

CMS Draft Analysis Note

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

2017/07/21

Head Id: 416253

Archive Id: 417280

Archive Date: 2017/07/14

Archive Tag: trunk

Search for the Standard Model Higgs Boson Produced in Association with W and Z and Decaying to Bottom Quarks

Michele de Gruttola¹, Caterina Vernieri², Pierluigi Bortignon³, David Curry³, Ivan Furic³, Jacobo Konigsberg³, Sean-Jiun Wang³, Paolo Azzurri⁴, Tommaso Boccali⁴, Andrea Rizzi⁴, Stephane Brunet Cooperstein⁵, James Olsen⁵, Christopher Palmer⁵, Pirmin Berger⁶, Lorenzo Bianchini⁶, Christoph Grab⁶, Gael Ludovic Perrin⁶, Luca Perrozzi⁶, and Silvio Donato⁷

¹ CERN, European Organization for Nuclear Research, Geneva, Switzerland

² Fermi National Accelerator Laboratory, Batavia, Illinois, USA

³ University of Florida, Gainesville, Florida, USA

⁴ INFN, Scuola Normale and Università di Pisa, Italy

⁵ Princeton University, Princeton, New Jersey, USA

⁶ Institute for Particle Physics, ETH Zurich, Zurich, Switzerland

⁷ University of Zurich, Zurich, Switzerland

Abstract

A search for the standard model Higgs boson is presented in the WH- and ZH- associated production channel with decay to bottom quarks. A data sample comprising up to 35.9/fb from the full 2016 data taking with $\sqrt{s} = 13$ TeV has been analyzed in five channels $Z(\mu\mu)H$, $Z(ee)H$, $Z(\nu\nu)H$, $W(\mu\nu)H$, $W(e\nu)H$, and 95% C.L. upper limits derived, for a Higgs boson mass of 125 GeV, yielding and observed upper limit of 1.87 times the SM prediction, where 0.70 is expected in the absence of a signal. The corresponding fitted signal strength is found to be $\mu = \sigma/\sigma_{SM} = 1.19^{+0.35}_{-0.34}$, while the observed (expected) local significance of the excess over background prediction is quantified as 3.29σ (2.82σ). Please note that the AN-15-168_v4 describes the same analysis using 2.2/fb 2015 data at $\sqrt{s} = 13$ TeV. Please also note that the AN-15-168_v8 describes the same analysis using 12.9/fb 2016 data at $\sqrt{s} = 13$ TeV.

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

PDFAuthor: VHbb team
PDFTitle: Search for the Standard Model Higgs Boson Produced in Association with W and Z and Decaying to Bottom Quarks
PDFSubject: CMS
PDFKeywords: CMS, physics, software, computing

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

DRAFT

1 Contents

1	1	Introduction	3
2	1.1	Analysis Strategy	4
3	2	Data and Simulated Samples	6
4	2.1	Data	6
5	2.2	Monte Carlo samples	6
6	3	Triggers	12
7	3.1	Dedicated trigger for $Z(\nu\nu)H$	13
8	4	Physics Objects	18
9	4.1	Primary vertex selection and pile-up treatment	18
10	4.2	Electrons	19
11	4.3	Muons	20
12	4.4	Lepton isolation	21
13	4.5	Jets	21
14	4.6	Identification of b jets	21
15	4.7	Efficiency for b-jet identification	22
16	4.8	Missing transverse energy	22
17	4.9	Additional “soft” hadronic activity	23
18	5	Vector Boson Reconstruction	24
19	5.1	Vector boson transverse momentum reweighting in V+jets events	24
20	5.2	EWK signal VH corrections	24
21	6	Higgs Boson Reconstruction	27
22	6.1	Regression	27
23	7	Residual DATA/MC corrections	42
24	7.1	Correcting Data/MC discrepancy in LO Drell-Yan m_{jj} by re-weighting to NLO	42
25	7.2	W boson transverse momentum reweighting and LO/NLO check	42
26	8	Event Selection	49
27	8.1	Signal and background characteristics	49
28	8.2	Discriminating variables	50
29	8.3	Reconstructing the top mass in events with a lepton and MET	52
30	8.4	Analysis preselection	52
31	9	Binned BDT Shape Analysis	54
32	10	Signal Efficiency	64
33	10.1	Lepton reconstruction and ID efficiencies	64
34	11	Background Control Samples	75
35	11.1	$Z(\ell\ell)H$ control regions	75
36	11.2	$Z(\nu\nu)H$ control regions	75
37	11.3	$W(\ell\nu)H$ control regions	82
38	11.4	Scale factor determination	106
39	12	Systematics	108
40	12.1	Jet p_T, η decorrelation in JEC, JER and B-tagging systematics	111
41	12.2	Implementation of PDF Uncertainty Systematics	111

44	13	Di-boson Validation Analysis	113
45	14	Results	121
46	14.1	Signal and Control Regions fits	121
47	14.2	Upper limit calculation	122
48	14.3	Results VH	122
49	15	Conclusions	130

DRAFT

1 Introduction

- 50 In the Standard Model (SM) [1] [2] [3] the Brout-Englert-Higgs mechanism [4] [5] [6] explains
 51 the electroweak symmetry breaking and allows electroweak gauge bosons to acquire mass.
 52 The mechanism predicts the existence of a Higgs scalar field, and its observation in 2012 with
 53 the LHC Run1 proton-proton collision data by both the CMS [7] and ATLAS [8] collaborations
 54 achieved one of the main goals of the LHC physics program.
 55
- 56 The observed Higgs boson mass is now precisely determined to be $m_H = 125.09 \pm 0.24$ GeV
 57 [9], i.e. with a 0.2% precision level. On the other hand, the observed properties and couplings
 58 are consistent with those of a minimal SM Higgs boson but only with a precision at the level of
 59 10% or worse [10]. In particular, while a SM Higgs boson with a mass of 125.1 GeV is predicted
 60 to have a dominant decay branching ratios to b-quarks (58.1% [11]), the LHC Run1 data did not
 61 yet clearly establish the coupling of the Higgs boson to b-quarks, nor in general to down-type
 62 quarks.

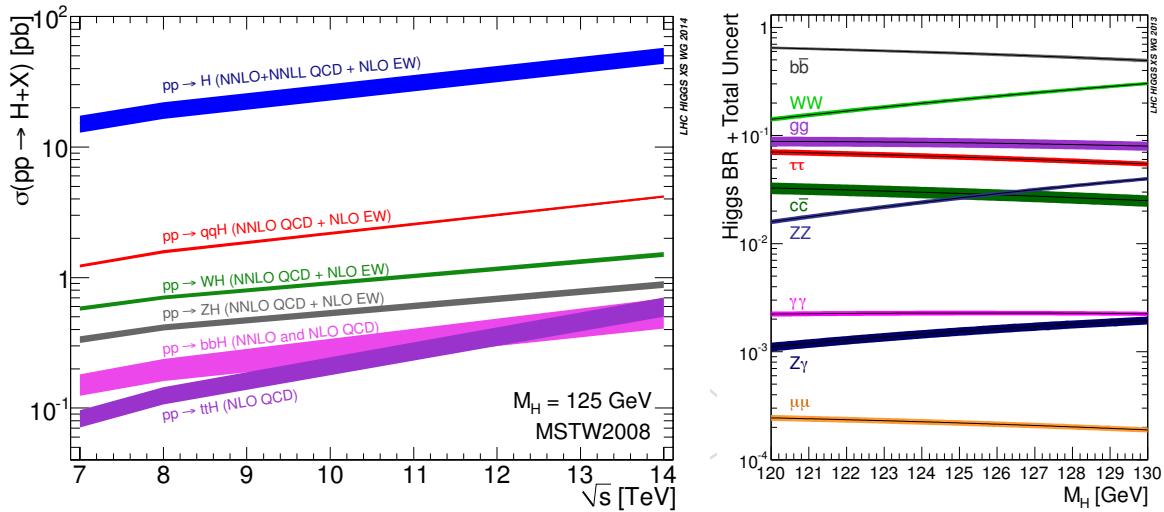


Figure 1: Minimal Standard Model Higgs production and decay at the LHC [12]. (left) Production cross-sections at $\sqrt{s} = 7$ –14 TeV, for $m_H = 125$ GeV. (right) Decay Branching Fractions for $m_H = 120$ –130 GeV.

- 63 Figure 1 shows the expected production cross sections as a function of the LHC proton-proton
 64 collision energy, and the expected decay modes Branching Fractions for minimal SM Higgs
 65 boson with $m_H \simeq 125$ GeV. The most abundant LHC channel for a SM Higgs boson after
 66 the gluon-fusion (GF) production is the Vector Boson Fusion (VBF) production followed by the
 67 vector boson associated (VH) production.
 68 The mass at which the Higgs boson has been observed is particularly challenging at hadron
 69 colliders. Figure 1 shows the Higgs boson branching fractions to accessible final states near
 70 its the measured mass. The dominant decay mode is $H \rightarrow b\bar{b}$, but if this is searched in the
 71 dominant GF production mode the resulting irreducible background from QCD production
 72 of b quarks is overwhelming, so that the the inclusive observation of the SM Higgs boson
 73 decaying to $b\bar{b}$ pairs is not imaginable in proton collisions.
 74 To search for the standard model decay $H \rightarrow b\bar{b}$ at the LHC the best method is to search for
 75 events where the Higgs boson is produced in association with a W or Z boson, and recoiling
 76 with large momentum transverse to the beam line [13]. The presence of a vector boson in the fi-
 77 nal state highly suppresses the QCD background, while also providing an efficient trigger path

when the W or Z decays to charged leptons. Requiring a large boost provides several additional advantages: it further reduces the large backgrounds from W and Z production in association with jets; it helps in reducing the large background from top-quark production in the signal channels including neutrinos; it makes accessible the $Z(\nu\nu)H$ channel via large missing transverse energy (E_T^{miss}); and it generally improves the mass resolution of the reconstructed Higgs candidates.

The first search results from CMS for five channels ($Z(\ell\ell)H$, $Z(\nu\nu)H$, $W(\ell\nu)H$, with $\ell = e, \mu$) using 1.1 fb^{-1} of data and a cut analysis technique were presented in 2011 [14, 15]. Subsequently, CMS published updated search results on the full 5 fb^{-1} 2011 dataset [16, 17] using similar techniques achieving sensitivity of roughly four times the SM expectation at $M_H = 125 \text{ GeV}$. These results were published in Ref. [18]. Further CMS updates were presented at the ICHEP conference in summer 2012 [19, 20] based on the full 7 TeV data, and 5.1 fb^{-1} of 2012 8 TeV data, and for the HCP 2012 conference [21, 22] based on the full 2011 7 TeV data and 12.1 fb^{-1} of 2012 8 TeV data. Results comprising the full Run1 data ($5 + 19 \text{ fb}^{-1}$ at 7 and 8 TeV) were produced for the LHCP 2013 conference [23, 24], and published [25]. The Run1 results yielded an excess of events observed above the expected background, with a local significance of 2.1 standard deviations for a Higgs boson mass of 125 GeV, consistent with the expectation from a SM Higgs boson production. The corresponding signal strength relative to the SM predicted one was fitted to $\mu = 1.0 \pm 0.5$.

The VH, $H \rightarrow b\bar{b}$ Run1 results have been corrected for the expected contribution of ZH productions initiated by gluon fusion, yielding a lower $\mu = 0.89 \pm 0.43$ fitted signal strength, but preserving the signal significance [26]. These results have been further combined with CMS $H \rightarrow b\bar{b}$ searches in the VBF production mode [27], and in the top quark pair $t\bar{t}H$ associated production mode [28, 29]. For $m_H = 125 \text{ GeV}$ the final CMS Run1 combination yielded a $H \rightarrow b\bar{b}$ signal strength $\mu = 1.03^{+0.44}_{-0.42}$ with a significance of 2.6 standard deviations.

This analysis note describes the search for the standard model Higgs boson with $H \rightarrow b\bar{b}$ decays, produced in association with a W or Z boson in LHC Run2, using $35.9/\text{fb}$ data collected in 2016 at $\sqrt{s} = 13 \text{ TeV}$. (Please note that the AN-15-168-v4 describes the same analysis using $2.2/\text{fb}$ 2015 data at $\sqrt{s} = 13 \text{ TeV}$, which are the very first low luminosity 13 TeV data collected by CMS.)

Analyses for the three standard topologies ($Z(\ell\ell)H$, $Z(\nu\nu)H$, $W(\ell\nu)H$, with $\ell = e, \mu$) are presented.

1.1 Analysis Strategy

The strategy follows closely the Run1 analysis described in detail in AN-12-181 [20], AN-12-349 [21] and AN-13-069 [23].

The event selection begins by identifying W and Z bosons decaying to charged leptons using standard methods, and Z bosons decaying to neutrinos by requiring large E_T^{miss} . The $H \rightarrow b\bar{b}$ decay is reconstructed as two separate particle-flow jets clustered using the anti- k_T algorithm with a size parameter of 0.4, and the Combined MVA v2 (CMVAv2) tagger is used to identify candidates that are consistent with the b-jet hypothesis. In the modes with larger backgrounds, $W(\ell\nu)H$ and $Z(\nu\nu)H$, the V and H candidates are required to have a larger boost ($\sim 100 \text{ GeV}$) transverse to the beam. In the $Z(\ell\ell)H$ channel, where the background from $t\bar{t}$ production is much less (and the effective signal cross section is also lower) lower boost regions are also considered. Further background rejection is achieved by exploiting signal properties as the resonant dijet mass, the back-to-back VH topology, and reduced additional hadronic activity.

123 Backgrounds arise from production of W and Z bosons in association with one or more jets,
124 singly and pair-produced top quarks, and dibosons. Smaller residual QCD background is
125 present in the $W(\ell\nu)H$ and $Z(\nu\nu)H$ channels. High-purity control regions for the $V + \text{udscg}$
126 and $t\bar{t}$ backgrounds are identified in data and used to estimate expected yields in the signal re-
127 gion. Samples orthogonal to the signal region in dijet mass and enhanced in $V + b\bar{b}$ production
128 are used to constrain the irreducible $b\bar{b}$ background from data. In the missing energy channel,
129 a $Z \rightarrow \nu\bar{\nu}$ control sample is isolated in data by removing the muons from a high-purity sample
130 of $Z(\mu\mu) + \text{jets}$ in data.

DRAFT

2 Data and Simulated Samples

2.1 Data

The 2016 13 TeV proton proton collision data taken with 25 ns minimum bunch spacing in the 06Feb2016 reminiAOD version have been used in this analysis. For $Z(\mu\mu)H$ channel we use DoubleMuon dataset, for $Z(ee)H$ we use DoubleEG, for $W(\mu\nu)H$ we use SingleMuon, for $W(e\nu)H$ we use SingleElectron while for $Z(\nu\nu)H$ channel we use MET dataset.

In Table 1 we summarizes taking as example the SingleMuon dataset the data samples used in the current analysis and the approximate luminosity as reported from the golden JSON

`Cert_271036 – 284044_13TeV_23Sep2016ReReco_Collisions16_JSON.txt`

file used in each analysis, with `brilcalc` and the `normtag`.

Table 1: List of 2016 data samples used for the SingleMuon dataset. An uncertainty of 2.5% is assigned for the 2016 luminosity, see hypernews

<https://hypernews.cern.ch/HyperNews/CMS/get/physics-announcements/4613.html>

Dataset	$\int \mathcal{L} (\text{fb}^{-1})$
SingleMuon_Run2016B-03Feb2017-v1	~ 5.9
SingleMuon_Run2016B-03Feb2017-v2	~ 2.7
SingleMuon_Run2016C-03Feb2017-v1	~ 4.3
SingleMuon_Run2016E-03Feb2017-v1	~ 4.1
SingleMuon_Run2016F-03Feb2017-v1	~ 3.2
SingleMuon_Run2016G-03Feb2017-v1	~ 3.8
SingleMuon_Run2016H-03Feb2017-v1	~ 11.8
Total Lumi	35.9

2.2 Monte Carlo samples

Monte Carlo samples in CMSSW 80X are taken from the RunIISummer16 productions re-miniAODv2 with the *Asympt25ns* conditions. The list of samples with the number of events, cross sections and integrated luminosities is given below for Higgs boson signal (Table 2), di-vector boson production (Table 4), vector boson plus jets (Table 5- 6), and $t\bar{t}$ with QCD multi-jet (Table 7).

Samples were produced using one or more of the following programs: PYTHIA 8 [30, 31], POWHEG [32], TAUOLA [33], and MADGRAPH 5_aMC@NLO [34, 35] with MLM merging [36] or FxFx merging scheme [37]. Parton shower and hadronisation are performed with PYTHIA 8 [31] using the CUETP8M1 tune [38]. The NNPDF3.0 parton distribution functions (PDF) [39] are used for all samples.

The production cross sections for $W+\text{jets}$ and $Z+\text{jets}$ are rescaled to next-to-next-to-leading-order (NNLO) cross sections calculated using the FEWZ 3.1 program [40–42]. The $t\bar{t}$ and single top quark samples are also rescaled to their cross sections based on NNLO calculations [43, 44].

152 **2.2.1 Signal simulation**

153 The list of signal Monte Carlo samples is given in Table 2, including $H \rightarrow b\bar{b}$ (HToBB) decays
 154 for Higgs bosons produced in vector boson associated processes initiated either by quark fusion
 155 or gluon fusion processes.

Table 2: Signal Monte Carlo samples with $M_H = 125$ GeV

Sample	Generator	m_H (GeV/ c^2)	σ (pb)	events	$\int \mathcal{L}$ (fb $^{-1}$)
/WplusH_HToBB_WToLNu_M125_13TeV_powheg_pythia8	POWHEG +PYTHIA 8	125	0.840 * 0.108535 * 0.5824	1 317 467	8039.05
/WminusH_HToBB_WToLNu_M125_13TeV_powheg_pythia8	POWHEG +PYTHIA 8	125	0.533 * 0.108535 * 0.5824	1 290 538	12410.58
/ZH_HToBB_ZToLL_M125_13TeV_powheg_pythia8	POWHEG +PYTHIA 8	125	(0.8839 - 0.1227) * 0.10974 * 0.5824	4 926 620	102213.80
/ZH_HToBB_ZToNuNu_M125_13TeV_powheg_pythia8/	POWHEG +PYTHIA 8	125	(0.8839 - 0.1227) * 0.20103 * 0.5824	1 205 831	13661.96
/ggZH_HToBB_ZToLL_M125_13TeV_powheg_pythia8	POWHEG +PYTHIA 8	125	0.1227 * 0.10974 * 0.5824	2 998 600	192975.96
/ggZH_HToBB_ZToNuNu_M125_13TeV_powheg_pythia8	POWHEG +PYTHIA 8	125	0.1227 * 0.20103 * 0.5824	2 396 838	168136.79

156 Table 3 summarizes the cross sections and branching fractions assumed for each signal channel
 157 and with $M_H = 125$ GeV. The cross sections are computed at NNLO, as described in Ref. [11,
 158 12, 45].

159 **2.2.2 Background simulations**

160 Tables 4, 5, 6, 7 summarize the simulated background samples and their equivalent luminosities
 161 and cross sections where applicable. Appropriate pile-up reweighting is applied when
 162 comparing to data, in order to represent the true primary vertex distribution in the different
 163 run ranges. An example of the effect of reweighting is shown in Figure 2, which shows the
 164 distribution of the number of primary vertices in data and reweighted simulation for the $t\bar{t}$
 165 control region described in Sec. 11. Due to dynamical tracker inefficiencies, the agreement is
 166 not expected to be excellent. The average energy density, ρ , in an event is shown in the right of
 167 Figure 2. This variable is used as a measure of pile-up. The scaling for luminosity to PU was
 168 constructed in part using this variable (centrally in CMS). The min bias cross section used for
 169 scaling luminosity into PU is 69.2 mb (as centrally prescribed by the Luminosity POG).

Table 3: Signal cross sections and branching ratios for $M_H = 125$ GeV at $\sqrt{s} = 13$ TeV [11].

process	σ (pb)	QCDScale	PDF
ggH	44.14	+7.6-8.1%	$\pm 3.1\%$
VBF H	3.782	+0.4-0.3%	$\pm 2.1\%$
WH	1.373 (0.840+0.533)	+0.5 -0.7%	$\pm 1.9\%$
ZH	0.8839	+3.8 -3.1%	$\pm 1.6\%$
decay	BR	Uncertainty	
$H \rightarrow b\bar{b}$	58.24%	+0.72-0.74%	

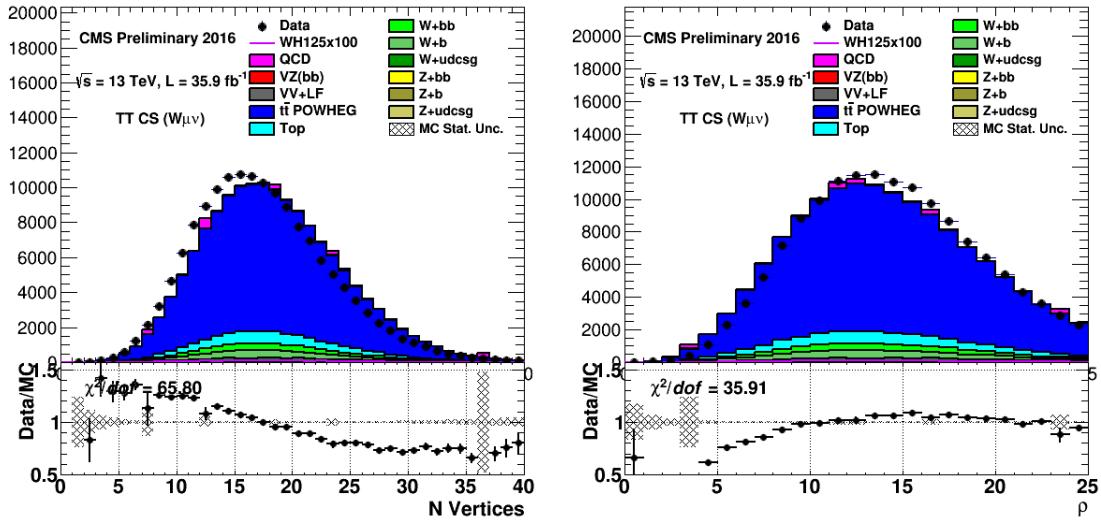


Figure 2: Distribution of the number of reconstructed primary vertices and ρ in data compared to simulation in the $t\bar{t}$ control region for $W(\mu\nu)H$. See Sec. 11 for the definition of this control region.

Table 4: List of Monte Carlo diboson samples

Sample	Generator	σ (pb)	events	$\int \mathcal{L} (fb^{-1})$
/WW_TuneCUETP8M1_13TeV-pythia8	PYTHIA 8	118.7	993 640	8.37
/WZ_TuneCUETP8M1_13TeV-pythia8	PYTHIA 8	47.13	1 000 000	21.22
/ZZ_TuneCUETP8M1_13TeV-pythia8	PYTHIA 8	16.523	985 600	59.65

170 2.2.3 Additional jet modeling in VZbb MC

171 The number of additional jets is a key variable in the $W(\ell\nu)H$ analysis—both in the defini-
 172 tion of the control regions and in the signal selection. The available $VZ(bb)$ MC samples are
 173 aMC@NLO, PYTHIA (LO, low-stat.), and POWHEG (low-stat.). In other analyses good agree-
 174 ment has been observed between data and the POWHEG prediction for the $VZ(bb)$ number
 175 of additional jets distribution. Figure 3 shows a comparison of the number of additional jets
 176 among the available $WZ(bb)$ samples. Although there is some small difference in shape pre-
 177 diction between the aMC@NLO and Powheg samples, the generators agree reasonably well
 178 within statistical uncertainties. Since the available statistical power of the aMC@NLO $VZ(bb)$
 179 samples is much better, this analysis uses aMC@NLO to predict the $VZ(bb)$ shape in the fitted
 180 regions for all sub-channels.

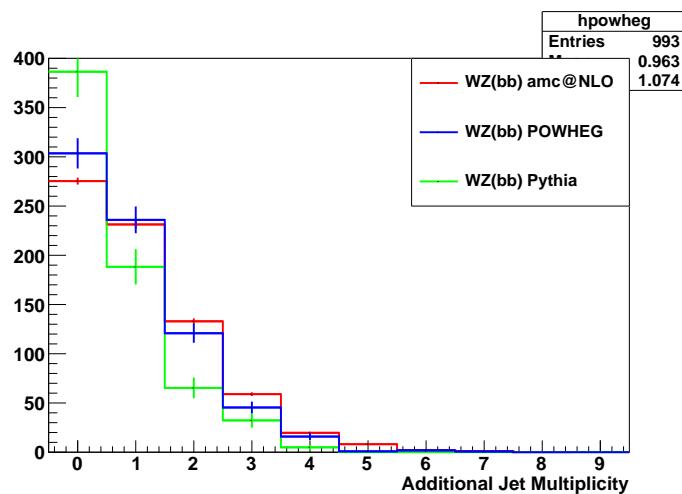


Figure 3: Number of additional jets with $p_T > 25$ GeV, $|\eta| < 2.9$ for WZ(bb) generated with aMC@NLO (red), LO PYTHIA (green), and WZ(bb) POWHEG (blue).

Table 5: List of Monte Carlo V + jets leading order samples

Sample	Generator	σ (pb)	events	$\int \mathcal{L} (\text{fb}^{-1})$
/DY1JetsToLL_M-10to50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	MADGRAPH 5+PYTHIA 8	725	39 800 000	54.5
/DY2JetsToLL_M-10to50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	MADGRAPH 5+PYTHIA 8	394.5	19 400 000	50.2
/DY3JetsToLL_M-10to50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	MADGRAPH 5+PYTHIA 8	96.47	4 960 000	52.2
/DYJetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	MADGRAPH 5+PYTHIA 8	4960*1.23	49 100 000	9.9
/DYJetsToLL_M-50_HT-100to200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	MADGRAPH 5+PYTHIA 8	147.40*1.23	10 610 000	72
/DYJetsToLL_M-50_HT-200to400_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	MADGRAPH 5+PYTHIA 8	40.99*1.23	9 652 000	235.4
/DYJetsToLL_M-50_HT-400to600_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	MADGRAPH 5+PYTHIA 8	5.678*1.23	10 010 000	1759
/DYJetsToLL_M-50_HT-600to800_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	MADGRAPH 5+PYTHIA 8	1.367*1.23	8 290 000	6111
/DYJetsToLL_M-50_HT-800to1200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	MADGRAPH 5+PYTHIA 8	0.6304*1.23	2 670 000	4280
/DYJetsToLL_M-50_HT-1200to2500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	MADGRAPH 5+PYTHIA 8	0.1514*1.23	596 000	3940
/DYJetsToLL_M-50_HT-2500toInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	MADGRAPH 5+PYTHIA 8	0.003565*1.23	399 000	109000
/DYBJetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	MADGRAPH 5+PYTHIA 8	71.77*1.23	1 470 000	20.9
/DYBJetsToLL_M-50_Zpt-100to200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	MADGRAPH 5+PYTHIA 8	3.027*1.23	4 080 000	1320
/DYBJetsToLL_M-50_Zpt-200toInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	MADGRAPH 5+PYTHIA 8	0.297*1.23	2 110 000	6670
/ZJetsToNuNu_HT-100To200_13TeV-madgraph	MADGRAPH 5+PYTHIA 8	280.35*1.23	5 240 199	15.20
/ZJetsToNuNu_HT-200To400_13TeV-madgraph	MADGRAPH 5+PYTHIA 8	42.75*1.23	5 032 927	95.7
/ZJetsToNuNu_HT-400To600_13TeV-madgraph	MADGRAPH 5+PYTHIA 8	10.73*1.23	954 435	72.32
/ZJetsToNuNu_HT-600ToInf_13TeV-madgraph	MADGRAPH 5+PYTHIA 8	4.116 *1.23	9 645 493	1905.21
/WJetsToLNu_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	MADGRAPH 5+PYTHIA 8	61526.7 *1.21	86 700 000	1.72
/WJetsToLNu_HT-100To200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	MADGRAPH 5+PYTHIA 8	1345 *1.21	79 300 000	58.9
/WJetsToLNu_HT-200To400_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	MADGRAPH 5+PYTHIA 8	359.7 *1.21	39 650 00	110.1
/WJetsToLNu_HT-400To600_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	MADGRAPH 5+PYTHIA 8	48.91*1.21	7 760 000	159.2
/WJetsToLNu_HT-600To800_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	MADGRAPH 5+PYTHIA 8	12.05*1.21	18 680 000	1543
/WJetsToLNu_HT-800To1200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	MADGRAPH 5+PYTHIA 8	5.501*1.21	6 200 000	1130
/WJetsToLNu_HT-1200To2500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	MADGRAPH 5+PYTHIA 8	1.329*1.21	6 875 000	5174
/WJetsToLNu_HT-2500ToInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	MADGRAPH 5+PYTHIA 8	0.03216*1.21	2 634 000	82200
/WBJetsToLNu_Wpt-100to200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	MADGRAPH 5+PYTHIA 8	6.004*1.21	3 979 072	662.7
/WBJetsToLNu_Wpt-200toInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	MADGRAPH 5+PYTHIA 8	0.8524*1.21	2 892 981	3393.3
/WJetsToLNu_BGenFilter_Wpt-100to200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	MADGRAPH 5+PYTHIA 8	26.1*1.21	6 690 000	256.3
/WJetsToLNu_BGenFilter_Wpt-200toInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	MADGRAPH 5+PYTHIA 8	3.545*1.21	11 650 000	3286.3

Table 6: List of Monte Carlo V + jets next-to-leading order samples

Sample	Generator	σ (pb)	events	$\int \mathcal{L} (\text{fb}^{-1})$
/DYJetsToLL_Pt-50To100_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	MADGRAPH 5+PYTHIA 8	369.3	2.19E+07	8.23E+00
/DYJetsToLL_Pt-100To250_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	MADGRAPH 5+PYTHIA 8	81.2	499 000	7.87
/DYJetsToLL_Pt-250To400_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	MADGRAPH 5+PYTHIA 8	2.99	1 609 000	73.6
/DYJetsToLL_Pt-400To650_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	MADGRAPH 5+PYTHIA 8	0.388	1 626 000	628
/DYJetsToLL_Pt-650ToInf_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	MADGRAPH 5+PYTHIA 8	0.0374	1 629 000	7360
/DYToLL_0J_13TeV-amcatnloFXFX-pythia8	MADGRAPH 5+PYTHIA 8	4760	49 600 000	6.98
/DYToLL_1J_13TeV-amcatnloFXFX-pythia8	MADGRAPH 5+PYTHIA 8			
/DYToLL_2J_13TeV-amcatnloFXFX-pythia8	MADGRAPH 5+PYTHIA 8	341	42 300 000	10.5

Table 7: Top and QCD Monte Carlo samples

Sample	Generator	σ (pb)	events	$\int \mathcal{L} (\text{fb}^{-1})$
/TT_TuneCUETP8M1_13TeV-powheg-pythia8	POWHEG +PYTHIA 8	831.76	187 626 200 + 97 994 442	343
/ST_tW_top_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1	POWHEG +PYTHIA 8	35.6	1 000 000	28.09
/ST_tW_antitop_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1	POWHEG +PYTHIA 8	35.6	999 400	28.07
/ST_t-channel_top_4f_leptonDecays_13TeV-powheg-pythia8	POWHEG +PYTHIA 8	136*0.325	999 400	22.6
/ST_t-channel_antitop_4f_leptonDecays_13TeV-powheg-pythia8	POWHEG +PYTHIA 8	81*0.325	1 695 400	64.4
/ST_s-channel_4f_leptonDecays_13TeV-amcatnlo-pythia8	POWHEG +PYTHIA 8	10.32	998 400	96.74
/QCD_HT100to200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	MADGRAPH 5+PYTHIA 8	27990000	82 095 800	0.003
/QCD_HT200to300_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	MADGRAPH 5+PYTHIA 8	1712000	18 784 379	0.011
/QCD_HT300to500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	MADGRAPH 5+PYTHIA 8	347700	54 267 650	0.16
/QCD_HT500to700_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	MADGRAPH 5+PYTHIA 8	2.94e4	19 542 847	0.66
/QCD_HT700to1000_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	MADGRAPH 5+PYTHIA 8	6831	45 100 675	6.60
/QCD_HT1000to1500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	MADGRAPH 5+PYTHIA 8	1207	15 193 645	12.59
/QCD_HT1500to2000_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	MADGRAPH 5+PYTHIA 8	119.9	3 939 077	32.85
/QCD_HT2000toInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	MADGRAPH 5+PYTHIA 8	25.42	1 961 774	77.2

181 3 Triggers

182 A mix of different triggers are used to collect events consistent with the signal hypothesis in
 183 each channel. Table 3 summarizes the triggers used in this analysis.

184 $W(\mu\nu)H$ and $W(e\nu)H$ channels utilize single lepton triggers. The $Z(\mu\mu)H$ and $Z(ee)H$ channels
 185 are based on di-lepton triggers which are more efficient for di-lepton signal. All of these triggers
 186 are emulated in the MC and that emulation is required by all MC events analyzed in these
 187 channels. Differences in trigger variables between data and MC require the trigger efficiency
 188 of MC to be corrected. Correction scale factors are derived using the tag-and-probe method,
 189 which utilizes di-lepton events from Z bosons. Because the tag lepton selection is very strict
 190 and the di-lepton invariant mass is consistent with the Z boson mass, the probe lepton is very
 191 pure with minimal selection.

192 The trigger efficiencies are measured after the application of offline lepton identification and
 193 isolation selections. For di-lepton triggers scale factors for each leg of the trigger must be com-
 194 puted separately because the selection of the two leptons is different. The efficiency correction
 195 scale factors are nearly 1 for single lepton triggers as well as for both legs of the di-lepton
 196 triggers. The correction scale factor and their uncertainties for electron triggers are shown in
 197 Figures 4 and 5 as a function of electron p_T and η .

198 Following the muon POG recommendation, the double muon trigger efficiencies for data and
 199 MC are studied separately for run B, C, D, E, F, G (affected by the HIPs inefficiency) and run H.

200 For run B, C, D, E, F, G corrections are applied on the HLT path `HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_v*`
 201 OR `HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_v*` by evaluating the data and MC efficiencies
 202 separately for each leg. The corresponding scale factors are shown in Figure 6 as distributions in
 203 p_T and η . For run H, corrections are applied on the `HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_v*`
 204 OR `HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ_v*`, which is the same trigger as for run
 205 B, C, D, E, F, G with an additional cut on the dZ between the tracks of the two muon can-
 206 didates. The corrections are factorised: a first scale factor is computed on the HLT path without
 207 the additional dZ cut on run H, using the same methode as run B, C, D, E, F, G. The corre-

Table 8: List of L1 and HLT triggers used for the 2016 data set, and the channels to which they apply.

Channel	L1 Seeds	HLT Paths
$W(\mu\nu)H$	<code>L1_SingleMu20</code>	<code>HLT_IsoMu24 OR</code> <code>HLT_IsoTkMu24</code>
$Z(\mu\mu)H$	<code>L1_SingleMu20</code>	<code>HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_v* OR</code> <code>HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_v* OR</code> <code>HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_v* OR</code> <code>HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ_v*</code>
$W(e\nu)H$	<code>L1_SingleIsoEG22er OR</code> <code>L1_SingleEG25</code>	<code>HLT_Ele27_WPTight_Gsf</code>
$Z(ee)H$	<code>L1_SingleEG30 OR</code> <code>L1_SingleIsoEG22er OR</code> <code>L1_SingleIsoEG24 OR</code> <code>L1_DoubleEG_15_10</code>	<code>HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ</code>
$Z(\nu\bar{\nu})H$	<code>L1_ETM50 L1_ETM60 L1_ETM70 L1_ETM80</code>	<code>HLT_PFMET110_PFMHT110_IDTight OR</code> <code>HLT_PFMET120_PFMHT120_IDTight OR</code> <code>HLT_PFMET170_NoiseCleaned OR</code> <code>HLT_PFMET170_HBHECleaned OR</code> <code>HLT_PFMET170_HBHE_BeamHaloCleaned</code>

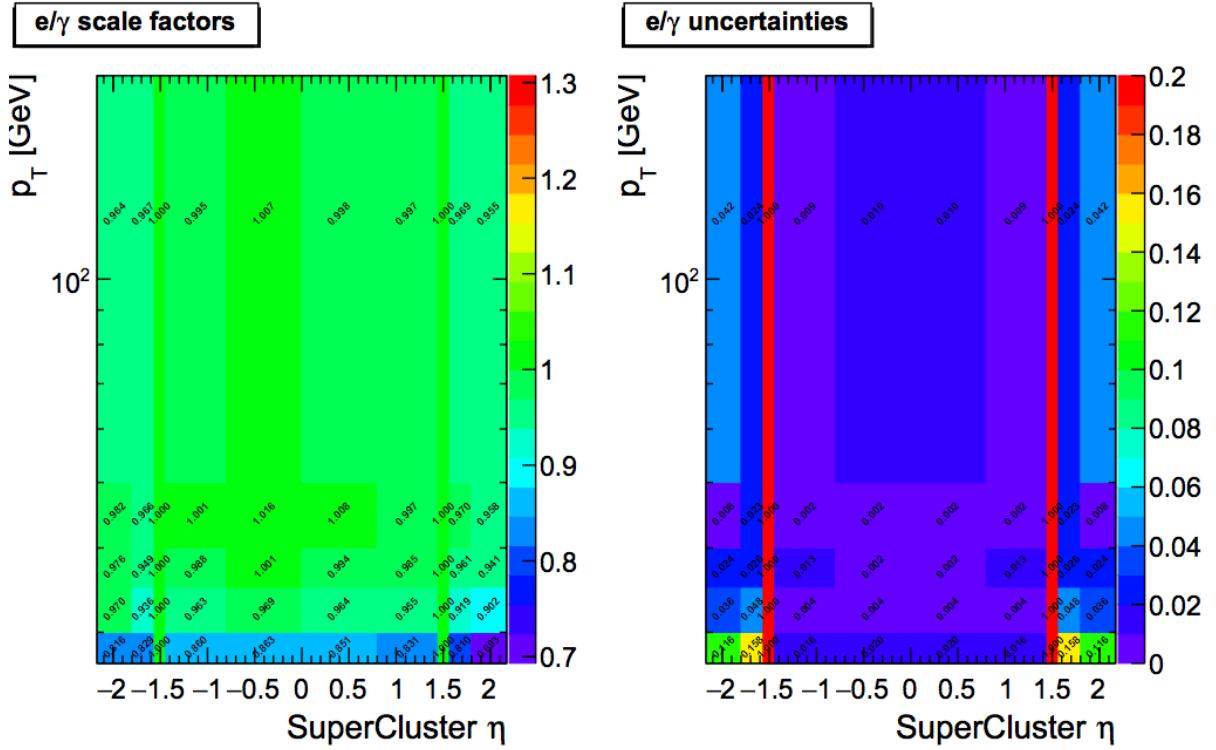


Figure 4: Distributions of HLT_Ele27_WPTight_Gsf efficiency as function of p_T and η for 2016 data. The efficiencies are measured after applying WP80 in the general purpose electron MVA IDs plus isolation selection. The turn-on can be seen as rising efficiency in p_T above 27 GeV.

sponding scale factors are shown in Figure 7 as distributions in p_T and η . A second scale factor is then computed to account for the dZ selection, requiring the numerator to pass the full HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_v* OR HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ_v* trigger path, and the denominator to pass the trigger path without any dZ requirement. The corresponding scale factors are shown in Figure 8.

For $Z(\nu\nu)H$, a dedicated trigger was developed and it is described in the following paragraph.

3.1 Dedicated trigger for $Z(\nu\nu)H$

The main trigger of the $Z(\nu\nu)H$ is HLT_PFMET110_PFMHT110_IDTight which is seeded at L1 by an OR or L1.ETM triggers with thresholds spanning from 50 to 80 depending on the instantaneous luminosity of the LHC. It cuts on both PF MET and PF MHT in order to reduce the rate and keep low thresholds. The online PF MHT is evaluated using PF corrected jets with $p_T > 20$ GeV, $|\eta| < 5.2$, and passing the following ID requirements:

- Neutral Hadronic Fraction < 0.9 ;
- Neutral Electromagnetic Fraction < 0.99
- Number of Constituents > 1 ;
- Charged Hadronic Fraction > 0 ;
- Charged multiplicity > 0 ;
- Charged Electromagnetic Fraction < 0.99 ;

The last three cuts are only applied to jets within the tracker acceptance ($|\eta| < 2.4$).

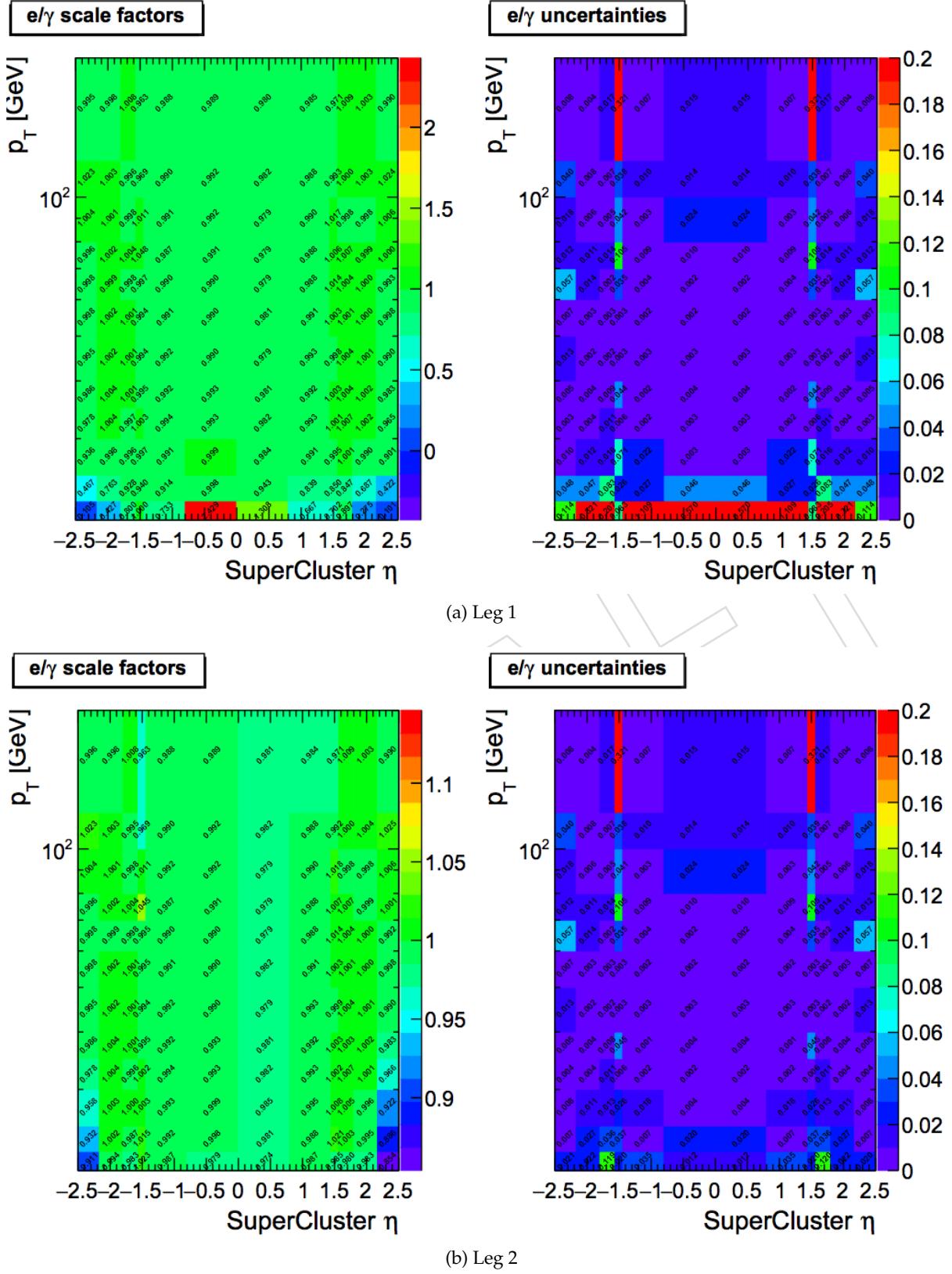


Figure 5: Distributions of HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ efficiency as function of p_T and η for 2016 data. The efficiencies are measured after applying WP90 in the general purpose electron MVA IDs plus isolation selection.

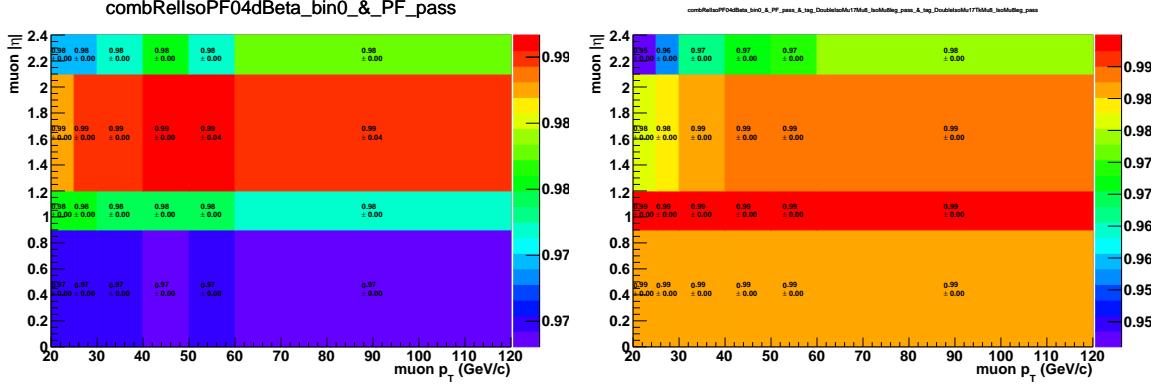


Figure 6: Distributions for each leg of the HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_v* OR HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_v* trigger as function of p_T and η for Run B, C, D, E, F, G. The left figure corresponds to the 8 GeV leg scale factor. The right figure correspond to the 17 GeV leg scale factor, requiering the second muon to pass the 8 GeV leg.

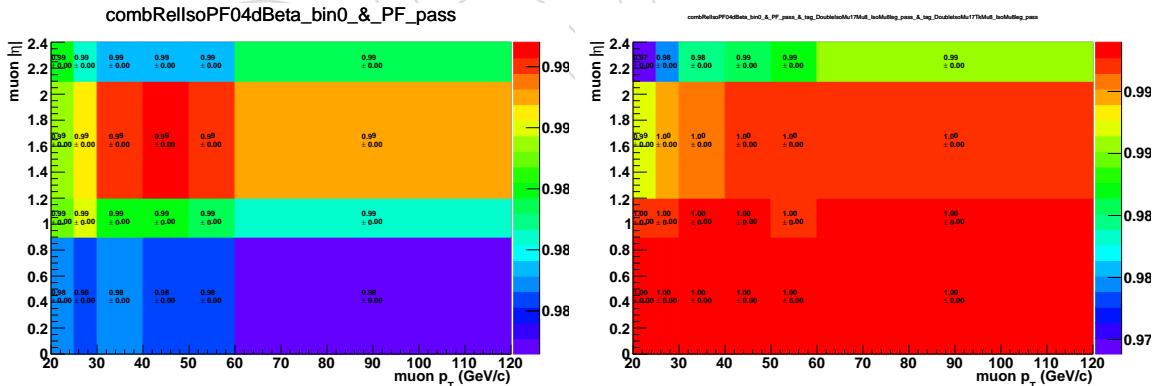


Figure 7: Distributions for each leg of the HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_v* OR HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_v* trigger as function of p_T and η for Run H. The left figure corresponds to the 8 GeV leg scale factor. The right figure correspond to the 17 GeV leg scale factor, requiering the second muon to pass the 8 GeV leg.

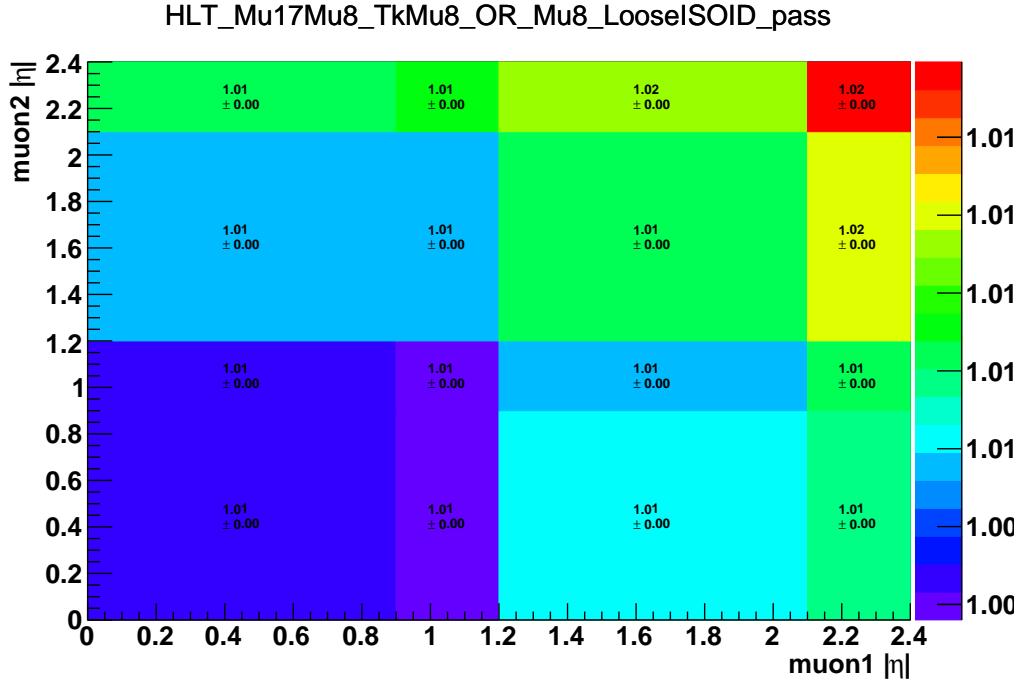


Figure 8: Distribution of the dZ requirement on top of the $\text{HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_v}^*$ OR $\text{HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_v}^*$ trigger as a function of the $|\eta|$ of each muon.

227 The `HLT_PFMET110_PFMHT110_IDTight` trigger cuts both on MET and MHT at the same
 228 threshold. For this reason, the trigger efficiency has been studied as a function of the offline
 229 $\min(\text{MET}, \text{MHT})$. The pure MET trigger `HLT_PFMET170` has been used to recover a possible
 230 small inefficiency due to the online MHT cut in the large MET region.

231 The overall trigger efficiency has been measured using the data collected by the single-muon
 232 and single-electron triggers and requiring the presence of two jets in the tracker acceptance in
 233 the event. In order to avoid bias from the L1 MET (calo-MET), the lepton is required not to
 234 be back to back with the reconstructed MET. The measured efficiency is then applied on the
 235 simulation.

236 Figure 9 shows the trigger efficiency for various triggers and the OR as function of the offline
 237 $\min(\text{MET}, \text{MHT})$ distributions obtained in data in the single-electron (right) datasets. The top
 238 plots have been obtained on the full dataset. The points are then fitted with an exponential
 239 function of the form

$$p2 * (1e0 - \exp(-p0 * (x - p1))), \quad (1)$$

240 where $p2$ represents the efficiency at plateau, $p1$ the x -value where the efficiency starts to rise and $p0$ gives a measure of how slow is the turn-on.

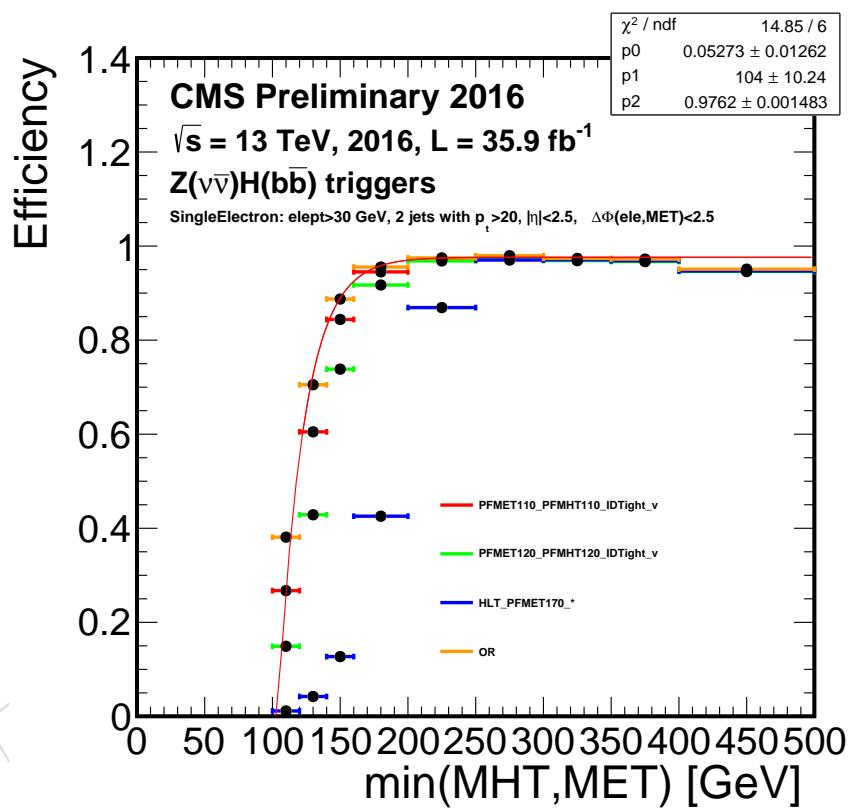


Figure 9: Distributions of trigger efficiency as function of $\text{min}(\text{MET}, \text{MHT})$ for the data in the single-electron for 2016 full dataset.

242 4 Physics Objects

243 This analysis uses standard physics objects provided in the miniAOD data format and ap-
 244 proved by the relevant POGs. The `heppy` framework [46], developed on GitHub, is used to
 245 process the events, select them and produce plain ROOT trees for all Data and Monte Carlo
 246 samples listed in Section 2.

247 Sections 4.1–4.8 describe the reconstruction, identification, and selection of electrons, muons,
 248 jets, b jets, and missing transverse energy. Since pile-up affects all of these physics objects, we
 249 begin with a description of the primary vertex selection and the methods applied to mitigate
 250 the effects of pile-up.

251 4.1 Primary vertex selection and pile-up treatment

252 Primary vertices positions where proton-proton collisions occurred are reconstructed using
 253 tracks clustered with the Deterministic Annealing algorithm [47]. We use the miniAOD
 254 “`offlineSlimmedPrimaryVertices`” collection. Reconstructed primary vertices are re-
 255 quired to be valid, to have a z position within 24 cm of the nominal detector center, a radial
 256 position within 2 cm of the beamspot axis, and a vertex fit exceeding four degrees of freedom.

257 During the run-1, the signal vertex was chosen as the primary vertex with the largest $\sum_{\text{tracks}} p_T^2$.
 258 This algorithm has been improved during the Long Shutdown 1. The idea is to choose the
 259 vertex depending on the $\sum p_T^2$ of the elementary particles produced in the collision (quarks,
 260 gluons, charged leptons, photons, and neutrinos). To do that, the $\sum p_T^2$ is evaluated for each
 261 primary vertex using track-jets, track-MET, and charged leptons, instead of tracks. Moreover,
 262 displaced tracks originating from a B hadron decay are associated to the proper vertex exploit-
 263 ing the respective jet direction. This improved algorithm increases the probability to choose the
 264 correct vertex as signal vertex especially in $Z(\nu\nu)H$ channel, where only two b -jets are present
 265 in the final state.

266 Given the collisions’ instantaneous luminosity the data sample contains a significant number
 267 of additional interactions per bunch crossing—an effect known as pile-up (PU). The number of
 268 reconstructed primary vertices is related to the number of PU interactions in each triggered
 269 event, with a primary vertex reconstruction efficiency that is roughly around 70%. The pile-up
 270 distribution in collected data decreases as a function of time during each LHC fill and varies
 271 along the year because of different LHC collisions settings. In 2016 during run periods B, C and
 272 D the average pile-up multiplicity spanned between around 40 at the beginning of most fills
 273 to below 15 in near the end of the longest fills where bunch size growth and proton depletion
 274 have the largest impact.

275 Over the course of the LHC operation, in-time PU as well as out-of-time pile-up increased. Ef-
 276 fects from out-of-time pile-up are particularly more relevant after the LHC switched to bunch
 277 filling schemes with 25 ns spacing (as opposed to 50 ns minimum spacing employed through-
 278 out Run1 collisions). Figure 2 shows the distribution of primary vertices in $t\bar{t}$ events recon-
 279 structed as $W(\mu\nu)H$ candidates. The presence of PU interactions affects the reconstruction of
 280 jets in general, in particular the jet momentum resolution and the reconstruction of the Higgs
 281 mass, and the use of vetoes on additional jet activity. It also affects the MET reconstruction,
 282 lepton isolation and b-tagging. There are two distinct approaches to address all these effects
 283 (apart from MET):

- 284 • **PFnoPU:** also known as Charged Hadron Subtraction (CHS), PFPU is an algorithm
 285 embedded in the PF jet processing chain that attempts to filter all charged hadrons
 286 that do not appear to originate from the primary interaction. This approach is very

287 effective but only works in the pseudorapidity region covered by the Tracker. Algorithms for tagging b jets are not impacted, since they apply their own track pre-
 288 filtering that is also designed to be PU-resistant.

- 290 • **Fastjet:** is an external software package from which CMSSW takes virtually all its
 291 jet reconstruction services [48]. In particular it provides the means to calculate the
 292 momentum density per unit area ρ due to PU for each event, which can be used to
 293 subtract the contamination of jets and lepton isolation cones based on their respec-
 294 tive areas. These methods are therefore referred to as "Fastjet Subtraction."

295 Ideally, charged hadrons from PU interactions are filtered from the event first before the appli-
 296 cation of Fastjet. In this analysis, both the PFnoPU and Fastjet Subtraction methods are applied
 297 consistently in the reconstruction and identification of jets, and in the calculation of lepton
 298 isolation.

299 The RunIISpring16 80X Monte Carlo samples were generated with the so-called

300 2016_25ns_SpringMC_PUScenarioV1_PoissonOOTPU

301 PU-distribution. That PU distribution was a conservative generated PU scenario to account
 302 for a range of PU evolution scenarios. In order to make MC match data we must reweight the
 303 simulated events by the number of generated PU. The standard reweighting technique [49] is
 304 used in this analysis. We use a central value of 69.2 mb for the total inelastic pp cross section,
 305 and we estimate the systematic uncertainties on the weights by changing the central value to
 306 66.02 and 72.38 mb, respectively.

307 4.2 Electrons

308 Electrons are reconstructed with the Gaussian Sum Filter algorithm (GSF Electrons) [50]. They
 309 are preselected by requiring $p_T > 7 \text{ GeV}$, $|\eta| < 2.4$, $d_{xy} < 0.05 \text{ cm}$, $d_z < 0.2 \text{ cm}$ (where both
 310 distances are taken with respect to the primary vertex), and a very loose relative isolation cut
 311 of 0.4, where the ρ -subtracted PF isolation in a cone of radius 0.3 is used (Sec. 4.4).

312 A tighter identification is then applied using a multivariate approach recommended by the
 313 electron-gamma (EGM) POG as documented here:

314 [https://twiki.cern.ch/twiki/bin/viewauth/CMS/
 315 MultivariateElectronIdentificationRun2](https://twiki.cern.ch/twiki/bin/viewauth/CMS/MultivariateElectronIdentificationRun2).

316 A dedicated multivariate discriminator is trained for electrons that pass a set of cuts meant to reproduce
 317 the detector cuts applied by the most common electron triggers. In this case, a set of offline cuts on
 318 ECAL-based electron quantities is applied on top of the multivariate discriminator to reproduce the
 319 conditions of the training sample:

```
320 pt>15 &
321 (abs(superCluster().eta)<1.4442 & full5x5_sigmaIetaIeta<0.012 &
322 hcalOverEcal<0.09 &
323 (ecalPFClusterIso/pt)<0.4 & (hcalPFClusterIso/pt)<0.25 &
324 (dr03TkSumPt/pt)<0.18 & abs(deltaEtaSuperClusterTrackAtVtx)<0.0095 &
325 abs(deltaPhiSuperClusterTrackAtVtx)<0.065) ||
326 (abs(superCluster().eta)>1.5660 & full5x5_sigmaIetaIeta<0.033 &
327 hcalOverEcal<0.09 &
328 (ecalPFClusterIso/pt)<0.45 & (hcalPFClusterIso/pt)<0.28 &
329 (dr03TkSumPt/pt)<0.18)
330 ).
```

331 Two cuts on the MVA ID discriminator [51] are applied defining two different working points based on

332 the expected selection efficiency of either 90% (loose, WP90) or 80% (tight, WP80).
 333 The loose WP90 working point is used for global event classification (based on vector boson type and
 334 decay), in the counting of additional leptons for the veto requirement, and in the event selection of the
 335 Z(ee)H channel.
 336 The tighter WP80 working point is used in the W($e\nu$)H channel to suppress the fake background in that
 337 final state. p_T thresholds of 30 and 20 GeV are applied in the W($e\nu$)H and Z(ee)H analyses, respectively,
 338 for the leading lepton. In the Z(ee)H channel, the trailing lepton is required to have p_T in excess of
 339 15 GeV.
 340 For 12.9/fb analysis we use mvaEleID-Spring15-25ns-Trig-V1-wpXX working points with chan-
 341 nel specific isolation cuts (0.06 for W($e\nu$)H). For the full dataset we anticipate using
 342 mvaEleID-Spring16-GeneralPurpose-V1-wpXX working points.

343 4.3 Muons

344 Muons are reconstructed from combined tracker and muon-chamber information (global muons) [52, 53]
 345 They are preselected by requiring the loose muon POG ID (see below), $p_T > 5$ GeV, $|\eta| < 2.4$, $d_{xy} <$
 346 0.5 cm, $d_z < 1.0$ cm (where both distances are taken with respect to the primary vertex), and a very loose
 347 relative isolation cut of 0.4, where the $\Delta\beta$ -subtracted PF isolation in a cone of radius 0.4 is used (Sec. 4.4).
 348 They are further required to pass standard criteria suggested by the Muon POG. Two WP are used, a
 349 loose and tight:

- 350 • Loose muon:
 - 351 • Particle-Flow Muon:
 isPFMuon()
 - 352 • is Global or Tracker Muon:
 isGlobalMuon() || isTrackerMuon()
- 355 • Tight muon:
 - 356 • the candidate is reconstructed as a Global Muon:
 isGlobalMuon()
 - 358 • Particle-Flow Muon:
 isPFMuon()
 - 360 • $\chi^2/ndof$ of the global-muon track fit:
 globalTrack() -> normalizedChi2() < 10.
 - 362 • at least one muon-chamber hit included in the global-muon track fit:
 globalTrack() -> hitPattern().numberOfValidMuonHits() > 0
 - 364 • muon segments in at least two muon stations; this implies that the muon is also an
 arbitrated tracker muon:
 numberOfMatchedStations() > 1
 - 367 • tracker track transverse impact parameter w.r.t. the primary vertex:
 fabs(muonBestTrack() -> dxy(vertex->position())) < 0.2
 - 369 • longitudinal distance of the tracker track wrt. the primary vertex:
 fabs(muonBestTrack() -> dz(vertex->position())) < 0.5
 - 371 • number of pixel hits:
 innerTrack() -> hitPattern().numberOfValidPixelHits() > 0
 - 373 • cut on number of tracker layers with hits:
 innerTrack() -> hitPattern().trackerLayersWithMeasurement() > 5

375 The same p_T threshold is applied to muons in the W($\mu\nu$)H and Z($\mu\mu$)H channels. For the latter, the
 376 trailing muon p_T threshold is lowered down to 10 GeV.

377 **4.4 Lepton isolation**

378 Lepton isolation is defined starting from the PF isolation equation:

$$R \equiv \frac{\sum_i [p_{T,i}(\text{chargedHadron}) + p_{T,i}(\text{neutralHadron}) + p_{T,i}(\text{Photon})]}{p_{T,\ell}} \quad (2)$$

379 adding a term for subtraction of PU energy and momentum. This subtraction is based on the per-event
 380 estimated neutral energy expected to enter a cone of radius 0.3 (0.4) around the electron (muon) momen-
 381 tum. For electrons the estimate is obtained by computing the standard ρ variable using only neutral par-
 382 ticle flow objects, and then multiplying by a POG-estimated effective area of the cone (`Spring15_25ns_v1`).
 383 For muons the correction is estimated from the deposit associated to charged tracks not belonging to the
 384 primary vertex, with calibration factors given by the muon POG.

385 Both muon and electron (WH) channels are required to have a relative isolation smaller than 0.06. They
 386 are the same by chance. In both cases the cuts were tightened to remove excess of data in poorly isolated
 387 distribution tails. Although the cuts are tight there is no loss in expected sensitivity and the data/mc
 388 agreement is very good in the bulk of the isolation distribution. In the $Z(\mu\mu)H$ channel, the threshold is
 389 relaxed to 0.12.

390 **4.5 Jets**

391 Jets are reconstructed from particle-flow candidates [54, 55], using the anti- k_T clustering algorithm
 392 with distance parameter $R = 0.4$ [56, 57]. Reconstructed jets require a small additional energy correction,
 393 mostly due to thresholds on reconstructed tracks and clusters in the PF algorithm and various recon-
 394 struction inefficiencies [58]. Loose jet identification criteria are also applied to reject misreconstructed
 395 jets resulting from detector noise, as well as jets heavily contaminated with pileup energy (clustering
 396 of energy deposits not associated with a parton from the primary pp interaction) [59]. Jets that overlap
 397 geometrically ($\Delta R < 0.4$) with preselected electrons or muons are discarded. Jets calibrated using the
 398 official JME POG prescriptions are considered for the analysis if $|\eta| < 2.5$. In the WH and $Z(\nu\nu)H$ chan-
 399 nels a minimum threshold of $p_T > 25$ GeV is used, while a looser selection ($p_T > 20$ GeV) is applied in
 400 the cleaner $Z(\ell\ell)H$ channels with lower dijet boost.

401 We smear the jet energy resolution using the prescription from the JetMET POG:

402 <https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetResolution>.

403 Based on the observed data/MC difference in energy resolution, we smear all jets in simulated samples
 404 that are matched to generator level jets using the standard geometrical PAT matching. The smear factors
 405 range between 6% and 12% when moving from the central part of the detector towards the endcaps, and
 406 are applied irrespectively of the jet p_T . These smeared jet energies are then used as default throughout
 407 the analysis for the both signal and background. The effect on the mass resolution in simulation is
 408 almost negligible (few % relative). On top of this smearing, we apply the standard scale and resolution
 409 systematic uncertainty procedures.

410 **4.6 Identification of b jets**

411 The identification of jets that originate from the hadronization of b quarks is done with the Combined
 412 MVA v2 (cMVAv2) algorithm [60] with PAT string:

413 `pfCombinedMVAV2BJetTags`

414 (while for the previous iteration of the analyses we have instead used the Combined Secondary Vertex
 415 CSVv2 algorithm). The CMVAv2 algorithm provides a continuous discriminator output combining in
 416 an optimal way the information about track impact parameters and identified secondary vertices within
 417 jets, even when full vertex information is not available, and information of any soft lepton present in the
 418 jet. Additional categories for jets where a “pseudo vertex” is found, or no vertex at all is identified, can

419 be defined and combined in a multivariate discriminant (Boosted Decision Tree) to provide maximal
 420 separation of b jets from the much larger background of jets arising from charm decay, and from the
 421 fragmentation of light quarks and gluons.

422 The CMVAv2 output that can be used to select optimal working points with respect to the VH analyses,
 423 in addition to the standard Loose/Medium/Tight working points defined by the BTV POG: CMVAv2L
 424 (> -0.5884), CMVAv2M (> 0.4432), and CMVAv2T (> 0.9432). Independent optimizations of the
 425 selection criteria in all five channels arrive at roughly the same optimal selection for the jet in the Higgs
 426 decay that has the higher value of the CMVAv2 output: very close to the CMVAv2T working point.
 427 For the second jet, the optimal selection typically falls between the CMVAv2L and CMVAv2M working
 428 points.

429 4.7 Efficiency for b-jet identification

430 The calibration of the CMVAv2 discriminator is determined using a tag-and-probe method as docu-
 431 mented in Ref. [61]. This method attempts at correcting the distribution of the CMVAv2 discriminator
 432 for simulated jets as to match the distribution observed in data control regions. These control regions
 433 are preselected by the requiring at least two opposite-sign leptons plus at least two jets. Two exclusive
 434 set of selections based on the dilepton mass and the CMVAv2 discriminator or a “tag” jet are further
 435 imposed to enrich the control regions in Z+jets or t̄t, respectively. The binned CMVAv2 distribution of
 436 the “probe” jet is the compared to the one expected from simulation. An iterative procedure is then
 437 initiated to scale every bin content in the simulation simultaneously for light and heavy flavour jets
 438 (hadronFlavour ()=0 and hadronFlavour ()=5, respectively). This procedure is carried out for
 439 various p_T and $|\eta|$ bins. The ratio between the re-scaled distribution in the simulated sample and the
 440 original one is used as a jet-by-jet weight, w_j , defined as

$$w_j(\text{CMVAv2}_j; p_{Tj}, |\eta_j|, \text{flavour}_j). \quad (3)$$

441 By construction, the weights average to one when sampling CMVAv2 using the distribution predicted
 442 by the simulation: $N^{-1} \sum_{i=1}^N w_i \rightarrow 1$, for $N \gg 1$ jets. In an event with N_{jet} selected jets whose CMVAv2
 443 discriminator is used in the analysis, an event weight w is defined starting from the jet weights as:

$$w = \prod_{j=1}^{N_{\text{jet}}} w_j(\text{CMVAv2}_j; p_{Tj}, |\eta_j|, \text{flavour}_j). \quad (4)$$

444 4.8 Missing transverse energy

445 The use of missing transverse energy is central to the analyses presented in this note. It is critical in
 446 the reconstruction of the $W \rightarrow \ell\nu$ and $Z \rightarrow \nu\bar{\nu}$ decays, and is used in the $Z(\mu\mu)H$ channel as a control
 447 sample to evaluate the comparison of data and simulation in the $Z(\nu\nu)H$ channel. It is also used in
 448 the $Z(\ell\ell)H$ to increase the purity of the $t\bar{t}$ control sample. For the offline analysis, missing transverse
 449 energy is computed from the list of particle-flow objects with the method described in Ref. [62]. The
 450 vector $\vec{E}_{\text{T}}^{\text{miss}}$ is calculated as the negative of the vectorial sum of transverse momenta of all particle-
 451 flow objects identified in the event, and the magnitude of this vector is referred to as “pfMET”. The
 452 estimation of the $\vec{E}_{\text{T}}^{\text{miss}}$ vector in simulated events is improved by correcting it for the difference between
 453 raw (i.e. uncorrected) and calibrated jets (Sec. 4.5), including the the scale and resolution corrections,
 454 with $p_T > 15 \text{ GeV}$, $|\eta| < 4.7$, and passing a set of filters meant to remove electron and muon candidates
 455 as explained in

456 <https://twiki.cern.ch/twiki/bin/viewauth/CMS/MissingETRun2Corrections>.

457 In addition, we apply a set of recommended filters to remove known issues of instrumental noise and
 458 problematic events as explained in

459 <https://twiki.cern.ch/twiki/bin/view/CMS/MissingETOptionalFiltersRun2>.

460 The pfMET value divided by the scalar sum of E_T of all particle-flow objects is referred to as the “pfMET
 461 significance.”. For the $Z(\nu\nu)H$ analysis, pfMET is required to be greater than 120 GeV.

462 In $Z(\nu\nu)H$ analysis, pfMHT is used too. It is defined as: $\text{pfMHT} = \sum_i \vec{p}_T(\text{jet})$, considering jets with
 463 $p_T > 30$ and $|\eta| < 2.4$.

464 4.9 Additional “soft” hadronic activity

465 For the searched VH signal events not much additional hadronic activity is expected after excluding the
 466 V and H decay products. Since the amount of additional radiation is expected to be small (soft) care can
 467 be taken to avoid the contributions from pileup interactions, and therefore make only use of charged
 468 tracks that clearly originate from the event main interaction point to monitor the additional radiation.
 469 As described above, the main interaction point in the event is defined as the “hardest” reconstructed
 470 primary vertex (PV), i.e. with the largest p_T^2 sum for the tracks that have been used to reconstruct it.

471 The additional soft activity is defined as follows. At first a collection of “Additional tracks” is built using
 472 reconstructed tracks that

- 473 • have a *high purity* quality flag ,
- 474 • have $p_T > 300$ MeV/ c ,
- 475 • are not associated to the vector decay leptons, nor to the selected two b-jets in the event
 (through the PF candidates components track references),
- 477 • make minimum $|d_z(\text{PV})|$ when associated to the event hardest primary vertex (PV),
- 478 • satisfy $|d_z(\text{PV})| < 2$ mm with respect to the hardest PV.

479 In addition to removing the tracks in the jets, also tracks in the region between the two b-jets are removed
 480 by defining a ellipse in the $\eta\phi$ plane around the two b-jets with axes $(a, b) = (\Delta R(\text{bb}) + 1, 1)$, and
 481 excluding all tracks pointing within the ellipse from the additional tracks collection.

482 After this track selection a collection of “soft track-jets” is build clustering the “Additional tracks” col-
 483 lection with the anti- k_T clustering algorithm [57] with distance parameter $R = 0.4$. The use of track-jets
 484 represents a clean and commissioned method [63] to reconstruct the hadronization of partons with very
 485 low energies, down to few GeV [64].

486 For the purpose of separating the signal from backgrounds, we make use of clustered soft TrackJets, and
 487 eventually consider

- 488 • the scalar $p_{T\text{sum}}$ of the soft TrackJets with transverse momentum $p_T > 1$ GeV, H_T^{soft} ;
- 489 • the soft TrackJet multiplicity N_2^{soft} with transverse momentum $p_T > 2$ GeV, N_2^{soft} ;
- 490 • the soft TrackJet multiplicity N_5^{soft} with transverse momentum $p_T > 5$ GeV, N_5^{soft} ;
- 491 • the soft TrackJet multiplicity N_{10}^{soft} with transverse momentum $p_T > 10$ GeV, N_{10}^{soft} ;

492 The soft hadronic activity is used as discriminating variable in the regression for all the channels, as
 493 discussed in Section 9.

494 5 Vector Boson Reconstruction

495 Reconstruction of W and Z bosons begins with the identification and selection of charged leptons and
 496 pfMET described in the previous section. Given the unique signature of a boosted vector boson recoiling
 497 from two jets, the dominant background is from real W and Z decays. Therefore, a minimal selection is
 498 sufficient to identify highly pure samples of V+jets events.

499 Candidate $Z \rightarrow \ell\ell$ decays are reconstructed by combining isolated electrons and muons and requiring
 500 the dilepton invariant mass to satisfy $75 < M_{\ell\ell} < 105$ GeV. This analysis is performed with one bin in
 501 $p_T(Z) (> 50$ GeV).

502 Candidate $W \rightarrow \ell\nu$ decays are identified primarily by the topology of a single isolated lepton and ad-
 503 ditional missing transverse energy. The transverse momentum $p_T(W)$ and mass M_T of the W candidate
 504 are computed as:

$$p_T(W) = \sqrt{(p_{\text{fMET}} + p_x^\ell)^2 + (p_{\text{fMET}} + p_y^\ell)^2}, \quad \text{and} \quad (5)$$

$$M_T = \sqrt{(p_{\text{fMET}} + p_T^\ell)^2 - p_T(W)^2}. \quad (6)$$

505 It is observed that in the boosted regime, where the QCD background is much reduced, simply requiring
 506 $p_T(W) > \sim 100$ GeV is sufficient to select a relatively clean sample of real W decays. This analysis is
 507 performed with one bin in $p_T(W) (> 100$ GeV).

508 For inclusive W production, the distribution of M_T reflects the characteristic Jacobian peak and is very
 509 effective at separating signal from the large background of generic QCD production at small values
 510 of the transverse mass. In contrast, for the high boost used in this analysis, the neutrino begins to
 511 overlap with the lepton in azimuth, creating a broad flat region in M_T between 0–50 GeV that reduces
 512 the effectiveness of this variable in rejecting QCD background. Therefore, no selection is applied on M_T
 513 in the reconstruction of W candidates in the signal region. However, M_T remains effective at reducing
 514 QCD background in the low-boost $Wb\bar{b}$ control region (see Sec. 11), and in generally cleaning up the
 515 background in all of the control regions for the electron mode.

516 Candidate $Z \rightarrow \nu\bar{\nu}$ decays are reconstructed simply requiring $\text{pfMET} > 120$ GeV and $\text{pfMHT} > 120$ GeV.
 517 The transverse momentum of the Z candidate is defined as $p_T(Z) = \min(\text{pfMET}, \text{pfMHT})$. The $Z(\nu\nu)H$
 518 analysis is performed in two $p_T(Z)$ bins: 120 GeV $< p_T(Z) < 150$ GeV and $p_T(Z) > 150$ GeV.

519 5.1 Vector boson transverse momentum reweighting in V+jets events

520 In an inclusive selection enriched in vector boson production a shape difference in the vector boson
 521 transverse momentum of the simulated LO samples with respect to data has been observed, both in W
 522 and Z events. The observed data has a softer spectrum than the data and a negative correction with
 523 increasing $p_T(V)$ is necessary to correct for the effect.

524 This negative correction is expected to stem from higher order QCD and electroweak corrections to the
 525 vector boson production. The former correction is derived for the different H_T bins while the latter is
 526 derived as a function of boson $p_T(V)$ at generator level. The correction factors are multiplicative factors,
 527 more details could be found in Ref. [65].

528 The function used for the electroweak correction is shown in Fig. 10. The effect of the correction is
 529 shown in Fig. 11, where an improvement of the shape description and normalization is found.

530 For the time being, no correction is applied to the signal samples.

531 5.2 EWK signal VH corrections

532 The signal Monte Carlo sample for $qqVH$ is produced with POWHEG+MiNLO and then rescaled to
 533 NNLO QCD. The total cross section σ^{VH} is given by [11]:

$$\sigma^{WH} = \sigma_{\text{NNLOQCD}}^{\text{WH,DY}}(1 + \delta_{\text{EW}}) + \sigma_{\text{t-loop}} + \sigma_\gamma, \quad (7)$$

$$\sigma^{ZH} = \sigma_{\text{NNLOQCD}}^{\text{ZH,DY}}(1 + \delta_{\text{EW}}) + \sigma_{\text{t-loop}} + \sigma_\gamma + \sigma^{ggZH} \quad (8)$$

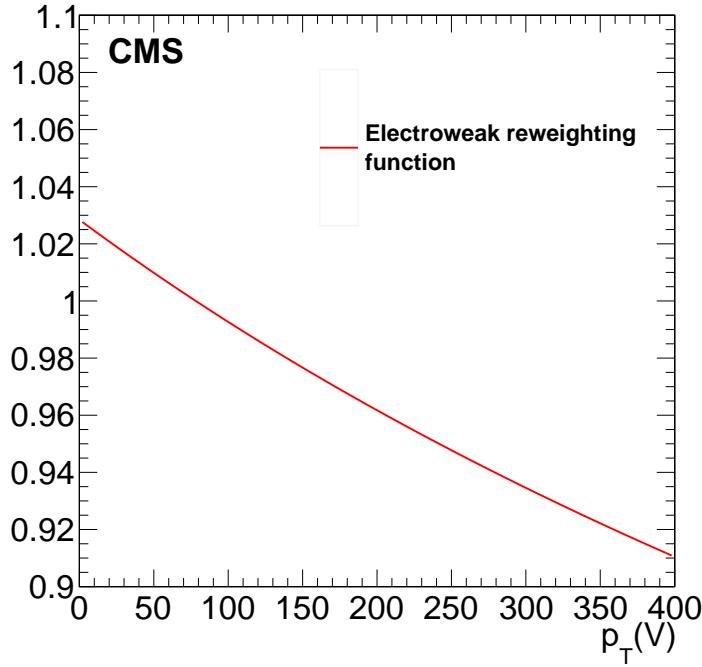


Figure 10: Electroweak correction as a function of boson $p_T(V)$.

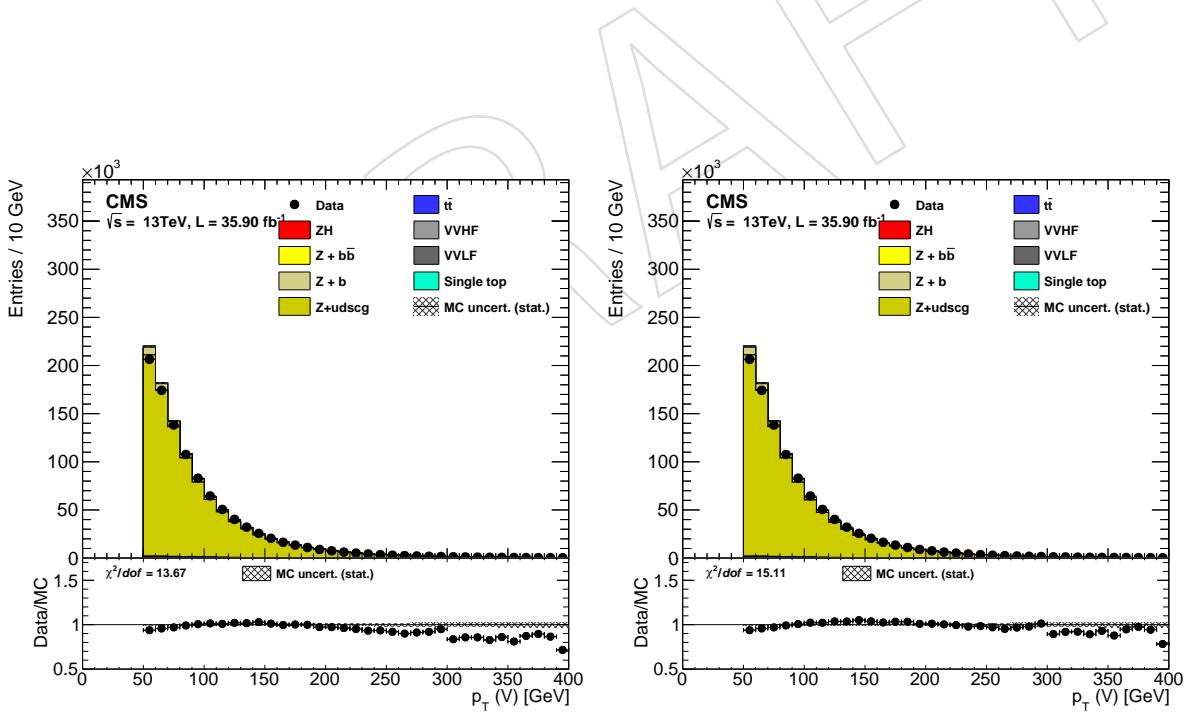


Figure 11: Impact of the p_T re-weighting on the vector boson transverse momentum at generator level in events with DY+light jets. Data and simulation comparison before re-weighting (left) and after it (right).

534 Since up to NLO, the electroweak corrections factorize, they can be applied differentially (in $p_T(V)$) as
 535 multiplicative weight $(1 + \delta_{EW})$ to the signal. Figure 12 shows the shape of these corrections for two of
 536 the channels.

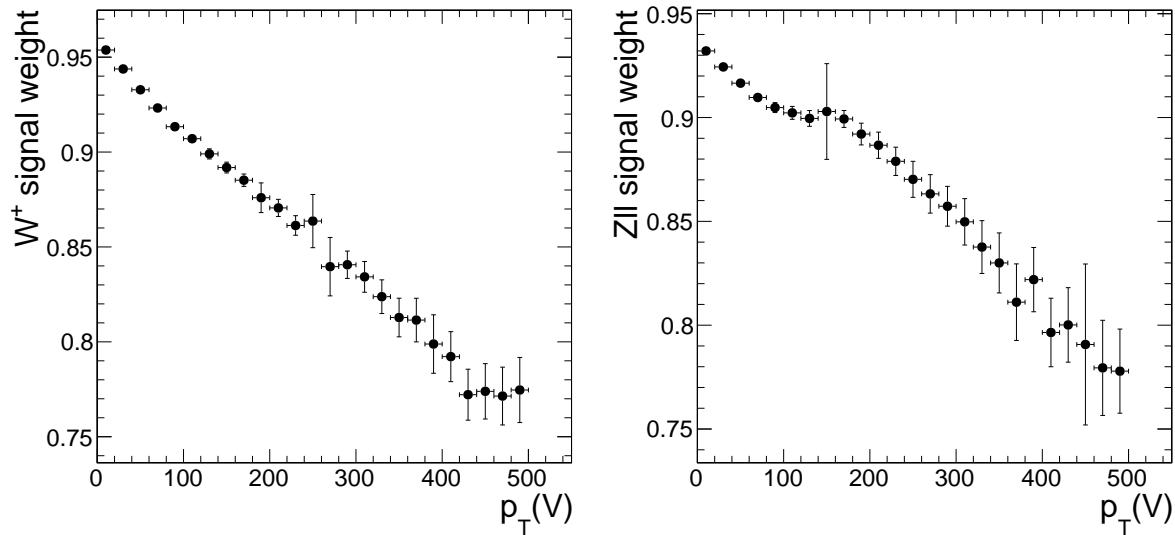


Figure 12: Multiplicative weights to apply the differential NLO electroweak signal correction.

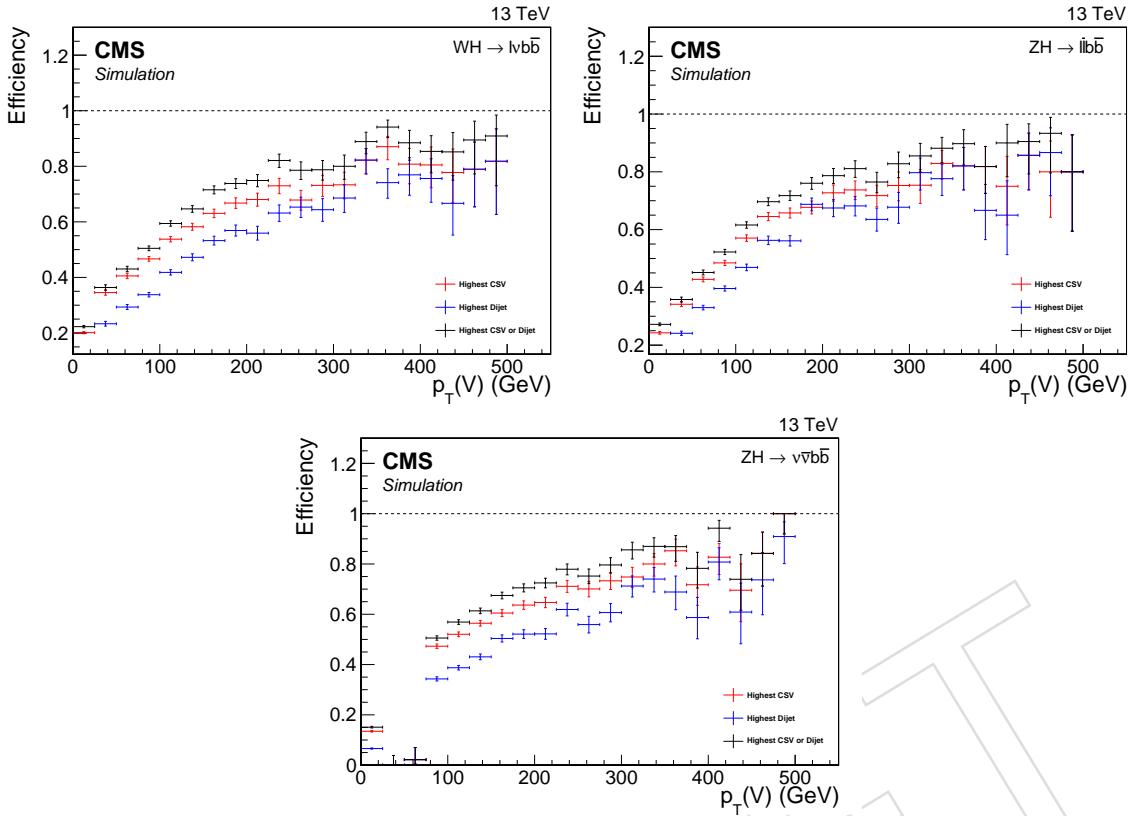


Figure 13: The efficiency of selecting correctly the jets from the Higgs boson decay to $b\bar{b}$ as a function of $p_T(V)$ for $W(\ell\nu)H$ (top left), $Z(\ell\ell)H$ (top right), and $Z(\nu\nu)H$ (bottom). The selection is considered correct if the reconstructed Higgs matches the generated Higgs within a cone of $\Delta R < 0.3$. The plots show the performance of selecting the Higgs jets as the two most b -tagged jets (red), as the two highest p_T jets (blue), and from the *OR* of the two previous strategies (black). The plots were obtained with the requirement that the two b quarks originating from the Higgs decay have $p_T > 20$ GeV and $|\eta| < 2.5$.

537 6 Higgs Boson Reconstruction

538 The reconstruction of the $H \rightarrow b\bar{b}$ decay is a fundamental element of this analysis. In order to maximise
 539 the signal sensitivity it is critical to achieve the best possible dijet mass resolution for the Higgs
 540 candidate, while optimizing the correct selection of the two signal jets, rejecting wrong combinations.

541 Requiring a boost for the dijet system or a high threshold separately on the p_T of each jet generally
 542 improves the dijet mass resolution. Candidate $H \rightarrow b\bar{b}$ decays are selected by identifying the dijet
 543 combination with the largest associated CMVA output values. Figure 13 compares the performance in
 544 simulation of a few possible algorithms of Higgs jets selection.

545 6.1 Regression

546 Resolution on the dijet mass is improved significantly using regression techniques previously developed
 547 at the Tevatron [66]. A correction is computed for individual jets that attempts to more accurately
 548 estimate the true b -jet energy, thus improving the resolution and reducing the scale bias on a per-jet
 549 basis. The resulting improvement in the dijet mass resolution is approximately 15–20% depending on
 550 the channel. All searched channels use a common set of input variables for the b -jet energy regression
 551 as explained in the following.

552 6.1.1 Configuration and training

553 The TMVA package [67] is used to perform the regression. We employ the `TMVARegression.C` driver
 554 configured to use the BDT algorithm and to target the ratio between the generator level p_T of jets in-
 555 cluding neutrinos (i.e., the p_T of the “genJet” associated with the reco jet) and the reconstructed jet. The
 556 regression is trained on ditop events in order to avoid biases towards the signal properties.

557 The complete set of variables used in the regression training are:

- 558 • p_T – transverse momentum of the jet after corrections;
- 559 • M_T – transverse mass of the jet after corrections;
- 560 • η – pseudorapidity of the jet;
- 561 • $ptLeadTrk$ – transverse momentum of the leading track in the jet;
- 562 • $vtx3dL$ – 3-d flight length of the jet secondary vertex;
- 563 • $vtx3deL$ – error on the 3-d flight length of the jet secondary vertex;
- 564 • $vtxPt$ – transverse momentum of the jet secondary vertex;
- 565 • $vtxNtrk$ – number of tracks associated with the jet secondary vertex;
- 566 • $neEmEf$ – fraction of jet constituents detected in the ECAL that have neutral charge;
- 567 • $neHEF$ – fraction of jet constituents detected in the HCAL that have neutral charge;
- 568 • $nPVs$ – number of primary vertices;
- 569 • $SoftLeptPtRel$ – relative transverse momentum of soft lepton candidate in the jet;
- 570 • $SoftLeptPt$ – transverse momentum of soft lepton candidate in the jet;
- 571 • $SoftLeptdR$ – distance in $\eta - \phi$ space of soft lepton candidate with respect to the jet axis;

572 All variables are common and used across all 5 modes.

573 The regression is trained on jets in signal MC events satisfying $p_T > 20\text{ GeV}$, $|\eta| < 2.4$, and CMVA
 574 greater than zero. The results are evaluated on independent single MC samples and shown in Fig. 14,
 575 where “before and after” plots are displayed. For each channel a different $p_T(jj)$ cut is applied as it will be
 576 described later, ($Z(\ell\ell)H$: $p_T(jj) > 50\text{ GeV}$, $Z(\nu\nu)H$: $p_T(jj) > 170\text{ GeV}$, $W(\ell\nu)H$: $p_T(jj) > 100\text{ GeV}$). The dijet
 577 invariant mass resolution before the regression depends on the Higgs boson p_T . Thus the improvement
 578 due to the additional corrections is modest for $Z(\nu\nu)H$ as it is evaluated by requiring a very large boost.
 579 The performance on the ZZ dijet invariant mass resonances is also shown in Fig. 15.

580 We find that both the resolution and the scale biases improve after the regression. Figures 16–20 show
 581 comparisons of data and MC for several control regions, demonstrating that backgrounds are not biased
 582 by the regression technique.

583 The most discriminating variables across all modes are kinematic, and this is due to the fact that most
 584 of the power of the regression is coming from the neutrinos present in semileptonic B decays, which
 585 appear as a mismeasurement of the b-jet energy. This effect is illustrated in Fig. 21, which shows the
 586 dijet mass before and after regression for the case where the b jet contains a muon, and when it does
 587 not, in $Z(\ell\ell)H$ signal events. This comparison shows essentially all of the improvement comes from the
 588 subset of events where the b jet contains an identified muon from semileptonic B decay. Therefore soft
 589 lepton variables are included in the regression across all modes.

590 6.1.2 Validation

591 Several studies have been performed to validate the regression technique. The ability of the simulation
 592 to reproduce each of the variables used as input to the regression is demonstrated in Figs. 22–26 using
 593 VH events in the $t\bar{t}$ control regions defined in Sec. 11. Due to the strip dynamic inefficiency, we do expect
 594 smaller SV reconstruction efficiency in data compared to the simulation, resulting in bad modeling of all
 595 the SV related observables. We will monitor these variables as soon as the re-reco and new MC samples
 596 will become available, which should provide a reasonable agreement.

597 Dijet balance in Z+jets

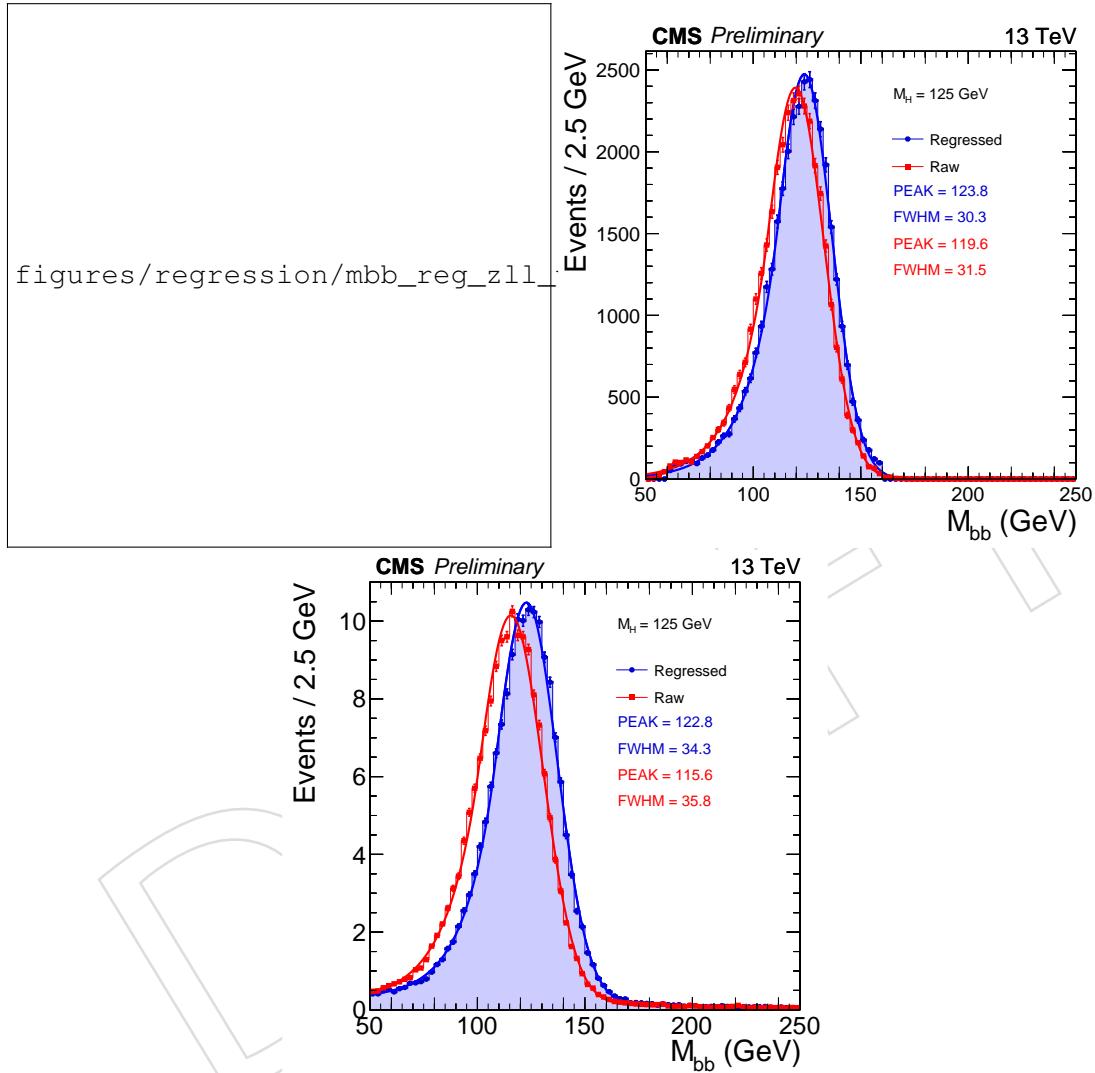


Figure 14: Distributions of dijet invariant mass in signal $Z(\ell\ell)H$ (top left), $Z(\nu\nu)H$ (top right), and $W(\ell\nu)H$ (bottom) events before and after the regression is applied. A combination of a Bernstein polynominal and a Crystal Ball is fit to the distribution, and the full-width half-maximum and peak are derived from the resulting fit.

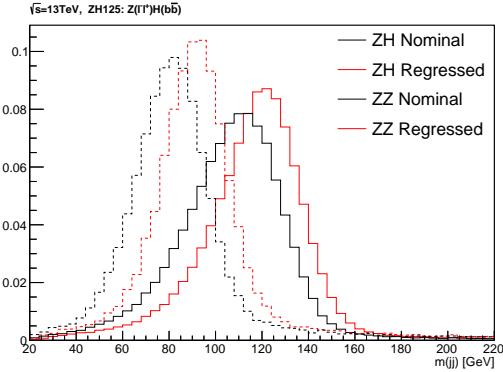


Figure 15: ZZ(bb) dijet invariant mass resonances in the $Z(\ell\ell')H$ channel before and after the regression is applied. Regression helps to improves the separation.

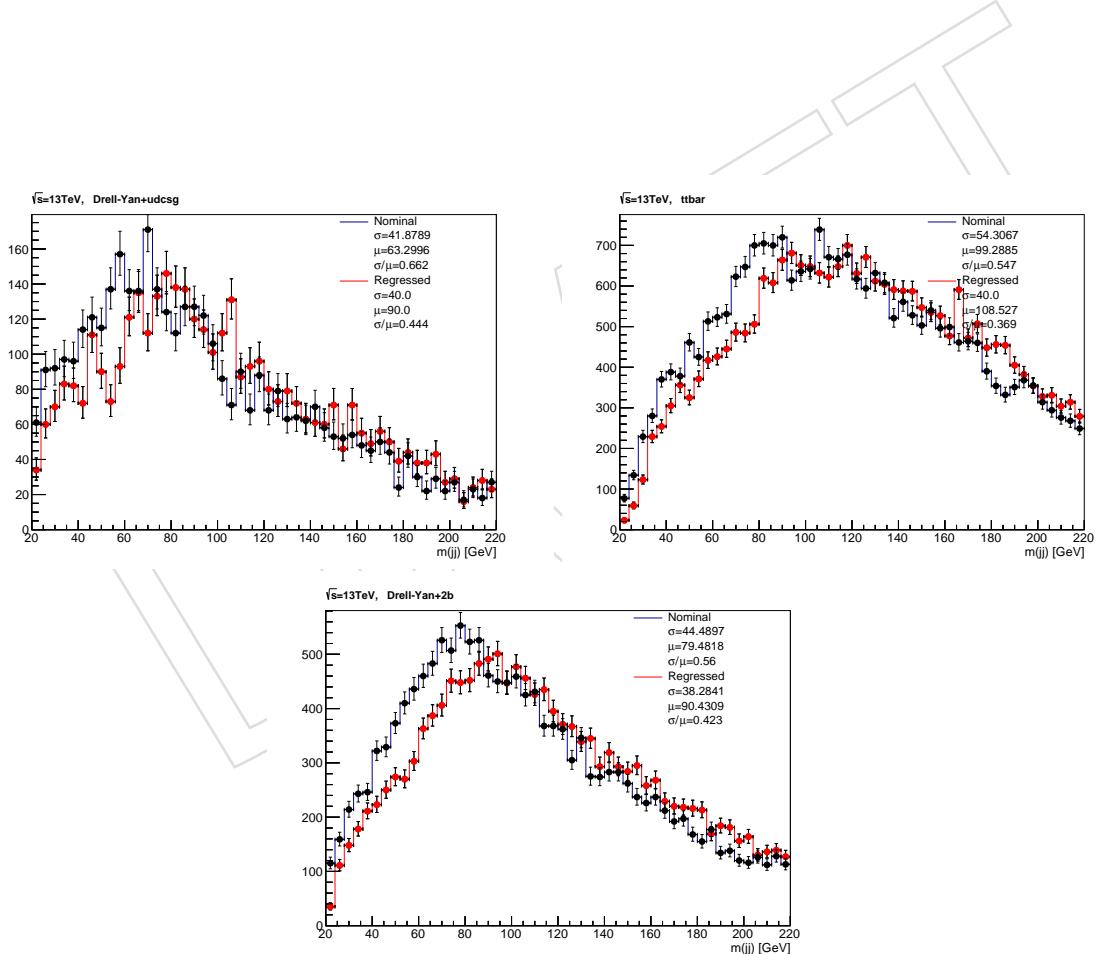


Figure 16: Distributions of dijet invariant mass in background control regions for $Z(\ell\ell')H$ enhanced in $Z + udscg$ (top), in $Z + b\bar{b}$ (middle) and $t\bar{t}$ (bottom) before (left) and after (right) the regression is applied.

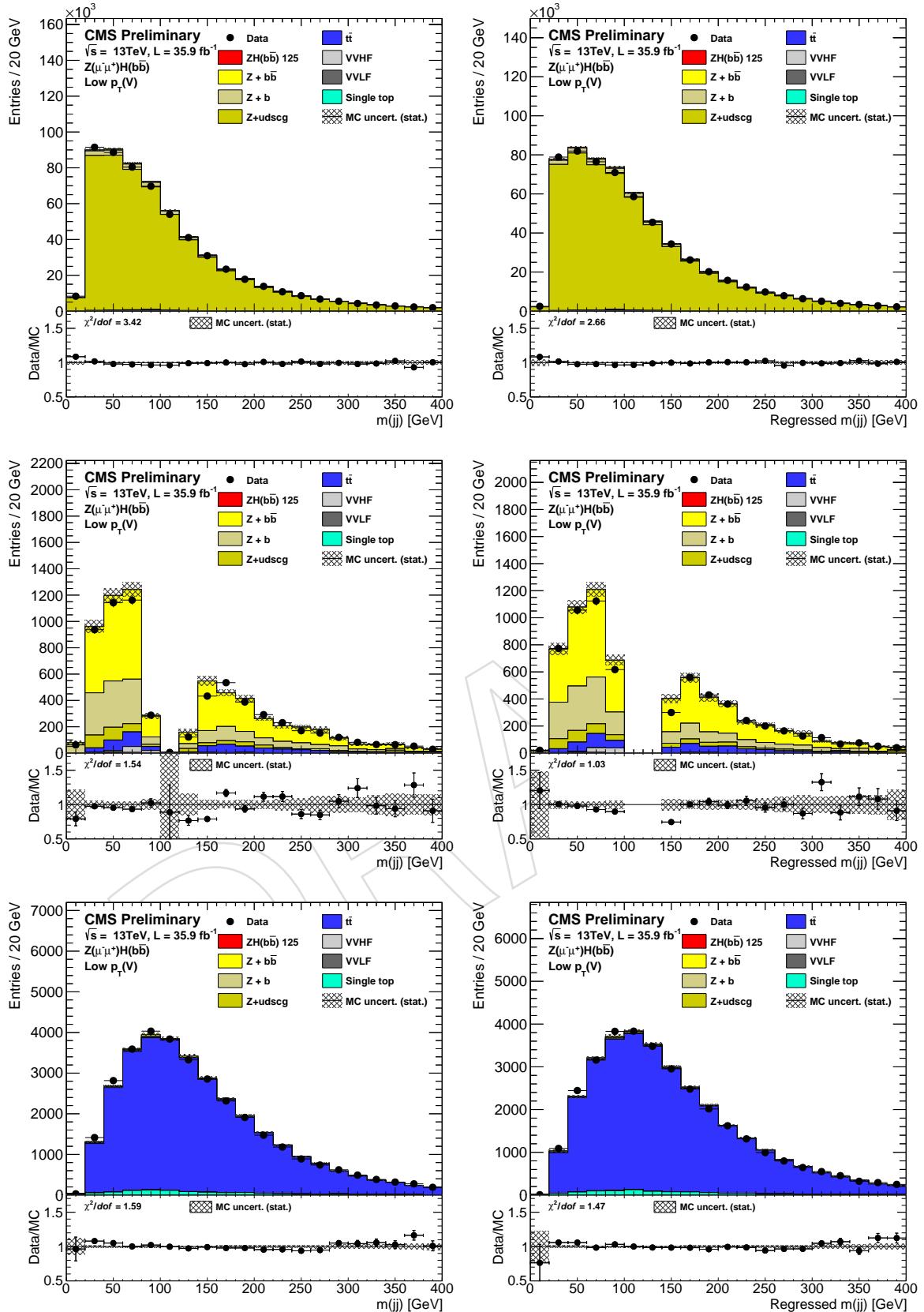


Figure 17: Distributions of dijet invariant mass in background control regions for $Z(\ell\ell)H$ enhanced in $Z + \text{udscg}$ (top), in $\text{t}\bar{\text{t}}$ (middle) and $Z + b\bar{b}$ (bottom) before (left) and after (right) the regression is applied.

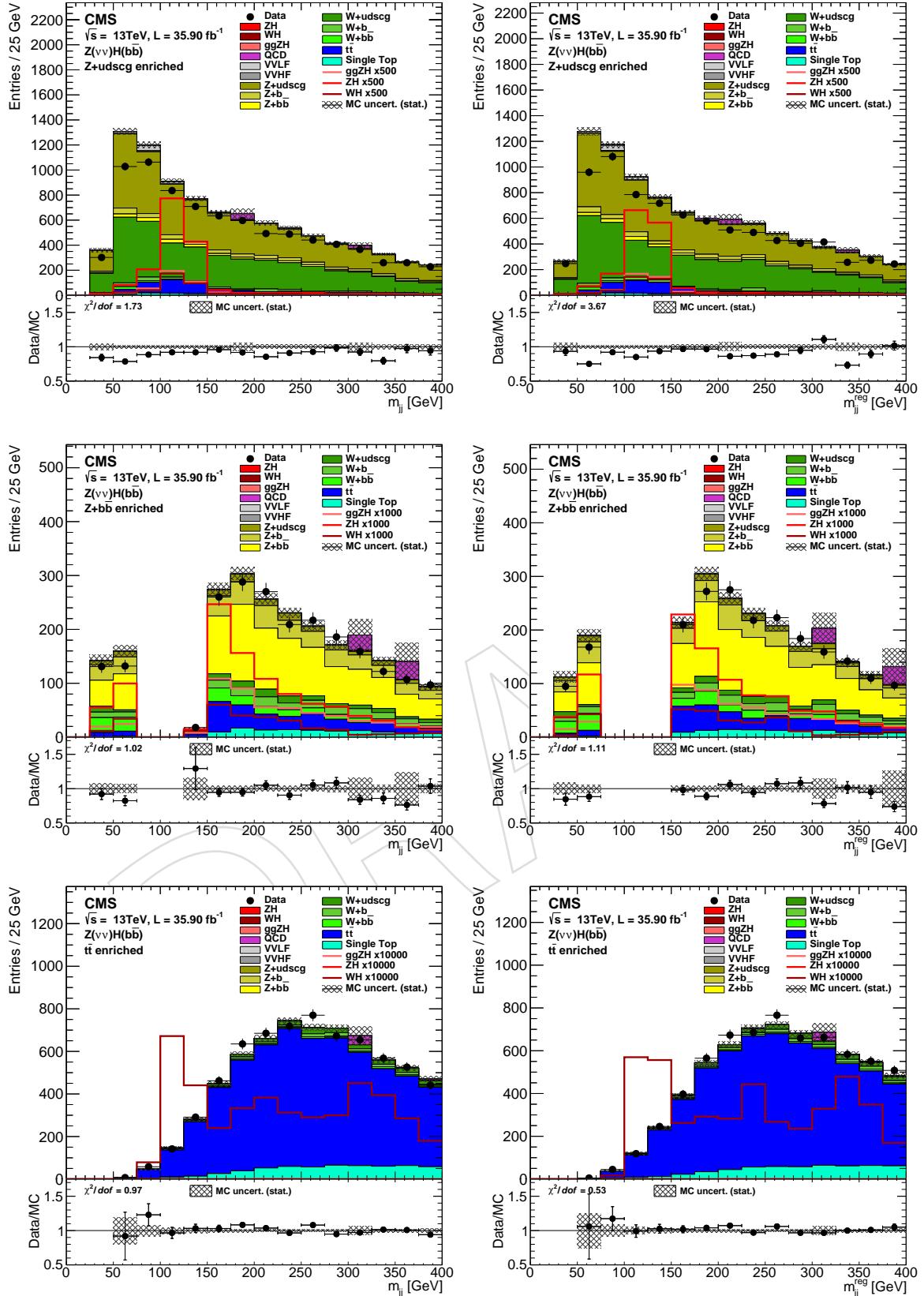


Figure 18: Distributions of dijet invariant mass in background control regions for $Z(vv)H$ enhanced in $Z + udscg$ (top), in $Z + b\bar{b}$ (middle) and $t\bar{t}$ (bottom) before (left) and after (right) the regression is applied.

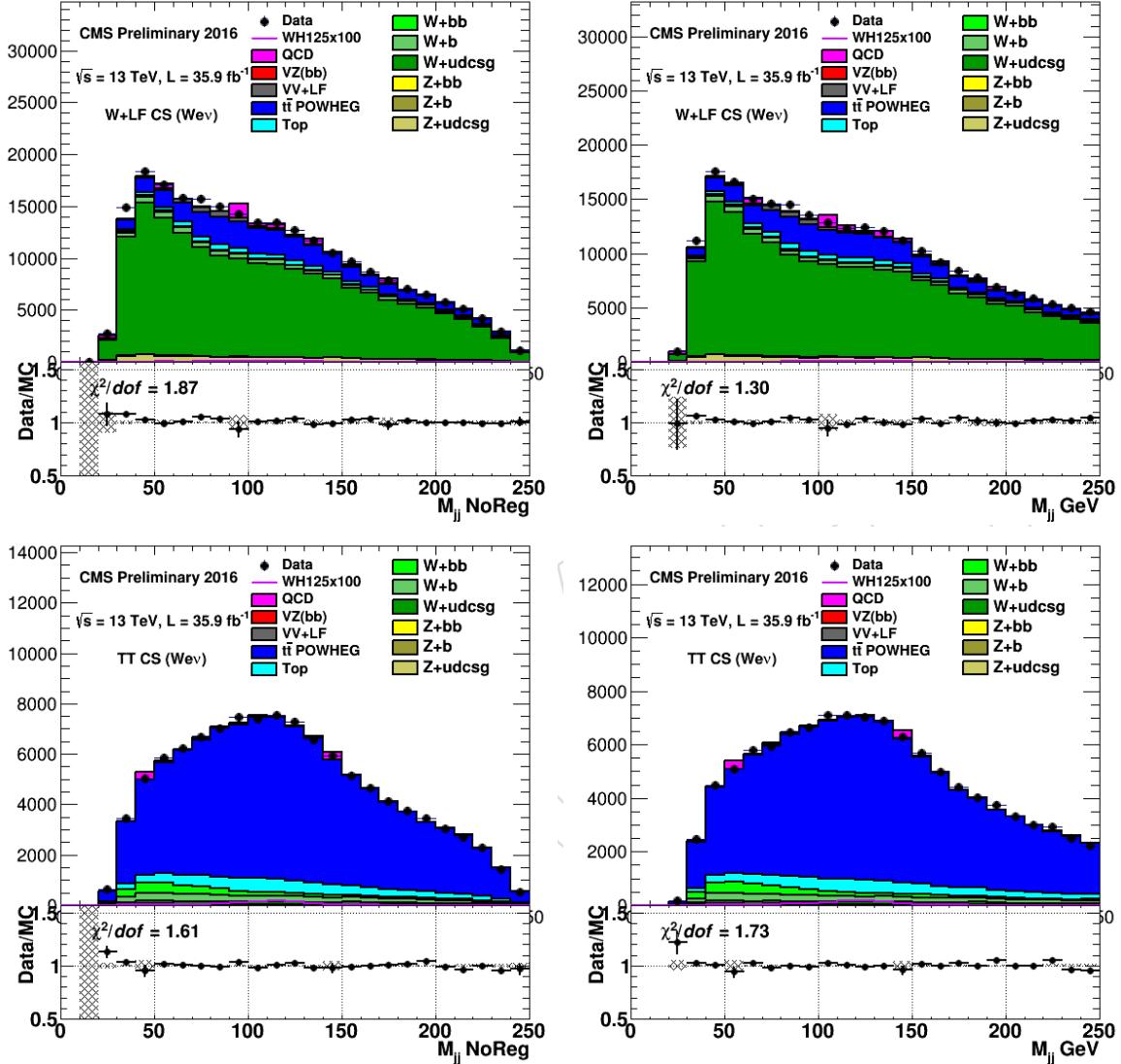


Figure 19: Distributions of dijet invariant mass in background control regions for $W(\ell\nu)H$ (electron channel) enhanced in $W + \text{udscg}$ (top) and $t\bar{t}$ (bottom) before (left) and after (right) the regression is applied.

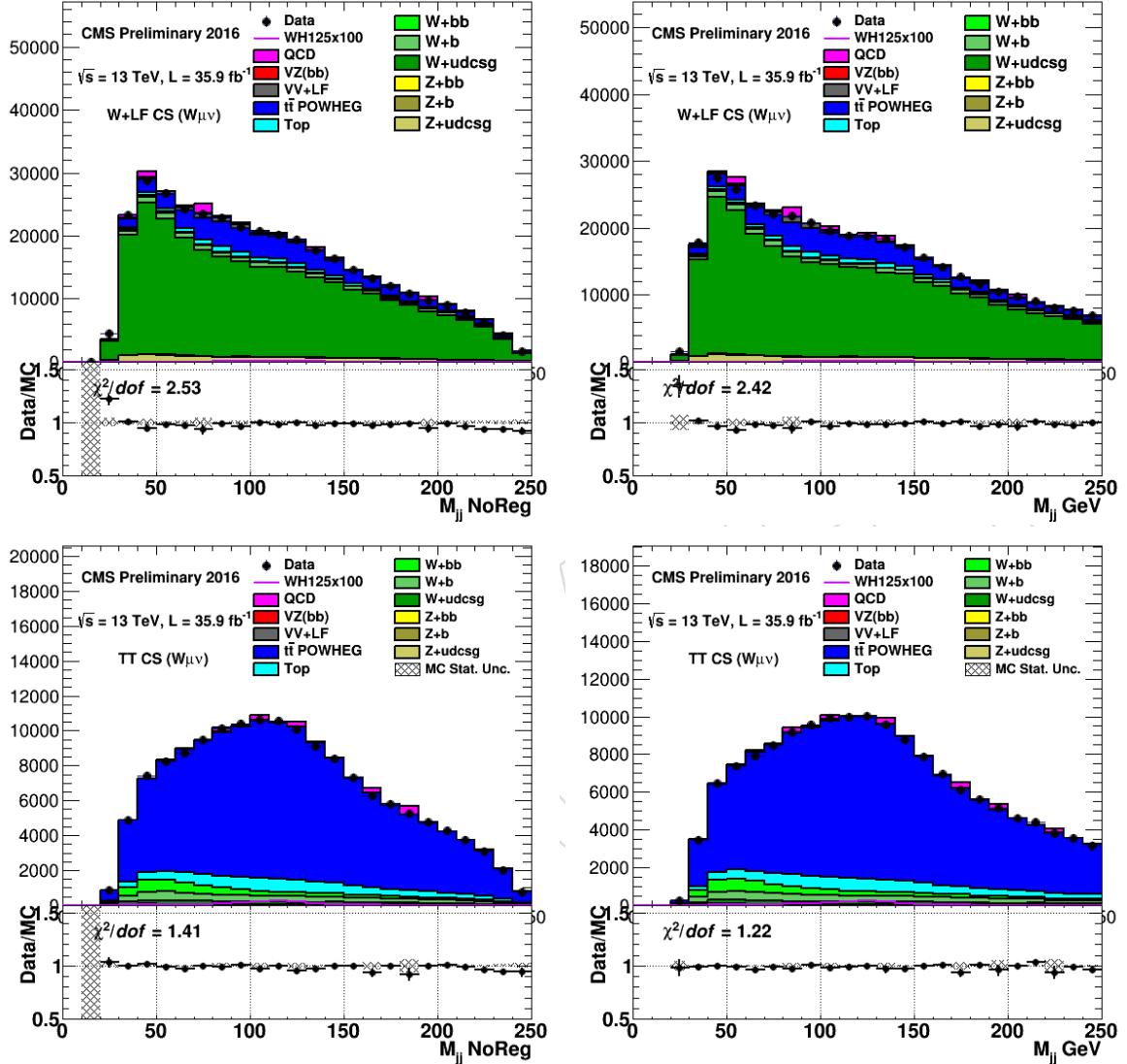


Figure 20: Distributions of dijet invariant mass in background control regions for $W(\ell\nu)H$ (muon channel) enhanced in $W + \text{udscg}$ (top) and $t\bar{t}$ (bottom) before (left) and after (right) the regression is applied.

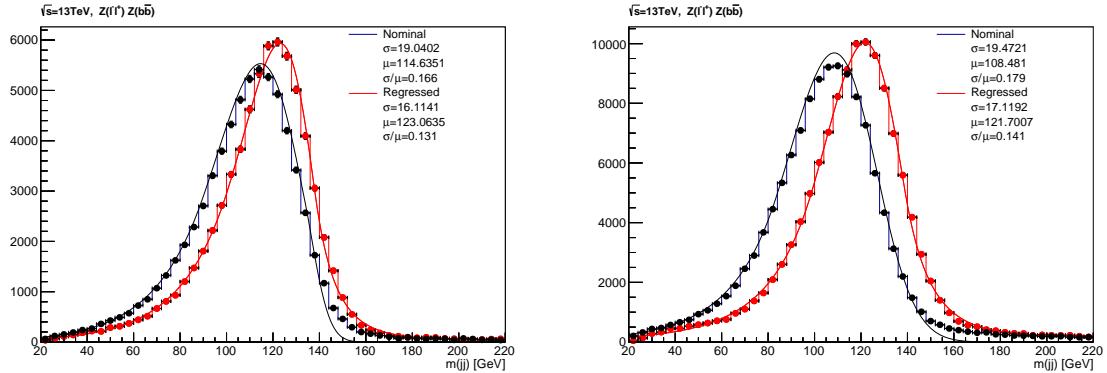


Figure 21: Comparison of the reconstructed dijet invariant mass for Higgs candidates in $Z(\ell\ell)H$ signal events before and after the regression. Separate plots are shown for the case where the b jet does not (left) or does (right) contain a muon from semileptonic B decay.

598 The regression technique is validated in a sample of data events using a typical $Z + b\bar{b}$ selection targeting
 599 the $Z(\ell\ell)H$ final state, combining the dimuon as well as the dielectron channel. Exactly two jets are
 600 required with both jets having a CMVA value of greater than 0.5. Candidate Z decays are constructed
 601 from the two leptons, which are required to have an invariant mass between $75 < M(\ell\ell) < 105$ GeV
 602 which is centered around the Z boson mass. The detailed cuts used are:

603 • Jet selection

- 604 • $|\eta| < 2.4$;
- 605 • $p_T > 20$ GeV;
- 606 • Loose PF Jet ID.

607 • Lepton selection

- 608 • Same as described in Sec. 4.

609 • Event-level selection of $Z + b\bar{b}$

- 610 • $N_\ell = 2$;
- 611 • $N_{jets} = 2$;
- 612 • $|M(\ell\ell) - M_Z| < 15$ GeV;
- 613 • $p_T(\ell\ell) > 50$. GeV;

614 In these events, with a $Z + b\bar{b}$ purity of approximately 70%, we use the distribution of the ratio between
 615 the $p_T(jj)$ and the p_T of the dilepton system

$$p_{T\text{Balance}} = \frac{p_T(jj)}{p_T(\ell\ell)} \quad (9)$$

616 to quantify the improvement coming from the regression. Figure 27 shows the resulting distribution
 617 (red before regression and green after regression), where the mean of the distribution is sensitive to the
 618 scale and the width is sensitive to the dijet pt resolution. A data to simulation comparison is shown in
 619 Fig. 27 before regression (upper left) and after regression (upper right). An improved resolution and
 620 scale is observed and the agreement between data and MC improves with the regression.

621 We extract the mean and RMS from the p_T balance(Fig. 27) and are tabulated in Table 9 for data and in
 622 Table 10 for simulated events. It is visible that the mean of the Gaussian is centered at unity after the
 623 regression.

624 **Top mass in single top events**

625 As further validation, we check the effect of the regression on the top mass in events selected to enhance
 626 single top production. Section 8.3 describes the top reconstruction procedure in the WH analysis. Fig-
 627 ure 28 shows the reconstructed top mass in this sample, comparing data and simulation before and after

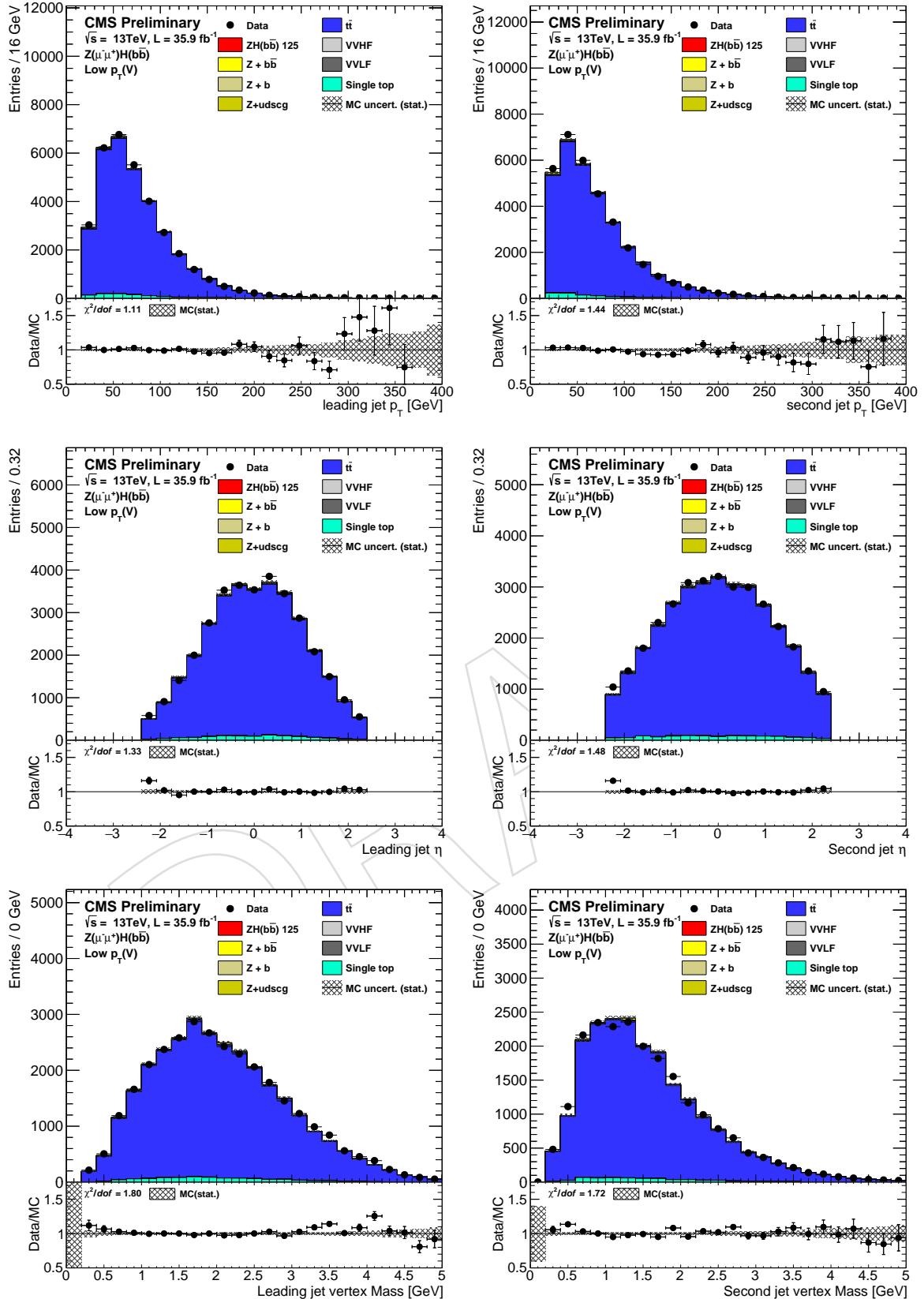


Figure 22: Comparison of input variables to the regression in the $Z(\ell\ell)\bar{t}t$ control region in data (leading jet on the left, sub-leading jet on the right, and variables from top to bottom): p_T , η , vertex mass

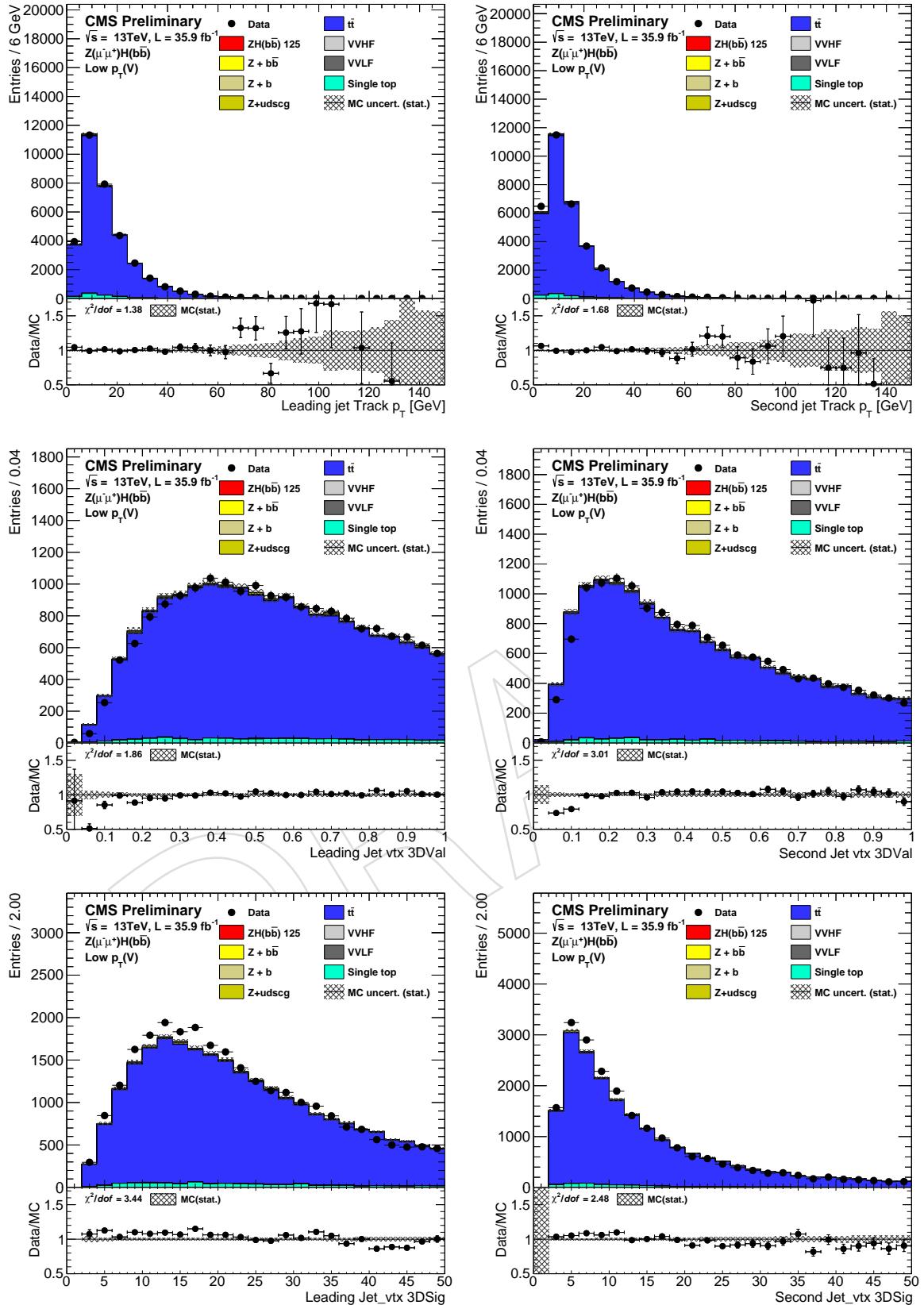


Figure 23: Comparison of input variables to the regression in the $Z(\ell\ell)\bar{t}\bar{t}$ control region in data (leading jet on the left, sub-leading jet on the right, and variables from top to bottom): lead track p_T , secondary vtx 3d decay length, secondary vtx error on 3d decay length.

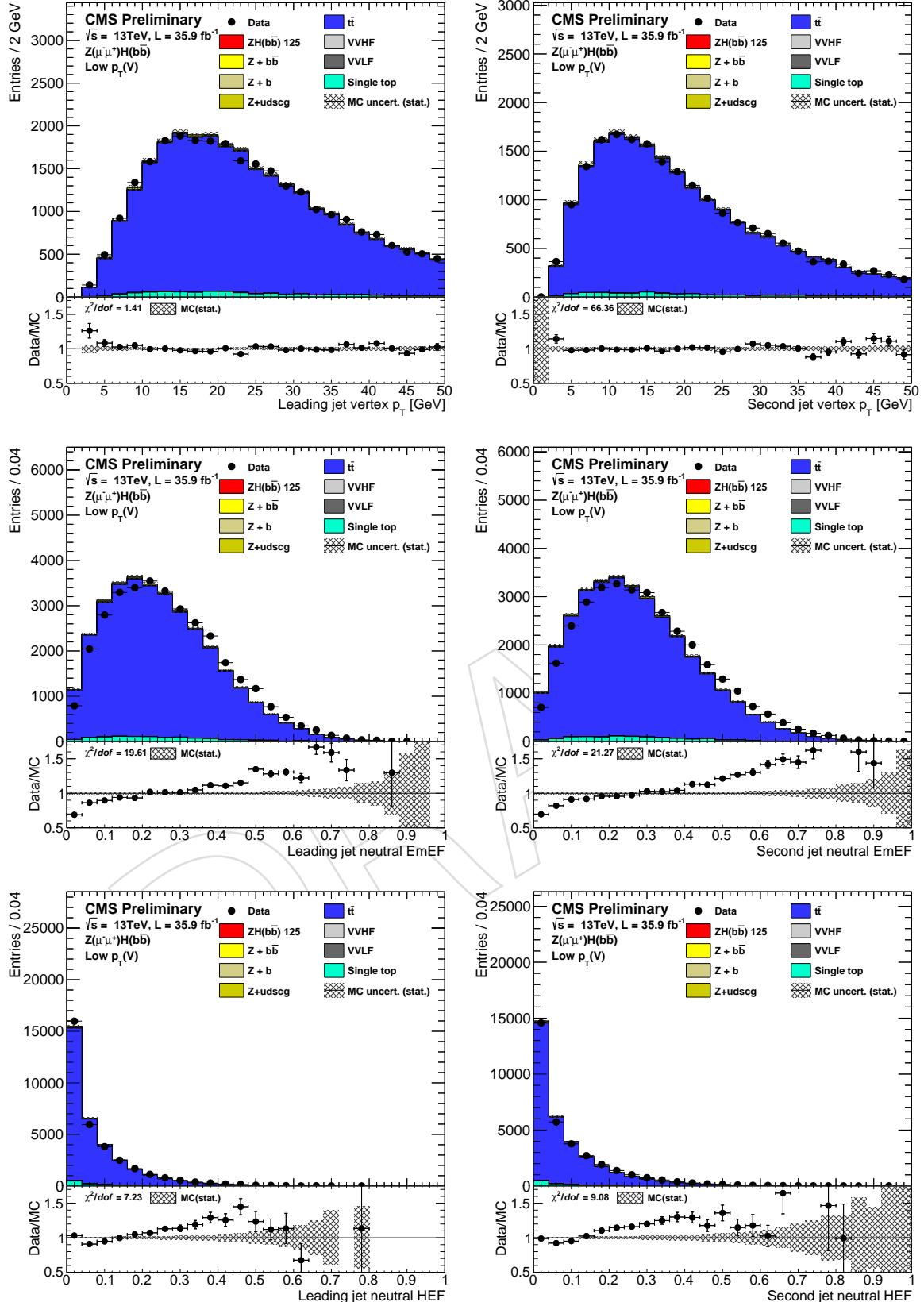


Figure 24: Comparison of input variables to the regression in the $Z(\ell\ell) t\bar{t}$ control region in data (leading jet on the left, sub-leading jet on the right, and variables from top to bottom): secondary vtx p_T , energy fraction of neutral jet constituents(ECAL), energy fraction of neutral jet constituents(HCAL).

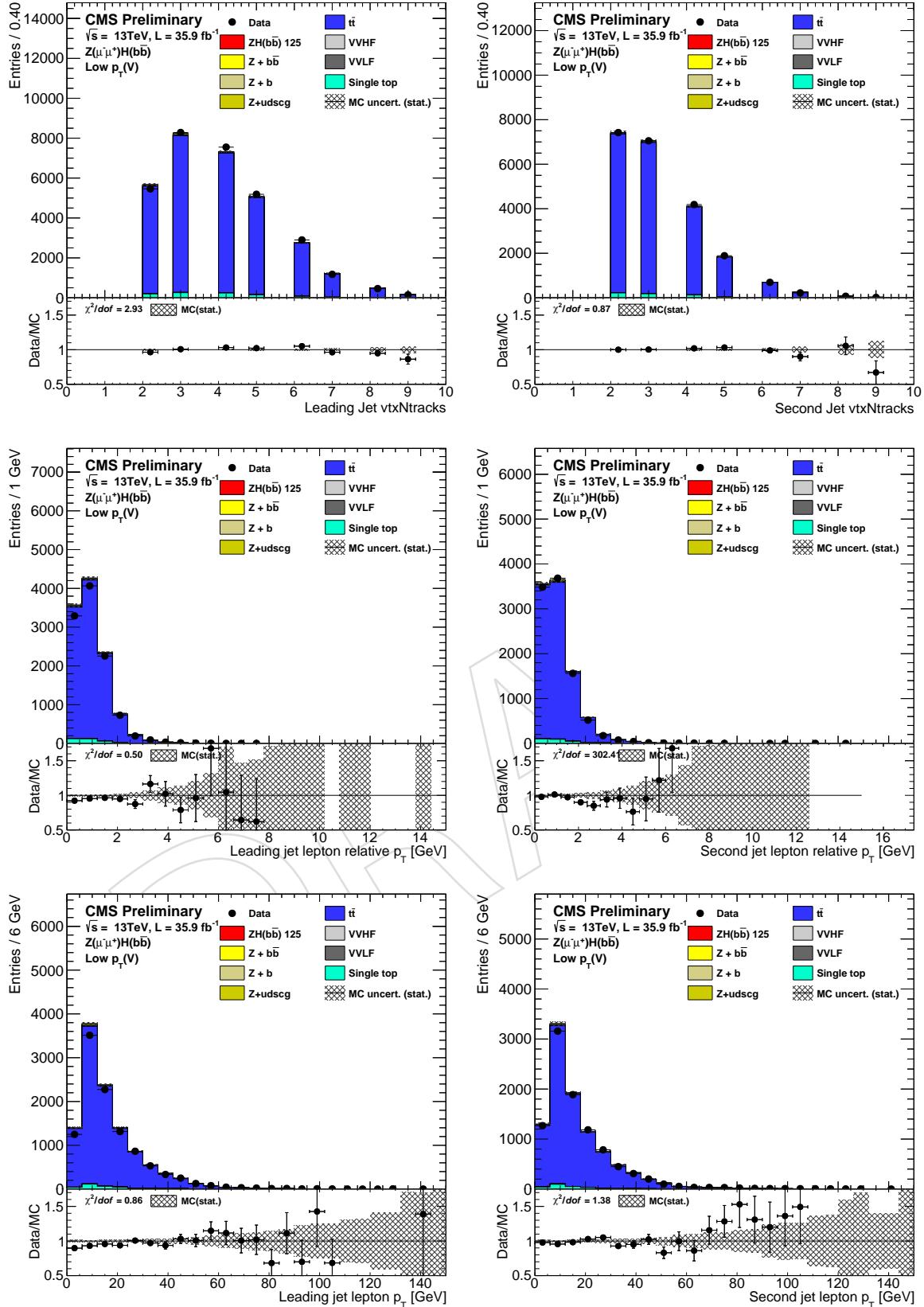


Figure 25: Comparison of input variables to the regression in the $Z(\ell\ell)$ $t\bar{t}$ control region in data (leading jet on the left, sub-leading jet on the right, and variables from top to bottom): number of vertex tracks, relative transverse momentum of soft lepton candidate in the jet, transverse momentum of soft lepton candidate in the jet.

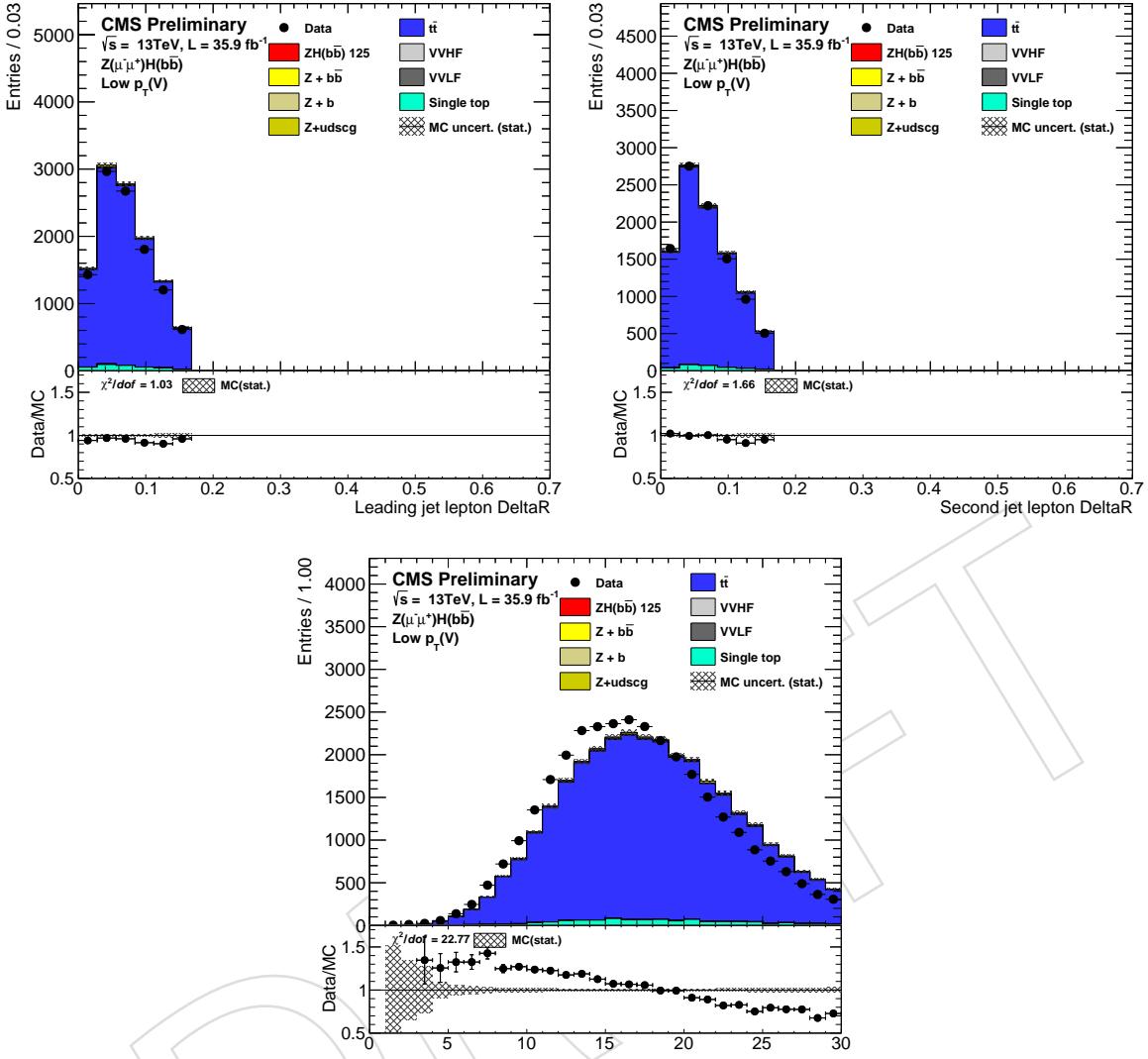


Figure 26: Comparison of input variables to the regression in the $Z(\ell\ell)\bar{t}\bar{t}$ control region in data (leading jet on the left, sub-leading jet on the right, and variables from top to bottom): distance in $\eta - \phi$ space of soft lepton candidate with respect to the jet axis, number of primary vertices.

Table 9: Fit parameters from fits to validate the regression in the $Z(\ell\ell)H$ channel in actual data.

	Data without regression	Data with regression
Mean	0.939 ± 0.007	0.987 ± 0.009
Sigma	0.327 ± 0.008	0.324 ± 0.007

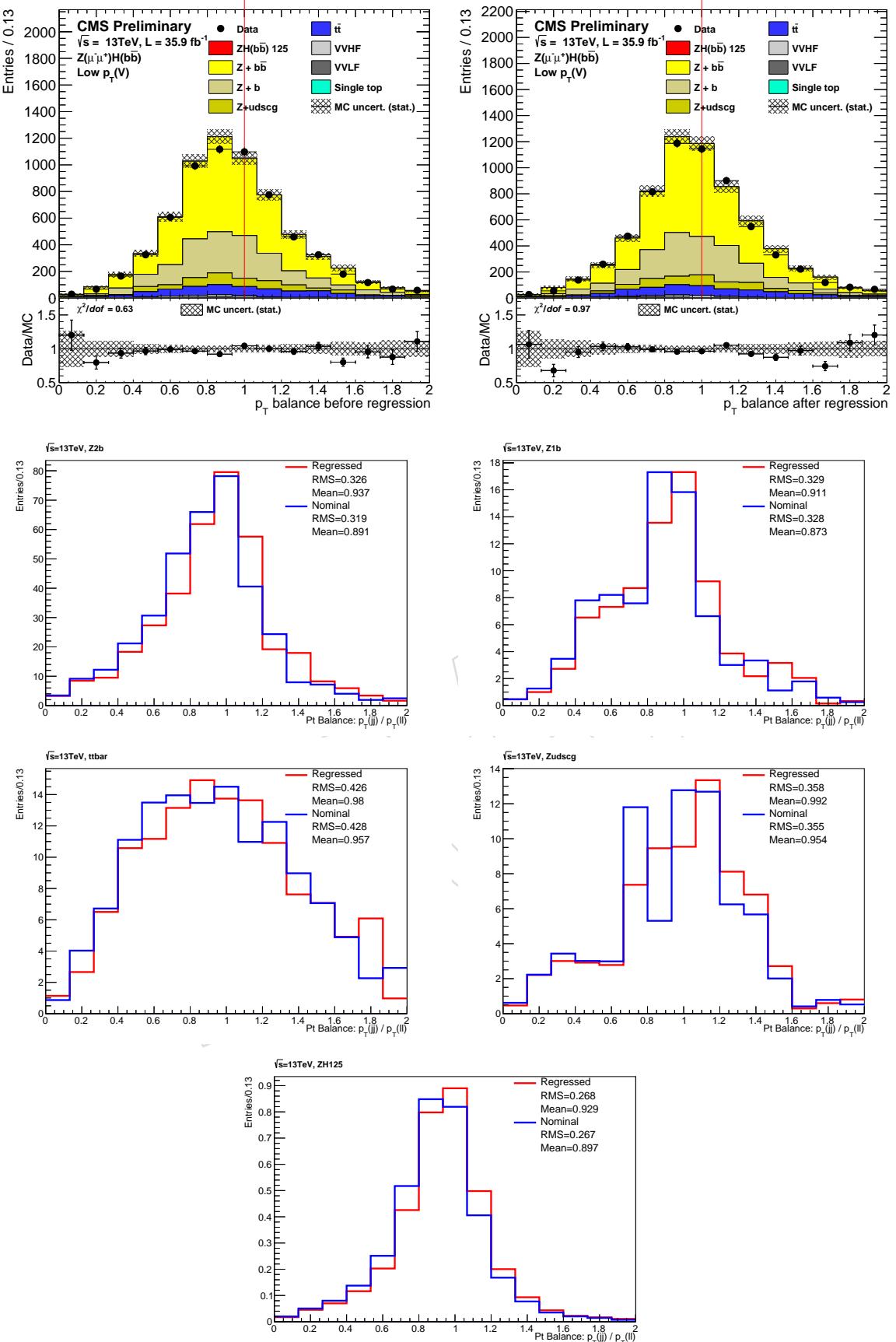


Figure 27: Distribution of the ratio between the $p_T(jj)$ and the p_T of the dilepton system on data versus MC before (top left) and after regression (top right). Comparison of the balance in data (lower middle) before (blue) and after (red) regression. Both mean and resolution of the fitted Gaussian improve with regression.

Table 10: Fit parameters from fits to validate the regression in the $Z(\ell\ell)H$ channel in simulation.

	MC without regression	MC with regression
Mean	0.925 ± 0.005	0.988 ± 0.007
Sigma	0.358 ± 0.006	0.341 ± 0.006

regression, and the individual histograms for the single top and $t\bar{t}$ events overlaying the distributions before/after the regression is applied. A clear improvement in the resolution is observed in the single top and $t\bar{t}$ events both in data and simulation.

7 Residual DATA/MC corrections

7.1 Correcting Data/MC discrepancy in LO Drell-Yan m_{jj} by re-weighting to NLO

We observe that the NLO DY+jets MC describes fairly well the Data, in particular the invariant mass of the di-jet system. However, the NLO samples are relatively low in statistics with an effective integrated luminosity of few fb^{-1} . When propagated to the final result, this statistical uncertainty severely affects the analysis performance.

On the other side, the LO sample is much larger with an effective integrated luminosity higher than data. Therefore, we decided to use the LO sample in the analysis.

The price to pay is a residual data to MC discrepancy showing up in the invariant mass in the LO sample. It is also evident in di-jet $\Delta\eta$, which is shown in Figure 29. Reconstruction level comparisons of NLO and LO $\Delta\eta$ are used to derive event re-weighting. The NLO/LO ratio calculated in the inclusive phase space is shown in Figure 30 for the $Z + 0b$, $Z + 1b$, $Z + 2b$ cases. The ratio is unchanged if generator level quantities are used and if the phase space is varied, even largely. After the reweighting is applied, the invariant mass distributions show better agreement in the LO MC, shown in 31. Other distributions are unaffected, apart from a slight improvement in the jet p_T distribution. No systematic uncertainty is assigned to this reweighting.

7.2 W boson transverse momentum reweighting and LO/NLO check

In the $W(\ell\nu)H$ control regions a downward slope in the data/MC ratio is observed for the reconstructed $p_T(V)$ after the full set of corrections and scale factors have been applied. Independent linear re-weighting functions are derived to correct this slope for $t\bar{t}$, $W + \text{udscg}$, and the combination of $W\bar{b}\bar{b}$ and single top via a simultaneous fit of the reconstructed $p_T(V)$ in the $W(\ell\nu)H$ control regions to data. The input PDF for the fit in each control region is a sum of the MC prediction for each process corrected by a linear function of the reconstructed $p_T(V)$ with a slope that is allowed to float in the fit. The relative composition of the fitted processes in each control region is fixed. Figure 32 shows the result of this simultaneous fit, and Table 11 lists the fitted slopes for each process with the uncertainties from RooFit.

Table 11: Linear correction factors obtained from a simultaneous fit to the $p_T(V)$ distribution in data in the $W(\ell\nu)H$ control regions.

Process	Fitted Slope
$t\bar{t}$	$0.000380 \pm 0.000089 \text{ GeV}^{-1}$
$W + \text{udscg}$	$0.000575 \pm 0.000046 \text{ GeV}^{-1}$
$W\bar{b}\bar{b} + \text{single top}$	$0.00167 \pm 0.00013 \text{ GeV}^{-1}$

The TOP group has observed similar MC mis-modelling of the reconstructed $p_T(V)$ distribution for $t\bar{t}$ events simulated with Powheg [<https://twiki.cern.ch/twiki/bin/view/CMS/TopPtReweighting>]. Their

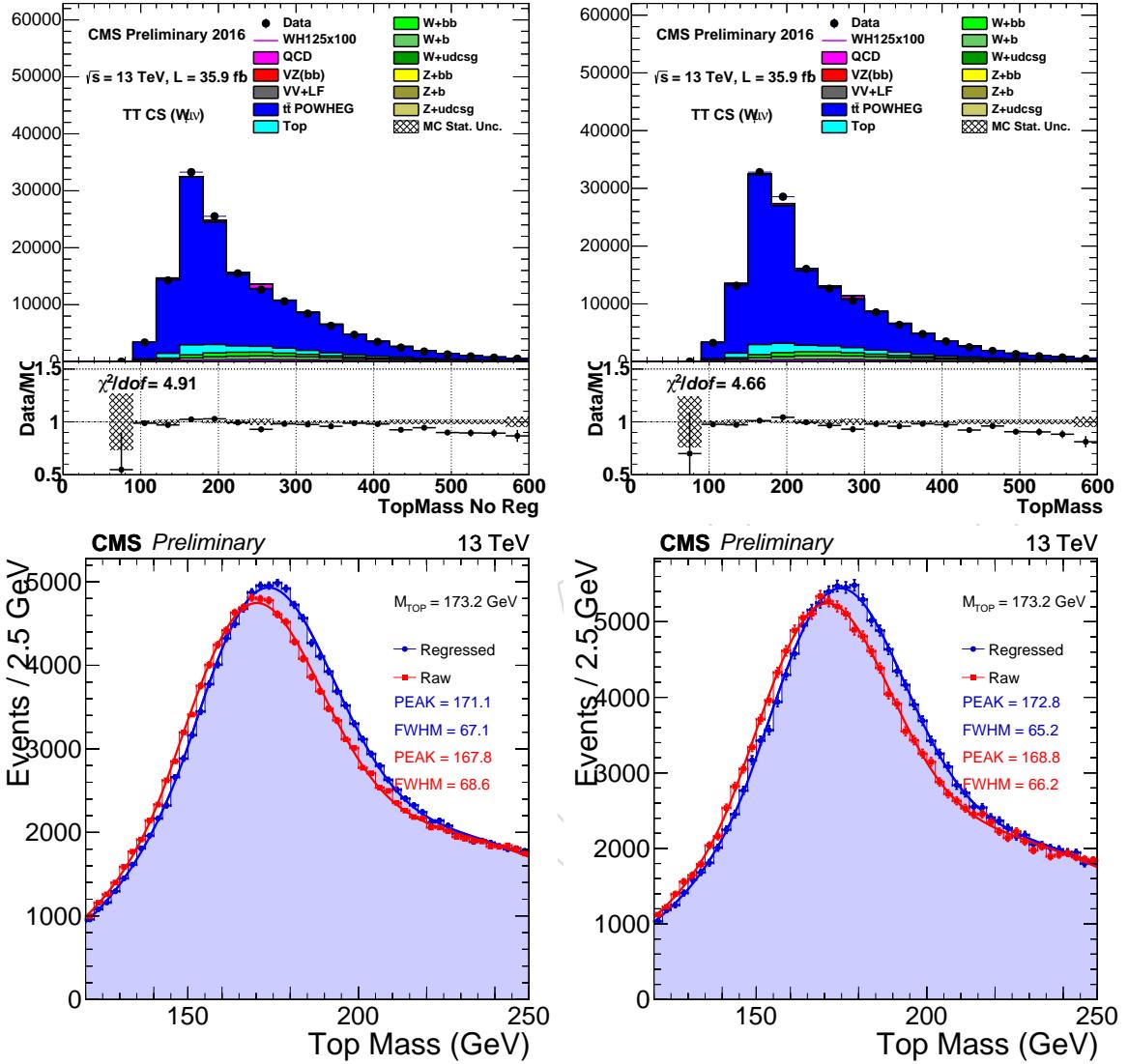


Figure 28: Distributions of the reconstructed top mass in a sample of events selected to enhance single top production. The top two plots show the data/MC agreement before (left) and after (right) the regression is applied. The bottom two plots show the before/after comparison in simulation (left) and actual data (right) events.

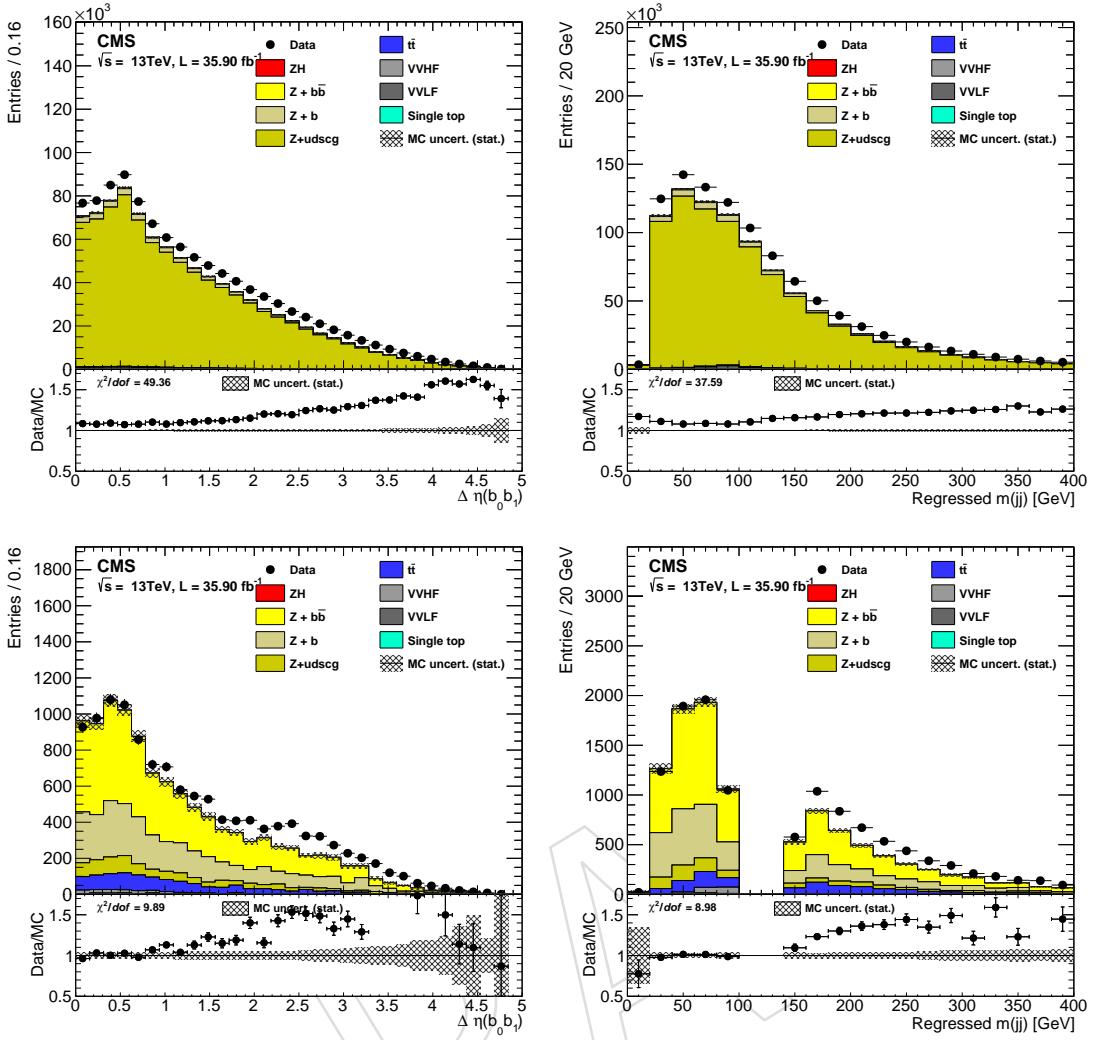


Figure 29: $\Delta\eta$ (left) and invariant mass (right) of di-jet system in the light (top) and heavy (bottom) flavor control regions utilizing LO MC.

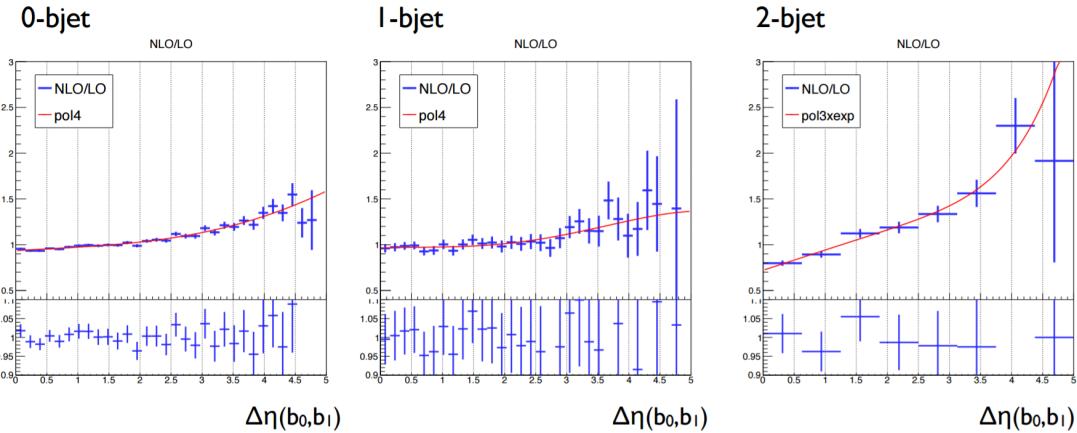


Figure 30: Ratio of the NLO/LO reconstruction level $\Delta\eta$ for DY+Jets MC for the $Z + 0b$ (left), $Z + 1b$ (center), $Z + 2b$ (right) cases.

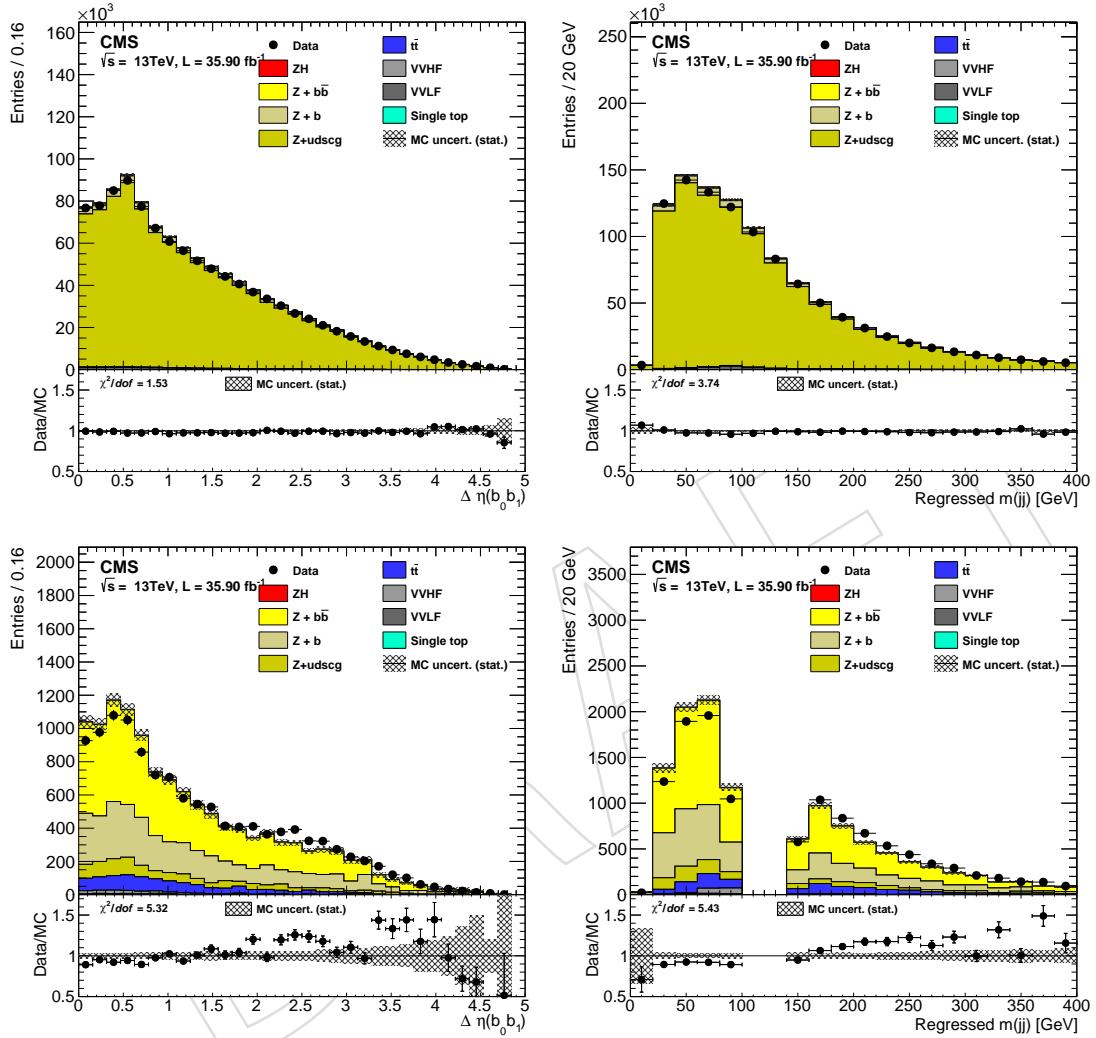


Figure 31: $\Delta\eta$ (left) and invariant mass (right) of di-jet system in the light (top) and heavy (bottom) flavor control regions (above) and the heavy flavor (below) utilizing LO MC after the $\Delta\eta$ reweighting.

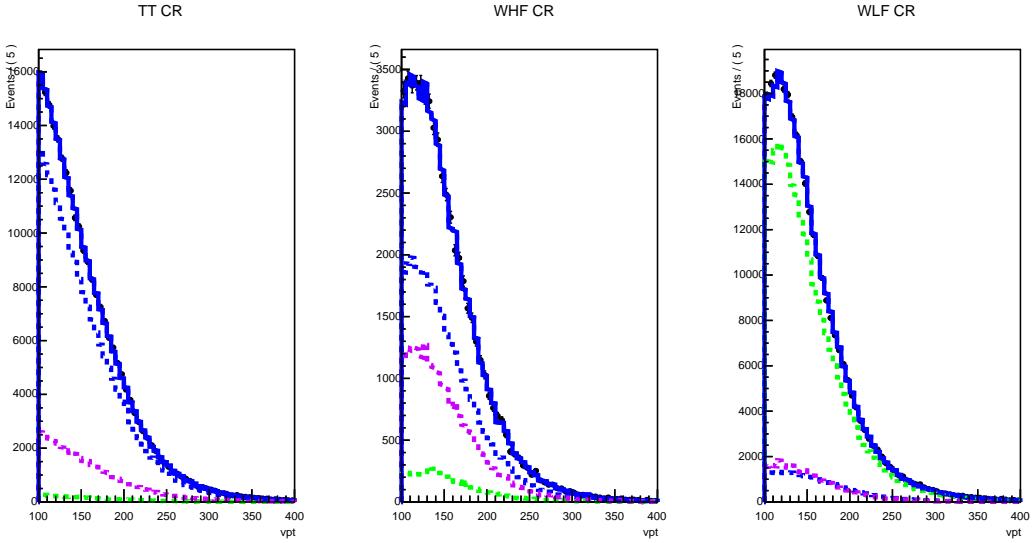


Figure 32: Result of a simultaneous fit of the reconstructed $p_T(V)$ distribution to data in the $W(\ell\nu)H$ control regions. $t\bar{t}$ (blue), $W + udscg$ (green) and the combination of $Wb\bar{b}$ and single top (purple) are treated as independent processes in the fit, each with a linear correction function that is allowed to float.

prescription for correcting this effect is per-event re-weighting as a function of the $p_T(\text{top})$ at generator level. Figure 33 shows a comparison of the TOP group's corrections as an alternative solution/closure test (green) with our $t\bar{t}$ linear correction of $p_T(V)$ derived via the simultaneous fit of the reconstructed $p_T(V)$ to data (black). The systematic uncertainty band for the linear correction is the error on the fitted slope from RooFit. The corrections for $t\bar{t}$ derived via the simultaneous fit are in very good agreement with the re-weighting prescribed by the TOP group (tested as an alternative). In the $Z(\nu\nu)H$ and $Z(\ell\ell)H$ channels the TOP re-weighting is applied to $t\bar{t}$ simulation. It has been verified that the result of simultaneous fit is not sensitive to changes in the definition of the fitted $W+HF$ control region such as loosening the additional jet multiplicity requirement or adjusting the $M(jj)$ selection.

The systematic uncertainties on the $p_T(V)$ corrections are taken from the uncertainties on the fitted slopes given by RooFit. This corresponds to a 23 percent uncertainty on the fitted slope for $t\bar{t}$ and an 8 percent uncertainty for both $W + udscg$ and $W+HF + \text{single top}$. It has been checked that this uncertainty band covers the residual data/MC differences in the $p_T(V)$ distribution after applying the corrections. The overall effect of this systematic on the $W(\ell\nu)H$ analysis sensitivity is less than three percent.

Although there is good agreement in the $M(jj)$ and $p_T(V)$ distributions between data and NLO $W+jets$ MC without any additional corrections needed, the statistical power of the available $W+jets$ NLO samples is significantly less than the available LO $W+jets$ MC. This difference in statistical power translates to at least a 10 percent reduction in expected sensitivity when using the NLO $W+jets$ samples as opposed to the LO $W+jets$ samples. Furthermore, issues with the stitching of the available NLO $W+jets$ samples have been observed which lead to mismodelling of the reconstructed $p_T(V)$ distribution for $p_T(V) < 150$ GeV. As a validation of the corrections applied to the LO $W+jets$, the prediction for each of the inputs to the BDT as well as the BDT score is compared for LO and NLO $W+jets$ in the $W(\ell\nu)H$ signal region with $p_T(V) > 150$ GeV. Figure 34 shows the comparison of the BDT score. Good agreement is observed for the BDT score as well as each of the individual BDT inputs.

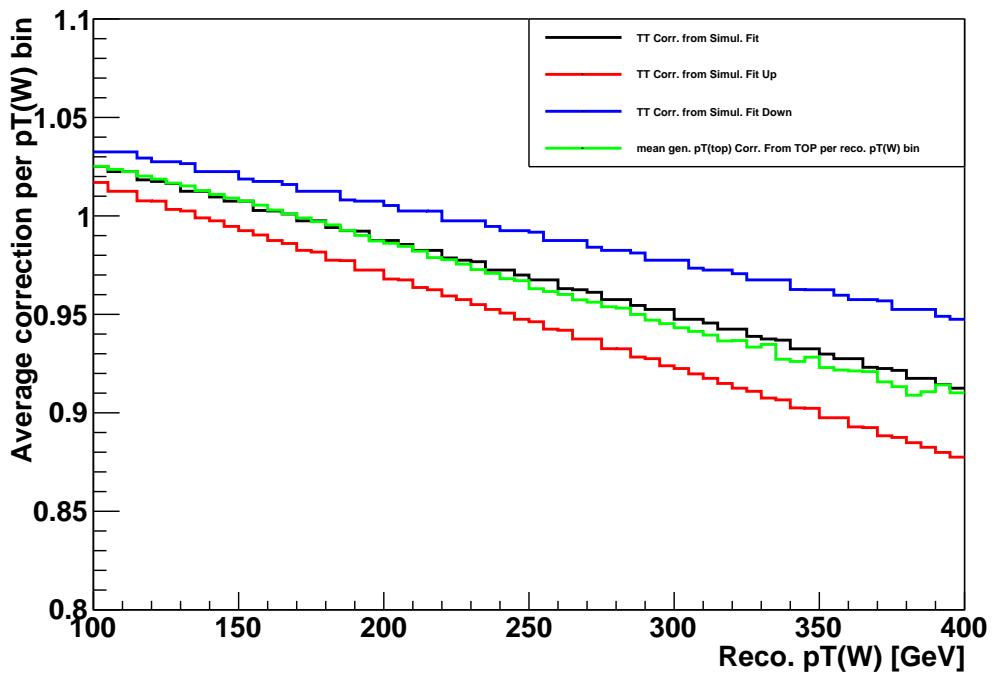


Figure 33: A comparison of the generator-level $p_T(\text{top})$ re-weighting prescribed by TOP (green) with the TT linear corrections obtained via the simultaneous fit to the $p_T(V)$ data distribution (black). The uncertainty band is taken from the uncertainty on the fitted slope given by RooFit. The two independently derived corrections are in very good agreement.

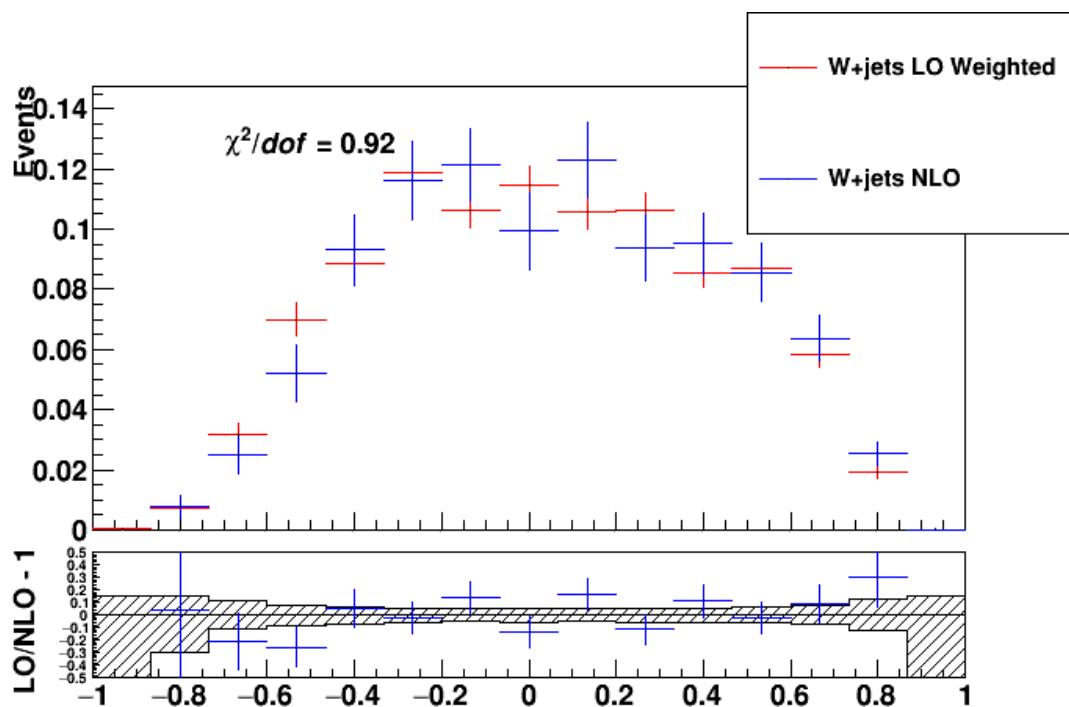


Figure 34: A comparison of the re-weighted LO (red) with the NLO (blue) W+jets MC for the BDT score in the $W(\ell\nu)H$ signal region with the additional requirement that the reconstructed $p_T(V)$ be greater than 150 GeV. Reasonable agreement is observed between the weighted LO and the NLO W+jets predictions.

683 8 Event Selection

684 The building blocks of the analysis described in the following section are discriminating variables whose
 685 distributions differ between signal and background. In this section we broadly describe characteristics
 686 of signal and background and the consequential reconstructed variables that allow us to separate the
 687 two.

688 Regardless of the final analysis' selection and statistical analysis, a preselection is applied at the begin-
 689 ning of the analysis. This preselection varies per channel and is based primarily on the characteristics of
 690 the vector boson's decay as well as limitation from the triggers, which also vary per channel.

691 8.1 Signal and background characteristics

692 Signal events are characterized by the presence of a vector boson recoiling from two b jets with an
 693 expected invariant mass in the range $110 < M_H < 150$ GeV, based on current experimental data and
 694 theory input. The transverse momentum of the b jets peaks at roughly $M_H/2$, and the dijet p_T spectrum
 695 is harder than the V+jets background spectrum. Isolated leptons not arising from the decay of a W or
 696 Z boson are expected to be negligible, and additional jet activity is reduced. The vector boson and dijet
 697 objects are expected to be central and back-to-back in the transverse plane, so that the azimuthal opening
 698 angle between them is peaked at π .

699 The dominant backgrounds encountered in this analysis arise from three general sources, each of which
 700 carries with it a distinctly different pattern in one or more of the variables characterizing signal events.

- 701 • **V+jets:** production of W and Z bosons in association with one or more jets. This background
 702 looks very much like signal topologically, but has a generally softer p_T spectrum, a sharply
 703 falling dijet mass distribution, and decay characteristics (effective spin and color radiation)
 704 that, in principle, differ significantly from the signal. The contributions from udscg jets domi-
 705 nates is much reduced after the application of tagging on both Higgs daughters. In the highest
 706 purity phase space of the analysis contributions from V + bb̄ can be nearly degenerate with
 707 signal.
- 708 • **Top quarks:** production of tt̄ pairs, as well as single top quarks in the tW, t-channel, and
 709 s-channel processes, represent a particularly challenging background at the Run2 13 TeV en-
 710 ergy, where the production cross section is over three times larger than at Run1 7-8 TeV. These
 711 backgrounds include one or two real W decays, at least two b jets (for tt̄), with intrinsic mass
 712 and momentum scales close to the Higgs mass scale. The primary handles to reduce the tt̄
 713 background are topological: the additional jet multiplicity extends well beyond two (typically
 714 from the hadronic decay of the second W), the azimuthal opening angle between the vector
 715 boson and dijet, which is more broadly distributed than in signal events, and finally (in the
 716 W(ℓν)H(bb̄) channels) the reconstructed top mass can be used to eliminate events with mass
 717 similar to the top. Single top events are more difficult to reject relative to signal, but the cross
 718 section is such that it typically represents only 10-20% of the total background in WH and
 719 even less in the other channels.
- 720 • **Dibosons (WW, WZ, ZZ):** production of vector-boson pairs is another important background,
 721 since they can produce a resonant dijet system having a mass within a few standard devia-
 722 tions of the signal. The dominant contribution arises when one boson decays leptonically
 723 (including $Z \rightarrow \nu\bar{\nu}$), and the second boson decays to jets. The ZZ and WZ modes can lead
 724 directly to a V + bb̄ combination with two real b jets coming from the hadronic $Z \rightarrow bb̄$ de-
 725 cay. This background is virtually indistinguishable from signal events, apart from the peak
 726 position of the dijet mass. Good mass resolution is the key handle to separate signal from this
 727 background.

728 8.1.1 Multijet QCD

729 In the Run 2 Z(νν)H analysis, we reported also a larger multijet QCD background compared to Run 1.
 730 In multijet events, MET is mainly produced by the mis-measurement of the energy of a jet (fake MET) or
 731 by the emission of a high- p_T neutrino originating from a hadron decay, as shown in Fig. 35. The fraction

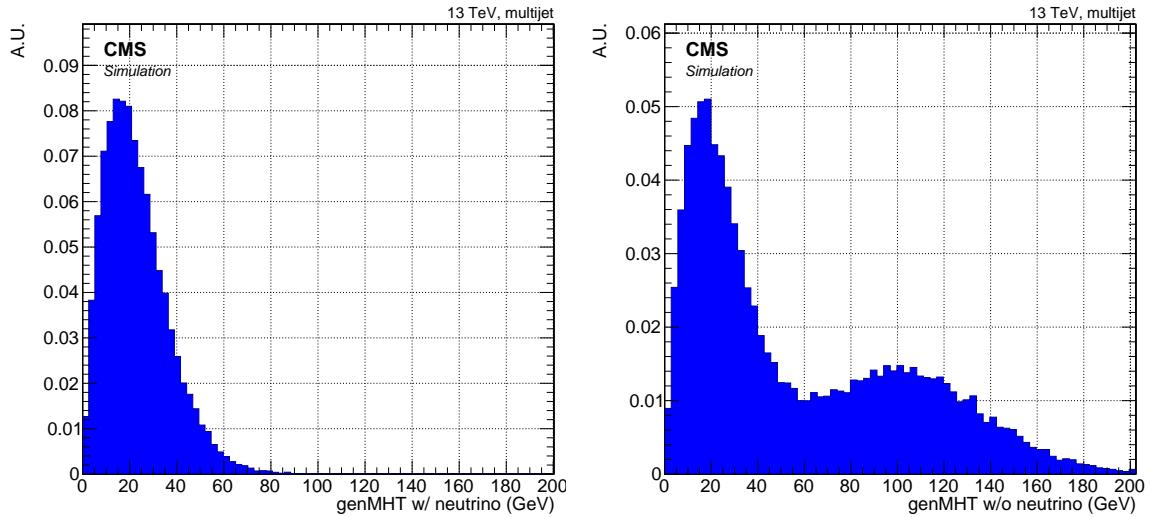


Figure 35: MET evaluated using generator level information in multijet events with simulated jets with $p_T > 20 \text{ GeV}$, including (left) and excluding (right) the neutrinos originating from the hadron decays in the computation of the MET. The plot has been produced using a multijet simulation with the sum of the generated transverse parton momenta within 500 GeV and 700 GeV and selecting the events with reconstructed MET $> 120 \text{ GeV}$. The plots show that there exist some portion of the multijet events with large, real MET due to the emission of a high- p_T neutrino. These events are found around $\sim 100 \text{ GeV}$ on the right plot. The remaining events have large MET but it is produced by the mis-measurement of the transverse momentum of a jet.

732 of events with high- p_T neutrinos is predicted with large uncertainty since it depends on production
 733 rates of heavy hadrons and branching ratios that are known with limited precision.

734 Fake MET can be produced by the mis-measurement of the transverse momentum of a jet.

735 In a multijets simulation with reconstructed MET $> 120 \text{ GeV}$ $\sim 50\%$ of events contains a neutrino with
 736 $p_T > 50 \text{ GeV}$ and $\sim 40\%$ contain an underestimated jet pointing towards to a dead ECAL cell. The two
 737 sets of events are roughly independent and the remaining $\sim 10\%$ of multijet events with large MET are
 738 due to other mismeasured jets.

739 The main feature of the multijet events is the presence of at least one jet close to the MET direction on
 740 the transverse plane. For this reason, the main variable used to reject the multijet events with fake MET
 741 is the minimum $\Delta\varphi(\text{jet}, \text{MET})$. Figure 36 shows the distribution of $\min(\Delta\varphi(\text{jet}, \text{MET}))$, in multijets and
 742 signal MC events (considering only the two jets with the highest p_T).

743 8.2 Discriminating variables

744 The following variables have been considered to be potentially useful in separating signal from back-
 745 ground:

- 746 • $M(\text{jj})$: dijet invariant mass; it peaks at M_H for signal and M_Z for diboson events, falls sharply
 747 for V+jets, and peaks broadly over the region 100–160 GeV for $t\bar{t}$ events.
- 748 • $p_T(\text{jj})$: transverse momentum of the Higgs candidate.
- 749 • p_{T_j} : transverse momentum of the Higgs candidate daughters.
- 750 • $p_T(V)$: vector boson transverse momentum, as defined in 5, which is highly correlated with
 751 $p_T(\text{jj})$ for signal and most backgrounds.
- 752 • CMVA: continuous output of the CMVA discriminant, optimized separately for the jet with
 753 the higher value (CMVA1), and the one with the lower value (CMVA2).
- 754 • M_t : the top mass reconstructed in events with a leptonic decaying W and one of the b-jets.

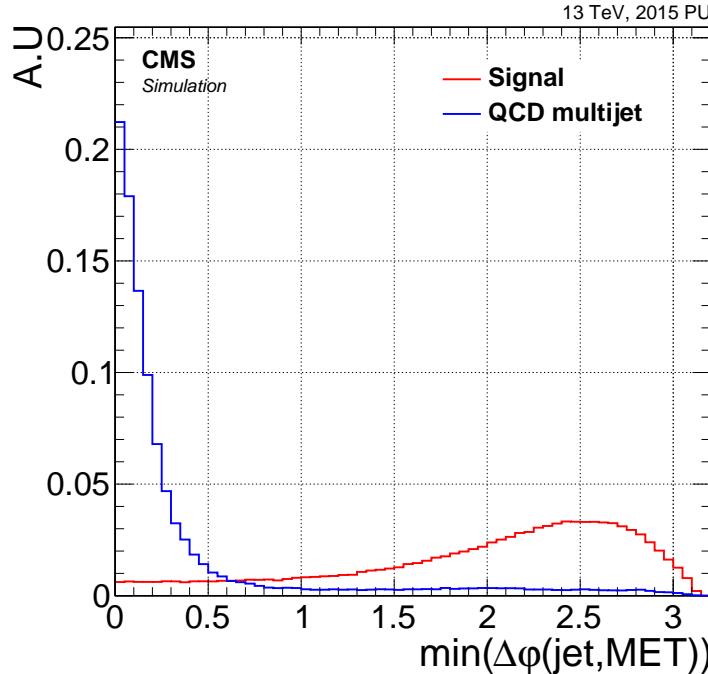


Figure 36: $\min(\Delta\phi(\text{jet}, \text{MET}))$ considering the two leading in p_T jets only for signal (red) and multijet (blue) simulation with $500 \text{ GeV} < \text{genHT} < 700 \text{ GeV}$. Events are required to have $\text{MET} > 120 \text{ GeV}$.

- $\Delta\phi(V, H)$: azimuthal opening angle between the momenta of the vector boson and the Higgs candidate.
- $\Delta\eta(jj)$: distance in pseudorapidity between the two jets comprising the Higgs candidate.
- $\Delta R(jj)$: distance in η - ϕ space between the two jets comprising the Higgs candidate.
- N_{aj} : number of additional jets in the event apart from the Higgs candidate. Only central jets with $|\eta| < 2.5$ are considered, but the p_T threshold and number of additional jets to allow are parameters to be optimized separately for each channel. In practice, the optimal threshold was found to be $p_T > 20 \text{ GeV}$ in all channels where a jet veto is applied.
- N_{al} : number of additional isolated leptons (as defined in Sec. 4) apart from those associated with the W or Z decay. Only leptons satisfying $p_T > 20 \text{ GeV}$ and $|\eta| < 2.5$ are considered in the count.
- pfMET : E_T^{miss} calculated with particle-flow objects. of the missing transverse energy.
- $\Delta\phi(\text{pfMET}, J)$: azimuthal opening angle between the pfMET vector direction and the transverse momentum of the closest central jet in azimuth. Only jets satisfying $p_T > 30 \text{ GeV}$ and $|\eta| < 2.5$ are considered. This variable helps in reducing residual QCD background in the $Z(\nu\nu)H$ channel, where the source of the missing transverse energy is typically from fluctuations in the measured energy of a single jet.
- $\Delta\phi(\text{pfMET}, \text{lept.})$: azimuthal opening angle between the pfMET vector direction and the leading lepton direction. This variable helps in reducing events of fully leptonic decay of $t\bar{t}$ in $Z(\nu\nu)H$ analysis.
- $\text{maxCMVA}_{\text{aj}}$: maximum CMVA of the additional jets in an event. This variable helps in reducing ttbar background in the $Z(\nu\nu)H$ and $W(\ell\nu)H$ channels.
- $\text{min}\Delta R(H, aj)$: minimum distance between an additional jet and the Higgs candidate. This variable helps in reducing ttbar background in the $Z(\nu\nu)H$ and $W(\ell\nu)H$ channels.
- $\text{soft - activity } N_5^{\text{soft}}$: number of additional soft track-jets with $p_T > 5 \text{ GeV}$ as defined in Sec. 4.9.

- 781 • Angular variables: Various kinematic and angular variables: HV system mass, Angle Z-Z*,
 782 Angle Z-l, Angle H-jet (only for $Z(\ell\ell)H$). These variables help to discriminate against ZZ and
 783 $Z + b\bar{b}$ backgrounds.

784 **8.2.1 Anti-QCD Cut**

785 In order to reject the multijet background, a cut of $\Delta\varphi(\text{jet}, \text{MET}) > 0.5$ is applied to all jets with $p_T >$
 786 30 GeV in the signal region. The cut rejects the $\sim 93\%$ of multijet events having $500 \text{ GeV} < \text{genHT} <$
 787 700 GeV , with a signal efficiency in the signal region of $\sim 96\%$. This cut is also used to increase the
 788 purity of the $Z + \text{light-jets}$ and $Z + b\text{-jets}$ in their respective control regions.

789 **8.3 Reconstructing the top mass in events with a lepton and MET**

790 In $W(\ell\nu)H(b\bar{b})$ channels the background from $t\bar{t}$ production is quite dominant except after the tightest
 791 selection where $Wb\bar{b}$ is nearly as significant. Furthermore, the $t\bar{t}$ background grows faster with \sqrt{s} than
 792 Higgs boson production, and so the 13 TeV $t\bar{t}$ cross section is 3.4 times larger, whereas Higgs boson
 793 production only increases a factor of 2. In the WH channels several variables were analyzed to help
 794 further discriminate against this background. In the end, the reconstructed top mass was included as an
 795 analysis variable because of its discriminating power.

796 $W(\ell\nu)H(b\bar{b})$ events are characterized by the presence of a well isolated lepton, MET and two b-jets. This
 797 is the same signature for semi-leptonic $t\bar{t}$ events. Furthermore, the lepton and the MET should both arise
 798 from the decay of the W . Using the constraint of the W mass, one can solve a quadratic equation with
 799 lepton p_T W mass, known p_T of the neutrino (assumed to be equal to MET) and unknown longitudinal
 800 neutrino momentum.

$$M_W^2 = (E_\nu + E_\ell)^2 - (\vec{p}_\nu + \vec{p}_\ell)^2 \quad (10)$$

801 There are always two solutions to this equation. When both are real, the solution with the smaller
 802 longitudinal neutrino momentum is selected. When the solutions are imaginary, the real part is taken as
 803 the longitudinal neutrino momentum.

804 Once this is done, energy-momentum 4-vectors of the neutrino, the lepton and the closest b-jet are added
 805 and that composition 4-vector is assumed to be a top quark. The mass is taken from this 4-vector.

806 **8.4 Analysis preselection**

807 In each channel it is necessary to construct a loose, basic selection in order to obtain the topology of the
 808 signal for optimization. In addition, there may be cuts applied to avoid problematic regions where, for
 809 instance, the trigger is inefficient or not fully efficient for signal. Table 12 lists the preselections for each
 810 channel.

Table 12: Preselection cuts for each channel to define the signal region.

Variable	$W(\ell\nu)H$	$Z(\ell\ell)H$	$Z(\nu\nu)H$
$p_T(V)$	> 100	$[50 - 150], > 150$	> 170
$m_{\ell\ell}$	–	$[75 - 105]$	–
p_T^ℓ	$(> 25, > 30)$	> 20	–
$p_T(j_1)$	> 25	> 20	> 60
$p_T(j_2)$	> 25	> 20	> 35
$p_T(jj)$	> 100	–	> 120
$M(jj)$	$[90 - 150]$	$[90 - 150]$	$[60 - 160]$
$CMVA_{\max}$	CMVAT	CMVAL	CMVAT
$CMVA_{\min}$	CMVAL	CMVAL	CMVAL
N_{aj}	< 2	–	–
N_{al}	$= 0$	–	$= 0$
E_T^{miss}	–	–	> 170
Anti-QCD	–	–	Yes
$\Delta\phi(V, H)$	> 2.5	> 2.5	> 2.0
$\Delta\phi(\text{pfMET, trkMET})$	–	–	< 0.5
$\Delta\phi(\text{pfMET, lep})$	< 2.0	–	–
Tightened Lepton Iso.	$(0.06, 0.06)$	–	–

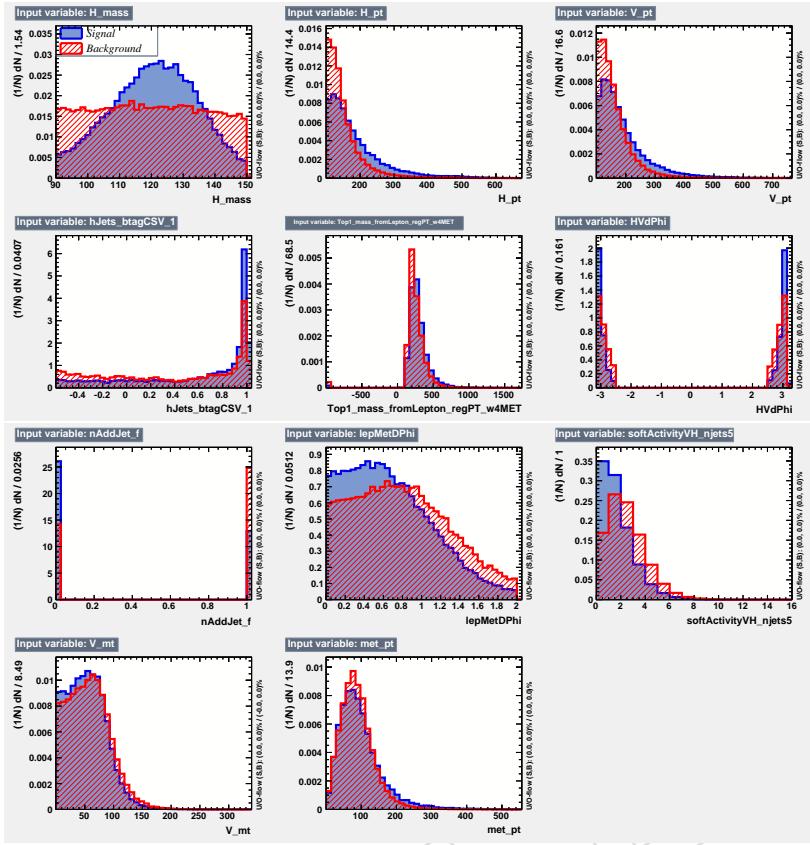


Figure 37: Distributions of signal (blue) and all backgrounds (red) for the variables trained in the BDT for $W(\ell\nu)H$.

811 9 Binned BDT Shape Analysis

812 Multivariate techniques are used to further enhance the statistical power of the analysis by making full
 813 use of correlations between discriminating variables in signal and background events. The Boosted
 814 Decision Tree (BDT) method is implemented in the TMVA framework [67].

815 Separate BDTs have been trained in each channel, after applying the b-jet energy regression described in
 816 Sec. 6, and electron modes are trained together with muon modes in $W(\ell\nu)H$ and $Z(\ell\ell)H$. Table 13 lists
 817 the discriminating variables used in the training of the BDT analyzers. Only events passing the relatively
 818 loose preselection, listed in Table 12, is used for training. All MC events have “event numbers”. The
 819 even events are used for training the BDTs, while the odds events are saved for the evaluation of the
 820 BDT performance.

821 Figures 37, 39, 40, and 42 show the separation between signal (blue) and all backgrounds (red) for the
 822 input variables for events used in each channels’ training (events passing preselection). Figures 38, 41,
 823 and 43 shows a comparison of correlation plots for the input variables in data and simulation in the $t\bar{t}$
 824 control region defined in Section 11.3.

825 To describe the performance of a BDT we observe the signal and background composition as a function
 826 of ever-tightening cuts on the BDT distribution. Essentially this is a sliding integral of the BDT distribu-
 827 tions in data and MC shown in Figure 44. Figure 45 shows the expected signal to background ratio in
 828 bins of BDT score for each channel.

Table 13: Variables used in the BDT training.

Variable	Channels utilizing
$M(jj)$: dijet invariant mass	All
$p_T(jj)$: dijet transverse momentum	All
$p_T(V)$: vector boson transverse momentum	All
$CMVA_{\max}$: value of CMVA for the Higgs daughter with largest CMVA value	$Z(\ell\ell)H, Z(\nu\nu)H$
$CMVA_{\min}$: value of CMVA for the Higgs daughter with second largest CMVA value	All
$CMVA_{\text{add}}$: value of CMVA for the additional jet with largest CMVA value	$Z(\nu\nu)H$
$\Delta\phi(V, H)$: azimuthal angle between V and dijet	All
$p_T(j)$: transverse momentum of each Higgs daughter	$Z(\ell\ell)H, Z(\nu\nu)H$
$p_T(\text{add.})$: transverse momentum of leading additional jet	$Z(\nu\nu)H$
$\Delta\eta(jj)$: difference in η between Higgs daughters	$Z(\ell\ell)H, Z(\nu\nu)H$
$\Delta R(jj)$: distance in η - ϕ between Higgs daughters	$Z(\ell\ell)H$
N_{aj} : number of additional jets N.B. definition slightly different per channel	$W(\ell\nu)H, Z(\ell\ell)H$
$p_T(jj) / p_T(V)$: p_T balance between Higgs candidate and vector boson	$Z(\ell\ell)H$
M_Z : Z boson mass	$Z(\ell\ell)H$
SA5: number of soft activity jets with $p_T > 5 \text{ GeV}$	All
M_t : reconstructed top mass	$W(\ell\nu)H$
$\Delta\varphi(\text{pfMET, lept.})$: azimuthal angle between E_T^{miss} and lepton	$W(\ell\nu)H$
E_T^{miss} : missing transverse energy	$W(\ell\nu)H, Z(\ell\ell)H$
$m_T(W)$: W transverse mass	$W(\ell\nu)H$
$\Delta\varphi(jj)$: difference in φ between Higgs daughters	$Z(\nu\nu)H$
$\Delta\varphi(\text{pfMET, jet.})$: azimuthal angle between E_T^{miss} and the closest jet with $p_T > 30 \text{ GeV}$	$Z(\nu\nu)H$

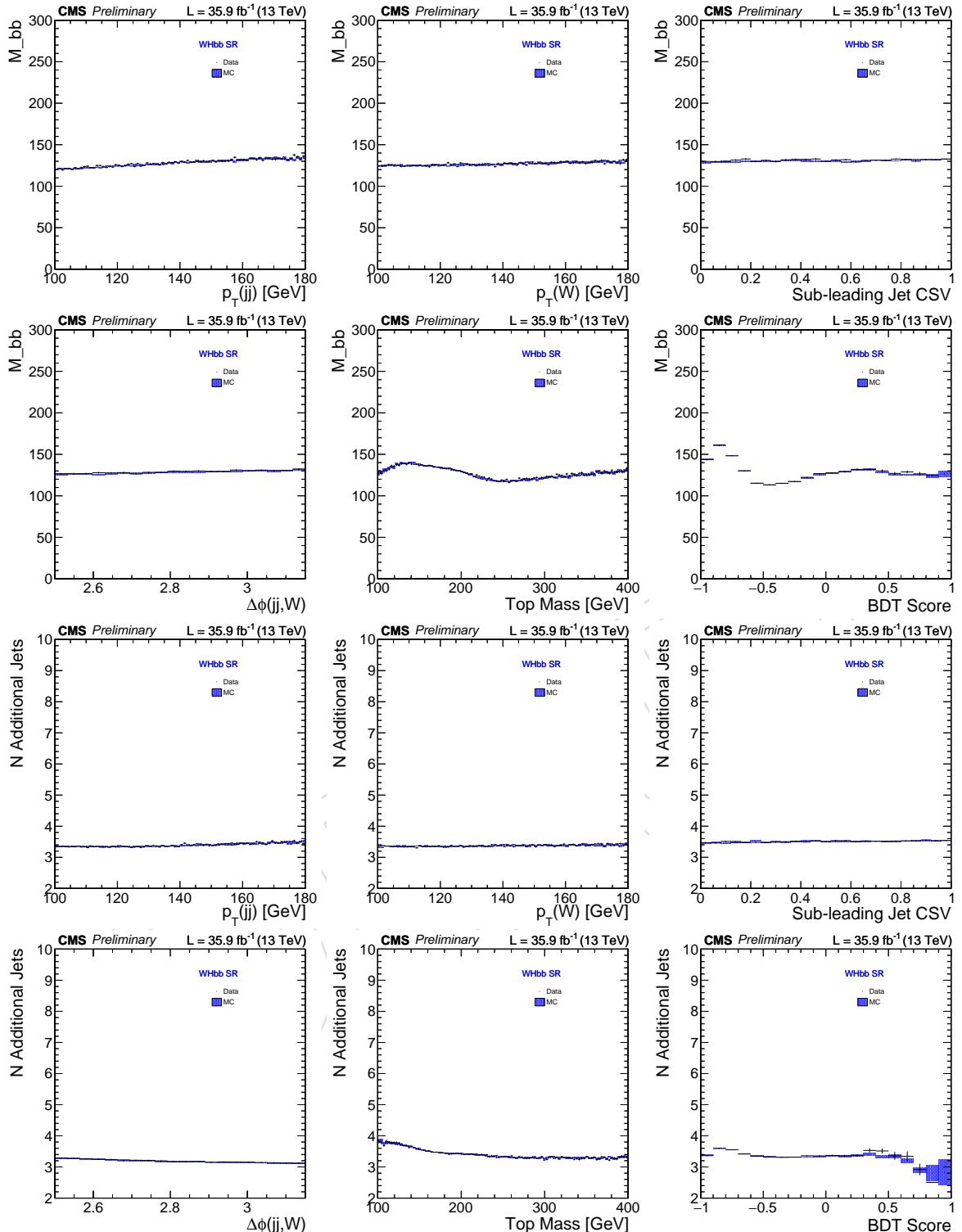


Figure 38: Correlations between BDT input variables as evaluated in the $t\bar{t}$ control region for the $W(\ell\nu)H$ channel (electron and muon channels are merged) for data (red) and simulated events (black). See Sec. 11 for the $t\bar{t}$ control sample definition. Left to right and top to bottom: $p_T(jj)$ vs. $M(jj)$, $p_T(W)$ vs. $M(jj)$, $\min(\text{CMVA1}, \text{CMVA2})$ vs. $M(jj)$, $\Delta\phi(W, H)$ vs. $M(jj)$, top mass vs. $M(jj)$, BDT vs. $M(jj)$, $p_T(jj)$ vs. N_{aj} , $p_T(W)$ vs. N_{aj} , $\min(\text{CMVA1}, \text{CMVA2})$ vs. N_{aj} , $\Delta\phi(W, H)$ vs. N_{aj} , top mass vs. N_{aj} , BDT vs. N_{aj} . Acceptable agreement is found for all pairs of variables, including the key correlation between the BDT output and the most discriminating variable, $M(jj)$.

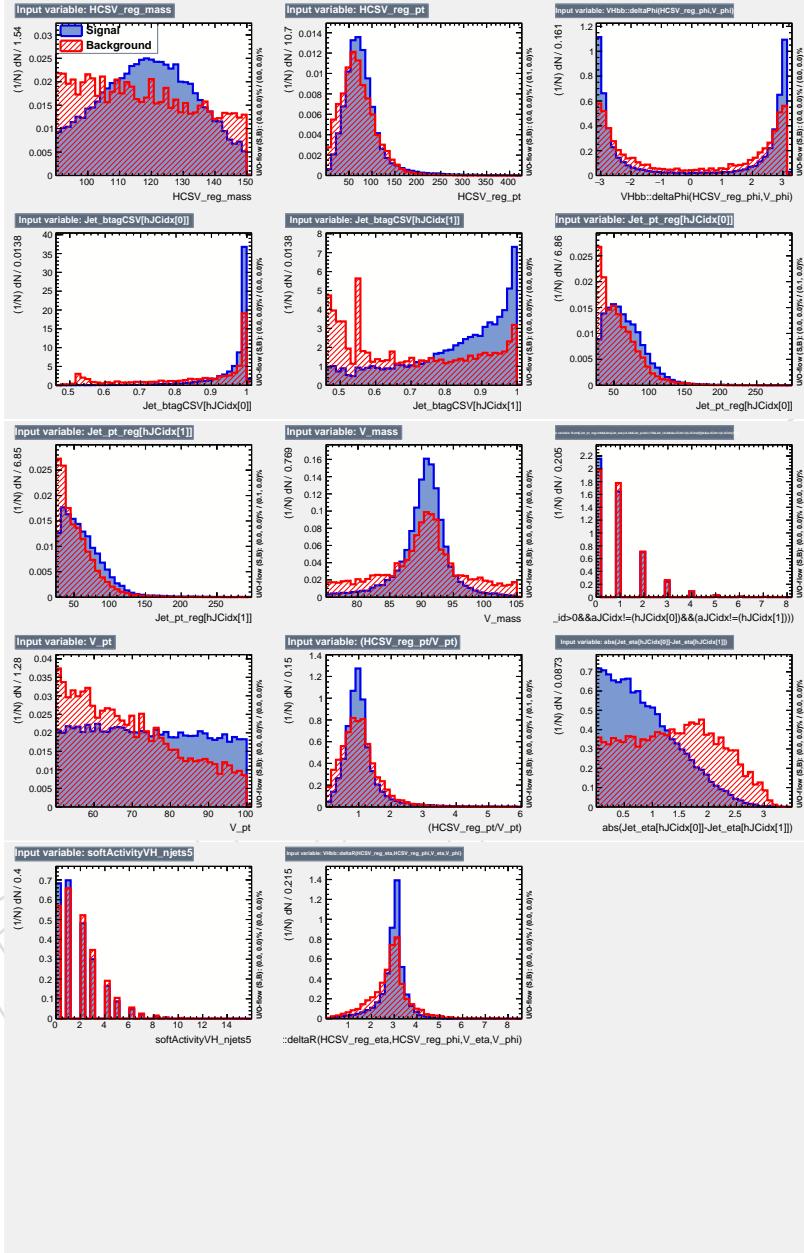


Figure 39: Distributions of signal (blue) and all backgrounds (red) for the variables trained in the BDT for $Z(\ell\ell)H$ in the low p_T bin.

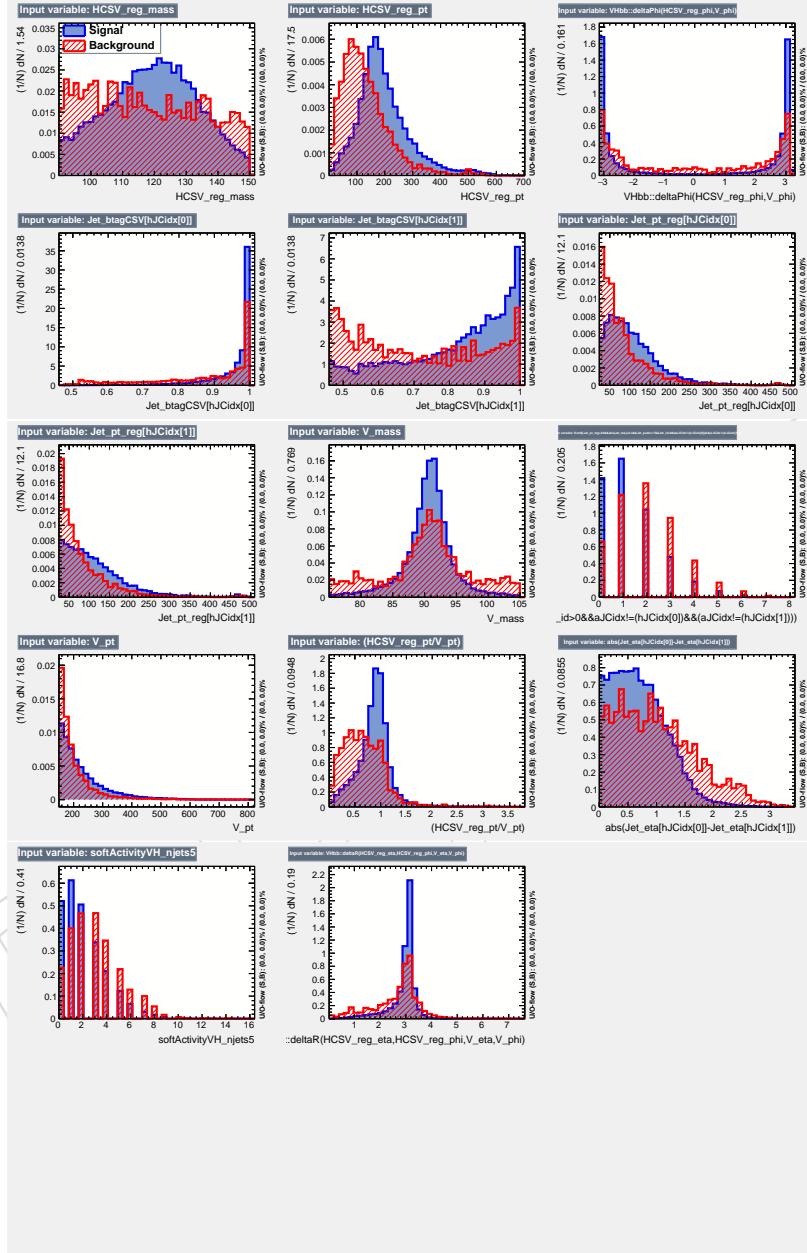


Figure 40: Distributions of signal (blue) and all backgrounds (red) for the variables trained in the BDT for $Z(\ell\ell)H$ in the high p_T bin.

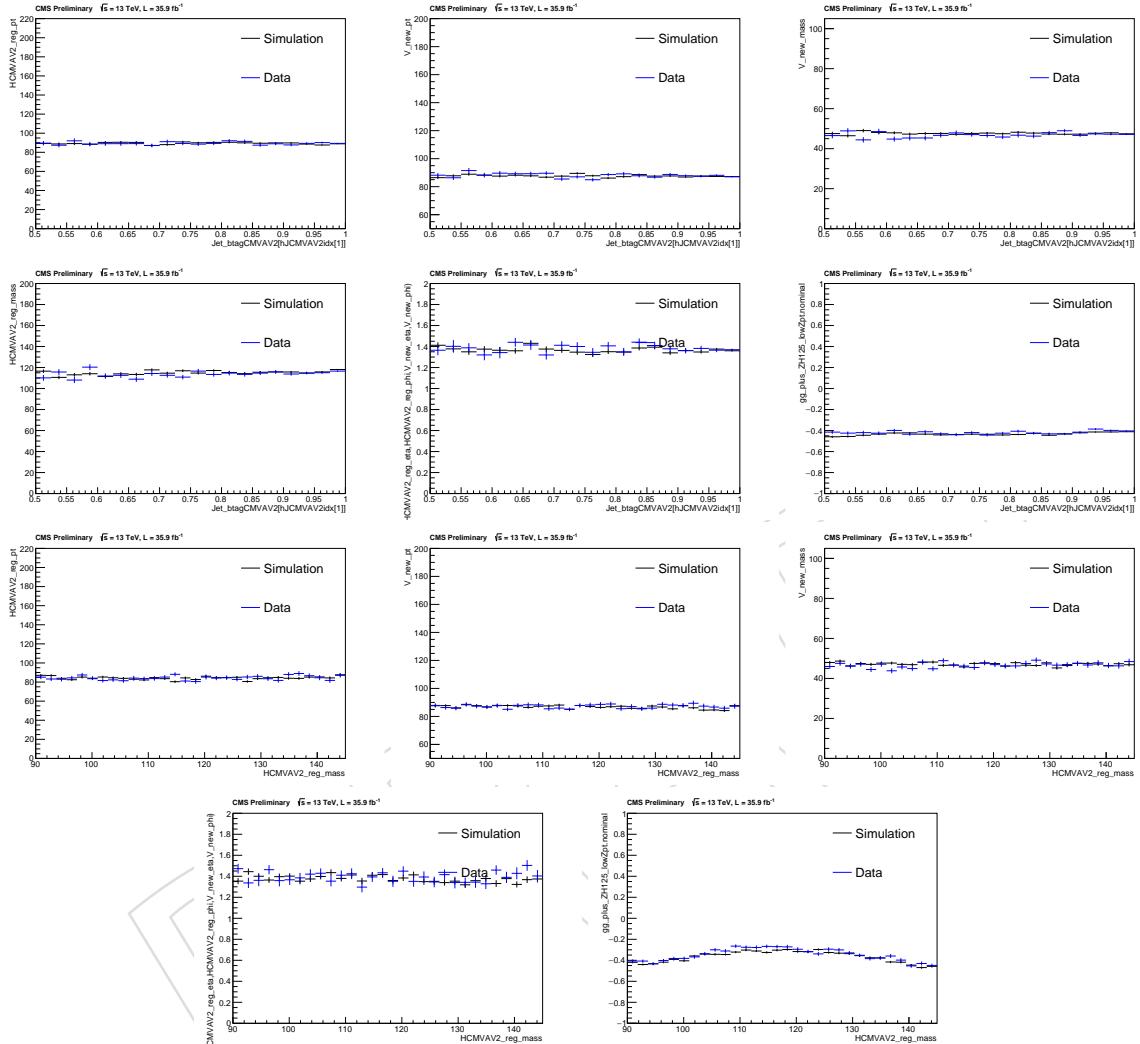


Figure 41: Correlations between BDT input variables as evaluated in the $t\bar{t}$ control region for the $Z(\ell\ell)H$ channel (electron and muon channels are merged) for data (red) and simulated events (black). See Sec. 11 for the $t\bar{t}$ control sample definition. Left to right and top to bottom: $p_T(jj)$ vs. minCMVA , $p_T(Z)$ vs. minCMVA , M_Z vs. minCMVA , $M(jj)$ vs. minCMVA , $\Delta\phi(Z, H)$ vs. minCMVA , BDT vs. minCMVA , $p_T(jj)$ vs. $M(jj)$, $p_T(Z)$ vs. $M(jj)$, M_Z vs $M(jj)$, $\Delta\phi(Z, H)$ vs. $M(jj)$, BDT vs. $M(jj)$. Acceptable agreement is found for all pairs of variables, including the key correlation between the BDT output and the most discriminating variable, $M(jj)$.

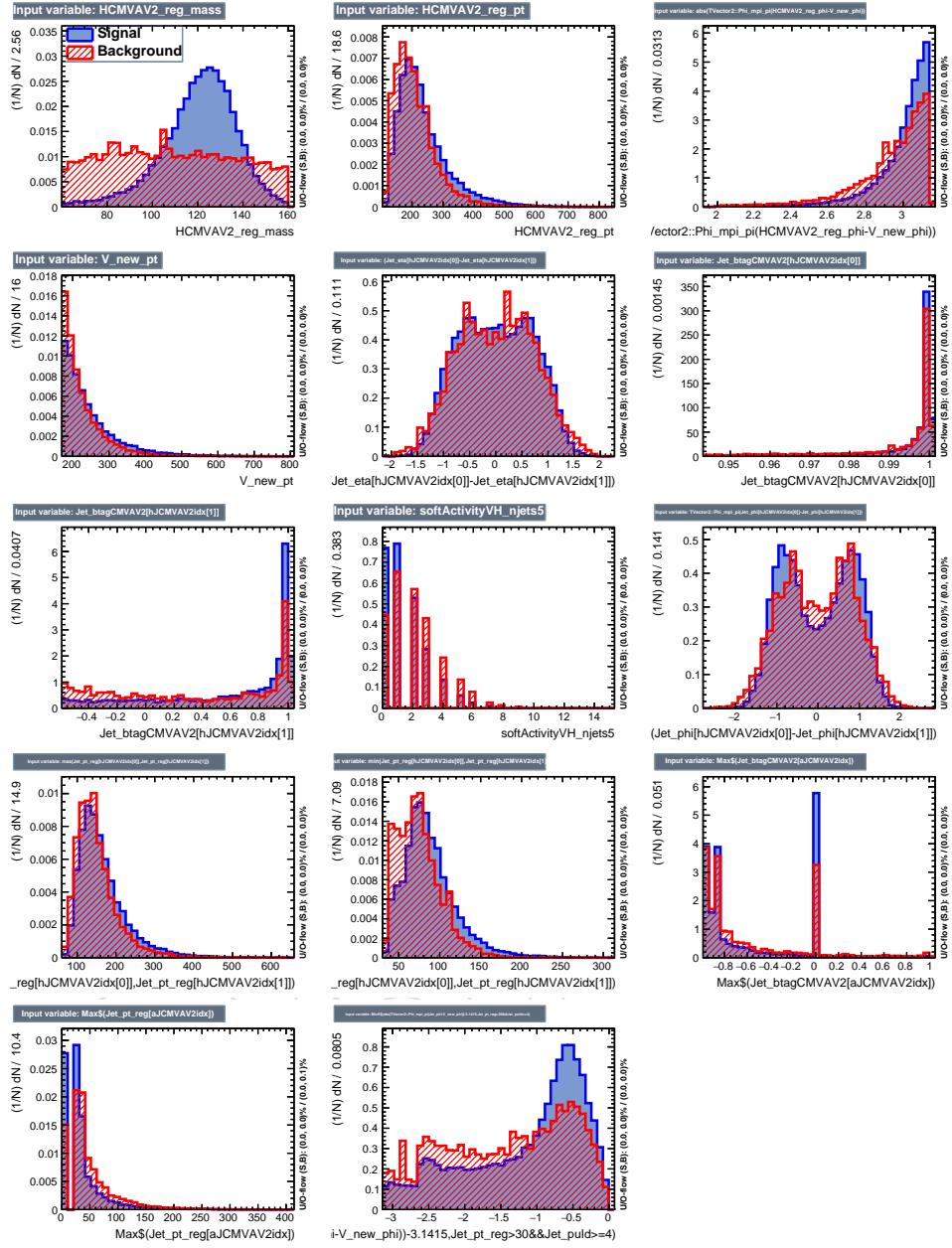


Figure 42: Distributions of signal (blue) and all backgrounds (red) for the variables trained in the BDT for $Z(\nu\nu)H$.

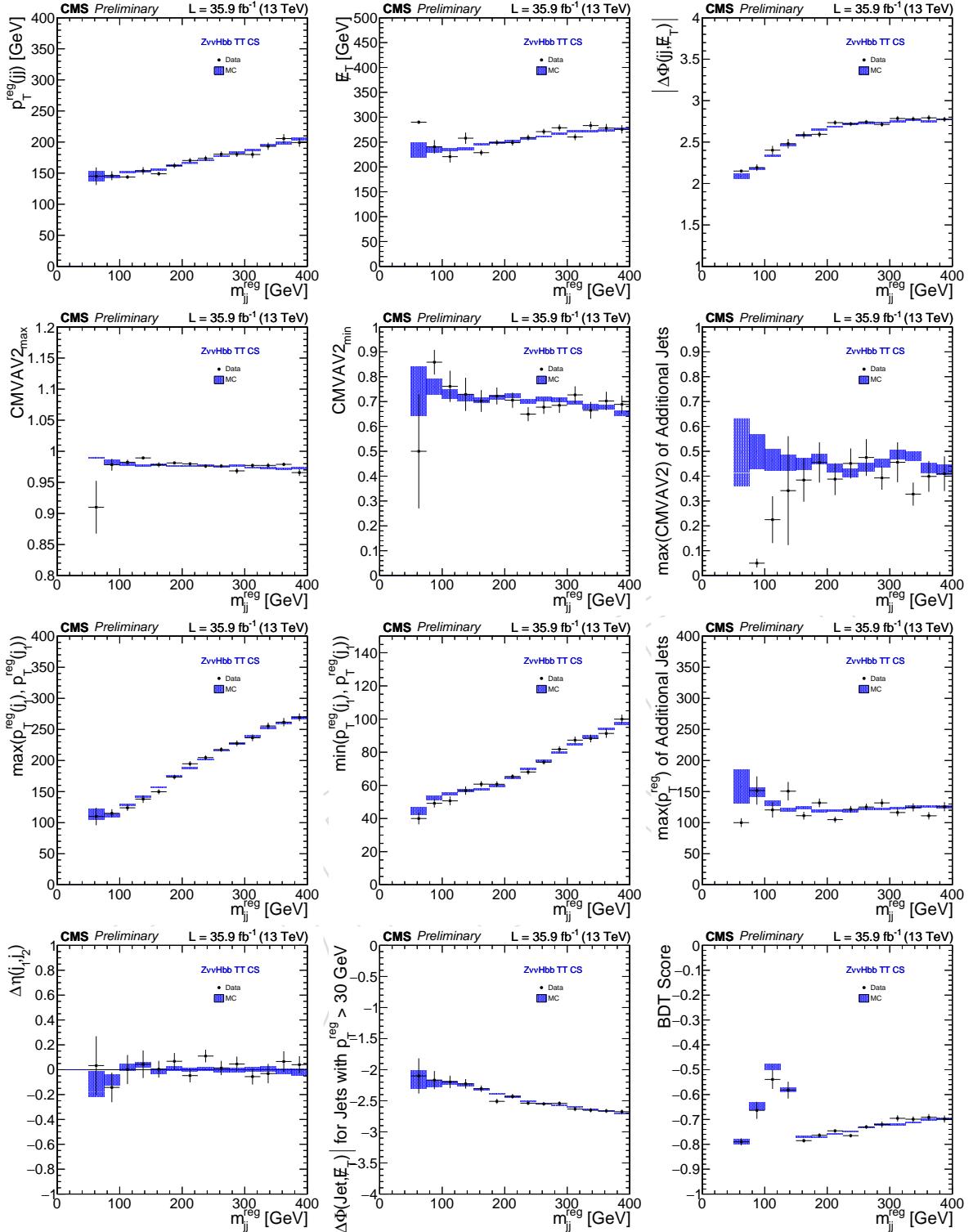


Figure 43: Correlations between BDT input variables as evaluated in the $t\bar{t}$ control region for the $Z(\nu\nu)H$ channel for data (blue) and simulated events (black). See Sec. 11 for the $t\bar{t}$ control sample definition. From left to right, top to bottom: $p_T(jj)$ vs. $M(jj)$, E_T^{miss} vs. $M(jj)$, $\Delta\phi(V, H)$ vs. $M(jj)$, CMVAV2_{\max} vs. $M(jj)$, CMVAV2_{\min} vs. $M(jj)$, $\text{CMVAV2}(\text{add})$ vs. $M(jj)$, $\max p_T(j)$ vs. $M(jj)$, $\min p_T(j)$ vs. $M(jj)$, $\max p_T(\text{add. jet})$ vs. $M(jj)$, $\Delta\eta(jj)$ vs. $M(jj)$, $\Delta\varphi(\text{pfMET}, \text{jet})$ vs. $M(jj)$, BDT vs. $M(jj)$. Acceptable agreement is found for all pairs of variables, including the key correlation between the BDT output and the most discriminating variable $M(jj)$.

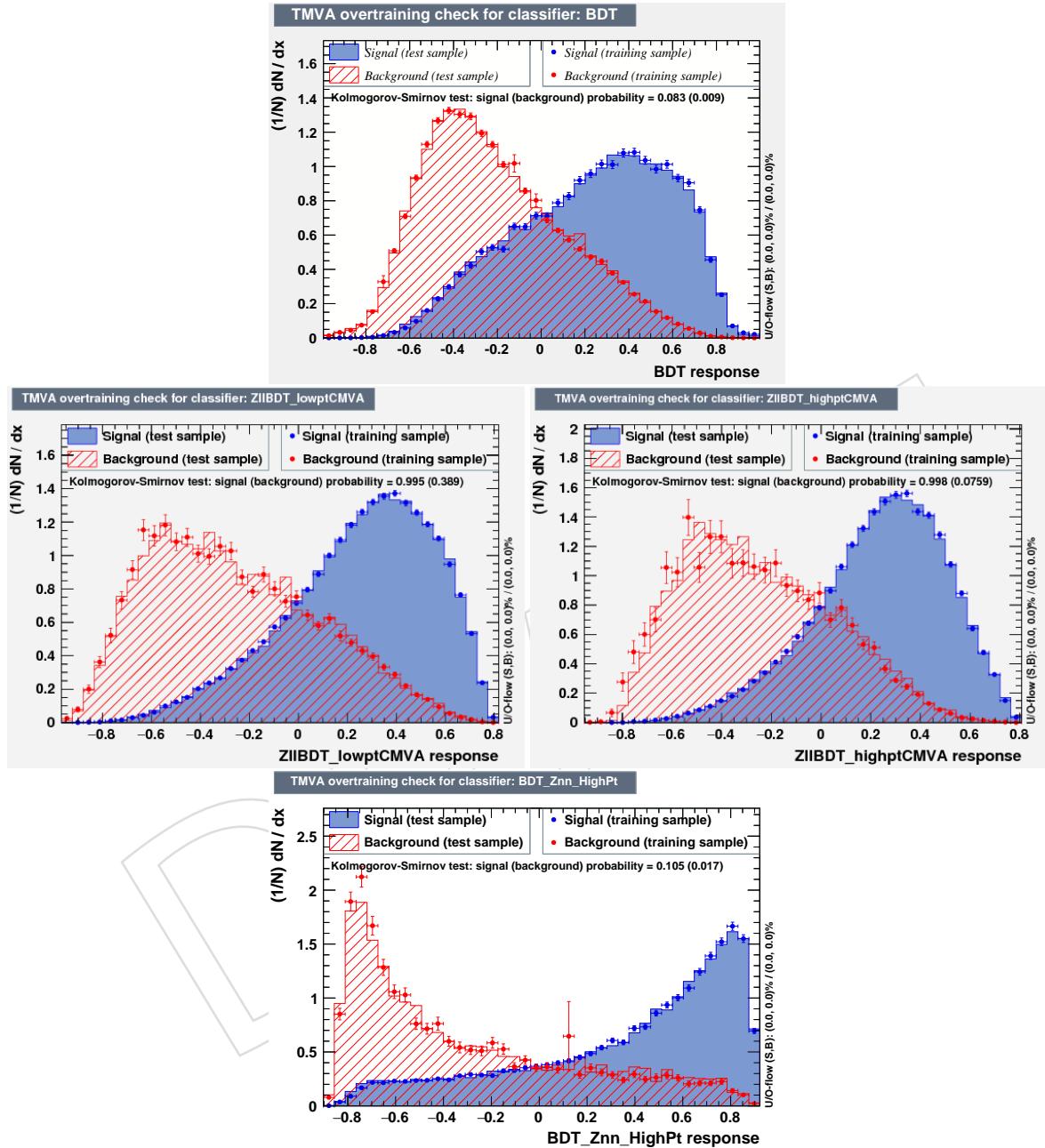


Figure 44: The plots above show the BDT output for signal and background simulation in each of the three analysis channels. From left to right, top to bottom: $W(\ell\nu)H$ $Z(\ell\ell)H$ (low- and high-boost), $Z(\nu\nu)H$.

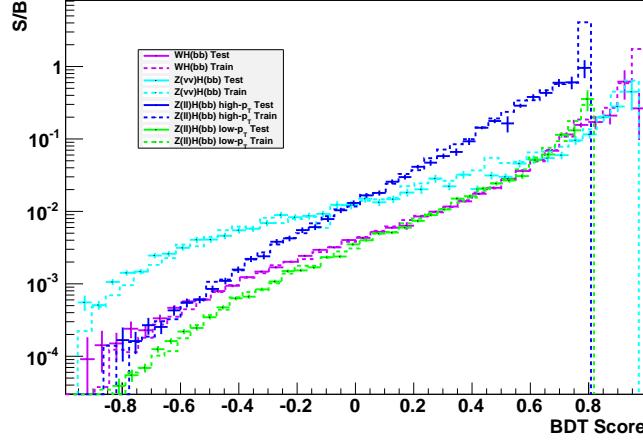


Figure 45: Expected signal over background yield vs. BDT score for each of the channels. The BDT's are trained on half of the MC statistics and evaluated on the remaining "test" half of the MC statistics to ensure no over-training. The BDT output is shown separately for training events (dashed histograms) and test events (points with solid error bars).

Table 14: Variable ranking for the BDT Classifiers (top variable is best ranked and has highest importance in the classifier).

Rank	$W(\ell\nu)H$	$Z(\nu\nu)H$	$Z(\ell\ell)H\text{low } p_T$	$Z(\ell\ell)H\text{high } p_T$
1	N_{aj}	Higgs mass	CMVA _{min}	Higgs p_T
2	Higgs p_T	$\min p_T(j)$	b-jets $\Delta\eta$	Z mass
3	E_T^{miss}	$\Delta\phi(\text{pfMET}, \text{jet})$	CMVA _{max}	Higgs mass
4	N_5^{soft}	CMVAV2 _{min}	$\Delta\phi(H, Z)$	CMVA _{max}
5	M_t	N_5^{soft}	$p_T(Z)$	CMVA _{min}
6	CMVA _{min}	$\Delta\eta(jj)$	Z mass	b-jets ΔR
7	Higgs mass	$\Delta\phi(V, H)$	Higgs mass	Leading jet p_T
8	$\Delta\phi(\text{pfMET}, \text{lept.})$	CMVAV2 _{max}	Higgs p_T	Soft activity
9	$\Delta\phi(V, H)$	$\Delta\phi(jj)$	b-jets ΔR	$p_T(Z)$
10	$p_T(V)$	max CMVAV2 _{add}	Soft activity	sub-leading jet p_T
11	$m_T(W)$	Higgs p_T	Leading jet p_T	N_{aj}
12	—	max $p_T(\text{add.jet})$	$p_T\text{balance}$	$p_T\text{balance}$
13	—	E_T^{miss}	sub-leading jet p_T	$\Delta\phi(H, Z)$
14	—	max $p_T(j)$	N_{aj}	b-jets $\Delta\eta$

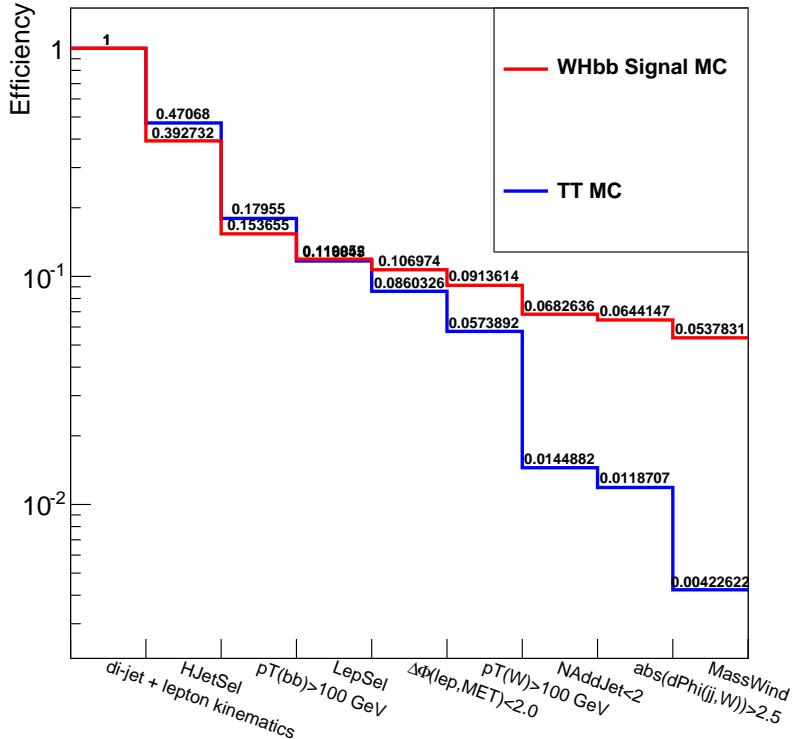


Figure 46: Efficiency and background reduction after each cut used in the signal region definition for the $W(\ell\nu)H$ channel, for signal and $t\bar{t}$, the dominant background. The plot has been obtained requiring at least two jets and one lepton with $p_T > 25 \text{ GeV}$.

829 10 Signal Efficiency

830 Figure 46 and Figure 47 shows the efficiency and the background reduction after each cut used in the
 831 signal region definition of the $W(\ell\nu)H$ and $Z(\ell\ell)H$ analysis respectively.

832 10.1 Lepton reconstruction and ID efficiencies

833 Muon reconstruction efficiency data/MC scale factors in 2016 data are computed in two steps. First,
 834 the efficiency for any global muon to be reconstructed as a loose (tight) muon is found, followed by the
 835 efficiency for a loose (tight) muon to pass the isolation cuts. The efficiencies are separated in two run
 836 periods: Run BCDEF and Run GH. The results for the loose identification and the loose isolation on top
 837 of a loose identified muon are shown in Figure 48, 49 for Run BCDEF and Figure 50, 51 for Run GH. The
 838 results for the tight identification and the tight isolation on top of a loose identified muon are shown in
 839 Figure 52, 53 for Run BCDEF and Figure 54, 55 for Run GH. For both working points, the scale factors
 840 in the ratio plot are not strongly dependent on p_T and appear only as a function of η , although these
 841 factors are parameterized in both variables when applied in the analysis.

842 Electron reconstruction efficiency data/MC scale factors in 2016 data are computed together for the
 843 MVA-based identification cut and for the isolation cut. Two working points are considered WP80 and
 844 WP90. The efficiency of the identification and isolation requirement for these working points appear in
 845 Figure 56.

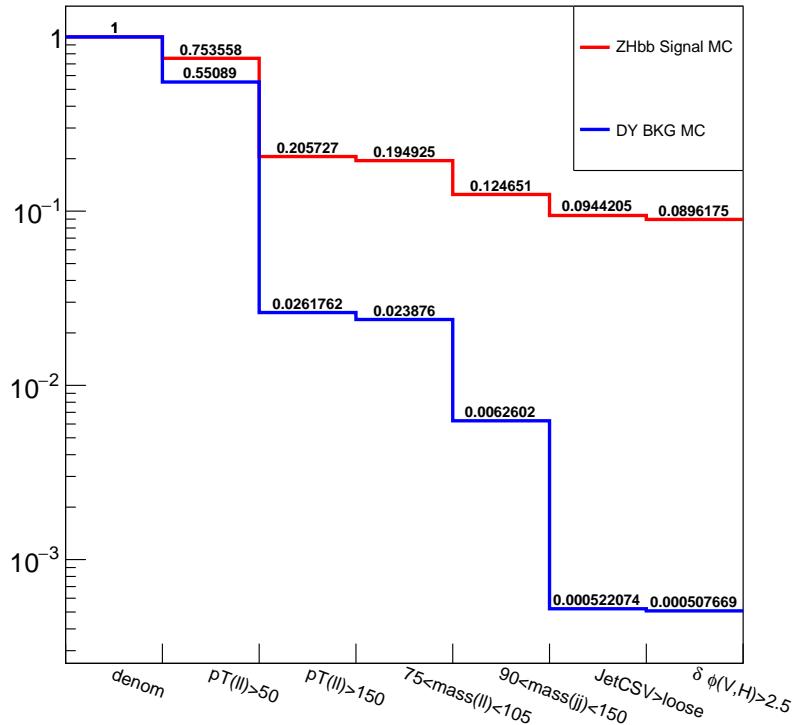


Figure 47: Efficiency and background reduction after each cut used in the signal region definition for the $Z(\ell\ell)H$ channel, for signal and $Z+jets$, the dominant background. The plot has been obtained requiring at least two jets and one lepton with $p_T > 20$ GeV.

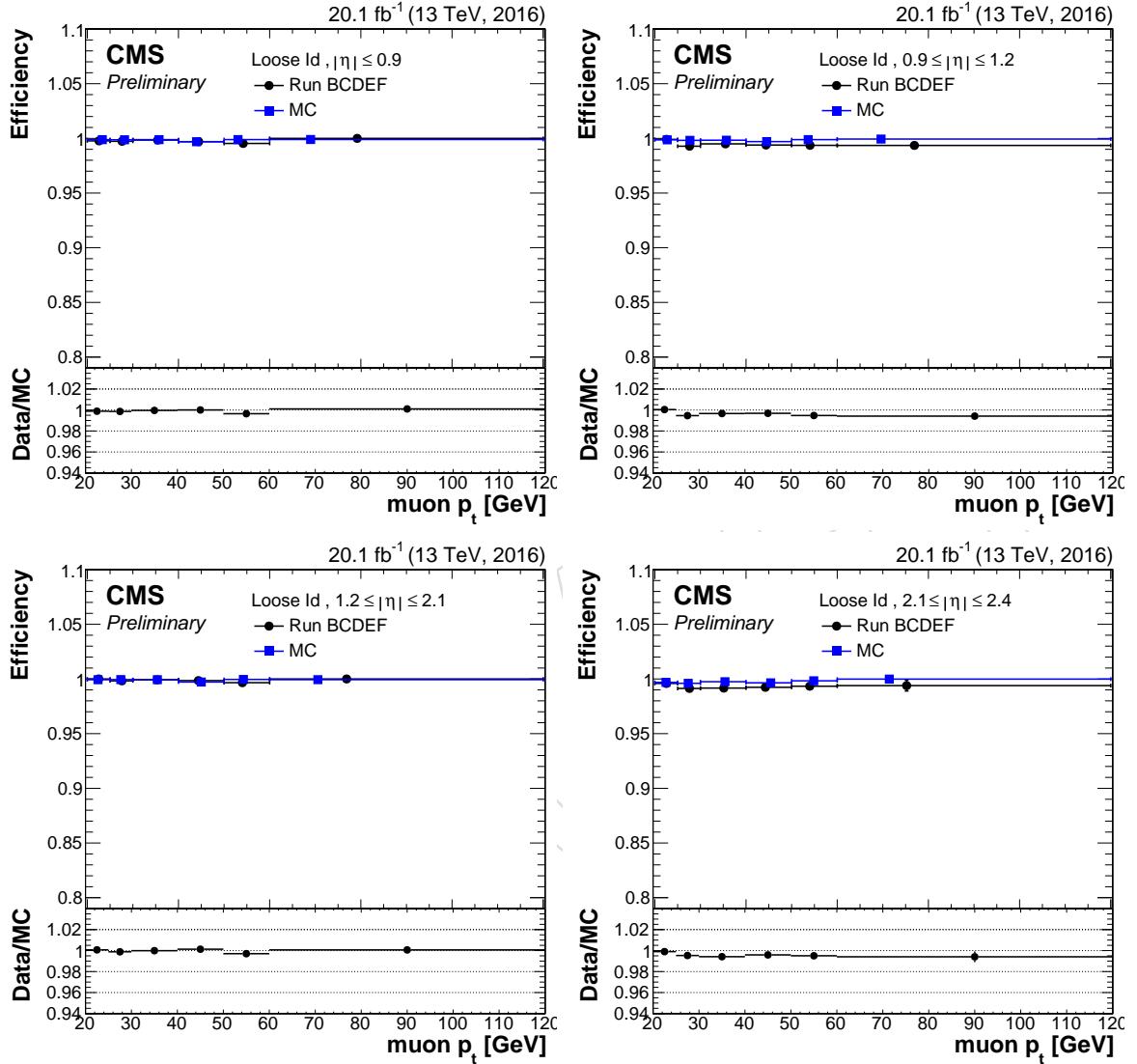


Figure 48: Data and MC efficiency for a muon to pass the loose ID, distributed in the muon p_T . The dataset correspond to the runs B, C, D, E and F. The efficiency have been computed in 4 $|\eta|$ bins. From left to right, top to bottom: $|\eta| < 0.9$, $0.9 < |\eta| < 1.2$, $1.2 < |\eta| < 2.1$, $2.1 < |\eta| < 2.4$

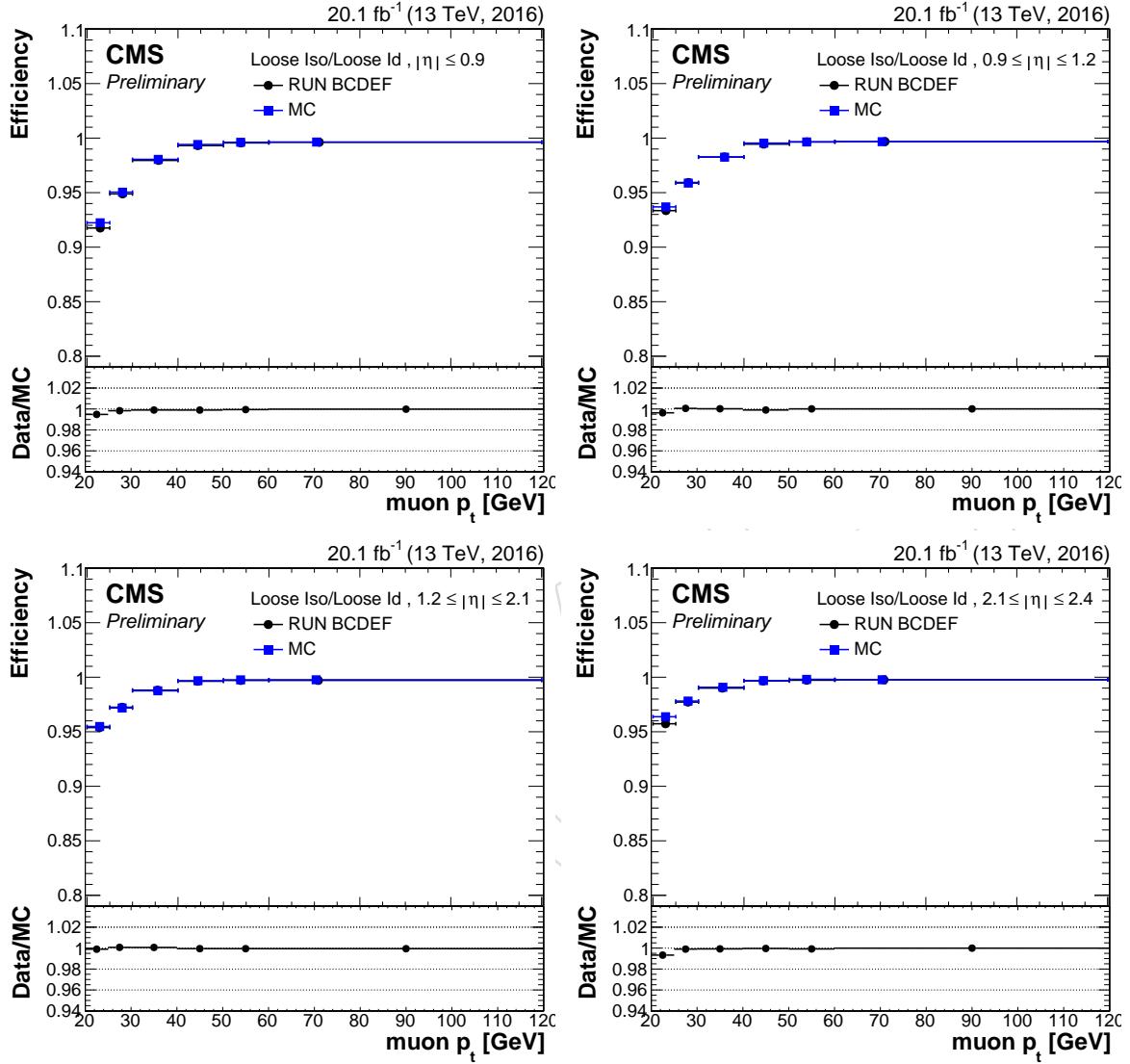


Figure 49: Data and MC efficiency for loose identified muon to pass the loose ISO, distributed in the muon p_T . The dataset correspond to the runs B, C, D, E and F. The efficiency have been computed in 4 $|\eta|$ bins. From left to right, top to bottom: $|\eta| < 0.9$, $0.9 < |\eta| < 1.2$, $1.2 < |\eta| < 2.1$, $2.1 < |\eta| < 2.4$

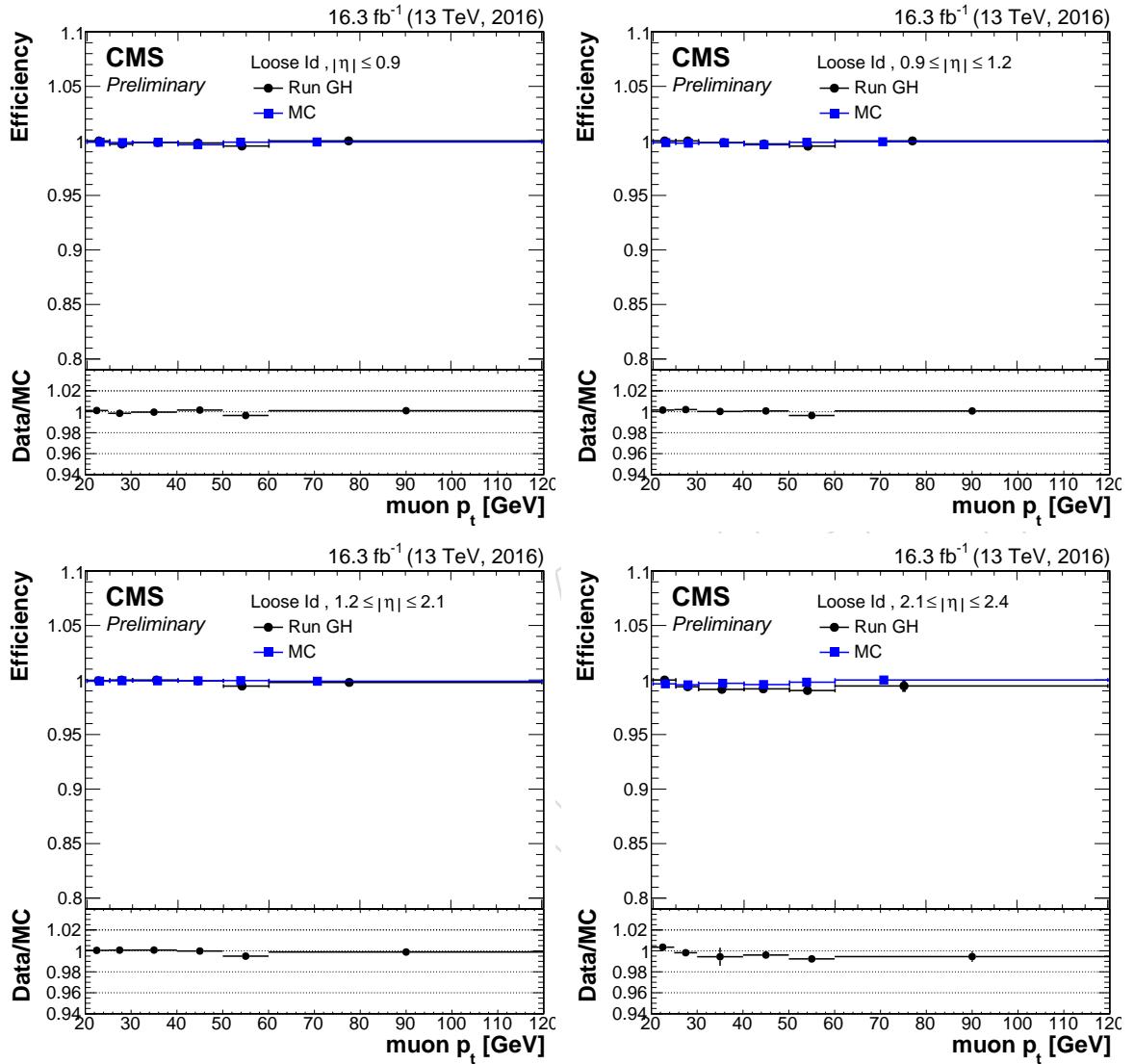


Figure 50: Data and MC efficiency for a muon to pass the loose ID, distributed in the muon p_T . The dataset correspond to the runs G and H. The efficiency have been computed in 4 $|\eta|$ bins. From left to right, top to bottom: $|\eta| < 0.9$, $0.9 < |\eta| < 1.2$, $1.2 < |\eta| < 2.1$, $2.1 < |\eta| < 2.4$

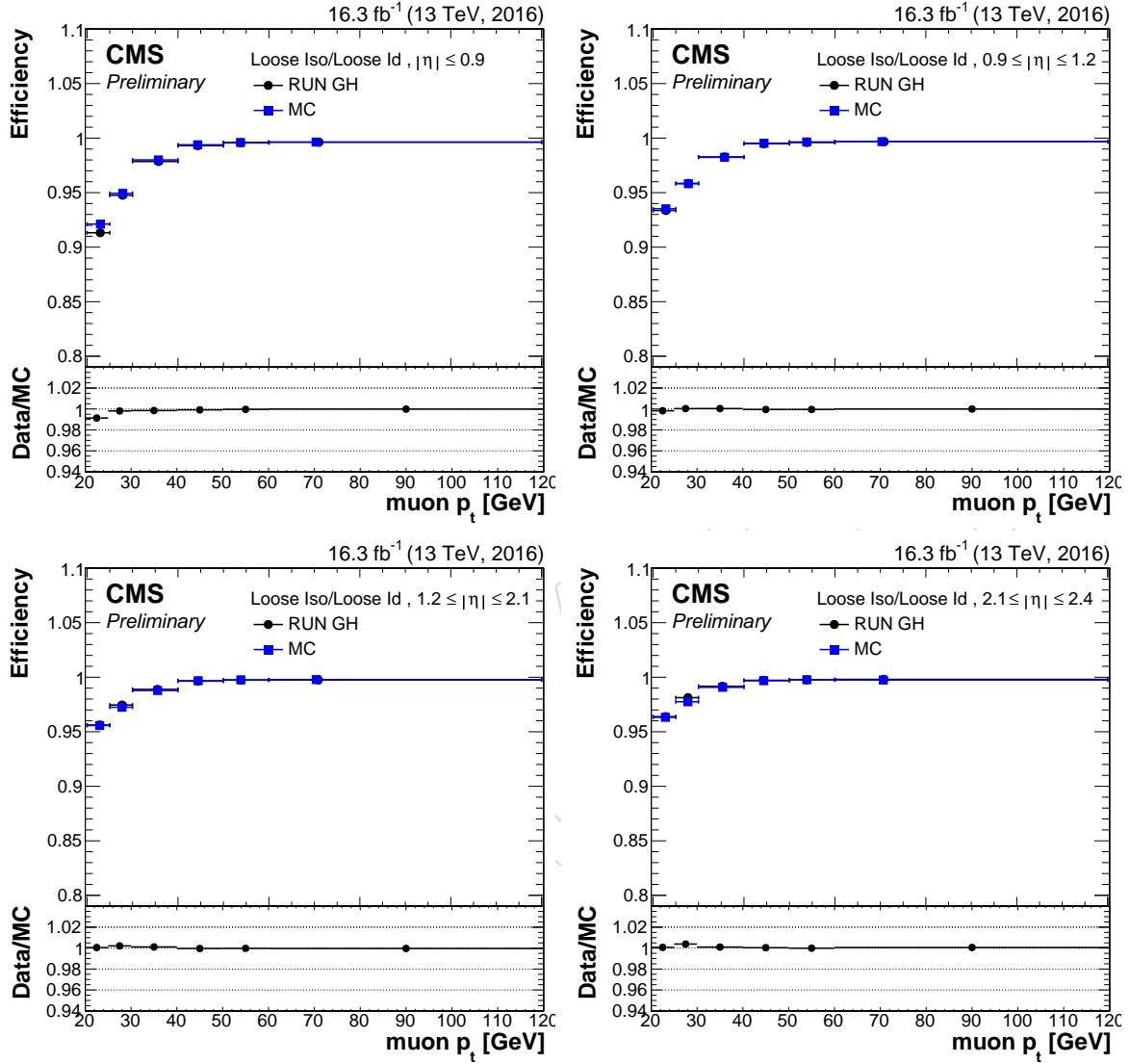


Figure 51: Data and MC efficiency for loose identified muon to pass the loose ISO, distributed in the muon p_T . The dataset correspond to the runs G and H. The efficiency have been computed in 4 $|\eta|$ bins. From left to right, top to bottom: $|\eta| < 0.9$, $0.9 < |\eta| < 1.2$, $1.2 < |\eta| < 2.1$, $2.1 < |\eta| < 2.4$

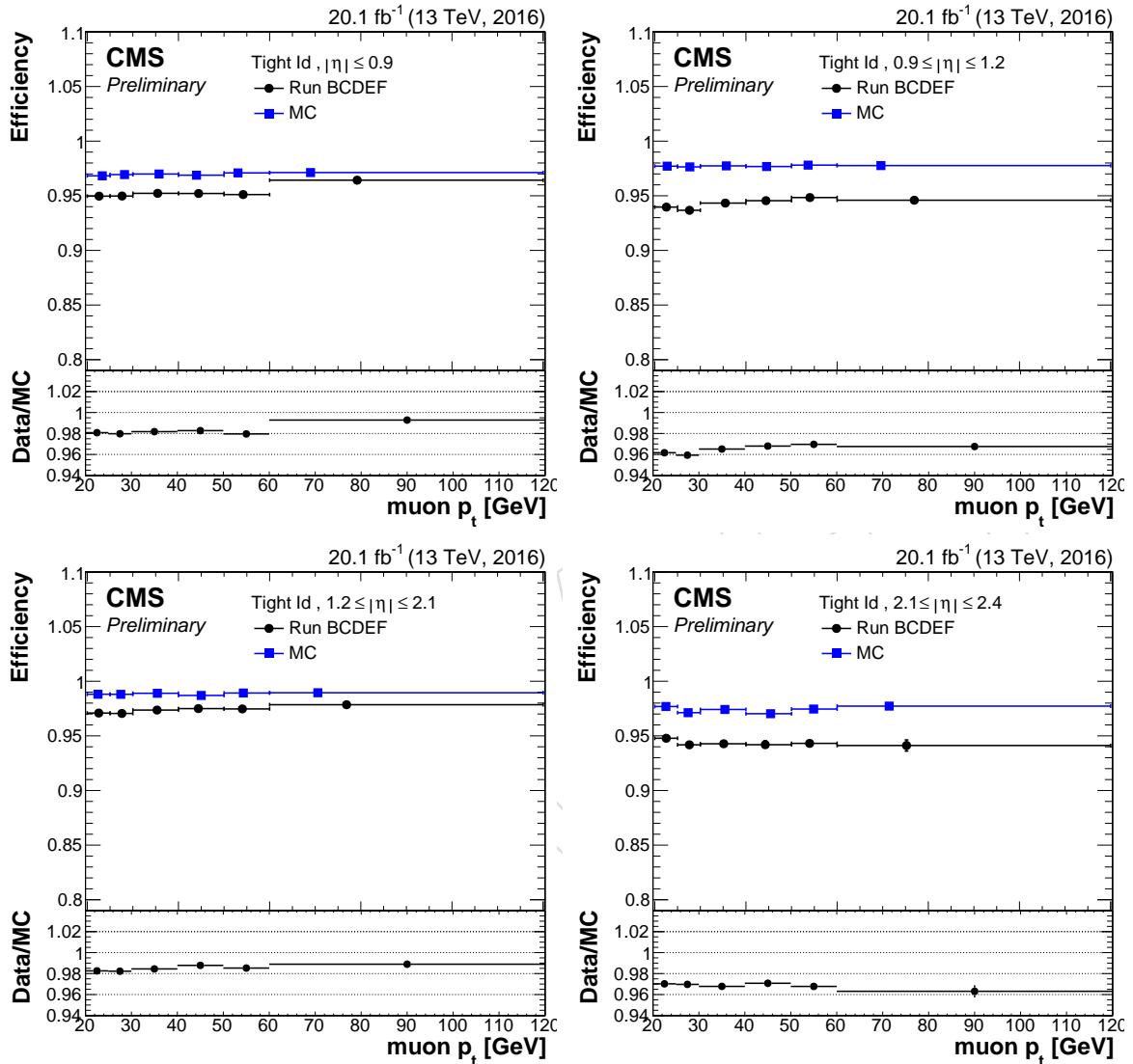


Figure 52: Data and MC efficiency for a muon to pass the tight ID, distributed in the muon p_T . The dataset correspond to the runs B, C, D, E and F. The efficiency have been computed in 4 $|\eta|$ bins. From left to right, top to bottom: $|\eta| < 0.9$, $0.9 < |\eta| \leq 1.2$, $1.2 < |\eta| \leq 2.1$, $2.1 < |\eta| \leq 2.4$

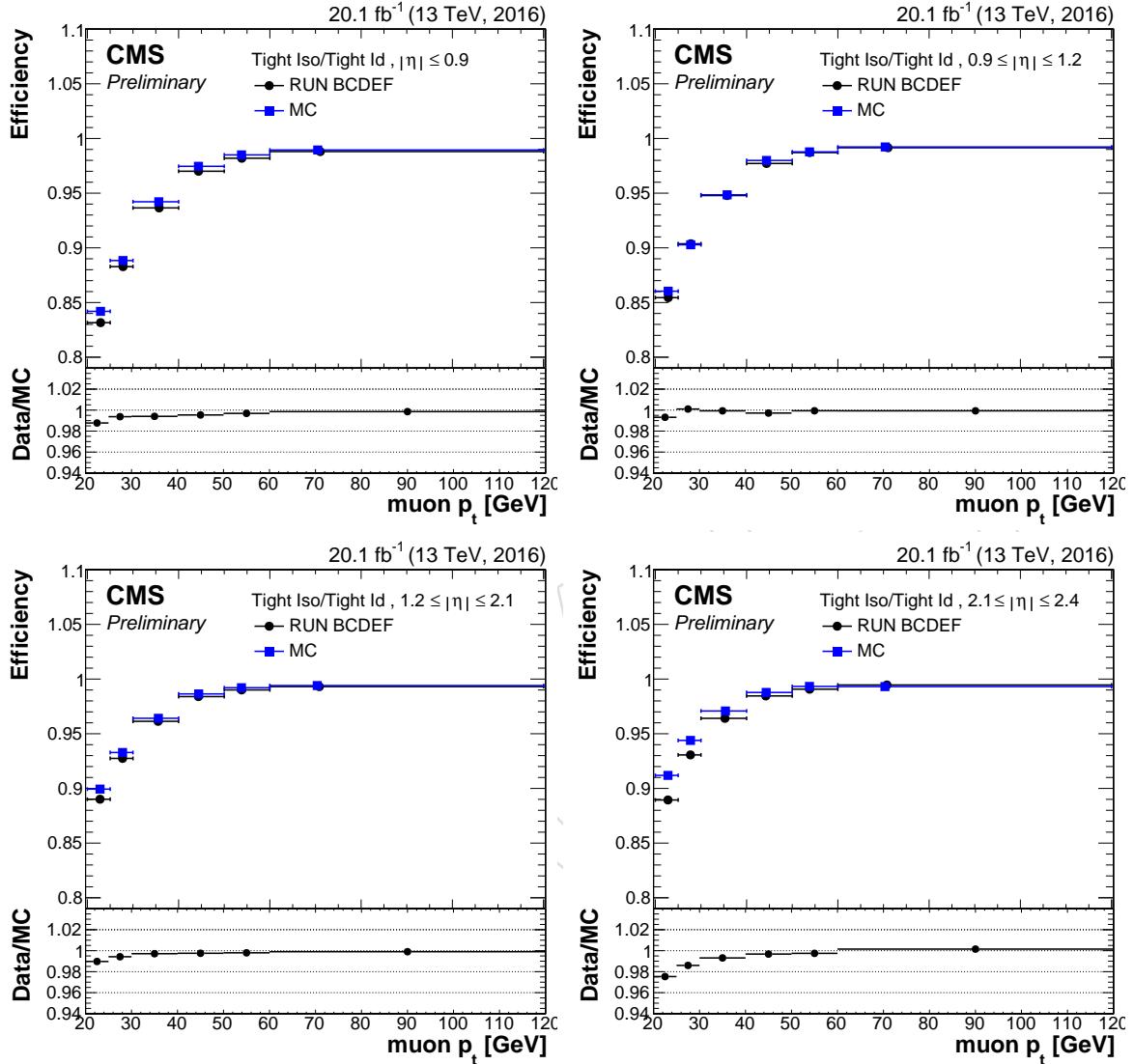


Figure 53: Data and MC efficiency for tight identified muon to pass the tight ISO, distributed in the muon p_T . The dataset correspond to the runs B, C, D, E and F. The efficiency have been computed in 4 $|\eta|$ bins. From left to right, top to bottom: $|\eta| < 0.9$, $0.9 < |\eta| < 1.2$, $1.2 < |\eta| < 2.1$, $2.1 < |\eta| < 2.4$

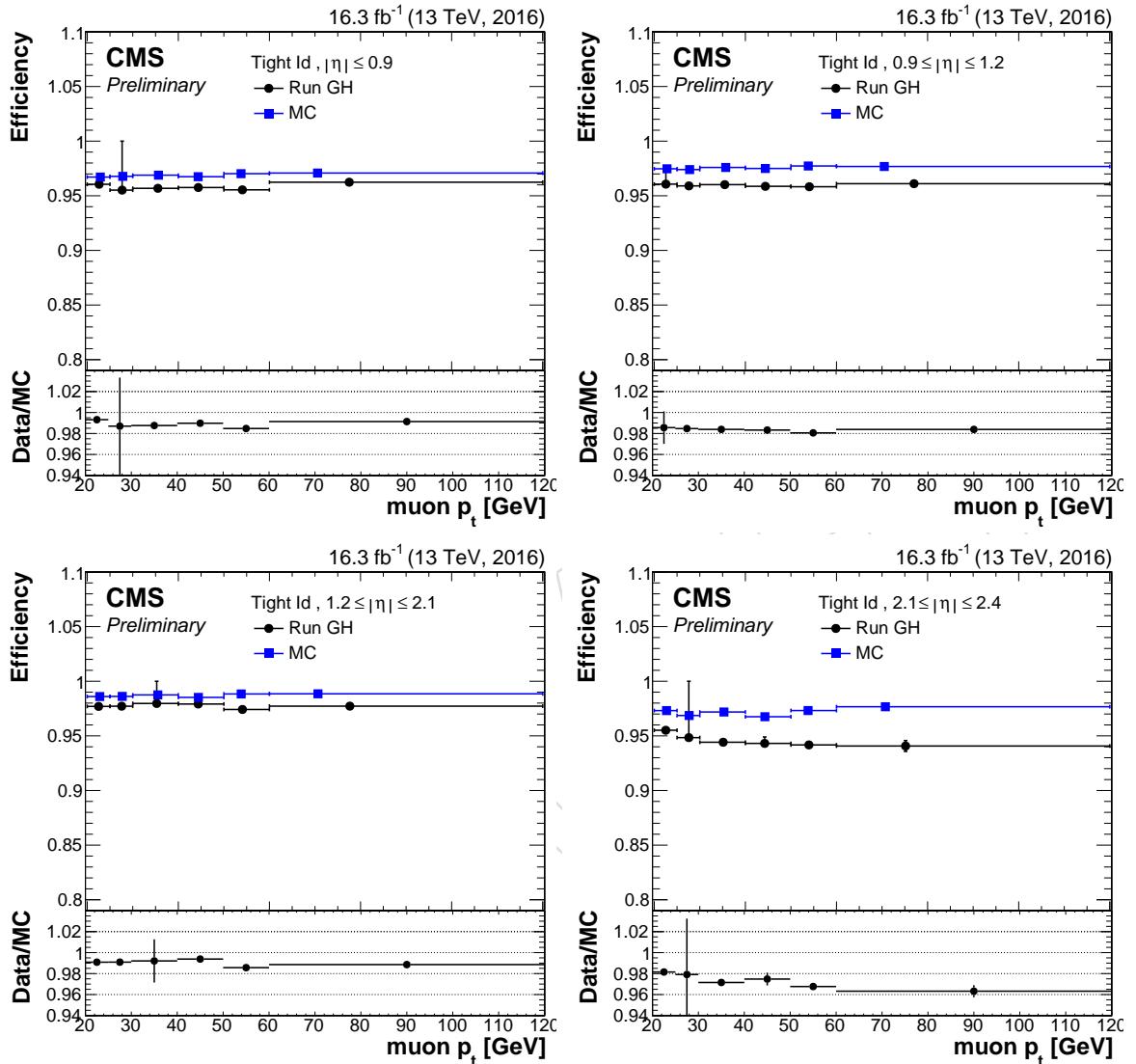


Figure 54: Data and MC efficiency for a muon to pass the tight ID, distributed in the muon p_T . The dataset correspond to the runs G and H. The efficiency have been computed in 4 $|\eta|$ bins. From left to right, top to bottom: $|\eta| < 0.9$, $0.9 < |\eta| < 1.2$, $1.2 < |\eta| < 2.1$, $2.1 < |\eta| < 2.4$

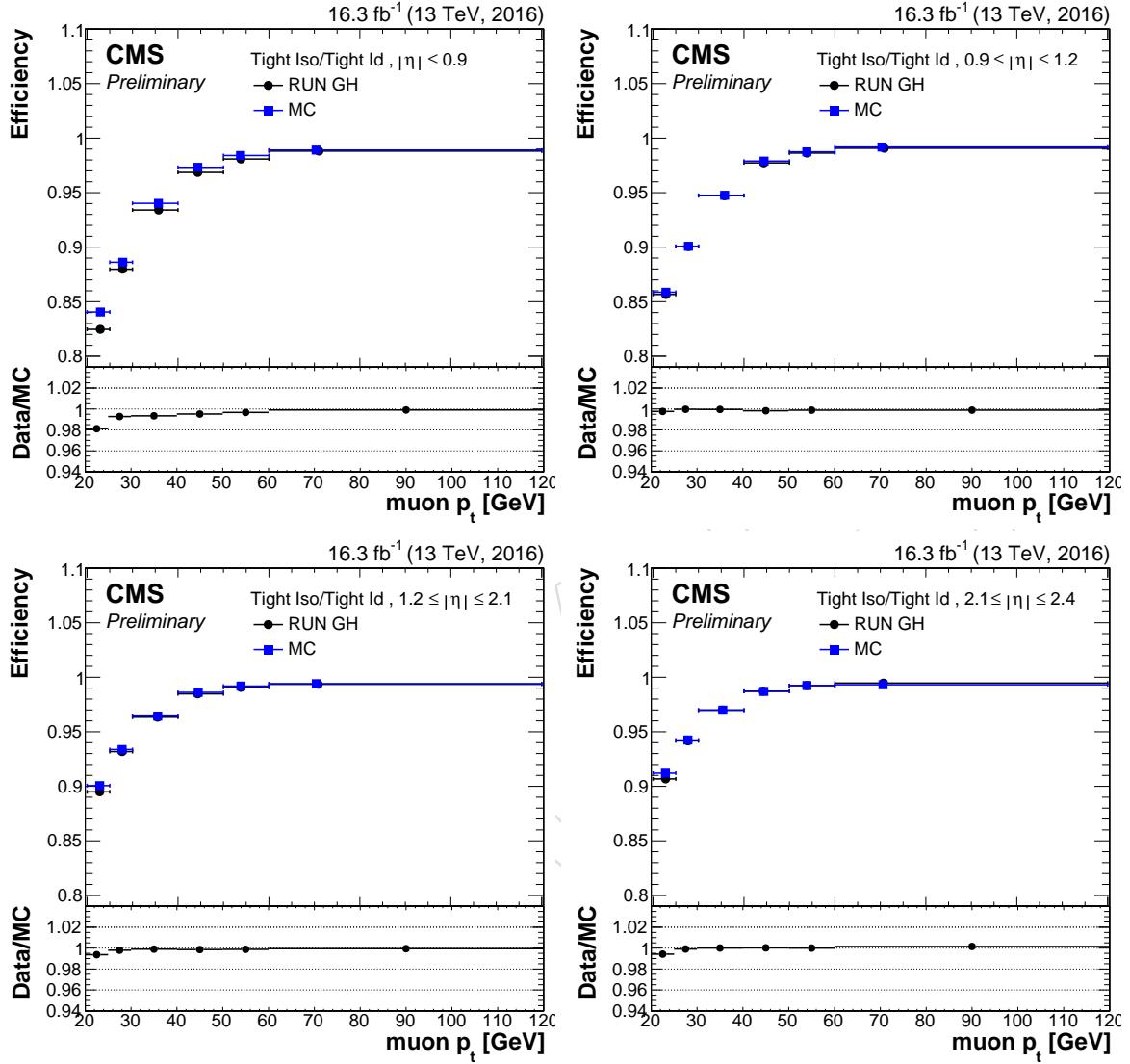


Figure 55: Data and MC efficiency for tight identified muon to pass the tight ISO, distributed in the muon p_T . The dataset correspond to the runs G and H. The efficiency have been computed in 4 $|\eta|$ bins. From left to right, top to bottom: $|\eta| < 0.9$, $0.9 < |\eta| < 1.2$, $1.2 < |\eta| < 2.1$, $2.1 < |\eta| < 2.4$

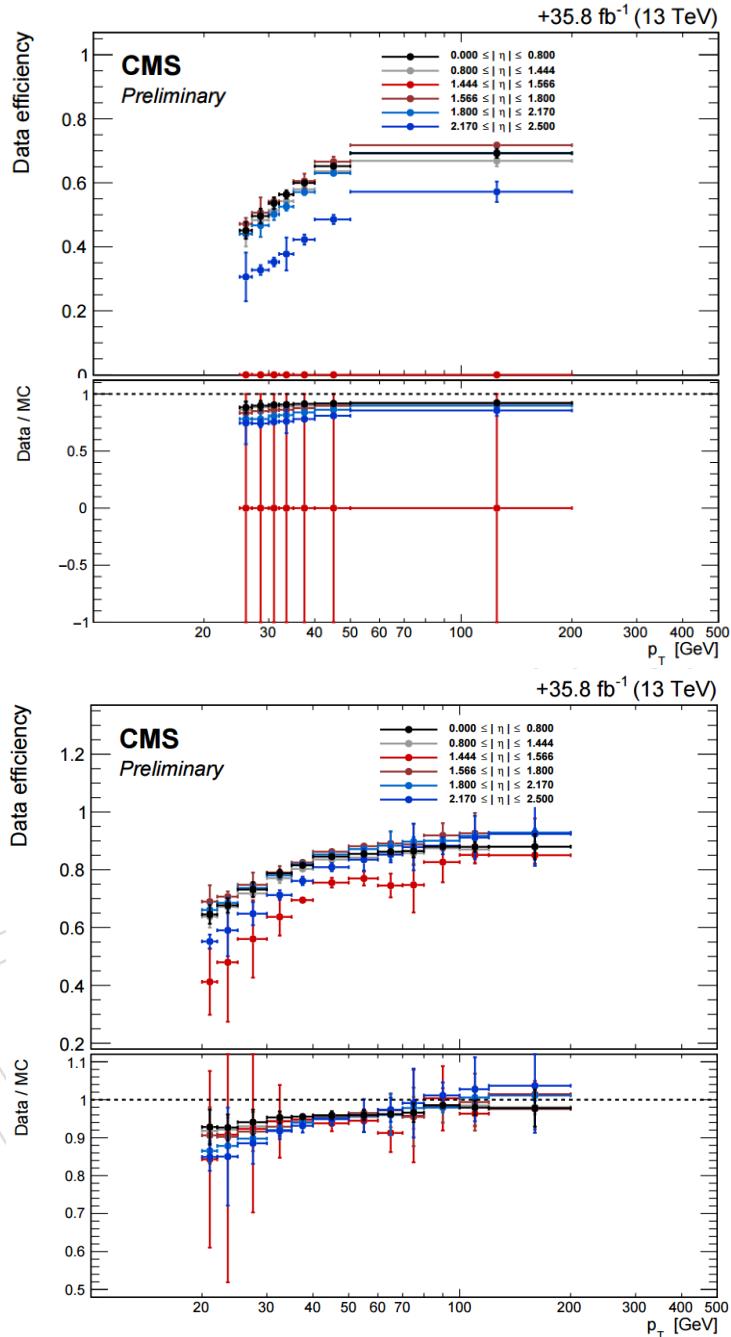


Figure 56: Efficiency for identification MVA cut for the electron selection for the WP80 (above) and WP90 (below) working points and isolation cuts in the 2016 sample, split into several regions in η .

Table 15: Definition of control regions for the $Z(\ell\ell)H$ channel.

Variable	Z+LF	$t\bar{t}$	Z+HF
$p_T(jj)$	> 100	> 100	–
$p_T(V)$	[50, 150], > 150	[50, 150], > 150	[50, 150], > 150
CMVAV2 _{max}	< CMVAV2 Tight	> CMVAV2 Tight	> CMVAV2 Tight
CMVAV2 _{min}	< CMVAV2 Loose	> CMVAV2 Loose	> CMVAV2 Loose
N_{aj}	–	–	–
N_{al}	–	–	–
E_T^{miss}	–	–	< 60
$\Delta\phi(V, H)$	–	–	> 2.5
$m_{\ell\ell}$	75–105	veto 0–10, 75–120	85–97
$M(jj)$	–	–	veto 90–150

846 11 Background Control Samples

847 Appropriate control regions are identified in data and used to adjust Monte Carlo estimates for several
 848 of the most important background processes, including production of W and Z bosons in association
 849 with jets, and $t\bar{t}$ production. A set of simultaneous fits is then performed to the control regions sepa-
 850 rately in each channel to obtain consistent data/MC scale factors. These scale factors account not only
 851 for cross section discrepancies, but also potential residual differences in physics object selection. There-
 852 fore, separate scale factors are used for each background source in the different channels. The final
 853 extrapolation of the predicted background yield from the control region (determined from data) to the
 854 fit region of the shape analysis is obtained from simulation, with appropriate systematic uncertainties
 855 to take into account potential shape differences between data and Monte Carlo samples.

856 The control regions are defined with orthogonal event selections and separate scale factor fits are per-
 857 formed in each dataset. Although the scale factors are not necessarily expected to be the same in each
 858 data sample (again, because these are not cross section measurements), consistency is a useful check that
 859 nothing significant has changed in the data or simulation.

860 11.1 $Z(\ell\ell)H$ control regions

861 This section describes the control regions for $Z + \text{udscg}$, $t\bar{t}$, and $Z + b\bar{b}$ production as reconstructed in
 862 the $Z(\ell\ell)H$ channel. Table 15 summarizes the selection criteria used to define each samples that enter
 863 the scale factor fit.

864 11.2 $Z(\nu\nu)H$ control regions

865 This section describes the control regions for multijet $V + \text{udscg}$, $t\bar{t}$, and $V + b\bar{b}$ production as recon-
 866 structed in the $Z(\nu\nu)H$ channel requiring $E_T^{\text{miss}} > 170 \text{ GeV}$.

867 The cuts used to defined the signal and control regions are reported in Tab. 16.

- 868 • The signal region is defined requiring $Z \rightarrow \nu\nu$ category, to be orthogonal with the other
 869 $VH(bb)$ searches. A lepton veto is applied and at least one medium b -tagged jet is required.
 870 A m_{jj} cut is used to separate the signal from the $Z + b$ -jets control region. The multijet back-
 871 ground is reduced applying the anti-QCD cut.
- 872 • The $Z + b$ -jets control region is similar to the signal region, but inverting the m_{jj} cut. The
 873 anti-QCD cut is used to increase the $Z + b$ -jets purity.
- 874 • The $Z + \text{light-jets}$ control region is defined inverting the b -tagging cut and removing the m_{jj}
 875 cut. The remaining cuts are identical to the $Z + b$ -jets control region.
- 876 • In the multijet control region, the anti-QCD cut is inverted.

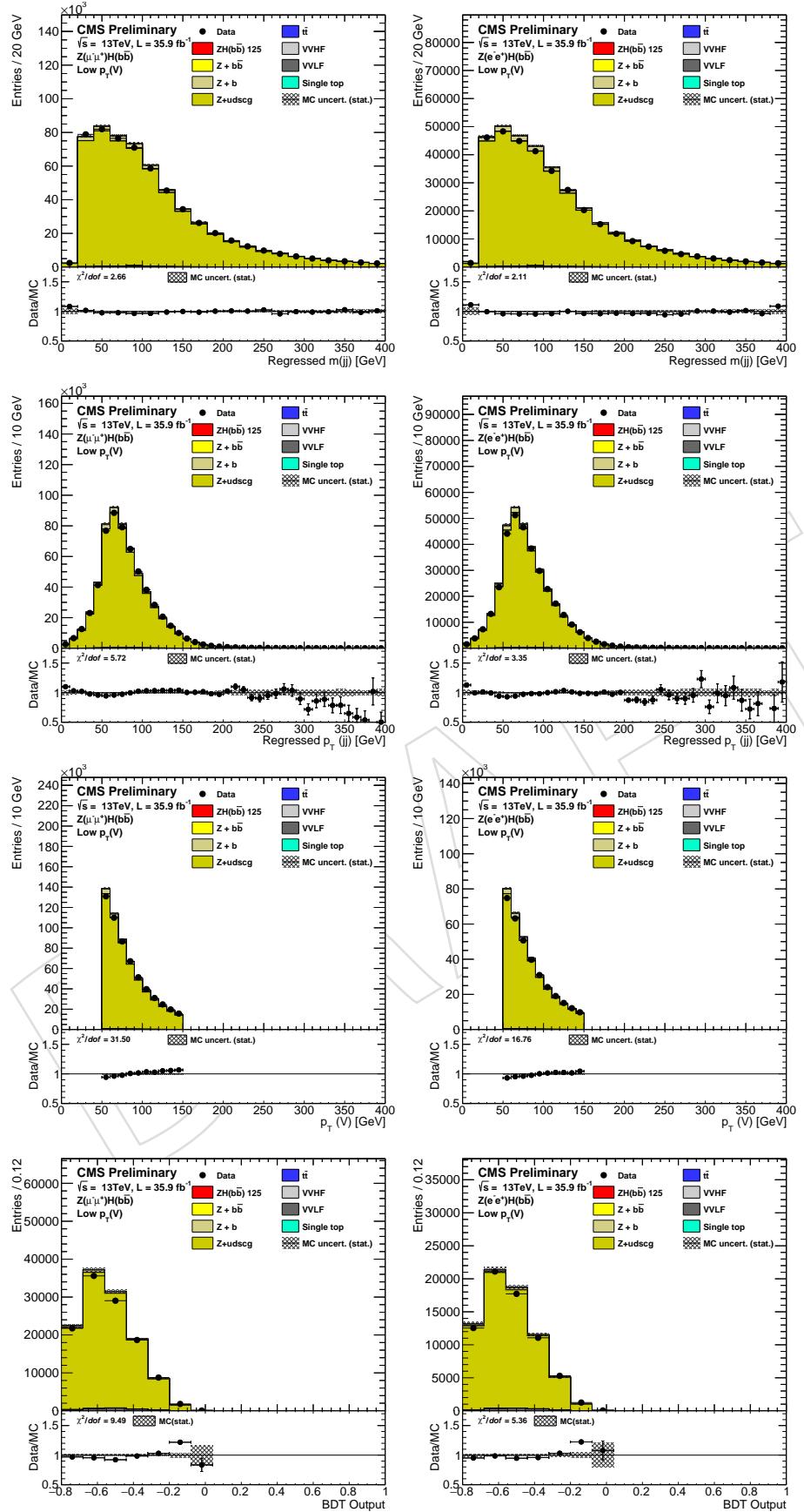


Figure 57: Distributions of variables in data and simulated samples in the $Z(\mu\mu) + \text{udscg}$ control region for $Z(\mu\mu)\text{H}$ (left) and $Z(\text{ee}) + \text{udscg}$ control region for $Z(\text{ee})\text{H}$ (right) in the low V p_{T} bin. From top to bottom: dijet invariant mass, dijet p_{T} , $p_{\text{T}}(Z)$, and BDT output. The plots are normalized using the SFs to facilitate shape comparison.

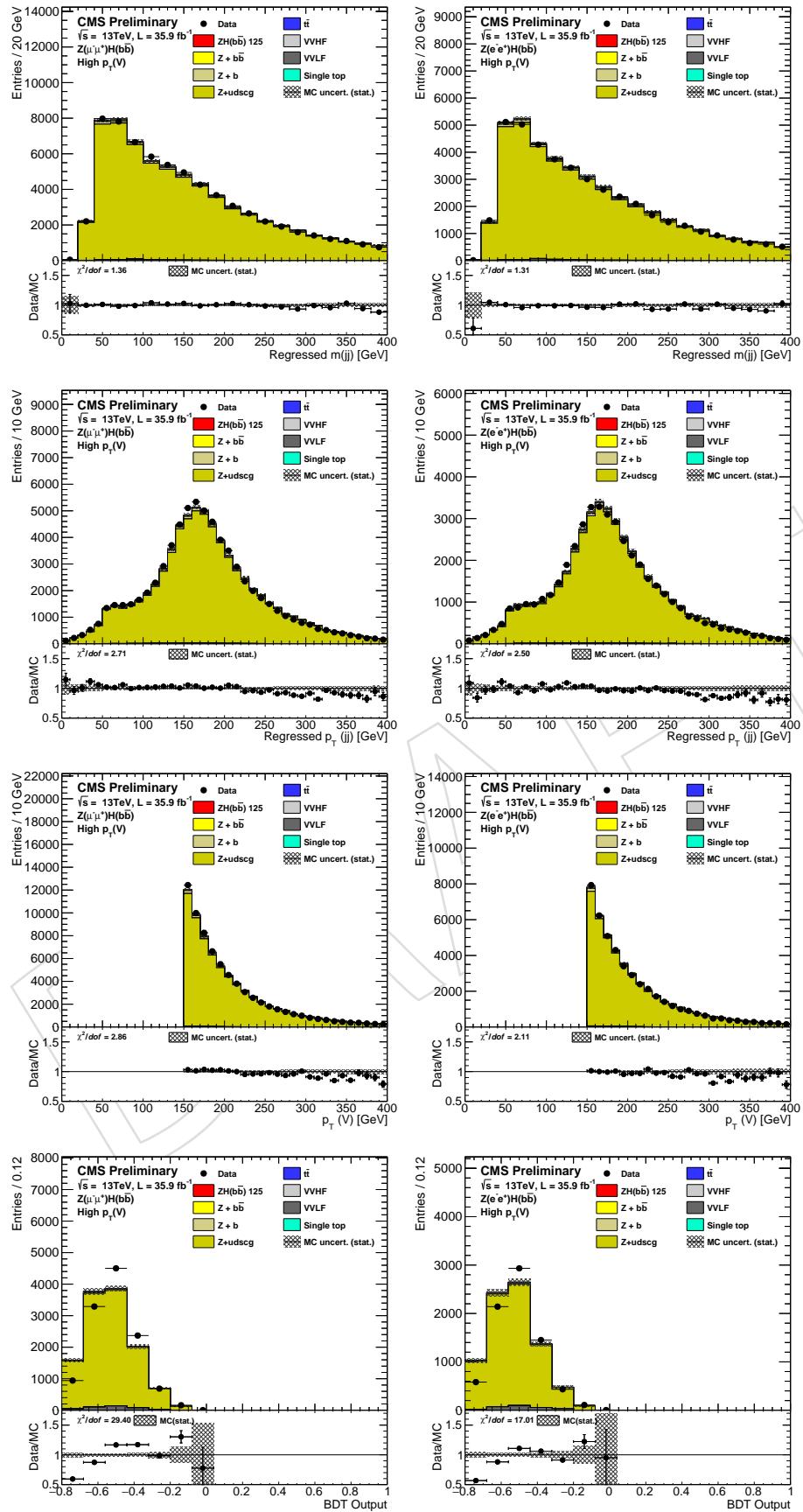


Figure 58: Distributions of variables in data and simulated samples in the $Z(\mu\mu) + \text{udscg}$ control region for $Z(\mu\mu)H$ (left) and $Z(ee) + \text{udscg}$ control region for $Z(ee)H$ (right) in the high V p_T bin. From top to bottom: dijet invariant mass, dijet p_T , $p_T(Z)$, and BDT output. The plots are normalized using the SFs to facilitate shape comparison.

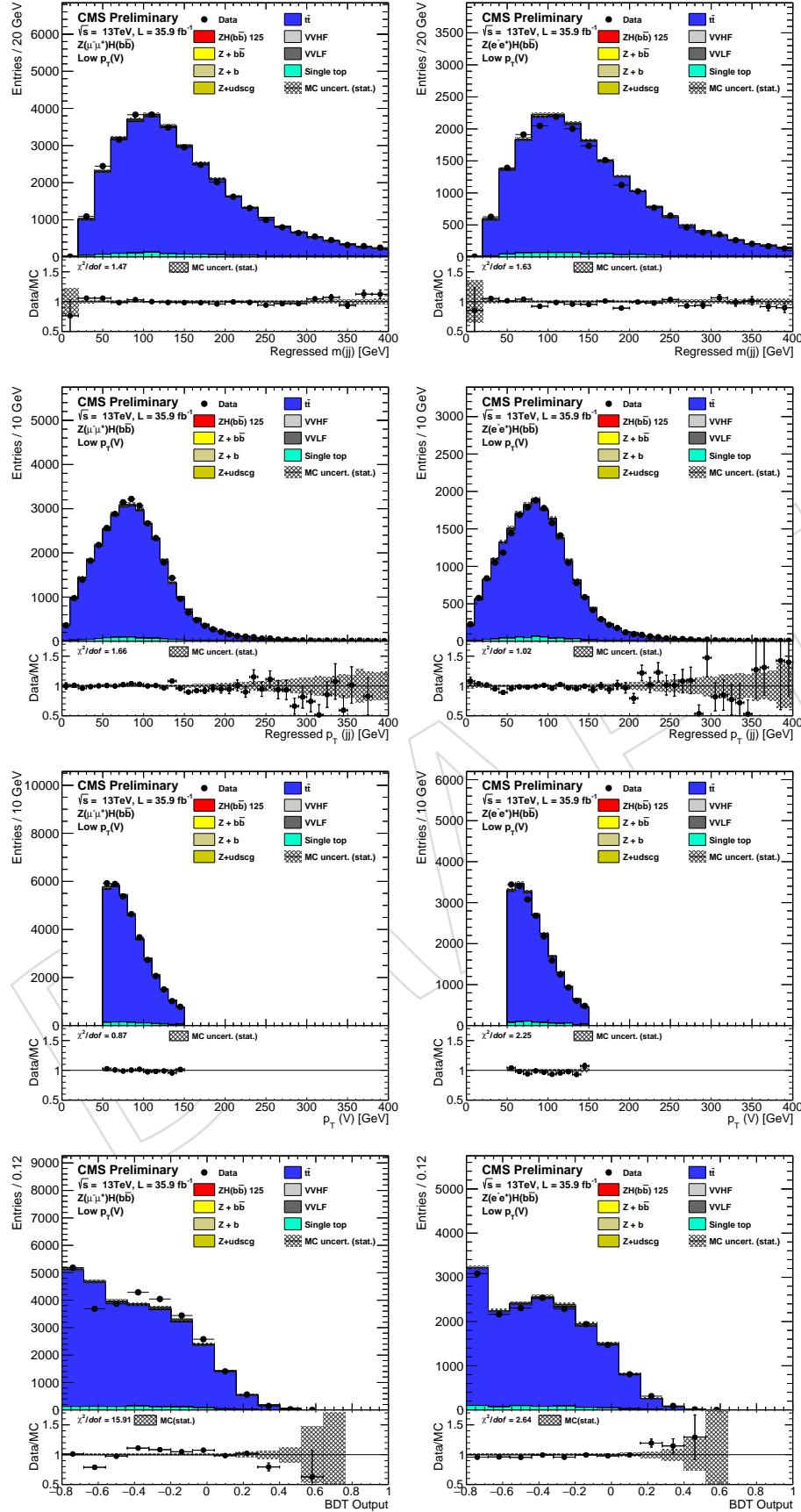


Figure 59: Distributions of variables in data and simulated samples in the $t\bar{t}$ control region for $Z(\mu\mu)H$ (left) and $Z(ee) + udscg$ control region for $Z(ee)H$ (right) in the low $V p_T$ bin. From top to bottom: dijet invariant mass, dijet p_T , $p_T(Z)$, and BDT output. The plots are normalized using the SFs to facilitate shape comparison.

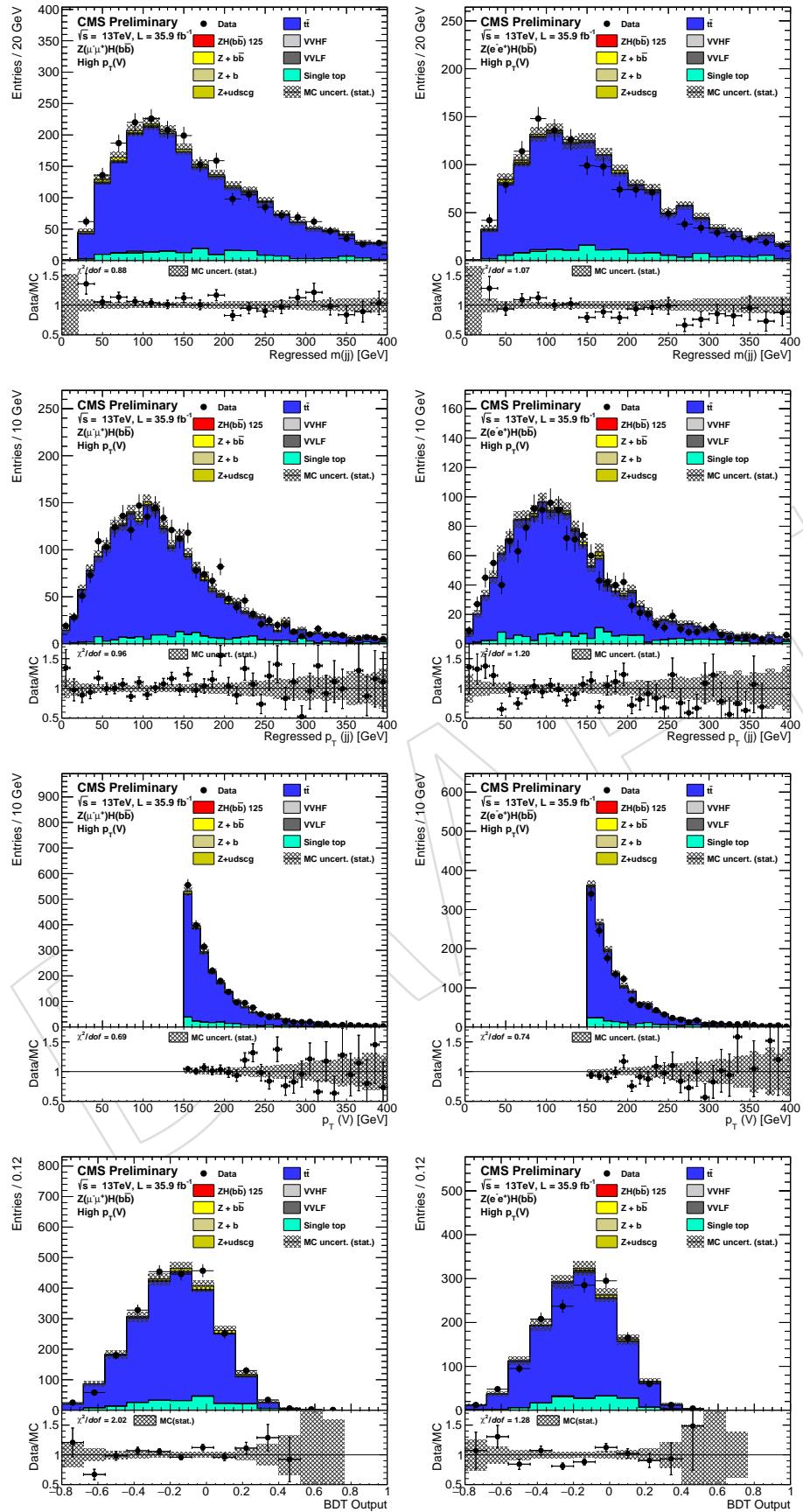


Figure 60: Distributions of variables in data and simulated samples in the $t\bar{t}$ control region for $Z(\mu\mu)H$ (left) and $t\bar{t}$ control region for $Z(ee)H$ (right) in the high $V p_T$ bin. From top to bottom: dijet invariant mass, dijet p_T , $p_T(Z)$, and BDT output. The plots are normalized using the SFs to facilitate shape comparison.

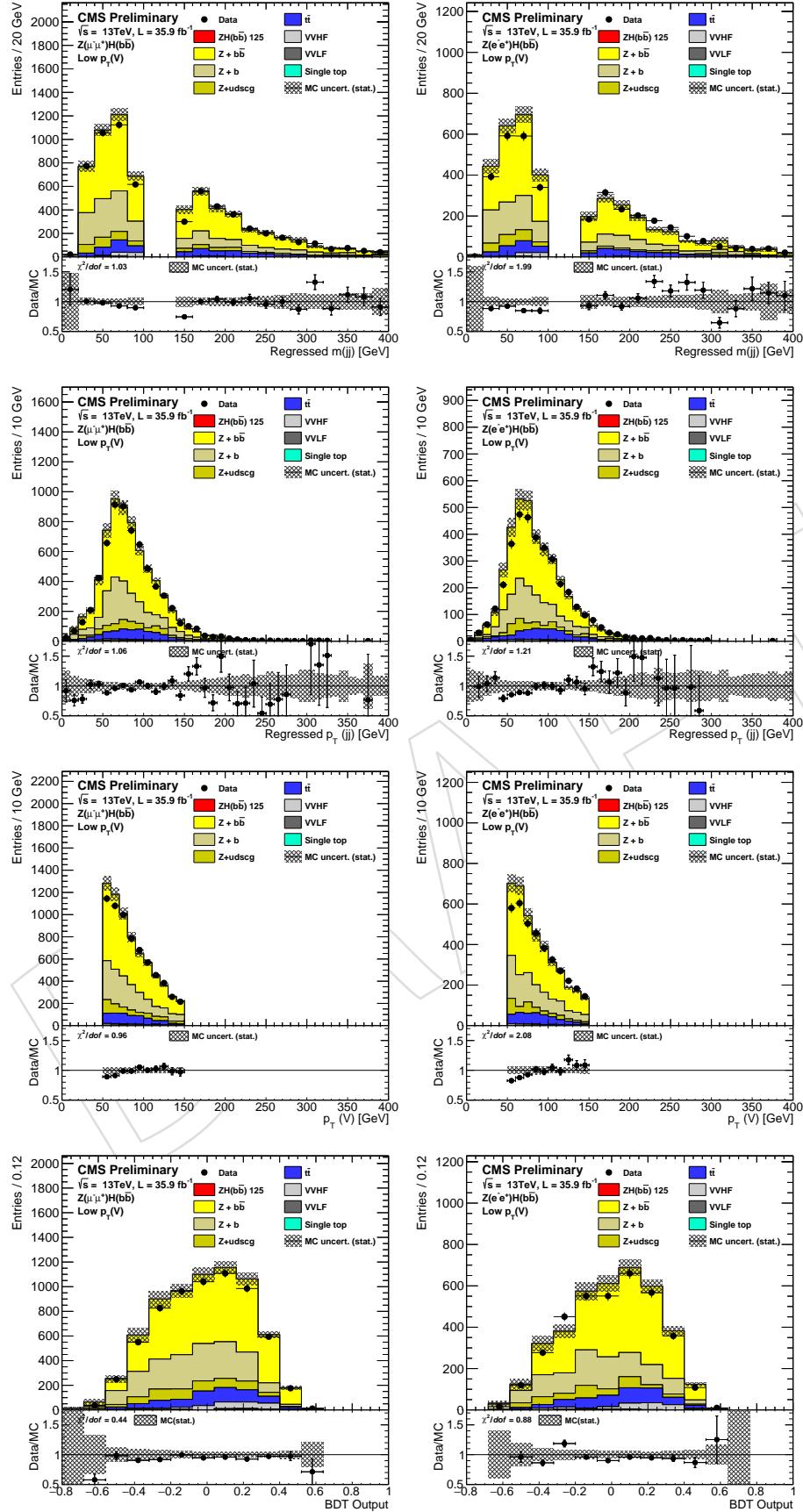


Figure 61: Distributions of variables in data and simulated samples in the $Z(\mu\mu) + b\bar{b}$ control region for $Z(\mu\mu)H$ (left) and $Z(ee) + b\bar{b}$ control region for $Z(ee)H$ (right) in the low $V p_T$ bin. From top to bottom: dijet invariant mass, dijet p_T , $p_T(Z)$, and BDT output. The plots are normalized using the SFs to facilitate shape comparison.

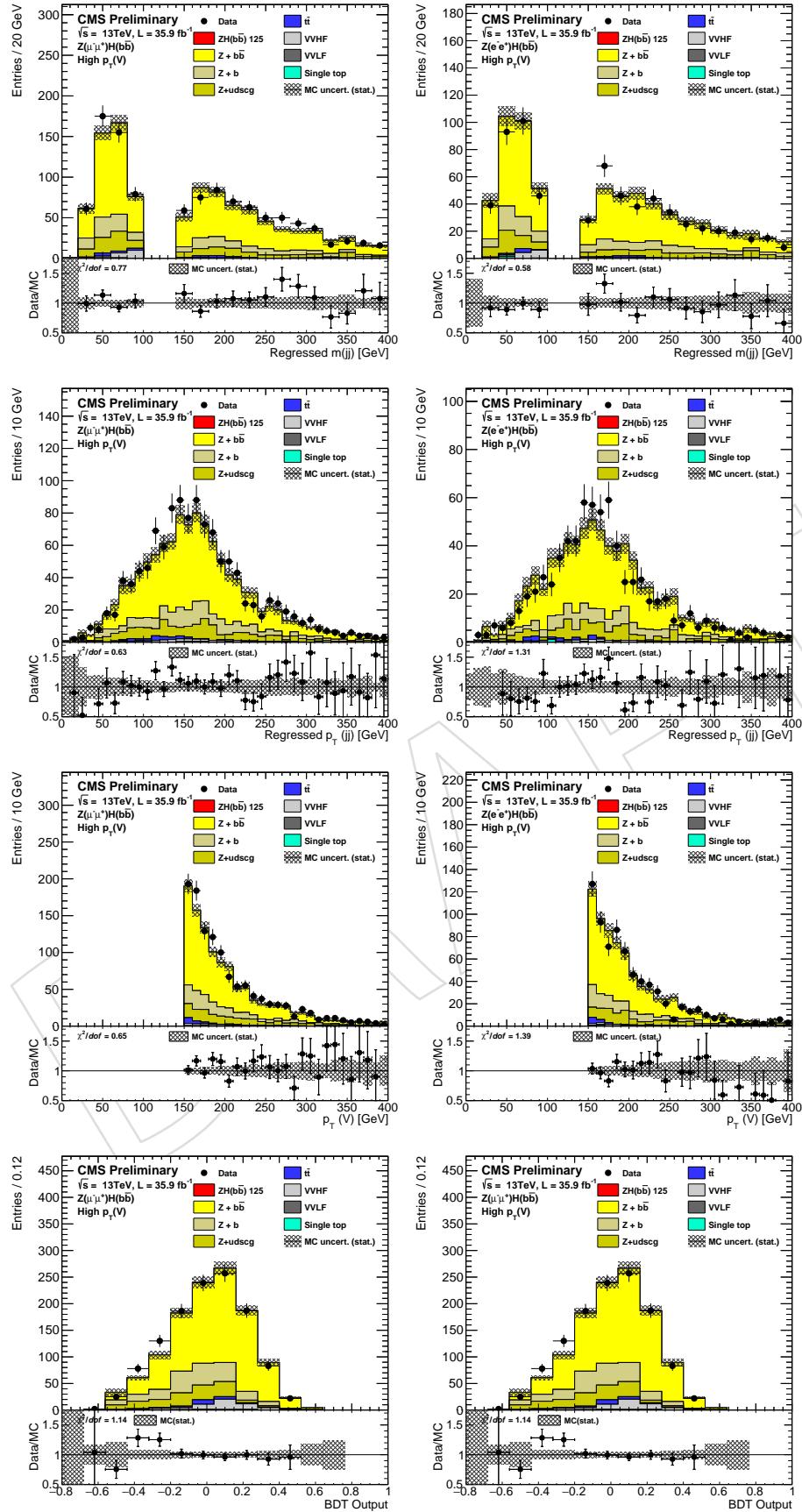


Figure 62: Distributions of variables in data and simulated samples in the $Z(\mu\mu) + b\bar{b}$ control region for $Z(\mu\mu)H$ (left) and $Z(ee) + b\bar{b}$ control region for $Z(ee)H$ (right) in the high $V p_T$ bin. From top to bottom: dijet invariant mass, dijet p_T , $p_T(Z)$, and BDT output. The plots are normalized using the SFs to facilitate shape comparison.

- The $t\bar{t}$ control region is defined by a medium b -tagged jet, at least one isolated lepton and four jets with $p_T > 30 \text{ GeV}$.
- The $W + b\text{-jets}$ control region is independent from the $t\bar{t}$ control region by inverting the four jets cut. In order to reduce the $t\bar{t}$ contamination, exactly one additional lepton is required and a cut of $\Delta\phi(\text{lept.,MET}) < \pi/2$ is used for rejecting fully leptonic $t\bar{t}$ decays. The cut $\Delta R_{jj} \leq \pi/3$ is used to reduce the $t\bar{t}$ contamination as well. An m_{jj} veto is used to remove the possible $WH \rightarrow \ell\nu bb$ contribution into this control region.
- In the $W + \text{light-jets}$ control region, the medium b -tagging cut is inverted and no m_{jj} veto is applied.

11.3 $W(\ell\nu)H$ control regions

This section describes the control regions for $W + \text{udscg}$, $t\bar{t}$, and $Wbb\bar{b}$ production as reconstructed in the $W(\ell\nu)H$ channel. For $W(e\nu)H$ and $W(\mu\nu)H$ Table 17 summarizes the selection criteria used to define samples enriched in the three main backgrounds. For $W + \text{udscg}$ events we loosen the b -tagging requirement to enhance light flavor jets. For $t\bar{t}$ we add one tight b tag and reverse the requirement on number of additional jets. Finally, for $Wbb\bar{b}$ we apply the same criteria on the jet with highest CMVA score as for $t\bar{t}$, but require that there are no additional jets in the event (besides the two b -jets). In addition, we apply a mass veto to remove the signal region (otherwise this sample would be enhanced in signal as well as $Wbb\bar{b}$).

Table 18 summarizes the predicted yields for each background in each control region for $W(e\nu)H$ and $W(\mu\nu)H$. Figures 67–84 show the data/MC agreement in several variables for all three control regions in the 13 TeV data.

Table 16: Definition of control regions for $Z(\nu\nu)H$.

Variable	$t\bar{t}$	Multijet	$Z + \text{light}$	$Z + bb$
V Decay Category	$W(\ell\nu)$	$Z(\nu\nu)$	$Z(\nu\nu)$	$Z(\nu\nu)$
$p_T(j_1)$	> 60	> 60	> 60	> 60
$p_T(j_2)$	> 35	> 35	> 35	> 35
$p_T(jj)$	> 120	> 120	> 120	> 120
E_T^{miss}	> 170	> 170	> 170	> 170
$\Delta\phi(V, H)$	> 2	< 2	> 2	> 2
N_{al}	≥ 1	$= 0$	$= 0$	$= 0$
$N_{\text{jets}}^{\text{central}}$	≥ 4	—	≤ 3	< 3
$M(jj)$	—	—	—	$\notin [60 - 160]$
CMVA2_{\max}	CMVA2M	CMVA2L	!CMVA2M	CMVA2T
CMVA2_{\min}	CMVA2L	CMVA2L	CMVA2L	CMVA2L
$\Delta\phi(j, E_T^{\text{miss}})$	—	< 0.4	> 0.5	> 0.5
$\Delta\phi(\text{tkMET}, E_T^{\text{miss}})$	—	—	< 0.5	< 0.5
$\min \Delta\phi(j, E_T^{\text{miss}})$	$< \pi/2$	—	—	—

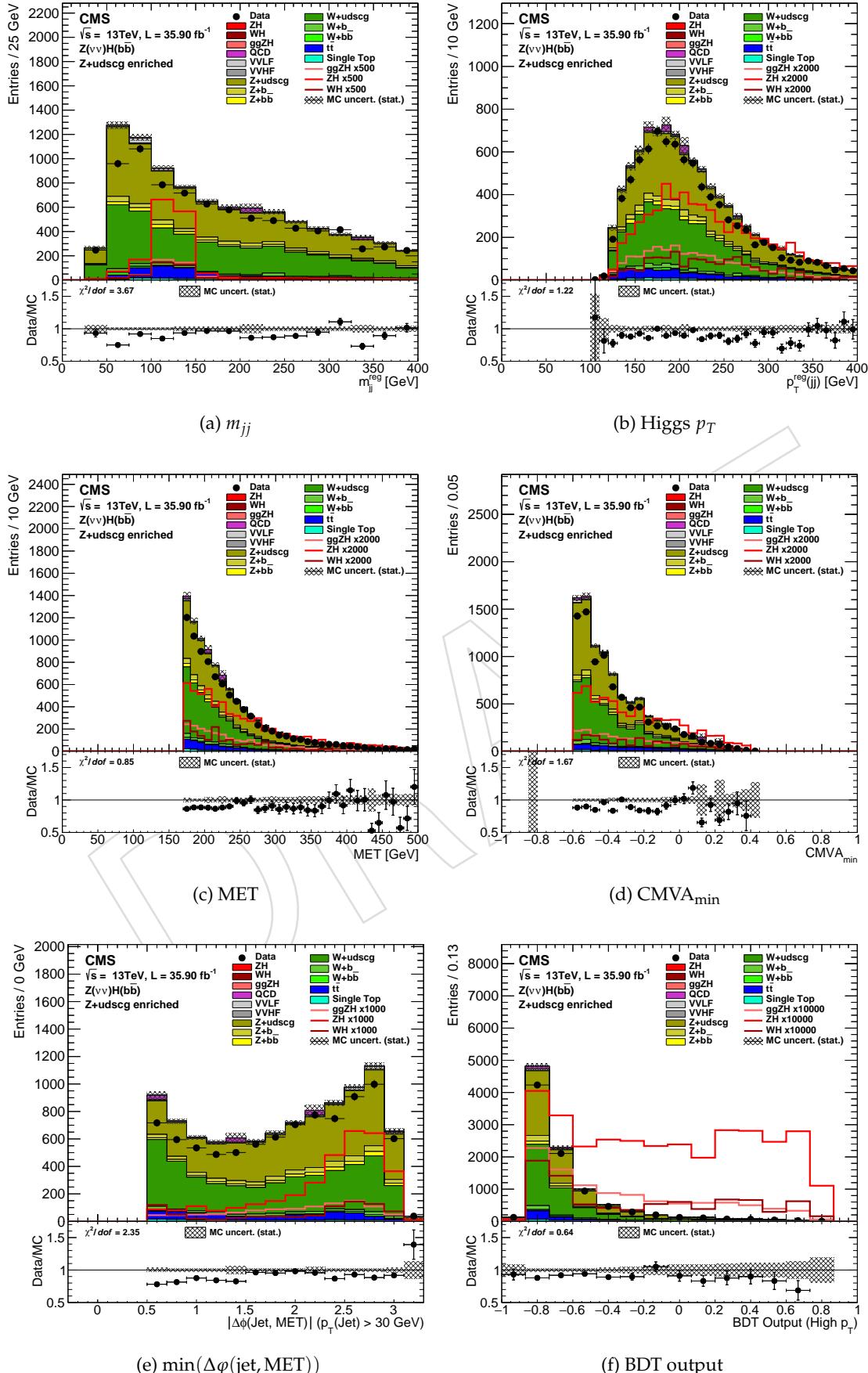


Figure 63: Distributions of some variables for the $Z + \text{light-jets}$ control region in the $Z(\nu\nu)H$ analysis. The plots use the preliminary scale factors reported in Sect. 11.4.

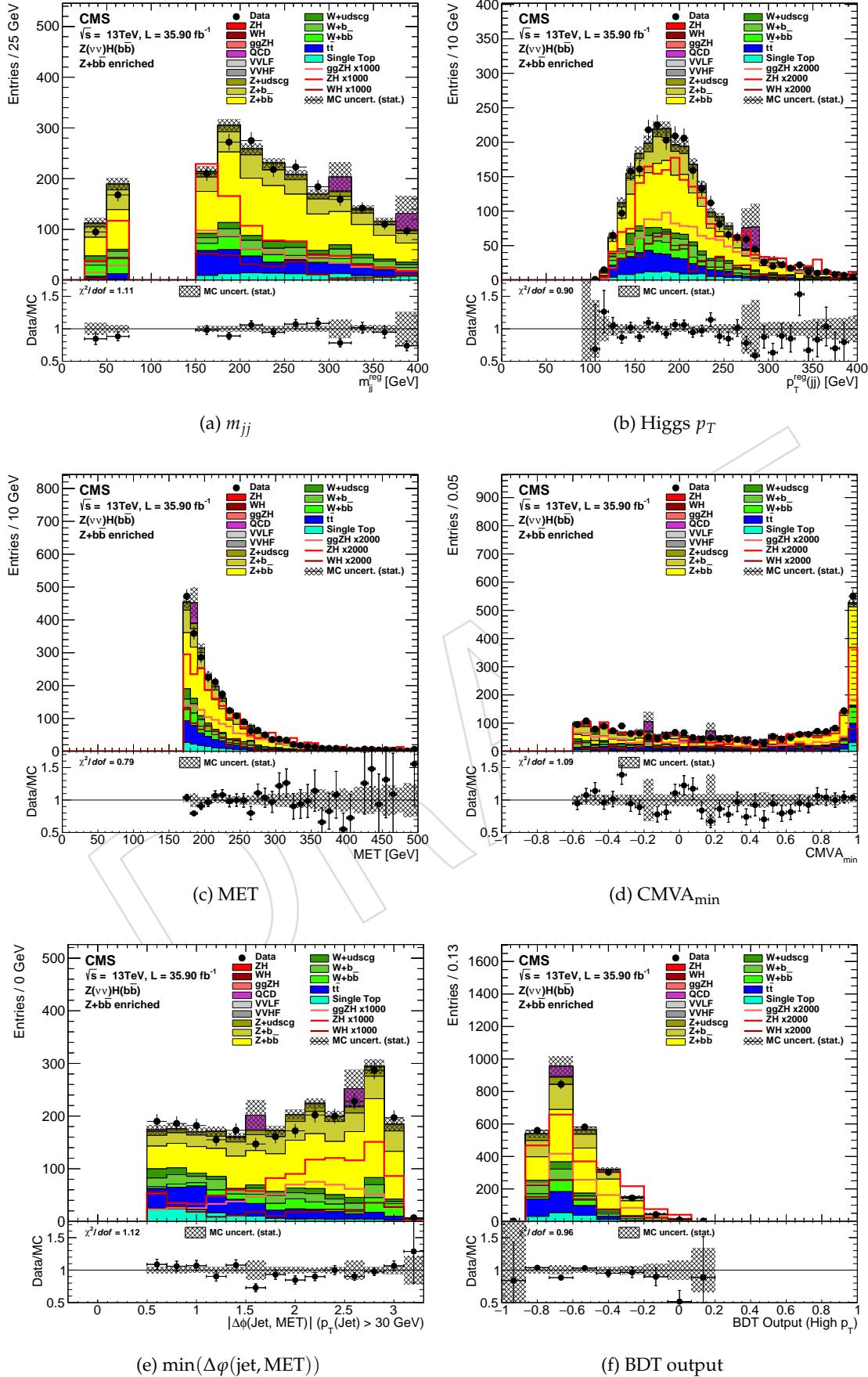


Figure 64: Distributions of some variables for the $Z+b$ -jets control region in the $Z(\nu\nu)$ analysis. The plots use the preliminary scale factors reported in Sect. 11.4.

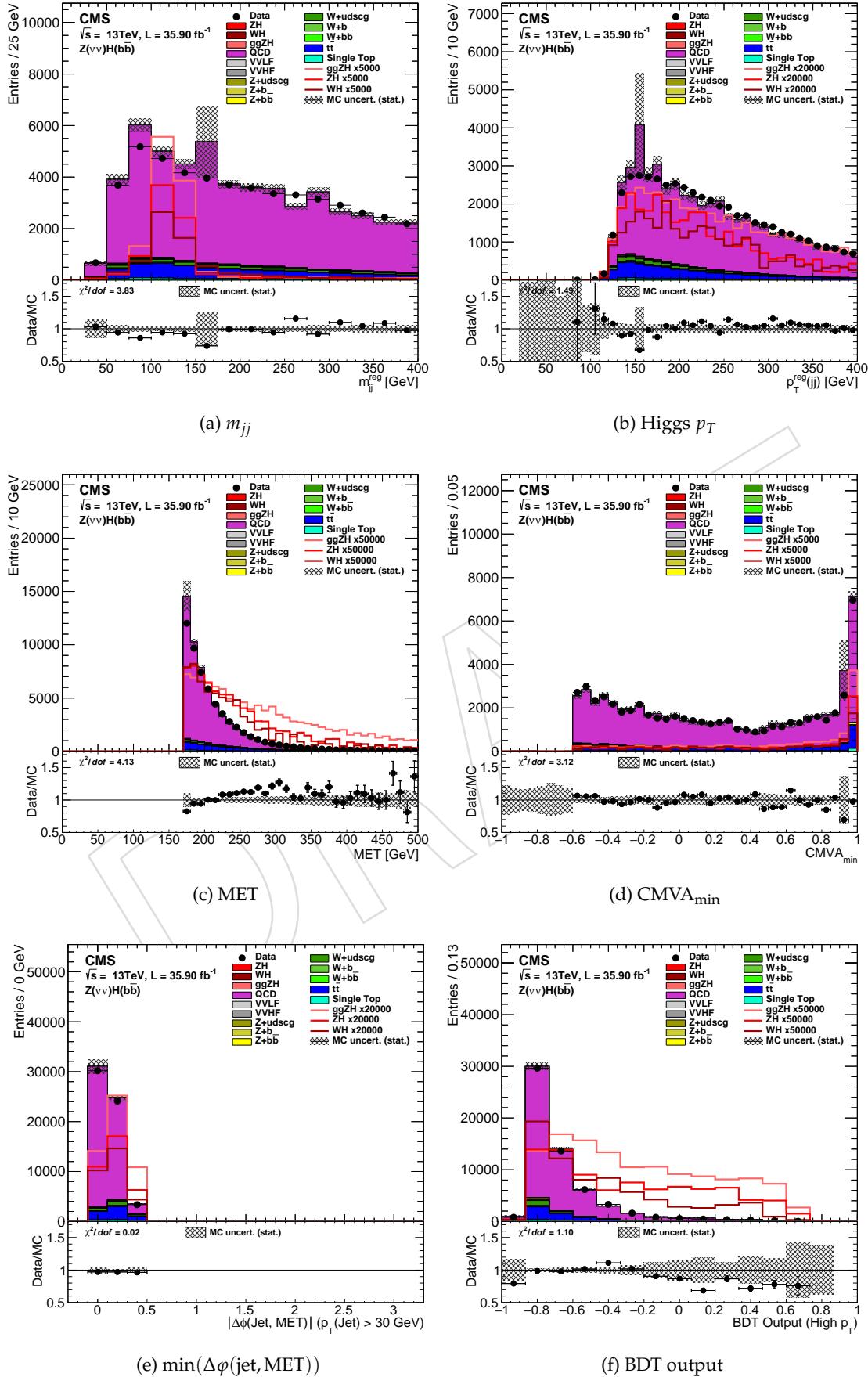


Figure 65: Distributions of some variables for the multijet control region in the $Z(vv)H$ analysis. The plots use the preliminary scale factors reported in Sect. 11.4.

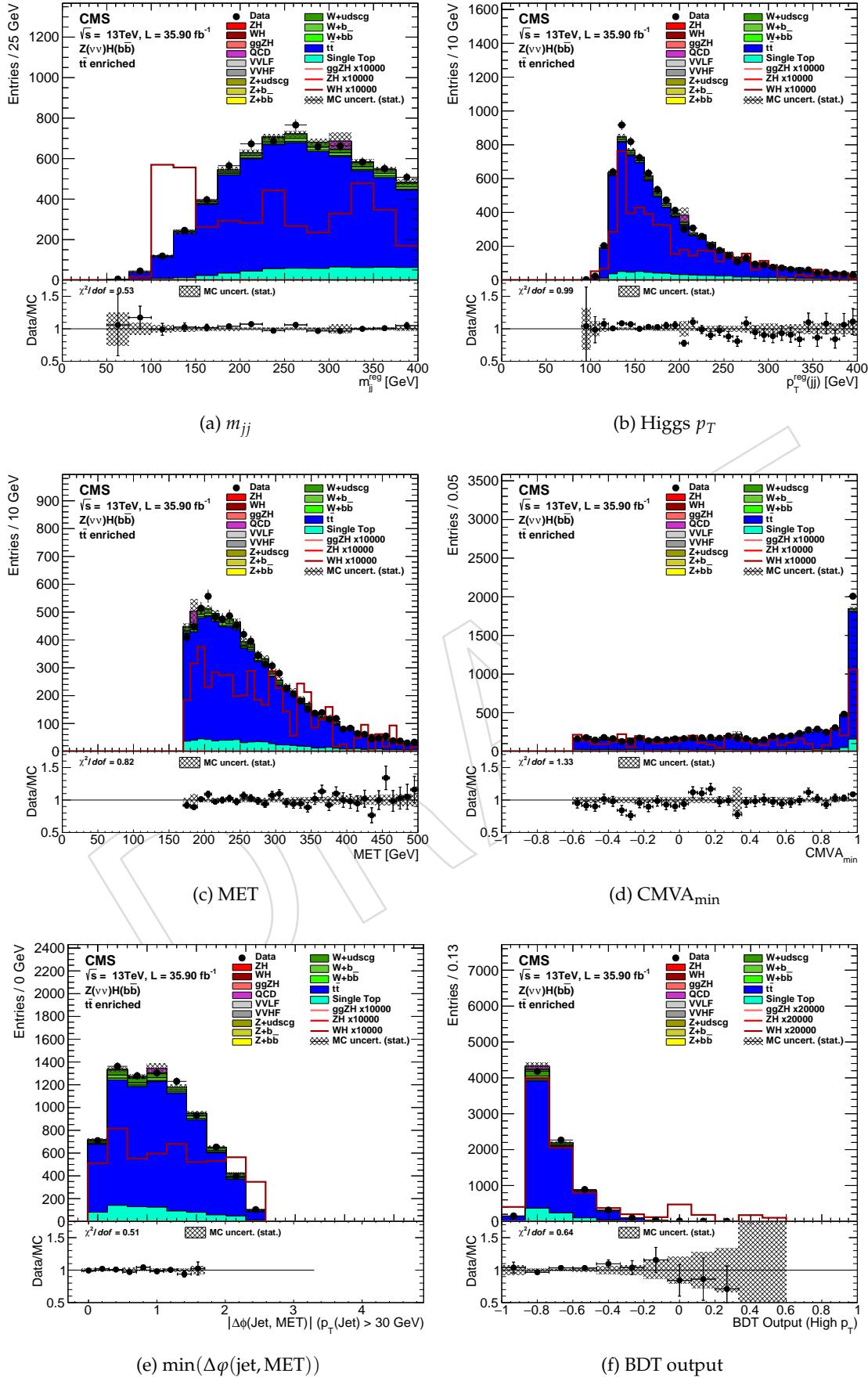


Figure 66: Distributions of some variables for the $t\bar{t}$ -enriched control region in the $Z(\nu\nu)H$ analysis. The plots use the preliminary scale factors reported in Sect. 11.4.

Table 17: Definition of control regions for the $W(e\nu)H$ and $W(\mu\nu)H$ channels for the low, intermediate, and high $p_T(V)$ regions. The values in parenthesis are used for the intermediate and high $p_T(V)$ regions. LF and HF refer to light- and heavy-flavor jets. N_{al} is the number of additional isolated leptons in the event. METsig is the significance of the E_T^{miss} . The values listed for kinematical variables are in units of GeV.

Variable	$W+LF$	$t\bar{t}$	$W+HF$
$p_T(j_1)$	> 25	> 25	> 25
$p_T(j_2)$	> 25	> 25	> 25
$p_T(jj)$	> 100	> 100	> 100
$p_T(V)$	> 100	> 100	> 100
CMVA _{max}	$[-0.5884 - 0.4432]$	> 0.9432	> 0.9432
N_{aj}	-	> 1	= 0
N_{al}	= 0	= 0	= 0
METsig	> 2.0	-	> 2.0
$\Delta\phi(\text{pfMET}, \text{lep})$	< 2	< 2	< 2
$M(jj)$	< 250	< 250	< 250, veto [90 – 150]

Table 18: 13 TeV predicted yield for each background reconstructed as $W(e\nu)H$ and $W(\mu\nu)H$ obtained in `Summer16` Monte Carlo samples and scaled to the appropriate data luminosity (see Table 1); total MC yield; and observed yield in data. Lepton trigger efficiency, pile-up, and btag reweighting have been applied to all MC predictions. The yields are extracted after a simultaneous MaximumLikelihoodFit of all $W(\ell\nu)H$ control regions. All uncertainties include both a statistical and a systematic component.

Process	W+light		$t\bar{t}$		W+heavy	
	$W(\mu\nu)H$	$W(e\nu)H$	$W(\mu\nu)H$	$W(e\nu)H$	$W(\mu\nu)H$	$W(e\nu)H$
Zj0b	13706.97 ± 117.08	8192.51 ± 90.51	189.31 ± 13.76	133.00 ± 11.53	164.61 ± 12.83	84.92 ± 9.22
Zj1b	688.83 ± 26.25	473.53 ± 21.76	450.49 ± 21.22	367.05 ± 19.16	502.72 ± 22.42	300.20 ± 17.33
Zj2b	142.46 ± 11.94	79.95 ± 8.94	450.85 ± 21.23	378.23 ± 19.45	209.81 ± 14.48	107.13 ± 10.35
Wj0b	272703.07 ± 522.21	173840.44 ± 416.94	2759.15 ± 52.53	1838.62 ± 42.88	4161.12 ± 64.51	2575.18 ± 50.75
Wj1b	9670.86 ± 98.34	6182.08 ± 78.63	5394.61 ± 73.45	3494.13 ± 59.11	5740.30 ± 75.76	3590.50 ± 59.92
Wj2b	1264.86 ± 35.56	804.28 ± 28.36	3071.95 ± 55.43	1980.86 ± 44.51	1607.39 ± 40.09	1145.21 ± 33.84
TT	42834.20 ± 206.96	30066.83 ± 173.40	122255.68 ± 349.65	88740.43 ± 297.89	11864.72 ± 108.93	8098.13 ± 89.99
ST	13938.12 ± 118.06	9665.17 ± 98.31	16349.61 ± 127.87	11880.95 ± 109.00	7318.26 ± 85.55	4954.34 ± 70.39
VVHF	151.90 ± 12.32	99.20 ± 9.96	256.36 ± 16.01	181.33 ± 13.47	132.79 ± 11.52	84.61 ± 9.20
VVLF	5590.51 ± 74.77	3672.98 ± 60.61	245.08 ± 15.66	180.12 ± 13.42	90.53 ± 9.51	49.05 ± 7.00
Total Background	360691.77 ± 600.58	233076.98 ± 482.78	151423.10 ± 389.13	109174.71 ± 330.42	31792.24 ± 178.30	20989.27 ± 144.88
Data	359285	234539	152209	108318	32320	20651

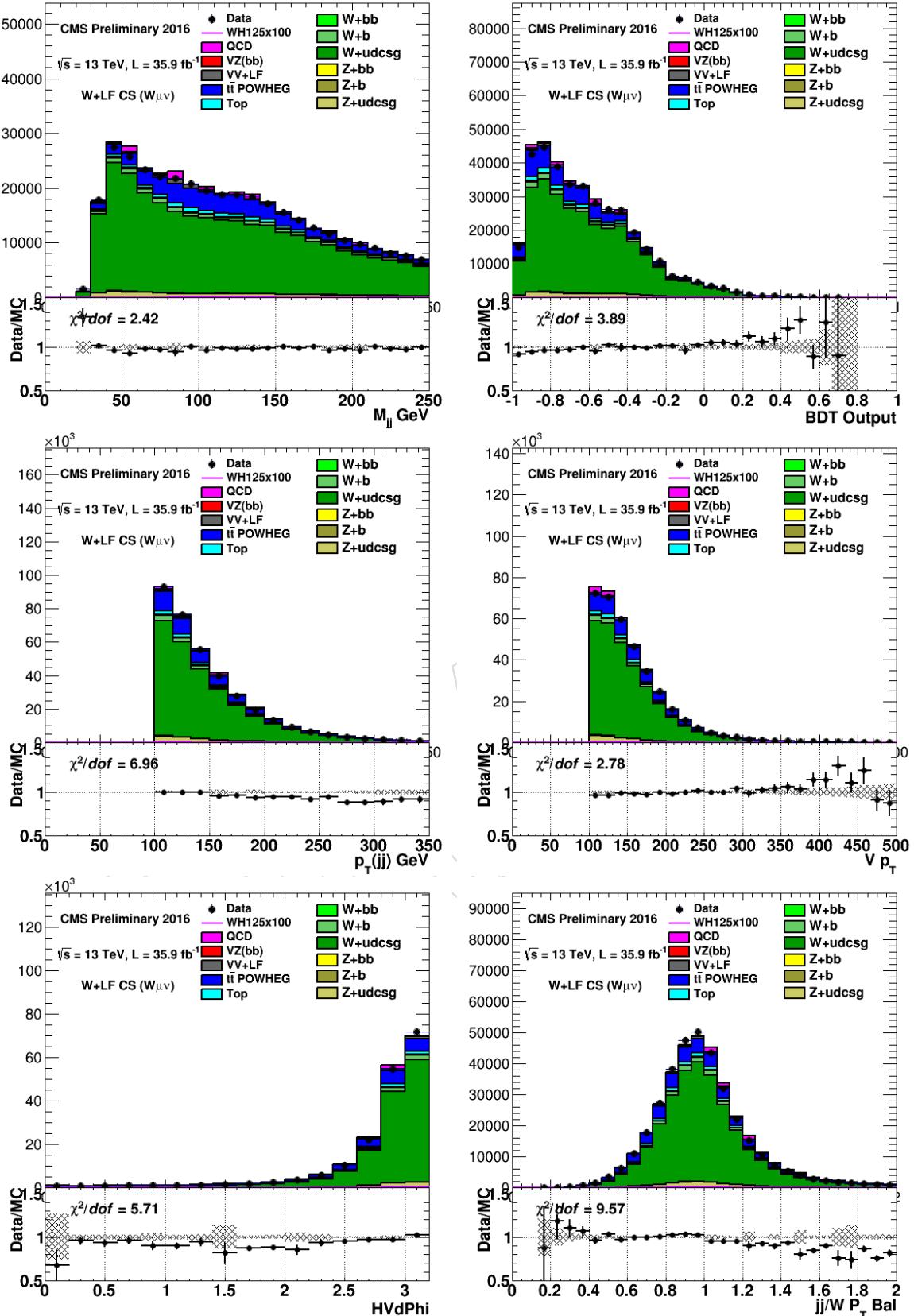


Figure 67: Distributions of data and simulated samples in the $W(\mu\nu) + \text{udcsg}$ control region. Left to right and top to bottom: dijet invariant mass, BDT output, $p_T(jj)$ of the dijet, $p_T(W)$ of the reconstructed W boson, azimuthal angle between the W and dijet, and the ratio of $p_T(jj)/p_T(W)$. To facilitate the shape comparison, the Scale Factors are applied to the simulation.

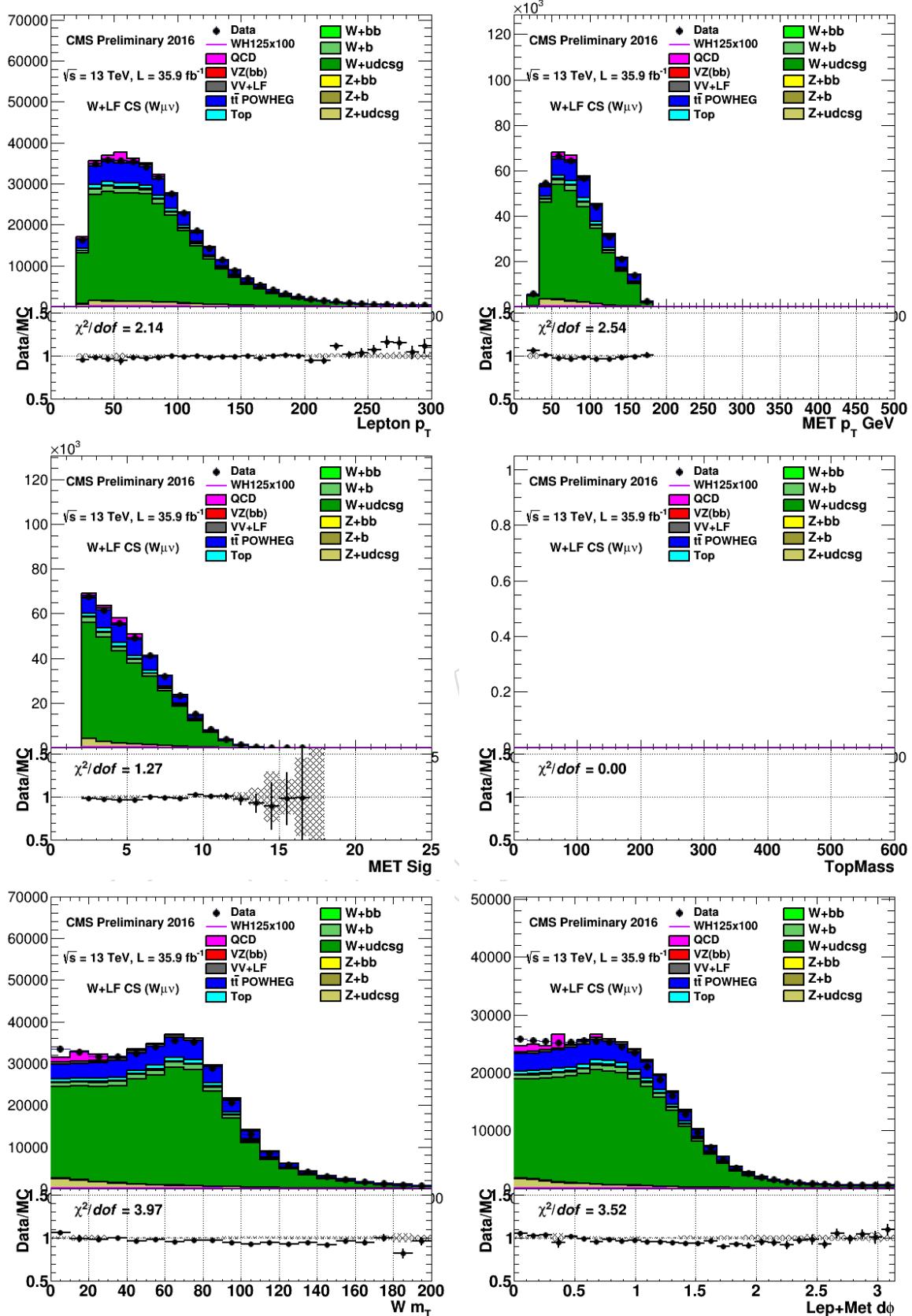


Figure 68: Distributions of data and simulated samples in the $W(\mu\nu) + \text{udscg}$ control region. Left to right and top to bottom: electron p_T , pfMET, pfMET significance, reconstructed top mass, the transverse mass of the reconstructed W boson, and the azimuthal angle between the lepton and missing energy. To facilitate the shape comparison, the Scale Factors are applied to the simulation.

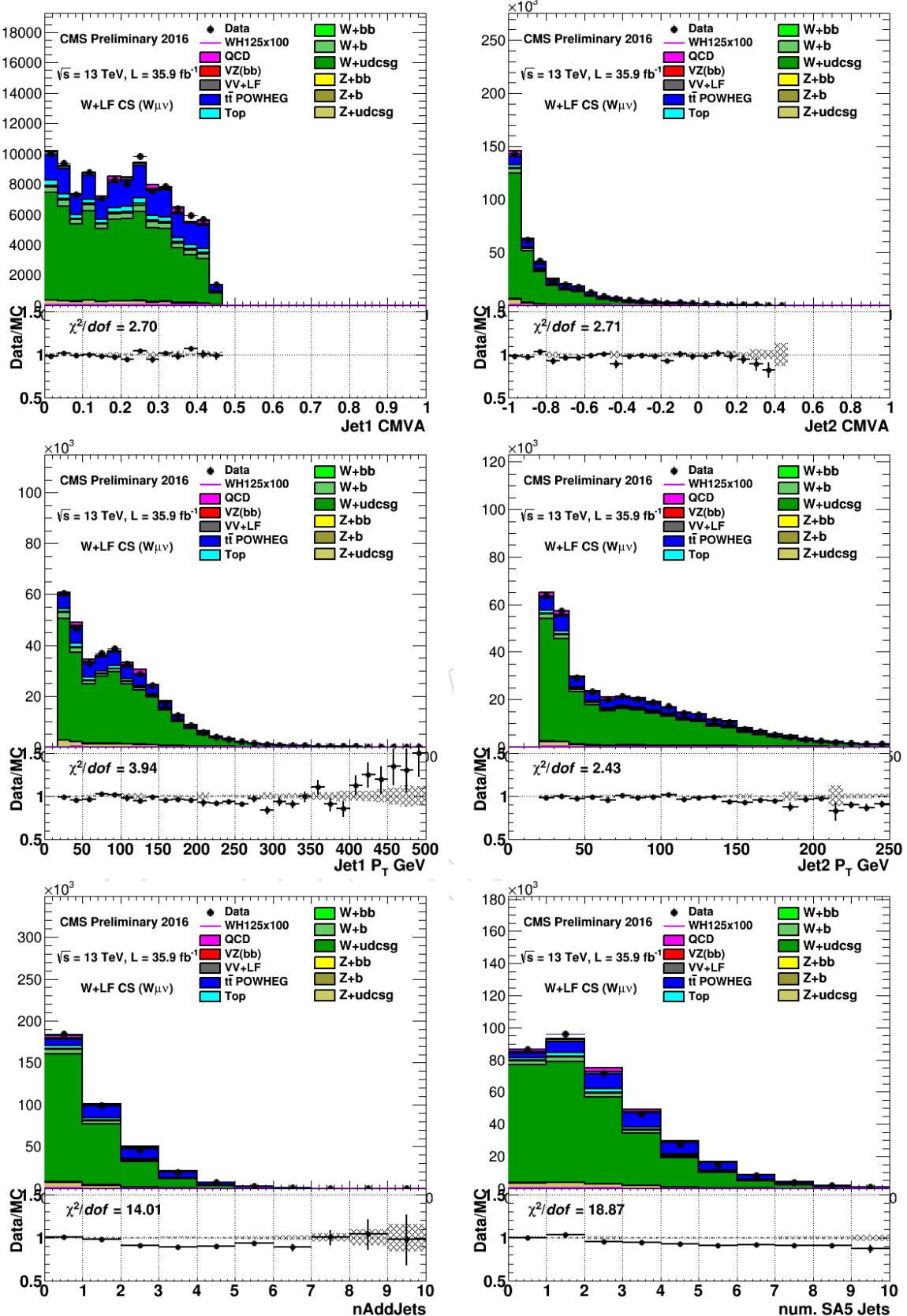


Figure 69: Distributions of data and simulated samples in the $W(\mu\nu) + \text{udcsg}$ control region. Left to right and top to bottom: the output of the CMVA for the highest and next highest scoring b-jets, the p_T distributions of those jets, the number of additional jets and the number of additional soft activity jets with $p_T > 5$ GeV. To facilitate the shape comparison, the Scale Factors are applied to the simulation.

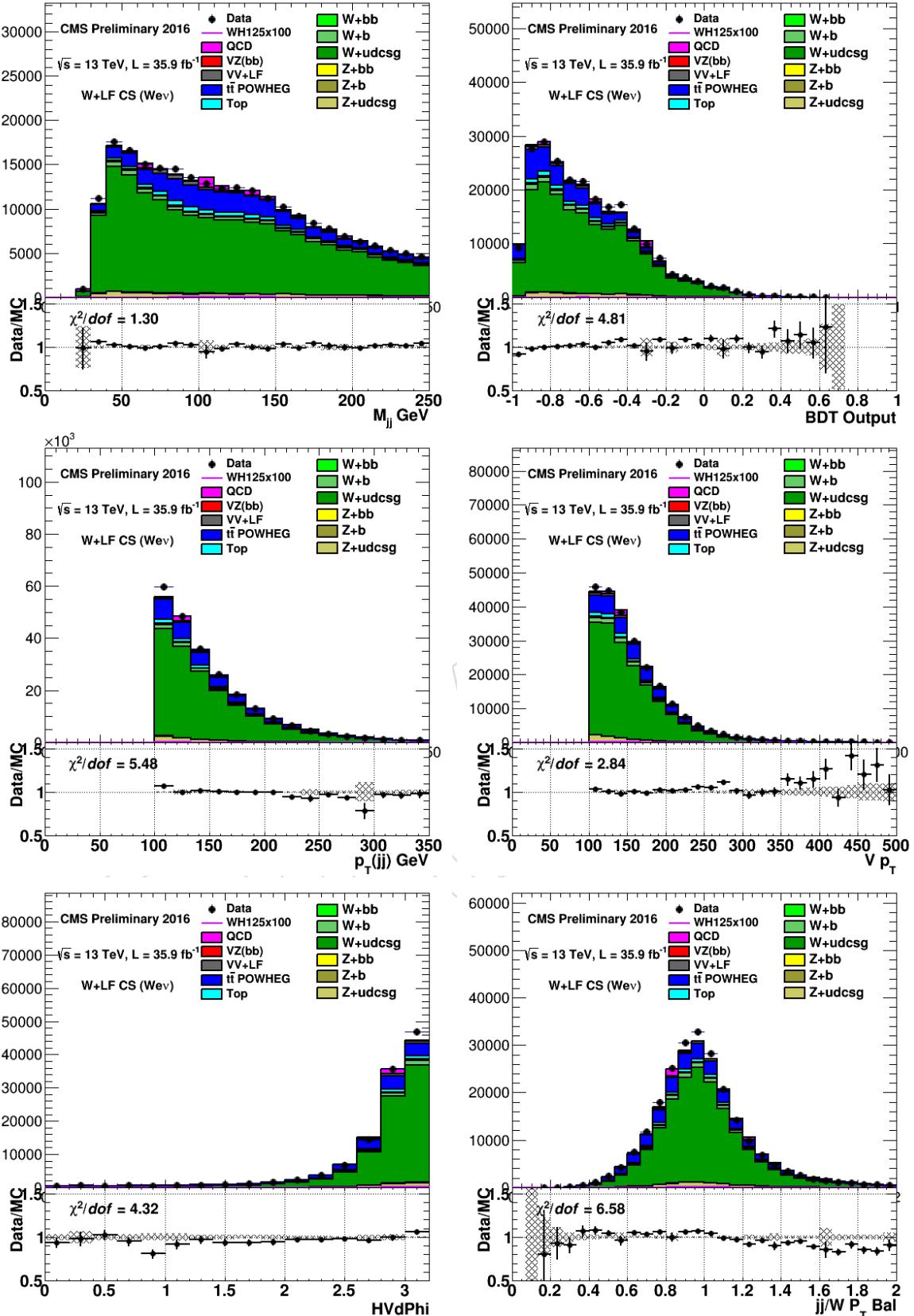


Figure 70: Distributions of data and simulated samples in the $W(e\nu) + \text{udscg}$ control region. Left to right and top to bottom: dijet invariant mass, BDT output, $p_T(\text{jj})$ of the dijet, $p_T(W)$ of the reconstructed W boson, azimuthal angle between the W and dijet, and the ratio of $p_T(\text{jj})/p_T(W)$. To facilitate the shape comparison, the Scale Factors are applied to the simulation.

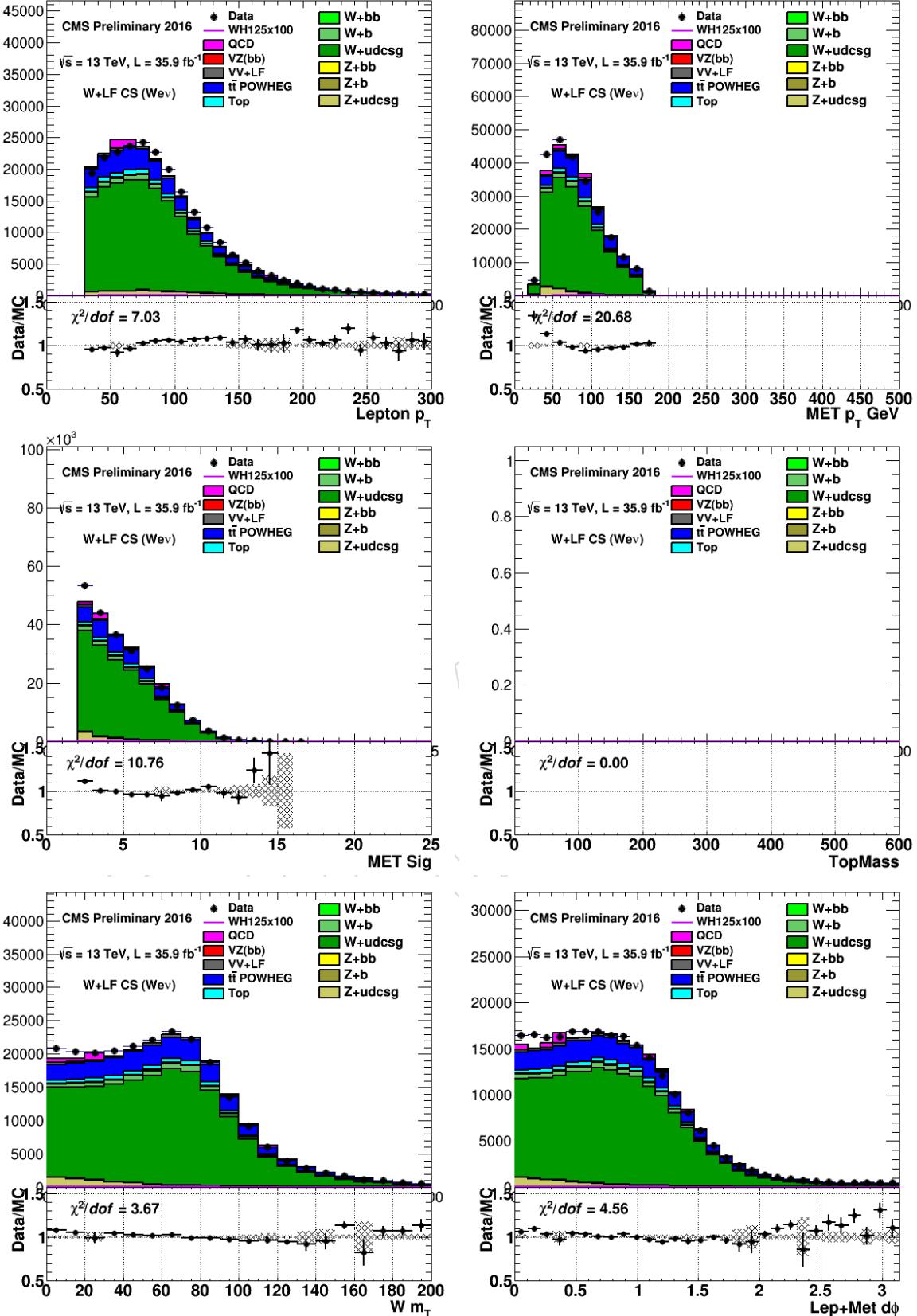


Figure 71: Distributions of data and simulated samples in the $W(e\nu) + udscg$ control region. Left to right and top to bottom: electron p_T , pfMET, pfMET significance, reconstructed top mass, the transverse mass of the reconstructed W boson, and the azimuthal angle between the lepton and missing energy. To facilitate the shape comparison, the Scale Factors are applied to the simulation.

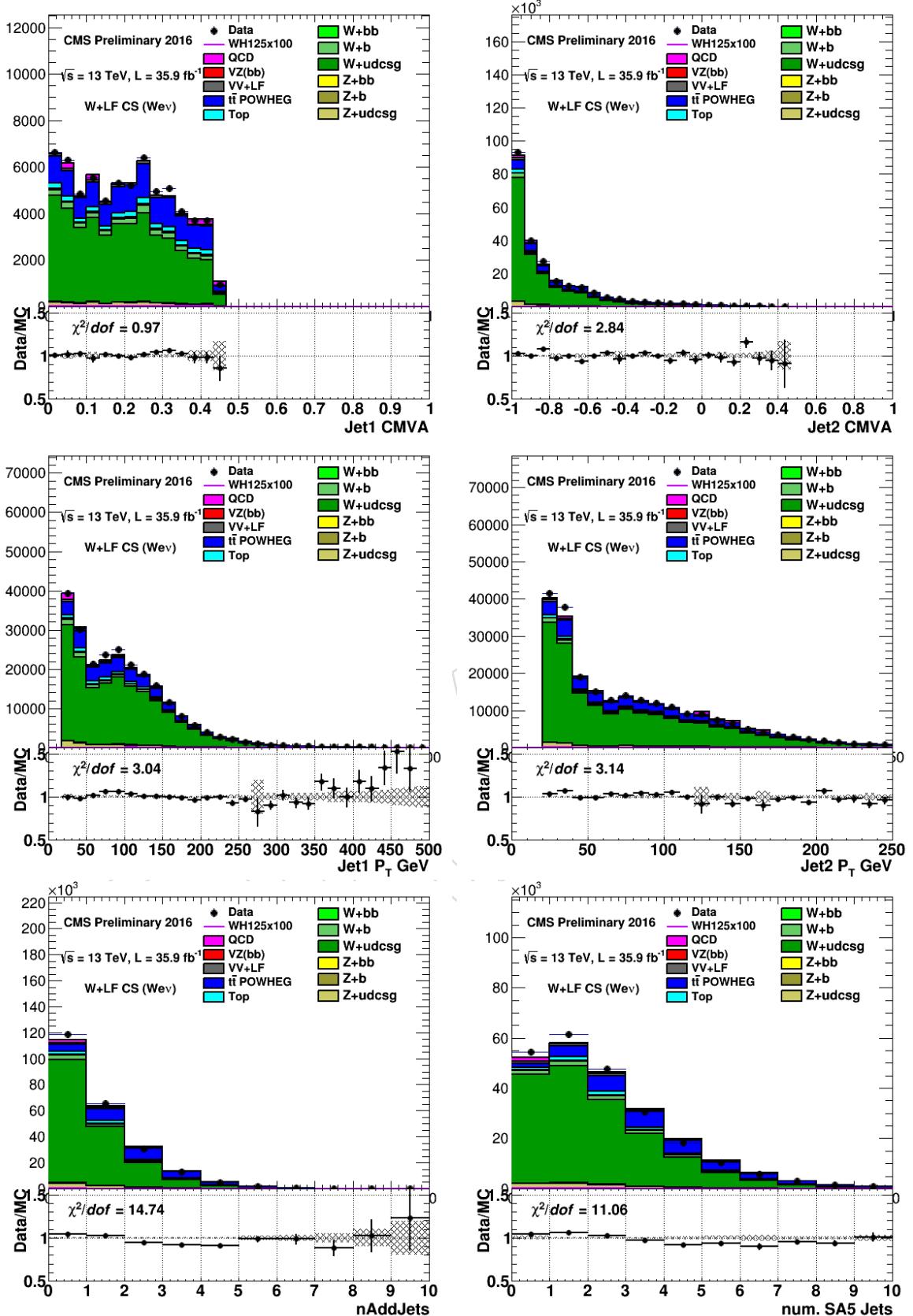


Figure 72: Distributions of data and simulated samples in the $W(e\nu) + \text{udscg}$ control region. Left to right and top to bottom: the output of the CMVA for the highest and next highest scoring b-jets, the p_T distributions of those jets, the number of additional jets and the number of additional soft activity jets with $p_T > 5$ GeV. To facilitate the shape comparison, the Scale Factors are applied to the simulation.

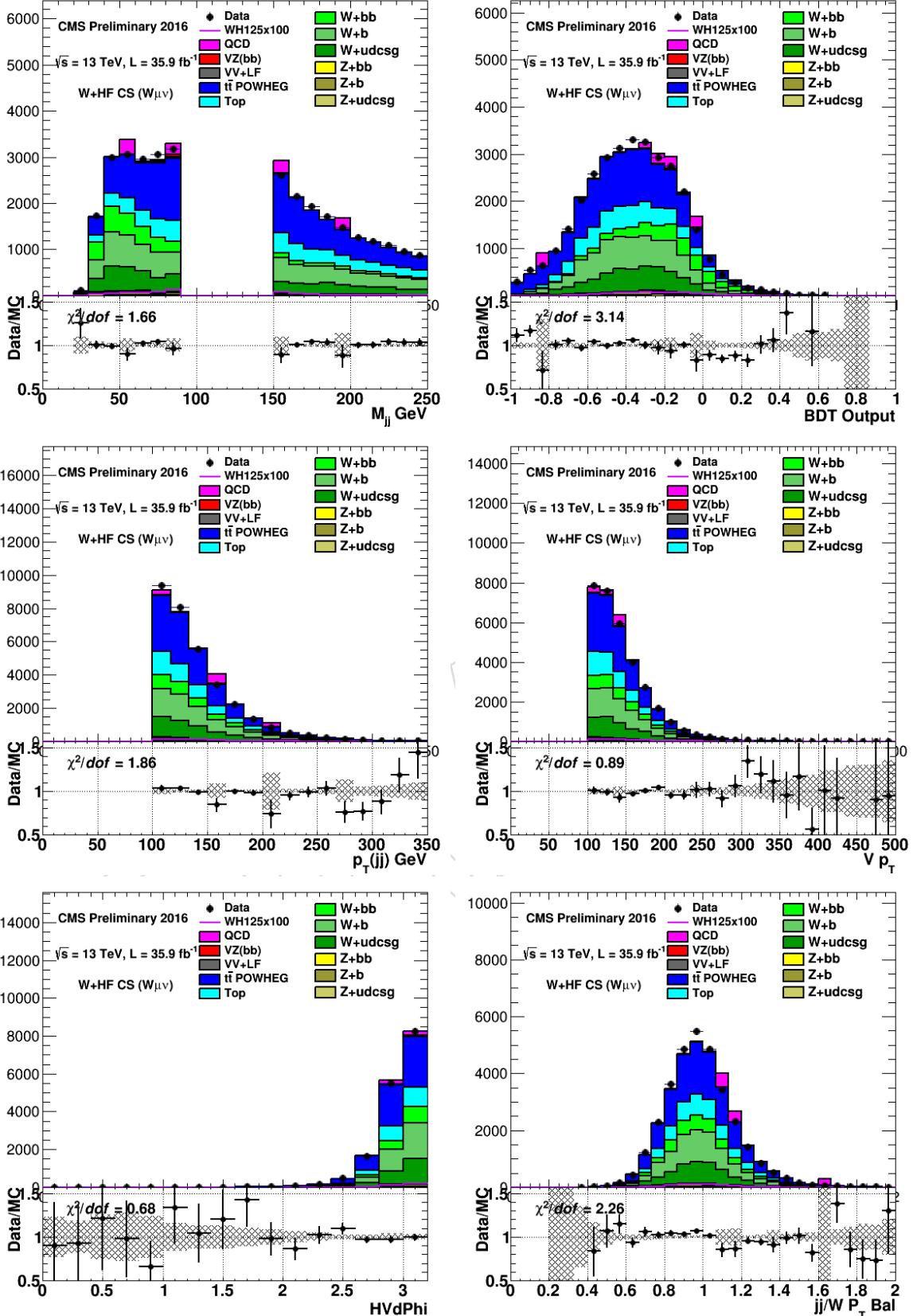


Figure 73: Distributions of data and simulated samples in the $W(\mu\nu)H\bar{W}b\bar{b}$ control region. Left to right and top to bottom: dijet invariant mass, BDT output, $p_T(jj)$ of the dijet, $p_T(W)$ of the reconstructed W boson, azimuthal angle between the W and dijet, and the ratio of $p_T(jj)/p_T(W)$. To facilitate the shape The simulation is adjusted by the scale factors in Table 19, but the overall normalization between data and simulation is not forced to be the same.

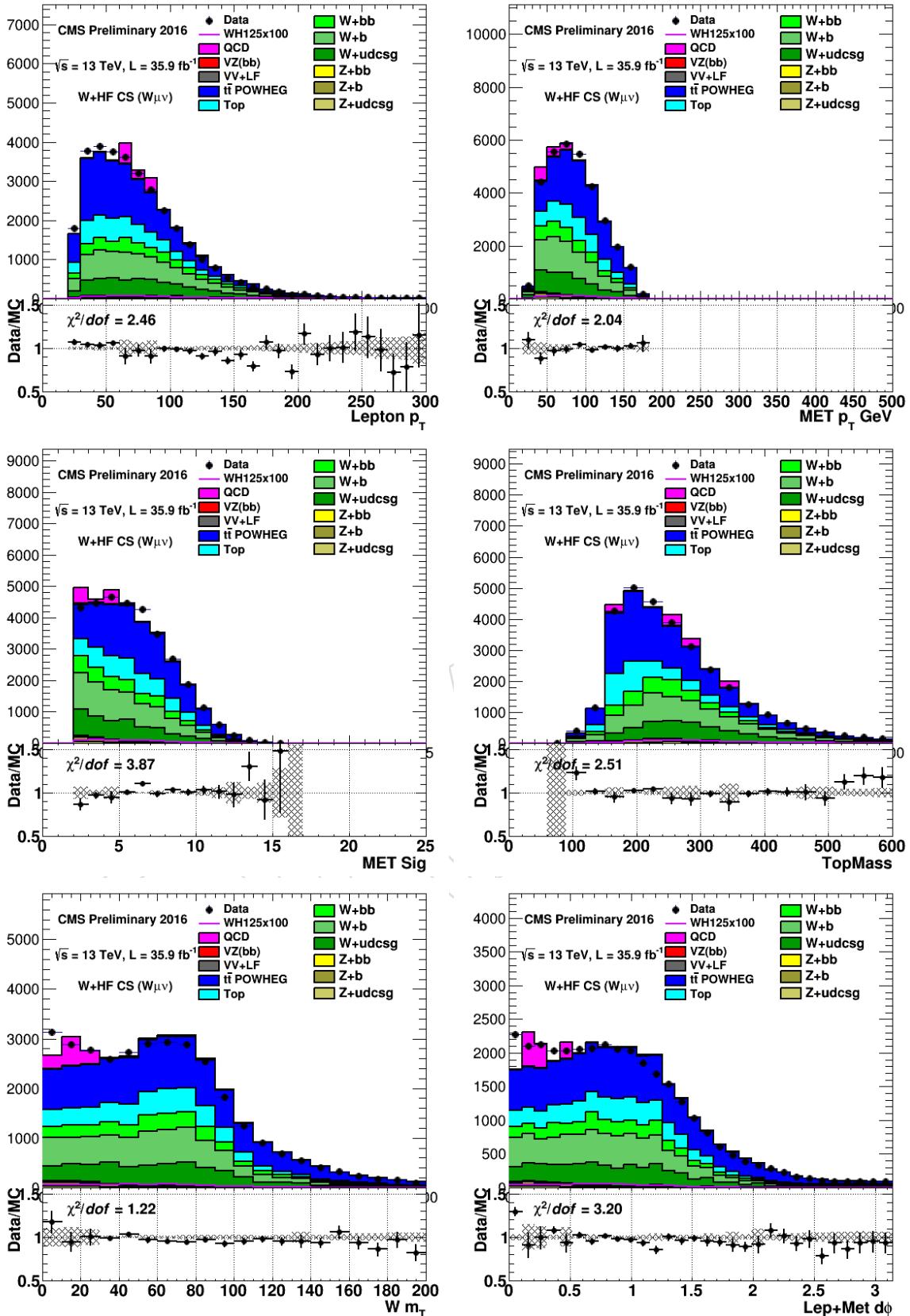


Figure 74: Distributions of data and simulated samples in the $W(\mu\nu)H$ $Wb\bar{b}$ control region. Left to right and top to bottom: muon p_T , pfMET, pfMET significance, reconstructed top mass, the transverse mass of the reconstructed W boson, and the azimuthal angle between the lepton and missing energy. To facilitate the shape comparison, the Scale Factors are applied to the simulation.

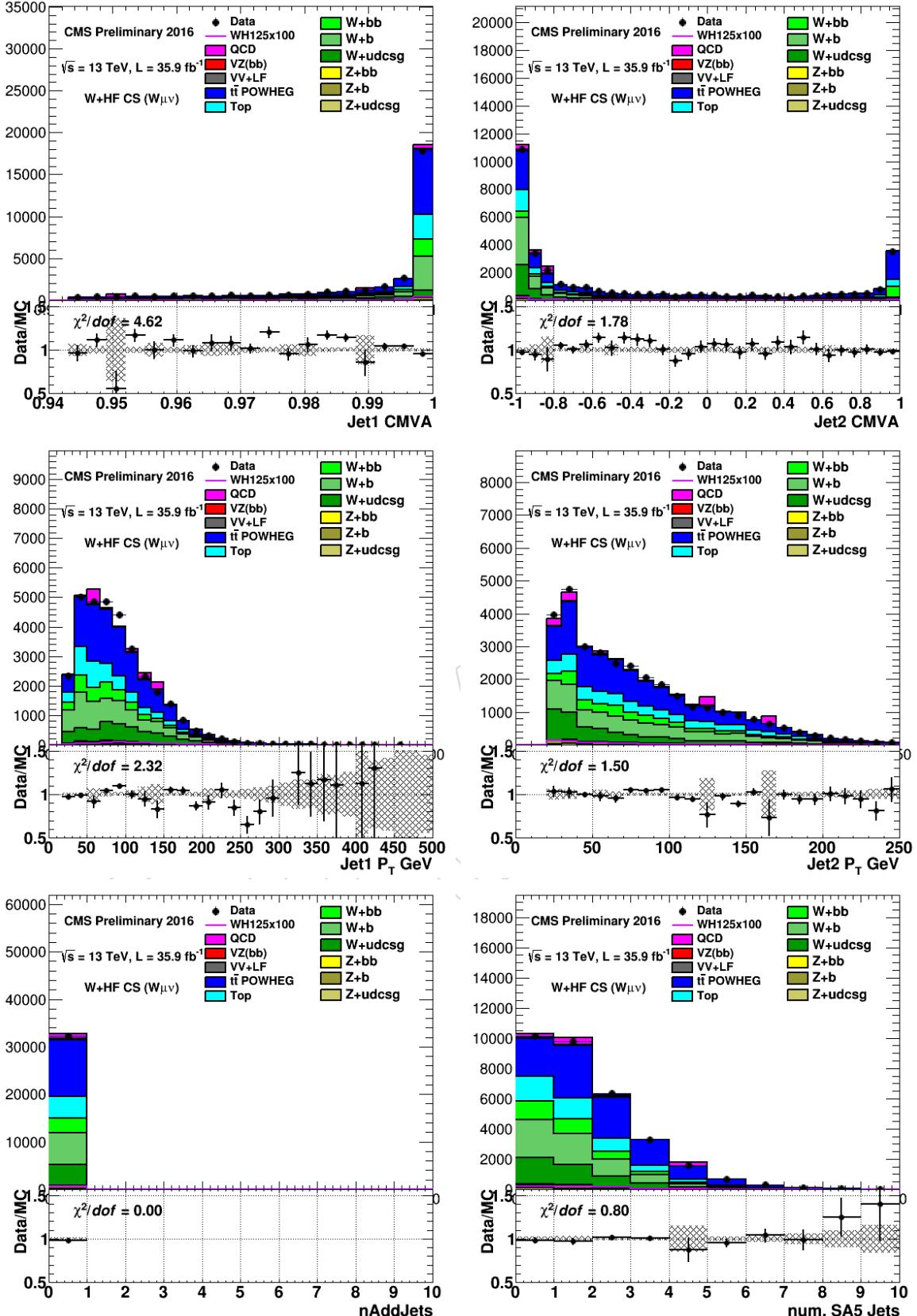


Figure 75: Distributions of data and simulated samples in the $W(\mu\nu)H$ $Wb\bar{b}$ control region. Left to right and top to bottom: the output of the CMVA for the highest and next highest scoring b-jets, the p_T distributions of those jets, the number of additional jets and the number of additional soft activity jets with $p_T > 5$ GeV. To facilitate the shape comparison, the Scale Factors are applied to the simulation.

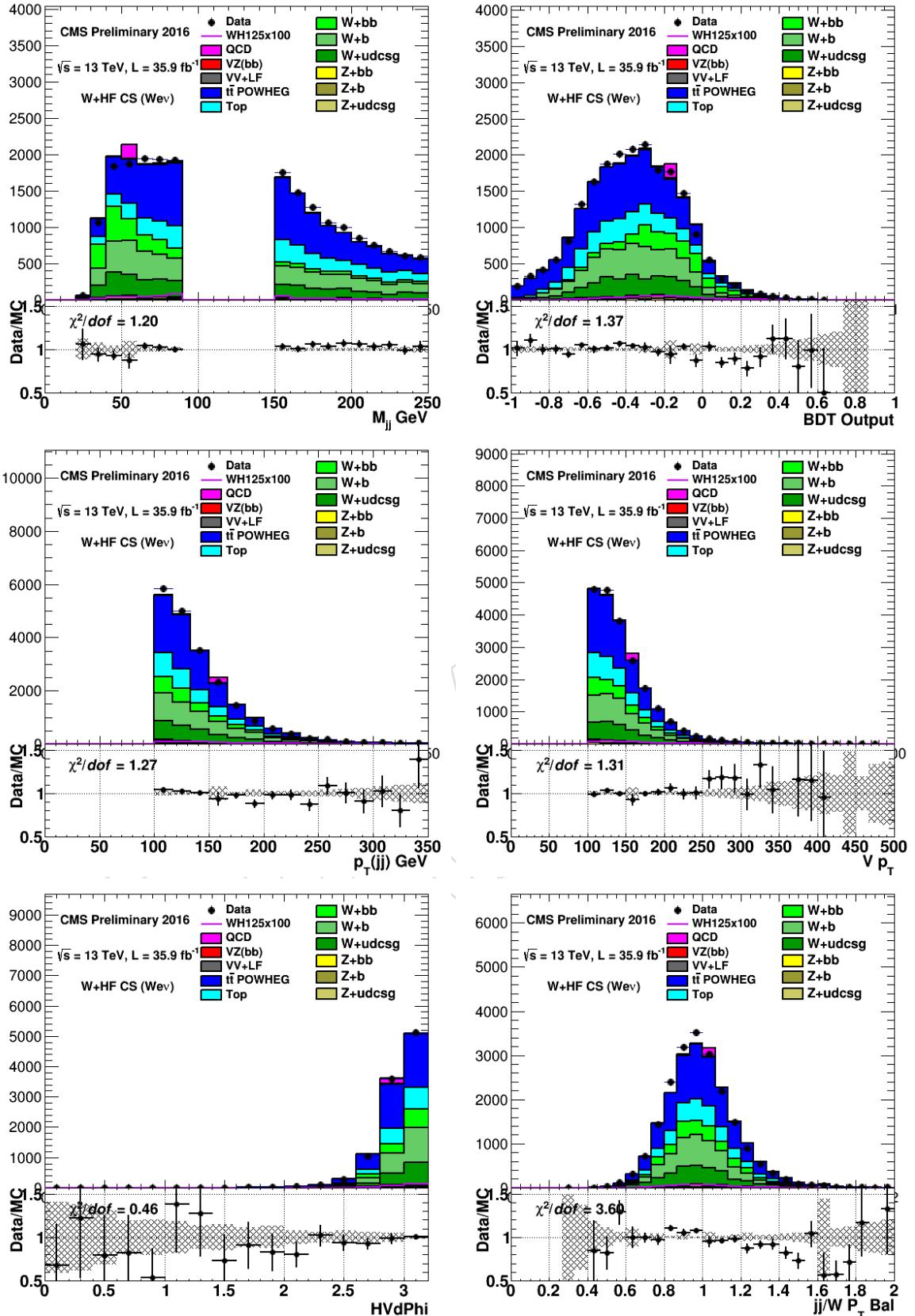


Figure 76: Distributions of data and simulated samples in the $W(e\nu)H$ $W\bar{b}b$ control region. Left to right and top to bottom: dijet invariant mass, BDT output, $p_T(jj)$ of the dijet, $p_T(W)$ of the reconstructed W boson, azimuthal angle between the W and dijet, and the ratio of $p_T(jj) / p_T(W)$. To facilitate the shape The simulation is adjusted by the scale factors in Table 19, but the overall normalization between data and simulation is not forced to be the same.

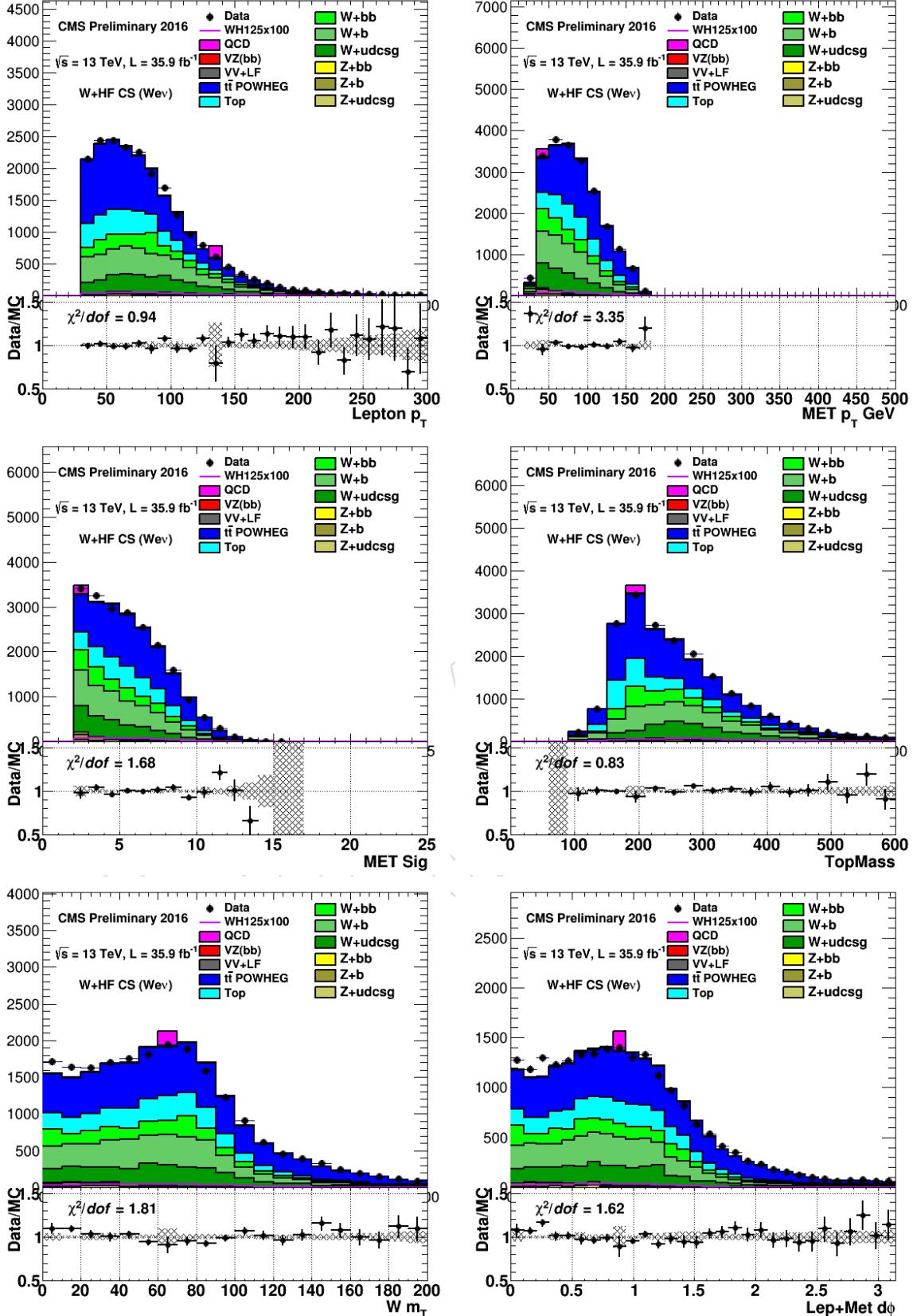


Figure 77: Distributions of data and simulated samples in the $W(e\nu)H$ $Wb\bar{b}$ control region. Left to right and top to bottom: electron p_T , pfMET, pfMET significance, reconstructed top mass, the transverse mass of the reconstructed W boson, and the azimuthal angle between the lepton and missing energy. To facilitate the shape comparison, the Scale Factors are applied to the simulation.

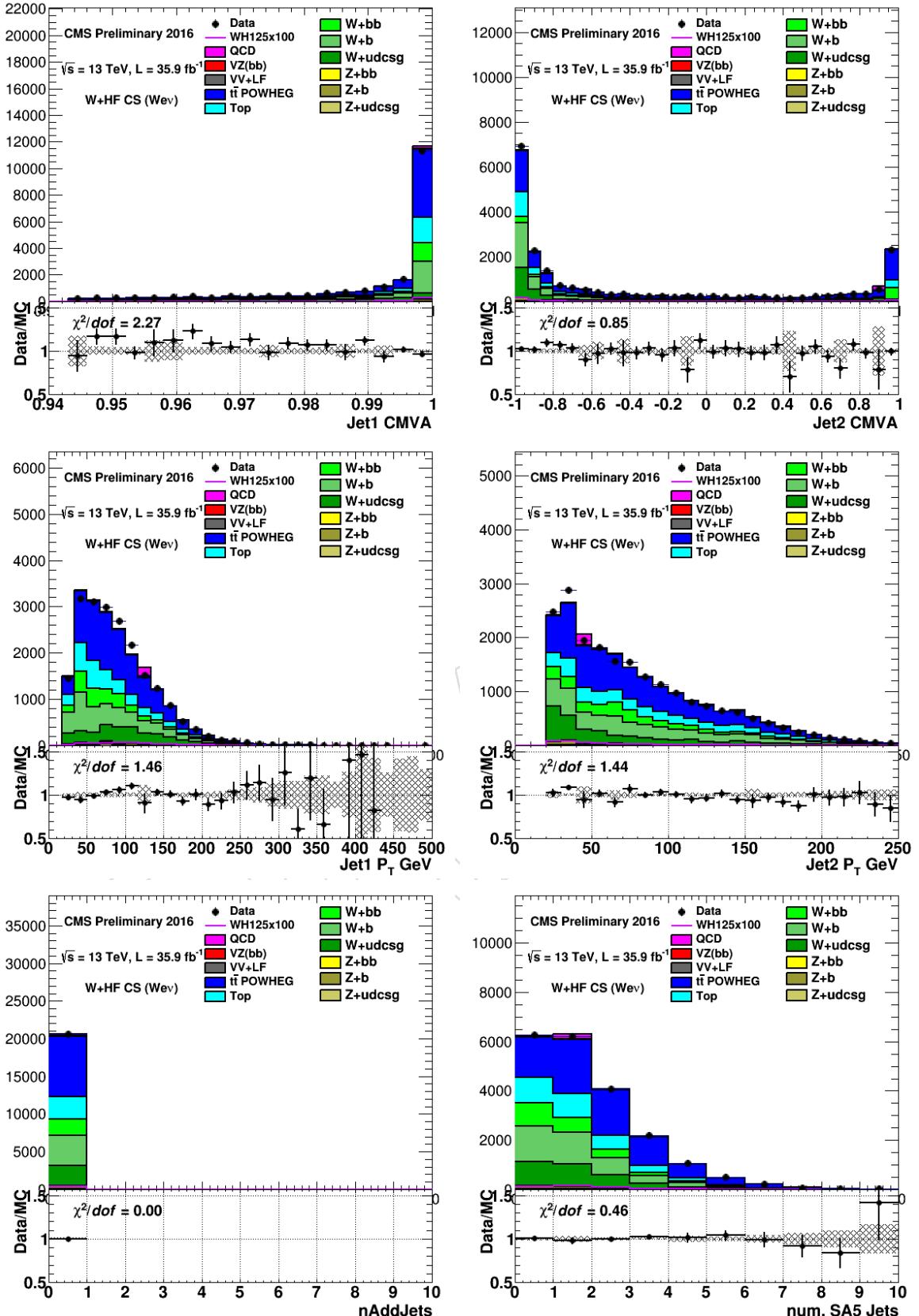


Figure 78: Distributions of data and simulated samples in the $W(e\nu)H$ $W\bar{b}b$ control region. Left to right and top to bottom: the output of the CMVA for the highest and next highest scoring b-jets, the p_T distributions of those jets, the number of additional jets and the number of additional soft activity jets with $p_T > 5$ GeV. To facilitate the shape comparison, the Scale Factors are applied to the simulation.

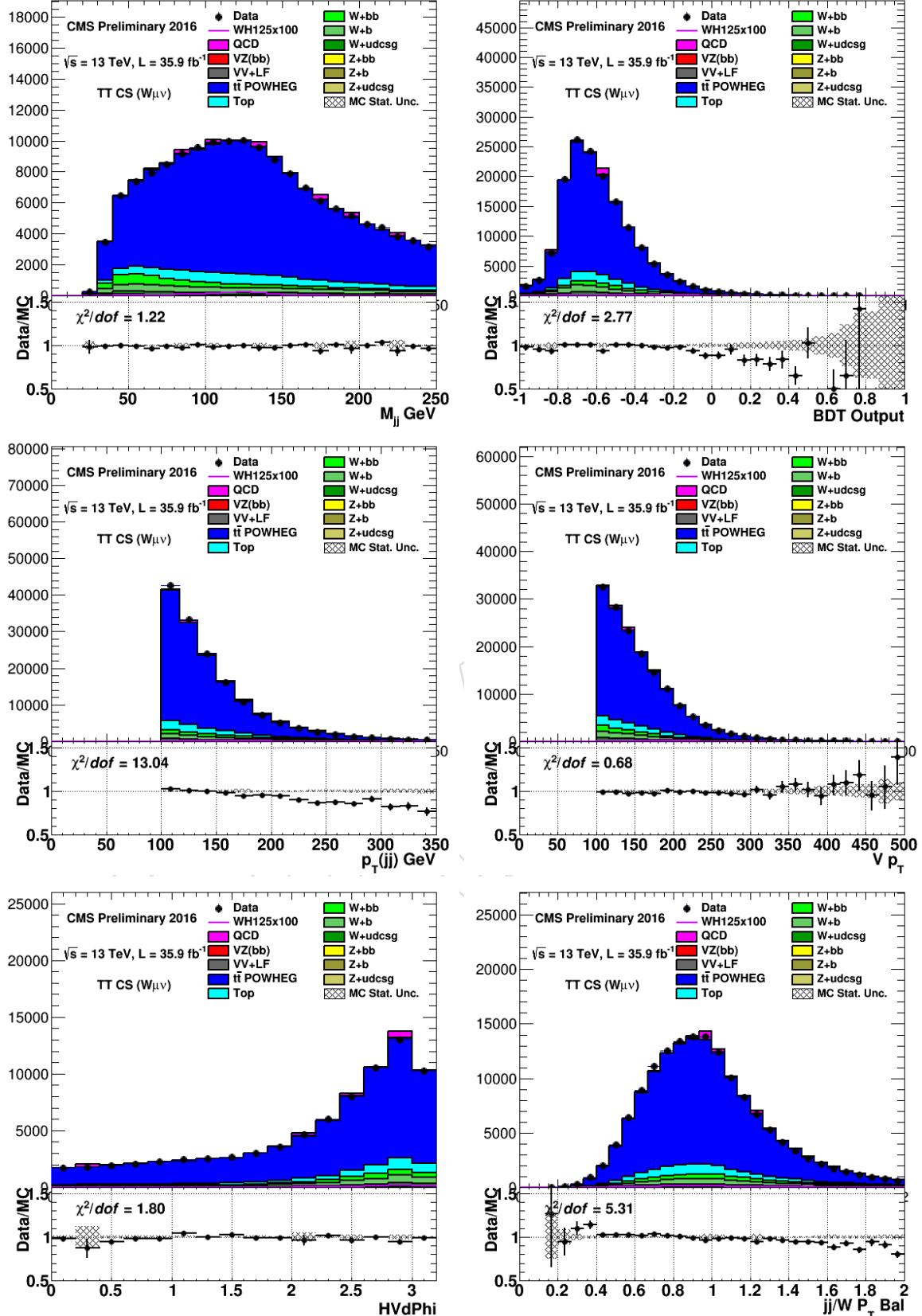


Figure 79: Distributions of data and simulated samples in the $W(\mu\nu)H\bar{t}\bar{t}$ control region. Left to right and top to bottom: dijet invariant mass, BDT output, $p_T(jj)$ of the dijet, $p_T(W)$ of the reconstructed W boson, azimuthal angle between the W and dijet, and the ratio of $p_T(jj)/p_T(W)$. To facilitate the shape The simulation is adjusted by the scale factors in Table 19, but the overall normalization between data and simulation is not forced to be the same.

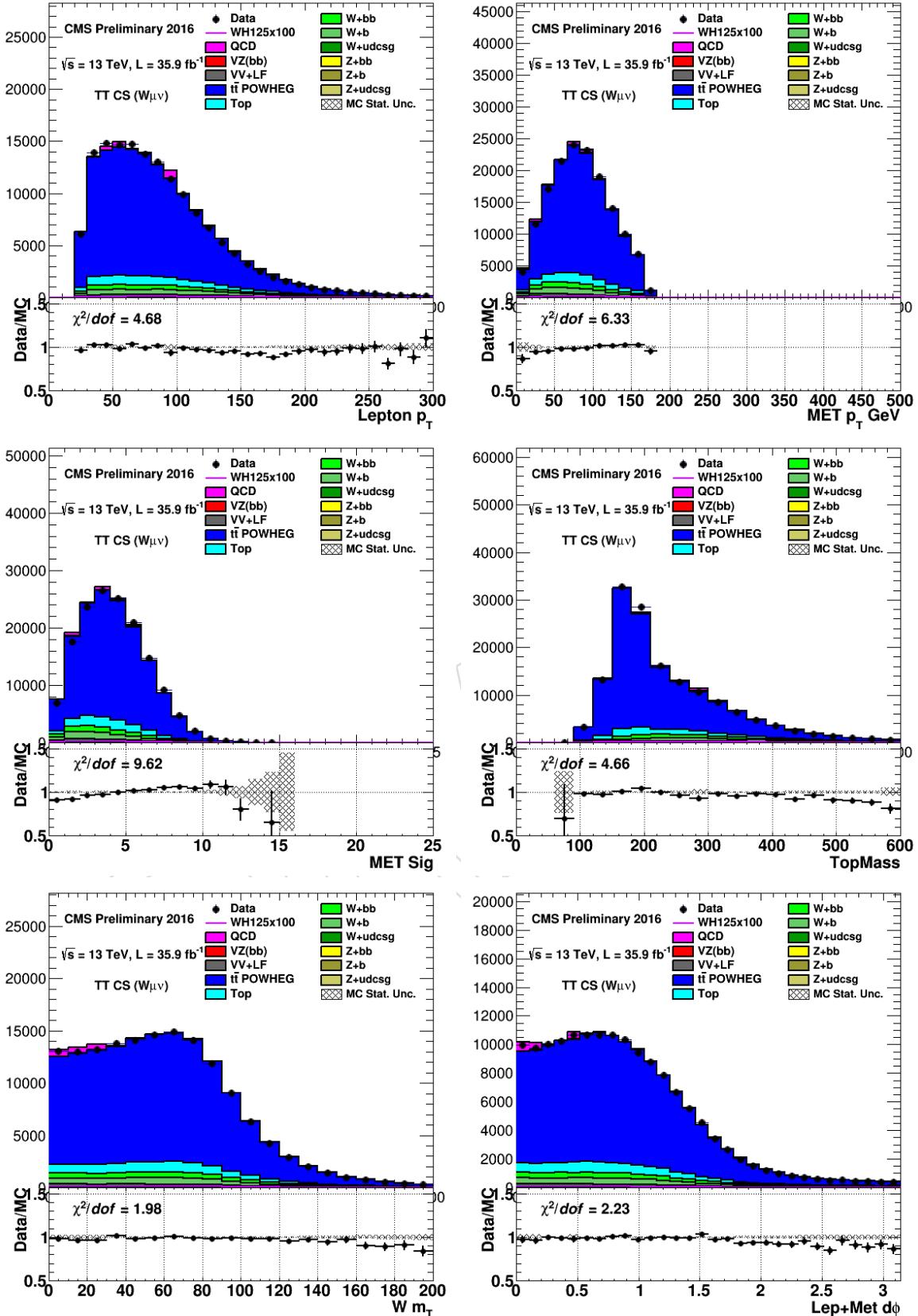


Figure 80: Distributions of data and simulated samples in the $W(\mu\nu)H\bar{t}\bar{t}$ control region. Left to right and top to bottom: muon p_T , pfMET, pfMET significance, reconstructed top mass, the transverse mass of the reconstructed W boson, and the azimuthal angle between the lepton and missing energy. To facilitate the shape comparison, the Scale Factors are applied to the simulation.

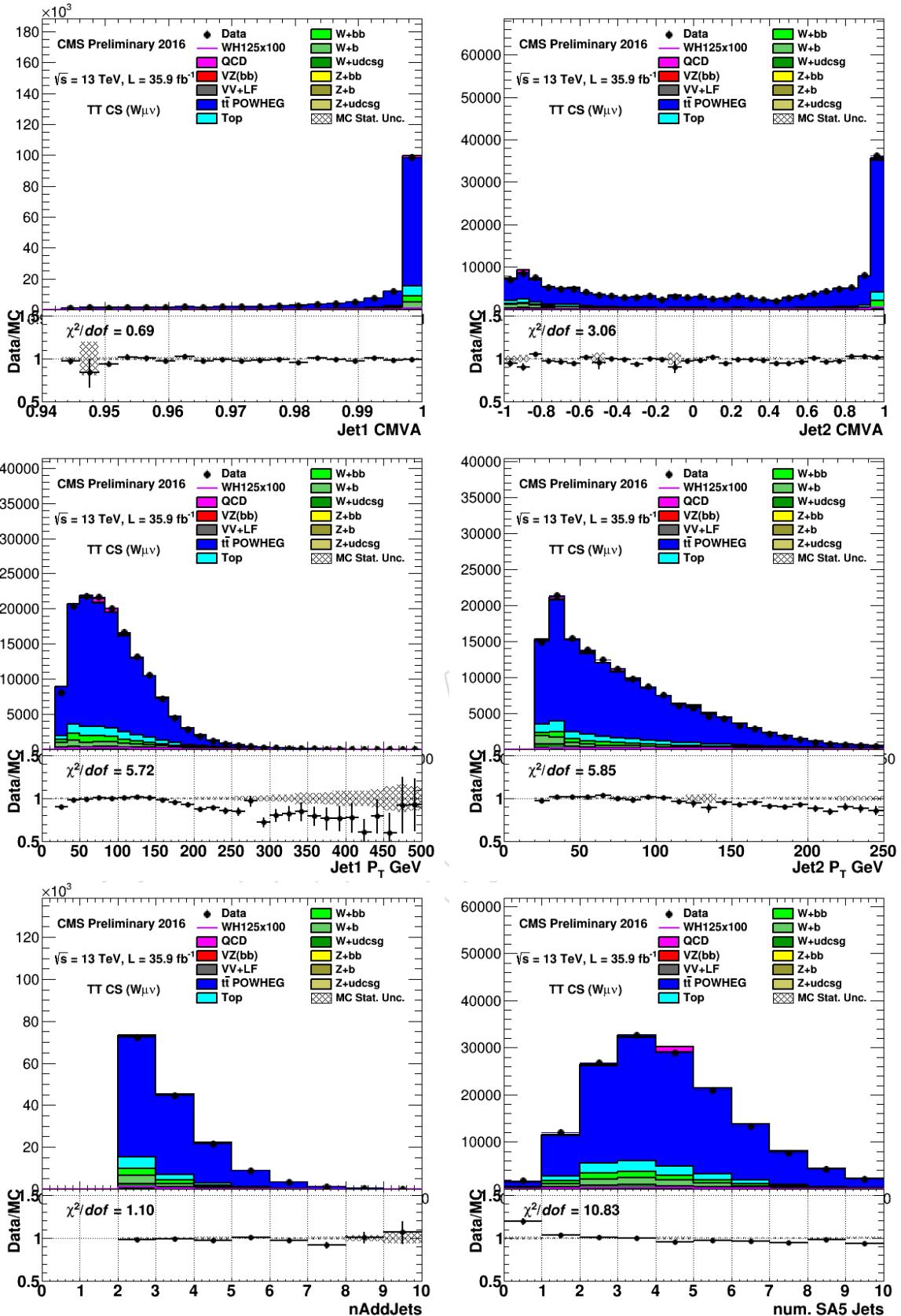


Figure 81: Distributions of data and simulated samples in the $W(\mu\nu)H\bar{t}t$ control region. Left to right and top to bottom: the output of the CMVA for the highest and next highest scoring b-jets, the p_T distributions of those jets, the number of additional jets and the number of additional soft activity jets with $p_T > 5$ GeV. To facilitate the shape comparison, the Scale Factors are applied to the simulation.

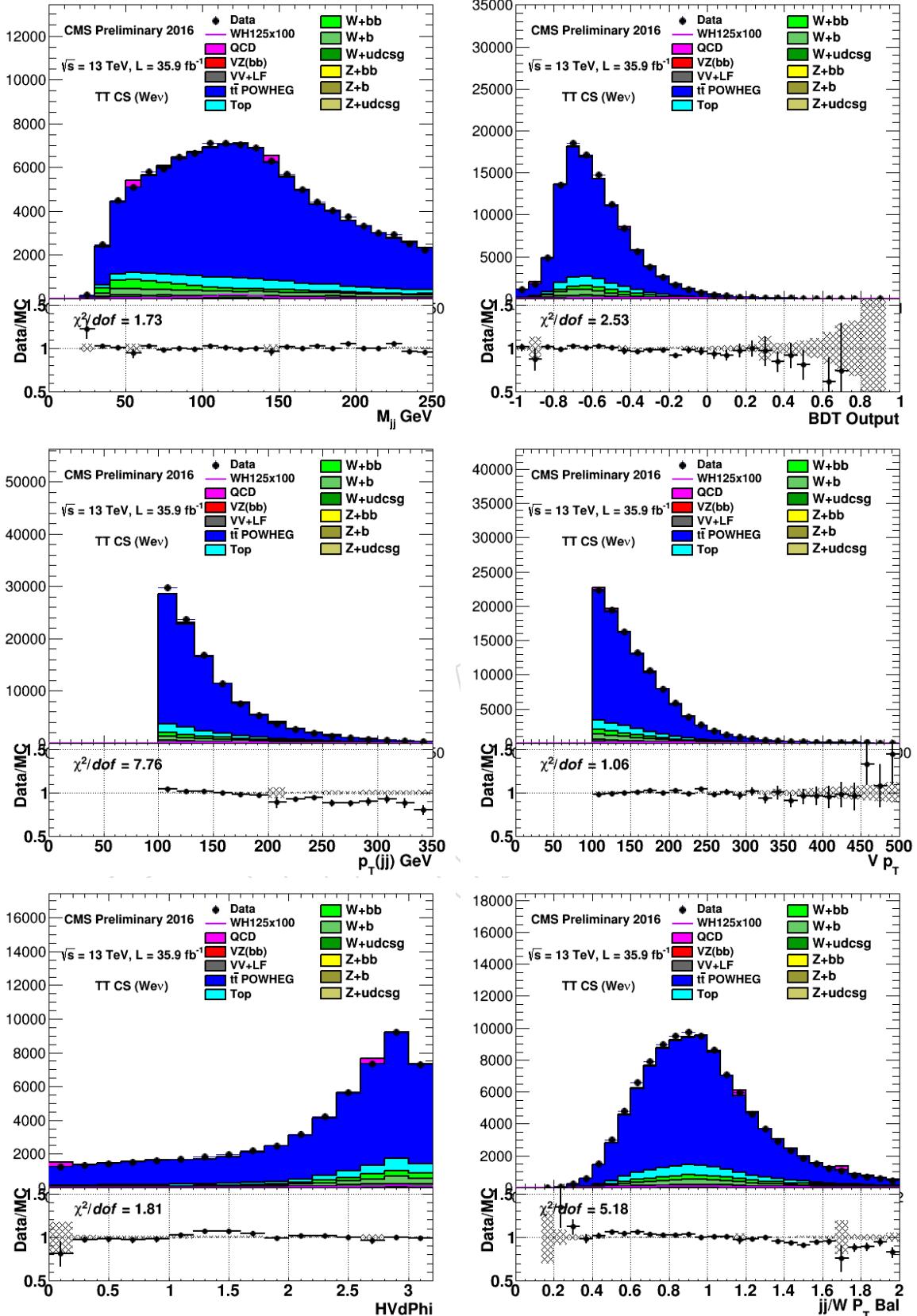


Figure 82: Distributions of data and simulated samples in the $W(e\nu)H\bar{t}t$ control region. Left to right and top to bottom: dijet invariant mass, BDT output, $p_T(jj)$ of the dijet, $p_T(W)$ of the reconstructed W boson, azimuthal angle between the W and dijet, and the ratio of $p_T(jj)/p_T(W)$. To facilitate the shape The simulation is adjusted by the scale factors in Table 19, but the overall normalization between data and simulation is not forced to be the same.

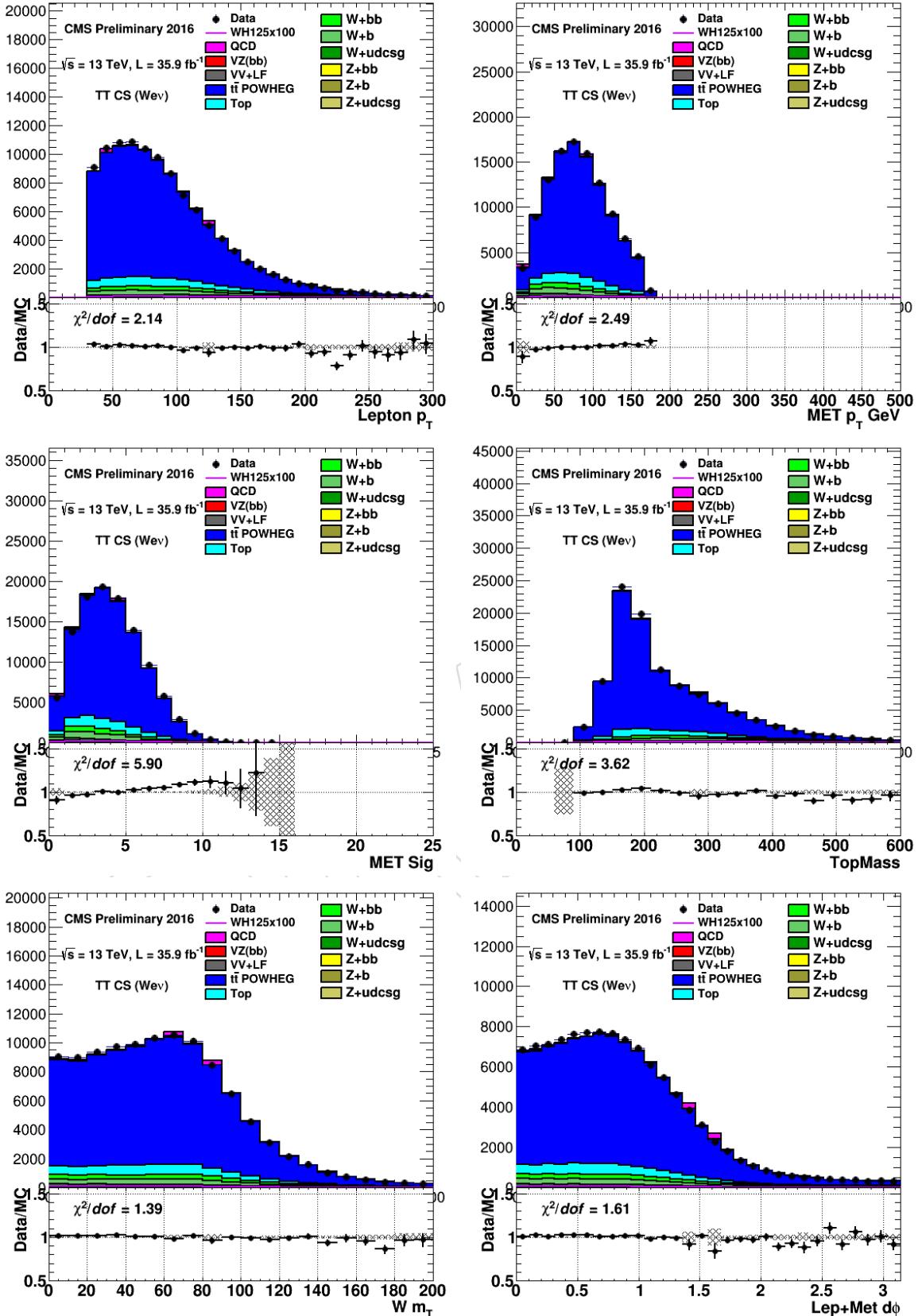


Figure 83: Distributions of data and simulated samples in the $W(\mu\nu)H\bar{t}\bar{t}$ control region. Left to right and top to bottom: electron p_T , pfMET, pfMET significance, reconstructed top mass, the transverse mass of the reconstructed W boson, and the azimuthal angle between the lepton and missing energy. To facilitate the shape comparison, the Scale Factors are applied to the simulation.

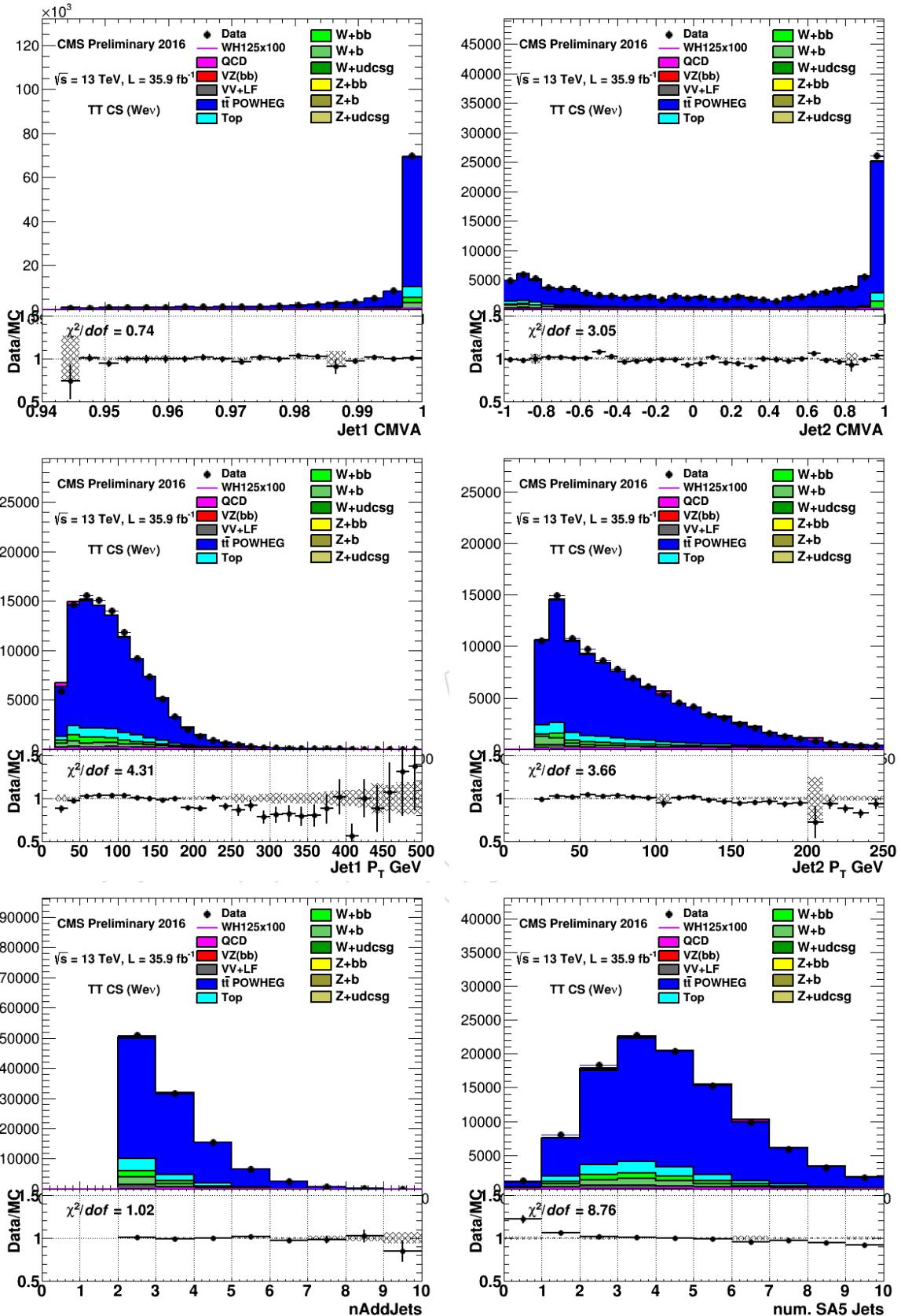


Figure 84: Distributions of data and simulated samples in the $W(e\nu)H\bar{t}\bar{t}$ control region. Left to right and top to bottom: the output of the CMVA for the highest and next highest scoring b-jets, the p_T distributions of those jets, the number of additional jets and the number of additional soft activity jets with $p_T > 5$ GeV. To facilitate the shape comparison, the Scale Factors are applied to the simulation.

Table 19: 13 TeV Data/MC scale factors for each control region in each decay mode for the SR + CRs fit. The errors include the statistical uncertainty from the fit, and a systematic uncertainty accounting for possible data/MC shape differences in the discriminating variables. Electron and muons samples in $Z(\ell\ell)H$ and $W(\ell\nu)H$ are fit simultaneously to determine average scale factors. The values refer to the simultaneous control regions plus signal region fit used for the signal extraction, and we checked that we get compatible fitted value for a fit in the control regions only.

Process	$W(\ell\nu)H$	$Z(\ell\ell)H$ low $Z p_T$	$Z(\ell\ell)H$ high $Z p_T$	$Z(\nu\nu)H$
W0b	1.14 ± 0.07	–	–	1.14 ± 0.07
W1b	1.66 ± 0.14	–	–	1.66 ± 0.14
W2b	1.49 ± 0.12	–	–	1.49 ± 0.12
Z0b	–	1.02 ± 0.06	1.02 ± 0.06	1.02 ± 0.07
Z1b	–	0.98 ± 0.06	0.99 ± 0.10	1.28 ± 0.17
Z2b	–	1.08 ± 0.07	1.29 ± 0.09	1.60 ± 0.10
$t\bar{t}$	0.91 ± 0.03	0.99 ± 0.03	1.02 ± 0.05	0.78 ± 0.05
multipjet	–	–	–	1.50 ± 0.25

898 11.4 Scale factor determination

- 899 The normalizations of the $t\bar{t}$, W + light-jets, W + b-jets, Z + light-jets, Z + b-jets and multijet backgrounds
900 are freely fitted in the signal extraction. However, preliminary background normalizations are needed
901 earlier, both to train the multivariate discriminant using the proper background normalizations and to
902 test the simulation modeling of the main variables distribution.
- 903 Within a given topology, scale factors for multijet, V+light, $t\bar{t}$, and V+heavy are obtained by fitting the
904 control regions simultaneously. channel, and can be different for each control region within a given
905 channel.
- 906 All the control regions are defined to be signal free, and to minimize the extrapolation uncertainties
907 from the signal region. In $Z(\ell\ell)H$ two specific regions have been defined: $Z + jets$ and $t\bar{t}$. These are
908 respectively the dijet and the dilepton mass sidebands of the signal regions. The exact definitions are
909 reported in Table 15. Muon and electron scale factors are fitted together considering the electron and
910 muon efficiencies as additional nuisances. For all control regions the variable fitted is the $CMVA_{min}$. See
911 Figure 97 for all post-fit distributions.
- 912 In the $W(\ell\nu)H$ heavy flavor control region a degeneracy between single top, W1b, and W2b is re-
913 moved by further splitting events into a Wbb -enriched low-mass category ($M(bb) < 90$ GeV) and a
914 Wbb -depleted high-mass category ($M(bb) > 150$ GeV). This categorization is gives another constraint to
915 the fit and stabilizes the fitted scale factors for W1b, W2b, and single top.
- 916 Electron and muon channels are fit simultaneously to determine average scale factors. In all cases,
917 binned maximum likelihood fits are performed using templates derived from simulation. Systematic
918 uncertainties on the fitted scale factors are determined by evaluating the effect on the template shapes
919 from various sources of systematics, including b-tagging and jet energy scale and resolution.
- 920 Table 19 summarizes the fit results in all three topologies for 13 TeV data for the SR+CRs fit. In general,
921 the fit results are very good and reliable scale factors are determined that are consistent with estimates
922 from previous iterations of this analysis. Tables 20–22 show the correlation matrix for each of the three
923 scale factor fits.

Table 20: Scale factors fit correlation matrix for the $Z(\ell\ell)H$ channel. The fit is performed including systematic uncertainties, not shown here and discussed in section 12.

	$t\bar{t}$ high	$Z + 0b$ high	$Z + 1b$ high	$Z + b\bar{b}$ high	$t\bar{t}$ low	$Z + 0b$ low	$Z + 1b$ low	$Z + b\bar{b}$ low
$t\bar{t}$ high	1.0	—	—	—	—	—	—	—
$Z + 0b$ high	0.80	1.0	—	—	—	—	—	—
$Z + 1b$ high	0.01	-0.03	1.0	—	—	—	—	—
$Z + b\bar{b}$ high	0.69	0.62	-0.41	1.0	—	—	—	—
$t\bar{t}$ low	0.79	0.90	0.16	0.54	1.0	—	—	—
$Z + 0b$ low	0.79	0.95	0.17	0.57	0.93	1.0	—	—
$Z + 1b$ low	0.46	0.54	0.10	0.33	0.52	0.50	1.0	—
$Z + b\bar{b}$ low	0.64	0.75	0.13	0.45	0.85	0.78	0.24	1.0

Table 21: Correlation matrix from the scale factor fit for the $Z(\nu\nu)H$ channel..

Variable	QCD	$t\bar{t}$	W+1b	W+2b	W+0b	Z+1b	Z+2b	Z+0b
QCD	1.000	—	—	—	—	—	—	—
$t\bar{t}$	-0.020	1.000	—	—	—	—	—	—
W+1b	0.083	-0.164	1.000	—	—	—	—	—
W+2b	-0.024	-0.065	-0.345	1.000	—	—	—	—
W+0b	0.249	0.242	-0.045	-0.015	1.000	—	—	—
Z+1b	0.069	0.176	-0.336	0.223	-0.000	1.000	—	—
Z+2b	-0.081	0.191	0.131	-0.397	-0.025	-0.405	1.000	—
Z+0b	0.081	0.371	0.167	-0.009	0.197	-0.316	0.158	1.000

Table 22: Scale factors fit correlation matrix for the $W(\ell\nu)H$ channel. The fit is performed including systematic uncertainties, not shown here and discussed in section 12.

	$t\bar{t}$	$W + 0b$	$W + 1b$	$W\bar{b}\bar{b}$
$t\bar{t}$	1.0	—	—	—
$W + 0b$	0.35	1.0	—	—
$W + 1b$	0.49	0.18	1.0	—
$W\bar{b}\bar{b}$	0.25	0.14	0.33	1.0

924 12 Systematics

925 The primary physics results described in this note are 1) an upper limit on the production of a standard
 926 model Higgs boson produced in association with a vector boson and decaying to a $b\bar{b}$ pair and 2) the
 927 significance of an signal like excess.. Systematic uncertainties on the expected signal and background
 928 yields and shapes affect the upper limit, and several sources have been considered:

- 929 • **Luminosity:** an uncertainty of 2.5% is assessed for 2016 luminosity.
- 930 • **Lepton Efficiency:** muon and electron trigger, reconstruction, and identification efficiencies
 931 are determined in data using the standard tag-and-probe technique with Z bosons. The sys-
 932 tematic uncertainty is evaluated from the statistical uncertainties in the bin-by-bin efficiencies
 933 and scale factors as applied to signal Monte Carlo samples. The total uncertainty is 1.6% per
 934 muon, and 1.5% per electron, which we take as a constant 2% per charged lepton. No shape
 935 uncertainties are considered.
- 936 • **Unclustered MET:** we follow the suggested procedure from the JetMET POG and find a 3%
 937 systematic uncertainty on the calibration of unclustered MET (ie, missing energy associated
 938 with particles not clustered into jets). No shape uncertainties are considered.
- 939 • **MET+jets Trigger:** the parameters describing the trigger efficiency curve have been varied
 940 within their statistical uncertainties. An uncertainty of 3% is estimated, and no shape uncer-
 941 tainties are considered.
- 942 • **Jet Energy Scale:** the jet energy scale for each jet is varied within one standard deviation
 943 based on p_T and η , and the efficiency of the analysis selection is recomputed to assess the
 944 systematic variation on the normalization of the signal and all background components. An
 945 uncertainty of 2% is found for signal, while background can vary up to 3%. For the shape vari-
 946 ation, the BDT output is recomputed after shifting the scale up and down, and the observed
 947 variation in the BDT is used to set shape variation. Each source of uncertainty assessed by
 948 the JETMET group is varied individually.
- 949 • **Jet Energy Resolution:** we smear the energy resolution for each jet using the JetMET pre-
 950 script as our default, and then assign a systematic uncertainty based on further smearing
 951 (up and down). An uncertainty of 3% is estimated for the normalization. For the shape un-
 952 certainty, the BDT is recomputed after the smearing and the modified output is used to define
 953 the shape variation.
- 954 • **B-jet Tagging:** official b-tagging scale factors are applied consistently to jets in signal and
 955 background events. An average systematic uncertainty of 6% per b jet, 12% per c jet, and 15%
 956 per fake tag (light quarks and gluons) are used to account for the normalization uncertainty.
 957 For the shape, we vary the reshaping of the CMVA output described in Sec. 4.7 based on the
 958 official uncertainties provided by the BTV POG. This gives two new shapes for the CMVA
 959 output (“up” and “down”) that we then input to the BDT. The resulting modified BDT output
 960 is then used as the shape variation due to b-tagging uncertainties. Decorrelation in p_T and η
 961 as described in Section 12.1 is implemented (in bins of BTV systematic assessments).
- 962 • **Cross section:** the total signal cross section has been calculated to next-to-next-to-leading
 963 order accuracy, and the total uncertainty is 4% [68], including the effect of scale and PDF
 964 variations.
- 965 • **Theoretical p_T Spectrum:** this analysis is performed in the boosted regime, and thus, po-
 966 tential differences in the p_T spectrum of the V and H between data and Monte Carlo gen-
 967 erators could introduce systematic effects in the signal acceptance and efficiency estimates.
 968 Recently, two calculations have become available that estimate the NLO electroweak [69–71]
 969 and NNLO QCD [72] corrections to VH production in the boosted regime. Both the EWK
 970 and NNLO QCD corrections have been applied to the signal MC samples. The estimated
 971 effect from NNLO electroweak corrections are 2% for ZH and 2% for WH [69–71]. For the
 972 remaining QCD correction an uncertainty of 5% for both ZH and WH is estimated.
- 973 • **$\Delta\phi(V,H)$:** systematic uncertainty on the jet angular resolution is assumed to be negligible,
 974 and this is confirmed by the good agreement observed in the control regions.

- N_{el} : the efficiency of the lepton veto is found to be 100% in the simulation, and no additional uncertainty is assigned.
- **Background Estimate:** a mix of data-driven methods, simulation, and theory uncertainties contribute to the total uncertainty on the background estimates. Correlated (luminosity, b-tagging, JEC/JER, and TrP efficiencies) and uncorrelated uncertainties (statistical, control region, and cross section) are combined separately. An uncertainty of 30% is assumed for single top (approximately the uncertainty on the measured cross section) and diboson (assumed to have the same uncertainty as the signal). The other backgrounds are taken directly from data, with the associated uncertainties from the control regions.
- **Monte Carlo Statistics:** the finite size of the signal and background MC samples are included in the normalization uncertainties. In addition, the shape of the BDT is allowed to vary within the bin-by-bin statistical uncertainties from the MC samples (in a coherent way), while also constraining the total integral within its uncertainty.
- **V+jets Monte Carlo model:** we consider the difference in the shapes output (BDT and Mjj) of a different Monte Carlo V+jets with respect to the nominal Madgraph MC. In particular Herwigpp high pt V+jets samples were compared with the Madgraph ones. The difference in BDT shape observed between these two generators is then symmetrized and used to define the shape variation.
- **PDF uncertainties:** the imperfect knowledge of the proton quark content is encoded in a set of NNPDF MC replicas. For each process, the RMS of all the variations is checked in each bin of the BDT distribution and the largest variation is used as normalization nuisance in the datacards. For more details refer to section 12.2.
- **QCD scale variations:** The QCD normalization and factorization scale variations 1/2 and 2 are considered separately for each process and taken as uncorrelated sources of systematic uncertainties (shape + normalization).
- **Shape uncertainties:** as noted above for particular sources of uncertainty, additional uncertainties on the BDT shape are computed. These uncertainties are implemented as “up and down” variations on the BDT output that arise due to the particular source of uncertainty. In the final limit determination, using RooStats, these additional uncertainties define the envelope within which the BDT shape is allowed to vary. This variation is performed in the following way: given a nominal shape $n(x)$ and two variations for “up” and “down”, $p(x)$ and $m(x)$, the code introduces a nuisance parameter “ t ”, and the final function $f(x)$ used to describe the variable is defined as:

$$t > 0 : f(x; t) = n(x) + (p(x) - n(x)) * a(t) \quad t < 0 : f(x; t) = n(x) + (m(x) - n(x)) * a(t) \quad (11)$$

Here, $a(t)$ is quadratic for $|t| < 1$; linear for $|t| \geq 1$; is everywhere continuous; and has continuous first and second derivatives. Appendix B shows the variations used for each source in each channel.

A summary plot showing the impact (and pulls) on signal strength of the most important systematic uncertainty is shown in fig. 85. The S+B fit to data is used to obtain this plot, while blinding the output signal strength.

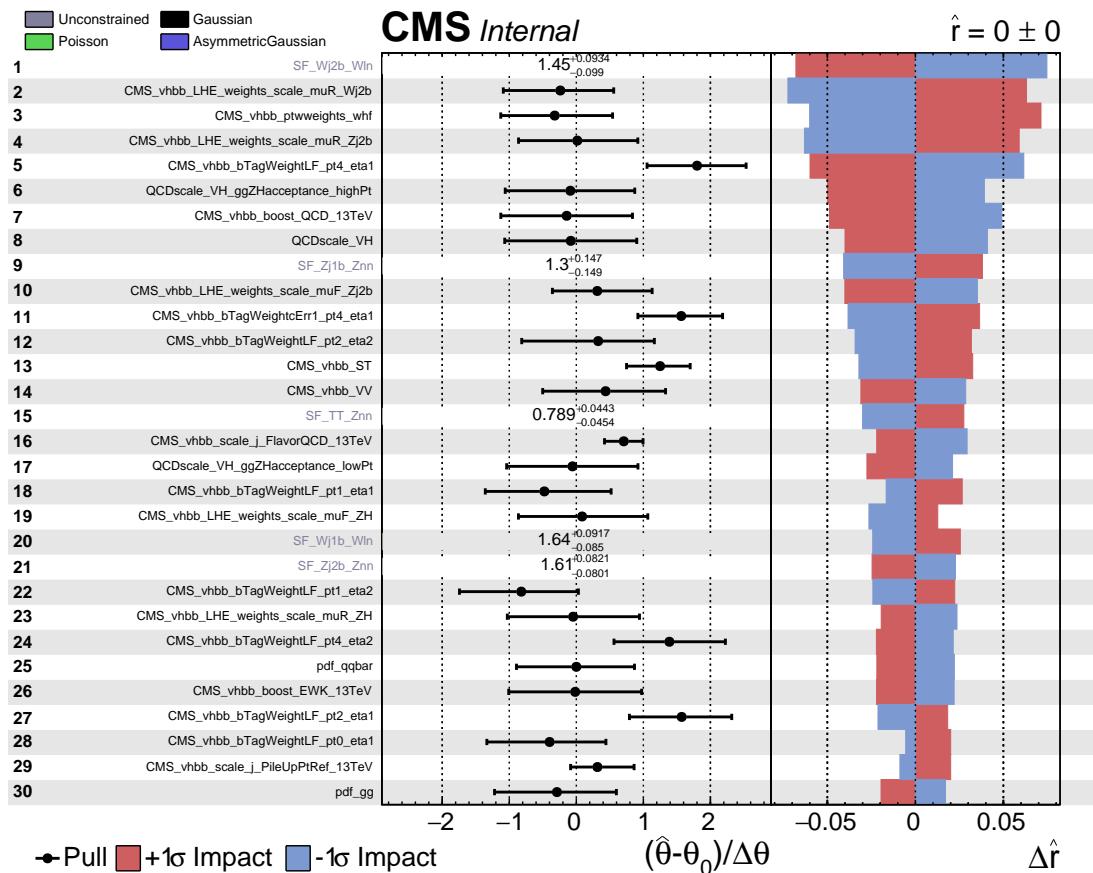


Figure 85: Impact and pulls on signal strength of the systematic sources affecting the most the signal strength. The asimov s+b fit is used to obtain this plot.

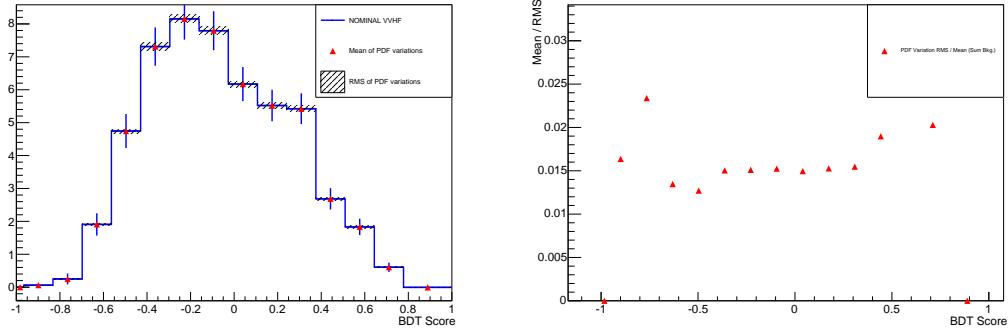


Figure 86: Left plot: nominal BDT score in the $W(\ell\nu)H(b\bar{b})$ signal region (blue histogram with errors) compared to the per-bin mean (red triangles) and RMS (shaded error band) over the 101 NNPDF MC replicas for the $WZ(bb)$ MC sample. Right plot: the RMS / mean per BDT score bin for the same distribution.

1014 12.1 Jet p_T , η decorrelation in JEC, JER and B-tagging systematics

1015 In prior iterations of this analysis, the nuisance parameters associated to jet energy scale and resolution,
 1016 as well as those associated to b-tagging, were prone to overconstraining in the signal extraction fit. We
 1017 attribute this to the fact that only one nuisance parameter was used for all jets regardless of p_T and η
 1018 (although the size of the nuisance varies as a function of p_T and η), and so of artificially fixing the pull
 1019 of these nuisances in the most signal-like regions from the high statistics background regions where
 1020 typically jets have lower p_T .

1021 However, this overconstraining behavior is symptomatic of enforcing full correlation when there is sig-
 1022 nificant uncorrelated behavior. In this edition of the analysis the recommendation of JETMET POG to
 1023 use 27 sources of uncertainties as described in <https://twiki.cern.ch/twiki/bin/viewauth/CMS/JECUncertaintySources>
 1024 is fully implemented for jet energy scale uncertainties. While for jet energy resolution a single nuisance
 1025 parameter is sufficient in our analysis (no significant constraining).

1026 For b-tagging systematics we de-correlate these systematics in five p_T bins and three η bins. The choice
 1027 of binning was selected to match the binning in which these systematics were originally assessed by the
 1028 BTV POG.

1029 12.2 Implementation of PDF Uncertainty Systematics

1030 The imperfect knowledge of the proton quark content is encoded in a set of 101 NNPDF MC replicas. For
 1031 each of these replicas, the BDT score in the signal region is re-calculated separately for each simulation
 1032 sample. The RMS and mean over the replicas is evaluated for each BDT score bin. Figure 86 (left plot)
 1033 shows the nominal BDT score in the $W(\ell\nu)H(b\bar{b})$ signal region (blue histogram with errors) compared
 1034 to the per-bin mean (red triangles) and RMS (shaded error band) over the replicas for the $WZ(bb)$ MC
 1035 sample. The right plot shows the RMS divided by the mean for each BDT score bin. The same plots are
 1036 made for each simulation sample separately, and an effective log-normal yield uncertainty is determined
 1037 by considering the largest variation. Table 23 shows the resulting effective log-normal uncertainty value
 1038 for each sample.

1039 This method approximates that the effect of each individual replica is a flat yield scaling. This approx-
 1040 imation was also made in the Run-1 VH($b\bar{b}$) analysis. However, it is now apparent that for some of
 1041 the MC replicas there is a clear shape dependence of the variation with respect to the nominal BDT
 1042 score. The log-normal implementation of this systematic is therefore not entirely correct, and needs to
 1043 be revised. The recommended prescription is to consider an independent shape nuisance for each of the
 1044 Hessian eigenvariations. Unfortunately, in the currently available production of ntuples only the repli-
 1045 cas are available. An overly conservative implementation of each of the 101 replicas as independent
 1046 shape nuisances resulted in approximately a five percent reduction in the analysis expected sensitivity.

Table 23: The effective log-normal yield uncertainty per channel due to the imperfect knowledge of the proton quark content. These values are derived by taking the maximum value of the RMS / mean of the 101 NNPDF MC replicas over all BDT score bins for each sample separately.

Sample	TT VV+HF	W+bb VV+udscg	W+b Z+bb	W+udscg Z+b	WH Z+udcsg	ZH
Effective Log-normal Yield Uncertainty (%)	0.5 2	2 3	3 2	5 3	1 5	1

¹⁰⁴⁷ Therefore it is not expected that the final implementation of this systematic will have a large effect on
¹⁰⁴⁸ the analysis sensitivity.

DRAFT

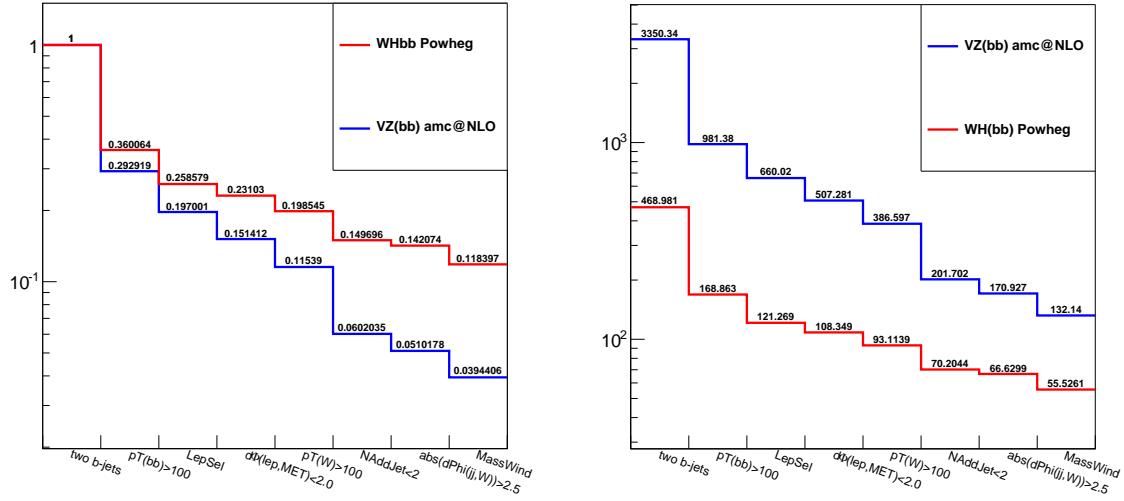


Figure 87: Cut-flow comparison of the WH and WZ analyses. The left shows the efficiency of the WH or WZ signal (amcAtNLO) with respect to the very loose kinematic preselection in the first bin. The right plot shows the yields expected with 12.9/fb up to the analysis preselection. The right most bin shows the yield entering the analysis.

1049 13 Di-boson Validation Analysis

1050 A small component of the final state in the VH search is from diboson, VV, events (where V can be a
 1051 W or Z). This background can be used to validate the methodology of the VH search by re-constructing
 1052 the analysis using VV signal as the target. Only very necessary changes are made to selection and
 1053 optimization in order to make this analysis as similar as possible to the primary analysis.

1054 These minimal changes include:

- 1055 • Moving mass window selection in signal region to 60-160 GeV
- 1056 • Moving mass window vetoes in control samples to 60-160 GeV
- 1057 • Retraining all BDTs using appropriate VV target samples

1058 Because of the difference in invariant mass the efficiency of several cuts change the VZ analysis efficiency
 1059 times acceptance (compared to the nominal VH analysis). Figure 87 shows that the cuts that enforce
 1060 boosted topology are less efficient in the WZ analysis. The number of additional jets is also notably less
 1061 efficient. Both result in a factor of three loss in relative efficiency. The figure also shows that there is still
 1062 substantially more (x2.4) signal in the WZ analysis.

1063 The results are summarized in Table 24. Results are currently blinded, but they are expected to be
 1064 very sensitive to the VV signal. The PostFit BDT distributions for the (blinded) data and simulation
 1065 for all channels are shown in Fig. 89. Control regions included in the final fit for Z($\ell\ell$)V, Z($\nu\nu$)V, and
 1066 W($\ell\nu$)V are shown in Figures 90, 91, and 92 respectively. Figure 93 shows the combination of the BDT
 1067 distributions for all sub-channels, with each sub-channel BDT bin weighted by the expected signal-to-
 1068 background ratio.

1069 The channel compatibility and signal strength is shown in Fig. 94.

1070 For this fit we show in Fig. 95 the list of the list of problematic nuisances among the 715 used for the
 1071 signal extraction. As this list is minimal, we consider this as a good test of the goodness of the systematics
 1072 treatment.

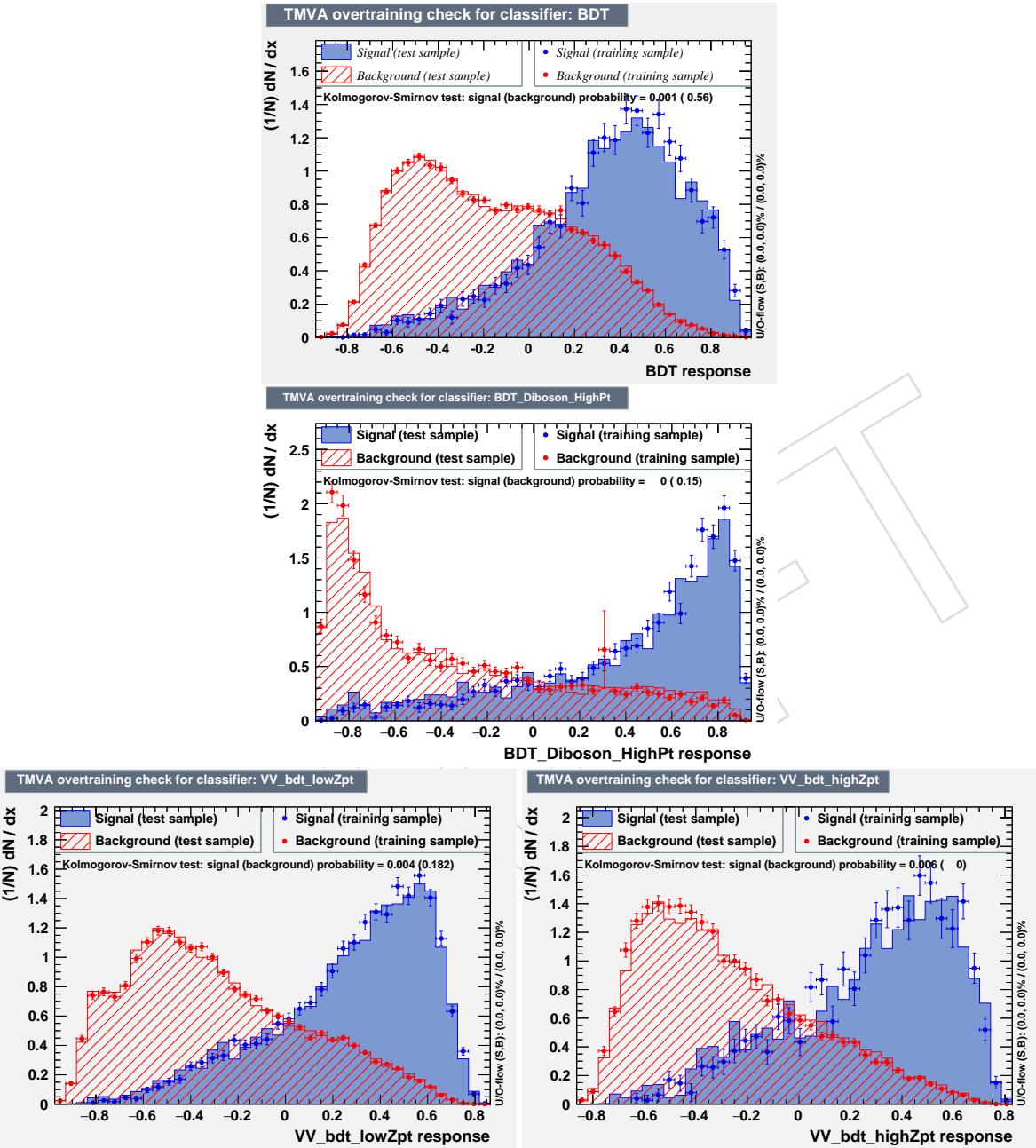


Figure 88: The testing and training samples for signal and background are shown evaluated on dedicated BDT trainings for the $W\ell\nu Z\bar{b}b$ (top), $Z\nu\nu Z\bar{b}b$ (middle), and $Z\ell\ell Z\bar{b}b$ (bottom, low and high pT regions) targets.

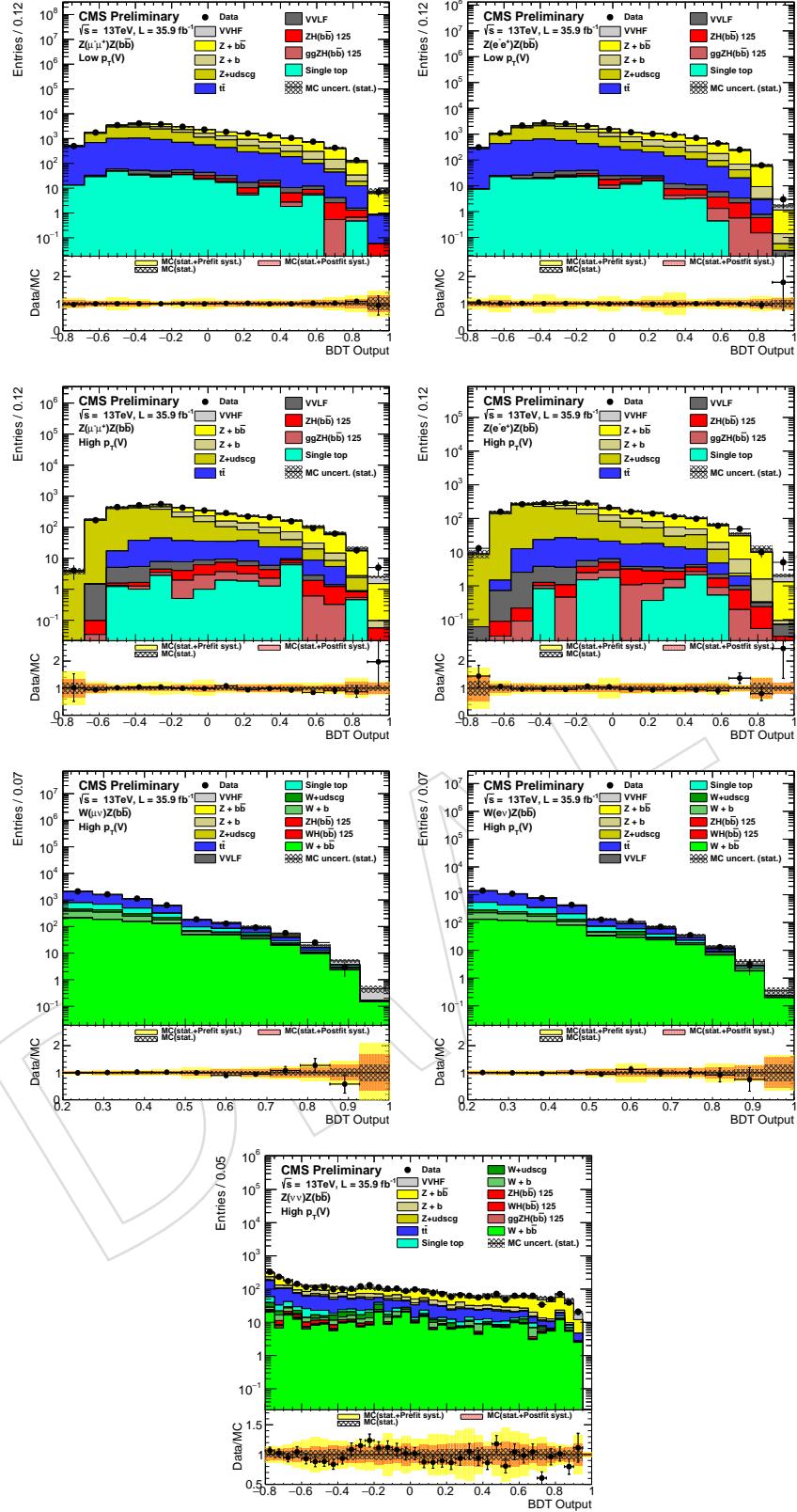


Figure 89: Post-fit distributions of blinded BDT output for combined $Z(\ell\ell)V$ (top), $W(\ell\nu)V$ (middle), and $Z(\nu\nu)V$ (bottom) channels for 13 TeV 2016 data (points with errors), after all selection criteria have been applied.

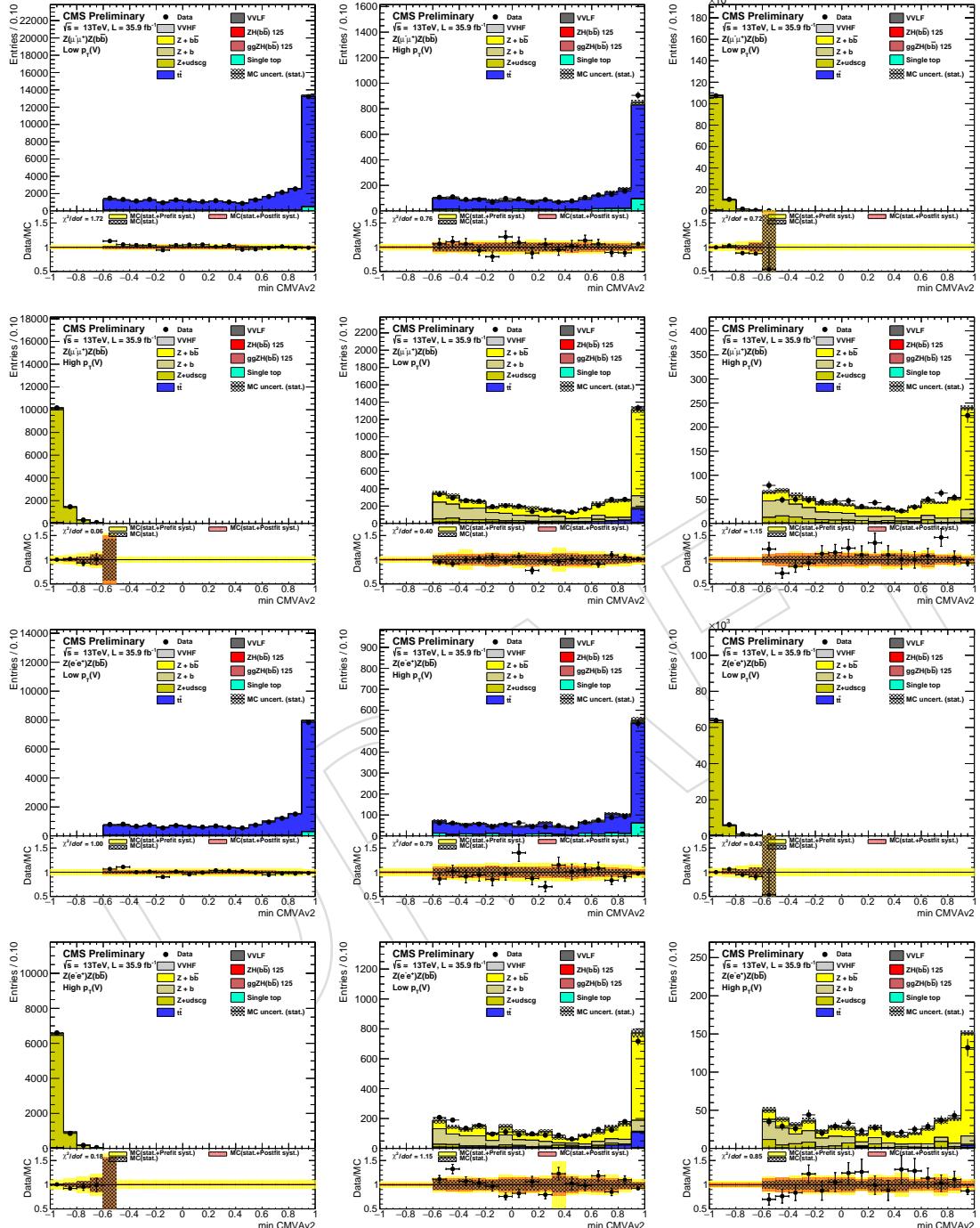


Figure 90: Post-fit distributions of the control regions for combined $Z(\ell\ell)V$ channels for 13 TeV 2016 data (points with errors), after all selection criteria have been applied.

Table 24: Results of diboson validation (unblinded). Expected and observed local significances of the excess of events above the estimated background on the diboson production in bb final states. Values are given in numbers of standard deviations. Also the value for the best-fit signal strength modifier μ is given.

Channel	$ZnnZbb$	$WlnZbb$	$ZllZbb$	Combined
Expected $\Delta\mu$	0.33	0.38	0.31	0.22
Obs μ	0.57	1.67	1.33	1.02
Obs $\Delta\mu$	0.32	0.47	0.34	0.22
Expected Signif. (postfit)	3.11	2.61	3.13	4.94
Obs Signif.	2.01	3.71	4.49	4.96

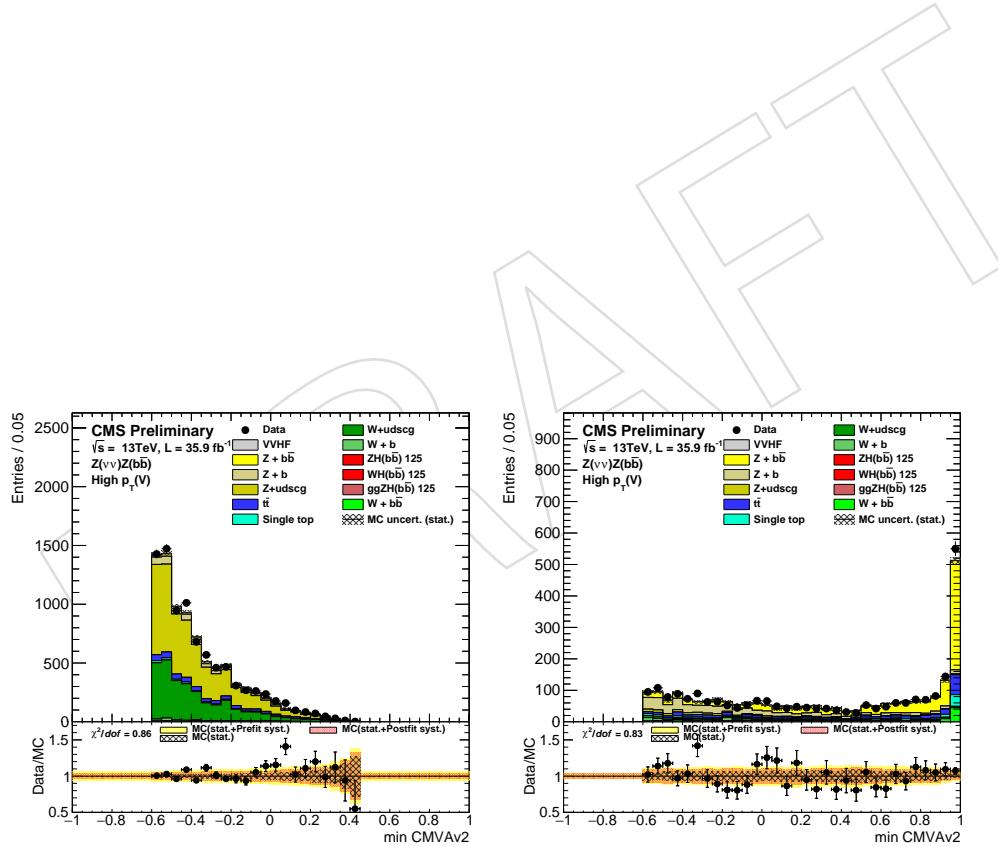


Figure 91: Post-fit distributions of the blinded BDT output in the signal region and $CMVA2_{min}$ in the $Z+udscg$, $Z+b$, $t\bar{t}$ control regions for $Z(vv)V$ with 13 TeV data (points with errors), all backgrounds, and signal after all selection criteria have been applied.

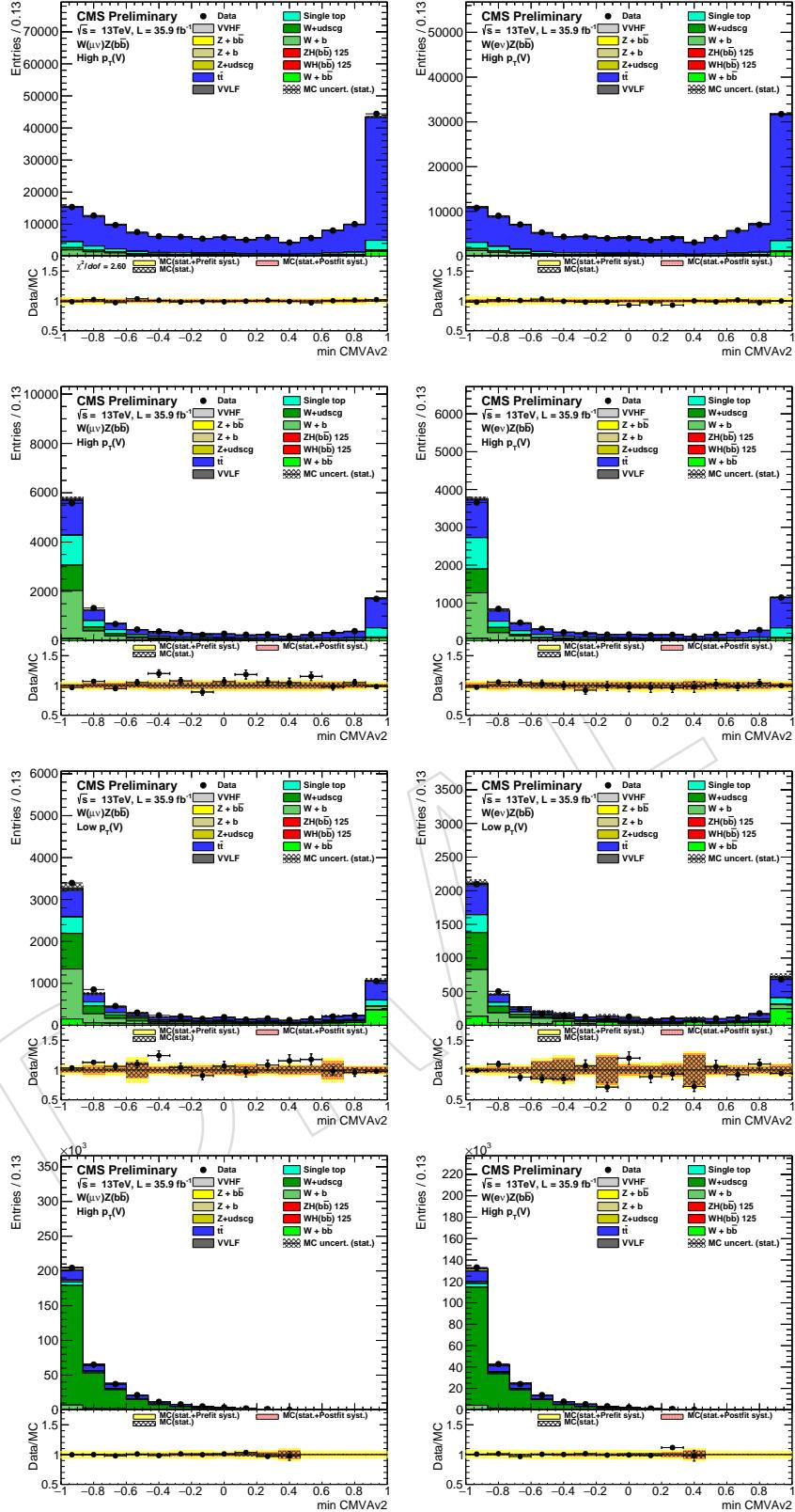


Figure 92: Post-fit distributions of blinded BDT output and min CMVA distributions from $t\bar{t}$, $W+\text{heavy flavor jets}$, and $W+\text{light flavor jets}$ for $W(\ell\nu)V$ on the left (right) in the signal region for 13 TeV 2016 data (points with errors), all backgrounds, and signal after all selection criteria have been applied.

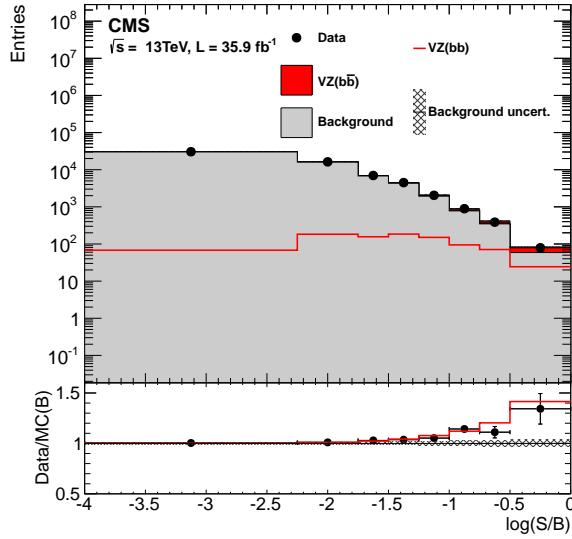


Figure 93: Combination of all channels into a single distribution. Events are sorted in bins of similar expected signal-to-background ratio, as given by the value of the output of their corresponding BDT discriminant (trained to extract a VZ(bb) signal). The bottom inset shows the ratio of the data to the predicted background, with a red line overlaying the expected Standard Model VZ(bb) signal contribution.

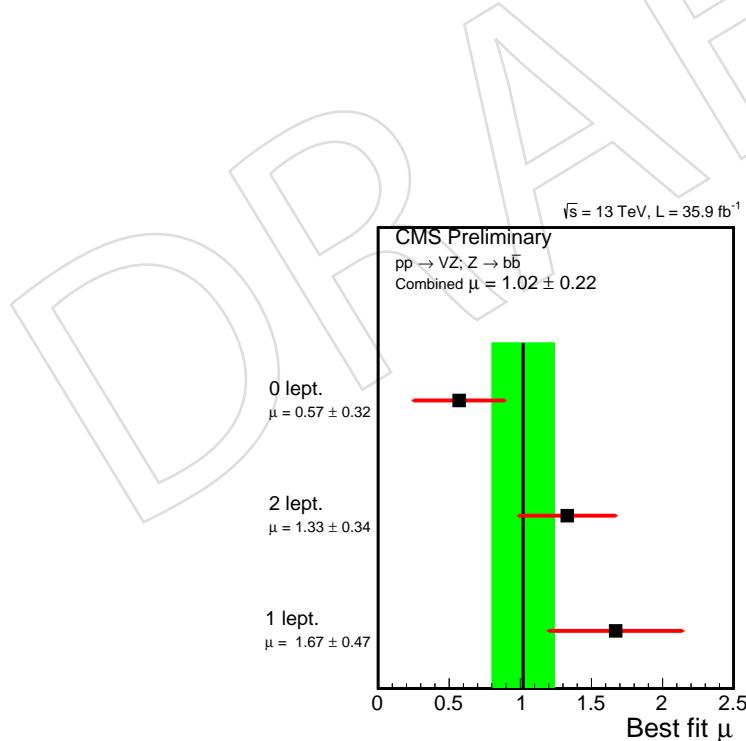


Figure 94: The most likely signal strength and uncertainty fitted in the three channels, together with the combination.

nuisance	background fit $\Delta x/\sigma_{in}, \sigma_{out}/\sigma_{in}$	signal fit $\Delta x/\sigma_{in}, \sigma_{out}/\sigma_{in}$	$p(\mu, \theta)$
CMS_vhbb_LHE_weights_scale_muF_TT	+0.08, 0.45	+0.14, 0.45	+0.02
CMS_vhbb_bTagWeightJES_pt0_eta1	-2.26, 0.87	-2.23, 0.87	+0.00
CMS_vhbb_bTagWeightJES_pt1_eta3	-2.33, 0.92	-2.22, 0.92	+0.02
CMS_vhbb_bTagWeightJES_pt2_eta2	-2.02, 0.94	-2.07, 0.94	-0.00
CMS_vhbb_bTagWeightJES_pt2_eta3	-1.99, 0.92	-2.01, 0.92	+0.00
CMS_vhbb_bTagWeightLF_pt2_eta2	-0.01, 1.52	+0.33, 0.95	+0.04
CMS_vhbb_bTagWeightLF_pt4_eta1	+2.07, 0.81	+1.75, 0.82	-0.08
CMS_vhbb_bTagWeightcErrl_pt0_eta1	+0.30, 0.40	+0.28, 0.45	-0.01
CMS_vhbb_bTagWeightcErrl_pt0_eta2	+0.28, 0.40	+0.29, 0.38	+0.01
CMS_vhbb_bTagWeightcErrl_pt0_eta3	+0.28, 0.33	+0.26, 0.33	-0.01
CMS_vhbb_bTagWeightcErrl_pt1_eta1	+0.41, 0.35	+0.42, 0.36	+0.01
CMS_vhbb_bTagWeightcErrl_pt1_eta2	+0.30, 0.43	+0.30, 0.42	+0.00
CMS_vhbb_bTagWeightcErrl_pt1_eta3	-0.27, 0.42	-0.29, 0.41	-0.01
CMS_vhbb_bTagWeightcErrl_pt2_eta1	+0.35, 0.43	+0.39, 0.44	+0.02
CMS_vhbb_bTagWeightcErrl_pt2_eta2	-0.43, 0.44	-0.36, 0.43	+0.02
CMS_vhbb_bTagWeightcErrl_pt4_eta1	+2.19, 0.70	+2.19, 0.69	-0.00
CMS_vhbb_eff_e_MVAID_Zll_13TeV	-0.45, 0.19	-0.48, 0.19	-0.02
CMS_vhbb_eff_e_Wln_13TeV	+0.58, 0.20	+0.57, 0.20	-0.01
CMS_vhbb_puWeight	+0.27, 0.18	+0.27, 0.17	+0.03
CMS_vhbb_res_j_13TeV	-0.13, 0.34	-0.35, 0.33	-0.03
CMS_vhbb_scale_j_AbsoluteMPFBias_13TeV	+0.28, 0.64	-0.41, 0.44	+0.02
CMS_vhbb_scale_j_AbsoluteScale_13TeV	+0.49, 0.54	+0.60, 0.44	+0.08
CMS_vhbb_scale_j_AbsoluteStat_13TeV	-0.07, 0.43	-0.10, 0.47	-0.01
CMS_vhbb_scale_j_FlavorQCD_13TeV	-0.52, 0.24	+0.40, 0.28	-0.03
CMS_vhbb_scale_j_Fragmentation_13TeV	-0.14, 0.31	-0.11, 0.32	-0.00
CMS_vhbb_scale_j_PileUpDataMC_13TeV	-0.02, 0.29	-0.03, 0.32	-0.00
CMS_vhbb_scale_j_PileUpPtBB_13TeV	-0.08, 0.36	-0.12, 0.43	-0.02
CMS_vhbb_scale_j_PileUpPtRef_13TeV	-0.11, 0.32	-0.15, 0.37	-0.01
CMS_vhbb_scale_j_RelativeFSR_13TeV	+0.10, 0.37	+0.05, 0.43	-0.01
CMS_vhbb_scale_j_RelativeJEREC1_13TeV	+0.11, 0.40	+0.10, 0.44	+0.01
CMS_vhbb_scale_j_RelativePtEC1_13TeV	+0.17, 0.47	+0.10, 0.51	+0.00
CMS_vhbb_scale_j_RelativeStatEC_13TeV	-0.09, 0.45	-0.12, 0.47	+0.01
CMS_vhbb_scale_j_RelativeStatFSR_13TeV	+0.08, 0.43	+0.07, 0.44	+0.00
CMS_vhbb_scale_j_SinglePionECAL_13TeV	+0.08, 0.31	+0.08, 0.33	+0.01
CMS_vhbb_scale_j_SinglePionHCAL_13TeV	+0.06, 0.29	+0.09, 0.31	+0.02

Figure 95: List of nuisances with either pulls or constraints for the maximum likelihood fit used to extract the VZbb signal strength.

1073 14 Results

1074 The final predicted amount of signal is determined with a simultaneous fit of the background enriched
 1075 control regions (CRs) distributions, namely the sub-leading CMVA distribution, where the regions are
 1076 described in Sec. 11, and the BDT trained in the signal region (SR). For $Z(\nu\nu)H$ and $W(\ell\nu)H$ we use
 1077 one single category, while in $Z(\ell\ell)H$ we split the SR in two bins of the vector boson p_T . This section
 1078 summarizes all of the final signal and background estimates in both sets.

1079 14.1 Signal and Control Regions fits

1080 The blinded distribution of the BDT output and the distributions in the CRs used for the final combined
 1081 fit of all channels are shown in the category used by each channel in Figures 96–99 for 13 TeV data from
 1082 2016. All scale factors from Sec. 11.4 have been applied to the pre-fit plots. The post-fit plots consider
 1083 the adjustments of all nuisance parameters in the final maximum likelihood fit to extract the signal. We
 1084 consider both shape and rate changes in the post-fit plots.

1085 Table 25 reports the expected signal and backgrounds in the signal region bins.

Table 25: The total number of events for partial combinations of channels in the most sensitive bins of their corresponding BDT for the expected backgrounds (B), for the 125 GeV SM Higgs boson VH signal (S), and for data. Also shown is the signal-to-background ratio (S/B).

Process	$Z(\ell\ell)H$		$W(\ell\nu)H$	$Z(\nu\nu)H$
	Low $p_T(V)$	High $p_T(V)$		
Vbb	578.73	87.12	102.5	133.4
Vb	143.37	64.33	19.9	26.3
V + udscg	70.03	1.90	9.8	15.9
t̄t	147.53	2.79	98.0	52.4
Single-top-quark	2.00	0.23	44.6	10.9
VV(udscg)	6.36	0.58	1.5	-0.8
VZ(b̄b)	23.05	3.81	6.9	10.3
Total backgrounds	971.07	160.76	283.3	248.4
VH	49.97	30.36	26.0	31.0
Data	1030.80	179.90	320	–
S/B (%)	5.1	18.8	9.2	12.5

1086 14.2 Upper limit calculation

1087 The primary technique for deriving exclusion limits is based on the so-called CL_s prescription, which
 1088 we use with the profile likelihood test statistic \tilde{q}_μ :

$$\tilde{q}_\mu = -2 \ln \frac{\mathcal{L}(\text{data}|\mu, \hat{\theta}_\mu)}{\mathcal{L}(\text{data}|\hat{\mu}, \hat{\theta})}, \quad \text{with a constraint } 0 \leq \hat{\mu} \leq \mu \quad (12)$$

1089 The likelihood is given by the product of the individual likelihoods for each channel

$$\mathcal{L}(\text{data} | \mu, \theta) = \text{Poisson}(N_i | \mu \cdot s_i(\theta) + b_i(\theta)) \cdot p(\tilde{\theta} | \theta). \quad (13)$$

1090 Here “data” is either the actual experimental *observation* or, as we do now before unblinding, *pseudo-data*
 1091 used to construct sampling distributions. The symbols N_i , s_i and b_i represent the observed, expected signal,
 1092 and expected background rates in bin i . The parameter μ is the signal strength modifier, $\mu=\sigma/\sigma_{SM}$,
 1093 and θ represents the full suite of nuisance parameters, with $\tilde{\theta}$ representing the best estimate of the nu-
 1094 iance prior to the data analysis. Poisson($N_i | \mu \cdot s_i(\theta) + b_i(\theta)$) stands for the Poisson probabilities to
 1095 observe N_i events given the expected event rate $\mu \cdot s_i(\theta) + b_i(\theta)$, with the understanding that some
 1096 analyses are unbinned and use the extended likelihood formalism. The probabilities $p(\tilde{\theta} | \theta)$ encode
 1097 information on the systematic errors. Since we are still blind, we use the expected background as our
 1098 “observed” event yields.

1099 The maximum likelihood estimates or best-fit-values of μ and θ are denoted $\hat{\mu}$ and $\hat{\theta}$, while $\hat{\theta}_\mu$ denotes
 1100 the conditional maximum likelihood estimate of all nuisance parameters with μ fixed. In this analysis
 1101 the range of μ is restricted to the physically meaningful regime, *i.e.* it is not allowed to be negative.

1102 Following the fully frequentist methodology, Monte Carlo pseudo-experiments that include pseudo-
 1103 data and $\tilde{\theta}$ values of the nuisance parameters are generated to construct the pdfs. The CL_s value is
 1104 calculated as the ratio of two probabilities:

$$\text{CL}_s(\mu) = \frac{P(\tilde{q}_\mu \geq \tilde{q}_\mu^{\text{obs}} | \mu, \hat{\theta}_\mu^{\text{obs}})}{P(\tilde{q}_\mu \geq \tilde{q}_\mu^{\text{obs}} | 0, \hat{\theta}_0^{\text{obs}})}. \quad (14)$$

1105 If for $\mu = 1$, $\text{CL}_s < 0.05$, the SM Higgs boson with a nominal production rate is said to be excluded at
 1106 95 % Confidence Level (C.L.).

1107 The results are obtained from the standard shape analysis implementation within RooStats, where the
 1108 BDT shapes are allowed to vary within an envelope defined by the plots in Appendix B.

1109 14.3 Results VH

1110 Preliminary 95% C.L. upper limits (in the absence of signal) are calculated for 35.9 fb^{-1} of 13 TeV data.
 1111 The exact commands used to produce the limits are the following:

1112 Asymptotic: `combine -M Asymptotic datacard.txt -t -1`

1113 For the 1- and 2- σ bands, the CLs frequentist calculation currently recommended by the LHC Higgs
 1114 Combination Group (summarized in the previous section) is employed [73].

1115 The three channels are combined to provide limits for $m_H=125 \text{ GeV}$.

1116 The limits are shown in Fig. 100. Numerical values shown in Tab. 26.

1117 The channel compatibility and signal strength is shown in Fig. 102. At 125 GeV the signal strength for all
 1118 modes is compatible.

1119 The corresponding fitted signal strength is found to be $\mu = \sigma/\sigma_{SM} = 1.00^{+0.XX}_{-0.XX}$, with a predicted uncer-
 1120 tainty with the hypothesis of signal strength equal to unity is 0.31. The observed (expected post-fit) local
 1121 significance of the excess over background prediction is quantified as X.XX σ (2.82 σ).

1122 The exact command used to get the signal strength and significance are:

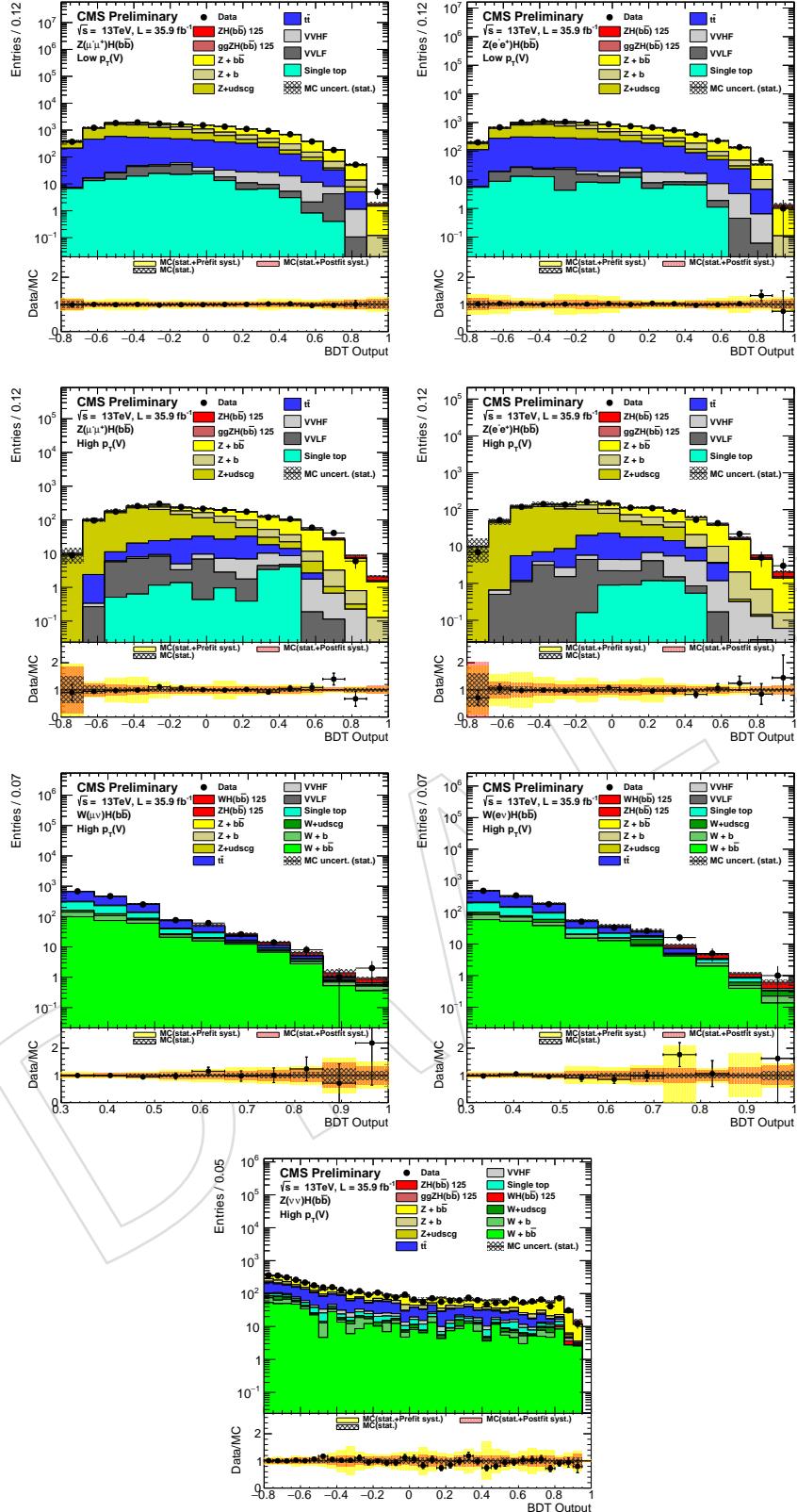


Figure 96: Post-fit distributions(VH signal is pre-fit) of BDT output for combined $Z(\ell\ell)H$ (top), $W(\mu\nu)H$ ($W(\nu\mu)H$) (middle), and $Z(\nu\nu)H$ (bottom) channels for 13 TeV 2016 data (points with errors), after all selection criteria have been applied.

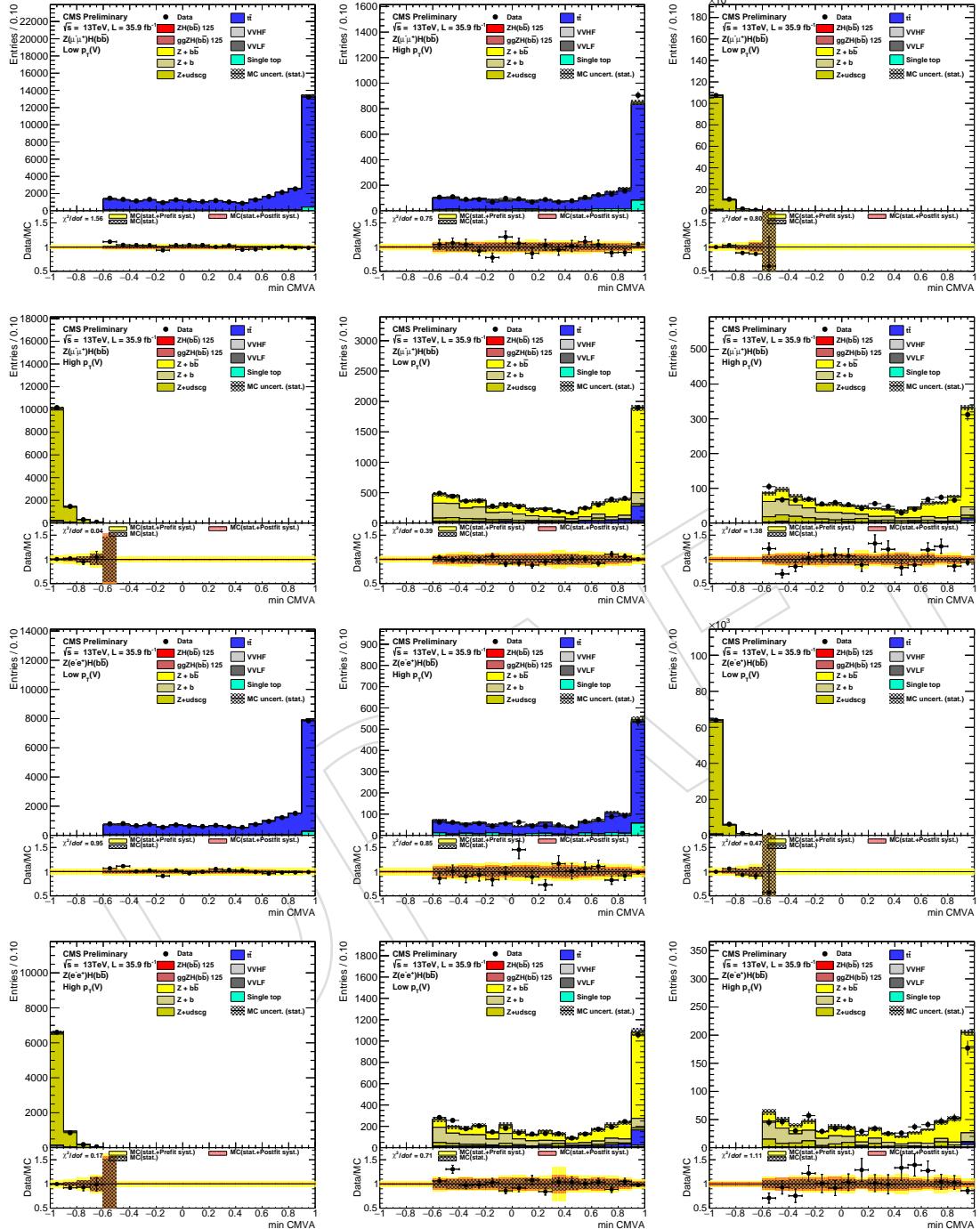


Figure 97: Post-fit distributions of the control regions for combined $Z(\ell\ell)H$ channels for 13 TeV 2016 data (points with errors), after all selection criteria have been applied.

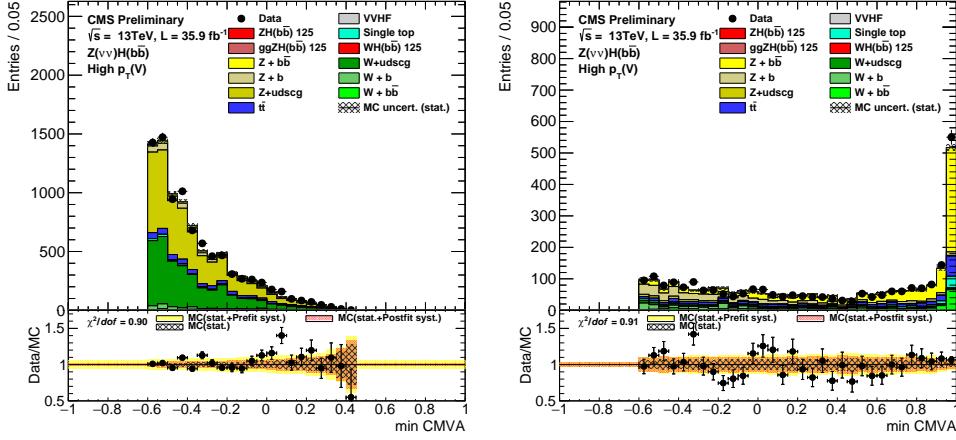


Figure 98: Post-fit distributions of the blinded BDT output in the signal region and $\text{CMVA}_{\text{V2 min}}$ in the $Z+\text{udscg}$, $Z+\text{b}$, $t\bar{t}$, and QCD control regions for $Z(\nu\nu)H$ with 13 TeV data (points with errors), all backgrounds, and signal after all selection criteria have been applied.

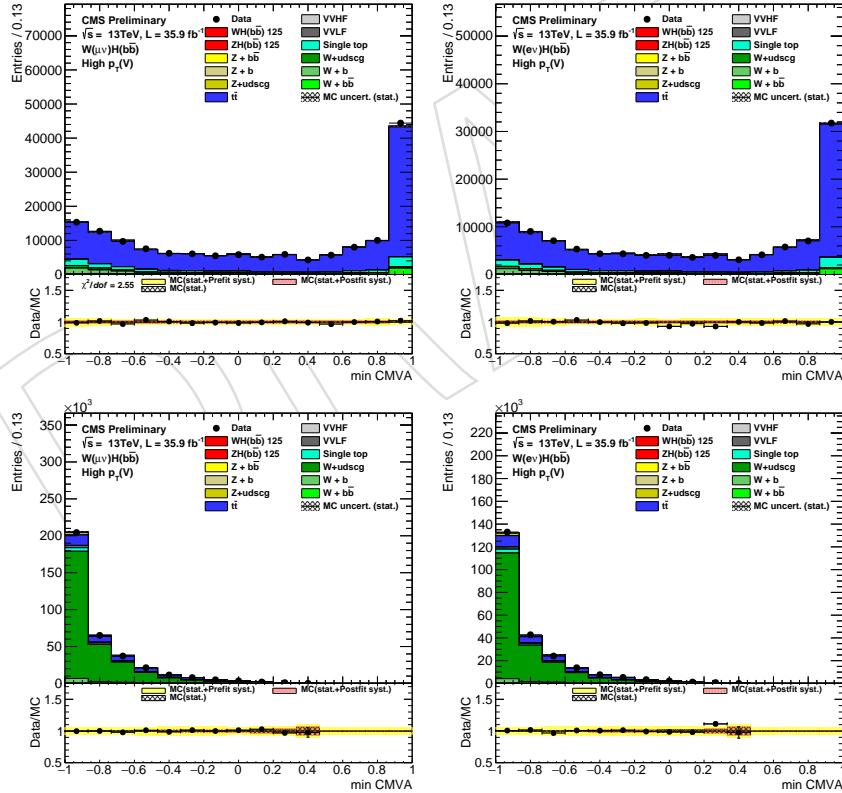


Figure 99: Post-fit distributions of blinded BDT output and min CMVA distributions from $t\bar{t}$, $W+\text{heavy flavor jets}$, and $W+\text{light flavor jets}$ for $W(\mu\nu)H$ ($W(e\nu)H$) on the left (right) in the signal region for 13 TeV 2016 data (points with errors), all backgrounds, and signal after all selection criteria have been applied.

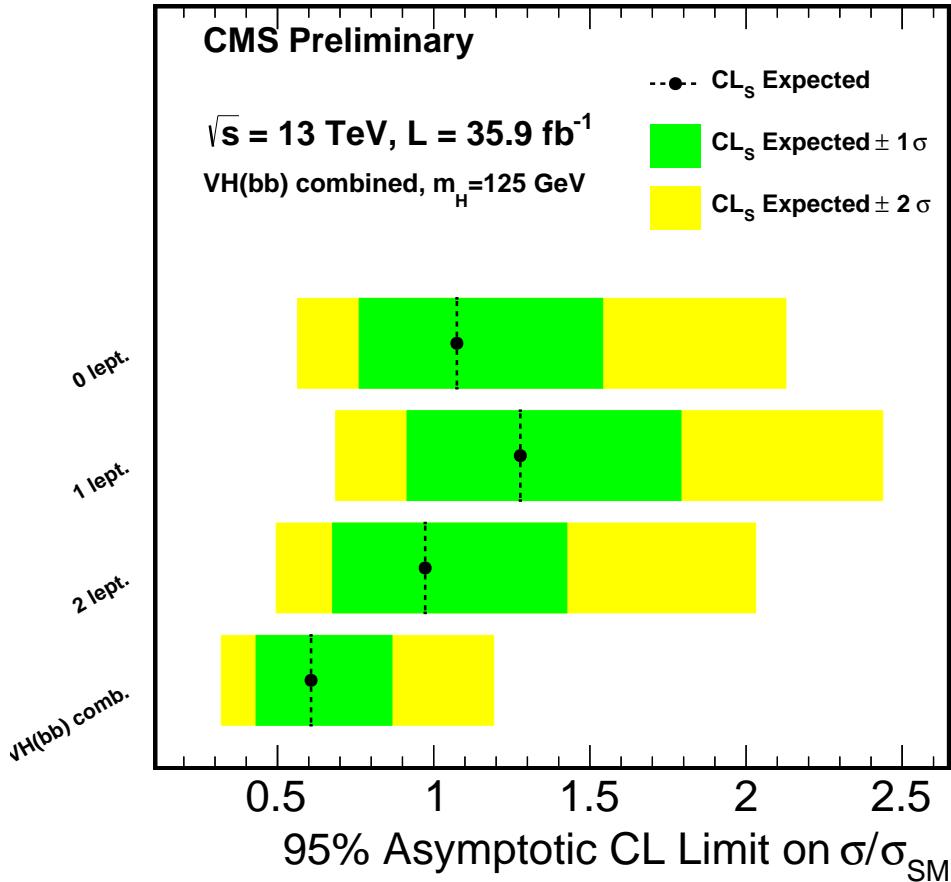


Figure 100: **BLINDED** Expected and observed 95% C.L. combined upper limits (in the absence of signal) on the ratio of VH $b\bar{b}$ production for the BDT analysis on the 13 TeV data. The median and observed expected limit and the 1- and 2- σ bands are obtained with the full LHC CLs method as implemented in `combine`.

Table 26: **BLINDED** Expected and observed 95% CL upper limits on the product of the VH production cross section times $H \rightarrow b\bar{b}$ branching ratio, with respect to the expectations for a standard model Higgs boson in the BDT analyses (13 TeV combination).

channel	$Z(\ell\ell)H$	$W(\ell\nu)H$	$Z(\nu\nu)H$	combo
BDT Exp.	1.19	1.34	1.31	0.70
BDT Obs.	3.13	3.38	1.22	1.87

```

1123 Significance: combine -M ProfileLikelihood -m 125 --signif --pvalue -t -1 --toysFreq
1124     --expectSignal=1 card
1125 Signal strength: combine -M MaxLikelihoodFit -m 125 -t -1 --expectSignal=1 --stepSize=0.05
1126     --rMin=-5 --rMax=5 --robustFit==1 --saveNorm --saveShapes --plots -v 3
1127     card
1128 For this fit we show in Fig. 101 the list of the list of problematic nuisances among the 715 used for
1129 the signal extraction. As this list is minimal, we consider this as a good test of the goodness of the
1130 systematics treatment.

```

Table 27: Expected and observed background fit (post-fit) significance for the VH production cross section times $H \rightarrow b\bar{b}$ branching ratio, for the combination and sub-channels separately.

channel	$Z(\ell\ell)H$	$W(\ell\nu)H$	$Z(\nu\nu)H$	combo
significance Exp.	1.76	1.50	1.51	2.81
significance Obs.	3.07	3.18	0	3.29

```

1131 Figure 103 combines the BDT outputs of all channels where the events are gathered in bins of similar
1132 expected signal-to-background ratio, as given by the value of the output of their corresponding BDT
1133 discriminant.

```

DRAFT

2017-6-22

Comparison of nuisances

Constrained nuisances for VHbb fit

CMS_vhbb_LHE_weights_scale_muF_TT	+0.07, 0.48	+0.10, 0.47	+0.03
CMS_vhbb_LHE_weights_scale_muF_Wj1b	+2.35, 0.92	+2.39, 0.92	+0.01
CMS_vhbb_bTagWeightHF_pt4_eta1	+0.01, 0.47	+0.04, 0.46	+0.01
CMS_vhbb_bTagWeightJES_pt0_eta1	-2.19, 0.87	-2.17, 0.87	+0.01
CMS_vhbb_bTagWeightLF_pt4_eta1	+2.41, 0.81	+2.01, 0.82	-0.17
CMS_vhbb_bTagWeightcErr1_pt0_eta1	-0.32, 0.24	-0.33, 0.24	-0.01
CMS_vhbb_bTagWeightcErr1_pt0_eta3	+0.30, 0.37	+0.30, 0.38	-0.00
CMS_vhbb_bTagWeightcErr1_pt1_eta1	+0.45, 0.43	+0.46, 0.43	+0.00
CMS_vhbb_bTagWeightcErr1_pt1_eta2	+0.38, 0.41	+0.37, 0.41	-0.00
CMS_vhbb_bTagWeightcErr1_pt3_eta1	-0.36, 0.46	-0.31, 0.47	+0.04
CMS_vhbb_eff_e_MVAID_Zll_13TeV	-0.51, 0.39	-0.51, 0.39	-0.01
CMS_vhbb_eff_e_Wln_13TeV	+0.55, 0.20	+0.54, 0.20	-0.02
CMS_vhbb_puWeight	+0.23, 0.19	+0.23, 0.19	+0.00
CMS_vhbb_scale_j_AbsoluteScale_13TeV	+0.09, 0.43	+0.06, 0.44	-0.02
CMS_vhbb_scale_j_FlavorQCD_13TeV	+0.71, 0.39	+0.75, 0.32	-0.06
CMS_vhbb_scale_j_Fragmentation_13TeV	+0.06, 0.39	+0.05, 0.38	-0.00
CMS_vhbb_scale_j_PileUpDataMC_13TeV	+0.06, 0.40	+0.04, 0.39	-0.01
CMS_vhbb_scale_j_PileUpPtBB_13TeV	+0.10, 0.36	+0.14, 0.37	+0.03
CMS_vhbb_scale_j_PileUpPtRef_13TeV	+0.20, 0.48	+0.21, 0.47	+0.03
CMS_vhbb_scale_j_RelativeFSR_13TeV	+0.22, 0.42	+0.20, 0.42	+0.00
CMS_vhbb_scale_j_TimePtEta_13TeV	+0.01, 0.39	-0.02, 0.39	-0.01

Figure 101: List of nuisances with either pulls or constraints for the maximum likelihood fit used to extract the VHbb signal strength.

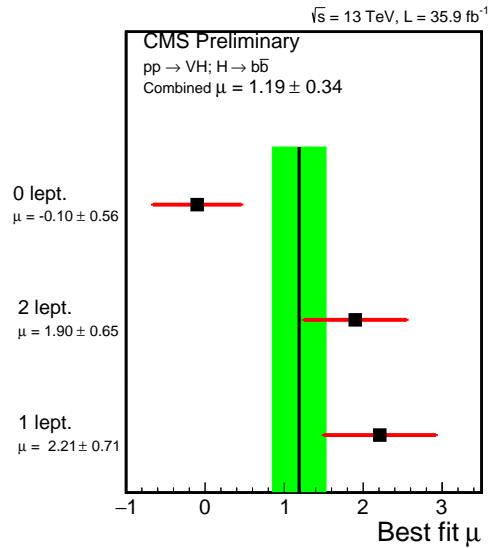


Figure 102: The most likely signal strength and uncertainty fitted in the three channels, together with the combination.

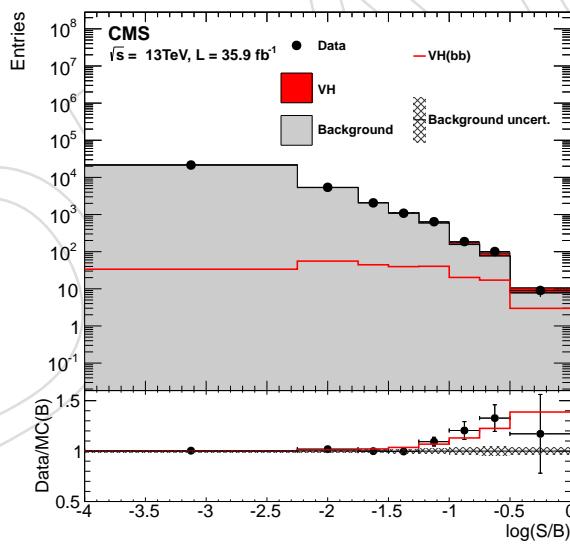


Figure 103: **BLINDED** Combination of all channels into a single distribution. Events are sorted in bins of similar expected signal-to-background ratio, as given by the value of the output of their corresponding BDT discriminant (trained with a Higgs boson mass hypothesis of 125 GeV). The bottom inset show the ratio of the data to the predicted sum of background and SM Higgs boson signal with a mass of 125 GeV. All MC templates are post-fit except for the signal template, which is pre-fit.

1134 15 Conclusions

1135 This note reports the search for the standard model Higgs production in association with vector bosons
1136 and decaying into $b\bar{b}$ pairs. A data sample of approximately 35.9 fb^{-1} corresponding to the full 2016
1137 running period has been analyzed in five modes ($Z(\mu\mu)H$, $Z(ee)H$, $Z(\nu\nu)H$, $W(\mu\nu)H$, $W(e\nu)H$) and
1138 95% C.L. upper limits are derived for the 125 GeV mass point. The expected upper limits in the absence
1139 of a signal is 0.70 times the SM prediction, while the observed upper limit is 1.87. The fitted signal
1140 strength at 125GeV is $\mu = \sigma/\sigma_{\text{SM}} = 1.19^{+0.35}_{-0.34}$, with an observed (expected) significance of 3.29 (2.81)
1141 standard deviations.

DRAFT

1142 References

- 1143 [1] A. Salam and J. C. Ward, "On a gauge theory of elementary interactions", *Nuovo Cim.* **19** (1961)
1144 165–170, doi:10.1007/BF02812723.
- 1145 [2] S. Glashow, "Partial symmetries of weak interactions", *Nucl.Phys.* **22** (1961) 579–588,
1146 doi:10.1016/0029-5582(61)90469-2.
- 1147 [3] S. Weinberg, "A model of leptons", *Phys.Rev.Lett.* **19** (1967) 1264–1266,
1148 doi:10.1103/PhysRevLett.19.1264.
- 1149 [4] F. Englert and R. Brout, "Broken symmetry and the mass of gauge vector mesons", *Phys. Rev. Lett.*
1150 **13** (1964) 321–323, doi:10.1103/PhysRevLett.13.321.
- 1151 [5] P. W. Higgs, "Broken symmetries and the masses of gauge bosons", *Phys. Rev. Lett.* **13** (1964)
1152 508–509, doi:10.1103/PhysRevLett.13.508.
- 1153 [6] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, "Global conservation laws and massless
1154 particleless", *Phys. Rev. Lett.* **13** (1964) 585–587, doi:10.1103/PhysRevLett.13.585.
- 1155 [7] CMS Collaboration, "Observation of a new boson at a mass of 125 gev with the cms experiment at
1156 the lhc", *Phys.Lett.* **B716** (2012) 30–61, doi:10.1016/j.physletb.2012.08.021,
1157 arXiv:1207.7235.
- 1158 [8] ATLAS Collaboration, "Observation of a new particle in the search for the standard model higgs
1159 boson with the atlas detector at the lhc", *Phys.Lett.* **B716** (2012) 1–29,
1160 doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.
- 1161 [9] ATLAS and CMS Collaboration, "Combined Measurement of the Higgs Boson Mass in pp
1162 Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments", *Phys. Rev. Lett.* **114**
1163 (2015) 191803, doi:10.1103/PhysRevLett.114.191803, arXiv:1503.07589.
- 1164 [10] The ATLAS and CMS Collaborations, "Measurements of the Higgs boson production and decay
1165 rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp
1166 collision data at $\sqrt{s} = 7$ and 8 TeV".
- 1167 [11] LHC Higgs Cross Section Working Group Collaboration, "Handbook of LHC Higgs Cross
1168 Sections: 4. Deciphering the Nature of the Higgs Sector", arXiv:1610.07922.
- 1169 [12] LHC Higgs Cross Section Working Group et al., "Handbook of lhc higgs cross sections: 3. higgs
1170 properties", CERN-2013-004 (CERN, Geneva, 2013) arXiv:1307.1347.
- 1171 [13] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam, "Jet substructure as a new Higgs
1172 search channel at the LHC", *Phys. Rev. Lett.* **100** (2008) 242001,
1173 doi:10.1103/PhysRevLett.100.242001, arXiv:0802.2470.
- 1174 [14] CMS Collaboration, "Search for the standard model Higgs boson decaying to bottom quarks and
1175 produced in association with a W or a Z boson", *CMS Physics Analysis Summary*
1176 **CMS-PAS-HIG-11-012** (2011).
- 1177 [15] VHbb Team, "Search for the standard model higgs boson produced in association with a w or z
1178 and decaying to bottom quarks", *CMS Note* **2011/240** (2011).
- 1179 [16] CMS Collaboration, "Search for the standard model Higgs boson decaying to bottom quarks",
1180 *CMS Physics Analysis Summary* **CMS-PAS-HIG-11-031** (2011).
- 1181 [17] VHbb Team, "Search for the standard model higgs boson produced in association with a w or z
1182 and decaying to bottom quarks", *CMS Note* **2011/430** (2011).
- 1183 [18] CMS Collaboration, "Search for the standard model Higgs boson decaying to bottom quarks in pp
1184 collisions at $\sqrt{s} = 7$ TeV", *Phys. Lett.* **B710** (2012) 284–306,
1185 doi:10.1016/j.physletb.2012.02.085, arXiv:1202.4195.
- 1186 [19] CMS Collaboration, "Search for the standard model Higgs boson produced in association with W
1187 or Z bosons, and decaying to bottom quarks (ICHEP 2012)", *CMS Physics Analysis Summary*
1188 **CMS-PAS-HIG-12-019** (2012).

- [20] Hbb subgroup, "Search for the standard model higgs boson decaying to bottom quarks and produced in association with a w or a z boson", *CMS Physics Analysis Note AN-12-181* (2011).
- [21] Hbb subgroup, "Search for the standard model higgs boson decaying to bottom quarks and produced in association with a w or a z boson (hcp)", *CMS Physics Analysis Note AN-12-349* (2012).
- [22] CMS Collaboration, "Search for the standard model Higgs boson produced in association with W or Z bosons, and decaying to bottom quarks", *CMS Physics Analysis Summary CMS-PAS-HIG-12-044* (2012).
- [23] Hbb subgroup, "Search for the standard model higgs boson decaying to bottom quarks and produced in association with a w or a z boson (lhcp 2013)", *CMS Physics Analysis Note AN-13-069* (2013).
- [24] CMS Collaboration, "Search for the standard model Higgs boson produced in association with W or Z bosons, and decaying to bottom quarks (LHCP 2013)", *CMS Physics Analysis Summary CMS-PAS-HIG-13-012* (2013).
- [25] CMS Collaboration, "Search for the standard model Higgs boson produced in association with a W or a Z boson and decaying to bottom quarks", *Phys. Rev.* **D89** (2014), no. 1, 012003, doi:10.1103/PhysRevD.89.012003, arXiv:1310.3687.
- [26] 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.* **C75** (2015), no. 5, 212, doi:10.1140/epjc/s10052-015-3351-7, arXiv:1412.8662.
- [27] CMS Collaboration, "Search for the standard model Higgs boson produced through vector boson fusion and decaying to $b\bar{b}$ ", *Phys. Rev.* **D92** (2015), no. 3, 032008, doi:10.1103/PhysRevD.92.032008, arXiv:1506.01010.
- [28] CMS Collaboration, "Search for the standard model Higgs boson produced in association with a top-quark pair in pp collisions at the LHC", *JHEP* **05** (2013) 145, doi:10.1007/JHEP05(2013)145, arXiv:1303.0763.
- [29] CMS Collaboration, "Search for the associated production of the Higgs boson with a top-quark pair", *JHEP* **09** (2014) 087, doi:10.1007/JHEP09(2014)087, 10.1007/JHEP10(2014)106, arXiv:1408.1682. [Erratum: JHEP10,106(2014)].
- [30] T. Sjöstrand, S. Mrenna, and P. Skands, "Pythia 6.4 physics and manual", *JHEP* **05** (2006) 026, doi:10.1088/1126-6708/2006/05/026, arXiv:hep-ph/0603175.
- [31] T. Sjostrand, S. Mrenna, and P. Z. Skands, "A Brief Introduction to PYTHIA 8.1", *Comput. Phys. Commun.* **178** (2008) 852–867, doi:10.1016/j.cpc.2008.01.036, arXiv:0710.3820.
- [32] P. Nason, "A new method for combining nlo qcd with shower monte carlo algorithms", *JHEP* **0411** (2004) 040, doi:10.1088/1126-6708/2004/11/040, arXiv:hep-ph/0409146.
- [33] P. Golonka et al., "The tauola photos f environment for the tauola and photos packages: Release. 2.", *Comput.Phys.Commun.* **174** (2006) 818–835, doi:10.1016/j.cpc.2005.12.018, arXiv:hep-ph/0312240.
- [34] J. Alwall et al., "Madgraph 5 : Going beyond", *JHEP* **1106** (2011) 128, doi:10.1007/JHEP06(2011)128, arXiv:1106.0522.
- [35] J. Alwall et al., "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations", *JHEP* **07** (2014) 079, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.
- [36] M. L. Mangano, M. Moretti, F. Piccinini, and M. Treccani, "Matching matrix elements and shower evolution for top-quark production in hadronic collisions", *JHEP* **01** (2007) 013, doi:10.1088/1126-6708/2007/01/013, arXiv:hep-ph/0611129.

- 1236 [37] R. Frederix and S. Frixione, "Merging meets matching in MC@NLO", *JHEP* **12** (2012) 061,
1237 doi:[10.1007/JHEP12\(2012\)061](https://doi.org/10.1007/JHEP12(2012)061), arXiv:[1209.6215](https://arxiv.org/abs/1209.6215).
- 1238 [38] CMS Collaboration, "Event generator tunes obtained from underlying event and multiparton
1239 scattering measurements", arXiv:[1512.00815](https://arxiv.org/abs/1512.00815).
- 1240 [39] NNPDF Collaboration, "Parton distributions for the LHC Run II", *JHEP* **04** (2015) 040,
1241 doi:[10.1007/JHEP04\(2015\)040](https://doi.org/10.1007/JHEP04(2015)040), arXiv:[1410.8849](https://arxiv.org/abs/1410.8849).
- 1242 [40] R. Gavin, Y. Li, F. Petriello, and S. Quackenbush, "FEWZ 2.0: A code for hadronic Z production at
1243 next-to-next-to-leading order", *Comput. Phys. Commun.* **182** (2011) 2388,
1244 doi:[10.1016/j.cpc.2011.06.008](https://doi.org/10.1016/j.cpc.2011.06.008), arXiv:[1011.3540](https://arxiv.org/abs/1011.3540).
- 1245 [41] Y. Li and F. Petriello, "Combining QCD and electroweak corrections to dilepton production in the
1246 framework of the FEWZ simulation code", *Phys. Rev. D* **86** (2012) 094034,
1247 doi:[10.1103/PhysRevD.86.094034](https://doi.org/10.1103/PhysRevD.86.094034), arXiv:[1208.5967](https://arxiv.org/abs/1208.5967).
- 1248 [42] R. Gavin, Y. Li, F. Petriello, and S. Quackenbush, "W physics at the LHC with FEWZ 2.1", *Comput.
1249 Phys. Commun.* **184** (2013) 208, doi:[10.1016/j.cpc.2012.09.005](https://doi.org/10.1016/j.cpc.2012.09.005), arXiv:[1201.5896](https://arxiv.org/abs/1201.5896).
- 1250 [43] M. Czakon, P. Fiedler, and A. Mitov, "Total Top-Quark Pair-Production Cross Section at Hadron
1251 Colliders Through $O(\alpha_S^4)$ ", *Phys. Rev. Lett.* **110** (2013) 252004,
1252 doi:[10.1103/PhysRevLett.110.252004](https://doi.org/10.1103/PhysRevLett.110.252004), arXiv:[1303.6254](https://arxiv.org/abs/1303.6254).
- 1253 [44] N. Kidonakis, "Differential and total cross sections for top pair and single top production", (2012).
1254 arXiv:[1205.3453](https://arxiv.org/abs/1205.3453).
- 1255 [45] LHC Higgs Cross Section Working Group et al., "Handbook of lhc higgs cross sections: 1.
1256 inclusive observables", CERN-2011-002 (CERN, Geneva, 2011) arXiv:[1101.0593](https://arxiv.org/abs/1101.0593).
- 1257 [46] CMS Physics Analysis Tools, "Heppy : a mini framework for HEP event processing in python".
1258 <https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideHeppy>.
- 1259 [47] CMS Collaboration, "Description and performance of track and primary-vertex reconstruction
1260 with the CMS tracker", *JINST* **9** (2014), no. 10, P10009,
1261 doi:[10.1088/1748-0221/9/10/P10009](https://doi.org/10.1088/1748-0221/9/10/P10009), arXiv:[1405.6569](https://arxiv.org/abs/1405.6569).
- 1262 [48] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet User Manual", *Eur. Phys. J.* **C72** (2012) 1896,
1263 doi:[10.1140/epjc/s10052-012-1896-2](https://doi.org/10.1140/epjc/s10052-012-1896-2), arXiv:[1111.6097](https://arxiv.org/abs/1111.6097).
- 1264 [49] CMS Physics Analysis Tools, "Pileup reweighting procedure". [https://twiki.cern.ch/...](https://twiki.cern.ch/)
- 1265 [50] CMS Collaboration, "Performance of Electron Reconstruction and Selection with the CMS
1266 Detector in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV", *JINST* **10** (2015), no. 06, P06005,
1267 doi:[10.1088/1748-0221/10/06/P06005](https://doi.org/10.1088/1748-0221/10/06/P06005), arXiv:[1502.02701](https://arxiv.org/abs/1502.02701).
- 1268 [51] CMS Electron/gamma Physics Object Group, "recipes for Run 2 ID criteria".
1269 <https://twiki.cern.ch/twiki/bin/view/CMS/EgammaIDRecipesRun2>.
- 1270 [52] CMS Collaboration, "Performance of muon identification in pp collisions at $\sqrt{s} = 7$ TeV", CMS
1271 Physics Analysis Summary CMS-PAS-MUO-10-002, 2010.
- 1272 [53] CMS Collaboration, "Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$
1273 TeV", *JINST* **7** (2012) P10002, doi:[10.1088/1748-0221/7/10/P10002](https://doi.org/10.1088/1748-0221/7/10/P10002), arXiv:[1206.4071](https://arxiv.org/abs/1206.4071).
- 1274 [54] CMS Collaboration, "Particle-flow event reconstruction in CMS and performance for jets, taus,
1275 and E_T^{miss} ", CMS Physics Analysis Summary CMS-PAS-PFT-09-001, CERN, 2009.
- 1276 [55] CMS Collaboration, "Commissioning of the particle-flow event reconstruction with the first lhc
1277 collisions recorded in the cms detector", CMS Physics Analysis Summary CMS-PAS-PFT-10-001,
1278 CERN, 2010.
- 1279 [56] M. Cacciari and G. P. Salam, "Dispelling the N^3 myth for the k_t jet-finder", *Phys. Lett. B* **641** (2006)
1280 57, doi:[10.1016/j.physletb.2006.08.037](https://doi.org/10.1016/j.physletb.2006.08.037), arXiv:[hep-ph/0512210](https://arxiv.org/abs/hep-ph/0512210).
- 1281 [57] M. Cacciari, G. P. Salam, and G. Soyez, "The anti- k_t jet clustering algorithm", *JHEP* **04** (2008) 063,
1282 doi:[10.1088/1126-6708/2008/04/063](https://doi.org/10.1088/1126-6708/2008/04/063), arXiv:[0802.1189](https://arxiv.org/abs/0802.1189).

- 1283 [58] CMS Collaboration, "Determination of Jet Energy Calibration and Transverse Momentum
1284 Resolution in CMS", *JINST* **6** (2011) P11002, doi:10.1088/1748-0221/6/11/P11002,
1285 arXiv:1107.4277.
- 1286 [59] CMS Collaboration, "Pileup jet identification", CMS Physics Analysis Summary
1287 CMS-PAS-JME-13-005, CERN, 2013.
- 1288 [60] CMS Collaboration, "Identification of b-quark jets with the CMS experiment", *JINST* **8** (2013)
1289 P04013, doi:10.1088/1748-0221/8/04/P04013, arXiv:1211.4462.
- 1290 [61] CMS BTV POG, "Event reweighting using scale factors calculated with a tag and probe method".
1291 <https://twiki.cern.ch/twiki/bin/view/CMS/BTagShapeCalibration>.
- 1292 [62] CMS Collaboration, "Performance of the CMS missing transverse momentum reconstruction in
1293 pp data at $\sqrt{s} = 8$ TeV", *JINST* **10** (2015), no. 02, P02006,
1294 doi:10.1088/1748-0221/10/02/P02006, arXiv:1411.0511.
- 1295 [63] CMS Collaboration, "Commissioning of trackjets in pp collisions at $\sqrt{s}=7$ TeV", CMS Physics
1296 Analysis Summary CMS-PAS-JME-10-006, CERN, 2010.
- 1297 [64] CMS Collaboration, "Performance of jet reconstruction with charged tracks only", CMS Physics
1298 Analysis Summary CMS-PAS-JME-08-001, CERN, 2009.
- 1299 [65] S. P. M. S. Stefan Kallweit, Jonas M. Lindert and P. Maierhfer, "Nlo qcd+ew predictions for v+jets
1300 including off-shell vector-boson decays and multijet merging", *Phys. Rev. Lett.* (Mar, 2016)
1301 doi:10.1007/JHEP04(2016)021.
- 1302 [66] CDF, D0 Collaboration, "Improved b -jet Energy Correction for $H \rightarrow b\bar{b}$ Searches at CDF",
1303 arXiv:1107.3026.
- 1304 [67] H. Voss, A. Höcker, J. Stelzer, and F. Tegenfeldt, "TMVA, the toolkit for multivariate data analysis
1305 with ROOT", in *XIth International Workshop on Advanced Computing and Analysis Techniques in*
1306 *Physics Research (ACAT)*, p. 40. 2007. arXiv:physics/0703039.
- 1307 [68] LHC Higgs Cross Section Working Group et al., "Handbook of lhc higgs cross sections: Inclusive
1308 observables", arXiv:1101.0593.
- 1309 [69] M. Ciccolini, A. Denner, and S. Dittmaier, "Strong and electroweak corrections to the production
1310 of Higgs+2jets via weak interactions at the LHC", *Phys. Rev. Lett.* **99** (2007) 161803,
1311 doi:10.1103/PhysRevLett.99.161803, arXiv:0707.0381.
- 1312 [70] M. Ciccolini, A. Denner, and S. Dittmaier, "Electroweak and QCD corrections to Higgs production
1313 via vector-boson fusion at the LHC", *Phys. Rev.* **D77** (2008) 013002,
1314 doi:10.1103/PhysRevD.77.013002, arXiv:0710.4749.
- 1315 [71] Stefan Dittmaier, talk at EPS 2011.
- 1316 [72] G. Ferrera, M. Grazzini, and F. Tramontano, "Associated WH production at hadron colliders: a
1317 fully exclusive QCD calculation at NNLO", arXiv:1107.1164. * Temporary entry *.
- 1318 [73] CMS Higgs Combination, "Search for standard model higgs boson in pp collisions at $\sqrt{s} = 7$ ",
1319 *CMS Physics Analysis Summary, HIG-11-011 HIG-11-011* (2011).