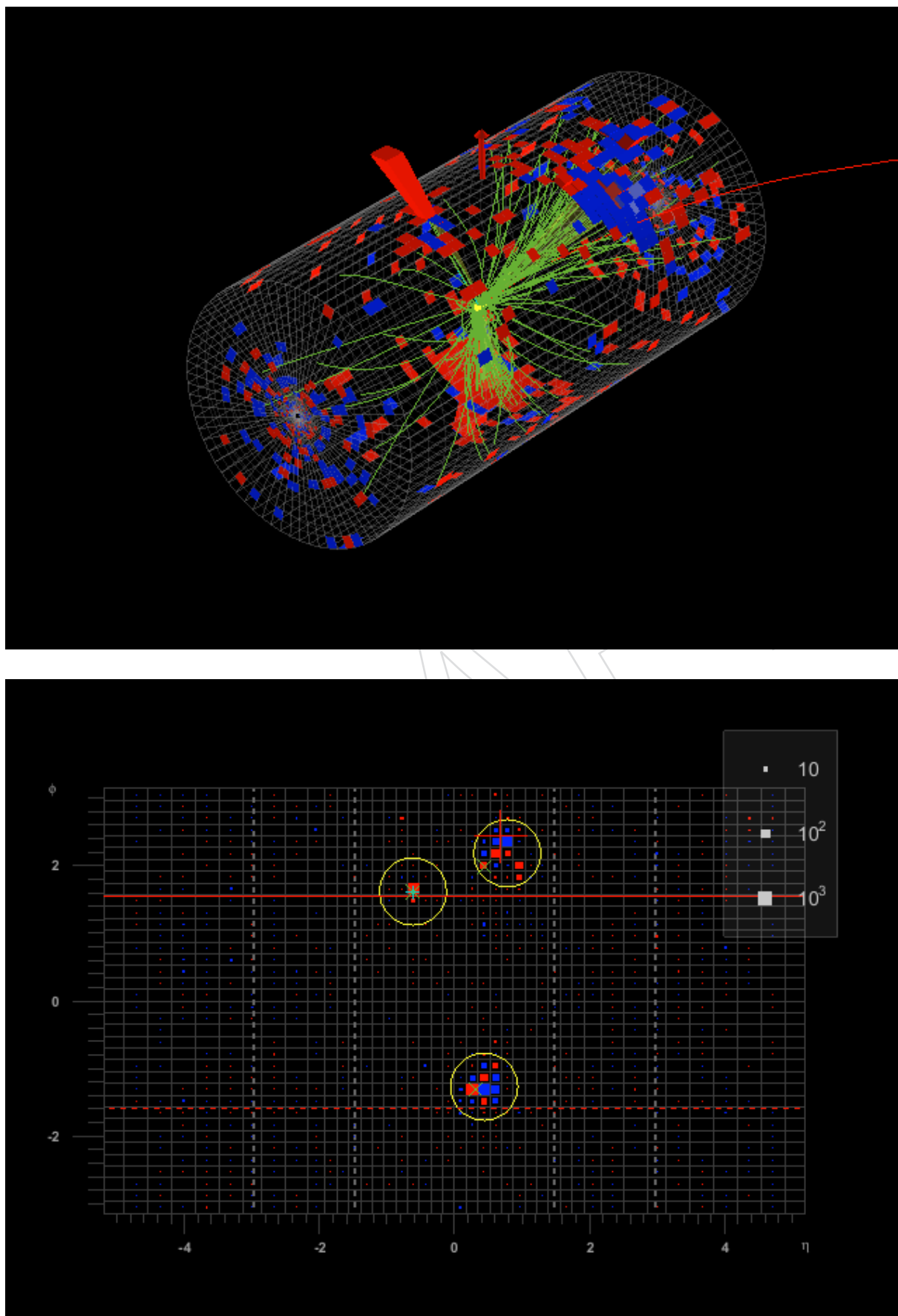


Figure 65: Event display of leptonic W plus Higgs bb-jet event. The leptonic W has a transverse momentum of 0.61 TeV, while the Higgs bb-jet has a transverse momentum of 1.1 TeV. The invariant mass of the WH system is 1.81 TeV.

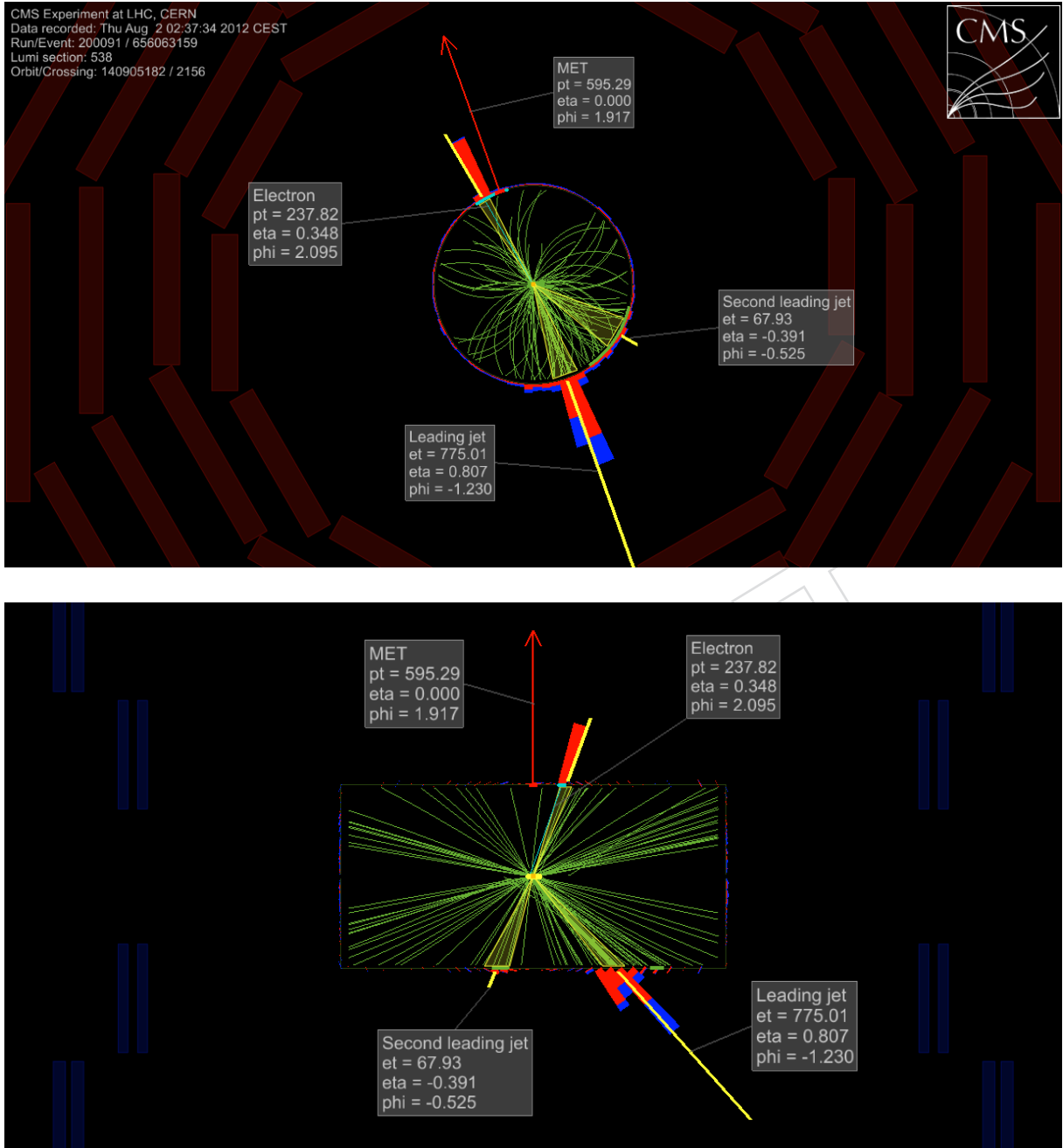


Figure 66: Event display of leptonic W plus Higgs bb-jet event. The leptonic W has a transverse momentum of 0.91 TeV, while the Higgs bb-jet has a transverse momentum of 0.87 TeV. The invariant mass of the WH system is 1.87 TeV.

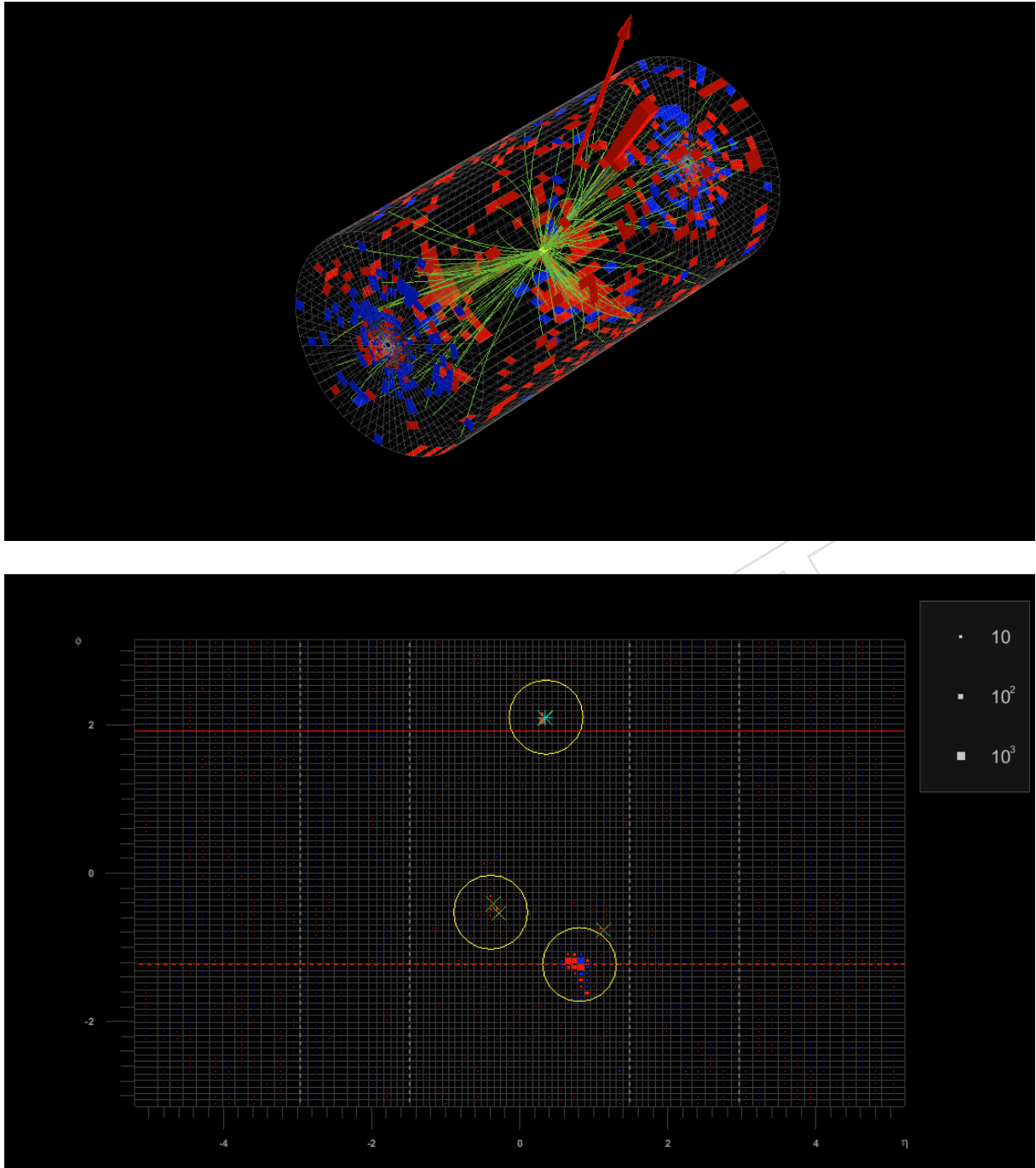


Figure 67: Event display of leptonic W plus Higgs bb-jet event. The leptonic W has a transverse momentum of 0.91 TeV, while the Higgs bb-jet has a transverse momentum of 0.87 TeV. The invariant mass of the WH system is 1.87 TeV.



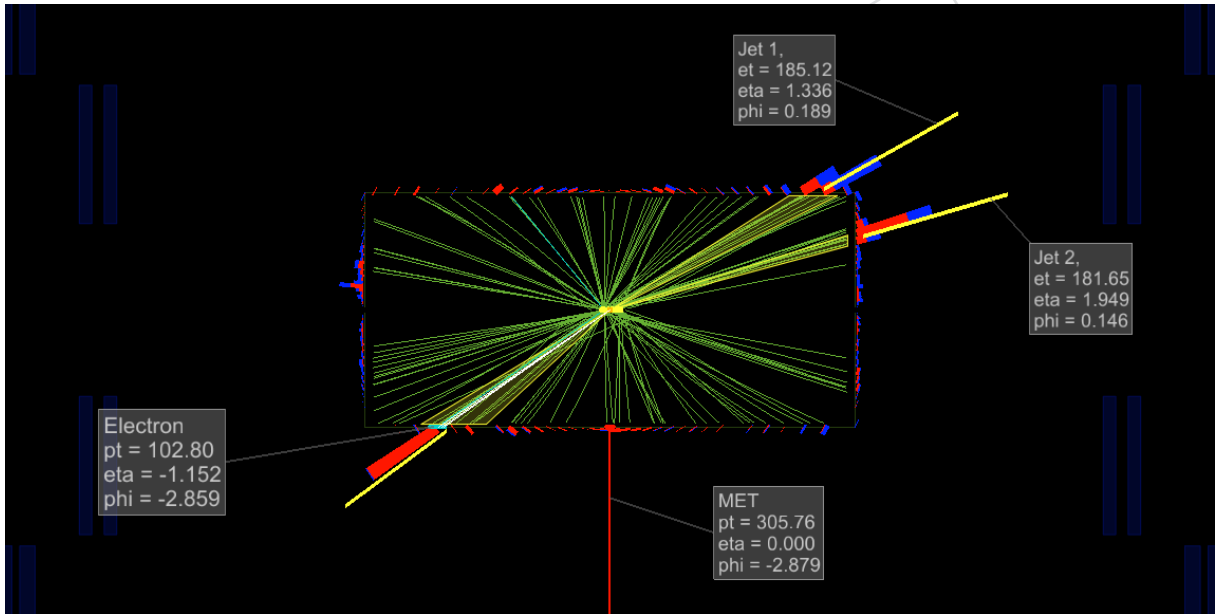
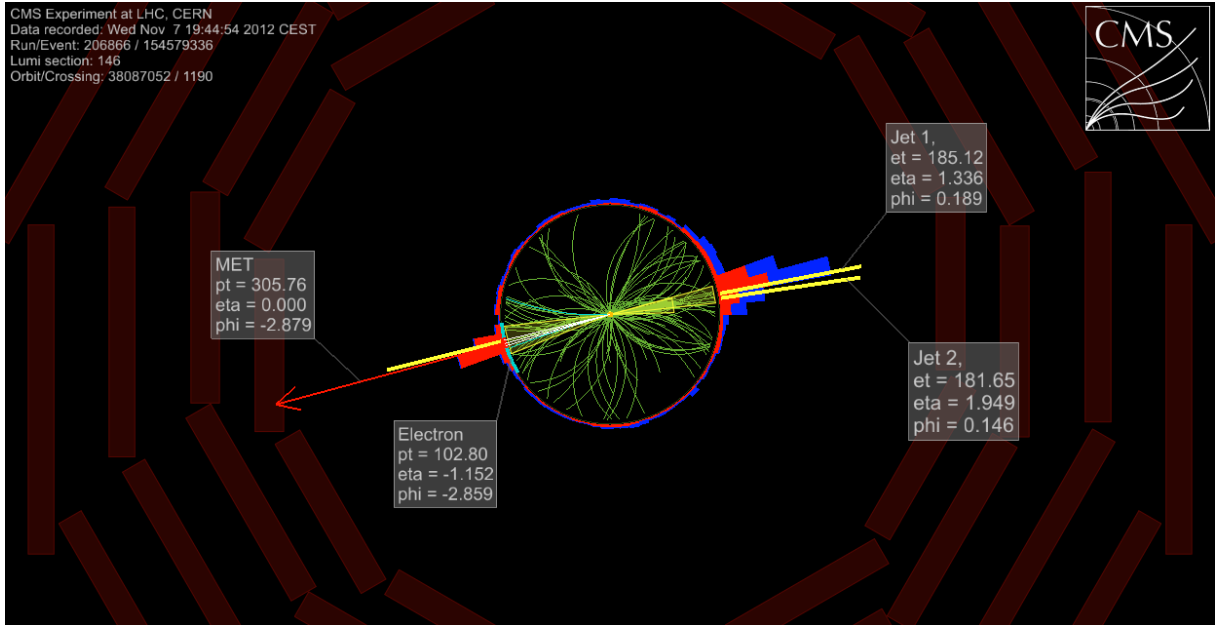


Figure 68: Event display of leptonic W plus Higgs bb-jet event. The leptonic W has a transverse momentum of 0.47 TeV, while the Higgs bb-jet has a transverse momentum of 0.38 TeV. The invariant mass of the WH system is 1.59 TeV.

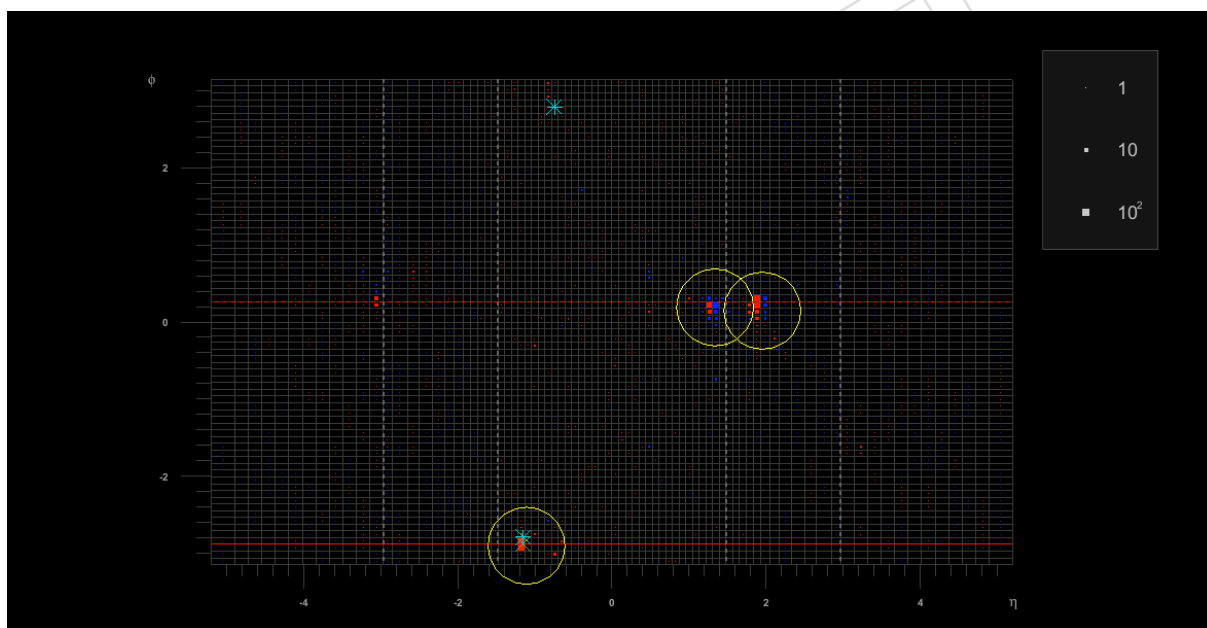
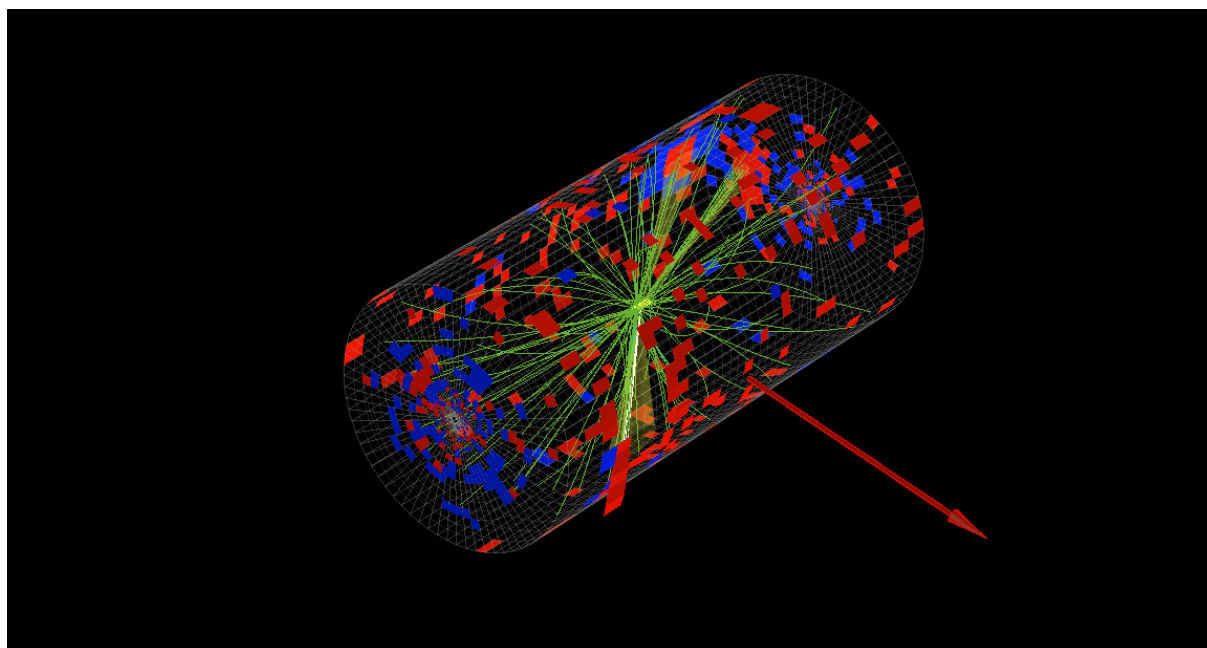


Figure 69: Event display of leptonic W plus Higgs bb-jet event. The leptonic W has a transverse momentum of 0.47 TeV, while the Higgs bb-jet has a transverse momentum of 0.38 TeV. The invariant mass of the WH system is 1.59 TeV.

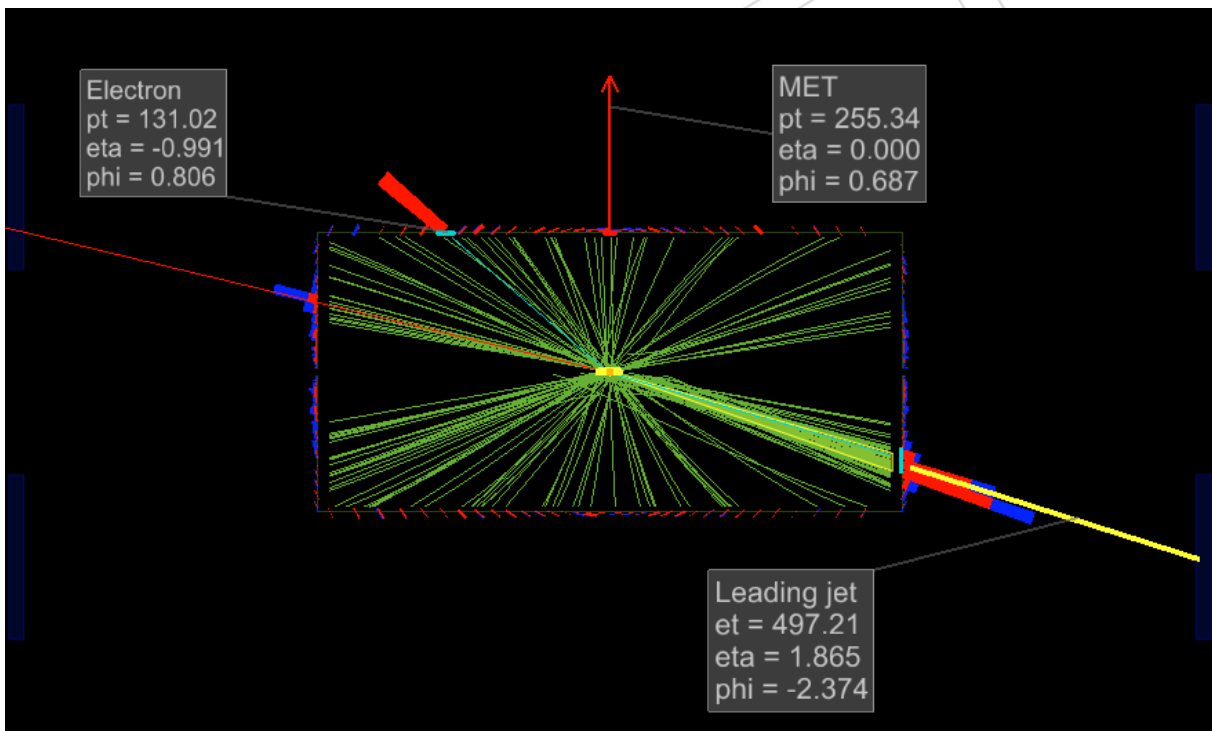
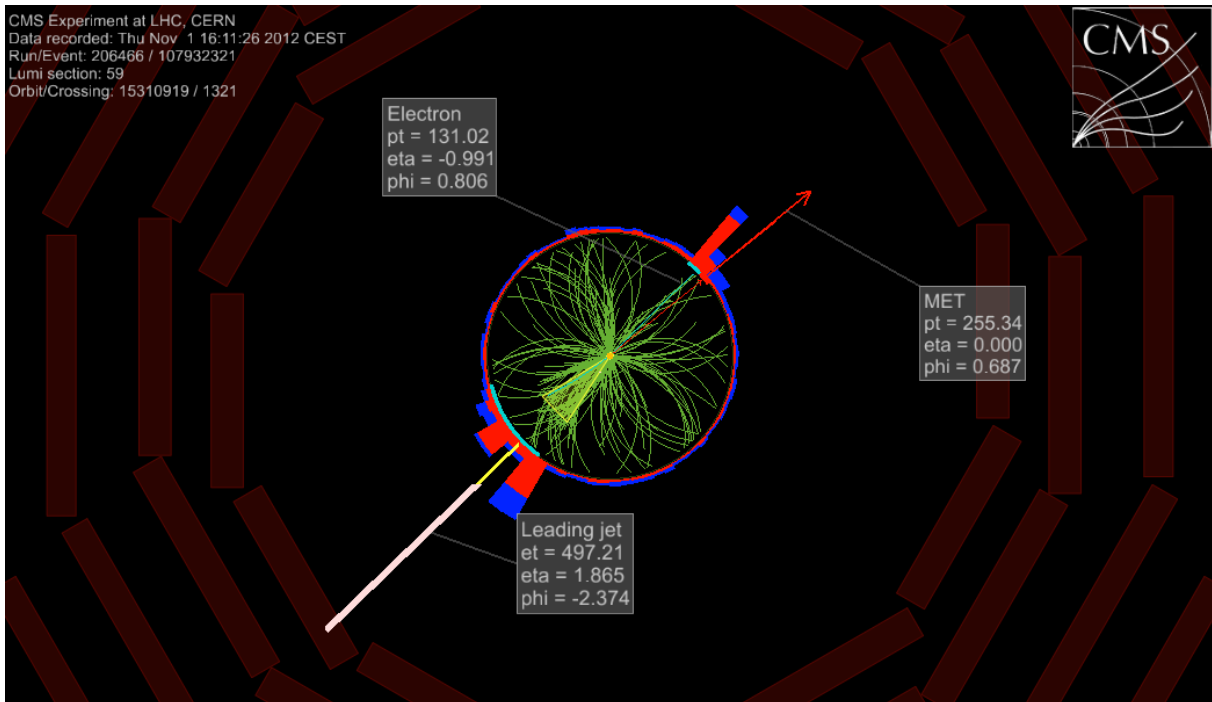


Figure 70: Event display of leptonic W plus Higgs bb-jet event. The leptonic W has a transverse momentum of 0.39 TeV, while the Higgs bb-jet has a transverse momentum of 0.52 TeV. The invariant mass of the WH system is 1.79 TeV.

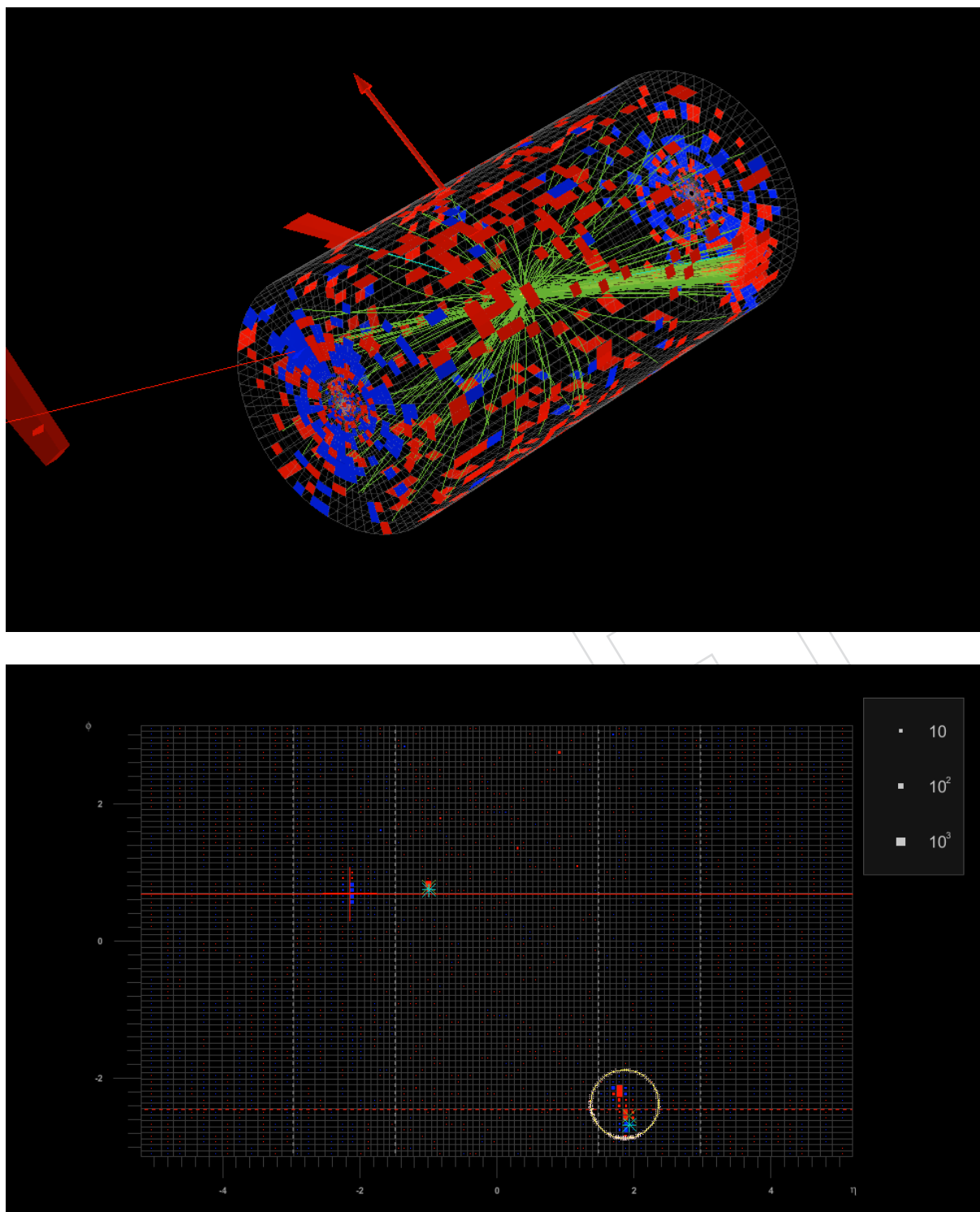


Figure 71: Event display of leptonic W plus Higgs bb-jet event. The leptonic W has a transverse momentum of 0.39 TeV, while the Higgs bb-jet has a transverse momentum of 0.52 TeV. The invariant mass of the WH system is 1.79 TeV.

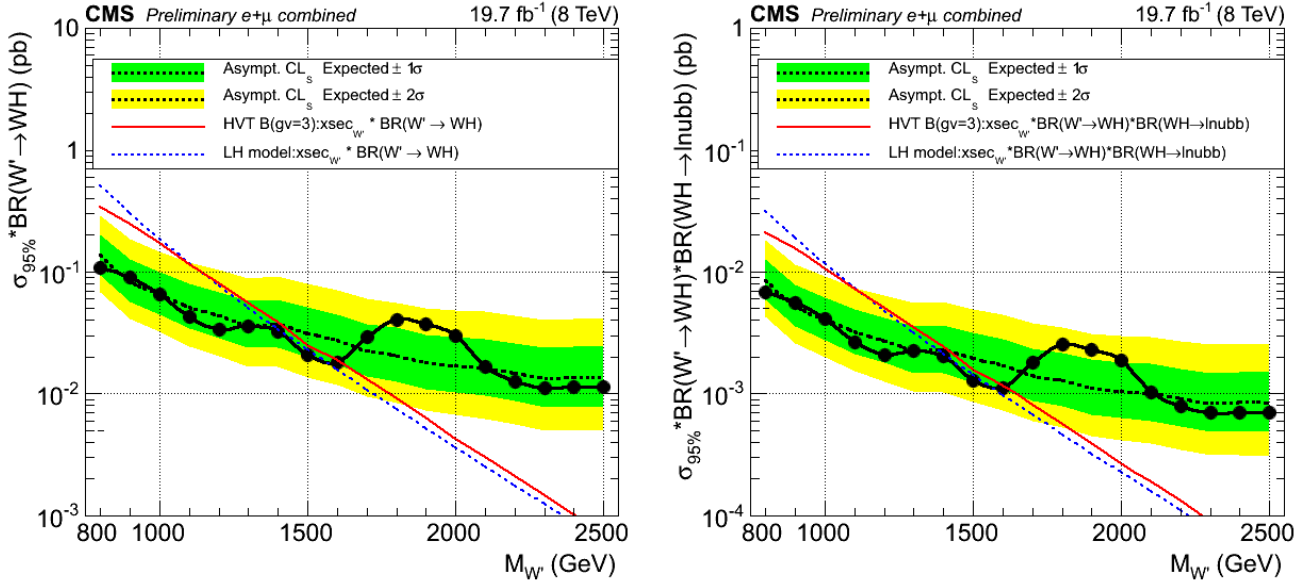


Figure 72: Method 1 - All analysis categories combined - Observed (dashed) and expected (solid) 95% CL upper limit (Asymptotic) on the product of the ' production cross section times the branching fraction of  $X \rightarrow WH$  (top), or times the branching fraction of  $X \rightarrow WH \rightarrow \lnubb$  (bottom) using  $19.7 \text{ fb}^{-1}$  of data. The 68% and 95% ranges of expectation are also shown with green and yellow bands. The theoretical cross section times the corresponding BR for  $X \rightarrow WH$  or  $X \rightarrow WH \rightarrow \lnubb$  are shown as a red solid (blue dashed) curve for HVT (LH) model.

## F Asymptotic Limits

In this part, we show the Asymptotic  $CL_S$  [55] limit results.

Figure 72 shows the 95% CL upper limits on the cross section of the process  $pp \rightarrow X \rightarrow WH$  and on the cross section of the process  $pp \rightarrow X \rightarrow WH \rightarrow \lnubb$  combining the two event categories as a function of the resonance mass, based on Asymptotic algorithm. For comparison, the production cross sections times branching ratio of  $X \rightarrow WH$  (and times branching ratio of  $X \rightarrow WH \rightarrow \lnubb$ ) are presented for the benchmark HVT and LH models, [14] and [12, 13].

Figure 73 shows the Asymptotic limits separately for the 2 categories considered in this analysis based on background estimation Method 1 (Section 7.2).

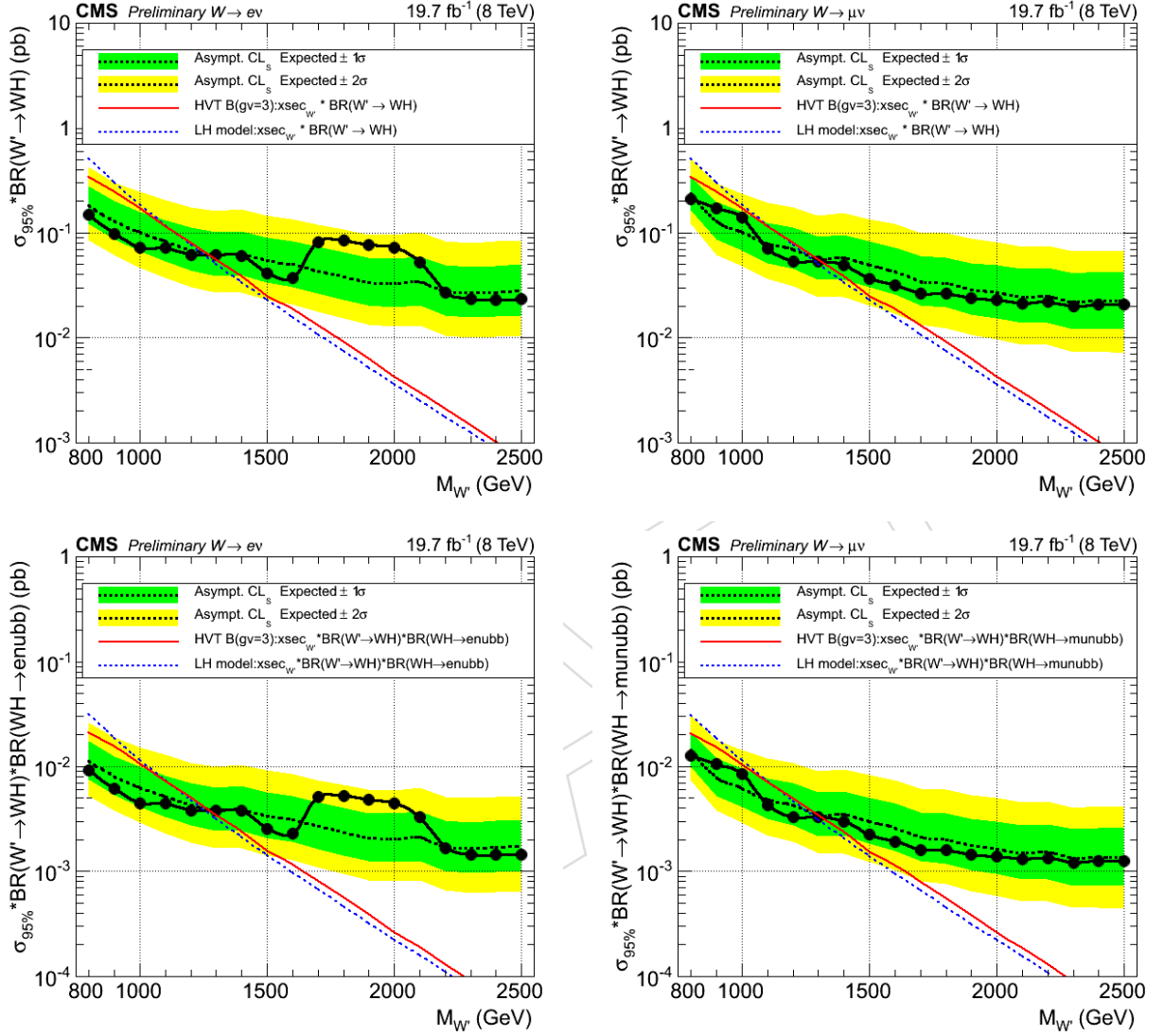


Figure 73: Method 1 - Ele (left), Mu (right) - Observed (dashed) and expected (solid) 95% CL upper limit (Asymptotic) on the product of the ' production cross section and the branching fraction of  $X \rightarrow WH$  (top) or the branching fraction of  $X \rightarrow WH \rightarrow l\nu bb$  (bottom) using 19.7 fb<sup>-1</sup> of data. The 68% and 95% ranges of expectation are also shown with green and yellow bands. The theoretical cross section times BR for  $X \rightarrow WH$  is shown as a red solid (blue dashed) curve for HVT (LH) model.

## References

- [1] ALEPH, CDF, D0, DELPHI, L3, OPAL, SLD Collaborations, the LEP Electroweak Working Group, the Tevatron Electroweak Working Group, and the SLD Electroweak and Heavy Flavour Groups, "Precision electroweak measurements and constraints on the Standard Model", (2010). [arXiv:1012.2367](#).
- [2] ATLAS Collaboration Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", *Physics Letters B* **716** (2012), no. 1, 1, [doi:10.1016/j.physletb.2012.08.020](#).
- [3] CMS Collaboration Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Physics Letters B* **716** (2012), no. 1, 30, [doi:10.1016/j.physletb.2012.08.021](#).
- [4] B. Bellazzini, C. Csaki, and J. Serra, "Composite Higgses", *EPJC* **74** (2014) 2766, [doi:10.1140/epjc/s10052-014-2766-x](#), [arXiv:1401.2457](#).
- [5] R. Contino, D. Marzocca, D. Pappadopulo, and R. Rattazzi, "On the effect of resonances in composite Higgs phenomenology", *JHEP* **1110** (2011) 081, [doi:10.1007/JHEP10\(2011\)081](#), [arXiv:1109.1570](#).
- [6] D. Marzocca, M. Serone, and J. Shu, "General Composite Higgs Models", *JHEP* **1208** (2012) 013, [doi:10.1007/JHEP08\(2012\)013](#), [arXiv:1205.0770](#).
- [7] CMS Collaboration, "Search for new physics in final states with a lepton and missing transverse energy in pp collisions at the LHC", *Phys. Rev. D* **87** (2013) 072005, [doi:10.1103/PhysRevD.87.072005](#).
- [8] ATLAS Collaboration, "Search for a heavy gauge boson decaying to a charged lepton and a neutrino in 1 fb<sup>-1</sup> of pp collisions at  $\sqrt{s} = 7$  TeV using the ATLAS detector", *Phys. Lett. B* **705** (2011) 28, [doi:10.1016/j.physletb.2011.09.093](#), [arXiv:1108.1316](#).
- [9] CMS Collaboration, "Search for a W' or  $\rho_{TC}$  decaying to WZ in pp collisions at  $\sqrt{s} = 7$  TeV", *Phys. Rev. Lett.* **109** (2012) 141801, [doi:10.1103/PhysRevLett.109.141801](#).
- [10] ATLAS Collaboration, "Search for resonant diboson production in the WW/WZ  $\rightarrow l\nu jj$  decay channels with the ATLAS detector at  $\sqrt{s} = 7$  TeV", *Phys. Rev. D* **87** (2013) 112006, [doi:10.1103/PhysRevD.87.112006](#), [arXiv:1305.0125](#).
- [11] ATLAS Collaboration, "Search for resonant WZ production in the WZ to l nu l' l' channel in  $\sqrt{s} = 7$  TeV pp collisions with the ATLAS detector", *Phys. Rev. D* **85** (2012) 112012, [doi:10.1103/PhysRevD.85.112012](#), [arXiv:1204.1648](#).
- [12] ATLAS Collaboration, "Search for the decays  $Z_H \rightarrow Zh$  and  $W_H \rightarrow Wh$  in the Little Higgs model assuming  $m(h) = 120$  GeV", Technical Report ATL-PHYS-2004-001, CERN, Geneva, 2004.
- [13] T. Han, H. E. Logan, B. McElrath, and L.-T. Wang, "Phenomenology of the little Higgs model", *Phys. Rev. D* **67** (2003) 095004, [doi:10.1103/PhysRevD.67.095004](#), [arXiv:0301040](#).
- [14] D. Pappadopulo, A. Thamm, R. Torre, and A. Wulzer, "Heavy Vector Triplets: Bridging Theory and Data", (2014). [arXiv:1402.4431](#).



- [15] CMS Collaboration, "Search for new resonances decaying to  $WW$  to  $l\nu q\bar{q}$  in the final state with a lepton, missing transverse energy, and single reconstructed jet", Technical Report CMS-PAS-EXO-12-021, CERN, Geneva, 2013.
- [16] CMS Collaboration, "Search for a narrow spin-2 resonance decaying to  $Z$  bosons in the semileptonic final state", Technical Report CMS-PAS-EXO-12-022, CERN, Geneva, 2013.
- [17] CMS Collaboration, "Search for heavy resonances in the  $W/Z$ -tagged dijet mass spectrum in  $pp$  collisions at 8 TeV", Technical Report CMS-PAS-EXO-12-024, CERN, Geneva, 2013.
- [18] CMS Collaboration, "Performance of  $b$  tagging at  $\sqrt{s} = 8$  TeV in multijet,  $t\bar{t}$  and boosted topology events", Technical Report CMS PAS BTV-13-001, CERN, Geneva, 2013.
- [19] CMS Collaboration, "Search for heavy resonances decaying to two Higgs bosons in hadronic final states: using  $b$ -tagging and counting experiment", Technical Report CMS AN 2014/124, CERN, Geneva, 2014.
- [20] J. Alwall et al., "MadGraph 5 : Going Beyond", *JHEP* **1106** (2011) 128, doi:10.1007/JHEP06(2011)128, arXiv:1106.0522.
- [21] "Hidden Abelian Higgs Model".  
<http://feynrules.irmp.ucl.ac.be/wiki/HiddenAbelianHiggsModel>.
- [22] GEANT4 Collaboration, "GEANT4: A Simulation toolkit", *Nucl.Instrum.Meth.* **A506** (2003) 250–303, doi:10.1016/S0168-9002(03)01368-8.
- [23] CMS Collaboration, "Search for a BSM resonance decaying to  $W$  vector bosons in the semileptonic final state", CMS Analysis Note CMS-AN-2013/045, CERN, 2013.
- [24] CMS Collaboration, "Search for a BSM resonance decaying to  $Z$  vector bosons in the semileptonic final state", CMS Analysis Note CMS-AN-2013/040, CERN, 2013.
- [25] CMS Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and  $E_T^{\text{miss}}$ ", *CMS Physics Analysis Summary* **CMS-PAS-PFT-09-001** (2009).
- [26] CMS Collaboration, "Commissioning of the Particle-Flow Reconstruction in Minimum-Bias and Jet Events from  $pp$  Collisions at 7 TeV", CMS Physics Analysis Summary CMS-PAS-PFT-10-002, 2010.
- [27] CMS Collaboration, "Search for a BSM resonance decaying to  $W$  vector bosons in the semileptonic final state", CMS Analysis Note CMS AN-13/045, 2013.
- [28] CMS Collaboration, "Search for new physics in the single lepton + MET final states with full 2012 dataset at  $\sqrt{s}=8$  TeV", CMS Analysis Note CMS-AN-2012/423, 2013.
- [29] CMS Collaboration, "Jet Energy Corrections determination at 7 TeV", CMS Physics Analysis Summary CMS-PAS-JME-10-010, 2010.
- [30] CMS Collaboration, "Search for anomalous  $t\bar{t}$  production in the highly-boosted all-hadronic final state", *JHEP* **09** (2012) 029, doi:10.1007/JHEP09(2012)029, arXiv:1204.2488. CMS-PAS-EXO-11-006.
- [31] CMS Collaboration, "Searches for new physics in the  $WW \rightarrow l\nu j$  final state with hadronic  $W$  bosons merged into a single jet", CMS Analysis Note CMS-AN-2012/381, 2012.



- [32] J. Bauer, "Prospects for the Observation of Electroweak Top-Quark Production with the CMS Experiment". PhD thesis, Karlsruher Institut für Technologie (KIT), May, 2010.
- [33] M. Gouzevitch et al., "Scale-invariant resonance tagging in multijet events and new physics in Higgs pair production", [arXiv:1303.6636](#).
- [34] ATLAS Collaboration, "Performance of large-R jets and jet substructure reconstruction with the ATLAS detector", Technical Report ATLAS-CONF-2012-065, CERN, Jul, 2012.
- [35] CMS Collaboration, "Studies of jet mass in dijet and  $W/Z$ +jet events", *JHEP* **05** (2013) 090, doi:10.1007/JHEP05(2013)090, [arXiv:1303.4811](#). CMS-PAS-SMP-12-019.
- [36] CMS Collaboration, "Search for heavy resonances in the  $W/Z$ -tagged dijet mass spectrum in pp collisions at 7 TeV", *PLB* **723** (2013) 280, doi:10.1016/j.physletb.2013.05.040, [arXiv:1212.1910](#). CMS-PAS-EXO-11-095.
- [37] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, "Recombination Algorithms and Jet Substructure: Pruning as a Tool for Heavy Particle Searches", *Phys.Rev.* **D81** (2010) 094023, doi:10.1103/PhysRevD.81.094023, [arXiv:0912.0033](#).
- [38] CMS Collaboration, "Identifying hadronically decaying vector bosons merged into a single jet", CMS Physics Analysis Summary CMS-PAS-JME-13-006, 2013.
- [39] G. Punzi, "Sensitivity of searches for new signals and its optimization", in *proceedings to PhyStat 2003 Conference, Stanford, Ca, USA, September 2003*, PSN MODT002. [arXiv:physics/0308063v2](#).
- [40] CMS Collaboration, "Search for heavy di-Higgs resonances decaying to 4 bottom quarks", CMS Analysis Note CMS AN-14/124, 2014.
- [41] "Monte Carlo particle numbering scheme". <http://pdg.lbl.gov/2002/montecarlopp.pdf>.
- [42] D. Pappadopulo, A. Thamm, R. Torre and A. Wulzer, "Tools for the study of heavy vector triplets".
- [43] G. Burdman, M. Perelstein, and A. Pierce, "Large Hadron Collider tests of a little Higgs model", *Phys.Rev.Lett.* **90** (2003) 241802, doi:10.1103/PhysRevLett.90.241802, [arXiv:hep-ph/0212228](#).
- [44] CMS Collaboration, ""Search for high mass exotic resonances decaying to  $WW$  in the semi-leptonic channel"", CMS Analysis Note CMS AN-13/139, 2013.
- [45] CMS Collaboration *PAS SMP-12-019* (2012).
- [46] "Measurement of the and  $\{ZZ\}$  production cross sections in pp collisions at", *Physics Letters B* **721** (2013), no. 45, 190 – 211, doi:<http://dx.doi.org/10.1016/j.physletb.2013.03.027>.
- [47] CMS Collaboration Collaboration, "CMS Luminosity Based on Pixel Cluster Counting - Summer 2013 Update", Technical Report CMS-PAS-LUM-13-001, CERN, Geneva, 2013.
- [48] CMS Collaboration, ""ECAL Detector Performance, 2011 Data"", CMS Performance Note CMS-DP-2012/007, 2012.

- [49] CMS Collaboration, “Search for High Mass Resonances Decaying to Electron Pairs at 8 TeV with the Full 2012 dataset”, CMS Analysis Note CMS-AN-2012/415, 2013.
- [50] CMS Collaboration, “8 TeV Jet Energy Scale performance plots based on 11 fb<sup>-1</sup> of 8 TeV data”, *CMS Detector Performance Plots* **DP-2013/011** (2013).
- [51] CMS Collaboration, “Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS”, *JINST* **6** (2011) P11002, doi:10.1088/1748-0221/6/11/P11002, arXiv:1107.4277.
- [52] A. L. Read, “Presentation of search results: The CL(s) technique”, *J. Phys.* **G28** (2002) 2693–2704, doi:10.1088/0954-3899/28/10/313.
- [53] A. L. Read, “Modified frequentist analysis of search results (the CLs method)”, technical report, 2000.
- [54] CMS Collaboration, “Combined results of searches for the standard model Higgs boson in pp collisions at  $\sqrt{s} = 7$  TeV”, *Phys. Lett. B* (2012), no. 710, arXiv:1202.1488.
- [55] G. Cowan, K. Cranmer, E. Gross, O. Vitells, “Asymptotic formulae for likelihood based tests of new physics”, *Eur. Phys. J.* **C71** (2011) 1554, arXiv:physics/1007.1727v2.
- [56] CMS Collaboration, “Search for massive resonances decaying into pairs of boosted bosons in semi-leptonic final states at  $\sqrt{s} = 8$  TeV”, *JHEP* **1408** (2014) 174, doi:10.1007/JHEP08(2014)174, arXiv:1405.3447.