

# CMS Draft Analysis Note

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

---

2016/02/25

Head Id: 322572

Archive Id: 328042:328048M

Archive Date: 2016/02/03

Archive Tag: trunk

## Search for heavy resonances decaying into a vector boson and a Higgs boson in the $(\ell\ell, \ell\nu, \nu\nu) b\bar{b}$ final state

Lisa Benato<sup>1</sup>, Yu-Hsiang Chang<sup>5</sup>, Ching-Wei Chen<sup>5</sup>, Michele de Gruttola<sup>4</sup>, Ji-Kong Huang<sup>5</sup>, Raman Khurana<sup>5</sup>, Stefano Lacaprara<sup>2</sup>, Yun-Ju Lu<sup>5</sup>, Jacopo Pazzini<sup>3</sup>, Maurizio Pierini<sup>4</sup>, Henry Yee-Shian Tong<sup>5</sup>, Jun-Yi Wu<sup>5</sup>, Shin-Shan Eiko Yu<sup>5</sup>, Marco Zanetti<sup>1</sup>, and Alberto Zucchetta<sup>3</sup>

<sup>1</sup> Università di Padova, Padova, Italy

<sup>2</sup> INFN sezione di Padova, Padova, Italy

<sup>3</sup> CERN, Meyrin, Switzerland and Università di Padova, Padova, Italy

<sup>4</sup> CERN, Meyrin, Switzerland

<sup>5</sup> National Central University, Chung-Li, Taiwan

### Abstract

This analysis note describes the search for heavy resonances decaying into a vector boson and a Higgs boson, where the vector boson can be either a Z or W boson decaying leptonically (electrons, muons or neutrinos) and the Higgs decaying hadronically into couple of b-quarks. Final states with either two, one or no leptons and a Higgs-tagged jet are probed. The search is performed in the boosted regime for resonances with mass larger than 1 TeV up to 4.5 TeV. Model-independent upper limits are derived as a function of the resonance mass and natural width, and are interpreted within the Heavy Vector Triplet theoretical model.

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

PDFAuthor:	A. Zucchetta, J. Pazzini, M. Zanetti, S.-S. Yu
PDFTitle:	Search for heavy resonances decaying into a vector boson and a Higgs boson in the “( $\ell\ell, \ell\nu, \nu\nu$ ) $b\bar{b}$ final state”
PDFSubject:	CMS
PDFKeywords:	CMS, physics, software, computing

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



## 1 Contents

1	1	Introduction . . . . .	3
2	1.1	Theoretical motivations . . . . .	3
3	2	Datasets and samples . . . . .	6
4	2.1	Signal . . . . .	6
5	2.2	Signal characterization . . . . .	6
6	2.3	Background samples . . . . .	10
7	2.4	V boson momentum corrections . . . . .	14
8	2.5	Data . . . . .	17
9	2.6	Trigger . . . . .	18
10	3	Physics objects . . . . .	21
11	3.1	Vertex and Pile-up . . . . .	21
12	3.2	Electrons . . . . .	21
13	3.3	Muons . . . . .	27
14	3.4	Mini-isolation studies . . . . .	31
15	3.5	Taus . . . . .	34
16	3.6	Photons . . . . .	34
17	3.7	Jets . . . . .	35
18	3.8	b-tagging . . . . .	38
19	3.9	Missing Energy . . . . .	41
20	4	Boson reconstruction . . . . .	43
21	4.1	Z boson to leptons . . . . .	43
22	4.2	Z boson to neutrinos . . . . .	43
23	4.3	W boson reconstruction . . . . .	43
24	4.4	Higgs to bb reconstruction . . . . .	44
25	5	Event selection . . . . .	54
26	5.1	Leptonic selection . . . . .	54
27	5.2	Hadronic selection . . . . .	55
28	5.3	Topology and event cleaning . . . . .	55
29	6	Data-simulation comparison . . . . .	58
30	6.1	Zero lepton channel . . . . .	58
31	6.2	Single lepton channel . . . . .	58
32	6.3	Dilepton channel . . . . .	67
33	7	Top control regions . . . . .	80
34	8	Alpha ratio background prediction . . . . .	83
35	8.1	Background normalization . . . . .	84
36	8.2	Background shape . . . . .	91
37	8.3	Background prediction . . . . .	103
38	8.4	Signal modeling . . . . .	106
39	8.5	Alpha method validation . . . . .	117
40	9	Systematic uncertainties . . . . .	119
41	9.1	Main background uncertainties . . . . .	119
42	9.2	Triggers and Leptons . . . . .	119

44	9.3	Jet momentum . . . . .	120
45	9.4	Higgs tagging efficiency . . . . .	120
46	9.5	b-tagging . . . . .	121
47	9.6	Missing Energy . . . . .	122
48	9.7	Pile-up . . . . .	122
49	9.8	QCD renormalization and factorization scale . . . . .	122
50	9.9	Summary . . . . .	123
51	10	Results . . . . .	126
52	10.1	Signal extraction strategy . . . . .	126
53	10.2	Results of the alpha method . . . . .	126
54	11	Conclusions . . . . .	132
55	A	Optimization of electron kinematic selection criteria . . . . .	137
56	B	Optimization of selection criteria based on event topology . . . . .	137

DRAFT

## 57 1 Introduction

58 This analysis searches for signal of heavy resonances decaying into a vector boson (denoted  
 59 as  $V=Z, W$ ) and a Higgs boson ( $h$ ). In turn, the vector boson is identified through its leptonic  
 60 decays ( $\ell = e, \mu$ ) or a pair of neutrinos ( $\nu\bar{\nu}$ ), which escape undetected. The Higgs boson  $h$  is  
 61 expected to hadronically decay primarily into a pair of b-quarks. The investigated final states  
 62 consists of two b-quarks and zero, just one or two charged leptons. In the case of the zero-  
 63 lepton channel, a large amount of missing energy is measured in the detector. In the leptonic  
 64 channels, leptons are identified in the detector and limit the presence of the background, while  
 65 the hadronic Higgs decay collects the largest possible fraction of Higgs events.

66 The search is performed by examining the distribution of the reconstructed ( $m_{\ell\ell bb}, m_{\ell\nu bb}$ ) or  
 67 transverse mass ( $m_{\nu\nu bb}^T$ ) for a localized excess. The signal strength and the background normal-  
 68 ization are determined from appropriate control regions for each channel studied.

### 69 1.1 Theoretical motivations

70 Although the Higgs boson discovery by the ATLAS and CMS [1–3] collaborations imposes  
 71 strong constraints on theories beyond the Standard Model (SM), the extreme fine tuning in  
 72 quantum corrections required to have a light fundamental Higgs boson with mass close to  
 73 125 GeV [4–7] suggests that the SM may be incomplete, and not valid beyond a scale of a few  
 74 TeV. Various dynamical electroweak symmetry breaking scenarios which attempt to solve this  
 75 naturalness problem, such as Minimal Walking Technicolor [], Little Higgs [8–10], or composite  
 76 Higgs models [11–13], predict the existence of new resonances decaying to a vector boson plus  
 77 a Higgs boson.

78 The result of the search is primarily interpreted in a model-independent way within a sim-  
 79 plified approach based on a phenomenological Lagrangian that incorporates Heavy Vector  
 80 Triplets (HVT) [14]. In these models, new heavy vector bosons ( $V^\pm, V^0$ ) that couple to the  
 81 Higgs and SM gauge bosons with the parameters  $g_V$  and  $c_H$  and to the fermions via the com-  
 82 bination  $(g_2/g_V)c_F$ . The parameter  $g_V$  represents the strength of the new vector boson interac-  
 83 tion, while  $c_H$  and  $c_F$  represent the couplings to the Higgs and the fermions respectively, and  
 84 are expected to be of order unity in most models.

The additional Lagrangian term resulting from the introduction of the heavy triplet has the form:

$$\begin{aligned}\mathcal{L}_V = & -\frac{1}{4}D_\mu V_\nu^a D^\mu V^{\nu a} + \frac{m_V^2}{2}V_\mu^a V^{\mu a} \\ & + ig_V c_H V_\mu^a H^\dagger \tau^a \bar{D}^\mu H + \frac{g^2}{g_V} c_F V_\mu^a \sum_f \bar{f}_L \gamma^\mu \tau^a f_L \\ & + \frac{g_V}{2} c_{VVV} \epsilon_{abc} V_\mu^a V_\nu^b D^\mu V^{\nu c} + \text{quadrilinear terms}\end{aligned}$$

85 Two benchmark models [14] are considered. In the first model, referred to as model A, the  
 86 branching fractions to fermions and gauge bosons are comparable, as in some extensions of the  
 87 SM gauge group []. In this weakly coupled model,  $g_V \sim g \sim 1, c_H = -g^2/g_V^2, c_F \sim 1$

88 For model B, fermionic couplings are suppressed, as for example in a composite Higgs model.  
 89 Vector bosons then are strongly coupled, leading to  $g_V \lesssim 4\pi, c_H \sim c_F \sim 1$ . This model B is  
 90 particularly interesting for the present search, since it predicts signal cross sections in the order  
 91 of fb for resonances up to  $2 \sim 3$  TeV, branching ratios to vector bosons close to the unity, and  
 92 thus being accessible at the LHC Run-II.

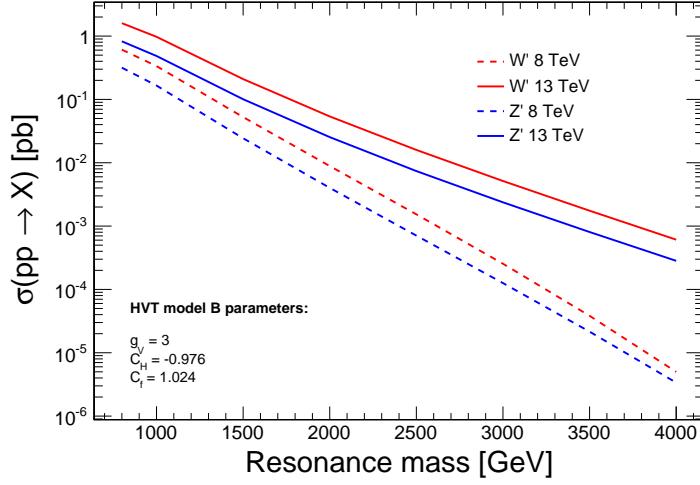


Figure 1: Resonance cross sections as a function of their mass for the HVT benchmark model B.

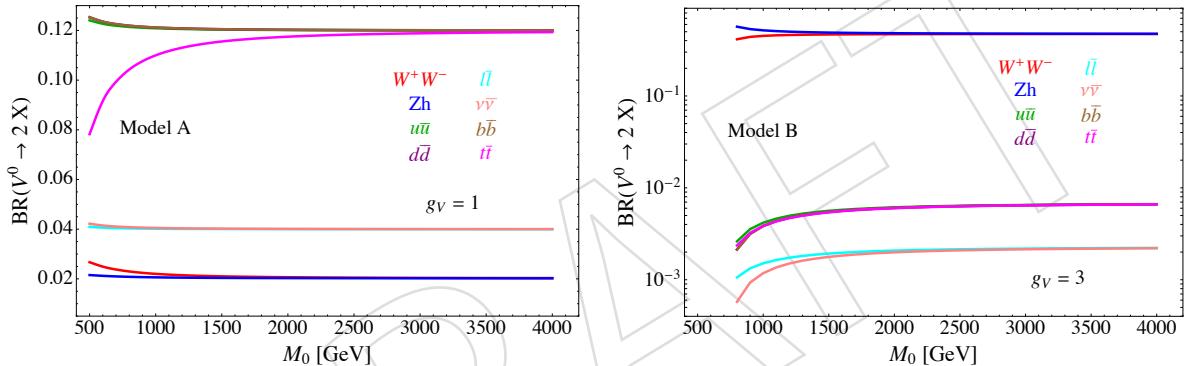


Figure 2: Branching ratios as a function of the resonance mass for the HVT benchmark model A (left) and model B (right).

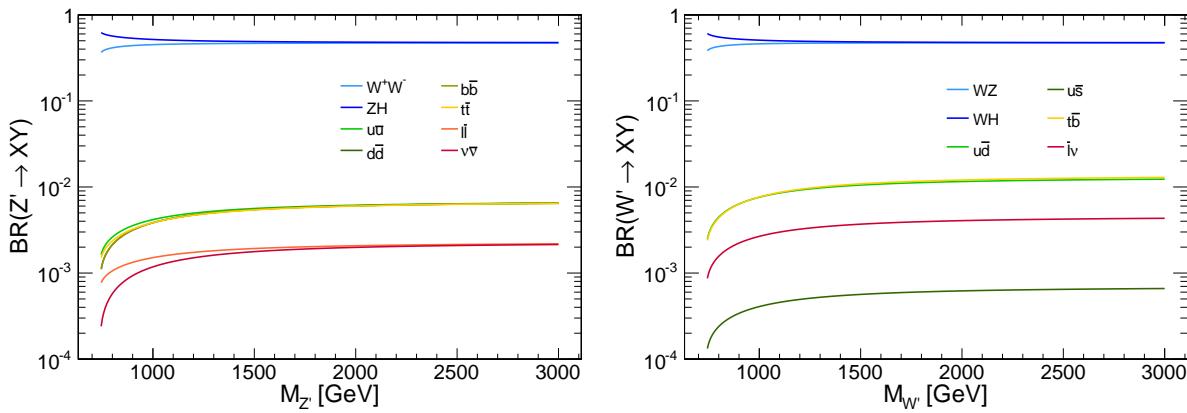


Figure 3: Branching ratios as a function of the resonance mass for a  $Z'$  (left) and  $W'$  (right) in the HVT model B.

- 93 As an alternative benchmark, the Minimal Walking Technicolor model (MWT) is considered, a  
94 model with strongly coupled dynamics. This model predicts two triplets of resonances,  $R_1^{\pm,0}$   
95 and  $R_2^{\pm,0}$ , one of which is a vector and the other an axial-vector, that couple to vector bosons  
96 with strength  $\tilde{g}$  and to fermions with  $g/\tilde{g}$ , where  $g$  is the weak SU(2) coupling constant. The  
97 bare axial-vector mass  $m_A$  determines the masses of  $R_1$  and  $R_2$ , with the lower mass resonance  
98  $R_1$  having a mass close to  $m_A$ . Lattice simulations in this model [] predict masses close to 2 TeV.  
99 The decay channel  $R_{1,2}^0 \rightarrow Z h$  is allowed and leads to the  $\nu\nu b\bar{b}$  and  $\ell\ell b\bar{b}$  final states.
- 100 The channels explored in this analysis have never been investigated at CMS at the TeV-scale  
101 regime. A low mass ( $m_X \leq 600$  GeV) search for  $m_{\ell\ell b\bar{b}}$  resonances has already been published  
102 by CMS with the full  $\sqrt{s} = 8$  TeV dataset [15].

A large, semi-transparent watermark consisting of the word "DRAFT" in a bold, sans-serif font. The letters are oriented diagonally, sloping upwards from the bottom-left towards the top-right. The watermark is rendered in a very light gray color, making it less prominent against the white background of the page.

## 103 2 Datasets and samples

### 104 2.1 Signal

105 Signal samples are generated with the MADGRAPH5 [16] LO generator, while hadronization  
 106 and fragmentation are handled by PYTHIA [17]. A full detector simulation and event recon-  
 107 struction has been performed with GEANT4 [18] and CMSSW. Samples and related properties  
 108 are reported in Table 1. All signal samples belong to the RunIISpring15MiniAODv2-74X\_mcRun2  
 109 campaign with the 25ns asymptotic conditions.

110 The following samples are produced assuming the narrow-width approximation, with the res-  
 111 onance width set to 0.1% of the resonance mass.

### 112 2.2 Signal characterization

113 This analysis is performed in a high mass region (TeV scale). The MADGRAPH generates the  
 114 hard process production in the collision with  $p_T = 0$ . In the next step of the simulation, during  
 115 the hadronization, PYTHIA adds the QCD ISR (initial state radiation) and consequently a reso-  
 116 nance  $p_T$  different from 0. The  $p_T$  and rapidity distributions of the heavy resonance after the  
 117 PYTHIA simulation is shown in Figure 4. The typical  $p_T$  is small compared to the mass of the  
 118 resonance, and about two thirds of the events have  $p_T$  smaller than 50 GeV. The  $X \rightarrow Vh$  pro-  
 119 cess is a two body decay and, in the heavy resonance reference frame, the energy of its decay  
 120 products Z and h are univocally defined. In Figure 4 the  $p_T$  distributions of Z and h at generator  
 121 level are reported. These distributions have a Jacobian peak close to  $m_X/2$ , that corresponds  
 122 to the value in  $p_T$  of a two body decay with massless products. In our case, the mass of the  
 123 produced particles, Z and Higgs bosons, is about 91 and 125 GeV respectively (Figure 4) and  
 124 their mass is small compared to the mass of the decayed resonance, which lies at the TeV scale.

The tail on the right of the peak of Figure 4 is due to non-zero  $p_T$  of the original resonance,  
 while the smooth shape at lower  $p_T$  is due to events with a sizable  $p_z$  component. Figure 5  
 shows the  $\Delta R$  distribution at generator level of the Higgs and Z decay products respectively,  
 where for the Higgs decay we simply consider the direction of the partons. Both are peaked at

$$\Delta R = 2 \frac{m_{Z,h}}{p_{T,Z,h}} \approx 4 \frac{m_{Z,h}}{m_X}$$

125 that corresponds to the configuration in which the final particles are emitted perpendicularly  
 126 to the direction of motion of the Higgs or Z; this configuration is preferred by the phase space  
 127 (Jacobian peak). For masses larger than 1000 GeV  $\Delta R$  is often smaller than 0.4, that is the size of  
 128 the jet cone. This consideration leads us to a particular event topology, where the jets produced  
 129 from the quarks or gluons hadronization merge in a single fat jet.

Sample name	Events	$\sigma$ (pb)
ZprimeToZhToZlephbb_narrow_M-600_13TeV-madgraph-v1	49400	-
ZprimeToZhToZlephbb_narrow_M-800_13TeV-madgraph-v1	48400	0.855309
ZprimeToZhToZlephbb_narrow_M-1000_13TeV-madgraph-v1	50000	0.509804
ZprimeToZhToZlephbb_narrow_M-1200_13TeV-madgraph-v1	50000	0.271104
ZprimeToZhToZlephbb_narrow_M-1400_13TeV-madgraph-v1	50000	0.146961
ZprimeToZhToZlephbb_narrow_M-1600_13TeV-madgraph-v1	50000	0.0822156
ZprimeToZhToZlephbb_narrow_M-1800_13TeV-madgraph-v1	50000	0.0473673
ZprimeToZhToZlephbb_narrow_M-2000_13TeV-madgraph-v1	50000	0.0279823
ZprimeToZhToZlephbb_narrow_M-2500_13TeV-madgraph-v1	50000	0.00815289
ZprimeToZhToZlephbb_narrow_M-3000_13TeV-madgraph-v1	50000	0.00257265
ZprimeToZhToZlephbb_narrow_M-3500_13TeV-madgraph-v1	49800	0.000850838
ZprimeToZhToZlephbb_narrow_M-4000_13TeV-madgraph-v1	50000	0.000288261
ZprimeToZhToZlephbb_narrow_M-4500_13TeV-madgraph-v1	50000	-
WprimeToWhToWlephbb_narrow_M-600_13TeV-madgraph-v1	50000	-
WprimeToWhToWlephbb_narrow_M-800_13TeV-madgraph-v1	47600	1.587885
WprimeToWhToWlephbb_narrow_M-1000_13TeV-madgraph-v1	48400	0.986533
WprimeToWhToWlephbb_narrow_M-1200_13TeV-madgraph-v1	50000	0.535394
WprimeToWhToWlephbb_narrow_M-1400_13TeV-madgraph-v1	50000	0.2955239
WprimeToWhToWlephbb_narrow_M-1600_13TeV-madgraph-v1	50000	0.1681478
WprimeToWhToWlephbb_narrow_M-1800_13TeV-madgraph-v1	50000	0.0984325
WprimeToWhToWlephbb_narrow_M-2000_13TeV-madgraph-v1	48800	0.058998
WprimeToWhToWlephbb_narrow_M-2500_13TeV-madgraph-v1	50000	0.01771031
WprimeToWhToWlephbb_narrow_M-3000_13TeV-madgraph-v1	50000	0.00567529
WprimeToWhToWlephbb_narrow_M-3500_13TeV-madgraph-v1	50000	0.001878491
WprimeToWhToWlephbb_narrow_M-4000_13TeV-madgraph-v1	50000	0.00062615
WprimeToWhToWlephbb_narrow_M-4500_13TeV-madgraph-v1	49800	-
ZprimeToZhToZinvhbb_narrow_M-600_13TeV-madgraph-v1	100000	-
ZprimeToZhToZinvhbb_narrow_M-800_13TeV-madgraph-v1	99800	0.855309
ZprimeToZhToZinvhbb_narrow_M-1000_13TeV-madgraph-v1	99200	0.509804
ZprimeToZhToZinvhbb_narrow_M-1200_13TeV-madgraph-v1	99200	0.271104
ZprimeToZhToZinvhbb_narrow_M-1400_13TeV-madgraph-v1	100000	0.146961
ZprimeToZhToZinvhbb_narrow_M-1600_13TeV-madgraph-v2	100000	0.0822156
ZprimeToZhToZinvhbb_narrow_M-1800_13TeV-madgraph-v1	100000	0.0473673
ZprimeToZhToZinvhbb_narrow_M-2000_13TeV-madgraph-v1	98400	0.0279823
ZprimeToZhToZinvhbb_narrow_M-2500_13TeV-madgraph-v1	98400	0.00815289
ZprimeToZhToZinvhbb_narrow_M-3000_13TeV-madgraph-v1	100000	0.00257265
ZprimeToZhToZinvhbb_narrow_M-3500_13TeV-madgraph-v1	100000	0.000850838
ZprimeToZhToZinvhbb_narrow_M-4000_13TeV-madgraph-v1	100000	0.000288261
ZprimeToZhToZinvhbb_narrow_M-4500_13TeV-madgraph-v1	100000	-

Table 1:  $X \rightarrow Vh$  signal samples and production cross sections in HVT model B. The cross section for each relative sample is obtained by multiplying the production cross section by the vector boson and Higgs branching fractions ( $\mathcal{B}(Z \rightarrow \ell\ell) = 0.101$ ,  $\mathcal{B}(Z \rightarrow \nu\nu) = 0.200$ ,  $\mathcal{B}(W \rightarrow \ell\nu) = 0.324$ ,  $\mathcal{B}(h \rightarrow b\bar{b}) = 0.577$ ).

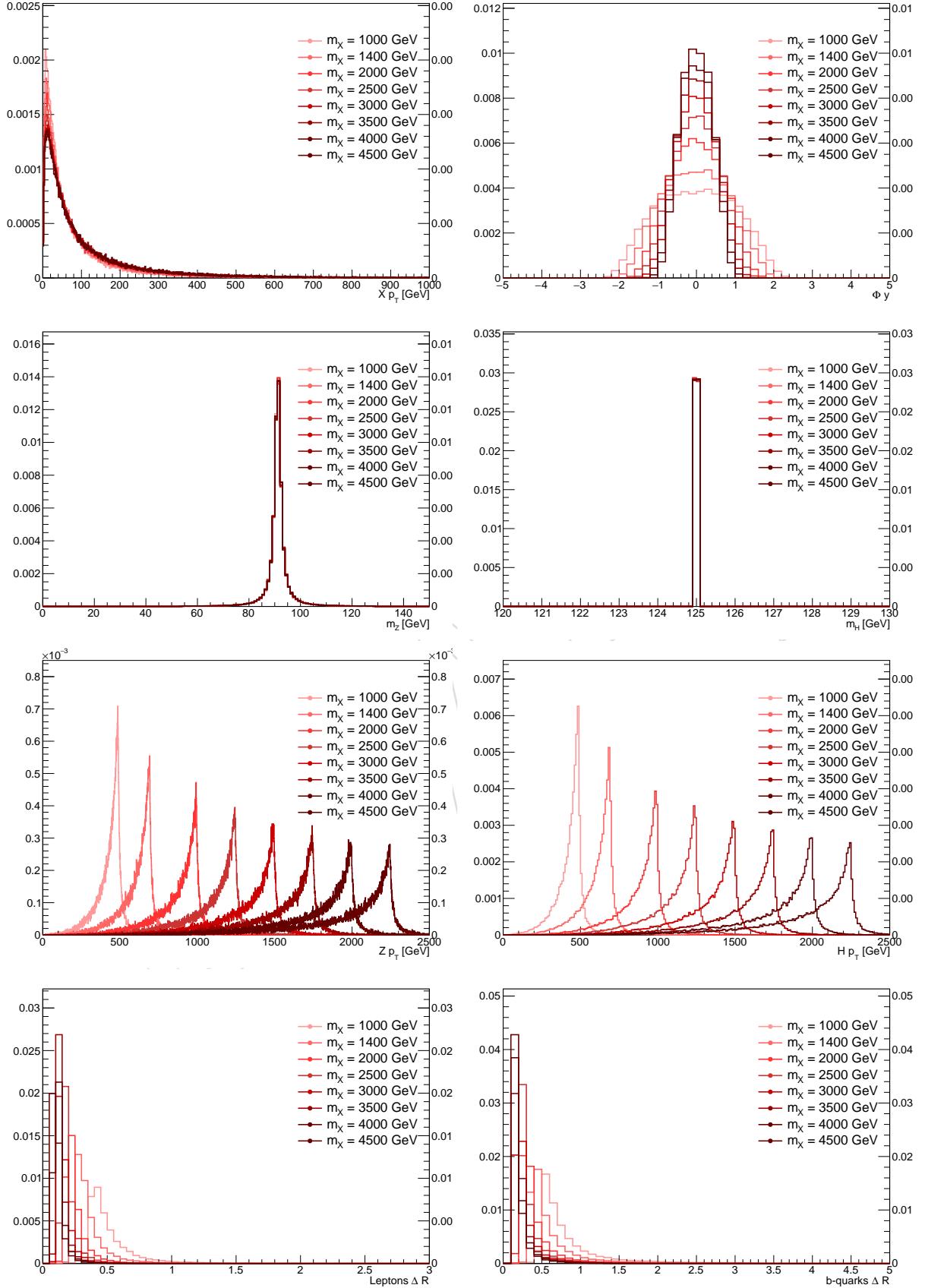


Figure 4: Main signal kinematic quantities at generation level after showering.

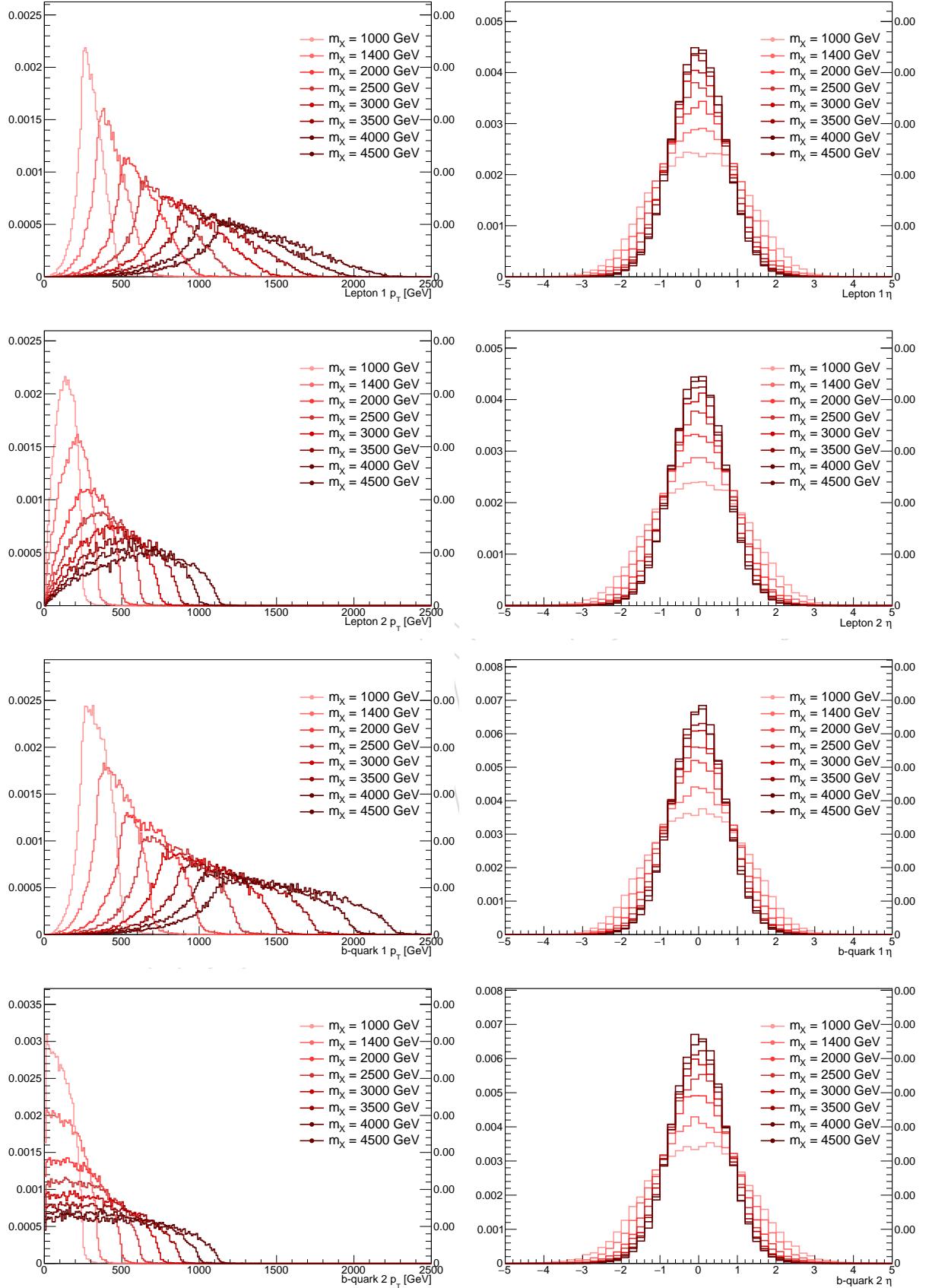


Figure 5: Main signal kinematic quantities at generation level after showering.

130 **2.3 Background samples**

131 All physics processes yielding final states with one or two leptons and a large missing trans-  
 132 verse momentum in association with one or two b quarks have to be considered as possible  
 133 sources of background for the analysis. The complete list of background datasets considered is  
 134 presented in Table 2 and Table 3, where the cross section used to normalize SM backgrounds  
 135 measured by CMS and/or calculated at (N)NLO by the Standard Model Cross Section Working  
 136 Group [] are also reported.

137 • **Z + jets:** this process represents the main irreducible background for the signal in  
 138 the 0- and 2-lepton final states given the large missing transverse momentum or the  
 139 presence of two resonating leptons in the final state respectively. The production of  
 140 single Z/ $\gamma^*$  bosons in association with one or more partons or gluons in the final  
 141 state is topologically similar to the searched signal, but its final state quarks fea-  
 142 ture a generally softer  $p_T$  spectrum, a non-resonant and rapidly falling di-jet mass  
 143 distribution, and other less distinctive characteristics (effective spin and color radia-  
 144 tion) that should theoretically distinguish it from the signal. Before b-quark tagging,  
 145 the contribution from *udscg* (light) partons dominates, while after the application  
 146 of b-tagging the primary contribution in the signal region is from Z + b(b). This  
 147 Z +jets background is produced in one single inclusive sample at NLO with the  
 148 AMC@NLO generator or several samples binned in HT (the sum of the  $p_T$  of the  
 149 hadrons at LHE level) starting from 100 GeV with the MADGRAPH LO generator.  
 150 The contribution of events with HT less than 100 GeV is found to be negligible af-  
 151 ter the selection requirements on the Z  $p_T$  or  $E_T^{\text{miss}}$ . The reconstructed dilepton  $p_T$   
 152 spectra is compared in the inclusive LO Z sample for events with  $H_T > 0$  GeV and  
 153 with  $H_T > 100$  GeV (Fig. 6). The contribution from events with  $H_T < 100$  GeV is  
 negligible after we require the dilepton  $p_T$  to be greater than 200 GeV (Section 5).

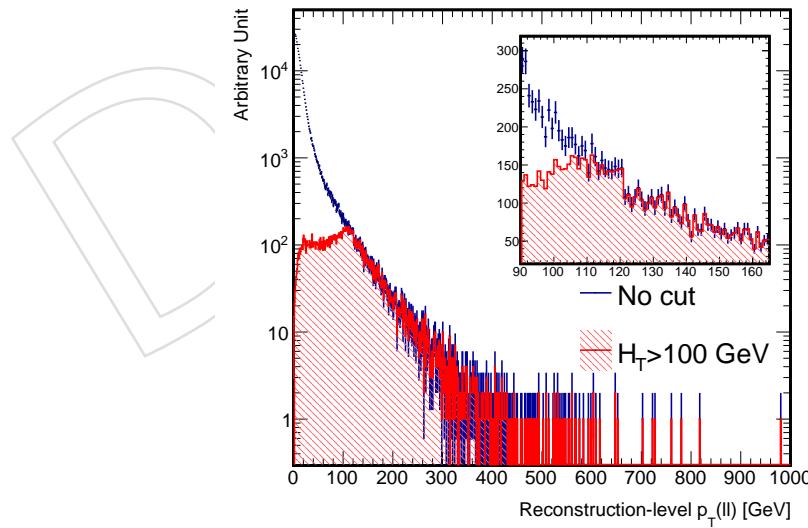


Figure 6: Reconstructed  $p_T(\ell\ell)$  from the LO inclusive Z MC sample: DYJetsToLL\_M-50\_TuneCUETP8M1\_13TeV-madgraphMLM-pythia8, for events without any  $H_T$  re-  
 quirement (markers) and for events with  $H_T > 100$  GeV (filled histogram).

154

155 • **W + jets:** the leptonic decay of a W boson can be an irreducible background in the  
 156 single-lepton channel, or in the zero-lepton channel in the case the charged lepton

escapes undetected (e.g. outside the detector acceptance) or fails the lepton identification requirements. The production of a W boson has a cross section larger by an order of magnitude with respect to the Z, and this makes the W +jets a relevant background also when a lepton veto is applied. Analogously to the Z +jets samples, an inclusive  $W(\rightarrow \ell\nu)$  sample has been produced with AMC@NLO, while an alternative inclusive sample and the HT-binned samples have been produced with the MADGRAPH LO generator.

- **t̄t and single-top:** production of t̄t pairs represents a particularly challenging background at the LHC, given its large production cross section. These events always contain two energetic b-jets and two W bosons which may decay to high  $p_T$ , isolated leptons. Both semi-leptonic and fully-leptonic t̄t decays have been considered, the latter being much more important than the former for the final state considered in this analysis. The primary handles to reduce the t̄t background are topological, such as its larger jet multiplicity and the azimuthal opening angle between the vector boson and the dijet system, which is more broadly distributed in top pair production than in signal events. In the dilepton final state, the most important cut to reduce t̄t is the candidate Z  $p_T$ . In t̄t production the dilepton  $p_T$  spectrum is sharply falling, given the absence of a single boosted resonance. Two generators are tested to simulate t̄t contribution: the NLO generator POWHEG and the LO MADGRAPH. This analysis makes use of t̄t samples based on the former, given the higher statistics and the more accurate prediction of the t̄t  $p_T$  spectra. An inclusive t̄t sample has been produced with POWHEG interfaced with PYTHIA, including all the possible decays of the W bosons. The s-channel single-top sample is produced in the 4-flavor scheme using AMC@NLO, while all the other single-top samples are produced with POWHEG interfaced with PYTHIA.
- **Diboson:** the production of two vector bosons in the SM is a rare process, with a similar kinematics to that of the signal. Furthermore, the boost of the bosons is generally high. The main handle to discriminate against V V backgrounds is a tight cut on the jet mass. The SM Higgs boson production (V h), instead, is virtually indistinguishable from the signal except for the mass resonance itself. However, the SM Higgs production cross section is much smaller than the one of the other diboson processes. All the diboson production processes (W W, W Z, Z Z, W h, Z h) and their corresponding (semi)-leptonic decay modes are considered, including those involving one or more neutrinos. These backgrounds are simulated with the NLO AMC@NLO generator.
- **multipjet (QCD):** despite its enormous cross section at LHC, the probability to produce final states with prompt, isolated leptons or large missing transverse momentum is very low. HT binned samples are generated with the MADGRAPH generator.

Table 2: Z, W +jets simulated samples. The cross section  $\times$  branching ratio is shown in pb.

Dataset	$\sigma \times \mathcal{B}$ [pb]	Events
DYjetsToLL_M-50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	6025.2	28747969
DYjetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	6025.2	9042031
DYjetsToLL_M-50_HT-100to200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	139.4	2725655
DYjetsToLL_M-50_HT-200to400_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	42.75	973937
DYjetsToLL_M-50_HT-400to600_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	5.497	1067758
DYjetsToLL_M-50_HT-600toInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	2.21	998912
DYjetsToNuNu_HT-100To200_13TeV-madgraph	12172.12	20500164
ZjetsToNuNu_HT-200To400_13TeV-madgraph	280.47	5154824
ZjetsToNuNu_HT-400To600_13TeV-madgraph	78.36	4998316
ZjetsToNuNu_HT-600ToInf_13TeV-madgraph	10.944	1018882
WjetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	4.203	1008333
WjetsToLNu_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	61526.7	24184766
WjetsToLNu_HT-100To200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1292.0	10152718
WjetsToLNu_HT-200To400_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	385.9	5221599
WjetsToLNu_HT-400To600_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	47.9	1745914
WjetsToLNu_HT-600ToInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	19.9	1039152

Table 3:  $t\bar{t}$ , dibosons and multijet simulated samples. The cross section  $\times$  branching ratio is shown in pb.

Dataset	$\sigma \times \mathcal{B}$ [pb]	Events
TT_TuneCUETP8M1_13TeV-powheg-pythia8	831.76	96834559
TTWJetsToLNu_M-10_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8	0.2043	252908
TTZToLLNuNu_M-10_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.2529	398000
ST_s-channel_4f_leptonDecays_13TeV-amcatnlo-pythia8_TuneCUETP8M1	3.65792	984400
ST_t-channel_antitop_4f_leptonDecays_13TeV-powheg-pythia8_TuneCUETP8M1	26.0659	1680200
ST_t-channel_top_4f_leptonDecays_13TeV-powheg-pythia8_TuneCUETP8M1	43.79844	3299800
ST_tW_antitop_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1	35.6	988500
ST_tW_top_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1	35.6	995600
WWTo1L1Nu2Q_13TeV_amcatnloFXFX_madspin_pythia8	49.997	5035754
WWTo1L1Nu2Q_13TeV_amcatnloFXFX_madspin_pythia8	10.71	24714550
WWTo1L3Nu_13TeV_amcatnloFXFX_madspin_pythia8	3.05	1696910
WWTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	5.595	31477411
ZZTo2Q2Nu_13TeV_amcatnloFXFX_madspin_pythia8	4.04	36840500
ZZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	3.22	18790122
ZZTo4L_13TeV-amcatnloFXFX-pythia8	1.212	10561099
ZH_HToBB_ZToLL_M125_13TeV_amcatnloFXFX_madspin_pythia8	0.0507	2154718
WH_HToBB_WToLNu_M125_13TeV_amcatnloFXFX_madspin_pythia8	0.25798824	2178916
ZH_HToBB_ZToNuNu_M125_13TeV_amcatnloFXFX_madspin_pythia8	0.1003	2115877
QCD_HT100to200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	2754000.0	81637494
QCD_HT200to300_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1735000.0	18718905
QCD_HT300to500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	366800.0	19826197
QCD_HT500to700_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	29370.0	19664159
QCD_HT700to1000_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	6524.0	15356448
QCD_HT1000to1500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1064.0	4963895
QCD_HT1500to2000_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	121.5	3868886
QCD_HT2000toInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	25.42	1912529

195 **2.4 V boson momentum corrections**

196 **2.4.1 NLO QCD**

197 In Run II, the use of next-to-leading order generators, such as aMCatNLO, allowed to have a  
 198 much better description of the vector bosons ( $Z$ ,  $W$ ) with respect to Run I, when only leading  
 199 order generators were available. This is confirmed by data/simulation comparison. Unfor-  
 200 tunately, NLO generators have not been used to generate large exclusive samples with the  
 201 high statistics needed for analyses in the high- $p_T$  regime. Instead, exclusive MADGRAPH sam-  
 202 ples are available. In these, the  $p_T$  spectra of the  $W$  and  $Z$  bosons is known to be non-perfectly  
 203 described, compared to data and the inclusive aMCatNLO sample, as seen in Fig. 7-left. Cor-  
 204 rections are derived to improve the simulation description. Instead of a correction function as  
 205 a function of the  $p_T$  at generation level, separate multiplicative factors (k-factors) are derived  
 206 for exclusive MADGRAPH starting from the inclusive aMCatNLO sample, in order to take into  
 207 account the effect of QCD NLO processes.

208 The procedure to derive the k-factors is the following. Since in the  $V$   $p_T$  spectrum there is  
 209 overlap between the HT-binned exclusive samples, one solution could be to separate the  $p_T$   
 210 spectrum in at least four regions and solve a linear system for the normalization of the samples  
 211 in each of these regions. However, the transverse momentum at generation level of a certain  
 212 exclusive sample never goes above upper HT threshold. This is always true before showering,  
 213 but even after taking showering into account the effect is very small (< 0.3%). The matrix  
 214 representing the linear system can thus be considered diagonal, and a simpler approach is  
 215 followed. The ratio between the inclusive and the higher-HT exclusive sample, normalized to  
 216 the corresponding cross sections and evaluated in the  $p_T$  range above the lower HT threshold,  
 217 is taken as the k-factor for the sample, and its normalization is then considered as fixed. In  
 218 the following steps, k-factors for the lower HT exclusive sample are evaluated with the same  
 219 procedure, but taking into account the fixed contribution of the exclusive samples with higher-  
 220 HT binning.

221 This k-factor extraction procedure is performed for DYJetsToLL, DYJetsToNuNu, and WJets-  
 222 sToLNu samples independently. Small differences arise, but generally the k-factor can be as  
 223 large as 1.5 in the lower part of the  $p_T$  spectrum, and very close to 1 in the higher-HT samples.  
 224 The numerical results are reported in Table 4. The agreement with the inclusive NLO sample  
 225 after the reweighting is flat at 1, and is reported in Figure 7 and 8.

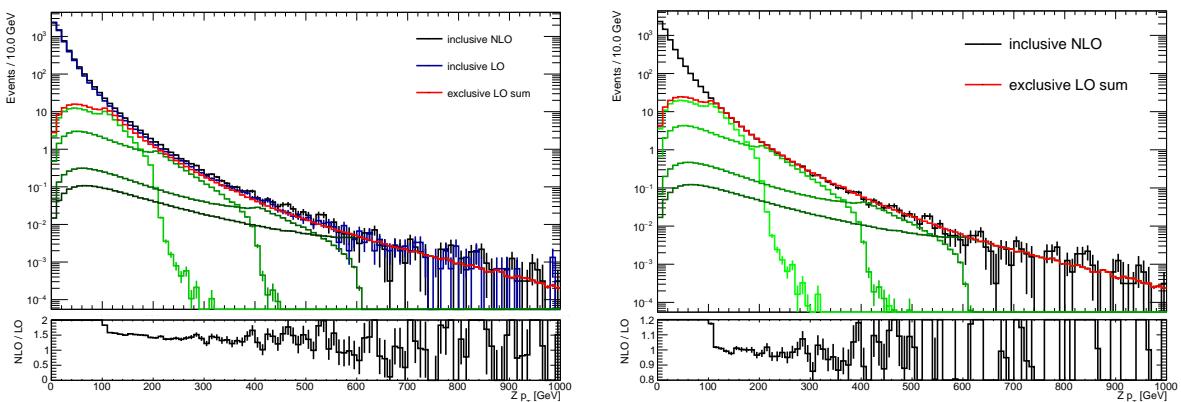


Figure 7:  $p_T$  spectrum for the inclusive NLO and exclusive LO samples for the  $Z \rightarrow \ell\ell$  process before (left) and after (right) the k-factor application.

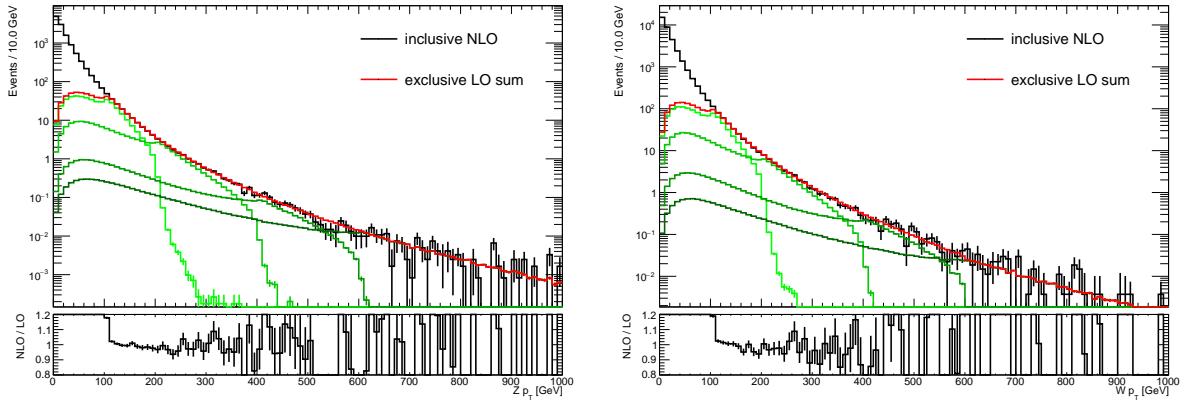


Figure 8:  $p_T$  spectrum for the inclusive NLO and exclusive LO samples for the  $Z \rightarrow \nu\nu$  (left) and  $W \rightarrow \ell\nu$  processes (right) after the k-factor application.

Dataset	k-factor
DYJetsToLL_M-50_HT-100to200	1.588
DYJetsToLL_M-50_HT-200to400	1.438
DYJetsToLL_M-50_HT-400to600	1.494
DYJetsToLL_M-50_HT-600toInf	1.139
ZJetsToNuNu_HT-100To200	1.626
ZJetsToNuNu_HT-200To400	1.617
ZJetsToNuNu_HT-400To600	1.459
ZJetsToNuNu_HT-600ToInf	1.391
WJetsToLNu_HT-100To200	1.459
WJetsToLNu_HT-200To400	1.434
WJetsToLNu_HT-400To600	1.532
WJetsToLNu_HT-600ToInf	1.004

Table 4: K-factors for the V+jets samples.

#### 226 2.4.2 NLO Electroweak

227 Further corrections to the V  $p_T$  spectrum comes from NLO electroweak contributions, that  
 228 become more and more important with the transverse momentum. These corrections, applied  
 229 on top of the k-factors, are effectively applied on a per-event basis depending on the  $p_T$  of the  
 230 vector boson at generation level. The calculation of these contributions is explained in Ref. [19].  
 231 Figure 9 shows the amount of the correction for the W and Z bosons.

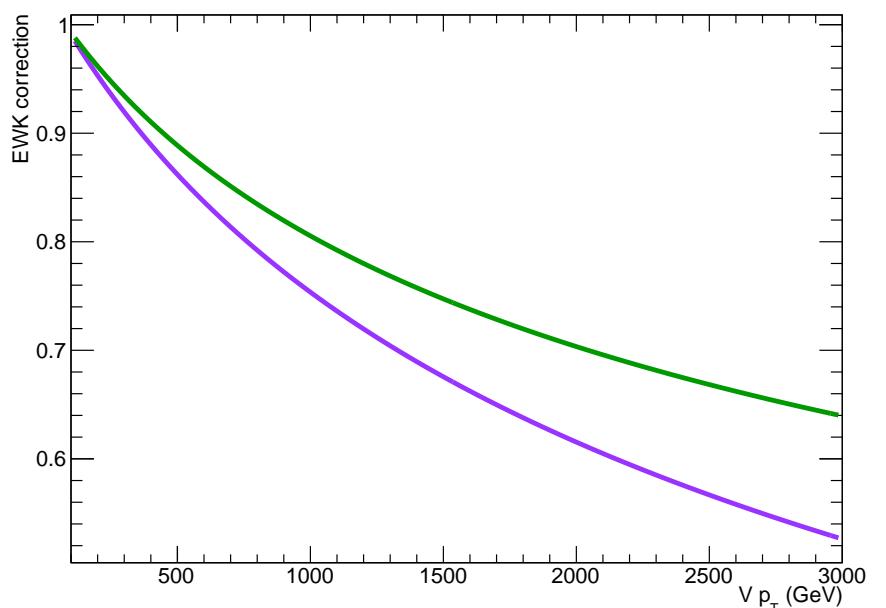


Figure 9: Electroweak corrections for the Z (green line) and W boson (purple line) as a function of the transverse momentum [19].

232 **2.5 Data**

233 Data samples used in this analysis have been collected during 2015 RunC and RunD, at a center-  
 234 of-mass energy of 13 TeV, in 25ns runs and with the magnetic field enabled. The MET primary  
 235 dataset is used for the regions where a missing energy trigger is required, and the SingleMuon  
 236 and SingleElectron are used for data selected with a single muon and electron trigger,  
 237 respectively. The full list of datasets used is shown in Table 5. Data is processed from 05Oct  
 238 re-miniAOD campaign when available, and from PromptReco otherwise.

239 Two JSON files are used in the analysis, depending on the channel:

240 **Golden:** Cert\_246908-258750\_13TeV\_PromptReco\_Collisions15\_25ns\_JSON.txt  
 241 includes all the runs certified as “good” for all subsystems. The integrated luminos-  
 242 ity amounts to  $2.11 \text{ fb}^{-1}$ .

243 **Silver:** Cert\_246908-258750\_13TeV\_PromptReco\_Collisions15\_25ns\_JSON\_Silver.txt  
 244 includes runs affected by a decrease in the HF occupancy, that in turn can spoil all  
 245 the physics objects reconstructed using HF. The integrated luminosity amounts to  
 246  $2.46 \text{ fb}^{-1}$ .

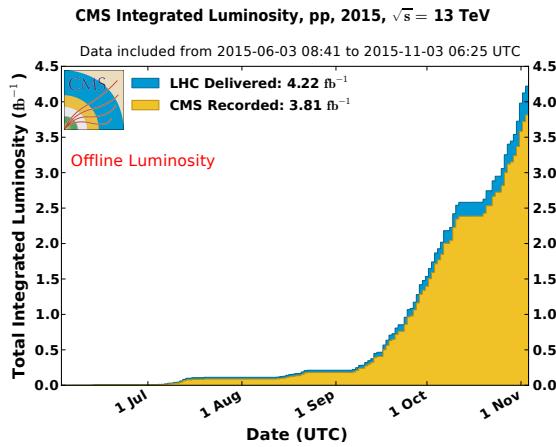


Figure 10: Cumulative luminosity versus day delivered to (blue), and recorded by CMS (orange) during stable beams and for pp collisions at 13 TeV centre-of-mass energy in 2015. The delivered luminosity accounts for the luminosity delivered from the start of stable beams until the LHC requests CMS to turn off the sensitive detectors to allow a beam dump or beam studies. Given is the luminosity as determined from counting rates measured by the luminosity detectors

247 The “golden” JSON file is applied for categories that require a full PF event description, and  
 248 make use of the PF missing energy; this is the case of the zero- and 1-lepton channels. In the  
 249 2-lepton channel, PF  $E_T^{\text{miss}}$  is not used, and a list of runs and lumisections are included in the  
 250 “silver” JSON file.

251 In order to remove problematic or noise-dominated events, the following list of filters have  
 252 been applied on data:

- 253   • HBHENoiseFilter
- 254   • HBHENoiseIsoFilter
- 255   • CSCTightHaloFilter
- 256   • eeBadScFilter

257 • goodVertices

Table 5: Datasets Run2015C and Run2015D.

Dataset
MET/Run2015D-PromptReco-v4
MET/Run2015D-05Oct2015-v1
MET/Run2015C_25ns-05Oct2015-v1
SingleMuon/Run2015D-PromptReco-v4
SingleMuon/Run2015D-05Oct2015-v1
SingleMuon/Run2015C_25ns-05Oct2015-v1
SingleElectron/Run2015D-PromptReco-v4
SingleElectron/Run2015D-05Oct2015-v1
SingleElectron/Run2015C_25ns-05Oct2015-v1

258 **2.6 Trigger**

259 Events are selected on-line by a two-stage trigger. The Level 1 (L1) trigger consists of hardware  
 260 processors that perform a very basic selection and counting of physics objects, and reduce the  
 261 rate from 40 MHz to up to 100 kHz. Events passing the L1 decision are acquired by the DAQ  
 262 system, and a complete and more accurate reconstruction is performed by the High Level Trig-  
 263 ger (HLT), which exploits similar but faster variations of the same algorithms used in the offline  
 264 event reconstruction. A trigger path is a string that identifies a list of selections performed at  
 265 HLT. Single lepton triggers, requiring at least one, non-isolated lepton, have been used to select  
 266 events for the one and two-lepton categories. Events with no genuine leptons, characterized by  
 267 large missing transverse momentum final states, pure  $E_T^{\text{miss}}$  triggers or  $E_T^{\text{miss}}$  triggers combined  
 268 with other event requirements are used.

269 The list of triggers used is reported in Table 6.

Table 6: HLT trigger paths used in the analysis.

HLT path
HLT_PFMETNoMu90_JetIdCleaned_PFMHTNoMu90_IDTight
HLT_PFMETNoMu90_NoiseCleaned_PFMHTNoMu90_IDTight
HLT_PFMETNoMu90_PFMHTNoMu90_IDTight
HLT_PFMETNoMu120_JetIdCleaned_PFMHTNoMu120_IDTight
HLT_PFMETNoMu120_NoiseCleaned_PFMHTNoMu120_IDTight
HLT_PFMETNoMu120_PFMHTNoMu120_IDTight
HLT_PFMET170_NoiseCleaned (backup)
HLT_Mu45_eta2p1
HLT_Mu50
HLT_Ele105_CaloIdVT_GsfTrkIdT
HLT_Ele115_CaloIdVT_GsfTrkIdT

270 The same set of trigger paths are required to be fired for both data and simulated events. The  
 271  $E_T^{\text{miss}}$  triggers are the logic OR of different trigger quantities, with thresholds on both the MET  
 272 and the HT computed using particle flow objects. These HLT triggers are all seeded at L1 by  
 273 L1\_ETM\_50, L1\_ETM\_60, L1\_ETM\_70, which are the lowest unprescaled L1 trigger common  
 274 to all the HLT  $E_T^{\text{miss}}$  paths. The efficiency of the  $E_T^{\text{miss}}$  triggers is measured selecting  $W \rightarrow \mu\nu$   
 275 events using a SingleMuon trigger HLT\_IsoMu20 or HLT\_IsoTkMu20, where the contribution  
 276 of the muon is subtracted from the  $E_T^{\text{miss}}$  computation as in the online algorithm. The trigger

277 path originally intended for the zero lepton category, `HLT_PFMET170_NoiseCleaned`, is now  
278 being ruled out and kept as a backup solution, because the logical OR of paths with lower MET  
279 thresholds allow a sharper turn-on curve, that reaches the plateau just above  $E_T^{\text{miss}} \approx 200 \text{ GeV}$ .  
280 The efficiencies of the single muon triggers are provided centrally by the Muon POG [20] with a  
281 tag and probe procedure by selecting  $Z \rightarrow \ell\ell$  events. Tight lepton ID requirements are applied  
282 to the *probes*. The trigger efficiency is then evaluated studying the *tag* lepton efficiency as a  
283 function of both  $p_T$  and  $\eta$  for both data and MC. The muon trigger scale factors are applied  
284 consistently to the simulation throughout the analysis. The average scale factor for the triggers  
285 used is about 95%.  
286 The single electron trigger turn-on curve is evaluated in the diboson VZ Analysis Note [21],  
287 and a flat  $1.002 \pm 0.017$  scale factor is applied in the electron channels. As a cross check, the  
288 trigger efficiency is re-derived using a tag-and-probe method. Since the muon trigger is in-  
289 dependent on the electron trigger, the muon sample can be used to measure the performance  
290 of the electron trigger in data and simulation. For this study, a sample with one muon (satis-  
291 fying trigger, id and isolation, with the corresponding scale factors applied) and one electron  
292 is selected. The electron is required to pass the loose working point, and to have a  $p_T$  larger  
293 than 135 GeV so that the single electron trigger is fully efficient. In order to obtain a clean  $t\bar{t}$   
294 sample, the electron and muon are additionally required to have opposite electric charge. The  
295 resulting sample consists of 581 events in data with 605.03 predicted events from simulation.  
296 The selected sample is a very clean  $t\bar{t}$  sample, with more than 98% of the events predicted to  
297 originate from  $t\bar{t}$  production, and a negligible fraction from  $W + \text{jets}$  and  $Z + \text{jets}$  production.  
298 The efficiency of the electron trigger is given by the fraction of events in this dilepton sample  
299 that fired the `HLT_Ele105_CaloIdVT_GsfTrkIdT` trigger. The number of events that fired the  
300 electron trigger is 557 in data, expecting 573.02 events from simulation. The resulting total effi-  
301 ciency in data and simulation is estimated to be 0.959 and 0.947 respectively. The comparison  
302 between the efficiencies measured in data and simulation leads to a scale factor of  $1.012 \pm 0.057$   
303 (statistical only), which is very close and within uncertainties with respect to the one found by  
304 the VZ analysis.  
305 The turn-on curve for the  $E_T^{\text{miss}}$  trigger used in this analysis is shown in Figure 11 as a func-  
306 tion of the offline reconstructed missing transverse momentum. A complete description of the  
307 method used to derive trigger efficiencies and scale factor can be found in Ref. [22]. Trigger  
308 scale factors are applied consistently in simulation to take into account residual efficiency dif-  
309 ferences.

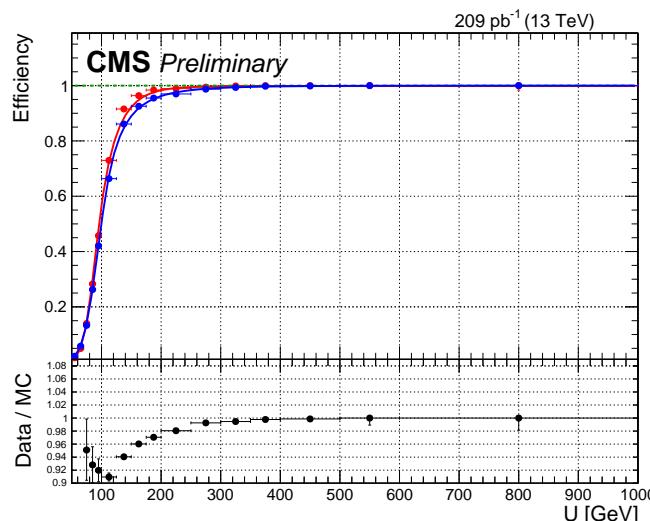


Figure 11: Trigger efficiency for the OR of the two HLT\_PFMETNoMu90\_PFMHTNoMu90 HLT\_PFMETNoMu120\_PFMHTNoMu120 HLT paths as function of the offline  $E_T^{\text{miss}}$  ( $U$  in the plot) for MC (red) and Data (blue) [22].

### 3 Physics objects

In this section, a list of the physics objects used in the analysis is presented, together with performance and validation plots.

The objects are selected according to the standard Run2 recommendations provided by the various POGs for the Spring15 (25ns).

The version of CMSSW used for the analysis is CMSSW\_7\_4\_14.

#### 3.1 Vertex and Pile-up

Due to pileup several primary vertices are typically reconstructed in an event. The primary vertex of the event is defined as the one with the highest sum of transverse momenta  $\sum p_T^2$  of the tracks associated to it, which passes the following selections:

- number of degrees of freedom  $N_{DoF} > 4$
- vertex position along the beampipe  $|z_{vtx}| < 24 \text{ cm}$
- vertex distance with respect the beam pipe  $d_0 < 2 \text{ cm}$

where  $z_{vtx}$  and  $d_0$  are the distance along and perpendicular to the beam line of the vertex with respect the nominal interaction point  $(0, 0, 0)$ .

The data sample contains a significant number of additional interactions per bunch crossing, an effect known as pileup (PU).

The Spring15 v2 MINIAOD Monte Carlo samples are generated simulating the PU conditions, using the 25ns asymptotic PU scenario. Nevertheless, the MC PU description do not match exactly the conditions in data, and there is therefore the need to reweight the simulated events in order to improve the agreement with the data.

The MC samples are reweighted using the standard CMS PU reweighting technique [23, 24] assuming a total inelastic cross section of  $\sigma_{in} = 69\,000\mu\text{b}$ . The pileup distribution is shown in Figure 12.

The comparison between the distributions of primary vertices in data and MC after the PU reweighting is applied is shown in Figure 13 for events selected in the Z peak and with a single-lepton enriched selections.

#### 3.2 Electrons

Electrons are reconstructed from energy deposits in the ECAL matched to tracks reconstructed in the silicon tracker. The electron trajectories are reconstructed using a dedicated modeling of the electron energy loss and fitted with a Gaussian sum filter. Electrons used in this analysis are required to pass the Particle Flow criteria, and to fall in the ECAL pseudorapidity fiducial range ( $|\eta| < 2.5$ ).

The electron identification used in this analysis is based on the “cut-based” Id defined by the EGamma POG for the Spring15 25ns [25]. Isolation cuts are already applied within the cut-based Id definitions, therefore no additional Isolation cut is required. In the isolation definition the effect of PU is considered by taking into account the energy deposits in the calorimeter, estimated through the so-called  $\rho$ -area method, by subtracting the median energy density in the event  $\rho$  multiplied by electron effective area. The isolation value is computed in a  $\Delta R$  cone of 0.3 centered along the lepton direction.

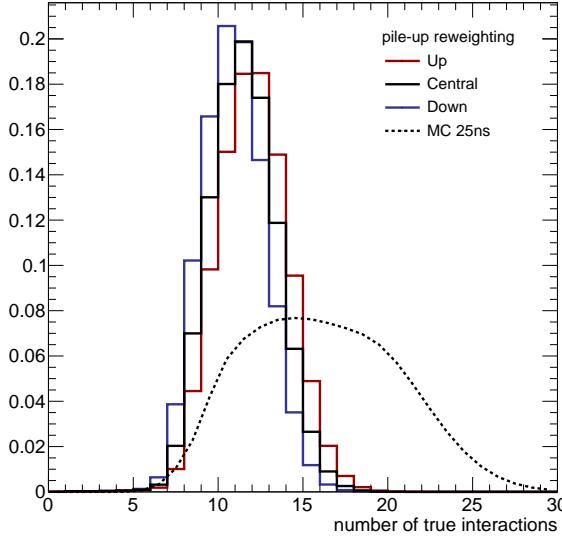


Figure 12: Pile-up distribution estimated from data assuming a total inelastic cross section of  $\sigma_{in} = 69000\mu b$ . The red and blue lines correspond to  $\pm 5\%$  variation of the cross section value.

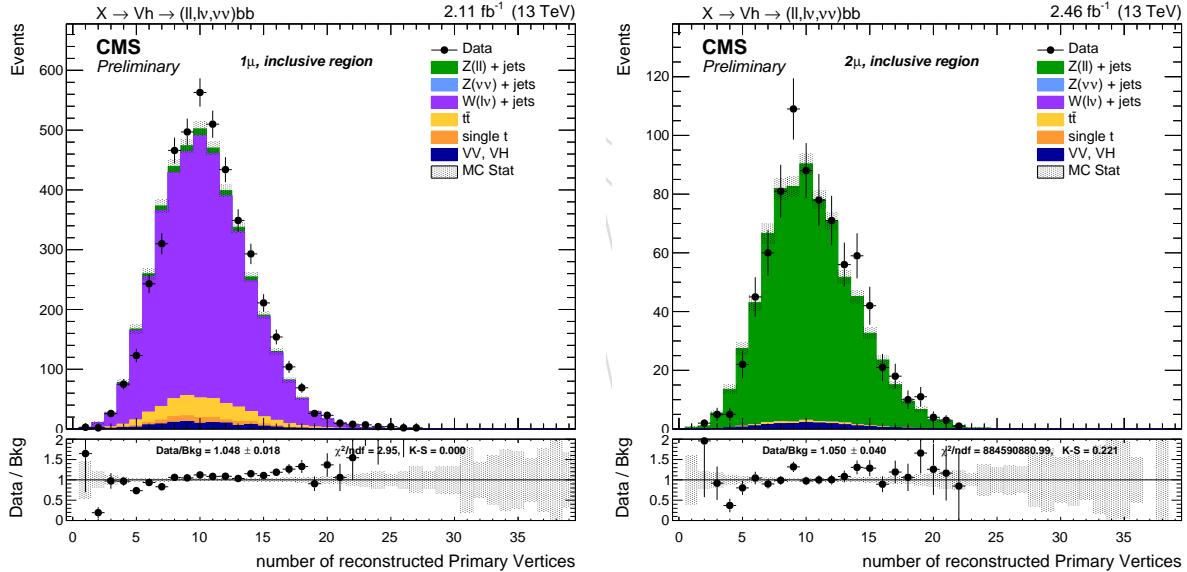


Figure 13: Primary vertices distributions after reweighting with the official recipe and  $\sigma_{in} = 69 000\mu b$ , in the  $1\mu$  (left) and  $2\mu$  selection (right).

Three different electron cut-based Id are used, `Veto`, `loose` and `tight`, depending on the channel.  
 The detailed set of cuts are reported in the Table 7.  $\Delta\eta_{in}^{seed}$  and  $\Delta\varphi_{in}$  are the difference in  $\eta$  and  $\varphi$  between the track position as measured in the inner layer, extrapolated to the interaction vertex and then extrapolated to the calorimeter and the  $\eta$  of the seed cluster or the  $\varphi$  of the supercluster,  $H/E$  is the ratio of the hadronic energy of the CaloTowers in a cone of radius 0.15 centred on the electron's position in the calorimeter to the electromagnetic energy of the electron's supercluster,  $\sigma_{\eta\eta\eta}$  is the spread in eta in units of crystals of the electrons energy in 5x5 block centred on the seed crystal,  $E^{2\times 5}/E^{5\times 5}$  is fraction of energy in 2x5 crystals around seed to the energy in 5x5 crystals around the seed, and  $1/E - 1/p$  is the difference of the inverse of the energy and the momentum.

Table 7: Spring15 cut-based selection for 25ns conditions. EB: barrel cuts ( $|\eta_{\text{supercluster}}| \leq 1.479$ ); EE: endcap cuts ( $|\eta_{\text{supercluster}}| > 1.479$ )

Electrons	Veto		loose		tight	
	EB	EE	EB	EE	EB	EE
$\sigma_{i\eta i\eta}$	<	0.0114	0.0352	0.0103	0.0301	0.0101
$\Delta\eta_{in}^{\text{seed}}$	<	0.0152	0.0113	0.0105	0.00814	0.00926
$\Delta\varphi_{in}$	<	0.216	0.237	0.115	0.182	0.0336
$H/E$	<	0.181	0.116	0.104	0.0897	0.0597
rellIso (EA)	<	0.126	0.144	0.0893	0.121	0.0354
$1/E - 1/p$	<	0.207	0.174	0.102	0.126	0.012
$ d_0 $	<	0.0564	0.222	0.0261	0.118	0.0111
$ d_z $	<	0.472	0.921	0.41	0.822	0.0466
missing hits	$\leq$	2	3	2	1	1
conversion veto	yes	yes	yes	yes	yes	yes

360 An alternative electron identification procedure (*not used in the analysis*) is investigated, based  
 361 on a different set of cuts defined by the EGamma POG and optimized for high- $p_T$  electrons  
 362 and called HEEP Id [26], currently used in the  $Z' \rightarrow ee$  analysis. However, the predefined  
 363 selections include also isolation cuts which are not suitable in case the electrons are too close to  
 364 each other, and the footprint of one electron tends to be included in the isolation computation  
 365 of the other, leading to a drop in efficiency. With the HEEP Id, electrons are required to have  
 366  $p_T > 35 \text{ GeV}$  and lie in the ECAL fiducial region ( $|\eta| < 2.5$ , excluding the transition region  
 367  $|\eta| < 1.4442$ ,  $|\eta| > 1.566$ ) the HEEP id v6.0 is then used only for the electron identification.  
 368 The official HEEP Id also includes two additional selections that depend on electron isolation.  
 369 The original HEEP Id is optimized for  $Z' \rightarrow ee$  searches, where the two electrons are emitted  
 370 back-to-back and isolated. This does not happen in the present case, where lepton isolation is  
 371 critical. In this analysis, the two isolation cuts are not included in the present HEEP definition,  
 372 which is then called HEEPNoIso. The efficiency for each cut in the HEEPNoIso selection is  
 373 shown in Figure 15. The list of selections, separated for electron's superclusters in the barrel  
 374 and endcaps regions, are:

- 375 • the electron is driven by calorimetric deposits (`isEcalDriven`)
- 376 •  $|\Delta\eta_{in}^{\text{seed}}| < 0.004$  in the barrel and  $< 0.006$  in the endcaps
- 377 •  $|\Delta\varphi_{in}| < 0.06$  in both the barrel and the endcaps
- 378 •  $H/E < 1/E + 0.05$  in the barrel,  $5/E + 0.05$  in the endcaps
- 379 •  $\sigma_{i\eta i\eta} < 0.03$  only for endcap electrons
- 380 •  $E^{2x5}/E^{5x5} > 0.94$  or  $E^{1x5}/E^{5x5} > 0.83$  for barrel electrons
- 381 • fraction of energy in  $2x5$  crystals around seed to the energy in  $5x5$  crystals around  
   the seed  $> 0.94$ , or fraction of energy in  $1x5$  crystals around seed to the energy in  $5x5$   
   crystals around the seed  $> 0.84$
- 384 • the number of inner tracker layer lost hits  $\leq 1$
- 385 •  $|d_{xy}|$  smaller than 0.02cm in the barrel and 0.05cm in the endcaps

386 The electron identification efficiency is calculated in signal samples with respect to the gener-  
 387 ation level electrons but factorizing out the acceptance,  $p_T$  threshold and matching terms. The  
 388 efficiency formula is then  $\epsilon = \frac{\text{matched electrons passing preselections and Id}}{\text{matched electrons passing preselections}}$ , and it is found to be close to  
 389 1 for very high momentum electrons, and always above 0.8 even for relatively low- $p_T$  electrons  
 390 around  $\sim 50 \text{ GeV}$ , as shown in Figure 15. Due to the boost of the  $Z$  boson, for high resonance

masses the electrons are quite close to each other, with an average  $\Delta R$  separation  $\lesssim 0.2$ . Therefore it is critical to check that there is no efficiency drop if the two electrons are very close to each other, as it happens for HEEP electrons (Figure 14). The standard electron *veto*, *loose* and *tight* working points have been tested as well. The *veto* and *loose* working points offer slightly better performances with respect to the HEEPNolIso selections, while the efficiency for *tight* electrons is noticeably lower (Figure 14).

After extensive testing of the benefits and performance of the various identification methods and working points, electrons in the present analysis are identified with the standard identification method, and the *veto* and *tight* working point. The former is used to count and identify electrons in all the leptonic categories. The latter is used only in the 1-electron category to identify the leading electron. The reason for this choice is that the standard *tight* Id and isolation offers a better QCD rejection than the HEEP selection. A looser working point is instead used to identify electrons in the 2-electrons category, where a sufficiently loose selection criterion is needed in order to select leptons arising from a boosted Z, potentially close in  $\Delta R$ . In this case the *loose* working point is used. The HEEP Id is not then used in the analysis, but it is kept only as a cross-check.

Validation of the electron object is performed with a inclusive  $Z \rightarrow e^+e^-$  selection, for events passing the trigger `HLT_Ele105_CaloIdVT_GsfTrkIdT`, and the leading (sub-leading) electron passing the *tight* (*veto*) Id, and a  $p_T$  threshold of 30 GeV (10 GeV).

Scale factors for the electron identification (including isolation) are derived through the tag-and-probe method on the Z mass peak for all the working points separately, as a function of the  $p_T$  and  $\eta$  of the electrons. The numerical values are reported in Ref. [22].

The data/simulation comparison is shown in Figures 16-18.

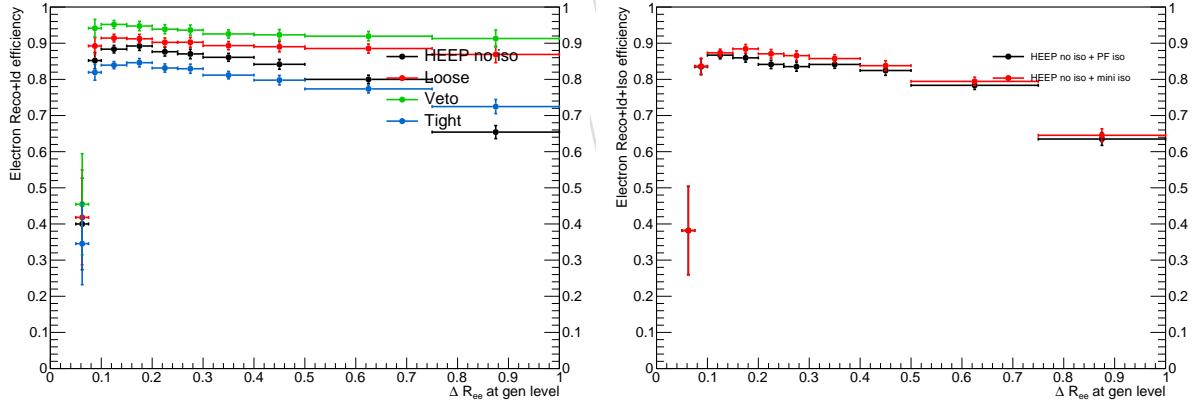


Figure 14: Electron identification efficiency (left) and combined identification and isolation efficiency (right) as a function of the  $\Delta R$  between the two electrons at generation level, after matching and  $p_T$  selections. Events from all the generated mass points are considered together.

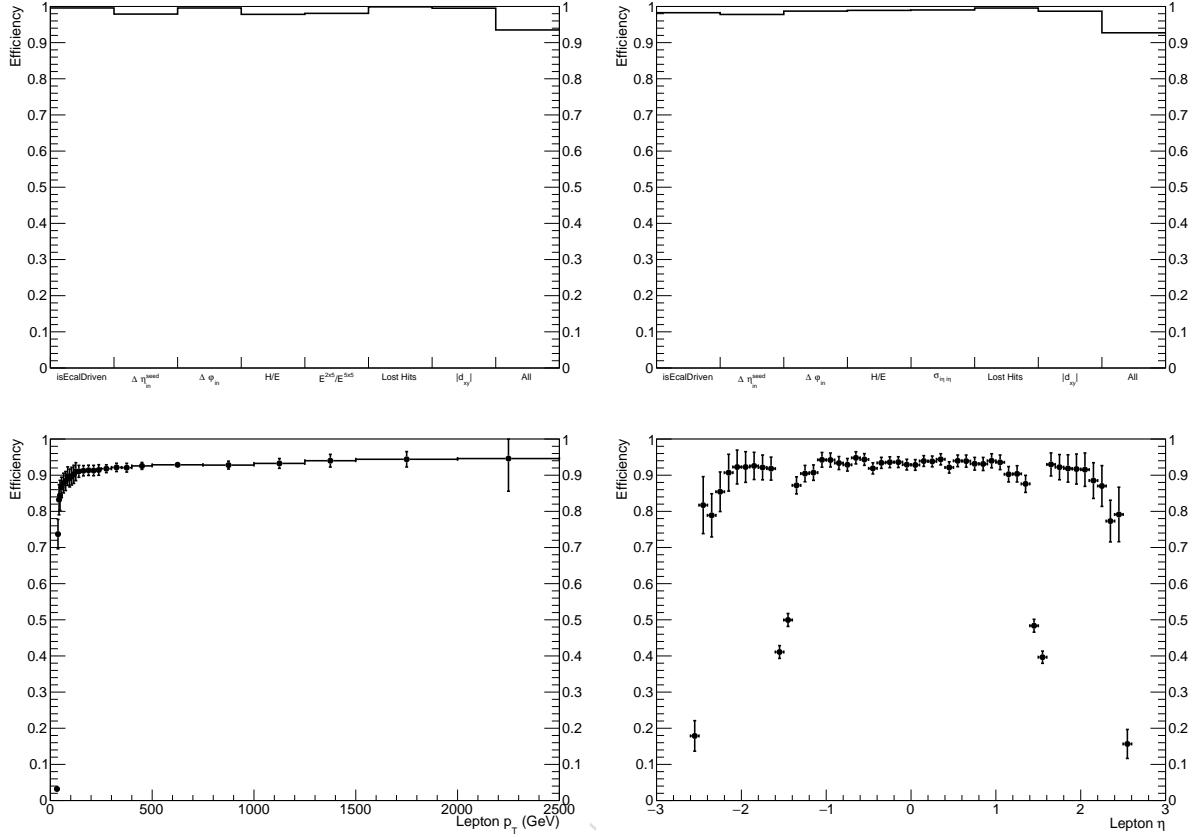


Figure 15: Top: efficiency of each single cut for the HEEP selections for barrel (left) and endcaps (right) for the  $m_X = 2000$  GeV signal sample. Bottom: efficiency as a function of the electron gen  $p_T$  (left) and  $\eta$  (right) for the HEEPNolIso selections, after matching and  $p_T$  selections. Events from all the generated mass points are considered together.

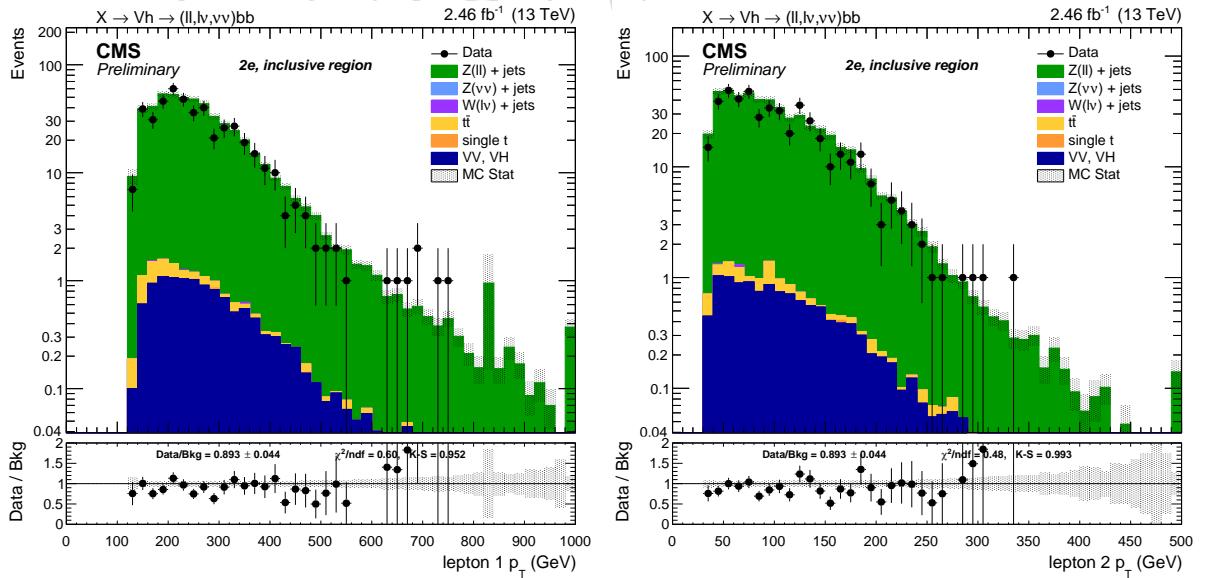


Figure 16: Leading (left) and sub-leading (right) electron  $p_T$  spectra after  $Z \rightarrow ee$  selections.

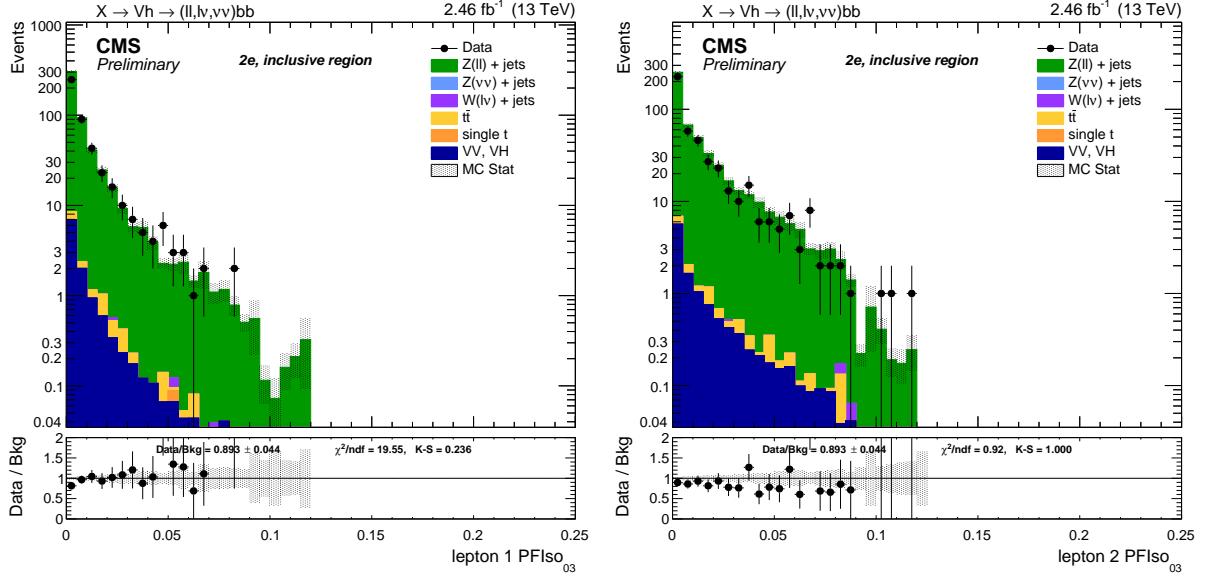


Figure 17: Leading (left) and sub-leading (right) electron isolation distribution after  $Z \rightarrow ee$  selections.

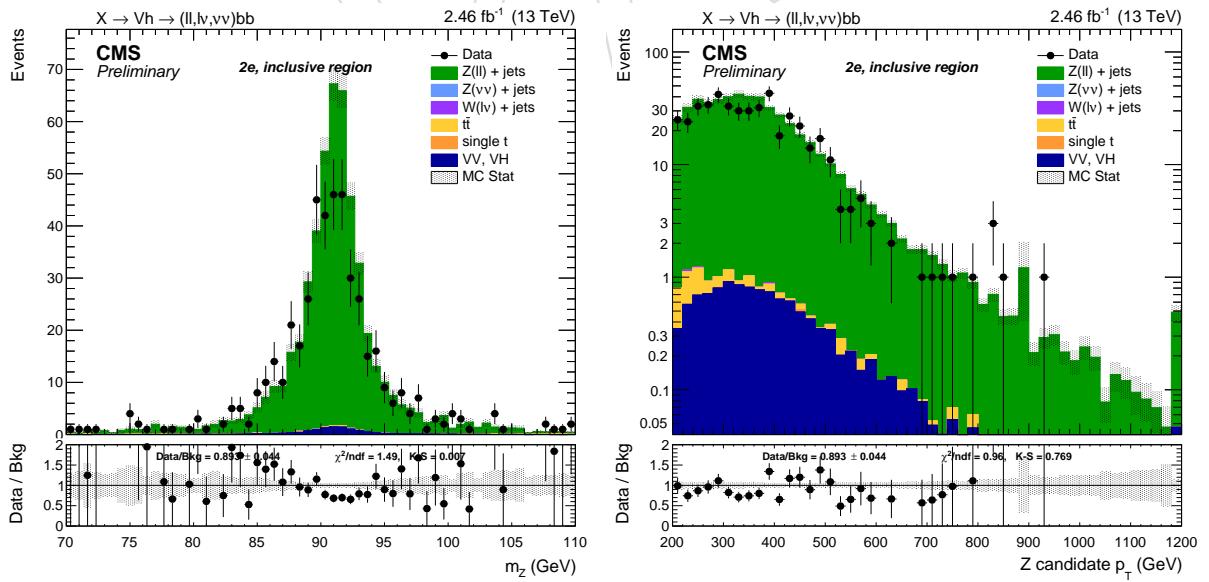


Figure 18: Reconstructed  $Z$  mass (left) and  $Z$   $p_T$  (right) after the inclusive  $Z \rightarrow ee$  selections.

### 414 3.3 Muons

415 In the standard CMS reconstruction for  $pp$  collisions, muon tracks are first reconstructed inde-  
 416 pendently in the inner tracker (tracker track) and in the muon system (standalone-muon track).  
 417 Based on these objects, two reconstruction approaches are used [27]: *Global Muon* (outside-in)  
 418 and *Tracker Muon* (inside-out).

419 *Global Muon reconstruction (outside-in)*: for each standalone-muon track, a matching tracker  
 420 track is found by comparing parameters of the two tracks propagated onto a com-  
 421 mon surface, and a global-muon track is fitted combining hits from the tracker track  
 422 and standalone-muon track, using the Kalman-filter technique [28]. At large trans-  
 423 verse momenta,  $p_T > 200 \text{ GeV}$ , the global-muon fit can improve the momentum  
 424 resolution compared to the tracker-only fit.

425 *Tracker Muon reconstruction (inside-out)*: in this approach, all tracker tracks with  $p_T > 0.5 \text{ GeV}$   
 426 and the total momentum  $p > 2.5 \text{ GeV}$  are considered as possible muon candidates  
 427 and are extrapolated to the muon system taking into account the magnetic field, the  
 428 average expected energy losses, and multiple scattering in the detector material. If  
 429 at least one muon segment (i.e., a short track stub made of DT or CSC hits) matches  
 430 the extrapolated track, the corresponding tracker track qualifies as a Tracker Muon.

431 Tracker Muon reconstruction is more efficient than the Global Muon reconstruction at low  
 432 momenta,  $p_T \lesssim 5 \text{ GeV}$ , because it requires only a single muon segment in the muon system,  
 433 whereas Global Muon reconstruction is designed to have high efficiency for muons penetrat-  
 434 ing through more than one muon station and typically requires segments in at least two muon  
 435 stations. Thanks to the high tracker-track efficiency and a very high efficiency of reconstruct-  
 436 ing segments in the muon system, about 99% of muons produced in  $pp$  collisions and having  
 437 sufficiently high momentum are reconstructed either as a Global Muon or a Tracker Muon, and  
 438 very often as both. Muons reconstructed only as standalone-muon tracks have worse momen-  
 439 tum resolution and less favorable collision muon to cosmic-ray muon ratio than the Global and  
 440 Tracker Muons and are usually not used in physics analyses.

441 Muons are usually based on the *Particle Flow Muon* selection, considering Global Muon or a  
 442 Tracker Muon candidates and by applying minimal requirements on the track components  
 443 in the muon system and taking into account a matching with small energy deposits in the  
 444 calorimeters. However, in the boosted  $Z \rightarrow \mu\mu$  regimes, muons have similar problems to  
 445 electrons. The Global muon reconstruction suffers a drop in efficiency as the  $\Delta R$  between the  
 446 muon decreases. This is a consequence of the seeding algorithm, which includes in the seed  
 447 some segments of the other muon, and after the final muon trajectory builder, a cleaning is  
 448 applied based on the number of segments and the  $\chi^2$  of the muon track. After the cleaning,  
 449 only one muon track is selected among the group of tracks that share segments, and this effect  
 450 avoids the reconstruction of the other muon (Figure 19).

451 The reconstruction of muons very close in  $\Delta R$  is also a problem for the PF algorithm, which  
 452 is based on the hypothesis that the muon is a minimum-ionizing particle. Another high- $p_T$   
 453 track close to a reconstructed muon can also fail to pass the PF identification of the nearby  
 454 muon, further lowering the efficiency at small angles, as shown in Figure 19. The adopted  
 455 compromise between efficiency and fake-muons rejection is then to require that *at least* one  
 456 of the two muons (not necessarily the highest- $p_T$ ) fulfills both the PF and the HighPt, while  
 457 the other has to satisfy a looser selection. This asymmetric selection ensures a high efficiency  
 458 ( $\gtrsim 95\%$ ) in the whole  $p_T$ ,  $\eta$  and  $\Delta R$  ranges (Figure 19). The HighPt Id is a set of cuts specifically  
 459 designed for high-momentum muons from the Muon POG [29]:

- 460 • to be reconstructed also as Global muon
- 461 • at least one muon chamber hit included in the global-muon track fit
- 462 • muon segments in at least two muon stations
- 463 • the track used to obtain the muon momentum needs to pass  $\delta p_T / p_T < 0.3$
- 464 • tracker track transverse impact parameter  $d_{xy} < 2\text{mm}$  w.r.t. the primary vertex
- 465 • longitudinal impact parameter  $d_z < 5\text{mm}$  w.r.t. the primary vertex
- 466 • number of pixel hits  $> 0$
- 467 • number of tracker layers with hits  $> 5$ .

468 The other muons in the event, if present, should be identified with a tracker-only selection.  
 469 A tight CustomTracker Id is defined using the same quality cuts on the muon track as the  
 470 HighPt, but dropping the requirements that force the muon to be Global or PF:

- 471 • to be reconstructed as a standard Tracker muon (was Global muon in HightPt Id)
- 472 • at least one muon chamber hit included in the global-muon track fit
- 473 • muon segments in at least two muon stations
- 474 • the track used to obtain the muon momentum needs to pass  $\delta p_T / p_T < 0.3$ .
- 475 • tracker track transverse impact parameter  $d_{xy} < 2\text{mm}$  w.r.t. the primary vertex
- 476 • longitudinal impact parameter  $d_z < 5\text{mm}$  w.r.t. the primary vertex
- 477 • number of pixel hits  $> 0$
- 478 • number of tracker layers with hits  $> 5$ .

479 On the other hand, the single-muon channel does not share the same problem. In this case, the  
 480 only muon reconstructed has to pass the HighPt Id, and a veto is applied to other muons that  
 481 fulfill the Tracker Id. Since the PF event description is required, the leading HighPt muon is  
 482 also required to be reconstructed by the PF algorithm.

483 In the zero-lepton channel, muons with  $p_T$  as low as 10 GeV are vetoed. In order to have an  
 484 effective veto, a standard *loose* id is applied in this case:

- 485 • to be reconstructed with the PF algorithm
- 486 • to be or Tracker or Global

For muons reconstructed using the PF algorithm, the standard muon isolation is defined as the ratio of the  $p_T$  sum of all charged and neutral particle-flow candidates in the event within a cone with a radius of  $\Delta R = 0.4$  centered along the lepton direction. Corrections in order to reduce the PU contamination are also applied, using the  $\Delta\beta$  method. Charged candidates falling into the cone that are not compatible with the primary vertex are removed from the sum. Additionally, the neutral contribution from PU is estimated to be half the one coming from charged candidates, and this quantity is also subtracted from the total. Eventually, the scalar sum is divided by the lepton  $p_T$  itself. The general formula for the standard *particle-flow* isolation is then:

$$I_{rel} = \left[ \sum p_T^{\text{ch had}} + \max(\sum p_T^{\text{neu had}} + \sum p_T^\gamma - 0.5 \cdot \sum p_T^{\text{pu ch had}}, 0) \right] / p_T^\ell$$

487 where  $\sum p_T^{\text{ch had}}$  is the sum of the transverse momenta of the charged hadrons,  $\sum p_T^{\text{neu had}}$  is  
 488 the sum of transverse energies of the neutral hadrons,  $p_T^\gamma$  is the sum of the transverse energy  
 489 of particle flow photons and  $\sum p_T^{\text{pu ch had}}$  is the sum of transverse momenta of the charged  
 490 particles in the cone of interest but with particles not originating from the primary vertex (for

491 pileup corrections).

492 In the case of muons not reconstructed using the PF algorithm, as in the case of the `CustomTracker`  
 493 id used to select muons in the 2-muons category, the MuonPOG provides a Tracker-based def-  
 494 inition of isolation. This isolation is computed as the scalar sum of the  $p_T$  of all the track from  
 495 the leading PV in the event within a cone with a radius of  $\Delta R = 0.3$  centered along the muon  
 496 direction, normalized by the muon  $p_T$ .

497 In the present analysis muons are identified with the identification method: *loose* and *HighPt*  
 498 standard Id defined by the MuonPOG, and the modified version of the traker Id (`CustomTracker`)  
 499 previously discussed in this section. Muons with  $p_T > 10$  GeV identified with the *loose* id and  
 500 passing the *loose* working point of the PFIsolation ( $< 0.25$ ) are vetoed in the 0-lepton category.

501 In the 1-muon category the *HighPt* Id is used to identify the leading muon, which is also  
 502 required to pass the tracker isolation cut ( $< 0.15$ ).

503 Finally, in the 2-muon category at least one muons is required to be identified with the *HighPt*  
 504 standard id, while the other is selected if passes the `CustomTracker` requirements. Given the  
 505 possibility of having one muon not identified with the PF algorithm, the *loose* working point  
 506 ( $< 0.1$ ) of the tracker-based isolation criterion is used to select the leptons. When evaluating  
 507 the tracker-isolation for each muon, if the other selected muon falls in the  $\Delta R < 0.3$  isolation  
 508 cone, the tracker-based isolation is recomputed after removing the inner tracker track  $p_T$  of the  
 509 other muon.

510 Scale factors for trigger, muon identification and isolation, are centrally provided as a function  
 511 of the muon  $p_T$  and  $\eta$  by the Muon POG [20], and applied consistently in the analysis.

512 The muon selection efficiency in the signal is defined in the same way as the electrons. Fig-  
 513 ure 19 show the muon selection efficiency as a function of the different Id selections applied  
 514 as a function of the distance between the two muons. Figure 21 show the data/simulation  
 515 comparison for the muons after the preselections.

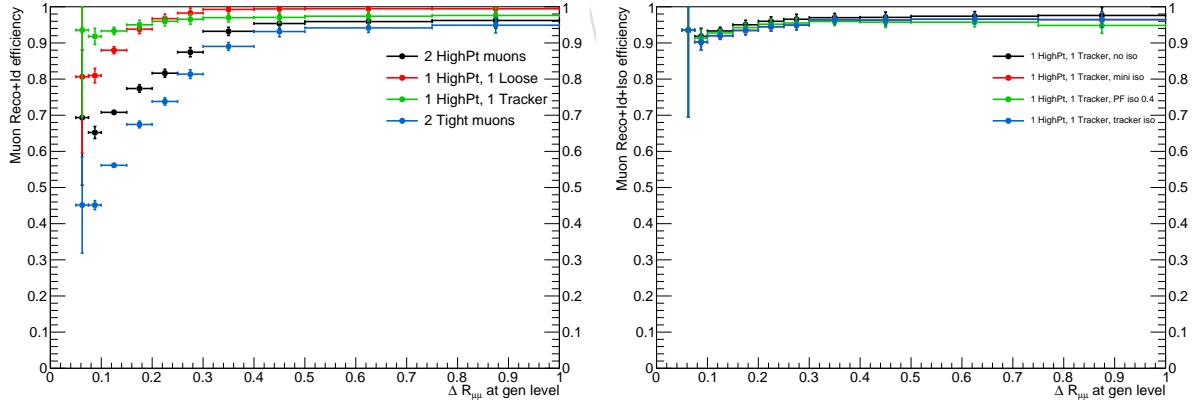


Figure 19: Identification (left) and combined identification and isolation (right) efficiency as a function of the  $\Delta R$  at generation level, after matching and  $p_T$  selections, when muons are re-  
 quired to pass different identification and isolation requirements. Events from all the generated  
 mass points are considered together.

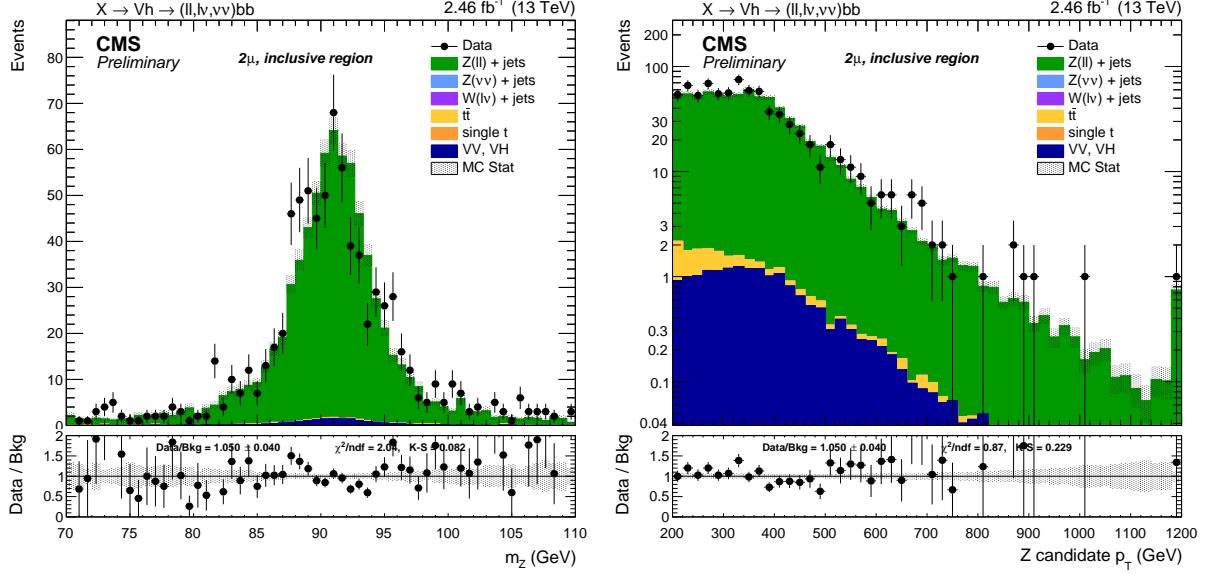


Figure 20: Reconstructed Z mass (left) and Z  $p_T$  (right) after the inclusive  $Z \rightarrow \mu\mu$  selections.

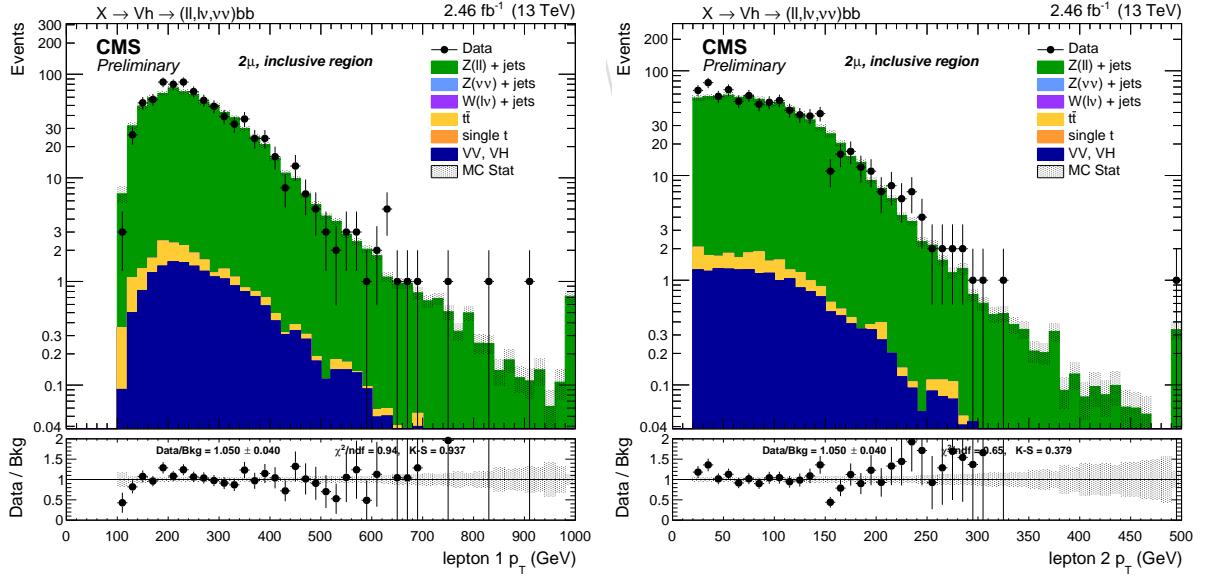


Figure 21: Leading (left) and sub-leading (right) muon  $p_T$  spectra after Z selections.

516 **3.4 Mini-isolation studies**

In boosted topologies, leptons are closer to each other than the usual cone size ( $\Delta R = 0.3$  or  $0.4$ ), and the second lepton can be included in the isolation sum, spoiling the final value. The particle-flow algorithm has been developed to avoid this problem for muons, while for electrons the energy deposits of another nearby electron are still not removed. The solution to this problem has been studied by the SUSY group, developing a modified isolation algorithm similar to the particle-flow but with a shrinking cone whose size depends from the lepton  $p_T$ , as shown in Figure 22. The cone size is set to be:

$$R = \begin{cases} 0.2 & \text{if } p_T < 50 \text{ GeV} \\ 10 \text{ GeV}/p_T & \text{if } 50 \leq p_T \leq 200 \text{ GeV} \\ 0.05 & \text{if } p_T > 200 \text{ GeV} \end{cases}$$

517  $R$  has been chosen to be as small as possible to reduce overlaps with nearby high- $p_T$  particles,  
 518 and to be large enough to contain  $b$ -decay products. Two reference working points have also  
 519 been proposed for single-lepton and multi-lepton topologies, the latter being  $I_{mini} < 0.2$  for  
 520 muons and  $I_{mini} < 0.1$  for electrons.

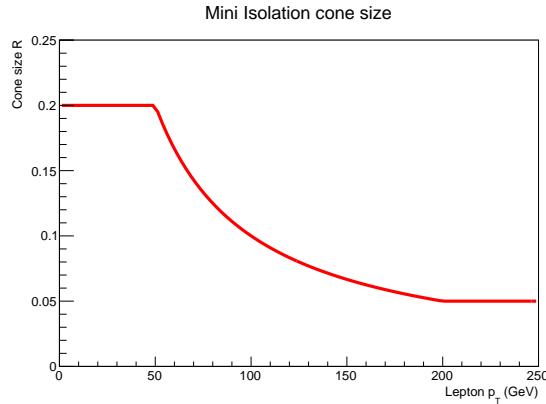


Figure 22: Mini Isolation cone as a function of the lepton  $p_T$ .

521 Figure 23 and Figure 24 show the combined identification and isolation efficiency with the two  
 522 different algorithms. For the particle-flow isolation, a loose cut is applied to both the leptons,  
 523 while for mini-isolation the cut is set to the suggested value of 0.1. Despite the fact that the  
 524 particle-flow isolation cut being Loose, the isolation efficiency is slightly lower than the one  
 525 obtained after the application of the mini-isolation. The latter has also a better behavior when  
 526 the leptons are very close to each other ( $\Delta R_{\ell\ell}$ ), due to the fact that the isolation cone at high  
 527 lepton  $p_T$  is very narrow.

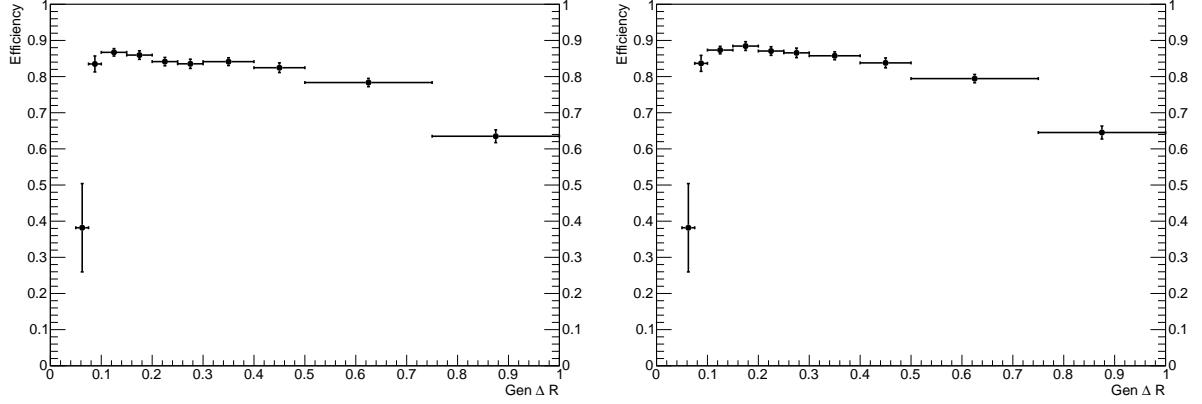


Figure 23: HEEPNoIso electron identification and isolation efficiency as a function of the  $\Delta R$  between the two electrons at generation level, after matching and  $p_T$  selections. The two considered isolation algorithms are compared: particle-flow isolation with a fixed cone  $R = 0.3$  and  $I_{rel} < 0.15$  (left), and mini-isolation with shrinking cone and  $I_{mini} < 0.1$  (right).

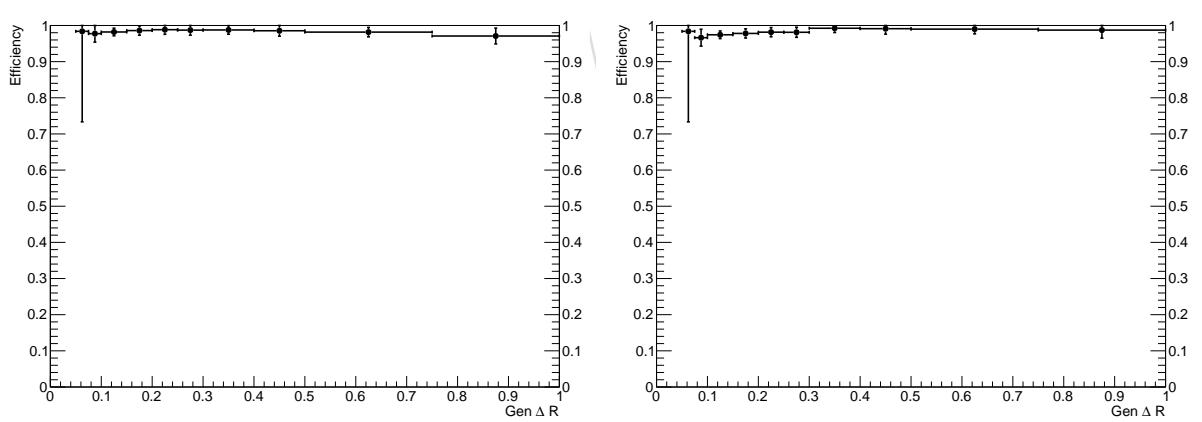


Figure 24: Muon identification and isolation efficiency for a HighPt and a Tracker muon as a function of the  $\Delta R$  between the two muons at generation level, after matching and  $p_T$  selections. The two considered isolation algorithms are compared: particle-flow isolation with a fixed cone  $R = 0.4$  and  $I_{rel} < 0.2$  (left), and mini-isolation with shrinking cone and  $I_{mini} < 0.1$  (right).

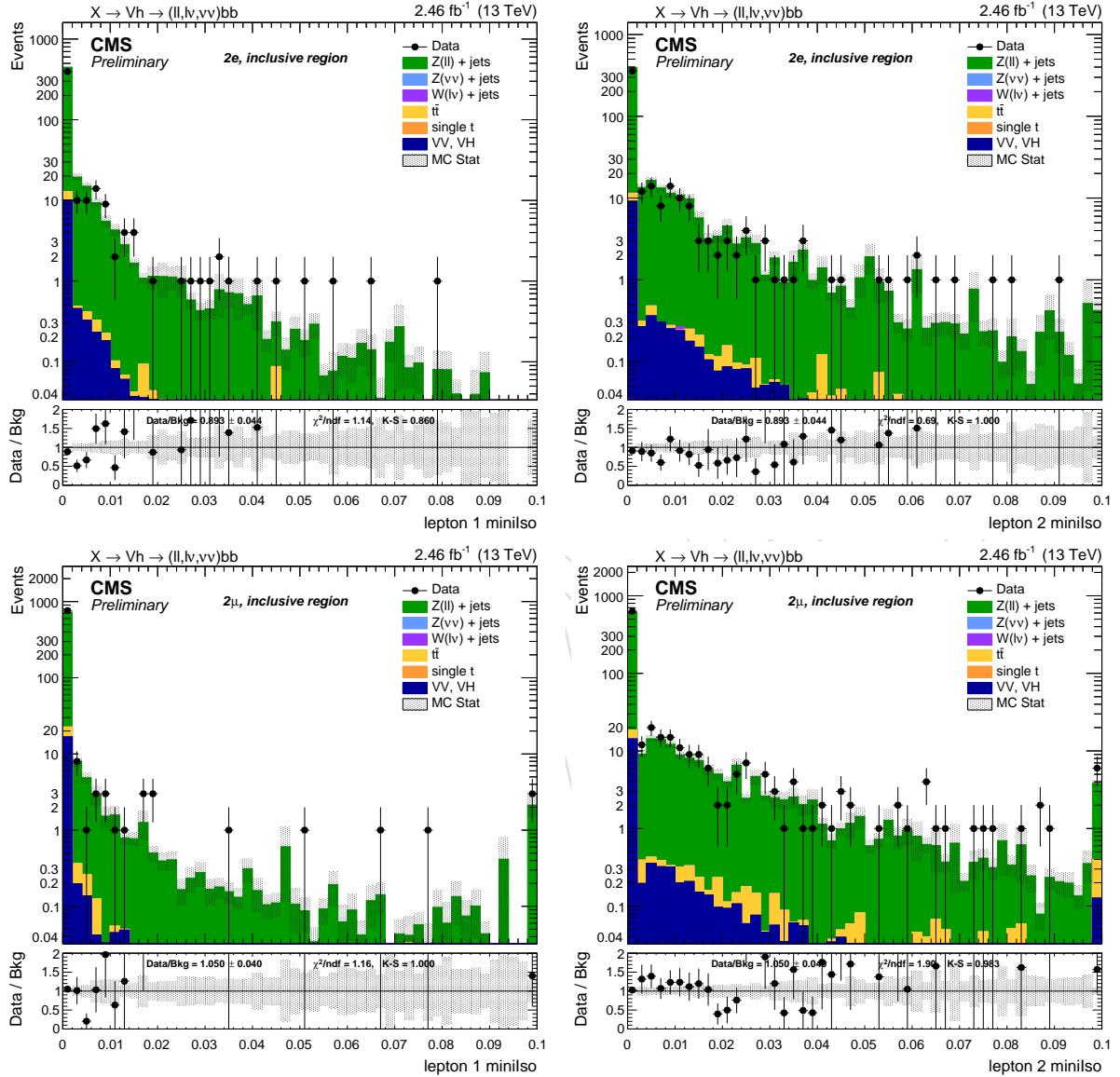


Figure 25: Mini-isolation distribution for the leading (left) and sub-leading (right) electrons (top) or muons (bottom).

528 **3.5 Taus**

529 The presence of hadronically-decaying taus only act as veto for the events both in the signal  
 530 and in the control regions to suppress electroweak backgrounds. The selection criteria for taus  
 531 are  $p_T > 18$  GeV and  $|\eta| < 2.3$ . The Run2 TauPOG recommended identification criteria [30]  
 532 (`decayModeFinding, byLooseCombinedIsolationDeltaBetaCorr3Hits`) are required  
 533 and applied in order to identify possible tau candidates.

534 **3.6 Photons**

535 As in the case of tau leptons, a photon veto is applied in the analysis both for the signal and  
 536 the control regions. Events are rejected if they contain one (or more) photon with  $p_T > 15$   
 537 GeV,  $|\eta| < 2.5$ , passing the Loose cut-based photon ID. The Loose photon Id is applied  
 538 as in the EGamma POG recommendations for Run2 analyses [31] (tuned on Spring15 25 ns  
 539 samples). The isolation cuts (using the rho-area method for the mitigation of the pileup) and  
 540 conversion safe electron veto are applied. The isolation value is computed in a  $\Delta R$  cone of  
 541 0.3 and is corrected for pileup by subtracting the event-by-event energy density ( $\rho$ ) times an  
 542 effective area. The applied cut-based definition of the Loose photon Id is reported in Table 8.

Table 8: Photon cut-based Id for Spring15 25ns conditions. EB: barrel cuts ( $|\eta_{\text{supercluster}}| \leq 1.479$ ); EE: endcap cuts ( $|\eta_{\text{supercluster}}| > 1.479$ )

Photons	Loose	
	EB	EE
$H/E$	< 0.05	0.05
$\sigma_{i\eta i\eta}$	< 0.0102	0.0274
PF ch.had.iso.( $\rho$ -corr)	< 3.32	1.97
PF neu.had.iso.( $\rho$ -corr)	< $1.92 + 0.014p_T + 0.000019p_T^2$	$11.86 + 0.0139p_T + 0.000025p_T^2$
PF photon iso.( $\rho$ -corr)	< $0.81 + 0.0053p_T$	$0.83 + 0.0034p_T$
conversion veto	yes	yes

### 543 3.7 Jets

544 Events in the CMS detector are reconstructed using the the particle-flow algorithm [32, 33],  
 545 which combines information from all sub-detectors in order to reconstruct stable particles  
 546 (muons, electrons, photons, neutral and charged hadrons). The charged hadron subtraction  
 547 algorithm (CHS) removes candidates not associated to the primary vertex in order to remove  
 548 contributions from pileup [34]. The remaining particles are used as input to jet clustering algo-  
 549 rithms to reconstruct particle-flow jets. The jets are clustered using the FASTJET package [35]  
 550 with the anti- $k_T$  jet clustering algorithm [36] with a clustering parameter of  $R = 0.8$  (“fat”-jets  
 551 or AK8 jets) or  $R = 0.4$  (“standard”-jets or AK4 jets). In order to avoid double-counting of PF  
 552 candidates, AK4 jets are considered only if the angular separation from the leading AK8 jet is  
 553 larger than  $\Delta R > 0.8$ . Several levels of jet energy corrections are applied to the momentum of  
 554 the clustered (raw) jets in order to obtain the energy value that is closer to the true energy of  
 555 the initial parton [37]:

556 *L1 Offset*: the pileup and electronic noise effects are removed. This correction can be esti-  
 557 mated using events collected by a random trigger, without any preconditions except  
 558 a beam crossing, referred as *zero bias* events. The offset contribution from pileup  
 559 is estimated by the FastJet method which relies on the definition of a jet area [35]  
 560 from which a median energy density ( $\rho$ , in GeV/Area) per event can be defined.  
 561 The correction subtracted to the jet  $p_T$  equals to  $\rho$  times the jet area. FastJet has the  
 562 advantage of being able to remove the out-of-time pileup component, but has the  
 563 disadvantage of subtracting the underlying event contribution as well.

564 *L2 Relative ( $\eta$ )*: the variation in jet response with  $\eta$  is flattened. The unbalance between  
 565 the jets transverse momentum that is observed on average, is due to the variation of  
 566 the jet response across the detector versus  $\eta$ .

567 *L3 Absolute ( $p_T$ )*: the calorimetric energy response varies as a function of the jet  $p_T$ . The  
 568 absolute correction removes these variations and makes the response equal to unity.  
 569 This correction is obtained from simulation using the Monte Carlo truth information.

570 *L2L3 Residual*: differences between data and simulation after L2 and L3 corrections are  
 571 removed by applying a specific calibration to data events. Residual corrections are  
 572 extracted from data using the transverse momentum balance in  $\gamma$ +jets and Z +jets  
 573 events [37].

574 The latest jet energy corrections are applied to AK4 and AK8 CHS jets, and the tags are Summer15\_25nsV6\_DA  
 575 and Summer15\_25nsV5\_MC for data and simulation, respectively.

576 In this analysis, jets are considered if the corrected  $p_T$  is larger than 30 GeV for AK4 jets and  
 577 200 GeV for AK8 jets, and lie in the tracker acceptance ( $|\eta| < 2.4$ ). Additionally, they are re-  
 578 quired to pass *loose* jet identification requirements defined by the JETMET POG for Run2 anal-  
 579 yses [38], listed in Table 9.

580 In the 0-leptons channel, a more robust jet selection is applied, requiring jets to pass the *tight*  
 581 jet identification criterion, also listed in Table 9.

582 Since it has been measured that the jet energy resolution (JER) is not the same in data and  
 583 MC, an additional smearing is applied in simulation, in order to get a better agreement. This  
 584 procedure is suggested by the JETMET POG [39], and the smearing coefficients and their errors  
 585 are reported in Tab. 10 for 2015 data.

PF Jet ID	loose	tight
Neutral Hadron Fraction	< 0.99	< 0.90
Neutral EM Fraction	< 0.99	< 0.90
Number of Constituents	> 1	> 1
Muon Fraction	-	-
Additionally, for $ \eta  < 2.4$		
Charged Hadron Fraction	> 0	> 0
Charged Multiplicity	> 0	> 0
Charged EM Fraction	< 0.99	< 0.99

Table 9: *Loose* and *Tight* jet identification requirements for Run2 (Spring15) 25ns conditions.

Jet $\eta$	SF
0.0 – 0.8	$1.061 \pm 0.023$
0.8 – 1.3	$1.088 \pm 0.029$
1.3 – 1.9	$1.106 \pm 0.030$
1.9 – 2.5	$1.126 \pm 0.094$
2.5 – 3.0	$1.343 \pm 0.123$
3.0 – 3.2	$1.303 \pm 0.111$
3.2 – 5.0	$1.320 \pm 0.286$

Table 10: Smearing coefficients and JER uncertainties

586 

### 3.7.1 Jet mass

587 The jet mass is the main observable in distinguishing a H-jet from a QCD jet. Jet grooming  
 588 consists in the suppression of uncorrelated UE/PU (underlying event and pile-up) radiation  
 589 from the target jet and improves the discrimination pushing the jet mass for QCD jets towards  
 590 lower values while maintaining the jet mass for V(H)-jets around the boson-mass.

591 Three different grooming algorithms were originally considered:

592 **Trimming:** Trimming is a technique that ignores regions within a jet that falls below a  
 593 minimum  $p_T$  threshold. Trimming reclusters the jet's constituents with a radius  $R_{sub}$   
 594 and then accepts only the subjets that have  $p_{T,sub} > f_{cut}$ , where  $f_{cut}$  is typically taken  
 595 proportional to  $H_T$ , the scalar sum of the  $p_T$  of all jet reconstructed in the event.

596 **Filtering:** This procedure provides a hierarchical structure for the clustering like the  $k_T$   
 597 algorithm, but in angles rather than in relative transverse momenta. It creates a se-  
 598 ries of n new subjects  $s_1, s_2, \dots, s_n$  ordered in descending  $p_T$ . The final jet is redefined  
 599 as the sum of the four-momenta of the three highest  $p_T$  subjets  $\sum_i^{\min(n,3)} s_i$ .

Pruning: The idea is to take a jet of interest and then to recluster it using a vetoed se-  
 quential clustering algorithm. Clustering is performed with the CA algorithm, but  
 particles are discarded if they are too far away in  $\Delta R$ :

$$\Delta R_{ij} = D_{cut} \alpha \frac{m_j}{p_{T,j}}$$

and the energy sharing is too asymmetric:

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

600 where  $z_{cut}$  and  $\alpha$  are parameters of the algorithm. If both these conditions are satis-  
 601 fied the softer of the two particles is not considered.

**Soft-drop:** The “soft drop declustering” is new jet substructure technique which recursively removes soft wide-angle radiation from a jet [40]. The soft drop algorithm depends on two parameters: a soft threshold  $z$  cut and an angular exponent  $\beta$ . Like any grooming method, soft drop declustering removes wide-angle soft radiation from a jet in order to mitigate the effects of contamination from initial state radiation (ISR), underlying event (UE), and multiple hadron scattering (pileup). Given a jet of radius  $R_0$  with only two constituents, the soft drop procedure removes the softer constituent unless:

$$\frac{\min(p_T^1, p_T^2)}{p_T^1 + p_T^2} > z_{cut} \left( \frac{\Delta R_{12}}{R_0} \right)^\beta$$

By construction, this condition fail for wide-angle soft radiation. The degree of jet grooming is controlled by  $z_{cut}$  and  $\beta$ , with  $\beta \rightarrow \infty$  returning back an ungroomed jet. The  $\beta = 0$  limit of the energy loss is particularly interesting, since it is largely insensitive to the value of the strong coupling constant. The default parameters used by CMS are  $\beta = 0$  and  $z_{cut} = 1$ .

In general, the filtering algorithm is the least aggressive grooming technique, with groomed jet masses close to the original case. The trimming algorithm is moderately aggressive and produces a much wider final mass distribution. Pruning is the most aggressive technique and a bimodal distribution begins to appear: in cases where the pruned jet mass is small, jets usually have most of their energy configured in core components with little gluon radiation, which leads to narrow jets. Instead, when the pruned jet mass is large, the jets are split more symmetrically.

### 3.7.2 Jet substructure

In order to further discriminate signal from background, it is useful to investigate the inner structure of the jet. Studying the distribution of the jet constituents with respect to the jet axis allows us to test the hypothesis of the existence of multiple substructures, that could be evidence of jets originated by more than one parton. This procedure proceeds as follows: the constituents of the jet are clustered again with the usual algorithm, however the procedure is stopped when one obtains  $N$  subjets. Then, a new variable, the  $N$ -subjettiness, is introduced. It is defined as:

$$\tau_N = \frac{1}{d_0} \sum_k p_{T,k} \min(\Delta R_{1,k}^\beta, \Delta R_{2,k}^\beta, \dots, \Delta R_{N,k}^\beta)$$

where  $\beta$  is an arbitrary parameter, the index  $k$  runs over the jet constituents and the distances  $\Delta R_{N,k}$  are calculated with respect to the axis of the  $N$ -th subjet.

The normalization factor  $d_0$  is calculated as  $d_0 = \sum_k p_{T,k} R_0^\beta$ , setting  $R_0$  to the radius of the original jet. The  $N$ -subjettiness is always included in the interval from 0 to 1 and represents the compatibility of the jet structure with an  $N$ -subjet hypothesis: small values correspond to high compatibility. Indeed,  $\tau_N$  weights the transverse momentum of the jet constituents by their angular distance to the closest subjet. In this analysis the  $N$ -subjettiness is calculated from the ungroomed jet with the parameter  $\beta = 1$ . The subjettiness related to the one and two subjet hypothesis is thus:

$$\tau_1 = \frac{1}{d_0} \sum_k p_{T,k} \Delta R_{1,k}$$

and

$$\tau_2 = \frac{1}{d_0} \sum_k p_{T,k} \min(\Delta R_{1,k}, \Delta R_{2,k})$$

617 In principle, these two quantities should allow us to distinguish the dipole-like nature of the  
 618 showering of the Higgs decay from the classic monopole structure of QCD jets. In particular,  
 619 the variable that best discriminates between H-jets and QCD jets is the ratio of 2-subjettiness  
 620 and 1-subjettiness,  $\tau_{21} = \tau_2 / \tau_1$ .

### 621 3.8 b-tagging

622 The presence of a pair of b-quarks from the  $h \rightarrow b\bar{b}$  decay is a very distinctive signature that  
 623 permits a strong discrimination against all the backgrounds that involve jets with light fla-  
 624 vors. The only background which cannot be reduced with this technique is the Z production  
 625 in association with one or two b-quarks, and the  $Z \rightarrow b\bar{b}$  decay which is also topologically  
 626 similar to the signal. The latter can be only reduced by applying a jet mass cut, as described in  
 627 Section 3.7.1.

628 Tagging b-jets in boosted topologies has been undergone several modifications with respect  
 629 to Run-I taggers. One of the most important differences is that the tagger algorithms are not  
 630 directly based on tracks. B-tagging algorithms are applied to both the fat-jet and the sub-jets,  
 631 independently. For subjets, run-II taggers are by default applied on the same charged particle-  
 632 flow candidate list that is used in the jet clustering (*explicit jet-to-track association*). Thanks to  
 633 the explicit jet-to-track association, the two sub-jets do not share any PF-constituent, avoiding  
 634 unintended correlations.

635 Several algorithms have been developed to tag jets from b-quarks. The recommended and best-  
 636 performing algorithm, used throughout this analysis, is the `pfCombinedInclusiveSecondaryVertexV2B`  
 637 often shortened to *combined secondary vertex* (CSV). This algorithm involves the use of secondary  
 638 vertices, together with other lifetime information, like the IP significance or decay lengths. Sec-  
 639 ondary vertices are reconstructed with the inclusive vertex finder algorithm, that does not re-  
 640 quire jets (and thus is independent on the jet size) and uses all tracks to reconstruct secondary  
 641 vertices [41]. In order to provide discrimination even when no secondary vertices are found, so  
 642 the maximum possible b-tagging efficiency is not limited by the secondary vertex reconstruc-  
 643 tion efficiency (50 ~ 60%). In many cases, tracks with an IP significance > 2 can be combined in  
 644 a so-called pseudo vertex, allowing for the computation of a subset of secondary vertex based  
 645 quantities even without an actual vertex fit. When even this is not possible, a no vertex cate-  
 646 gory reverts simply to track based variables similarly to the jet probability algorithm. The list  
 647 of variables feeded as input to an Artificial Neural Network is:

- 648 • the vertex category (real, pseudo, or no vertex)
- 649 • 2D flight distance significance
- 650 • vertex mass
- 651 • number of tracks at the vertex
- 652 • ratio of the energy carried by tracks at the vertex with respect to all tracks in the jet
- 653 • the pseudo-rapidity of the tracks at the vertex with respect to the jet axis
- 654 • 2D IP significance of the first track that raises the invariant mass above the charm
- 655 threshold of 1.5 GeV when subsequently summing up tracks ordered by decreasing
- 656 IP significance
- 657 • 3D signed IP significances for all tracks in the jet
- 658 • number of tracks in the jet
- 659 •  $\Delta R$  between the secondary vertex flight direction and the jet axis
- 660 • number of secondary vertices associated to the jet or sub-jet

661 The jet or sub-jet is considered as tagged if the discriminator value is above some threshold  
 662 value, often referred to as the cut value, and the efficiency is defined as the number of jets  
 663 which have a discriminator value that is above that cut divided by the total number of jets  
 664 (of the same flavor). The typical b-tagging efficiency is between 40% and 70% while keeping  
 665 the rate of mis-identified light-flavor jets between 0.1% and 10%. Three working points are  
 666 usually defined for each algorithm, defining cuts in the discriminators based on the level of  
 667 mis-tagging. The cut values and the corresponding mis-tagging for light-flavor jets relative to  
 668 the CSV algorithm are reported in Table 11.

Working point	Cut	$\varepsilon_{light}$
Loose	0.605	$\sim 10\%$
Medium	0.890	$\sim 1\%$
Tight	0.970	$\sim 0.1\%$

Table 11: CSV official working points.

669 It is known that b-tagging efficiency is not the same in data and MC. In order to take into  
 670 account this shortcoming, the BTV POG provides collections of b-tagging scale factors for b-jets  
 671 and mistagged light jets, measured for different physics processes, for the supported tagging  
 672 algorithms and the three standard working points [42]. Usually, a weight is calculated on a per-  
 673 event basis as a function of the b-tagging status of the jets in the event [43]; this is a simple and  
 674 effective method if there are a small number of possible combinations of jets and b-tagged jets.  
 675 Unfortunately, this is not the case of the present analysis. Other techniques allow to take into  
 676 account the scale factors provided by BTV by *reshaping* the discriminator output; this method  
 677 has been already successfully applied in the SM VH analysis [44], and is also described in  
 678 detail in Ref. [15, 45]. The reshaping method does not sensibly increase the b-tagging scale  
 679 factors uncertainty with respect to using single operating point SFs or other techniques used to  
 680 apply the same scale factors.  
 681 The CSV discriminator output has been reshaped in the simulation taking into account the  
 682 official SF provided by the POG. The procedure is tested in a  $t\bar{t}$ -enriched sample, obtained by  
 683 requiring one tight and isolated muon and a tight electron and at least one jet above 30 GeV.  
 684 The original and reshaped CSV distribution of the leading and sub-leading jet in the event are  
 685 reported in Fig. 26 and Fig. 27.

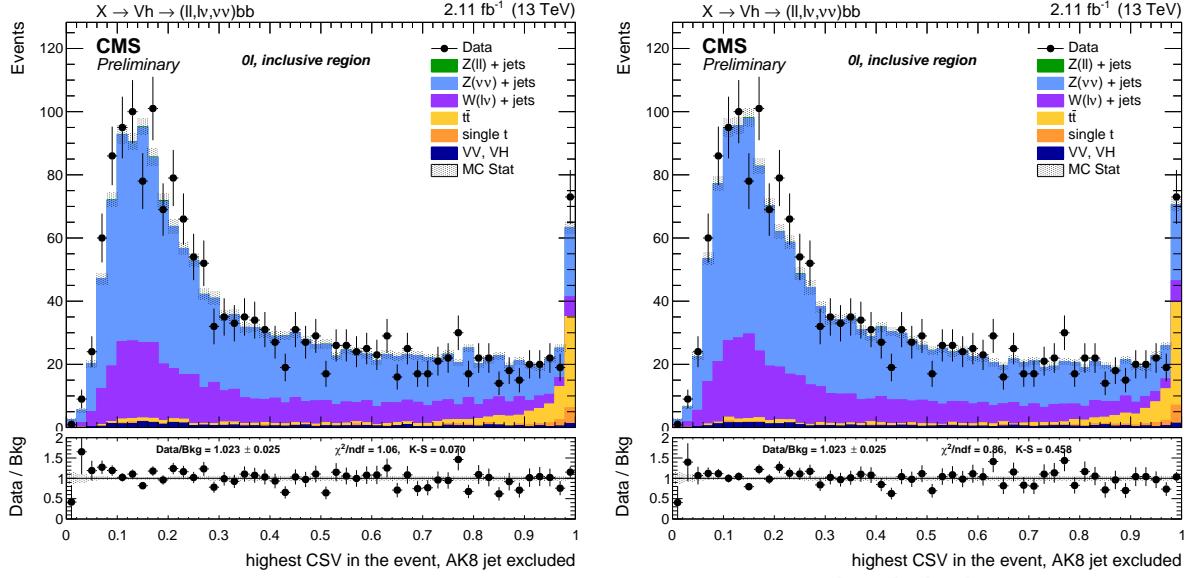


Figure 26: Combined Secondary Vertex discriminator before (left) and after (right) the reshaping procedure applied to the highest-CSV AK4 jet in the event in the 0-lepton channel.

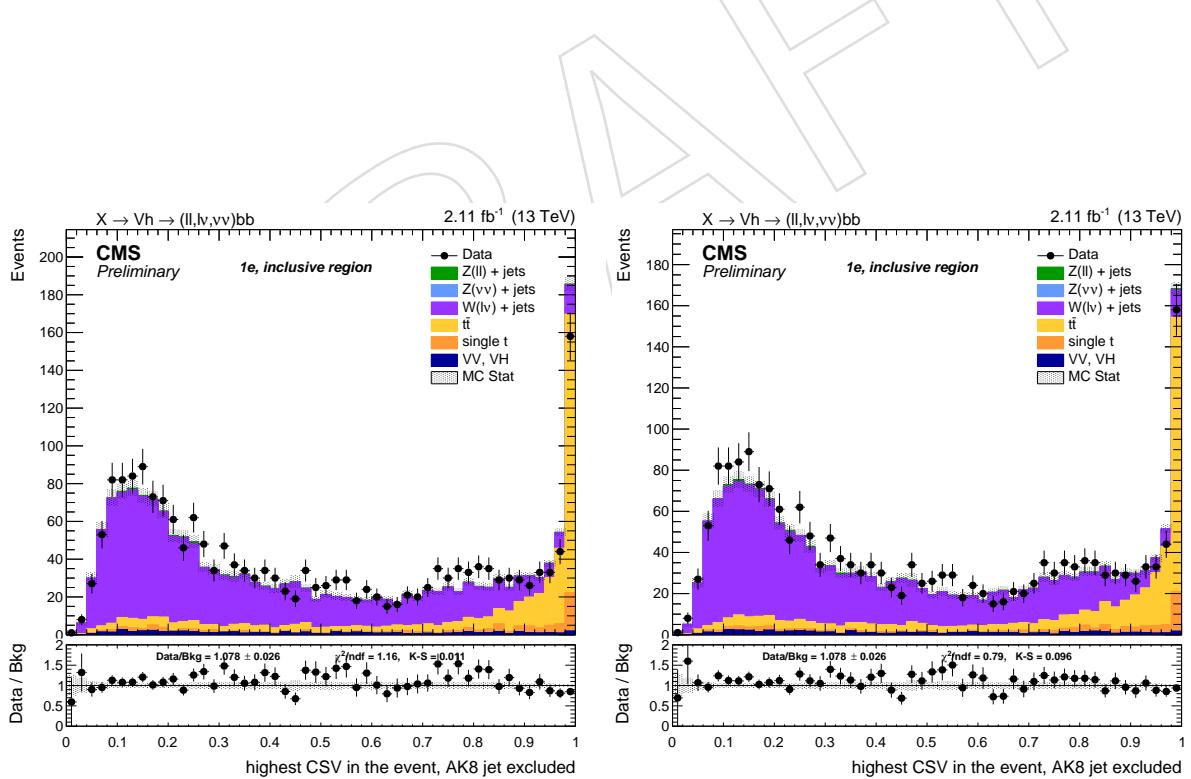


Figure 27: Combined Secondary Vertex discriminator before (left) and after (right) the reshaping procedure applied to the highest-CSV AK4 jet in the event in the 1-lepton channel.

### 686 3.9 Missing Energy

687 The  $E_T^{\text{miss}}$  is the imbalance in the transverse energy of all visible particles, and it is reconstructed  
 688 with the particle flow algorithm [32]. The raw  $E_T^{\text{miss}}$  is defined as the inverse vectorial sum of  
 689 the transverse momentum of all the reconstructed charged and neutral particle flow candi-  
 690 dates:  $\vec{E}_T^{\text{miss}} = -\sum_{i=0}^{\text{all}} \vec{p}_{T,i}$ . The raw  $E_T^{\text{miss}}$  is systematically different from true  $E_T^{\text{miss}}$ , for many  
 691 reasons including the non-compensating nature of the calorimeters and detector misalignment.  
 692 To better estimate the true  $E_T^{\text{miss}}$ , corrections can be applied:

693 Type-0: a mitigation for the degradation of the  $E_T^{\text{miss}}$  reconstruction due to the pileup in-  
 694 teractions, by applying the CHS algorithm. However, the  $E_T^{\text{miss}}$  contribution from  
 695 pileup neutral particles cannot be easily subtracted; the assumption is that the  $E_T^{\text{miss}}$   
 696 contribution term of charged and neutral pileup particles are the same, and cancel-  
 697 lation at the true level is exact:  $\sum_{\text{neuPU}} \vec{p}_{T,i}^{\text{true}} + \sum_{\text{chPU}} \vec{p}_{T,i}^{\text{true}} = 0$ . An additional  $E_T^{\text{miss}}$   
 698 term is then added to the raw  $E_T^{\text{miss}}$  to take into account the neutral PU contribu-  
 699 tion, which is equal to the charged one with a multiplicative scale factor taking into  
 700 account calorimeter mismeasurements of low- $p_T$  energy deposits.

701 Type-1: propagation of the jet energy corrections (JEC) to MET. The Type-I correction re-  
 702 places the vector sum of transverse momenta of particles which can be clustered  
 703 as jets with the vector sum of the transverse momenta of the jets to which JEC is  
 704 applied.

705 Particle flow  $E_T^{\text{miss}}$  with type-1 corrections applied is currently the default one used by CMS  
 706 physics analyses; Figure 28 show the  $E_T^{\text{miss}}$  distribution for data and Monte Carlo after the  
 707 corrections.

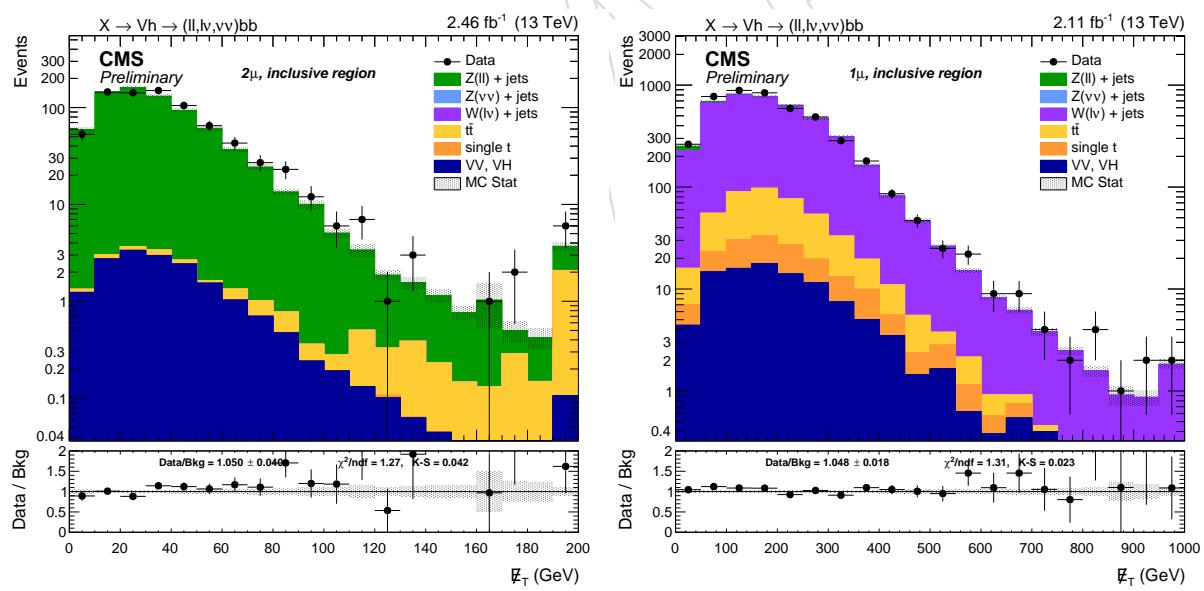


Figure 28: Type-1 corrected  $E_T^{\text{miss}}$  in  $Z \rightarrow \mu\mu$  (left) and  $W \rightarrow \ell\nu$  (right) events.

#### 708 3.9.1 Recoil corrections

709 The missing transverse momentum is known to be not perfectly described in MC, hence leading  
 710 to a non-perfect data-MC agreement. A correction can be derived from independent  $Z \rightarrow \mu\mu$   
 711 and  $\gamma$ +jets samples, as described in [46] and exploited in the mono-V search performed by CMS  
 712 with the Run1 dataset [47], as well as in other current Run2 analyses [22, 48].

The corrections are applied by smearing and shifting the recoil in MC to match the recoil in data. The hadronic recoil  $\vec{U}$ , is defined in as:

$$\vec{U} = -M\vec{ET} - \sum \vec{p}_T(\mu_i/\gamma) \quad (1)$$

713 The recoil is typically parametrized in the longitudinal and transverse component of the bo-  
 714 son  $p_T$  (taken as a vector in the transverse plane of the detector), defined here as  $U_{||}$ , and  $U_{\perp}$   
 715 respectively.

716 The distributions of  $U_{||}$  and  $U_{\perp}$  are fitted, for both data and MC, with double Gaussians with  
 717 common mean in bins of the  $p_T$  of the boson. From the fit results the mean  $\mu$  and the com-  
 718 bined width  $\sigma$  are extracted, determinig the  $U_{||}/U_{\perp}$  response and resolution, respectively. A  
 719  $p_T$ -dependent scale factor between data and MC is then computed by correcting for both the  
 720 response and resolution discrepancies. The resulting distribution of  $U_{||}$  after the application  
 721 of the smearing on the MC is shown in Figure 29, where a good agreement is found after the  
 722 corrections.

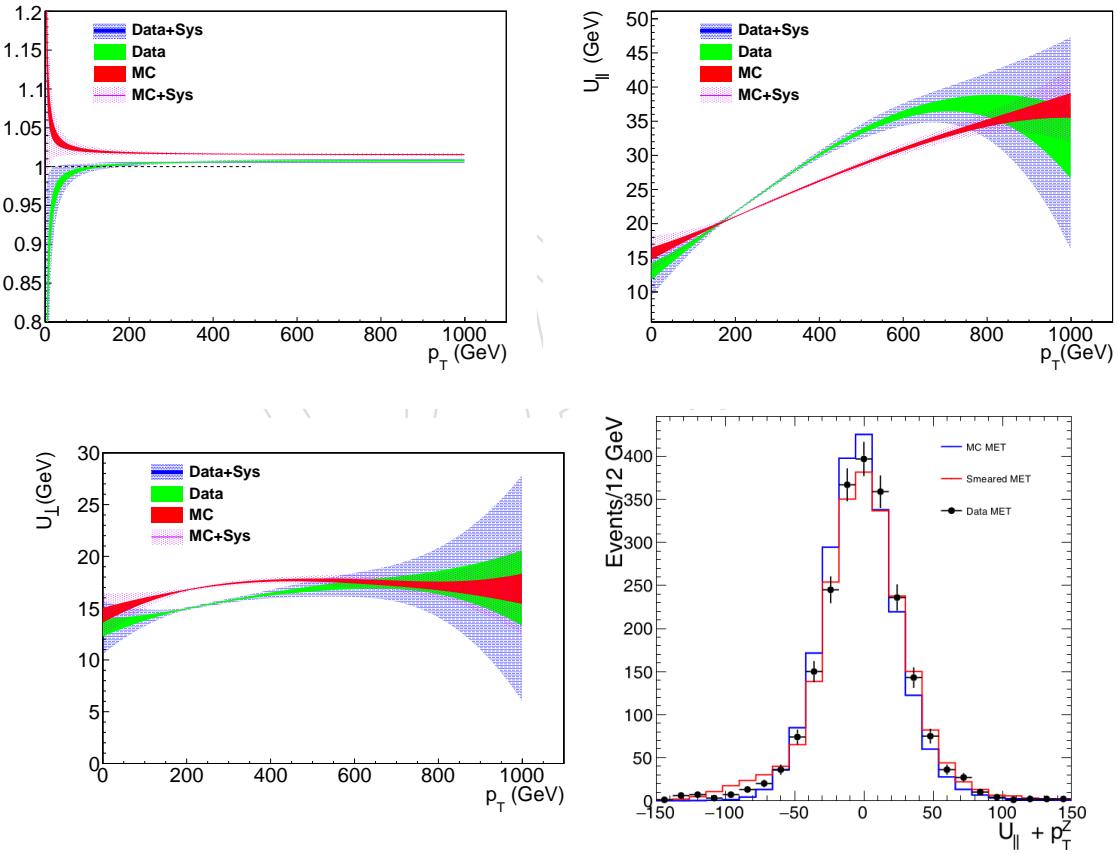


Figure 29:  $u_{||}$  response and resolution (top) and  $u_{\perp}$  resolution component of the missing energy in data and simulation. The recoil in MC before and after smearing is shown and compared to the recoil in data. The smeared recoil agrees better with data in both mean and RMS.

723 Uncertainties on the  $E_T^{\text{miss}}$  shape can be obtained by propagating the fit parameters through the  
 724 full set of corrections.

725 A complete description of the method used to derive recoil corrections is found in Ref. [22].

## 726 4 Boson reconstruction

### 727 4.1 Z boson to leptons

728 In leptonic channels ( $X \rightarrow Zh \rightarrow eeb\bar{b}$  and  $X \rightarrow Zh \rightarrow \mu\mu b\bar{b}$ ) the candidate Z boson is recon-  
 729 structed from the pair of selected leptons. The two leptons are chosen as the pair of same flavor  
 730 and opposite sign lepton pair with the highest  $p_T$  in the event. Additional requirements are  
 731 applied to the Z candidate:

732 **Mass:**  $70 \leq m_{\ell\ell} \leq 110 \text{ GeV}$

733  **$p_T$ :**  $p_T^Z > 200 \text{ GeV}$

### 734 4.2 Z boson to neutrinos

735 If the Z boson decays into a pair of neutrinos, no products are visible in the detector. A simple  
 736 and effective solution, adopted in this analysis, is to consider the transverse mass of the X  
 737 candidate using the jet and  $E_T^{\text{miss}}$  kinematics defined by the following formula:

$$m_{\nu\nu bb}^T = \sqrt{2E_T^j E_T^{\text{miss}} \cdot (1 - \cos \Delta\varphi(j, E_T^{\text{miss}}))}$$

738 An alternative approach, explored but not used in this analysis, is based on the so-called “recoil  
 739 formula”. Assuming an heavy resonance, and that the heavy object is produced at rest (as if the  
 740 fraction of the momentum carried by the involved partons  $x \rightarrow 1$ ),  $m_X$  can be estimated with  
 741 the following (“recoil”) second-degree equation:

$$m_X^2 - 2E_j m_X + m_j^2 - m_Z^2 = 0$$

742 Solving for  $m_X$ , this leads to two solutions:

$$m_X^\pm = E_j \pm \sqrt{E_j^2 - (m_j^2 - m_Z^2)}$$

743 and the chosen solution is chosen to be the largest one. However, the mass peak resolution is  
 744 found to be worse than the transverse mass, and the background shape is also shifted to higher  
 745 values of  $m_X$ , making this approach less effective.

### 746 4.3 W boson reconstruction

747 The  $W \rightarrow \ell\nu$  candidate is reconstructed from the charged lepton and the  $E_T^{\text{miss}}$  reconstructed  
 748 in the event. The neutrino produced in semileptonic  $W \rightarrow \ell\nu$  decays escapes undetected from  
 749 the detector, making a precise reconstruction of the W boson impossible. However, the three-  
 750 momenta of the neutrino can be estimated with kinematic reconstruction techniques [49]. Under  
 751 the assumption that there is no other source of genuine  $E_T^{\text{miss}}$  in the event, the  $x$  and  $y$   
 752 components of the reconstructed  $E_T^{\text{miss}}$  can be considered as the  $p_x$  and  $p_y$  components of the  
 753 neutrino. The third component  $p_z$  can be recovered by using imposing the W mass constraint  
 754 ( $m_W = 80.4 \text{ GeV}$ ) on the invariant mass of the so-built neutrino (with  $m_\nu = 0$ ) and the lepton  
 755 four-momenta. The resulting second-degree equation is:

$$m_W^2 = m_\ell^2 + m_\nu^2 + 2(E_\ell E_\nu - p_{x\ell} p_{x\nu} - p_{y\ell} p_{y\nu} - p_{z\ell} p_{z\nu})$$

756 The two solutions of the equation are:

$$pz_v^\pm = \frac{-B \pm \sqrt{\Delta}}{2A}$$

757 where:

$$a = m_W^2 - m_\ell^2 + 2px_\ell px_v + 2py_\ell py_v$$

$$A = 4(E_\ell E_\ell - pz_\ell pz_\ell)$$

$$B = -4apz_\ell$$

$$C = 4E_\ell E_\ell(px_v^2 + py_v^2) - a^2$$

$$\Delta = B^2 - 4AC$$

758 If  $\Delta > 0$ , there are two real solutions and  $pz^\nu$  is chosen as the solution with the smaller absolute  
 759 value. In a generic  $t\bar{t}$  event, this choice has been found to be the correct one in about 60% of the  
 760 events [49]. On the other hand, if  $\Delta < 0$ , there is no real solution but two complex solutions  
 761 instead. In this case,  $pz^\nu$  is set to the real part of the two solutions ( $pz^\nu = -B/2A$ ). Figure 30  
 762 shows the mass of the candidate W boson after reconvering the neutrino  $p_z$ . An additional  
 763 requirement on the W  $p_T$  is applied in order to select boosted topologies:

764  $p_T: p_T^W > 200 \text{ GeV}$

#### 765 4.4 Higgs to bb reconstruction

766 The SM V+jets production represents the main background of the analysis. These events have  
 767 the same topology but the jet is generated by different processes: jets from background events  
 768 are produced by one single parton, while jets from the signal samples are generated by a pair of  
 769 quarks or gluons. As a consequence, it becomes important to distinguish as much as possible  
 770 jets produced by QCD interactions from merged jets produced in the Higgs decay. Being the  
 771 jets composite objects, their mass and internal structure contain valuable information. The  
 772 jet mass is defined as the invariant mass of all the objects contained inside the jet: the pion  
 773 mass is associated to charged hadronic tracks, while the reconstructed photons are considered  
 774 massless. The identification of jets produced by Higgs decay is based on three ideas:

775 **Jet mass:** jets produced by the decay of a massive particle should have the invariant mass  
 776 around the nominal mass of the original particle. Oppositely, jets originated by  
 777 QCD radiation are produced by the emission of quarks or gluons and typically have  
 778 smaller invariant mass.

779 **Jet substructure:** looking inside the structure of jets can help the discrimination of the  
 780 original seed of the jet. Indeed, H-jets are produced by two partons merged into a  
 781 single fat jet.

782 **Sub-jet b-tagging:** one of the most characteristic features to distinguish jets originating  
 783 from  $h \rightarrow b\bar{b}$  from V+jets backgrounds is to check for the presence of one or two  
 784 b-quarks inside the jet.

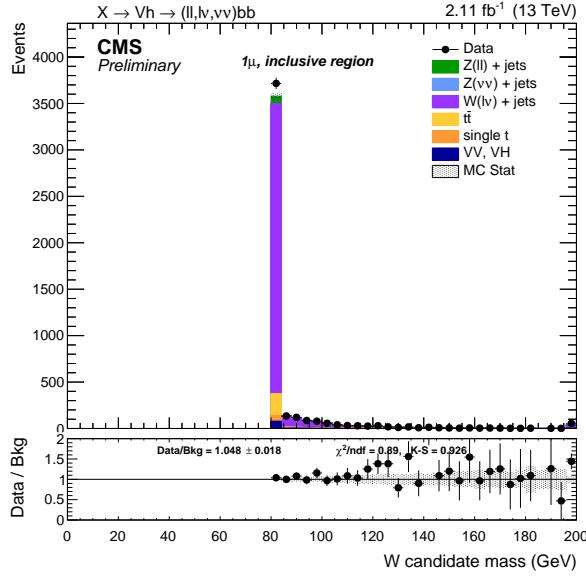


Figure 30: Mass of the W candidate after reconstructing the neutrino  $p_z$ . The events at  $m_W = 80.4$  GeV correspond to those with a real solution of the second-degree equation. The other events, scattered at different mass values, correspond to those events where there is no real solution of the equation.

For the reconstruction of the mass of the heavy resonance  $m_X$ , the kinematics of the ungroomed jets are used instead. The reason for this choice is that jet energy corrections are not available for groomed jets, so the original calibrated jet is used to have a proper description of the event kinematics.

#### 4.4.1 Jet mass

The mass cut applied on the jet mass is a critical step of the analysis, and it has to fulfill three purposes: it has to provide the maximum signal significance (best compromise between signal efficiency and background reduction), it has to avoid overlaps with the  $Z \rightarrow b\bar{b}$  mass window, and it has to provide a sufficient data and simulation statistics for the control regions (the regions outside the mass cut). The pruned and soft-drop jet mass distributions are shown in Figure 31. Ref. [50] and Fig. 31 show that the soft-drop mass variable has stronger  $p_T$  and pileup dependence, and, therefore, the pruned mass variable is recommended for the first round of all diboson analyses. The uncorrected pruned mass of the higgs jet peaks at the lower mass region populated with more background. The standard L2RELATIVE and L3ABSOLUTE (with L2L3RESIDUAL for data) jet energy corrections are applied to the pruned mass and shift the peak position from  $\approx 110$  GeV to  $\approx 120$  GeV. See Fig 32 for a comparison of the corrected pruned mass from the DY+Jets background and various mass variables from the X signal.

An optimization procedure is then performed to choose the best mass cut window. The chosen figure of merit is the *Punzi significance*, which has the advantage to be independent on the signal normalization, and is defined as:

$$\mathcal{P} = \frac{\varepsilon_S}{1 + \sqrt{B}}$$

where  $\varepsilon_S$  is the signal efficiency and  $B$  the number of background events. Both of them are evaluated by counting events within a  $2\sigma$ -width window in the  $m_X$  mass spectrum centered around the peak mean value. Here, the  $\varepsilon_S$  is the signal efficiency with its denominator being

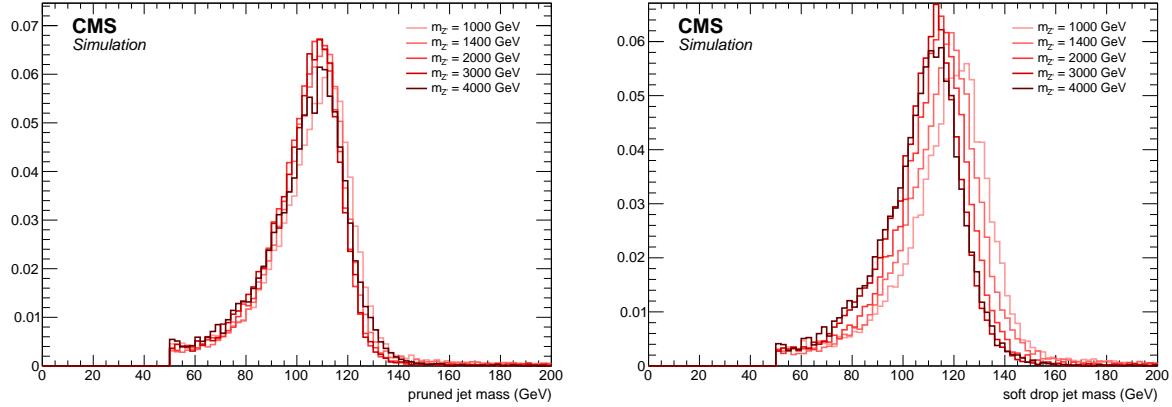


Figure 31: Jet mass distribution for different signal samples for the pruned (left) and soft drop mass (right).

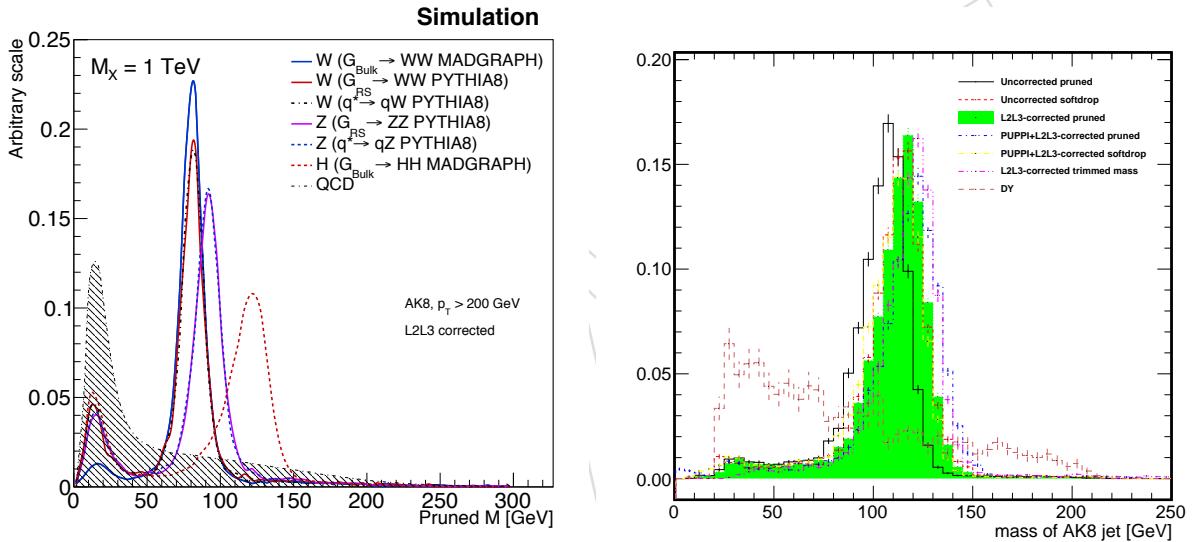


Figure 32: left: L2L3-corrected pruned mass for hadronic  $W$ ,  $Z$  and Higgs decays and a resonance mass of 1 TeV. The shape of the multijet background is also displayed for comparison. right: Comparison of the L2L3-corrected pruned mass from the DY+jets background and various mass variables from the 2-TeV  $Z'$  signal. These mass variables are computed with various grooming techniques, including softdrop, pruning, and trimming, and two different pileup subtraction methods: charge hadron subtraction (CHS, not explicitly labeled in the figure) and pileup per particle identification (labeled with "PUPPI").

805 the number of signal events passing the preselection criteria and the  $B$  is the number of back-  
 806 ground events normalized to  $5 \text{ fb}^{-1}$  of integrated luminosity. Figure 33 shows the best Punzi  
 807 significance and its corresponding signal/background efficiencies, as a function of  $Z'$  mass.  
 808 Figure 34 shows the ratio of best significance relative to the significance when applying a mass  
 809 cut of 105-135 GeV (common window of the diboson group); in addition, the efficiency ratios  
 810 are also shown. Table 12 lists the input numbers of Punzi significance for various mass win-  
 811 dows and for  $Z'$  mass at 1, 2, 3, and 4 TeV respectively. Overall, about 20–30% of significance is

<sup>812</sup> reduced by using the mass window 105–135 GeV.

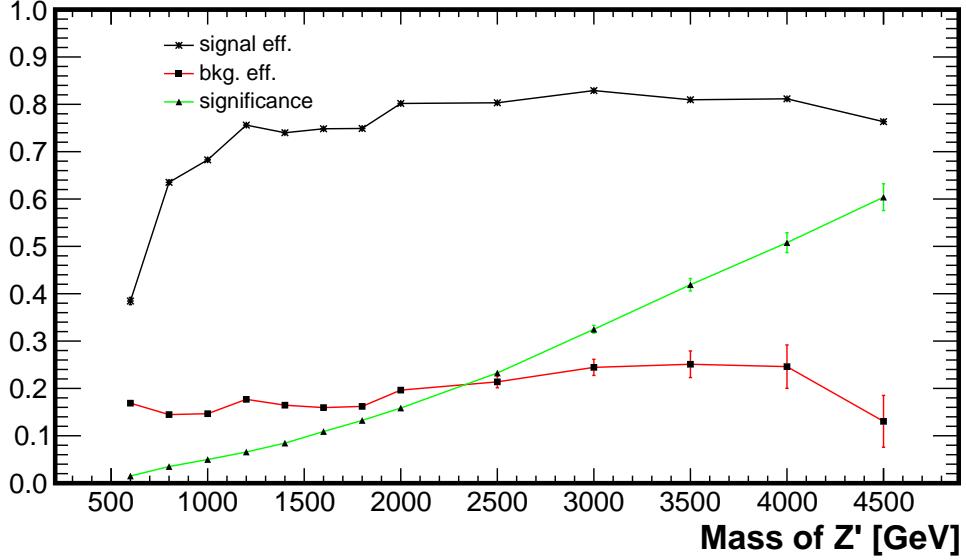


Figure 33: The best Punzi significance for the selection on L2L3-corrected pruned mass and the corresponding signal/background efficiencies, as a function of  $Z'$  mass.

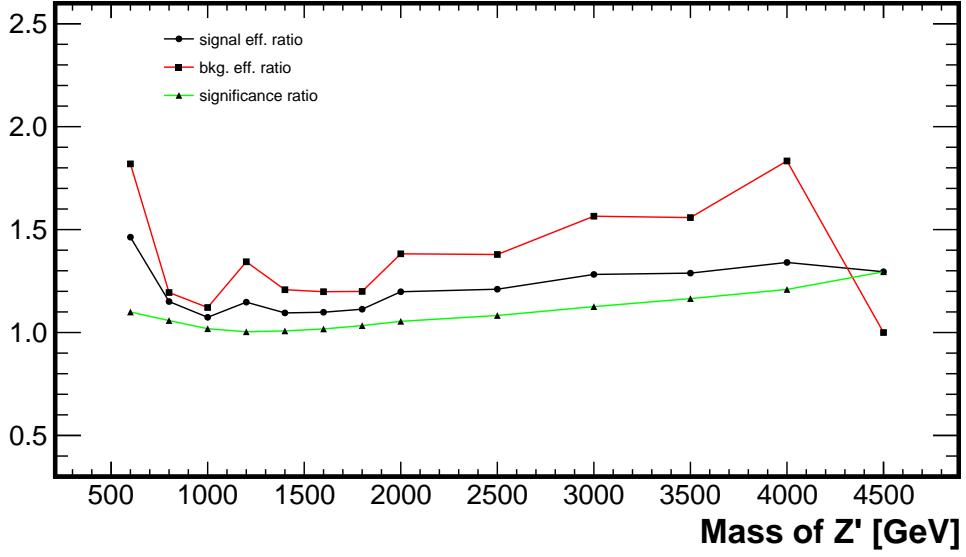


Figure 34: The ratio of the best significance relative to the significance with mass window of 105–135 GeV, as a function of  $Z'$  mass. The corresponding efficiency ratios for the signal and background are also shown.

<sup>813</sup> Alternatively, we compute the  $\epsilon_S$  and the  $B$  differently by counting events within a  $2\sigma$ -width window in the  $m_{\ell\ell\text{bb}}$  or  $m_{\nu\nu\text{bb}}^T$  mass spectrum centered around the peak mean value. The outcome of the test is shown in Figure 35. For the pruned mass, the lower cut is much more important than in the soft drop case, and it is the real limitation to the signal efficiency.

<sup>817</sup> In the present analysis, the signal region window chosen is not optimized. In order to comply

Table 12: Punzi significance, the signal efficiency and the number of background events for various mass windows. The numbers in bold font correspond to the mass window with best significance for this analysis. The numbers for the common window 105–135 GeV and the best window for all-hadronic channel are also shown.

$M_{Z'}$	corr. $M_{\text{pruned}}$ [GeV]	$\mathcal{P}$	$\varepsilon_S$	$B$
1 TeV	<b>105–140</b>	<b>0.0498</b>	<b>0.6827</b>	<b>161.6</b>
	105–135	0.0489	0.6357	144.1
2 TeV	<b>95–135</b>	<b>0.1586</b>	<b>0.8018</b>	<b>16.45</b>
	105–135	0.1504	0.6693	11.90
	100–130	0.1561	0.7090	12.55
3 TeV	<b>90–135</b>	<b>0.3249</b>	<b>0.8290</b>	<b>2.408</b>
	105–135	0.2886	0.6466	1.539
	100–130	0.3062	0.6991	1.646
4 TeV	<b>90–145</b>	<b>0.5078</b>	<b>0.8116</b>	<b>0.3580</b>
	105–135	0.4200	0.6055	0.1952
	95–135	0.4903	0.7449	0.2696

818 with the mass cut used in other analyses [50], only the  $m_{\text{pruned}} > 105$  GeV window is available;  
 819 as a consequence, the chosen cut for the signal region is  $105 < m_{\text{pruned}} < 135$  GeV.

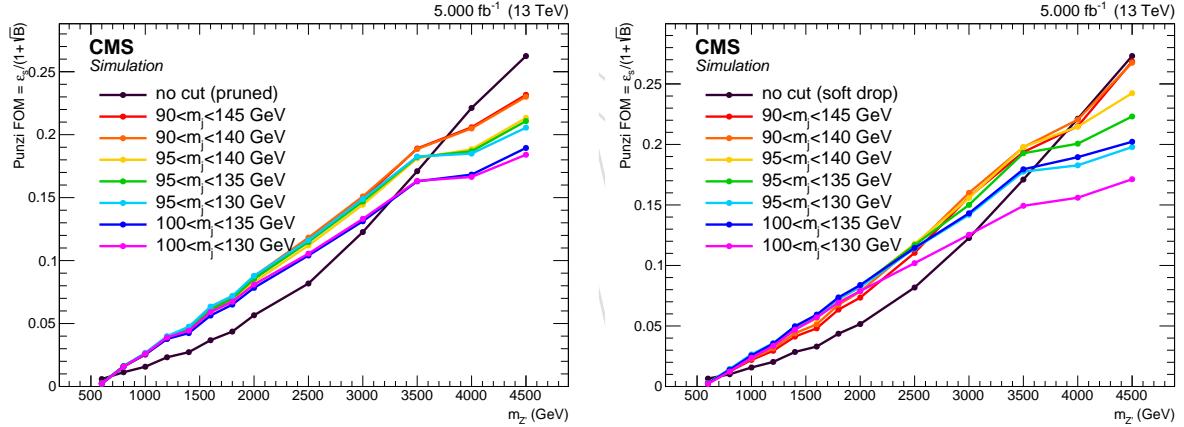


Figure 35: Signal significance as a function of the mass cut applied on the fat-jet pruned (left) and soft drop mass (right).

#### 820 4.4.2 Sub-jet b-tagging

821 The cut applied to the b-tagging discriminator can vary dramatically the sensitivity of the anal-  
 822 ysis. The best solution is to find a compromise between signal efficiency and background re-  
 823 duction. A complete scan of the b-tagging operating points and fat-jet/sub-jet choices has been  
 824 performed. The chosen figure of merit is the *Punzi significance*, which has the advantage to be  
 825 independent on the signal normalization where  $\varepsilon_S$  is the signal efficiency and  $B$  the number of  
 826 background events.

827 All the possible working points are tested for tagging the fat-jet, or both the sub-jets indepen-  
 828 dently<sup>1</sup>. Figure 36 shows that fat-jet b-tagging has the best performance for high resonance

<sup>1</sup>The following short notation will be used to indicate the btagging working point tested: CSVX meaning the X

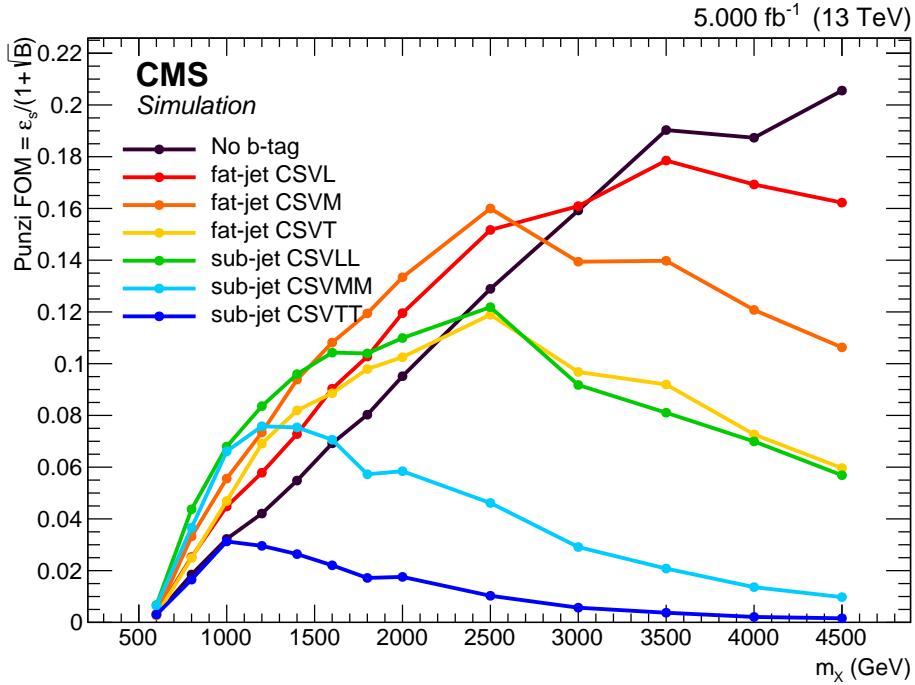


Figure 36: Signal significance as a function of the b-tagging working point, for both fat-jet and sub-jet b-tagging.

masses, but it is outclassed by the sub-jet b-tagging at lower masses ( $m_X \lesssim 2$  TeV). On the contrary, sub-jet b-tagging is not the best choice at high masses, mainly for two different reasons:

- in the highly boosted regime, background is very small and plays a minor role in the significance, which is dominated by the signal efficiency
- sub-jet b-tagging tend to lose efficiency due to the proximity of the two sub-jets. At  $m_X \approx 4$  TeV, the angular separation between the two b-quarks is  $\Delta R \sim 0.12$

One of the possible solutions for the former issue is defining a *looser-than-loose* working point, denoted here as *Ultra-loose* (U). This working point is not official, and its value is set to an arbitrary value of 0.3 (for reference, the *loose* working point is 0.605). Figure 37 show that although a small increase of the significance is obtained in the high end of the mass spectrum, for a large part of the mass points no clear advantage over the *loose* working point is achieved. This *Ultra-loose* working point is then not used in the analysis, preferring the *loose* working point.

A possible solution in order to address the latter issue is to switch between fat-jet and sub-jet b-tagging depending on the separation between the two sub-jets. This method, used in some Run-I analyses, allows to have a better significance over the whole mass spectrum as seen in Figure 38, without splitting the signal samples in two parts. On the other hand, switching between two different b-tagging methods is not the best choice when systematic uncertainties have to be taken into account. Furthermore, the CSV algorithm is trained on a sample with only one b-quark in the jet, and applying it to a fat-jet that contains two b-quarks is not an optimal choice.

working point for the tagging of the fat-jet; CSVXY meaning the X and Y working points for the tagging of the two sub-jets. Three standard working points are defined: *loose* (L), *medium* (M), *tight* (T).

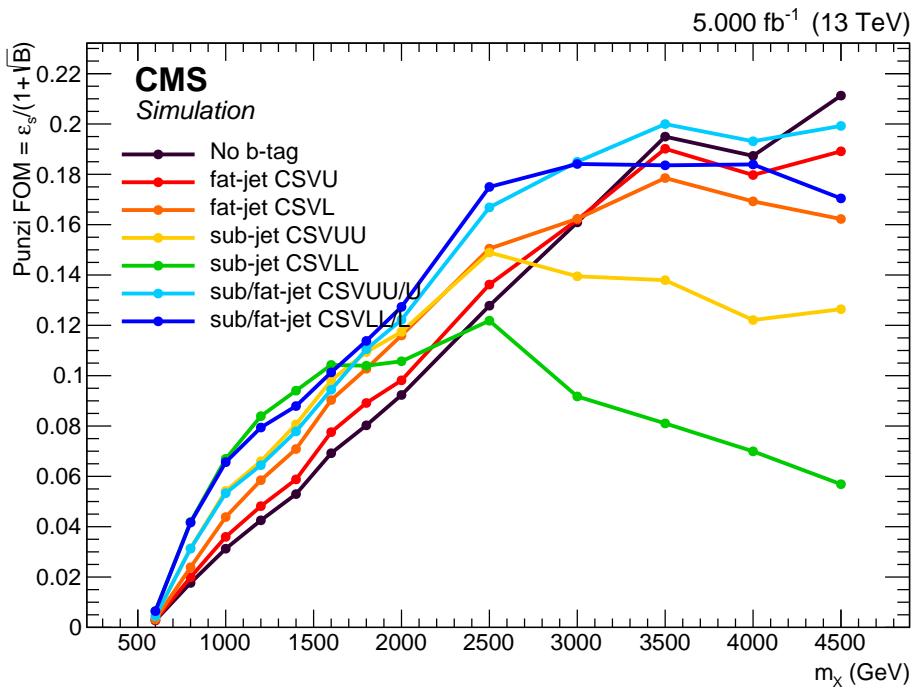


Figure 37: Signal significance as a function of the b-tagging working point, for both fat-jet and sub-jet b-tagging.

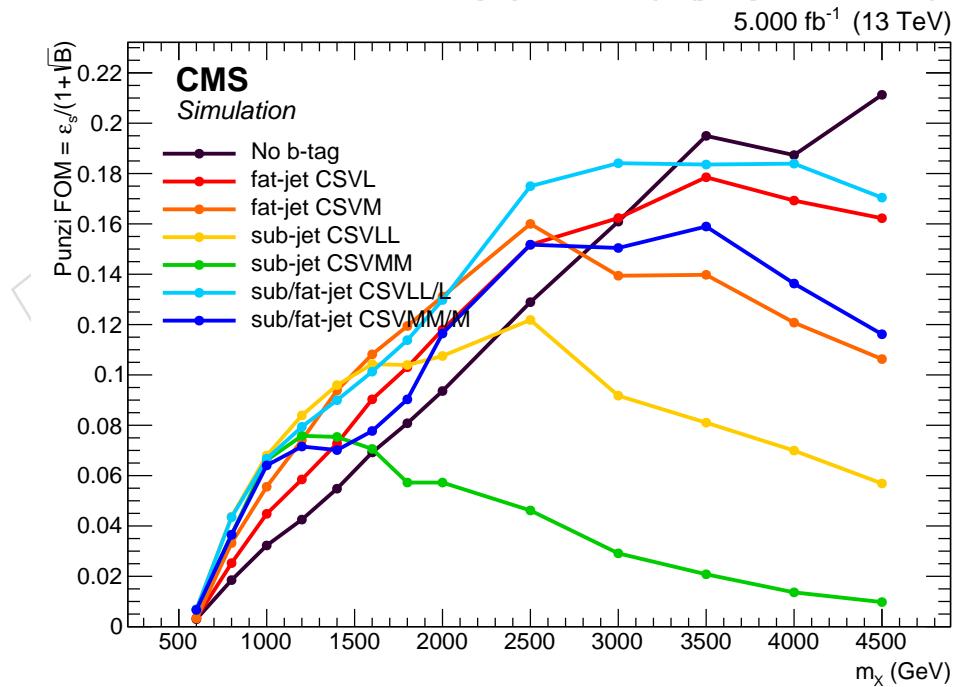


Figure 38: Signal significance as a function of the b-tagging working point, for both fat-jet and sub-jet b-tagging, and for switching between the two options for  $\Delta R \leq 0.3$ .

850 An alternative solution is to avoid the use of fat-jet b-tagging, and focus only on sub-jets. In the  
 851 high-boosted regime, the signal efficiency can be recovered by applying asymmetric working

points on the two sub-jets, or even dropping any b-tagging requirement on the sub-leading CSV jet. This option is considered in the test in Figure 39, and allow to have comparable or even better results than the fat-jet/sub-jet switching method with a more coherent approach. Figure 40 show that varying the  $\Delta R$  threshold does not change the overall signal significance, until the value drops under 0.2, and confirms that 0.3 is an optimal value.

For the analysis development, two b-tagging categories are made, as a result of the studies described in this Section. The 2 b-tag category, for which both the 2 sub-jets pass the cut provided by the *loose* working point, offers a much better signal over background ratio, and dominates the sensitivity in the lower mass spectrum. The 1 b-tag category, where only 1 sub-jet is tagged by the *loose* working point, allows to recover the efficiency lost at high mass, but it is not competitive at low mass due to the large background contamination.

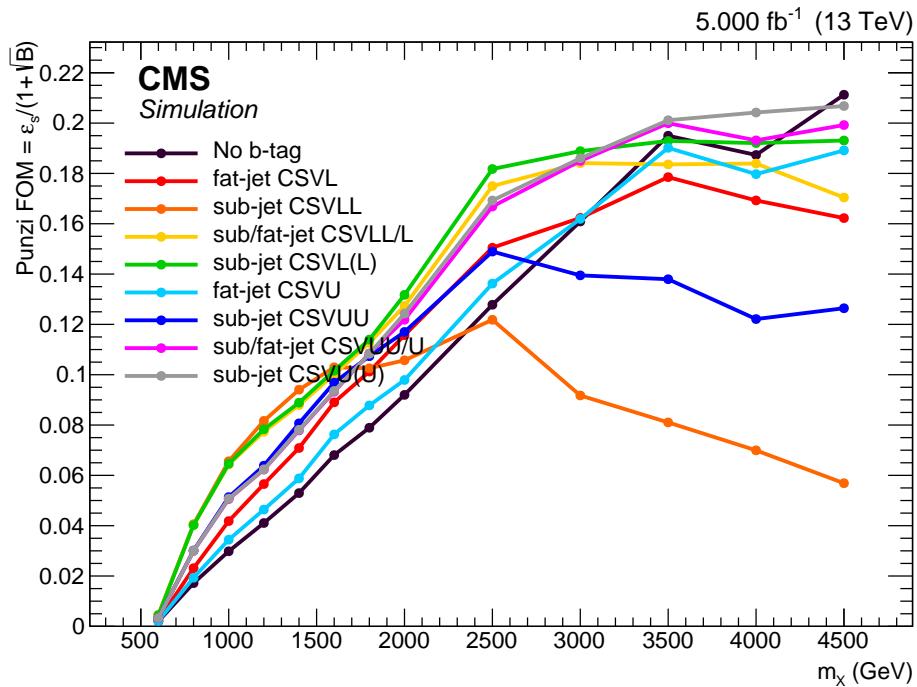


Figure 39: Signal significance as a function of the sub-jet b-tagging working point, and switching between the 1-jet and 2-jet tagged requirement for  $\Delta R \leq 0.3$ .

#### 4.4.3 Jet substructure

The  $\tau_{21}$  is a potentially discriminating variable (described in Section 3.7.2) between H-jets and QCD-jets, so a test to verify how much it can improve the signal significance is performed. The figure of merit is the same as the one described in Section 4.4.1, and it is computed after applying pre-selection cuts, V and h  $p_T$  and mass cuts. A number of possible cuts are tested, ranging from  $\tau_{21} < 0.3$  to  $\tau_{21} < 0.7$ , and compared between each other and the null cut option. Figure 41 shows the discriminating power of  $\tau_{21}$  for all sources of SM background compared to the signal, Higgs di-subjet decays. The significance for each selection is shown in Figure 42, demonstrating that cutting on  $\tau_{21}$  has no advantage with the current luminosity. The  $\tau_{21}$  variable is therefore not used in the present analysis.

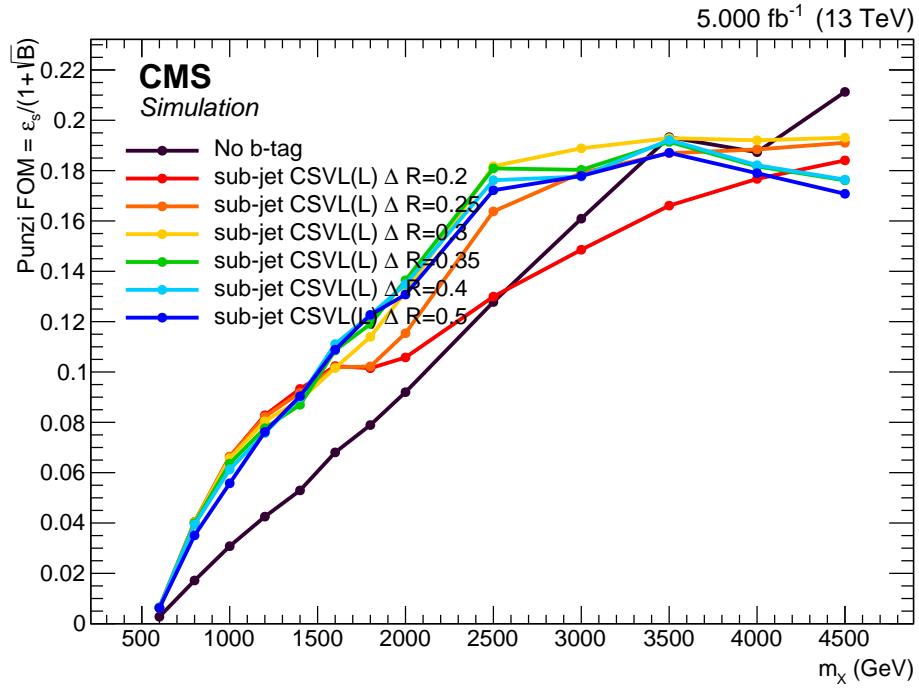


Figure 40: Signal significance requiring just one or both sub-jets b-tagged as a function of the  $\Delta R$  between the two sub-jets.

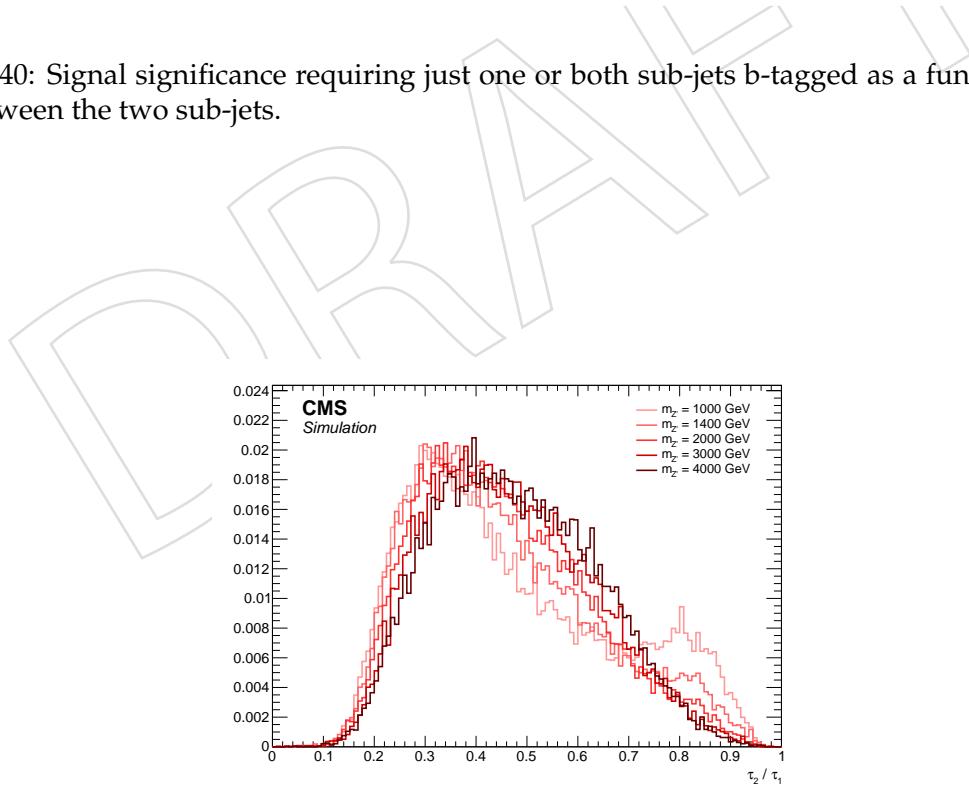
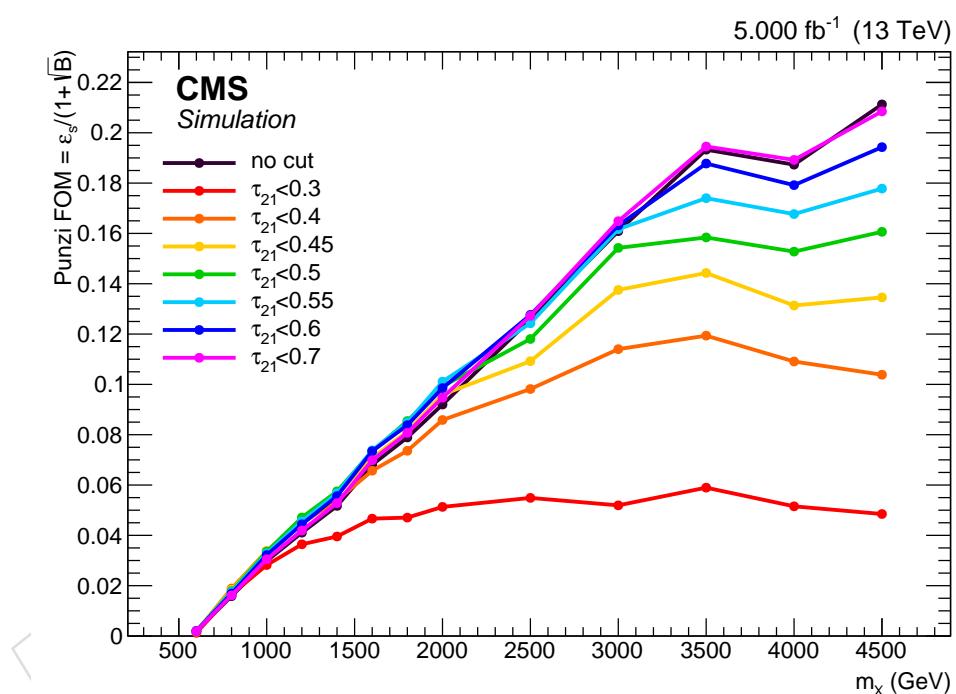


Figure 41: Jet  $\tau_{21}$  distribution for different signal samples after pre-selection and jet mass cuts.

Figure 42: Signal significance as a function of the  $\tau_{21}$  cut applied.

## 873 5 Event selection

874 Events considered in this analysis have to pass a certain number of selections before being  
 875 considered as suitable signal candidates, identically in both data and simulation. The selections  
 876 are reported below and in Table 13. The final signal efficiency is shown separately depending  
 877 on the number and flavor of the leptons and b-tagged subjets in Figure 43.

### 878 5.1 Leptonic selection

879 The leptonic *pre-selection* cuts are different by the final state, selected physics objects and candi-  
 880 date reconstruction:

881 **Di-electron channel:** ( $X \rightarrow Zh \rightarrow ee\bar{b}\bar{b}$ )

882 **Trigger:** HLT\_Ele105\_CaloIdVT\_GsfTrkIdT\_v\*

883 **p<sub>T</sub>:** at least two PF electrons with  $p_T > 135\text{ GeV}$  and  $35\text{ GeV}$  for leading and  
 884 sub-leading, respectively

885 **J:**  $|\eta_{SC}| < 1.4442$  or  $|\eta_{SC}| < 1.566$  and  $|\eta_{SC}| < 2.5$

886 **Id:** both electrons identified with the *loose* working point

887 **Iso:** included in Id requirement

888 **Z mass:**  $70 \leq m_{ee} \leq 110\text{ GeV}$

889 **Zp<sub>T</sub>:**  $p_T^Z > 200\text{ GeV}$

890 **Di-muon channel:** ( $X \rightarrow Zh \rightarrow \mu\mu\bar{b}\bar{b}$ )

891 **Trigger:** HLT\_Mu45\_eta2p1\_v\* or HLT\_Mu50\_v\*

892 **p<sub>T</sub>:** at least two muons with  $p_T > 55\text{ GeV}$  and  $20\text{ GeV}$  for leading and sub-  
 893 leading muon, respectively

894 **J:**  $|\eta| < 2.4$

895 **Id:** at least one muon identified as HighPt, the other with the *custom tracker*  
 896 Id (Section 3.3)

897 **Iso:** tracker-isolation  $< 0.15$

898 **Z mass:**  $70 \leq m_{\mu\mu} \leq 110\text{ GeV}$

899 **Zp<sub>T</sub>:**  $p_T^Z > 200\text{ GeV}$

900 **Single-electron channel:** ( $X \rightarrow Wh \rightarrow e\nu\bar{b}\bar{b}$ )

901 **Trigger:** HLT\_Ele105\_CaloIdVT\_GsfTrkIdT\_v\*

902 **p<sub>T</sub>:** exactly one PF electrons with  $p_T > 135\text{ GeV}$

903 **J:**  $|\eta_{SC}| < 1.4442$  or  $|\eta_{SC}| < 1.566$  and  $|\eta_{SC}| < 2.5$

904 **Id:** electron identified with the *tight* working point

905 **Iso:** included in Id requirement

906 **Wp<sub>T</sub>:**  $p_T^W > 200\text{ GeV}$

907 **Single-muon channel:** ( $X \rightarrow Wh \rightarrow \mu\nu\bar{b}\bar{b}$ )

908 **Trigger:** HLT\_Mu45\_eta2p1\_v\* or HLT\_Mu50\_v\*

909 **p<sub>T</sub>:** exactly one muon with  $p_T > 55\text{ GeV}$

910 **J:**  $|\eta| < 2.4$

911 **Id:** muon identified as HighPt and PF

912 **Iso:** tracker-isolation  $< 0.15$

913 **Wp<sub>T</sub>:**  $p_T^W > 200\text{ GeV}$

914 **Neutrino channel:** ( $X \rightarrow Zh \rightarrow \nu\nu\bar{b}\bar{b}$ )

915       **Trigger:** HLT\_PFMETNoMu90\_PFMHTNoMu90\_IDTight  
 916           or HLT\_PFMETNoMu120\_PFMHTNoMu120\_IDTight

917  
 918       **MET:** Type-1 corrected missing energy  $E_T^{\text{miss}} > 200 \text{ GeV}$

## 919   **5.2 Hadronic selection**

920   The selections of the hadronic part are exactly the same for all the three lepton categories, with  
 921   the exception of the jet Id in the specific case of the 0 leptons final state, for which a tighter cut  
 922   is applied:

923   **Hadronic selection:** (all categories)

924       **p<sub>T</sub>:** at least one AK8 PFJet with  $p_T > 200 \text{ GeV}$

925       **j:**  $|\eta| < 2.4$

926       **Id:** *loose* particle-flow Id for the single- and double-lepton categories; *tight*  
 927       particle-flow Id for the 0 lepton channel

928       **Lepton cleaning:** minimal separation between jet and isolated leptons  $\Delta R_{jet-\ell} >$   
 929       0.8

930       **Mass:** pruned mass with L2L3 JEC (Section 3.7.1)

931       **Substructure:** **no** selection on  $\tau_{21}$

932       **b-tagging:** **one or two** sub-jets b-tagged with the pfCombinedInclusiveSecondaryVertexV2BJet  
 933       depending on the category (Section 3.8)

## 934   **5.3 Topology and event cleaning**

935   The following additional selections are applied to reject noisy events, reject multijet or  $t\bar{t}$  back-  
 936   grounds, and assure orthogonality between all channels:

937       **Di-electron and di-muon channel:** ( $X \rightarrow Zh \rightarrow ee\bar{b}\bar{b}$  and  $X \rightarrow Zh \rightarrow e\mu\bar{b}\bar{b}$ )

938       **Bkg discrimination:**  $\Delta\eta(Z, jet) < 5$

939       **Cleaning:**  $\Delta\varphi(Z, jet) > 2.5$

940       **Single-electron and single-muon channels:** ( $X \rightarrow Wh \rightarrow e\nu\bar{b}\bar{b}$  and  $X \rightarrow Wh \rightarrow \mu\nu\bar{b}\bar{b}$ )

941       **Other isolated  $e, \mu$ :** veto

942       **Hadronic taus:** veto

943       **Cleaning:**  $\Delta\varphi(\ell, E_T^{\text{miss}}) < 2, \Delta\varphi(jet, E_T^{\text{miss}}) > 2$

944       **Top rejection:** anti b-tag *loose* on other AK4 jets

945       **QCD rejection:**  $E_T^{\text{miss}} > 80 \text{ GeV}$  (electron channel only)

946       **Neutrino channel:** ( $X \rightarrow Zh \rightarrow \nu\nu\bar{b}\bar{b}$ )

947       **Isolated  $e, \mu$ :** veto

948       **Hadronic taus:** veto

949       **Cleaning:**  $\Delta\varphi(jet, E_T^{\text{miss}}) > 2$

950       **Top rejection:** anti b-tag *loose* on other AK4 jets

951       **QCD rejection:** min  $\Delta\varphi > 0.5$  between  $E_T^{\text{miss}}$  and all AK4 jets

952       **Noise rejection:** AK8 jet *tight* id

	$Z \rightarrow ee$	$Z \rightarrow \mu\mu$	$W \rightarrow e\nu$	$W \rightarrow \mu\nu$	$Z \rightarrow \nu\nu$
Trigger	HLT_Ele105	HLT_Mu45_eta2p1 or HLT_Mu50	HLT_Ele105	HLT_Mu45_eta2p1 or HLT_Mu50	HLT_PFMETNoMu90_PFMHTNoMu90 HLT_PFMETNoMu120_PFMHTNoMu120
Leptons	$2e$ loose WP inc. in Id	$1\mu$ HighPt, $1\mu$ custom tracker trkIso < 0.1	$1e$ tight inc. in Id	$1\mu$ HighPt trkIso < 0.1	Veto
V boson $p_T$					
V boson mass	$70 < m_{\ell\ell} < 110 \text{ GeV}$	$\rho_T^V > 200 \text{ GeV}$	-	$E_T^{\text{miss}} > 200 \text{ GeV}$	-
H-jet					
H mass		$105 < m_j < 135 \text{ GeV}$ (SR), $30 < m_j < 65 \text{ GeV}$ , $m_j > 135 \text{ GeV}$ (SB)			
H b-tag		1 or 2 subjets b-tagged loose			
Top rejection	-		max CSV AK4 anti-loose		
QCD rejection	-	$MET > 80 \text{ GeV}$	-	$\min \Delta\phi_{j, E_T^{\text{miss}}} > 0.5$	
Noise cleaning	$\Delta\eta(Z, jet) < 5$ $\Delta\varphi(Z, jet) > 2.5$	$\Delta\varphi(\ell, E_T^{\text{miss}}) < 2$ $\Delta\varphi(jet, E_T^{\text{miss}}) > 2$	$\Delta\varphi_{jet, E_T^{\text{miss}}} > 2$ jet tight Id		

Table 13: Summary of the selection cuts for all the five channels of the  $VH \rightarrow b\bar{b}$  analysis.

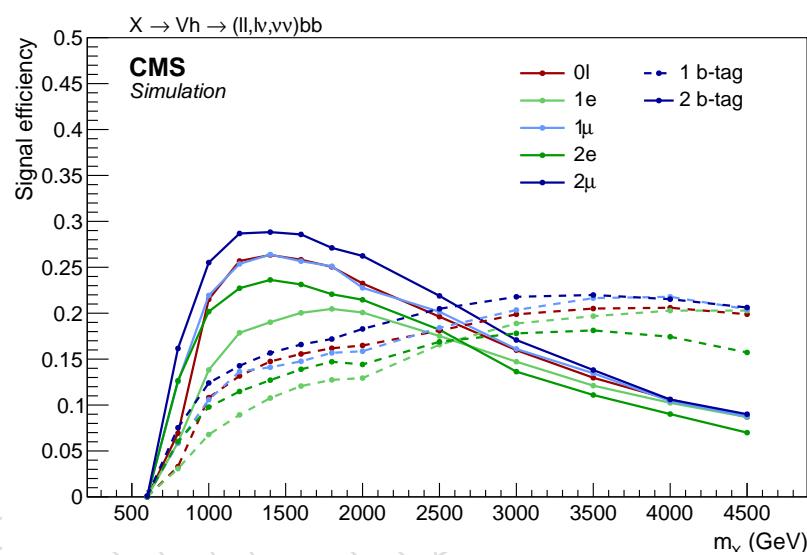


Figure 43: Signal efficiency separated by final state and b-tagging multiplicity after the signal region selections.

## 953 6 Data-simulation comparison

954 In this section, a comparison between data and simulation is reported for various kinematic  
 955 observables. It can be seen that the dominant background contribution comes from the V+jets  
 956 production (V being Z or W depending on the channel), while sub-leading contributions from  
 957  $t\bar{t}$  and dibosons can be non-negligible especially in the zero and 1-lepton channels.

958 In the following plots, the comparison is performed in three different regions. On top of the  
 959 preselections defined in Section 5, additional regions are defined for each category:

960 **Inclusive:** no selection is performed on top of the preselections, except for a veto on the  
 961 jet mass  $65 < m_j < 135 \text{ GeV}$  to avoid signal contamination from VV and VH signals

962 **Sidebands (SB):** Only events in the sidebands, defined the interval between  $30 < m_j <$   
 963  $65 \text{ GeV}$  and  $m_j > 135 \text{ GeV}$  are considered. This region can be considered as signal-  
 964 depleted. The main difference with the previous regions is that the bulk of the jet  
 965 mass distribution, peaking at  $m_j \sim 20 \text{ GeV}$ , is not included. The region selected is  
 966 thus much closer kinematically to the signal region.

967 **Signal region (SR):** Represents the phase space where signal is expected. The signal  
 968 region is considered **blind**, so data is not shown in these plots.

969 For visualization purposes, the background normalization is rescaled in SB and SR to the num-  
 970 ber of events observed in data in the corresponding SB, with the  $t\bar{t}$  normalization fixed from the  
 971 appropriate control regions described in Section 7. In the inclusive regions, all the backgrounds  
 972 are normalized to luminosity.

### 973 6.1 Zero lepton channel

### 974 6.2 Single lepton channel

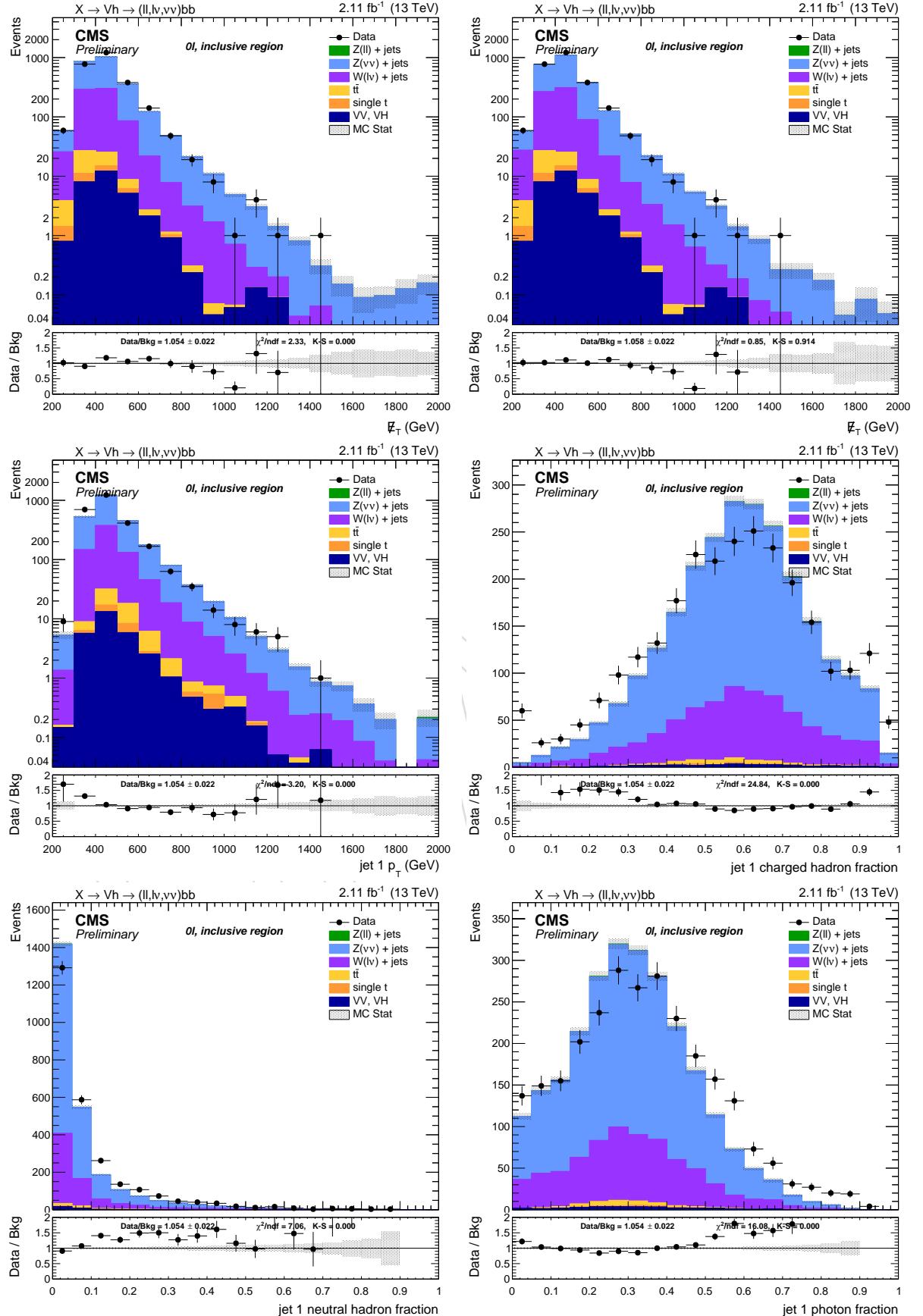


Figure 44: Top:  $E_T^{\text{miss}}$  before (left) and after (right) the recoil corrections. Center: leading AK8 jet  $p_T$  jet (left) and charged hadron fraction (right). Bottom: jet neutral hadron fraction (left) and photon fraction (right). Events are selected with the *inclusive* selection, and simulated backgrounds are normalized to luminosity.

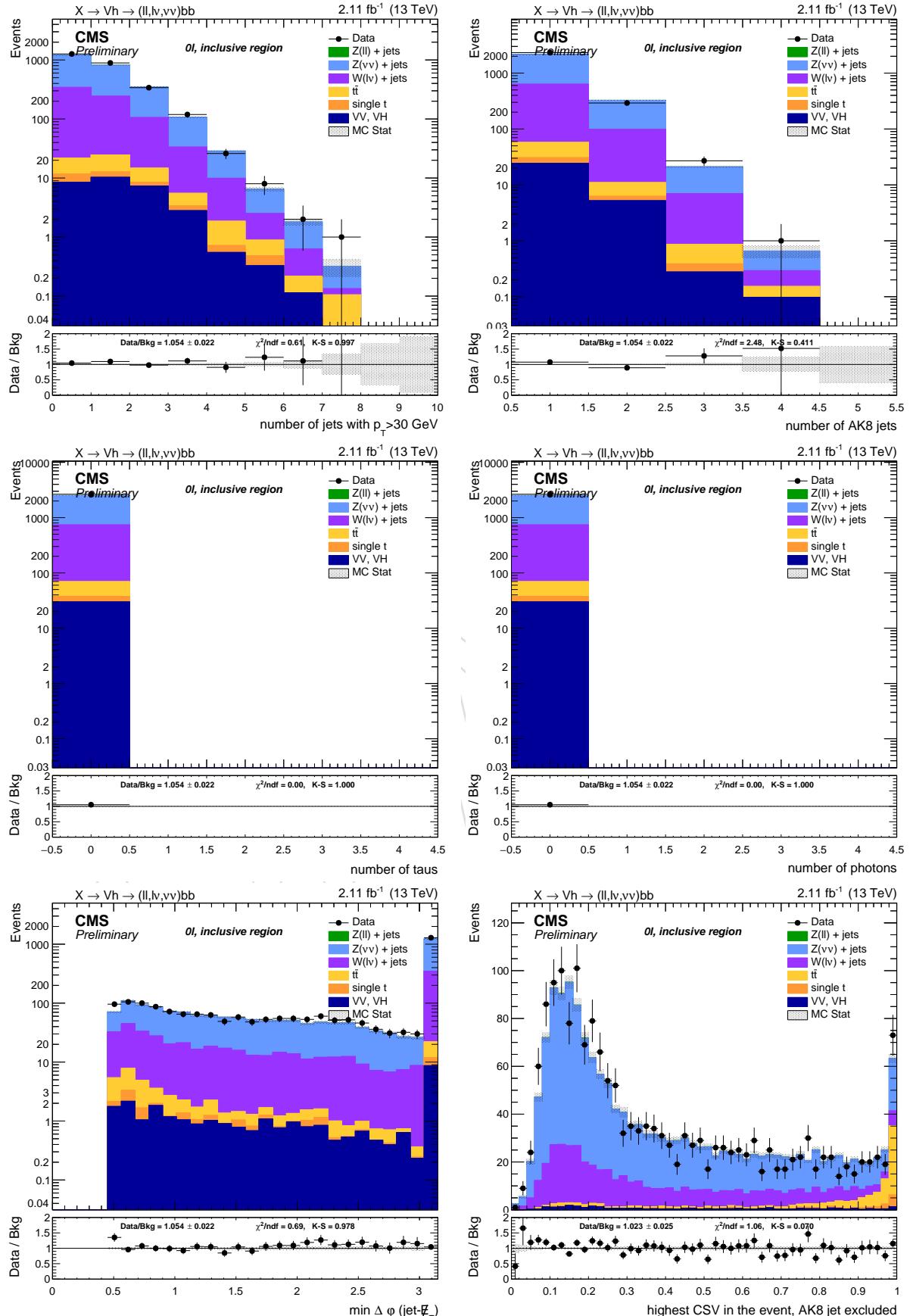


Figure 45: Top: number of AK4 jets (left) and AK8 jets (right). Center: number of hadronic taus (left) and photons (right). Bottom: minimum  $\Delta\phi$  between the AK4 and AK8 jets in the event (left) and highest CSV of the AK4 jets in the event (right). Events are selected with the *inclusive* selection, and simulated backgrounds are normalized to luminosity.

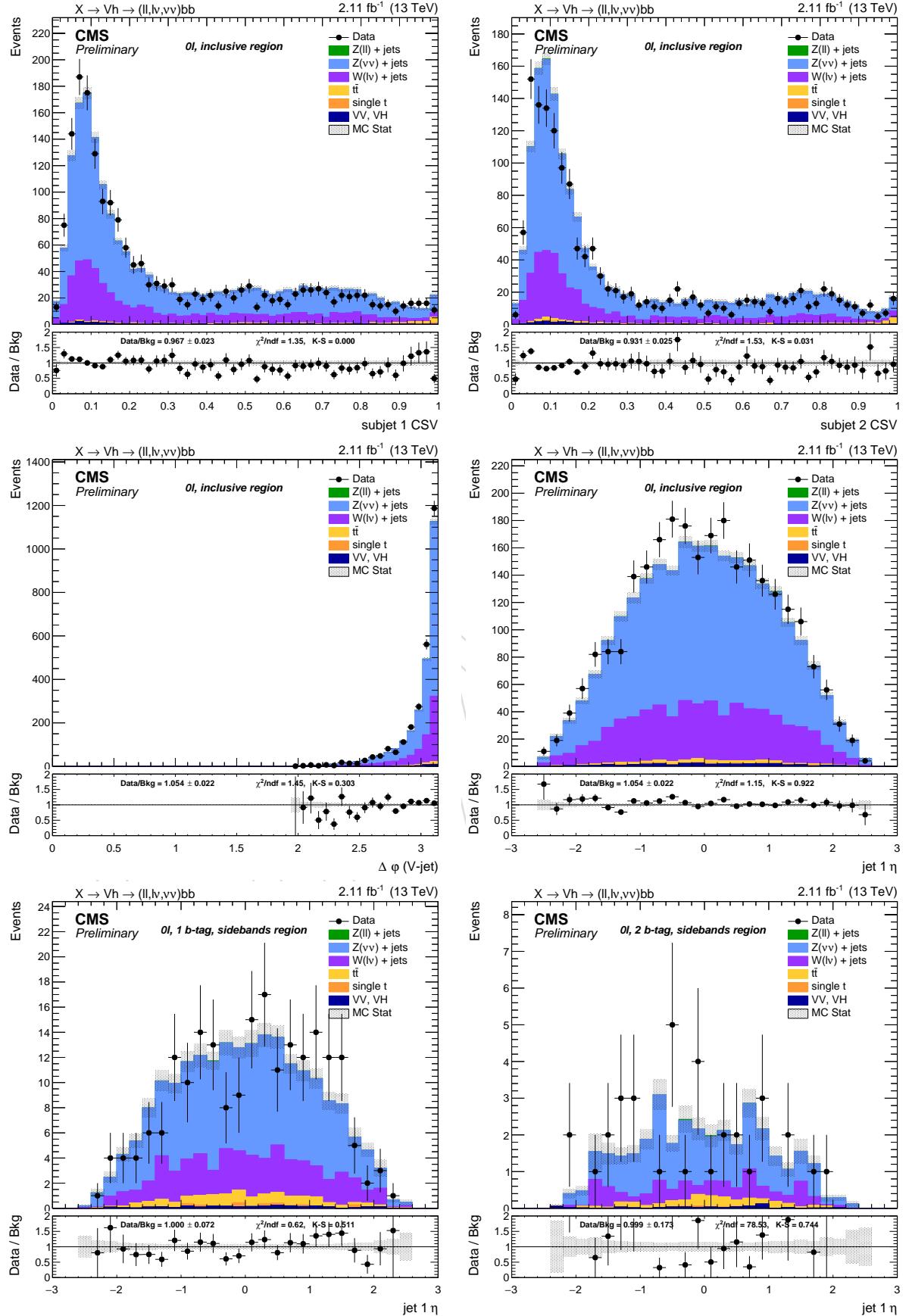


Figure 46: Top: leading AK8 jet CSV of the leading (left) and sub-leading (right) subjets. Center:  $\Delta\phi$  between the jet and  $E_T^{\text{miss}}$  (left) and  $\Delta R$  between sub-jets (right). Bottom: AK8 jet  $\eta$  in the 1 (left) and 2 b-tag category (right). Scale factors are applied in the SB regions.

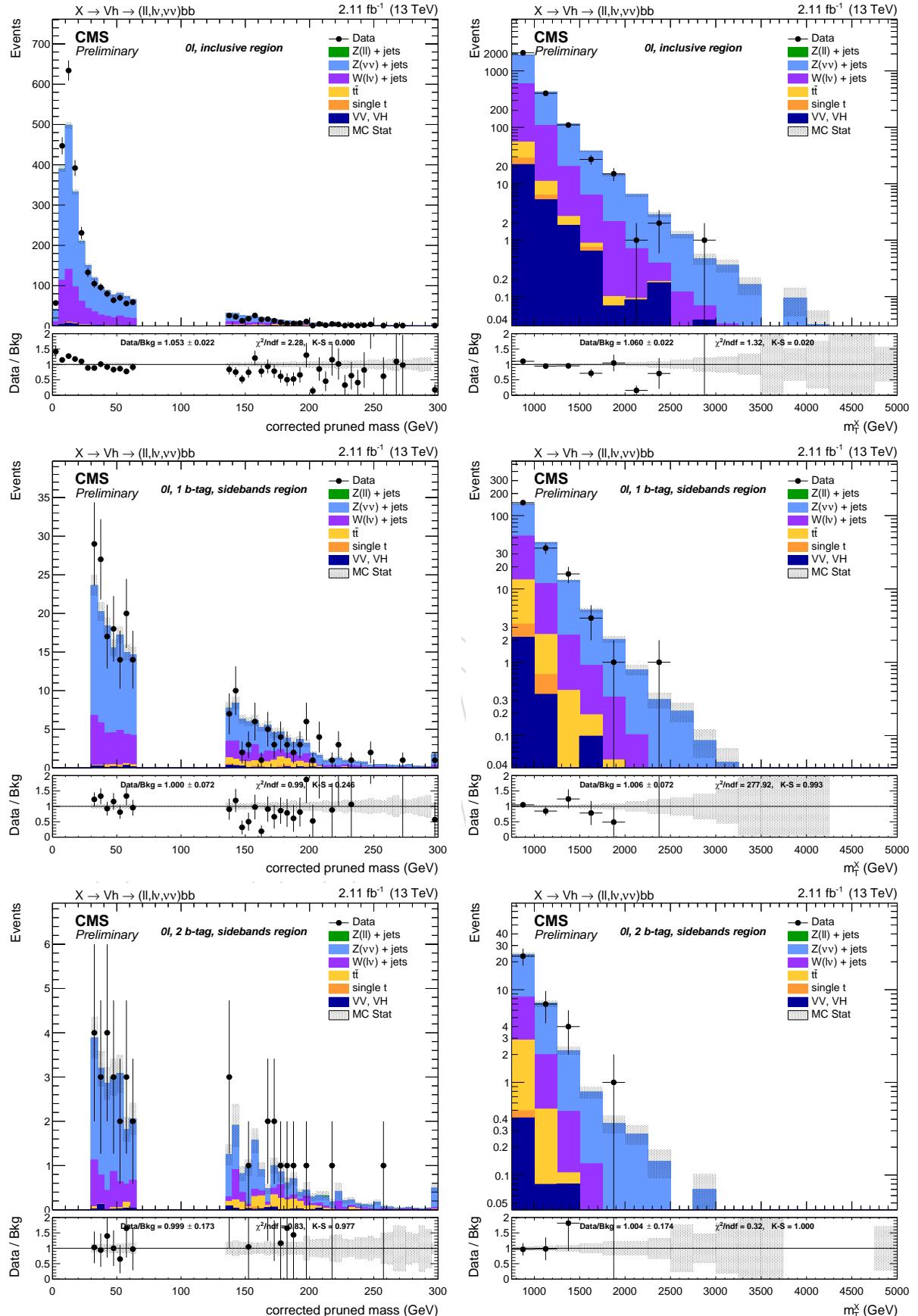


Figure 47: Leading AK8 jet pruned mass with L2L3 corrections (left) and resonance transverse mass (right). Top: inclusive region, no b-tagging nor jet mass selection applied. Center: jet mass sidebands in the 1 b-tag category. Bottom: mass sidebands and 2 b-tag category. Scale factors are applied in the SB regions.

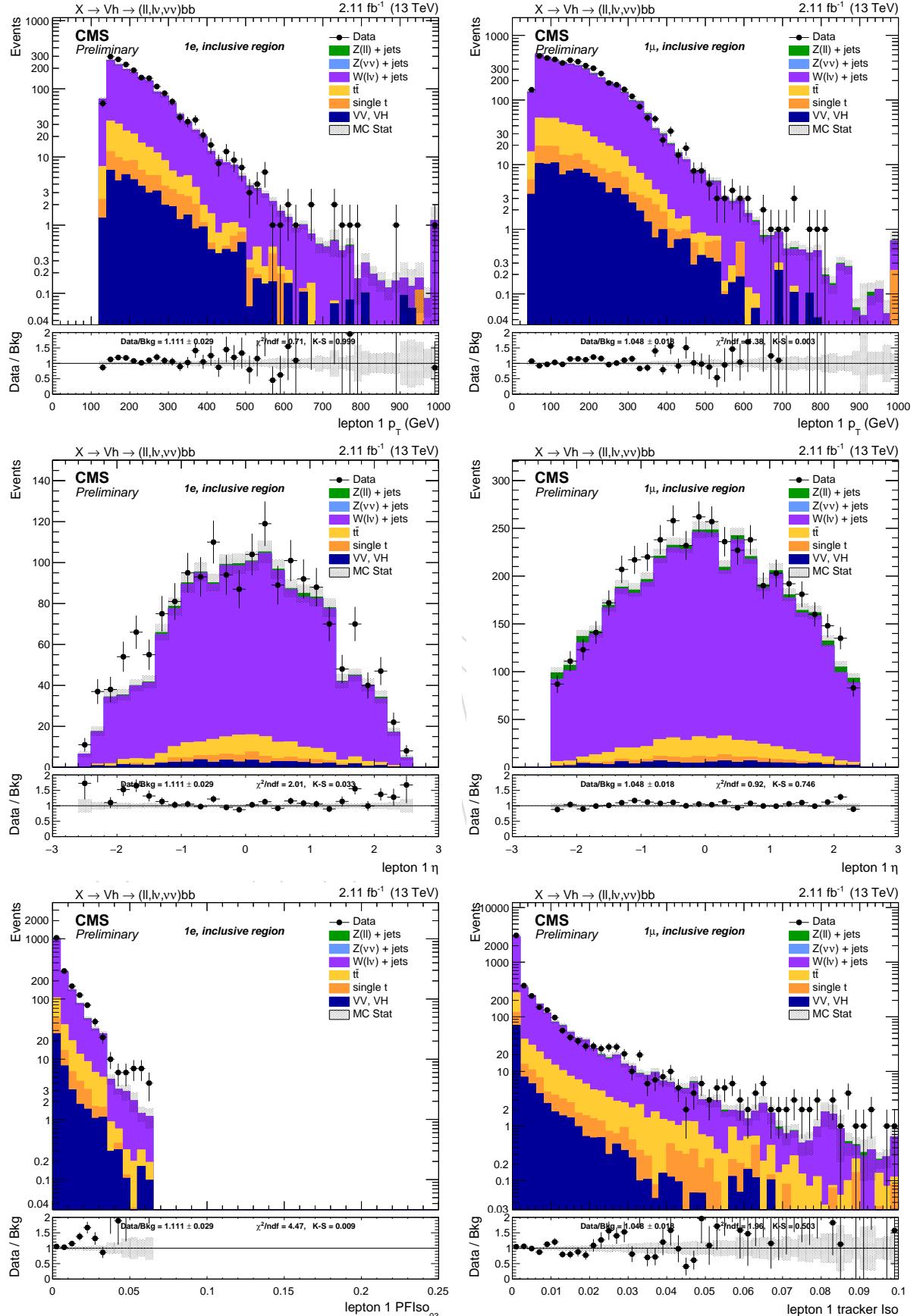


Figure 48: Lepton  $p_T$  (top),  $\eta$  (center), PF and tracker isolation (bottom). left: electron channel. right: muon channel. Events are selected with the *inclusive* selection, and simulated backgrounds are normalized to luminosity.

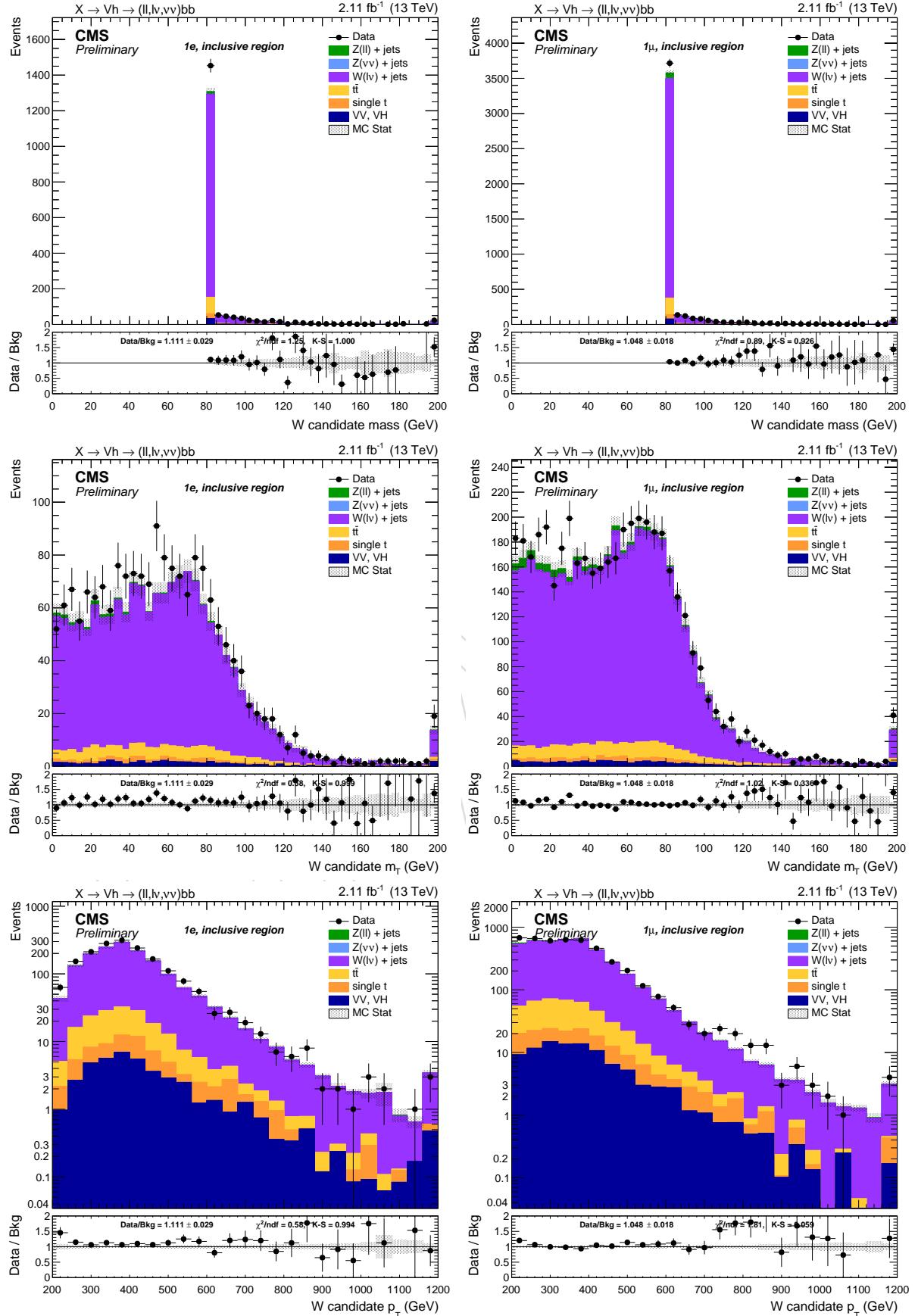


Figure 49: Top: W candidate mass. Center: W candidate transverse mass. Bottom: W candidate  $p_T$ . Left: electron channel. right: muon channel. Events are selected with the *inclusive* selection, and simulated backgrounds are normalized to luminosity.

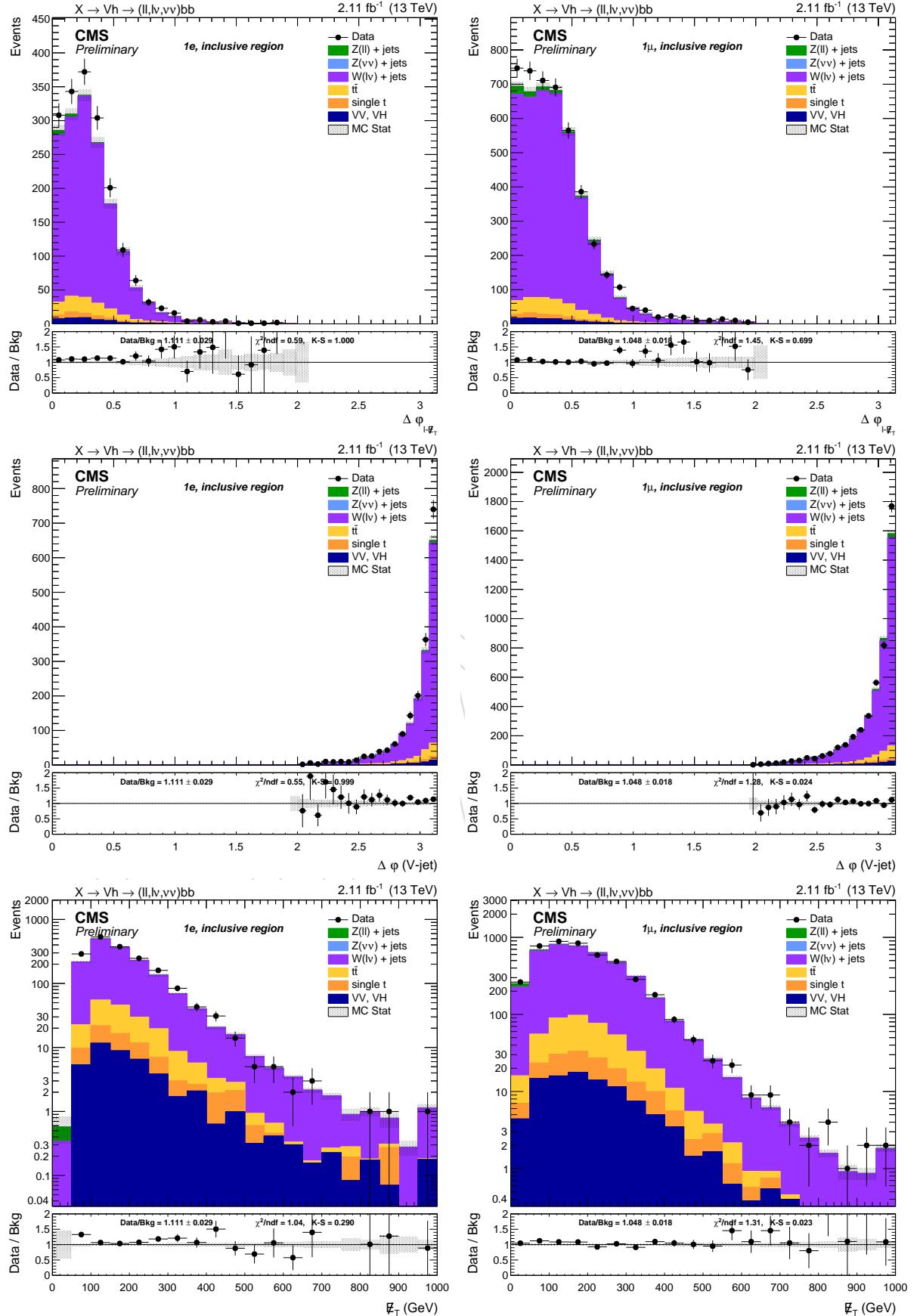


Figure 50: Top:  $\Delta\phi$  between the lepton and the  $E_T^{\text{miss}}$ . Center:  $\Delta\phi$  between the W candidate and the AK8 jet. Bottom: type-1  $E_T^{\text{miss}}$ . left: electron channel. right: muon channel. Events are selected with the *inclusive* selection, and simulated backgrounds are normalized to luminosity.

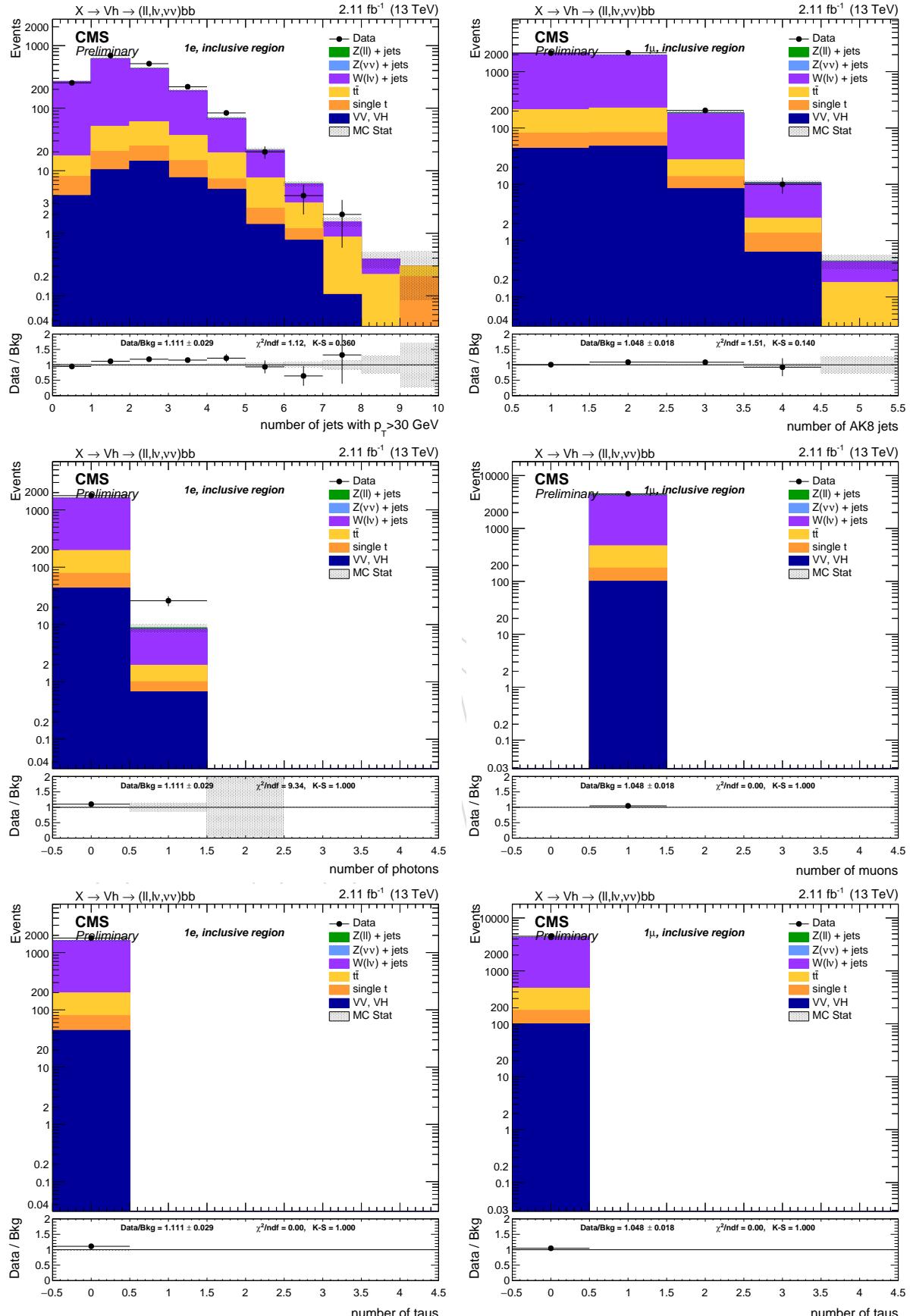


Figure 51: Top: number of AK4 jets (left) and AK8 jets (right). Center: number of photons (left) and muons (right). Bottom: number of hadronic taus. left: di-electron channel. right: di-muon channel. Events are selected with the *inclusive* selection, and simulated backgrounds are normalized to luminosity.

975 **6.3 Dilepton channel**

DRAFT

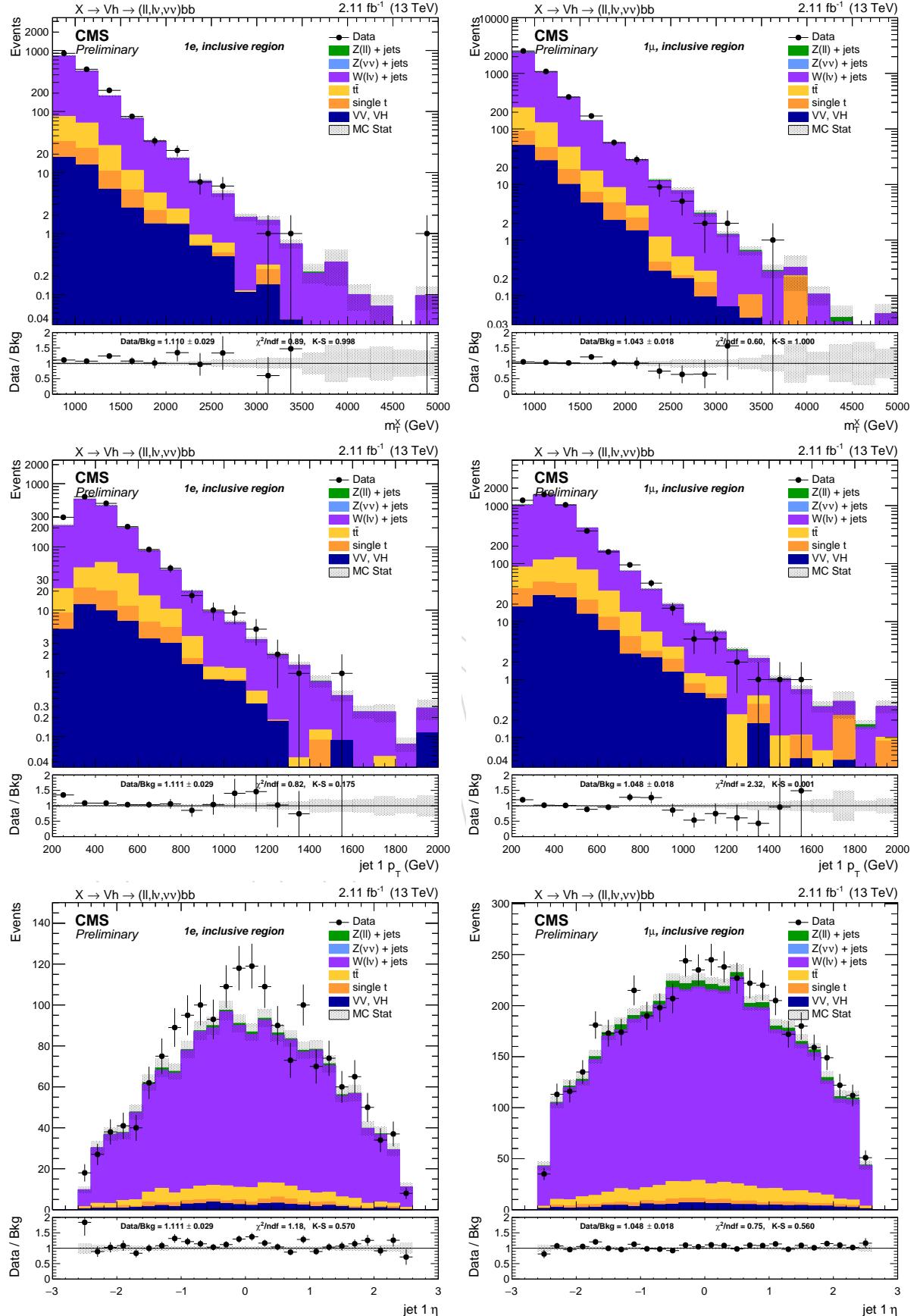


Figure 52: Top: resonance transverse mass. Center: leading AK8 jet  $p_T$ . Bottom: leading AK8 jet  $\eta$ . left: di-electron channel. right: di-muon channel. Events are selected with the *inclusive* selection, and simulated backgrounds are normalized to luminosity.

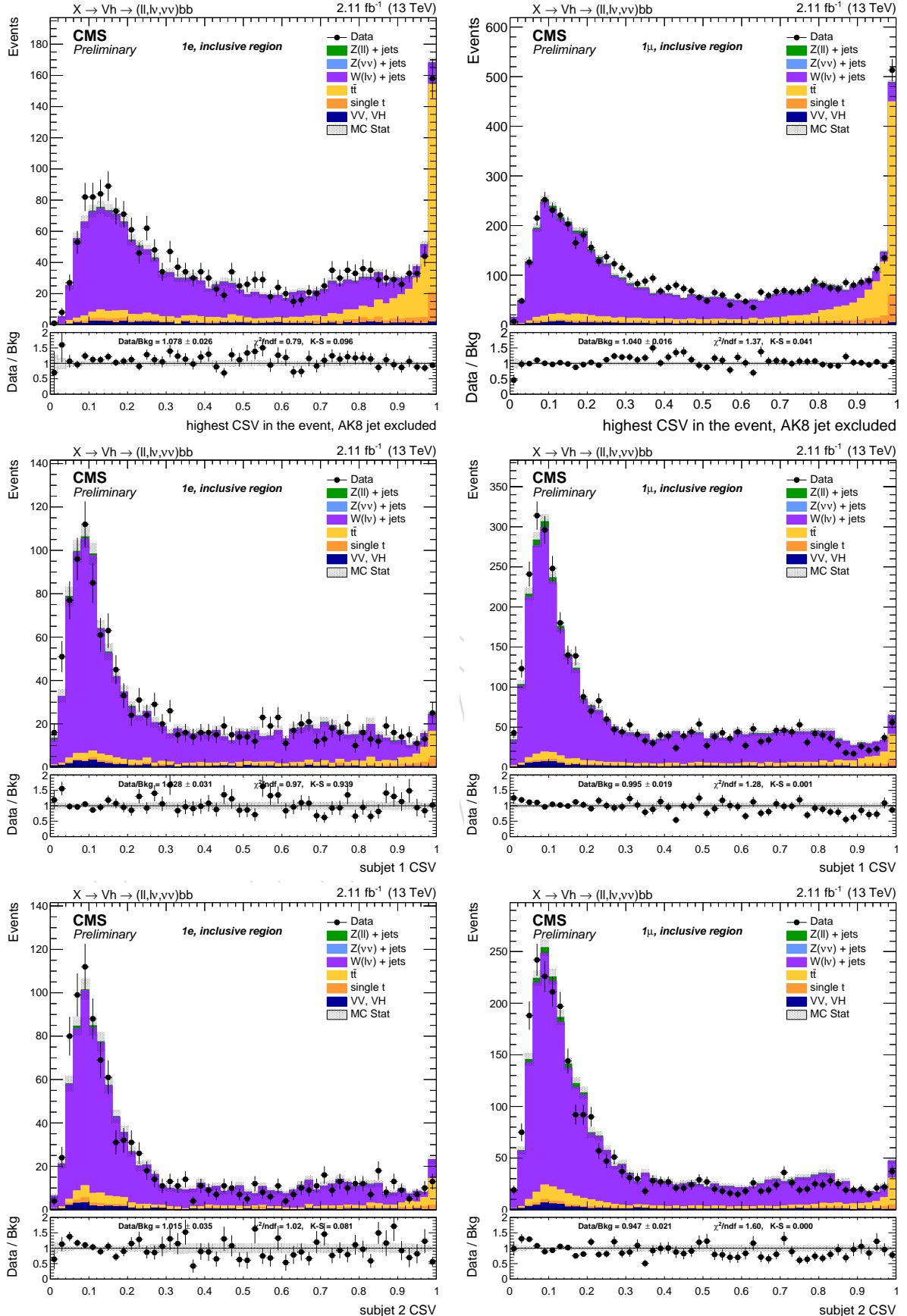


Figure 53: Top: highest CSV of the AK4 jets in the event. Center: CSV distribution of the leading AK8 sub-jet. Bottom: CSV distribution of the sub-leading AK8 sub-jet. left: di-electron channel. right: di-muon channel. Events are selected with the *inclusive* selection, and simulated backgrounds are normalized to luminosity.

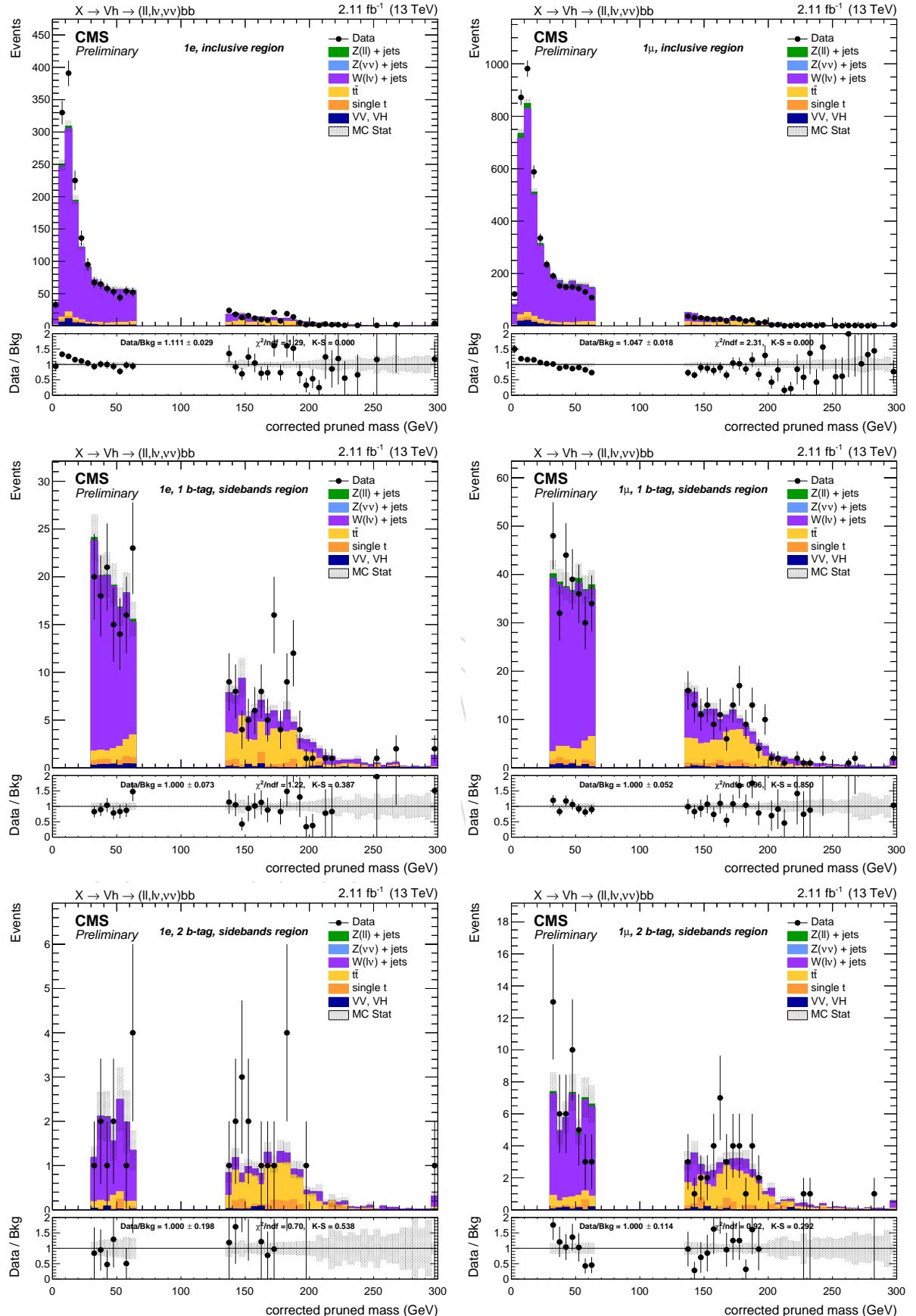


Figure 54: Leading AK8 jet pruned mass with L2L3 corrections. Top: inclusive region, no b-tagging nor jet mass selection applied. Center: jet mass sidebands in the 1 b-tag category. Bottom: mass sidebands and 2 b-tag category. right: di-muon channel. left: di-electron channel. right: di-muon channel. Scale factors are applied in the SB regions.

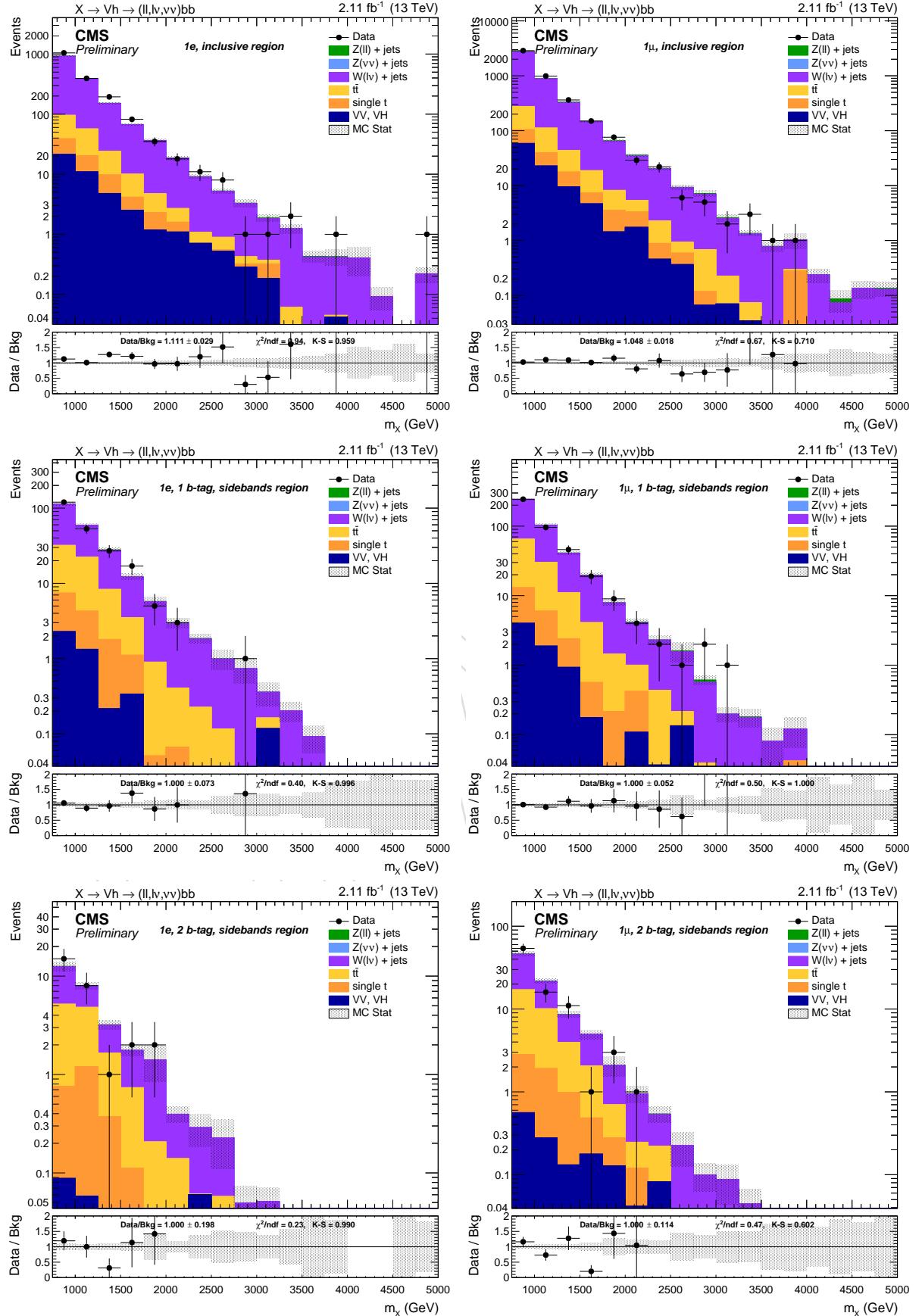


Figure 55: Resonance candidate mass. Top: inclusive region, no b-tagging nor jet mass selection applied. Center: jet mass sidebands in the 1 b-tag category. Bottom: mass sidebands and 2 b-tag category. right: di-muon channel. Scale factors are applied in the SB regions.

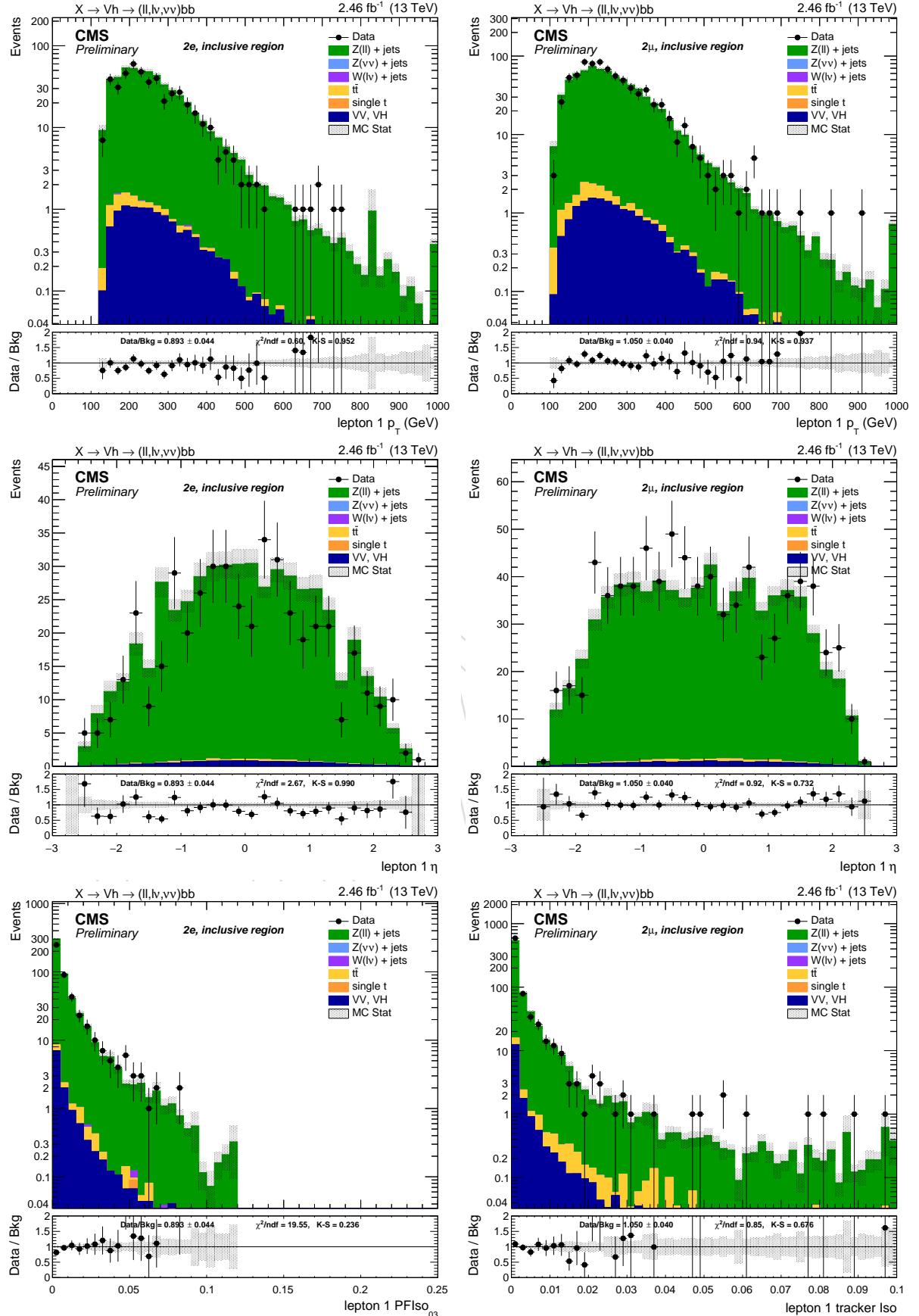


Figure 56: Leading lepton  $p_T$  (top),  $\eta$  (center), PF and tracker isolation (bottom). left: di-electron channel. right: di-muon channel. Events are selected with the *inclusive* selection, and simulated backgrounds are normalized to luminosity.

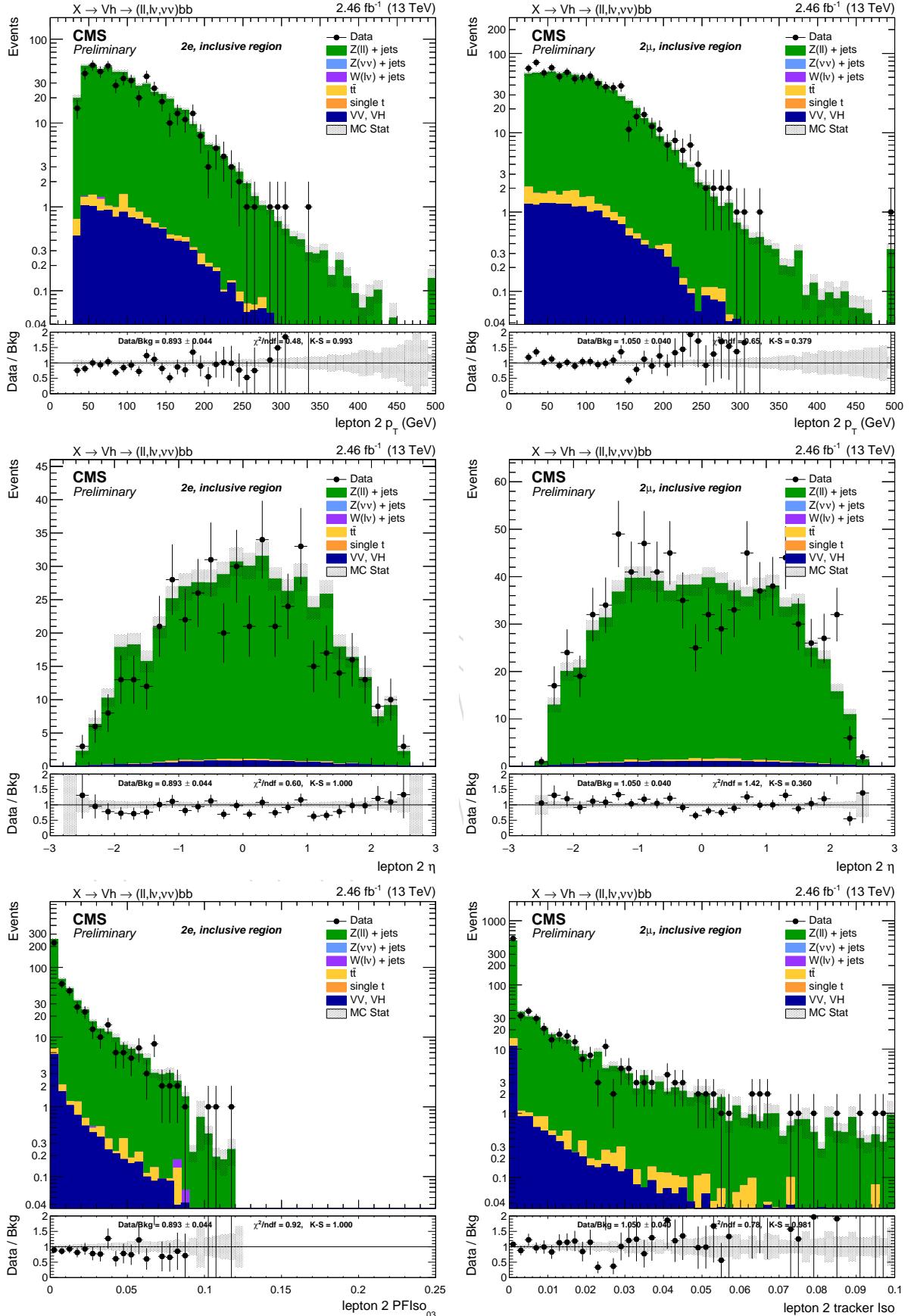


Figure 57: Sub-leading lepton  $p_T$  (top),  $\eta$  (center), PF and tracker isolation (bottom). left: di-electron channel. right: di-muon channel. Events are selected with the *inclusive* selection, and simulated backgrounds are normalized to luminosity.

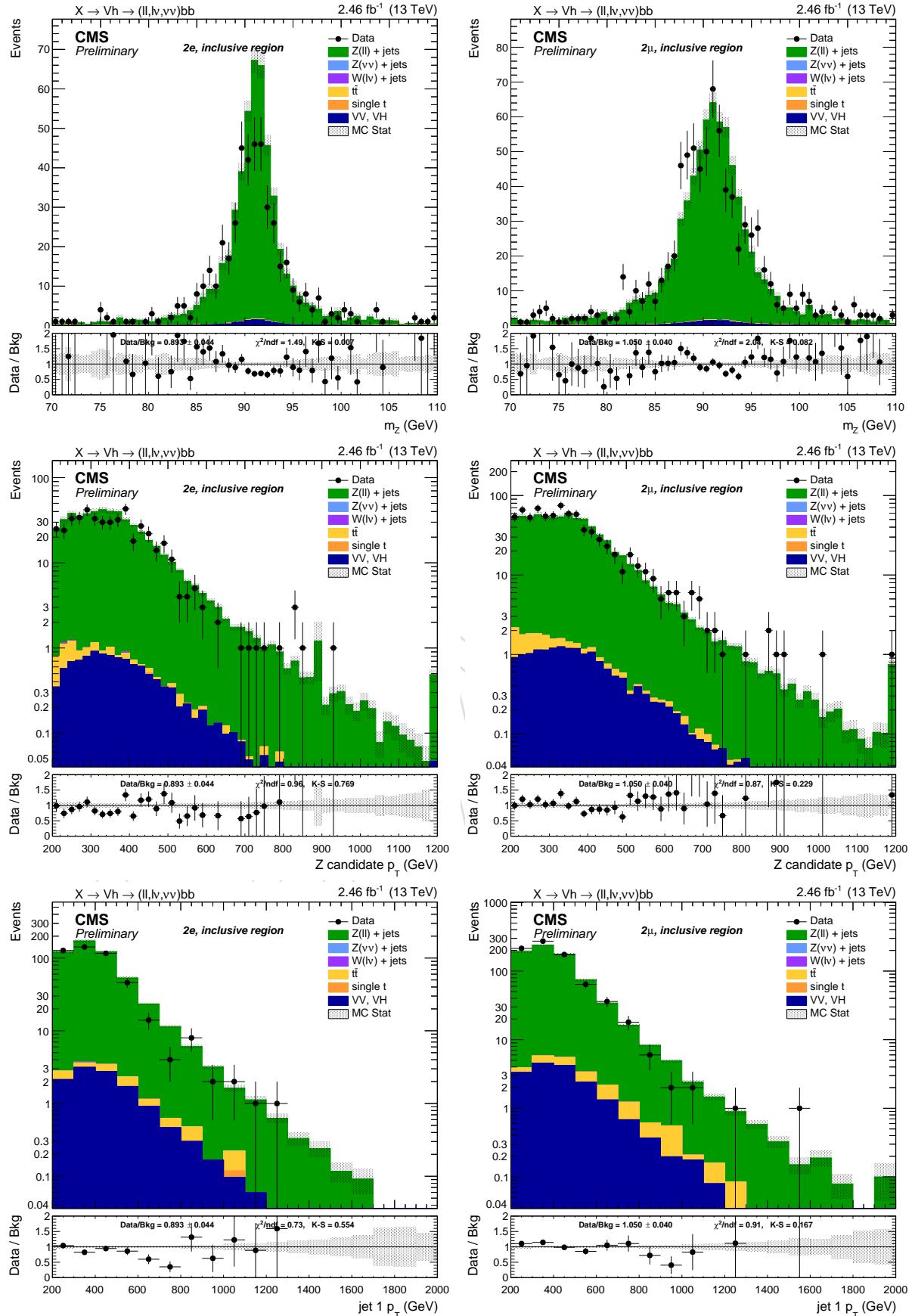


Figure 58: Top:  $Z$  candidate mass. Center:  $Z$  candidate  $p_T$ . Bottom: leading AK8 jet  $p_T$ . left: di-electron channel. right: di-muon channel. Events are selected with the *inclusive* selection, and simulated backgrounds are normalized to luminosity.

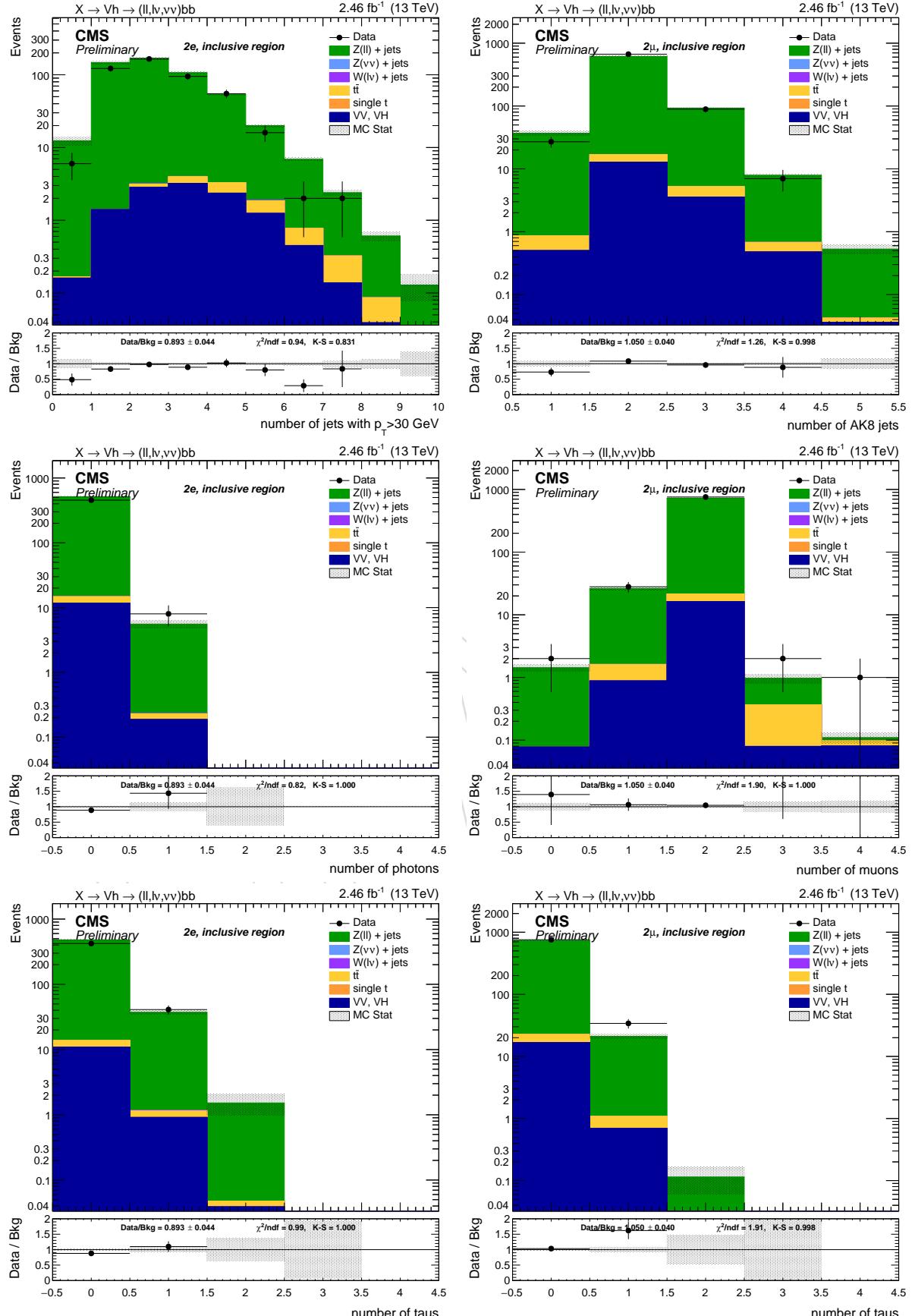


Figure 59: Top: number of AK4 jets (left) and AK8 jets (right). Center: number of photons (left) and muons (right). Bottom: number of hadronic taus. left: di-electron channel. right: di-muon channel. Events are selected with the *inclusive* selection, and simulated backgrounds are normalized to luminosity.

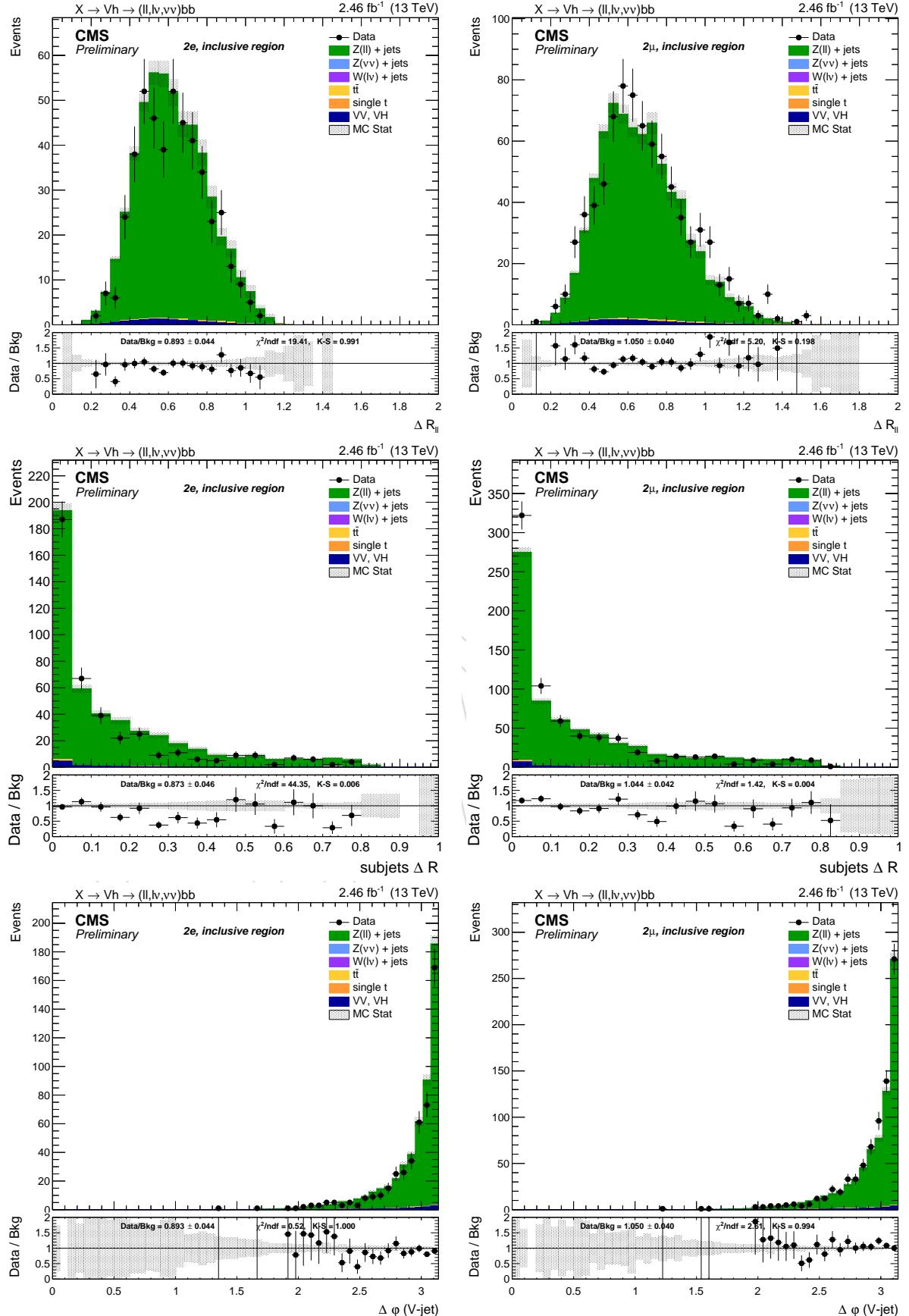


Figure 60: Top:  $\Delta R$  between the two leptons. Center:  $\Delta R$  between the AK8 sub-jets. Bottom:  $\Delta\phi$  between the  $Z$  and  $h$  candidates. left: di-electron channel. right: di-muon channel. Events are selected with the *inclusive* selection, and simulated backgrounds are normalized to luminosity.

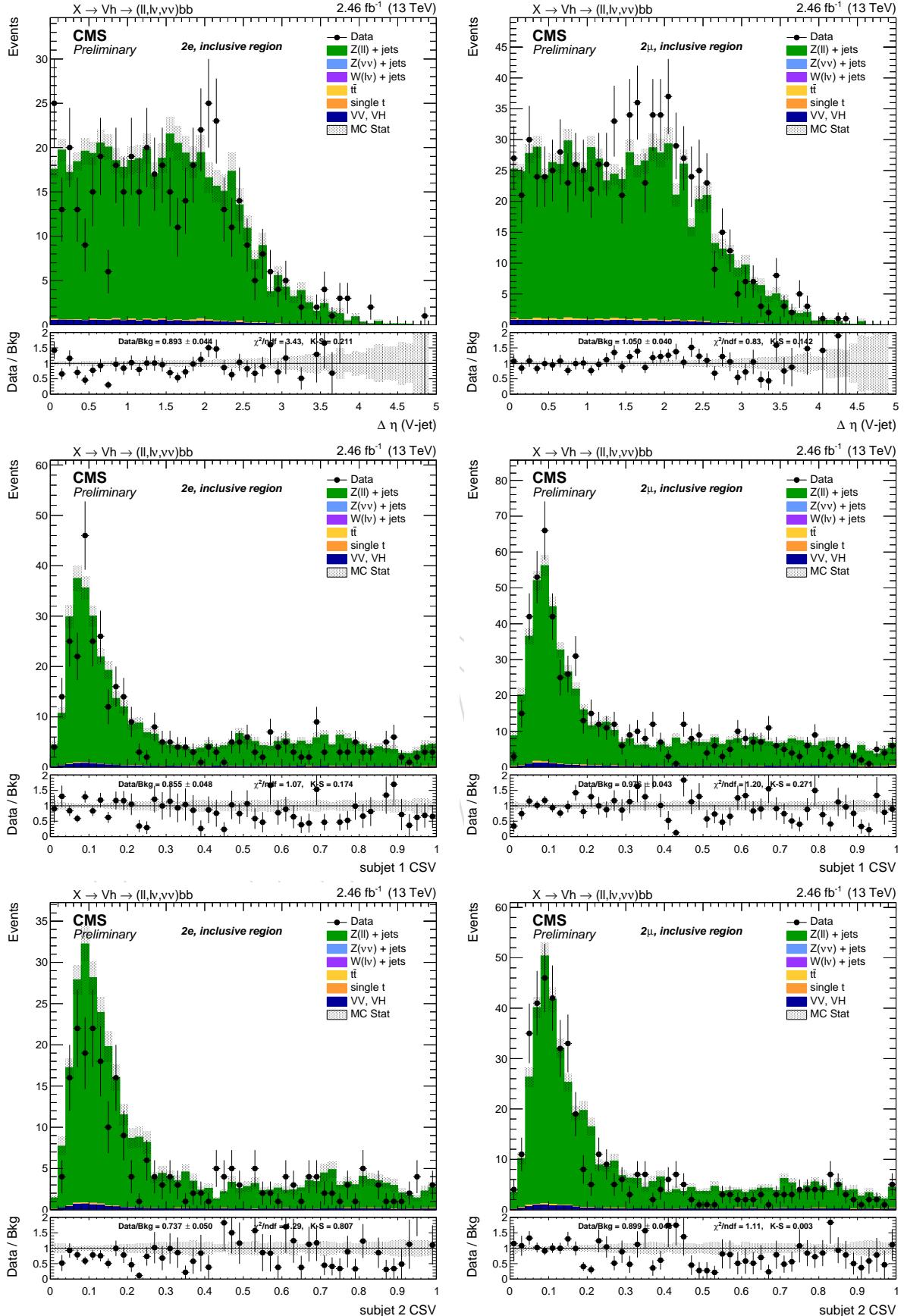


Figure 61: Top:  $\Delta\eta$  between the  $Z$  and  $h$  candidates. Center: CSV distribution of the leading AK8 sub-jet. Bottom: CSV distribution of the sub-leading AK8 sub-jet. left: di-electron channel. right: di-muon channel. Events are selected with the *inclusive* selection, and simulated backgrounds are normalized to luminosity.

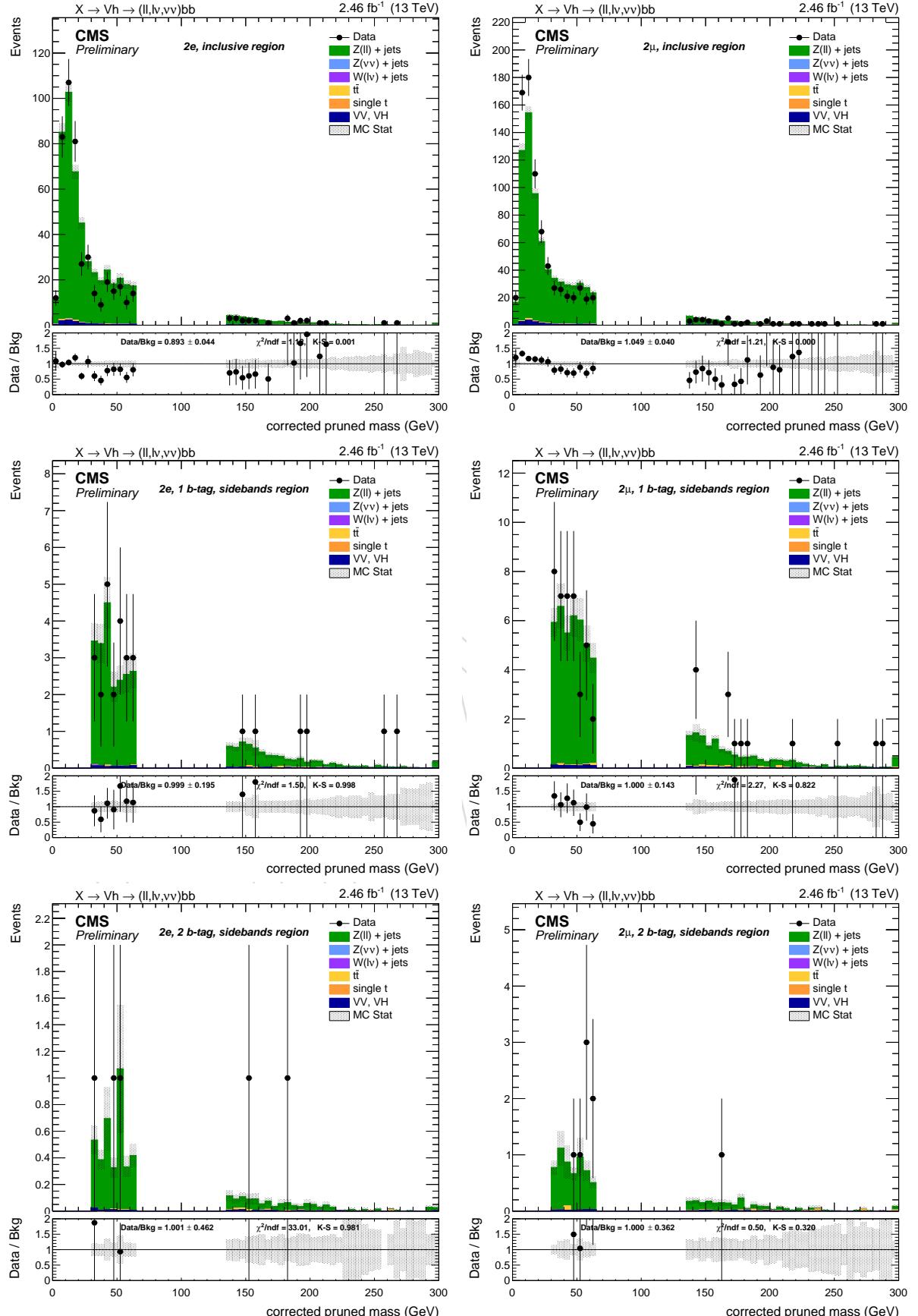


Figure 62: Leading AK8 jet pruned mass with L2L3 corrections. Top: inclusive region, no b-tagging nor jet mass selection applied. Center: jet mass sidebands in the 1 b-tag category. Bottom: mass sidebands and 2 b-tag category. Right: di-muon channel. Left: di-electron channel. Right: di-muon channel. Scale factors are applied in the SB regions.

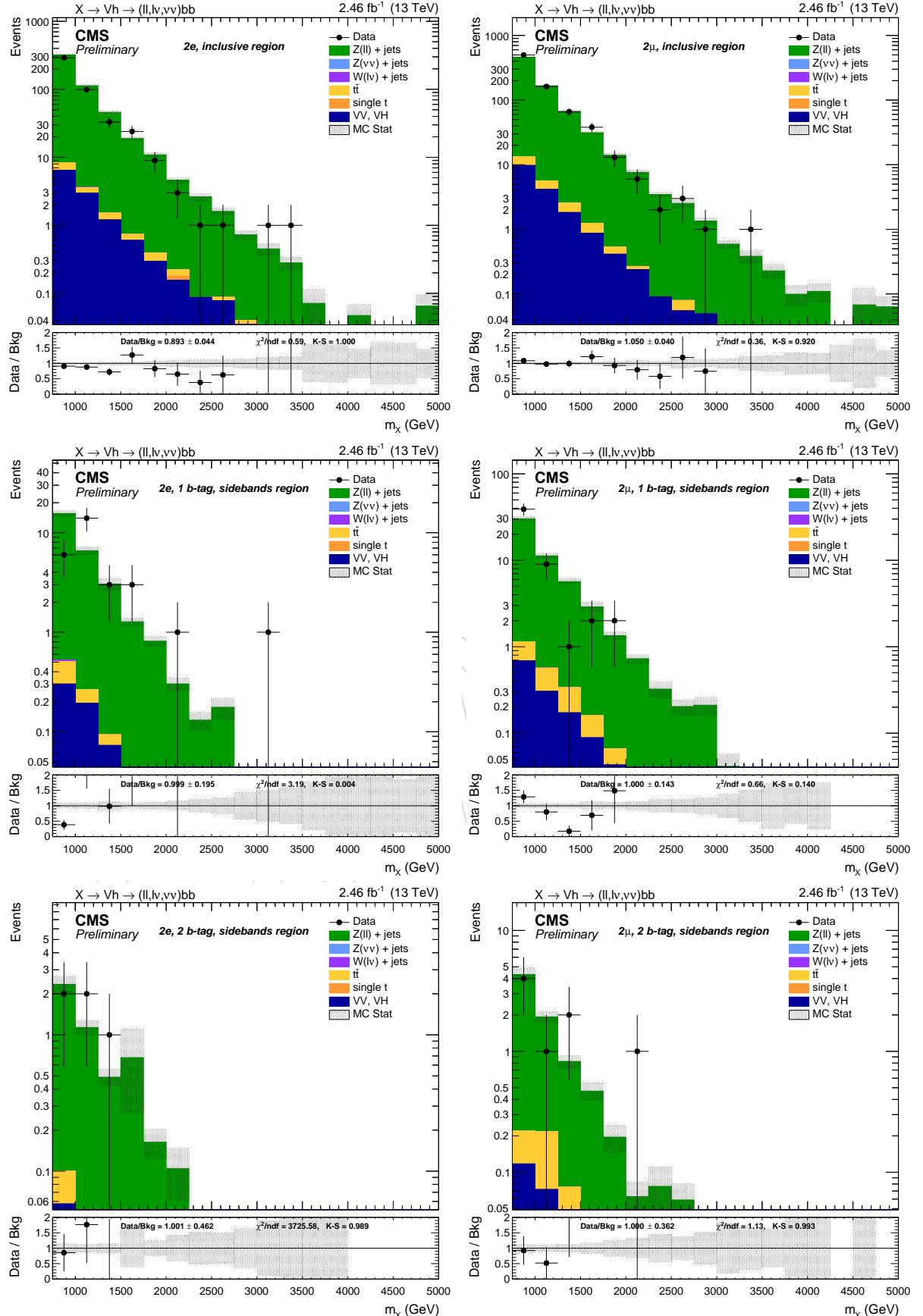


Figure 63: Resonance candidate mass. Top: inclusive region, no b-tagging nor jet mass selection applied. Center: jet mass sidebands in the 1 b-tag category. Bottom: mass sidebands and 2 b-tag category. left: di-electron channel. right: di-muon channel. Scale factors are applied in the SB regions.

## 976 7 Top control regions

977 In the present analysis, there is no dominant background in all the channels. The single lepton and 0-lepton channels, after the b-tagging requirement, contain a significant amount of  
 978  $t\bar{t}$  events, which may eventually be larger than the “main” background, the  $W + \text{jets}$ , when a  
 979 double b-tagging requirement is applied.

981 Because of the possibility to extract only one data-driven background, an alternative way to  
 982 evaluate and estimate the other backgrounds should be considered. The diboson contamination  
 983 and cross section is small, and there is no handle to evaluate the  $VV$  normalization and  
 984 shape from data, at least with the current luminosity. However, the  $t\bar{t}$  and single top produc-  
 985 tion are sizable, and they can be checked directly on data.

986 Appropriate Top control regions (CR) are defined separately for 0-lepton and single-lepton  
 987 channel, and also 1 and 2 b-tags, selecting an almost pure  $t\bar{t}$  sample. Electron and muon cate-  
 988 gories are merged together to have more statistical power. The 1 and 2 b-tag categories are still  
 989 separated because of the quite large difference in the scale factors between them. In this case,  
 990 the dependence on the sub-jet b-tagging discriminator in simulation is reduced. The only cut  
 991 to be modified is the anti b-tag veto, which is reversed and tightened in order to have an high  
 992  $t\bar{t}$  purity. The four Top control regions are then selected with the same criteria as the respective  
 993 sidebands regions, with one additional b-tagged AK4 jet in the event passing the CSV *tight*  
 994 working point. Scale factors for b-tagging are applied also in this case.

995 Figure 64-65 report the data and simulation distributions in the four Top control regions. Multi-  
 996 plicative scale factors for  $t\bar{t}$  and single top are derived for each category, and applied consis-  
 997 tently in the background prediction method.

category		Top SF	stat.	syst.
1 b-tag	$1\ell$	0.82	$\pm 0.03$	$\pm 0.04$
	$0\ell$	0.85	$\pm 0.06$	$\pm 0.04$
2 b-tag	$1\ell$	0.83	$\pm 0.07$	$\pm 0.04$
	$0\ell$	0.54	$\pm 0.13$	$\pm 0.02$

Table 14: Top normalization scale factors, reported independently for each channel. Electron and muon categories are merged.

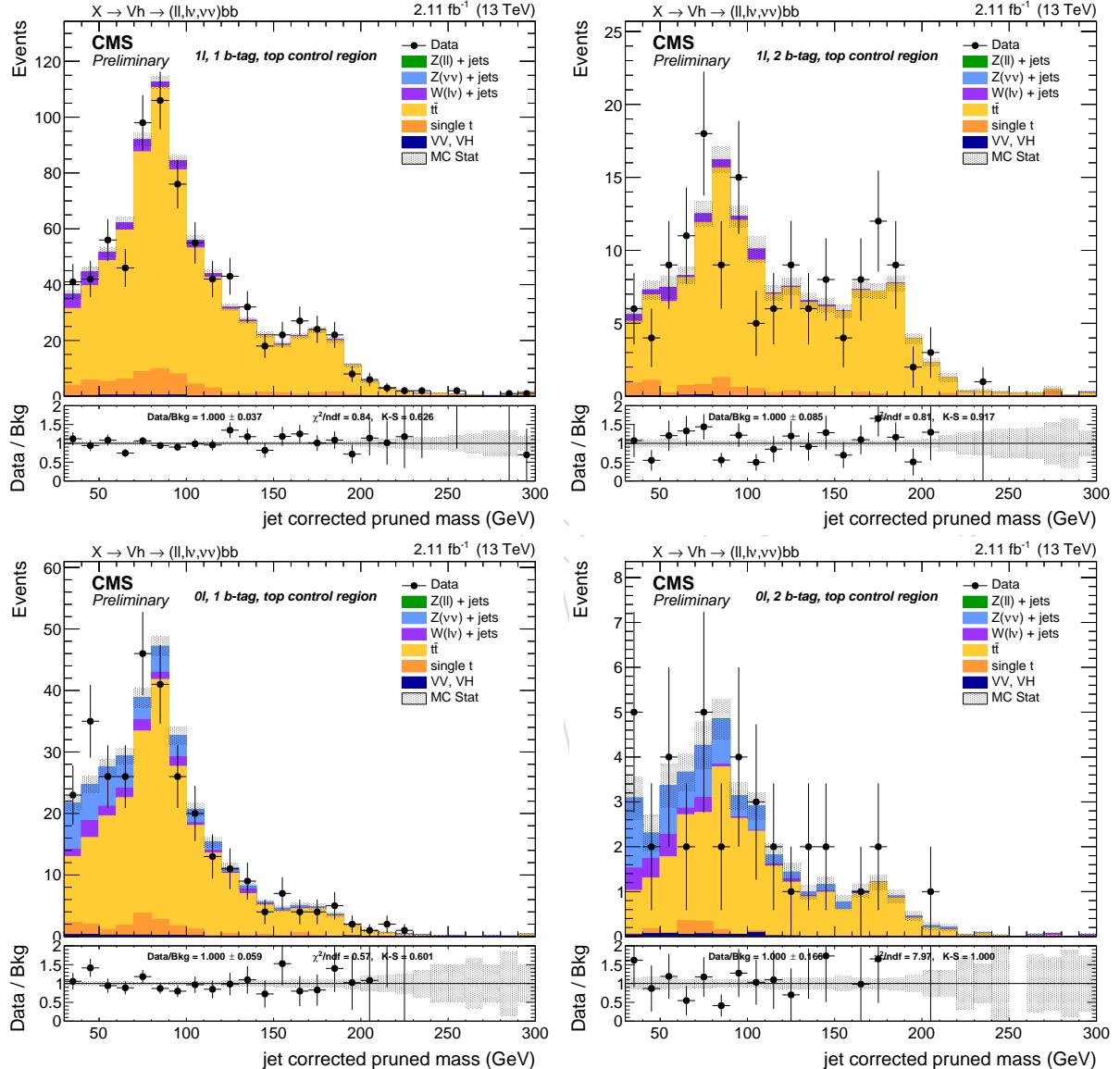


Figure 64: Jet mass distribution in the Top control regions in the jet mass sidebands.

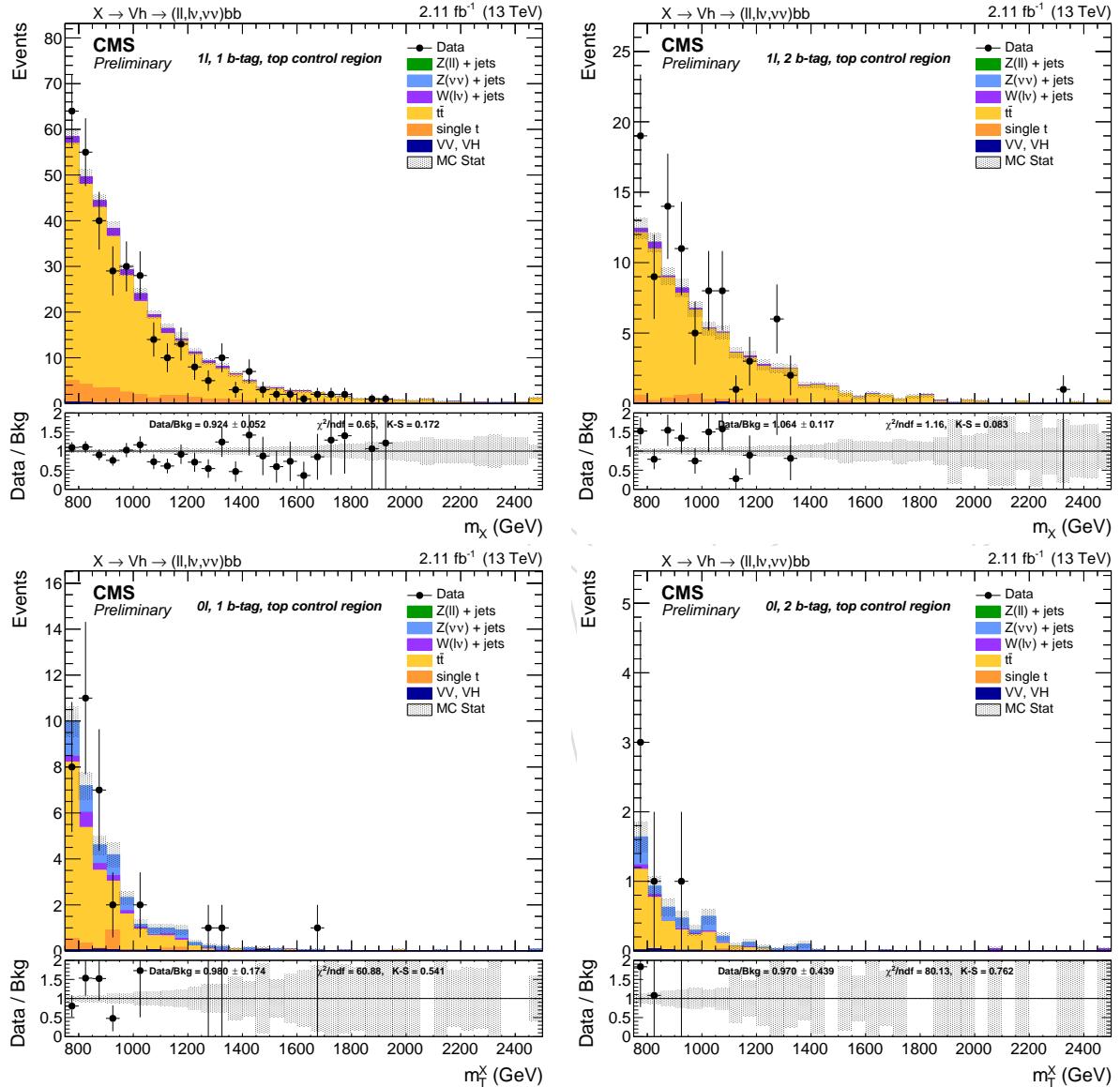


Figure 65:  $m_X$  distribution in the Top control regions in the jet mass sidebands.

## 998 8 Alpha ratio background prediction

999 The goal of the analysis is to look for localised excesses in the  $m_X$  spectrum. The *alpha method*  
1000 was the background estimation method used for the Run 1 diboson searches, and it has been  
1001 introduced to be less dependent on the MC simulation for the background  $m_X$  estimation, due  
1002 to the many sources of systematic uncertainties that are hard to understand and control. Two  
1003 exclusive regions, named *signal region* (SR) and *sidebands region* (SB), are defined in order to  
1004 select a signal enriched or signal depleted phase space, respectively. First, the background  
1005 normalization is extracted from data in the SB. Then, the alpha method extracts a predicted  
1006 shape from the data in the SB to the SR using a transfer function (the  $\alpha$  function) derived from  
1007 simulation. The method relies on the assumption that the correlation between  $m_X$  and the  
1008 pruned jet mass is reasonably well reproduced by the MC. The  $\alpha$  ratio is deemed to be more  
1009 trustworthy since many systematic uncertainties would approximately cancel in the ratio.

The  $\alpha$  function is defined as the ratio of the two functions describing the simulated  $m_X$  shape  
in the SR and SB:

$$\alpha(m_X) = \frac{N_{SR}^{MC,bkg}(m_X)}{N_{SB}^{MC,bkg}(m_X)}$$

and the background distribution in the SR is thus estimated as the product of  $\alpha(m_X)$  with the  
shape in the data SB:

$$N_{bkg}(m_X) = N_{SB}(m_X) \times \alpha(m_X)$$

1010 Notice that in the above description does not include the definition of the SB and SR. Ideally,  
1011 the best choice would be a variable such that the distribution of  $m_X$  in the Signal and Sideband  
1012 regions are similar. In this analysis, the pruned corrected jet mass  $m_j$  (see Section 3.7.1) is  
1013 chosen as the control variable, and the cut values are those reported in Section 4.4.1. All the  
1014 selections used in the  $\alpha$  method background prediction are the same reported in Section 5. In  
1015 order to respect the blinding policy of the diboson VV searches, the Z/W mass window is not  
1016 used neither for the background estimation (it can be potentially contaminated by a signal) nor  
1017 as signal region.

1018 In a real case scenario, the background is not purely composed of one single process neither in  
1019 the SR nor in the SB. The background composition is assumed to be dominated by one single  
1020 process (V+jets) whose modeling in simulation is considered not to be trustworthy. Other sub-  
1021 dominant backgrounds ( $t\bar{t}$  including single top and V V) generally have smaller contributions,  
1022 and are considered quite well understood and modeled by MC generators. However, due to  
1023 the b-tagging and lepton selections, in some channels the  $t\bar{t}$  can be of the same order of mag-  
1024 nitude or even larger than the main background. The treatment of sub-dominant backgrounds  
1025 thus require a dedicated treatment.

1026 Top control regions are defined separately for 0-lepton and single-lepton channel, and also 1  
1027 and 2 b-tags, selecting an almost pure  $t\bar{t}$  sample (Section 7). In the dilepton categories, the  $t\bar{t}$   
1028 contamination is so small that an additional CR would be useless due to the lack of events,  
1029 and the normalization in this case is taken form simulation. The normalization of the  $t\bar{t}$  process  
1030 is then fixed in the CR, and a multiplicative scale factor is applied to the normalization in  
1031 the subsequent steps of the analysis. The shape and normalization of the diboson production,  
1032 instead, is taken from the simulation. The sub-dominant backgrounds are then subtracted from  
1033 the V+jets contribution when fitting this template to data.

1034 A different background prediction is derived for each category separately, thus dividing ele-  
1035 cron and muon channels, and single and double b-tag categories, in order to reduce systematic  
1036 uncertainty due to leptonic triggers, identification, isolation, and b-tagging efficiencies.

1037 **8.1 Background normalization**

The first step in the background prediction consists in a proper estimation of the background normalization. The three main backgrounds ( $V + \text{jets}$ ,  $t\bar{t}$  and single top, and  $VV$  including  $VH$ ) are considered separately due to the different shape in the jet mass distribution. The three contributions are described with functional forms determined by fits on the simulated backgrounds. The number of expected events in the SR is extracted through the same equation:

$$N_{SR}^{data} = \left[ N_{SB}^{data} - N_{SB}^{Top} - N_{SB}^{VV} \right] \times \left[ \frac{N_{SR}^{Vjet}}{N_{SB}^{Vjet}} \right] + N_{SR}^{Top} + N_{SR}^{VV}$$

1038 where in this case  $N$  are the number of events, and not functions.

1039 The empirical functional forms for each background are chosen reflect the physics properties  
 1040 of the samples. The  $V + \text{jets}$  background has a smoothly falling background with no peaks. On  
 1041 top of the jet mass spectrum, the  $VV$  has one peak, corresponding to the vector boson hadronic  
 1042 decay reconstruction, and possibly a second one due to the presence of the Higgs. In certain  
 1043 channels, the latter is not visible due to the smaller cross section with respect to the former.  
 1044 The  $t\bar{t}$  and single top backgrounds are considered together, because they both have two peaks  
 1045 corresponding to the  $W \rightarrow jj$  decays and all-hadronic top quark decays  $t \rightarrow Wb \rightarrow jjb$ . In some  
 1046 channels, and particularly those with 2 b-tags, the AK8 jets clusters both the b-quarks from the  
 1047 top decay, so the  $W \rightarrow jj$  is not clearly visible. The functional forms chosen to build the jet mass  
 1048 templates are:

**Exp:** an exponential function:

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

**Pol:** a third order polynomial:

$$F_{Pol}(x) = a_0 \cdot x + a_2 \cdot x^2 + a_3 \cdot x^3$$

**ExpErf:** an “error function”, that consists of an exponential multiplied by an Erf:

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

**Gaus:** one gaussian:

$$F_{Gaus}(x) = e^{2(x-a)^2/b}$$

**Gaus2:** two gaussians:

$$F_{Gaus2}(x) = f_0 \cdot e^{2(x-a)^2/b} + (1 - f_0) \cdot e^{2(x-c)^2/d}$$

**Gaus3:** three gaussians:

$$F_{Gaus3}(x) = f_0 \cdot e^{2(x-a)^2/b} + f_1 \cdot e^{2(x-c)^2/d} + (1 - f_0 - f_1) \cdot e^{2(x-e)^2/g}$$

**ExpGaus:** an exponential plus one gaussian:

$$F_{ExpGaus}(x) = f_0 \cdot e^{ax} + (1 - f_0) \cdot e^{2(x-b)^2/c}$$

**ExpGaus2:** an exponential plus two gaussians:

$$F_{ExpGaus2}(x) = f_0 \cdot e^{ax} + f_1 \cdot e^{2(x-b)^2/c} + (1 - f_0 - f_1) \cdot e^{2(x-d)^2/e}$$

**ErfExpGaus:** an error function plus one gaussian:

$$F_{\text{ExpGaus}}(x) = f_0 \cdot F_{\text{ErfExp}}(x, a, b, c) + (1 - f_0) \cdot e^{2(x-d)^2/e}$$

**ErfExpGaus2:** an error function plus two gaussians:

$$F_{\text{ExpGaus2}}(x) = f_0 \cdot F_{\text{ErfExp}}(x, a, b, c) + f_1 \cdot e^{2(x-d)^2/e} + (1 - f_0 - f_1) \cdot e^{2(x-f)^2/g}$$

1049 The choice of the functions is channel-dependent, and it depends on the background shape and  
 1050 the available statistics, and is summarized in Table 15.

	category	V +jets	alt. V +jets	t̄t	VV
1 b-tag	0ℓ	ErfExp	Pol	Gaus2	ExpGaus
	1e	ErfExp	Pol	Gaus3	ExpGaus
	1μ	ErfExp	Pol	Gaus3	ExpGaus
	2e	ErfExp	Pol	Gaus	ExpGaus2
	2μ	ErfExp	Pol	Gaus	ExpGaus2
2 b-tag	0ℓ	Exp	Pol	Gaus2	ExpGaus
	1e	Exp	Pol	Gaus2	ExpGaus
	1μ	Exp	Pol	Gaus2	ExpGaus
	2e	Exp	Pol	Gaus	ExpGaus2
	2μ	Exp	Pol	Gaus	ExpGaus2

Table 15: Chosen functions to fit the jet mass distribution for each channel.

1051 The following plots (Figure 66-80) show the mass fits to the jet mass in the different channels.  
 1052 Table 16 summarizes the expected background yield in the signal region.

	category	Expected	Stat.	Syst.	Alt. function	Observed
1 b-tag	0ℓ	49.741	±6.402	±0.165	±5.256	47
	1e	77.158	±10.204	±0.287	±14.012	57
	1μ	129.182	±7.863	±0.400	±12.527	119
	2e	4.460	±1.076	±0.036	±0.987	7
	2μ	12.510	±1.682	±0.055	±0.739	19
2 b-tag	0ℓ	7.307	±1.279	±0.091	±1.115	6
	1e	6.689	±1.102	±0.258	±0.399	7
	1μ	20.432	±2.196	±0.343	±1.969	14
	2e	1.030	±0.371	±0.016	±0.066	2
	2μ	1.759	±0.613	±0.010	±0.325	1

Table 16: Expected background yield in the SR ( $105 < m_j < 135$  GeV) and relative uncertainties.

## 8 Alpha ratio background prediction

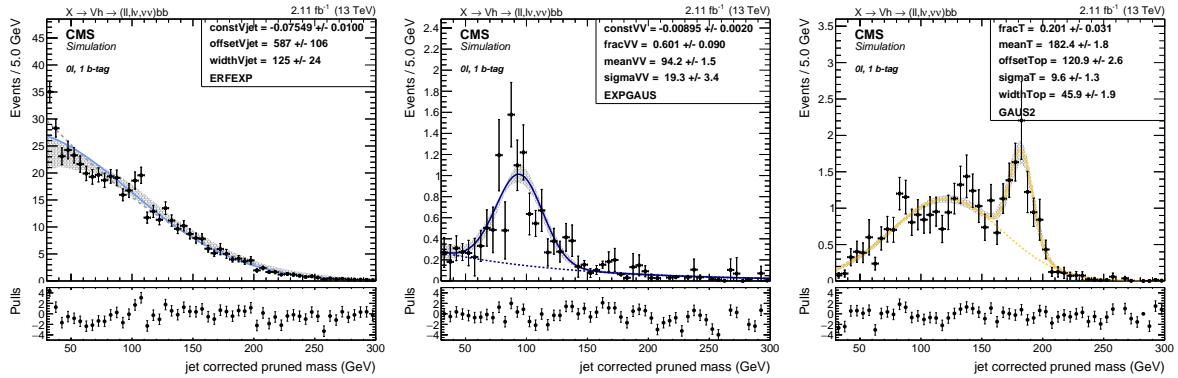


Figure 66: Fit to the simulated  $m_j$  in the 0 lepton, 1 b-tag category for the three backgrounds: V+jets (left), VV (center), Top (right).

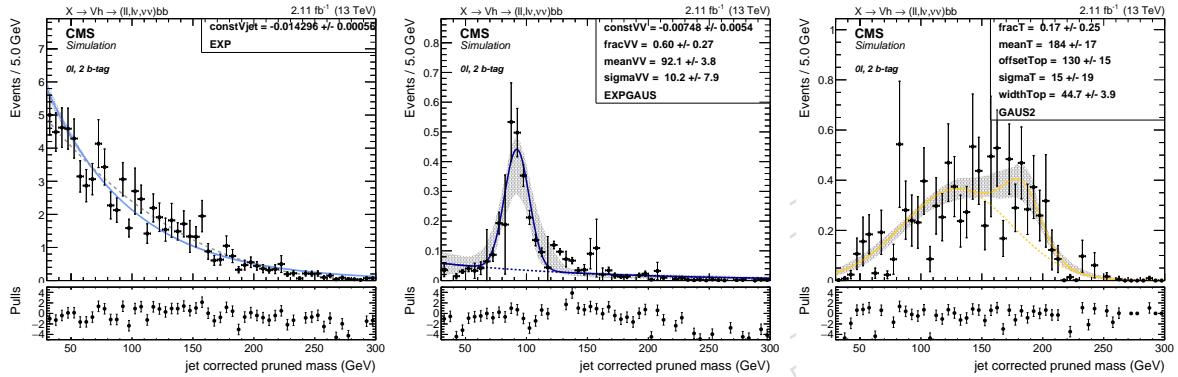


Figure 67: Fit to the simulated  $m_j$  in the 0 lepton, 2 b-tag category for the three backgrounds: V+jets (left), VV (center), Top (right).

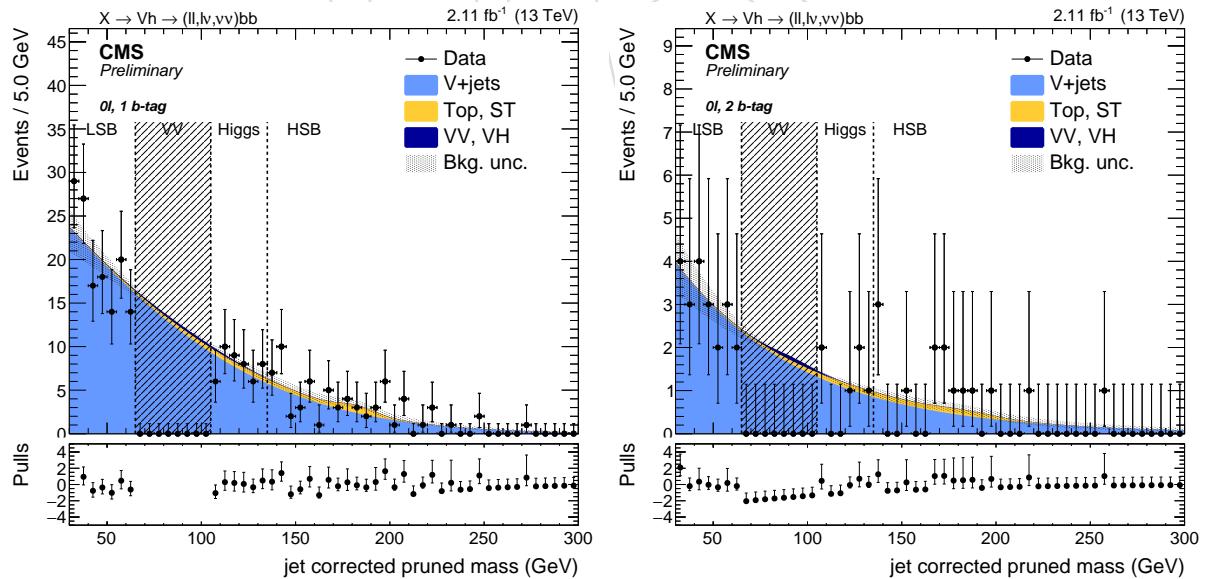


Figure 68: Fit to data  $m_j$  in the 0 lepton, 1 b-tag (left) and 2 b-tag category (right).

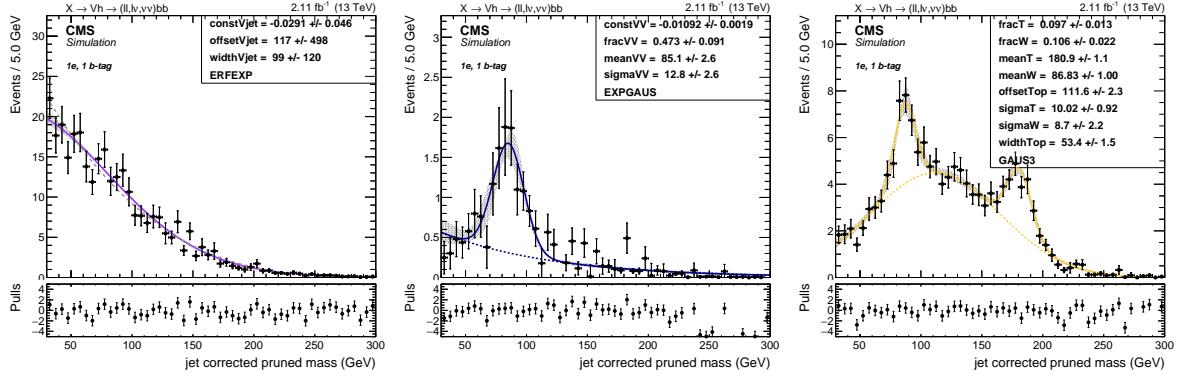


Figure 69: Fit to the simulated  $m_j$  in the 1 electron, 1 b-tag category for the three backgrounds: V +jets (left), VV (center), Top (right).

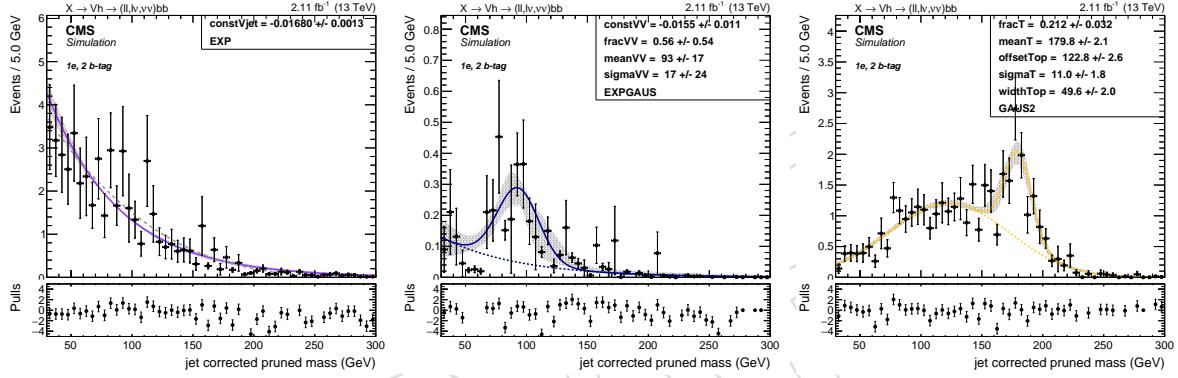


Figure 70: Fit to the simulated  $m_j$  in the 1 electron, 2 b-tag category for the three backgrounds: V +jets (left), VV (center), Top (right).

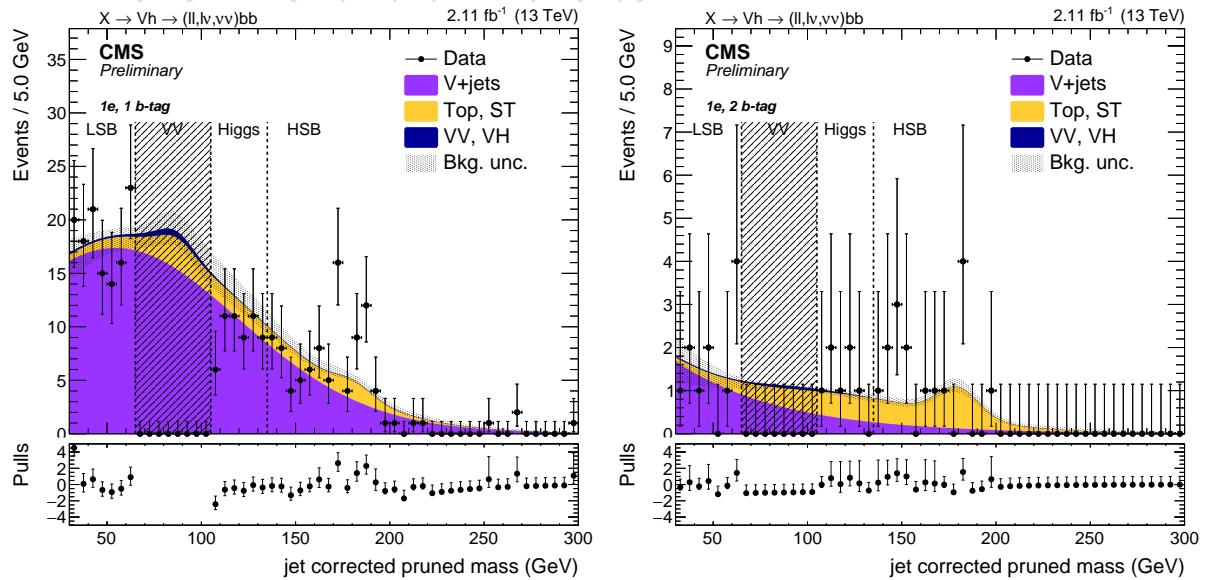


Figure 71: Fit to data  $m_j$  in the 1 electron, 1 b-tag (left) and 2 b-tag category (right).

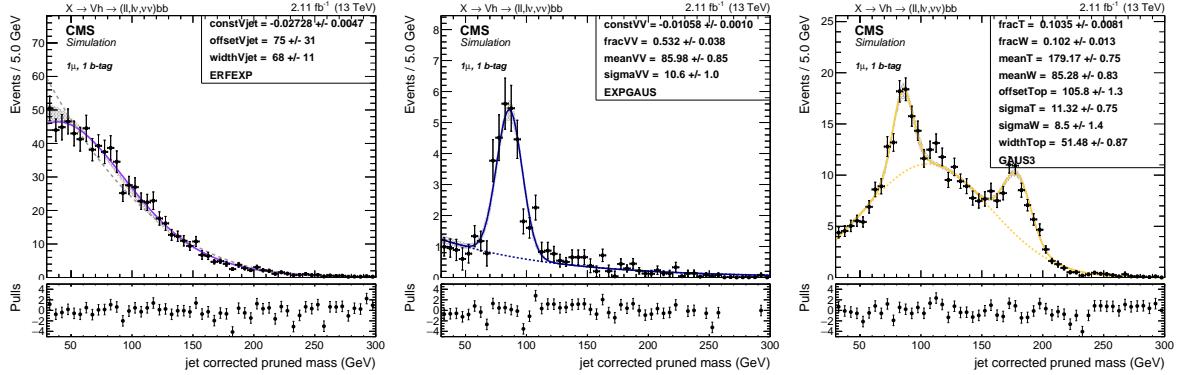


Figure 72: Fit to the simulated  $m_j$  in the 1 muon, 1 b-tag category for the three backgrounds: V+jets (left), VV (center), Top (right).

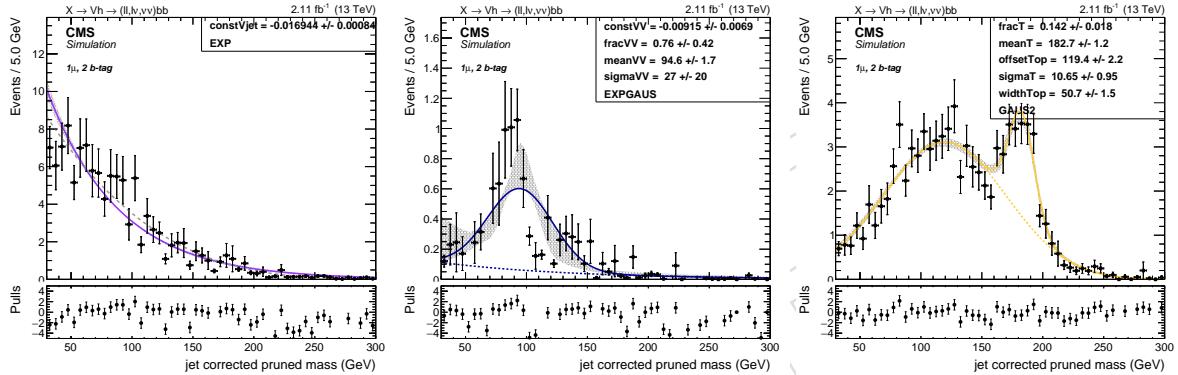


Figure 73: Fit to the simulated  $m_j$  in the 1 muon, 2 b-tag category for the three backgrounds: V+jets (left), VV (center), Top (right).

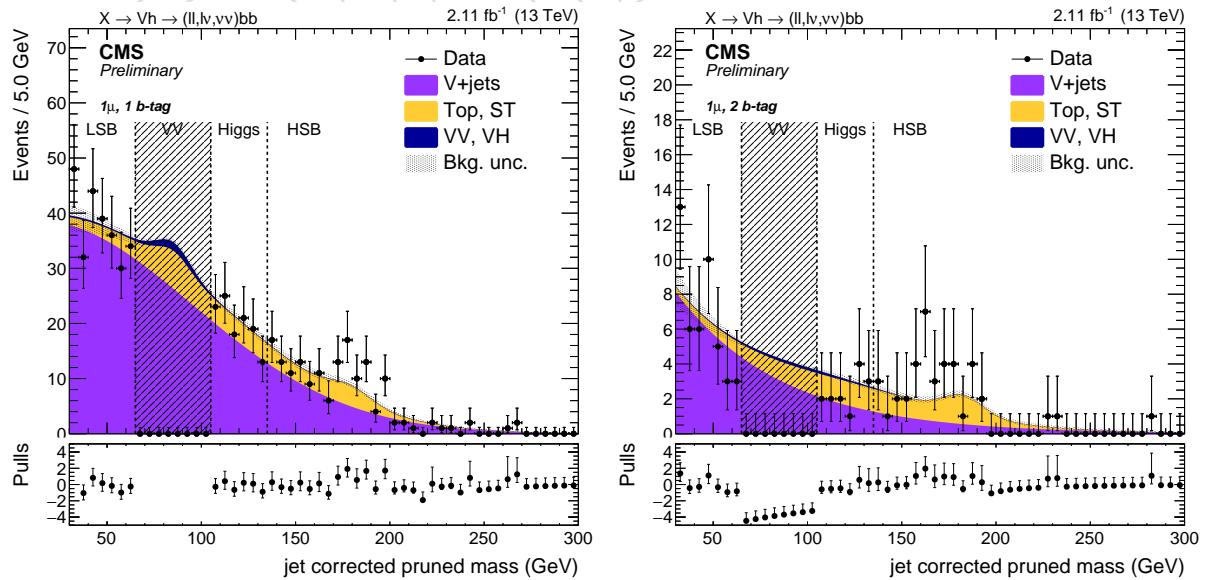


Figure 74: Fit to data  $m_j$  in the 1 muon, 1 b-tag (left) and 2 b-tag category (right).

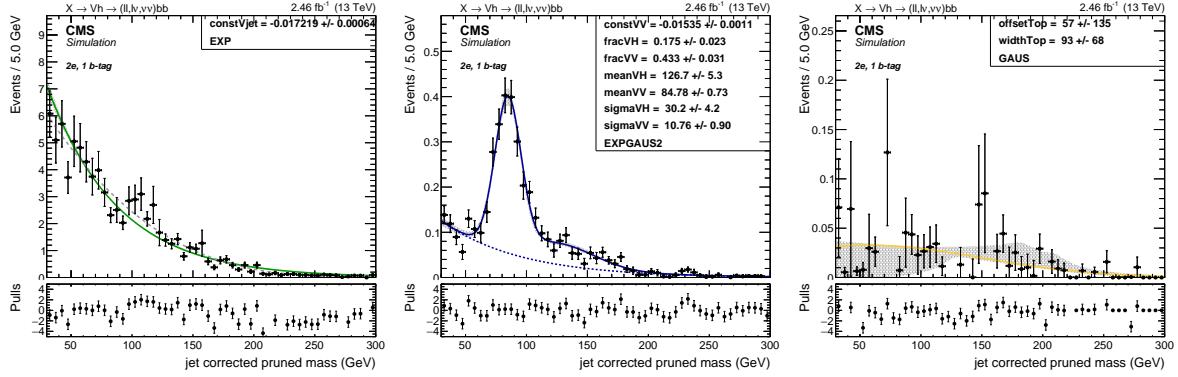


Figure 75: Fit to the simulated  $m_j$  in the 2 electrons, 1 b-tag category for the three backgrounds: V + jets (left), VV (center), Top (right).

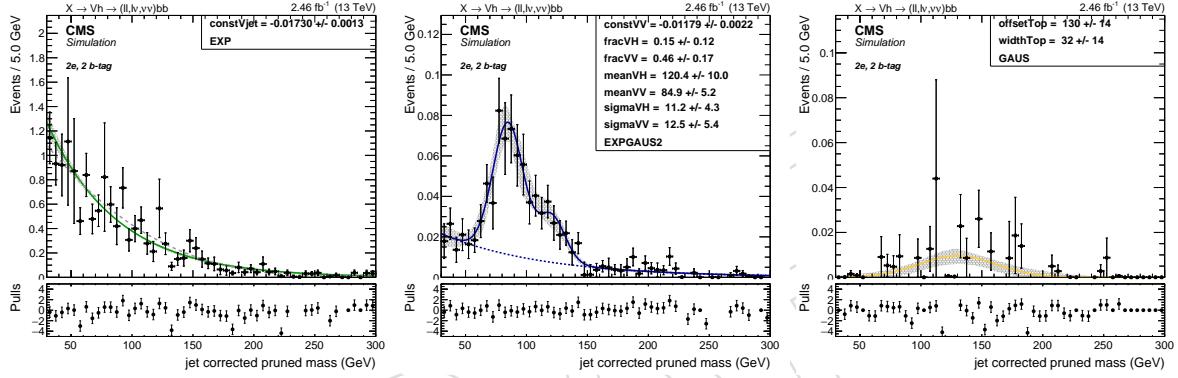


Figure 76: Fit to the simulated  $m_j$  in the 2 electrons, 2 b-tag category for the three backgrounds: V + jets (left), VV (center), Top (right).

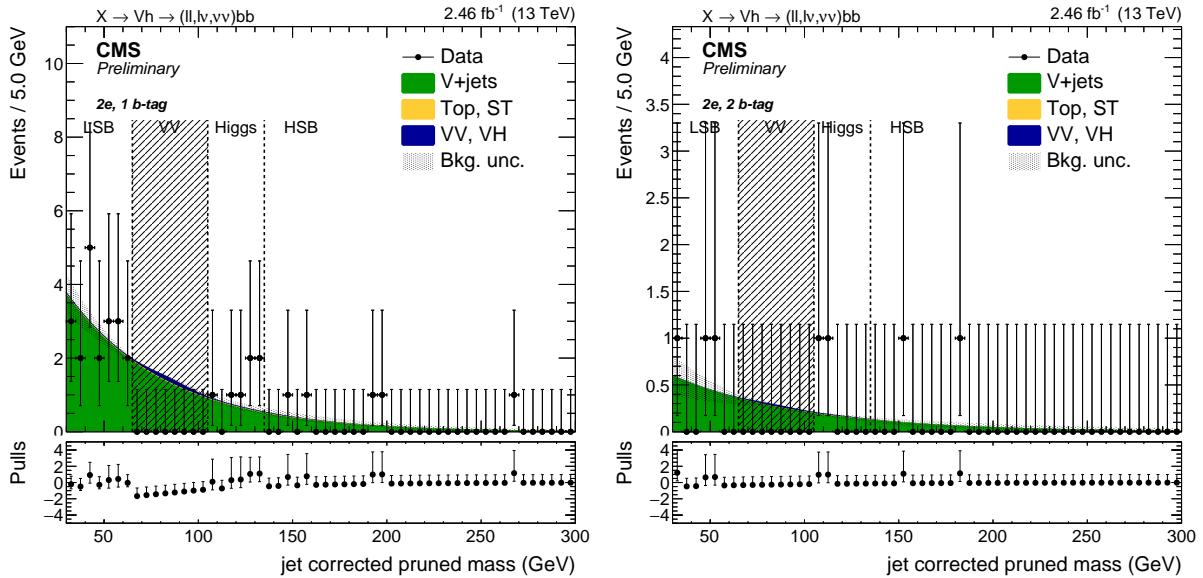


Figure 77: Fit to data  $m_j$  in the 2 electrons, 1 b-tag (left) and 2 b-tag category (right).

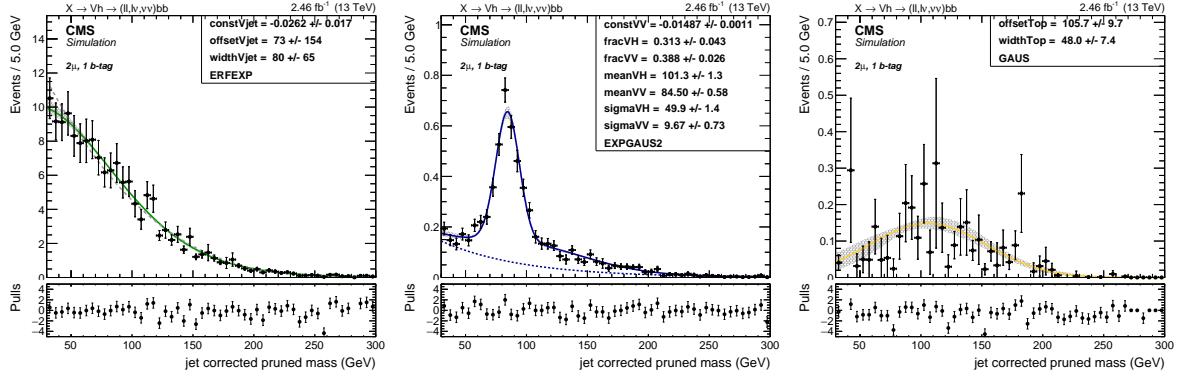


Figure 78: Fit to the simulated  $m_j$  in the 2 muons, 1 b-tag category for the three backgrounds:  $V + \text{jets}$  (left),  $VV$  (center),  $\text{Top}$  (right).

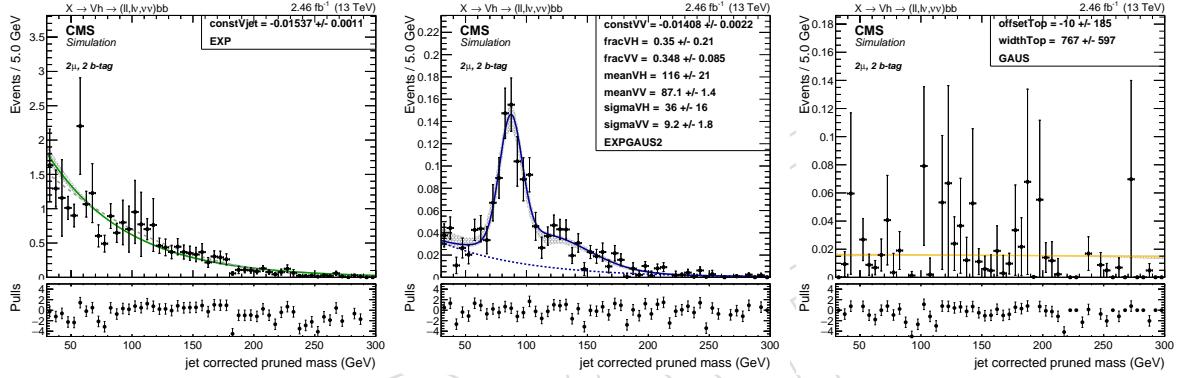


Figure 79: Fit to the simulated  $m_j$  in the 2 muons, 2 b-tag category for the three backgrounds:  $V + \text{jets}$  (left),  $VV$  (center),  $\text{Top}$  (right).

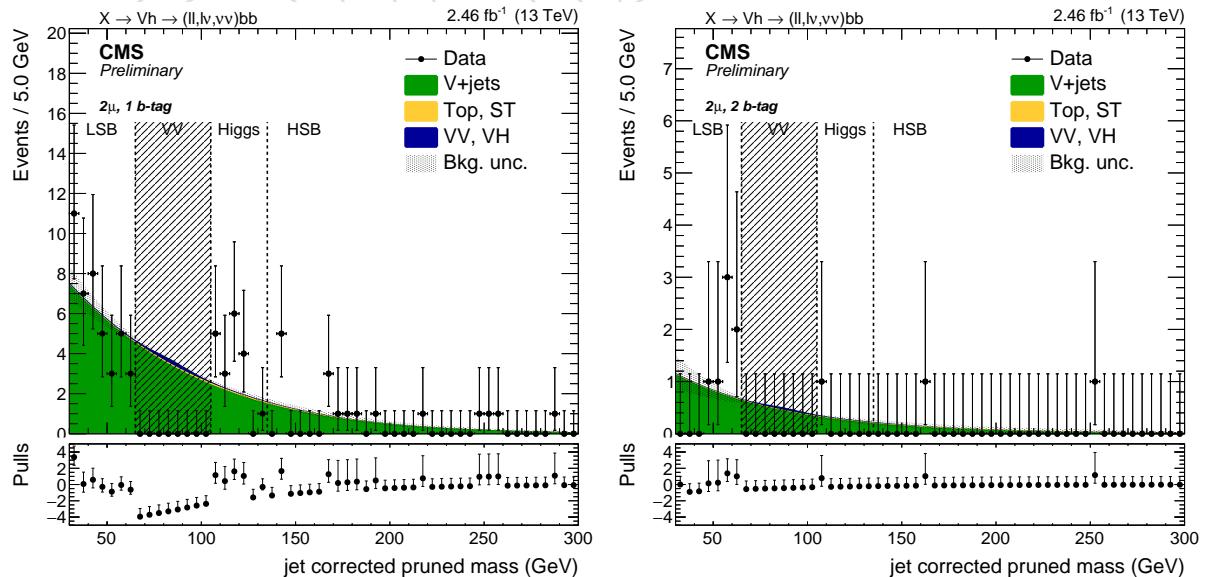


Figure 80: Fit to data  $m_j$  in the 2 muons, 1 b-tag (left) and 2 b-tag category (right).

1053 **8.2 Background shape**

The mass of the resonance candidate ( $m_X$ ) is parametrized separately for the V+jets ( $N_{SR}^{Vjet}(m_X)$ ,  $N_{SB}^{Vjet}(m_X)$ ),  $t\bar{t}$  ( $N_{SR}^{Top}(m_X)$ ,  $N_{SB}^{Top}(m_X)$ ), and dibosons ( $N_{SR}^{VV}(m_X)$ ,  $N_{SB}^{VV}(m_X)$ ). These functions are extracted fitting the simulated  $m_X$  (or  $m_t^X$  in the zero-lepton channel) spectrum in SR and SB, respectively. The  $t\bar{t}$  is weighted to match the normalization in the Top CR, and the diboson is normalized to luminosity. The V+jets functions are used to extract the  $\alpha$ -function:

$$\alpha(m_X) = \frac{N_{SR}^{Vjet}(m_X)}{N_{SB}^{Vjet}(m_X)}$$

1054 The main background is extracted through a fit to data in the SB, after subtracting the corre-  
 1055 sponding the Top and VV contribution from data. The resulting shape is then multiplied by  
 1056 the  $\alpha$ -function in order to get the main background expectation in the SR. Finally, the Top and  
 1057 diboson contribution in the SR is added to the main background estimation.

1058 In formulas, the procedure used to extract the total background prediction is the following:

$$N_{SR}^{main}(m_X) = N_{SR}^{main}(m_X) \times \alpha(m_X)$$

$$N_{SR,SB}^{bkg}(m_X) = N_{SR,SB}^{main}(m_X) + N_{SR,SB}^{Top}(m_X) + N_{SR,SB}^{VV}(m_X)$$

$$N_{SR}^{data}(m_X) = \left[ N_{SB}^{data}(m_X) - N_{SB}^{Top}(m_X) - N_{SB}^{VV}(m_X) \right] \times \left[ \frac{N_{SR}^{Vjet}(m_X)}{N_{SB}^{Vjet}(m_X)} \right] + N_{SR}^{Top}(m_X) + N_{SR}^{VV}(m_X)$$

1059 The functions used to parametrize the  $m_X$  distributions are:

**Exp:** a simple exponential function. Its simplicity is balanced by the limited possibility to model the  $m_X$  tails in some channels:

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

**Exp2:** a double exponential function. It has better description of the tails, but introduces two new parameters:

$$F_{\text{Exp2}}(x) = (1 - f_0) \cdot e^{ax} + f_0 \cdot e^{bx}$$

**ExpN:** a product of two exponentials:

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

**ExpTail:** a modified exponential function with an additional parameter to model the exponential tails:

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

**Pow:** a second-order power function:

$$F_{\text{Pow}}(x) = a_0 \cdot x + a_1 \cdot x^2$$

1060 The functions chosen to parametrize the main background and extract the  $\alpha$ -function are re-  
 1061 ported in Table 17 for each category. The diboson and Top are always parametrized with an  
 1062 exponential (Exp) function.

1063 As a cross-check for the main  $\alpha$ -function used in the background estimation, an additional  $\alpha$ -  
 1064 function is extracted with alternative function choices. Table 17 reports both the main function  
 1065 and the alternative function.

category		Main bkg function	Main bkg alternative	diboson	top
1 b-tag	0 $\ell$	ExpN	ExpTail	ExpN	ExpN
	1e	ExpTail	ExpN	ExpN	ErfExp
	1 $\mu$	ExpN	ExpTail	ExpN	ErfExp
	2e	ExpTail	Pow	ExpTail	Exp
	2 $\mu$	ExpTail	Pow	ExpTail	Exp
2 b-tag	0 $\ell$	ExpN	ExpTail	ExpN	ExpN
	1e	ExpN	Pow	ExpN	ErfExp
	1 $\mu$	ExpN	Pow	ExpN	ErfExp
	2e	ExpTail	Pow	ExpTail	Exp
	2 $\mu$	ExpTail	Pow	ExpTail	Exp

Table 17: Main and alternative functions chosen to parametrize the main background contribution in the  $m_X$  distribution for each channel.

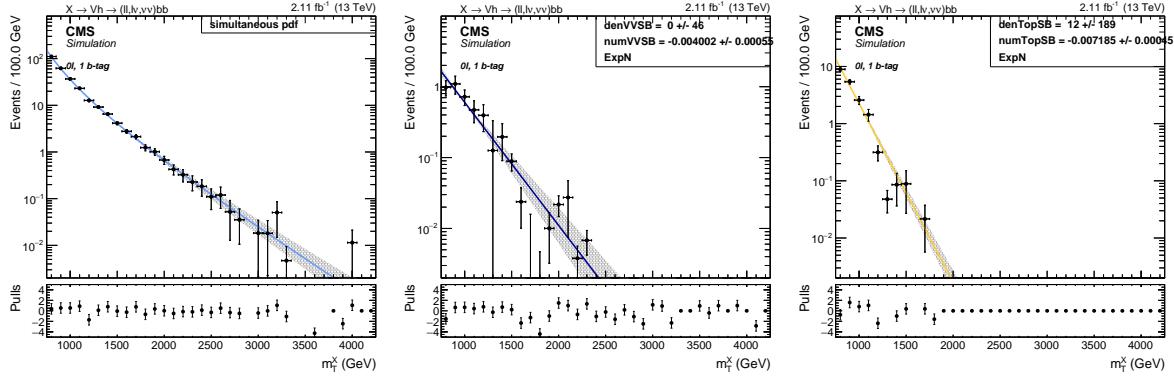


Figure 81: 0 leptons, 1 b-tag channel. Fits to the simulated background components V +jets (left), VV (center), Top (right) in the sidebands (SB).

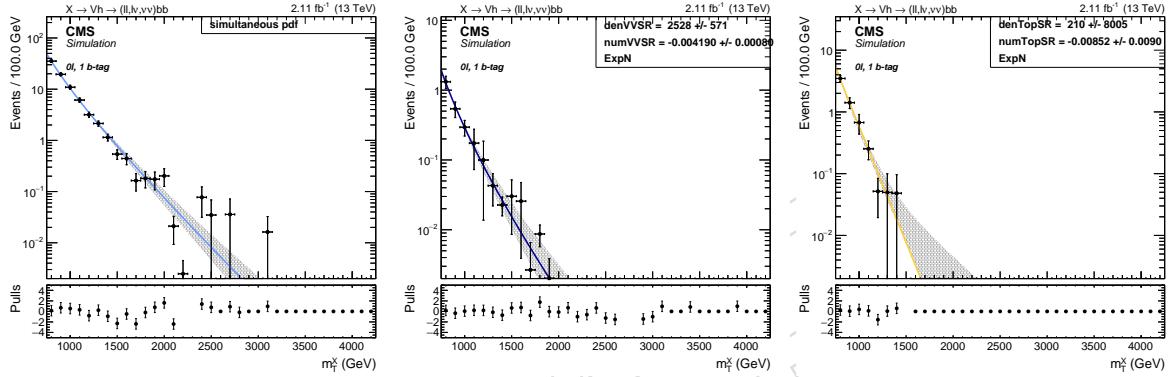


Figure 82: 0 leptons, 1 b-tag channel. Fits to the simulated background components V +jets (left), VV (center), Top (right) in the signal region (SR).

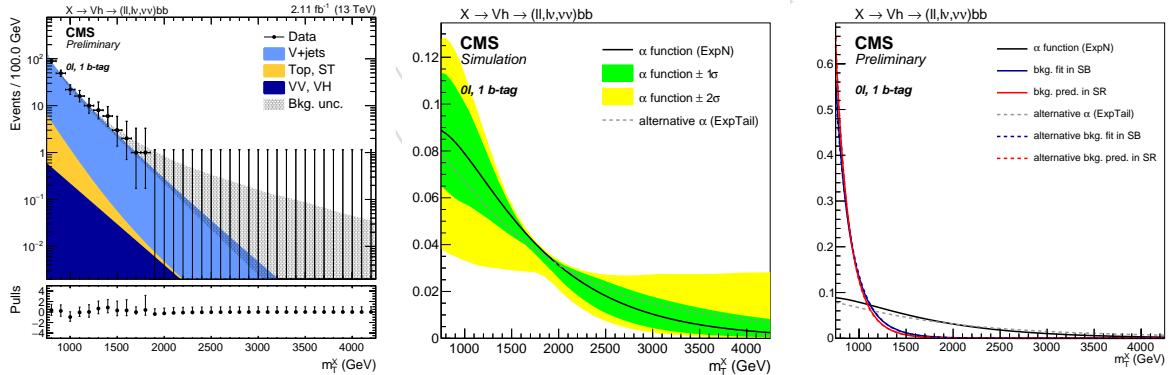


Figure 83: 0 leptons, 1 b-tag channel. Fit to data in the SB (left), alpha function (center), and alpha function compared to the background shape in both SB and SR (right). The black line, with the corresponding 1 $\sigma$  (green) and 2 $\sigma$  (yellow) uncertainty bands, represents the  $\alpha$ -function. The gray line is the alternative  $\alpha$ -function. The blue and lines represent the estimated background in the SB and SR, respectively, with both the main (solid line) and alternative (dotted line) parametrization.

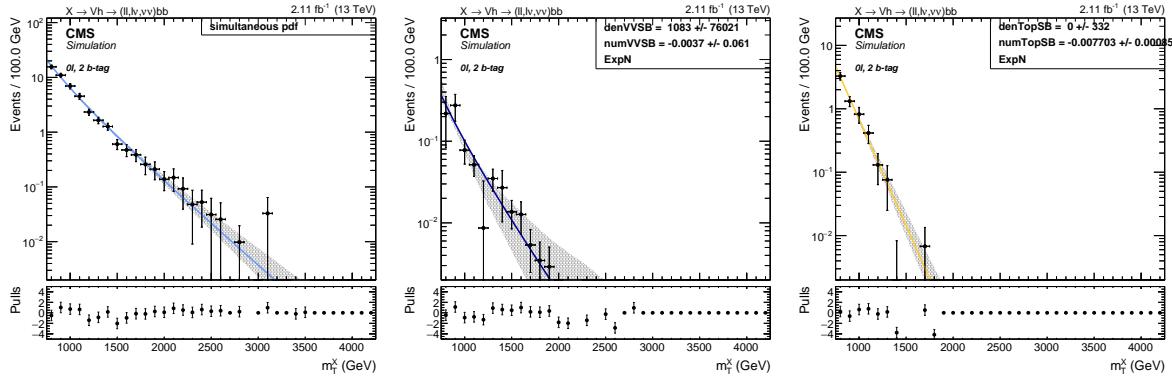


Figure 84: 0 leptons, 2 b-tag channel. Fits to the simulated background components V +jets (left), VV (center), Top (right) in the sidebands (SB).

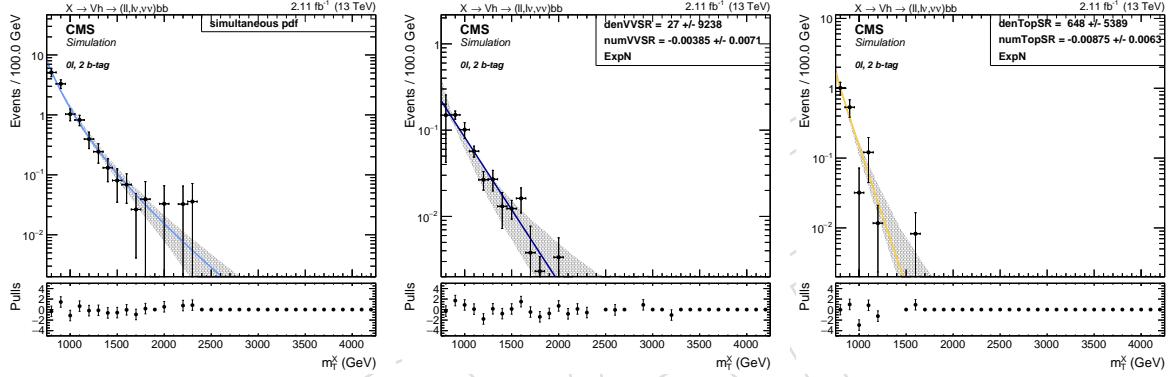


Figure 85: 0 leptons, 2 b-tag channel. Fits to the simulated background components V +jets (left), VV (center), Top (right) in the signal region (SR).

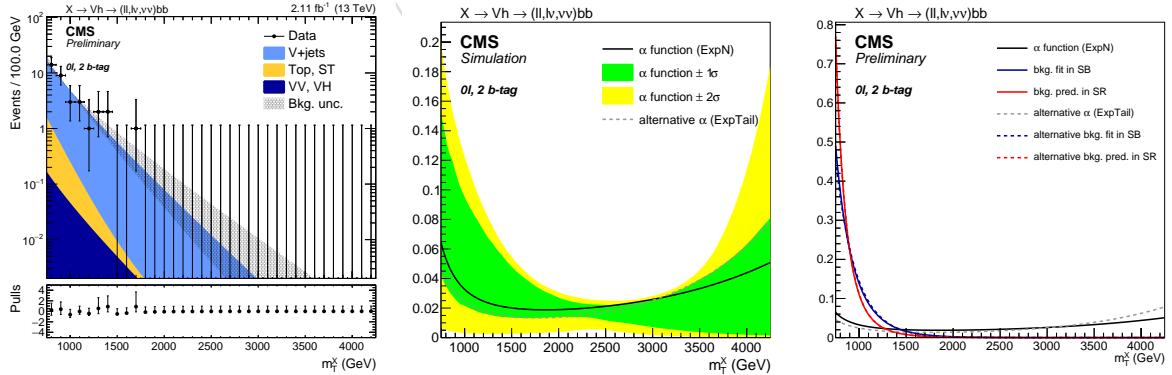


Figure 86: 0 leptons, 2 b-tag channel. Fit to data in the SB (left), alpha function (center), and alpha function compared to the background shape in both SB and SR (right). The black line, with the corresponding  $1\sigma$  (green) and  $2\sigma$  (yellow) uncertainty bands, represents the  $\alpha$ -function. The gray line is the alternative  $\alpha$ -function. The blue and red lines represent the estimated background in the SB and SR, respectively, with both the main (solid line) and alternative (dotted line) parametrization.

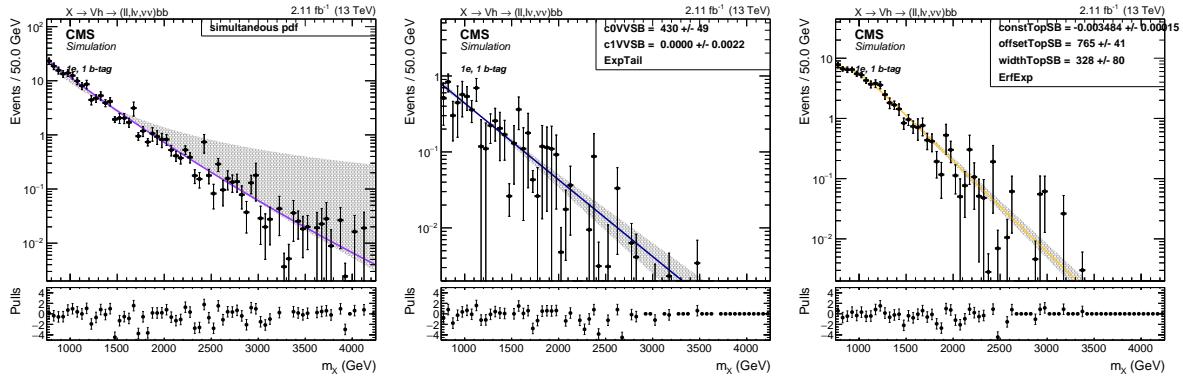


Figure 87: 1 electron, 1 b-tag channel. Fits to the simulated background components V +jets (left), VV (center), Top (right) in the sidebands (SB).

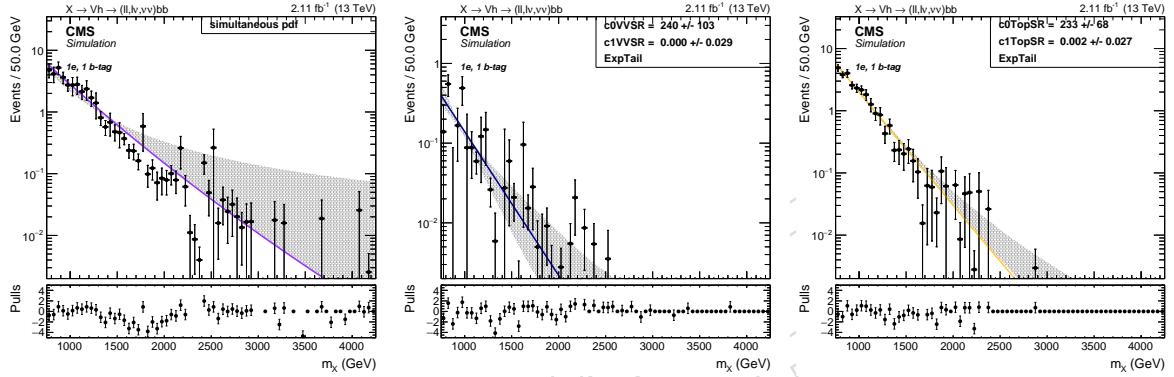


Figure 88: 1 electron, 1 b-tag channel. Fits to the simulated background components V +jets (left), VV (center), Top (right) in the signal region (SR).

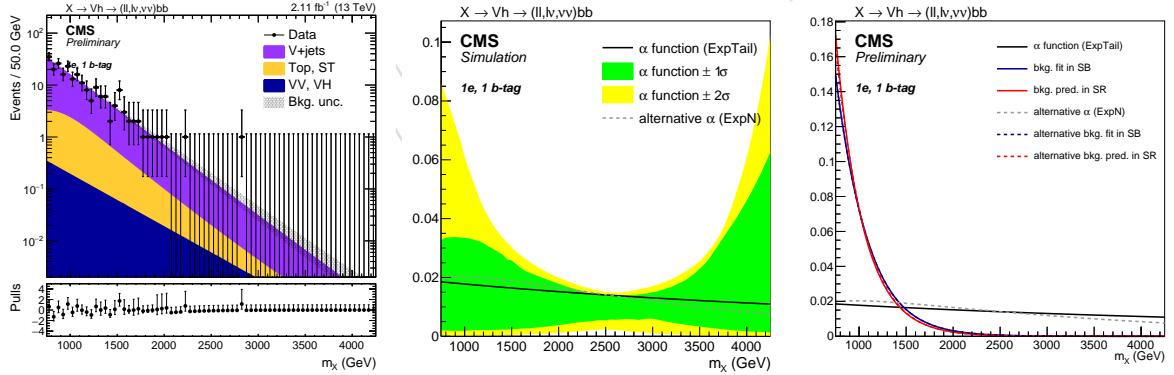


Figure 89: 1 electron, 1 b-tag channel. Fit to data in the SB (left), alpha function (center), and alpha function compared to the background shape in both SB and SR (right). The black line, with the corresponding  $1\sigma$  (green) and  $2\sigma$  (yellow) uncertainty bands, represents the  $\alpha$ -function. The gray line is the alternative  $\alpha$ -function. The blue and lines represent the estimated background in the SB and SR, respectively, with both the main (solid line) and alternative (dotted line) parametrization.

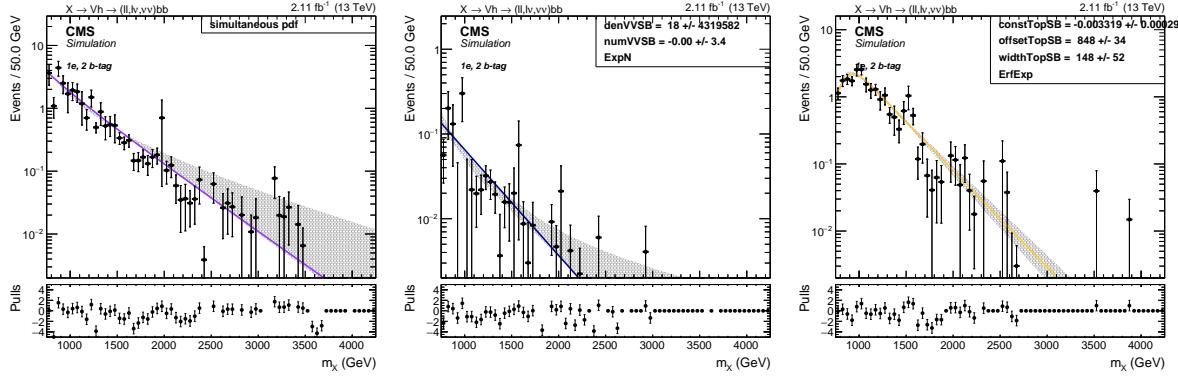


Figure 90: 1 electron, 2 b-tag channel. Fits to the simulated background components  $V + \text{jets}$  (left),  $VV$  (center),  $\text{Top}$  (right) in the sidebands (SB).

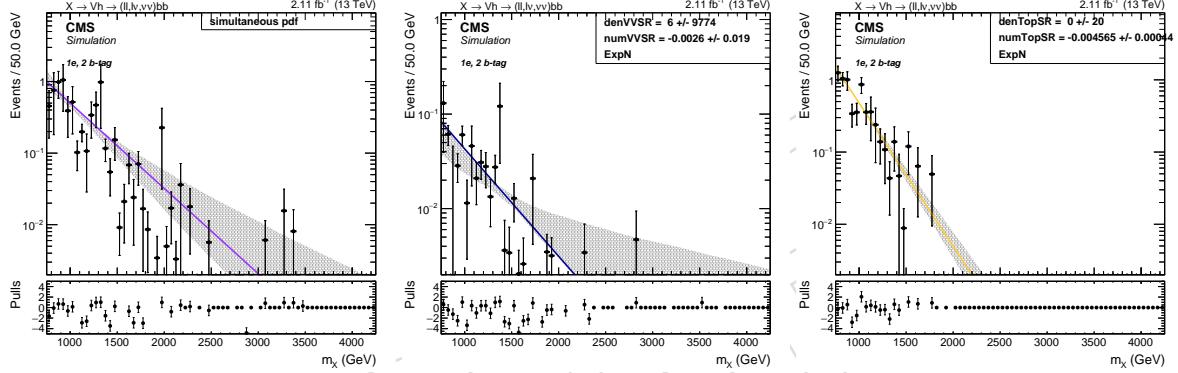


Figure 91: 1 electron, 2 b-tag channel. Fits to the simulated background components  $V + \text{jets}$  (left),  $VV$  (center),  $\text{Top}$  (right) in the signal region (SR).

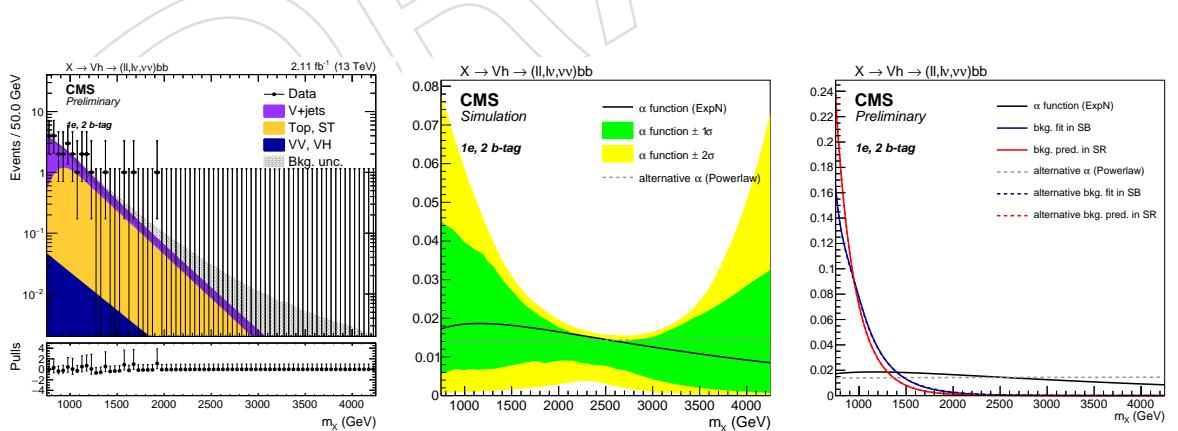


Figure 92: 1 electron, 2 b-tag channel. Fit to data in the SB (left), alpha function (center), and alpha function compared to the background shape in both SB and SR (right). The black line, with the corresponding  $1\sigma$  (green) and  $2\sigma$  (yellow) uncertainty bands, represents the  $\alpha$ -function. The gray line is the alternative  $\alpha$ -function. The blue and lines represent the estimated background in the SB and SR, respectively, with both the main (solid line) and alternative (dotted line) parametrization.

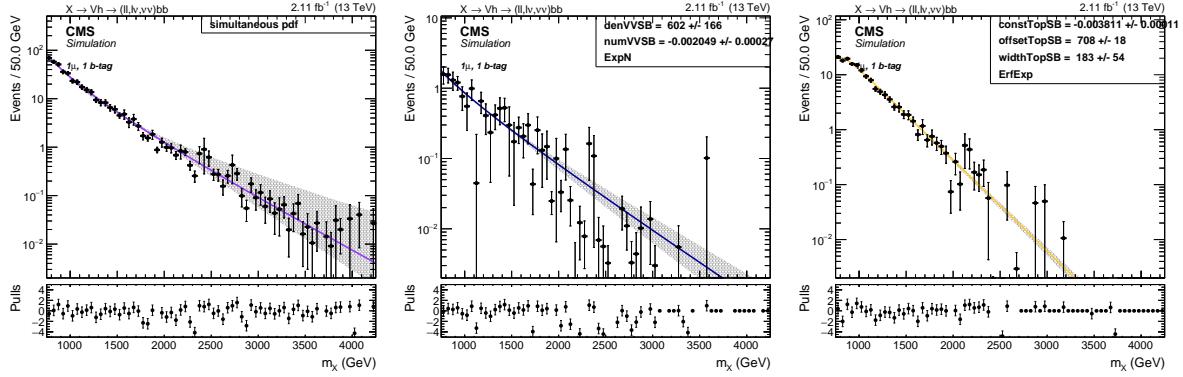


Figure 93: 1 muon, 1 b-tag channel. Fits to the simulated background components  $V + \text{jets}$  (left),  $VV$  (center),  $\text{Top}$  (right) in the sidebands (SB).

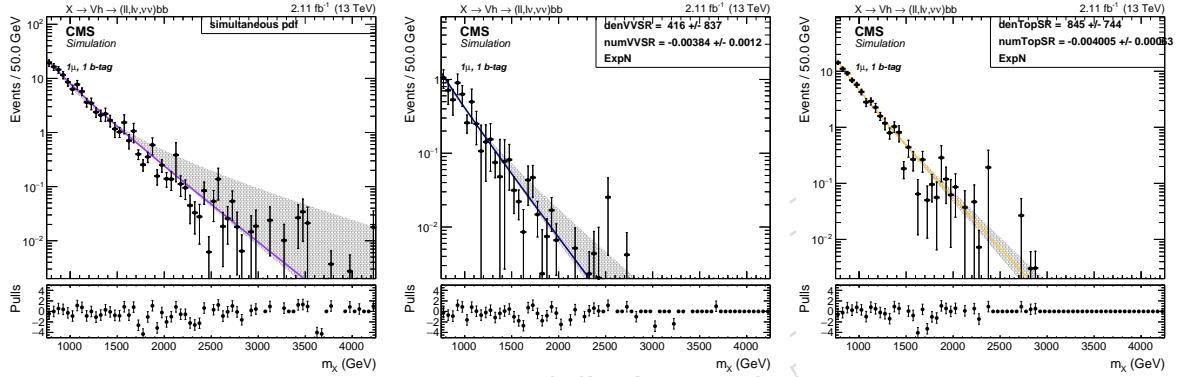


Figure 94: 1 muon, 1 b-tag channel. Fits to the simulated background components  $V + \text{jets}$  (left),  $VV$  (center),  $\text{Top}$  (right) in the signal region (SR).

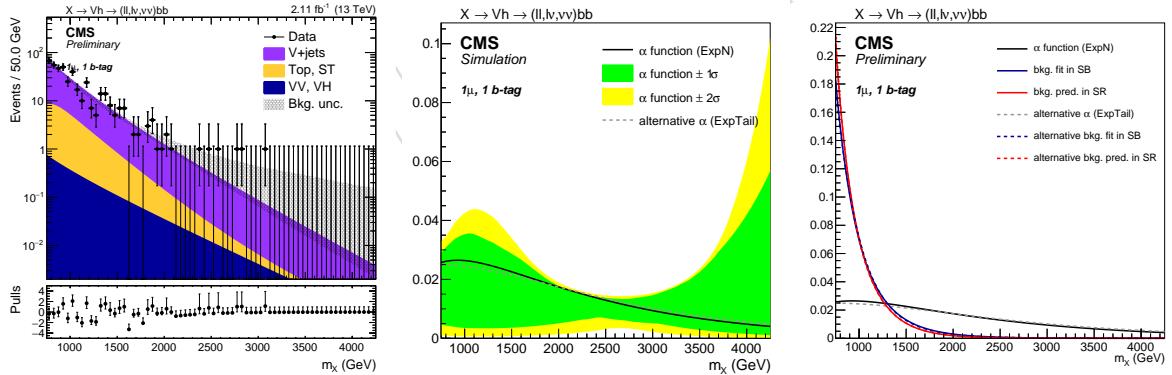


Figure 95: 1 muon, 1 b-tag channel. Fit to data in the SB (left), alpha function (center), and alpha function compared to the background shape in both SB and SR (right). The black line, with the corresponding  $1\sigma$  (green) and  $2\sigma$  (yellow) uncertainty bands, represents the  $\alpha$ -function. The gray line is the alternative  $\alpha$ -function. The blue and lines represent the estimated background in the SB and SR, respectively, with both the main (solid line) and alternative (dotted line) parametrization.

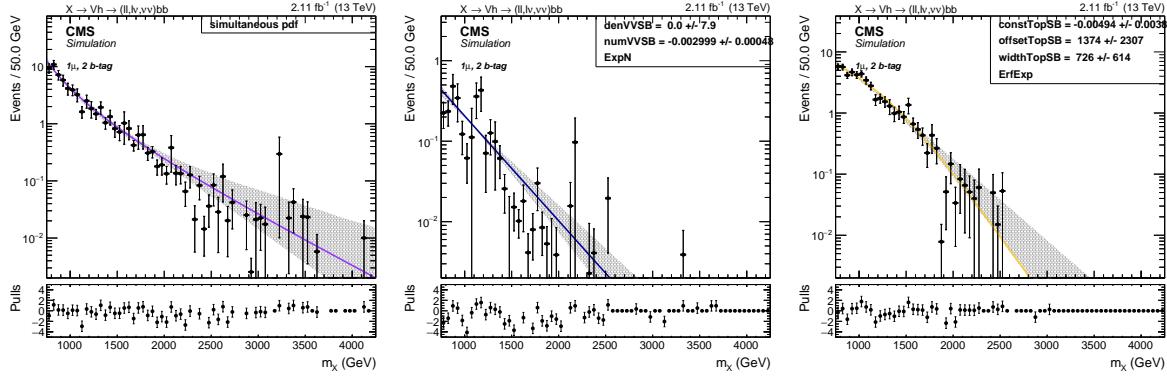


Figure 96: 1 muon, 2 b-tag channel. Fits to the simulated background components  $V + \text{jets}$  (left),  $VV$  (center), Top (right) in the sidebands (SB).

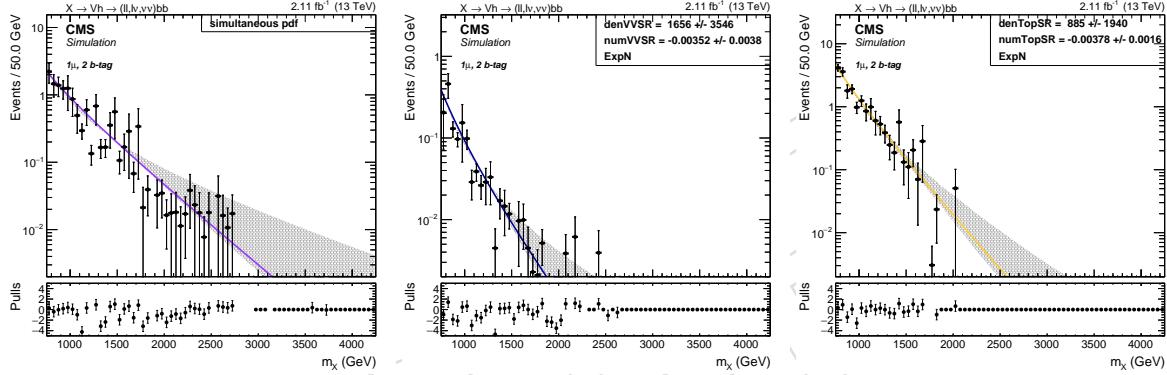


Figure 97: 1 muon, 2 b-tag channel. Fits to the simulated background components  $V + \text{jets}$  (left),  $VV$  (center), Top (right) in the signal region (SR).

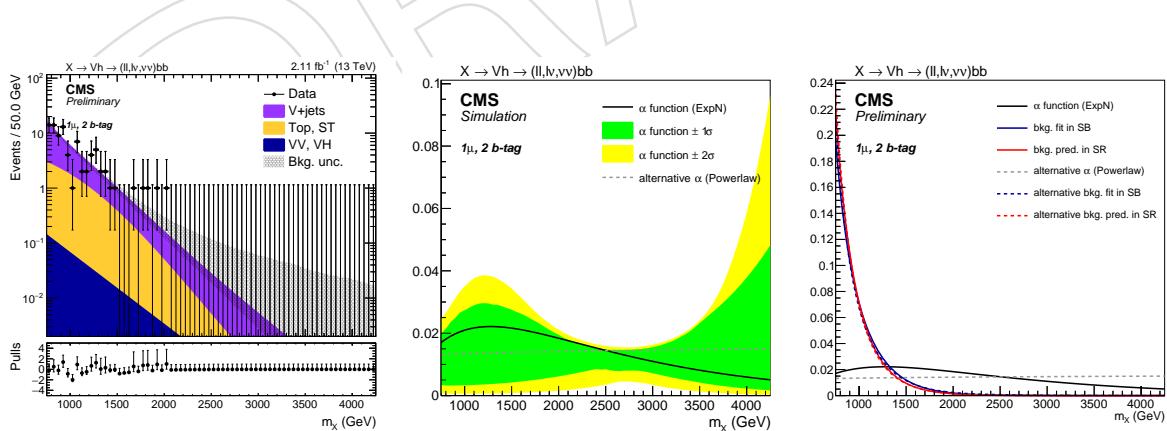


Figure 98: 1 muon, 2 b-tag channel. Fit to data in the SB (left), alpha function (center), and alpha function compared to the background shape in both SB and SR (right). The black line, with the corresponding  $1\sigma$  (green) and  $2\sigma$  (yellow) uncertainty bands, represents the  $\alpha$ -function. The gray line is the alternative  $\alpha$ -function. The blue and lines represent the estimated background in the SB and SR, respectively, with both the main (solid line) and alternative (dotted line) parametrization.

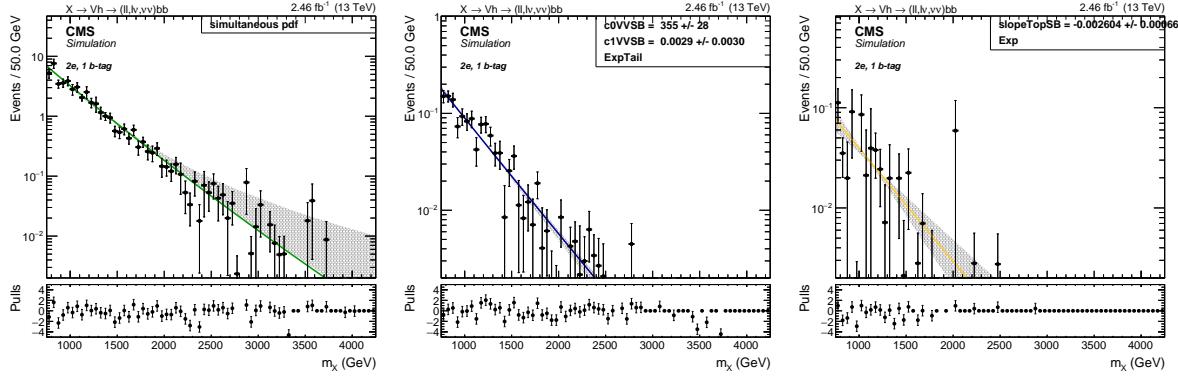


Figure 99: 2 electrons, 1 b-tag channel. Fits to the simulated background components V +jets (left), VV (center), Top (right) in the sidebands (SB).

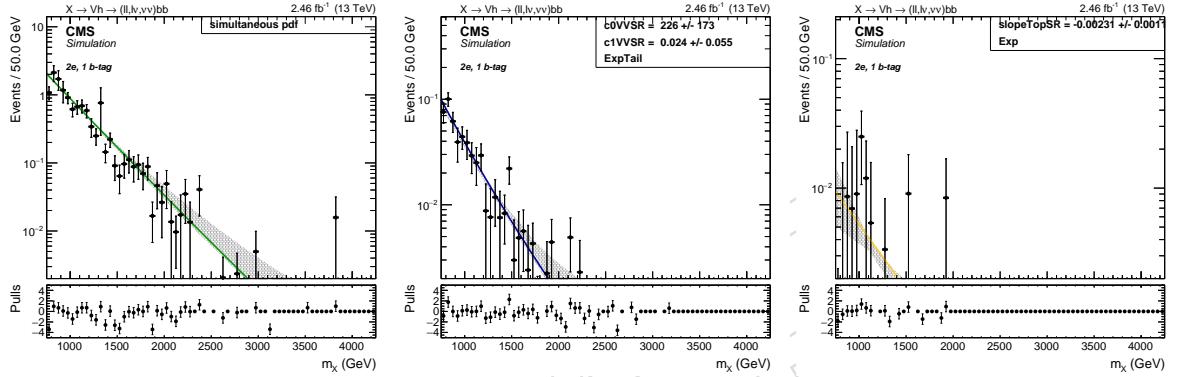


Figure 100: 2 electrons, 1 b-tag channel. Fits to the simulated background components V +jets (left), VV (center), Top (right) in the signal region (SR).

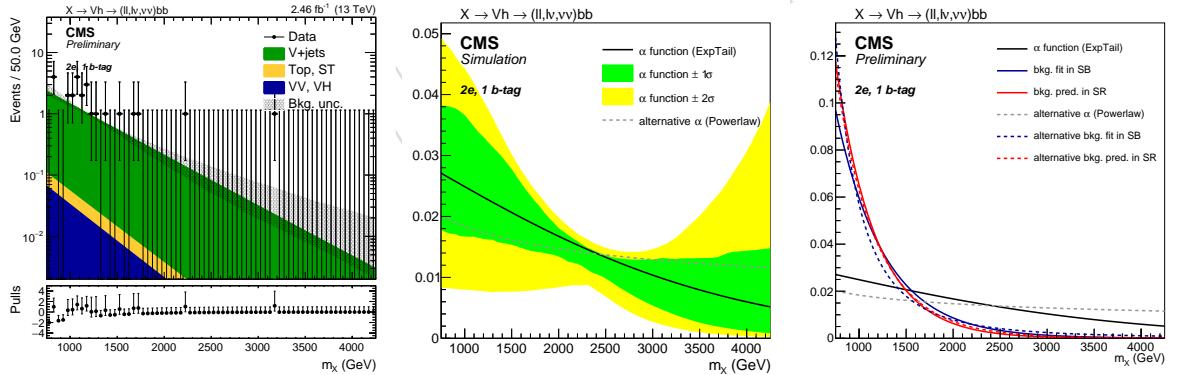


Figure 101: 2 electrons, 1 b-tag channel. Fit to data in the SB (left), alpha function (center), and alpha function compared to the background shape in both SB and SR (right). The black line, with the corresponding  $1\sigma$  (green) and  $2\sigma$  (yellow) uncertainty bands, represents the  $\alpha$ -function. The gray line is the alternative  $\alpha$ -function. The blue and lines represent the estimated background in the SB and SR, respectively, with both the main (solid line) and alternative (dotted line) parametrization.

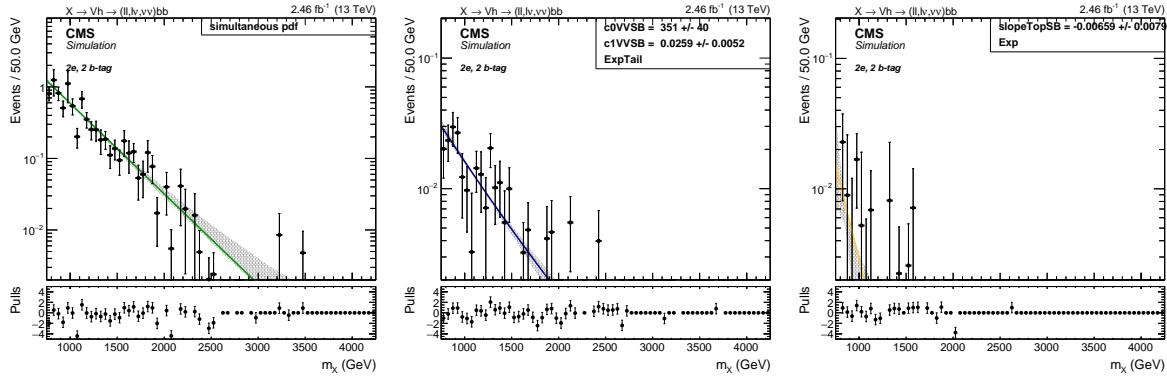


Figure 102: 2 electrons, 2 b-tag channel. Fits to the simulated background components  $V + \text{jets}$  (left),  $VV$  (center), Top (right) in the sidebands (SB).

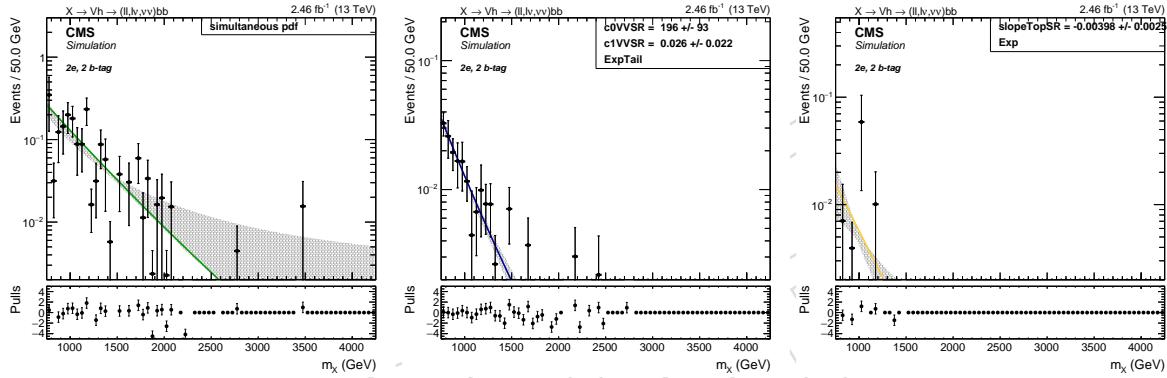


Figure 103: 2 electrons, 2 b-tag channel. Fits to the simulated background components  $V + \text{jets}$  (left),  $VV$  (center), Top (right) in the signal region (SR).

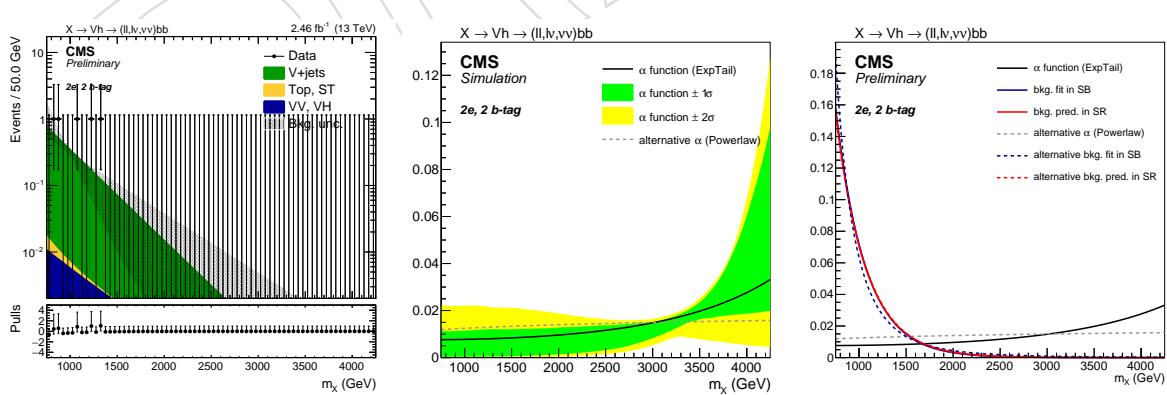


Figure 104: 2 electrons, 2 b-tag channel. Fit to data in the SB (left), alpha function (center), and alpha function compared to the background shape in both SB and SR (right). The black line, with the corresponding  $1\sigma$  (green) and  $2\sigma$  (yellow) uncertainty bands, represents the  $\alpha$ -function. The gray line is the alternative  $\alpha$ -function. The blue and lines represent the estimated background in the SB and SR, respectively, with both the main (solid line) and alternative (dotted line) parametrization.

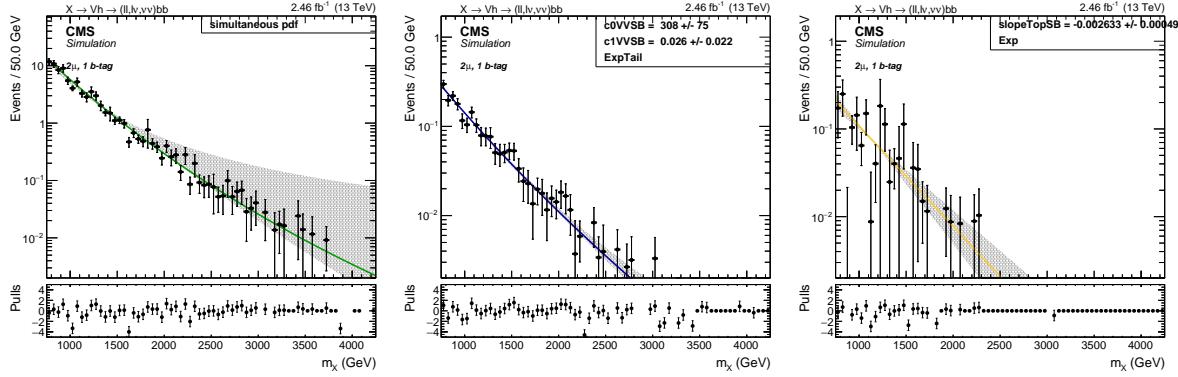


Figure 105: 2 muons, 1 b-tag channel. Fits to the simulated background components V +jets (left), VV (center), Top (right) in the sidebands (SB).

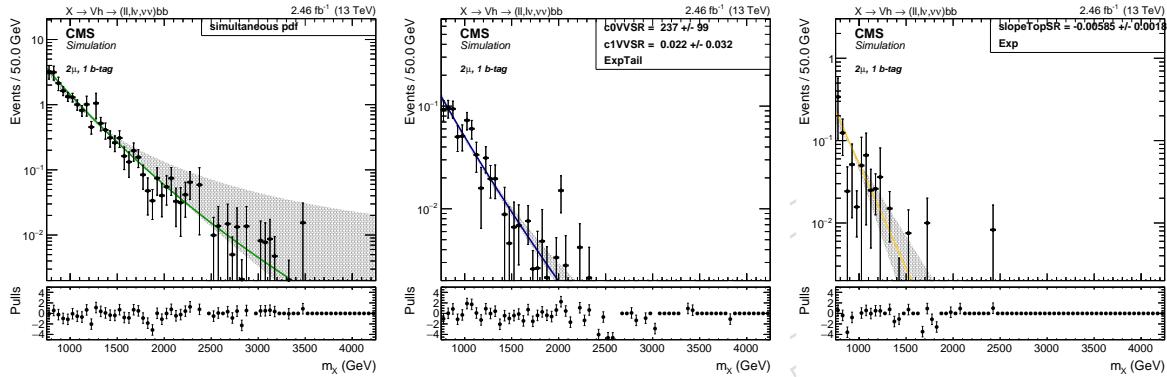


Figure 106: 2 muons, 1 b-tag channel. Fits to the simulated background components V +jets (left), VV (center), Top (right) in the signal region (SR).

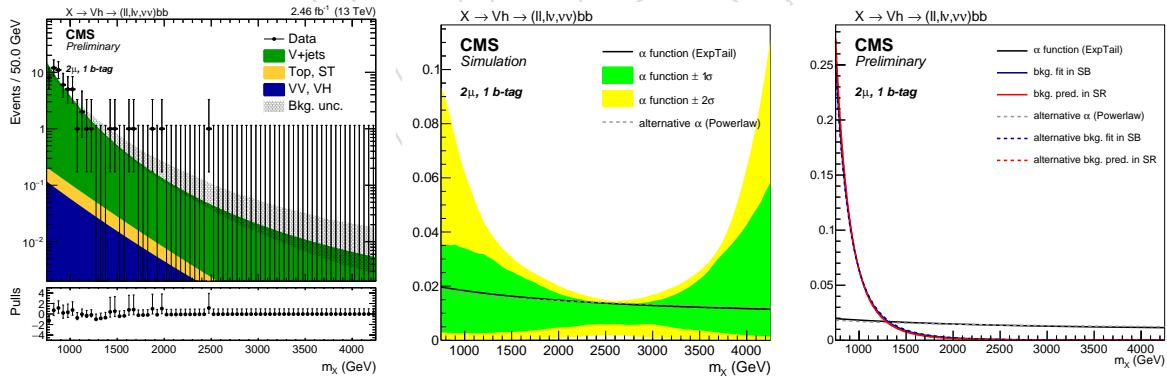


Figure 107: 2 muons, 1 b-tag channel. Fit to data in the SB (left), alpha function (center), and alpha function compared to the background shape in both SB and SR (right). The black line, with the corresponding  $1\sigma$  (green) and  $2\sigma$  (yellow) uncertainty bands, represents the  $\alpha$ -function. The gray line is the alternative  $\alpha$ -function. The blue and lines represent the estimated background in the SB and SR, respectively, with both the main (solid line) and alternative (dotted line) parametrization.

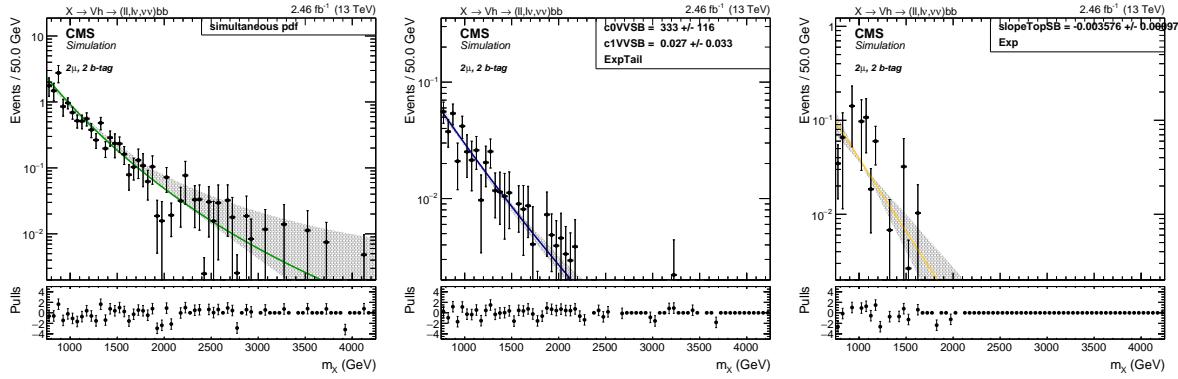


Figure 108: 2 muons, 2 b-tag channel. Fits to the simulated background components  $V + \text{jets}$  (left),  $VV$  (center), Top (right) in the sidebands (SB).

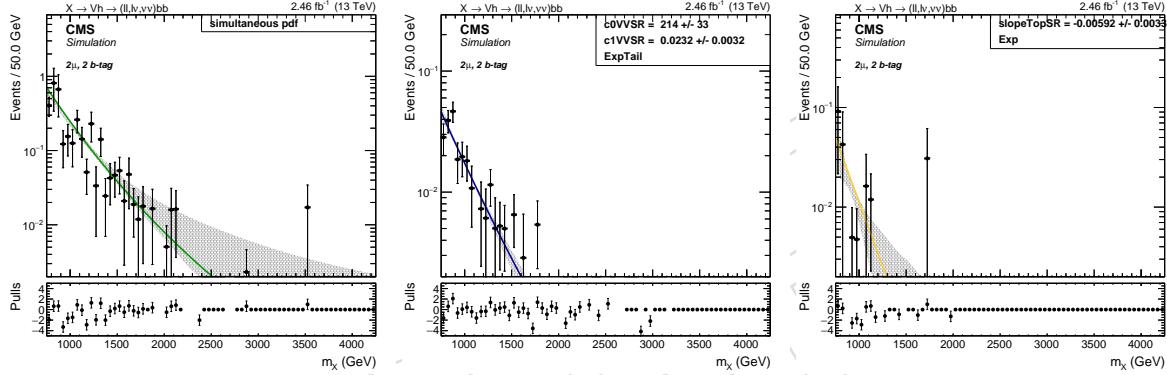


Figure 109: 2 muons, 2 b-tag channel. Fits to the simulated background components  $V + \text{jets}$  (left),  $VV$  (center), Top (right) in the signal region (SR).

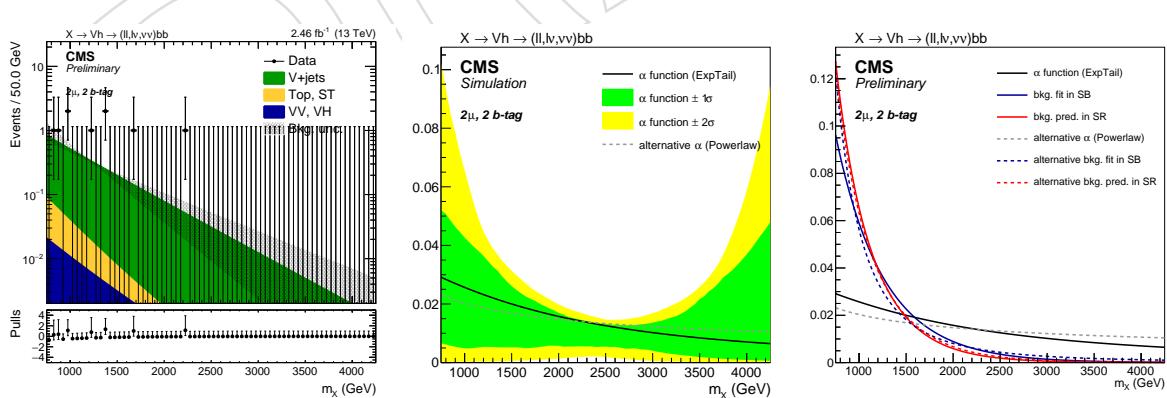


Figure 110: 2 muons, 2 b-tag channel. Fit to data in the SB (left), alpha function (center), and alpha function compared to the background shape in both SB and SR (right). The black line, with the corresponding  $1\sigma$  (green) and  $2\sigma$  (yellow) uncertainty bands, represents the  $\alpha$ -function. The gray line is the alternative  $\alpha$ -function. The blue and lines represent the estimated background in the SB and SR, respectively, with both the main (solid line) and alternative (dotted line) parametrization.

1066 **8.3 Background prediction**

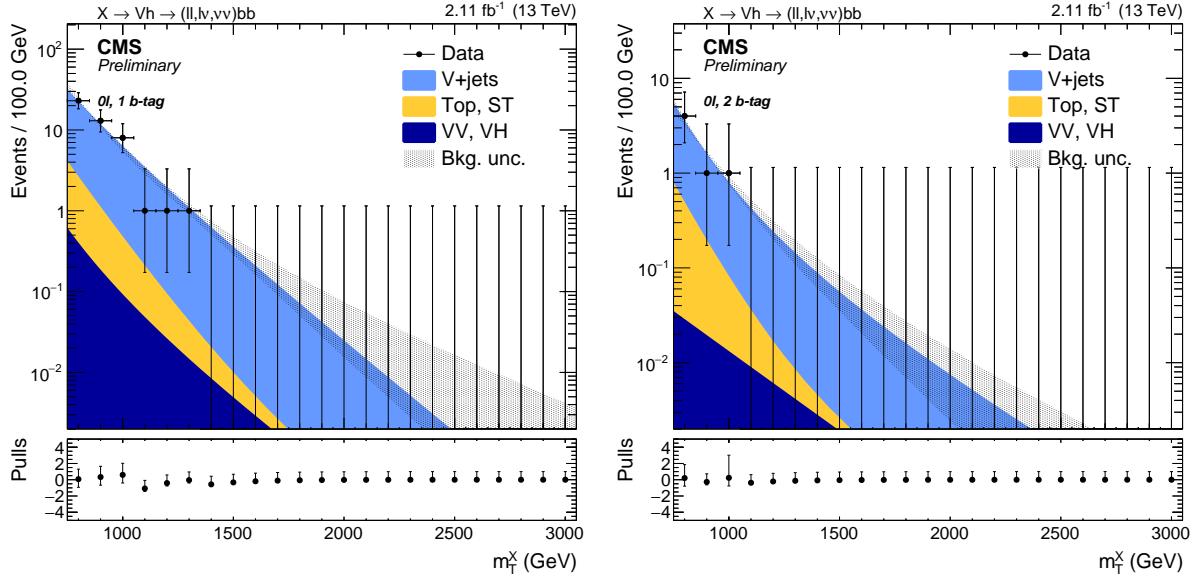


Figure 111: Expected background with the  $\alpha$  method in the 0 lepton, 1 b-tag (left) and 2 b-tag category (right).

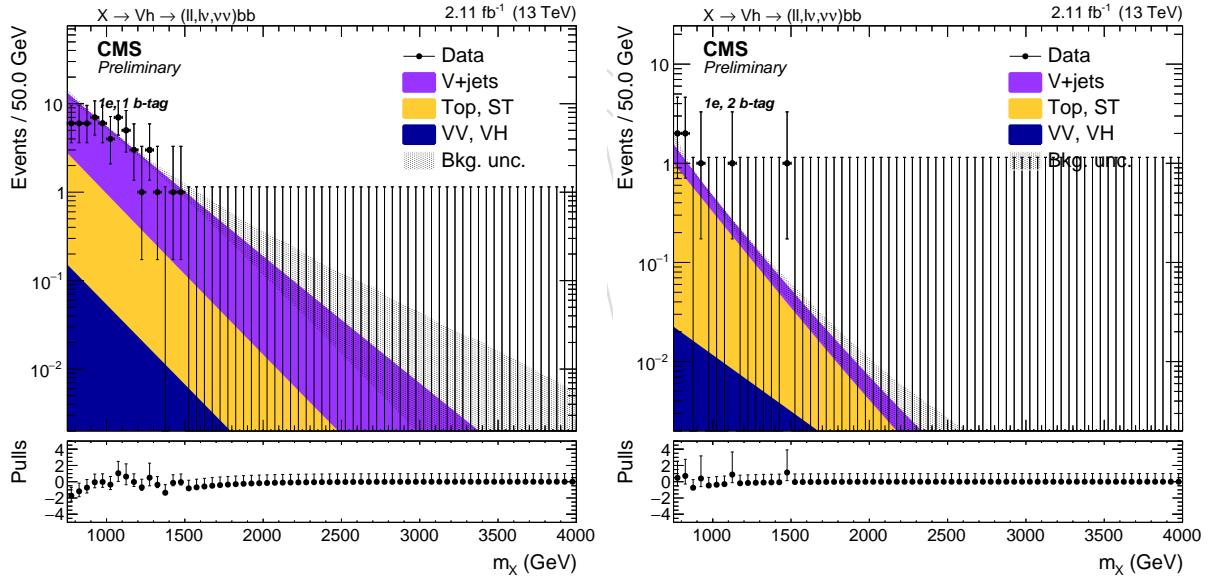


Figure 112: Expected background with the  $\alpha$  method in the 1 electron, 1 b-tag (left) and 2 b-tag category (right).

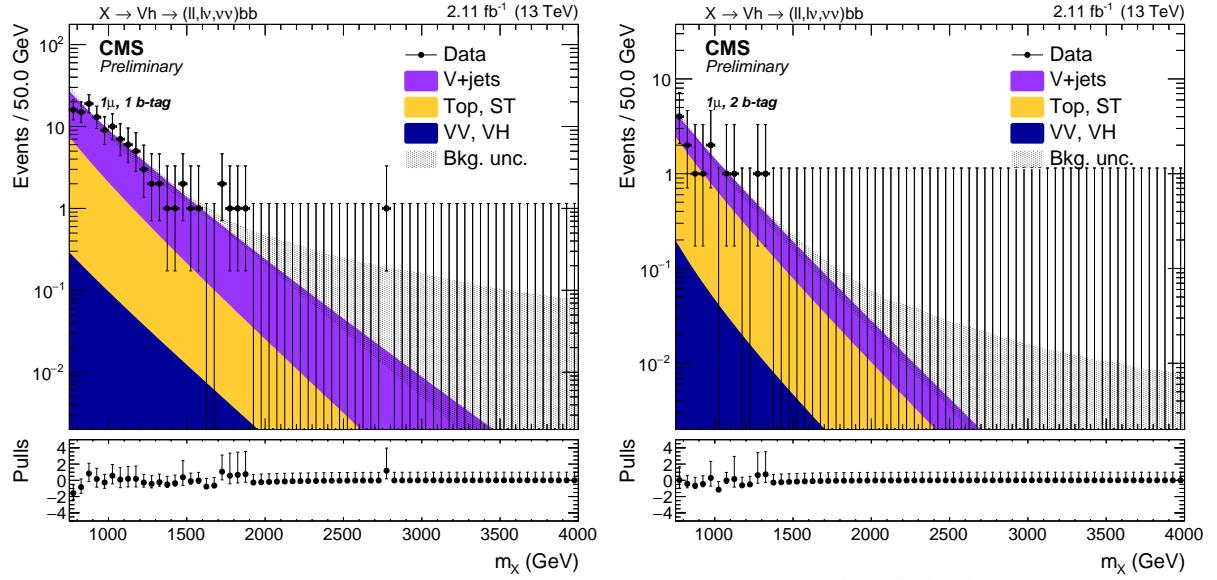


Figure 113: Expected background with the  $\alpha$  method in the 1 muon, 1 b-tag (left) and 2 b-tag category (right).

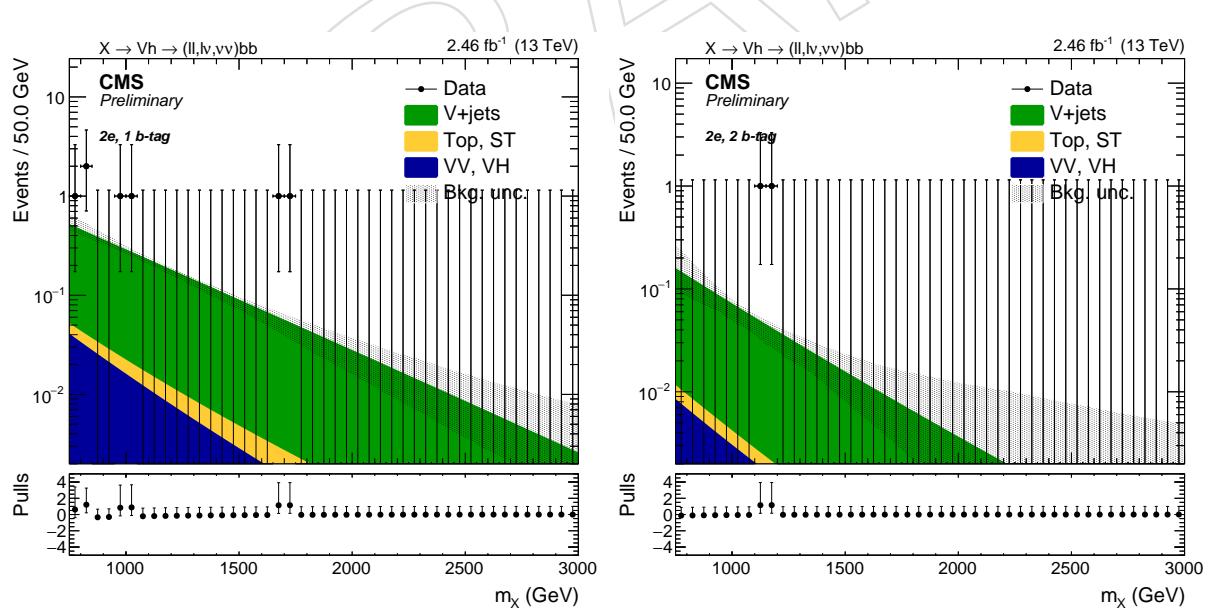


Figure 114: Expected background with the  $\alpha$  method in the 2 electrons, 1 b-tag (left) and 2 b-tag category (right).

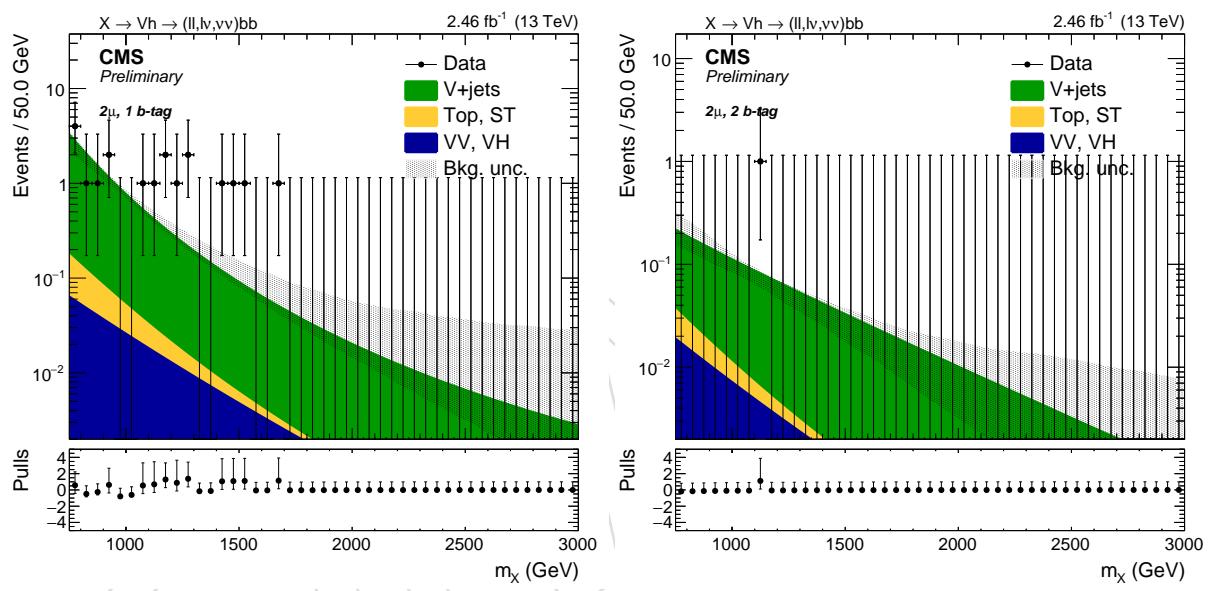


Figure 115: Expected background with the  $\alpha$  method in the 2 muons, 1 b-tag (left) and 2 b-tag category (right).

1067 **8.4 Signal modeling**

1068 The simulated signal mass points are fitted in the SR with an empiric function in order to be  
1069 able to perform an unbinned likelihood fit for the signal extraction. The signal model is made  
1070 of a gaussian-like peak plus a tail towards lower values. The function chosen to model the  
1071 signal is the *Crystal Ball*, which consists in a gaussian core and a power function that describes  
1072 the low-end tail, below a certain threshold.

DRAFT

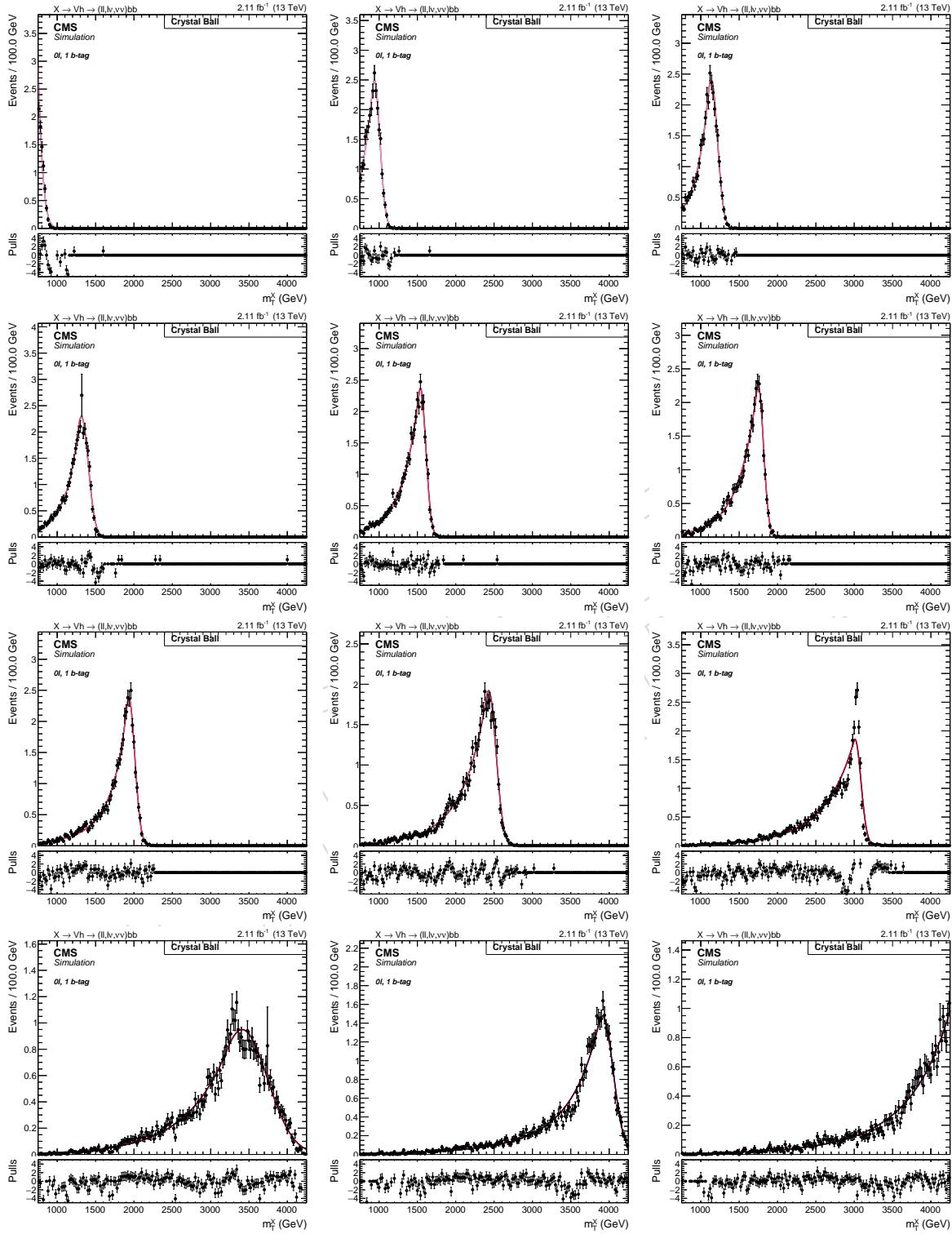


Figure 116: Fit to the signal samples in the 0 leptons, 1 b-tag category.

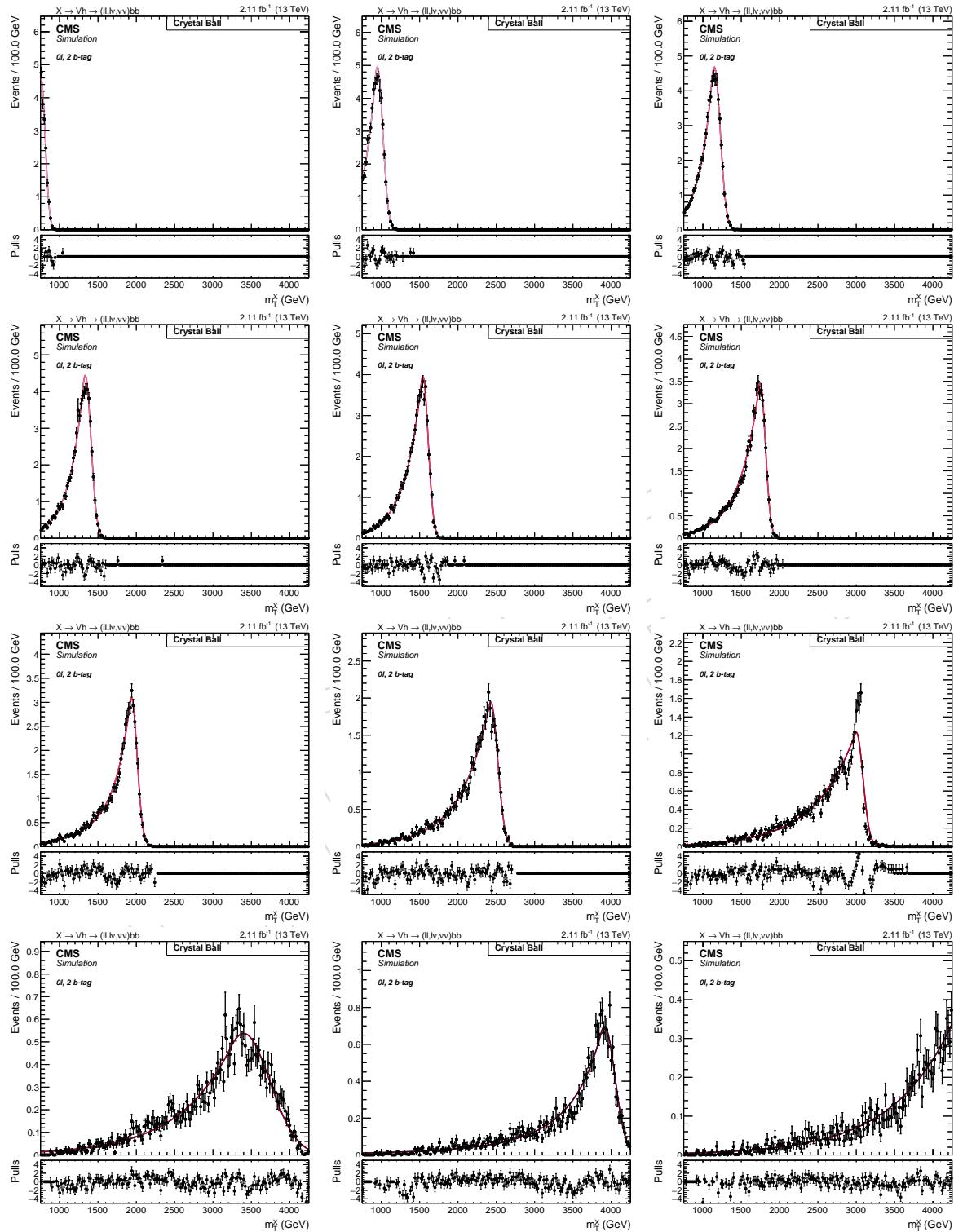


Figure 117: Fit to the signal samples in the 0 leptons, 2 b-tag category.

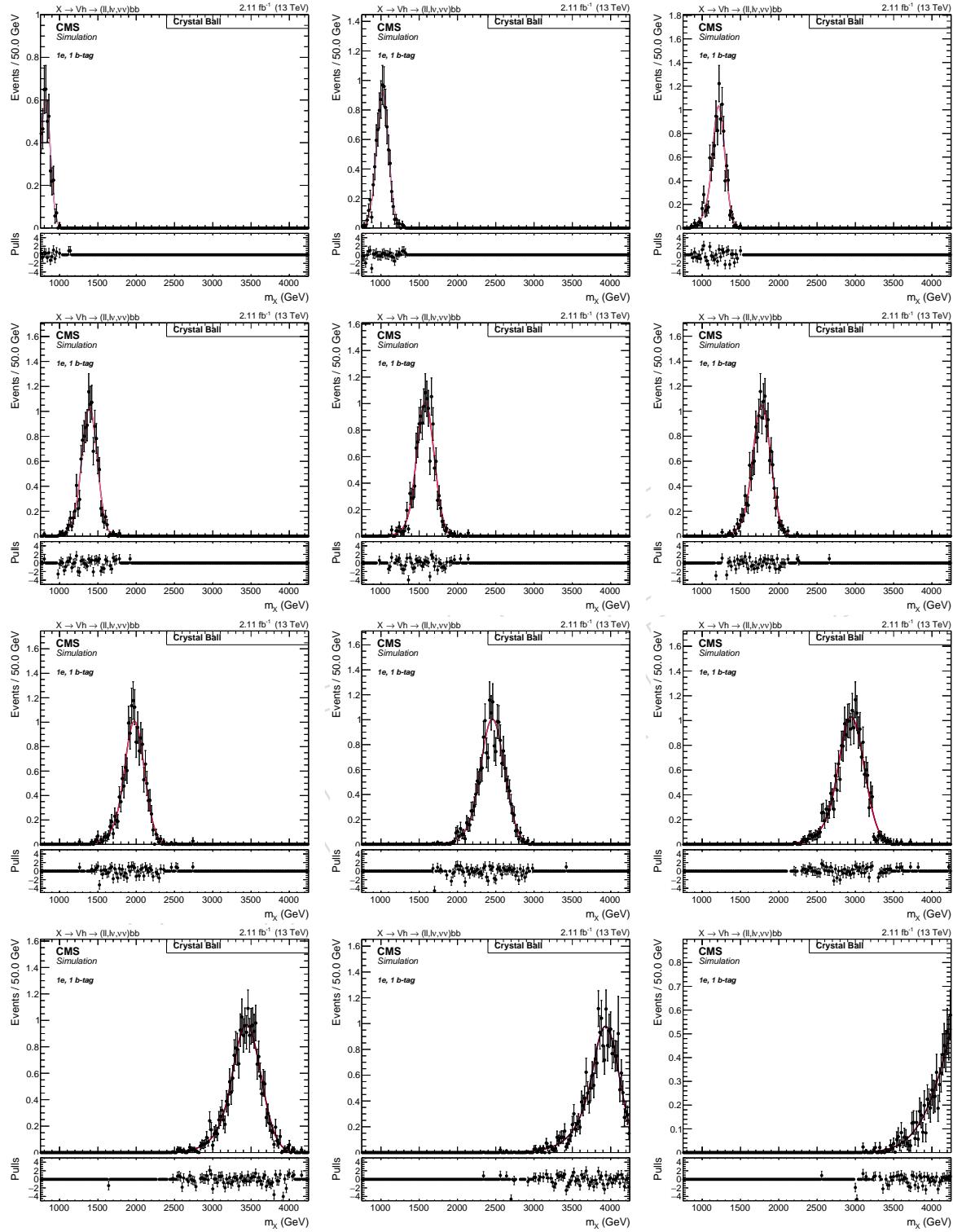


Figure 118: Fit to the signal samples in the 1 electron, 1 b-tag category.

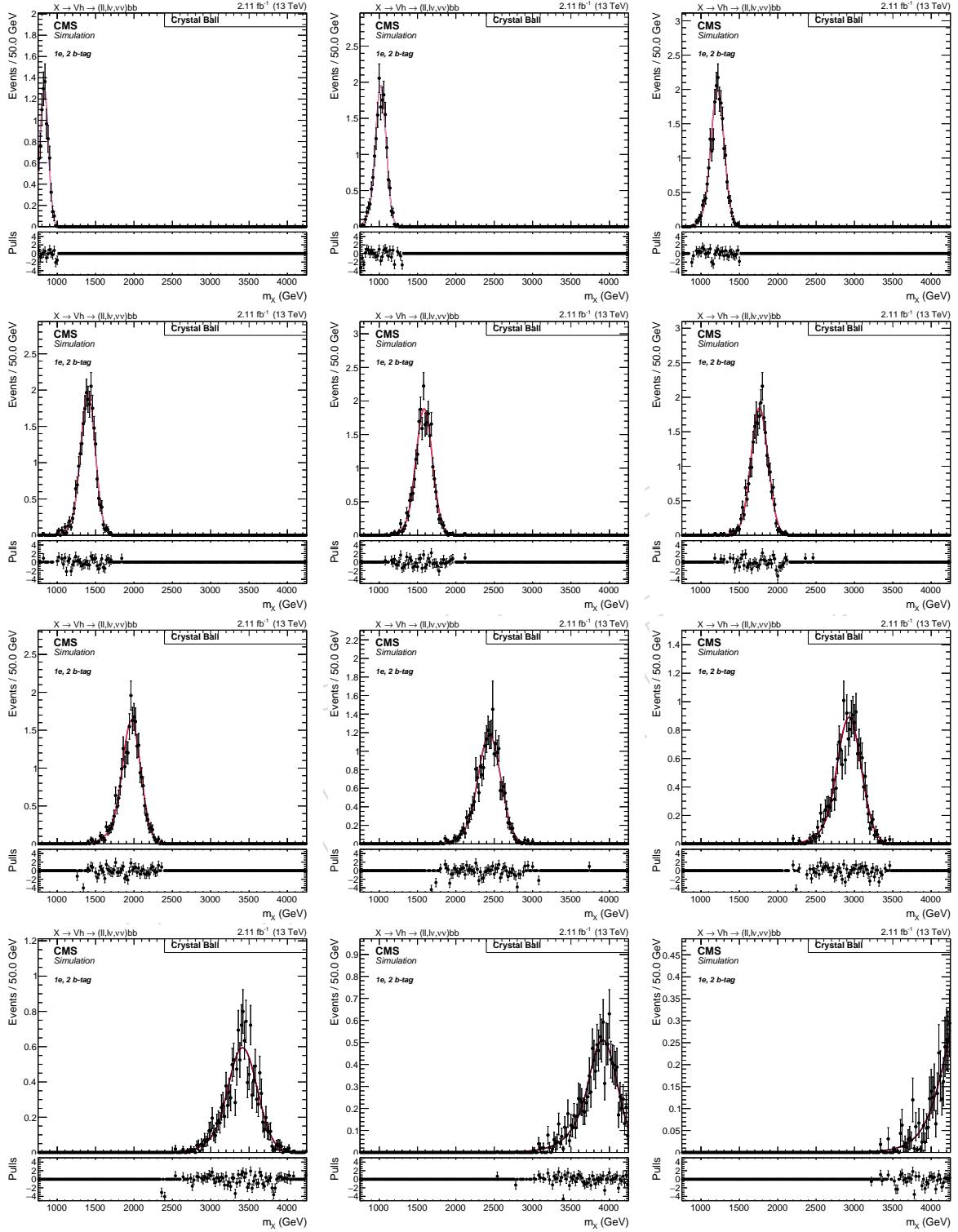


Figure 119: Fit to the signal samples in the 1 electron, 2 b-tag category.

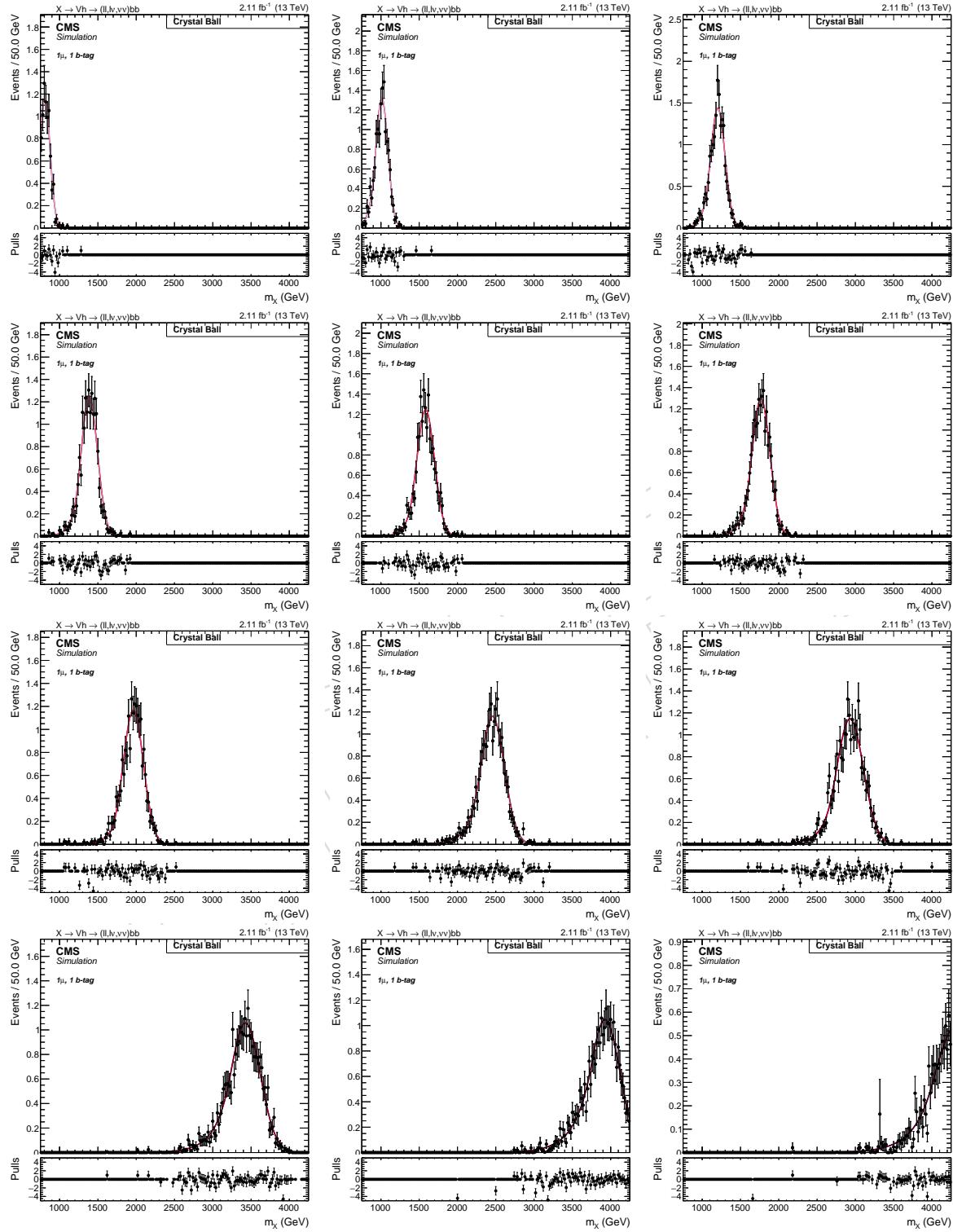


Figure 120: Fit to the signal samples in the 1 muon, 1 b-tag category.

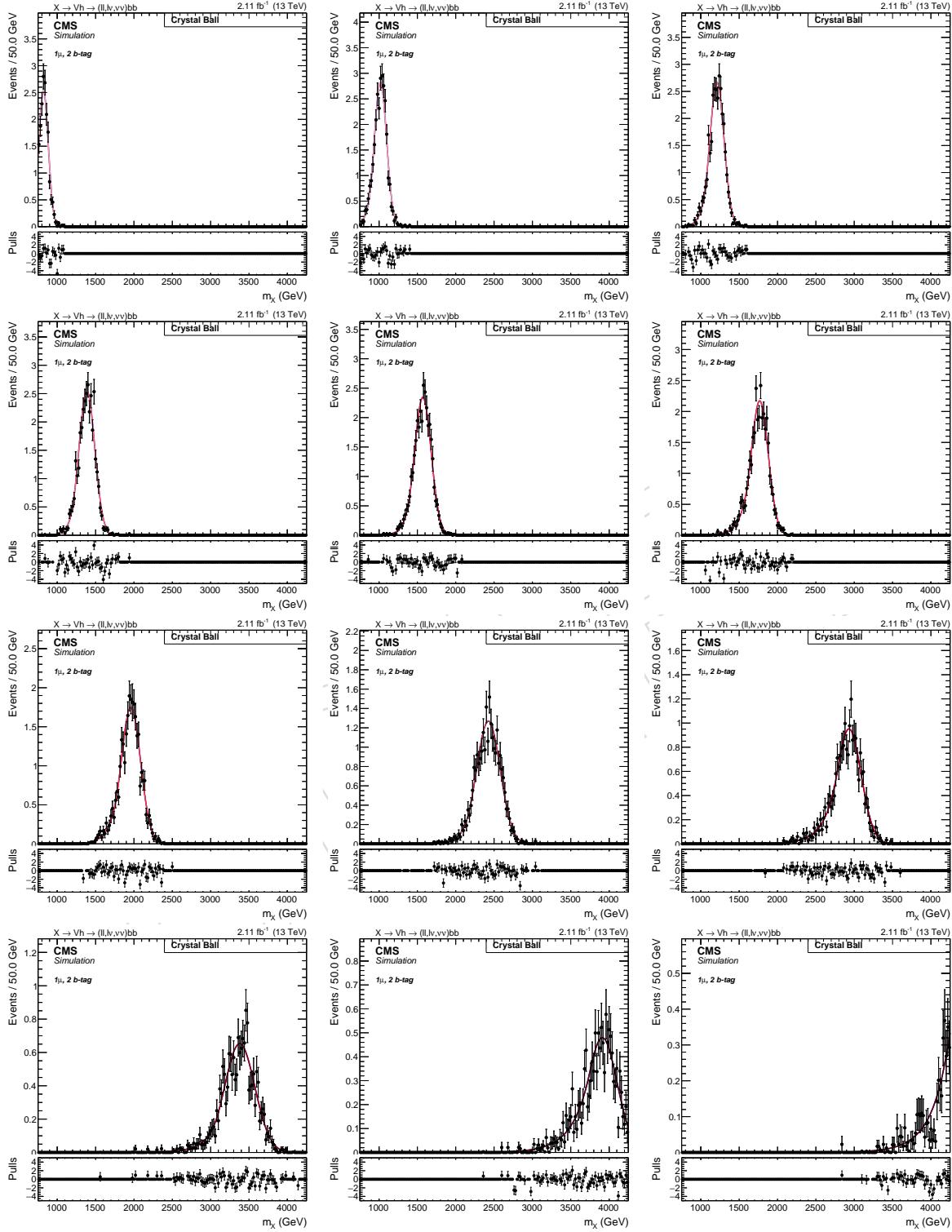


Figure 121: Fit to the signal samples in the 1 muon, 2 b-tag category.

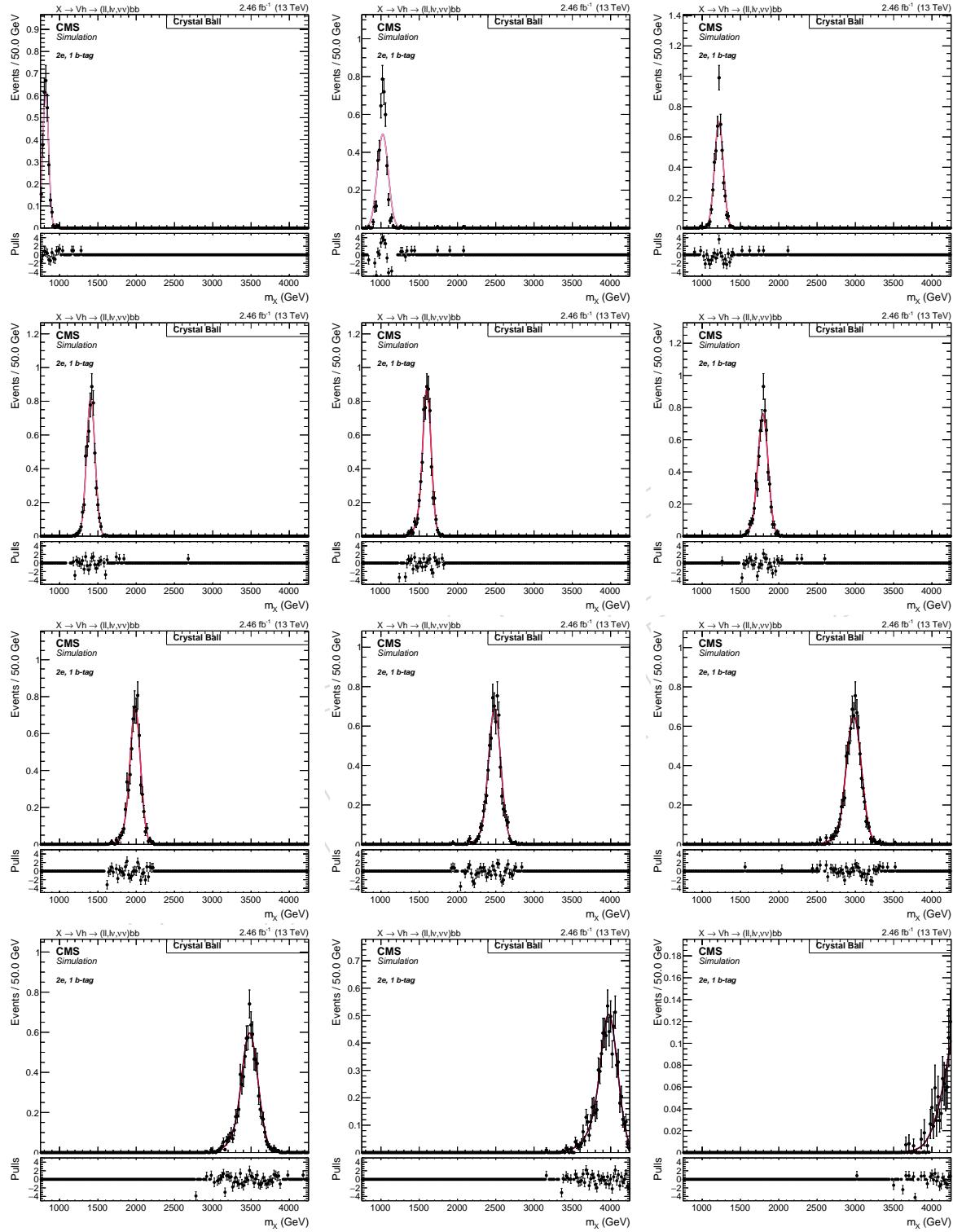


Figure 122: Fit to the signal samples in the 2 electrons, 1 b-tag category.

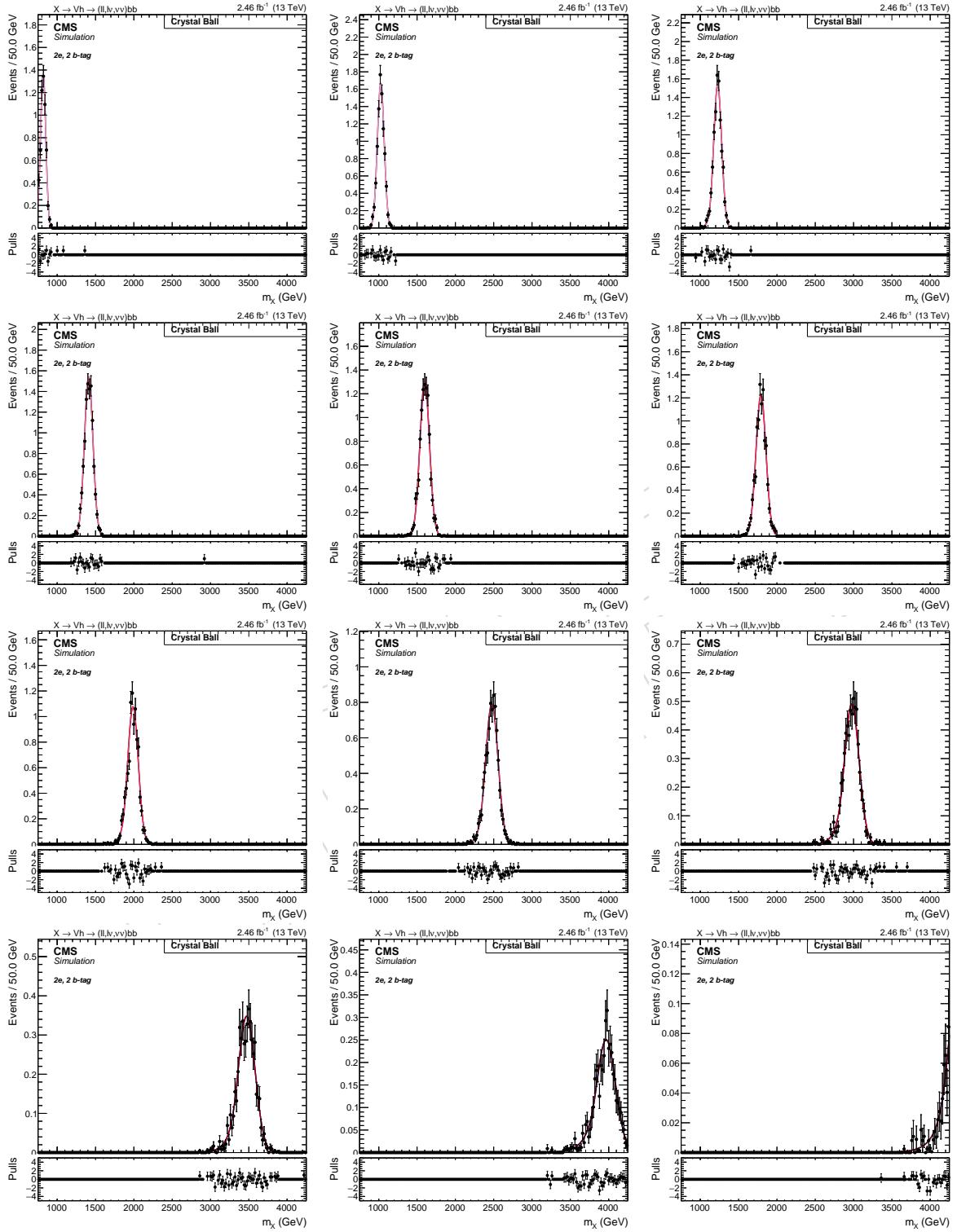


Figure 123: Fit to the signal samples in the 2 electrons, 2 b-tag category.

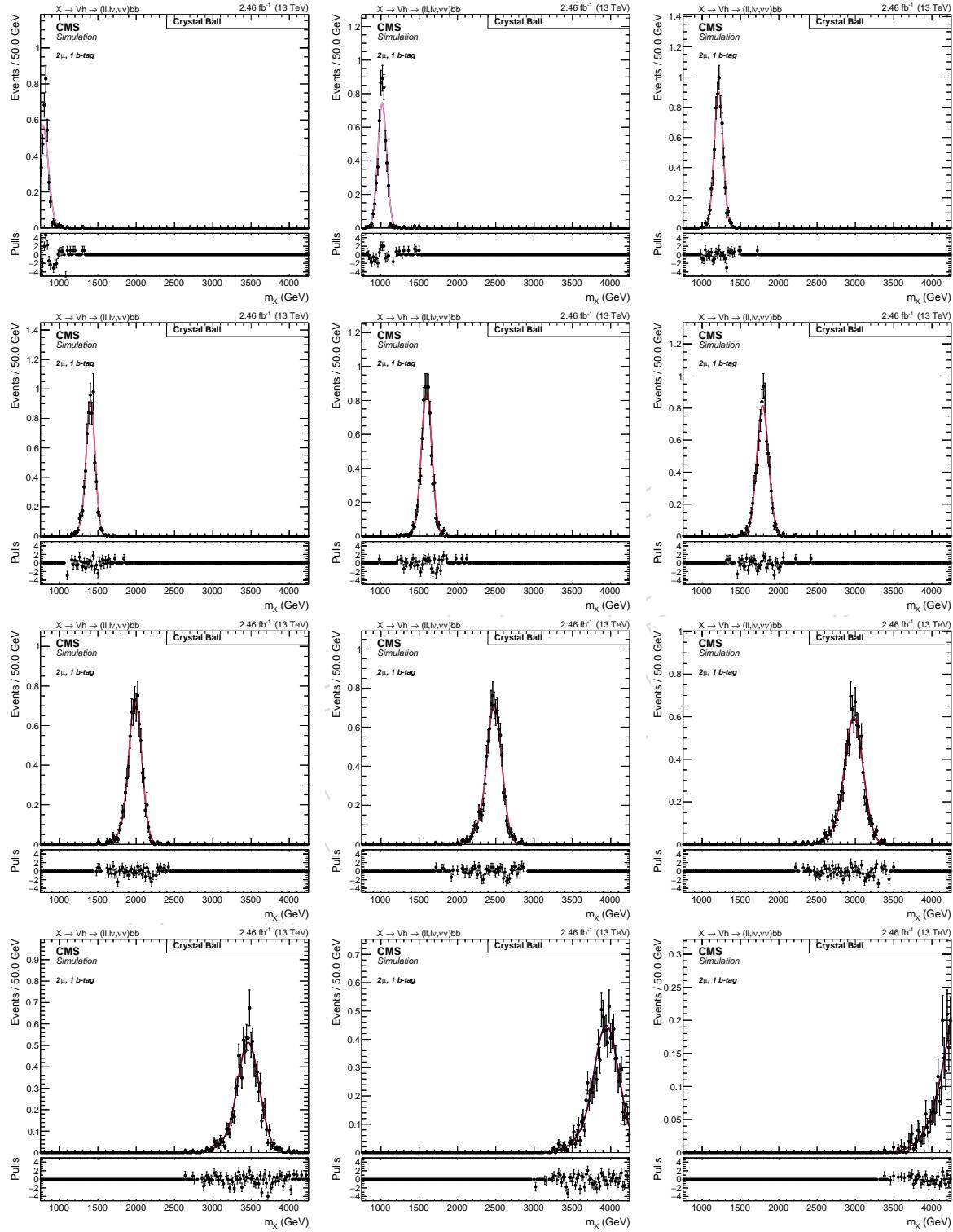


Figure 124: Fit to the signal samples in the 2 muons, 1 b-tag category.

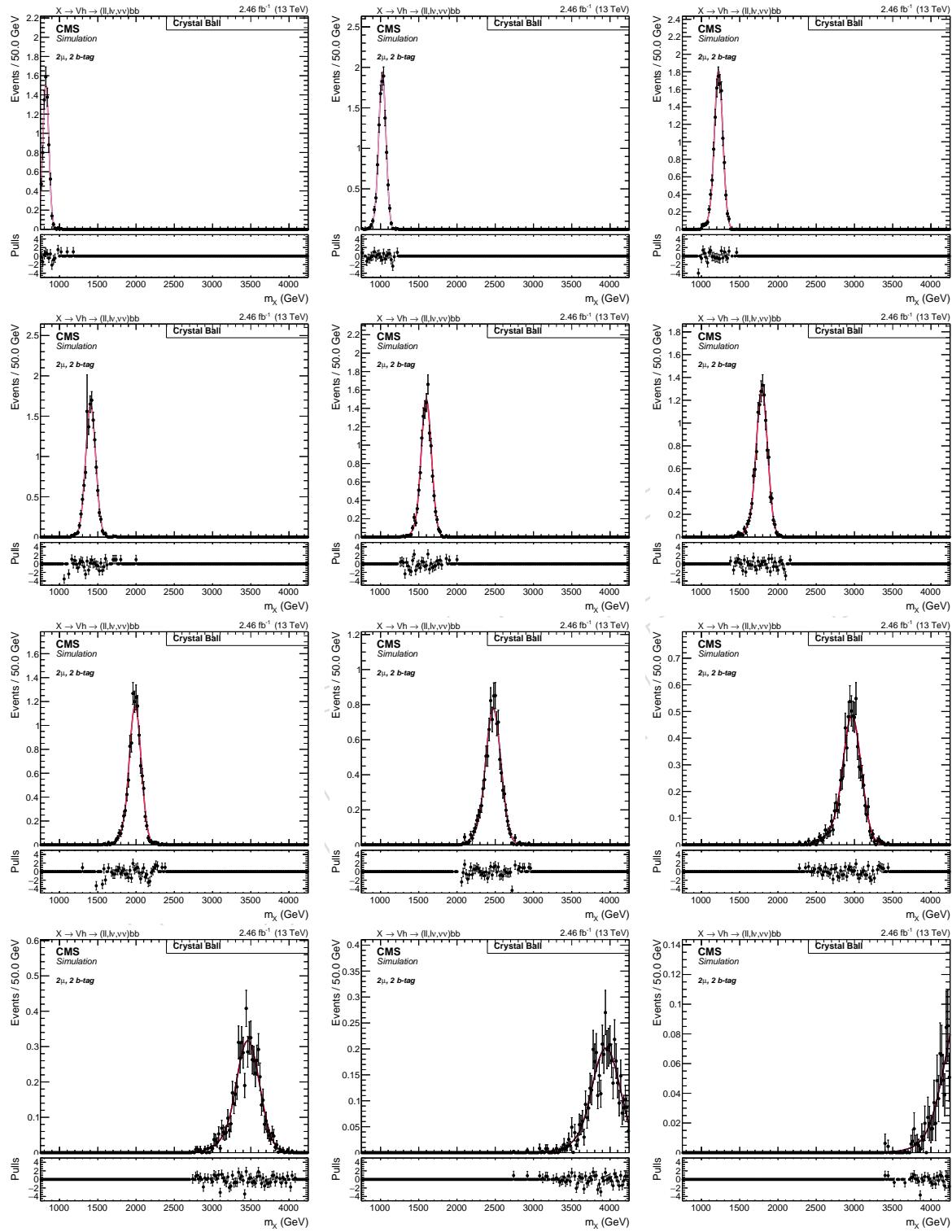


Figure 125: Fit to the signal samples in the 2 muons, 2 b-tag category.

1073 **8.5 Alpha method validation**

1074 As a validation of the  $\alpha$ -ratio method described in this Section, a closure test is performed  
 1075 on data. Instead of predicting the background in the real Higgs SR form both the lower and  
 1076 the upper jet mass sidebands, the SB and SR are redefined for the purposes of this test. The  
 1077 new pseudo-signal region (not blind) is now the upper jet sideband, and the new SB is just  
 1078 the lower sideband. With this configuration, the prediction of the background in the upper  
 1079 sideband is estimated exclusively in the lower sideband, and checked with data for both shape  
 1080 and normalization.

1081 In Figure 126 and Table 18, the predicted shape and normalizations are compared to the ob-  
 1082 served ones in data. This cross check confirms that the alpha method to extract the V +jets  
 1083 background is reliable and can be used to search for the signal extraction.

1084 The test is still in progress. In particular, the extrapolation from the lower to upper sidebands  
 1085 is more aggressive than extrapolating from the two SB in the SR, and the  $\alpha$  function is much  
 1086 more complicated to describe.

category		Expected	Observed
1 b-tag	0 $\ell$	$88.740 \pm 13.172$	68
	1e	$94.093 \pm 8.064$	98
	1 $\mu$	$192.464 \pm 13.723$	161
	2e	$8.112 \pm 2.743$	6
	2 $\mu$	$15.755 \pm 3.877$	18
2 b-tag	0 $\ell$	$16.508 \pm -$	14
	1e	$22.866 \pm 1.385$	16
	1 $\mu$	$56.102 \pm 6.536$	40
	2e	$1.227 \pm -$	2
	2 $\mu$	$3.579 \pm -$	2

Table 18: Expected and observed background yield in the upper jet mass sideband ( $m_j > 135$  GeV), predicted from the lower one ( $30 < m_j < 65$  GeV).

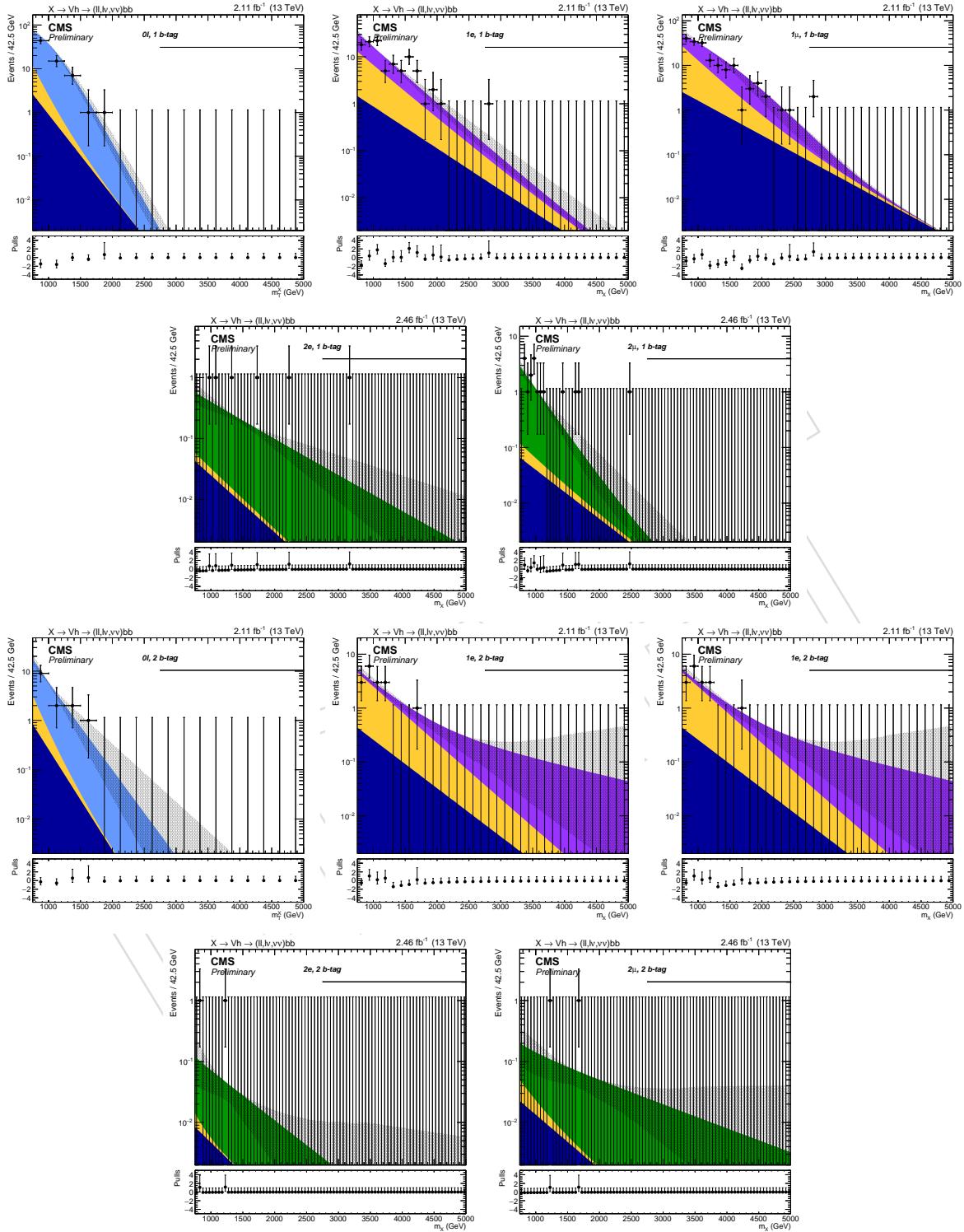


Figure 126: Expected and observed number of events in the upper jet mass sideband, extrapolated from the lower mass sideband.

## 1087 **9 Systematic uncertainties**

1088 The background and signal predictions are affected by systematic uncertainties that have to be  
 1089 estimated and taken into account in the signal fit. This section includes a list of the relevant  
 1090 systematic uncertainties for this analysis and how they are estimated. Most of the systematic  
 1091 uncertainties are dedicated to samples that are not normalized on data or the shape is taken  
 1092 from simulation.

### 1093 **9.1 Main background uncertainties**

#### 1094 **9.1.1 Normalization**

1095 The main background, V +jets, takes both the normalization and the shape from data. The  
 1096 normalization is extracted from fits to the jet mass sidebands with arbitrary functions tested  
 1097 on simulation. The contribution of the sub-dominant backgrounds is also taken into account,  
 1098 for both the normalization and the shape. The shape of these backgrounds is extracted from  
 1099 simulation, and then the parameters are fixed in the fit to data. The uncertainty of the main  
 1100 background normalization is then obtained by propagating the uncertainties affecting all the  
 1101 terms of the equation reported in Section 8.1 ( $N_{SB}^{Top}$ ,  $N_{SB}^{VV}$ ,  $N_{SR}^{Vjet}$ ,  $N_{SB}^{Vjet}$ ,  $N_{SR}^{Top}$ ,  $N_{SR}^{VV}$ ) that in turn  
 1102 depend on the parameters and the errors of the functions modeling the jet mass.

1103 The limited number of events in data in the sidebands ( $N_{SB}^{data}$ ) is treated separately as a source  
 1104 of statistical uncertainty.

1105 An additional uncertainty on the main background comparison comes from the fit with the  
 1106 alternative function. In this case, the difference in predicted number of events between the two  
 1107 function choices is taken as a systematic uncertainty.

1108 Numerical values are reported channel by channel in Table 16.

1109 The t̄ and single top normalization uncertainty, in the zero and single-lepton categories, orig-  
 1110 inates by the limited statistics in the control regions and from the variations of the b-tagging  
 1111 scale factor used to veto/select events in the Top CR. The values are reported in Table 14.

1112 The diboson normalization uncertainty, and the t̄ normalization in the double-lepton cate-  
 1113 gories, depends on the knowledge of the cross sections of these processes in the considered  
 1114 phase-space, and it is estimated to be 20%.

#### 1115 **9.1.2 Shape**

1116 The shape uncertainties are determined with the  $\alpha$ -method, explained in Section 8.2. The uncer-  
 1117 tainties on the parameters of the main background prediction in the signal regions are affected  
 1118 by the parameter error of the simultaneous fit to  $m_X$  in data in the jet mass sidebands, and from  
 1119 the  $\alpha$ -function itself, which depends on the  $m_X$  fits to the simulated V +jets distributions in SR  
 1120 and SB. These uncertainties are propagated to the shape of the main background in the signal  
 1121 region. Before being provided to the likelihood fit, these parameters are decorrelated through  
 1122 a linear transformation.

## 1123 **9.2 Triggers and Leptons**

1124 Trigger uncertainty due to the limited statistics is evaluated by shifting by one standard de-  
 1125 viation the trigger scale factors, as reported in Section 2.6. Additionally, a flat 2% systematic  
 1126 uncertainty is assigned for the electron trigger and 0.5% for the muon trigger as suggested by  
 1127 the corresponding POGs [20]. Identification and isolation systematics are evaluated by moving

1128 up and down the scale factors for identification and isolation by their uncertainties (Section 3)  
 1129 provided by POGs. For muons, additional flat uncertainties for identification and isolation,  
 1130 accounting for 1% and 1% respectively, are applied following the Muon POG prescription [20].  
 1131 Muon trigger, reconstruction, identification and isolation are responsible for a 7.5% and 8.5%  
 1132 normalization uncertainty in the single and double muon channels.  
 1133 For electrons, a 6.2% and 7.4% normalization uncertainties are estimated in the single and dou-  
 1134 ble electron channels accounting for trigger, reconstruction, identification and isolation.  
 1135 The diboson background uncertainty has similar values than those reported for the signal. No  
 1136 shape uncertainty is considered for trigger and lepton energy scale and resolution.  
 1137 The uncertainty assigned to signal samples and dibosons is 1% and 0.3% respectively for the  
 1138 HLT\_PFMETNoMu90\_NoiseCleaned\_PFMHTNoMu90\_IDTight trigger in the 0-lepton categories,  
 1139 after the  $E_T^{\text{miss}} > 200 \text{ GeV}$  selections. The trigger uncertainties for electron, muon and  $E_T^{\text{miss}}$  trig-  
 1140 gers are considered as uncorrelated.

### 1141 9.3 Jet momentum

1142 Jets uncertainties are evaluated in the signal regions by moving up and down by one standard  
 1143 deviation the source of the uncertainty. The two sources are the uncertainty on the jet energy  
 1144 correction, also identified as jet energy scale (JES), and the uncertainty due to the different jet  
 1145 momentum resolution (JER). For the jet energy scale the  $p_T$  of the jets are shifted by the error  
 1146 value of the jet energy corrections. Since jet energy corrections are applied to the jet mass, the  
 1147 effect of the JES uncertainty is evaluated also for this quantity. The difference in the normaliza-  
 1148 tion of the jet energy correction is propagated to the signal region, without taking into account  
 1149 the effect on the  $E_T^{\text{miss}}$  (its uncertainty is estimated separately). The resulting normalization un-  
 1150 certainty is  $\pm 5\%$  for diboson background, and ranges from  $\pm 1\%$  (1 TeV) to  $\pm 3\%$  (4 TeV) for the  
 1151 signal samples.  
 1152 The JEC uncertainties also impact the signal shape, and specifically the mean and width of the  
 1153 Crystal Ball. The uncertainty is found to be stable with the resonance mass between channels  
 1154 and categories, and is found to be 0.3% for the mean, and 1.0% for the width.

### 1155 9.4 Higgs tagging efficiency

1156 An uncertainty of 7% is associated to the uncertainty on the way to tag a Higgs boson. It does  
 1157 not include the btagging scale factors (discussed in Section 9.5) and the pileup uncertainty. It  
 1158 includes the uncertainty due to selection on the pruned mass of the Higgs-jet only.  
 1159 The uncertainties on the W and top mass tagging are well estimated using semileptonic  $t\bar{t}$   
 1160 sample; the semileptonic  $t\bar{t}$  sample provides a good source of boosted hadronic W and boosted  
 1161 top. The W mass peak in MC has been found to be consistent with the mass peak in data within  
 1162 statistical uncertainty; the data/MC scale factor is  $0.992 \pm 0.005$  [50]. On the contrary we have  
 1163 no pure source of boosted Higgs. Therefore, we use the following technique to have a first  
 1164 estimate of the associated uncertainty.

1165 We perform a double ratio estimate between bulk graviton going to WW and hh. Due to the  
 1166 lack of HERWIG++ MC samples with HVT models, we choose to use the bulk graviton sam-  
 1167 ples for this study. We choose a mass window for W and h. We calculate the ratio of W-mass  
 1168 and h-mass efficiencies for PYTHIA and HERWIG showering algorithms and obtain  $R_{\text{HERWIG}}$  and  
 1169  $R_{\text{PYTHIA}}$ , respectively. Then we calculate the double ratio  $R_{\text{HERWIG}}/R_{\text{PYTHIA}}$ . The double ratio pro-  
 1170 vides an estimate how different showering algorithms reacts on the difference between W and

1171 h decays to jets.

1172 The values of the double ratios are provided in Table 19 for W mass  $65 < m_W < 105$  GeV  
 1173 (large window), and Table 20 for W mass  $65 < m_W < 85$  GeV (window defined to separate  
 1174 W from  $Z^0$ ). In the first case the difference is approximately of 7% per jet. While in second  
 1175 case a smaller difference of 2% is observed. One may observe that the second case corresponds  
 1176 to a more similar efficiency between W and h tagging. The source of the difference may be  
 1177 observed in the Fig. 127: HERWIG predicts a larger right tail than PYTHIA. We verified that this  
 1178 difference does not come from the L2L3 corrections.

1179 Consequently we assign an overall uncertainty of 7% to cover the difference between the po-  
 1180 tential differences between W tagging point, where the SF was derived, and h tagging point.

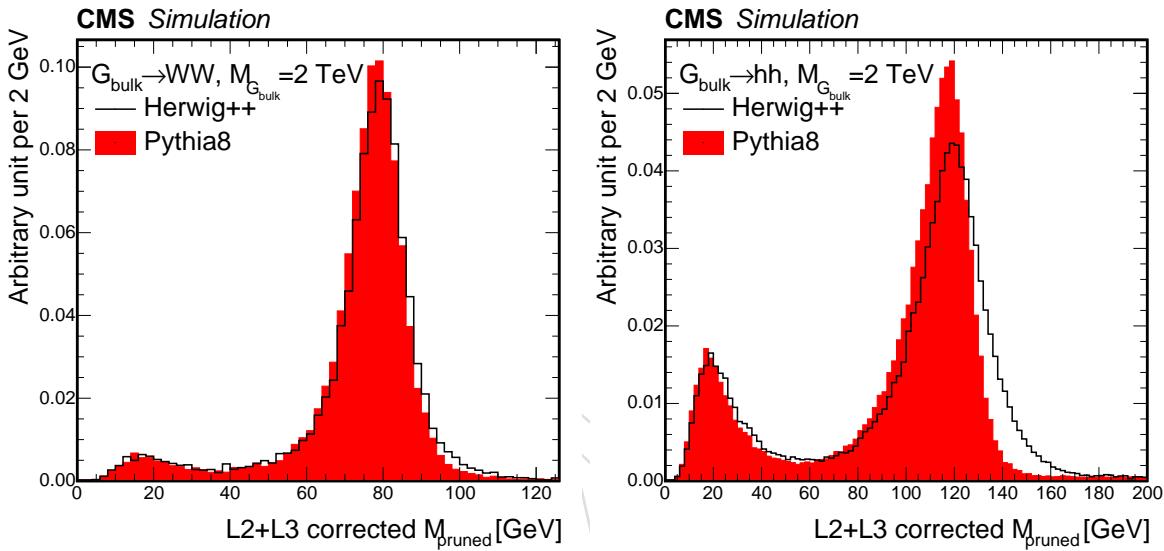


Figure 127: Comparison of W (left) and h (right) pruned masses after L2L3 corrections for PYTHIA and HERWIG at  $M_{\text{bulk}} = 2$  TeV.

## 1181 9.5 b-tagging

1182 The b-tagging is a fundamental tool for the present analysis, and its uncertainty represent the  
 1183 main systematic uncertainty source for samples that are not normalized on data (especially  
 1184 signal). The impact of b-tagging uncertainty is evaluated by moving the CSV scale factors  
 1185 provided by the BTV POG [43] in the reshaping procedure (Section 3.8). An average systematic  
 1186 uncertainty of 6% per b-jet, 12% per c-jet, and 15% per fake tag (light quarks and gluons) are  
 1187 used; the exact value is assigned for each jet as a function of its  $p_T$  and  $\eta$ . In order to evaluate  
 1188 the impact of the scale factors uncertainties, the reshaping procedure is repeated with the scale  
 1189 factors shifted up and down by one standard deviation. No relevant shape variations have  
 1190 been observed; the uncertainty on the overall normalization is fairly linear on the signal mass  
 1191 hypothesis, and corresponds to approximately  $\pm 15\%$  at 1 TeV up to 4% at 4 TeV for the single  
 1192 b-tagged category, and reveals an opposite trend for the double b-tagged category, where it ranges  
 1193 between 8% and 30%. A 5% and 12% uncertainty is assigned to diboson samples in the single  
 1194 and double b-tagged categories, respectively. As the other uncertainties, b-tagging variations  
 1195 are considered as correlated. No shape uncertainty due to b-tagging is considered.

1196 **9.6 Missing Energy**

1197 The  $E_T^{\text{miss}}$  is a composed object, built upon all the reconstructed particles in the detector. There-  
 1198 fore, it is affected by the energy scale and resolution of all the reconstructed objects, i.e. charged  
 1199 and neutral hadrons (clustered in jets and non-clustered), muons, electrons, photons and hadronic  
 1200 taus. Dedicated uncertainties have been derived by propagating the original object scales and  
 1201 resolutions to the  $E_T^{\text{miss}}$  itself. However, it is not clear if these effects can effectively cover the  
 1202 possible  $E_T^{\text{miss}}$  discrepancies between data and simulation.

1203 The alternative approach adopted in this analysis relies on data to derive  $E_T^{\text{miss}}$  corrections and  
 1204 uncertainties, as described in section 3.9.1. The corrections are evaluated in dedicated regions  
 1205 where a Z boson can be reconstructed in dileptonic final states, and the  $E_T^{\text{miss}}$  is decomposed in  
 1206 two components parallel ( $U_{\parallel}$ ) and orthogonal ( $U_{\perp}$ ) to the flight direction of the vector boson.  
 1207 The uncertainty on the scale and resolution of the  $U_{\parallel}$  component, together with the uncertainty  
 1208 on the  $U_{\perp}$  resolution, are then propagated to the missing energy corrections, and finally to the  
 1209 candidate mass in the 0- and 1-lepton channels. After the recoil corrections, the  $E_T^{\text{miss}}$  scale and  
 1210 resolution uncertainties are found to be small. (to be checked)

1211 **9.7 Pile-up**

1212 An additional source of systematic error is the limited knowledge of the total inelastic cross-  
 1213 section at 8 TeV, used to get the expected primary vertices distribution used for pile-up reweight-  
 1214 ing. A 5% uncertainty is assumed for the default value of 69 mb [23], and the expected primary  
 1215 vertices distributions are varied accordingly. Changing the pile-up weight varies also the MC  
 1216 normalization in the signal regions, and the relative difference is estimated to be 3% for the  
 1217 diboson sample, and a flat 0.5% for signal samples. No shape uncertainties are considered for  
 1218 PU.

1219 **9.8 QCD renormalization and factorization scale**

1220 Per-event weights are provided for a variation of the QCD renormalization and factorization  
 1221 scales by a factor 2. The two scales can be varied separately and independently, or together  
 1222 assuming 100% correlation. The weight is propagated up to the final distributions, accounting  
 1223 for both shape and normalization uncertainties. The envelope of all the considered variations  
 1224 is then considered as the template for the scale uncertainty.

1225 The normalization uncertainty due to the QCD scales amounts to 5% for the diboson back-  
 1226 ground, and ranges between 4% at 1 TeV and 12% at 4 TeV for signal samples. It is also ob-  
 1227 served that the uncertainty in signal are very consistent between all the different channels. The  
 1228 QCD scales also have negligible effect on the mean and sigma of the Crystal Ball (< 0.1%), and  
 1229 no shape uncertainties are considered for the signal.

1230 **9.8.1 PDF**

1231 Systematic uncertainties coming from PDF uncertainties have been considered for this analysis,  
 1232 according to the PDF4LHC prescriptions, and using the NNPDF3.0 set. The 100 weights have  
 1233 been considered together, by calculating the envelope of the weight distribution around the  
 1234 central value for each distribution, and propagated as a normalization and shape uncertainty  
 1235 to the final distributions. The effect of the PDF uncertainty on the acceptance is found to be  
 1236 consistent between all the  $X \rightarrow Zh$  and  $Z \rightarrow Wh$  signal samples, and it starts from 2.5% at  
 1237 1 TeV, grows up to 3.5% at 2 TeV, and then decreases to 1.5% at 4 TeV. A flat 3% is then taken  
 1238 for all the signal samples, and 0.5% for dibosons.

**1239 9.9 Summary**

1240 A summary of all systematics is listed in Tab. 21. In addition to those described in the pre-  
1241 vious sections, a 15% uncertainty on normalization is assumed on the cross-section value of  
1242 non-data-driven Monte Carlo backgrounds from CMS measurements (see Sec. 2.3), and a 4.6%  
1243 uncertainty for luminosity.

DRAFT

Table 19: The per-jet efficiency of requiring the mass of Higgs (W) jets to be within 105–135 (65–105) GeV. The efficiency is evaluated with the  $G_{\text{bulk}} \rightarrow hh(\text{WW})$  samples. Each AK8 jet is required to match to the generator-level boson within a  $\Delta R$  of 0.4.

$M_{G_{\text{bulk}}} [\text{GeV}]$	$\epsilon_{hh}^{\text{HERWIG}}$	$\epsilon_{WW}^{\text{HERWIG}}$	$\epsilon_{hh}^{\text{HERWIG}} / \epsilon_{WW}^{\text{HERWIG}}$	$\epsilon_{hh}^{\text{PYTHIA}}$	$\epsilon_{WW}^{\text{PYTHIA}}$	$\epsilon_{hh}^{\text{PYTHIA}} / \epsilon_{WW}^{\text{PYTHIA}}$	$R_{\text{HERWIG}} / R_{\text{PYTHIA}}$
1000	$0.454 \pm 0.002$	$0.811 \pm 0.003$	$0.559 \pm 0.0028$	$0.494 \pm 0.002$	$0.817 \pm 0.002$	$0.605 \pm 0.0027$	$0.924 \pm 0.0062$
2000	$0.491 \pm 0.002$	$0.792 \pm 0.003$	$0.619 \pm 0.0031$	$0.534 \pm 0.002$	$0.8 \pm 0.002$	$0.667 \pm 0.0028$	$0.928 \pm 0.0061$
3000	$0.462 \pm 0.002$	$0.765 \pm 0.003$	$0.604 \pm 0.0032$	$0.505 \pm 0.002$	$0.774 \pm 0.003$	$0.653 \pm 0.003$	$0.925 \pm 0.0065$

Table 20: The per-jet efficiency of requiring the mass of Higgs (W) jets to be within 105–135 (65–85) GeV. The efficiency is evaluated with the  $G_{\text{bulk}} \rightarrow hh(\text{WW})$  samples. Each AK8 jet is required to match to the generator-level boson within a  $\Delta R$  of 0.4.

$M_{G_{\text{bulk}}} [\text{GeV}]$	$\epsilon_{hh}^{\text{HERWIG}}$	$\epsilon_{WW}^{\text{HERWIG}}$	$\epsilon_{WW}^{\text{HERWIG}} / \epsilon_{hh}^{\text{HERWIG}}$	$\epsilon_{hh}^{\text{PYTHIA}}$	$\epsilon_{WW}^{\text{PYTHIA}}$	$\epsilon_{hh}^{\text{PYTHIA}} / \epsilon_{WW}^{\text{PYTHIA}}$	$R_{\text{HERWIG}} / R_{\text{PYTHIA}}$
1000	$0.454 \pm 0.002$	$0.54 \pm 0.004$	$0.84 \pm 0.0065$	$0.494 \pm 0.002$	$0.581 \pm 0.003$	$0.851 \pm 0.0053$	$0.987 \pm 0.0098$
2000	$0.491 \pm 0.002$	$0.632 \pm 0.004$	$0.777 \pm 0.0051$	$0.534 \pm 0.002$	$0.674 \pm 0.003$	$0.792 \pm 0.0041$	$0.981 \pm 0.0082$
3000	$0.462 \pm 0.002$	$0.591 \pm 0.004$	$0.781 \pm 0.0055$	$0.505 \pm 0.002$	$0.631 \pm 0.003$	$0.801 \pm 0.0045$	$0.976 \pm 0.0088$

Table 21: Summary of systematic uncertainties for the backgrounds and signal samples.

	shape	Main	Top	VV	Signal
$\alpha$ -function	✓	✓	-	-	-
Bkg. normalization		10 – 35%	-	-	-
Top scale factors		4 – 25%	-	-	-
jet energy scale	✓	-	✓	5%	1 – 3%
Higgs tagging		-	-	-	7%
b-tagging		-	4.9%	5% (1b), 12% (2b)	15 – 4% (1b), 8 – 30% (2b)
MET scale and res.	✓	-	-	1%	1%
leptons and trigger		-	-	7.5 – 8.5% ( $\mu$ ), 6.2 – 7.5% ( $e$ ), 1% ( $0\ell$ )	0.5%
pile-up		-	-	3%	
QCD scales	✓	-	✓	5%	4 – 12%
PDF		-	-	0.5%	3%
luminosity		-	-	20%	4.6%
cross section		-	-		-

## 1244 10 Results

1245 The  $CL_s$  criterion [51, 52] is used to determine the 95% confidence-level limit on the signal  
 1246 contribution in the data, using the `RooStats` package [53]. In order to extract the limit on the  
 1247 production cross section times the branching ratios, the CMS standard `combine` tool [54] has  
 1248 been used. The `Asymptotic` method is used to calculate preliminary 95% C.L. upper limits  
 1249 with  $1\sigma$  and  $2\sigma$  bands using the CLs frequentist calculation currently recommended by the LHC  
 1250 Higgs Combination Group [55]. The `ProfileLikelihood` method is used for significance  
 1251 and the background p-value; finally, the `MaxLikelihoodFit` method allows to get the signal  
 1252 Best Fit Ratio, the fit pulls and the pre/post fit distributions.

### 1253 10.1 Signal extraction strategy

1254 There are two possible alternatives to extract the signal. Both are fully independent, complete  
 1255 and covered with the appropriate systematic uncertainties.

- 1256 1. **Alpha method (A):** this is the *main extraction method*. The background prediction is es-  
 1257 timated with the  $\alpha$  ratio method, explained in Section 8. The background normalization  
 1258 is taken from data, and the shape from the data in the SB and multiplied by a transfer  
 1259 function extracted from simulation. Most of the systematic uncertainties are covered and  
 1260 included in the  $\alpha$  function. The signal extraction fit is unbinned.
- 1261 2. **Binned method (B):** this represents the *backup and cross-check* method. The background  
 1262 normalization is taken from data, but the shape are entirely taken from simulation, and  
 1263 validated in appropriate control regions. The complete set of systematic uncertainties on  
 1264 the shape is applied. The signal extraction fit is performed on histograms.

1265 The selected method to extract the signal strength or the exclusion limit is the **alpha ratio**  
 1266 **method**. The likelihood fit is performed on each channel (different lepton flavour and num-  
 1267 ber, b-tag multiplicity), and finally all the  $X \rightarrow Zh$  and  $X \rightarrow Wh$  channels are combined in  
 1268 order to present a global limit on the production cross section of the two signal processes.

### 1269 10.2 Results of the alpha method

1270 The expected exclusion limits, with the  $\pm 1$  and  $\pm 2\sigma$  bands, obtained with the  $\alpha$  method (A) are  
 1271 reported in Figure 136 in terms of upper limits on the signal cross section.

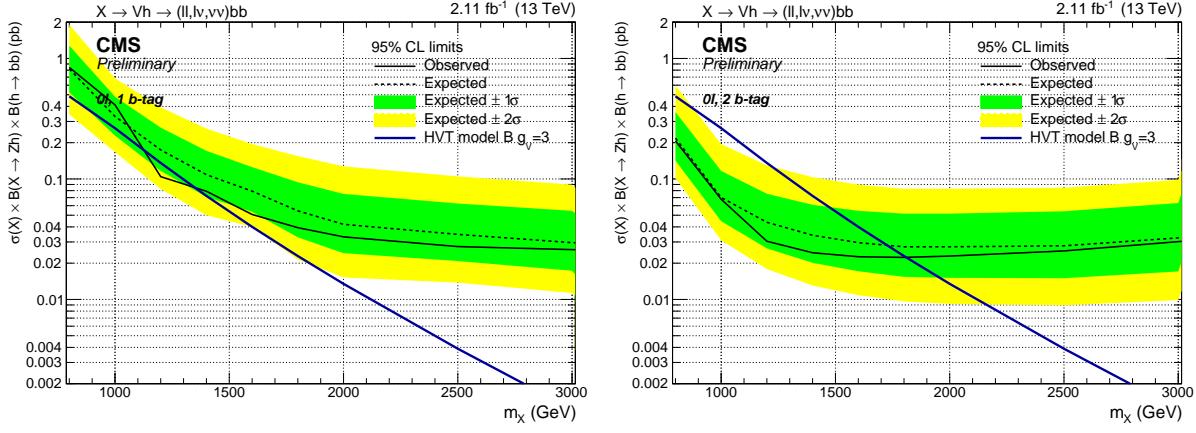


Figure 128: Observed and expected (with  $\pm 1(2)\sigma$  band) 95% C.L. upper limit on  $\sigma \times \mathcal{B}(X \rightarrow Zh)$  in the zero lepton, 1 b-tag (left) and 2 b-tag (right) categories, including all statistical and systematics uncertainties. Results are extracted with the  $\alpha$  method.

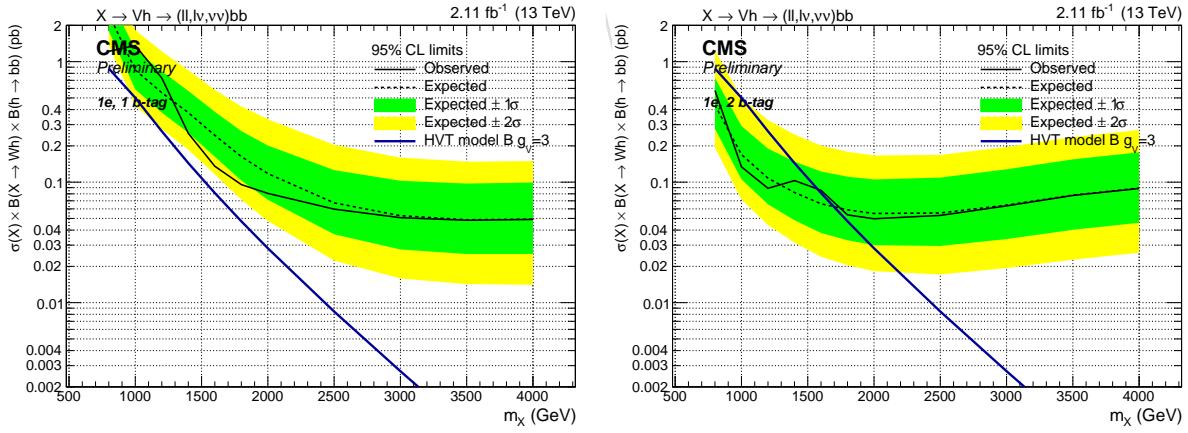


Figure 129: Observed and expected (with  $\pm 1(2)\sigma$  band) 95% C.L. upper limit on  $\sigma \times \mathcal{B}(X \rightarrow Zh)$  in the 1-electron, 1 b-tag (left) and 2 b-tag (right) categories, including all statistical and systematics uncertainties. Results are extracted with the  $\alpha$  method.

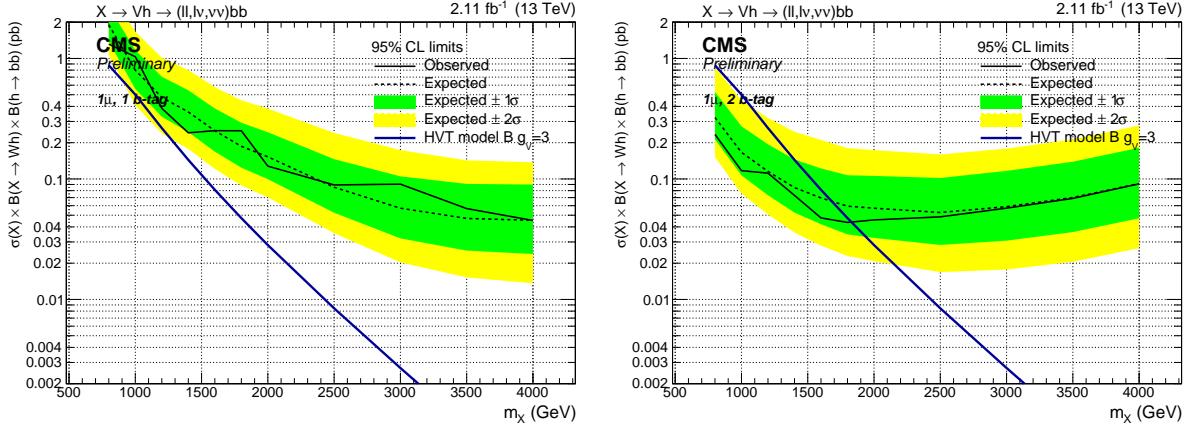


Figure 130: Observed and expected (with  $\pm 1(2)\sigma$  band) 95% C.L. upper limit on  $\sigma \times \mathcal{B}(X \rightarrow Zh)$  in the 1-muon, 1 b-tag (left) and 2 b-tag (right) categories, including all statistical and systematics uncertainties. Results are extracted with the  $\alpha$  method.

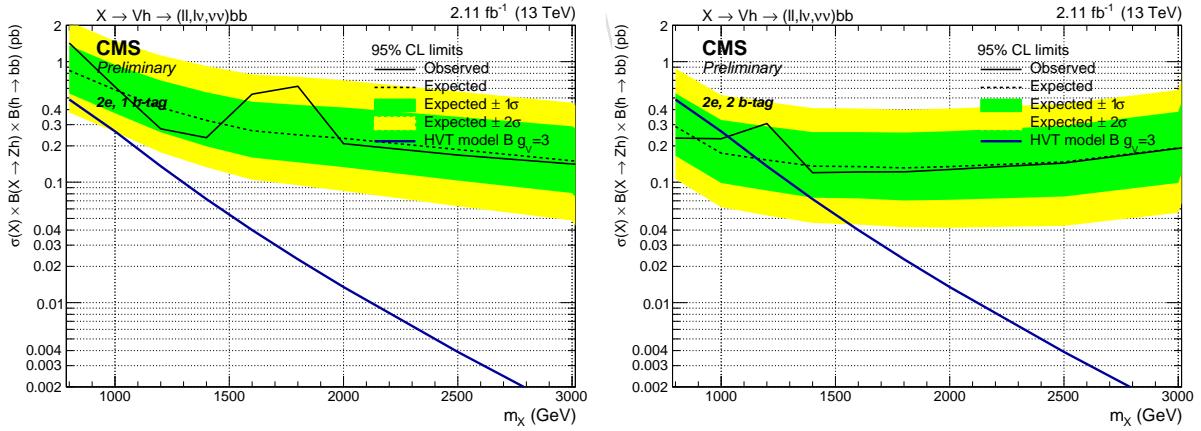


Figure 131: Observed and expected (with  $\pm 1(2)\sigma$  band) 95% C.L. upper limit on  $\sigma \times \mathcal{B}(X \rightarrow Zh)$  in the 2-electrons, 1 b-tag (left) and 2 b-tag (right) categories, including all statistical and systematics uncertainties. Results are extracted with the  $\alpha$  method.

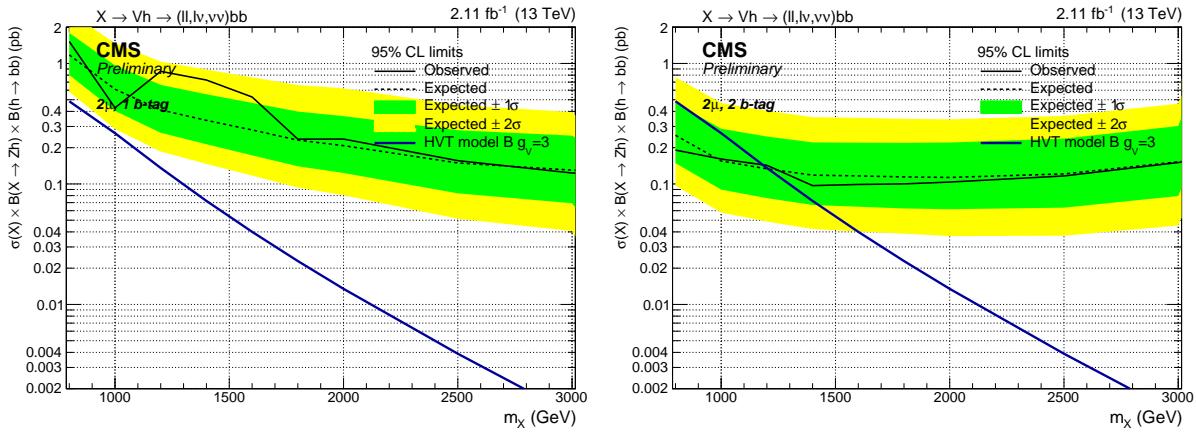


Figure 132: Observed and expected (with  $\pm 1(2)\sigma$  band) 95% C.L. upper limit on  $\sigma \times \mathcal{B}(X \rightarrow Zh)$  in the 2-muons, 1 b-tag (left) and 2 b-tag (right) categories, including all statistical and systematics uncertainties. Results are extracted with the  $\alpha$  method.

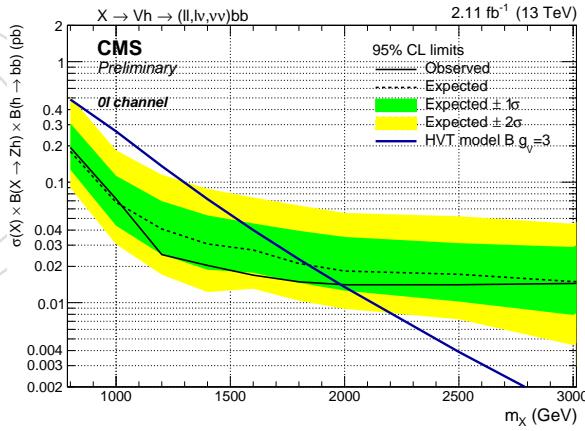


Figure 133: Observed and expected (with  $\pm 1(2)\sigma$  band) 95% C.L. upper limit on  $\sigma \times \mathcal{B}(X \rightarrow Zh)$  for the combination of 1 and 2 b-tag categories in the 0-leptons final state as a function of  $m_X$ , including all statistical and systematics uncertainties. Results are extracted with the  $\alpha$  method.

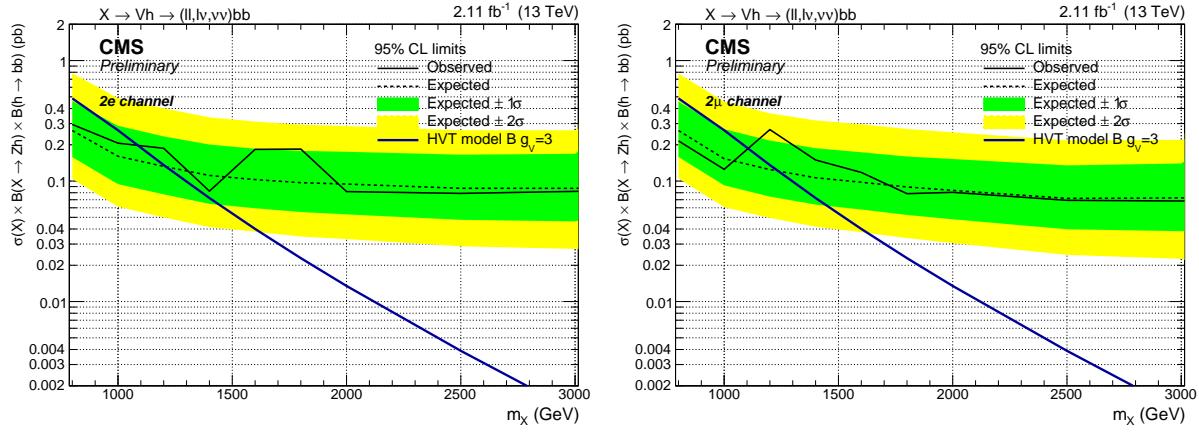


Figure 134: Observed and expected (with  $\pm 1(2)\sigma$  band) 95% C.L. upper limit on  $\sigma \times \mathcal{B}(X \rightarrow Zh)$  for the combination of the 1 and 2 b-tag categories in the 2-electrons (left) and 2-muons (right) final state as a function of  $m_X$ , including all statistical and systematics uncertainties. Results are extracted with the  $\alpha$  method.

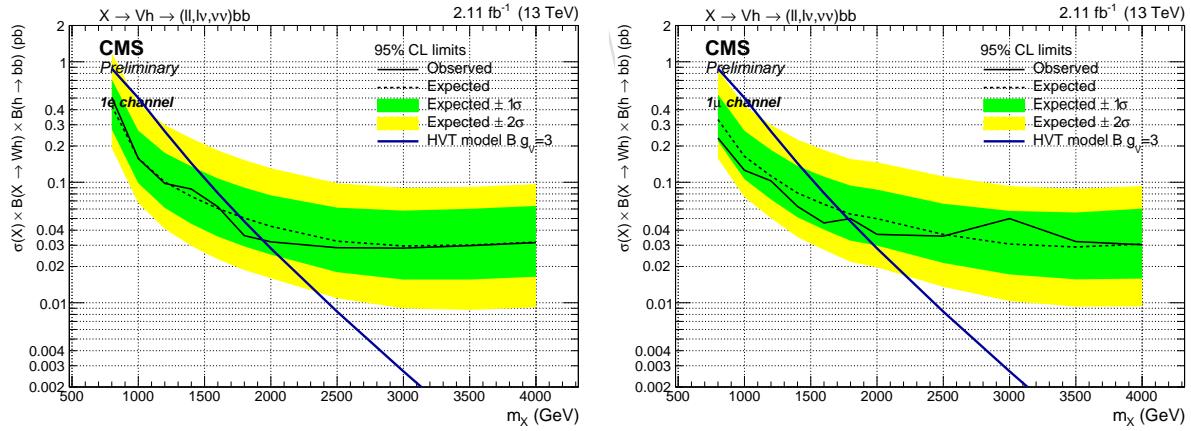


Figure 135: Observed and expected (with  $\pm 1(2)\sigma$  band) 95% C.L. upper limit on  $\sigma \times \mathcal{B}(X \rightarrow Wh)$  for the combination of the 1 and 2 b-tag categories in the 1-electron (left) and 1-muon (right) final state as a function of  $m_X$ , including all statistical and systematics uncertainties. Results are extracted with the  $\alpha$  method.

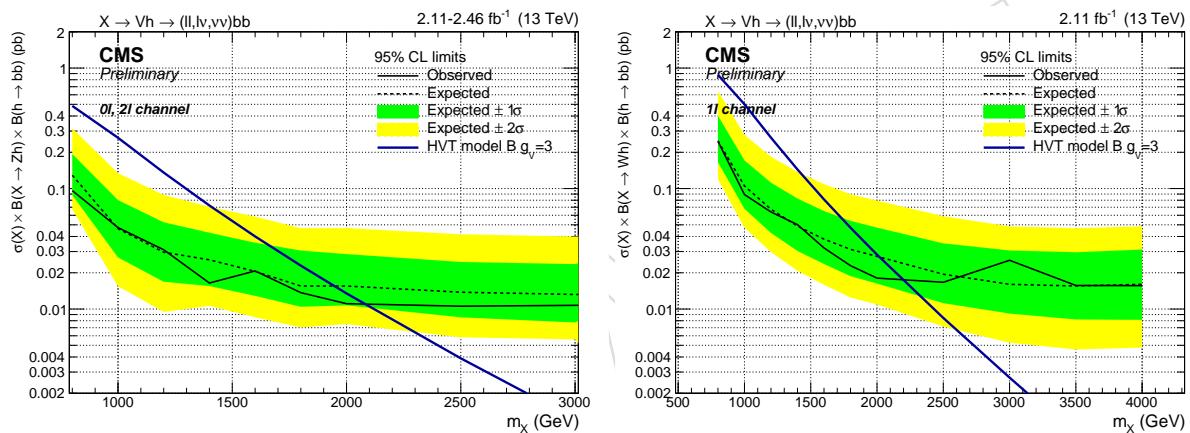


Figure 136: Observed and expected (with  $\pm 1(2)\sigma$  band) 95% C.L. upper limit on  $\sigma \times \mathcal{B}(X \rightarrow \text{Zh})$  (left) and  $\sigma \times \mathcal{B}(X \rightarrow \text{Wh})$  (right) as a function of  $m_X$ , including all statistical and systematics uncertainties. Results are extracted with the  $\alpha$  method.

**1272 11 Conclusions**

1273 This Analysis Note described a search for heavy resonances with mass larger than 1 TeV decay-  
1274 ing into a vector boson and a Higgs boson, using the data collected at  $\sqrt{s} = 13$  TeV during the  
1275 2015 operations by the CMS experiment at LHC Run-2. The final states explored include all the  
1276 leptonic decay modes of the vector boson, in events with zero ( $Z \rightarrow \nu\bar{\nu}$ ), exactly one  $W \rightarrow \ell\nu$   
1277 and two  $Z \rightarrow \ell\ell$  leptons. The only considered Higgs decay channel is the hadronic decay into  
1278 a pair of b-quarks. Using the available dataset of  $2.1 - 2.5 \text{ fb}^{-1}$ , an upper limit is to  $200 \sim 20 \text{ fb}$   
1279 on the cross section times branching ratio of exotic resonances, depending on the mass.

DRAFT

## 1280 References

- [1] ATLAS Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", *Phys. Lett. B* **716** (2012) 1, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.
- [2] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys. Lett. B* **716** (2012) 30, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.
- [3] CMS Collaboration, "Observation of a new boson with mass near 125 GeV in pp collisions at  $\sqrt{s} = 7$  and 8 TeV", *JHEP* **06** (2013) 081, doi:10.1007/JHEP06(2013)081, arXiv:1303.4571.
- [4] CMS Collaboration, "Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV", *Eur. Phys. J. C* **75** (2015) 212, doi:10.1140/epjc/s10052-015-3351-7, arXiv:1412.8662.
- [5] ATLAS Collaboration, "Measurement of the Higgs boson mass from the  $H \rightarrow \gamma\gamma$  and  $H \rightarrow ZZ^* \rightarrow 4\ell$  channels in  $pp$  collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector", *Phys. Rev. D* **90** (2014) 052004, doi:10.1103/PhysRevD.90.052004, arXiv:1406.3827.
- [6] ATLAS Collaboration, "Evidence for the spin-0 nature of the Higgs boson using ATLAS data", *Phys. Lett. B* **726** (2013) 120, doi:10.1016/j.physletb.2013.08.026, arXiv:1307.1432.
- [7] CMS and ATLAS Collaborations, "Combined Measurement of the Higgs Boson Mass in  $pp$  Collisions at  $\sqrt{s} = 7$  and 8 TeV with the ATLAS and CMS Experiments", (2015). arXiv:1503.07589. Submitted to *Phys. Rev. Lett.*
- [8] 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:hep-ph/0301040.
- [9] M. Schmaltz and D. Tucker-Smith, "LITTLE HIGGS THEORIES", *Annual Review of Nuclear and Particle Science* **55** (2005), no. 1, 229–270, doi:10.1146/annurev.nucl.55.090704.151502, arXiv:<http://dx.doi.org/10.1146/annurev.nucl.55.090704.151502>.
- [10] M. Perelstein, "Little Higgs models and their phenomenology", *Progress in Particle and Nuclear Physics* **58** (2007), no. 1, 247–291, doi:<http://dx.doi.org/10.1016/j.ppnp.2006.04.001>.
- [11] R. Contino, D. Pappadopulo, D. Marzocca, and R. Rattazzi, "On the effect of resonances in composite Higgs phenomenology", *Journal of High Energy Physics* **2011** (2011), no. 10, 1–50, doi:10.1007/JHEP10(2011)081.
- [12] D. Marzocca, M. Serone, and J. Shu, "General composite Higgs models", *Journal of High Energy Physics* **2012** (2012), no. 8, 1–52, doi:10.1007/JHEP08(2012)013.
- [13] B. Bellazzini, C. Csaki, and J. Serra, "Composite Higgses", *Eur. Phys. J. C* **74** (2014), no. 5, 2766, doi:10.1140/epjc/s10052-014-2766-x, arXiv:1401.2457.

- [14] D. Pappadopulo, A. Thamm, R. Torre, and A. Wulzer, “Heavy vector triplets: bridging theory and data”, *Journal of High Energy Physics* **2014** (2014), no. 9, 1–50, doi:10.1007/JHEP09(2014)060.
- [15] CMS Collaboration, “Search for a pseudoscalar boson decaying into a Z boson and the 125 GeV Higgs boson in  $\ell^+\ell^-b\bar{b}$  final states”, *Phys. Lett. B* **748** (2015) 221–243, doi:10.1016/j.physletb.2015.07.010, arXiv:1504.04710.
- [16] J. Alwall et al., “MadGraph 5 : Going Beyond”, *JHEP* **1106** (2011) 128, arXiv:1106.0522.
- [17] T. Sjstrand, S. Mrenna, and P. Skands, “PYTHIA 6.4 physics and manual”, *Journal of High Energy Physics* **2006** (2006), no. 05, 026.
- [18] S. Agostinelli et al., “Geant4a simulation toolkit”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **506** (2003), no. 3, 250 – 303, doi:10.1016/S0168-9002(03)01368-8.
- [19] S. Kallweit et al., “NLO QCD+EW predictions for V+jets including off-shell vector-boson decays and multijet merging”, arXiv:1511.08692.
- [20] CMS Collaboration, “Reference muon id, isolation and trigger efficiencies for Run-II - <https://twiki.cern.ch/twiki/bin/view/CMS/MuonReferenceEffsRun2>”,.
- [21] EXO diboson team, “Search for new VZ resonances in semileptonic final states at  $\sqrt{s} = 13$  TeV”, technical report, AN-2015/200, (2015).
- [22] C. Collaboration, “Search for New Physics in the Monojet Final State”, CMS Note 2015/072, (2015).
- [23] “Pileup Reweighting Procedure”, 2011.
- [24] “Utilities for Accessing Pileup Information for Data”, 2011.
- [25] <https://twiki.cern.ch/twiki/bin/view/CMS/CutBasedElectronIdentificationRun2> CMS Collaboration, “EGamma POG - Cut based electron ID for Run2 - <https://twiki.cern.ch/twiki/bin/view/CMS/CutBasedElectronIdentificationRun2>”,.
- [26] <https://twiki.cern.ch/twiki/bin/viewauth/CMS/HEEPIdentificationRun2> CMS Collaboration, “HEEP Electron ID and isolation - <https://twiki.cern.ch/twiki/bin/viewauth/CMS/HEEPIdentificationRun2>”,.
- [27] CMS Collaboration, “Performance of CMS muon reconstruction in pp collision events at  $\sqrt{s} = 7$  TeV”, (2012). arXiv:1206.4071. Submitted to *J. Inst.*
- [28] R. Frhwirth, “Application of Kalman filtering to track and vertex fitting”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **262** (1987), no. 23, 444 – 450, doi:[http://dx.doi.org/10.1016/0168-9002\(87\)90887-4](http://dx.doi.org/10.1016/0168-9002(87)90887-4).
- [29] CMS Collaboration, “Muon POG - Muon ID and Isolation for Run2 - <https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideMuonIdRun2>”,.
- [30] CMS Collaboration, “Tau POG - Tau ID for Run2 - <https://twiki.cern.ch/twiki/bin/viewauth/CMS/TauIDRecommendation13TeV>”,.

- [31] <https://twiki.cern.ch/twiki/bin/view/CMS/CutBasedPhotonIdentificationRun2> CMS Collaboration, “EGamma POG - Cut based photon ID for Run2 - <https://twiki.cern.ch/twiki/bin/view/CMS/CutBasedPhotonIdentificationRun2>”,.
- [32] CMS Collaboration, “Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and  $E_T^{\text{miss}}$ ”,.
- [33] CMS Collaboration, “Commissioning of the Particle-flow Event Reconstruction with the first LHC collisions recorded in the CMS detector”,.
- [34] CMS Collaboration, “Pileup Removal Algorithms”, Technical Report CMS-PAS-JME-14-001, CERN, Geneva, (2014).
- [35] R. Salam, “FastJet package”, technical report, CNRS, (2011).
- [36] M. Cacciari, G. P. Salam, and G. Soyez, “The anti- $k_t$  jet clustering algorithm”, *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.
- [37] T. C. collaboration, “Determination of jet energy calibration and transverse momentum resolution in CMS”, *Journal of Instrumentation* **6** (2011), no. 11, P11002.
- [38] CMS Collaboration, “JetMET POG - Jet ID for Run2 - <https://twiki.cern.ch/twiki/bin/view/CMS/JetID>”,.
- [39] “Jet Energy Resolution Smearing”, 2013.
- [40] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler, “Soft Drop”, *JHEP* **05** (2014) 146, doi:10.1007/JHEP05(2014)146, arXiv:1402.2657.
- [41] CMS Collaboration, “Measurement of  $B\bar{B}$  angular correlations based on secondary vertex reconstruction at  $\sqrt{s} = 7 \text{ TeV}$ ”, *JHEP* **03** (2011) 136, doi:10.1007/JHEP03(2011)136.
- [42] CMS Collaboration, “Identification of b-quark jets with the CMS experiment”, *JINST* **8** (2013) P04013, doi:10.1088/1748-0221/8/04/P04013, arXiv:1211.4462.
- [43] “Methods to apply b-tagging efficiency scale factors”, 2013.
- [44] CMS Collaboration, “Search for the Standard Model Higgs Boson Produced in Association with W and Z and Decaying to Bottom Quarks (LHC 2013)”, CMS Analysis Note 2013/069, CERN, (2013).
- [45] CMS Collaboration, “Search for new physics in the A to Zh to llbb channel”, CMS Analysis Note 2013/302, CERN, (2013).
- [46] “Modeling of  $W \rightarrow \ell\nu E_T$  with Boson Recoil”, CMS Note 2010/322.
- [47] CMS Collaboration, “Search for New Physics in the V/jet + MET final state”, CMS PAS EXO-12-055, CERN, Geneva, (2015).
- [48] C. Collaboration, “Search for Dark Matter produced in association with top quark pairs”, CMS Note 2015/120, (2015).
- [49] Julia Bauer, “Prospects for the Observation of Electroweak Top-Quark Production with the CMS Experiment”, technical report, TS-2010/007, (2010).

- 1396 [50] EXO diboson team, “Common note on searches for new diboson resonances in  
1397 semileptonic and hadronic final states at  $\sqrt{s} = 13$  TeV”, technical report, AN-2015/196,  
1398 (2015).
- 1399 [51] A. L. Read, “Presentation of search results: the CLs technique”, *J. Phys.* **G28** (2002) 2693,  
1400 doi:10.1088/0954-3899/28/10/313.
- 1401 [52] T. Junk, “Confidence level computation for combining searches with small statistics”,  
1402 *Nucl. Instrum. Meth.* **A434** (1999) 435, arXiv:9902006.
- 1403 [53] L. Moneta [http://pos.sissa.it/archive/conferences/093/057/ACAT2010\\_057.pdf](http://pos.sissa.it/archive/conferences/093/057/ACAT2010_057.pdf) et al.,  
1404 “The RooStats Project”, in *13<sup>th</sup> International Workshop on Advanced Computing and Analysis*  
1405 *Techniques in Physics Research* (ACAT2010). SISSA, 2010. arXiv:1009.1003.  
1406 PoS(ACAT2010)057.
- 1407 [54] <https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideHiggsAnalysisCombinedLimit>  
1408 “CMS Higgs Combine Tool”, 2013.
- 1409 [55] CMS Collaboration, “Search for standard model Higgs boson in pp collisions at  
1410  $\sqrt{s} = 7$  TeV”, Physics Analysis Summary CMS-PAS-HIG-11-011, CERN, (2011).
- 1411 [56] S. Bolognesi et al., “Search for a SM Higgs or BSM Boson  $H \rightarrow ZZ^{(*)} \rightarrow q\bar{q}\ell^-\ell^+$ ”,  
1412 technical report, AN-2011/388, (2011).
- 1413 [57] L. Borrello et al., “Search for the standard model Higgs boson in the decay channel  
1414  $H \rightarrow ZZ \rightarrow 2\ell 2b$ ”, technical report, AN-2011/399, (2011).

## 1415 A Optimization of electron kinematic selection criteria

In this section, we study the best selection on the  $p_T$ 's of leading and sub-leading electrons, respectively. Particularly, given that the minimum  $p_T$  threshold of leading electron is limited by the single electron trigger, we would like to see how low we could go in the  $p_T$  of sub-leading electron. The chosen figure of merit is the “Punzi significance”, which has the advantage to be independent of the signal normalization and is defined as:

$$\mathcal{P} = \frac{\varepsilon_S}{1 + \sqrt{B}}.$$

1416 Here, the  $\varepsilon_S$  is the signal efficiency with its denominator being the number of signal events  
 1417 passing the preselection criteria in Table 22, and the  $B$  is the number of background events  
 1418 normalized to  $1 \text{ fb}^{-1}$  of integrated luminosity. When optimizing the selection on the  $p_T$  of  
 1419 sub-leading electron, the requirement of “ $p_T^{\text{sub}} > 35 \text{ GeV}$ ” is removed.

1420 Figures 137–139 show (i) the signal and background distributions of the  $p_T$  of leading electron,  
 1421 and (ii) the Punzi significance, signal and background efficiencies as a function of minimum  
 1422  $p_T$  thresholds, for  $Z'$  mass at 1, 2, and 3 TeV, respectively. Figure 140 shows, as a function of  
 1423  $Z'$  mass, (i) the best Punzi significance, its corresponding signal/background efficiencies, and  
 1424 (ii) the ratio of best significance relative to the significance with preselection only. Although an  
 1425 improvement of 1–14% of significance could be gained by applying a  $p_T$  threshold tighter than  
 1426 the preselection (at 115 GeV), one also see that the threshold strongly depends on the  $Z'$  mass  
 1427 and the best selection corresponds to about 1/4 of the  $Z'$  mass.

1428 Figures 141–143 show (i) the signal and background distributions of the  $p_T$  of sub-leading elec-  
 1429 tron, and (ii) the Punzi significance, signal and background efficiencies as a function of mini-  
 1430 mum  $p_T$  thresholds, for  $Z'$  mass at 1, 2, and 3 TeV, respectively. Figure 144 shows, as a function  
 1431 of  $Z'$  mass, (i) the best Punzi significance, its corresponding signal/background efficiencies,  
 1432 and (ii) the ratio of best significance relative to the significance with preselection only. The  
 1433 improvement is at most 5 %. Therefore, we choose to stay with the preselection  $p_T$  threshold.

1434 Note that due to the small size of simulated samples, very few background events satisfy the  
 1435 preselection, particularly the  $M_{Zh}$  requirement. For example, only one DY+jets background  
 1436 event survives the pre-selection for the optimization at  $M_{Z'} = 4.5 \text{ TeV}$ . The results will be up-  
 1437 dated once a larger size of simulated samples are produced.

## 1438 B Optimization of selection criteria based on event topology

1439 While the  $p_T$ 's of leading and sub-leading electrons strongly depend on the mass of  $Z'$  (Ap-  
 1440 pendix A), several kinematic variables have little dependence on the  $Z'$  mass and are sensitive  
 1441 to the spin of the particles and the squared matrix element. These kinematic variables are:

- 1442 •  $\cos \theta^*$ , the angle between the momentum of one daughter (either higgs or  $Z$  here)  
 1443 as measured in the  $Z'$  rest frame and the flight direction of the  $Z'$  in the lab frame  
 1444 ( $z$ -axis), a variable sensitive to spin of  $Z'$  and the squared matrix element
- $\Delta y_{ZH}$ , the rapidity difference between the higgs and the leptonic- $Z$ , this variable can  
 be approximated by the pseudo-rapidity difference  $\Delta\eta_{ZH}$ , which is related to  $\cos \theta^*$ :

$$\cos \theta^* = \tanh \left( \frac{\Delta\eta_{ZH}}{2} \right).$$

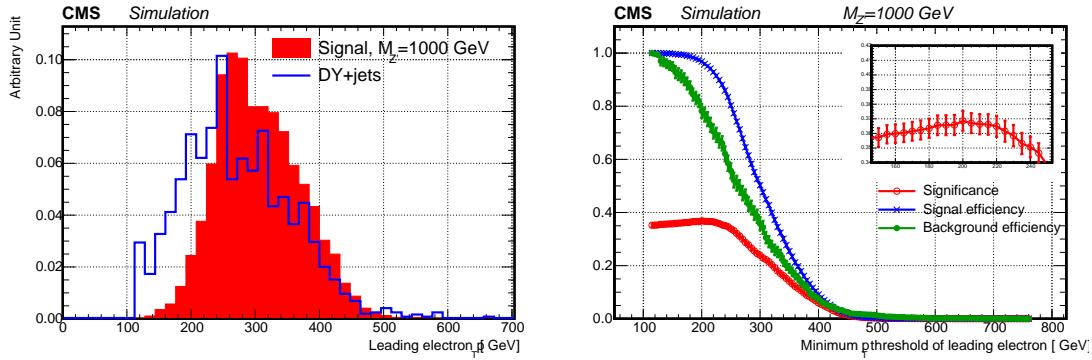


Figure 137: The signal and background distributions of leading electron  $p_T$  (left) and the Punzi significance, the signal, and the background efficiencies as a function of minimum  $p_T$  threshold (right). The  $Z'$  mass is set to 1 TeV. The signal and background distributions on the left are normalized to the same integrated area. The uncertainties on the right are correlated between different bins.

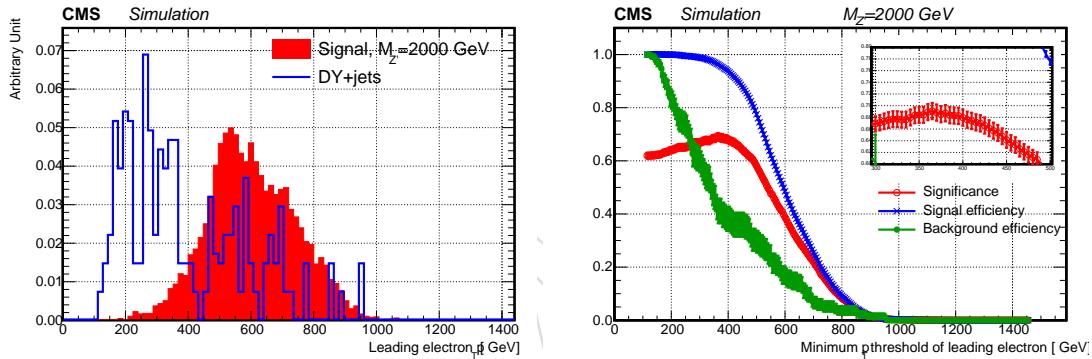


Figure 138: The signal and background distributions of leading electron  $p_T$  (left) and the Punzi significance, the signal, and the background efficiencies as a function of minimum  $p_T$  threshold (right). The  $Z'$  mass is set to 2 TeV. The signal and background distributions on the left are normalized to the same integrated area. The uncertainties on the right are correlated between different bins.

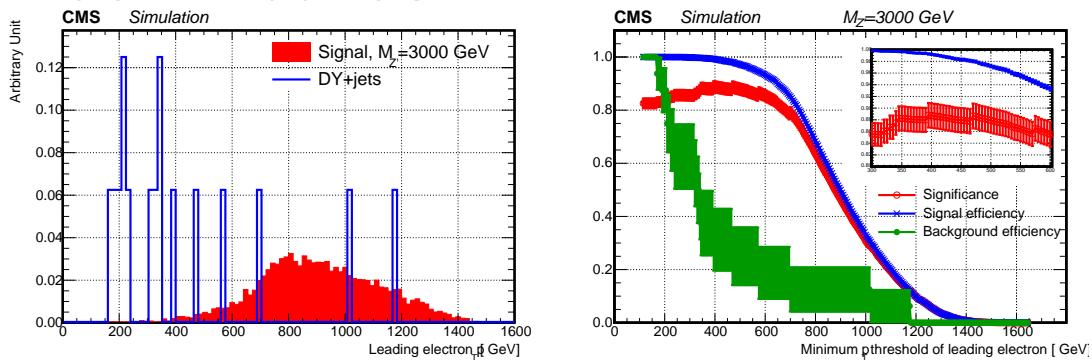


Figure 139: The signal and background distributions of leading electron  $p_T$  (left) and the Punzi significance, the signal, and the background efficiencies as a function of minimum  $p_T$  threshold (right). The  $Z'$  mass is set to 3 TeV. The signal and background distributions on the left are normalized to the same integrated area. The uncertainties on the right are correlated between different bins.

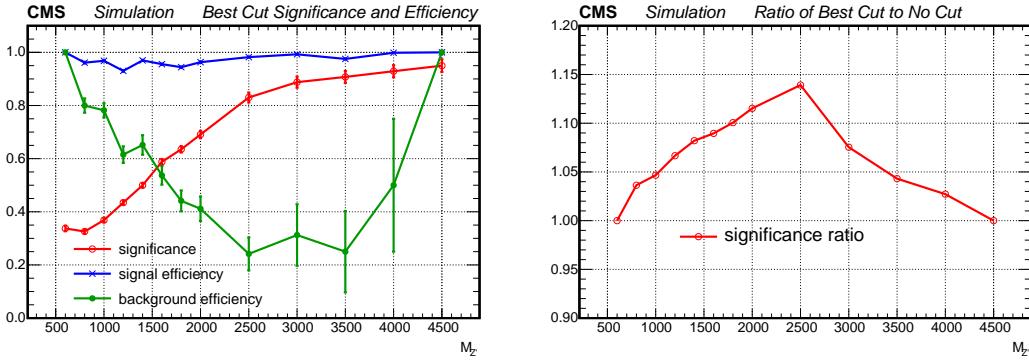


Figure 140: The best Punzi significance for the selection on leading electron  $p_T$  and the corresponding signal/background efficiencies (left), and the ratio of the best significance relative to the significance with only preselection (right), as a function of  $Z'$  mass. Uncertainties on the efficiencies are binomial.

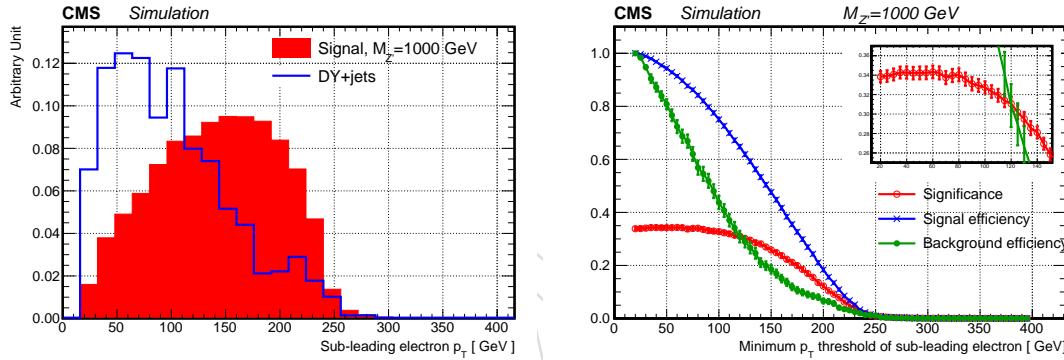


Figure 141: The signal and background distributions of sub-leading electron  $p_T$  (left) and the Punzi significance, the signal, and the background efficiencies as a function of minimum  $p_T$  threshold (right). The  $Z'$  mass is set to 1 TeV. The signal and background distributions on the left are normalized to the same integrated area. The uncertainties on the right are correlated between different bins.

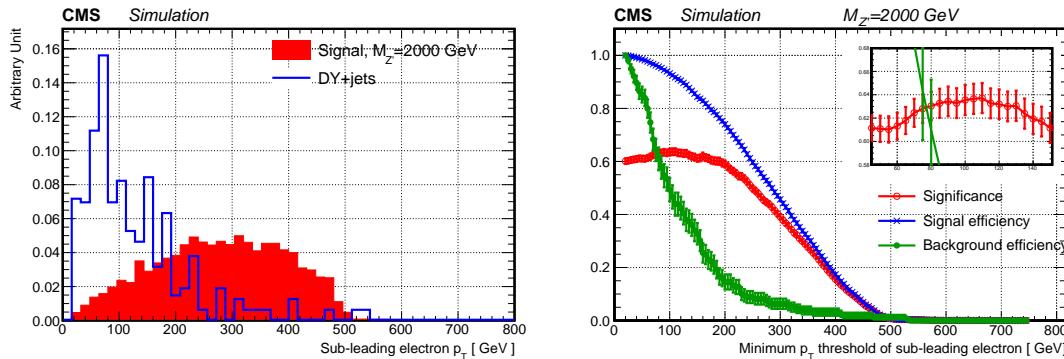


Figure 142: The signal and background distributions of sub-leading electron  $p_T$  (left) and the Punzi significance, the signal, and the background efficiencies as a function of minimum  $p_T$  threshold (right). The  $Z'$  mass is set to 2 TeV. The signal and background distributions on the left are normalized to the same integrated area. The uncertainties on the right are correlated between different bins.

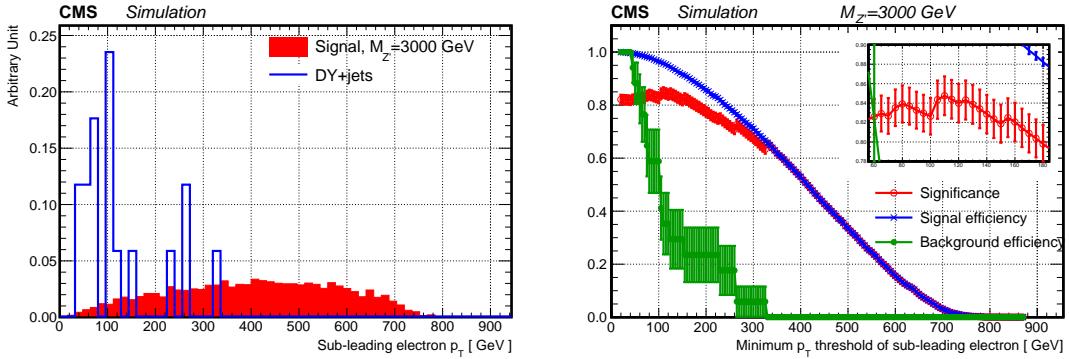


Figure 143: The signal and background distributions of sub-leading electron  $p_T$  (left) and the Punzi significance, the signal, and the background efficiencies as a function of minimum  $p_T$  threshold (right). The  $Z'$  mass is set to 3 TeV. The signal and background distributions on the left are normalized to the same integrated area. The uncertainties on the right are correlated between different bins.

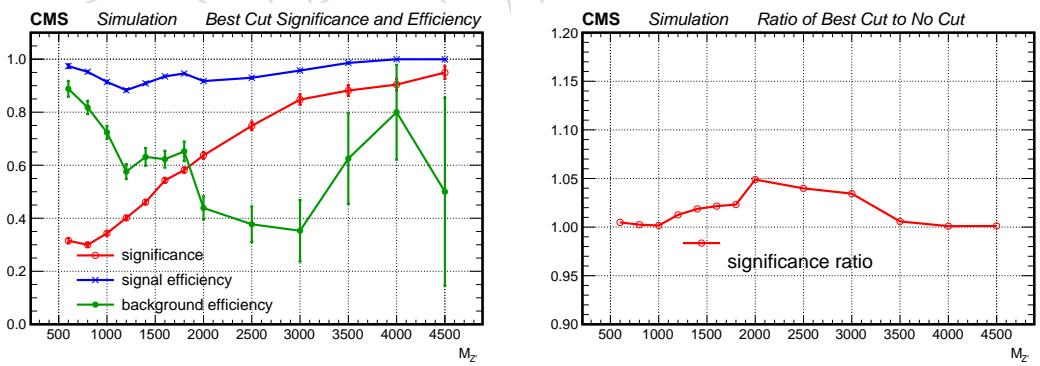


Figure 144: The best Punzi significance for the selection on sub-leading electron  $p_T$  and the corresponding signal/background efficiencies (left), and the ratio of the best significance relative to the significance with only preselection (right), as a function of  $Z'$  mass. Uncertainties on the efficiencies are binomial.

Table 22: Pre-selection for the  $X \rightarrow Zh$  analysis.

Physics Quantity	Cut Value
HLT Path	HLT_ELE105_CALOVIDVT_GSFTRKIDT_*
Good vertex	$\geq 1$
	Electron
supercluster $\eta$	$ \eta_{SC}  < 1.4442$ or $1.566 <  \eta_{SC}  < 2.5$
ID	HEEP NOISO v6.0
miniIso	$< 0.1$
	Di-electron pair
$p_T^{\text{lead}}$	$> 115 \text{ GeV}$
$p_T^{\text{sub}}$	$> 35 \text{ GeV}$
Oppositely charged	Yes
$M_{\ell\ell}$	70–110 GeV
$p_T^{\ell\ell}$	$> 200 \text{ GeV}$
	AK8 jet
$\Delta R(e, j)$	$> 0.8$
$p_T$	$> 200 \text{ GeV}$
$ \eta $	$< 2.4$
Uncorrected pruned mass	95–130 GeV
$M_{Zh}$	$\pm 15\%$ of mean

- 1445 •  $\cos \theta_{1,2}$ , the angle between the momentum of one daughter of higgs (leptonic-Z) as  
 1446 measured in the higgs (leptonic-Z) rest frame and the flight direction of the higgs  
 1447 (leptonic-Z) boson in the lab frame, a variable sensitive to the spin of higgs (leptonic-  
 1448 Z) and the squared matrix element

1449 Figures 145–146 show the distributions of these kinematic variables from  $Z'$  mass of 800 GeV to  
 1450  $Z'$  mass of 4500 GeV. Little difference is seen between different mass values. Figure 147 shows  
 1451 the angular distributions of the major background DY+jets events from Refs. [56, 57]. The  
 1452 shapes of the signal and the background are similar for the  $\cos \theta_{1,2}$ . Nevertheless, the  $\cos \theta^*$  of  
 1453 the  $Z'$  signal has a parabola shape (frowny face) with most of the events around  $\cos \theta^* \sim 0$ ,  
 1454 while the  $\cos \theta^*$  of DY+jets has a shape of smiley face with peaks at around  $|\cos \theta^*| \sim 1$ . Given  
 1455 that one could see a distinctive shape difference between signal and background in the  $\cos \theta^*$   
 1456 distribution, we choose to optimize the variable  $\Delta\eta_{ZH}$ , which is a variable closely related to  
 1457  $\cos \theta^*$  and is easier to compute<sup>2</sup>.

1458 We first find out the best selection criteria by optimizing on the Punzi significance as detailed  
 1459 in Appendix A. Figures 148–150 show (i) the signal and background distributions of  $\Delta\eta_{ZH}$ , and  
 1460 (ii) the Punzi significance, signal and background efficiencies as a function of maximum  $\Delta\eta_{ZH}$   
 1461 thresholds, for  $Z'$  mass at 1, 2, and 3 TeV, respectively. Figure 151 shows, as a function of  $Z'$   
 1462 mass, (i) the best Punzi significance and its corresponding signal/background efficiencies, and  
 1463 (ii) the ratio of best significance relative to the significance with preselection only. One could  
 1464 see that an additional 1–13% improvement of significance (maximum at  $M_{Z'} = 2 \text{ TeV}$ ) could be  
 1465 gained by applying the  $\Delta\eta_{ZH}$  selection. Note that due to the small size of simulated samples,  
 1466 very few background events satisfy the preselection, particularly the  $M_{Zh}$  requirement. For  
 1467 example, only one DY+jets background event survives the pre-selection for the optimization at

<sup>2</sup>The pseudorapidity difference is smaller for the  $s$ -channel processes such as the  $Z'$  signal model than for the  $t$ -channel processes dominating DY+jets background.

<sup>1468</sup>  $M_{Z'} = 4.5 \text{ TeV}$ . The results will be updated once a larger size of simulated samples are produced.

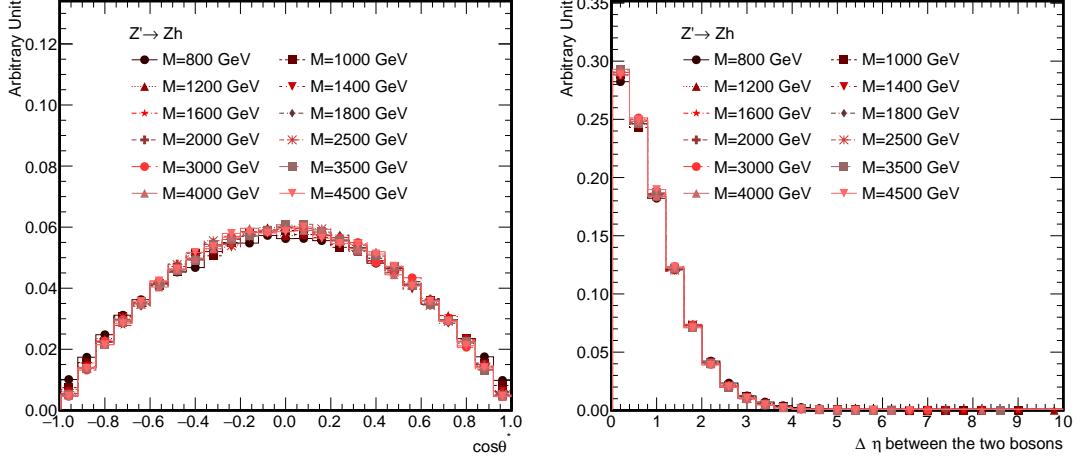


Figure 145: Comparison of the LHE-level  $\cos\theta^*$  (left) and  $\Delta\eta_{ZH}$  (right) for 12  $Z'$  mass values, from 800 GeV to 4500 GeV.

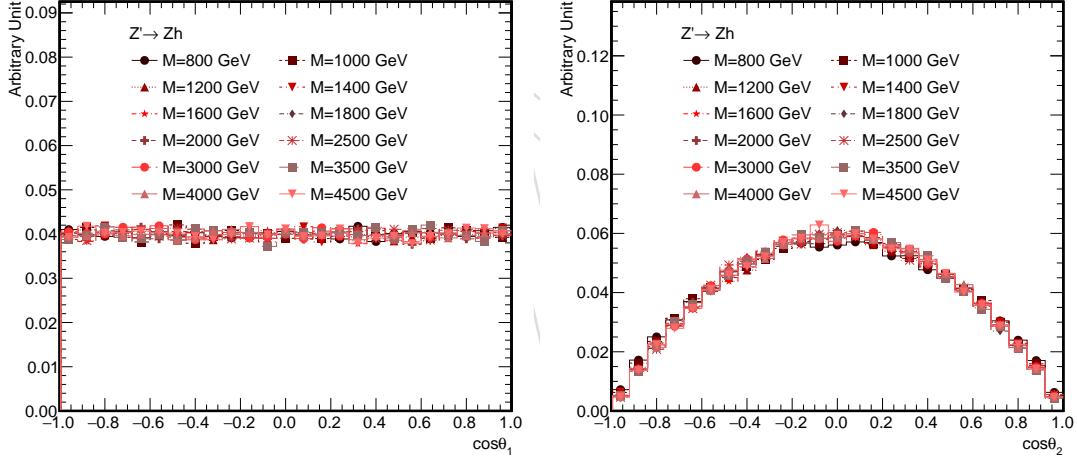


Figure 146: Comparison of the LHE-level  $\cos\theta_1$  (left) and  $\cos\theta_2$  (right) for 12  $Z'$  mass values, from 800 GeV to 4500 GeV. The  $\cos\theta_1$  ( $\cos\theta_2$ ) is computed with the four-momentum of the higgs (leptonic-Z) daughters.

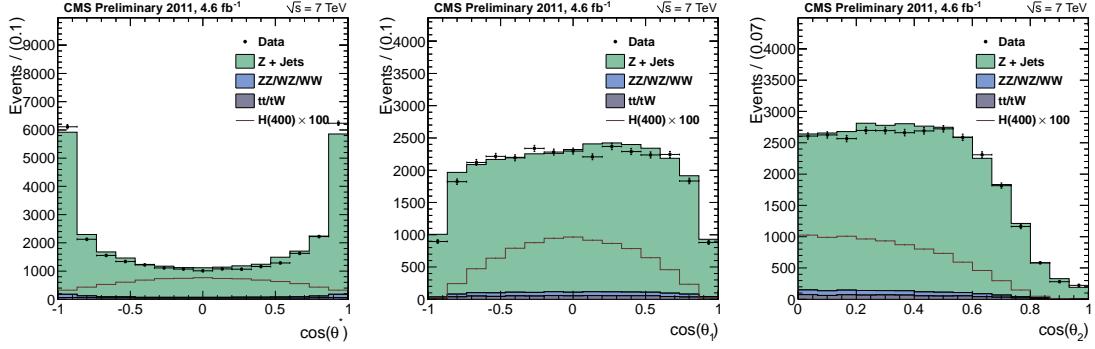


Figure 147: The  $\cos \theta^*$  (left),  $\cos \theta_1$  (middle) and  $\cos \theta_2$  (right) from DY+jets events in Refs. [56, 57].

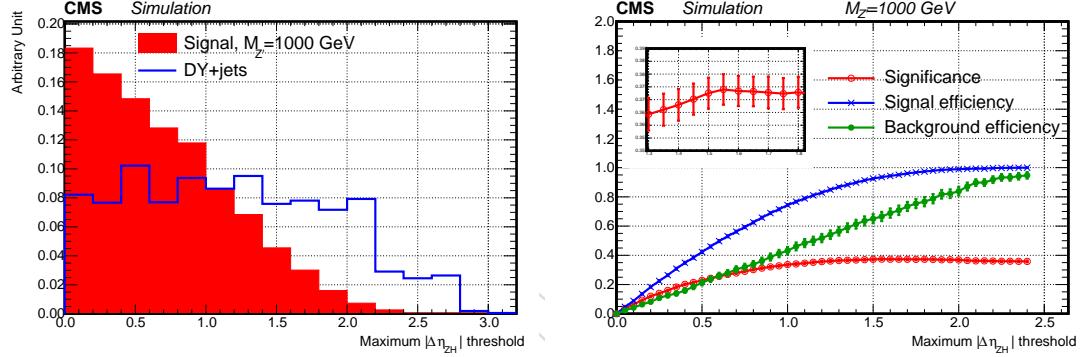


Figure 148: The signal and background distributions of  $\Delta\eta_{ZH}$  (left) and the Punzi significance, the signal, and the background efficiencies as a function of  $\Delta\eta_{ZH}$  selection:  $\Delta\eta_{ZH} < x$  (right). The  $Z'$  mass is set to 1 TeV. The signal and background distributions on the left are normalized to the same integrated area. The uncertainties on the right are correlated between different bins.

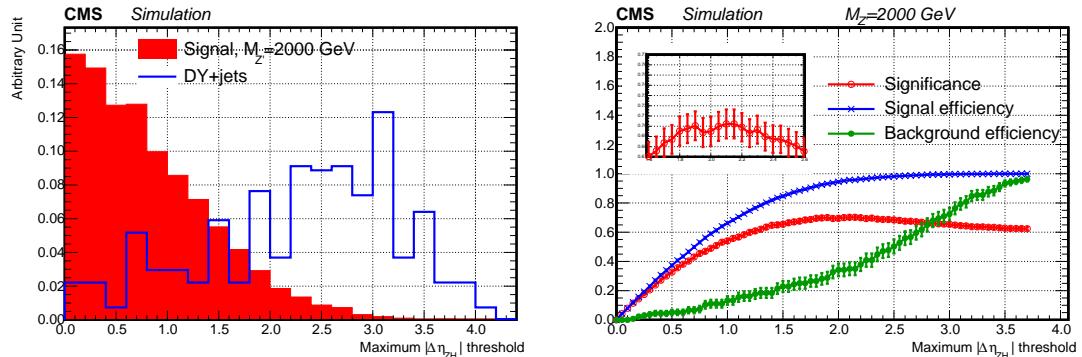


Figure 149: The signal and background distributions of  $\Delta\eta_{ZH}$  (left) and the Punzi significance, the signal, and the background efficiencies as a function of  $\Delta\eta_{ZH}$  selection:  $\Delta\eta_{ZH} < x$  (right). The  $Z'$  mass is set to 2 TeV. The signal and background distributions on the left are normalized to the same integrated area. The uncertainties on the right are correlated between different bins.

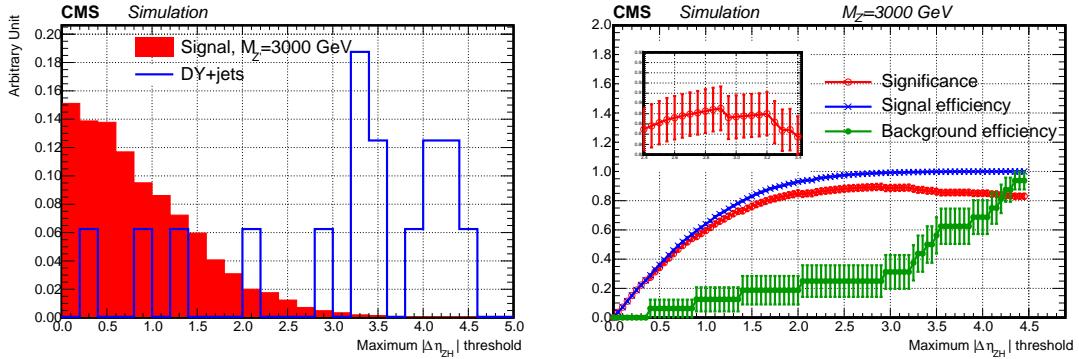


Figure 150: The signal and background distributions of  $\Delta\eta_{ZH}$  (left) and the Punzi significance, the signal, and the background efficiencies as a function of  $\Delta\eta_{ZH}$  selection:  $\Delta\eta_{ZH} < x$  (right). The  $Z'$  mass is set to 3 TeV. The signal and background distributions on the left are normalized to the same integrated area. The uncertainties on the right are correlated between different bins.

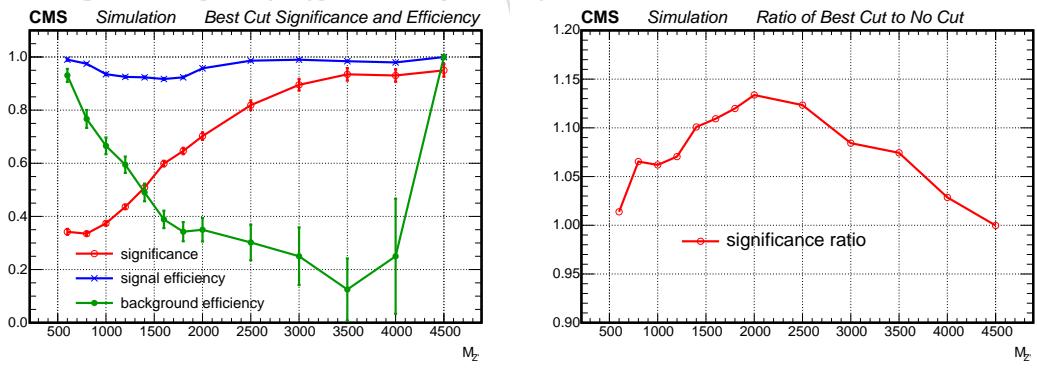


Figure 151: The best Punzi significance for the  $\Delta\eta_{ZH}$  selection and the corresponding signal/background efficiencies (left), and the ratio of the best significance relative to the significance with only preselection (right), as a function of  $Z'$  mass. Uncertainties on the efficiencies are binomial.